Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

In silico genomic characterization of putative metal-resistance and metal- homeostasis determinants in Achromobacter strain H21

Abdallah Saeed

s.abdallah@uds.edu.gh

University for Development Studies

Research Article

Keywords: Achromobacter strain H21, heavy-metal resistance, metal homeostasis, arsenic resistance, genomic analysis, conserved domains, RND efflux transporters, P-type ATPases

Posted Date: November 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11364033/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

In silico genomic characterization of putative metal-resistance and metal-homeostasis determinants in *Achromobacter* strain H21

Abdallah Saeed

Department of Biotechnology and Molecular Biology, Faculty of Biosciences, University for Development Studies, Tamale, Ghana

Correspondence author email: s.abdallah@uds.edu.gh

Abstract

Heavy-metal contamination presents significant environmental challenges due to the persistence and toxicity of metallic and metalloid pollutants. This study aimed to identify and characterize putative heavy-metal resistance and metal-homeostasis determinants in *Achromobacter* strain H21 using integrated bioinformatics approaches. The draft genome (NCBI RefSeq GCF_054172185.1) was investigated through genome annotation, TBLASTN and BLASTN sequence-similarity searches, conserved-domain analysis, and genomic-neighbourhood reconstruction. The genome comprised 6,184,534 bp across 65 scaffolds, with 66.12% GC content, 5,915 annotated gene features, 5,940 coding-sequence features and 5,851 predicted proteins. Annotation screening identified candidate genes associated with arsenic, chromate and divalent-metal transport, alongside regulatory and stress-response proteins. TBLASTN analysis of 16 protein queries generated 90 alignments, while BLASTN screening of 578 nucleotide references produced 349 high-scoring segment pairs. Following filtering at ≥70% identity, ≥70% query coverage and E-value $\leq 1 \times 10^{-5}$, 17 alignments representing 15 unique proteins were retained. Conserved-domain analysis identified recognizable domains in all 15 proteins, supporting the classification of ArsB-like transporters, an ArsH-family flavoprotein, CzcA- and MexI-like RND transporters, ZntA- and GolT-like P-type ATPases, and regulatory proteins. Genomic-neighbourhood analysis revealed distinct arsenic-associated regions and a MexI-like locus containing putative tripartite efflux components. Preliminary taxonomic comparison indicated approximately 90.31% average nucleotide identity with the tested *Achromobacter denitrificans* reference, warranting cautious species assignment. Collectively, these findings indicate diverse putative metal-associated determinants in H21, providing a genomic foundation for future experimental validation of metal-resistance mechanisms.

Keywords: *Achromobacter* strain H21; heavy-metal resistance; metal homeostasis; arsenic resistance; genomic analysis; conserved domains; RND efflux transporters; P-type ATPases.

1. Introduction

Heavy metals have become a serious environmental challenge affecting various aspects of the environment, including land, water, and living things. Human activities such as mining, industrial activities, and waste disposal account for the introduction of these heavy metals int o the environment. These metals are characterized as being persistent and bioavailable, which is responsible for their devastating effects. Heavy metals promote the generation of reactive oxygen species (ROS) and reactive nitrogen species (RNS), upsetting redox balance and causing damage to lipids, proteins, and DNA. This oxidative assault is repeatedly described as a core driver of toxicity for metals such as mercury, lead, arsenic, cadmium, chromium, and nickel (Salazar-Flores et al., 2019; Jomová et al., 2024). These environmental challenges, as well as the health-related risks, account for a serious need for studies on these heavy metals.

According to Hrynkiewicz et al. (2014), Khalifa et al. (2020), and Lamb & Burheim (2021), heavy metals may interfere with membrane-associated activities, causing metal inflow or cellular content leakage. This can lead to growth suppression and altered physiology in bacteria under metal stress. Bacteria use a number of strategies, including biofilm formation, sequestration, and active efflux pumps, to withstand and tolerate heavy metals. Under metal-stressed environments, these pathways increase the likelihood of bacterial survival.

Several bacterial genes have been associated with metal resistance. For instance, the classic arsRBC operon encodes ArsR (sensor/repressor), ArsB (arsenite efflux pump), and ArsC (arsenate reductase), which form a more efficient ArsAB pump accounting for arsenic resistance across diverse bacterial taxa (Sánchez-Corona et al., 2025). Also, in many bacteria, ChrA mediates Cr(VI) efflux, reducing intracellular Cr toxicity, while ChrB responds to Cr (VI) concentrations to regulate expression, with variants such as chrABC or chrBAC described in multiple species, including Pseudomonas and Alcaligenes derivatives (Ding et al., 2024; Hoque et al., 2024).

Achromobacter is a genus of aerobic, non-fermenting Gram-negative bacilli increasingly recognized as emerging pathogens in cystic fibrosis (CF) and other clinical settings, as well as other environmental reservoirs, including domestic and built environments, and have been isolated from various niches outside the human host. Accurate identification at the species level is challenging with many routine methods, and misidentifications are well documented across

MALDI-TOF MS, 16S rRNA sequencing, and housekeeping gene approaches (Gomila et al., 2014). Intra-genus diversity revealed by MLSA and nrdA sequencing correlates with differences in metabolic capabilities and ecological roles, informing how different species contribute to environmental processes and resilience in ecosystems (Coward et al.,2016).

Comparative genomics and specialized bioinformatics workflows enable the identification of metal resistance genes (MRGs) across bacterial genomes. Robust pipelines integrate multiple data layers (Genomes, MAGs, Virulence and Resistance databases) and employ both reference-based and de novo gene discovery to capture known and novel metal resistance determinants (PitaGaleana et al., 2025; Correia, 2026). Databases commonly used include BacMet for antibacterial biocide and metal resistance genes, and MEGARes for resistance gene annotation in metagenomic datasets. For cross-species and environmental studies, integrating multiple repositories improves discovery of both known and novel MRGs, including co-localized efflux, sequestration, and enzymatic detoxification genes (Passi et al., 2021).

Although genome sequencing and automated annotation have improved the identification of bacterial metal-associated genes, assigning precise biological functions remains challenging. Automated genome annotation pipelines frequently misassign or miss protein functions due to limitations in homology-based predictions, incomplete or biased databases, and the complexity of protein functions. This error arises because large public protein databases contain many entries annotated by computational inference rather than experimental validation, leading to systematic misannotation that can spread across related proteins and genomes (Schnoes et al., 2009)

Using an integrated bioinformatics approach, the current work sought to discover and describe potential factors of metal homeostasis and heavy-metal resistance in Achromobacter strain H21. The research specifically looked at the genome size, GC content, gene content, and assembly quality of H21. BLASTN and TBLASTN sequence-similarity searches and genome annotation were used to find potential metal-associated genes. The predicted functions of selected proteins were further evaluated using conserved domain analysis, while their genomic organization was examined to identify potential associations among transporters, regulators and neighbouring genes. The study sought to provide a genomic foundation for understanding potential metal-homeostasis mechanisms in strain H21 and prioritizing candidate genes for future experimental validation.

2. MATERIALS AND METHODS

2.1. Study design and genomic data sources

A genome-based, computational analysis was conducted using the publicly deposited draft genome of Achromobacter strain H21 (RefSeq assembly GCF_054172185.1; GenBank assembly GCA_054172185.1; WGS project JBSYXR01). The NCBI genome annotation files, including genomic feature coordinates (GFF) and predicted amino-acid sequences (protein FASTA), were used as the primary sources for candidate identification, conserved-domain analysis and genomic-context mapping. The work evaluated predicted functions; no phenotypic metal-tolerance or transport assays were performed as part of the analyses described here.

2.2. Genome assembly and annotation characterization

Assembly-level metrics, including genome length, number of scaffolds and GC content, were compiled from the genome record and associated assembly assessment. Annotated gene and coding-sequence (CDS) features were enumerated from the genomic GFF, and protein records were counted from the protein FASTA file. The counts of CDS features and unique protein sequences were treated separately because they are not necessarily identical. Previously reported genome quality estimates (CheckM completeness and contamination) were retained as assembly metadata rather than described as newly generated measurements.

2.3. Annotation-based identification of metal-associated candidate genes

The predicted protein descriptions and genomic feature annotations were screened for terms associated with arsenic, chromate, copper, cadmium, cobalt, nickel, zinc, lead, and general metal transport or regulation. Candidate annotations included ars-family proteins, ChrA-like transporters, Czc components, metal-translocating P-type ATPases, RcnA-family transporters, CorA-like transporters, ABC transport components, multicopper oxidases, and associated regulators. Hits were manually reviewed in genomic context to distinguish plausible metal-homeostasis determinants from nonspecific matches such as metalloenzymes. Annotation terms were used for candidate prioritization and were not considered proof of substrate specificity or resistance phenotype.

2.4. Reference sequence-similarity assessment

Protein-based and nucleotide-based reference searches were treated as distinct analyses. In the protein-query analysis, 16 selected H21 protein sequences were compared with a nucleotide metal-resistance reference collection using TBLASTN; 90 local alignments were recorded. Protein identity, alignment coverage and E-values were examined, with reference coverage distinguished from query coverage. Exact software version, reference database composition and command-line settings for this analysis remain to be verified from the original logs.

BLASTN was used to screen 578 metal-associated nucleotide reference sequences against the H21 genome for the nucleotide-query analysis that followed. 133 reference queries were represented by the 349 local high-scoring segment pairs (HSPs) that were produced. 17 HSPs were obtained from 16 reference queries using a filtering process that maintained nucleotide identity ≥70%, query coverage ≥70%, and E-value $\leq 1 \times 10^{-2}$. Hits were reconciled with the H21 genome annotation to obtain 15 distinct protein accessions. HSPs, reference queries, and unique H21 proteins were counted separately to avoid inflating candidate numbers. The BLASTN software version and any additional search parameters should be checked against the original command history.

2.5. Conserved-domain identification and interpretation

Predicted protein sequences were assessed using the NCBI Conserved Domain Database (CDD) through Batch CD-Search. The 15 distinct proteins mapped from the filtered BLASTN results were part of the recorded batch analysis (batch identifier QM3-qcdsearch-1C6109358FF7417F). Family assignments, domain coordinates, significance values, and differences between particular and superfamily matches were analyzed in conserved-domain records. There was at least one documented domain hit for each of the 15 proteins. 427 clusters of 727 overlapping hits were found in the output; these results were regarded as database output counts rather than as the number of distinct functional domains.

Domain evidence was reviewed for ArsB-family permeases, ArsH, RND transporters, P-type ATPases, CueR-like and related regulators, ModB, PstB/PstC, Dps, RuvB and SodB. Family-level predictions were not interpreted as experimental proof of metal specificity, gene expression, or resistance. The earlier candidate-specific CDD searches, and their recorded significance threshold, were retained separately from the later 15-protein batch analysis.

2.6. Genomic-neighborhood reconstruction

CDS records were extracted from the genomic GFF using scaffold identifiers and genomic coordinate windows around candidate proteins. For each feature, the locus tag, protein accession, annotated product, start and end positions, and strand were examined. Neighboring genes were evaluated for shared orientation, short intergenic spacing, coding-sequence overlap, and potential functional association. Multiple CDS records carrying the same protein identifier were handled as one annotated protein for overview diagrams where appropriate, while preserving awareness of split CDS features and programmed frameshifts. Genomic proximity and co-orientation were not interpreted as definitive evidence of operon membership.

Additional focused genomic-neighbourhood inspections were conducted for the second ArsB-family candidate WP_457986626.1 (ACXIUT_RS15630) on NZ_JBSYXR010000009.1 and the MexI/MexW-like RND transporter WP_457985789.1 (ACXIUT_RS10815) on NZ_JBSYXR010000005.1. Annotated gene start and end coordinates, strand orientation, predicted products, and adjacent intergenic distances were used to draw strand-aware neighbourhood diagrams. For the ArsB-like candidate, the approximately 10-kb flanking interval examined was screened for annotated arsenic-associated genes. For the RND locus, neighbouring periplasmic adaptor and outer-membrane efflux component annotations were assessed for possible tripartite efflux-system organization. Physical proximity and orientation were interpreted as evidence of genomic organization only, not confirmation of operons, co-transcription or substrate specificity.

2.7. Visualization of metal-associated genomic regions

Genomic-neighborhood schematics were generated in R using ggplot2 and gggenes. Gene arrows were placed at annotated start and end coordinates and oriented according to the recorded strand; fill colors were assigned by predicted protein category. The visualizations included two arsenic-associated regions, chromate transport candidates, a CzcABC-associated region, a copper-associated region, ZntA- and GolT-like ATPase regions, and an RcnA-associated transport region. Figures were exported in raster (PNG) and vector (PDF) formats. Gene diagrams are schematic representations of annotated loci and do not independently validate gene function or co-transcription.

2.8. Characterization of additional metal-transport candidates

The RcnA-associated region on scaffold NZ_JBSYXR010000002.1 was investigated using domain searches of WP_457984698.1 and its neighboring ABC transport components (WP_457984700.1–WP_457984702.1). A separate CorA-like candidate (WP_457986084.1) on scaffold NZ_JBSYXR010000006.1 was assessed by conserved-domain analysis and neighborhood inspection. RcnA, TroA-like, ZnuB, ZnuC and MtCorA-like assignments were recorded according to their respective CDD evidence. ABC transporter substrate specificity and CorA-mediated nickel/cobalt transport were not assumed solely from gene names.

2.9. Preliminary antimicrobial-resistance and taxonomic analyses

Preliminary screening for antibiotic resistance and taxonomic comparisons are also included in the Results. The precise databases, reference versions, instruments, identity/coverage criteria, and processes used in the initial workflow should be used to define these analyzes. This part is kept as a reporting placeholder rather than providing unverified methods since those parameters are not included in the accessible Results text. In the absence of species-level resolution, the taxonomic classification is conservatively given as Achromobacter strain H21.

2.10. Data handling and interpretation

R was used for data extraction, filtering, and summary tabulation. Stable protein accessions and genomic coordinates were used to track candidate proteins. E-values shown as zero were interpreted as values below the output reporting precision. Numerical evidence from CDD and sequence similarity was provided as produced by the corresponding services. Without experimental data, we categorized the results as computational predictions and deduced no direct evidence of metal resistance, gene expression, transport direction, or regulatory interaction.

2.13. Overview of the computational workflow

The computational workflow comprised acquisition of the H21 draft-genome resources, assembly and annotation characterization, candidate-gene discovery, sequence-similarity screening, conserved-domain validation, genomic-neighbourhood reconstruction, and evidence-based reporting. We evaluated annotation-derived and similarity-derived candidates together; conserved domains and gene context strengthened functional assignments without treating predicted

functions as experimentally verified resistance phenotypes. Figure 1 summarizes the principal analytical stages.

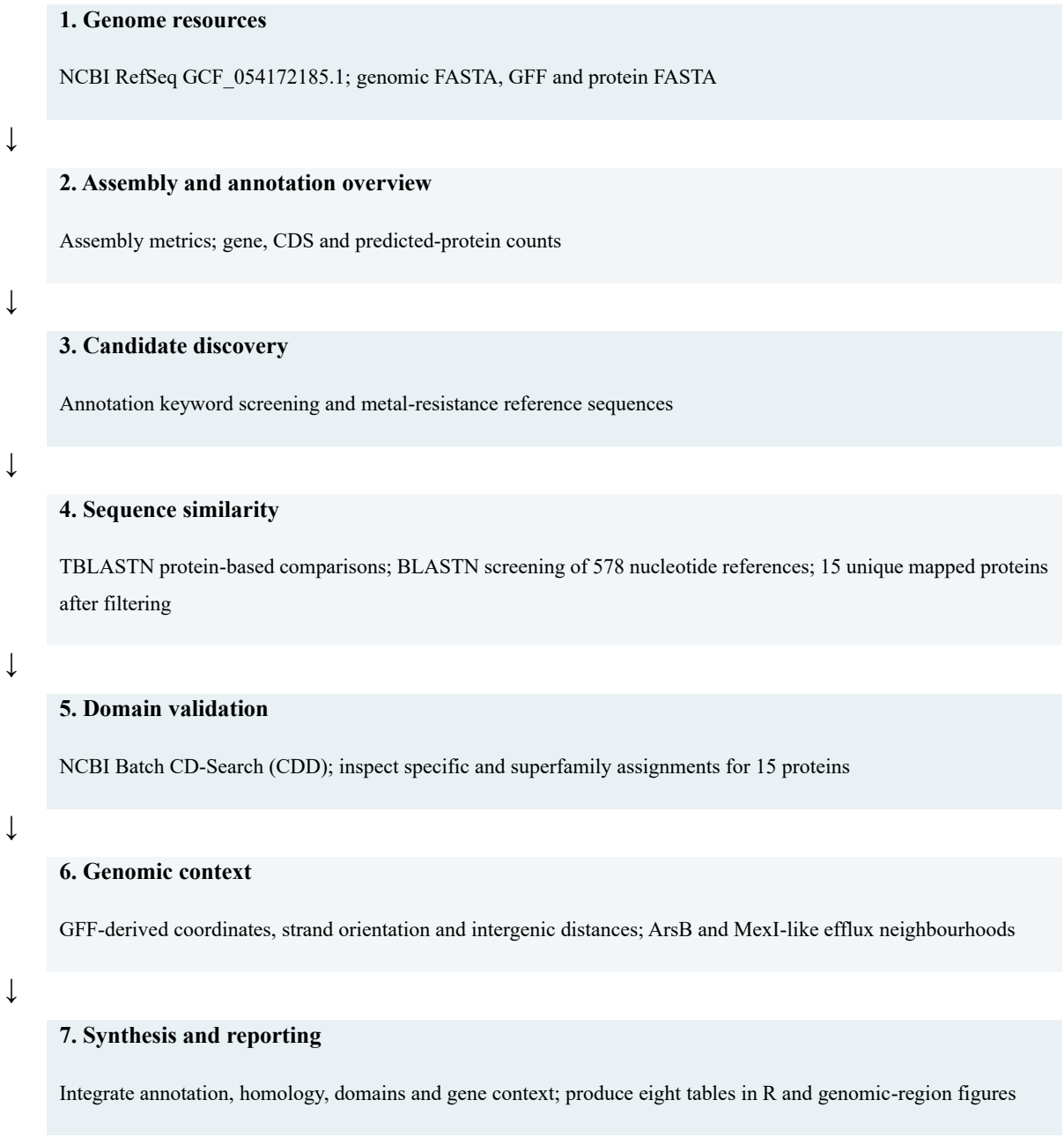


Figure 1. Computational workflow for identification and characterization of putative metal-resistance and metal-homeostasis determinants in Achromobacter strain H21.

The workflow depicts analytical order rather than independent experimental validation. Candidate selection was iterative, and the sequence-similarity, conserved-domain and neighbourhood findings were cross-checked before assigning tentative functional interpretations.

Method details to verify before journal submission

NCBI download date and annotation release used for GFF/protein FASTA.

Exact R and package versions, BLAST+ version, TBLASTN settings and reference database source.

Methods and cutoffs used for antimicrobial-resistance screening and taxonomic comparison (including ANI).

The analytic summaries were used to create eight manuscript tables, which were then imported into R for verification and display. Each file's row and column counts were confirmed. Tables were formatted and unique PDF layouts with consistent headers, row coloring, and text wrapping were exported using a base-R script. The underlying sequence-similarity and conserved-domain findings were unaffected by the table-design stage, which was a presentation technique. The final analytic environment should be used to document the precise R version and script revision.

2.11. Integration of sequence, domain and genomic-context evidence

Sequence-similarity evidence was compared with projected protein descriptions, CDD family designations, and, if available, gene-neighborhood context for the 15 BLASTN-mapped protein accessions. Instead of using individual BLAST HSP or overlapping domain hits, the evidence was arranged according to accession. Although candidate prioritizing was supported by concordance across sources, this was not regarded as experimental functional validation.

3. RESULTS

3.1. Genome assembly characteristics and functional annotation

Achromobacter strain H21's draft genome (GenBank GCA_054172185.1; RefSeq GCF_054172185.1) has a GC content of 66.12% and had 6,184,534 bp spread over 65 scaffolds. With a mean of 95,146.68 bp and a N50 of 240,512 bp, scaffold lengths varied from 521 to 635,563 bp. The assembly has 124 contigs, a reported sequencing coverage of 98.31×, and a contig N50 of around 98.9 kb. The evaluation of genome quality revealed 4.71% contamination and 94.02% completeness.

5,940 coding-sequence (CDS) features and 5,915 gene characteristics were found by genome annotation. 5,851 predicted protein sequences with a mean of 314.15 and a median of 284 amino acids, ranging from 26 to 2,525 amino acids, were included in the protein FASTA collection. There were 7,167 unclear nucleotide locations in the assembly. Metal-associated candidates were screened and specific genomic areas were examined using these genome resources (Table 1).

Table 1. Genome assembly and annotation characteristics of strain H21.

Characteristic	Value
Genome size (bp)	6,184,534
Scaffolds	65
Contigs	124
GC content (%)	66.12
Scaffold N50 (bp)	240,512
Contig N50 (kb)	98.9
Longest scaffold (bp)	635,563
Coverage (×)	98.31
Completeness (%)	94.02
Contamination (%)	4.71
Annotated gene features	5,915
CDS features	5,940
Predicted proteins	5,851

3.2. Identification of putative heavy-metal resistance-associated genes

Annotation-based screening identified candidate proteins associated with metal efflux, transport, sensing, regulation, and intracellular metal handling. These included RND-family transporters, heavy-metal-translocating P-type ATPases, arsenic transporters, arsenical pump-driving ATPases, ArsD-family metallochaperones, chromate transporters, copper-associated proteins, and metal-responsive transcriptional regulators. Additional annotations included Nramp-, ZIP-, nickel/cobalt-, and CorA-family transporters. As several of these families participate in essential metal homeostasis rather than resistance, their annotations were not considered direct evidence of metal-tolerance phenotypes. Sixteen selected candidates were examined by sequence-similarity analysis.

3.3. Sequence-similarity analysis of selected metal-resistance candidates

TBLASTN comparison of 16 H21 protein sequences against the nucleotide MetalResistance reference collection produced 90 local alignments. Thirteen queries yielded matches at an E-value threshold of 10^{-5}, whereas WP_457988408.1, WP_457985772.1, and WP_457985774.1 did not yield qualifying matches under these conditions. The 90 alignments should not be interpreted as 90 distinct resistance genes.

The strongest matches included WP_373980100.1 to CzcA (78.49% identity; 99% reference coverage), WP_457984032.1 to CzcR (61.67%; 97%), WP_457988399.1 to ArsB (77.00%; 99%), and WP_457988400.1 to ArsA (67.30%; 98%). WP_457985413.1 matched PbrA (64.94%; 86%) and WP_457986202.1 matched GolT (71.39%; 89%). Chromate-associated matches had comparatively lower identities or reference coverage (Table 2; Figure 1).

Table 2. Selected TBLASTN matches to metal-resistance-associated reference sequences

H21 accession	Reference	Identity (%)	Ref. cov. (%)	E-value
WP_373980100.1	CzcA	78.49	99	0
WP_457984032.1	CzcR	61.67	97	5.13×10^{-89}
WP_457985413.1	PbrA	64.94	86	0
WP_457986202.1	GolT	71.39	89	0
WP_457986626.1	ArsB	73.46	96	0
WP_457988399.1	ArsB	77.00	99	0
WP_457988400.1	ArsA	67.30	98	0
WP_457984031.1	CusS	36.63	81	8.80×10^{-70}
WP_457986782.1	ArsB ref.	42.78	95	1.18×10^{-83}
WP_457987913.1	SrpC	31.31	57	1.03×10^{-10}
WP_457987914.1	ChrA	32.76	59	6.42×10^{-12}
WP_457988352.1	ChrA1	38.54	23	2.09×10^{-13}

Coverage refers to reference coverage, not query coverage. An E-value reported as 0 is below the software reporting precision. The full 13-hit dataset should be verified and supplied as supplementary data.

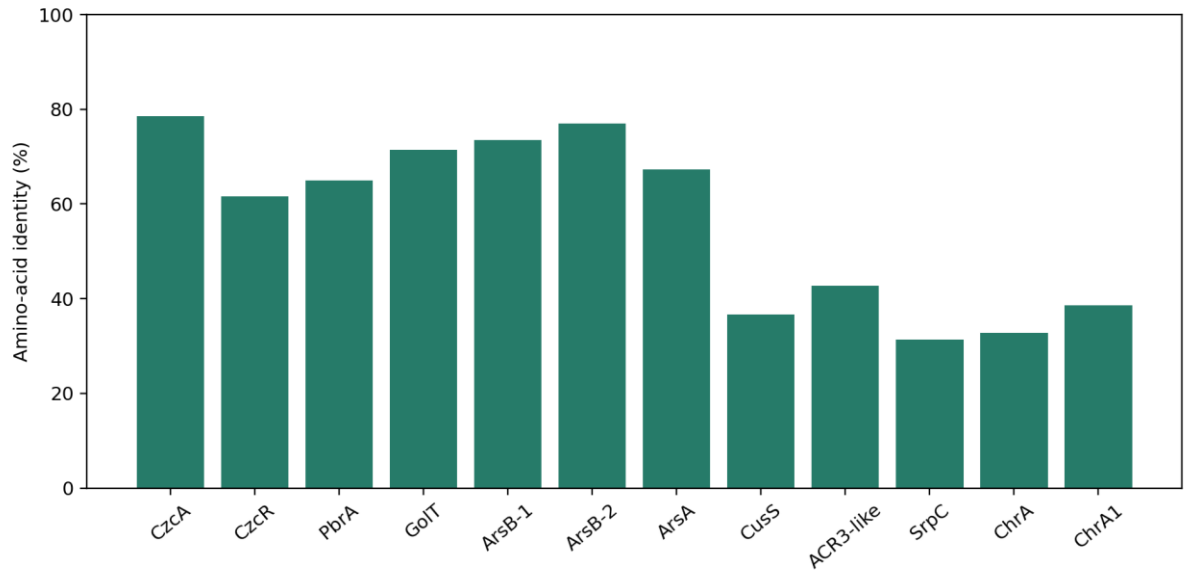


Figure 2. Amino-acid identity of selected H21 proteins relative to their matched metal-resistance reference proteins. Protein identities are presented for descriptive comparison and must be interpreted with alignment coverage and reference annotation (Table 2).

3.3a. Nucleotide-reference homology screening and mapping to annotated CDS

Using a set of 578 metal-associated nucleotide reference sequences (MetalResistance.fa), a complementary nucleotide-level homology search was conducted against the RefSeq genome assembly GCF_054172185.1. BLASTN (NCBI BLAST+ v2.17.0+) was run with the blastn task, an E-value threshold of 1×10^{-5}, DUST filtering enabled, and a maximum of 100 target sequences per query. The screen returned 349 local alignments involving 133 distinct reference sequences (23.0% of the collection). These counts describe sequence matches, not experimentally confirmed resistance genes.

Individual high-scoring segment pairs (HSPs) had a median nucleotide identity of 71.15% (range 63.88–96.88%) and a median per-HSP query coverage of 14% (range 1–99%). Exploratory filtering for identity ≥70%, per-HSP query coverage ≥70%, and E-value $\leq 1 \times 10^{-5}$ retained 17 alignments involving 16 reference sequences. These alignments were allocated to 15 different H21 protein accessions based on coordinate overlap with identified CDS characteristics. Two separate ruvB references mapped to WP_457985600.1 and two independent czcA references mapped to WP_373980100.1 at the same genomic coordinates; neither pair was counted as two unique H21 genes.

The ArsB-like references among the preserved loci matched to two distinct areas (WP_457988399.1 on scaffold 25 and WP_457986626.1 on scaffold 9), while an ArsH-like reference mapped to WP_457986785.1 on scaffold 10. Additional matches included MexI-like WP_457985789.1, GolT-like WP_457986202.1, PbrA-like WP_457985413.1, and CzcA-like WP_373980100.1. Although they do not independently prove Pb or Au selectivity, the pbrA and golT matches show ATPase-family similarity. Similarly, further functional evaluation is needed for the MexI-like assignment. The remaining loci, which should not all be classified as direct metal-resistance determinants, comprised transport, control, DNA repair, and proteins linked to oxidative stress.

Table 3. Summary of BLASTN alignments retained after exploratory sequence identity and query coverage filtering, grouped by unique protein accession in Achromobacter strain H21.

Reference	H21 protein	Scaffold	Identity (%)	Coverage (%)	Alignment (bp)
arsB	WP_457988399.1	25	70.49	97	60,010–61,265
arsB	WP_457986626.1	9	70.40	96	106,756–107,990
arsH	WP_457986785.1	10	71.90	82	37,313–37,884
czcA	WP_373980100.1	20	70.56	97	93,664–96,771
dpsA	WP_457984776.1	3	75.45	80	12,801–13,189
golT	WP_457986202.1	7	71.38	82	102,853–104,724
hmrR	WP_457984846.1	3	71.28	84	96,096–96,468
irlR	WP_457984032.1	1	73.46	77	337,039–337,566
mexI	WP_457985789.1	5	74.27	99	324,422–327,480
modB	WP_457986023.1	6	70.84	73	215,286–215,794
pbrA	WP_457985413.1	4	72.34	73	265,235–266,990
pstB	WP_457987455.1	14	74.45	94	73,578–74,305
pstC	WP_457987457.1	14	70.53	78	75,338–76,084
ruvB	WP_457985600.1	5	72.58	86	109,266–110,175
sodB	WP_129239398.1	6	72.36	93	55,446–55,984

The protein-based and conserved-domain analyzes mentioned above are complementary to the BLASTN screen; however, since the query types and analytical methods are different, the findings should not be combined with previous TBLASTN counts. Instead of 15 confirmed metal-resistance genes, the 15 loci are screening hits. Candidate functions are still theoretical unless they are verified experimentally and at the domain level.

3.4. Genomic organization of the putative arsenic-resistance region

Genomic neighborhood analysis identified five closely arranged arsenic-associated genes on scaffold NZ_JBSYXR010000025.1, spanning approximately 60,003–64,458 bp. These encoded

an ArsB-like arsenite transporter (WP_457988399.1; 60,003–61,286 bp), ArsA arsenical pump-driving ATPase (WP_457988400.1; 61,339–63,129 bp), ArsD metallochaperone (WP_171662646.1; 63,140–63,499 bp), ArsC arsenate reductase (WP_008163574.1; 63,578–64,084 bp), and ArsR/SmtB-family regulator (WP_457988401.1; 64,081–64,458 bp).

All five genes occurred on the reverse strand. The intergenic gaps were 52 bp between the ArsB-like gene and arsA, 10 bp between arsA and arsD, and 78 bp between arsD and arsC; arsC and arsR overlapped by 4 bp. A coordinate-scaled R visualization illustrated the shared orientation and compact organization of the region (Figure 2). These observations are consistent with a potential arsenic-detoxification locus but do not demonstrate co-transcription or regulation.



Figure 3. Genomic organization of the putative arsenic-resistance region in Achromobacter strain H21. Arrows indicate the five annotated genes and their reverse-strand orientation on scaffold NZ_JBSYXR010000025.1. Gene functions are predicted from annotation and sequence-based evidence.

An additional arsenic-associated annotation (ACXIUT_RS26145) was located approximately 9 kb upstream of this region and was not included in the five-gene cluster. A separate ArsB-family candidate, WP_457986626.1, was located on scaffold NZ_JBSYXR010000009.1 (106,717–107,997 bp).

3.5. Sequence and conserved-domain characterization of the ArsR-family regulator

The 125-amino-acid protein WP_457988401.1 exhibited 99.20% amino-acid identity and 100% query coverage to an ArsR/SmtB-family regulator from Achromobacter deleyi (WP_171662377.1; E-value 9×10^{-84}). Conserved-domain analysis identified an ArsR-family DNA-binding domain (COG0640; CDD accession 440405) spanning residues 2–91, corresponding to 72.0% of the protein. Together, sequence similarity, conserved-domain annotation, and local genomic context support its predicted classification as an ArsR/SmtB-family regulator. Metal responsiveness,

DNA-binding targets, and transcriptional activity were not experimentally established (Supplementary Figure S1).

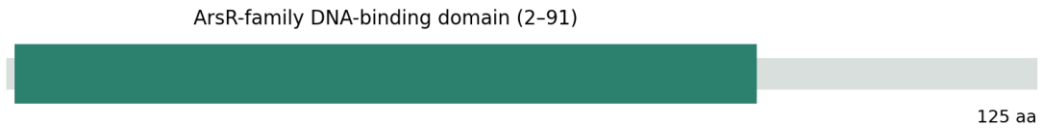


Figure 4. Schematic domain architecture of WP_457988401.1 showing the ArsR-family conserved domain (residues 2–91) within the 125-amino-acid protein.

3.6. Additional metal-associated transport and regulatory candidates

Beyond the arsenic-associated locus, H21 encoded metal-associated candidates including a CzcA-like RND transporter (WP_373980100.1), a PbrA-like ATPase (WP_457985413.1), and a GolT-like ATPase (WP_457986202.1). On scaffold NZ_JBSYXR010000020.1, the CzcA-like transporter (93,622–96,771 bp; reverse strand) was adjacent to a predicted RND adaptor (WP_373980099.1; 96,784–98,256 bp) and a TolC-family protein (WP_373980098.1; 98,269–99,528 bp), both on the reverse strand. A nearby heavy-metal response regulator (WP_373980097.1; 99,889–100,578 bp) and sensor histidine kinase (WP_421601212.1; 100,578–101,939 bp) were oriented on the forward strand. Conserved-domain results supported the broad transporter, adaptor, outer-membrane protein, and regulatory protein classifications, but direct regulation and metal-substrate specificity were not established. A separate copper-associated region on scaffold NZ_JBSYXR010000001.1 included a copper resistance protein B, multicopper oxidase, cupredoxin-domain protein, sensor histidine kinase, and response regulator (WP_457984028.1–WP_457984032.1; 332,335–337,698 bp). These genes were not co-localized with the CzcA-like efflux region.

3.7. Genomic organization and conserved-domain analysis of putative chromate transporters

Three putative chromate-transporter proteins were identified across two genomic regions of Achromobacter strain H21. On scaffold NZ_JBSYXR010000018.1, WP_457987913.1 (174 amino acids; 67,620–68,144 bp) and WP_457987914.1 (191 amino acids; 68,141–68,716 bp) were encoded by adjacent reverse-strand genes with a four-base-pair overlap. The surrounding neighborhood contained two universal stress proteins (WP_457987917.1 and WP_457987918.1),

a sensor histidine kinase (WP_457987919.1), and a response regulator (WP_457987920.1). The presence of these neighboring genes does not establish chromate-specific regulation.

A third gene, annotated as chrA and encoding WP_457988352.1 (418 amino acids), occurred separately on scaffold NZ_JBSYXR010000025.1 at 9,075–10,331 bp on the forward strand. Nearby annotations included an FMN-dependent NADH-azoreductase, GNAT-family N-acetyltransferase, enoyl-CoA hydratase, multidrug MFS transporter, and other transport/metabolic proteins. The two genomic neighborhoods are shown in Figure 3.

NCBI Batch CD-Search detected a specific ChrA-family domain (COG2059; residues 1–171; E = 4.66×10^{-41}) in WP_457987913.1 and a specific chromate-transporter domain (pfam02417; residues 21–177; E = 3.24×10^{-31}) in WP_457987914.1; both were assigned to superfamily cl17781. WP_457988352.1 matched the 2A51 transporter superfamily (cl36776; residues 35–404; E = 8.02×10^{-67}), a broader assignment that does not independently establish chromate specificity (Table 3).

Table 4. Conserved-domain assignments for three putative chromate-transporter proteins in Achromobacter strain H21.

Protein	Length (aa)	Domain	Residues	E-value
WP_457987913.1	174	COG2059 (ChrA)	1–171	4.66×10^{-41}
WP_457987914.1	191	pfam02417 (Chromate_transp)	21–177	3.24×10^{-31}
WP_457988352.1	418	cl36776 (2A51 superfamily)	35–404	8.02×10^{-67}

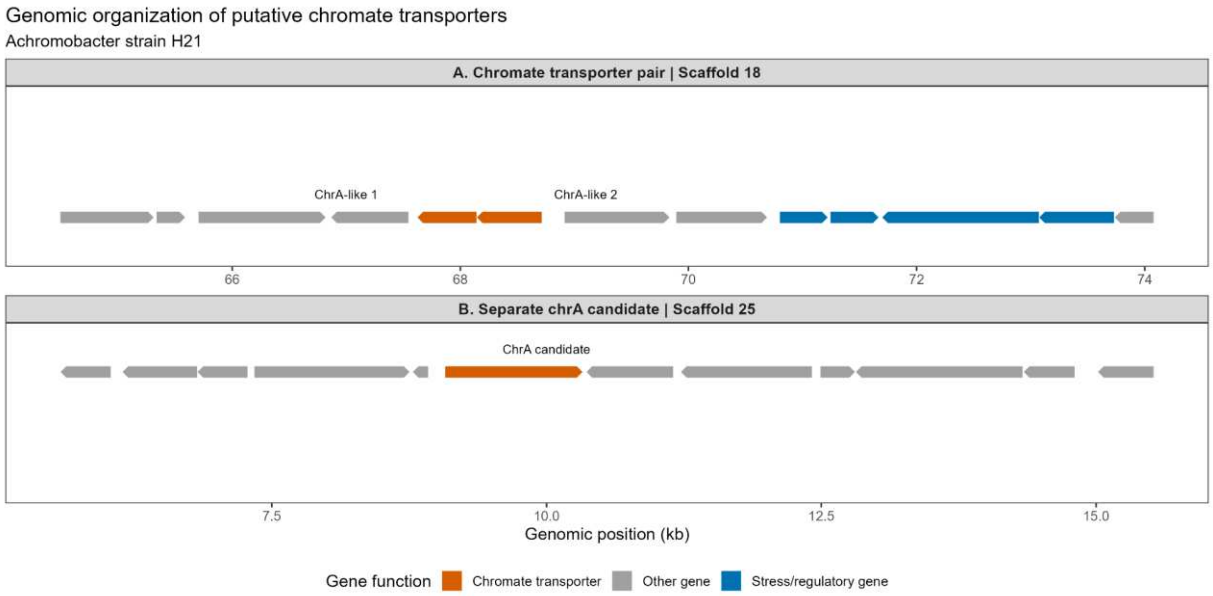


Figure 5. Genomic organization of putative chromate-transporter genes in Achromobacter strain H21. (A) Adjacent reverse-strand genes encoding WP_457987913.1 and WP_457987914.1 on scaffold NZ_JBSYXR010000018.1. (B) Separate forward-strand chrA-annotated candidate WP_457988352.1 on scaffold NZ_JBSYXR010000025.1. Orange arrows indicate predicted chromate transporters; blue arrows indicate neighboring stress-associated or regulatory proteins; grey arrows represent other coding sequences. Gene functions are computational predictions.

3.8. Genomic organization and conserved-domain characterization of a putative ZntA–CadR/PbrR metal-resistance locus

A putative metal-resistance-associated locus was identified on scaffold NZ_JBSYXR010000004.1. The region contained a heavy-metal-translocating P-type ATPase (WP_457985413.1; 806 amino acids), encoded at positions 265,211–267,631 bp on the reverse strand, and a neighboring Cd(II)/Pb(II)-responsive transcriptional regulator (WP_457985414.1; 144 amino acids), encoded at positions 267,731–268,165 bp on the forward strand. The genes are divergently oriented and separated by a 99-bp intergenic interval (Figure 4).

NCBI conserved-domain analysis identified a specific ZntA-family domain (COG2217; residues 105–804; E-value reported as 0) in the ATPase and a specific CadR/PbrR-family helix–turn–helix domain (cd04784; residues 1–128; $E = 1.68 \times 10^{-64}$) in the adjacent regulator (Table 4). The corresponding domain spans represent approximately 86.8% and 88.9% of the respective protein lengths. An earlier sequence-similarity comparison of the ATPase to a PbrA reference yielded 64.94% amino-acid identity and 86% reference coverage. These findings support the identification

of a candidate divalent-metal-associated transporter–regulator locus; they do not independently establish Pb(II) specificity, transcriptional regulation, or phenotypic resistance.

Table 5. Conserved-domain assignments for the ZntA-like ATPase and adjacent CadR/PbrR-family regulator in Achromobacter strain H21.

Protein accession	Protein length (aa)	Specific domain	Domain span (aa)	E-value
WP_457985413.1	806	COG2217 (ZntA)	105–804	0 (reported)
WP_457985414.1	144	cd04784 (HTH_CadR-PbrR)	1–128	1.68×10^{-64}

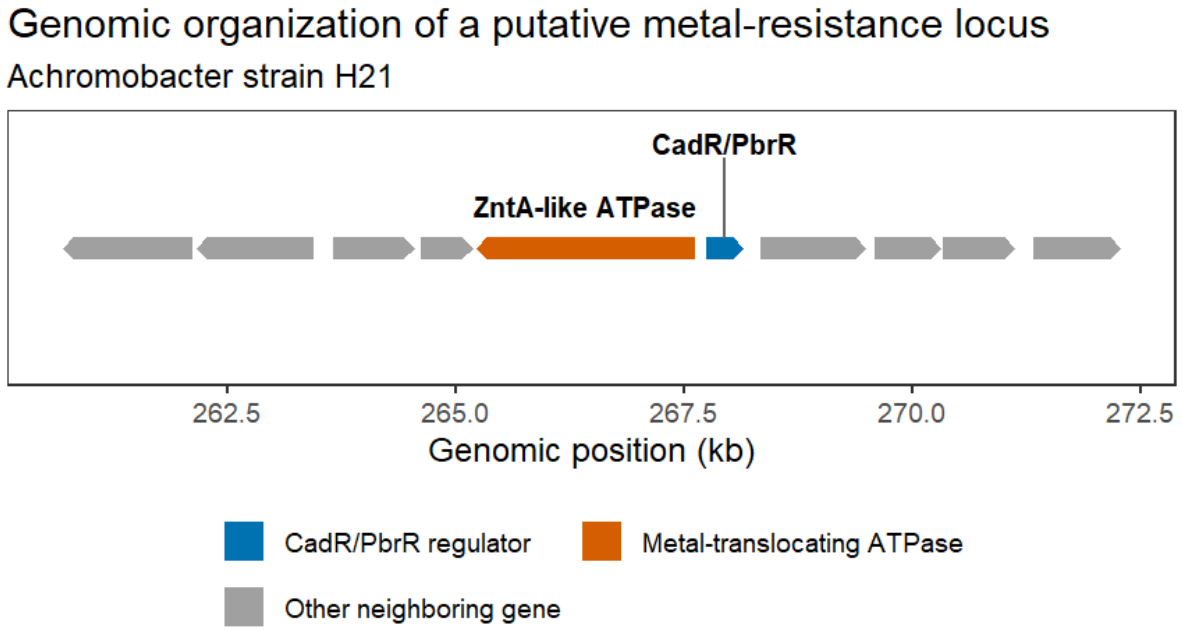


Figure 6 Genomic organization of a putative metal-resistance-associated locus in Achromobacter strain H21. The region on scaffold NZ_JBSYXR010000004.1 contains a ZntA-family heavy-metal-translocating P-type ATPase (WP_457985413.1) and an adjacent CadR/PbrR-family regulator (WP_457985414.1), separated by a 99-bp intergenic region. The orange and blue arrows indicate the ATPase and regulator, respectively; grey arrows indicate neighboring genes. Arrow directions show predicted transcriptional orientations. Functional assignments are computational predictions.

3.9. Genomic organization and conserved-domain characterization of a GolT-like metal-translocating ATPase

A heavy-metal-translocating P-type ATPase candidate (WP_457986202.1; locus tag ACXIUT_RS12915) was identified on scaffold NZ_JBSYXR010000007.1 at positions 102,262–104,739 bp on the forward strand. The predicted protein comprised 825 amino acids and previously showed approximately 71.39% sequence identity to a GolT reference protein. Genomic-neighborhood analysis identified an upstream putative transporter small subunit (WP_457986201.1; 101,972–102,094 bp), a more distant LysR-family regulator (WP_457986198.1; 97,566–98,438 bp), and neighboring metabolic and stress-associated genes. The transporter small subunit and ATPase were separated by 167 bp; their functional association has not been established.

NCBI conserved-domain analysis identified a specific copper-type P-type ATPase domain (cd02094; residues 166–817; E-value reported as 0), along with two specific N-terminal CopZ-like metal-binding domain matches (COG2608; residues 78–144, $E = 3.45 \times 10^{-20}$; and residues 11–56, $E = 2.05 \times 10^{-14}$). These domain assignments support classification as a GolT-like, copper-type metal-translocating ATPase candidate but do not establish gold transport, metal substrate specificity, or a shared regulatory system with neighboring genes (Table 5; Figure 5).

Table 6. Conserved-domain assignments for the GolT-like ATPase WP_457986202.1 (825 amino acids).

Domain accession	Domain designation	Residues	E-value
cd02094	P-type_ATPase_Cu-like	166–817	0 (reported)
COG2608	CopZ-like metal-binding	78–144	3.45×10^{-20}
COG2608	CopZ-like metal-binding	11–56	2.05×10^{-14}

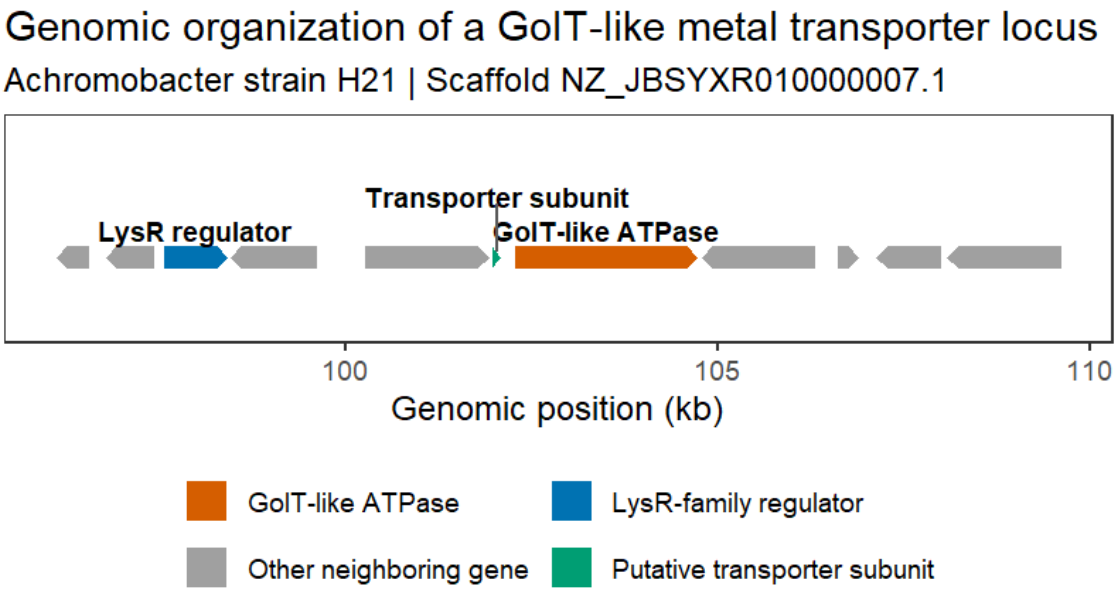


Figure 7. Genomic organization of a GolT-like metal-translocating ATPase locus in Achromobacter strain H21. Scaffold NZ_JBSYXR010000007.1 contains a GolT-like ATPase (WP_457986202.1), an upstream putative transporter small subunit (WP_457986201.1), and a LysR-family transcriptional regulator (WP_457986198.1). Orange, green, and blue arrows denote these features, respectively; grey arrows represent other neighboring genes. Arrowheads indicate annotated transcriptional orientation. Conserved-domain analysis of the ATPase identified cd02094 and two N-terminal COG2608 matches; functional relationships among the neighboring genes remain unverified.

3.10. Genomic organization of a second arsenic-associated region

A second arsenic-associated region was identified on scaffold NZ_JBSYXR010000010.1 at positions 33,139–37,929 bp. Seven adjacent genes, all oriented on the forward strand, encoded an ArsR/SmtB-family regulator (WP_457986779.1), an ArsI/CadI-like metalloenzyme annotation (WP_457986780.1), an ArsC-family protein (WP_457986781.1), an ACR3-family transporter (WP_457986782.1), an MFS transporter (WP_457986783.1), a reductase-like protein (WP_457986784.1), and an ArsH-family protein (WP_457986785.1).

Specific conserved-domain matches supported the ArsR (COG0640; $E = 5.49 \times 10^{-21}$), ArsC (cd16345; $E = 3.71 \times 10^{-69}$), ACR3 (COG0798; $E = 8.36 \times 10^{-144}$), and ArsH (TIGR02690; $E = 9.34 \times 10^{-132}$) assignments. In contrast, the ArsI/CadI-like annotation matched the broad VOC

superfamily (cl14632), the MFS transporter matched a broad MFS superfamily (cl28910), and the reductase-like protein matched a thioredoxin-like superfamily (cl00388). ArsI activity, arsenic transport via the MFS protein, or arsenate reductase activity by the reductase-like protein are not established by these more general matches. A single operon is not independently shown by the tight spacing and common orientation.

3.11. Extended CzcABC-associated genomic region

The CzcABC-associated region on scaffold NZ_JBSYXR010000020.1 extended from positions 93,622 to 109,570 bp. It contained the CzcA-like RND transporter WP_373980100.1 (93,622–96,771 bp, negative strand), CzcB-like membrane-fusion component WP_373980099.1 (96,784–98,256 bp, negative strand), and CzcC-like outer-membrane channel WP_373980098.1 (98,269–99,528 bp, negative strand). A response regulator (WP_373980097.1) and sensor kinase (WP_421601212.1) occurred immediately downstream on the positive strand. Domain analysis identified an ACR_tran superfamily match for CzcA (cl47191; E-value reported as 0), a cl36796 match for CzcB, and a TolC-family match (COG1538) for CzcC.

The extended interval also contained an IS3-family transposase (WP_421601239.1) represented by two CDS records associated with a programmed frameshift, an additional response regulator (WP_373980095.1), a porin (WP_373980094.1), three DUF4148-family proteins, and a heavy-metal-associated P-type ATPase (WP_373980090.1; 107,615–109,570 bp, positive strand). The ATPase matched the specific ZosA/PfeT-like domain cd07551 (residues 15–646; E-value reported as 0). Its substrate specificity remains unresolved. The presence of these genes indicates a broader metal-associated neighborhood but does not demonstrate that all components are co-regulated or participate in a single efflux system.

3.12. RcnA-associated metal transport region and CorA-like transporter

An RcnA-associated transport region was identified on scaffold NZ_JBSYXR010000002.1 at positions 455,208–459,488 bp. Five closely spaced genes on the negative strand encoded an RcnA-family nickel/cobalt-associated transporter (WP_457984698.1), a DUF1007-family protein (WP_457984699.1), a TroA-like metal-binding component (WP_457984700.1), a ZnuB-like permease (WP_457984701.1), and a ZnuC-like ATP-binding component (WP_457984702.1). Conserved-domain searches identified COG2215 (RcnA; residues 66–276; $E = 6.29 \times 10^{-59}$), cl00262 (TroA-like; residues 29–336; $E = 1.06 \times 10^{-116}$), COG1108 (ZnuB; residues 14–275; E =

5.55 × 10^{-54}), and COG1121 (ZnuC; residues 10–244; E = 6.49 × 10^{-100}). The RcnA domain alignment was marked C-terminal incomplete by CDD; this does not alone establish protein truncation. These results support the presence of an RcnA-family transporter adjacent to a putative metal-ion ABC transport system, but do not establish the metal specificity or transport direction of the ABC system, co-transcription, or physiological resistance.

A separate CorA-like transporter (WP_457986084.1; 286,717–287,718 bp, negative strand) was identified on scaffold NZ_JBSYXR010000006.1. It exhibited a specific MtCorA-like domain match (cd12830; residues 43–333; E = 2.68 × 10^{-129}). Its neighboring region included three ABC transporter-associated proteins (WP_457986080.1–WP_457986082.1), a thermonuclease-family protein, an ATP-dependent helicase, a DUF72 protein, and a PdxR-family regulator. The CorA-like assignment supports a divalent-cation transporter-family classification but does not specifically demonstrate nickel or cobalt resistance.

3.13. Consolidated conserved-domain evidence for metal-associated candidates

Table 7 summarizes selected conserved-domain assignments supporting the major metal-associated candidates. Specific domain matches and broader superfamily matches are distinguished to avoid overinterpreting predicted substrate specificity. Proteins with only annotation-based evidence remain described in the corresponding genomic-neighborhood subsections.

Table 7. Conserved-domain assignments of selected putative metal-resistance and metal-homeostasis proteins in Achromobacter strain H21.

System	Protein accession	CDD assignment	Hit type	E-value	Putative function
Arsenic	WP_457986779.1	COG0640 (ArsR)	Specific	5.49 × 10^{-21}	Regulator candidate
Arsenic	WP_457986781.1	cd16345 (ArsC)	Specific	3.71 × 10^{-69}	Arsenate reductase family
Arsenic	WP_457986782.1	COG0798 (ACR3)	Specific	8.36 × 10^{-144}	Arsenite efflux family
Arsenic	WP_457986785.1	TIGR02690 (ArsH)	Specific	9.34 × 10^{-132}	ArsH family

Chromate	WP_457987913.1	COG2059 (ChrA)	Specific	4.66×10^{-41}	Chromate transporter family
Chromate	WP_457987914.1	pfam02417	Specific	See Table 3	Chromate transporter family
Czc	WP_373980100.1	cl47191 (ACR_tran)	Superfamily	0	RND transporter
ATPase	WP_457985413.1	COG2217 (ZntA)	Specific	0	Metal P-type ATPase
Regulator	WP_457985414.1	cd04784 (CadR/PbrR)	Specific	1.68×10^{-64}	HTH regulator
ATPase	WP_457986202.1	cd02094 (Cu-type)	Specific	0	Cu-type P-type ATPase
Czc region	WP_373980090.1	cd07551 (ZosA/PfeT-like)	Specific	0	Metal P-type ATPase
Ni/Co	WP_457984698.1	COG2215 (RcnA)	Specific	6.29×10^{-59}	Ni/Co-associated transporter
Metal ABC	WP_457984700.1	cl00262 (TroA-like)	Superfamily	1.06×10^{-116}	Metal-binding component
Metal ABC	WP_457984701.1	COG1108 (ZnuB)	Specific	5.55×10^{-54}	ABC permease
Metal ABC	WP_457984702.1	COG1121 (ZnuC)	Specific	6.49×10^{-100}	ABC ATPase
CorA	WP_457986084.1	cd12830 (MtCorA-like)	Specific	2.68×10^{-129}	Divalent-cation transporter

3.14. Preliminary antimicrobial-resistance screening

CARD/RGI screening yielded inconsistent results between analyses. One protein-based run returned no Perfect, Strict, or Loose matches, while a separate genome-based run reported three

Strict hits, comprising two adeF-associated entries and one qacG-associated entry. The adeF matches were associated with RND-family antimicrobial efflux categories, and qacG with a small multidrug-resistance-family efflux category. Because the genome-based input filename and analysis settings require verification, these findings **are** preliminary and are not interpreted as evidence of phenotypic antimicrobial resistance.

3.15. Taxonomic assessment of strain H21

Available average nucleotide identity comparisons returned approximately 90.31% identity to the examined Achromobacter denitrificans reference assembly and 90.73% to the examined A. veterisilvae reference assembly, with limited alignment coverage. These results do not support a confident species-level assignment. Accordingly, the organism is referred to here as Achromobacter sp. H21 pending further genome-based taxonomic analysis.

3.17. Batch conserved-domain validation of the 15 BLASTN-mapped proteins

All 15 distinct H21 proteins retained after BLASTN screening and coding-sequence mapping were submitted to NCBI Batch CD-Search (search ID QM3-qcdsearch-1C6109358FF7417F). The search reported 15 valid queries, no failed queries, no queries without domain hits, 727 total domain hits and 427 clusters, using an E-value cutoff of 0.01, composition-corrected scoring, and no low-complexity filtering. The uploaded alignment export was labelled “Full Results (Superfamily only)”; consequently, the domain-family assignments below are interpreted alongside the specific-hit information provided separately in the NCBI search summary. The 727 hits represent overlapping domain-model matches, not 727 different domains or genes.

Table 8. Conserved-domain family and superfamily assignments of 15 distinct Achromobacter strain H21 proteins identified through BLASTN analysis and coding sequence (CDS) mapping.

H21 protein	Observed domain families	Cautious functional interpretation
WP_457988399.1	ArsB_NhaD_permease / ArsB	ArsB-like arsenite transport candidate
WP_457986626.1	ArsB_NhaD_permease / ArsB	Second ArsB-like candidate
WP_457986785.1	FMN_red; specific TIGR02690 resist_ArsH (search summary)	ArsH-family flavoprotein candidate

WP_373980100.1	ACR_tran; CusA/AcrB	RND efflux transporter, CzcA-like
WP_457984776.1	Dps; ferritin-like	DNA protection / oxidative-stress response
WP_457986202.1	P-type ATPase, Cu-like; related ZntA/CopA models	Cu-type metal-transporting ATPase; substrate uncertain
WP_457984846.1	HTH_MerR-SF / CueR	Metal-responsive transcriptional regulator
WP_457984032.1	cztR_silR_copR / OmpR / REC	Two-component response regulator
WP_457985789.1	ACR_tran; AcrB/CusA	RND efflux transporter; metal specificity unconfirmed
WP_457986023.1	ModB; TM_PBP2	Molybdate-family ABC permease
WP_457985413.1	ZntA; ATPase-IB_hvy / ATPase-IB2_Cd	ZntA-like divalent-metal P-type ATPase candidate
WP_457987455.1	PstB; P-loop NTPase	Phosphate transporter ATP-binding component
WP_457987457.1	PstC / PstA; TM_PBP2	Phosphate transporter permease
WP_457985600.1	RuvB; AAA / P-loop NTPase	DNA recombination/repair ATPase
WP_129239398.1	SodA / Sod_Fe_C / Sod_Fe_N	Superoxide dismutase-family oxidative-stress protein

The two ArsB-like candidates shared ArsB_NhaD_permease/ArsB family assignments, supporting two separate genomic loci. WP_457986785.1 matched the ArsH-specific TIGR02690 domain in the NCBI sample output. The CzcA-like and MexI-like proteins both matched the broad ACR_tran RND-efflux superfamily; these results support transporter-family classification but do not by themselves establish the transport substrate. WP_457985413.1 matched ZntA-related P-type ATPase superfamilies, consistent with the previously observed ZntA-like annotation, while WP_457986202.1 was assigned a Cu-like P-type ATPase-specific domain in the NCBI sample output. Functional assignment of individual metal substrates requires additional sequence-level or experimental evidence.

Other mapped proteins had domain families consistent with metal-responsive regulation (CueR and cztR/silR/copR-like regulators), phosphate or molybdate transport (PstB, PstC and ModB), DNA repair (RuvB), and oxidative-stress protection (Dps and superoxide dismutase). These should not be counted as experimentally verified metal-resistance genes. No functional assays, gene expression measurements, or metal-resistance phenotypes were performed in this analysis.

Source: NCBI Batch CD-Search job QM3-qcdsearch-1C6109358FF7417F; uploaded "aligndata.txt" alignment export and NCBI on-screen search statistics and sample hits.

3.18. Genomic neighbourhood of the second ArsB-like transporter

The second ArsB-like candidate, WP_457986626.1 (ACXIUT_RS15630), was located on scaffold NZ_JBSYXR010000009.1 at 106,717–107,997 bp on the positive strand. Examination of coding sequences within 10 kb on either side identified a transporter-rich neighbourhood. The immediately upstream genes encoded a DMT-family transporter (WP_457986624.1; 104,841–105,734 bp) and an AEC-family transporter (WP_457986625.1; 105,762–106,706 bp), while the downstream gene encoded a YeeE/YedE-family protein (WP_457986627.1; 108,151–109,377 bp). All four transport-associated genes were co-oriented on the positive strand. The gaps between successive coding sequences were 27, 10, and 153 bp, respectively. A GNAT-family N-acetyltransferase (WP_457986623.1; 104,280–104,723 bp) was located upstream on the negative strand.

The ArsB-family assignment and arsenic-transporter annotation support WP_457986626.1 as a candidate arsenite transport protein. No genes annotated as canonical arsR, arsC, arsA or arsD were found in the examined 10-kb flanking region. Although the neighbouring transport-related genes may indicate local functional organization, neither their arsenic specificity nor co-transcription with the ArsB-like gene can be inferred from genomic proximity alone.

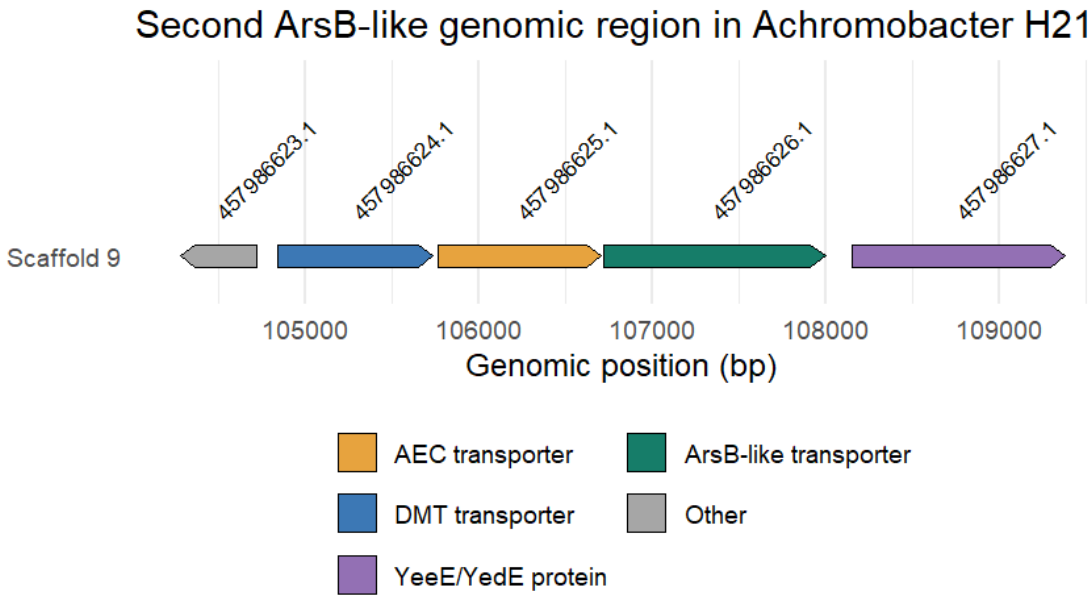


Figure 8. Genomic neighbourhood of the second ArsB-like transporter on scaffold NZ_JBSYXR010000009.1. Arrows show CDS positions and orientations; the highlighted ArsB-like candidate is flanked by DMT-, AEC- and YeeE/YedE-family proteins. Protein identifiers correspond to RefSeq accessions.

3.19. Two adjacent RND efflux-associated modules surrounding a MexI-like transporter

The MexW/MexI-family RND transporter WP_457985789.1 (ACXIUT_RS10815) was located on scaffold NZ_JBSYXR010000005.1 at 324,422–327,526 bp on the positive strand. Two adjacent coding sequences encoded an RND periplasmic adaptor (WP_457985788.1; 323,275–324,405 bp) and an efflux outer-membrane subunit (WP_457985790.1; 327,539–328,990 bp). These three genes were co-oriented on the positive strand and separated by 16 and 12 bp, respectively, consistent with the structural components of a putative tripartite RND efflux system.

An additional, oppositely oriented RND-associated module was identified upstream. It comprised an outer-membrane efflux subunit (WP_457985823.1; 314,075–315,529 bp), an RND permease (WP_457985783.1; 315,532–318,699 bp), and a periplasmic adaptor (WP_457985784.1; 318,696–319,883 bp), all on the negative strand. The outer-membrane and permease CDSs were separated by 2 bp, whereas the permease and adaptor CDSs overlapped by 4 bp. The two modules were separated by several other genes, including predicted regulators and transport-associated proteins. Their architecture supports two putative RND efflux systems but does not establish their substrates or confirm an operon structure.

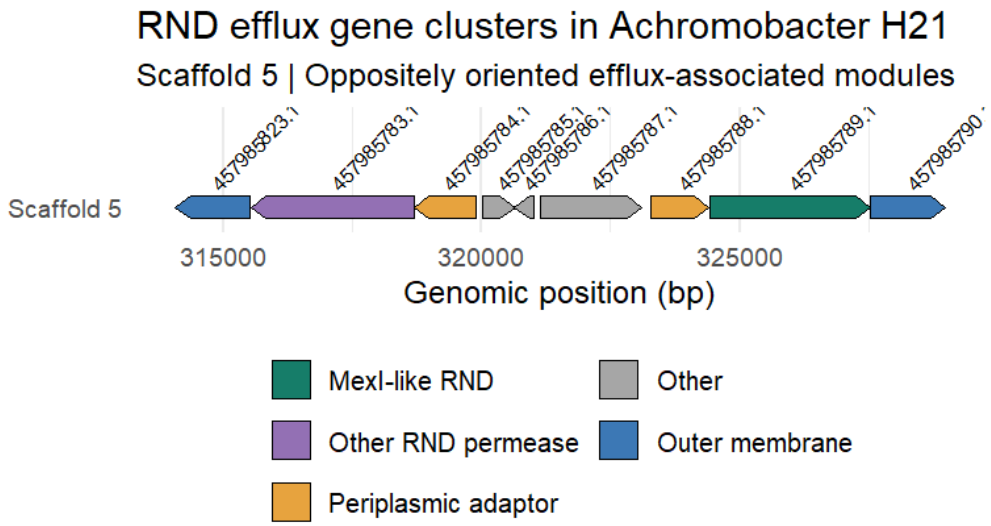


Figure 9. Genomic organization of two oppositely oriented RND efflux-associated modules on scaffold NZ_JBSYXR010000005.1. The negative-strand module contains an outer-membrane subunit, RND permease and periplasmic adaptor; the positive-strand module contains the corresponding components, including the MexW/MexI-like transporter WP_457985789.1. Gene orientation and coordinates are derived from RefSeq GFF annotations.

Interpretive synthesis. Together, these neighbourhood analyses provide additional genomic-context support for a second ArsB-like transport locus and two adjacent RND efflux-associated modules. The arsenic-transporter annotation is more directly connected to a specific metalloid-associated function, whereas the MexI-like RND assignment indicates efflux potential without establishing heavy-metal specificity. Functional assays, substrate-specific transport measurements or transcriptomic data are needed to test the inferred roles and co-regulation of neighbouring genes.

4. DISCUSSION

4.1. Genome assembly characteristics and functional annotation

The draft genome of Achromobacter strain H21, comprising 6,184,534 bp across 65 scaffolds and containing 5,851 predicted proteins, provides a substantial sequence resource for exploring putative metal-associated functions. The reported GC content of 66.12% and annotation of 5,915 gene features and 5,940 CDS features describe the assembly and annotation framework used for subsequent screening. It is not appropriate to interpret the various gene counts, CDS characteristics, and protein data as equivalent indicators of gene content. In instance, the quantity or functional activity of metal-resistance systems cannot be determined only by the number of annotated characteristics. (Schnoes et al., 2009; Jeukens et al., 2017).

The reported completeness estimate of 94.02% suggests broad, but incomplete, recovery of the genome, while the reported contamination estimate of 4.71% warrants caution when assigning individual loci to the organism. Fragmentation into multiple scaffolds can also interrupt gene clusters and obscure longer-range associations. Therefore, it is possible to interpret genomic neighborhoods found inside a scaffold as local arrangements, but the apparent lack of further route genes or regulators must be regarded as tentative. The interpretation of distributed arsenic-associated loci and the taxonomic confusion that follows are particularly pertinent to these quality issues.

When understanding this data, assembly quality and annotation provenance are important considerations. While Cornet et al. (2022) outline repeatable microbial-genomics methods, Cornet and Baurain (2022) address the difficulties of identifying contamination in genome datasets. Farooq and Rafique (2025) also examine composite measures of bacterial genome quality. These publications provide methodological context rather than independent confirmation of the quality of the H21 assembly. Amoutzias et al. (2022) place bacterial genomics in a wider analytical context, and Venter et al. (2011) illustrate how proteogenomic evidence can refine genome annotation; neither study experimentally verifies the H21 gene calls.

4.2. Annotation-based identification of metal-associated candidates

The initial annotation search revealed many predicted proteins associated with cellular stress responses, transcriptional regulation, metal sensing, and membrane transport. P-type ATPases, chromate-transporter candidates, RND transporters, Ars-family proteins, and other cation-transport families are all found together, suggesting that H21 encodes a variety of pathways that might be crucial for metal homeostasis. Some of these mechanisms may indirectly protect cells from secondary stress, others may export toxic ions, and still others may maintain essential metal homeostasis. These mechanisms do not, however, perform the same functions. (Ayangbenro & Babalola, 2020; Kommu et al., 2024).

This distinction limits the interpretation of keyword-based annotation. Proteins from CorA, ZIP, Nramp, ABC, and related families may have physiological roles in ion homeostasis or nutrient acquisition even when their homologues are present in metal-stress situations. Therefore, the annotation screen served to generate hypotheses rather than provide a simple inventory of confirmed resistance genes. Additional homology, domain, and neighborhood studies allow for the evaluation of these broad candidate assignments at different levels of accuracy.

Transport annotations should be considered in addition to overall metal homeostasis. Jiménez (2009) reviews transition-metal homeostasis; Hogle et al. (2016) investigate trace-metal acquisition processes in marine heterotrophic bacteria; and Fichant et al. (2006) describe an ABC transporter repertoire identification resource. These sources support reading transporter annotations as functional hypotheses rather than automatically classifying all transport-related proteins as resistance determinants.

4.3. Protein-based sequence-similarity evidence

Thirteen of the 90 local alignments obtained from the TBLASTN comparison of 16 chosen H21 proteins with the nucleotide reference collection met the specified E-value criteria of 10^{-1}. CzcA-like WP_373980100.1, ArsB-like WP_457988399.1, ArsA-like WP_457988400.1, and the CzcR-like regulator WP_457984032.1 were the strongest chosen matches. Homology to the related protein families is supported by high amino acid identity and wide reference coverage, especially when the same assignments are in line with local gene organization and conserved-domain evidence. (Schnoes and others, 2009).

However, reference coverage should not be mistaken for query coverage, and local alignment counts should not be equated with the number of different genes. The lack of qualifying matches for the three proteins that were chosen for this search does not mean that there is no biological function; rather, it suggests that the comparison was not strong enough. Similarly, the similarities between PbrA and GolT do not prove the transported metal, but they do support the ATPase-family relatedness. In order to prevent a comparative sequence result from becoming a functional claim, several qualifiers are required.

These assignments may be framed by comparative genomic analysis without developing strain-specific characteristics. While Lorková et al. (2025) looked at environmental effects on genomic characteristics in potential new Streptomyces species, Sorouri et al. (2023) studied variance in bacterial properties across environments and phylogenetic clades. Their significance in this case stems from comparative methods rather than proof that H21 picked up specific loci via environmental selection.

4.4. Nucleotide-reference BLASTN screening and CDS mapping

By matching 578 nucleotide reference sequences against the H21 assembly, the complementary BLASTN analysis expanded the screening beyond the chosen protein queries. Of 349 local alignments from 133 reference queries, exploratory filtering for at least 70% identity, at least 70% per-HSP query coverage and an E-value no greater than 10^{-5} retained 17 alignments from 16 reference queries. Mapping the retained coordinates to annotated CDS features consolidated these matches into 15 distinct protein accessions. This reduction illustrates why alignment counts, reference-query counts and unique genomic loci must be reported separately. (Schnoes et al., 2009).

A number of references, such as paired czcA references mapping to WP_373980100.1 and paired ruvB references mapping to WP_457985600.1, converged on the same identified protein. The evidence for sequence homology may be strengthened by such convergence, but the number of genomic genes is not increased. DNA repair proteins, oxidative-stress proteins, and phosphate and molybdate transport components are also included in the final collection. Their inclusion in the filtered hits supports the idea that the 15 proteins should be seen as a diverse candidate collection rather than as 15 metal-resistance determinants that have been verified via experimentation. The raw numbers of BLASTN and TBLASTN should stay apart as they have different query types and alignment behaviors.

4.5. Compact arsenic-associated gene cluster

One of the most obvious instances of pathway-like genomic structure in H21 is represented by the five closely spaced genes on scaffold NZ_JBSYXR010000025.1. The area is oriented on the reverse strand and contains an ArsB-like transporter, an ArsA-like ATPase, an ArsD-family metallochaperone, an ArsC-like reductase, and an ArsR/SmtB-family regulator. Their tight structure, which includes an overlap between arsC and arsR and small intergenic gaps, is consistent with the coordinated organization of functions related to arsenic. At the functional-model level, an ArsC-like enzyme could contribute to arsenate reduction, while ArsB/ArsA-family components could be associated with arsenite export and ArsR/ArsD-family proteins with regulation or intracellular handling. (Li et al., 2012; Jeukens et al., 2017).

Nevertheless, this mechanistic model remains an inference from gene-family identities and local context. Shared orientation and proximity do not establish a single operon, co-transcription, enzyme activity or arsenic-dependent induction. The additional arsenic-associated annotation located outside the five-gene region also suggests that not all potentially relevant determinants need to reside in the same cluster. The combination of complementary predicted components makes this locus a priority for experimental follow-up.

The arsenic-related interpretation should remain distinct from evidence obtained in other organisms and biological systems. Potdukhe et al. (2018) examined transcripts associated with arsenic accumulation in the plant Pteris vittata; that work offers a contrasting arsenic-response model, but it cannot be used to infer bacterial ArsB activity in H21.

4.6. ArsR/SmtB-family regulator characterization

WP_457988401.1's designation as a potential transcriptional regulator is supported by its high sequence similarity to a reference ArsR/SmtB-family protein, an ArsR-family DNA-binding domain, and closeness to the arsenic-associated cluster. In addition to the anticipated transport and detoxifying genes, this offers a tenable regulatory component. The 125-amino-acid protein's conserved domain makes up a significant chunk of it, supporting the family-level assignment. (Brown et al., 2003; Jung & Lee, 2019).

It is unclear what metal signal WP_457988401.1 senses, what promoter sequences it recognizes, and how it affects neighboring genes. To prove arsenic-specific repression or derepression, an ArsR/SmtB-family classification is insufficient. To test the suggested regulatory function, promoter analysis, transcriptional profiling under arsenic exposure, and DNA-binding experiments would be required.

Similarly, regulatory comparison calls for care. While Wu and Outten (2011) investigated Fur- and YqjI-dependent regulation of a ferric reductase gene in Escherichia coli, Prabhakaran et al. (2018) described a cadmium-associated cadR promoter and gene in Pseudomonas aeruginosa. These investigations show different regulatory topologies, but neither one shows the H21 ArsR/SmtB candidate's ligand specificity or promoter targets. The cyanobacterial regulator RexT's redox sensing was studied by Li et al. (2022), highlighting the possibility that even well-known transcription-factor families might react to cues other than metals.

4.7. Additional transport and regulatory candidates

Beyond the major arsenic locus, CzcA-like, PbrA-like, and GolT-like candidates have been identified, suggesting that H21 may encode various transport topologies instead of depending on a single metal-associated route. Because neighboring annotations for the CzcA-like protein contain an RND adaptor and a TolC-family outer-membrane protein—a combination consistent with a tripartite efflux assembly—it is especially instructive. Although nearby response-regulator and sensor-kinase genes provide possible regulatory context, their physical closeness and divergent orientation do not demonstrate direct regulation of the efflux genes. (Riccardi et al., 2008; Blencowe & Morby, 2003).

A different local functional theme is suggested by the unique copper-associated neighborhood, which includes a copper resistance protein, multicopper oxidase, cupredoxin-domain protein, and regulatory elements. Annotation alone cannot demonstrate a common regulatory circuit or copper detoxification, but such genomic arrangement may represent specific metal-handling systems. When taken as a whole, these areas provide focused comparisons of metal-dependent expression, gene organization, and transporter-family domains.

Complex genomic environments may also include metal-associated loci. A precedent for studying mobility-associated neighborhoods was established by Adhikarla et al. (2009), who revealed a metal-resistance determinant in Pseudomonas putida bordered by metal-inducible mobile elements. Chung et al. (2022) looked at possible mechanisms linked to gallium bioleaching in bacterial isolates using genome mining. In H21, neither sample shows movement or gallium transfer.

4.8. Chromate-transporter candidates and genomic context

On scaffold NZ_JBSYXR010000018.1, the two neighboring short proteins WP_457987913.1 and WP_457987914.1 contain particular ChrA/chromate-transporter conserved-domain support, which makes their family assignment more targeted than a broad membrane-transporter annotation. Close local organization is suggested by their four-base-pair coding sequence overlap and reverse-strand orientation. Although they don't prove chromate-specific activation or direct interaction with the transporter genes, nearby universal stress proteins and a sensor–response regulator pair provide a more comprehensive genomic context linked to stress. (Schnoes and others, 2009).

The two short proteins and the distinct 418-amino-acid chrA-annotated candidate WP_457988352.1 have different genomic settings and lengths. Rather than presuming that all three choices reflect the same transport systems, the interpretation should take into account any differences in domain-level specificity. To verify these theories, comparative phylogenetic research and chromium-exposure tests would be necessary. However, their location across two areas raises questions regarding different evolutionary origins or physiological functions.

4.9. ZntA-like ATPase and adjacent CadR/PbrR-family regulator

The ZntA-like ATPase WP_457985413.1, together with the adjacent CadR/PbrR-family regulator, constitutes a credible locus for metal-related transport and control. Conserved-domain classifications bolster the ATPase family and regulatory protein categorization, offering evidence that extends beyond the initial annotation. The juxtaposition of a membrane transport ATPase with a putative metal-responsive regulator is biologically coherent and identifies a useful candidate region for further investigation. (Brown et al., 2003; Jung & Lee, 2019).

Even so, P-type ATPases can transport different metal ions, and family-level similarity to PbrA or ZntA does not demonstrate lead, cadmium or zinc specificity in H21. Nor does proximity establish that the regulator controls the ATPase. Functional interpretation should therefore remain at the level of a candidate divalent-metal-associated ATPase and neighbouring putative regulator pending metal-dependent expression and transport measurements.

Specificity cannot be determined only from adjacent regulatory annotations: Prabhakaran et al. (2018) examined cadR-related regulation in an other bacteria, while the literature on metal sensing and transport underscores the need for organism-specific validation.

4.10. GolT-like metal-translocating ATPase

The GolT-like candidate WP_457986202.1 contains a conserved Cu-like P-type ATPase domain and two N-terminal CopZ-like metal-binding domain assignments. This structure facilitates its classification as a metal-transporting P-type ATPase with possible metal-binding properties. The domain evidence has significant importance since the GolT annotation by itself can be misinterpreted as conclusive proof of gold transmission. (Blencowe & Morby, 2003).

The examined domain combination does not clarify whether the protein facilitates the transfer of copper, gold, or an alternative substrate under physiological settings. Metal-binding motifs, transporter phylogeny, and local genomic context might assist in prioritizing rival theories; nonetheless, functional confirmation is essential for determining substrate specificity. Consequently, the candidate is more appropriately characterized as a GolT-like, Cu-type ATPase instead of a confirmed gold-resistance factor.

4.11. Second arsenic-associated genomic region

A secondary arsenic-related area on scaffold NZ_JBSYXR010000010.1 has annotations for ArsR-, ArsC-, ACR3-, and ArsH-family proteins inside seven forward-strand genes. The existence of an alternative proposed arsenite transporter family, ACR3, in conjunction with a reductase and regulator indicates a supplementary potential arsenic-processing arrangement apart from the ArsB–ArsA-linked area on scaffold 25. An ArsH-family flavoprotein could have a role in arsenic-related metabolic processes or stress reactions, although its exact physiological substrate and reaction cannot be deduced only from its family categorization. (Jeukens et al., 2017).

The scattered arsenic-related areas could suggest several genomic adaptations to arsenic-induced stress, although the preliminary assembly hinders definitive insights into their overall structure or evolutionary background. The significant finding is the presence of many anticipated arsenic-related elements, rather than a confirmed enhancement in arsenic tolerance.

4.12. Extended CzcABC-associated region

The expanded CzcABC-related area integrates anticipated RND transport elements with regulatory and supplementary metal-related annotations, including a sensor-response regulator duo, porin-like components, a transposase, and ZosA-related characteristics. The fundamental configuration of an RND permease, periplasmic adaptor, and outer-membrane element aligns with a tripartite efflux mechanism. Supplementary adjacent genes expand the biological framework but should not be indiscriminately included into the same functional unit. (Kommu et al., 2024; Riccardi et al., 2008).

The existence of a transposase may need the investigation of potential local genomic rearrangements or mobility; nonetheless, a solitary transposase annotation is inadequate to confirm horizontal gene transfer. Likewise, anticipated regulators in proximity to the cluster might affect expression, although regulatory connections need concrete proof. The locus is best seen as a potential Czc-type efflux region with supplementary, inadequately defined auxiliary roles.

Environmental resistance might be linked to mobile elements, although this connection has to be verified rather than assumed. Adhikarla et al. (2009) provide an instance of a metal-resistance determinant associated with adjacent metal-inducible mobile elements; Gaeta et al. (2022) delineate the genomic characteristics of an environmental Escherichia coli strain demonstrating mercury tolerance and multidrug resistance. Neither discovery demonstrates horizontal transmission or co-selection in H21.

4.13. RcnA-associated region and CorA-like transporter

The RcnA-family candidate WP_457984698.1 occurs near proteins annotated as TroA-like, ZnuB-like and ZnuC-like ABC transport components. Their close arrangement suggests local enrichment of metal-ion transport functions, but it does not demonstrate that the proteins form a single transport complex or share nickel/cobalt specificity. The reported incomplete C-terminal domain alignment for RcnA should also not be taken as evidence of a truncated protein without inspection of the underlying sequence and gene model. (Iwig et al., 2006; Dokpikul et al., 2016).

The distinct CorA-like protein WP_457986084.1 has a particular MtCorA-like domain alignment, reinforcing its categorization within the divalent-cation transporter family. Membership in the CorA family alone does not determine a function in nickel or cobalt resistance, since these proteins can be involved in crucial ion homeostasis. Collectively, these results expand the array of potential transporters while highlighting the distinction between recognizing an ion-transporter family and determining its true physiological substrate.

The physiological distinctions among metal transport systems are important. Fantino et al. (2010) analyzed cobalt stress responses in Escherichia coli; Kim et al. (2015) studied nickel–cobalt efflux regulation in Streptomyces coelicolor; Li and Zamble (2009) reviewed nickel homeostasis and regulation; and Okamoto and Eltis (2011) discussed cobalt trafficking. Macomber and Hausinger (2011) reviewed mechanisms of nickel toxicity, while Macomber et al. (2011) identified a specific nickel-sensitive enzyme target in E. coli. Together, these studies justify testing metal-specific phenotypes, but do not establish the substrates of H21 RcnA- or CorA-like proteins.

4.14. Consolidated conserved-domain evidence

The consolidated CDD analysis provides an independent layer of functional classification for many of the candidate proteins identified through annotation and sequence homology. Specific conserved-domain matches are particularly informative for assigning proteins to established families, whereas broader superfamily matches generally support less precise classifications. Concordance between an annotated gene name, a high-coverage homology match and a relevant domain increases confidence that a candidate belongs to a particular protein family. (Schnoes et al., 2009).

This data ought not to be construed as confirmation of in vivo metal resistance. Transporters that are closely related may exhibit variations in substrate affinity, and certain domains are found in proteins with varied biological functions. Domain E-values shown as 0 signify a measure of reporting accuracy rather than an actual probability of zero. The integrated study therefore supports prioritizing and functional theories, rather than definitive substrate allocations.

A comparable disparity between genomic capability and confirmed function is seen in metal-resistance genome analyzes. Dissanayake et al. (2025) delineated the genetic characteristics of a multi-metal-resistant isolate of Staphylococcus warneri. Kommu et al. (2023) conducted a prior metagenome-assembled-genome examination of metal resistance in polluted soils; the subsequent peer-reviewed study by Kommu et al. (2024) references the 2023 preprint, which should be regarded as associated preliminary research rather than an autonomous replication.

4.15. Preliminary antimicrobial-resistance screening

The variable results of CARD/RGI restrict the ability to draw inferences on antimicrobial resistance in H21. A protein-centric investigation yielded no qualifying results, but an independent genome-focused examination identified three exact matches linked to the adeF and qacG classifications. These discrepancies could indicate varying inputs, database setups, or analytical parameters; the current records do not clarify which rationale is applicable. Therefore, the results need not to be regarded as a confirmed antimicrobial-resistance profile. (Scoffone et al., 2021; Schnoes et al., 2009).

The evident correlation between RND-related annotations and antimicrobial-efflux classifications merits biological examination, however cross-resistance between antibiotics and heavy metals cannot be deduced only from transporter-family affiliation. Prior to dissemination, the initial input files, RGI version, CARD database version, and detection parameters must be validated, and assertions should be confined to replicable sequence-level data.

Genomic data indicating antibiotic resistance should not be equated with a susceptibility phenotype. Singh et al. (2018) performed comparative genomic studies on multidrug-resistant Enterobacter bugandensis isolates, exemplifying genome-centric antimicrobial-resistance research in a different species. It fails to authenticate the H21 CARD/RGI forecasts or confirm metal–antibiotic co-resistance.

4.16. Taxonomic interpretation

The conducted ANI comparisons demonstrated around 90.31% identity to the analyzed A. denitrificans reference and 90.73% to the scrutinized A. veterisilvae reference, although with restricted alignment coverage. These findings do not support the allocation of H21 to any of the examined species. Designating the bacterium as Achromobacter sp. H21 is hence more justifiable

than use a species name derived from the present comparisons. (Gomila et al., 2014; Jeukens et al., 2017).

Taxonomic ambiguity is significant for comparative analysis due to the variability of metal-associated gene collections across related strains. A more extensive comparison using well-chosen reference genomes, alignment-fraction analysis, and supplementary phylogenomic methodologies is essential prior to drawing species-specific evolutionary or ecological inferences.

Species assignment also benefits from orthogonal approaches. Saitta et al. (2025) investigated a 16S rRNA-based molecular identification protocol for Achromobacter, providing a complementary taxonomic perspective; this does not replace genome-wide comparisons for resolving the present H21 ANI ambiguity.

4.17. Integration of the initial genomic finding

Taken together, the initial genome annotation, protein-based homology comparisons, conserved-domain assignments and local gene maps reveal several distinct classes of candidate metal-associated determinants. The most coherent configurations include the compact ArsB–ArsA–ArsD–ArsC–ArsR region, the Czc-type tripartite efflux neighbourhood, and the ZntA-like ATPase with an adjacent putative regulator. Other candidates, including chromate transporters, a GolT-like ATPase and RcnA/CorA-family proteins, provide additional hypotheses that vary in the specificity of their supporting evidence. (Ayangbenro & Babalola, 2020; Kommu et al., 2024).

The strength of this integrated approach is the ability to distinguish well-supported family assignments from weaker predictions. Its limitation is that the analyses remain observational and sequence based. The appropriate overarching conclusion is the presence of a diverse putative metal-homeostasis repertoire rather than verified resistance to a defined panel of metals.

Follow-up based on expression is crucial; yet, transcriptional reactions are contingent on environment. Kies et al. (2026) examined the transcriptional modifications of Staphylococcus aureus under several sulfur-supplementation scenarios. Despite its lack of connection to arsenic efflux or H21, this instance exemplifies the wider experimental significance of condition-specific transcriptional assessments and should not be referenced as direct proof of metal-resistance gene regulation.

4.18. Batch CDD validation of 15 BLASTN-mapped proteins

All 15 distinct proteins preserved during BLASTN-to-CDS mapping yielded CDD correspondences. The 727 reported hits and 427 clusters stem from overlapping domain models and must not be construed as several discrete domains or functional genes. The two ArsB-like proteins were classified within the ArsB family, but the ArsH candidate was clearly supported by the ArsH family. The CzcA-like and MexI-like proteins corresponded to the extensive ACR_tran RND-efflux superfamily, therefore confirming their categorization at the family level without clarifying substrate specificity. (Schnoes et al., 2009).

The residual proteins demonstrate the diversity of the evaluated collection. CueR-like and response-regulator domains indicate possible regulatory functions; ModB and Pst components are related to the transport of molybdate or phosphate; Dps and superoxide dismutase are connected to the defense against cellular stress; and RuvB is linked to DNA repair mechanisms. Their identification could pertain to overarching stress physiology; nonetheless, their categorization among nucleotide-reference hits does not adequately qualify them as direct metal-efflux or detoxifying proteins. This evidence-driven distinction enhances the biological accuracy of the candidate list.

4.19. Genomic neighbourhood of the second ArsB-like transporter

The second ArsB-like transporter, WP_457986626.1, occurs on scaffold NZ_JBSYXR010000009.1 within a transporter-rich region containing DMT-, AEC- and YeeE/YedE-family annotations. The positive-strand orientation of these neighbouring genes and their relatively short intergenic separations indicate close local organization. In contrast to the five-gene arsenic-associated region on scaffold 25, no canonical arsR, arsC, arsA or arsD annotation was identified within the examined 10-kb flanks of this second ArsB-like gene. (Li et al., 2012; Jeukens et al., 2017).

This contrast suggests that the two ArsB-family candidates occur in different genomic contexts, which may be relevant to their regulation or physiological roles. However, the surrounding DMT, AEC and YeeE/YedE proteins cannot be assigned arsenic transport functions from proximity alone, and an ArsB-like domain does not prove transport activity. The second locus is therefore a useful target for comparative promoter analysis, arsenic-dependent expression profiling and direct transporter characterization.

4.20. Two RND efflux-associated modules around the MexI-like transporter

The MexI-like protein WP_457985789.1 is situated between a hypothesized periplasmic adaptor and an outer-membrane efflux element on the positive strand. The 16- and 12-base pair intergenic intervals and their common orientation support a credible tripartite RND efflux structure. An independent upstream module on the negative strand similarly includes an outer-membrane subunit, an RND permease, and a periplasmic adaptor, with closely positioned or overlapping coding regions. These configurations provide structural proof for two neighboring but oppositely aligned potential RND systems. (Scoffone et al., 2021; Jeukens et al., 2017).

The discovery is significant for comprehending the arrangement of local efflux systems; nevertheless, it does not confirm that any module facilitates the export of heavy metals. MexI/MexW-like and other RND-family transporters may possess diverse substrate ranges, and the ACR_tran domain is too expansive to differentiate them. Neither can shared strand orientation nor brief gaps validate co-transcription. Focused transcriptional assessments and efflux experiments under specified metal and antimicrobial conditions are essential to ascertain whether these modules are involved in metal regulation, antimicrobial expulsion, or alternative transport roles.

DECLARATIONS

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Ethics, Consent to Participate, and Consent to Publish declarations

Not applicable.

REFERENCES

Adhikarla, H., Tiwari, A., Patnala, K., Rodrigue, A., Mohan, P. M., & Singh, S. S. (2009). MrdH, a Novel Metal Resistance Determinant of Pseudomonas putida KT2440, Is Flanked by Metal-Inducible Mobile Genetic Elements. Journal of Bacteriology, 191(19), 5976–5987. https://doi.org/10.1128/jb.00465-09

Amoutzias, G. D., Nikolaidis, M., & Hesketh, A. (2022). The Notable Achievements and the Prospects of Bacterial Pathogen Genomics. Microorganisms, 10(5), 1040. https://doi.org/10.3390/microorganisms10051040

Ayangbenro, A. S., & Babalola, O. O. (2020). Genomic analysis of Bacillus cereus NWUAB01 and its heavy metal removal from polluted soil. Scientific Reports, 10. https://doi.org/10.1038/s41598-020-75170-x

Blencowe, D. K., & Morby, A. P. (2003). Zn(II) metabolism in prokaryotes. FEMS Microbiology Reviews, 27(2–3), 291–311. https://doi.org/10.1016/S0168-6445(03)00041-X

Brown, N. L., Stoyanov, J. V., Kidd, S. P., & Hobman, J. L. (2003). The MerR family of transcriptional regulators. FEMS Microbiology Reviews, 27(2–3), 145–163. https://doi.org/10.1016/S0168-6445(03)00051-2

Chung, A. P., Francisco, R., Morais, P. V., & Branco, R. (2022). Genome mining to unravel potential metabolic pathways linked to gallium bioleaching ability of bacterial mine isolates. Frontiers in Microbiology, 13. https://doi.org/10.3389/fmicb.2022.970147

Cornet, L., & Baurain, D. (2022). Contamination detection in genomic data: more is not enough. Genome Biology, 23(1). https://doi.org/10.1186/s13059-022-02619-9

Cornet, L., Durieu, B., Baert, F., D'hooge, E., Colignon, D., Meunier, L., Lupo, V., Cleenwerck, I., Daniel, H., Rigouts, L., Sirjacobs, D., Declerck, S., Vandamme, P. A. R., Wilmotte, A., Baurain, D., & Becker, P. (2022). The GEN-ERA toolbox: unified and reproducible workflows for research in microbial genomics. Gigascience, 12. https://doi.org/10.1093/gigascience/giad022

Correia, L. N. (2026). Exploring the oral resistome: From metagenomics to precision oral health. Microorganisms, 14(9), 2033. https://doi.org/10.3390/microorganisms14092033

Coward, A., Kenna, D., Perry, C., Martin, K., Doumith, M., & Turton, J. F. (2016). Use of nrdA gene sequence clustering to estimate the prevalence of different Achromobacter species among cystic fibrosis patients in the UK. Journal of Cystic Fibrosis, 15(4), 479–485. https://doi.org/10.1016/j.jcf.2015.09.005

Ding, C., Ding, Z., Liu, Q., Liu, W., & Chai, L. (2024). Advances in mechanism for the microbial transformation of heavy metals: Implications for bioremediation strategies. Chemical Communications, 60(85), 12315–12332. https://doi.org/10.1039/D4CC03722G

Dissanayake, D., Chandrasekharan, N. V., & Wijayarathna, C. D. (2025). Multi-metal-resistant Staphylococcus warneri strain TWSL_1: revealing heavy metal-resistant genomic features by whole-genome sequencing and analysis. Access Microbiology, 7(5). https://doi.org/10.1099/acmi.0.000954.v5

Dokpikul, T., Chaoprasid, P., Saninjuk, K., Sirirakphaisarn, S., Johnrod, J., Nookabkaew, S., Sukchawalit, R., & Mongkolsuk, S. (2016). Regulation of the cobalt/nickel efflux operon dmeRF in Agrobacterium tumefaciens and a link between the iron-sensing regulator RirA and cobalt/nickel resistance. Applied and Environmental Microbiology, 82(15), 4732–4742. https://doi.org/10.1128/AEM.01262-16

Fantino, J., Py, B., Fontecave, M., & Barras, F. (2010). A genetic analysis of the response of Escherichia coli to cobalt stress. Environmental Microbiology, 12(10), 2846–2857. https://doi.org/10.1111/j.1462-2920.2010.02265.x

Farooq, A., & Rafique, A. (2025). Composite Genome Quality Index for Pathogenic Bacterial Genomes. Applied Microbiology, 5(4), 144. https://doi.org/10.3390/applmicrobiol5040144

Fichant, G., Basse, M.-J., & Quentin, Y. (2006). ABCdb: an online resource for ABC transporter repertories from sequenced archaeal and bacterial genomes. Fems Microbiology Letters, 256(2), 333–339. https://doi.org/10.1111/j.1574-6968.2006.00139.x

Gaeta, N. C., Carvalho, D. U. de, Fontana, H., Sano, E., Moura, Q., Fuga, B., Muñoz, P., Gregory, L., & Lincopán, N. (2022). Genomic features of a multidrug-resistant and mercury-tolerant environmental Escherichia coli recovered after a mining dam disaster in South America. The Science of the Total Environment, 823, 153590. https://doi.org/10.1016/j.scitotenv.2022.153590

Gomila, M., Prince-Manzano, C., Svensson-Stadler, L., Busquets, A., Erhard, M., Martínez, D. L., Lalucat, J., & Moore, E. R. B. (2014). Genotypic and phenotypic applications for the differentiation and species-level identification of Achromobacter for clinical diagnoses. PLOS ONE, 9(12), e114356. https://doi.org/10.1371/journal.pone.0114356

Hogle, S., Thrash, J. C., Dupont, C. L., & Barbeau, K. A. (2016). Trace Metal Acquisition by Marine Heterotrophic Bacterioplankton with Contrasting Trophic Strategies. Applied and Environmental Microbiology, 82(5), 1613–1624. https://doi.org/10.1128/aem.03128-15

Hoque, M. N., Mannan, A. B. A., Hossian, A. K. M. N., Faisal, G. M., Hossain, M. A., & Sultana, M. (2024). Arsenotrophic Achromobacter aegrifaciens strains isolated from arsenic contaminated tubewell water and soil sources shared similar genomic potentials. BMC Microbiology, 24(1). https://doi.org/10.1186/s12866-024-03676-9

Hrynkiewicz, K., Złoch, M., Kowalkowski, T., Baum, C., Niedojadło, K., & Buszewski, B. (2014). Strain-specific bioaccumulation and intracellular distribution of Cd^{2+} in bacteria isolated from the rhizosphere, ectomycorrhizae, and fruitbodies of ectomycorrhizal fungi. Environmental Science and Pollution Research, 22(4), 3055–3067. https://doi.org/10.1007/s11356-014-3489-0

Iwig, J. S., Rowe, J. L., & Chivers, P. T. (2006). Nickel homeostasis in Escherichia coli—the rcnR–rcnA efflux pathway and its linkage to NikR function. Molecular Microbiology, 62(1), 252–262. https://doi.org/10.1111/j.1365-2958.2006.05369.x

Jeukens, J., Freschi, L., Vincent, A. T., Emond-Rhéault, J.-G., Kukavica-Ibrulj, I., Charette, S. J., & Lévesque, R. C. (2017). A pan-genomic approach to understand the basis of host adaptation in Achromobacter. Genome Biology and Evolution, 9(4), 1030–1046. https://doi.org/10.1093/gbe/evx061

Jiménez, A. R. (2009). Transition Metal Homeostasis. Ecosal Plus, 3(2). https://doi.org/10.1128/ecosalplus.5.4.4.3

Jomová, K., Alomar, S. Y., Nepovimová, E., Kuča, K., & Valko, M. (2024). Heavy metals: Toxicity and human health effects. Archives of Toxicology, 99(1), 153–209. https://doi.org/10.1007/s00204-024-03903-2

Jung, J., & Lee, S. J. (2019). Biochemical and biodiversity insights into heavy metal ion-responsive transcription regulators for synthetic biological heavy metal sensors. Journal of Microbiology and Biotechnology, 29(10), 1522–1542. https://doi.org/10.4014/jmb.1908.08002

Kies, P. J., Zimpel, C. K., Lensmire, J. M., Major, M. R., Burtchett, T., Wischer, M. R., & Hammer, N. D. (2026). Defining the transcriptional adaptation of Staphylococcus aureus to a range of nutritional sulfur supplementation. https://doi.org/10.64898/2026.05.20.726469

Kim, H. M., Ahn, B., Lee, J.-H., & Roe, J. (2015). Regulation of a nickel–cobalt efflux system and nickel homeostasis in a soil actinobacterium Streptomyces coelicolor. Metallomics, 7(4), 702–709. https://doi.org/10.1039/c4mt00318g

Kommu, N., Stothard, P., Chukwujindu, C., Pathak, A. K., & Chauhan, A. (2024). Utilizing a metagenome assembled genome approach revealed further insights into microbially mediated heavy-metal resistance in soils from a former nuclear materials production facility. Applied Microbiology, 4(1), 376–389. https://doi.org/10.3390/applmicrobiol4010026

Kommu, N., Stothard, P., Chukwujindu, C., Pathak, A., & Chauhan, A. (2023). Metagenome Assembled Genomes (MAG) Facilitate a Better Understanding of Microbially-mediated Heavy Metal Resistance in Soils from a Former Nuclear Materials Production Facility. https://doi.org/10.1101/2023.10.20.563326

Li, B., Jo, M., Liu, J., Tian, J., Canfield, R. E., & Bridwell-Rabb, J. (2022). Structural and mechanistic basis for redox sensing by the cyanobacterial transcription regulator RexT. Communications Biology, 5(1). https://doi.org/10.1038/s42003-022-03226-x

Li, X., Hu, Y., Gong, J., Lin, Y., Johnstone, L., Rensing, C., & Wang, G. (2012). Genome sequence of the highly efficient arsenite-oxidizing bacterium Achromobacter arsenitoxydans SY8. Journal of Bacteriology, 194(5), 1243–1244. https://doi.org/10.1128/JB.06667-11

Li, Y., & Zamble, D. B. (2009). Nickel Homeostasis and Nickel Regulation: An Overview. Chemical Reviews, 109(10), 4617–4643. https://doi.org/10.1021/cr900010n

Lorková, Z., Cimermanová, M., Piknová, M., Adhvaryu, S. G., Pristaš, P., & Kisková, J. (2025). Environmental impact on the genome shaping of putative new Streptomyces species. BMC Microbiology, 25(1). https://doi.org/10.1186/s12866-025-03779-x

Macomber, L., & Hausinger, R. P. (2011). Mechanisms of nickel toxicity in microorganisms. Metallomics, 3(11), 1153. https://doi.org/10.1039/c1mt00063b

Macomber, L., Elsey, S. P., & Hausinger, R. P. (2011). Fructose-1,6-bisphosphate aldolase (class II) is the primary site of nickel toxicity in Escherichia coli. Molecular Microbiology, 82(5), 1291–1300. https://doi.org/10.1111/j.1365-2958.2011.07891.x

Okamoto, S., & Eltis, L. D. (2011). The biological occurrence and trafficking of cobalt. Metallomics, 3(10), 963. https://doi.org/10.1039/c1mt00056j

Passi, A., Tibocha-Bonilla, J. D., Kumar, M., Tec-Campos, D., Zengler, K., & Zúñiga, C. (2021). Genome-scale metabolic modeling enables in-depth understanding of big data. Metabolites, 12(1), 14. https://doi.org/10.3390/metabo12010014

Pita-Galeana, M. A., Ruhle, M., López-Vázquez, L., Anda-Jáuregui, G. de, & Lemus, E. H. (2025). Computational metagenomics: State of the art. International Journal of Molecular Sciences, 26(18), 9206. https://doi.org/10.3390/ijms26189206

Potdukhe, R. M., Bedi, P., Sarangi, B. K., Pandey, R., & Thul, S. T. (2018). Root transcripts associated with arsenic accumulation in hyperaccumulator Pteris vittata. Journal of Biosciences, 43(1), 105–115. https://doi.org/10.1007/s12038-018-9735-8

Prabhakaran, R., Rajkumar, S. N., Ramprasath, T., & Selvam, G. S. (2018). Identification of promoter P cadR , in silico characterization of cadmium resistant gene cadR and molecular cloning of promoter P cadR from Pseudomonas aeruginosa BC1. Toxicology and Industrial Health, 34(12), 819–833. https://doi.org/10.1177/0748233718795934

Riccardi, G., Milano, A., Pasca, M. R., & Nies, D. H. (2008). Genomic analysis of zinc homeostasis in Mycobacterium tuberculosis. FEMS Microbiology Letters, 287(1), 1–7. https://doi.org/10.1111/j.1574-6968.2008.01320.x

Saitta, G., Veschetti, L., Feletti, R., Sandri, A., Boaretti, M., Melotti, P., Carelli, M., Lleò, M. M., Malerba, G., & Signoretto, C. (2025). Development of a Simple and Accurate Molecular Protocol Using 16SrRNA for Species-Specific Identification of Achromobacter spp. Pathogens, 14(3), 271. https://doi.org/10.3390/pathogens14030271

Salazar-Flores, J., Torres-Jasso, J. H., Rojas-Bravo, D., Villela, Z. M. R., & Torres-Sánchez, E. D. (2019). Effects of mercury, lead, arsenic and zinc to human renal oxidative stress and functions: A review. Journal of Heavy Metal Toxicity and Diseases, 4(1). https://doi.org/10.21767/2473-6457.10027

Sánchez-Corona, C. G., Gonzalez-Avila, L. U., Hernández-Cortez, C., Rojas-Vargas, J., Castro-Escarpulli, G., & Castelán-Sánchez, H. G. (2025). Impact of heavy metal and resistance genes on antimicrobial resistance: Ecological and public health implications. Genes, 16(6), 625. https://doi.org/10.3390/genes16060625

Schnoes, A. M., Brown, S. D., Dodevski, I., & Babbitt, P. C. (2009). Annotation error in public databases: Misannotation of molecular function in enzyme superfamilies. PLOS Computational Biology, 5(12), e1000605. https://doi.org/10.1371/journal.pcbi.1000605

Scoffone, V. C., Trespidi, G., Barbieri, G., Irudal, S., Perrin, E., & Buroni, S. (2021). Role of RND efflux pumps in drug resistance of cystic fibrosis pathogens. Antibiotics, 10(7), 863. https://doi.org/10.3390/antibiotics10070863

Singh, N. K., Bezdan, D., Sielaff, A. C., Wheeler, K., Mason, C. E., & Venkateswaran, K. (2018). Multi-drug resistant Enterobacter bugandensis species isolated from the International Space Station and comparative genomic analyses with human pathogenic strains. BMC Microbiology, 18(1). https://doi.org/10.1186/s12866-018-1325-2

Sorouri, B., Rodriguez, C. I., Gaut, B. S., & Allison, S. (2023). Variation in Sphingomonas traits across habitats and phylogenetic clades. Frontiers in Microbiology, 14. https://doi.org/10.3389/fmicb.2023.1146165

Venter, E., Smith, R., & Payne, S. (2011). Proteogenomic Analysis of Bacteria and Archaea: A 46 Organism Case Study. Plos One, 6(11), e27587. https://doi.org/10.1371/journal.pone.0027587

Wu, Y., & Outten, F. W. (2011). Fur and the Novel Regulator YqjI Control Transcription of the Ferric Reductase Gene yqjH in Escherichia coli. Journal of Bacteriology, 193(2), 563–574. https://doi.org/10.1128/jb.01062-10