



Research article

Comparative genomics based exploration of xenobiotic degradation patterns in *Glutamicibacter*, *Arthrobacter*, and *Pseudarthrobacter* isolated from diverse ecological habitats

Nisha Ghimire^{a,b}, Byeollee Kim^{a,b,c}, So-Ra Han^{a,b,c}, Tae-Jin Oh^{a,b,c,d,*}^a Department of Life Science and Biochemical Engineering, Graduate School, SunMoon University, Asan, 31460, Republic of Korea^b Bio Big Data-based Chungnam Smart Clean Research Leader Training Program, SunMoon University, Asan, 31460, Republic of Korea^c Genome-based BioIT Convergence Institute, Asan, 31460, Republic of Korea^d Department of Pharmaceutical Engineering and Biotechnology, SunMoon University, Asan, 31460, Republic of Korea

ARTICLE INFO

Keywords:

Xenobiotic
KEGG pathway
Arthrobacter
Pseudarthrobacter
Glutamicibacter

ABSTRACT

Xenobiotics pose a substantial threat to environmental integrity by disrupting normal ecosystems. The genus *Arthrobacter*, known for its metabolic versatility can degrade several xenobiotic compounds. *Arthrobacter* strains have also undergone frequent taxonomic revisions and reclassifications to strains including *Pseudarthrobacter* and *Glutamicibacter*. Here, we present the complete genome sequence of *Glutamicibacter protophormiae* strain NG4, isolated from a coastal area surrounded by chemical plants. Further, through comparative genomics involving 55 strains from *Glutamicibacter*, *Arthrobacter*, and *Pseudarthrobacter*, we elucidated taxonomic relationships and xenobiotic degradation potential. Our genomics-based findings revealed a generally even distribution of xenobiotic-degrading genes and pathways among the studied strains. *Glutamicibacter* species emerged as potential candidate for steroid degradation. A significant number of host-specific and environmental isolates predominantly possessed pathways for 4-hydroxybenzoate (4-HB) degradation and only the environmental isolates possessed benzoate degradation pathway. Certain *Arthrobacter* and *Pseudarthrobacter* species isolated from the environmental settings were identified as potential degraders of toluene, xylene, and phenanthrene. Notably, most strains contained pathways for azathioprine, capecitabine, and 5-fluorouridine pharmaceutical drug metabolism. Overall, our findings shed light on microbial metabolic diversity among 55 strains isolated from diverse sources and hint the importance of strict environmental monitoring. Further, for the application of the putative xenobiotic degrading strains, experimental validation is required in the future.

1. Introduction

Due to the proliferation of chemical compound synthesis in the industrially revolutionizing world, xenobiotics pose a significant environmental threat due to their toxicity, persistence, and accumulation in the environment. Microbial degradation of xenobiotics is an area of interest owing to the metabolic diversity and ubiquitous presence of microbes across diverse environments [1–3]. Microbial

* Corresponding author. Department of Pharmaceutical Engineering and Biotechnology, SunMoon University, Asan, 31460, Republic of Korea.
E-mail address: tjoh3782@sunmoon.ac.kr (T.-J. Oh).

<https://doi.org/10.1016/j.heliyon.2024.e40280>

Received 6 August 2024; Received in revised form 5 November 2024; Accepted 7 November 2024

Available online 8 November 2024

2405-8440/© 2024 Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

activity is commonly regarded as the primary process responsible for the mineralization or full biodegradation of organic compounds found in water and soil in many cases [4]. Microbes have evolved various genetic adaptations to degrade xenobiotic compounds, including gene mutation, duplication, transposition, insertional activation, and DNA rearrangements [5]. Horizontal gene transfer also plays a vital role in microbial adaptation to xenobiotics [6].

Members of the genus *Arthrobacter* exhibit remarkable metabolic versatility and are renowned for their ability to degrade a broad spectrum of xenobiotic compounds in the environment (Supplementary Table S1), including aromatic compounds [7,8], polycyclic aromatic hydrocarbons (PAH) [9,10], pesticides [11,12], lignin [13] and more. *Arthrobacter* species demonstrate a ubiquitous presence across various environments, ranging from extreme habitats to soil and sites contaminated with environmental pollutants (Supplementary Table S2). These characteristics highlight the importance of comprehending the genomic features prevalent in these metabolically adaptable species. The xenobiotic degradation category in Kyoto Encyclopedia of Genes and Genomes (KEGG) [14] encompasses the analysis of diverse categories of xenobiotic compound degradation genes and pathways, offering insights into the xenobiotic degradation pattern across a wide range of microorganisms at a large scale.

The taxonomy of *Arthrobacter* species has undergone frequent revisions and reclassifications, leading to the emergence of *Pseudarthrobacter* and *Glutamicibacter* based on differences in their quinone system, polar lipid profiles, and peptidoglycan composition [15]. However, several reports for the degradation of xenobiotic compounds like Phthalic acid, o-nitrobenzoate and

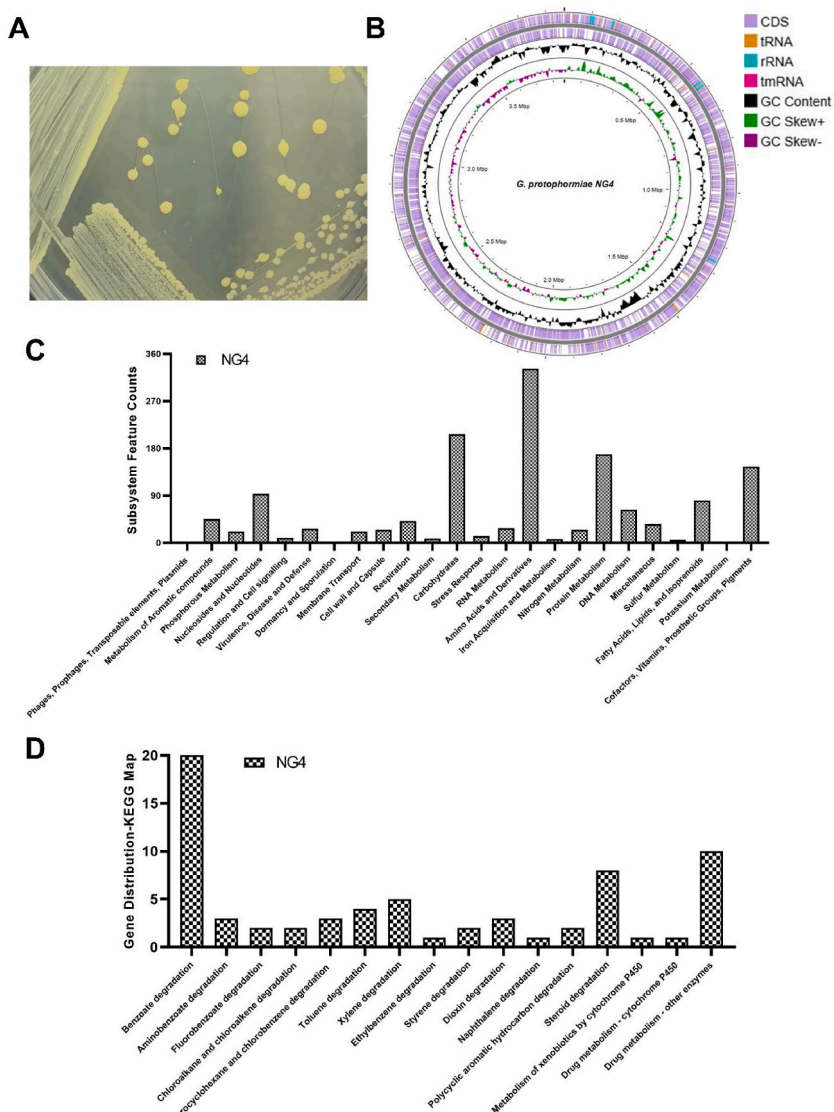


Fig. 1. Isolation and whole genome analysis of *G. protophormiae* strain NG4. (A) Pure isolate of the strain NG4. (B) Circular map with the whole genome information of the strain NG4. The outer circle to inner circle includes, NG4 coding sequence obtained from Prokka annotation (purple color), GC content, GC skew +, and GC skew -. (C) Depiction of subsystem feature counts of NG4 obtained from RAST annotation, and distribution of genes in the xenobiotics degradation category of KEGG pathway analysis (D).

p-hydroxybenzoate have also been reported for *Pseudarthrobacter* [16,17] and *Glutamicibacter* [18,19] strains (Supplementary Table S1). *G. protophormiae* NG4 from this study was isolated from the west coastal area surrounded by chemical industries. Originally described as *Brevibacterium protophormiae* [20], this strain was later reclassified as *Arthrobacter protophormiae* in 1983 [21], and further reclassified as *G. protophormiae* in 2016 [15]. The xenobiotic degradation potential of *G. protophormiae* species has not been extensively explored in prior studies. This study presents a comprehensive analysis and comparison of xenobiotic degradation genes and pathways among strains of *Arthrobacter*, *Pseudarthrobacter*, and *Glutamicibacter*, delineating their relationship and potential role in xenobiotic degradation.

2. Results and discussion

2.1. Isolation and complete genomic characterization of *G. protophormiae* NG4

G. protophormiae strain NG4 was isolated from a coastal area surrounded by chemical plants, suggesting potential exposure to aromatic pollutants. Analysis of the 16S rRNA sequence using EzBioCloud classified the strain as *G. protophormiae*, 100 % identical to *G. protophormiae* DSM 20168. In-silico taxonomy analysis using the whole genome in TYGS identified *G. protophormiae* DSM 20168 and JCM 1973 as the closest type strains of NG4 strain (Supplementary Fig. S1). Further, 16S rRNA-based phylogenetic tree was constructed to compare various bacterial genera within the order *Micrococcales*, including representatives from different families such as *Micrococcaceae*, *Dermacoccaceae*, *Cellulomonadaceae*, and others (Supplementary Fig. S2). *G. protophormiae* NG4, which belongs to the order *Micrococcales* and the family *Micrococcaceae*, clustered with other genera like *Paeniglutamicibacter*, *Zhihengliuella*, *Renibacterium*,

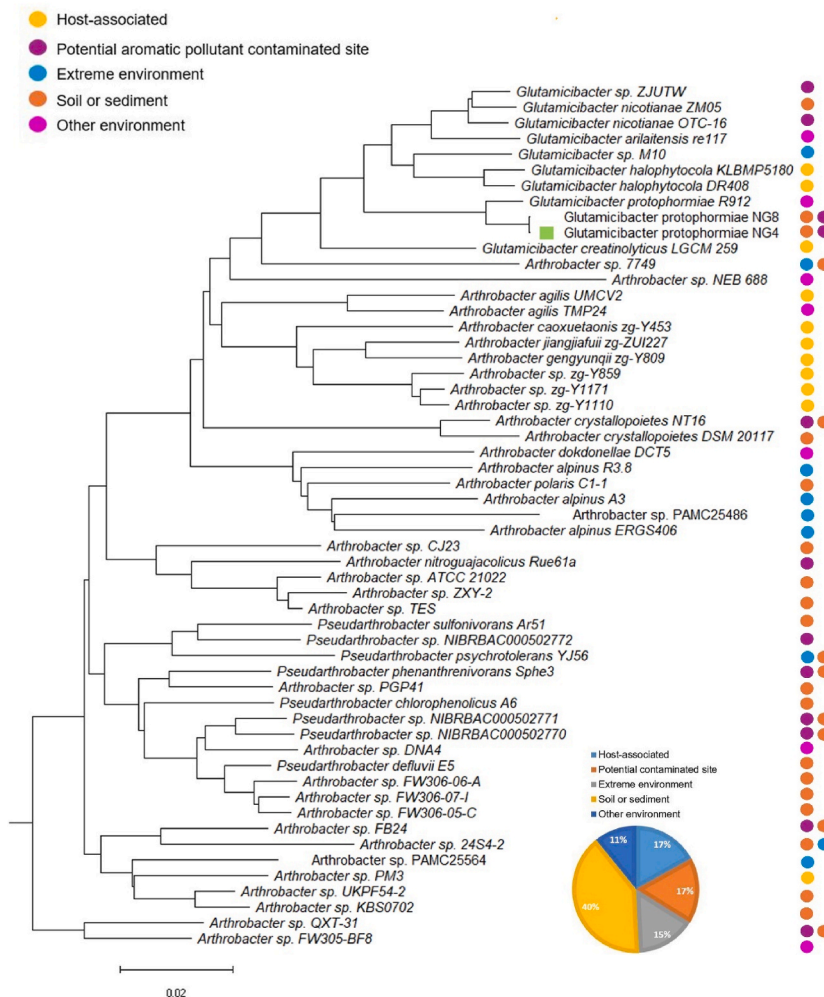


Fig. 2. Pan/core phylogeny of the 55 selected strains for pan-genome analysis from three different genera and the designation of their isolation source. NG4 strain from this study is designated with a green rectangular box. Each strain consists of the designation of their isolation environment to the right corner using the circles with respective labels. The pie-chart represents the distribution of the selected strains among various isolation sources.

Arthrobacter, *Pseudarthrobacter*, *Paenarthrobacter*, and *Micrococcus*, all of which are also members of the *Micrococcaceae* family. Meanwhile, other species from the same family were grouped into a separate clade, indicating a closer relationship within the species in the first clade (Supplementary Fig. S2). However, lower bootstrap values (less than 50 %) at certain branch points suggest limited statistical support for the precise relationships among some of these strains.

The genome sequencing and quality analysis of *G. protophormiae* strain NG4 yielded a complete assembly with a genome size of 3.9 Mb, comprising one chromosome and one scaffold, both measuring 3.9 Mb. The assembly contains 3613 genes, including 3496 protein-coding genes, and exhibits a GC content of 64 % with a genome coverage of 100.0x. Quality assessment using CheckM (v1.2.2) indicated a completeness of 98.99 % (88th percentile) and a low contamination rate of 0.92 %. Detailed assembly statistics are summarized in Supplementary Table S3. Fig. 1A depicts the pure isolate of the strain NG4. The circular map (Fig. 1B) illustrates coding sequences (CDS) and other genomic information of NG4. Further, the genome annotation of the strain was performed using Rapid Annotation Subsystem Technology (RAST) and Kyoto Encyclopedia of Genes and Genomes (KEGG). In the RAST annotation, 24 % of the coding genes fell within the subsystem coverage, with the majority allocated to carbohydrates, amino acids, and derivatives metabolism categories. Notably, genes related to metabolism of aromatic compounds were also present (Fig. 1C). Further analysis of KEGG pathways involved in xenobiotics degradation and metabolism revealed significant gene distribution in KEGG subcategories such as benzoate metabolism, steroid degradation, and drug metabolism by other enzymes (Fig. 1D).

Further, the elucidation of xenobiotic degradation genes and pathways in *G. protophormiae* NG4 was conducted alongside a comparative genomic analysis with closely related genera like *Arthrobacter* and *Pseudarthrobacter*. These genera, along with *Glutamicibacter*, belong to the *Micrococcaceae* family, and Supplementary Fig. S2 shows the clustering of these species within the same clade. Notably, *Arthrobacter* and *Pseudarthrobacter* were selected for comparison because of their well-established roles in xenobiotic degradation compared to other strains in the clade, even though some strains in the clade including *Paenarthrobacter* and *Micrococcus* have also shown possibility in xenobiotic degradation. The close phylogenetic relationship of these genera is also reflected in the frequent taxonomic reclassification of *Arthrobacter* strains, which have been reassigned over time to other genera such as *Glutamicibacter*, *Pseudarthrobacter*, *Paeniglutamicibacter*, *Pseudoglutamicibacter*, and *Paenarthrobacter* [15]. For comparative study, a total of 11 *Glutamicibacter*, 36 *Arthrobacter*, and 8 *Pseudarthrobacter* species were compared (Fig. 2). To ensure a consistent basis for comparison, only strains with complete genome sequences were selected, excluding draft genomes.

2.2. Pan-genome analysis and pan-phylogeny construction

Using a 70 % cutoff value for clustering, the analysis revealed 72 core genes, 27,400 accessory genes, and 36,781 unique genes among the strains. The core-pan plot, distribution of the gene families among the 55 genomes and the number of new genes among the strains are shown in Supplementary Figs. S3A–C respectively. The constructed pan phylogeny positioned NG4 strain among *Glutamicibacter* and *Arthrobacter* species, while several *Arthrobacter* strains were closely grouped with *Pseudarthrobacter* strains (Fig. 2). Notably, while the speciation point of *Glutamicibacter* strains was clearly delineated, the speciation points for all *Pseudarthrobacter* strains appeared less distinct, with *Arthrobacter* strains (PGP41, DNA4, FW306-06-A, FW306-07-I, and FW306-05-C) interspersed among *Pseudarthrobacter* species.

Recently, several *Arthrobacter* strains were reclassified into five distinct genera, including *Pseudarthrobacter* and *Glutamicibacter*, based on differences in quinone systems, peptidoglycan types, and polar lipid profiles [15]. Previous studies have described the reclassification of *Arthrobacter* species such as *protophormiae*, *nicotianae*, *arilaitensis*, and *creatinolyticus* to *Glutamicibacter* strains, as well as the reclassification of species like *sulfonivorans*, *phenanthrenivorans*, *chlorophenolicus*, and *defluvii* to *Pseudarthrobacter* [15]. Our findings, based on pan/core phylogeny analysis, indicate some *Arthrobacter* and *Pseudarthrobacter* strains closely cluster together in phylogeny. Previous proposals for reclassifying bacterial taxonomy have been based on phylogenomic tree analysis, including core genes of the strains [22]. This study suggests a revision of taxonomic classification between *Arthrobacter* and *Pseudarthrobacter* strains, particularly for those where the speciation event in the phylogenetic tree is not clear. Furthermore, in the ANI analysis, *Arthrobacter* species which clustered together with *Pseudarthrobacter* strains in the phylogenetic tree had more nucleotide identity percentages to *Pseudarthrobacter* than *Arthrobacter* strains (Supplementary Table S4). However, it's important to note that ANI analysis is primarily used to estimate genetic relatedness and aid in species delineation [23] where, the species that shares ≥ 95 % ANI value are considered as the same species [24,25]. Through ANI analysis, this study identified an uncharacterized species *G. sp. ZJUTW* as *G. nicotianae* with ANI value of 100 % with *G. nicotianae* ZM05 and OTC-16 (Supplementary Table S4) thus proposing the species delineation of *G. sp. ZJUTW* to *G. nicotianae*. Similarly, ANI analysis also showed *G. protophormiae* NG4 and *G. protophormiae* NG8 to be an identical strain with 100 % identity, however, further analysis is required to confirm the identical nature of these two strains.

Besides the phylogenetic relationship, understanding the relation of the metabolic capacity of a strain and their isolation environment is beneficial to predict certain environmental conditions. Therefore, the isolation sources of the strains under study were determined, indicating that the majority were isolated from soil, followed by host-associated strains, contaminated sites, extreme environmental conditions, and other environments such as cheese, poultry farms, and water (Fig. 2 and Supplementary Table S2). In Fig. 2, it is evident that most *Arthrobacter* strains isolated from host-specific sources clustered together, as did strains isolated from soil and extreme environments. However, there was no distinct separation based on isolation source, with strains often clustering randomly among multiple groups within the same isolation environment.

2.3. Comparative analysis of xenobiotic degradation pathway genes among the three genera

The xenobiotic degradation categories in the KEGG pathway were thoroughly analyzed across the 55 strains. It was observed that

nearly all strains exhibited a higher abundance of genes (some genes involved in multiple reactions within the same degradation pathway) for benzoate degradation, chlorocyclohexane-chlorobenzene degradation, and drug metabolism by other enzymes (Fig. 3). Additionally, certain strains displayed a notable abundance of genes in categories such as fluorobenzoate degradation and xylene degradation, while a few genes consistently appeared among most species in categories like aminobenzoate degradation, toluene degradation, and chloroalkane—chloroalkene degradation. Intriguingly, four *Glutamicibacter* species exhibited the presence of genes and pathways for steroid degradation, a feature absent among all other species. Two strains Sphe3 and ATCC21022 possessed quite a few genes for polycyclic hydrocarbon degradation among all the strains.

2.4. Steroid degradation pathway is unique to *Glutamicibacter* among the three genera

Steroids, particularly naturally occurring steroid hormones, are recognized as common environmental pollutants, notably in wastewater or sewage treatment plants, where they can potentially exert endocrine-disrupting effects on humans and animals [26,27]. Various strategies have been employed to mitigate their presence in the environment [28,29], and several bacterial strains have been reported to degrade steroids [30,31]. The presence of the steroid degradation pathway and associated genes exclusively among *Glutamicibacter* species prompted further investigation. Thus, all *Glutamicibacter* species were selected for analysis, regardless of their genomic status (complete or draft genome) and isolation source unavailability.

Initially, the steroid degradation genes among the selected *Glutamicibacter* strains NG4, NG8, R912, and ZM05 were examined. It was discovered that all of them harbored the androstenedione degradation pathway (Fig. 4A, Table 1), an intermediate in the degradation of epiandrosterone and cholesterol. Additionally, NG4, NG8, and R912 strains possessed the cholesterol oxidase enzyme, which serves as the initial key enzyme for cholesterol degradation (Fig. 4A). Subsequently, all *Glutamicibacter* strains available in the JGI IMG database were further analyzed and categorized based on species and isolation source (Fig. 4B). It was observed that all *protophormiae* species, including NG4, NG8, R912, JCM1973, and DSM20168, possessed both cholesterol oxidase and the androstenedione degradation pathway. These strains were predominantly isolated from environmental settings such as soil and water, with NG4, NG8, and R912 specifically identified from these sources, while information for JCM1973 was unavailable, and DSM20168 was isolated from flies. Similarly, both *mysorens* species DSM 12798 and NBRC103060, isolated from sewage, exhibited the androstenedione degradation pathway. Additionally, *soli* NHPC-3, isolated from rhizosphere soil, possessed genes for the androstenedione degradation pathway along with cholesterol oxidase, while *nicotianae* ZM05 and *arilaitensis* LJ1H9, both isolated from soil, contained only the androstenedione degradation pathway. Conversely, MNS18, isolated from a fermentation bed, possessed cholesterol oxidase alone. The remaining *Glutamicibacter* strains did not demonstrate the potential for steroid degradation, irrespective of the isolation environment.

Previously, a comparative genomic study analyzed steroid-degrading genes among 8000 microorganisms using the hidden Markov model [32]. Among the strains investigated in our study, only one *Arthrobacter* genus (species: *gangotriensis*, strain: Lz1y) was identified as a potential candidate by their study, although this strain has since been reclassified into *Paeniglutamicibacter*. Our study is the first to present the potential role of *Glutamicibacter* in steroid degradation. Their study also revealed that out of the 8000 microorganisms analyzed, only 265 putative steroid-degrading strains were identified, underscoring the fewer number of strains involved in steroid degradation.

G.protophormiae harbors the androstenedione degradation pathway, an intermediate in the degradation of epiandrosterone. The

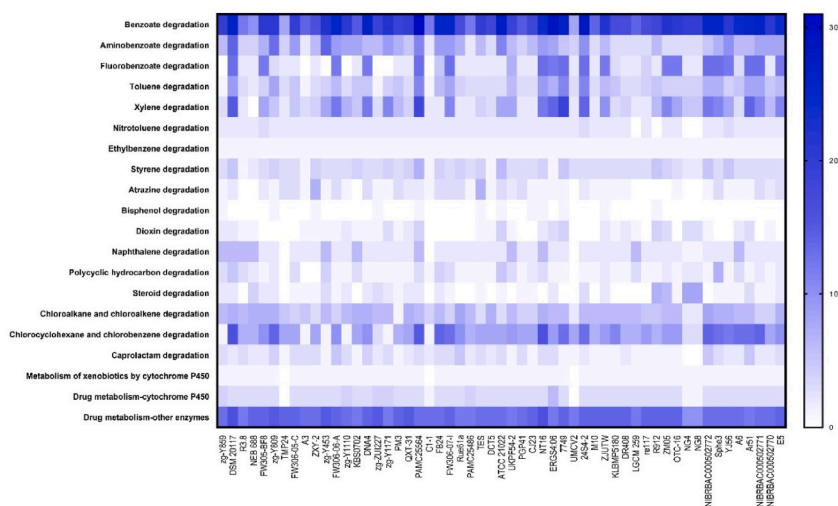


Fig. 3. Heat-map representation of the distribution of xenobiotics degradation related genes from KEGG pathway among all the 55 strains under this study. The number of genes involved in the pathway map for each category and the same gene acting in different reactions within the same pathway were taken separately.

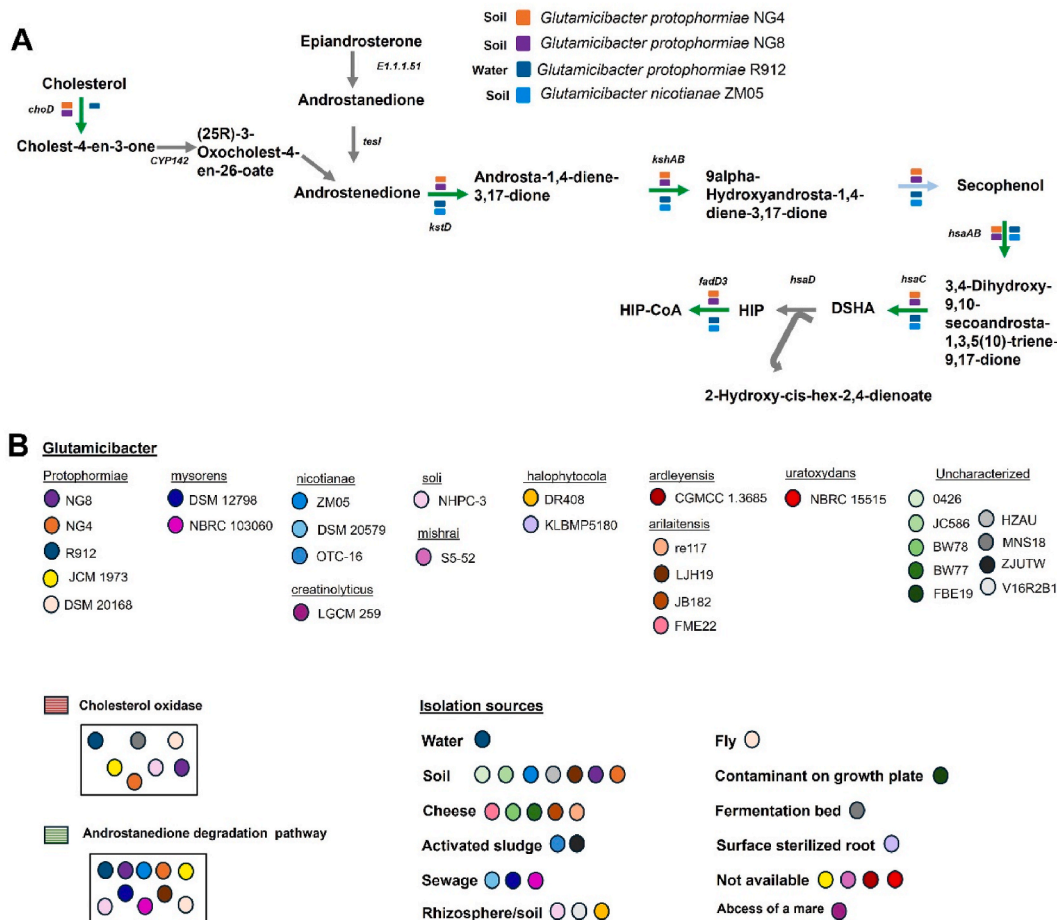


Fig. 4. Cholesterol oxidase and androstenedione degradation related genes distributed among *Glutamicibacter* strains isolated from diverse environments. (A) Steroid degradation pathway as shown by KEGG map for 4 strains that showed potential for steroid degradation among 55 selected strains for analysis. The green arrow represents the presence of the gene; blue color represent spontaneous reaction while grey color represents the absence. (B) Analysis of the steroid degradation potential among all the strains of *Glutamicibacter* available in JGI IMG database and their categorization based on isolation environments.

consistent presence of this pathway among all *protophormiae* species suggests its importance throughout their evolutionary history. *G. protophormiae* was initially isolated from a fly named *Protophormia terraenovae*, from which the species name was derived for the genus [15,20]. While it was initially expected that the distribution of the pathway among *Glutamicibacter* species might be specific to certain species, it was also observed in single strains of the species *nicotianae* and *arilaitensis*. Furthermore, a correlation between the pathway and the isolation environment was analyzed among putative steroid-degrading strains, revealing that species falling into this category were predominantly related to environmental isolates, particularly water, sewage, and soil, although some host-associated strains were also included.

2.5. Benzoate degradation pathway among the environmental isolates

Generally, benzoate and 4-HB are categorized as phenolic acids, commonly found in lignin-derived compounds [33,34], or as components of plant autotoxins [35,36] deposited in the rhizosphere by plants. However, due to their introduction through industrial processes, synthetic production, or use as preservatives in food, cosmetics, and pharmaceuticals, these compounds resemble xenobiotic compounds specially found as wastewater contaminants [37–39]. The presence of degradation pathways in an organism and their correlation with the isolation environment was observed for putative benzoate-degrading strains. The benzoate degradation category was highly prevalent among almost all the 55 strains, with few exceptions (Fig. 3). This category encompasses two major pathways: degradation of benzoate itself and 4-HB.

Most strains demonstrated complete pathways for the degradation of either benzoate, 4-HB, or both, leading to the citrate cycle (TCA cycle) intermediates. These compounds have multiple degradation routes (Fig. 5A, Table 1). 4-HB degradation route followed by beta oxidation (No. 1 and 5, Fig. 5A and B) was most abundant among the species. Interestingly, only *Arthrobacter* sp. ERGS4:06 possessed *ligAB*-mediated degradation route (No. 1 and 3, Fig. 5A and B) while no species showed the degradation pathway through

Table 1

Predicted genes involved in xenobiotics degradation based on KEGG pathway analysis (Figs. 4–7).

KEGG xenobiotic degradation pathway/ KEGG symbol	Strain name	KEGG protein name/KO	Reference gene/ Uniport or PDB accession	Identity/query cover (%)
<u>Steroid</u>	<i>Glutamicibacter protophormiae</i> NG4	cholesterol oxidase/K03333	Nr-database/ TDW30780.1	cholesterol oxidase (55.68/97)
choD				
kstD		3-oxosteroid 1-dehydrogenase/K05898	Swissprot/Q04616.3	3-oxosteroid 1-dehydrogenase/ (41.92/97)
kshA/kshB		3-ketosteroid 9alpha-monooxygenase subunit A and B/K15982 and K15983	Swissprot/B6V6V5.1 Swissprot/AOR525.1	3-ketosteroid 9alpha-monooxyge- nase subunit A/(69.40/92) 3-ketosteroid 9alpha-monooxyge- nase subunit B/(59.01/96)
hsaA		3-hydroxy-9,10-secoandrosta-1,3,5 (10)-triene-9,17-dione monooxygenase/K16047	Nr-database/ WP_204676060.1	Hydroxylase/(100/100)
hsaC		3,4-dihydroxy-9,10-secoandrosta-1,3,5 (10)-triene-9,17-dione 4,5- dioxygenase/K16049	Nr-database/ ALD65327.1	Extradiol dioxygenase/(96.44/100)
fadD3		HIP—CoA ligase/K18687	Swissprot// Q0S7V5.1	HIP—CoA ligase/(49.16/97)
<u>Benzoate</u>	<i>Arthrobacter</i> sp. 7749			
benA, benB and benC		benzoate 1,2-dioxygenase alpha, beta and reductase subunit/K05549, K05550, K05784	Swissprot/P07769.2 Swissprot/P07770.1 Swissprot/P23101.1	benzoate 1,2-dioxygenase alpha/ (67.89/93) benzoate 1,2-dioxygenase beta (63.98/89) Toluate 1,2-dioxygenase electron transfer component/(54.49/64)
benD		dihydroxycyclohexadiene carboxylate dehydrogenase/K05783	Swissprot/P07772.2	1,6-dihydroxycyclohexa-2,4-diene- 1-carboxylate dehydrogenase/ (59.92/92)
catA		catechol 1,2-dioxygenase/K03381	Swissprot/P95607.1	catechol 1,2-dioxygenase/(56.99/ 95)
catC		muconolactone delta-isomerase/ K03464	Swissprot/O33947.1	Muconolactone Delta-isomerase (59.78/100)
pcaD		3-oxoadipate enol-lactonase/K01055	Swissprot/P00632.3	3-oxoadipate enol-lactonase/ (35.29/87)
pcaJ		3-oxoadipate CoA-transferase alpha and beta subunit/K01031 and K01032	Swissprot/Q01103.2 Swissprot/P0A101.2	3-oxoadipate CoA-transferase alpha/(65.26/95) 3-oxoadipate CoA-transferase beta (59.15/96)
fadA, fadI catE		acetyl-CoA acyltransferase/K00632 catechol 2,3-dioxygenase/K07104	Swissprot/Q5E8X7.2 Swissprot/P54721.2	3-ketoacyl-CoA thiolase (42.35/94) catechol 2,3-dioxygenase/(34.93/ 91)
dmpD, xylF		2-hydroxymuconate-semialdehyde hydrolase/K10216	Swissprot/P19076.1	2-hydroxymuconate-semialdehyde hydrolase/(40.38/89)
mhpD		2-keto-4-pentenoate hydratase/K02554	Swissprot/Q13VU0.1	2-keto-4-pentenoate hydratase (41.41/97)
mhpE		4-hydroxy 2-oxovalerate aldolase/ K01666	Swissprot/Q0RXC3.1	4-hydroxy 2-oxovalerate aldolase (65.45/96)
mhpF		acetaldehyde dehydrogenase/K04073	Swissprot/Q0S4F9.1	acetaldehyde dehydrogenase (59.80/96)
<u>4-hydroxybenzoate</u>	<i>Glutamicibacter protophormiae</i> NG4			
pobA		p-hydroxybenzoate 3-monooxygenase/ K00481	Swissprot/8WEQ_A	p-hydroxybenzoate 3- monooxygenase/(63.36/98)
pcaG and pcaH		protocatechuate 3,4-dioxygenase, alpha and beta subunit/K00448 and K00449	Swissprot/P20371.3 Swissprot/P15110.1	protocatechuate 3,4-dioxygenase, alpha subunit/(41.50/93) protocatechuate 3,4-dioxygenase, beta subunit/(56.81/76)
pcaB		3-carboxy-cis,cis-muconate cycloisomerase/K01857	Swissprot/Q9I6Q8.1	3-carboxy-cis,cis-muconate cycloisomerase/(42.27/92)
pcaC		4-carboxymuconolactone decarboxylase/K01607	Swissprot/P20370.2	4-carboxymuconolactone decarboxylase/(44.53/92)
pcaD		3-oxoadipate enol-lactonase/K01055	Swissprot/P00632.3	3-oxoadipate enol-lactonase/(35/ 90)
pcaJ		3-oxoadipate CoA-transferase alpha and beta subunit/K01031 and K01032	Swissprot/Q01103.2 Swissprot/P0A101.2	3-oxoadipate CoA-transferase alpha subunit/(67.61/95) 3-oxoadipate CoA-transferase beta subunit/(64/92)
fadA, fadI		acetyl-CoA acyltransferase/K00632	Swissprot/	

(continued on next page)

Table 1 (continued)

KEGG xenobiotic degradation pathway/ KEGG symbol	Strain name	KEGG protein name/KO	Reference gene/ Uniport or PDB accession	Identity/query cover (%)
ligA and ligB	<i>Arthrobacter alpinus</i> ERGS4:06	protocatechuate 4,5-dioxygenase alpha and beta subunit/ K04100 and K04101	Swissprot/P22635.1 Swissprot/P22636.1	protocatechuate 4,5-dioxygenase alpha/(56.91/93) protocatechuate 4,5-dioxygenase beta/(60.21/100)
ligC		4-carboxy-2-hydroxymuconate semialdehyde dehydrogenase/ K10219	Swissprot/Q9KWL3.1	4-carboxy-2-hydroxymuconate semialdehyde dehydrogenase (67.41/99)
ligI		2-pyrone-4,6-dicarboxylate hydrolase/ K10221	Swissprot/Q93PS7.1	2-pyrone-4,6-dicarboxylate hydrolase/(63.46/99)
galD		4-oxalomesaconate tautomerase/ K16514	Swissprot/Q88JY0.1	4-oxalomesaconate tautomerase/(57.94/98)
ligJ		4-oxalomesaconate hydratase/ K10220	Swissprot/G2IQQ5.1	2-keto-4-carboxy-3-hexenedioate hydratase/(75.52/100)
ligK, galC		4-carboxy-4-hydroxy-2-oxoadipate aldolase/ K10218	Swissprot/G2IQQ8.1	4-carboxy-4-hydroxy-2-oxoadipate aldolase/(59.01/97)
Toluene E1.1.1.90	<i>Arthrobacter</i> sp. 7749	aryl-alcohol dehydrogenase/ K00055	Swissprot/P39849.1	aryl-alcohol dehydrogenase/(50.14/99)
xyIC		benzaldehyde dehydrogenase (NAD+)/ K00141	Swissprot/P43503.1	benzaldehyde dehydrogenase (NAD+)/(49.69/99)
benA-xyIX, benB-xyIY, and benC-xyIZ		benzoate 1,2-dioxygenase alpha, beta and reductase subunit/ K05549 , K05550 , K05784	Swissprot/P07769.2 Swissprot/P07770.1 Swissprot/P23101.1	benzoate 1,2-dioxygenase alpha/(67.89/93) benzoate 1,2-dioxygenase beta (63.98/89) Toluate 1,2-dioxygenase electron transfer component/(54.49/64)
benD-xyIL		dihydroxycyclohexadiene carboxylate dehydrogenase/ K05783	Swissprot/P07772.2	1,6-dihydroxycyclohexa-2,4-diene-1-carboxylate dehydrogenase/(59.92/92)
catA-		catechol 1,2-dioxygenase/ K03381	Swissprot/P95607.1	catechol 1,2-dioxygenase/(56.99/95)
catC		muconolactone delta-isomerase/ K03464	Swissprot/O33947.1	Muconolactone Delta-isomerase (59.78/100)
pcaD		3-oxoadipate enol-lactonase/ K01055	Swissprot/P00632.3	3-oxoadipate enol-lactonase/(35.29/87)
pcaIJ		3-oxoadipate CoA-transferase alpha and beta subunit/ K01031 and K01032	Swissprot/Q01103.2 Swissprot/P0A101.2	3-oxoadipate CoA-transferase alpha/(65.26/95) 3-oxoadipate CoA-transferase beta (59.15/96)
fadA, fadI catE, dmpB-xyIE		acetyl-CoA acyltransferase/ K00632 catechol 2,3-dioxygenase/ K07104	Swissprot/Q5E8X7.2 Swissprot/P54721.2	3-ketoacyl-CoA thiolase (42.35/94) catechol 2,3-dioxygenase/(34.93/91)
dmpD, xyIF		2-hydroxymuconate-semialdehyde hydrolase/ K10216	Swissprot/P19076.1	2-hydroxymuconate-semialdehyde hydrolase/(40.38/89)
mhpD		2-keto-4-pentenoate hydratase/ K02554	Swissprot/Q13VU0.1	2-keto-4-pentenoate hydratase (41.41/97)
mhpE		4-hydroxy 2-oxovalerate aldolase/ K01666	Swissprot/Q0RXC3.1	4-hydroxy 2-oxovalerate aldolase (65.45/96)
mhpF		acetaldehyde dehydrogenase/ K04073	Swissprot/Q0S4F9.1	acetaldehyde dehydrogenase (59.80/96)
p-Xylene E1.1.1.90	<i>Arthrobacter</i> sp. 7749	aryl-alcohol dehydrogenase/ K00055	Swissprot/P39849.1	aryl-alcohol dehydrogenase/(50.14/99)
xyIC		benzaldehyde dehydrogenase (NAD+)/ K00141	Swissprot/P43503.1	benzaldehyde dehydrogenase (NAD+)/(49.69/99)
benA-xyIX, benB-xyIY, and benC-xyIZ		Benzoate/toluate 1,2-dioxygenase alpha, beta and reductase subunit/ K05549 , K05550 , K05784	Swissprot/P07769.2 Swissprot/P07770.1 Swissprot/P23101.1	benzoate 1,2-dioxygenase alpha/(67.89/93) benzoate 1,2-dioxygenase beta (63.98/89) Toluate 1,2-dioxygenase electron transfer component/(54.49/64)
benD-xyIL		dihydroxycyclohexadiene carboxylate dehydrogenase/ K05783	Swissprot/P07772.2	1,6-dihydroxycyclohexa-2,4-diene-1-carboxylate dehydrogenase/(59.92/92)
dmpB, xyIE		3,4-dihydroxyphenylacetate 2,3-dioxygenase/ K00446	Pdb/1Q0C_A	homoprotocatechuate 2,3-dioxygenase/(86.26/98)

(continued on next page)

Table 1 (continued)

KEGG xenobiotic degradation pathway/ KEGG symbol	Strain name	KEGG protein name/KO	Reference gene/ Uniport or PDB accession	Identity/query cover (%)
dmpD, xylF		2-hydroxymuconate-semialdehyde hydrolase/ K10216	Swissprot/P19076.1	2-hydroxymuconate-semialdehyde hydrolase/(40.38/89)
<u>o-Xylene/m-Xylene</u> E1.1.1.90	<i>Arthrobacter</i> sp. 7749	aryl-alcohol dehydrogenase/ K00055	Swissprot/P39849.1	aryl-alcohol dehydrogenase/ (50.14/99)
xylC		benzaldehyde dehydrogenase (NAD+)/ K00141	Swissprot/P43503.1	benzaldehyde dehydrogenase (NAD+)/(49.69/99)
benA-xylX, benB-xylY, and benC-xylZ		Benzoate/toluate 1,2-dioxygenase alpha, beta and reductase subunit/ K05549 , K05550 , K05784	Swissprot/P07769.2 Swissprot/P07770.1 Swissprot/P23101.1	benzoate 1,2-dioxygenase alpha/ (67.89/93) benzoate 1,2-dioxygenase beta (63.98/89) Toluate 1,2-dioxygenase electron transfer component/(54.49/64)
benD-xylL		dihydroxycyclohexadiene carboxylate dehydrogenase/ K05783	Swissprot/P07772.2	1,6-dihydroxycyclohexa-2,4-diene- 1-carboxylate dehydrogenase/ (59.92/92)
dmpB, xylE		3,4-dihydroxyphenylacetate 2,3- dioxygenase/ K00446	Pdb/1Q0C_A	homoprotocatechuate 2,3- dioxygenase/(86.26/98)
mhpD		2-keto-4-pentenoate hydratase/ K02554	Swissprot/Q13VU0.1	2-keto-4-pentenoate hydratase (41.41/97)
mhpE		4-hydroxy 2-oxovalerate aldolase/ K01666	Swissprot/Q0RXC.3.1	4-hydroxy 2-oxovalerate aldolase (65.45/96)
mhpF		acetaldehyde dehydrogenase/ K04073	Swissprot/Q0S4F9.1	acetaldehyde dehydrogenase (59.80/96)
<u>Phenanthrene</u>	<i>Pseudarthrobacter</i>			
nidA	<i>phenanthrenivorans</i> Sphe3	ring-hydroxylating dioxygenase, large terminal subunit/ K11943	Pdb/2B1X_A	naphthalene dioxygenase large subunit/(74.41/96)
phdI		1-hydroxy-2-naphthoate 1,2- dioxygenase/ K11948	Swissprot/O24721.2	1-hydroxy-2-naphthoate 1,2- dioxygenase/(85.53/100)
phdJ		4-(2-carboxyphenyl)-2-oxobut-3- enoate aldolase/ K11949	Swissprot/Q79EM8.1	Trans-2'-carboxybenzalpyruvate hydratase-aldolase (94.28/96)
phdK		2-formylbenzoate dehydrogenase/ K18275	Swissprot/Q79EM7.1	2-formylbenzoate dehydrogenase/ (73.54/99)
phtAb		phthalate 3,4-dioxygenase subunit beta/ K18252	Nr-database/ MCU1563476.1	phthalate 3,4-dioxygenase subunit beta/(89.05/100)
phtC		3,4-dihydroxyphthalate decarboxylase/ K18256	Nr-database/ MCW2683780.1	3,4-dihydroxyphthalate decarboxylase (61.25/99)
pcaG and pcaH		protocatechuate 3,4-dioxygenase, alpha and beta subunit/ K00448 and K00449	Swissprot/P00436.3 Swissprot/P15110.1	protocatechuate 3,4-dioxygenase alpha chain (39.90/97) protocatechuate 3,4-dioxygenase beta chain (55.35/73)
<u>Azathioprene</u> (prodrug)	<i>Arthrobacter</i> <i>crystallopoietes</i> DSM 20117			
hprT, hpt, HPRT1		hypoxanthine phosphoribosyltransferase/ K00760	Swissprot/Q839B2.1	hypoxanthine phosphoribosyltransferase/(52.81/ 97)
IMPDH, guaB		inosine-5'-monophosphate dehydrogenase/ K00088	Swissprot/Q8G3N6.1	inosine-5'-monophosphate dehydrogenase/(65.65/98)
guaA, GMPS		GMP synthase (glutamine- hydrolysing)/ K01951	Nr-database/ WP_074702029.1	glutamine amidotransferas (100/ 100)
<u>Capecitabine</u> (Prodrug)	<i>Arthrobacter</i> <i>crystallopoietes</i> DSM 20117			
cdd, CDA		cytidine deaminase/ K01489	Swissprot/P9WPH2.1	cytidine deaminase/(56.35/75)
deoA, TYMP		thymidine phosphorylase/ K00758	Swissprot/P9WFS0.1	thymidine phosphorylase/(62.59/ 98)
tdk, TK		thymidine kinase/ K00857	Swissprot/ Q5QWC7.1	thymidine kinase/(45.45/98)
ndk, NME		nucleoside diphosphate kinase/ K00940	Swissprot/P36010.1	nucleoside diphosphate kinase (61.19/97)
DPYS, dht, hyDA UPB1, pydC		Dihydropyrimidinase/ K01464 N-carbamoylputrescine amidase/ K01431	Swissprot/Q14117.1 Swissprot/Q9XGI9.1	Dihydropyrimidinase (32.46/93) N-carbamoylputrescine amidase/ (40.78/96)
5-Fluorouridine udk, UCK	<i>Arthrobacter</i> <i>crystallopoietes</i> DSM 20117	uridine kinase/ K00876	Nr-database/ WP_074702196.1	uridine kinase/(100/100)
ndk, NME		nucleoside diphosphate kinase/ K00940	Swissprot/P36010.1	nucleoside diphosphate kinase/ (61.19/97)

routes 2 and 4 (Fig. 5). Among the 55 studied strains, 9 *Arthrobacter* species (R3.8, NEB 688, TMP24, A3, ZXY-2, C1-1, PAMC25486, DCT5, and UMCV2) and 1 *Pseudarthrobacter* sp. NIBRBAC000502772 did not possess any routes for 4-HB degradation, whereas all *Glutamicibacter* strains contained at least one pathway for 4-HB degradation.

In contrast, the benzoate degradation pathway, either through the *catA* or *dmpB*-mediated route, was found in very few species (No. 6–8, Fig. 5A). 10 *Arthrobacter* species (24S4-2, 7749, ERGS4:06, NT16, FW306-07-I, PAMC25564, DNA4, FW306-06-A, FW305-BF8, and DSM 20117), 6 *Pseudarthrobacter* species (E5, NIBRBAC000502771, Ar51, YJ56, Sphe3, and NIBRBAC000502772) and 3 *Glutamicibacter* species (OTC-16, ZM05, and ZJUTW) possessed *catA* mediated degradation route (No. 6 and 7, Fig. 5B). Among these strains having at least one route for benzoate degradation, some *Arthrobacter* (7749, ERGS4:06, and PAMC25564), *Pseudarthrobacter* (Ar51) and *Glutamicibacter* (ZM05) contained both the routes for benzoate degradation (No. 6–8, Fig. 5B).

When comparing the isolation sites of strains possessing the benzoate and 4-HB degradation pathways, it was observed that most strains involved in benzoate degradation were isolated from soil/sediment, followed by pollutant sites, while very few strains were isolated from other environments (Fig. 5C and Supplementary Table S2). Similarly, strains possessing the 4-HB degradation pathway were primarily isolated from soil/sediment, followed by host-associated and pollutant sites (Fig. 5D). Notably, none of the host-

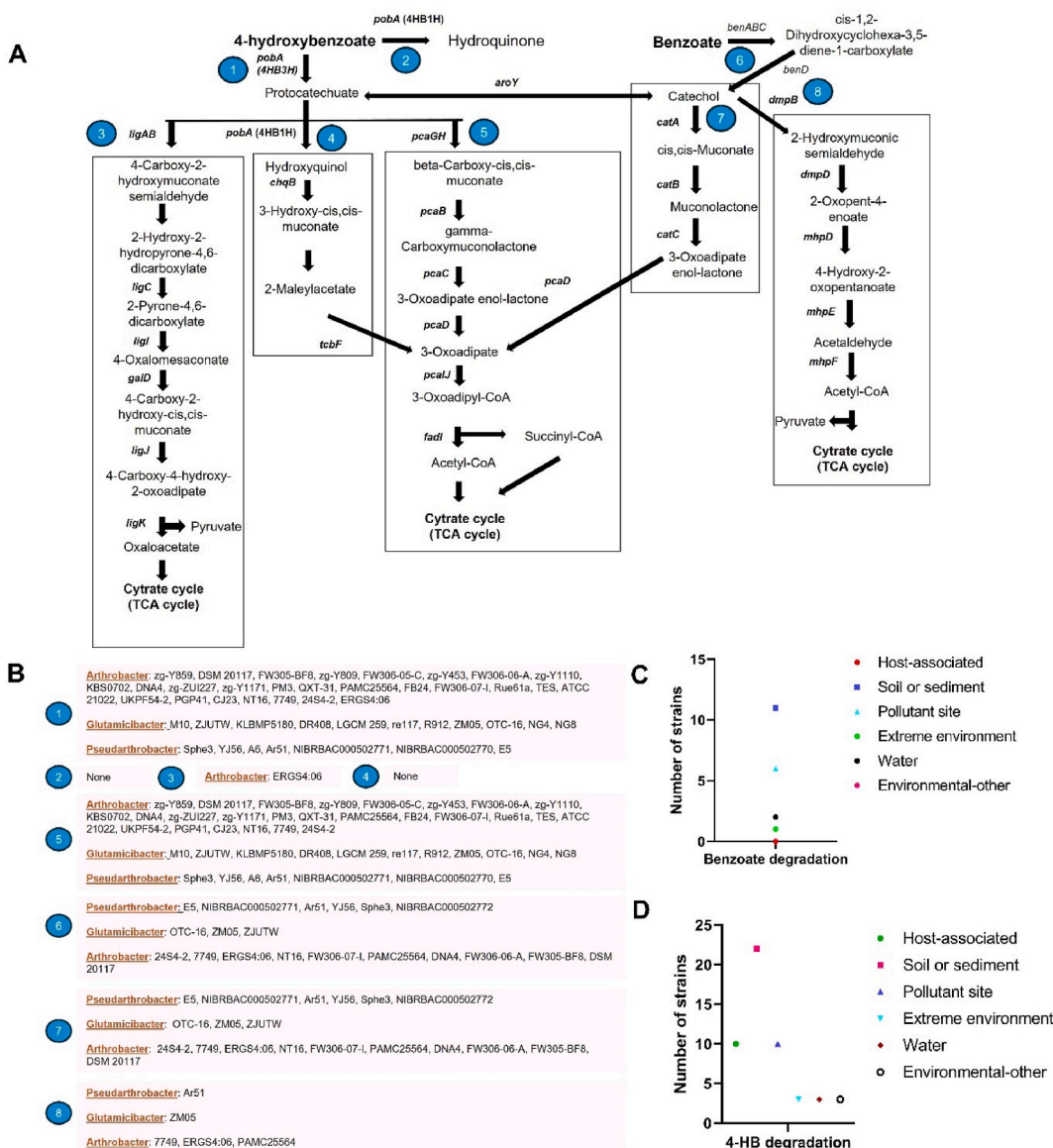


Fig. 5. Benzoate and 4-HB degradation pathway widely distributed among the three genera. (A) Different routes of benzoate and 4-HB degradation pathway as shown by KEGG map. (B) Categorization of the presence of each route for benzoate and 4-HB degradation among all the 55 selected strains. (C and D) Graph representation of the correlation between degradation pathway and the isolation source among the strains having the potential genes for benzoate (C) and 4-HB (D) degradation.

associated strains exhibited the benzoate degradation pathway, suggesting that this pathway is predominantly distributed among environmental isolates. While both the aerobic degradation pathways for benzoate and 4-HB are widely distributed among all three genera, the 4-HB degradation pathway was nearly ubiquitous across species, with few exceptions, whereas benzoate was only present in certain environmental isolates. A study showed the comparatively higher possibility for 4-HB deposition in soil as a result of its persistence and higher secretion potential [36]. This characteristic may render 4-HB more readily available for bacterial degradation, thus leading to the evolution and maintenance of 4-HB degradation pathways across a wider range of bacterial species.

Moreover, both benzoate and 4-HB serve as peripheral intermediates during the degradation of more toxic aromatic pollutants contaminating the environment. Environmental bacteria are exposed to a wide range of aromatic pollutants in their natural habitats, which may have driven the evolution of metabolic pathways for utilizing benzoate and 4-HB as carbon and energy sources. Benzoate typically forms as a peripheral intermediate during the degradation of toxic aromatic compounds such as biphenyl, xylene, toluene, and pyrene, while 4-HB is formed during the degradation of pyrene, p-cresol, and toluene [40–43]. Overall, both compounds play crucial roles in the bioremediation of toxic pollutants from the environment.

2.6. Some *Arthrobacter* and *Pseudarthrobacter* possess potential genes to degrade toxic aromatic pollutants

Among all the strains, only a few species were found to harbor the necessary set of genes for degrading toxic aromatic compounds such as toluene, xylene, and phenanthrene (Fig. 6, Table 1). The toluene degradation route through benzoate as a peripheral intermediate and catechol as the central intermediate was identified in three *Arthrobacter* strains, although the first enzyme was not annotated (Fig. 6A). Interestingly all these isolates were obtained from extreme environments: PAMC25564 was isolated from cryoconite in the Ötztal Alps, 24S4-2 from soil in Antarctica, and 7749 from marine sediment in Antarctica. Similarly, most of the *o*-xylene and *m*-xylene degradation genes were abundant in the strains PAMC25564 and 7749 (Fig. 6B). *Pseudarthrobacter phenanthrenivorans* Sphe3 isolated from creosote polluted soil in Greece possesses the initial gene for phenanthrene degradation. Additionally, it contains the peripheral pathway intermediate phthalate and the central pathway intermediate protocatechuate, facilitating complete degradation and utilization of phenanthrene as a carbon source by the degrading strain (Fig. 6C).

Certain strains, particularly from the *Arthrobacter* and *Pseudarthrobacter* genera, harbored putative toluene, xylene, and

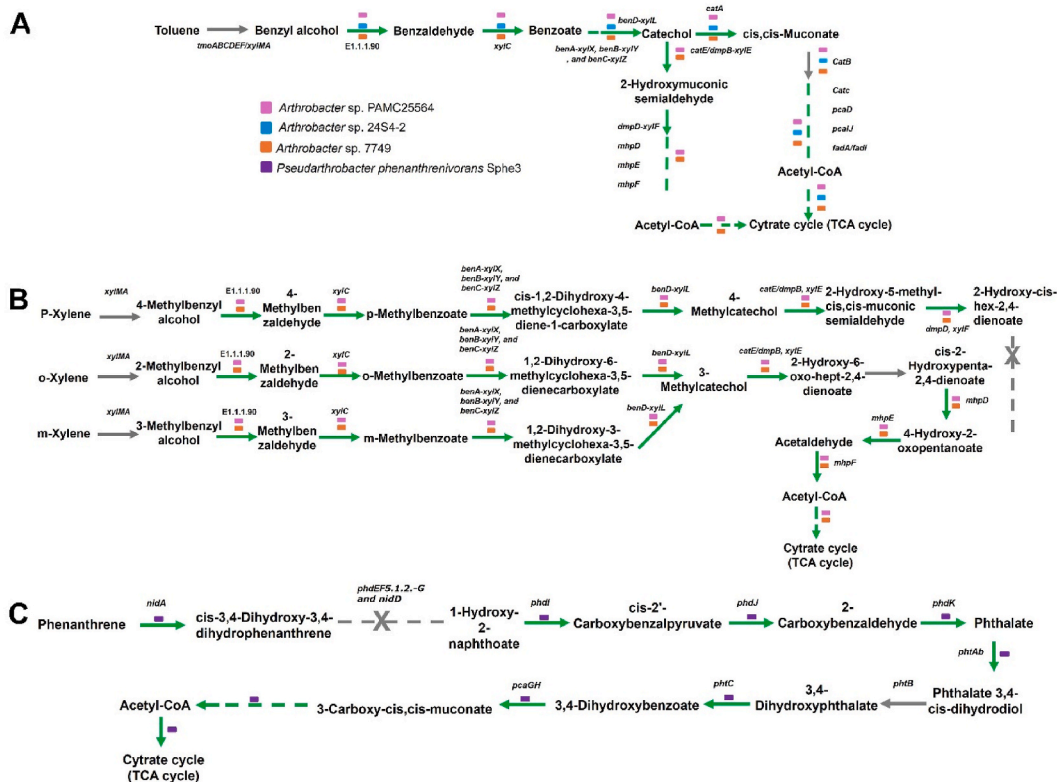


Fig. 6. Distribution of the toxic aromatic compounds degrading genes among the 55 strains. The green color represents the presence of the gene while grey color represents the absence. (A) KEGG pathway map of toluene degradation and the distribution of the genes among three *Arthrobacter* species. (B) KEGG pathway map for xylene degradation distributed among the two *Arthrobacter* strains. (C) KEGG pathway map for phenanthrene degradation in the *Pseudarthrobacter phenanthrenivorans* Sphe3. The strains possessing the maximum set of genes for the degradation pathway among all the strains are selected for the representation.

phenanthrene degradation genes. While many other species exhibited a decent distribution of such genes, we focused on identifying the most promising candidates with the maximum availability of genes for further study. *Arthrobacter* species such as PAMC25564, 24S4-2, and 7749 were found to possess the toluene degradation pathway, which involves benzyl alcohol-mediated degradation [44], leading to the peripheral intermediate benzoate and ultimately to the TCA cycle intermediate. Although direct reports on toluene biotransformation by *Arthrobacter*, *Pseudarthrobacter*, and *Glutamicibacter* species are lacking, the heterologous expression of phenol hydroxylase from *Arthrobacter* has demonstrated toluene degradation [45].

On the other hand, efficient degradation of *p*-xylene and phenanthrene has been reported by *Arthrobacter* sp. YC-RL1 [10]. However, the detection of pathways for xylene degradation in strains PAMC25564 and 7749 suggests the potential for *o*-xylene and *m*-xylene degradation, highlighting the significance of these strains for this purpose. Similarly, the putative phenanthrene-degrading strain *Pseudarthrobacter phenanthrenivorans* Sphe3 has been previously characterized for phenanthrene degradation by other research groups [46,47]. The strains identified in this study as potential candidates for degrading toxic pollutants could serve as valuable resources for bioremediation efforts targeting toxic aromatic pollutants.

2.7. Some drugs metabolic pathways that are widely distributed among all three genera

Interestingly, the drug metabolic pathway was prevalent among all the studied genera and strains. Three types of drug metabolic pathways were identified, including azathioprine, capecitabine, and 5-fluorouridine (Fig. 7A, Table 1). Among them, the azathioprine degradation pathway was present in all studied strains (Fig. 7B–D). Capecitabine degradation-related genes were also detected in all *Arthrobacter* species, while some *Glutamicibacter* and *Pseudarthrobacter* species did not possess them. Conversely, genes related to 5-fluorouridine degradation were less common across all three genera.

Drug metabolism by human gut microbiome is the most studied field of research [48], but the drug metabolism by the



Fig. 7. Distribution of the drug metabolism-related genes among the 55 strains. (A) KEGG pathway representation for drug metabolism from the strain *Arthrobacter crystallopoietes* DSM 20117 possessing the highest number of genes in the category among 55 strains. The green color represents the presence of the gene, grey color represents the absence, blue color represents the non-enzymatic reaction, and the dashed black color represents the unknown pathway. (B) Categorization of the number of strains possessing degradation pathway genes among three types of drug metabolism pathway for *Arthrobacter*, *Pseudarthrobacter*, and *Glutamicibacter*.

environmental isolate is a rarely discussed phenomenon. However, bacterial enzymes from strains isolated from diverse environments are gaining attention for their potential role in drug metabolism. Notably, cytochrome P450 monooxygenase from environmental isolates have been widely studied for the metabolism of drugs like steroids [49], alkaloids [50], macrolides [51] and more. Our genome-scale study provided an overview of drug-metabolizing pathways distributed among host-specific and environmental isolates.

It is noteworthy that all the studied strains possess at least one pathway for metabolizing azathioprine, an immunosuppressant drug [52]. Additionally, they harbor genes involved in metabolizing drugs like capecitabine and 5-fluorouridine, which are commonly used in cancer therapy [53]. The widespread distribution of the drug metabolic pathway presents an alarming effect of anthropogenic contamination in the ecological system. The presence of these drug metabolism-related genes in environmental isolates underscores the importance of understanding their distribution and functional significance in environmental bacteria. This knowledge could provide valuable insights into microbial ecology and the potential for microbial-mediated transformations of pharmaceutical compounds in natural ecosystems.

3. Conclusions

This study focused on 55 strains from three different genera that were formerly grouped under the same taxonomic genus, *Arthrobacter*. Primarily, we conducted a comparative genomics study examining the xenobiotic degradation capabilities of these strains isolated from diverse environments. Through the comparative genomic process, we also determined and suggested revisions in the taxonomic classification of certain *Arthrobacter* species and suggested species delineation of a *Glutamicibacter* species, supported by pan/core phylogeny and ANI analysis, respectively. Additionally, our comparative analysis highlights the putative role of *Glutamicibacter* species isolated from diverse niches in steroid degradation. The widespread distribution of benzoate and 4-HB degradation pathways among all three genera, with 4-HB being the most common, underscores the environmental adaptation of microorganisms and the persistence of the compounds in the environment. Moreover, we identified candidate strains for the degradation of toxic aromatic compounds, emphasizing their potential role in bioremediation strategies. The correlation between the isolation source and the presence of specific xenobiotic degradation pathways is particularly evident in the case of benzoate and toxic pollutant degradation, which were exclusively identified among environmental isolates. Furthermore, our study sheds light on the rarely discussed phenomenon of drug metabolism by environmental isolates, revealing a widespread distribution of drug metabolizing genes and pathways among the studied strains suggesting their ecological implications and impact on environmental health. Overall, our comprehensive analysis provides insights into the relationship between three genera and their metabolic versatility in xenobiotic degradation. The presence of the xenobiotic degradation pathway in strains isolated from both contaminated and non-contaminated areas suggests the ubiquitous presence of xenobiotics in the environment, possibly through natural and anthropogenic activities. Finally, the study calls for experimental validations to explore the potential applications of these candidate strains in environmental bioremediation processes.

4. Materials and method

4.1. Sample collection

The soil sample was collected from the west coastal area in Seosan city, South Korea, which is characterized by the presence of chemical plants, including pharmaceutical and petrochemical industries. Using a shovel, the sample was collected from a depth of 10 cm and transferred to sterilized tubes, which were immediately sealed with caps to prevent contamination. The sample was then transported to the laboratory in an icebox and stored at -50°C until further experimentation. On-site measurements were taken to estimate the temperature of the seawater during collection, while pH measurements were conducted in the laboratory.

4.2. Isolation of bacteria and 16s rRNA analysis

1 g of soil sample was mixed with 100 mL of mineral salt medium (20.0 g NaCl, 7.0 g $\text{Na}_2\text{HPO}_4 \cdot 12\text{H}_2\text{O}$, 1 g KH_2PO_4 , 10 mg $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$, 1 mg FeCl_3 , 20 mg $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 1 g $(\text{NH}_4)_2\text{SO}_4$, and 1000 mL of distilled water; pH 7.0) containing several aromatic pollutants and PAHs like phenanthrene, pyrene, naphthalene, and biphenyl separately. The compounds were dissolved in acetone and 10 mg of each compound was placed in the sterilized flasks and the acetone was left to dry on the clean bench followed by the addition of 100 mL of MSB medium. The mixture was incubated for 7 days, and 10 mL of the culture was transferred to 90 mL of the media following the same protocol and incubated for 7 days more. Finally, serial dilution was performed using the culture and spread on marine agar medium, R2A and TSB medium subsequently and incubated at 25°C for a week. Distinct colonies were selected, and pure culture was obtained (NG4 from R2A media). The 16S rRNA gene was amplified using universal primers: 27F (5'-AGA GTT TGA TCM TGG CTC AG -3') and 1492R (5'-GGT TAC CTT GTT ACG ACT T -3'). The obtained sequence was matched and compared with other species using EzBioCloud database [54]. The 16S rRNA sequence was used to create a phylogenetic tree using MEGA X neighbor-joining tree with 500 bootstrap replicates value [55]. For the type strain prediction, the whole genome was uploaded to Type strain genome server (TYGS) [56] and the intergenomic distances was used to create a balanced minimum evolution tree with FASTME 2.1.4 with 100 pseudo-bootstrap replicates [57]. The average nucleotide identity was performed by pyani v 0.2 using ANiB method [58,59]. The pyani was used to perform ANiB analysis of 55 strains using their *fna* files, with BLAST+ and default parameters. The results were visualized using heatmap2 in R package.

4.3. Genome sequencing, assembly, and annotation of NG4 strain

Before the sequencing process, quantity and purity were determined using an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) to obtain complete genome sequencing. Genome sequencing was performed using PacBio RS II single-molecule real-time (SMRT) sequencing technology (Pacific Biosciences, Menlo Park, CA, USA). Library preparation and sequencing were conducted by DNALink, an authorized service provider located in Seoul, South Korea (DNA Link Inc.). SMRTbell library inserts (20 kb) were sequenced using SMRT cells. Using PacBio sequencing raw data, FALCON-Unzip (v0.5) was first used to conduct de novo assembly. In addition, Raw sequence data were generated, and de novo assembled using the hierarchical genome assembly process (HGAP) protocol and RS HGAP4 Assembly in SMRT Analysis Software (ver. 2.3; Pacific Biosciences, SMRT Link 4.0.0) [60,61]. Annotation was performed with the NCBI Prokaryotic Genome Annotation Pipeline (PGAP). Information about PGAP can be found here: https://www.ncbi.nlm.nih.gov/genome/annotation_prok/. We submitted NG4 strain to NCBI and received an accession number as CP139423.1. For genome annotation, the complete genomes of NG4 strain was subjected to RAST annotation server [62]. For predicting the pathways involved in xenobiotic degradation, the genomes of NG4 strain was subjected to BlastKOALA [63] using the KEGG database. The circular map of the genomic DNA was prepared using the CGView tool [64].

4.4. Selection of *Glutamicibacter*, *Arthrobacter*, and *Pseudarthrobacter* genomes for comparison

For comparison of *G. protophormiae* NG4 strain with other strains from the same genus along with *Arthrobacter* and *Pseudarthrobacter*, only the completely sequenced genomes were selected from joint genome institute (JGI) integrated microbial genomes (IMG) database [65]. Complete genomes are of high quality to draft genomes as they contain lesser gaps, less error and contain sequenced assemblies in high percentage unlike draft genome [66]. Further, the strains lacking the proper information about the isolation site and the genomes showing sequence error while processing genome analysis tools used in this study were discarded. Finally, 55 complete genomes were selected for further analysis. The selected genomes were isolated from diverse environments including the host-associated and environmental along with pollutant contaminated and non-contaminated areas. Detailed information of the selected strains with their isolation environment is available in [Supplementary Table S2](#).

4.5. Pan-genome analysis and core-phylogeny construction of the selected strains of *Glutamicibacter*, *Arthrobacter*, and *Pseudarthrobacter*

For pan-genome analysis, the fasta sequences of complete genomes involving 55 selected strains were subjected to BPGA software package [67]. USEARCH Clustering Algorithm was used with 70 % as sequence identity cut-off value. The core/pan phylogeny was performed as an automated downstream analysis and the phylogenetic tree was obtained using MUSCLE and neighbor-joining method [68].

4.6. Comparative analysis for xenobiotic degradation genes and pathways among the selected strains

For comparison of the xenobiotic degradation genes and pathways among the 55 selected genomes, the KEGG annotation results available in the JGI IMG database was used. JGI IMG database consists of all the information of protein coding genes associated with KEGG pathways, which were used for pathway mapping in the strains. For strains unavailable in the database, BlastKOALA tool [63] within KEGG database as described above was used by uploading the amino acid sequence of the strain. Protein coding genes connected to KEGG pathways were tracked by reconstructing the pathway. Further, the predicted genes were subjected to NCBI BLAST against database like UniProtKB/Swiss-Prot, Protein data bank (pdb) and non-redundant protein sequences (Nr-database) for further analysis. The priority database was in the order UniProtKB/Swiss-Prot- Protein data bank (pdb) and then Nr-database to support the protein coding genes predicted by KEGG. The xenobiotic degradation category in KEGG database was detailly analyzed and the information obtained was processed for comparative purposes.

CRediT authorship contribution statement

Nisha Ghimire: Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization. **Byeollee Kim:** Writing – original draft, Methodology, Formal analysis. **So-Ra Han:** Writing – original draft, Methodology, Formal analysis. **Tae-Jin Oh:** Writing – review & editing, Writing – original draft, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Data availability statement

The data generated in this study are available in the manuscript along with supplementary files.

Declaration of competing interest

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

Acknowledgement

This work was supported by the Technology Innovation Program (20018705, Development of masking and commercialization of biodegradable technology in an urban residential environment using rancid odor-reducing microorganisms and its fragrances) funded by the Ministry of Trade, Industry & Energy (MOTIE, Korea).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.heliyon.2024.e40280>.

References

- [1] J.C. Hunter-Cevera, The value of microbial diversity, *Curr. Opin. Microbiol.* 1 (1998) 278–285.
- [2] A.K. Priya, M. Muruganandam, A. Kumar, N. Senthilkumar, M. Shkir, B. Pandit, M. Imran, C. Prakash, M. Ubaidullah, Recent advances in microbial-assisted degradation and remediation of xenobiotic contaminants; challenges and future prospects, *J. Water Proc. Eng.* 60 (2024) 105106.
- [3] D.T. Marghade, A.D. Chahande, M.S. Tiwari, P.D. Patil, Microbial degradation of xenobiotic compounds. *Recent Advances in Microbial Degradation*, 2021, pp. 173–217.
- [4] M. Alexander, Biodegradation of chemicals of environmental concern, *Science* 211 (1981) 132–138.
- [5] J.R. Van Der Meer, W.M. De Vos, S. Harayama, A.J. Zehnder, Molecular mechanisms of genetic adaptation to xenobiotic compounds, *Microbiol. Rev.* 56 (1992) 677–694.
- [6] D. Springael, E.M. Top, Horizontal gene transfer and microbial adaptation to xenobiotics: new types of mobile genetic elements and lessons from ecological studies, *Trends Microbiol.* 12 (2004) 53–58.
- [7] M. Gil, A. Haidour, J.L. Ramos, Degradation of o-methoxybenzoate by a two-member consortium made up of a gram-positive *Arthrobacter* strain and a gram-negative *Pantoea* strain, *Biodegradation* 11 (2000) 49–53.
- [8] R. Margesin, P. Bergauer, S. Gander, Degradation of phenol and toxicity of phenolic compounds: a comparison of cold-tolerant *Arthrobacter* sp. and mesophilic *Pseudomonas putida*, *Extremophiles* 8 (2004) 201–207.
- [9] J.-S. Seo, Y.-S. Keum, Y. Hu, S.-E. Lee, Q.X. Li, Phenanthrene degradation in *Arthrobacter* sp. P1-1: initial 1, 2-, 3, 4-and 9, 10-dioxygenation, and meta-and ortho-cleavages of naphthalene-1, 2-diol after its formation from naphthalene-1, 2-dicarboxylic acid and hydroxyl naphthoic acids, *Chemosphere* 65 (2006) 2388–2394.
- [10] L. Ren, Y. Jia, R. Zhang, Z. Lin, Z. Zhen, H. Hu, Y. Yan, Insight into metabolic versatility of an aromatic compounds-degrading *Arthrobacter* sp. YC-RL1, *Front. Microbiol.* 9 (2018) 2438.
- [11] X. Wang, L. Sun, H. Wang, H. Wu, S. Chen, X. Zheng, Surfactant-enhanced bioremediation of DDTs and PAHs in contaminated farmland soil, *Environ. Technol.* 39 (2018) 1733–1744.
- [12] L.I. Qingyan, L.I. Ying, Z.H.U. Xikun, C.A.I. Baoli, Isolation and characterization of atrazine-degrading *Arthrobacter* sp. AD26 and use of this strain in bioremediation of contaminated soil, *J. Environ. Sci.* 20 (2008) 1226–1230.
- [13] C. Jiang, H. Yan, X. Shen, Y. Zhang, Y. Wang, S. Sun, H. Jiang, H. Zang, X. Zhao, N. Hou, Genome functional analysis of the Psychrotrophic lignin-degrading bacterium *Arthrobacter* sp. C2 and the role of DyP in catalyzing lignin degradation, *Front. Microbiol.* 13 (2022) 921549.
- [14] M. Kanehisa, S. Goto, KEGG: kyoto encyclopedia of genes and genomes, *Nucleic Acids Res.* 28 (2000) 27–30.
- [15] H.-J. Busse, Review of the taxonomy of the genus *Arthrobacter*, emendation of the genus *Arthrobacter sensu lato*, proposal to reclassify selected species of the genus *Arthrobacter* in the novel genera *Glutamicibacter* gen. nov., *Paeniglutamicibacter* gen. nov., *Pseudoglutamicibacter* gen. nov., *Paenarthrobacter* gen. nov. and *Pseudarthrobacter* gen. nov., and emended description of *Arthrobacter roseus*, *Int. J. Syst. Evol. Microbiol.* 66 (2016) 9–37.
- [16] J. Li, W. Peng, X. Yin, X. Wang, Z. Liu, Q. Liu, Z. Deng, S. Lin, R. Liang, Identification of an efficient phenanthrene-degrading *Pseudarthrobacter* sp. L1SW and characterization of its metabolites and catabolic pathway, *J. Hazard Mater.* 465 (2024) 133138.
- [17] J. Lamba, S. Anand, J. Dutta, S. Chatterjee, S. Nagar, S.M. Celin, P.K. Rai, Study on aerobic degradation of 2, 4, 6-trinitrotoluene (TNT) using *Pseudarthrobacter chlorophenolicus* collected from the contaminated site, *Environ. Monit. Assess.* 193 (2021) 1–11.
- [18] C. Ren, Y. Wang, Y. Wu, H.-P. Zhao, L. Li, Complete degradation of di-n-butyl phthalate by *Glutamicibacter* sp. strain 0426 with a novel pathway, *Biodegradation* 35 (2024) 87–99.
- [19] H. Li, H. Zhou, L. Fan, L. Meng, Y. Zhao, L. Zhao, B. Wang, *Glutamicibacter nicotianae* AT6: a new strain for the efficient biodegradation of tilimicosin, *J. Environ. Sci.* 142 (2024) 182–192.
- [20] O. Lysenko, The occurrence of species of the genus *Brevibacterium* in insects, *J. Insect Pathol.* 1 (1959) 34–42.
- [21] E. Stackebrandt, V.J. Fowler, F. Fiedler, H. Seiler, Taxonomic studies on *arthrobacter nicotianae* and related taxa: description of *arthrobacter uratoxydans* sp. nov. and *arthrobacter sulfureus* sp. nov. and reclassification of *Brevibacterium protophormiae* as *arthrobacter protophormiae* comb. Nov., *Syst. Appl. Microbiol.* 4 (1983) 470–486.
- [22] H. Liao, X. Lin, Y. Li, M. Qu, Y. Tian, Reclassification of the taxonomic framework of orders cellvibrionales, oceanospirillales, pseudomonadales, and alteromonadales in class gammaproteobacteria through phylogenomic tree analysis, *mSystems* 5 (2020) 10–1128.
- [23] C. Jain, L.M. Rodriguez-R, A.M. Phillippy, K.T. Konstantinidis, S. Aluru, High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries, *Nat. Commun.* 9 (2018) 5114.
- [24] J. Goris, K.T. Konstantinidis, J.A. Klappenbach, T. Coenye, P. Vandamme, J.M. Tiedje, DNA–DNA hybridization values and their relationship to whole-genome sequence similarities, *Int. J. Syst. Evol. Microbiol.* 57 (2007) 81–91.
- [25] K.T. Konstantinidis, J.M. Tiedje, Genomic insights that advance the species definition for prokaryotes, *Proc. Natl. Acad. Sci. USA* 102 (2005) 2567–2572.
- [26] Y.G. Ying GuangGuo, R.S. Kookana, R.Y. Ru YingJun, Occurrence and Fate of Hormone Steroids in the Environment, 2002.
- [27] Y.R. Chiang, S.T.S. Wei, P.H. Wang, P.H. Wu, C.P. Yu, Microbial degradation of steroid sex hormones: implications for environmental and ecological studies, *Microb. Biotechnol.* 13 (2020) 926–949.
- [28] W. Shi, L. Wang, D.P.L. Rousseau, P.N.L. Lens, Removal of estrone, 17 α -ethinylestradiol, and 17 β -estradiol in algae and duckweed-based wastewater treatment systems, *Environ. Sci. Pollut. Control Ser.* 17 (2010) 824–833.
- [29] L. Peng, X. Dai, Y. Liu, J. Sun, S. Song, B.-J. Ni, Model-based assessment of estrogen removal by nitrifying activated sludge, *Chemosphere* 197 (2018) 430–437.
- [30] M. Horinouchi, T. Kurita, T. Hayashi, T. Kudo, Steroid degradation genes in *Comamonas testosteroni* TA441: isolation of genes encoding a $\Delta 4$ (5)-isomerase and 3α - and 3β -dehydrogenases and evidence for a 100 kb steroid degradation gene hot spot, *J. Steroid Biochem. Mol. Biol.* 122 (2010) 253–263.
- [31] Q. Zhang, C. Xue, G. Owens, Z. Chen, Isolation and identification of 17 β -estradiol degrading bacteria and its degradation pathway, *J. Hazard Mater.* 423 (2022) 127185.
- [32] L.H. Bergstrand, E. Cardenas, J. Holert, J.D. Van Hamme, W.W. Mohn, Delineation of steroid-degrading microorganisms through comparative genomic analysis, *mBio* 7 (2016) 10–1128.

- [33] Z.Z.A. Kuzyk, M.A. Goñi, G.A. Stern, R.W. Macdonald, Sources, pathways and sinks of particulate organic matter in Hudson Bay: evidence from lignin distributions, *Mar. Chem.* 112 (2008) 215–229.
- [34] P. Lu, W. Wang, G. Zhang, W. Li, A. Jiang, M. Cao, X. Zhang, K. Xing, X. Peng, B. Yuan, Isolation and characterization marine bacteria capable of degrading lignin-derived compounds, *PLoS One* 15 (2020) e0240187.
- [35] C.N. He, W.W. Gao, J.X. Yang, W. Bi, X.S. Zhang, Y.J. Zhao, Identification of autotoxic compounds from fibrous roots of *Panax quinquefolium* L, *Plant Soil* 318 (2009) 63–72.
- [36] Y. Wang, C. Li, Q. Wang, H. Wang, B. Duan, G. Zhang, Environmental behaviors of phenolic acids dominated their rhizodeposition in boreal poplar plantation forest soils, *J. Soils Sediments* 16 (2016) 1858–1870.
- [37] S. Krishnani, Vineet, N.M. Tripathi, R. Kaur, G.K. Aseri, N. Khare, D. Sharma, D. Singh, Isolation and identification of benzoate degrading bacteria from food Industry effluent, *J. Water Chem. Technol.* 44 (2022) 191–197.
- [38] S. Wang, M. Bilal, H. Hu, W. Wang, X. Zhang, 4-Hydroxybenzoic acid—a versatile platform intermediate for value-added compounds, *Appl. Microbiol. Biotechnol.* 102 (2018) 3561–3571.
- [39] W. Wang, K. Kannan, Fate of parabens and their metabolites in two wastewater treatment plants in New York State, United States, *Environ. Sci. Technol.* 50 (2016) 1174–1181.
- [40] P.S. Phale, H. Malhotra, B.A. Shah, Degradation strategies and associated regulatory mechanisms/features for aromatic compound metabolism in bacteria, *Adv. Appl. Microbiol.* 112 (2020) 1–65.
- [41] A. Wongbunmak, S. Khaiwan, M. Suphantharika, T. Pongtharangkul, BTEX biodegradation by *Bacillus amyloliquefaciens* subsp. *plantarum* W1 and its proposed BTEX biodegradation pathways, *Sci. Rep.* 10 (2020) 17408.
- [42] Y.-C. Chang, K. Takada, D. Choi, T. Toyama, K. Sawada, S. Kikuchi, Isolation of biphenyl and polychlorinated biphenyl-degrading bacteria and their degradation pathway, *Appl. Biochem. Biotechnol.* 170 (2013) 381–398.
- [43] Y. Lyu, W. Zheng, T. Zheng, Y. Tian, Biodegradation of polycyclic aromatic hydrocarbons by *Novosphingobium pentaromativorans* US6-1, *PLoS One* 9 (2014) e101438.
- [44] M. Suzuki, T. Hayakawa, J.P. Shaw, M. Rekik, S. Harayama, Primary structure of xylene monooxygenase: similarities to and differences from the alkane hydroxylation system, *J. Bacteriol.* 173 (1991) 1690–1695.
- [45] F. Ma, S.-N. Shi, T.-H. Sun, A. Li, J.-T. Zhou, Y.-Y. Qu, Biotransformation of benzene and toluene to catechols by phenol hydroxylase from *Arthrobacter* sp, W1, *Applied microbiology and biotechnology* 97 (2013) 5097–5103.
- [46] A. Kallimanis, K. Kavakiotis, A. Perisynakis, C. Sproer, R. Pukall, C. Drains, A.I. Koukkou, *Arthrobacter phenanthrenivorans* sp. nov., to accommodate the phenanthrene-degrading bacterium *Arthrobacter* sp. strain Sphe3, *Int. J. Syst. Evol. Microbiol.* 59 (2009) 275–279.
- [47] E. Vandera, M. Samiotaki, M. Parapouli, G. Panayotou, A.I. Koukkou, Comparative proteomic analysis of *Arthrobacter phenanthrenivorans* Sphe3 on phenanthrene, phthalate and glucose, *J. Proteomics* 113 (2015) 73–89.
- [48] R.G. Nichols, J.M. Peters, A.D. Patterson, Interplay between the host, the human microbiome, and drug metabolism, *Hum. Genom.* 13 (2019) 1–10.
- [49] K.-H. Kim, H. Do, C.W. Lee, P. Subedi, M. Choi, Y. Nam, J.H. Lee, T.-J. Oh, Crystal structure and biochemical analysis of a cytochrome P450 steroid hydroxylase (BaCYP106A6) from *Bacillus* species, *J. Microbiol. Biotechnol.* 33 (2023) 387.
- [50] B.D. Pardhe, H. Do, C.S. Jeong, K.H. Kim, J.H. Lee, T.J. Oh, Characterization of high-H₂O₂-tolerant bacterial cytochrome P450 CYP105D18: insights into papaverine N-oxidation, *IUCr* 8 (2021) 684–694.
- [51] M.C. Betlach, J.T. Kealey, M.C. Betlach, G.W. Ashley, R. McDaniel, Characterization of the macrolide P-450 hydroxylase from *Streptomyces venezuelae* which converts narbomycin to picromycin, *Biochemistry* 37 (1998) 14937–14942.
- [52] A.J. Garcia-Fernández, Ecotoxicology, avian, *Encyclopedia of toxicology* 2 (2014) 289–294.
- [53] C.M. Walko, C. Lindley, Capecitabine: a review, *Clin. Therapeut.* 27 (2005) 23–44.
- [54] S.-H. Yoon, S.-M. Ha, S. Kwon, J. Lim, Y. Kim, H. Seo, J. Chun, Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies, *Int. J. Syst. Evol. Microbiol.* 67 (2017) 1613–1617.
- [55] S. Kumar, G. Stecher, M. Li, C. Knyaz, K. Tamura, Mega X: molecular evolutionary genetics analysis across computing platforms, *Mol. Biol. Evol.* 35 (2018) 1547.
- [56] J.P. Meier-Kolthoff, M. Göker, TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy, *Nat. Commun.* 10 (2019) 2182.
- [57] V. Lefort, R. Desper, O. Gascuel, FastME 2.0: a comprehensive, accurate, and fast distance-based phylogeny inference program, *Mol. Biol. Evol.* 32 (2015) 2798–2800.
- [58] M. Richter, R. Rosselló-Móra, Shifting the genomic gold standard for the prokaryotic species definition, *Proc. Natl. Acad. Sci. USA* 106 (2009) 19126–19131.
- [59] L. Pritchard, R.H. Glover, S. Humphris, J.G. Elphinstone, I.K. Toth, Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens, *Anal. Methods* 8 (2016) 12–24.
- [60] C.-S. Chin, D.H. Alexander, P. Marks, A.A. Klammer, J. Drake, C. Heiner, A. Clum, A. Copeland, J. Huddleston, E.E. Eichler, Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data, *Nat. Methods* 10 (2013) 563–569.
- [61] C.-S. Chin, P. Peluso, F.J. Sedlazeck, M. Nattestad, G.T. Concepcion, A. Clum, C. Dunn, R. O'Malley, R. Figueroa-Balderas, A. Morales-Cruz, Phased diploid genome assembly with single-molecule real-time sequencing, *Nat. Methods* 13 (2016) 1050–1054.
- [62] R. Overbeek, R. Olson, G.D. Pusch, G.J. Olsen, J.J. Davis, T. Disz, R.A. Edwards, S. Gerdes, B. Parrello, M. Shukla, The SEED and the Rapid annotation of microbial genomes using subsystems technology (RAST), *Nucleic Acids Res.* 42 (2014) D206–D214.
- [63] M. Kanehisa, Y. Sato, K. Morishima, BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences, *J. Mol. Biol.* 428 (2016) 726–731.
- [64] J.R. Grant, P. Stothard, The CGView Server: a comparative genomics tool for circular genomes, *Nucleic Acids Res.* 36 (2008) W181–W184.
- [65] I.M.A. Chen, K. Chu, K. Palaniappan, A. Ratner, J. Huang, M. Huntemann, P. Hajek, S. Ritter, N. Varghese, R. Seshadri, The IMG/M data management and analysis system v. 6.0: new tools and advanced capabilities, *Nucleic Acids Res.* 49 (2021) D751–D763.
- [66] L.E. Palmer, W.R. McCombie, On the importance of being finished, *Genome Biol.* 3 (2002) 1–4.
- [67] N.M. Chaudhari, V.K. Gupta, C. Dutta, BPGA—an ultra-fast pan-genome analysis pipeline, *Sci. Rep.* 6 (2016) 24373.
- [68] R.C. Edgar, MUSCLE: multiple sequence alignment with high accuracy and high throughput, *Nucleic Acids Res.* 32 (2004) 1792–1797.