Hungarian University of Agriculture and Life Sciences

Doctoral School of Environmental Sciences

AEROBIC AND MICROAEROBIC ENRICHMENT OF XYLENE DEGRADING BACTERIA

Theses of doctoral (Ph.D.) dissertation

Sinchan Banerjee
GÖDÖLLŐ, HUNGARY
2023

Hungarian University of Agriculture and Life Sciences, Hungary

Name of Doctoral School: Doctoral School of Environmental Sciences

Discipline: Molecular Ecology, Environmental Microbiology

Head of Doctoral School: Csákiné Dr. Erika Michéli, Ph.D.
Professor, Head of the department
Department of Soil science and Agrochemistry
Faculty of Agricultural and Environmental Sciences
Institute of Environmental Sciences

Supervisor : Dr. András Táncsics , Ph.D
Head of the department, Senior Research Fellow
Department of Molecular Ecology
Institute of Aquaculture and Environmental Safety

……. …….

Approval of Head of Doctoral School **Approval of Supervisor**

Table of Contents

1 Background and aim

Despite of several alternative energy sources available for mankind but unfortunatelytill date , fossil fuels are the main sources of energy. The use of fossil fuels leads to several consequences mainly related to global warming. Moreover, air pollution caused by the burning of fossil fuels is responsible for nearly one in every five deaths worldwide. Hence, the world is trying to turn down the use of fossil fuels. Even in the 26th UN Climate Conference (COP26) in Glasgow (2021), world leaders explicitly plan to reduce the use of fossil power (coal, oil, gas). Due to this widely used nature, accidental and negligent contamination of soil and aquifers is also very common. There are several incidents of such contaminations that took place during the last century. As one of the rapidly growing economies of the European Union, Hungary is also experiencing petroleum derivatives as the most common environmental pollutants like the rest of the world. However, there are several laws in place to limit such contaminations, but unfortunately, they cannot completely restrict such incidents. That is why we need to be prepared with advanced processes in place to treat such conditions to save the ecosystem and mankind from their toxic and carcinogenic effects. The mono-aromatic hydrocarbons, such as the carcinogenic benzene, and others like toluene, ethylbenzene and *o*-, *m*-, and *p*- xylene are among the most frequent groundwater contaminants in Hungary, as well. Due to their relatively high-water solubility, soil contamination of such kind is always a threat to the ground water reserve. Therefore, the cleanup of the polluted ecosystems is always obligatory by law in Hungary. Since microbial degradation of these mono-aromatic hydrocarbons readily occurs under aerobic conditions, oxygen supply into the contaminated ecosystem is a frequently applied method to enhance aerobic degradation of the contaminants. This is necessary because hydrocarbons stimulate the metabolic activity of indigenous aerobic microbial populations resulting in a rapid depletion of dissolved oxygen in subsurface environments. Thus, the development of extensive microaerobic to anaerobic zones in these environments can be observed. Although a broad diversity of microorganisms can degrade mono-aromatic hydrocarbons even in the absence of oxygen, benzene, along with *p*-and *o*-xylene, are among the least degradable BTEX compounds. Though there are several studies have been conducted regarding both aerobic and anaerobic degradation, somehow, the microaerobic degradation aspect got neglected. As the contaminated subsurface environment is generally oxygen limited in nature, so it is always worth studying the microaerobic degraders that harbour specific enzymes encoded by unique

functional genes needed to transform contaminants under low dissolved oxygen concentration. Unfortunately, very little is known about how different microbial communities evolve under different levels of dissolved oxygen availability in xylene contaminated ecosystems . It is known that trace amounts of oxygen can support aerobic degradation and certain functional genes (catechol dioxygenases) play a key role in the degradation of aromatic compounds under microaerobic conditions. Microbial communities of BTEX-contaminated microaerobic ecosystems are usually dominated by Betaproteobacteria. Besides, in these environments, catechol 2,3- dioxygenase (*C23O*) genes that encode subfamily I.2.C-type extradiol dioxygenase enzymes are represented in large diversity. Therefore, it is conceivable that a much smaller level of aeration of the contaminated subsurface environment would be sufficient to enhance aerobic degradation of the contaminants. In light of environmental sustainability, it would be beneficial to reveal those structural changes in the microbial communities induced by the contamination and concurrent processes, such as decreased oxygen availability.

Key questions and aims of the research are:

- One of the main research questions of this research is whether the low availability of dissolved oxygen is the main factor which causes the dominancy of Betaproteobacteria in xylene contaminated subsurface environments. Could other bacterial groups and other types of *C23O*s become dominant with oxygen limitation?

- The second main aim of the project is to isolate bacterial strains from the xylene-degrading enrichment cultures and to look for new *C23O* and other relevant functional genes (e.g., toluene monooxygenases) possibly taking part in the degradation of the contaminants either microaerobic or fully aerobic conditions.

- Thirdly, to find out how distinct microbial communities evolve in case of a xylene contamination under hypoxic or aerobic condition.

2 Materials and methods

To answer the research questions mentioned above, we have performed two enrichment experiments during the tenure of my doctoral study. The following diagram is a simplified schematic representation of the material and methods used in general for the whole study.

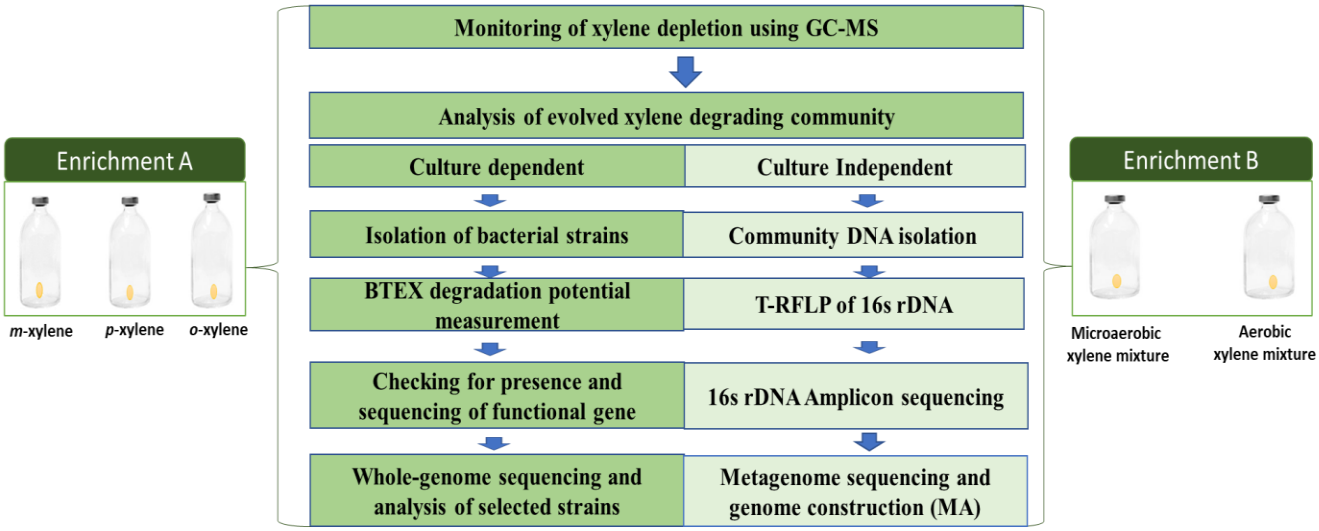

Figure 1: Schematic diagram of the overall processes flow: Enrichment A: aerobic enrichment with individual xylene isomer as carbon source, Enrichment B: Aerobic and microaerobic enrichment with xylene isomers mixture as carbon source.

2.1 Microbial enrichment experiment

In the first experiment, we established aerobic (7-8 mg/L dissolved oxygen) enrichment using *m*-, *p*- and *o*-xylene individually as sources of carbon and energy. For our second enrichment experiment, aerobic and microaerobic ($\leq$ 0.5 mg/L dissolved O_2) enrichments were set up, in which a mixture of xylene isomers (*m*-, *p*-, *o*-xylene) was used as the sole source of carbon and energy. Details of the enrichment set up explained using a simplified diagram in **Fig. 2**.

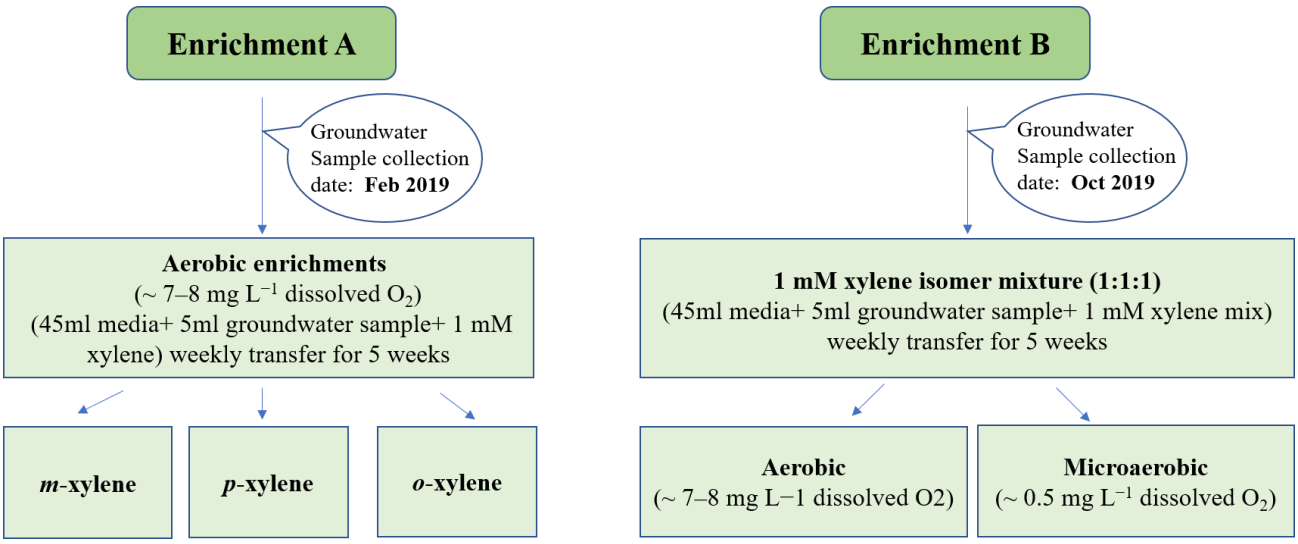

Figure 2: Schematic diagram of enrichment experiments

For enrichment A, we worked with two parallel samples each, and for Enrichment B, three parallel samples were made. As a culture medium, modified minimal mineral salts (MS) medium developed by Fahy et al. (2006) was used. *m*-, *p*- or *o*-xylene was added individually in enrichment A and a xylene mixture (1:1:1) in enrichment B as sole carbon and energy source at a final concentration of 1mM. For both the enrichment contaminated groundwater sample collected from the BTEX contaminated site of Siklos, Hungary was used as initial inoculum. The microcosms were incubated for 7 days , then 5 mL inoculum from each microcosm was transferred into a fresh medium. Similar transfers were made for consecutive five weeks. To monitor the consumption of xylene as the sole carbon source, the concentration of *m*-, *p*- and *o*-xylene was measured for 5th-week enrichment microcosms at 24h intervals by headspace analysis on GC–MS.

2.2 Culture dependent Analyses

Cultivable bacterial strains were isolated from fifth week samples of xylene induced enrichment cultures using conventional methods. Colonies of different morphology were isolated and maintained on R2A plates. In order to identify the strains, the genomic DNA of isolates was extracted from pure cultures grown on R2A agar using the UltraClean Microbial DNA Kit (Qiagen. Germany). The 16S rRNA and *C23O* functional genes were PCR amplified and sequenced following Sanger sequencing method to reveal the identity of the strains and look for the presence of desired functional gene. BTEX

degradation potential of the isolated strains were measured under both aerobic and microaerobic conditions with particular emphasis on xylene degradation as the strains were isolated from a xylene degrading enrichments. Measurements were carried out in triplicates, along with triplicates of negative controls by headspace analysis on an ISQ Single Quadrupole gas chromatography-mass spectrometer (GC–MS). Whole genome sequencing of *Hydrogenophaga* sp. strains D2P1^{T} and D2P3 isolated from enrichment A and *Pseudomonas* sp. MAP12^{T} , and *Sphingobium* sp. AS12, from enrichment B were performed by SeqOmics Ltd. Hungary. These xylene degrading yet undescribed strains were investigated using polyphasic approach to describe them as new species of the genus using molecular biological and chemotaxonomic studies suggested by J. B. Tindall et al. (2010).

2.3 Culture Independent Analyses

To get an overview of the selectively enriched xylene degrading microbial community in the enrichments, the genomic DNA of the community was isolated. That was performed with the fifth week samples after five consecutive transfers. Terminal restriction fragment polymorphism (T-RFLP) fingerprinting was performed to get an overview of the microbial community composition. In the first step, 16s rDNA was amplified by PCR using fluorescently labelled (VIC label, Life Technologies ™, USA) forward primers. In our examination, the VIC-labelled 16S rDNA were digested with RsaI (GT↓AC) (Thermo Fischer Scientific Inc. USA). To acquire in depth knowledge about the microbial community composition, 16S rRNA gene amplicon sequencing was performed for the enrichment sample using the isolated community DNA sample. The variable V3 and V4 regions of the 16S rDNA were amplified on Illumina MiSeq platform by Seqomics Kft. The metagenome sequencing and metagenome-associated genome analysis were performed to identify the uncultivated xylene degraders and look for new functional genes involve in xylene degradation. For the genome-resolved metagenomics approach metagenome sequencing was performed on the Illumina MiSeq platform by SeqOmics Kft. using an Agilent 2200 Tapestation system.

3 Results and discussion

3.1 Microbial community analysis of aerobic m-, p- and o-xylene degrading enrichments (Enrichment A)

3.1.1 Microbial community composition of xylene degrading enriched communities revealed by T-RFLP and 16S rRNA gene amplicon sequencing

The xylene degradation of enrichments reflected that highly competent aerobic xylene degrading communities evolved in the enrichments by the fifth week. Among the enrichments, *m*-xylene degradation was the fastest, followed by *p*-xylene and *o*-xylene. The primary data of T-RFLP analysis of enrichments and their duplicates revealed the fact about the similarity of the replicates and their community representation. The bar plot of T-RFLP fingerprints clearly showed that the composition of the bacterial communities of *m*-, *p*-, and *o*-xylene-degrading enrichments were distinctly different, and the community composition was relatively similar for replicates (**Fig. 3**). The most dominant T-RFs in *m*-xylene degrading enrichments were 841-bp (approx.72%), 418-bp (approx. 5%), in *p*-xylene degrading enrichments 841-bp (approx. 40%), 418-bp (20%), and in *o*-xylene degrading enrichments 117-bp (10%), 416-bp (44%). The most prominent T-RF that was detectable in all three types of enrichments in different fractions was the 841-bp. The major difference creating T-RFs were 411-bp, 435-bp, 460-bp, and 466-bp. These peaks were present in *o*-xylene degrading enrichments and missing in the others enrichments.

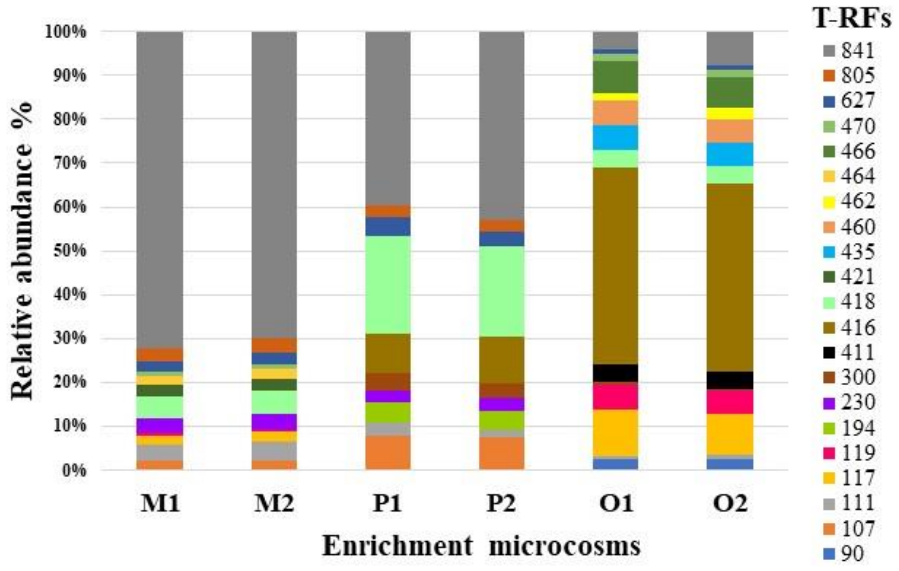

Figure 3. Bacterial community structure of *meta*- (M1, M2), *para*- (P1, P2) and *ortho*-xylene-degrading (O1, O2) enrichment cultures as revealed by 16S rRNA gene based T-RFLP.

Based on the T-RFLP results, enrichment samples, namely M1, P1 and O1, were selected for Illumina 16S rRNA gene amplicon sequencing as representative of the community diversity. A closer, in-depth look into the individual enrichments at genus levels helped us to understand the community diversity and composition. At the class level, members of the class *Gammaproteobacteria* dominated the bacterial community in most of the enrichments. Enrichment M1 showed the utmost order-based diversity. Both M1 and P1 enrichment communities were dominated by *Gammaproteobacteria* and *Bacteroidia*, and the exception was the O1 enrichment sample, where *Actinomycetia* showed more than 15% relative abundance in the community, which made the community different from the other two enrichments. (**Fig. 4**).

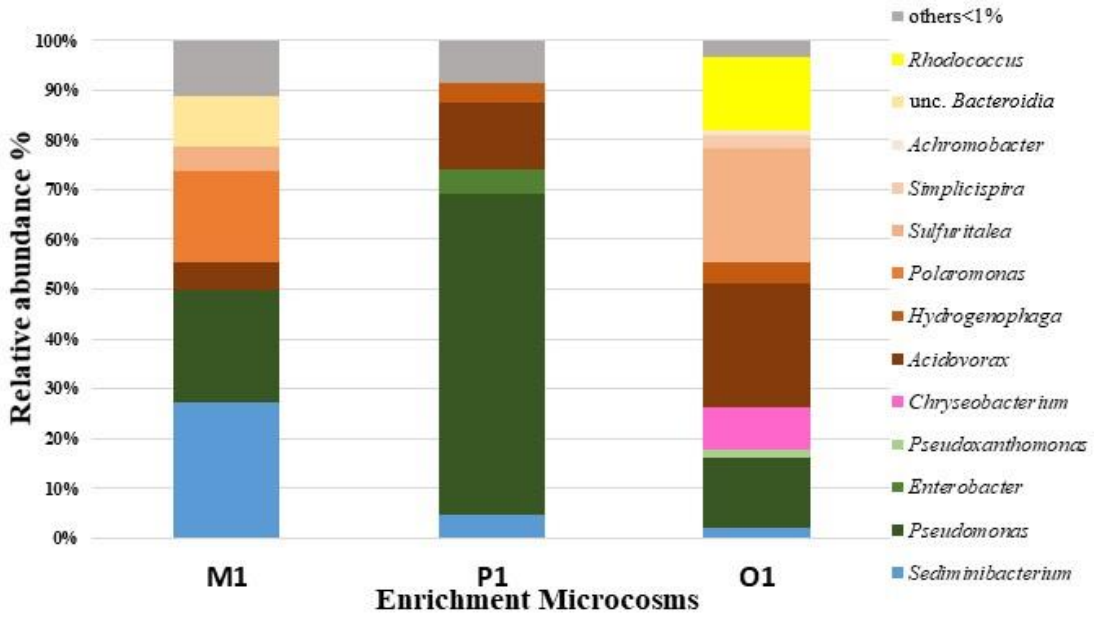

Figure 4**.** Genus level bacterial community structure of enrichments M1, P1 and O1 as revealed by Illumina paired-end 16S rRNA gene amplicon sequencing. Only taxa contributing more than 1% abundance were depicted.

The bacterial community of the *m*-xylene-degrading enrichment (M1) was dominated by members of the genera *Sediminibacterium* (27.1%), *Pseudomonas* (22.8%) and *Polaromonas* (18.4%). Whereas in *p*-xylene-degrading enrichment (P1), members of the genus *Pseudomonas* overwhelmingly dominated the community by showing 64% relative abundance (**Fig. 4**). Besides, members of some other genera like *Acidovorax* (13.2%), *Enterobacter* (5.1%) *Sediminibacterium* (4.6%) and *Hydrogenophaga* (3.9%) were detectable with prominent abundance. However, the *o*-xylene-degrading enrichment (O1) showed an altogether different community structure. *Pseudomonas*-related bacteria were present in the community but only with 14% abundance. *Acidovorax* (24.9%)

was the most dominant genus, along with *Sulfuritalea* (22.8%), *Rhodococcus* (14.6%), *Chryseobacterium* (8.4%) and *Hydrogenophaga* (4.4%), making it the most different enrichment based on community diversity in comparison to the other two enrichments. The *p*-xylene-degrading enrichment, which had a high abundance of sequence reads affiliated with genera *Pseudomonas* and *Acidovorax*, enabled us to assume that these bacteria had played a role in the aerobic degradation of *p*-xylene. Another major group was *Hydrogenophaga* present in all of the enrichments, which have been selected to study further to reveal their role in the enrichment community. Likewise, *m*-xylene degrading enrichment was also dominated by *Pseudomonas, Sediminibacterium* and *Acidovorax,* making the diversity of the enrichment similar to the *p*-xylene degrading enrichment. The difference was the introduction of the genus *Polaromonas,* which is reported to be a genus with potential BTEX degraders. Contrarily, the *o*-xylene-degrading bacterial community was considerably different due to the presence of the genera *Sulfuritalea* and *Rhodococcus,* along with *Pseudomonas* and *Acidovorax* as the major players in the community. Moreover, the presence of *Chryseobacterium, Simplicispira, Pseudoxanthomonas, Achromobacter* in noticeable fractions makes it an interesting and quite different community compared to the other enrichments. In addition to this, *Rhodococcus* was the only dominant genus that was not witnessed in other enrichments, assuming it as the leading player for *o*-xylene degradation. Subsequently, this hypothesis was proven in the BTEX degradation experiment as the isolated *Rhodococcus* strain could effectively degrade *o*-xylene .

3.1.2 Details of strains isolated from enrichments and their BTEX degradation capability

Based on different colony morphology and growth pattern total of 21 strains have been isolated, among which six isolates from *m*-xylene-degrading enrichment, eight isolates from *p*-xylene-degrading enrichment and seven isolates from *o*-xylene-degrading enrichment. Strains isolated from the *m*-xylene-degrading enrichment (M1) belonged to the genera of *Acidovorax, Pseudacidovorax, Polaromonas, Lysobacter, Pseudomonas* and *Achromobacter*. Amid those strains, *Acidovorax, Pseudacidovorax, Polaromonas, and Pseudomonas* possessed subfamily I.2.C-type *C23O* gene, which was sequenced further. Strains isolated from the *p*-xylene-degrading enrichment (P1), were members of the genera *Hydrogenophaga, Acidovorax, Mycolicibacterium, Pseudomonas* and *Enterobacter*. The screening of subfamily I.2.C-type *C23O* gene showed that *Hydrogenophaga, Acidovorax and Pseudomonas* strains harboured such a gene. The lowest diversity of isolates was observable in the case of the *o*-xylene-degrading enrichment O1. The seven isolates obtained from

this enrichment belonged to four genera. Besides, none of them harboured subfamily I.2.C-type *C23O* gene. The degradation potential of the isolated *Rhodococcus imtechensis* strain, D2O4 was assessed, which led us to the observation that it could degrade *o*-xylene, toluene and benzene, which explains its presence in the *o*-xylene-degrading enrichment community as a key player of *o*-xylene utilisation.

3.1.3 Comparative whole-genome analysis of two BTEX degrading strains Hydrogenophaga sp. D2P1^T and D2P3, isolated from xylene degrading enrichment culture

Among the isolates, two *Hydrogenophaga* strains, D2P1^T and D2P3, were the most interesting for us since they were able to degrade at least one of the xylene isomers and possessed subfamily I.2.C-type *C23O* genes. The two strains shared identical 16S rRNA genes, but Sanger-sequencing revealed that they encode entirely different subfamily I.2.C-type *C23O* genotypes. Moreover, it could be assumed that strain D2P1^T harbours more than one genotype of the corresponding *C23O* gene. Based on the 16S rRNA gene similarity, strains D2P1^T and D2P3 were closely related to *H. taeniospiralis* (~99.1% homology). strain D2P1^T is described as the type strain of the new species *Hydrogenophaga aromaticivorans*, and the analysis of the whole genome of strain D2P1^T revealed that it has three different subfamily I.2.C-type *C23O* genes. Besides, a large gene cluster (~28 kbp) was identified, which encoded all of the genes (e.g. xylene monooxygenase and benzoate 1,2-dioxygenase) required for the transformation of *p*-, and *m*-xylene to 3-, and 4-methylcatechol, respectively. This cluster contained one of the subfamily I.2.C-type *C23O* genes (locus tag F3K02_21385). Since strain D2P3 was not able to utilize *p*-, and *m*-xylene, but could use *o*-xylene as sole source of carbon and energy, we have sequenced its whole genome as well. Subsequently, the two genomes were aligned to each other and analyzed. The dDDH value between strain D2P1^T and D2P3 was 79.7%, while the OrthoANI value was 97.6%, which clearly indicated that they belong to the same genomic species. Besides, they had highly similar genome size (5.63 and 5.80 Mb, respectively). Alignment of the genome sequences revealed that strain D2P3 lacks the xylene-degradation gene cluster, which was observable in the case of strain D2P1^T. Both strains harbour a phenol degradation gene cluster, encoding a multicomponent phenol hydroxylase (mPH) together with a complete *meta*-cleavage pathway. However, these gene clusters are different in structure, and the corresponding genes show only ~80-90% sequence similarity to each other (**Fig. 5**). This difference can be the key to understand the different xylene-degrading ability of strains D2P1^T and D2P3. It was observed in case of *Pseudomonas stutzeri* strain OX1, which is prominent toluene and *o*-xylene-degrading bacterium, that a phenol degradation gene

cluster plays a crucial role in its *o*-xylene-degrading ability. The structure of this phenol degradation operon shows similarity to that was observed in strain D2P3.

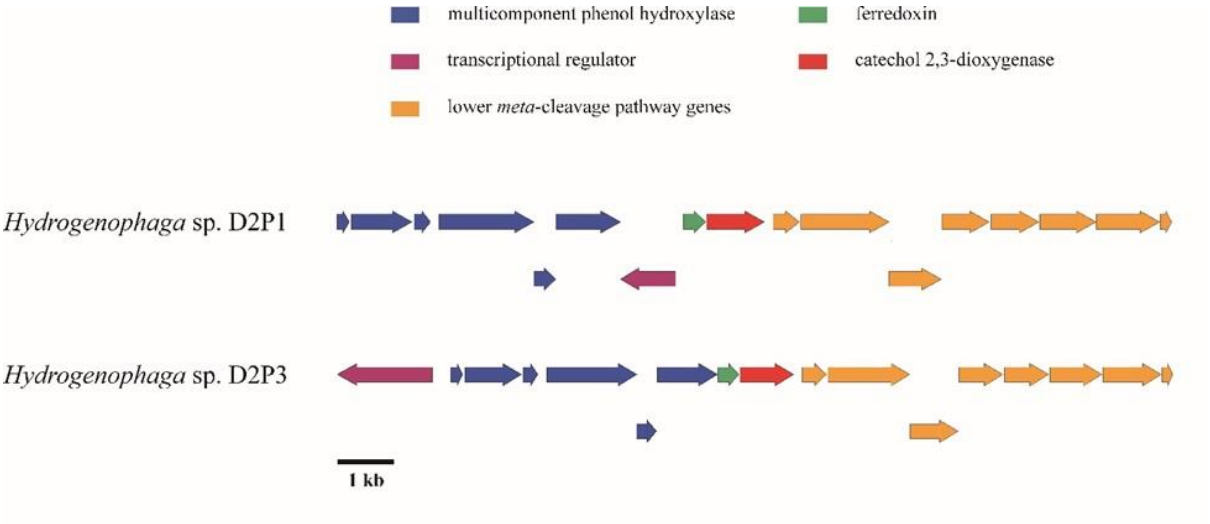

Figure 5. Physical map of the phenol degradation gene clusters of Hydrogenophaga sp. strains D2P1^T and D2P3.

Both operons are regulated by a σ^{54}-interacting transcriptional regulator, and the organization of the mPH and lower *meta*-cleavage genes is similar (at least at the known parts of the operon in the case of *P. stutzeri* strain OX1). However, it is still a question how the *o*-xylene is converted into 3,4-dimethylphenol, since toluene-*o*-xylene monooxygenase was not found in the genome of strain D2P3. One possible explanation is that the mPH is responsible for both the hydroxylation of *o*-xylene and the subsequent hydroxylation of 3,4-dimethylphenol to 3,4-dimethylcatechol. Nevertheless, transcriptomic analysis will be necessary to answer this question. In the case of strain D2P1^T, a LysR-type transcriptional regulator gene was found wedged between the mPH and the ferredoxin gene. Besides, no σ^{54}-interacting transcriptional regulator gene was found upstream of the mPH. Instead, an *orf* encoding an IS5 family transposase was found in the corresponding position, hinting at the possibility that this gene cluster was acquired through horizontal gene transfer (HGT) by strain D2P1^T. Moreover, it can also be speculated that this gene cluster functions only partially since the mPH lacks its own transcriptional regulator.

3.1.4 Description of a novel aerobic xylene degrading bacterial species isolated from para-xylene degrading enrichment (*Hydrogenophaga aromaticivorans* sp. nov)

Strain D2P1^T was isolated from *p*-xylene degrading enrichment. Based on our preliminary 16S rRNA gene sequencing results, this strain was identified as a new species of the genus *Hydrogenophaga*. Using a polyphasic taxonomic approach, including whole genome sequencing and physiological, biochemical and chemotaxonomic analyses its taxonomic position was established. D2P1^T, can degrade benzene, *p*-, and *m*-xylene. Its a Gram-stain-negative, aerobic, yellow-pigmented bacterium. Phylogenetic analyses based on 16S rRNA genes showed that D2P1^T shares a distinct phyletic lineage within the genus *Hydrogenophaga* and shows highest 16S rDNA sequence similarity to *Hydrogenophaga taeniospiralis* NBRC 102512^T (99.2%) and *Hydrogenophaga palleronii* NBRC 102513^T (98.3%). The draft genome sequence of D2P1^T is 5.63 Mb in length and the genomic DNA G+C content is 65.5 mol%. Orthologous average nucleotide identity (OrthoANI) and digital DNA-DNA hybridization (dDDH) analyses confirmed low genomic relatedness to its closest relatives (OrthoAni<86%; dDDH<30%). D2P1^T contains ubiquinone 8 (100%) as the respiratory quinone and phospholipid, phosphatidylglycerol, phosphatidylethanolamine, diphosphatidylglycerol as major polar lipids. The main whole-cell fatty acids of D2P1^T are summed feature 3 (C16:1 ω7c/C16:1 ω6c), C16:0 and summed feature 8 (C18:1 ω7c/C18:1 ω6c). The polyphasic taxonomic results indicated that strain D2P1^T represents a novel species of the genus *Hydrogenophaga*, for which the name *Hydrogenophaga aromatica* sp. nov. is proposed. The type strain is D2P1^T (=LMG 31780 =NCAIM B 02655).

3.2 Comparative analysis of aerobic and microaerobic xylene degrading bacterial enrichments (Enrichment B)

3.2.1 Bacterial community composition of aerobic and microaerobic xylene-degrading enrichmentsrevealed by 16S rDNA amplicon sequencing

The result of xylene degradation measured by GC-MS, reflected that by the fifth week, the enrichments had developed extremely competent aerobic and microaerobic xylene degrading microbial communities that could degrade xylene effectively. In the presence of oxygen *m*-, *p*-xylene took 24h of incubation, and *o*-xylene took 48h of incubation for the complete degradation. Likewise, under hypoxic conditions, degradation was a bit slower, where it took 48h for *m*- and *p*-xylene degradation and 96h for complete elimination of *o*-xylene. Results of 16S rRNA gene amplicon

sequencing revealed that the aerobic and microaerobic xylene degrading enrichments harboured noticeably different bacterial communities. The aerobic xylene degrading enrichments were predominantly dominated by the members of the genus *Pseudomonas* with relative abundance values between 47-59%, followed by *Sphingobium* (20-30%) and *Acidovorax* (4-6%). Noticeable minor community players were: *Flavobacterium*(6%) and *Roseimicrobium* (4%). On the other hand, in the case of microaerobic enrichments, members of the genus *Pseudomonas* were overwhelmingly dominant in the community by showing ~50% abundance value. But interestingly, *Sphingobium* was not detectable in any of the microaerobic enrichments. **(Fig.6)**. Members of the genus *Sphingobium* were completely replaced with either members of the genus *Azovibrio* (25% relative abundance in enrichment MIC1), or *Rhodoferax* (20 and 31% relative abundances in enrichments MIC2 and MIC3, respectively)

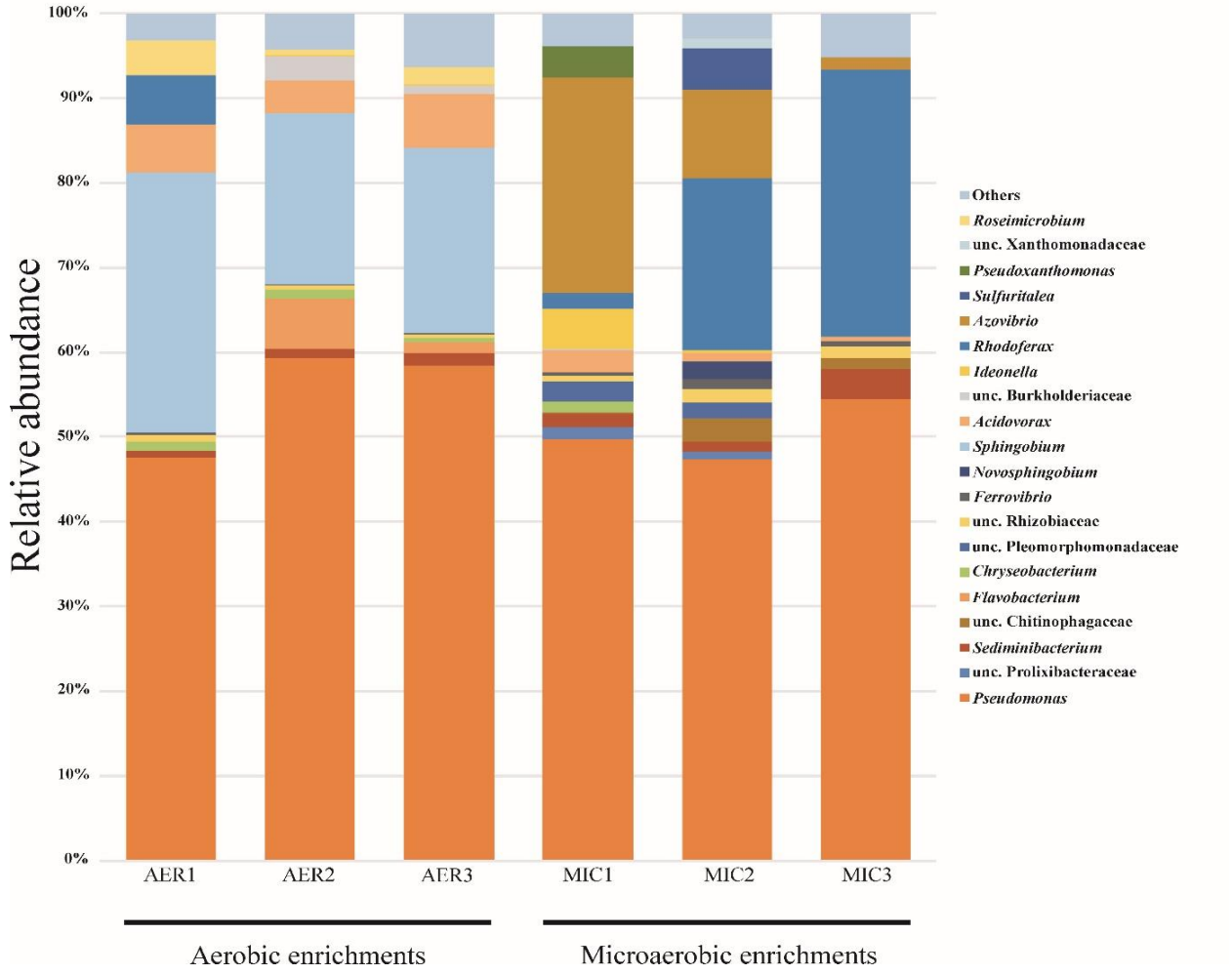

Figure 6. Genus level bacterial community structure of aerobic and microaerobic xylene-degrading enrichments as revealed by Illumina paired-end 16S rRNA gene amplicon sequencing. Only taxa contributing more than 1% abundance were depicted.

Besides these dominant groups, there were several other genera also present in relative abundance of more than 0.5% in aerobic and microaerobic enrichments like *Sediminibacterium, Flavobacterium, Chryseobacterium, Ferrovibrio, Ideonella, Sulfuritalea*, *Pseudoxanthomonas*. Minor community players of the microaerobic enrichments (e.g. *Pseudoxanthomonas, Sulfuritalea, Ideonella*) were typically present only in one of the triplicate enrichments. Other than *Pseudomonas*, members of the genus *Sediminibacterium* were the only ones that could be found in all of the enrichments, both in aerobic and microaerobic settings. The absence of *Sphingobium*-related sequence reads in the microaerobic enrichments indicates that members of the genus *Sphingobium* only contribute to xylene degradation under strict aerobic conditions. This statement could be supported by the fact that the *Sphingobium* strains isolated from the aerobic xylene degrading enrichment could effectively degrade all three xylene isomers aerobically. Overall, the composition of the bacterial community was distinctly different from microaerobic enrichments in general. The three parallel microaerobic enrichment cultures show similarity in microbial composition but with one enrichment (MIC1) bit different from the others in the percentage of *Azovibrio* and *Rhodoferax*. Interestingly a competition among these two groups could be observed where the proportion of one increases with the decrease of another. It is well known that a certain phyletic lineage of the genus *Rhodoferax* is associated with aromatic hydrocarbon-contaminated subsurface environments. However, so far, no cultivable representative belonging to this lineage has been isolated from a petroleum hydrocarbon-contaminated environment. A genome resolved metagenomics study of enrichment MIC3 was conducted in order to elucidate the genome of the *Rhodoferax*-related bacteria abundant in the microaerobic enrichments.

3.2.2 Metagenome assembled genome analysis

From metagenome data of enrichment MIC3, in total, twelve high-quality metagenome-assembled genomes (MAG) were reconstructed; among them, one *Comamonadaceae*-related MAG was found. The phylogenomic analyses revealed that this MAG Bin1 represented a *Rhodoferax* genus-related bacterial lineage. The MAG was 4.70 Mb in size with a G+C content of 56.45% but did not contain SSU sequences, although its completeness was >97%. The recovered genome was found to be a member of the genus *Rhodoferax* as indicated by both MiGA and the UBCG pipelines (**Fig.7**), and showed the highest phylogenetic relatedness with *Rhodoferax* sp. strain MIZ03, *R. fermentans and R. ferrireducens*, respectively.

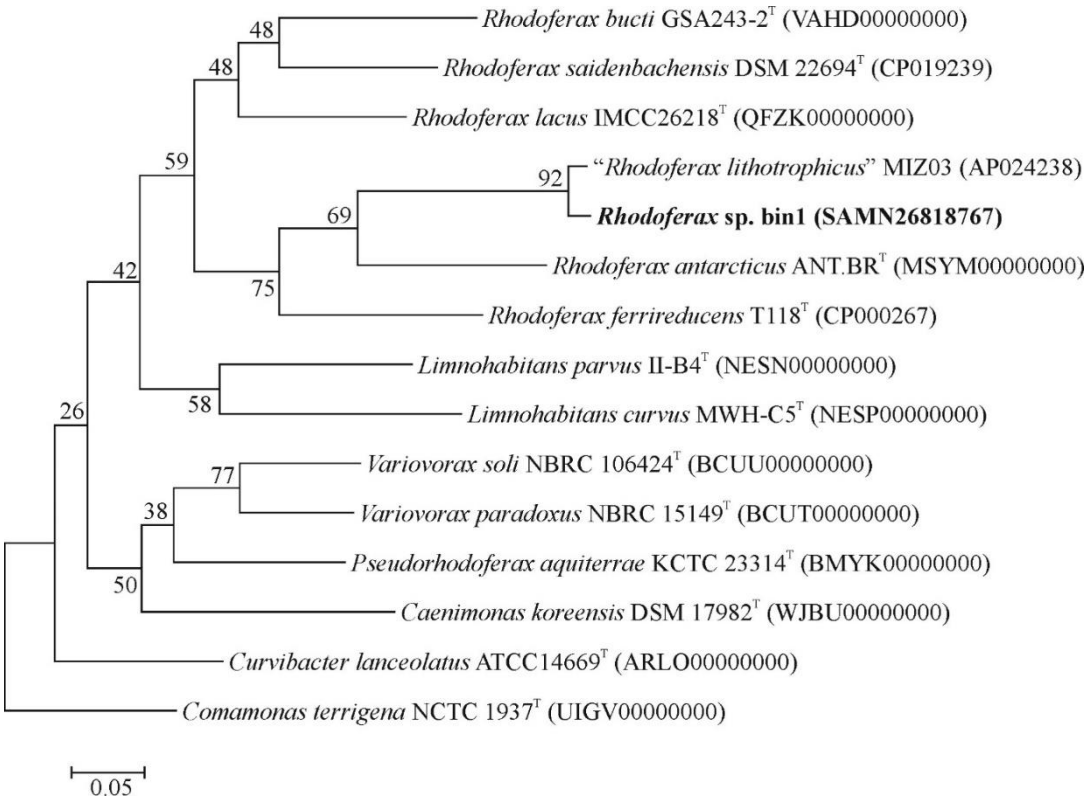

Figure 7. Phylogenomic tree constructed using UBCGs (concatenated alignment of 92 core genes) showing the phylogenetic position of *Rhodoferax sp*. bin1. Bar, 0.05 substitution per nucleotide position.

During the genome analysis, the presence of a gene cluster encoding a toluene-4-monooxygenase, by genes *tmoABCDEF* was uncovered. The mentioned gene cluster harboured a gene that encodes an outer membrane transport protein which is most likely involved in the movement of monoaromatic hydrocarbons from the extracellular environment to the periplasm. Another gene cluster coding for a multicomponent phenol hydroxylase (mPH) and a full *meta*-cleavage pathway was detected around 20 kilobases downstream of the gene cluster coding for Tmo (mPH1 cluster). In addition to that, another contig was also found to encode a second mPH and *meta*-cleavage enzymes (mPH2 cluster). Important to mention that these two mPH clusters showed completely different gene organizations (**Fig. 8**)

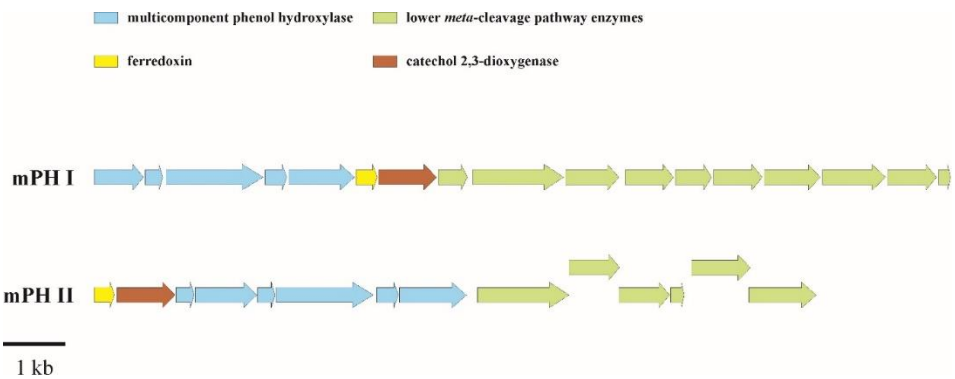

Figure 8 Physical maps of the multicomponent phenol hydroxylase (mPH) gene clusters in the MAG of *Rhodoferax* sp. bin1. The map was created by using CLC Genomics Workbench Tool version 2.1. Arrows represent predicted ORFs, while the direction of the arrows represents the direction of transcription. The different colours denote different functional groups of genes involved in aromatic hydrocarbon degradation.

Interesting to note that neither of the mPH clusters coded for subfamily I.2.C-type extradiol ring-cleavage dioxygenase (EDO) enzymes. On the phylogenetic tree, they actually formed a distinct but as-yet-undescribed subfamily of EDO enzymes. (**Fig. 9**).

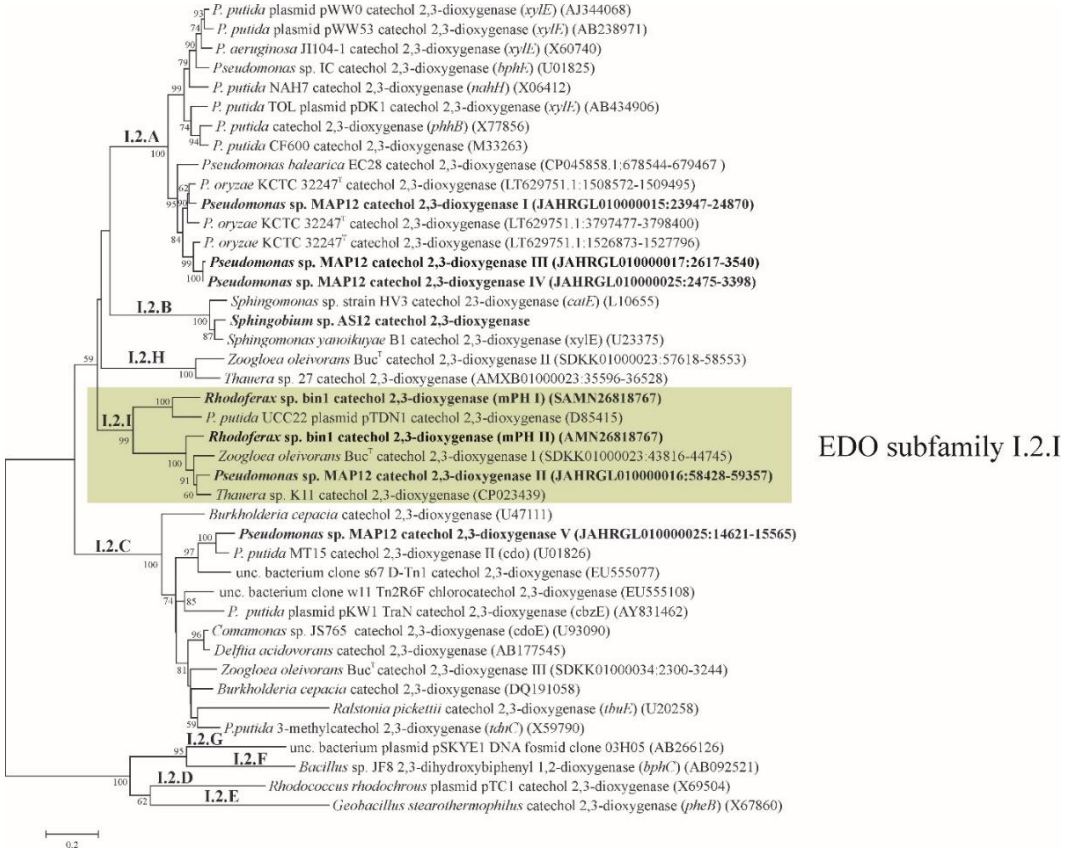

Figure 9. Neighbor-joining phylogenetic tree of deduced amino acid sequences of catechol 2,3-dioxygenase (*C23O*) genes encoding class I EDO enzymes. The tree shows the phylogenetic position of the newly defined class I EDO subfamily I.2.I. The C23Os revealed by the present study are highlighted with boldface type. Class I EDO subfamilies are indicated at the main branches of the tree. Only bootstrap values >50 are indicated. Scale bar, 0.2 substitutions per amino acid position.

Based on the evolutionary classification of extradiol dioxygenases, this new subfamily was assigned as I.2.I. (**Fig. 9**).The genome of *Rhodoferax* sp. strain MIZ03 (tentatively but not validly designated as "*R. lithotrophicus*") showed the presence of all of the aromatic hydrocarbon-degrading gene clusters that were found in *Rhodoferax sp.* bin1. This strain was described as the first single bacterium capable of both Fe(II) oxidation and Fe(III) reduction at circumneutral pH, isolated from a wetland in Japan. The genomic relatedness value between *Rhodoferax sp.* bin1 and *Rhodoferax sp*. strain MIZ03 obtained by OrthoANI (97.5%) indicated that they belong to the same genomic species. Hence, it may be assumed that this species has had the capacity to degrade aromatic hydrocarbons, and this property is not acquired recently. In addition, it can also be concluded that *Rhodoferax* lineage is strongly associated with aromatic hydrocarbon-contaminated subsurface environments. An in-depth analysis of the *Rhodoferax sp*. bin1 genome revealed the presence of a complete *narGHIJ* operon, encoding the respiratory nitrate reductase and upstream of this operon gene encoding a nitric oxide reductase complex. The existence of these genes may imply that this bacterium can respire with nitrate to convert it to nitrite. Since the MAG lacked nitrite and nitrous-oxide reductase genes, thus, its ability to carry out full denitrification remains unclear.

3.2.3 Description of strains isolated from aerobic and microaerobic enrichment cultures

Several strains were isolated, among them strains isolated from aerobic enrichment belonging to 6 genera and strains isolated from microaerobic enrichments belonging to 9 genera**.** Among aerobic and microaerobic strains belonging to common genera are *Pseudomonas, Variovorax,* and *Rhizobium*. The majority of the isolated strains from aerobic enrichments belonged to genus *Variovorax* (~40% of the isolates) along with representatives of genera *Flavobacterium*, *Sphingobium* and *Pseudomonas*. But in contrast, strains isolated from microaerobic enrichments were majorly found to be members of the genus *Pseudomonas* (~60% of the isolates). Based on the highest 16S rRNA gene similarity, the isolated *Pseudomonas* isolates could be divided into two groups *P. moorei* and P. *sagittaria/linyingensis* lineages, respectively. Results of genome-resolved metagenomics also supported this fact. Noticeably, the strains belonging to *Pseudomonas sagittaria/linyingensis* lineages were only found in microaerobic enrichments. Apart from *Pseudomonas,* strains belong to the genera *Pinisolibacter* and *Variovorax* were also isolated from microaerobic enrichments. Strain belonged to the genus *Pinisolibacter* found to represent a new lineage of the genus. Hence, this isolated strain named as MA2-2 was described as a new species of the genus and designated as *Pinisolibacter aquiterrae* has the ability to degrade aromatic hydrocarbons. The whole genome analysis of *Pi.*

aquiterrae also indicated towards its genetic ability to break down monoaromatic hydrocarbons due to the presence of catechol 2,3-dioxygenase (*C23O*) gene in a gene cluster, encoding a partial *meta*-cleavage pathway. As *Pinisolibacter* is the genus that shows the closest relatedness with the genus *Siculibacillus*, thus, the genome relatedness between these two was analysed. To do so, Bin10 MAG, which was identified as *Siculibacillus sp*. by the MiGA pipeline and *Pinisolibacter* strain MA2-2 were used. The Ortho ANI results revealed more than 99.9% relatedness between those two genomes, proving their degree of similarity. Among the isolated strains, we have identified some strains that represent a new lineage of genus *Pseudomonas* and only present in microaerobic enrichments. Besides, strains that presumably belong to a new lineage of genus *Sphingobium* and only present in aerobic enrichments as a representative of the most dominant community members were also isolated. These groups of organisms create the main difference between aerobic and microaerobic enrichment community structures. Therefore, it is interesting and important to analyse genomes of *Sphingobium sp*. strain AS12 and *Pseudomonas sp.* strain MAP12^{T}. Strain AS12 is capable of degrading all isomers of xylene, unlike its closest relatives. A rather rapid xylene degradation was observed under strict aerobic conditions, whereas in comparison, the degradation process was much slower under microaerobic (0.5 mg/L DO) conditions. This finding might help to explain why *Sphingobium* species were only found in aerobic enrichments. Strain AS12 was deposited at the National Collection of Agricultural and Industrial Microorganisms (NCAIM, Budapest, Hungary) with the accession number NCAIM B.02669. The result of xylene degradation revealed that MAP12^{T} was unable to utilize *o*-xylene as the sole source of carbon and energy but could only use *p*- and *m*-xylene. Furthermore, it could be concluded that oxygen availability does not affect the degradation efficiency of *m*-xylene as the degradation rate was similar for both aerobic and microaerobic experiments. **(Fig. 10)**. However, in the case of *p*-xylene degradation, microaerobic degradation was slower after 24hrs compared to aerobic degradation, though complete degradation occurred after 48hrs in both cases.

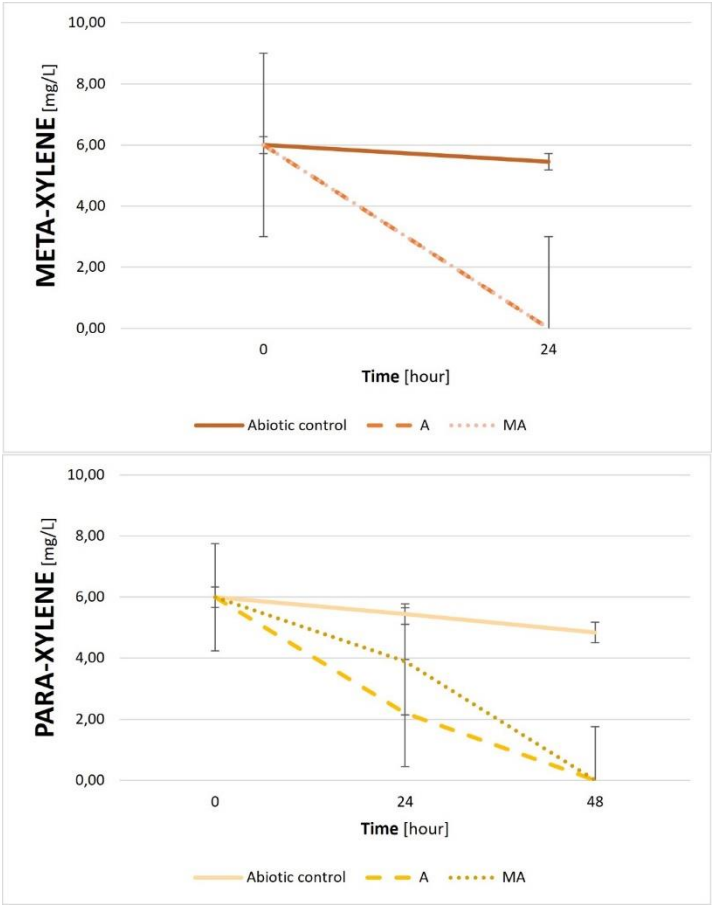

Figure 10**.** Aerobic and microaerobic degradation of meta- and para-xylene by *Pseudomonas sp*. strain MAP12^{T} . The concentration of xylene isomers was measured by GC-MS analysis for every 24 hours. The averages of triplicate experiments ± standard errors of the means, indicated by error bars, are shown.

3.2.4 Description of novel microaerobic xylene degrader *Pseudomonas aromaticivorans* sp. nov.

This strain MAP12^{T} was isolated from microaerobic xylene degrading enrichment. Primary 16s rDNA-based studies indicated that this strain represents a yet undescribed lineage of the genus *Pseudomonas*, which was the most dominant bacterial group in the xylene degrading microaerobic enrichments. To clarify the taxonomic status of this presumably novel member of the genus *Pseudomonas* using a polyphasic taxonomic approach, whole genome sequencing, physiological, biochemical, and chemotaxonomic analyses and evaluation of its microaerobic xylene degradation capability was done. The annotated genome of 4,392,784 bp in length was deposited to NCBI with accession number JAHRGL000000000. The MAP12^{T} genome has an average G+C content of 65.67 % and contains 4199 protein-coding sequences. In addition, the genome harbours the genes that encode at least five different C23O enzymes that break down aromatic molecules, among which one of them encoded a subfamily I.2.C-type extradiol dioxygenase, the ring–cleavage enzymes are known

to play a critical role under hypoxic degradation and one newly defined I.2.I type C23O that probably play a key role in xylene degradtion. The OrthoANI values between strain MAP12^T and phylogenetically closest neighbours were between 78-88% ,which is much lower than the threshold value of 95-96%, recommended for species level delineation. The dDDH values ranging from 22–52% were also less than the specified threshold (70%) recommended for species demarcation. The MAP12^T cells are rod-shaped, Gram-stain-negative, aerobic and motile with a polar flagellum. The cells were approximately 2.2-2.5 µm in length and 0.6–0.8 µm in diameter . Strain MAP12^T formed white, glossy, opaque, and circular colonies after 24 hrs of incubation on R2A at 28 °C. The major fatty acids (>5%) of strain MAP12^T were summed feature 3 (C16:1ω6c and/or C16:1ω7c), C16:0, summed feature 8 (C18:1ω6c and/or C18:1ω7c) and C12:0. The major respiratory quinones were ubiquinone Q9 (86.7%) and Q8 (13.3%). It was found that the strain was capable of efficiently degrading benzene, toluene, *m*- and *p*-xylene both aerobically and microaerobically. It could be concluded from the results of phylogenetic, genomic and chemotaxonomic studies that the strain MAP12^T represents a new species within the genus *Pseudomonas*, which was named *Pseudomonas aromaticivorans* sp. nov.

4 Conclusion and recommendations

Xylene is considered as one of the most common organic volatile toxic environmental contaminant found in the world, as well as in Hungary. Because of their relatively high-water solubility and least degradability, these compounds are often considered as a threat to the environment. It was observed that indigenous bacterial populations could naturally degrade such contaminants rapidly under strict aerobic conditions. But this rapid aerobic degradation leads to the formation of hypoxic conditions in the contaminated subsurface environment. Certain groups of bacteria that could adapt to these hypoxic environments via their special enzyme system play a role in the degradation of monoaromatic hydrocarbons. The main purpose of our study was to reveal how oxygen availability affects the development of xylene degrading bacterial communities and the functional genes involved in xylene degradation. Moreover, to find potential strains that belong to the most abundant communities and harbour subfamily I.2.C *C23O* genes, which are believed to have a key role in the hypoxic degradation of aromatic hydrocarbons. To achieve our research goal, two enrichments were established and investigated accordingly. For the first experiment, aerobic enrichments were set up with xylene-contaminated groundwater supplemented with *m*-, *p*-, and *o*-xylene as the sole carbon and energy source, respectively. Results revealed that distinctly different bacterial communities played role in the degradation of the different xylene isomers. *Pseudomonas* and *Acidovorax* were found to be the among abundant members present in all kinds of xylene degrading enrichments, but genera *Rhodococcus* and *Chryseobacterium* were the noticeable players in the aerobic *o*-xylene degrading community. Members of the genus *Hydrogenophaga* were also found to be present in abundance in *p*-and *o*-xylene degrading enrichments. Two representative strains of genus *Hydrogenophaga* were isolated from *p*-xylene-degrading enrichments and turned out that they were capable to degrade different isomers of xylene and harboured I.2.C *C23O* genes. The whole-genome analysis of these two *Hydrogenophaga* strains led us to the opinion that strains belonging to the same species with different xylene degrading potentials may coexist in the same environment. Moreover, it was also found that three or more subfamily I.2.C-type *C23O* genotypes can be linked to one single strain of hydrocarbon degrading bacterium. This observation allowed us to conclude that a high diversity of subfamily I.2.C-type *C23O* genes does not always mean a high diversity of degraders in a

contaminated environment. To understand the oxygen-mediated changes in the xylene degrading bacterial community composition, strictly aerobic and microaerobic enrichment microcosms were set up using a xylene mixture as a sole carbon source. The results showed that there was a significant difference in the microbial community structure. Though members of the genus *Pseudomonas* were the most dominant bacterial community present in both type of the enrichments but the rest of the community composition was noticeably different. In the case of aerobic enrichment, the genus *Sphingobium* was the second most abundant bacterial population involved in aerobic xylene degradation, whereas in microaerobic enrichment, members of the genera *Rhodoferax* or *Azovibrio* were the second most dominant community members and possibly participated in the microaerobic degradation of xylene. To support this statement, with the help of metagenomics and genome binning analysis, a yet unknown and uncultivated species of genus *Rhodoferax* was identified from the microaerobic xylene degrading enrichment harbouring a unique type of *C23O* gene (subfamily I.2.I.) which might play a crucial role in the microaerobic breakdown of xylene. As a part of our research goal, we have successfully isolated several bacterial strains that could be useful for bioremediation purposes in future. From the first experiment, two strains of xylene degrading strains represented a novel species of the genus *Hydrogenophaga* were isolated and described as *Hydrogenophaga aromaticivorans* sp. nov. In the second enrichment, we have isolated a microaerobic xylene degrading strain belonging to a yet undescribed species of the genus *Pseudomonas*. This strain was taxonomically positioned and named as *Pseudomonas aromaticivorans* sp.nov. This strain harbours five type of *C23O* genes including I.2.I.type *C23O* gene . From the aerobic enrichment, a strain of *Sphingobium* was isolated, which showed ability to degrade all three xylene isomers effectively. Phylogenomic analysis revealed that, presumably, this strain represents novel species of the genus *Sphingobium*.

5 New scientific results

Enrichment A:

- Study of *m*-, *p*- and *o*-xylene degrading aerobic enrichment cultures revealed that though members of the genera *Pseudomonas* and *Acidovorax* are abundant in all kinds of xylene isomer degrading enrichments but genera *Rhodococcus* and *Chryseobacterium* are the major players in the aerobic *o*-xylene degrading community. Moreover, members of the genus *Hydrogenophaga*, containing I.2.C *C23O* gene, can be prominent xylene-degraders.

- From the aerobic *p*-xylene degrading enrichment, a new bacterial strain designated as D2P1^T was isolated, belonging to yet undescribed species of the genus *Hydrogenophaga*. This strain showed the ability to degrade benzene, *m*- and *p*- xylene aerobically. According to international requirements, it was investigated to assign a taxonomic position, and the species was named *Hydrogenophaga aromaticivorans*.

- Whole-genome analysis of two *H.aromaticivorans* strains (D2P1^T and D2P3) isolated from the *p*-xylene degrading enrichments showed that different subpopulations of the same species might coexist in the same environment with different xylene degrading potential. Moreover, it was also found that three or more subfamily I.2.C-type *C23O* genotypes can be linked to one single hydrocarbon degrading strain. Accordingly, the high diversity of subfamily I.2.C-type *C23O* genes does not always guarantee a high diversity of degraders in a contaminated environment.

Enrichment B:

- A bacterial strain, designated as MAP12^T, has been isolated from microaerobic xylene degrading enrichment and identified as a novel bacterial species of the genus *Pseudomonas*. It

was analysed using polyphasic taxonomic approach and named *Pseudomonas aromaticivorans*. This newly described strain has the ability to degrade benzene-, toluene-, *m*- and *p*-xylene under both aerobically and microaerobically; in addition, it harbours five catechol 2,3-dioxygenase (*C23O*) genes including a newly defined subfamily of extradiol dioxygenases, designated as I.2.I.

- A new bacterial strain belonging to the genus *Sphingobium* was isolated from aerobic xylene degrading enrichment. Phylogenomic analysis of the strain disclosed that the strain presumably represents a yet undescribed lineage of the genus *Sphingobium.* The strain was able to degrade all three xylene isomers (*m*-, *p*- and *o*-xylene). It's the first report of any *Sphingobium* strain that can degrade all three xylene isomers along with toluene and ethylbenzene as a pure culture.

- With the help of metagenome sequencing and metagenome-associated genome analysis, a hitherto unknown and uncultivated bacterial species belonging to the genus *Rhodoferax* was identified, which was only predominant in microaerobic xylene degrading enrichments. The genome of this bacterium coded two catechol 2,3-dioxygenase enzymes which belonged to a newly defined subfamily of extradiol dioxygenases, designated as I.2.I. that might play a key role in the microaerobic degradation of xylene.

6 Publications

Táncsics A, **Banerjee S**, Soares A, Bedics A, Kriszt B. (2023) Combined Omics Approach Reveals Key Differences between Aerobic and Microaerobic Xylene-Degrading Enrichment Bacterial Communities: Rhodoferax─ A Hitherto Unknown Player Emerges from the Microbial Dark Matter. *Environmental Science & Technology*, **57**:2846–2855, doi:10.1021/acs.est.2c09283.

Bedics A, Táncsics A, Tóth E, **Banerjee S**, Harkai P, Kovács B, Bóka K, Kriszt B. (2022) Microaerobic enrichment of benzene-degrading bacteria and description of *Ideonella benzenivorans* sp. nov., capable of degrading benzene, toluene and ethylbenzene under microaerobic conditions. *Antonie van Leeuwenhoek*, **115**:1113-28. doi:10.1007/s10482-022-01759-z.

Banerjee S, Bedics A, Tóth E, Kriszt B, Soares AR, Bóka K, Táncsics A. (2022), Isolation of *Pseudomonas aromaticivorans* sp. nov. from a hydrocarbon-contaminated groundwater capable of degrading benzene-, toluene-, m-and p-xylene under microaerobic condition. *Frontiers in Microbiology*, 13:929128-929128. doi:10.3389/fmicb.2022.929128

Banerjee S, Bedics A, Harkai P, Kriszt B, Alpula N, Táncsics A. (2022). Evaluating the aerobic xylene-degrading potential of the intrinsic microbial community of a legacy BTEX-contaminated aquifer by enrichment culturing coupled with multi-omics analysis: uncovering the role of *Hydrogenophaga* strains in xylene degradation. *Environmental Science and Pollution Research*, **29**:28431-28445. doi: 10.1007/s11356-021-18300-w.

Bedics A, **Banerjee S**, Bóka K, Tóth E, Benedek T, Kriszt B, Táncsics A. (2022). Pinisolibacter aquiterrae sp. nov., a novel aromatic hydrocarbon-degrading bacterium isolated from benzene-, and xylene-degrading enrichment cultures, and emended description of the genus Pinisolibacter. *International Journal Of Systematic and Evolutionary Microbiology*. **72**:005229, doi: 10.1099/ijsem.0.005229.

Banerjee S, Táncsics A, Tóth E, Révesz F, Bóka K, Kriszt B. (2021). *Hydrogenophaga aromaticivorans* sp. nov., isolated from a para-xylene-degrading enrichment culture, capable of degrading benzene, meta-and para-xylene. *International Journal of Systematic and Evolutionary Microbiology*,**71**:004743, doi:10.1099/ijsem.0.004743.