Report

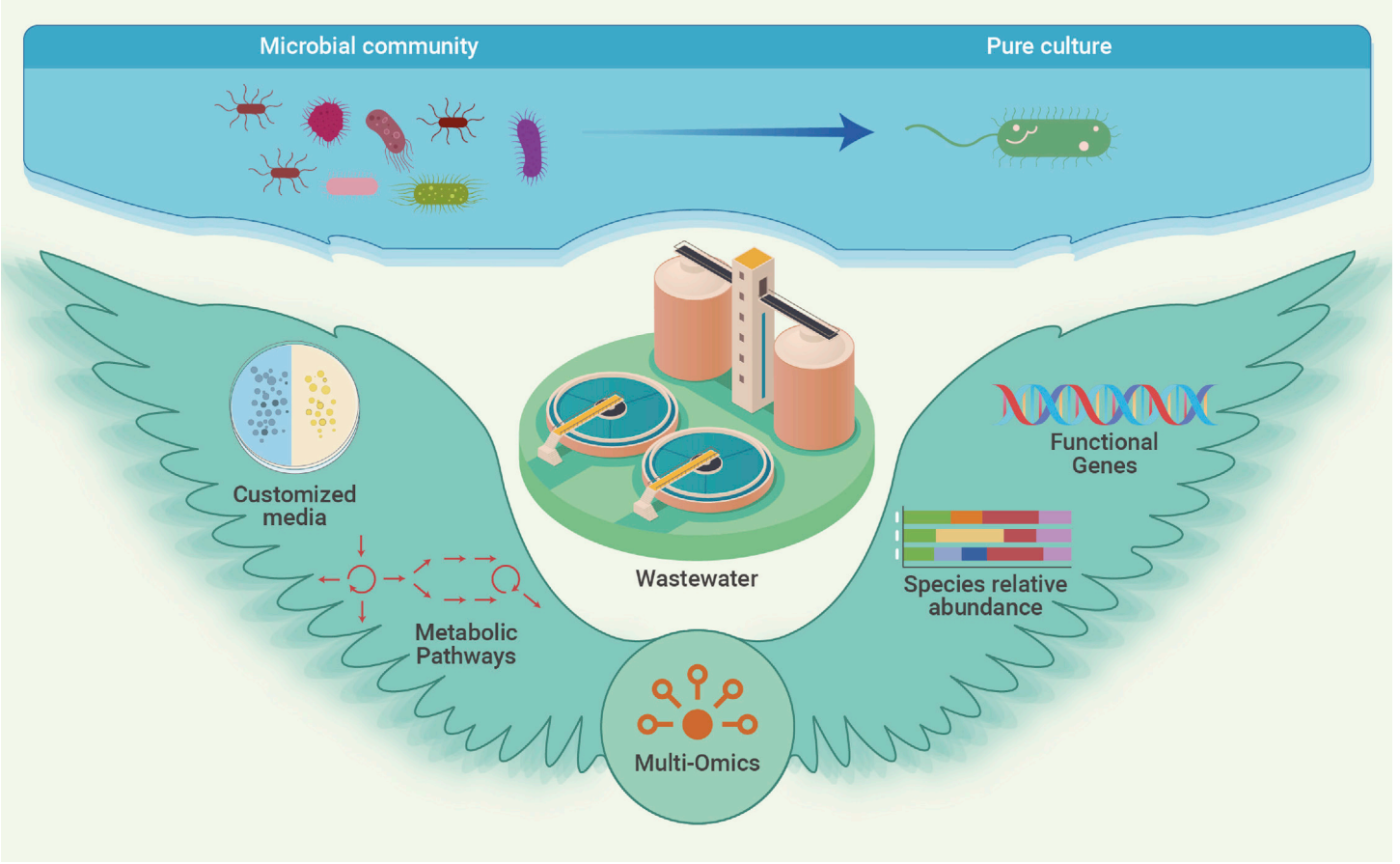
www.the-innovation.org

Community-integrated multi-omics facilitates screening and isolation of the organohalide dehalogenation microorganism

Yiqun Huang,[1] Lingyu Wen,[1] Lige Zhang,[1] Jijun Xu,[2] Weiwei Wang,[1] Haiyang Hu,[1] Ping Xu,[1] Zhao Li,[2] and Hongzhi Tang[1,*]
*Correspondence: tanghongzhi@sjtu.edu.cn

Received: April 3, 2022; Accepted: November 17, 2022; Published Online: November 22, 2022; https://doi.org/10.1016/j.xinn.2022.100355

GRAPHICAL ABSTRACT



PUBLIC SUMMARY

- Many pollutant-degrading microorganisms are hidden in complex environmental microbial communities
- The culture-independent multi-omics approach provides information about the potential key microorganisms and functional genes in industrial saponification wastewater
- The multi-omics approach guides the design of a culture method to obtain the organohalide dehalogenation microorganism

Report

Community-integrated multi-omics facilitates screening and isolation of the organohalide dehalogenation microorganism

Yiqun Huang,[1] Lingyu Wen,[1] Lige Zhang,[1] Jijun Xu,[2] Weiwei Wang,[1] Haiyang Hu,[1] Ping Xu,[1] Zhao Li,[2] and Hongzhi Tang[1,*]

[1]State Key Laboratory of Microbial Metabolism, Joint International Research Laboratory of Metabolic and Developmental Sciences, and School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China

[2]Befar Group Co., Ltd., Binzhou, Shandong 256619, China

*Correspondence: tanghongzhi@sjtu.edu.cn

Received: April 3, 2022; Accepted: November 17, 2022; Published Online: November 22, 2022; https://doi.org/10.1016/j.xinn.2022.100355

Citation: Huang Y., Wen L., Zhang L., et al., (2023). Community-integrated multi-omics facilitates screening and isolation of the organohalide dehalogenation microorganism. The Innovation **4(1)**, 100355.

A variety of anthropogenic organohalide contaminants generated from industry are released into the environment and thus cause serious pollution that endangers human health. In the present study, we investigated the microbial community composition of industrial saponification wastewater using 16S rRNA sequencing, providing genomic insights of potential organohalide dehalogenation bacteria (OHDBs) by metagenomic sequencing. We also explored yet-to-culture OHDBs involved in the microbial community. Microbial diversity analysis reveals that *Proteobacteria* and *Patescibacteria* phyla dominate microbiome abundance of the wastewater. In addition, a total of six bacterial groups (*Rhizobiales*, Rhodobacteraceae, *Rhodospirillales*, *Flavobacteriales*, *Micrococcales*, and *Saccharimonadales*) were found as biomarkers in the key organohalide removal module. Ninety-four metagenome-assembled genomes were reconstructed from the microbial community, and 105 hydrolytic dehalogenase genes within 42 metagenome-assembled genomes were identified, suggesting that the potential for organohalide hydrolytic dehalogenation is present in the microbial community. Subsequently, we characterized the organohalide dehalogenation of an isolated OHDB, *Microbacterium* sp. J1-1, which shows the dehalogenation activities of chloropropanol, dichloropropanol, and epichlorohydrin. This study provides a community-integrated multi-omics approach to gain functional OHDBs for industrial organohalide dehalogenation.

INTRODUCTION

Organohalides are a series of organic compounds containing fluorine, chlorine, or bromine.[1] Many organohalides are produced in industrial and agricultural processes, including trichloroethene, trichloroethanes, and triclosan.[2,3] Organohalides can cause pollution due to their release into the atmosphere or discharge into the environment in industrial wastewater. Organohalides can bioaccumulate through the food chain to damage the human body. Organohalides present a toxic risk because of their endocrine-disrupting effects on animals, making them a serious threat to human health and the environment.[3]

Several organohalides are produced as byproducts (eg, dichloropropane, dichloroisopropyl ether, and chloropropanol) of propylene oxide production.[4] Propylene oxide saponification wastewater also has the characteristics of strong alkalinity, high salt, and high solid suspended matter, which makes the degradation of organohalides in saponification wastewater quite challenging.

A two-phase biotechnological treatment technology, which consists of activated sludge and contact oxidation, is used to treat organohalides that occur in propylene oxide saponification wastewater (Figure 1A). The aeration tank provides sufficient oxygen for the growth of the activated microbial community in wastewater and the degradation of organic matter. Microorganisms are allowed to flow into the contact oxidation tank to form a biofilm, which continuously reduces the organic matter of the wastewater. An inappropriate microbial community will directly influence the final wastewater treatment efficiency.[5]

Classical cultivation methods (eg, selective nutrient or inhibitors, special physicochemical conditions, and size- or density-based separation) are commonly used to isolate microorganisms from the environment. However, these approaches have failed to identify some populations of microorganisms that are either unculturable or difficult to cultivate. The limitations of classical cultivation methods include lack of knowledge about metabolism (eg, substrate, growth factor, and dormancy), interaction (eg, symbiotic interdependencies), and abundance in the microbial community (eg, rare species).[6]

Culture-independent amplicon and metagenomic sequencing techniques can improve our understanding of environmental microbial communities. For example, the global core bacterial community of wastewater treatment plants includes 28 operational taxonomic units (OTUs) that were identified using 16S rRNA analysis.[7] In addition, metagenomic analysis has been used to identify novel dehalogenases and cofactor synthetic enzymes present in the trichloroethene-degrading microbial community.[8] The near-complete or draft genomes of uncultured bacteria were obtained using *de novo* assembled metagenomic sequences, which may provide the possibility to link taxa with metabolic genes in environmental microbial communities.[9] The genome-centric investigation of microorganisms can broaden our understanding of their potential metabolic functions; however, it is also necessary to isolate and culture microorganisms in order to define their microbial features with respect to metabolism and physiology in the real environment for bioremediation.

To circumvent the limitations associated with current cultivation methods, we expanded a multi–omics-guided design of selective media and built a simulated environment suitable for the growth of potential organohalide dehalogenation bacteria (OHDBs); the design was based on the substrate demands and growth factors predicted from the draft genomes of OHDBs in combination with the metabolomic data during a complete wastewater treatment process. In this study, we focused on potential OHDBs for organohalide dehalogenation from the wastewater of a propylene oxide manufacturing plant in Shandong Province, China. 16S rRNA gene sequencing was used to profile the dynamics of the microbial community and to distinguish the microbes that were dominant in key treatment modules. Previous studies of diverse environments like sewage effluent and animal rumen established analyses based on the existence of functional genes for specific substances like antibiotics or lignocellulosic polymers.[10,11] These studies provide analysis directions for understanding the degradation capacity of metagenome-assembled genomes (MAGs). In our present study, MAGs were constructed based on metagenomic reads from various phases of the wastewater treatment process, which were used to evaluate the microbial community's functional capability. An OHDB strain was isolated from the wastewater microbial community via an expanded selective media strategy. Moreover, we performed a series of organohalide dehalogenation capacity characterizations on the OHDB to uncover the potential for organohalide dehalogenation.

RESULTS AND DISCUSSION

Spatiotemporal variability of wastewater microbial community

Samples were taken from five distinct phases of the wastewater treatment process. *Proteobacteria*, *Patescibacteria*, *Actinobacteria*, and *Bacteroidetes* were the dominant phyla in the treatment process (Figure 2A). The bacterial diversity communities of the wastewater treatment process were analyzed using the Chao1 index (Figure 2C). Relatively higher microbial alpha diversity values were found for the raw water (RW) and final sedimentation tank (FST) phases; a sharp decrease in alpha diversity occurred when the wastewater moved from RW to aeration tank (AT). Significant differences in microbial abundance were found in the different phases of wastewater (Adonis, $p < 0.001$); this was also reflected in the detected variations in community composition (β-diversity; Figure S2).

Changes in environmental factors will influence the structure of microbial communities,[12] which were reflected upon the change of temperature (50°C to 35°C–38°C) and pH (11.7 to 6.5–6.9) between RW and AT. Within the top 15 ranking microbial orders (Figure 2B), *Saccharimonadales* was the most abundant *Patescibacteria* species during the AT phase. Studies on the specific physiological

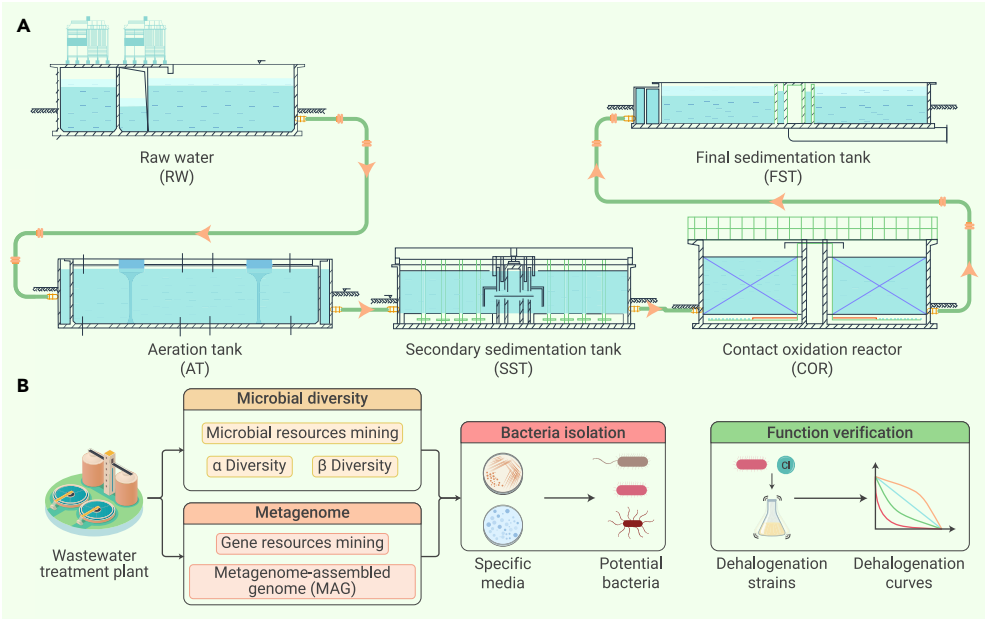


Figure 1. Overview of the wastewater treatment process and experimental design for the multi-omics analyses (A) Schematic of the two-phase biotechnological treatment technology used for saponification wastewater. The aeration tank (AT) used activated sludge for the degradation of organic matter. The secondary sedimentation tank (SST) was used to recycle activated sludge from the aeration tank. The contact oxidation reactor (COR) utilized biofilms to degrade organic matter. RW, raw water; FST, final sedimentation tank. (B) Multi-omics analysis workflow: microbial diversity and metagenomics analyses of the wastewater treatment plant were performed to obtain microbial and genetic profiles that were used as a guide for isolating potential organohalide dehalogenation bacteria and to carry out assays to evaluate their dehalogenation capacity.

ecology of *Saccharimonadales* are scarce, with only a few reports about phosphorus-mining bioindicators and sulfonamide-resistant bacteria available.[13,14]

Compared with other phases (RW, secondary sedimentation tank [SST], and FST), AT and contact oxidation reactor (COR) showed a more efficient chemical oxygen demand removal ability (Table S1). Thus, identifying the functional taxa of the AT and COR microbial communities should reveal microbial resources for degrading organic chemicals. The linear discriminant analysis effect size (LEfSe) method was used to identify bacterial taxa with significant differential abundance throughout the various phases of the wastewater treatment process (Figure 2D). A total of 17 bacterial taxa were found statistically differentially abundant in the AT and COR phases: five bacteria taxa were found in the AT phase, and 12 bacterial taxa were found in the COR phase. Phylogenetic analysis based on the 16S rRNA gene revealed that these taxa belong to six groups of microorganisms: *Flavobacteriales*, *Rhodospirillales*, *Rhizobiales*, *Micrococcales*, *Rhodobacteraceae*, and *Saccharimonadales* (Figure 2D). We classified these six groups as putative key strains, which were the focus of our follow-up metagenomic sequencing and genome assembly efforts seeking to characterize their dehalogenation potential. The LEfSe results for the taxa of other phases are presented in Figure S1.

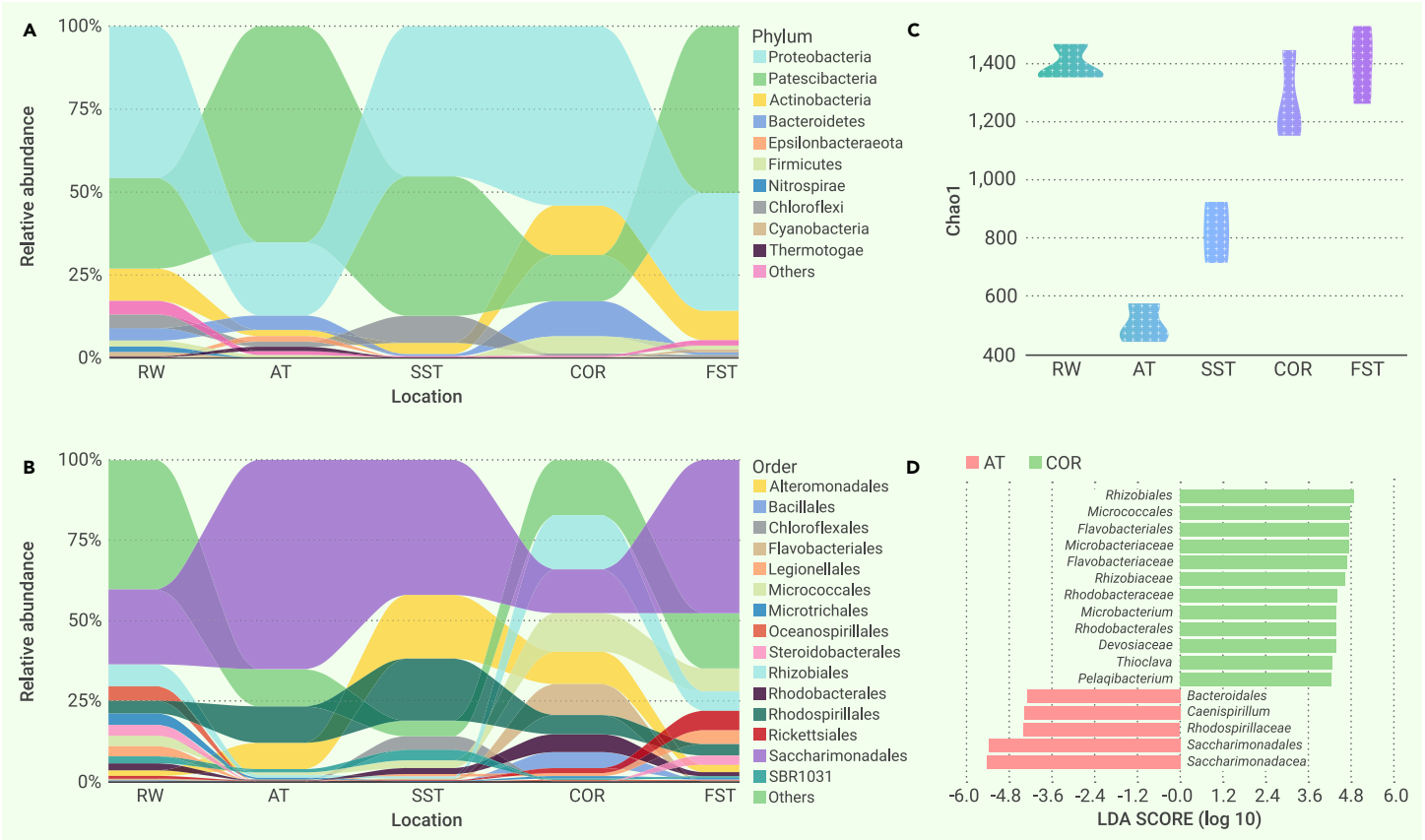


Figure 2. Microbial diversity in distinct phases of the wastewater treatment process (A) Spatiotemporal variability of the top 10 ranking (relative abundance) phyla during the wastewater treatment process (n = 3). (B) Spatiotemporal variability of the top 15 ranking (relative abundance) orders during the wastewater treatment process (n = 3). (C) Alpha diversity of the microbial community during the different phases of the wastewater treatment process. (D) Linear discriminant analysis effect size (LEfSe) analysis of the microbial community during the wastewater treatment process (linear discriminant analysis score>4.2, $p < 0.05$). These genus with significant differential abundance among distinct phases of the wastewater treatment process.

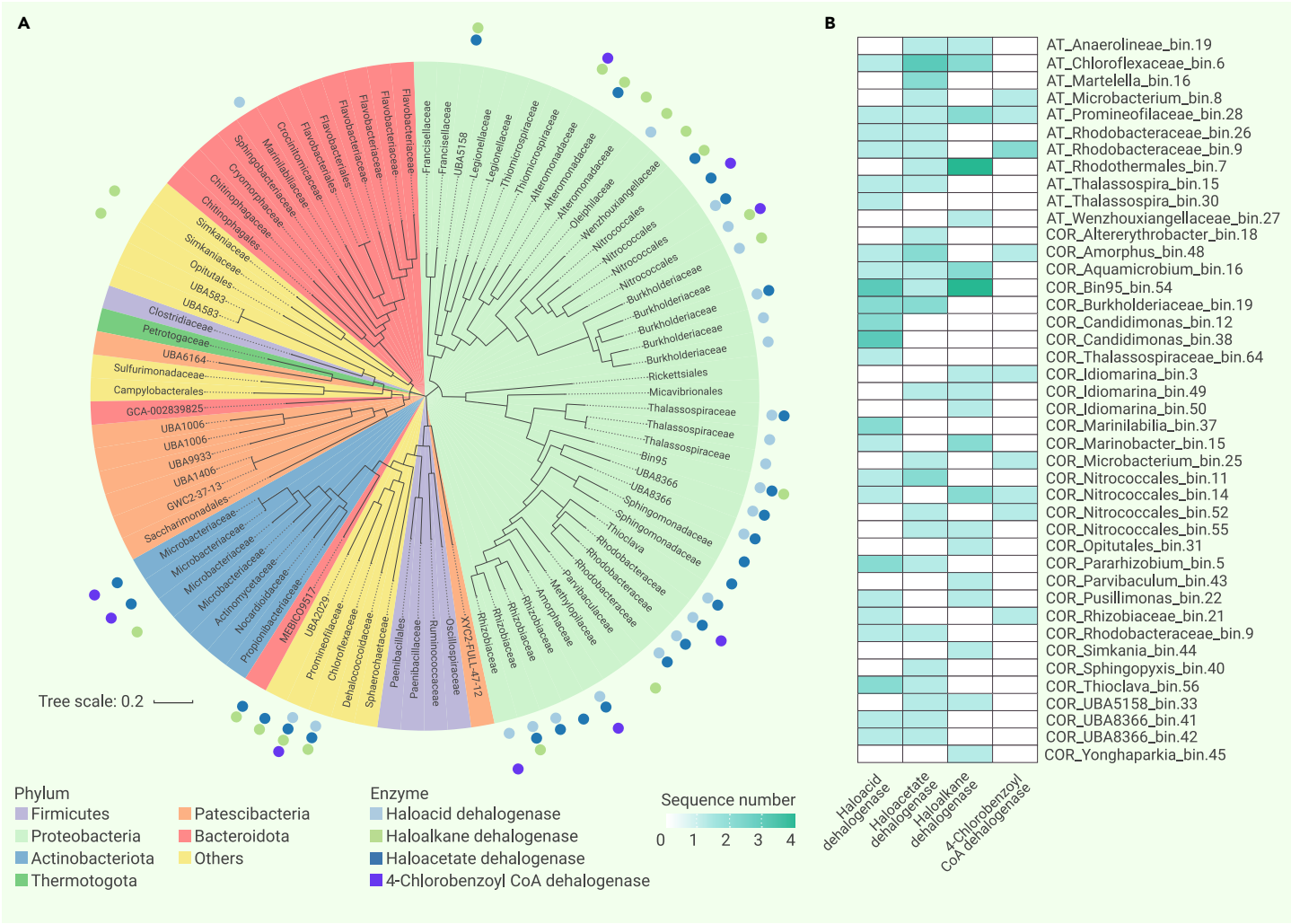


Figure 3. The phylogeny of MAGs and identification of functional genes (A) Phylogenetic identification of metagenome-assembled genomes (MAGs). These MAGs were selected using CheckM (completeness >70%, contamination <5%) and identified using the Genome Taxonomy Database. Distribution of potential dehalogenases is illustrated by the dots. (B) Potential dehalogenase sequences from the MAG analysis. The numbers of potential dehalogenase sequences in the various MAGs are indicated with the heatmap values (from 0 to 4).

OHDBs previously reported under anaerobic conditions (*Dehalococcoides*, *Dehalogenimonas*, and *Desulfitobacterium*) were searched for within the different phases of the wastewater treatment process.[15–17] In the COR phase, we found that *Dehalogenimonas* was the most abundant microorganism among these known anaerobic OHDBs; however, its abundance was still less than 0.03%. This indicates that reductive dehalogenation under anaerobic conditions is apparently not the main process through which organohalides are degraded in this wastewater treatment.

Phylogeny of MAGs and identification of functional genes

A total of 94 MAGs were obtained and filtered based on metagenomic reads (Figure 3A), which contained MAGs of AT (30) and COR (64). In addition, we found 20 near-perfect MAGs (completeness >95%, contamination <1%) within these MAGs. All MAGs (completeness >70%, contamination <5%) were used for subsequent functional gene and metabolism analyses. The detailed information for completeness and contamination of each MAG quality is shown in Table S3.

The Genome Taxonomy Database toolkit was used to analyze MAG taxonomic assignments, and it was found that *Proteobacteria*, *Bacteroidetes*, *Patescibateria*, and *Actinobacteria* were the dominant phyla (Figure 3A; Tables S4 and S5).[18] A phylogenetic tree was constructed using the UBCG pipeline,[19] which uses a bacterial core gene set to replace 16S rRNA genes, and the tree was visualized using iTOL,[20] with the coloring of the tree clades based on phylum-level taxonomy.

The chloroalkanes and chlorohydrins are the main pollutants in the saponification wastewater treatment process.[4] It is important to characterize dehalogenases and to isolate novel OHDBs for improving the organohalide bioremediation capacity. Compared with the substrate promiscuity of oxidative dehalogenase, hydrolytic dehalogenase shows a better substrate specificity, which makes it more favorable for investigation.[21,22] Hydrolytic dehalogenases include haloacid dehalogenases; haloalkane dehalogenases; haloacetate dehalogenases; 4-chlorobenzoyl coenzyme A (4-CBCoA) dehalogenases; and halohydrin dehalogenases.[21,22] To infer the metabolic potential of MAGs, these hydrolytic dehalogenase genes and specific metabolism were scanned via Prokka and KAAS analysis.[23,24]

A total of 105 candidate hydrolytic dehalogenase genes were detected in 42 MAGs (Table S6), including 33 identified as haloacid dehalogenases; 32 as haloacetate dehalogenases; 30 as haloalkane dehalogenases; and 10 as 4-CBCoA dehalogenases; halohydrin dehalogenase was not present. The highest number of candidate dehalogenases was from *Proteobacteria*, indicating that *Proteobacteria* may play an important role in organohalide removal (Figure 3A). The detailed distribution of dehalogenases observed were mainly encoded by *Rhizobiaceae*, *Rhodobacteraceae*, *Rhodospirillales*, *Micrococcales*, *Sphingomonadales*, *Chloroflexia*, and *Burkholderiales* and are shown in Figure 3B and Table S6.

Three kinds of predicted dehalogenases were found among the *Rhizobiaceae* MAGs: three haloacid dehalogenases, one haloacetate dehalogenase, and one 4-CBCoA dehalogenase. *Rhizobiaceae* reportedly participates in organohalide degradation.[25] Examination of three *Rhodobacterales* MAGs revealed five putative haloacid dehalogenases, four haloacetate dehalogenases, and two 4-CBCoA dehalogenases. The L-haloacid dehalogenase gene (*DehRhb*) and the

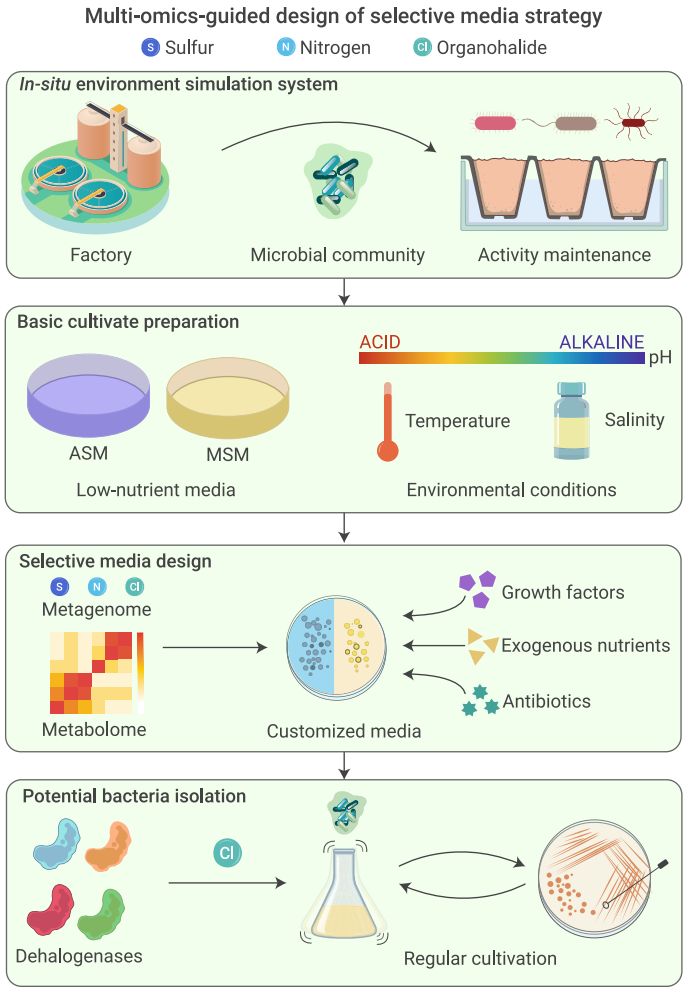


Figure 4. The OHDBs from the multi–omics-guided design of selective media strategy
Stage one established an *in situ* environment simulation system to maintain the activity of the microbial community. Stage two included the preparation of the culture media and physicochemical conditions for further customization. Stage three included the addition of supplementary substances (eg, growth factor, nutrient, and antibiotic) to the media under the guidance of multi-omics data. In stage four, the potential substrates of dehalogenases and the microbial community from the environment simulation system were added to the customized media for subsequent regular cultivation.

chlorothalonil hydrolytic dehalogenase gene (*chd*) were identified in *Rhodobacterales* and showed organohalide dehalogenation activity with brominated acid, chlorinated acid, and chlorothalonil.[26,27] *Rhodospirillales* include three MAGs that contained three putative haloacid dehalogenases and one haloacetate dehalogenase. Pentachlorophenol was degraded by *Rhodospirillales* (*Azospirillum*) under aerobic conditions.[28] There were two putative haloacetate dehalogenases, one haloalkane dehalogenase, and two 4-CBCoA dehalogenases in our *Micrococcales* MAGs. None of the *Flavobacteriales* MAGs contained any dehalogenase sequences. *Flavobacteriales* were found as a significant genus of floc-forming bacteria in flocculent activated sludge.[29]

Many microorganisms exist in the environmental microbial populations with low abundance, but they may still have a significant influence on substrate degradation; the growth rate of strains and substrate degradation efficiency are not necessarily related.[6] Some bacteria with relatively low abundance were found to contain putative dehalogenases. For example, four kinds of dehalogenases were found in MAGs among *Hyphomicrobiales*, *Sphingomonadales*, and *Chloroflexia* (Table S6). Notably, we detected eight predicted haloacid dehalogenases, two haloacetate dehalogenases, and one haloalkane dehalogenase in *Burkholderiales* MAGs. *Chloroflexia* are organohalide-respiring bacteria that utilize organohalides via reductive dehalogenases rather than hydrolytic dehalogenases.[30,31] Several studies also reported the presence of hydrolytic dehalogenases within *Hyphomicrobiales*, *Sphingomonadales*, and *Burkholderiales*, particularly haloalkane dehalogenases (*LinB*) and haloacid dehalogenases.[32–36] These microorganisms may function as organohalide degraders during the wastewater treatment process.

Multi–omics-guided design of selective media strategy

Our predictive analysis of the MAGs revealed potential dehalogenation capacity for *Rhizobiaceae*, *Rhodobacteraceae*, *Rhodospirillales*, *Micrococcales*, *Chloroflexia*, *Sphingomonadales*, and *Burkholderiales*. Experimental validation of potential OHDBs is required to confirm dehalogenation capacity. There remains a necessity to overcome the obstacles of pure culture isolation, which include (1) the identification of substrates and environmental conditions (temperature, salinity, pH); (2) a requirement for certain growth factors (eg, specific sulfur and nitrogen compounds, metal ions, etc.); and (3) the enrichment of some rare functional taxa in the environment that may have symbiotic interdependencies.[6,37,38] The *in silico* metabolic reconstruction from MAG data can help in identifying potential deficiencies or requirements for successful culture of targeted microorganisms. That is, assessing the genome of a strain can enable inference of which combinations of targeted nutrients should be included in an artificial selective media. For example, the absence of amino acid biosynthesis genes or sulfate metabolic pathways would suggest a requirement for specific nutrients.[39] Moreover, metabolomics can be an informative method for investigating metabolites in the environment that can support identification of intermediate metabolites or biomarkers and provide insights into microbial physiology.[40] These approaches can reduce trial and error and thereby reduce the time and effort required to determine suitable OHDBs culture conditions.

An illustration of the four stages multi–omics-guided design of selective media strategy we employed is shown in Figure 4. During stage one, an *in situ* environment simulation system was established to maintain the microbial activity of the *in situ* microbial community in the original wastewater samples. Activated sludge was added to the activated sludge bioreactor with an aeration device, after which fresh wastewater was periodically added (based on sludge retention time) to remove additional sludge and maintain the stability of the simulated system. These samples were used for subsequent strain isolation. During stage two, the low-nutrient medium (eg, artificial seawater medium, mineral salt medium) was used as the basic medium to support further alterations toward a customized medium. The culture conditions also matched the environmental conditions of the wastewater treatment phases, including 4% salinity, pH 6.5–6.9, and 30°C or 37°C (Table S1).

In stage three, the nitrogen and sulfur pathways based on MAGs were used to predict the potential metabolic capabilities of bacteria.[9] The metabolic capabilities of potential dehalogenation MAGs for nitrogen and sulfur metabolism are shown in Figure 5A. All 42 of the potential dehalogenation MAGs displayed potential for nitrogen fixation, nitrification, and thiosulfate oxidation. Notably, the Rhodobacteraceae_bin.26 genome contained all the relevant genes of the denitrification pathway. The majority of the potential dehalogenation MAGs (>70%) encoded genes related to nitrate or sulfur reduction.

The absence of metabolic pathway genes within identified MAGs may be due to metabolism deficiency of natural microorganisms or draft MAGs. We analyzed metabolic pathways of the potential OHDBs to identify the missing pathways and to determine whether supplemental nutrients should be added to the basic medium (Figure 5A). The SAR11 alpha-proteobacterial clade added exogenous sulfur sources necessary for growth, as evidenced by the deficient assimilation of sulfate reduction in metagenomic datasets. Additional sulfur metabolism supplementation (eg, methionine, cysteine, or thiosulphate) can be added to a customized medium.[41]

The antibiotic resistance genes detected in the genomes can be used to isolate target organisms, and specific antimicrobial compounds can be used to inhibit the rapid growth of non-target microorganisms.[42] The potential MAG antibiotic resistance genes were predicted using CARD (Table S8). The *qacJ* and *qacG* (multi-drug resistance efflux pump of quaternary ammonium compounds) genes were found in some MAGs of potential OHDBs (*Rhodospirillales*, *Rhodobacterales*, and *Burkholderiales*).[43,44] In contrast, there was an absence of antibiotic resistance genes in the MAGs of *Micrococcales*, *Hyphomicrobiales*, and *Chloroflexia*.

A non-targeted metabolomic was used to detect different metabolites in the wastewater treatment process, with Spearman's rank correlation coefficient values used to assess the relationships between samples from different phases (Figure 5B). These data indicate that tyrosine metabolism and oxidative phosphorylation are significantly different when wastewater was transferred from

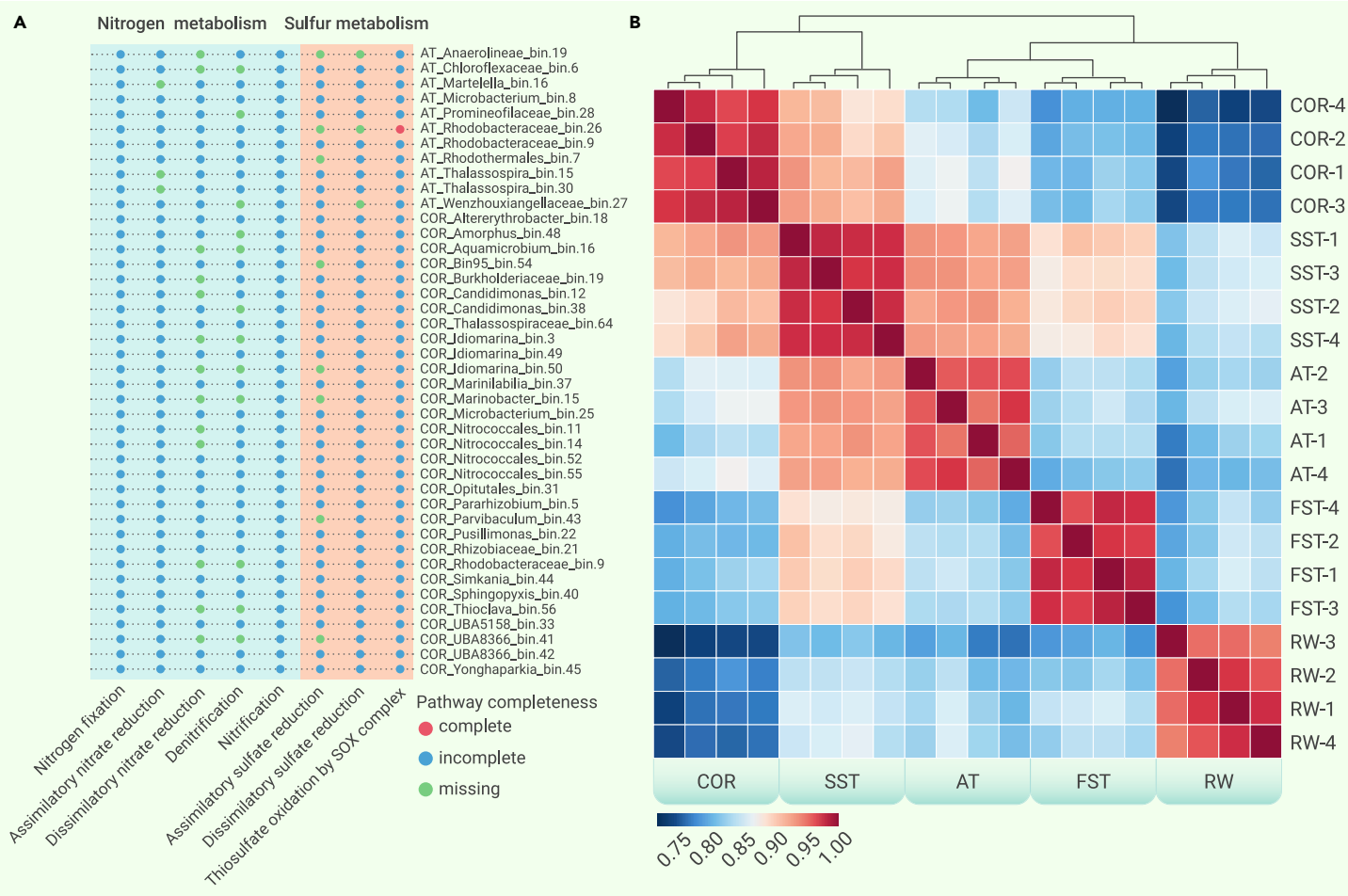


Figure 5. The assessment of MAG metabolic capacities (A) Nitrogen and sulfur metabolism content in MAGs. KEGG pathways are shown in dot form: red dot, the pathway is complete; blue dot, the pathway is incomplete; green dot, the pathway is missing. (B) Spearman's correlation analyses of the various phases of the wastewater treatment process. Cluster analysis was performed on each group of samples, which is displayed at the top of the figure (n = 4).

RW to AT. The data also indicate that pantothenate and CoA biosynthesis, as well as alanine, aspartate, and glutamate metabolism, were significantly different as the wastewater was moved from AT to COR (Table S7). Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis of the differential metabolite data indicate that succinate is an influencing factor in both the AT and COR phases.

At stage four, we considered known or suspected substrates of hydrolytic dehalogenases.[21,45–47] The organohalide substrate in our attempts to enrich and isolate OHDBs included dichloropropane, monochloroacetic acid, methyl chloroacetate, and dichloropropanol. The customized medium with specific organohalides was inoculated with the wastewater sample for shaking culture, followed by streak plate on specific media plates. The bacterial clones grown on the plates were picked for the next round of cultivation.

Characterization for organohalide dehalogenation of the isolated OHDB strain

Using this multi-omics strategy to adjust the selective medium to support the growth of OHDBs, *Microbacterium* sp. J1-1 was isolated from a customized mineral salt medium that contained sodium thiosulfate, disodium succinate, and a substrate mixture including chloroalkane and chlorohydrin without additional antibiotic. The phylogenetic tree of *Microbacterium* sp. J1-1 based on 16S rRNA is shown in Figure 6A. The morphological characteristics of *Microbacterium* sp. J1-1 are shown via a scanning electron microscope micrograph (Figure 6B).

We sequenced the genome of *Microbacterium* sp. J1-1 and evaluated the presence of candidate hydrolytic dehalogenase genes (Figure 6C). Our assembled genome for *Microbacterium* sp. J1-1 had each of the hydrolytic dehalogenase types that were annotated in the *Micrococcales* MAG. Additionally, two haloacid dehalogenases and one halohydrin dehalogenase were found in *Microbacterium* sp. J1-1. The genome of pure culture OHDB shows more dehalogenases. This may be related to (1) the incompleteness of *Micrococcales* MAGs (AT_bin.8, COR_bin.25, and COR_bin.45), as we obtained <95% completeness and thus might have missed some functional genes; (2) a portion of the functional genes remain uncharacterized in reference databases or are instead annotated as belonging to a broad protein family (eg, halohydrin dehalogenase belongs to the short-chain dehydrogenase/reductases)[48]; and (3) the results of the genome binning tools (eg, MetaBat2, MaxBin2, and CONCOCT) are affected by the complexity of the metagenomic dataset and generate a different number and quality of bins for subsequent analyses.[49]

The dehalogenation ability of *Microbacterium* sp. J1-1 was determined. Specifically, *Microbacterium* sp. J1-1 possesses organohalide dehalogenation capacities for chloropropanol, dichloropropanol, and epichlorohydrin (Figure 6D). The putative halohydrin dehalogenase HheA_J1 (Figure S5) from *Microbacterium* sp. J1-1 shows a high amino acid identity with HheA (98.77%, GenBank: BAA14361.1) and HheA_AD2 (98.36%, GenBank: AAK92100.1), two proteins with reported capacity to degrade chlorohydrin.[50,51] We also used organohalides, which are the potential substrates of other dehalogenases in the dehalogenation assay, including monochloroacetic acid, 2-chloropropionic acid, methyl chloroacetate, methyl 2-chloropropionate, and methyl 4-chlorobenzoate. However, none of these were degraded by *Microbacterium* sp. J1-1. *Microbacterium* sp. J1-1 is a candidate in the application of further saponification wastewater treatment process.

Conclusion

In the present study, we profiled the dynamics of the microbial community and identified microbes that were dominant in the wastewater treatment process. *Proteobacteria*, *Patescibacteria*, *Actinobacteria*, and *Bacteroidetes* were the

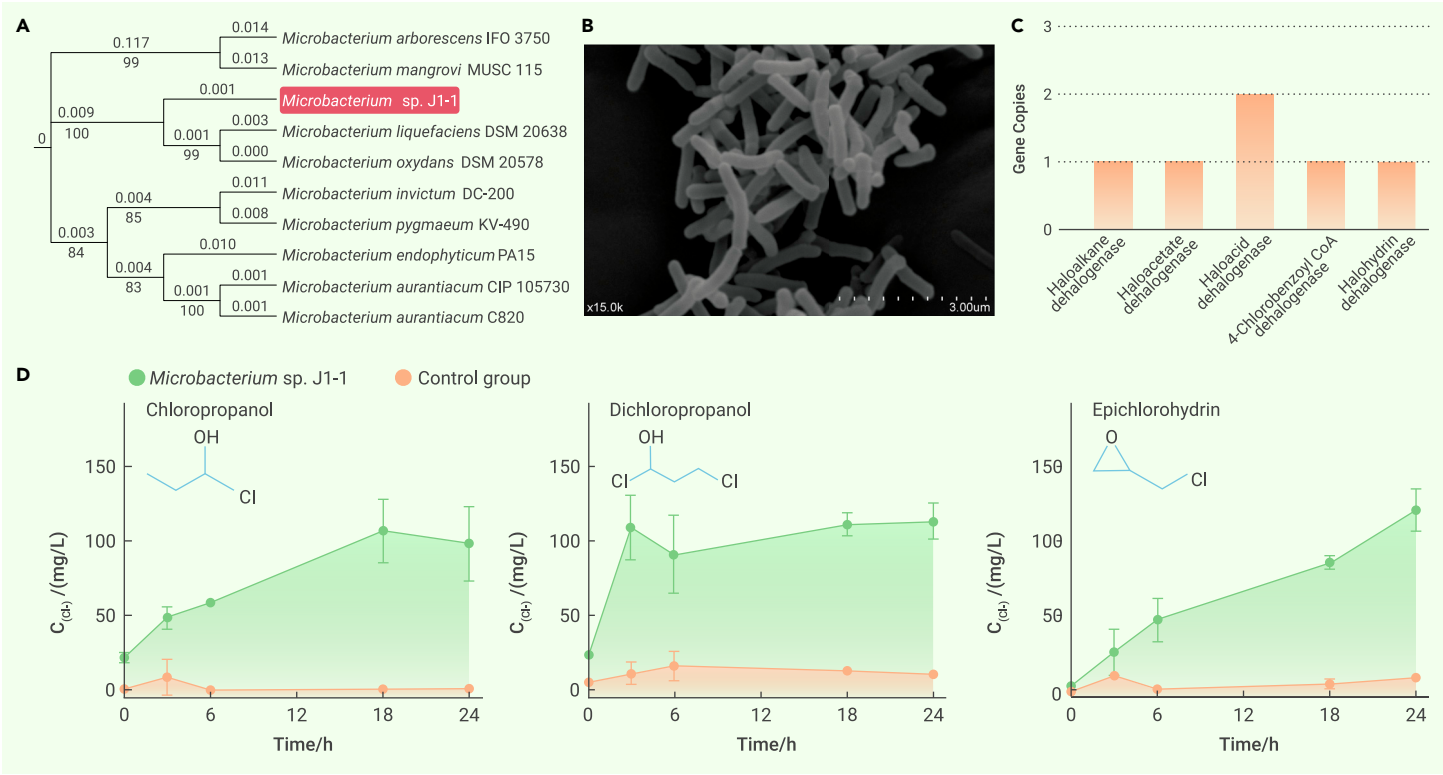


Figure 6. Organohalide dehalogenation assay of *Microbacterium* sp. J1-1 (A) Phylogenetic tree based on 16S rRNA sequencing. (B) Scanning electron microscope micrograph of *Microbacterium* sp. J1-1. (C) The distribution of dehalogenases in *Microbacterium* sp. J1-1. (D) The dehalogenation assay of the resting cell system. Green: the dehalogenation of 1000 μl/l organohalide using *Microbacterium* sp. J1-1. Orange: the dehalogenation of 1000 μl/l organohalide without bacteria (n = 3, error bar represents the standard error of the mean).

dominant phyla throughout the overall wastewater treatment process. Seven bacterial groups' MAGs (*Rhizobiaceae*, Rhodobacteraceae, *Rhodospirillales*, *Micrococcales*, *Chloroflexia*, *Sphingomonadales*, and *Burkholderiales*), which have various hydrolytic dehalogenase genes, show potential for dehalogenation in propylene oxide wastewater. Moreover, we subsequently harnessed the multi-omics data and developed a four-stage selective media strategy to facilitate the isolation of OHDBs, which we used to successfully isolate *Microbacterium* sp. J1-1. After genome assembly for the isolated OHDB, we found that the strain has diverse hydrolytic dehalogenases and possesses the capacity for organohalide dehalogenation. The OHDB is a candidate for further saponification wastewater treatment process. Combining multi-omics prediction information and microbial cultivation will be helpful for obtaining strain resources and genetic resources in more environments in the future.

MATERIALS AND METHODS

Sample collection and DNA extraction

Samples from a WWTP in Shandong Province of China were collected from the following five phases: RW, AT, SST, COR, and FST (Figure 1A). The AT and SST phases comprise the activated sludge module, while the COR and FST phases comprise the contact oxidation module. The wastewater samples have two characteristics of interest: (1) the samples contain various chlorinated organic substances, such as dichloropropane, dichloroisopropyl ether, and chloroacetone, and (2) the water samples have a chloride ion concentration of 23 000–26 000 mg/l, and this level is maintained throughout the wastewater treatment process. More detailed physical and chemical information for each of the different treatment phases can be found in Table S1.

Genomic DNA was extracted from the samples using a DNeasy PowerSoil Kit (QIAGEN), and a NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific) was used to measure and evaluate the quantity and quality of the extracted DNA. DNA samples were stored at −20°C.

16S rRNA gene sequencing and metagenome shotgun sequencing

For 16S rRNA gene sequencing, the forward primer 515F (5′-GTGCCAGCMGCCGCGGTAA-3′) and the reverse primer 907R (5′-CCGTCAATTCMTTTRAGTTT-3′) were used to target the V4–V5 region. For multiplex sequencing, specific 7-bp barcodes were incorporated into the primers for the various samples. 300-bp paired-end sequencing of the libraries was conducted on the Illumina MiSeq platform using a MiSeq Reagent Kit v.3 (Illumina).

For metagenome shotgun sequencing, the Illumina TruSeq Nano DNA LT Library Preparation Kit was used to construct metagenome shotgun sequencing libraries. The libraries were then sequenced using the PE150 strategy on the Illumina HiSeq X-ten platform. The multi-omics analysis workflow is shown in Figure 1B and was created with BioRender.com.

Sequence analysis

For 16S rRNA gene sequencing, the Quantitative Insights into Microbial Ecology (v.1.8.0) tool was used to denoise and filter the reads data. The OTUs were clustered with a 97% qualified sequence identity threshold.[52,53] The taxonomic classifications of OTUs were selected, and the best hit came from the Greengenes Database (release 13.8).[54] OTUs with relative abundance below 0.001% among all samples were removed.

For metagenomic shotgun sequencing, raw reads were processed by FastQC (v.0.11.8), Cutadapt (v.1.2.1), and BWA (v.0.7.17).[55] Metagenomes were constructed based on qualified reads using IDBA-UD (release 1.1.3). Metagenomic scaffolds (length >300 bp) were used to predict the coding regions based on MetaGeneMark (v.3.26).[56,57] With a sequence identity (>90%), coding regions were clustered by CD-HIT (v.4.7) to generate a non-redundant gene catalog.[58] DIAMOND (v.0.8.32.94) and KEGG were used to annotate the non-redundant genes catalog for functional predictions.[59,60]

Microbial diversity analysis

The relative abundance of OTUs at the phylum level were visualized by ggplot2 (v.3.3.3) and ggalluvial (v.0.12.3). The Vegan (v.2.5-7) and Picante (v.1.8.2) R packages were used to calculate the samples' alpha diversity, including the Chao1 richness estimator, the Shannon diversity index, the Abundance-based coverage estimator index, and the Simpson index. The beta diversity of samples was assessed using Bray-Curtis dissimilarity metrics by non-metric multi-dimensional scaling analysis.[61]

Identification of discriminative OTUs

LEfSe analysis was conducted to identify discriminative OTUs among the different treatment phases.[62] The logarithmic linear discriminant analysis score cutoff was set at 4.2; otherwise, default parameters were used throughout the whole pipeline. The exhaustive LEfSe result is shown in Figure S1.

Generation of MAGs

Metagenomic reads were binned into MAGs via MetaWRAP (v.1.2).[63] Initial MAGs were refined and assessed in MetaWRAP. MAGs that met the cutoff of completeness >70% and contamination <5% were used for analysis. The detailed information about completeness and contamination for the MAGs is presented in Table S3.

Phylogenetic analysis

The Genome Taxonomy Database toolkit (v.1.2) was used to make taxonomic assignments of the MAGs based on marker genes.[18] Phylogenetic trees were constructed based on the core gene set using UBCG (v.3) and visualized by iTOL.[19,20]

Metabolism annotation analysis

KO (KEGG Ortholog) of MAGs were assigned by KAAS (v.2.1), and the template gene dataset included the prokaryote representative set and closely related species.[24] The presence of specific metabolic pathways was analyzed by KEGG pathway maps (eg, nitrogen metabolism, sulfur metabolism). The antibiotic resistance genes were identified using RGI (v.5.2.1) and CARD (v3.2.3) using the "Perfect and Strict hits only" selection criteria. Functional genes (haloacid dehalogenase, haloalkane dehalogenase, haloacetate dehalogenase, and 4-CBCoA dehalogenase) were obtained from Prodigal (v.2.6.3) and Prokka (v.1.14.5).[23,64] Halohydrin dehalogenase was identified from the short-chain dehydrogenase/reductase via MAFFT (v.7.490) based on the special catalytic triads (Ser-Tyr-Arg) and the cofactor binding motif.[48] The hypothetical proteins were removed, and the predicted annotations related to dehalogenation were retained for further analysis. The metabolomics data were analyzed via MetaboAnalyst (v.5.0) and KEGG.[65]

Enrichment and isolation of OHDB

Mineral salt medium consisted of 2.0 g/l KH_2PO_4, 3.28 g/l $Na_2HPO_4{\cdot}12H_2O$, 0.1 g/l $MgSO_4$, 1.0 g/l $(NH_4)_2SO_4$, and 0.25 g/l $FeCl_3$.[66] The optional additional components included disodium succinate (10 mg/l) and sodium thiosulfate (0.1 g/l). Luria-Bertani medium for the resting cell system assay consisted of 10 g/l tryptone, 5 g/l yeast extract, and 10 g/l NaCl. Culture enrichments were incubated at 30°C and 37°C with shaking (200 RPM). Multi–omics-guided design of the selective media strategy is shown in Figure 4 and was created with BioRender.com, adapted from "Microbiology Agar Plates" (2021).

Organohalide dehalogenation determination

The strain was transferred to Luria-Bertani medium supplemented with specific organic halides to induce the dehalogenase genes at 30°C and shaking at 200 RPM. The scanning electron microscope micrograph of *Microbacterium* sp. J1-1 was obtained via Hitachi S3400N SEM. The bacterial medium was collected to prepare the resting cell system for subsequent dehalogenation experiments. The reaction medium was centrifuged, and the supernatant was transferred into a 25-ml volumetric flask. 2 ml saturated mercury thiocyanate in ethanol and 2 ml 0.25 M ferric ammonium sulfate in nitric acid were added to initiate the chromogenic reaction, and a UV spectrophotometer was used to detect the dehalogenation of organohalides.[67]

REFERENCES

1. Wang, S., Qiu, L., Liu, X., Xu, G., Siegert, M., Lu, Q., Juneau, P., Yu, L., Liang, D., He, Z., and Qiu, R. (2018). Electron transport chains in organohalide-respiring bacteria and bioremediation implications. Biotechnol. Adv. ***36***, 1194–1206.
2. Jugder, B.-E., Ertan, H., Bohl, S., Lee, M., Marquis, C.P., and Manefield, M. (2016). Organohalide respiring bacteria and reductive dehalogenases: key tools in organohalide bioremediation. Front. Microbiol. ***7***, 249.
3. Dann, A.B., and Hontela, A. (2011). Triclosan: environmental exposure, toxicity and mechanisms of action. J. Appl. Toxicol. ***31***, 285–311.
4. Nijhuis, T.A., Makkee, M., Moulijn, J.A., and Weckhuysen, B.M. (2006). The production of propene oxide: catalytic processes and recent developments. Ind. Eng. Chem. Res. ***45***, 3447–3459.
5. LaMartina, E.L., Mohaimani, A.A., and Newton, R.J. (2021). Urban wastewater bacterial communities assemble into seasonal steady states. Microbiome ***9***, 116.
6. Lewis, W.H., Tahon, G., Geesink, P., Sousa, D.Z., and Ettema, T.J.G. (2021). Innovations to culturing the uncultured microbial majority. Nat. Rev. Microbiol. ***19***, 225–240.
7. Wu, L., Ning, D., Zhang, B., Li, Y., Zhang, P., Shan, X., Zhang, Q., Brown, M.R., Li, Z., Van Nostrand, J.D., et al. (2019). Global diversity and biogeography of bacterial communities in wastewater treatment plants. Nat. Microbiol. ***4***, 1183–1195.
8. Brisson, V.L., West, K.A., Lee, P.K.H., Tringe, S.G., Brodie, E.L., and Alvarez-Cohen, L. (2012). Metagenomic analysis of a stable trichloroethene-degrading microbial community. ISME J. ***6***, 1702–1714.
9. Pan, J., Zhou, Z., Béjà, O., Cai, M., Yang, Y., Liu, Y., Gu, J.D., and Li, M. (2020). Genomic and transcriptomic evidence of light-sensing, porphyrin biosynthesis, Calvin-Benson-Bassham cycle, and urea production in *Bathyarchaeota*. Microbiome ***8***, 43.
10. Marathe, N.P., Berglund, F., Razavi, M., Pal, C., Dröge, J., Samant, S., Kristiansson, E., and Larsson, D.G.J. (2019). Sewage effluent from an Indian hospital harbors novel carbapenemases and integron-borne antibiotic resistance genes. Microbiome ***7***, 97.
11. Gharechahi, J., Vahidi, M.F., Bahram, M., Han, J.L., Ding, X.Z., and Salekdeh, G.H. (2021). Metagenomic analysis reveals a dynamic microbiome with diversified adaptive functions to utilize high lignocellulosic forages in the cattle rumen. ISME J. ***15***, 1108–1120.
12. Seib, M.D., Berg, K.J., and Zitomer, D.H. (2016). Influent wastewater microbiota and temperature influence anaerobic membrane bioreactor microbial community. Bioresour. Technol. ***216***, 446–452.
13. Zhu, L., Xu, H., Xiao, W., Lu, J., Lu, D., Chen, X., Zheng, X., Jeppesen, E., Zhang, W., and Wang, L. (2020). Ecotoxicological effects of sulfonamide on and its removal by the submerged plant *Vallisneria natans* (Lour.) Hara. Water Res. ***170***, 115354.
14. Mason, L.M., Eagar, A., Patel, P., Blackwood, C.B., and DeForest, J.L. (2020). Potential microbial bioindicators of phosphorus mining in a temperate deciduous forest. J. Appl. Microbiol. ***130***, 109–122.
15. Krzmarzick, M.J., and Novak, P.J. (2014). Removal of chlorinated organic compounds during wastewater treatment: achievements and limits. Appl. Microbiol. Biotechnol. ***98***, 6233–6242.
16. Zhao, S., Rogers, M.J., and He, J. (2020). Abundance of organohalide respiring bacteria and their role in dehalogenating antimicrobials in wastewater treatment plants. Water Res. ***181***, 115893.
17. Jing, R., and Kjellerup, B.V. (2020). Predicting the potential for organohalide respiration in wastewater: comparison of intestinal and wastewater microbiomes. Sci. Total Environ. ***705***, 135833.
18. Chaumeil, P.-A., Mussig, A.J., Hugenholtz, P., and Parks, D.H. (2019). GTDB-Tk: a toolkit to classify genomes with the genome taxonomy database. Bioinformatics ***36***, 1925–1927.
19. Na, S.-I., Kim, Y.O., Yoon, S.-H., Ha, S.M., Baek, I., and Chun, J. (2018). UBCG: up-to-date bacterial core gene set and pipeline for phylogenomic tree reconstruction. J. Microbiol. ***56***, 280–285.
20. Letunic, I., and Bork, P. (2021). Interactive Tree of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. ***49***, W293–W296.
21. Oyewusi, H.A., Wahab, R.A., and Huyop, F. (2020). Dehalogenase-producing halophiles and their potential role in bioremediation. Mar. Pollut. Bull. ***160***, 111603.
22. Agarwal, V., Miles, Z.D., Winter, J.M., Eustáquio, A.S., El Gamal, A.A., and Moore, B.S. (2017). Enzymatic halogenation and dehalogenation reactions: pervasive and mechanistically diverse. Chem. Rev. ***117***, 5619–5674.
23. Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. Bioinformatics ***30***, 2068–2069.
24. Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A.C., and Kanehisa, M. (2007). KAAS: an automatic genome annotation and pathway reconstruction server. Nucleic Acids Res. ***35***, W182–W185.
25. Wang, X., Xie, Z., Yan, J., He, X., Liu, W., and Sun, Y. (2019). Enhancement of the thermostability of halohydrin dehalogenase from *Agrobacterium radiobacter* AD1 by constructing a combinatorial smart library. Int. J. Biol. Macromol. ***130***, 19–23.
26. Novak, H.R., Sayer, C., Isupov, M.N., Paszkiewicz, K., Gotz, D., Spragg, A.M., et al. (2013). Marine *Rhodobacteraceae* L-haloacid dehalogenase contains a novel His/Glu dyad that could activate the catalytic water. FEBS J. ***280***, 1664–1680.
27. Yue, W., Xiong, M., Li, F., and Wang, G. (2015). The isolation and characterization of the novel chlorothalonil-degrading strain *Paracoccus* sp. XF-3 and the cloning of the chd gene. J. Biosci. Bioeng. ***120***, 544–548.
28. Kumari, P., Kaushik, R., Tripathi, B.M., et al. (2014). Genetic and metabolic profiling of pentachlorophenol utilizing bacteria from agricultural soil irrigated with pulp and paper mill effluent. Indian J. Agric. Sci. ***84***, 124–130.
29. Lv, J., Wang, Y., Zhong, C., Li, Y., Hao, W., and Zhu, J. (2014). The microbial attachment potential and quorum sensing measurement of aerobic granular activated sludge and flocculent activated sludge. Bioresour. Technol. ***151***, 291–296.
30. Yang, Y., Sanford, R., Yan, J., Chen, G., Cápiro, N.L., Li, X., and Löffler, F.E. (2020). Roles of organohalide-respiring *Dehalococcoidia* in carbon cycling. mSystems ***5***. e00757-19.
31. Krzmarzick, M.J., McNamara, P.J., Crary, B.B., and Novak, P.J. (2013). Abundance and diversity of organohalide-respiring bacteria in lake sediments across a geographical sulfur gradient. FEMS Microbiol. Ecol. ***84***, 248–258.
32. Skopelitou, K., Georgakis, N., Efrose, R., Flemetakis, E., and Labrou, N.E. (2013). Sol–gel immobilization of haloalkane dehalogenase from *Bradyrhizobium japonicum* for the remediation 1, 2-dibromoethane. J. Mol. Catal. B Enzym. ***97***, 5–11.
33. Adamu, A., Wahab, R.A., Aliyu, F., Aminu, A.H., Hamza, M.M., and Huyop, F. (2020). Haloacid dehalogenases of *Rhizobium* sp. and related enzymes: catalytic properties and mechanistic analysis. Process Biochem. ***92***, 437–446.
34. Sharma, P., Raina, V., Kumari, R., Malhotra, S., Dogra, C., Kumari, H., Kohler, H.P.E., Buser, H.R., Holliger, C., and Lal, R. (2006). Haloalkane dehalogenase LinB is responsible for β- and δ-hexachlorocyclohexane transformation in *Sphingobium indicum* B90A. Appl. Environ. Microbiol. ***72***, 5720–5727.
35. Ohkouchi, Y., Koshikawa, H., and Terashima, Y. (2000). Cloning and expression of DL-2-haloacid dehalogenase gene from *Burkholderia cepacia*. Water Sci. Technol. ***42***, 261–268.
36. Diez, A., Alvarez, M.J., Prieto, M.I., Bautista, J.M., and Garrido-Pertierra, A. (1995). Monochloroacetate dehalogenase activities of bacterial strains isolated from soil. Can. J. Microbiol. ***41***, 730–739.
37. Zengler, K., and Zaramela, L.S. (2018). The social network of microorganisms - how auxotrophies shape complex communities. Nat. Rev. Microbiol. ***16***, 383–390.

38. Stams, A.J.M., and Plugge, C.M. (2009). Electron transfer in syntrophic communities of anaerobic bacteria and archaea. Nat. Rev. Microbiol. ***7***, 568–577.
39. Carini, P., Steindler, L., Beszteri, S., and Giovannoni, S.J. (2013). Nutrient requirements for growth of the extreme oligotroph '*Candidatus* Pelagibacter ubique' HTCC1062 on a defined medium. ISME J. ***7***, 592–602.
40. Tang, T., Xu, Y., Wang, J., Tan, X., Zhao, X., Zhou, P., Kong, F., Zhu, C., Lu, C., and Lin, H. (2021). Evaluation of the differences between biofilm and planktonic *Brucella abortus* via metabolomics and proteomics. Funct. Integr. Genomics ***21***, 421–433.
41. Tripp, H.J., Kitner, J.B., Schwalbach, M.S., Dacey, J.W.H., Wilhelm, L.J., and Giovannoni, S.J. (2008). SAR11 marine bacteria require exogenous reduced sulphur for growth. Nature ***452***, 741–744.
42. Gutleben, J., Chaib De Mares, M., van Elsas, J.D., Smidt, H., Overmann, J., and Sipkema, D. (2018). The multi-omics promise in context: from sequence to microbial isolate. Crit. Rev. Microbiol. ***44***, 212–229.
43. Bjorland, J., Steinum, T., Sunde, M., Waage, S., and Heir, E. (2003). Novel plasmid-borne gene *qacJ* mediates resistance to quaternary ammonium compounds in equine *Staphylococcus aureus, Staphylococcus simulans*, and *Staphylococcus intermedius*. Antimicrob. Agents Chemother. ***47***, 3046–3052.
44. Heir, E., Sundheim, G., and Holck, A.L. (1999). The *qacG* gene on plasmid pST94 confers resistance to quaternary ammonium compounds in *staphylococci* isolated from the food industry. J. Appl. Microbiol. ***86***, 378–388.
45. Koonin, E.V., and Tatusov, R.L. (1994). Computer analysis of bacterial haloacid dehalogenases defines a large superfamily of hydrolases with diverse specificity. J. Mol. Biol. ***244***, 125–132.
46. Janssen, D.B. (2004). Evolving haloalkane dehalogenases. Curr. Opin. Chem. Biol. ***8***, 150–159.
47. Koudelakova, T., Chovancova, E., Brezovsky, J., Monincova, M., Fortova, A., Jarkovsky, J., and Damborsky, J. (2011). Substrate specificity of haloalkane dehalogenases. Biochem. J. ***435***, 345–354.
48. Schallmey, M., Koopmeiners, J., Wells, E., Wardenga, R., and Schallmey, A. (2014). Expanding the halohydrin dehalogenase enzyme family: identification of novel enzymes by database mining. Appl. Environ. Microbiol. ***80***, 7303–7315.
49. Yue, Y., Huang, H., Qi, Z., Dou, H.M., Liu, X.Y., Han, T.F., Chen, Y., Song, X.J., Zhang, Y.H., and Tu, J. (2020). Evaluating metagenomics tools for genome binning with real metagenomic datasets and CAMI datasets. BMC Bioinf. ***21***, 334.
50. van Hylckama Vlieg, J.E., Tang, L., Lutje Spelberg, J.H., Smilda, T., Poelarends, G.J., Bosma, T., van Merode, A.E., Fraaije, M.W., and Janssen, D.B. (2001). Halohydrin dehalogenases are structurally and mechanistically related to short-chain dehydrogenases/reductases. J. Bacteriol. ***183***, 5058–5066.
51. Yu, F., Nakamura, T., Mizunashi, W., and Watanabe, I. (1994). Cloning of two halohydrin hydrogen–halide–lyase genes of *Corynebacterium* sp. strain N-1074 and structural comparison of the genes and gene products. Biosci. Biotechnol. Biochem. ***58***, 1451–1457.
52. Lozupone, C.A., Hamady, M., Kelley, S.T., and Knight, R. (2007). Quantitative and qualitative β diversity measures lead to different insights into factors that structure microbial communities. Appl. Environ. Microbiol. ***73***, 1576–1585.
53. Edgar, R.C. (2010). Search and clustering orders of magnitude faster than BLAST. Bioinformatics ***26***, 2460–2461.
54. DeSantis, T.Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E.L., Keller, K., Huber, T., Dalevi, D., Hu, P., and Andersen, G.L. (2006). Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl. Environ. Microbiol. ***72***, 5069–5072.
55. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics ***25***, 1754–1760.
56. Peng, Y., Leung, H.C.M., Yiu, S.M., and Chin, F.Y.L. (2012). IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. Bioinformatics ***28***, 1420–1428.
57. Noguchi, H., Park, J., and Takagi, T. (2006). MetaGene: prokaryotic gene finding from environmental genome shotgun sequences. Nucleic Acids Res. ***34***, 5623–5630.
58. Fu, L., Niu, B., Zhu, Z., Wu, S., and Li, W. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics ***28***, 3150–3152.
59. Buchfink, B., Xie, C., and Huson, D.H. (2015). Fast and sensitive protein alignment using DIAMOND. Nat. Methods ***12***, 59–60.
60. Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. ***28***, 27–30.
61. Clarke, K.R. (1993). Non-parametric multivariate analyses of changes in community structure. Austral Ecol. ***18***, 117–143.
62. Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W.S., and Huttenhower, C. (2011). Metagenomic biomarker discovery and explanation. Genome Biol. ***12***, R60.
63. Uritskiy, G.V., DiRuggiero, J., and Taylor, J. (2018). MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. Microbiome ***6***, 158.
64. Hyatt, D., Chen, G.-L., LoCascio, P.F., Land, M.L., Larimer, F.W., and Hauser, L.J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinf. ***11***, 119.
65. Pang, Z., Chong, J., Zhou, G., de Lima Morais, D.A., Chang, L., Barrette, M., Gauthier, C., Jacques, P.É., Li, S., and Xia, J. (2021). MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. Nucleic Acids Res. ***49***, W388–W396.
66. Wang, W., Li, Q., Zhang, L., Cui, J., Yu, H., Wang, X., Ouyang, X., Tao, F., Xu, P., and Tang, H. (2021). Genetic mapping of highly versatile and solvent-tolerant *Pseudomonas putida* B6-2 (ATCC BAA-2545) as a 'superstar' for mineralization of PAHs and dioxin-like compounds. Environ. Microbiol. ***23***, 4309–4325.
67. Bergmann, J.G., and Sanik, J. (1957). Determination of trace amounts of chlorine in naphtha. Anal. Chem. ***29***, 241–243.

ACKNOWLEDGMENTS

This work was supported by a grant (2021YFA0909500) from the National Key R&D Program of China; by the Shanghai Excellent Academic Leaders Program (20XD1421900); by the "Shuguang Program" (17SG09) from the Shanghai Education Development Foundation and Shanghai Municipal Education Commission; and by a grant (32030004) from the National Natural Science Foundation of China.

AUTHOR CONTRIBUTIONS

Y.H. and H.T. designed the study. J.X. and Z.L. collected the samples. W.W. and L.Z. prepared amplicon sequencing and metagenomic sequencing materials. Y.H. performed the amplicon data and metagenomic data analyses. Y.H. and L.W. designed the strain media and conducted the assay. Y.H., L.W., H.H., H.T., and P.X. drafted and reviewed the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.xinn.2022.100355.

WEB RESOURCES

The raw sequence data of the 16S rRNA gene are deposited in NCBI SRA (NCBI: PRJNA785735). The metagenome data are deposited in NCBI SRA (NCBI: PRJNA786644). The bacterial genome of *Microbacterium* sp. J1-1 (NCBI: PRJNA785372) is also available from NCBI. The metabolomic data are deposited in Metabolights (Database: MTBLS3916).

LEAD CONTACT WEBSITE

https://life.sjtu.edu.cn/teacher/HongzhiTang

The Innovation, Volume *4*

Supplemental Information

Community-integrated multi-omics facilitates screening and isolation of the organohalide dehalogenation microorganism

Yiqun Huang, Lingyu Wen, Lige Zhang, Jijun Xu, Weiwei Wang, Haiyang Hu, Ping Xu, Zhao Li, and Hongzhi Tang

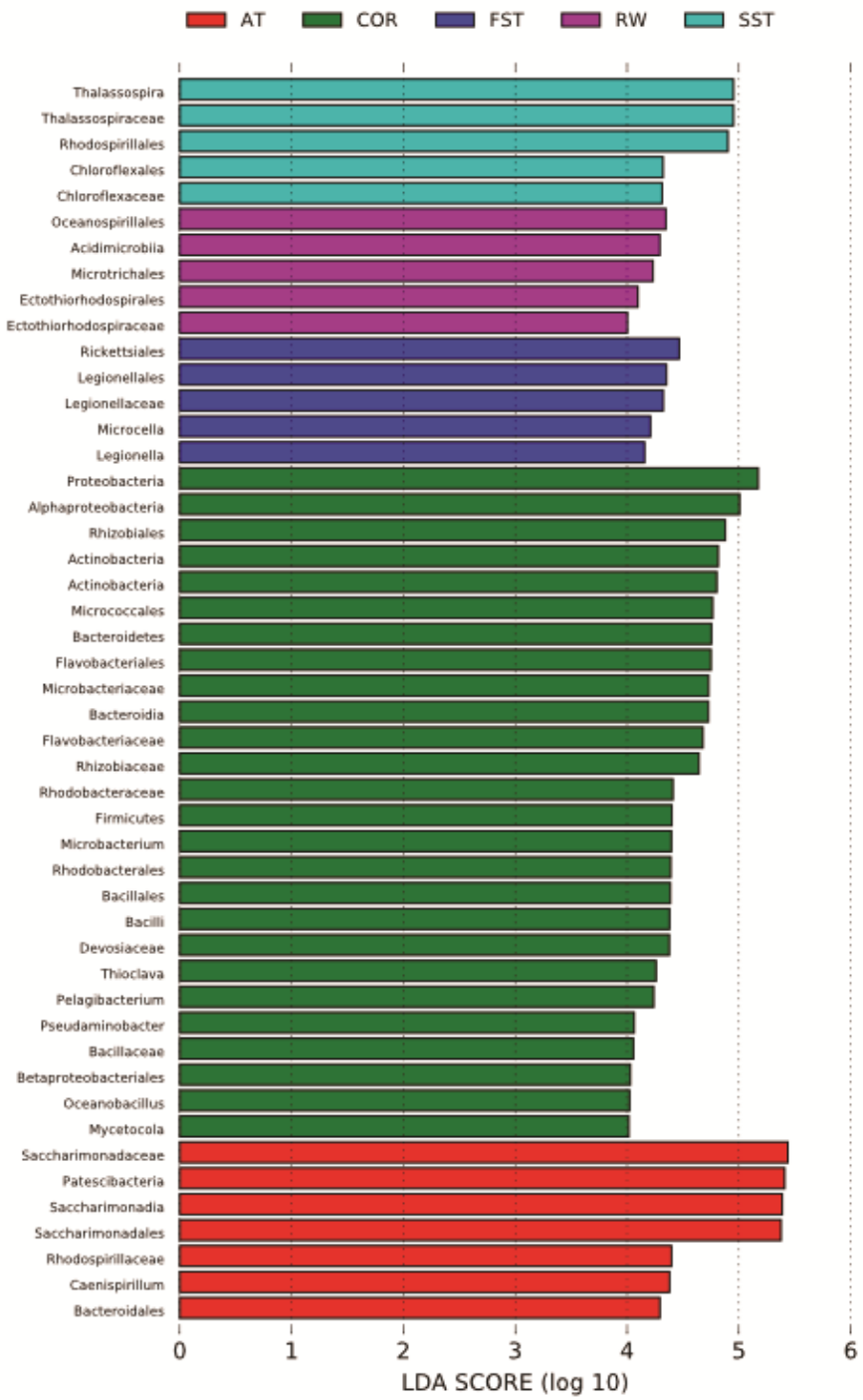


Figure S1. LEfSe results of wastewater treatment process.
Linear discriminant analysis effect size (LEfSe) analysis of the wastewater treatment process (LDA score>4, p<0.05). The indicated genera having statistically significant differences between distinct phases of the full process of wastewater treatment.

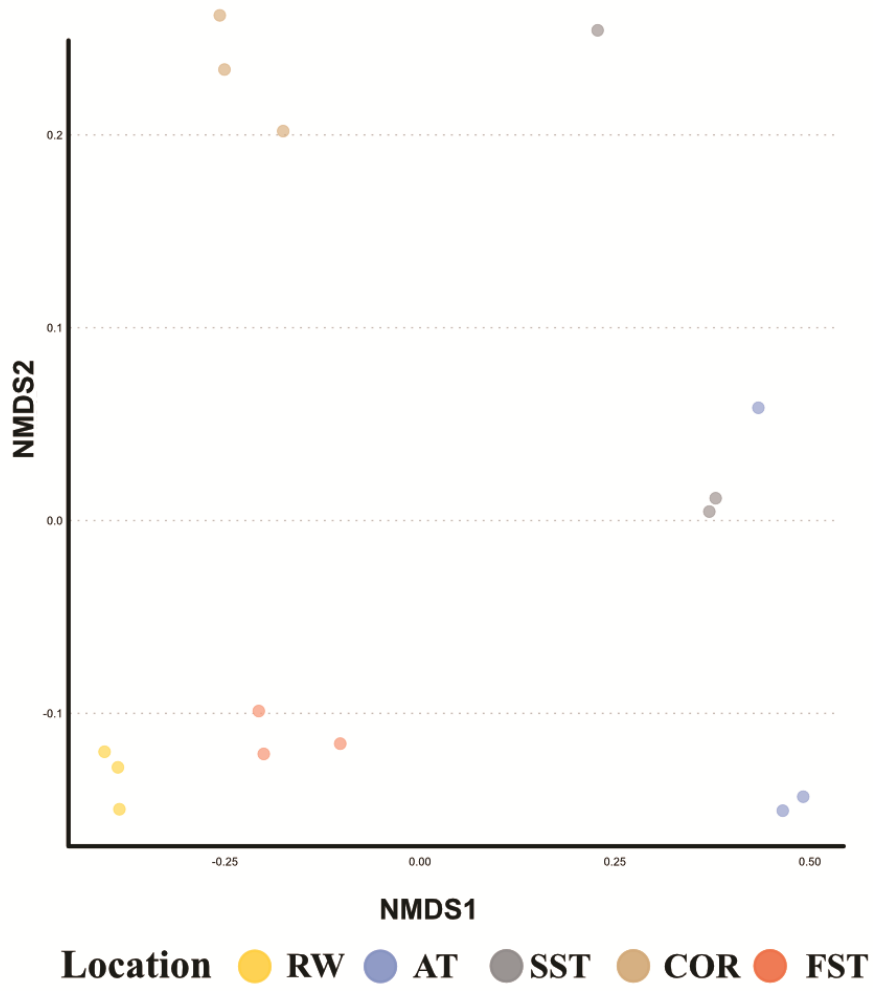


Figure S2. Beta diversity of wastewater treatment process.
Beta diversity of microbial community during the wastewater treatment process. The separation of wastewater treatment phases shows the difference in the composition of the microbial community.

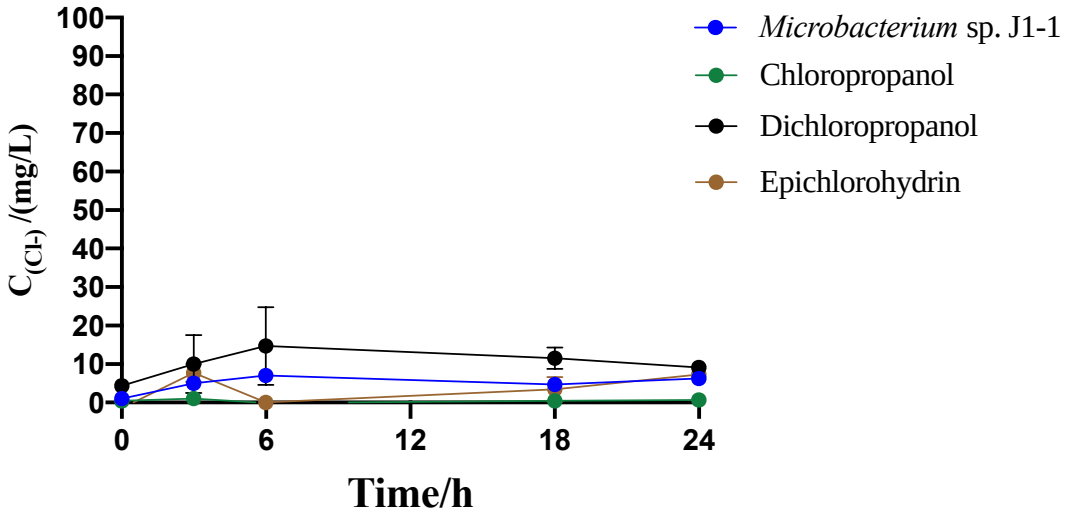


Figure S3. The halogen ion concentration of control group in 24 hours.
Blue: The halogen ion concentration of *Microbacterium* sp. J1-1 without organohalide substrates in the reaction system. Green, Black, Brown: The halogen ion concentration of organohalides without *Microbacterium* sp. J1-1 in the reaction system.

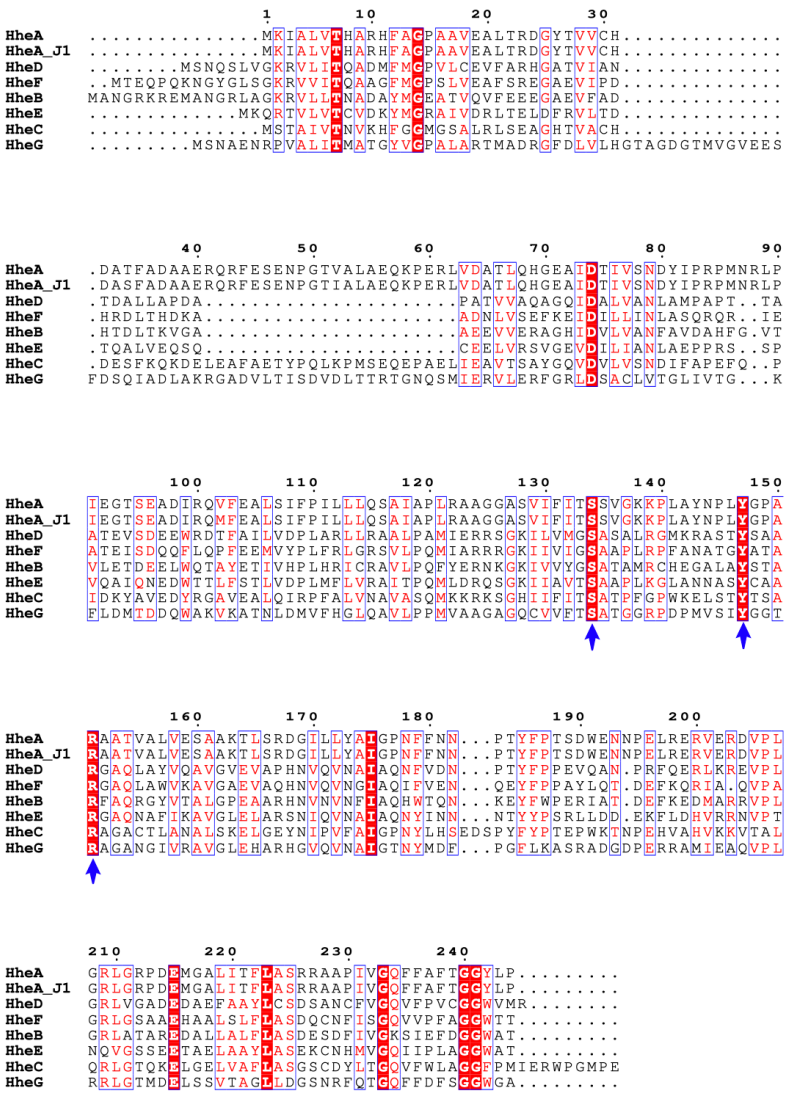


Figure S4 Amino acid sequence alignment of halohydrin dehalogenases. HheA: The halohydrin dehalogenase from *Corynebacterium* sp. Strain N-1074 (GenBank Accession BAA14361.1). HheB: The halohydrin dehalogenase from *Mycobacterium* sp. GP1 (GenBank Accession AAK73175.1). HheC: The halohydrin dehalogenase from *Agrobacterium tumefaciens* (GenBank Accession AAK92099.1). HheA_J1: The putative halohydrin dehalogenase from *Microbacterium* sp. J1-1. HheD: Halohydrin dehalogenase (GenBank Accession AMQ13565.1). HheE: Halohydrin dehalogenase (GenBank Accession AMQ13570.1). HheF: Halohydrin dehalogenase (GenBank Accession AMQ13575.1). HheG: Halohydrin dehalogenase (GenBank Accession AMQ13576.1). Blue arrow: the conserve catalytic triad (Ser-Tyr-Arg) of halohydrin dehalogenases.

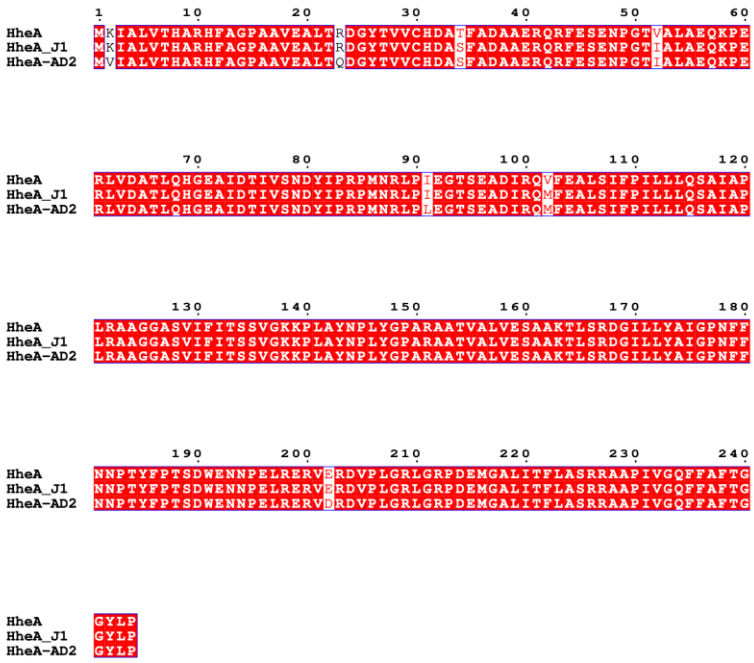


Figure S5 Amino acid sequence alignment of halohydrin dehalogenases of HheA halohydrin dehalogenases.

HheA: The halohydrin dehalogenase from *Corynebacterium* sp. Strain N-1074 (GenBank Accession BAA14361.1). HheA-AD2: The halohydrin dehalogenase from *Arthrobacter* sp. AD2 (GenBank Accession AAK92100.1). HheA_J1: The putative halohydrin dehalogenase from *Microbacterium* sp. J1-1.

Table S1 Organohalide sewage information

The physical and chemical data of organohalide sewage at different locations.

Location	Temperature (℃)	pH	COD (mg/L)	Cl^-(mg/L)
Raw water	50	>11.7	1,100–1,200	24,000–26,000
Aeration tank	35–38	6.5–6.9	280–330	23,000–24,000
Secondary sedimentation tank	33–36	6.5–6.9	230–250	23,000–24,000
Contact oxidation reactor	28–30	6.5–6.9	60–100	23,000–24,000
Final sedimentation tank	25–30	6.5–6.9	30–50	23,000–24,000

COD (Chemical oxygen demand): The amount of oxygen required for the chemical oxidation of organic compounds to inorganic products.

Table S2. Alpha diversity of organohalide sewage

The alpha diversity data of organohalide sewage at different locations.

	Simpson	Chao1	ACE	Shannon
	0.986774	1353.00	1353.00	8.36
Raw water	0.985574	1371.00	1371.00	8.35
	0.988694	1468.00	1468.00	8.49
	0.537006	583.29	589.89	2.97
Aeration tank	0.582560	496.33	525.56	3.17
	0.689960	451.00	451.00	3.61
	0.701489	818.62	821.35	3.77
Secondary sedimentation tank	0.735157	713.46	719.90	3.82
	0.942809	923.96	947.80	6.22
	0.982210	1152.00	1152.00	7.52
Contact oxidation reactor	0.974757	1447.00	1447.00	7.84
	0.982094	1224.00	1224.00	7.71
	0.973534	1529.66	1625.21	7.55
Final sedimentation tank	0.971725	1262.00	1262.00	7.54
	0.953481	1404.33	1374.26	6.70

Table S3. Quality evaluation of MAGs
The detailed information for completeness, contamination of MAG quality via MetaWRAP.

Location	Bin	Completeness	Contamination
Aeration tank	bin.10	95.93	0.949
Aeration tank	bin.11	86.17	2.482
Aeration tank	bin.12	99.18	1.434
Aeration tank	bin.13	91.39	1.728
Aeration tank	bin.14	98.65	0.223
Aeration tank	bin.15	93.25	3.184
Aeration tank	bin.16	99.41	0.904
Aeration tank	bin.17	75.86	0
Aeration tank	bin.18	78.31	1.41
Aeration tank	bin.19	92.18	4.545
Aeration tank	bin.1	74.14	0
Aeration tank	bin.20	73.93	0.561
Aeration tank	bin.21	77.43	1.442
Aeration tank	bin.22	99.5	0
Aeration tank	bin.23	98.75	0.069
Aeration tank	bin.24	74.82	0
Aeration tank	bin.25	99.05	1.342
Aeration tank	bin.26	93.93	0.202
Aeration tank	bin.27	86.24	4.267
Aeration tank	bin.28	73.48	3.636
Aeration tank	bin.29	93.76	1.006
Aeration tank	bin.2	76.3	2.808
Aeration tank	bin.30	99	2.093
Aeration tank	bin.3	77.41	0.172
Aeration tank	bin.4	72.36	1.123
Aeration tank	bin.5	94.15	0.802
Aeration tank	bin.6	88.36	2.075
Aeration tank	bin.7	98.3	4.538
Aeration tank	bin.8	93.75	2.875
Aeration tank	bin.9	93.83	2.525
Contact oxidation reactor	bin.10	70.29	1.98
Contact oxidation reactor	bin.11	96.16	2.003
Contact oxidation reactor	bin.12	95.49	0.315
Contact oxidation reactor	bin.13	94.36	1.692
Contact oxidation reactor	bin.14	95.59	3.048
Contact oxidation reactor	bin.15	98.87	0.277
Contact oxidation reactor	bin.16	98.77	0.081

Contact oxidation reactor	bin.17	77.95	1.638
Contact oxidation reactor	bin.18	81.96	4.186
Contact oxidation reactor	bin.19	99.05	0.599
Contact oxidation reactor	bin.1	74.14	0
Contact oxidation reactor	bin.20	90.52	3.478
Contact oxidation reactor	bin.21	94.2	3.62
Contact oxidation reactor	bin.22	96.99	1.227
Contact oxidation reactor	bin.23	74.88	2.247
Contact oxidation reactor	bin.24	95.94	0.19
Contact oxidation reactor	bin.25	89.98	1.287
Contact oxidation reactor	bin.26	84.45	1.724
Contact oxidation reactor	bin.27	100	0.54
Contact oxidation reactor	bin.28	99.28	0
Contact oxidation reactor	bin.29	98.24	0.66
Contact oxidation reactor	bin.2	97.81	0.757
Contact oxidation reactor	bin.30	72.93	1.123
Contact oxidation reactor	bin.31	96.62	0
Contact oxidation reactor	bin.32	80.78	2.777
Contact oxidation reactor	bin.33	92.69	1.298
Contact oxidation reactor	bin.34	96.32	1.308
Contact oxidation reactor	bin.35	96.44	1.421
Contact oxidation reactor	bin.36	95.29	1.219
Contact oxidation reactor	bin.37	74.68	3.097
Contact oxidation reactor	bin.38	93.73	1.698
Contact oxidation reactor	bin.39	94.76	0.048
Contact oxidation reactor	bin.3	96.87	3.415
Contact oxidation reactor	bin.40	93.85	0.433
Contact oxidation reactor	bin.41	75.86	0
Contact oxidation reactor	bin.42	99.91	0.144
Contact oxidation reactor	bin.43	95.65	1.536
Contact oxidation reactor	bin.44	86.09	1.047
Contact oxidation reactor	bin.45	93.29	0.42
Contact oxidation reactor	bin.46	93.01	1.208
Contact oxidation reactor	bin.47	70.8	0.102
Contact oxidation reactor	bin.48	90.13	2.452
Contact oxidation reactor	bin.49	92.41	0.505
Contact oxidation reactor	bin.4	72.28	1.937
Contact oxidation reactor	bin.50	97.53	1.498
Contact oxidation reactor	bin.51	81.46	1.477
Contact oxidation reactor	bin.52	99.64	2.456
Contact oxidation reactor	bin.53	71	1.724
Contact oxidation reactor	bin.54	72.41	1.724
Contact oxidation reactor	bin.55	97.01	3.755

Contact oxidation reactor	bin.56	97.27	0.414
Contact oxidation reactor	bin.57	75.33	4.286
Contact oxidation reactor	bin.58	93.74	1.403
Contact oxidation reactor	bin.59	99.41	1.462
Contact oxidation reactor	bin.5	88.5	4.57
Contact oxidation reactor	bin.60	93.66	0.706
Contact oxidation reactor	bin.61	95.79	0
Contact oxidation reactor	bin.62	77.01	2.945
Contact oxidation reactor	bin.63	94.48	0.755
Contact oxidation reactor	bin.64	97.34	0.497
Contact oxidation reactor	bin.6	76.89	0
Contact oxidation reactor	bin.7	98.45	1.111
Contact oxidation reactor	bin.8	90.61	0
Contact oxidation reactor	bin.9	96.78	0.301

Table S4. Taxonomic assignment results of AT MAGs

The taxonomic assignment results of AT MAGs via GTDB-tk

Genome	Phylum	Class	Order	Family	Genus	Species
bin.5	*Actinobacteriota*	*Actinobacteria*	*Propionibacteriales*	*Propionibacteriaceae*	*Tessaracoccus*	*lapidicaptus*
bin.3	*Thermotogota*	*Thermotogae*	*Petrotogales*	*Petrotogaceae*	*Petrotoga*	*mobilis*
bin.13	*Bacteroidota*	*Bacteroidia*	*Flavobacteriales*	koll-22	UBA5081	sp002415785
bin.14	*Firmicutes_A*	*Clostridia*	*Oscillospirales*	*Ruminococcaceae*	*Pygmaiobacter*	NA
bin.8	*Actinobacteriota*	*Actinobacteria*	*Actinomycetales*	*Microbacteriaceae*	*Microbacterium*	NA
bin.23	*Actinobacteriota*	*Actinobacteria*	*Actinomycetales*	*Actinomycetaceae*	KH17	NA
bin.20	*Patescibacteria*	*Gracilibacteria*	BD1-5	UBA6164	UBA7396	NA
bin.4	*Patescibacteria*	*Microgenomatia*	UBA1406	GWC2-37-13	NA	NA
bin.19	*Chloroflexota*	*Anaerolineae*	SBR1031	UBA2029	NA	NA
bin.28	*Chloroflexota*	*Anaerolineae*	*Promineofilales*	*Promineofilaceae*	NA	NA
bin.6	*Chloroflexota*	*Chloroflexia*	*Chloroflexales*	*Chloroflexaceae*	NA	NA
bin.24	*Spirochaetota*	*Spirochaetia*	*Sphaerochaetales*	*Sphaerochaetaceae*	*Sphaerochaeta*	NA
bin.22	*Bacteroidota*	*Bacteroidia*	*Chitinophagales*	ND	NA	NA
bin.7	*Bacteroidota*	*Rhodothermia*	*Rhodothermales*	MEBICO9517	NA	NA
bin.27	*Proteobacteria*	*Gammaproteobacteria*	*Xanthomonadales*	*Wenzhouxiangellaceae*	GCA-2722315	NA
bin.17	*Proteobacteria*	*Gammaproteobacteria*	NA	NA	NA	NA
bin.25	*Proteobacteria*	*Alphaproteobacteria*	*Rickettsiales*	NA	NA	NA
bin.29	*Firmicutes_A*	*Clostridia*	*Oscillospirales*	*Oscillospiraceae*	*Intestinimonas*	NA
bin.21	*Firmicutes_A*	*Clostridia*	*Clostridiales*	*Clostridiaceae*	NA	NA
bin.2	*Patescibacteria*	*Saccharimonadia*	*Saccharimonadales*	SZUA-47	NA	NA

bin.16	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	*Rhizobiaceae*	*Martelella*	NA
bin.26	*Proteobacteria*	*Alphaproteobacteria*	*Rhodobacterales*	*Rhodobacteraceae*	NA	NA
bin.9	*Proteobacteria*	*Alphaproteobacteria*	*Rhodobacterales*	*Rhodobacteraceae*	UBA996	NA
bin.30	*Proteobacteria*	*Alphaproteobacteria*	*Rhodospirillales*	*Thalassospiraceae*	*Thalassospira*	NA
bin.15	*Proteobacteria*	*Alphaproteobacteria*	*Rhodospirillales*	*Thalassospiraceae*	*Thalassospira*	NA
bin.18	*Campylobacterota*	*Campylobacteria*	*Campylobacterales*	NA	NA	NA
bin.12	*Campylobacterota*	*Campylobacteria*	*Campylobacterales*	*Sulfurimonadaceae*	*Sulfurimonas*	NA

Unclassified Family/Genus/Species are indicated as 'NA', not available.

Table S5. Taxonomic assignment results of COR MAGs

The taxonomic assignment results of COR MAGs via GTDB-tk.

Genome	Phylum	Class	Order	Family	Genus	Species
bin.46	*Bacteroidota*	*Bacteroidia*	*Flavobacteriales*	koll-22	UBA5081	sp002415785
bin.34	*Proteobacteria*	*Gammaproteobacteria*	*Francisellales*	*Francisellaceae*	*Allofrancisella*	*guangzhouensis*
bin.23	*Patescibacteria*	*Paceibacteria*	UBA9983_A	UBA1006	UBA1006	NA
bin.26	*Patescibacteria*	*Paceibacteria*	UBA9983_A	UBA1006	UBA1006	NA
bin.47	*Patescibacteria*	*Paceibacteria*	UBA6257	UBA9933	WO2-47-17b	NA
bin.10	*Patescibacteria*	ABY1	BM507	XYC2-FULL-47-12	NA	NA
bin.30	*Patescibacteria*	*Microgenomatia*	UBA1406	GWC2-37-13	NA	NA
bin.6	*Chloroflexota*	*Dehalococcoidia*	*Dehalococcoidales*	*Dehalococcoidaceae*	*Dehalogenimonas*	NA
bin.31	*Verrucomicrobiota*	*Verrucomicrobiae*	*Opitutales*	UBA3534	NA	NA
bin.44	*VerrucomicrobiotaA*	*Chlamydiia*	*Parachlamydiales*	*Simkaniaceae*	*Simkania*	NA
bin.63	*VerrucomicrobiotaA*	*Chlamydiia*	*Parachlamydiales*	*Simkaniaceae*	NA	NA
bin.37	*Bacteroidota*	*Bacteroidia*	*Bacteroidales*	*Marinilabiliaceae*	*Marinilabilia*	NA
bin.28	*Bacteroidota*	*Bacteroidia*	*Flavobacteriales*	*Cryomorphaceae*	NA	NA
bin.7	*Bacteroidota*	*Bacteroidia*	*Sphingobacteriales*	*Sphingobacteriaceae*	NA	NA
bin.51	*Bacteroidota*	*Bacteroidia*	*Chitinophagales*	*Chitinophagaceae*	*Taibaiella_B*	NA
bin.49	*Proteobacteria*	*Gammaproteobacteria*	*Enterobacterales*	*Alteromonadaceae*	*Idiomarina*	NA
bin.3	*Proteobacteria*	*Gammaproteobacteria*	*Enterobacterales*	*Alteromonadaceae*	*Idiomarina*	NA
bin.33	*Proteobacteria*	*Gammaproteobacteria*	UBA5158	UBA5158	2-12-FULL-43-28	NA
bin.62	*Proteobacteria*	*Gammaproteobacteria*	NA	NA	NA	NA
bin.55	*Proteobacteria*	*Gammaproteobacteria*	*Nitrococcales*	NA	NA	NA

bin.11	*Proteobacteria*	*Gammaproteobacteria*	*Nitrococcales*	NA	NA	NA
bin.52	*Proteobacteria*	*Gammaproteobacteria*	*Nitrococcales*	NA	NA	NA
bin.14	*Proteobacteria*	*Gammaproteobacteria*	*Nitrococcales*	NA	NA	NA
bin.22	*Proteobacteria*	*Gammaproteobacteria*	*Burkholderiales*	*Burkholderiaceae*	*Pusillimonas*	NA
bin.5	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	*Rhizobiaceae*	*Pararhizobium_A*	NA
bin.48	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	*Amorphaceae*	*Amorphus*	NA
bin.57	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	*Methylopilaceae*	Ga0077545	NA
bin.43	*Proteobacteria*	*Alphaproteobacteria*	*Parvibaculales*	*Parvibaculaceae*	*Parvibaculum*	NA
bin.18	*Proteobacteria*	*Alphaproteobacteria*	*Sphingomonadales*	*Sphingomonadaceae*	*Altererythrobacter_A*	NA
bin.40	*Proteobacteria*	*Alphaproteobacteria*	*Sphingomonadales*	*Sphingomonadaceae*	*Sphingopyxis*	NA
bin.54	*Proteobacteria*	*Alphaproteobacteria*	Bin95	Bin95	NA	NA
bin.20	*Proteobacteria*	*Alphaproteobacteria*	*Micavibrionales*	UBA2020	NA	NA
bin.42	*Proteobacteria*	*Alphaproteobacteria*	UBA8366	GCA-2696645	NA	NA
bin.41	*Proteobacteria*	*Alphaproteobacteria*	UBA8366	GCA-2696645	NA	NA
bin.8	*Proteobacteria*	*Alphaproteobacteria*	NA	NA	NA	NA
bin.15	*Proteobacteria*	*Gammaproteobacteria*	*Pseudomonadales*	*Oleiphilaceae*	*Marinobacter*	*subterrani*
bin.59	*Proteobacteria*	*Gammaproteobacteria*	*Francisellales*	*Francisellaceae*	*Francisella*	*hispaniensis*
bin.36	*Proteobacteria*	*Gammaproteobacteria*	*Thiomicrospirales*	*Thiomicrospiraceae*	*Hydrogenovibrio*	sp000711305
bin.56	*Proteobacteria*	*Alphaproteobacteria*	*Rhodobacterales*	*Rhodobacteraceae*	*Thioclava*	*marina*
bin.17	*FirmicutesI*	*Bacilli_A*	*Paenibacillales*	*Paenibacillaceae*	NA	NA
bin.39	*FirmicutesI*	*Bacilli_A*	*Paenibacillales*	NA	NA	NA
bin.45	*Actinobacteriota*	*Actinobacteria*	*Actinomycetales*	*Microbacteriaceae*	*Yonghaparkia*	NA
bin.2	*Actinobacteriota*	*Actinobacteria*	*Actinomycetales*	*Microbacteriaceae*	73-13	NA
bin.25	*Actinobacteriota*	*Actinobacteria*	*Actinomycetales*	*Microbacteriaceae*	*Microbacterium*	NA
bin.13	*Actinobacteriota*	*Actinobacteria*	*Propionibacteriales*	*Nocardioidaceae*	*Aeromicrobium*	NA

bin.24	*Bacteroidota*	*Bacteroidia*	*Flavobacteriales*	*Flavobacteriaceae*	NA	NA
bin.29	*Bacteroidota*	*Bacteroidia*	*Flavobacteriales*	*Flavobacteriaceae*	*Galbibacter*	NA
bin.60	*Bacteroidota*	*Bacteroidia*	*Flavobacteriales*	*Flavobacteriaceae*	*Flavobacterium*	NA
bin.32	*Bacteroidota*	*Bacteroidia*	*Flavobacteriales*	*Flavobacteriaceae*	*Flavobacterium_A*	NA
bin.27	*Bacteroidota*	*Bacteroidia*	*Flavobacteriales*	*Crocinitomicaceae*	UBA5422	NA
bin.61	*Bacteroidota*	*Kapabacteria*	*Kapabacteriales*	GCA-002839825	PGYR01	NA
bin.50	*Proteobacteria*	*Gammaproteobacteria*	*Enterobacterales*	*Alteromonadaceae*	*Idiomarina*	NA
bin.4	*Proteobacteria*	*Gammaproteobacteria*	*Legionellales*	*Legionellaceae*	*Legionella_D*	NA
bin.58	*Proteobacteria*	*Gammaproteobacteria*	*Legionellales*	*Legionellaceae*	*Legionella_A*	NA
bin.53	*Proteobacteria*	*Gammaproteobacteria*	*Thiomicrospirales*	*Thiomicrospiraceae*	*Hydrogenovibrio*	NA
bin.12	*Proteobacteria*	*Gammaproteobacteria*	*Burkholderiales*	*Burkholderiaceae*	*Candidimonas*	NA
bin.38	*Proteobacteria*	*Gammaproteobacteria*	*Burkholderiales*	*Burkholderiaceae*	*Candidimonas*	NA
bin.35	*Proteobacteria*	*Gammaproteobacteria*	*Burkholderiales*	*Burkholderiaceae*	NA	NA
bin.19	*Proteobacteria*	*Gammaproteobacteria*	*Burkholderiales*	*Burkholderiaceae*	NA	NA
bin.16	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	*Rhizobiaceae*	*Aquamicrobium_A*	NA
bin.21	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	*Rhizobiaceae*	*Aliihoeflea*	NA
bin.9	*Proteobacteria*	*Alphaproteobacteria*	*Rhodobacterales*	*Rhodobacteraceae*	NA	NA
bin.64	*Proteobacteria*	*Alphaproteobacteria*	*Rhodospirillales_A*	*Thalassospiraceae*	*Thalassospira*	NA

Unclassified Order/Family/Genus/Species are indicated as 'NA', not available.

Table S6. The distribution of dehalogenase in MAGs
Gene numbers of four types of dehalogenases in different MAGs.

Genome	**2-haloacid dehalogenase**	**haloacetate dehalogenase**	**haloalkane dehalogenase**	**4-chlorobenzoyl coenzyme A dehalogenase**
AT_Chloroflexaceae_bin.6	1	3	2	0
AT_Rhodothermales_bin.7	0	1	4	0
AT_Microbacterium_bin.8	0	1	0	1
AT_Rhodobacteraceae_bin.9	1	1	0	2
AT_Thalassospira_bin.15	1	1	0	0
AT_Martelella_bin.16	0	2	0	0
AT_Anaerolineae_bin.19	0	1	1	0
AT_Rhodobacteraceae_bin.26	1	1	0	0
AT_Wenzhouxiangellaceae_bin.27	0	0	1	0
AT_Promineofilaceae_bin.28	1	1	2	1
AT_Thalassospira_bin.30	1	0	0	0
COR_Idiomarina_bin.3	0	0	1	1
COR_Pararhizobium_bin.5	2	1	0	0
COR_Rhodobacteraceae_bin.9	1	1	0	0
COR_Nitrococcales_bin.11	1	2	0	0
COR_Candidimonas_bin.12	2	0	0	0
COR_Nitrococcales_bin.14	1	0	2	1
COR_Marinobacter_bin.15	1	0	2	0
COR_Aquamicrobium_bin.16	1	1	2	0
COR_Altererythrobacter_bin.18	0	1	0	0
COR_Burkholderiaceae_bin.19	2	2	0	0
COR_Rhizobiaceae_bin.21	1	0	0	1
COR_Pusillimonas_bin.22	1	0	1	0
COR_Microbacterium_bin.25	0	1	0	1
COR_Opitutales_bin.31	0	0	1	0
COR_UBA5158_bin.33	0	1	1	0
COR_Marinilabilia_bin.37	2	0	0	0
COR_Candidimonas_bin.38	3	0	0	0
COR_Sphingopyxis_bin.40	0	1	0	0
COR_UBA8366_bin.41	1	1	0	0
COR_UBA8366_bin.42	1	1	0	0
COR_Parvibaculum_bin.43	0	0	1	0
COR_Simkania_bin.44	0	0	1	0
COR_Yonghaparkia_bin.45	0	0	1	0
COR_Amorphus_bin.48	1	2	0	1
COR_Idiomarina_bin.49	0	1	1	0

COR_Idiomarina_bin.50	0	0	1	0
COR_Nitrococcales_bin.52	0	1	0	1
COR_Bin95_bin.54	3	1	4	0
COR_Nitrococcales_bin.55	0	1	1	0
COR_Thioclava_bin.56	2	1	0	0
COR_Thalassospiraceae_bin.64	1	0	0	0

Table S7. The significantly different metabolism pathways between locations
The significantly different metabolism pathways between locations based on the non-targeted metabolomics data.

	Pathway Name	**Pathway ID**	**Pvalue**	**Compounds**
RW/AT	Tyrosine metabolism	edi00350	0.0238	C03964|(R)-3-(4-Hydroxyphenyl) lactate; C00042|Succinate
RW/AT	Oxidative phosphorylation	edi00190	0.049	C00042|Succinate
AT/COR	Tyrosine metabolism	edi00350	0.00498	C03964|(R)-3-(4-Hydroxyphenyl) lactate; C00042|Succinate
AT/COR	Oxidative phosphorylation	edi00190	0.0229	C00042|Succinate
AT/COR	Pantothenate and CoA biosynthesis	edi00770	0.0398	C01053|(R)-4-Dehydropantoate
AT/COR	Alanine, aspartate, and glutamate metabolism	edi00250	0.0398	C00042|Succinate
AT/COR	Pyruvate metabolism	edi00620	0.044	C00042|Succinate
AT/COR	Sulfur metabolism	edi00920	0.0468	C00042|Succinate

Table S8. The antibiotic resistance ontology of the MAGs.
The potential ARO genes of the MAGs were predicted via RGI and CARD.

MAG	Taxonomy_MAG	Name	ARO	Resistance Mechanism
AT_bin.12	*Campylobacterales*	*qac*J	3007014	antibiotic efflux
AT_bin.14	*Oscillospirale*	*nim*I	3007111	antibiotic inactivation
AT_bin.15	*Rhodospirillales*	*qac*J	3007014	antibiotic efflux
AT_bin.16	*Rhizobiales*	*ade*F	3000777	antibiotic efflux
AT_bin.16	*Rhizobiales*	*qac*J	3007014	antibiotic efflux
AT_bin.21	*Clostridiales*	*van*R	3003728	antibiotic target alteration
AT_bin.21	*Clostridiales*	*gyr*B	3004562	antibiotic target alteration
AT_bin.26	*Rhodobacterales*	*qac*G	3007015	antibiotic efflux
AT_bin.30	*Rhodospirillales*	*qac*J	3007014	antibiotic efflux
AT_bin.30	*Rhodospirillales*	*ade*F	3000777	antibiotic efflux
AT_bin.30	*Rhodospirillales*	*ade*F	3000777	antibiotic efflux
COR_bin.5	*Rhizobiales*	*qac*G	3007015	antibiotic efflux
COR_bin.9	*Rhodobacterales*	*qac*G	3007015	antibiotic efflux
COR_bin.12	*Burkholderiales*	*qac*G	3007015	antibiotic efflux
COR_bin.12	*Burkholderiales*	*ade*F	3000777	antibiotic efflux
COR_bin.12	*Burkholderiales*	*qac*G	3007015	antibiotic efflux
COR_bin.14	*Nitrococcales*	*ade*F	3000777	antibiotic efflux
COR_bin.15	*Pseudomonadales*	*rsm*A	3005069	antibiotic efflux
COR_bin.15	*Pseudomonadales*	*qac*G	3007015	antibiotic efflux
COR_bin.16	*Rhizobiales*	*qac*G	3007015	antibiotic efflux
COR_bin.19	*Burkholderiales*	*ade*F	3000777	antibiotic efflux
COR_bin.19	*Burkholderiales*	*qac*G	3007015	antibiotic efflux
COR_bin.21	*Rhizobiales*	*qac*G	3007015	antibiotic efflux
COR_bin.22	*Burkholderiales*	*qac*G	3007015	antibiotic efflux
COR_bin.27	*Flavobacteriales*	*ade*F	3000777	antibiotic efflux
COR_bin.29	*Flavobacteriales*	*ade*F	3000777	antibiotic efflux
COR_bin.31	*Opitutales*	*ade*F	3000777	antibiotic efflux
COR_bin.35	*Burkholderiales*	*ade*F	3000777	antibiotic efflux
COR_bin.35	*Burkholderiales*	*ade*F	3000777	antibiotic efflux
COR_bin.35	*Burkholderiales*	*qac*G	3007015	antibiotic efflux
COR_bin.36	*Thiomicrospirales*	*ade*F	3000777	antibiotic efflux
COR_bin.36	*Thiomicrospirales*	*ade*F	3000777	antibiotic efflux
COR_bin.38	*Burkholderiales*	*qac*J	3007014	antibiotic efflux
COR_bin.38	*Burkholderiales*	*qac*G	3007015	antibiotic efflux
COR_bin.40	*Sphingomonadales*	*ade*F	3000777	antibiotic efflux
COR_bin.40	*Sphingomonadales*	*ade*F	3000777	antibiotic efflux
COR_bin.43	*Parvibaculales*	*ade*F	3000777	antibiotic efflux
COR_bin.48	*Rhizobiales*	*qac*G	3007015	antibiotic efflux
COR_bin.49	*Enterobacterales*	*rsm*A	3005069	antibiotic efflux

COR_bin.50	*Enterobacterales*	*rsm*A	3005069	antibiotic efflux
COR_bin.52	*Nitrococcales*	*ade*F	3000777	antibiotic efflux
COR_bin.53	*Thiomicrospirales*	*ade*F	3000777	antibiotic efflux
COR_bin.55	*Nitrococcales*	*ade*F	3000777	antibiotic efflux
COR_bin.56	*Rhodobacterales*	*qac*G	3007015	antibiotic efflux
COR_bin.57	*Rhizobiales*	*ade*F	3000777	antibiotic efflux
COR_bin.59	*Francisellales*	*qac*J	3007014	antibiotic efflux
COR_bin.60	*Flavobacteriales*	*ade*F	3000777	antibiotic efflux
COR_bin.60	*Flavobacteriales*	*ade*F	3000777	antibiotic efflux
COR_bin.64	*Rhodospirillales*	*ade*F	3000777	antibiotic efflux
COR_bin.64	*Rhodospirillales*	*qac*J	3007014	antibiotic efflux
COR_bin.64	*Rhodospirillales*	*ade*F	3000777	antibiotic efflux