Substrate promiscuity of xenobiotic-transforming hydrolases from stream biofilms impacted by treated wastewater

Yaochun Yu[1]*, Niklas Ferenc Trottmann[2]*, Milo R. Schärer[2], Kathrin Fenner[1,3], Serina L. Robinson✉[2]

[1] Department of Environmental Chemistry, Swiss Federal Institute of Aquatic Science and Technology (Eawag), Dübendorf 8600, Switzerland

[2] Department of Environmental Microbiology, Swiss Federal Institute of Aquatic Science and Technology (Eawag), Dübendorf 8600, Switzerland

[3] Department of Chemistry, University of Zürich, 8057 Zürich, Switzerland

*co-first authors contributed equally to this work

✉corresponding author: serina.robinson@eawag.ch

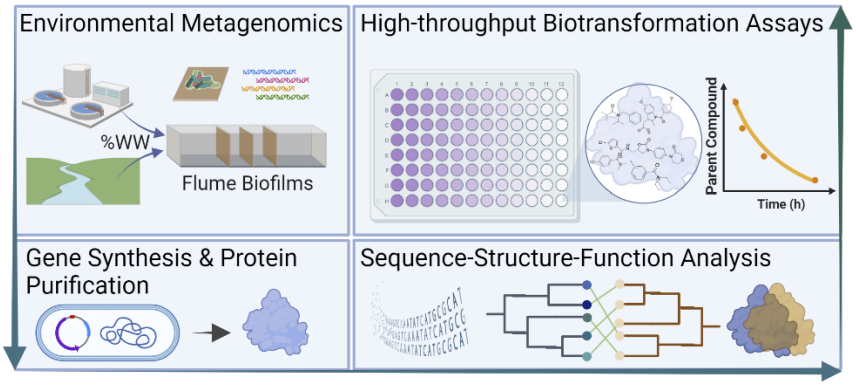
Keywords

treated wastewater, organic contaminants, biotransformations, aquatic biofilms, hydrolases, metagenomic sequencing

Abstract

Organic contaminants enter aquatic ecosystems from various sources, including wastewater treatment plant effluent. Freshwater biofilms play a major role in the removal of organic contaminants from receiving water bodies, but knowledge of the molecular mechanisms driving contaminant biotransformations in complex stream biofilm (periphyton) communities remains limited. Previously, we demonstrated that biofilms in experimental flume systems grown at higher ratios of treated wastewater (WW) to stream water displayed an increased biotransformation potential for a number of organic contaminants. We identified a positive correlation between WW percentage and biofilm biotransformation rates for the widely-used insect repellent, *N,N*-diethyl-meta-toluamide (DEET). Here, we conducted deep shotgun sequencing of flume biofilms and identified a positive correlation between WW percentage and metagenomic read abundances of DEET hydrolase (DH) homologs. To test the causality of this association, we constructed a targeted metagenomic library of DH homologs from flume biofilms. We screened our complete metagenomic library for activity with four different substrates and a subset thereof with 183 WW-related organic compounds. The majority of active hydrolases in our library preferred aliphatic and aromatic ester substrates while, remarkably, only a single reference enzyme was capable of DEET hydrolysis. Of the 626 total enzyme-substrate combinations tested, approximately 5% were active enzyme-substrate pairs. Metagenomic DH family homologs revealed a broad substrate promiscuity spanning 22 different compounds when summed across all enzymes tested. We biochemically characterized the most promiscuous and active enzymes identified based on metagenomic analysis from uncultivated *Rhodospirillaceae* and *Planctomycetaceae*. In addition to characterizing new DH family enzymes, we exemplified a framework for linking metagenome-guided hypothesis generation with experimental validation. Overall, this study expands the scope of known enzymatic contaminant biotransformations for metagenomic hydrolases from WW-receiving stream biofilm communities.

46 **Graphical abstract**



53 **Abbreviations**

54 Trace organic contaminants: TrOCs

55 *N,N*-diethyl-meta-toluamide: DEET

56 Treated wastewater: WW

57 DEET hydrolase: DH

58 Trichloroacetic acid: TCA

59 Dimethylsulfoxide: DMSO

60 Ultra-High-Performance Liquid Chromatography Coupled to High-Resolution Tandem Mass

61 Spectrometry: UHPLC-HRMS/MS

62 Fast protein liquid chromatography: FPLC

63 **Highlights**

64 - Metagenomic DEET hydrolase abundances higher in biofilms grown in treated

65 wastewater.

66 - Eleven out of 64 metagenomic hydrolases tested exhibited hydrolysis activity.

67 - Related enzymes in a single family of DEET hydrolases biotransform 20+ contaminants.

68 - Reference DEET hydrolase shows a substrate preference for benzamide moieties.

69 - ‘True’ DEET hydrolases are in low abundance even in biofilms that degrade DEET.

70

1. Introduction

Trace organic contaminants (TrOCs) released due to anthropogenic activities such as industrial production and agricultural practices are exerting increasing pressures on ecosystems and human health (Caliman and Gavrilescu, 2009; Daughton and Ternes, 1999; Evich et al., 2022; Fenner et al., 2013; Halden, 2011; Schwarzenbach et al., 2006). Many TrOCs and their transformation products exert toxicity through bioaccumulation, mixture effects, and specific toxic interactions (Altenburger et al., 2018; Carles et al., 2021; Jin et al., 2023). In aquatic systems, microbes are the dominant players in the biotransformation and removal of various TrOCs through metabolic and co-metabolic processes (Che et al., 2021; Fenner et al., 2013; Han et al., 2023, 2019; Xing et al., 2018; Yu et al., 2018). However, apart from a few specific pathways, the diversity of enzymes involved in microbial biotransformation processes remain largely unknown. Understanding the scope and diversity of microbial enzyme-substrate interactions within environmental microbiomes is critical to improve prediction of biotransformations and promote contaminant removal (Fenner et al., 2021).

Freshwater biofilms are ideal systems for investigating TrOC biotransformation processes due to their complex community composition and critical ecological functions in aquatic environments (Battin et al., 2016; Desiante et al., 2021). Most studies on biofilms have been conducted using amplicon sequencing to profile taxonomic community shifts that correlate with TrOC biotransformations in treated wastewater (Carles et al., 2022, 2021; Desiante et al., 2022; Mansfeldt et al., 2020; Tien et al., 2013). Metagenomics and metatranscriptomics analyses have recently deepened the resolution of community analysis to identify functional shifts in TrOC-exposed engineered wetland biomats (Vega et al., 2023, 2022). However, the number of studies conducted on the molecular mechanisms of TrOC biotransformations in freshwater biofilms or biomats is small relative to other systems such as enriched soil consortia (Vasileiadis et al., 2022) or activated sludge (Rich and Helbling, 2023). Activated sludge studies have used metagenomics and metatranscriptomics to correlate microbial taxa, functional genes, and pollutant

97 biotransformations (Achermann et al., 2020; Athanasakoglou and Fenner, 2022; Kennes-Veiga
98 et al., 2022; Stadler et al., 2018), but also have remained limited in uncovering experimentally-
99 validated enzymes involved in TrOC biotransformations. Many enzymes involved in the
100 biotransformation of contaminants remain a 'black box.'

101 In our previous study, stream biofilms cultivated at different ratios of stream water and
102 either ultrafiltered or non-ultrafiltered treated wastewater (WW) exhibited varying TrOC
103 biotransformation capacities (Desiante et al., 2022). Among 75 TrOCs tested, seven compounds
104 (*N,N*-diethyl-meta-toluamide (DEET), acesulfame K, caffeine, cyclamate, levetiracetam,
105 saccharin, and valsartan, Figure S1) showed clear evidence for increased biotransformation rates
106 in biofilms grown at higher WW percentages, as compared to controls grown in stream water or
107 in ultrafiltered WW. This phenomenon, previously defined as the *downstream effect* (Desiante et
108 al., 2022), strongly suggested that microorganisms present in the WW might affect the
109 biotransformation capacity of biofilms downstream of wastewater treatment plants. Notably, all
110 compounds that exhibited the *downstream effect* contained hydrolyzable moieties, primarily
111 sulfonamide- and amide-functional groups (Figure S1). This motivated us to initiate a large-scale
112 investigation into the potential roles of environmental hydrolases from wastewater effluent-
113 impacted biofilms in degrading TrOCs.

114 Here we employed a metagenome mining framework to investigate contaminant
115 biotransformations and enzyme-substrate specificity in complex stream biofilm communities. We
116 observed that increasing biotransformation rates of DEET (an insect repellant and one of the
117 seven *downstream effect* compounds) corresponded with higher metagenomic abundances of
118 DEET hydrolase (DH) homologs in stream biofilms with increasing proportions of treated
119 wastewater (WW). Because of the similarity in hydrolyzable moieties amongst all *downstream*
120 *effect* compounds, we hypothesized that diverse metagenomic DH family homologs detected in
121 WW-impacted biofilms have an expanded substrate specificity not only for DEET, but also for
122 other WW-relevant contaminants. Guided by protein structural modeling from metagenomic

123 sequences, we designed, constructed, and tested a targeted metagenomic library of 64 DH
124 homologs. This work maps the sequence-structure-substrate relationships of DH family enzymes
125 encoded by ten different bacterial phyla from complex stream biofilm communities. The results
126 not only expand our fundamental understanding of how enzyme diversity relates to
127 biotransformation processes, but also have implications for the identification of novel and
128 functional TrOC-degrading enzymes from environmental microbiomes.

129 **2. Materials and Methods**

130 **2.1. Biofilm sampling**

131 The biofilm samples which in this study were used for metagenomic sequencing and
132 hydrolase screening were previously analyzed by 16S and 18S rRNA gene amplicon sequencing
133 (Carles et al., 2021) and organic contaminant biotransformation experiments (Desiante et al.,
134 2022). Briefly, the experimental flume systems (Burdon et al., 2020) for biofilm growth consisted
135 of two independent experiments described by (Desiante et al., 2022), with 16 (for experiment 1)
136 and 20 water channels (for experiment 2) fed with different proportions of the following sources:
137 stream water mixed in proportions with either 10%, 30%, or 80% WW (for experiment 1) or 30%
138 or 80% ultrafiltered (pore size 0.4 μm) WW in experiment 2. Biofilms were grown on glass slides
139 placed in these flow-through channels for four weeks. Biofilms from different glass slides within
140 the same channel were scraped independently to form biological replicates (n=4 for each
141 treatment) and suspended in 400 mL of artificial river water (Carles et al., 2022, 2021; Desiante
142 et al., 2022).

143 **2.2. DEET hydrolase metagenomic library construction**

144 Metagenomic sequencing and standard pre-processing and assembly pipelines used are
145 described by Attrah et al. (2023). Protein FASTA files from the 75 metagenomes (Table S1) were
146 queried for hits to the sole characterized reference DEET hydrolase (A3R4E0_PSEPU). Filtering
147 steps for metagenomic library selection are described in the SI. The 64 selected metagenomic
148 DH homolog sequences were codon-optimized for *E. coli* expression using the build-optimization

149 software, BOOST (Oberortner et al., 2017), and cloned by the U.S. Department of Energy Joint
150 Genome Institute with C-terminal 6x-His tags, a tobacco etch virus cleavage site, and additional
151 flexible linker into the first multiple cloning site of pCDF-Duet vectors retaining the NdeI and HindIII
152 restriction sites. Sequence-verified constructs were transformed into T7 Express Competent *E.*
153 *coli* cells (NEB C2566I).

154 **2.3. Sequence and structure bioinformatic analysis**

155 Read mapping and coverage calculations for the *P. putida* reference DH and full-length
156 DH homologs from the above-described metagenomic library were computed as described in the
157 SI. DH phylogenetic analysis is further described in the SI. Full-length sequences of the 64
158 metagenomic DH proteins and reference enzymes were used as inputs for structural modeling
159 using AlphaFold 2.0 ColabFold v1.3.0 (Mirdita et al., 2022) with default parameters. Structural
160 alignment methods are described in the SI.

161 **2.4. Colorimetric activity screening of *P. putida* DH and DH homologs**

162 Activity screening for the 64 metagenomic DH homologs and the *P. putida* reference
163 sequence was performed using a previously established colorimetric screening protocol for
164 hydrolase activity (Robinson et al., 2020; Smith et al., 2020) and described in the SI.

165 **2.5. HPLC analysis of *P. putida* DH and metagenomic DH homologs hydrolysis activity**

166 The enzyme activity of *P. putida* DH and DH homologs with DEET and methylbenzoate
167 (MB) were tested in crude extracts containing soluble proteins. DEET and MB (final conc. 250
168 μM) were added to individual vials containing crude extract separately. Subsequently, reaction
169 progress was monitored by sampling and quenching the reaction with an equal volume of
170 acetonitrile at defined time intervals at 0, 1, 3, 5 and 24h and measured by HPLC (Summit HPLC
171 system, Dionex). Detailed information regarding the HPLC method is included in the SI.

172 **2.6. Purification of *P. putida* DH and metagenomic DH homologs**

173 Induction and expression of DH homologs, cell lysis, and harvesting of cell pellets were
174 performed as described in the SI with 400 mL Terrific broth culture media (12 g/L tryptone, 24 g/L

175 yeast extract, 0.4% (v/v) glycerol) supplemented with 10% (vol/vol) phosphate buffer (0.17 M
176 KH_2PO_4 and 0.72 M K_2HPO_4). Protein purification was performed using benchtop purification for
177 initial tests and FPLC (ÄKTA Pure 25, Cytiva) for data and figures reported in the manuscript.
178 Fractions obtained from both benchtop and FPLC purifications were subjected to analysis by
179 SDS-PAGE (Figure S2 – S4). The detailed purification methods can be found in the SI.

180 **2.7. TrOCs selection**

181 A total of 183 TrOCs were chosen for biotransformation experiments with purified *P. putida*
182 DH and the highly active DH homolog p055 selected as described in the Results & Discussion.
183 These TrOCs are commonly used pesticides, pharmaceuticals, and artificial sweeteners and were
184 selected to equally cover a number of relevant enzymatic transformation reactions. Stock
185 solutions of TrOCs were individually prepared in DMSO at a concentration of 10 mM and stored
186 at −20°C until use. Detailed information including the suppliers of each compound can be found
187 in Table S4 and S5

188 **2.8. High-throughput biotransformation experiments of TrOCs in 96-well plates**

189 Purified *P. putida* DH and the metagenomic DH homolog p055 were tested for their activity
190 with 14 sub-mixtures composed of the 183 TrOCs instead of testing individual compounds. To
191 avoid potential cumulative toxicities and competitive inhibition, we first conducted a hierarchical
192 clustering of all compounds based on their chemical fingerprints and reaction rules triggered by
193 enviPath (Wicker et al., 2016). Sub-mixtures were then designed to contain compounds with
194 different annotated reaction rules. Each of the 14 sub-mixtures (200 μM for each TrOC) was
195 prepared via appropriate dilution of the stock solutions in EtOH (Table S4 and S5) and stored at
196 −20°C for later biotransformation experiments. The capability of TrOCs biotransformation by
197 different enzymes was investigated in a clear-bottomed 96-well plate (Greiner). Individual TrOC
198 sub-mixtures were first added to an empty 96-well plate and the organic solvent EtOH was
199 completely evaporated before 245 μL 10 mM Tris-HCl buffer (pH=8) was added to each well. The
200 plate was shaken at 50 rpm for 15 min to redissolve the TrOCs. Individually purified enzymes

201 (final concentration of ~190 nM) were then added to a well plate and mixed thoroughly. Samples
202 for LC-HRMS/MS analysis were collected at 0, 1, 6, and 24h. At each sampling event, a 20 μL
203 sample was taken and mixed with the same amount of pure methanol to completely denature the
204 enzyme, 160 μL buffer was then added to the sample mixture for an approximately 10 times
205 dilution of the original sample. After centrifugation (13,000 × g, 4°C for 15 min), the supernatant
206 was collected and stored at 4°C for LC-HRMS/MS analysis. An enzyme-free abiotic fresh buffer
207 control and trichloroacetic acid (TCA) treated inactivated enzyme control were set up in the same
208 way as the biological experiment described above. All experiments were conducted in duplicate.
209 Samples were submitted to LC-HRMS/MS analysis within two days after collection. Matrix match
210 calibration standards were used for quantification. Quantification was based on concentrations
211 calculated from peak area integration using TraceFinder 5.1 (Thermo Fisher Scientific).

212 **2.9. UHPLC-HRMS/MS analysis**

213 The parent compounds and transformation products (TPs) were analyzed by UHPLC-
214 HRMS/MS (Q Exactive, Thermo Fisher Scientific) using a method as previously described but
215 with slight modifications (Yu et al., 2022). Detailed methods for the LC method can be found in
216 the SI. Xcalibur 4.0 (Thermo Fisher Scientific) was used for data acquisition and analysis. Suspect
217 screening for DEET transformation products was conducted using TraceFinder 5.1. ChemDraw
218 Professional 20.0 and MarvinSketch (v19.20.0, ChemAxon, http://www.chemaxon.com) were
219 used for drawing, displaying, and characterizing chemical structures. The minimum incubation
220 time for different control groups at which the maximum removal percentage of individual
221 compounds was achieved was calculated based on Equation 1. The results were included in
222 Table S4 (for DEET hydrolase) and Table S5 (for p055). Significant biological removal was
223 defined as following criteria: (1) > 20% removal in enzymatic control, and (2) at least 10% more
224 removal compared to both abiotic buffer control and inactivated enzyme control.

225 Equation 1. $Removal\ Percentage\ (\%) = \frac{C_0 - C_{final}}{C_0} \times 100\%$

226 Where C_0 is the initial concentration measured by LC-HRMS/MS and C_{final} is the final
227 concentration after biotransformation experiments.

228 **3. Results and Discussion**

229 **3.1. Metagenomic sequencing and analysis of WW-impacted biofilm communities**

230 We initiated this study to identify enzymes involved in the biotransformation of TrOCs
231 exhibiting the *downstream effect* in WW-impacted stream biofilms. Our previous study employed
232 16S rRNA gene analysis to identify differentially-abundant amplicon sequence variants correlating
233 with these TrOC biotransformations (Desiante et al., 2022). However, since 16S rRNA gene
234 sequencing doesn't capture functional genes, here we used deeply-sequenced metagenomes (up
235 to 200 million reads per sample) from the same biofilms (Attrah et al., 2023) and searched for
236 genes linked to the biodegradation of *downstream effect* TrOCs. Among the proposed
237 biotransformation pathways for the seven TrOCs showing a *downstream effect* (Desiante et al.,
238 2022), only one compound (DEET) could be linked to an experimentally validated hydrolase
239 sequence at the time of our analysis (Rivera-Cancel et al., 2007). The sequence of the protein
240 involved in hydrolysis of the *downstream effect* compound cyclamate remained unknown although
241 its partial purification was reported (Nimura et al., 1974). Several other pathways for *downstream*
242 *effect* compounds were predicted (Desiante et al., 2022), but only one additional validated
243 hydrolase for the *downstream effect* compound, acesulfame K, was reported during the course of
244 this analysis (Bonatelli et al., 2023). Using the reported acesulfame K hydrolase gene as a query
245 (Bonatelli et al., 2023), we detected no significant BLAST hits in our assembled metagenomes.
246 However, in the same metagenomes we detected more than 3,500 DEET hydrolase (DH)
247 homologs. This diversity and abundance motivated us to investigate experimentally the link
248 between DH homologs and the DEET biotransformation phenotype observed in the biofilms.

249 Next, we calculated the abundances of DH homologs in assembled metagenomes of
250 biofilms grown in different WW percentages. We used the *Psuedomonas putida* DH as a query
251 since this is the only characterized enzyme known to have DH activity. We quantified sequence

252 counts in the biofilms grown in 0%, 10%, 30%, and 80% treated WW mixed with stream water for
253 two independent flume experiments (Figure 1A). Notably, in two separate, independent flume
254 experiments, we observed a positive correlation between the metagenomic DH abundances in
255 biofilms and the percentage of unfiltered WW the biofilms were grown in (Figure 1A, R^2 = 0.77 for
256 Exp. 2, R^2 = 0.42 for Exp. 1). We also calculated the coverage of metagenomic reads mapping to
257 the *P. putida* DH reference in the stream, filtered WW, and unfiltered WW samples and identified
258 significantly more reads mapping in WW relative to stream water (Figure S5, *p*-value < 2e-16).

259 **3.2. Enzyme-mediated biotransformation of TrOCs**

260 The positive association between DH homolog abundances and treated wastewater
261 percentages suggested that DEET hydrolysis was an enzyme-mediated biotransformation
262 reaction occurring in WW and WW-impacted stream biofilms. We detected the expected DEET
263 hydrolysis transformation product (TP137) in biofilms grown at 80% WW (Figure 1C). The
264 transformation products were consistent with the higher abundance of DH homologs found in
265 these biofilms (Figure 1A). Besides the DEET hydrolysis product, we also detected the product of
266 tolyl group oxidation (i.e., TP222) in 80% WW-grown biofilms, indicating that oxidative
267 biotransformation of DEET also occurred. A previous study on microbial degradation of DEET
268 using activated sludge exclusively reported transformation products of oxidative transformation
269 reactions i.e., amide *N*-deethylation and tolyl group oxidation (Helbling et al., 2010b) (Figure 1B).
270 The *N*-deethylation product, TP164, was not detected in the biofilms probably due to its transient
271 behavior as described previously (Helbling et al., 2010b). In contrast, the amide hydrolysis product
272 DEET detected here (Figure 1B) was only previously reported from a pure culture of *P. putida*
273 isolated from activated sludge through enrichments on 2.6 mM DEET (Rivera-Cancel et al., 2007).

274 To further understand the substrate specificity of this reference *P. putida* DH, we cloned,
275 heterologously expressed, and purified the *P. putida* DH as a reference enzyme, and tested its
276 reactivity against a panel of 183 wastewater-related TrOCs. To avoid cumulative inhibitory effects,
277 the 183 WW-relevant compounds were pooled into 14 sub-mixtures (Table S4 and S5).

278 Compounds were separated into sub-mixtures using a pooling strategy that aimed to maximize
279 differences in initial preferred biotransformation reactions as predicted by enviPath (Wicker et al.,
280 2016; Zhang and Fenner, 2023). Results showed that the *P. putida* DH transformed 13 out of the
281 183 TrOCs tested compared to abiotic and inactivated enzyme controls (Figure 1D). With the
282 exceptions of mesalazine, duloxetine, and mefenamic acid, all other compounds biotransformed
283 by the *P. putida* DH contained at least one hydrolyzable moiety (Figure S6). Of those, one
284 compound contained an ester group as sole hydrolyzable moiety, four more compounds
285 contained a combination of esters and one or more amide groups, and five compounds contained
286 exclusively secondary or tertiary amides as hydrolyzable moieties. These results expand on
287 previous findings (Rivera-Cancel et al., 2012) that in addition to the ability of the *P. putida* DH to
288 catalyze hydrolysis of certain tertiary amides, it also exhibits promiscuous activity towards ester
289 and secondary amide-containing contaminants (Figure 1D and Figure S6). Here we use the term
290 'promiscuous activity' to refer to substrate promiscuity, which is the capability of an enzyme to
291 accept other substrates beyond its primary substrate. Of the five compounds for which hydrolysis
292 could exclusively have taken place at the amide functional group (i.e., DEET, MMclB,
293 moclobemide, indomethacin, and tebutam), four (i.e., DEET, MMclB, moclobemide, and
294 indomethacin) contained benzamide functional groups. This points to DH preferring, but not being
295 limited to, the presence of the benzamide moiety in its substrates. Notably, of the over 80 amide-
296 containing compounds tested in our library, five more compounds—enzalutamide, apremilast,
297 tebufenozide, bentazone, and apalutamide—contained benzamide substructures yet did not
298 undergo significant biotransformation. Closer inspection of their molecular structures revealed
299 that all of these compounds contained an ortho-substitution on the benzamide ring with the
300 exception of tebufenozide. This structural characteristic has been previously described to hinder
301 hydrolysis of secondary amides in activated sludge microbial communities due to steric effects
302 (Helbling et al., 2010a).

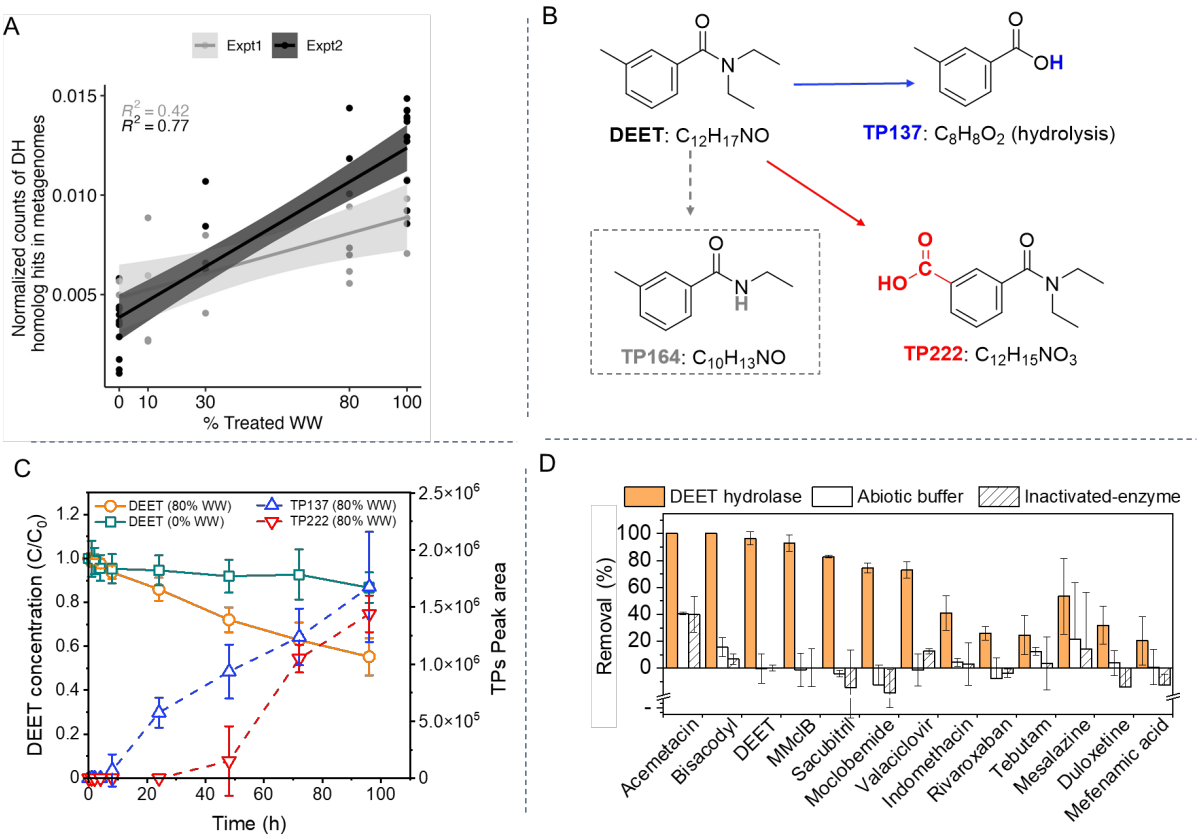


Figure 1: A) Normalized counts of hits in biofilm metagenomes bearing sequence similarity to the *P. putida* DH recovered from assembled metagenomic sequences fitted with a linear regression model (95% confidence interval) across two independent experiments: experiment 1 (light gray) and experiment 2 (dark gray). B) Documented DEET biotransformation pathways and transformation products (TPs) detected in this study. The product TP164 in the dashed-gray box was reported previously (Helbling et al., 2010b) but not detected in this study. C) DEET degradation (solid line) and detected transformation products (TPs) (dashed line) in 80% WW and 0% WW (no TP was detected in 0% WW samples). D) Purified *P. putida* DH substrate specificity for active enzyme (in orange) relative to an inactivated enzyme (hashed lines) and abiotic buffer control (white). Detailed selection criteria and complete biotransformation data are available in the Methods and Table S4. Compounds structures are displayed in Figure S6.

3.3. Construction of a metagenomic library of DEET hydrolase family homologs.

Previous studies on the *P. putida* DH revealed it is a di-domain CocE/NonD family enzyme containing both alpha/beta hydrolase and Xaa-Pro dipeptidyl-peptidase protein domains (Rivera-Cancel et al., 2012, 2007). Referred here for simplicity as the 'DH family', enzymes in this diverse family were also previously shown to cleave cocaine (Bresler et al., 2000) and α-amino acid esters (Barends et al., 2006). The broad substrate range of DH family enzymes beyond DEET hydrolysis

322 along with our results from the *P. putida* DH (Figure 1D) suggested the potential for DH homologs
323 in WW-grown stream biofilms to transform other TrOCs beside DEET. Therefore, we turned our
324 attention to uncharacterized DH homologs detected in the 75 metagenomes from this study (Table
325 S1). After metagenomic assembly and gene calling (Attrah et al., 2023), we selected a library of
326 64 full-length DH homologs based on taxonomic and sequence diversity (Table S2, see SI).
327 Taxonomic annotations for the DH sequences were assigned to total 10 different bacterial phyla:
328 *Acidobacteriota*, *Actinobacteriota*, *Bacteriodota*, *Chloroflexota*, *Firmicutes* (*Bacillota*),
329 *Gemmatimonadota*, *Myxococcota*, *Planctomycetota*, *Proteobacteria* (*Pseudomonadota*), and
330 *Verrucomicrobiota*. Surprisingly, a full-length DH sequence with amino acid sequence identity
331 greater than 40% to the *P. putida* DH reference (Rivera-Cancel et al., 2007) could not be
332 recovered from the assemblies, although we did recover DH homologs from *Pseudomonas* and
333 other related genera. Full-length DH homologs recovered from metagenomes with start and stop
334 codons intact had an average of only 25% amino acid sequence identity with the *P. putida* DH
335 (Table S2). The high level of sequence diversity and general conservation of DH homologs across
336 phyla suggested a more conserved role of DH homologs *e.g.*, in primary metabolism, beyond
337 specialized DEET hydrolysis. In order to test this hypothesis experimentally, we next investigated
338 sequence-structure-function relationships of the metagenomic DH family enzymes (Figure 2).

339 **3.4. Sequence-structure-function relationships among DH family enzymes**

340 We reconstructed a maximum-likelihood phylogeny of metagenomic DH homologs relative
341 to the reference *P. putida* DH (Figure 2A). Although enzymes from many folds are often sequence-
342 diverse, they still often share conserved protein structures (Illergård et al., 2009). Thus, to further
343 test structural similarity of our DH homologs, we used AlphaFold2 (Jumper and Hassabis, 2022)
344 to construct structural models for the 64 metagenomic DH homologs in our library. Structural
345 model comparisons revealed that enzymes shared an average of 1.73 ± 0.43 angstrom (Å) root-
346 mean squared deviation (RMSD) of atomic positions with the *P. putida* reference DH (Table S6).
347 A RMSD less than two indicates relatively high structural similarity and is comparable to values

348 reported in structural alignments of other related members of contaminant-degrading enzyme
349 families (Perez-Garcia et al., 2021; Sidhu et al., 2019). We then compared our sequence
350 phylogeny (Figure 2B) with a neighbor-joining tree generated from the RMSD matrix of all-vs-all
351 structural alignments in our AlphaFold2 library (Figure 2C). This head-to-head tree comparison
352 enabled us to view how clades of sequence-similar DH proteins grouped generally-with
353 structurally-related DH proteins. This approach also enabled us to identify 'structural deviants'
354 among the DH homologs discussed further in section 3.6.

355 Important for this comparison, we verified that our DH homolog structural models were of
356 high confidence with an average predicted local distance difference test (pLDDT) score >80
357 inclusive of unstructured N- and C-termini (Table S6). Structural alignments of catalytic residues
358 in our library with the *P. putida* DH active site revealed 100% conservation of the catalytic serine
359 and greater than 90% conservation of a complete Ser-His-Asp catalytic triad (Table S6). Our
360 metagenomic library contained more variation (Table S6) with respect to oxyanion hole residues
361 (*P. putida* DH: Y84 and W167) and two other known substrate interacting residues M170 and
362 W214 (Rivera-Cancel et al., 2012). We hypothesized that these active site variants occurring in
363 nature might correspond to altered substrate specificity of the DH homologs in our metagenomic
364 library.

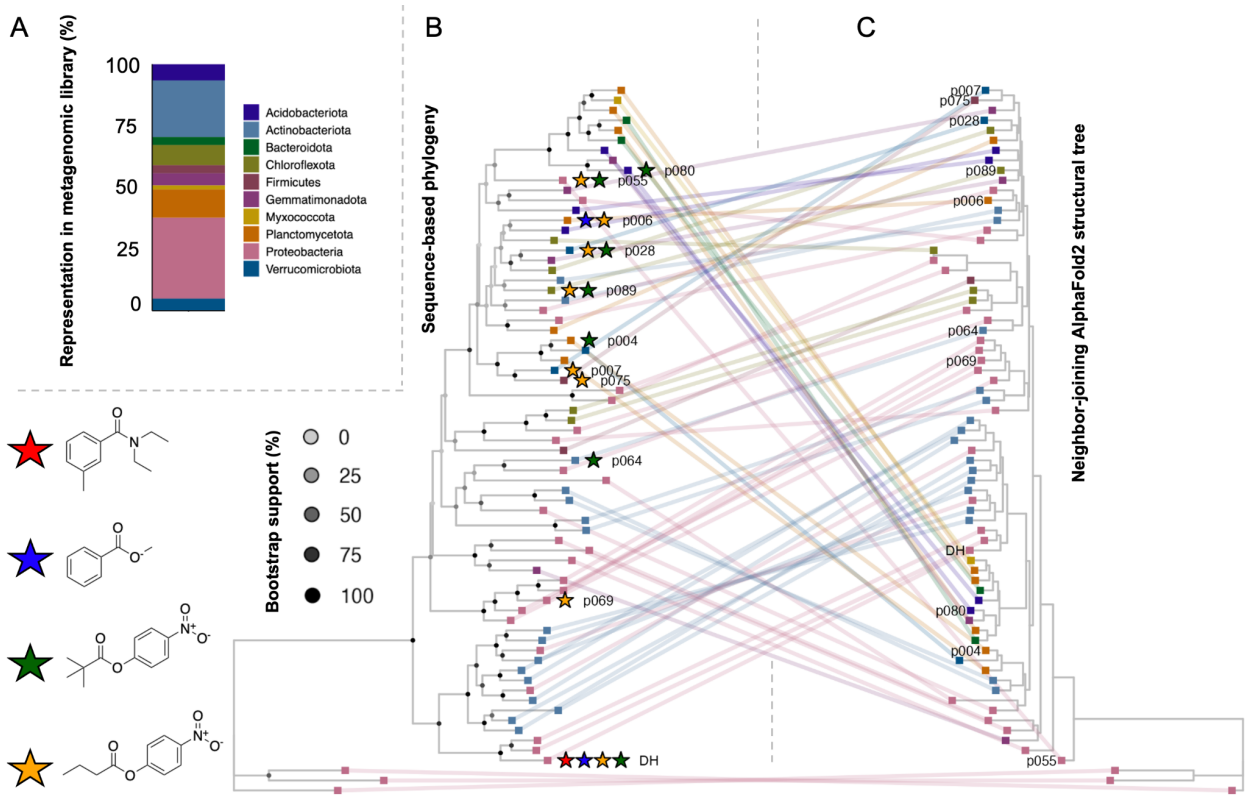


Figure 2. A) Representation of bacterial phyla as a relative abundance within the metagenomic DH homolog library (n = 64) B) Maximum-likelihood sequence-based phylogeny and C) neighbor-joining structural tree of DH family hydrolases. Tips with the same identity are linked between trees. Bootstrap support (1,000 replicates) for internal nodes is indicated by shading. Tips are colored based on the phylum of the assigned taxonomic origin of metagenomic hydrolases. Star colors correspond to activity measured with 4 substrates: Red = DEET (reference compound), Blue = methylbenzoate (aromatic ester), Green = 4-nitrophenyl trimethylacetate (branched aliphatic), Gold = 4-nitrophenyl butyrate (linear aliphatic). Functionally-divergent alpha/beta hydrolases (PDB IDs: 1MJ5, 3HI4 and 6QHV) were used as an outgroup.

3.5. Functional screening of the metagenomic DH homolog library

In order to test the substrate specificity of our metagenomic DH homolog library, we codon optimized, cloned and and expressed each of the 64 metagenomic sequences in *E. coli*. We screened for activity in each of the 64 hydrolase-expressing strains with four different ester or amide-containing substrates (Figure 2): DEET, methylbenzoate, 4-nitrophenyl butyrate, and 4-nitrophenyl trimethylacetate using a combination of HPLC and a previously-optimized 96-well plate protocol (Robinson et al., 2020) (Methods in SI). Eleven out of 64 metagenomic library enzymes tested were identified as active with at least one substrate, primarily with the linear and

384 branched 4-nitrophenyl aliphatic esters (Figure 3). Substrates with ester linkages were preferred
385 by metagenomic DH homologs likely because esters are generally susceptible to nucleophilic
386 attack, e.g., by the catalytic serine of the α/β-fold hydrolase triad conserved among all of the 64
387 DH homologs.

388 Out of the 64 enzymes screened, one metagenomic hydrolase (denoted p006) hydrolyzed
389 methylbenzoate and 4-nitrophenyl butyrate but not DEET (Figure 4). The substrate specificity of
390 p006 is striking in light of previous reports (Rivera-Cancel et al., 2012) confirmed by our own
391 results shown here (Figure 4) that the reference *P. putida* DH hydrolyzes both methylbenzoate
392 and DEET. Compounds containing tertiary amide groups, such as DEET, are more sterically
393 hindered than carboxylic acid esters. Additionally, after testing all 64 DH homologs, surprisingly
394 only the reference *P. putida* DH was capable of DEET hydrolysis (Figure 4). This may be due to
395 the fact that the *P. putida* DH contains active site residues which occur in a unique combination
396 relative to active site residues in the metagenomic DH homologs. Specifically, Y84, W167 and
397 M170 exclusively co-occur in the *P. putida* DH. Swapping each of these residues individually for
398 alanine negatively impacts the kinetics of DEET hydrolysis (Rivera-Cancel et al., 2012). Each of
399 these three residues (Y84, W167 and M170) occurred individually or in pairs in different proteins
400 in our metagenomic DH homolog library. However, all three residues co-occurring together were
401 not identified in any DH homologs (Table S6) in the full dataset or in public databases apart from
402 in the reference *P. putida* DH. This suggests that epistatic interactions between multiple amino
403 acids in the *P. putida* DH active site is critical for DEET hydrolysis. Moreover, in our metagenomic
404 analysis, we identified a naturally-occurring and diverse pool of genes encoding for various
405 combinations of the active site residues Y84, W167 and M170. Such observed genetic diversity
406 within microbial communities indicates a potential for developing specialized functions, especially
407 when exposed to selective pressure from contaminants. Specifically, we propose that different
408 active site variants in this diverse gene pool may contribute to the emergence of specialized

degradation enzymes, such as the *P. putida* DH which was identified under DEET enrichment conditions (Rivera-Cancel et al., 2007).

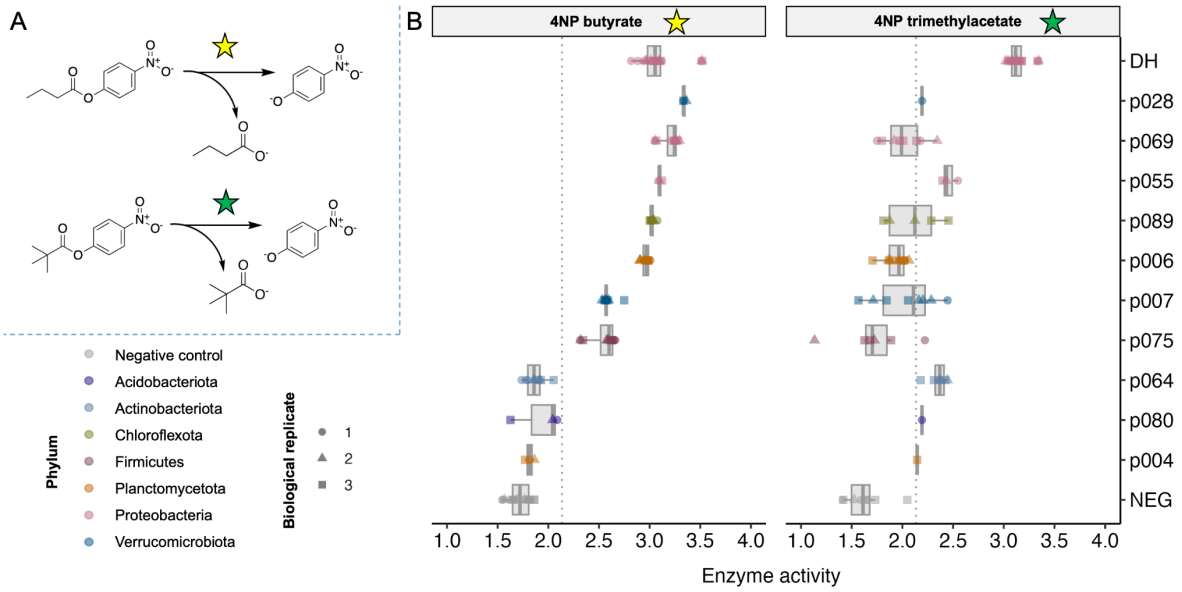


Figure 3. A) Linear and branched 4-nitrophenyl (4NP) ester substrates tested: 4NP butyrate (yellow star) and 4NP trimethylacetate (green star) both release 4NP as a chromophore B) The ten most active metagenomic enzymes in our library tested in a colorimetric screening assay for 4NP hydrolysis activity. Units for enzyme activity are reported as $\log_{10}$ of nanomoles of 4NP released per hour by an *E. coli* cell culture with an optical density unit (OD) of 1.0, enabling direct comparison with the activity values reported from previous enzymatic screening studies (Robinson et al., 2020; Smith et al., 2020). Enzyme activity is measured compared to negative control empty vector (NEG) and reference *P. putida* DEET hydrolase (DH). The average activity level calculated across all enzymes and substrates is drawn as a dotted gray line.

3.6. Biochemical characterization of a new metagenomic hydrolase p055 and its reactivity with TrOCs

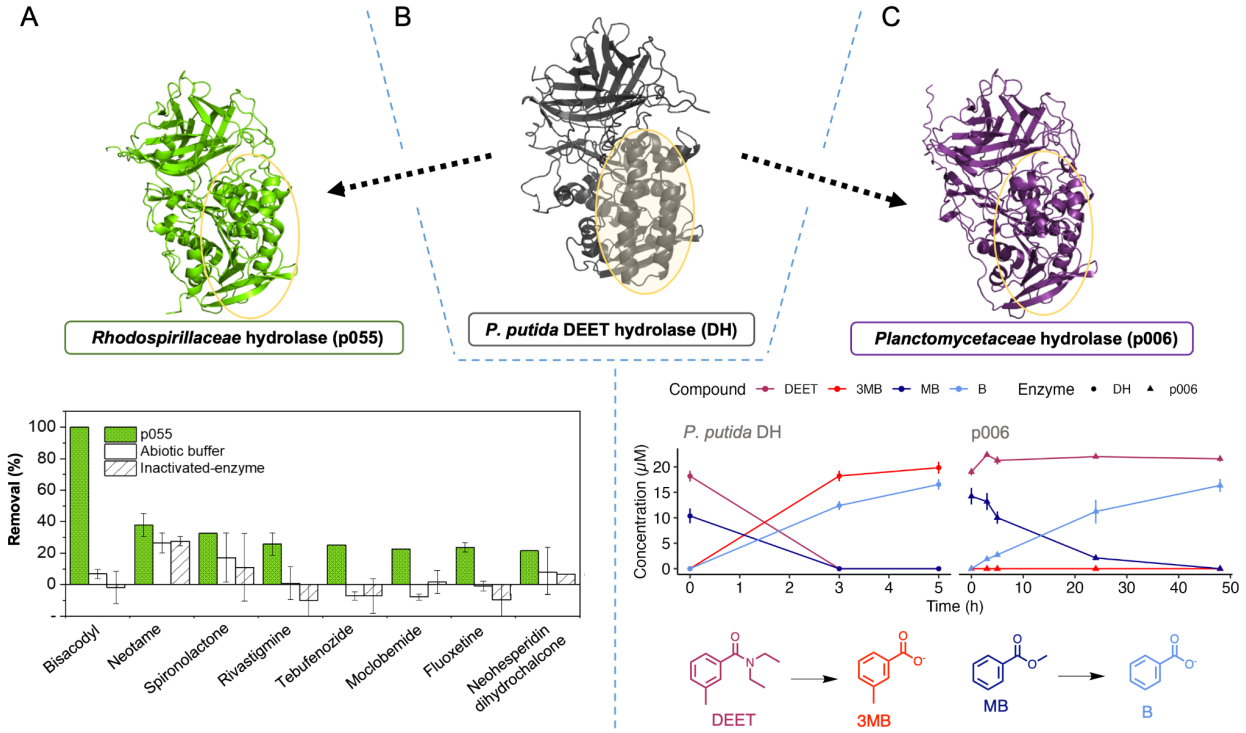
Among the 64 enzymes in our functional metagenomic library, p055 had the highest average combined enzyme activity with 4-nitrophenyl trimethylacetate and 4-nitrophenyl butyrate. Analysis of the DNA and translated protein sequences of the metagenomic sequence p055 could only resolve the taxonomic origin at the family level as belonging to purple nonsulfur bacteria assigned to the family *Rhodospirillaceae*. By mapping metagenomic raw reads to all full-length DH homologs in the library, we identified p055 as the only active enzyme in our library which also showed a clear increasing trend in metagenomic abundance with WW % (Figure S7).

433 Furthermore, p055 was ‘structurally deviant’ from the *P. putida* DH and other DH homologs in the
434 metagenomic library (Figure 2). With respect to primary amino acid sequence identity, p055
435 grouped with other metagenomic hydrolases (Figure 2B), but it was clearly differentiated from the
436 rest of the DH homologs on the basis of its structure (Figure 2C). Closer analysis revealed p055
437 has much shorter alpha helices in one of its domains relative to reference *P. putida* DH (Figure
438 4). Of note, the helical regions are in a high-confidence region of the AlphaFold2 models (pLDDT >
439 95) suggesting the structural deviation was not a modeling artifact. We calculated the overall
440 alpha helix content of the reference *P. putida* DH and found it had higher alpha helix content than
441 all but one other enzyme among the active hydrolases identified in the metagenomic DH homolog
442 library (Table S7). Since the alpha helices in p055 (encircled region, Figure 4) are smaller relative
443 to the *P. putida* DH helices in the vicinity of the active site, we hypothesized p055 structural
444 deviation might also correspond to substrate variation.

445 To test p055 substrate specificity, we expressed and purified the metagenome-derived
446 enzyme to homogeneity as a 66 kDa protein with a yield of 640 µg/mL (Figure S4). We calculated
447 the specific enzymatic activity in buffers over a range of different pHs and determined p055 had
448 a pH optimum between 8.5 – 9 (Figure S8). The enzyme p055 had an optimal temperature of
449 35°C (Figure S9) and displayed activity up to 45°C which corresponds well with the optimal growth
450 temperatures of the *Rhodospirillaceae* as the predicted source (Baldani et al., 2014). Following
451 incubation for one hour at 40°C, the enzyme lost 50% of its activity (Figure S10), indicating p055
452 is not highly thermostable. Loss of activity at 40°C is not unexpected since this enzyme originates
453 from an environmental biofilm grown in WW and stream water where temperatures ranged from
454 12°C - 14°C (Desiante et al., 2022).

455 A screening of the substrate promiscuity of this environmental hydrolase p055 with the
456 183 TrOCs revealed reactivity towards 8 substrates (Figure 5C and Figure S11). By comparing
457 p055 substrate specificity with the *P. putida* DH reference, we identified 20 substrates which were
458 different between the enzymes and two substrates which were shared. Namely, both p055 and

459 *P. putida* DH catalyze the hydrolysis of bisacodyl and moclobemide. Bisacodyl is known as a
460 prodrug that can be metabolized by intestinal and bacterial enzymes to its deacetylated active
461 metabolite. As active esterases, *P. putida* DH and p055 are both capable of facilitating this
462 hydrolytic process.



463

464 **Figure 4.** A) AlphaFold2 model of DH homolog p055 and substrate specificity as measured by
465 LC-MS/MS, B) AlphaFold2 model of *P. putida* DH. The yellow encircled region corresponds to an
466 alpha-helix rich region in the vicinity of the active site which has higher alpha helix content than
467 most other active metagenomic DH homologs (Table S7). C) AlphaFold2 model of p006 and
468 HPLC measurements of DEET and methylbenzoate (MB) and product formation of 3-
469 methylbenzoate (red, denoted as "3MB") and benzoate (light blue, denoted as "B"), respectively,
470 by the *P. putida* DH and p006 (C). Detailed selection criteria and complete biotransformation data
471 are available in Methods and Table S5. Recent analysis of the amino acid sequence p006
472 suggests a more accurate taxonomic placement to the phylum *'Candidatus Hydrogenedentes*'
473 based on 93.6% amino acid identity and 100% query cover with proteins from a metagenome-
474 assembled genome (MAG) from wastewater (Lin et al., 2021).
475
476 Overall, we detected activity in approximately 5% of the 626 distinct enzyme-substrate
477 combinations tested. The activity rate with our targeted metagenome mining approach was higher
478 than most untargeted functional metagenomics screens, where on average less than 0.08% of
479 colonies screened are active, depending on the activity screening method (Ferrer et al., 2016).

480 This result highlights the benefits of *in silico* analysis mapping sequence-structure-function
481 relationships to prioritize candidates for metagenomic libraries.

482 The fact that many of the metagenomic DH homologs tested were active with other organic
483 substrates, yet not active with DEET, also raises the question of whether there are additional
484 responsible enzyme(s) for DEET hydrolysis in stream biofilms. While DH homologs with high
485 similarity to the reference *P. putida* DH were not recovered in the metagenomic assembly
486 process, metagenomic short reads mapped to the *P. putida* DH (Figure S5). Therefore 'true' DHs
487 with activity similar to the *P. putida* DH are therefore likely present in low abundance. This also
488 emphasizes the limitations of metagenomic assembly-based approaches for capturing rare, low-
489 abundance genes involved in TrOC biotransformations. Here, we examined the only enzyme
490 family for which we could match an observed TrOC biotransformation to a validated reference
491 sequence detected in our metagenomes. In the environment, we also expect diverse hydrolase
492 families beyond DH and its homologs to contribute to hydrolysis of diverse TrOCs (Hübner et al.,
493 2021), including *downstream effect* compounds. Other DEET biotransformation products e..g,
494 formed from oxidative *N*-deethylation, were observed but the responsible enzymes are currently
495 not well-understood and represent future directions.

496 It is also worth noting that, within microbial communities, enzymes and their producers
497 exist in a mixture and certain transformation products of a TrOC may be further transformed by
498 enzymes produced by different community members (Wang et al., 2022). All the evidence points
499 to the importance of maintaining (bio)diversity in natural water systems to effectively cope with
500 and challenges by novel entities and anthropogenic activities. In our study, this was exemplified
501 by characterizing new DH homologs spanning ten different microbial phyla from complex WW-
502 impacted biofilm communities. Future studies investigating diverse enzyme families are
503 necessary to expand reference sets for TrOC biodegradation. This will not only deepen our
504 understanding of the fate of TrOCs in environments, but also facilitate the development of more

505 accurate prediction models for assessing TrOC biotransformations in natural and engineered
506 aquatic systems.
507
508 **4. Conclusions**
509 Through combined metagenome mining and experimental characterization, this work
510 yielded new insights into the involvement of DH family enzymes in TrOC biotransformations.
511 Several major conclusions arose from this study:
512 - This work provided lessons for the computational prediction of responsible enzymes for
513 TrOC removal including the importance of setting a high sequence similarity for assigning
514 protein functions. Depending on enzyme class, functions may or may not be conserved
515 when sequence similarity is low (e.g., less than 40%), even when structural similarity is
516 high as in the case of the DH homologs. Here we characterized enzymes with low amino
517 acid identity to the reference DH. Enzymes bearing low similarity to reference sequences
518 also come from unexplored regions of sequence space thereby presenting opportunities
519 for the discovery of new active metagenomic hydrolases such as p055 with altered
520 substrate specificity for other TrOCs.
521 - Activity-based screening approaches provide complementary and separate information
522 from sequence- and structure-based approaches to uncover enzymes involved in TrOC
523 biotransformations, as exemplified here. Additional strategies such as machine learning
524 based enzyme-substrate promiscuity predictions or (meta)transcriptomics to identify
525 candidates with differential expression patterns may point towards candidates not
526 detected by shifts in metagenomic abundances.
527 - Finally, our findings serve as a warning that correlations in ‘omics data do not equal
528 causation. Experimental validation is often critical to test hypotheses generated from
529 microbiome analyses. Metagenomic DH homologs were identified by bioinformatics
530 methods as being active in DEET biotransformation but were not active with DEET,

despite being active with other substrates. In addition to sequence similarity, further knowledge on specific amino acids (e.g., active site), secondary structural features (e.g., alpha helices as analyzed here), and experimental or environmental conditions (e.g., pH, temperature) must be incorporated to improve the prediction of TrOC removal in environmental systems.

CRediT Authorship Contribution Statement

Y.Y.: Conceptualization, Methodology, Investigation, Formal analysis, Writing - Original Draft, Writing - Review & Editing, Visualization; N.F.T.: Methodology, Investigation, Formal analysis, Visualization, Writing - Review & Editing; M.R.S.: Methodology, Formal analysis, Visualization, Writing - Review & Editing; K.F.: Conceptualization, Writing - Review & Editing, Supervision, Funding acquisition; S.L.R.: Conceptualization, Writing - Original Draft, Writing - Review & Editing, Supervision, Funding acquisition.

Declaration of Competing Interest

The authors declare no competing financial interest.

Data Availability

Sequencing data are available on NCBI under the BioProject accession number PRJNA1008123. Scripts for data analysis and figures in this manuscript are available on github.com/MSM-group/DEET-hydrolase.

Acknowledgments

We acknowledge Thomas Fleischmann, René Gall, Max Hofland, and Florian Felder for their technical support. We thank Elia Ceppi and Philipp Longree for assisting with the development of UHPLC-HRMS/MS methods, and Martina Kalt and Miguel Heussi for preparing chemical stock solutions. We thank Kunyang Zhang for assisting with cheminformatics analyses. We thank Giulia Gionchetta, Robert Niederdorfer, and Helmut Bürgmann for their contributions in the metagenomic analysis. We also acknowledge Yasuo Yoshikuni, Ian Blaby, and Miranda Harmon-Smith for their support with metagenomic library constructs. The work (proposal:

10.46936/10.25585/60008420) conducted by the U.S. Department of Energy Joint Genome Institute (https://ror.org/04xm1d337), a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. Department of Energy operated under Contract No. DE-AC02-05CH11231. This work was supported by Eawag Discretionary Funding (for S.L.R. and K.F.) and Swiss National Science Foundation (project no. 200021L_201006, for Y.Y. and K.F.; PZPGP2_209124 for S.L.R.). The graphical abstract was created with a licensed version of BioRender.com.

565 **References**

566 Achermann, S., Mansfeldt, C.B., Müller, M., Johnson, D.R., Fenner, K., 2020. Relating
567 Metatranscriptomic Profiles to the Micropollutant Biotransformation Potential of Complex
568 Microbial Communities. Environ. Sci. Technol. 54, 235–244.

569 Altenburger, R., Scholze, M., Busch, W., Escher, B.I., Jakobs, G., Krauss, M., Krüger, J., Neale,
570 P.A., Ait-Aissa, S., Almeida, A.C., Seiler, T.-B., Brion, F., Hilscherová, K., Hollert, H.,
571 Novák, J., Schlichting, R., Serra, H., Shao, Y., Tindall, A., Tollefsen, K.E., Umbuzeiro, G.,
572 Williams, T.D., Kortenkamp, A., 2018. Mixture effects in samples of multiple contaminants -
573 An inter-laboratory study with manifold bioassays. Environ. Int. 114, 95–106.

574 Athanasakoglou, A., Fenner, K., 2022. Toward characterizing the genetic basis of trace organic
575 contaminant biotransformation in activated sludge: The role of multicopper oxidases as a
576 case study. Environ. Sci. Technol. 56, 313–324.

577 Attrah, M., Schaerer, M.R., Esposito, M., Gionchetta, G., Buergmann, H., Lens, P., Fenner, K.,
578 van de Vossenberg, J., Robinson, S.L., 2023. Disentangling abiotic and biotic effects of
579 treated wastewater on stream biofilm resistomes enables the discovery of a new
580 planctomycete beta-lactamase. *bioRxiv*. https://doi.org/10.1101/2023.11.20.567610

581 Baldani, J.I., Videira, S.S., dos Santos Teixeira, K.R., Reis, V.M., de Oliveira, A.L.M., Schwab,
582 S., de Souza, E.M., Pedraza, R.O., Baldani, V.L.D., Hartmann, A., 2014. The family
583 *Rhodospirillaceae*, in the Prokaryotes. Springer Berlin Heidelberg, Berlin, Heidelberg, pp.
584 533–618.

585 Barends, T.R.M., Polderman-Tijmes, J.J., Jekel, P.A., Williams, C., Wybenga, G., Janssen,
586 D.B., Dijkstra, B.W., 2006. *Acetobacter turbidans* alpha-amino acid ester hydrolase: how a
587 single mutation improves an antibiotic-producing enzyme. J. Biol. Chem. 281, 5804–5810.

588 Battin, T.J., Besemer, K., Bengtsson, M.M., Romani, A.M., Packmann, A.I., 2016. The ecology
589 and biogeochemistry of stream biofilms. Nat. Rev. Microbiol. 14, 251–263.

590 Bonatelli, M.L., Rohwerder, T., Popp, D., Liu, Y., Akay, C., Schultz, C., Liao, K.-P., Ding, C.,

591 Reemtsma, T., Adrian, L., Kleinsteuber, S., 2023. Recently evolved combination of unique
592 sulfatase and amidase genes enables bacterial degradation of the wastewater
593 micropollutant acesulfame worldwide. Front. Microbiol. 14, 1223838.
594 Bresler, M.M., Rosser, S.J., Basran, A., Bruce, N.C., 2000. Gene cloning and nucleotide
595 sequencing and properties of a cocaine esterase from *Rhodococcus* sp. strain MB1. Appl.
596 Environ. Microbiol. 66, 904–908.
597 Burdon, F.J., Bai, Y., Reyes, M., Tamminen, M., Staudacher, P., Mangold, S., Singer, H.,
598 Räsänen, K., Joss, A., Tiegs, S.D., Jokela, J., Eggen, R.I.L., Stamm, C., 2020. Stream
599 microbial communities and ecosystem functioning show complex responses to multiple
600 stressors in wastewater. Glob. Chang. Biol. 26, 6363–6382.
601 Caliman, F.A., Gavrilescu, M., 2009. Pharmaceuticals, Personal Care Products and Endocrine
602 Disrupting Agents in the Environment – A Review. Clean Soil Air Water 37, 277–303.
603 Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L.,
604 2009. BLAST+: architecture and applications. BMC Bioinformatics 10, 421.
605 Carles, L., Wullschleger, S., Joss, A., Eggen, R.I.L., Schirmer, K., Schuwirth, N., Stamm, C.,
606 Tlili, A., 2022. Wastewater microorganisms impact microbial diversity and important
607 ecological functions of stream periphyton. Water Res. 225, 119119.
608 Carles, L., Wullschleger, S., Joss, A., Eggen, R.I.L., Schirmer, K., Schuwirth, N., Stamm, C.,
609 Tlili, A., 2021. Impact of wastewater on the microbial diversity of periphyton and its
610 tolerance to micropollutants in an engineered flow-through channel system. Water Res.
611 203, 117486.
612 Che, S., Jin, B., Liu, Z., Yu, Y., Liu, J., Men, Y., 2021. Structure-Specific Aerobic Defluorination
613 of Short-Chain Fluorinated Carboxylic Acids by Activated Sludge Communities. Environ Sci
614 Technol Lett 8, 668–674.
615 Daughton, C.G., Ternes, T.A., 1999. Pharmaceuticals and personal care products in the
616 environment: agents of subtle change? Environ. Health Perspect. 107 Suppl 6, 907–938.

Desiante, W.L., Carles, L., Wullschleger, S., Joss, A., Stamm, C., Fenner, K., 2022. Wastewater microorganisms impact the micropollutant biotransformation potential of natural stream biofilms. Water Res. 217, 118413.

Desiante, W.L., Minas, N.S., Fenner, K., 2021. Micropollutant biotransformation and bioaccumulation in natural stream biofilms. Water Res. 193, 116846.

Evich, M.G., Davis, M.J.B., McCord, J.P., Acrey, B., Awkerman, J.A., Knappe, D.R.U., Lindstrom, A.B., Speth, T.F., Tebes-Stevens, C., Strynar, M.J., Wang, Z., Weber, E.J., Henderson, W.M., Washington, J.W., 2022. Per- and polyfluoroalkyl substances in the environment. Science 375, eabg9065.

Fenner, K., Canonica, S., Wackett, L.P., Elsner, M., 2013. Evaluating pesticide degradation in the environment: blind spots and emerging opportunities. Science 341, 752–758.

Fenner, K., Elsner, M., Lueders, T., McLachlan, M.S., Wackett, L.P., Zimmermann, M., Drewes, J.E., 2021. Methodological advances to study contaminant biotransformation: new prospects for understanding and reducing environmental persistence? ACS ES&T Water 1, 1541–1554.

Ferrer, M., Martínez-Martínez, M., Bargiela, R., Streit, W.R., Golyshina, O.V., Golyshin, P.N., 2016. Estimating the success of enzyme bioprospecting through metagenomics: current status and future trends. Microb. Biotechnol. 9, 22–34.

Halden, R., 2011. Contaminants of Emerging Concern in the Environment: Ecological and Human Health Considerations. OUP USA.

Han, P., Rios-Miguel, A.B., Tang, X., Yu, Y., Zhou, L.-J., Hou, L., Liu, M., Sun, D., Jetten, M.S.M., Welte, C.U., Men, Y., Lücker, S., 2023. Benzimidazole fungicide biotransformation by comammox *Nitrospira* bacteria: Transformation pathways and associated proteomic responses. J. Hazard. Mater. 445, 130558.

Han, P., Yu, Y., Zhou, L., Tian, Z., Li, Z., Hou, L., Liu, M., Wu, Q., Wagner, M., Men, Y., 2019. Specific Micropollutant Biotransformation Pattern by the Comammox Bacterium. Environ.

643 Sci. Technol. 53, 8695–8705.
644 Helbling, D.E., Hollender, J., Kohler, H.-P.E., Fenner, K., 2010a. Structure-based interpretation
645 of biotransformation pathways of amide-containing compounds in sludge-seeded
646 bioreactors. Environ. Sci. Technol. 44, 6628–6635.
647 Helbling, D.E., Hollender, J., Kohler, H.-P.E., Singer, H., Fenner, K., 2010b. High-throughput
648 identification of microbial transformation products of organic micropollutants. Environ. Sci.
649 Technol. 44, 6621–6627.
650 Hübner, U., Wolff, D., Achermann, S., Drewes, J.E., Wick, A., Fenner, K., 2021. Analyzing
651 (Initial) Biotransformation Reactions as an Organizing Principle for Unraveling the Extent of
652 Trace Organic Chemical Biotransformation in Biofiltration Systems. ACS EST Water. 1,
653 1921–1931
654 Illergård, K., Ardell, D.H., Elofsson, A., 2009. Structure is three to ten times more conserved
655 than sequence--a study of structural response in protein cores. Proteins 77, 499–508.
656 Jin, B., Liu, H., Che, S., Gao, J., Yu, Y., Liu, J., Men, Y., 2023. Substantial defluorination of
657 polychlorofluorocarboxylic acids triggered by anaerobic microbial hydrolytic dechlorination.
658 Nature Water 1, 451–461.
659 Jumper, J., Hassabis, D., 2022. Protein structure predictions to atomic accuracy with AlphaFold.
660 Nat. Methods 19, 11–12.
661 Kennes-Veiga, D.M., Trueba-Santiso, A., Gallardo-Garay, V., Balboa, S., Carballa, M., Lema,
662 J.M., 2022. Sulfamethoxazole enhances specific enzymatic activities under aerobic
663 heterotrophic conditions: A metaproteomic approach. Environ. Sci. Technol. 56, 13152–
664 13159.
665 Lin, Y., Wang, L., Xu, K., Li, K., Ren, H., 2021. Revealing taxon-specific heavy metal-resistance
666 mechanisms in denitrifying phosphorus removal sludge using genome-centric
667 metaproteomics. Microbiome 9, 67.
668 Mansfeldt, C., Deiner, K., Mächler, E., Fenner, K., Eggen, R.I.L., Stamm, C., Schönenberger,

669 U., Walser, J.-C., Altermatt, F., 2020. Microbial community shifts in streams receiving
670 treated wastewater effluent. Sci. Total Environ. 709, 135727.
671 Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., Steinegger, M., 2022.
672 ColabFold: making protein folding accessible to all. Nat. Methods 19, 679–682.
673 Nimura, T., Tokieda, T., Yamaha, T., 1974. Partial purification and some properties of cyclamate
674 sulfamatase. J. Biochem. 75, 407–417.
675 Oberortner, E., Cheng, J.-F., Hillson, N.J., Deutsch, S., 2017. Streamlining the Design-to-Build
676 Transition with Build-Optimization Software Tools. ACS Synth. Biol. 6, 485–496.
677 Perez-Garcia, P., Kobus, S., Gertzen, C.G.W., Hoeppner, A., Holzscheck, N., Strunk, C.H.,
678 Huber, H., Jaeger, K.-E., Gohlke, H., Kovacic, F., Smits, S.H.J., Streit, W.R., Chow, J.,
679 2021. A promiscuous ancestral enzyme´s structure unveils protein variable regions of the
680 highly diverse metallo-β-lactamase family. Commun Biol 4, 132.
681 Rich, S.L., Helbling, D.E., 2023. Daily Monitoring at a Full-Scale Wastewater Treatment Plant
682 Reveals Temporally Variable Micropollutant Biotransformations. Environ. Sci. Technol. 57,
683 10404–10414.
684 Rivera-Cancel, G., Bocioaga, D., Hay, A.G., 2007. Bacterial degradation of *N,N*-diethyl-m-
685 toluamide (DEET): cloning and heterologous expression of DEET hydrolase. Appl. Environ.
686 Microbiol. 73, 3105–3108.
687 Rivera-Cancel, G., Sanders, J.M., Hay, A.G., 2012. Kinetics of hydrolysis and mutational
688 analysis of N,N-diethyl-m-toluamide hydrolase from *Pseudomonas putida* DTB. FEBS J.
689 279, 1044–1053.
690 Robinson, S.L., Smith, M.D., Richman, J.E., Aukema, K.G., Wackett, L.P., 2020. Machine
691 learning-based prediction of activity and substrate specificity for OleA enzymes in the
692 thiolase superfamily. Synth. Biol. 5, ysaa004.
693 Salazar, G., Ruscheweyh, H.-J., Hildebrand, F., Acinas, S.G., Sunagawa, S., 2021. mTAGs:
694 taxonomic profiling using degenerate consensus reference sequences of ribosomal RNA

genes. Bioinformatics 38, 270–272.

Schwarzenbach, R.P., Escher, B.I., Fenner, K., Hofstetter, T.B., Johnson, C.A., von Gunten, U., Wehrli, B., 2006. The challenge of micropollutants in aquatic systems. Science 313, 1072–1077.

Sidhu, C., Solanki, V., Pinnaka, A.K., Thakur, K.G., 2019. Structure elucidation and biochemical characterization of environmentally relevant novel extradiol dioxygenases discovered by a functional metagenomics approach. mSystems 4, e00316-19.

Smith, M.D., Robinson, S.L., Molomjamts, M., Wackett, L.P., 2020. Assay reveals microbial OleA thiolases initiating hydrocarbon and β-Lactone biosynthesis. MBio 11, 00111-20.

Stadler, L.B., Delgado Vela, J., Jain, S., Dick, G.J., Love, N.G., 2018. Elucidating the impact of microbial community biodiversity on pharmaceutical biotransformation during wastewater treatment. Microb. Biotechnol. 11, 995–1007.

Tien, C.-J., Lin, M.-C., Chiu, W.-H., Chen, C.S., 2013. Biodegradation of carbamate pesticides by natural river biofilms in different seasons and their effects on biofilm community structure. Environ. Pollut. 179, 95–104.

Vasileiadis, S., Perruchon, C., Scheer, B., Adrian, L., Steinbach, N., Trevisan, M., Plaza-Bolaños, P., Agüera, A., Chatzinotas, A., Karpouzas, D.G., 2022. Nutritional inter-dependencies and a carbazole-dioxygenase are key elements of a bacterial consortium relying on a *Sphingomonas* for the degradation of the fungicide thiabendazole. Environ. Microbiol. 24, 5105–5122.

Vega, M.A.P., Scholes, R.C., Brady, A.R., Daly, R.A., Narrowe, A.B., Bosworth, L.B., Wrighton, K.C., Sedlak, D.L., Sharp, J.O., 2022. Pharmaceutical Biotransformation is Influenced by Photosynthesis and Microbial Nitrogen Cycling in a Benthic Wetland Biomat. Environ. Sci. Technol. 56, 14462–14477.

Vega, M.A.P., Scholes, R.C., Brady, A.R., Daly, R.A., Narrowe, A.B., Vanzin, G.F., Wrighton, K.C., Sedlak, D.L., Sharp, J.O., 2023. Methane-Oxidizing Activity Enhances

721 Sulfamethoxazole Biotransformation in a Benthic Constructed Wetland Biomat. Environ.
722 Sci. Technol. 57, 7240–7253.
723 Wang, M., Chen, X., Liu, X., Fang, Y., Zheng, X., Huang, T., Tang, Y.-Q., Ackermann, M., Nie,
724 Y., Wu, X.-L., 2022. Even allocation of benefits stabilizes microbial community engaged in
725 metabolic division of labor. Cell Rep. 40, 111410.
726 Wicker, J., Lorsbach, T., Gütlein, M., Schmid, E., Latino, D., Kramer, S., Fenner, K., 2016.
727 enviPath--The environmental contaminant biotransformation pathway resource. Nucleic
728 Acids Res. 44, D502–8.
729 Xing, Y., Yu, Y., Men, Y., 2018. Emerging investigators series: occurrence and fate of emerging
730 organic contaminants in wastewater treatment plants with an enhanced nitrification step.
731 Environ. Sci.: Water Res. Technol. 4, 1412–1426.
732 Yu, Y., Che, S., Ren, C., Jin, B., Tian, Z., Liu, J., Men, Y., 2022. Microbial Defluorination of
733 Unsaturated Per- and Polyfluorinated Carboxylic Acids under Anaerobic and Aerobic
734 Conditions: A Structure Specificity Study. Environ. Sci. Technol. 56, 4894–4904.
735 Yu, Y., Han, P., Zhou, L.-J., Li, Z., Wagner, M., Men, Y., 2018. Ammonia monooxygenase-
736 mediated cometabolic biotransformation and hydroxylamine-mediated abiotic
737 transformation of micropollutants in an AOB/NOB coculture. Environ. Sci. Technol. 52,
738 9196–9205.
739 Zhang, K., Fenner, K., 2023. enviRule: An End-to-end System for Automatic Extraction of
740 Reaction Patterns from Environmental Contaminant Biotransformation Pathways.
741 Bioinformatics. 39, btad407.