Function and Global Distribution of Polyethylene Terephthalate (PET)-Degrading Bacteria and Enzymes in Marine and Terrestrial Metagenomes

Dissertation

for obtaining the degree of

Doctor rerum naturalium (Dr. rer. nat.)

at the Department of Biology

Subdivision at the Faculty of Mathematics, Informatics and Natural Sciences

of the Universität Hamburg

submitted by

Dominik Danso

from Hamburg

Hamburg 2018

Genehmigt vom Fachbereich Biologie der Universität Hamburg

auf Antrag von Prof. Dr. W. Streit

Weiterer Gutachter der Dissertation: Prof. Dr. Bodo Saake

Tag der Disputation: 25.01.2019

GEOMAR
Helmholtz-Zentrum für Ozeanforschung Kiel

GEOMAR | Wischhofstraße 1-3 | 24148 Kiel | Germany

Prof. Dr. Mirjam Perner
Geomicrobiology

RU Marine Geosystems

Tel +49 431 600-2837
Fax +49 431 600-2941
mperner@geomar.de

English language declaration

16. Oktober 2018

As a native speaker, I hereby declare that I have checked the thesis „The function and global distribution of polyethylene terephthalate (PET) degrading bacteria and enzymes “by Dominik Danso for grammatically correct English and the scientific accuracy of the language.

Sincerely,

Mirjam Perner

GEOMAR
Helmholtz-Zentrum für
Ozeanforschung Kiel

Wischhofstraße 1-3
24148 Kiel | Germany

Tel +49 431 600-0
Fax +49 431 600-2805
www.geomar.de

Deutsche Bank AG Kiel
BLZ 210 700 24
Kto. 144 8000

SWIFT/BIC DEUTDEDB210
IBAN DE 69210700240144800000

Steuernummer 2029745781
USt.-IdNr. DE281295378

Eidesstattliche Versicherung

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Declaration on oath

I hereby declare, on oath, that I have written the present dissertation by my own and have not used any other than the acknowledged resources and aids.

Hamburg, 08.02.2019

_____

Dominik Danso

Table of Contents

List of Figures

List of Tables

Abstract

Polyethylene terephthalate (PET) is one of the most used polymers, with an annual production capacity of almost 30 million tons worldwide in 2015 (PlasticsEurope, 2017). It accumulates in the environment and pollutes the oceans as microplastic and nanoparticles. Due to its poor conversion in nature it is considered as an emerging threat for the environment. It is estimated that a common PET beverage bottle needs at least decades or even several hundred years until it is so far degraded that it is considered as microplastic (Barnes *et al.*, 2009; Ioakeimidis *et al.*, 2016; Li, W. C. *et al.*, 2016; Shaw & Day, 1994; Webb *et al.*, 2013).

Within this framework there is a high demand for identifying novel microorganisms and affiliated enzymes that can degrade this persistent polymer. Further developing very active enzymes for the modification of PET fibers used in textile and other industries is of great interest. Crucial for the success of finding such tools are a detailed understanding of the mechanisms and enzymes involved in biological breakdown or modification of PET are important. Therefore, the goals of this work were in particular (i) enhancing the known biodiversity of PET hydrolase containing microorganisms, since a clear overrepresentation of sequences from the phylum *Actinobacteria* was observed (Hu *et al.*, 2010; Oeser *et al.*, 2010; Thumarat *et al.*, 2015; Thumarat *et al.*, 2012; Wei, Oeser, Then*, et al.*, 2014; Wei, Oeser, & Zimmermann, 2014), (ii) mapping the results on a global scale therefore, giving the possibility to recognize a greater context and (iii) the production of selected PET hydrolases for subsequent biochemical characterizations. Enrichment and functional screening methods were applied in order to identify enzymes with activity on polymeric substrates. The combination of enrichment of environmental samples and functional screening of 78,897 metagenomic fosmid clones resulted in a single open reading frame (ORF) which proofed to encode a hydrolase capable of diethyl phthalate breakdown. In further tests the enzyme did not show any activity on PET. Therefore, an alternative screening method had to be applied. For this a sequence based deep data mining for putative PET hydrolase sequences was established. As initial tool for the sequence mining a HMM was constructed based on alignments of nine known PET hydrolase sequences. The model was applied on the NCBI and UniProt databases which resulted in 1,187 and 504 sequences respectively. These sequences were affiliated with the phyla *Actinobacteria, Proteobacteria, Bacteroidetes, Deinococcus-Thermus* and *Cyanobacteria.* The new sequences were filtered according to their scores (>180) and reintegrated into the alignment for the HMM. Besides the database derived sequences, metagenomic sequences were screened as well. A total of 133 Public available metagenomic datasets (>16 Gb of assembled DNA sequence) were obtained from IMG and

used to map potential hits to the samplesites geographical data. The phylogenetic distribution was in this case slightly larger containing sequences from the before mentioned phyla of the domain *Bacteria* with the addition of *Verrucomicrobia, Spirochaetes, Lentisphaerae, Firmicutes and Acidobacteria* as well as phyla from *Archaea* (*Euryarchaeota, Candidatus Thorarchaeota, Candidatus Lokiarchaeota*) and *Eukaryota* (*Pelagophyceae, Haptophyceae, Chromerida*). The metagenomic data of terrestrial and marine derived samples were compared and revealed differences in the phylogenetic composition. In marine samples *Bacteroidetes* was the dominant phyla whereas *Actinobacteria* dominated terrestrial habitats. Further in the terrestrial dataset the sequence data of a crude oil associated sample was outstanding in terms of putative PET hydrolase sequence abundance. This samplesite contained 135 potential PET hydrolase sequences accounting to a frequency of 1.5 hits/Gb suggesting enrichment of such enzymes and bacteria whereas all other sampled metagenomes showed an frequency of less than 1 hit/Gb. To verify the HMM on a functional level four database derived sequences were chosen due to their phylogenetic origin. The sequences were cloned and the expressed in E. *coli* resulting in sufficient amount of active enzyme (PET2, PET5, PET6 and PET12) for an initial characterization. These PET hydrolases showed activity on diethyl phthalate (DEP), polycaprolactone (PCL) as well as on PET nanoparticles. To our surprise, PET2 and PET6 also showed activity at elevated temperatures of up to 90 °C, along with some solvent stability, these enzymes are good candidates for industrial applications.

Altogether the applied self-constructed Hidden Markov model (HMM) allowed the identification of conserved motifs connected to PET activity of hydrolases. Using combined approaches the biodiversity of known PET degrading bacterial phyla was significantly enriched. Furthermore the PET hydrolase distribution was displayed in a global context. The obtained data indicate that PET hydrolases also occur in other phyla besides the already PET hydrolase associated *Actinobacteria*, whereby *Bacteroidetes* and *Actinobacteria* are the dominant phyla for marine and terrestrial metagenomes respectively, followed in both cases by *Proteobacteria*. The cloned and expressed enzymes showed activity on PET nanoparticles and other substrates, while being highly stable at elevated temperatures and under the influence of solvents, detergents and other inhibitors. The in this work developed strategies can now be used as the basis for further development of recycling-technologies and other methods intending to modify or partial degrade PET or other synthetic as well as natural polymers.

Zusammenfassung

Polyethylenterephthalat (PET) ist eines der meistverwendeten Polymere mit einer jährlichen Produktionskapazität von fast 30 Millionen Tonnen weltweit im Jahr 2015 (PlasticsEurope, 2017). Es reichert sich in der Umwelt an und verschmutzt die Ozeane als Mikro- und Nanopartikel. Aufgrund seiner geringen Abbauraten in der Natur gilt es als eine stetig wachsende Bedrohung für die Umwelt. Es wird geschätzt, dass eine handelsübliche PET-Getränkeflasche mindestens Jahrzehnte oder sogar mehrere hundert Jahre benötigt, bis sie so weit abgebaut ist, dass sie als Mikroplastik angesehen wird (Barnes *et al.*, 2009; Ioakeimidis *et al.*, 2016; Li, W. C. *et al.*, 2016; Shaw & Day, 1994; Webb *et al.*, 2013).

In diesem Rahmen besteht eine große Nachfrage nach der Identifizierung neuer Mikroorganismen und assoziierter Enzyme, mit der Fähigkeit dieses beständige Polymer abbauen zu können. Die Entwicklung sehr aktiver Enzyme für die Modifizierung von PET-Fasern, die in der Textilindustrie und anderen Industriezweigen verwendet werden, ist von großem Interesse. Entscheidend für den Erfolg solcher Instrumente ist ein detailliertes Verständnis der Mechanismen und Enzyme, die am biologischen Abbau oder der Modifikation von PET beteiligt sind. Die Ziele dieser Arbeit waren daher insbesondere (i) die Erhöhung der bekannten Biodiversität von PET-Hydrolase assoziierten Mikroorganismen, da in der Vergangenheit eine deutliche Überrepräsentation von Sequenzen aus dem Phylum *Actinobacteria* beobachtet werden konnte (Hu *et al.*, 2010; Oeser *et al.*, 2010; Thumarat *et al.*, 2015; Thumarat *et al.*, 2012; Wei, Oeser, Then*, et al.*, 2014; Wei, Oeser, & Zimmermann, 2014), (ii) die Ergebnisse in einem globalen Kontext zu stellen, wodurch es möglich ist größere Zusammenhänge zu erkennen und (iii) ausgewählte potentielle neue PET hydrolasen in *E. coli* zu produzieren, für anschließende biochemische Charakterisierungen. Um diese Ziele zu erreichen wurden zwei Lösungsansätze gewählt. Anreicherungen und funktionelle Screening-Methoden wurden angewendet um Enzyme mit Aktivität auf polymeren Substraten zu identifizieren. Die Kombination von Anreicherungen von Umweltproben und funktionellem Screening von 78.897 metagenomischen Fosmidklonen resultierte in einem einzelnen ORF, der nachweislich für eine Hydrolase kodiert mit der Fähigkeit Diethylphthalat abzubauen. In weiteren Tests zeigte das Enzym keine Aktivität auf PET. Daher wurde eine alternative Screening-Methode angewendet. Hierzu wurde ein sequenzbasiertes „Deep Data Mining“ für mögliche PET-Hydrolase Sequenzen etabliert. Als Grundlage für das Sequenz-Mining wurde ein HMM, basierend auf Alignments von neun bekannten PET-Hydrolase-Sequenzen, konstruiert. Das Modell wurde auf die NCBI- und UniProt-Datenbanken angewendet, was zu 1.187 bzw. 504

Sequenzen führte. Diese Sequenzen wurden den Phyla *Actinobacteria*, *Proteobacteria*, *Bacteroides*, *Deinococcus-Thermus* und *Cyanobacteria* zugeordnet. Die neuen Sequenzen wurden nach „Bit-scores“ gefiltert (>180) und in die bestehenden Alignments für das HMM reintegriert. Neben den aus Datenbanken bezogenen Sequenzen wurden auch metagenomische Sequenzen durchsucht. Insgesamt 133 öffentlich verfügbare metagenomische Datensätze (>16 Gb assemblierte DNA-Sequenz) wurden von IMG bezogen um mit Hilfe der geographischen Metadaten potentielle PET Hydrolyse Treffer in einer Weltkarte abzubilden. Die phylogenetische Verteilung war in diesem Fall etwas größer und enthielt Sequenzen aus den zuvor erwähnten Phyla der Domäne *Bacteria*. Zusätzlich waren *Verrucomicrobia*, *Spirochaetes*, *Lentisphaerae*, *Firmicutes* und *Acidobacteria* vertreten. Außerdem Phyla aus *Archaea* (*Euryarchaeota*, *Candidatus Thorarchaeota*, *Candidatus Lokiarchaeota*) und *Eukaryota* (*Pelagophyceae*, *Haptophyceae*, *Chromerida*). Die metagenomischen Daten von terrestrischen und marinen Proben wurden verglichen und zeigten Unterschiede in der phylogenetischen Zusammensetzung. In marinen Metagenomen war *Bacteroidetes* das dominante Phylum, während das Phylum *Actinobacteria* terrestrische Habitate dominierte. Auffällig war ein terrestrisches Metagenom einer Rohöl assoziierten Probe im Hinblick auf die Häufigkeit von PET-Hydrolase Sequenzen. Dieses Metagenom enthielt 135 potentielle PET-Hydrolase Sequenzen, was einer Häufigkeit von 1,5 Hits/Gb entspricht. Diese im Vergleich zu den anderen Metagenomen (<1 Hit/Gb) hohe Anzahl von putativen PET-Hydrolase Sequenzen könnte auf eine Anreicherung solcher Enzyme und Bakterien hindeuten. Um das HMM auf funktionaler Ebene zu verifizieren, wurden aufgrund ihrer phylogenetischen Herkunft vier Sequenzen ausgewählt. Die Sequenzen wurden kloniert und in *E. coli* exprimiert, was zu einer ausreichenden Menge an aktivem Enzym (PET2, PET5, PET6 und PET12) für eine anfängliche Charakterisierung führte. Diese PET-Hydrolasen zeigten Aktivität auf DEP, PCL sowie auf PET-Nanopartikeln. Zu unserer Überraschung zeigten PET2 und PET6 auch Aktivität bei erhöhten Temperaturen von bis zu 90 °C, zusammen mit einer gewissen Lösungsmittelstabilität machen diese Eigenschaften die beiden Enzyme zu guten Kandidaten für industrielle Anwendungen.

Insgesamt ermöglichte das verwendete selbst konstruierte Hidden-Markov-Modell (HMM) die Identifizierung von konservierten Motiven, die mit der Aktivität von PET-Hydrolasen verbunden sind. Mit verschiedenen experimentellen Methoden wurde die Biodiversität bekannter PET-abbauender Bakterienstämme signifikant erhöht. Darüber hinaus wurde die Verteilung von PET-Hydrolase in einem globalen Kontext dargestellt. Die gewonnenen Daten zeigen, dass PET-Hydrolasen auch in anderen Phyla neben den bereits mit PET-Hydrolase assoziierten

Actinobacteria vorkommen, wobei *Bacteroidetes* und *Actinobacteria* die dominanten Phyla für marine bzw. terrestrische Metagenome sind, gefolgt in beiden Fällen von *Proteobacteria*. Die heterolog produzierten Enzyme zeigten eine Aktivität auf PET-Nanopartikeln und anderen Substraten, während sie bei erhöhten Temperaturen und unter dem Einfluss von Lösungsmitteln, Detergenzien und anderen Inhibitoren hohe Stabilität zeigten. Die in dieser Arbeit entwickelten Strategien können nun als Grundlage für die Weiterentwicklung von Recycling-Technologien und anderen Methoden verwendet werden, die darauf abzielen, PET oder andere synthetische sowie natürliche Polymere zu modifizieren oder teilweise abzubauen.

1 Introduction

Polyethylene terephthalate (PET) was first discovered by John Rex Whinfield in 1941 (Rex & Tennant, 1949). The synthetic polymer is produced by a polycondensation of its monomers terephthalic acid (TPA) and ethylene glycol (EG) (Figure 1). It's a solid thermoplast which is nearly insoluble in water and has a melting temperature of 250 - 260 C° (Lim, 2017). Since its discovery many variants of this polymer were produced and found wide applications in a variety of products like food container, foil or synthetic yarn. The available variants differ mainly in their grade of crystallinity or in their melting temperature (Rieckmann & Völker, 2004). PET is just one of many plastics which are produced worldwide. The overall plastic production amounted to 335 million tons in 2016 and it is still growing (PlasticsEurope, 2017). PET resin is mainly used for the production of plastic bottles which are used for carbonated drinks. As well as the production volume of PET bottles, the recycling rate is steadily rising in some countries and, with around 93%, is currently the best in Germany (GVM, 2016). It's important to mention that energy production by simply burning the raw material is also considered as recycling (Chilton *et al.*, 2010). Another positive effect of the incineration of plastic waste is, in addition to the generation of energy in the form of heat, the avoidance of landfills that take up valuable space. Simultaneously, the incineration releases potential toxic substances into the atmosphere and eliminates the chance of recovering valuable plastic monomers (Sinha *et al.*, 2010; Zhang *et al.*, 2004). Nevertheless, the recycling rate is far below the production rate, which means that most of the plastics are not further used and can be considered as waste which is in most cases stored in above-mentioned landfills. In fact, PET continuously accumulates in the environment, especially in oceans and other aquatic habitats, but also on land (e.g. landfills). Due to a lack of information about each type of synthetic polymer, most statistical data is based on plastics in general. As an evaluation aid, it should be noted that the demand for PET in relation to the general demand for plastics in Europe is 7%.

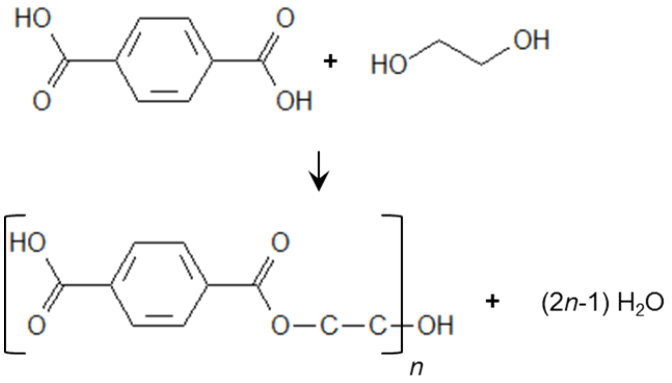

Figure 1: Polycondensation of terephthalic acid (TPA) and ethylene glycole (EG) for the production of polyethylene terephthalate (PET). Chemical structures were drawn with Chemograph Plus (version 6.50; DigiLab Software GmbH, Altenholz, Germany).

1.1 Plastic as an environmental pollutant

As mentioned above, plastic accumulates in the environment and especially in aquatic systems. It is found in all major ocean basins where it is carried by currents into gyres which are distributed over the entire planet. Although the density and size of the different plastic debris can differ they are often transported to similar locations where they can accumulate to a so called “garbage patch”. It is estimated that roughly 250,000 tons of plastic are currently floating at sea (Eriksen *et al.*, 2014). Similar studies show that most of the oceans plastics (>80%) originates from land (Jambeck *et al.*, 2015). Plastic debris in oceans is often eaten by fish, other sea creatures or birds. In extreme cases the whole gastric system and especially the stomach is completely filled with plastic, which causes starvation of the organism due to incapability of food uptake (Derraik, 2002). Entanglements within larger plastic products like bags or six-pack yokes must be mentioned here as well. In aquatic as well as in terrestrial systems water soluble components of the different plastic types can be taken up by organisms and have negative effect on their health. In case of human health it is reported that phthalates, which are often associated with PET, can act as endocrine disruptors and therefore should be considered as harmful (Manikkam *et al.*, 2013; Sax, 2010).

1.2 Degradation processes of plastic

Petroleum based plastics are in general highly stable and durable, hence it is commonly said that plastics do not degrade in nature to a large degree (Webb *et al.*, 2013) but still, once released into the environment plastic waste must somehow be degraded. Although the degradation processes described are all very slow and it is shown that for example a PET bottle remains up to 48 years in the ocean until it is decomposed, degradation definitely takes place (Muller *et al.*, 2001). In some cases mechanical treatment caused by waves, wind or the movement of solids across the plastic surface simply break down the debris into much smaller pieces which cannot be easily detected and therefore be falsely considered as completely degraded. The resulting microplastic show a much higher surface area compared to the bigger fragments which may contributes to further degradation by abiotic or biotic factors. In case of PET, the mechanisms of abiotic depletion are described but not understood in their full complexity. The so called “weathering” or “photodegradation” is considered to be the main force for initial depletion of PET. Due to the aromatic ring of the terephthalic acid, PET has a strong chromophore which makes it vulnerable to UV light. It was observed that UV light causes changes in properties of PET which can be observed by a color change from clear to slightly yellow. Other important factors are the ambient temperature and the temperature of the plastic

piece itself. Another factor is the humidity. In case of PET, which can be hydrolyzed, the surface moisture must also be considered in conjunction to the UV radiation (Fagerburg & Clauberg, 2004).

1.2.1 Depletion of PET by enzymatic hydrolysis

The main biotic factors for PET or plastics degradation in general are bacteria and other microorganisms. It was shown that several species of the genus *Thermobifida* and *Thermomonospora* expresses enzymes which can partially hydrolyze PET (Wei, Oeser, & Zimmermann, 2014). Later, such enzymes where described in other *Actinobacteria* like *Saccharomonospora viridis* (Kawai *et al.*, 2014), but also in bacteria of the genus *Bacillus* (Ribitsch *et al.*, 2011). Such enzymes are not restricted to bacteria as fungi like *Fusarium oxysporum* or *Humicola insolens* were affiliated with PET degradation as well (Carniel *et al.*, 2017; Vertommen *et al.*, 2005). The depletion is also highly influenced by the degree of crystallinity. Since most petroleum based plastics are semi-crystalline polymers, they contain amorphous, and thus easily accessible regions, as well as highly crystalline regions which are more or less resistant to enzymatic attacks (Webb *et al.*, 2013). The whole degradation mechanism is not fully resolved. Due to the chemical and physical properties of PET, it is assumed that the microorganisms must secrete the hydrolases into their environment. PET has a large molecular mass and therefore cannot be transported into the cells (Table 1).

Table 1: Physical properties of PET. Values with ranges indicate properties, which vary depending on crystallinity and degree of polymerization (Awaja & Pavel, 2005).

Property	Value
Average molecular weight	30,000–80,000 g mol^{-1}
Density	1.41 g cm^{-3}
Melting temperature	255–265 °C
Glass transition temperature	69–115 °C
Young's modulus	1700 MPa
Water absorption (24 h)	0.5%

Likewise, the secreted proteins are too large to penetrate deep into the plastics and therefore can only act on the surface layer, making the hydrolysis of plastics by extracellular enzymes a surface-erosion process (Figure 2). The resulting smaller intermediates are water soluble and can migrate into the bacterial cells where they are further metabolized (Mueller, 2006). In 2016, a full hydrolysis of PET was described by Yoshida *et al.* It was experimentally shown that two enzymes from the bacterium *Ideonella sakaiensis* are capable of fully degrading PET by a

twostep process. *I. sakaiensis* uses a PET hydrolase for the initial hydrolysis and a so called MHETase for the further breakdown of mono-(2-hydroxyethyl) terephthalate (MHET) which is considered as the major intermediate (Yoshida *et al.*, 2016). Nevertheless, PET degradation seems to be restricted to a few phyla given that more than 60 phyla in the bacterial domain are known so far (Lasken & McLean, 2014), but only three are described for PET hydrolyzing enzymes, and in case of the phylum *Proteobacteria* with just a single representative. It is assumed that besides the mentioned examples a multitude of highly useful potential PET hydrolases is still unidentified.

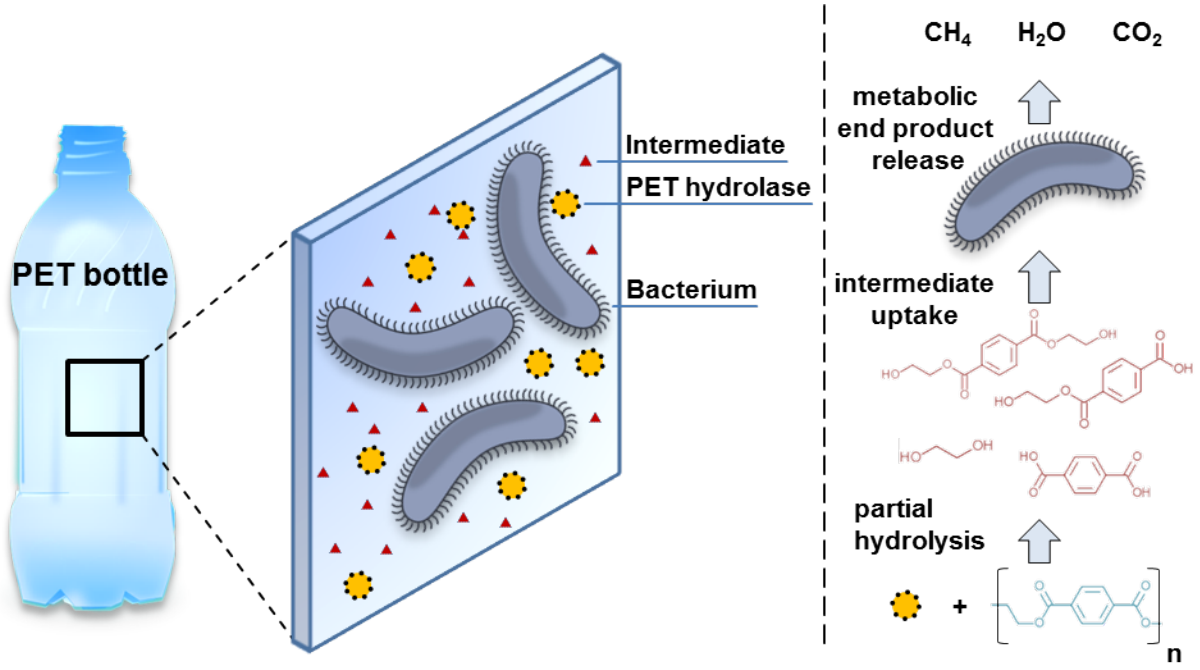

Figure 2: Schematic overview of the partial PET degradation by bacteria. Original figure of (Bornscheuer, 2016). PET hydrolases are secreted into the environment and act on their substrate. The resulting hydrolysis products are taken up by the bacteria and are further metabolized.

1.2.2 Polyethylene terephthalate (PET) degrading α/β hydrolases

Carboxylic ester hydrolases which act on ester bonds (EC 3.1.1.) are frequently used enzymes in several biotechnological processes. Triacylglycerol lipases (EC 3.1.1.3) for example are often used in feed and food production but also as additives in washing powder or as catalyst for pharmaceutic and fine chemical production (Hasan *et al.*, 2006; Houde *et al.*, 2004). Other members of the enzyme class of hydrolases like carboxylesterases (EC 3.1.1.1.), cutinases (EC 3.1.1.74) or phospholipases (EC 3.1.1.4) are as well very popular for more or less similar processes (Borrelli & Trono, 2015; Panda & Gowrishankar, 2005; Pio & Macedo, 2009). The

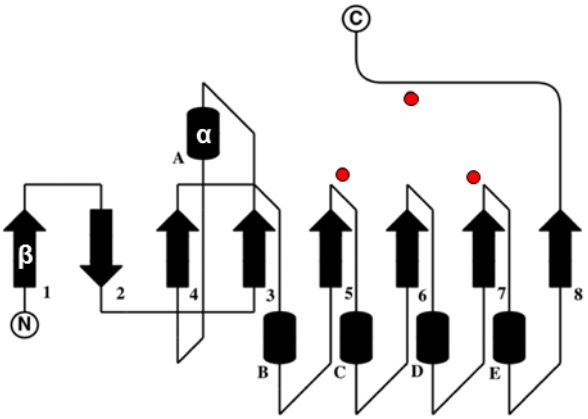

Figure 3: The canonical structure of α/β-hydrolases. Red circles indicate catalytic residues (Ser, His, Asp) (original figure by David and Paul (2009)). As known so far the core of all α/β-hydrolases shares this typical fold consisting of 8 β-sheets connected by α-helices.

hydrolases of this group all share the typical α/β-fold and a catalytic triad which is usually comprised of the amino acids serine, histidine and aspartic acid. The eight β-strands are arranged in a typical order and interconnected by six α-helices (Figure 3) (David & Paul, 2009). PET hydrolyzing enzymes have the EC number 3.1.1.101 and are representing a novel group of hydrolases with currently just one representative entry, the PET hydrolase of *I. sakaiensis* (as of February 2018). The best described enzymes affiliated with PET degradation are the above mentioned examples from *Actinobacteria* like *Thermobifida* or *Thermomonospora*. Crystal structural analysis of those proteins showed that they have a typical α/β-hydrolase fold consisting of a central twisted β-sheet that is flanked by α-helices on both sides. Published data indicate that the proteins are monomers in solution and are probably acting as such on their substrate (Kitadokoro *et al.*, 2012). In contrast to above mentioned carboxylesterases from different *Bacillus* and *Thermobifida* species (Billig *et al.*, 2010; Oeser *et al.*, 2010; Ribitsch *et al.*, 2011) real lipases display only poor or no activity at all against PET and other petroleum-based plastics. The typical lid structure of lipases which promotes activity on long chain substrates like lipids (chain length $>C_{10}$) (Khan *et al.*, 2017) interferes with the bulky PET as substrate and therefore prevents activity, as long as no compound is added which initiates interfacial activation of the protein (Eberl *et al.*, 2009; Guebitz & Cavaco-Paulo, 2008). The most active PET hydrolase was found in *Thermomyces* (*Humicola*) *insolens.* The protein was able to almost completely hydrolyze PET film indicating an efficient degradation of even crystalline parts at elevated temperatures of 70 °C, over a period of 96 hours (Ronkvist *et al.*, 2009). Temperatures close to the glass transition temperature (t_g) of PET, which is around 65 °C are always beneficial during hydrolysis. Amorphous parts within the substrate become flexible and therefore more accessible for the protein (Alves *et al.*, 2002; Parikh *et al.*, 1992). Taking the structural changes of the substrate into account, thermostable enzymes are clearly beneficial in future PET degrading processes. Positive effects on enzyme hydrolysis were also observed by addition of bivalent metal ions like Ca^{2+} or Mg^{2+}. They seem to promote thermostability as shown for actinobacterial enzymes (Kawai *et al.*, 2014; Miyakawa *et al.*, 2015; Sulaiman *et al.*, 2014; Thumarat *et al.*, 2012). A similar effect was observed for phosphate anions (Jensen *et al.*, 1995; Park *et al.*, 2001). It was possible to show that

kosmotropic phosphate anions stabilized the metagenome derived “leaf compost cutinase” (LCC) and the cutinase “TfCut2” from *T. fusca* (Schmidt *et al.*, 2016). In contrary to the beneficial effects described above, the water soluble hydrolysis products MHET and BHET had negative effects on some PET hydrolases. At concentrations between 0.5 and 2 mM MHET and BHET are competitive inhibitors with similar binding constants (Barth *et al.*, 2015). Hence, an elimination of product inhibition and maximization of efficiency, by a process consisting of a thermostable PET hydrolase and parallel extraction of released intermediates within a buffer system supplying phosphor and/or bivalent metal ions would be desirable. A further obstacle is the hydrophobic nature of PET. It represents a barrier which hampers the effective adsorption of enzymes to the surface (Atthoff & Hilborn, 2007). Initial adsorption of native PET hydrolase occurs at hydrophobic areas near the catalytic site which is rather inefficient (Acero *et al.*, 2011). Unlike proteins hydrolyzing natural polymers such as cellulose (Atthoff & Hilborn, 2007) proteins associated with PET hydrolysis do not show additional binding domains (Chen, S. *et al.*, 2013; Kitadokoro *et al.*, 2012).To bypass these limitations, attempts of fusing binding domains or fungal hydrophobins to a PET hydrolase were made and showed a clear positive effect (Espino-Rammer *et al.*, 2013; Ribitsch *et al.*, 2015; Ribitsch *et al.*, 2013). Multi enzyme reaction systems are also a possibility to increase PET hydrolysis and especially terephthalic acid yields. It was shown that the lipase CalB from *Candida antarctica* had a synergistic effect on the fungal PET hydrolase HiC. Experimental data suggest that CalB efficiently converts resulting MHET to terephthalic acid, protecting HiC from product inhibition (Carniel *et al.*, 2017). Similar results were obtained with the carboxylesterase TfCa from *T. fusca* KW3 (Barth *et al.*, 2016).

1.3 Metagenome screening for novel biocatalysts

New biocatalysts are used nowadays for a variety of applications and the demand is still increasing due to the cost reducing properties, eco-friendliness and simplicity of enzyme driven reactions (Li, S. *et al.*, 2012; Singh *et al.*, 2016). Bringing an unknown organism with a desired potential biocatalyst to culture can be very challenging. Cultivation success in defined medium, under laboratory conditions and adjusted parameters is highly dependent on a multitude of unknown parameters and often fails completely. Therefore, information on potential new biocatalysts is extracted by metagenome screening methods today. A multitude of screening strategies were established, providing tools for the identification of enzymes with improved activity or other beneficial properties (Steele *et al.*, 2009). Depending on the chemical properties and function of the enzyme, individual functional screening assays were developed (Streit, 2010 #240). Hence, there is no possibility to gain information about all potential biocatalysts in a habitat at once. Additionally, the functional screening methods are almost exclusively based on

heterologous expression in suitable hosts. The most prominent expression system is *E. coli* in combination with specially designed plasmids (Charles *et al.*, 2017). The need of expression for subsequent functional testing and the identification of the responsible protein is the major limitation of functional screening. As a result, metagenome sequencing has been established more and more, making it possible to completely relinquish on an *in vivo* expression for the initial identification. A similar trend can be observed for cultivation approaches. Phylogenetic marker genes can be easily identified in metagenome sequencing data. Alternatively, amplicon sequencing of target genes (e.g. 16S-rRNA genes) can be performed. Large sequencing projects and a multitude of published data proved the usefulness of metagenome sequencing. Whereby, habitat affiliated analysis of metabolic potential and taxonomic composition plays a major role (Hug *et al.*, 2016; Methe *et al.*, 2012; Pettersson *et al.*, 2009; Thompson *et al.*, 2017; Venter *et al.*, 2004). Considering PET hydrolyzing enzymes, traditional microbiological cultivation and functional metagenome screening approaches delivered a rather low taxonomic diversity (Barth *et al.*, 2015; Billig *et al.*, 2010; Danso *et al.*, 2018; Kale *et al.*, 2015; Kawai *et al.*, 2014; R.A. & L., 2017; Ribitsch *et al.*, 2012; Ribitsch *et al.*, 2011; Roth *et al.*, 2014; Sulaiman *et al.*, 2014; Wei, Oeser, Then*, et al.*, 2014; Yoshida *et al.*, 2016). Sequence based approaches may circumvent this bottleneck and help understanding the diversity of PET hydrolyzing organisms and enzymes.

1.4 Intention of this work

The main goal of this study was to obtain an overview and a better understanding of the bacterial PET degradation potential. Therefore a Hidden Markov model should be constructed and be used in deep metagenome mining of publicly available datasets. It was expected that the obtained data contributes new knowledge about the taxonomic diversity of PET degrading bacterial communities. Additionally, the overall geographic distribution and abundance of potential degraders were in focus and thus the used datasets had to contain a diversity of marine and terrestrial habitats. Potential PET degrading enzymes found during metagenome analysis were supposed to be characterized and functionally tested. Attention was payed onto enzymes with high degrading velocity at elevated temperature and stability under denaturing factors (e.g. strong organic solvents or detergents), which may have high potential in biotechnological processes (Figure 4).

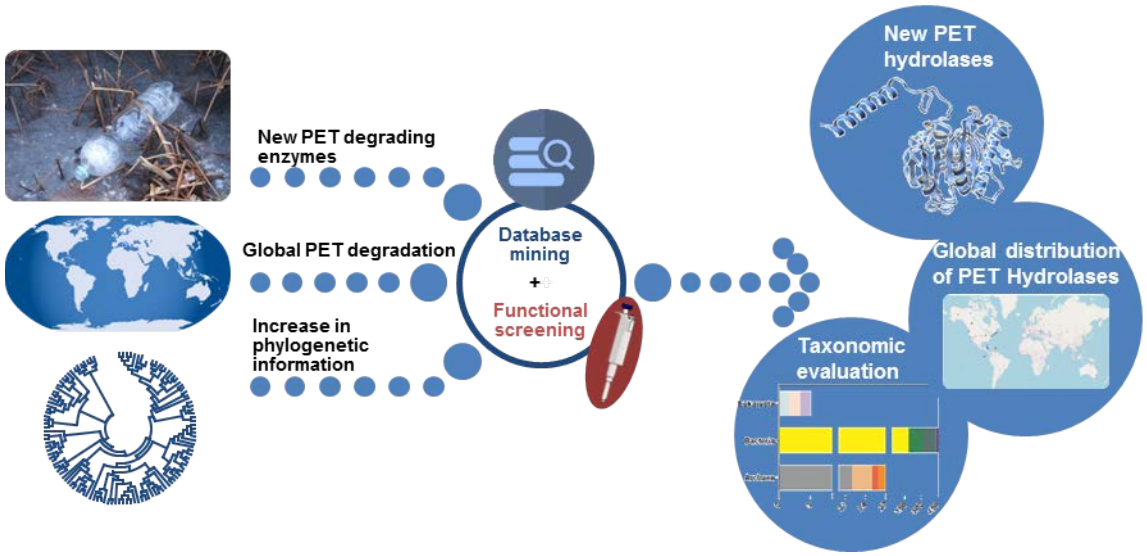

Figure 4: Schematic workflow of this study. Obtaining new PET hydrolases and demonstrating their importance in a global context with respect to their taxonomical affiliation was done with the help of sequence-based and function-based metagenomics. Pictograms were either self-prepared or available under creative commons license (CC0).

By using a sequence-based approach for the identification of novel enzymes from metagenome data and the subsequent functional analysis and activity verification, this work contributes to the general understanding of polyethylene terephthalate degrading enzymes.

The comprehensive analyses of bacteria affiliated with PET hydrolysis, based on metagenomic data, directly resulted in the following publication (see Appendix):

Danso *et. al*. 2018. Appl Environ Microbiol doi:10.1128/aem.02773-17.

New insights into the function and global distribution of polyethylene terephthalate (PET) degrading bacteria and enzymes in marine and terrestrial metagenomes.

2 Material and Methods

2.1 Bacterial strains, plasmids and respective cultivation conditions

2.1.1 Bacterial strains

Name, characteristics and the source of used strains within this study are shown in Table 2.

Table 2: Bacterial strains used in this study.

Strain	Characteristics[1]	Reference/ Source
E. coli T7Shuffle	**Expression strain for Protein purification:** F´ *lac*, *pro*, *lacI^q* / *Δ(ara-leu)7697 araD139 fhuA2 lacZ::T7 gene1 Δ(phoA)PvuII phoR ahpC* galE (or U) galK λatt*::pNEB3-r1-*cDsbC* (Spec^R, *lacI^q*) *ΔtrxB rpsL150*(Str^R) *Δgor Δ(malF)3*	NEB (Frankfurt am Main, Germany)
E. coli DH5α	**Cloning and subcloning strain:** F$^-$ *endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG purB20* φ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169, hsdR17($r_K^-m_K^+$), λ$^-$	Invitrogen (Karlsruhe, Germany;(Hanahan, 1983)
Deinococcus maricopensis	type strain, DSM-21211	DSMZ (Braunschweig, Germany)
Vibrio gazogenes	type strain, DSM-21264	DSMZ (Braunschweig, Germany)
Polyangium brachysporum	type strain, DSM-7029	DSMZ (Braunschweig, Germany)

[1] Abbreviations describing geno- and phenotypes were made according to Bachmann (1983).

2.1.2 Cultivating *D. maricopensis*, P. *brachysporum* and *V. gazogenes*

All ordered strains were cultivated as recommended from DSMZ. *D. maricopensis,* P. *brachysporum* and *V. gazogenes* were cultivated in DSMZ Medium 736, DSMZ Medium 67 and DSMZ Medium 514 respectively at 28 °C and 25 °C in case of P. *brachysporum*.

2.1.3 Cultivating *E. coli* strains

E. coli cultures were grown on lysogeny broth (LB) medium (Bertani, 1951) at 37 °C over night. Required supplements were added after autoclaving and cooling of the medium to at least 60 °C. For LB plates 1.5% agar was added to the medium prior to autoclaving. The different Stock solutions, concentrations and final concentrations in the media are summarized in Table 3.

Table 3: Antibiotics and supplements used in this study.

Substance	Stock solution	Final concentration	Solvent	Treatment
Antibiotics				
Ampicillin	100 mg/ml	100 µg/ml	70% EtOH	-
Kanamycin	25 mg/ml	25 µg/ml	H_2O	-
Other suplements				
Arabinose	0.1	0.0001	H_2O	autoclaved
IPTG	100 mg/ml	100 µg/ml	H_2O	filtered sterile

2.1.4 Enrichment of environmental samples containing PET degrading organisms

Enrichment cultures were prepared in 300 ml Erlenmeyer flasks containing 100 ml mineral medium (M9) (Harwood & Cutting, 1990). Cultures were incubated at 22 °C or 28 °C under continuous shaking. PET yarn, cut into 1 cm long pieces, was used as sole carbon source. A total of ten such pieces were used for each enrichment culture. DNA was isolated from fractions of the culture with the peqGOLD Bacterial DNA Mini Kit from VWR (Darmstadt, Germany) and the 16S-rRNA gene was amplified using standard 16S primers (Table 4). The amplified fragments were Sanger sequenced afterwards at Eurofins (Elsberg, Germany).

2.2 Vectors, primers and Constructs used in this work

Primers, vectors and constructs used in this study are shown in Table 4, Table 5 and Table 6.

Table 4: Primers used in this study.

Primer	Sequence (5' → 3')	Length (bp)	Tm (°C)	Source
T7 promotor	TAATACGACTCACTATAGGG	20	53.2	Eurofins (Elsberg, Germany)
T7 terminator	CTAGTTATTGCTCAGCGGT	19	54.5	Eurofins (Elsberg, Germany)
616v (16S)	AGAGTTTGATYMTGGCTC	18	60	(Juretschko *et al.*, 1998)
1492r (16S)	GGYTACCTTGTTACGACTT	19	60	(Frank *et al.*, 2008)
PET5_for	CGCCGCCATATGAATAAATC TATTCTAAAAAAACTCTC	38	68	this work
PET5_rev	CGATTCGGCGGCCGCGTAA TTACATGTGTCACGG	34	77	this work
PET6_for	CGTAGTCATATGGTACCGT GTTCGGACTG	29	69	this work
PET6_rev	CAGCGGCCGCCTAATAGTA ACTACAGTTGTCTCG	34	73	this work
PET12_for	CGCCATATGCAGACCAACC CCTACCAGCGAGGCCC	35	80	this work
PET12_rev	CTTGCGGCCGCTCAGTACG	39	84	this work

	GGCAGCTCTCGCGGTACTCC			
PET14_for	GGCCATATGCAGTGTGTTGGCGTGAAT	27	71	this work
PET14_rev	GGCGTCGACTCACCCACCATTCGGCAGATTG	31	76	this work

Table 5: Vectors used in this study.

Vector	properties	Reference/source
pET21a(+)	**Expression vector**: *lacI*, Amp^R, T7-*lac*-promoter, C-terminal His_6-tag coding sequence	Novagen/Merck (Darmstadt, Germany)
pET28a(+)	**Expression vector**: *lacI*, Amp^R, T7-*lac*-promoter, C- terminal and N-terminal His_6-tag coding sequence	Novagen/Merck (Darmstadt, Germany)
pEX-A2	**Cloning vector**: Amp^R, P_{lac}, *lacZ*, pUC *ori*	Eurofins (Ebersberg, Germany)
pBluescript II SK(+)	**Cloning vector**: Amp^R, P_{lac}, *lacZ*, pUC *ori*	Agilent (Santa Clara, USA)

Table 6: Constructs created in this study.

Construct	Vector	Insert (bp)	Characteristics
pET21a(+)::PET2	pET21a(+)	925	His-6-tag containing PET hydrolase
pET21a(+)::PET5	pET21a(+)	930	His-6-tag containing PET hydrolase
pET28a(+)::PET6	pET28a(+)	825	His-6-tag containing PET hydrolase
pET28a(+)::PET12	pET28a(+)	813	His-6-tag containing PET hydrolase
pET28a(+)::PET14	pET28a(+)	813	His-6-tag containing PET hydrolase
pEX-A2::PET2	pEX-A2	925	Synthesized gene (PET2) in a standard cloning vector from Eurofins (Elsberg, Germany)

2.3 Transmission electron microscopy (TEM) and Scanning electron microscopy (SEM)

Slices were prepared using the microtome Reichert-Jung Ultracut E. Fixation was done with 2% glutaraldehyde in 75 mM cacodylate buffer (pH 7.0). Afterwards, samples were supplied with 2% agar in 75 mM cacodylate buffer (pH 7.0) and fixed with 1% OsO4 in 50 mM cacodylate buffer (pH 7.0). Washing was performed with 75 mM cacodylate buffer (pH 7.0), residual water was removed with acetone and the samples were treated with Spurr resin (Polysciences, Warrington, PA, USA). TEM pictures were taken on a LEO 906 E using a Gatan 794 camera and the software Digital micrograph (Gatan GmbH, Munich, Germany). SEM images were taken with a LEO 1525. Samples were incubated in 1% paraformaldehyde and 0.25% glutaraldehyde. The

samples were dehydrated by ascending alcohol series, and dried at the critical point with a Balzers CPD 030 Critical Point Dryer (Bal-Tec, Schalksmühle, Germany). Finally samples were coated with gold using a SCD050 sputter coater (Bal-Tec).

2.4 Function based screening of metagenome libraries

In order to identify PET hydrolases in metagenome libraries functional screenings on indicator plates were performed. Clones were transferred from 96 Well plates on indicator plates (0) supplemented with antibiotics and arabinose. The plates were incubated over night at 37 °C. Halo producing clones were further tested. Existing metagenome libraries from the work group of Prof. Dr. Streit were used in this study (Table 7). Additional a subset of previous, functional tested hydrolase containing fosmid clones (Esterase-lipase-box) were tested on indicator plates.

Table 7: Metagenome libraries used in this study.

Metagenome library	No. of clones	Origin
Elephant faeces	20,064	Based on Ilmberger *et al.* (2014)
Teufelsbrück, river Elbe sediment	20,256	Prof. Streit lab
Glückstadt, river Elbe sediment	23,520	Prof. Streit lab
Algae photobioreactor biofilm	14,976	Krohn-Molt *et al.* (2013)
Esterase-lipase-box (in house collection of metagenome-derived enzymes)	81	Prof. Streit lab
Total = 78,897		

2.4.1 Preparation of indicator plates

Indicator plates were prepared with LB medium supplemented with antibiotics and either IPTG (100µg/µl) or arabinose (1%) for induction.

2.4.1.1 PCL containing indicator plates

A 1.5% solution of polycaprolactone (PCL) in acetone was prepared and incubated at 60 °C until the polymer was completely dissolved. The complete solution was poured into 500 ml sterile LB

medium under constant stirring. After complete evaporation of acetone under a fume hood other supplements were added as needed (Wei, Oeser, Barth*, et al.*, 2014).

2.4.1.2 PET-nanoparticle containing indicator plates

Polyethylene terephthalate (PET) nanoparticles were prepared as previously published (Welzel *et al.*, 2002). Particles were used without size estimation. PET nanoparticle plates were provided by Evocatal GmbH (Monheim am Rhein, Germany).

2.4.1.3 DEP containing indicator plates

For diethyl phthalate (DEP) containing LB plates 0.75 g gummi arabicum was partially solved in 7.5 ml H_2O. A total of 5 ml DEP was added and the two phases containing mixture was sonicated for 1 minute at maximum amplitude with 0.5 sec duty cycles. The resulting homogenized mixture was transferred into preheated medium and autoclaved. After addition of supplements the DEP containing LB medium was poured into petri dishes.

2.4.2 Subcloning of potential PET hydrolase ORF's from fosmid clones

In order to identify the active protein coding sequenced within a bigger metagenomic insert, the fosmid DNA was isolated using the Geneaid Presto™ Mini Plasmid Kit (Geneaid, New Taipei City, Taiwan) and digested with different commercial available restriction enzymes (Thermo Fisher Scientific, Waltham, USA) to obtain smaller DNA fragments. The fragments were purified using the Geneaid Gel/PCR DNA Fragments Kit. DNA fragments were ligated into the equally digested cloning vector pBluescript SK II(+) (Table 5). The construct was then transferred into chemical competent *E. coli* DH5α cells (Table 2) via heat shock method (Sambrook & Russell, 2001). Transformed cells were spread out on indicator plates containing IPTG (0) and incubated at 37 °C for at least 24 hours, ensuring enough time for halo formation. Positive clones were further cultivated and plasmids isolated. Construct size was checked with restriction enzymes on an agarose gel (0.8% Tris-Acetat-EDTA agarose gel). The smallest, still active constructs were chosen for further subcloning. The procedure was repeated until fragment sizes of < 2,000 bp were reached. In case of longer sequences a primer walk approach was used to obtain the active protein coding sequence. After Sanger sequencing at Eurofins (Elsberg, Germany) primer were constructed in order to obtain minimum size inserts with appropriate restriction sites for directional cloning into an expression vector (2.22.1).

2.5 Sequence based screening of metagenome libraries and databases

In order to identify novel PET hydrolases within big sequence datasets a Hidden Markov model (HMM) was constructed on the basis of known and functionally tested examples. The obtained sequences of initial search rounds were added to the HMM to refine the algorithm. A subset of sequences was chosen for synthesis or amplification in order to functionally verify their activity. Later, only the positive tested PET hydrolases contributed to the HMM construction. The general procedure is shown in Figure 5.

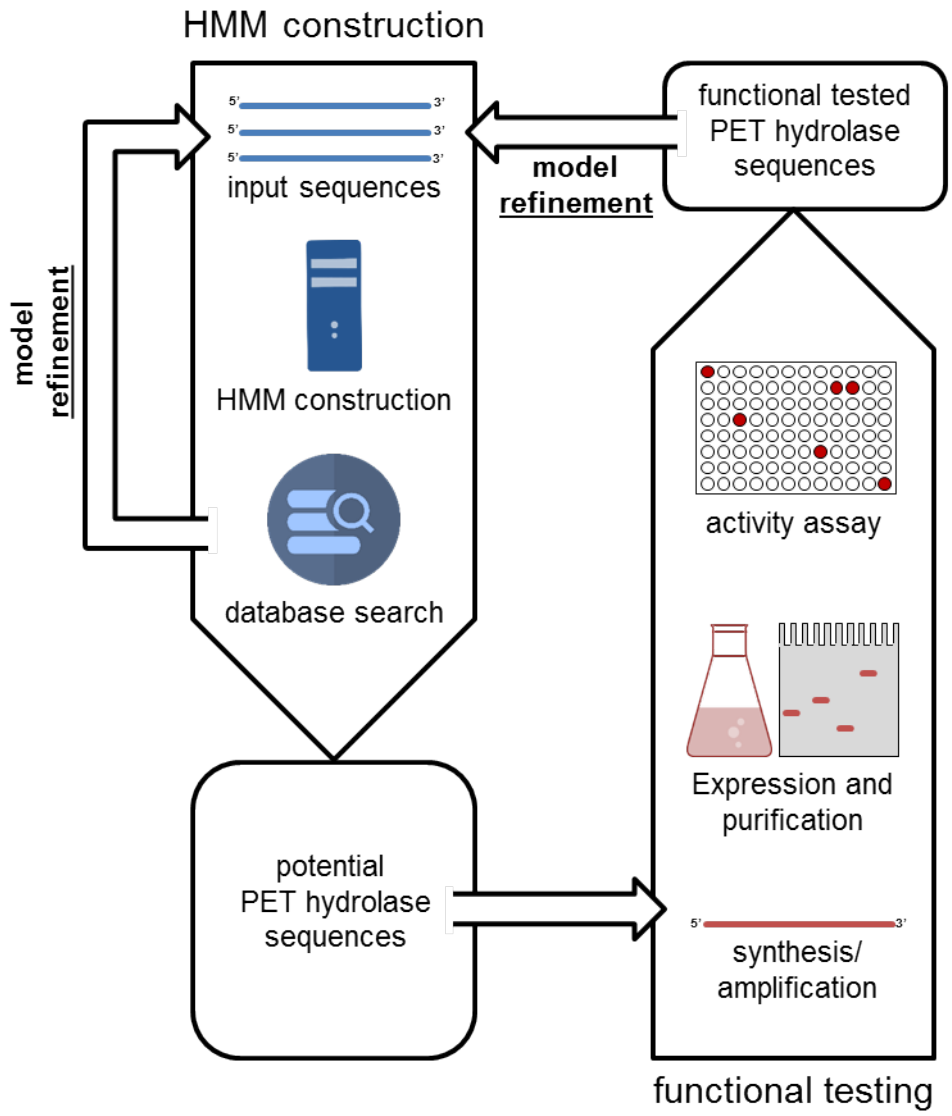

Figure 5: General procedure of constructing/refining the HMM and the subsequent search for PET hydrolases. The initial search rounds relied on sequence information as input. With subsequent functional testing of putative PET hydrolase sequences the model was refined and optimized. Pictograms were either self-prepared or available under creative commons license (CC0).

2.5.1 Construction of a Hidden Markov model

Construction of the HMM was done with the HMMer package for Unix based operating systems (Finn *et al.*, 2011). After an alignment in accurate mode with T-Coffee (Notredame *et al.*, 2000), containing nine described and functional tested PET hydrolases (Table 8) the program "hmmbuild" uses the alignment as input to generate the HMM. Parameters were left on default. The HMM was visualized by the HMM logo generator (Wheeler *et al.*, 2014).

Table 8: Protein sequences used for the initial construction of a PET hydrolase HMM.

Number	UniProt entry	Organism	Source
1	W0TJ64	*Saccharomonospora viridis*	(Kawai *et al.*, 2014)
2	E9LVI0	*Thermobifida fusca (Thermomonospora fusca)*	(Dresler *et al.*, 2006)
3	E5BBQ3	*Thermobifida fusca (Thermomonospora fusca)*	(Chen, S. *et al.*, 2008)
4	D1A9G5	*Thermomonospora curvata*	(Wei, Oeser, Then*, et al.*, 2014)
5	E9LVH7	*Thermobifida alba*	(Hu *et al.*, 2010)
6	H6WX58	*Thermobifida halotolerans*	(Ribitsch *et al.*, 2012)
7	E9LVH9	*Thermobifida celluloysilityca*	(Acero *et al.*, 2011)
8	A0A0K8P6T7	*Ideonella sakaiensis*	(Yoshida *et al.*, 2016)
9	G9BY57	uncultured bacterium	(Sulaiman *et al.*, 2012)

2.6 PET hydrolase expression in *E. coli*

The PET hydrolases PET2, PET5, PET6 and PET12 (sequences are listed under Appendix) were cloned into expression vectors (2.2). Chemical competent *E. coli* cells (2.1.1) were transformed via heat shock (Sambrook & Russell, 2001) with the obtained constructs and used for heterologous expression of the proteins.

2.6.1 Cloning of PET2, PET5, PET6 and PET12 in *E.coli*

In Case of PET5, 6 and 12 genomic DNA was isolated from the source organisms with the peqGOLD Bacterial DNA Mini Kit by VWR International GmbH (Darmstadt, Germany) and used to amplify the coding genes with specific primers (Table 4) and a modified protocol for touchdown PCR (Don *et al.*, 1991) (Table 9). PET2 was obtained from a metagenomic dataset, therefore the nucleotide sequence was synthesized and optimized for expression in *E. coli*

T7SHuffle cells at Eurofins (Ebersberg, Germany). The obtained fragments were cut with restriction enzymes (see Table 4) and the overlapping ends were used to ligate the inserts into pET21a(+) or pET28a(+) (Sambrook & Russell, 2001) (Lund *et al.*, 1996). Constructs were checked after restriction with *Nde*I and *Sal*I, size estimation was done on an agarose gel (0.8% TAE agarose gel, 120 V, 25 min). Additional the constructs were sequenced at Eurofins (Ebersberg, Germany) using Sanger sequencing and compared to the original sequences in order to check for unwanted mutations.

Table 9: Touch-Down PCR conditions for the amplification of PET hydrolase genes.

Initial denaturing	**95 °C**	**5 min**	
Denaturing	95 °C	0.5 min	16 x with decreasing Tann (-0.5°C/cycle) 14 x at calculated Tann (Tann = Tm-5°C)
Annealing	T_m + 3°C	0.25 min	
Elongation	72 °C	1 min/kb	
Final elongation	72 °C	5 min	

2.6.2 Expression and purification

Expression of PET hydrolases was done in auto induction medium ZYM-5052 without trace metals (1% tryptone/peptone, 0.5% yeast extract, 25 mM Na_2HPO_4, 25 mM KH_2PO_4, 50 mM NH_4Cl, 5 mM Na_2SO_4, 2 mM $MgSO_4$, 0.5% glycerol, 0.05% glucose, 0.2% lactose) (Studier, 2005). The expression culture was inoculated with 1% of an overnight culture and grown at 37 °C under constant aeration until OD_{600} = 1. Protein production took place at 17 °C under constant aeration and for at least 16 h. Purification was done with the Protino Ni-NTA agarose from Machery-Nagel (Düren, Germany) and the corresponding protocol for purification under native conditions. The expression culture was centrifuged (5000 g, 15 min) in order to obtain a cell pellet. The pellet was washed and ether resuspended in an appropriate amount of NPI-10 buffer for further processing, or frozen at -20 °C for long term storage. Cells were disrupted with a french pressure cell and two additional sonication rounds (Amplitude height was 60% and duty cycle set at 0.5 sec.) in order to degrade high molecular weight DNA. Protein containing elution fractions were checked by SDS-PAGE (Laemmli, 1970; Sambrook & Russell, 2001), western blot (Towbin *et al.*, 1979) and activity assay using *p*NP substrates (2.7.1) and indicator plates (2.4). Protein fractions were pooled and the elution buffer was exchanged against sodium

phosphate buffer (pH 7.0) using a centrifugal filter unit from Satorius (Göttingen, Germany).The centrifugation steps were carried out at 8 °C and at manufacturer recommended speed in case of buffer exchange.

2.7 Characterization of PET2, PET5, PET6 and PET12

The biochemical properties of PET hydrolases PET2, PET5, PET6 and PET12 were tested. Tests were executed with purified protein (2.6.2). The data was normalized to percentage of relative activity. In assays for reduction or promotion of activity an untreated sample was used as standard (100% relative activity).

2.7.1 *para*-nitrophenol (*p*NP) ester assay

The detection of hydrolytic active enzymes was accomplished by usage of *p*NP esters (Table 10). The colorless substrate releases chromogenic *para*-nitrophenol when the ester bond is hydrolyzed, which can then be detected photometric at a wave length of 405 nm. Tests were performed in 96-well plates with a maximum volume of 200 µl and measured in the Synergy HT plate reader from BioTek (Winooski, USA). If not indicated else, a total amount of 0.1 µg to 1 µg purified enzyme was used with 1 mM *p*NP-octanoate and sodium phosphate buffer (pH 8.0). The reaction mixture was incubated for 15 min at a defined temperature, pH and immediately measured afterwards.

Table 10: *p*NP substrates used in this work.

Chain length	Substrate name
C2	4-nitrophenyl acetate
C4	4-nitrophenyl butyrate
C6	4-nitrophenyl hexanoate
C8	4-nitrophenyl octanoate
C10	4-nitrophenyl decanoate
C12	4-nitrophenyl dodecanoate
C14	4-nitrophenyl myristate
C16	4-nitrophenyl palmitate
C18	4-nitrophenyl stearate

2.7.1.1 Substrate specificity

Long chain and short chain *p*NP esters were used to determine substrate specificity. Substrates with fatty acid chains containing even numbers of carbon atoms (C2 – C16) were used. Sodium phosphate buffer at pH 8.0 and 1 µg Enzyme was used (15 min at 37 °C).

2.7.1.2 Temperature optimum and stability

Temperature optima were tested from 17 °C to 90 °C. The amount of enzyme and incubation time was kept as previously mentioned. After preincubation of the enzyme and the buffer for 5 min at indicated temperatures, the reaction was performed with *p*NP octanoate as described above (2.7.1). In order to determine temperature stability of PET2 the enzyme was incubated at 60 °C. The temperature stability of PET2 was determined by incubation of the enzyme at 60 °C. A fraction of the heat treated enzyme was taken each hour and used for a reaction at 60 °C with *p*NP-decanoate (2.6.1).

2.7.1.3 pH optimum

The pH optimum was investigated by incubation in different buffers at different pH. Therefore, citrate buffer (pH 3.0 – 5.0), sodium phosphate buffer (pH 6.0 – 8.0) and carbonate-bicarbonate buffer (pH 9.2 and 10.2) was used.

2.7.1.4 Effect of metal ions, inhibitors, detergents and solvents

Whether or not the enzymes are cofactor dependent was tested at 1 mM and 10 mM concentration. Ca^{2+}, Co^{2+}, Cu^{2+}, Fe^{3+}, Mg^{2+}, Mn^{2+}, Rb^{2+} and Zn^{2+} were used as metal ions in sodium phosphate buffer (pH 8.0). The tested inhibitors phenylmethanesulfonyl fluoride (PMSF), ethylenediaminetetraacetic acid (EDTA) and dithiothreitol (DTT) were added to the reaction buffer at concentrations of 1 mM and 10 mM. The detergents sodium dodecyl sulfate (SDS), Triton X-100 and Tween 80 were used to investigate their influence on the activity on PET2 and PET6. Concentrations of 1% and 5% were used in the reaction buffer. Solvent stability was measured for acetone, acetonitrile, dimethylformamide (DMF), dimethyl sulfoxide (DMSO), ethanol, Isopropyl alcohol and methanol. All organic solvents were used in concentrations of 10% and 30% in sodium phosphate buffer (pH 8.0). All results were compared to an untreated control.

2.7.2 High-performance liquid chromatography (HPLC) measurements

The verification of PET hydrolysis products by HPLC was done to show PET hydrolase activity over time. A modified protocol from Sulaiman *et al.* was used to determine terephthalic acid (TPA), mono(2-hydroxyethyl)terephthalate (MHET) and bis(2-hydroxyethyl) terephthalate (BHET). A total of 7 mg to 14 mg of amorphous PET film with a thickness of 250 µm purchased from Goodfellow Ltd. (Bad Nauheim, Germany), was incubated at 55 °C in 500 µl 100 mM Tris-HCl buffer (pH 7.5), under constant shaking at 500 rpm, for 24 h. The reaction mixture was

diluted with an equal volume ice cold methanol in order to precipitate the enzyme. The samples were centrifuged at max speed (20,800 rcf) in a benchtop centrifuge at 4 °C and measured by HPLC LaChrom Elite® system from Hitachi (Tokyo, Japan) and a Lichrospher® 100 RP-18 endcapped (5 µm) column (VWR International GmbH, Darmstadt, Germany). A total of 99 µl were injected by an auto sampler. Isocratic elution was performed at a concentration of solvent B in solvent A of 20% (v/v) over 20 min. Solvent A was H_2O supplemented with 0.01% (v/v) trifluoroacetic acid and solvent B was acetonitrile. The flow rate was set at 1.0 ml/min. The degradation products of PET were detected by an UV detector set at a wavelength of 241 nm. The column was constantly heated to 45 °C. For the correct assignment of appearing peaks calibration curves were prepared with bis(2-hydroxyethyl) terephthalate (BHET) and terephthalic acid (TPA) (Figure 6). The retention time of TPA was at 8 min and at 10.5 min for BHET. Retention time of MHET was estimated with help of published example data (Yoshida *et al.*, 2016).

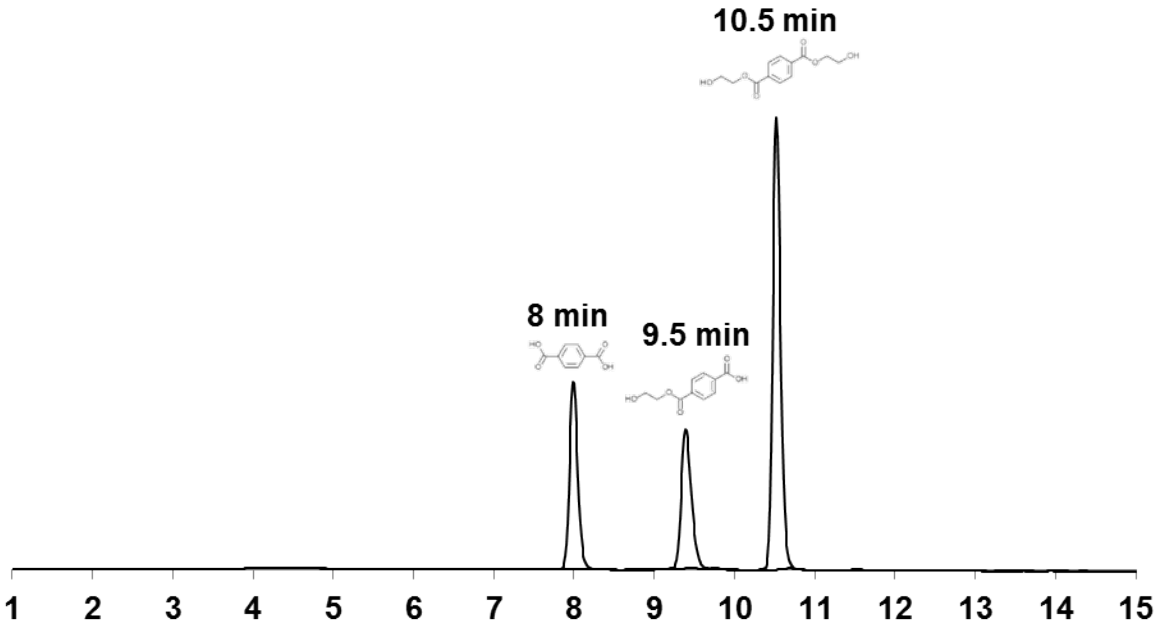

Figure 6: Overlaid HPLC chromatogram from isocratic separation of TPA and BHET. Retention times were 8 min and 10.5 min for TPA and BHET respectively. No standard was available for MHET. MHET retention time was estimated after comparison with published data of similar experiments.

2.8 Bioinformatic software

Sequence data (alignments, sequences) was edited and analyzed with BioEdit version 7.1.3.0 (Hall, 1999), Clone Manager Professional 9 or MEGA6 (Tamura *et al.*, 2013), which was used for calculation of phylogenetic trees as well. Alignments of up to 200 different amino acid

sequences were performed with T-Coffee (Notredame *et al.*, 2000) in accurate mode, to assure highly accurate alignments with structure information. If the desired sequence templates exceeded this limit the default mode of T-Coffee was used due to data processing limitations. The construction of Hidden Markov models was accomplished by the HMMer v3.1b2 package (Finn *et al.*, 2011) (Eddy, 1998) and for visualization of HMM the webpage https://www.skylign.org (Wheeler *et al.*, 2014) was used. After comparison to the HMMer webpage tool default parameters and its resulting output, locally obtained HMM search hits with scores above 180 were considered as significant and e-values for Inclusion and reporting -threshold were set to 10^{-10}. HMM searches were performed with the HMMer package as well on a local machine. DIAMOND-BLAST (Buchfink *et al.*, 2015) and MEGAN6 (Huson *et al.*, 2016) were used to assign taxonomy to the identified new PET hydrolase sequences by first blasting sequences against the complete non-redundant database of NCBI and subsequently taking the top five hits as taxonomical backbone. Sequences without taxonomical assignment were omitted for the phylogenetic analysis (3.3.3). The Global distribution map of PET hydrolases was generated using QGis Desktop 2.18.5 (http://www.qgis.org). Structure Ginzu prediction was performed with the Rosetta software and the Robetta server (Kim *et al.*, 2004). Substrate docking was performed with SwissDock (Grosdidier *et al.*, 2011) and substrate files were obtained from the Zinc[12] database (Irwin, 2017). Structure visualization and analysis was done with UCSF Chimera (Pettersen *et al.*, 2004). The used databases are listed in Table 11. Metagenomes mined for this work are listed in Table A1 and Table A2.

Table 11: Databases used in this study

Databases	Source
NCBI/RefSeq	(Coordinators, 2017)
IMG	(Markowitz *et al.*, 2012)
UniProt	(The UniProt, 2017)

3 Results

3.1 Enrichment of PET degrading organisms

Selective enrichment for PET degrading microorganism was performed in mineral salt medium (M9). Erlenmeyer flasks containing M9 medium were supplemented with PET yarn as sole carbon source, inoculated with environmental samples and incubated at either 22 °C or 28 °C (2.1.4). An enrichment culture, inoculated with mud of the Hamburg harbor basin, showed growth of a thick biofilm, attached to the PET yarn (Figure 7). Transmission electron microscopy images showed morphologic uniform cells attached to the filaments of the PET yarn. Scanning electron microscope images revealed that the cells were directly attached on the surface trough filamentous structures and formed bigger aggregates. The cells were rod shaped with occasionally flat cell poles. A 16S rRNA gene analysis and comparison to the NCBI database showed that the organism is highly similar to *Comamonas testosteroni* (identity >99%).

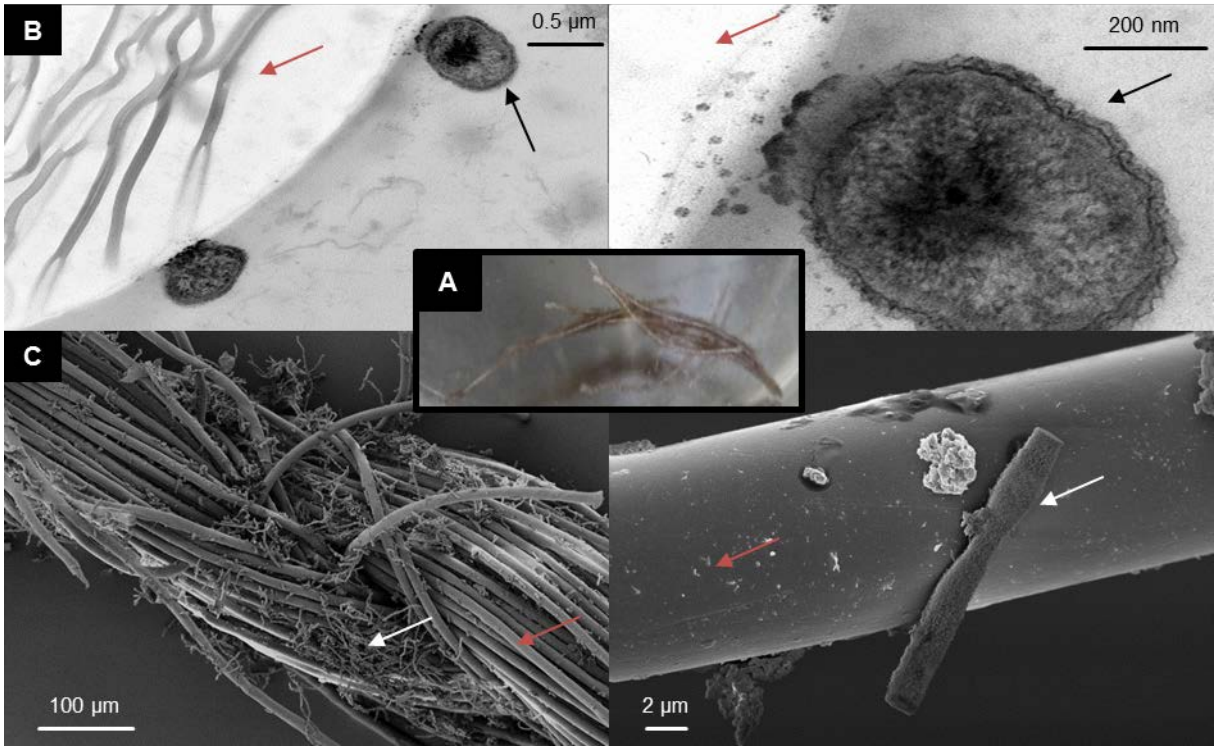

Figure 7: Microscopic imaging of *Comamonas* sp. isolate from Hamburg harbor basin enrichment culture. Red arrows indicate PET filaments. Black and white arrows indicate bacterial cells. **(A)** Biofilm containing PET yarn in M9 medium. **(B left)** Transmission electron microscopy image of a PET filament with attached *Comamonas* cells. **(B right)** Close-up of an attached cell. **(C left)** Scanning electron microscope image of PET yarn with a multitude of cells. **(C right)** Close-up of a single cell on the surface of a PET filament. Used microscopes and sample preparation are described under 2.3.

The DNA was also used for Illumina sequencing in order to obtain a draft genome. Amino acid comparison with the PET hydrolase specific HMM didn't show any PET hydrolase like sequences. A further KEGG database comparison showed that the *Comamonas* strain seems to be unable of metabolizing terephthalic acid.

3.2 Function based screening of metagenomic libraries with indicator-plates

Metagenomic libraries (2.3) listed in Table 7 were functionally tested on indicator plates (0). A total of 78,897 clones were tested, corresponding to about 1.9 to 3.8 Gb of DNA, assuming that the fosmid vector pCC1FOS contained 20 to 40 Kb of DNA. Assays were performed on DEP and PCL to find active hydrolases with the potential of PET degradation. Screening on DEP indicator plates resulted in a single clone capable of efficiently degrading DEP with a visible halo (Figure 8). The fosmid clone AL103.F12 was obtained from a photobioreactor metagenome library. To identify the protein coding sequence within the fosmid, a primer-walk analysis of the sequence was performed (2.4.2). The resulting construct (2.2) was used with *E. coli* M15 [pREP4] host cells to produce a 37.9 kDa his-tagged protein. Further tests with PCL indicator plates (2.4.1.1) showed that the protein cannot degrade complex polymers. Screening on PCL showed no active clones in the existing metagenomic libraries.

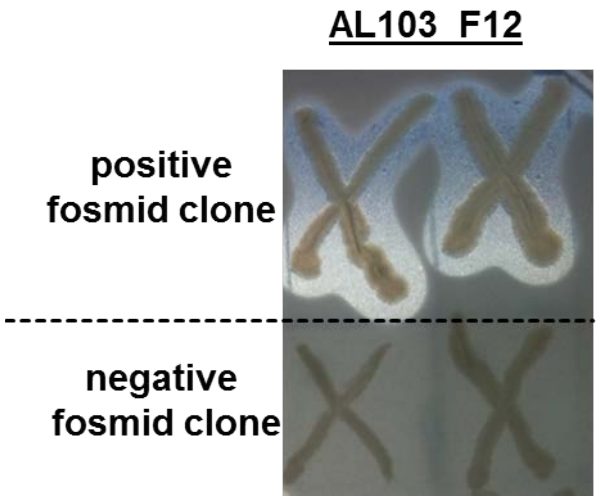

Figure 8: DEP hydrolase from the fosmid AL103_F12, subcloned and expressed in *E. coli* on LB/agar plates containing DEP. Halo formed after incubation at 37 °C over night. LB plates contained 1.5% polycaprolactone (PCL).

3.3 Sequence based screening via a self-constructed HMM

HMM based screening was performed with metagenomic datasets and public protein databases as input. A total of 1536 potential PET hydrolase sequences were obtained by this approach. Twelve new potential PET hydrolases were used for further analysis and a subset of four hydrolases was functionally tested.

3.3.1 Alignment of known PET hydrolase sequences and construction of the HMM

For construction of a PET hydrolase specific HMM (2.5.1), an alignment of nine known and functionally tested PET hydrolase sequences was used as template (Table 8) (Figure A 1). Following the visualization of the HMM, eight highly conserved regions were identified, suggesting that they are critical for PET hydrolase activity. Within these motifs, specific residues may be essential for the formation of the final protein structure (Figure 9) (Table 12).

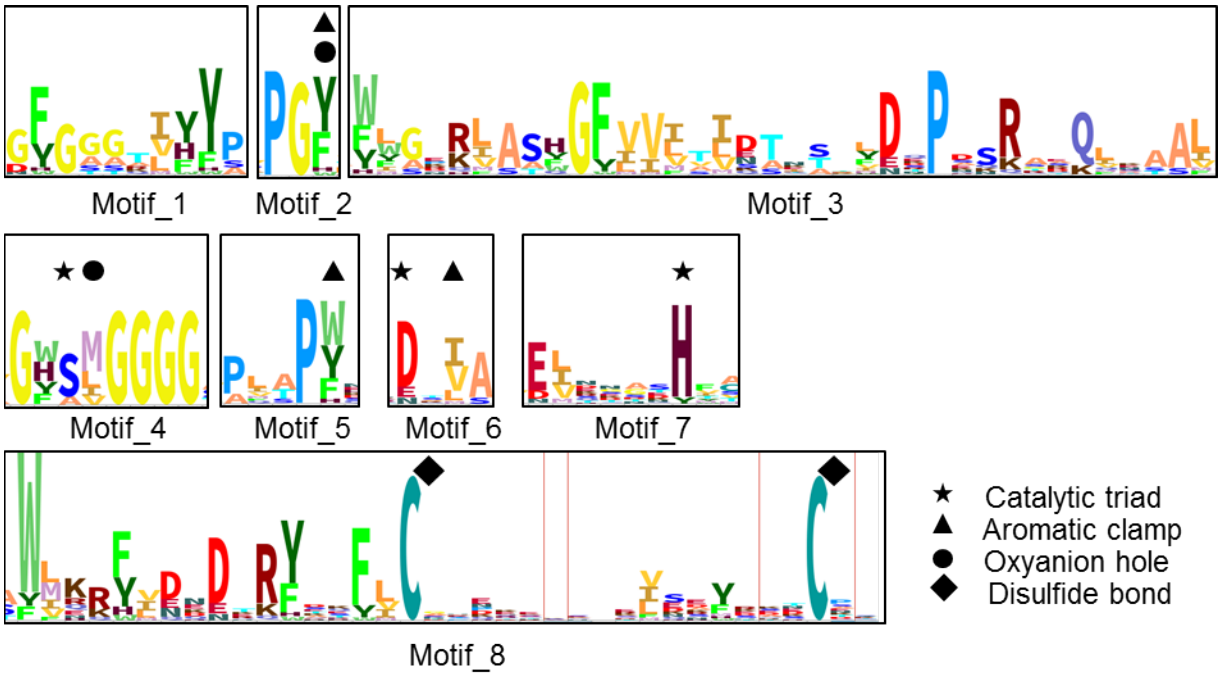

Figure 9: HMM logos of PET hydrolysis relevant motifs. Marked residues are highly conserved and fulfill specific functions like substrate binding, thermostability and substrate hydroxylation (see Table 12). The HMM was constructed on basis of a multiple alignment of described PET hydrolases using the HMMer software package.

Table 12: Determined search criteria for the identification of PET hydrolase candidates. The letter x indicates a non-conserved position within the sequence pattern. Brackets indicate a less conserved position within the sequence pattern. Numbering of amino acids and motifs is according to the HMM (see Figure 9).

Motif	Search criteria	Function
2	F,Y62	Aromatic amino acid for oxyanion hole formation and aromatic clamp
4	GxSMGGGG	Serine of catalytic triad and methionine for oxyanion hole formation
5	W,Y157	Aromatic amino acid for oxyanion hole formation and aromatic clamp
6	I,V180	Aromatic amino acid for oxyanion hole formation and aromatic clamp, aspartic acid of catalytic triad
7	H210	histidine of catalytic triad
8	DxDxR(Y)xxF(L), C245, C262	Conserved sequence prior to C-terminal cysteine residues for thermostability

3.3.2 Seeking and classification of potential PET hydrolases from databases

The published and previously described database search only included the UniprotKB database (Danso *et al.*, 2018). For an even more extensive search of new PET hydrolases the NCBI non-redundant protein database was used together with the specific HMM. The search with a reporting score threshold of >180 resulted in 1187 hits which surpassed the previous output more than twice. A phylogenetic assignment showed that 1044 sequences belong to the phylum of *Actinobacteria,* which complies for 87.96% of all found sequences. The second most abundant is the phylum of *Proteobacteria* with 118 sequences and a share of 9.93%. *Bacteroidetes, Deinococcus-Thermus* and *Cyanobacteria* sequences are represented by 13, four and one sequences respectively which account altogether to little more than 1.5%. Seven sequences were not assigned to any phylum. Due to the huge amount of sequence data it was impossible to group the sequences into bigger clusters like previously done with the much smaller dataset.

3.3.3 Seeking PET hydrolases within metagenomic datasets

Parallel to the database search, a metagenome-based search was performed. The metagenome records in Table A 1and Table A 2 were obtained from Integrated Microbial Genomes (IMG). A total of 108 marine and 25 terrestrial metagenomes were analyzed. Marine sequences were filtered in terms of sample depth and only metagenomes from a maximum depth of two meters were used. Potential PET hydrolase sequences were found in 31 marine and 11 terrestrial records. Identical datasets were reduced. The size of the assembled metagenomic data in case of marine metagenomes ranged from 10.85 Mb to 7.99 Gb. In case of terrestrial metagenomes,

the number of assembled bases ranged from 58 Mb to 9.2 Gb, all together 16 Gb of assembled DNA was analyzed using this approach. In absolute terms, analyzed metagenomic data suggest that at least 16 phyla have the potential of PET degradation (Figure 10).

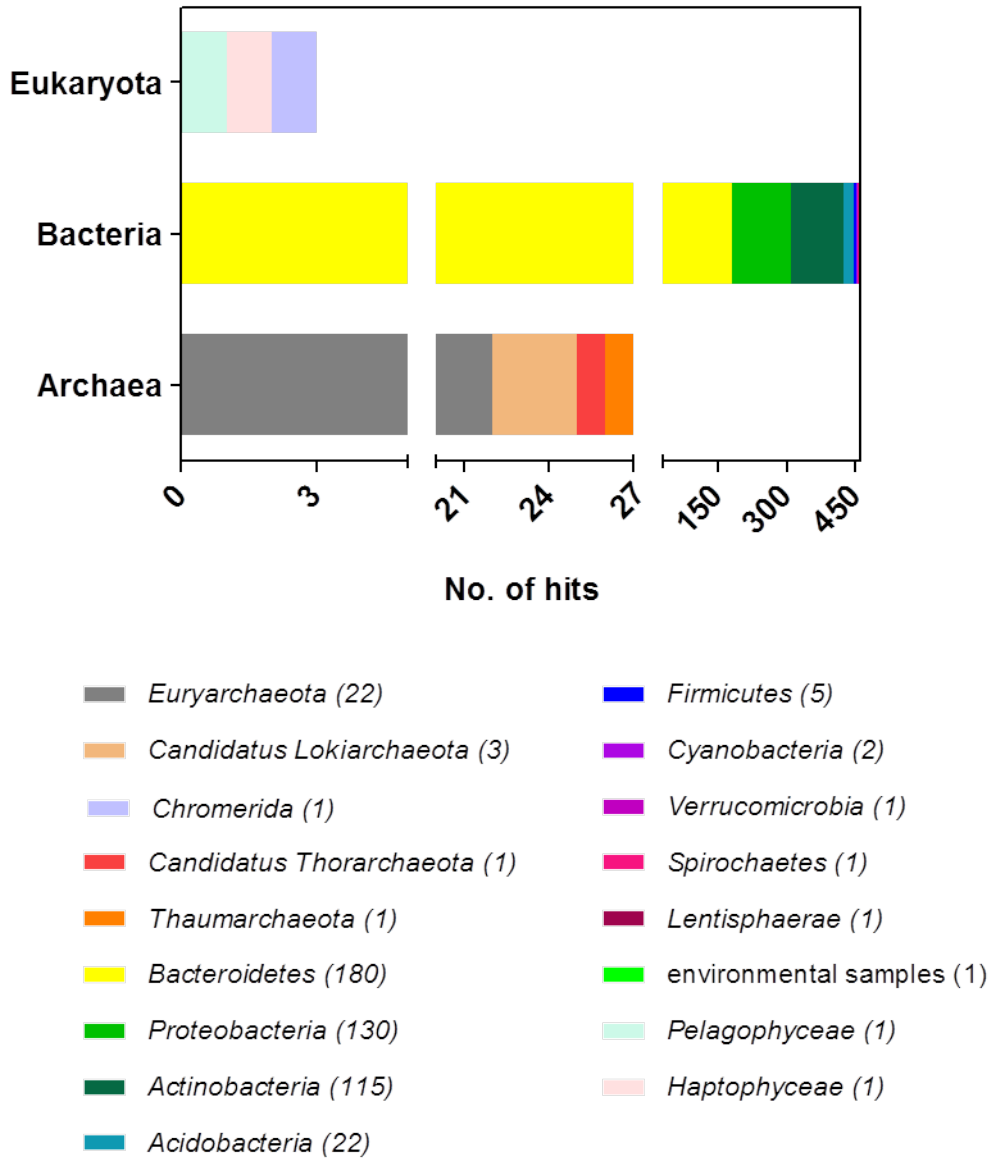

Figure 10: Combined phylogenetic assignment of metagenomic PET hydrolase hits. All found hits were assigned taxonomically on phylum level and grouped according to the three domains of life. The metagenomic sequences were obtained from IMG and are listed in Table A 1 and Table A 2.

Bacteria were represented by nine phyla and one environmental sample without annotation, Archaea by four and the domain *Eukaryota* by three. Within the Bacteria, *Bacteroidetes* was the most abundant phylum (180 hits), followed by *Proteobacteria* (130 hits) and *Actinobacteria*

(115 hits). The domain *Eukaryota* was represented by the phyla *Pelagophyceae*, *Haptophyceae* and *Chromerida* with one hit each. The four phyla associated with *Archaea* were *Euryarchaeota* (22 hits), *Thaumarchaeota* (1 hit) and the two phyla *Thorarchaeota* (1 hit) and *Lokiarchaeota* (3 hits), which have candidatus status. Percentage distribution shows no critical fluctuations between both metagenome datasets in case for phyla shared by both datasets. The only exceptions are the phyla *Bacteroidetes* and *Actinobacteria*. Whereby the *Actinobacteria* PET hydrolase hits are more abundant in terrestrial metagenomes and the *Bacteroidetes* PET hydrolase hits are dominant within marine metagenome sequences. The calculated ratio of *Actinobacteria* hits to *Bacteroidetes* hits within marine and terrestrial sequences is nearly one to one (Figure 11). Analyzed data showed that bacterial hits were by far the most abundant with a total of 458 hits, followed by archaeal hits which complied for 27 assigned sequences and eukaryotic sequences with just three hits. A total of 89.4% and 75.7% of all sequences were assigned taxonomically for marine and terrestrial metagenomes respectively. Remaining sequences could not be assigned unambiguously.

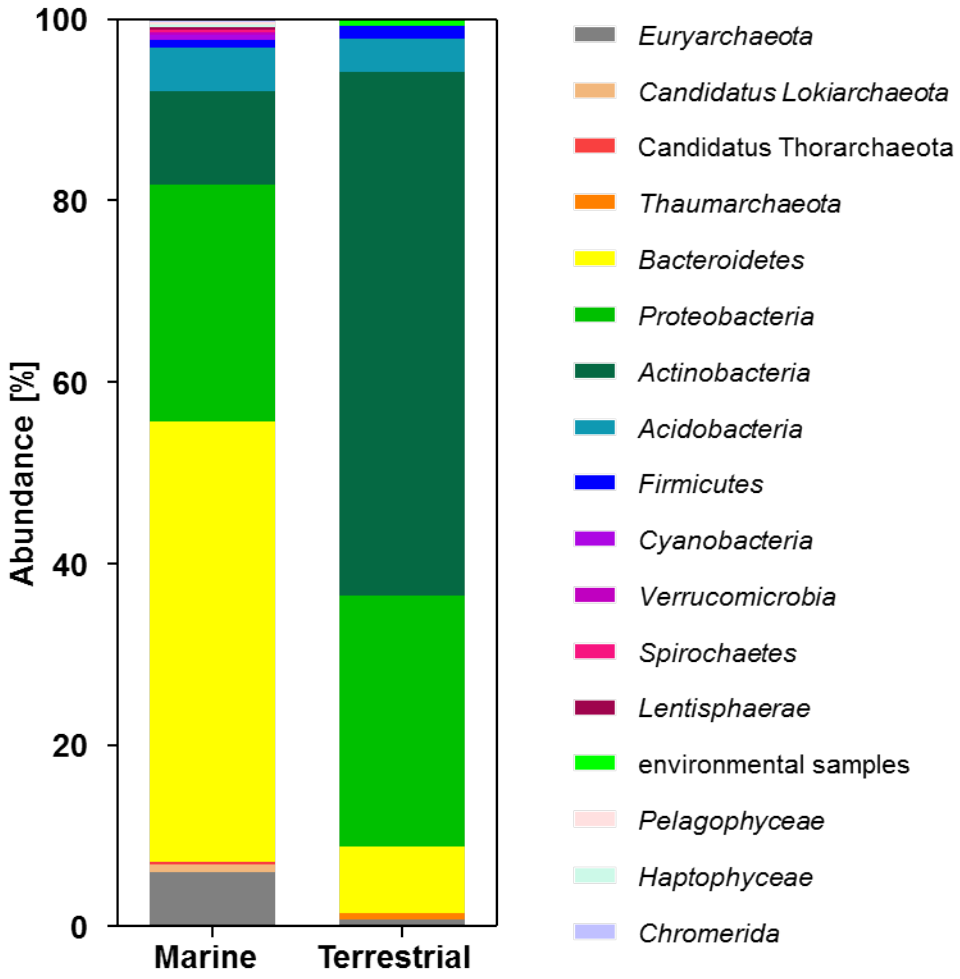

Figure 11: Taxonomical distribution on phylum level of PET hydrolase hits from metagenome records. Hits are shown for terrestrial and marine metagenome data independently. The metagenomic sequences were obtained from IMG and are listed in Table A 1 and Table A 2.

3.3.3.1 Geographical distribution of PET hydrolase harboring organisms from metagenome datasets

In addition to a taxonomic analysis, metadata such as the geographic distribution of the sample sites were linked to the frequency of PET hydrolase hits/Mb (Figure 12). A total of 108 marine and 25 terrestrial metagenome datasets were analyzed. After the reduction of records without hits, 11 terrestrial and 31 marine datasets remained. This corresponded to 44% and 28%, respectively, for terrestrial and marine metagenome records. Most PET hydrolases (31 hits) in marine data sets were found in a sample from Delaware Coast (USA) (IMG_ID: 3300000101). In terrestrial datasets, most hits (135) occurred in a metagenome from a sediment core of a heavy oil reservoir (Alberta, Canada; IMG_ID: 13680). In the case of terrestrial habitats, the heavy oil

metagenome also had the highest frequency (1.5 hits/Mb). For marine data sets, a metagenome from La Parguera in Puerto Rico (IMG_ID: 3300000444) had the highest frequency (0.092 hits/Mb) with only one hit. This dataset is nearly 60 times smaller than the record with most PET hydrolase hits.

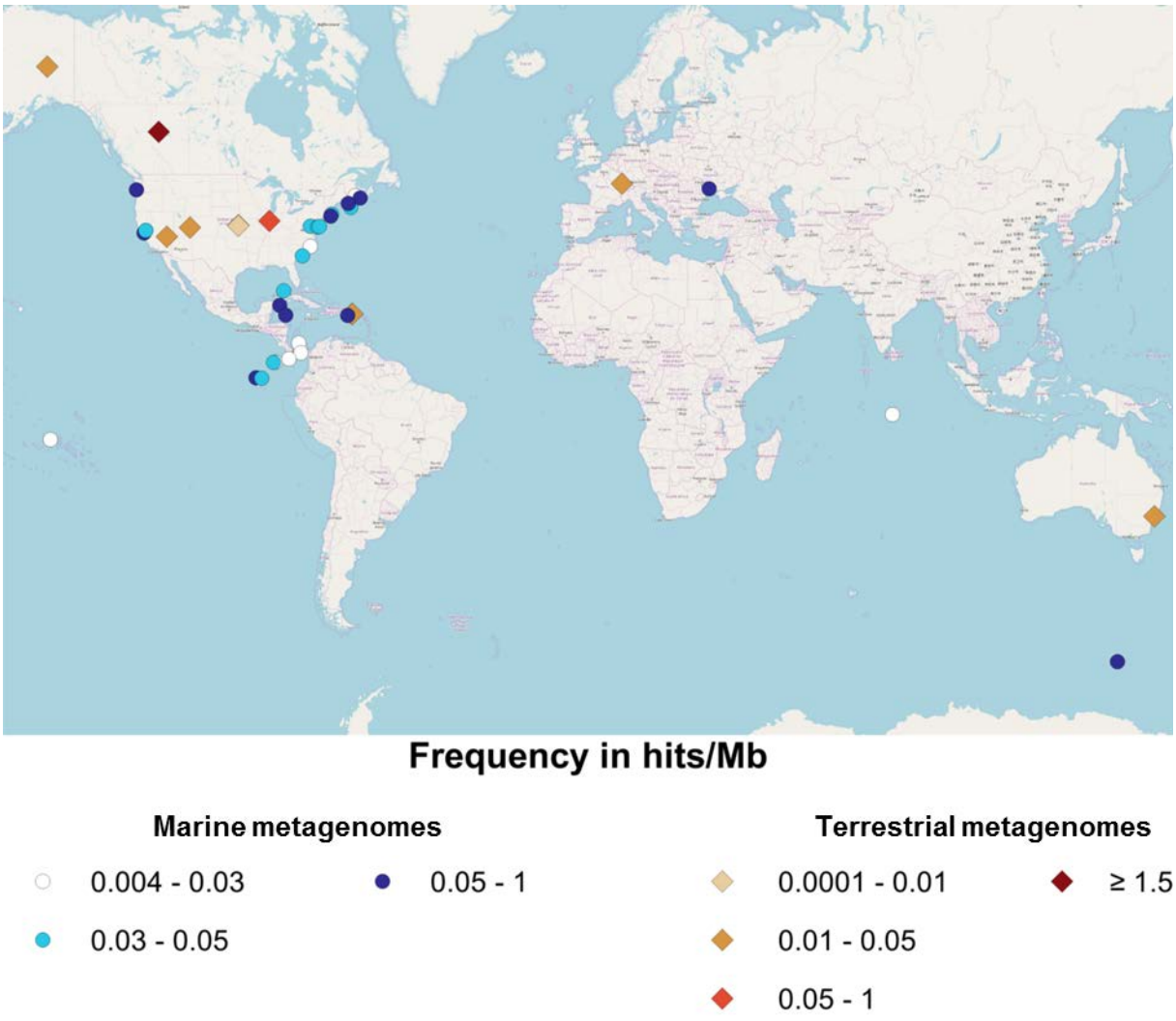

Figure 12: Geographical distribution of all used metagenomic datasets containing significant PET hydrolase hits. The map was taken from Openstreetmap.org and is available under database contents License (DbCL) 1.0. Markings for samplesites were added subsequently with QGis desktop version 2.18.5.

3.4 Characterization of the potential PET hydrolases PET2, PET5, PET6 and PET12

The genes of PET hydrolases, PET2, PET5, PET6 and PET12 obtained by HMM search were synthesized or amplified from the corresponding organisms, cloned into expression vectors and expressed by *E. coli* T7SHuffle cells (2.6). The resulting partially pure enzymes (Figure 13) were

used for several characterization experiments. PET2 was a metagenomic sequence synthesized at Eurofins. PET5, PET6 and PET12 were amplified with the primers listed in Table 4. The source organisms were cultivated according to the DSMZ protocol for each bacterium (2.1.2).

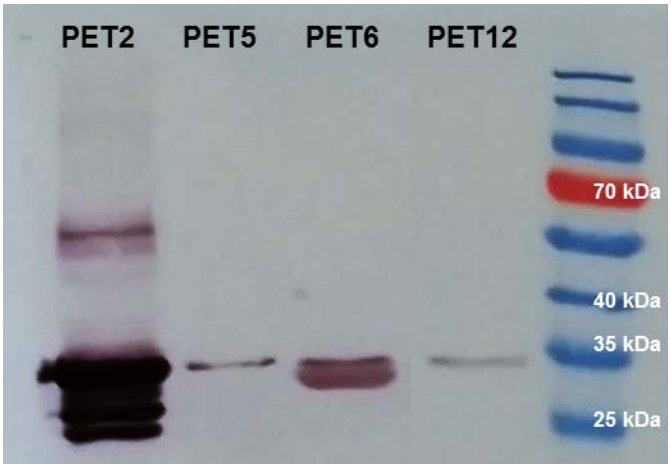

Figure 13: Western blot analysis of partially purified PET hydrolases found in this study. An equal volume of 10 µL was used for all samples. Samples were chosen from elution fractions containing most protein (100-2000 µg/µl) and detected with specific penta-his antibodies. Loaded reference marker is the prestained protein ladder "PageRuler™" from Thermo-Fisher.

3.4.1 Plate assay

Plate assays with PET-nanoparticles (2.4.1.2) and PCL (2.4.1.1) were performed in parallel for PET2, 5, 6 and 12 (Figure 14). Compared to a negative control consisting of *E. coli* T7SHuffle cell raw extract and a positive control from heterologously expressed Tfcut2 (Roth *et al.*, 2014), all His-tag purified enzymes showed activity. A volume of 50 µL containing 0.5 mg of protein (2.6.2) was spotted on each plate. After incubation at 50 °C or room temperature over night for PCL and PET plates respectively, halo formation could be observed. The biggest halo on PCL was created by the positive control Tfcut2, followed by PET6 and PET2. On PET nanoparticle plate, the positive control showed better activity as well followed again by PET2 and PET6. PET5 and PET12 produced much smaller and less transparent halos. For this reason further experiments focused on PET 2 and PET6.

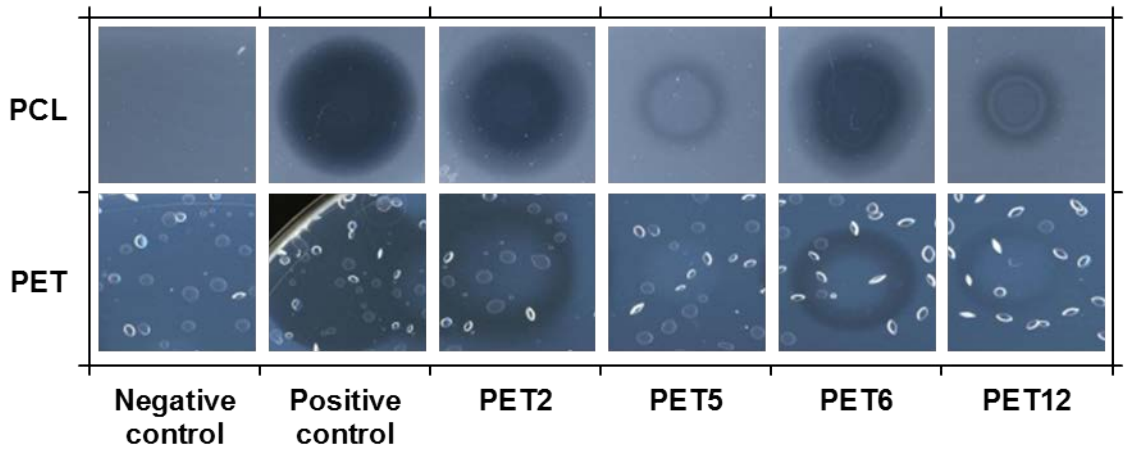

Figure 14: Plate assay with purified PET2, PET5, PET6 and PET12. The purified enzymes were spotted in equal amounts (0.5 mg) on a PCL and PET-nanoparticle containing plate respectively. After incubation at 50°C and room temperature, for PET nanoparticles and PCL respectively, halo formation was observed. The negative control was *E. coli* crude extract prepared with sodium phosphate buffer. The positive control contained Tfcut2 purified in a similar manner.

3.4.2 Substrate specificity, temperature optimum and temperature stability

The substrate specificity test (2.7.1.1) clearly showed that PET2 and PET6 both favor short-chain substrates such as *p*NP-butyrate (C4) and *p*NP-acetate (C2) (Figure 15). Tests with longer substrates showed that PET6 is capable of hydrolyzing ester bonds up to a chain length of 18 carbon atoms. PET2 was able to hydrolyze a chain length of C14, but beyond that none of the substrates. The optimum temperature (2.7.1.2) for PET6 was about 55 °C (Figure 16). Remaining activity of up to 80% was measured at 70 °C and even at 90 °C a residual activity of nearly 30% was maintained in the case of PET6. PET2 had its highest activity between 60 °C and 70 °C and was able to tolerate even elevated temperatures of up to 90 °C. Temperature optimum assays suggested high thermostability for PET2 therefore the thermostability of PET2 was further elucidated. Stability measurements of PET2 at 60 °C over more than 5 hours resulted in no significant decrement of activity (Figure 17), suggesting a highly thermostable enzyme.

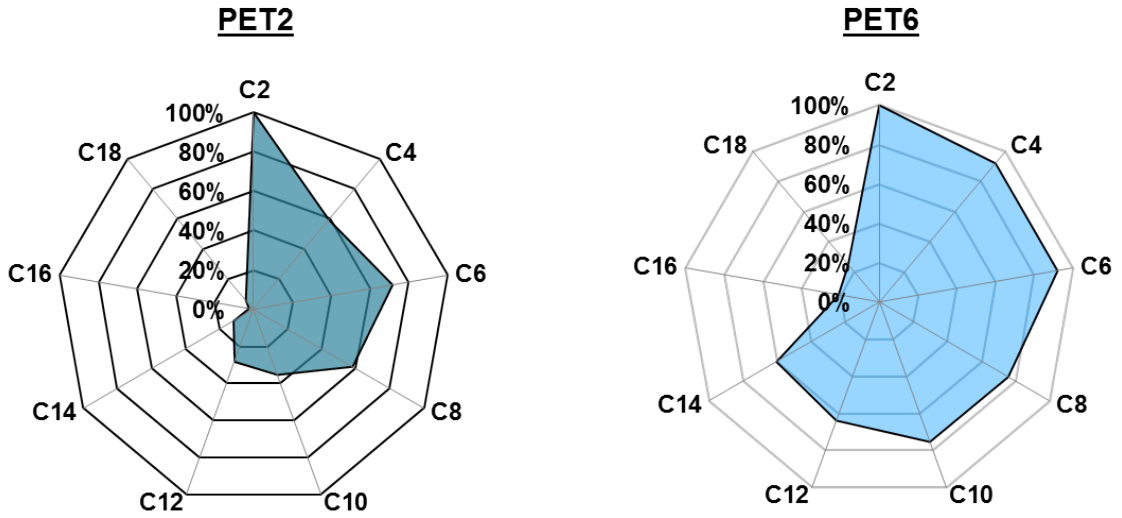

Figure 15: Substrate specificity of PET2 and PET6, calculated as percentage of relative activity. The specificity for several *p*NP-substrates with different carbon side chains (C2-C18) was tested. The standard deviation was below 0.15 and 0.1 for PET2 and PET6 respectively. Data are mean values of three independent measurements.

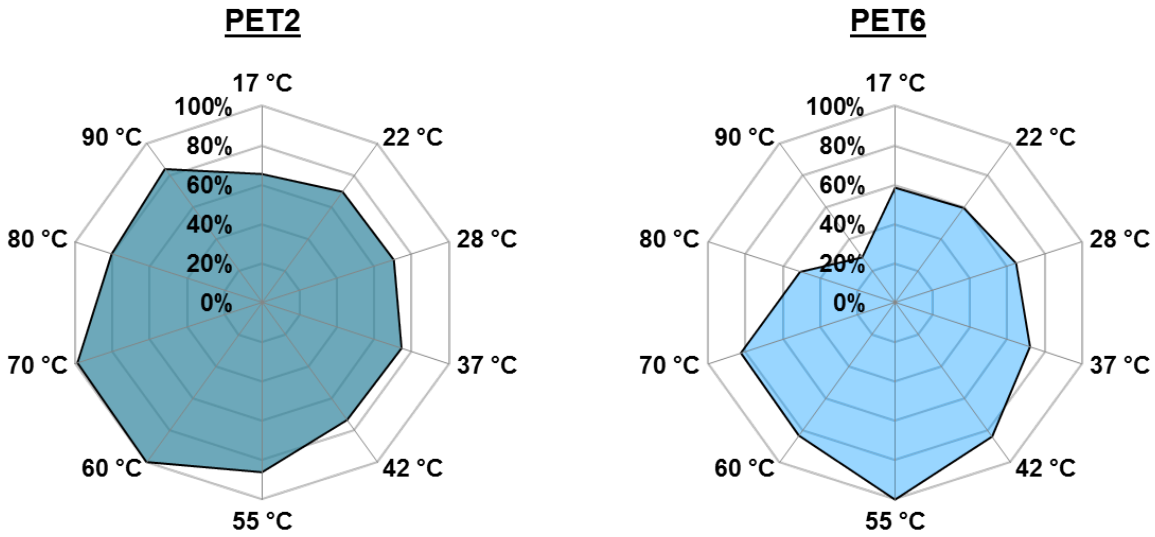

Figure 16: Temperature optimum of PET2 and PET6, calculated as percentage of relative activity. Tested temperatures ranged from 17 °C to 90 °C. Optimal temperatures were 60 °C and 55 °C for PET2 and PET6 respectively. The standard deviation was below 0.2 for PET2 and PET6. Data are mean values of three independent measurements.

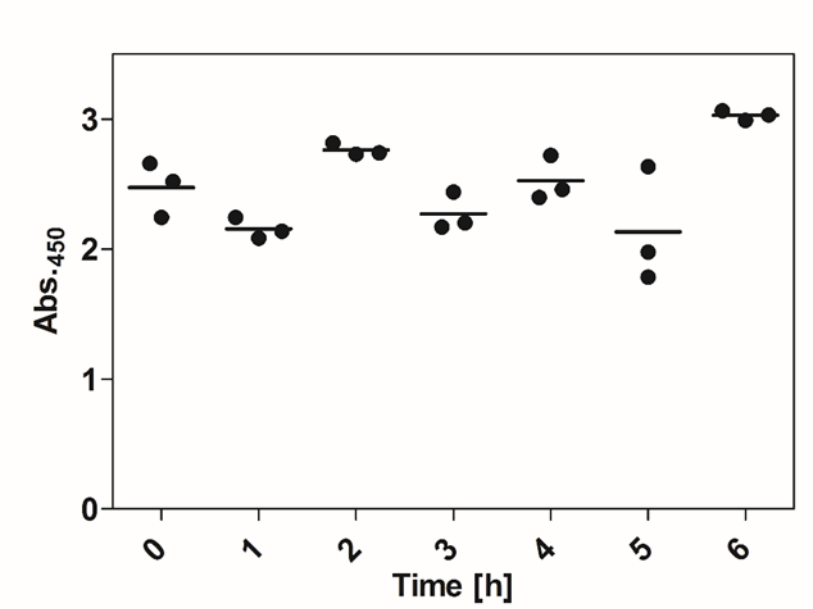

Figure 17: Thermostability of PET2. In an interval of one hour a fraction of the heat-treated enzyme was used in an activity assay (2.7.1) to verify thermostability at 60 °C. Data are mean values of three independent measurements represented by black dots.

3.4.3 Optimum pH

Buffers with pH between 3 and 10.2 were used to discover the optimal pH conditions for PET2 and PET6 (2.7.1.3). With *p*NP substrates it was shown that PET2 is much more active under alkaline conditions (pH >8.0). PET6 had a much narrower pH spectrum but had a comparable pH optimum as PET2 at 8.0 (Figure 18). Both enzymes lost their activity nearly completely at pH values below 7.0. Only 6% and 10% activity compared to the highest value in this experiment were reached at pH 6.0.

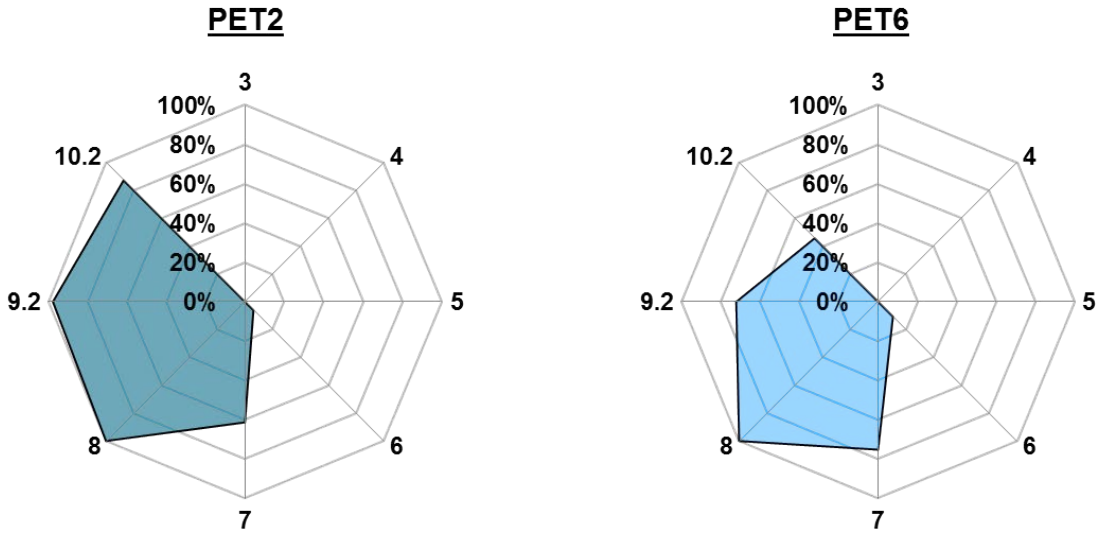

Figure 18: Calculated pH optimum of PET2 and PET6 as percentage of relative activity. Both hydrolases prefer slightly alkaline pH conditions from pH 8.0 to pH 9.2. The standard deviation was below 0.1 for PET2 and PET6. Data are mean values of three independent measurements.

3.4.4 Effect of metal ions, inhibitors detergents and solvents

PET2 showed enhanced activity with all tested metal ions (Figure 19). In general, a concentration of 1 mM enhanced the activity by up to 85%. For higher concentrations, rubidium had the greatest effect on activity with an enhancement of 90% compared to the control. In case of calcium and magnesium either concentration showed comparable results. The smallest effect was observed by adding 10 mM iron to the reaction with an activity increment of 33%. The activity of PET6 was positively influenced by metal ions only at lower concentrations (1 mM). Most beneficial was the substitution of 1 mM rubidium, which resulted in a rather small activity increment of 0.8%. Higher concentrations of metal ions (10 mM) minimized the activity of PET6.

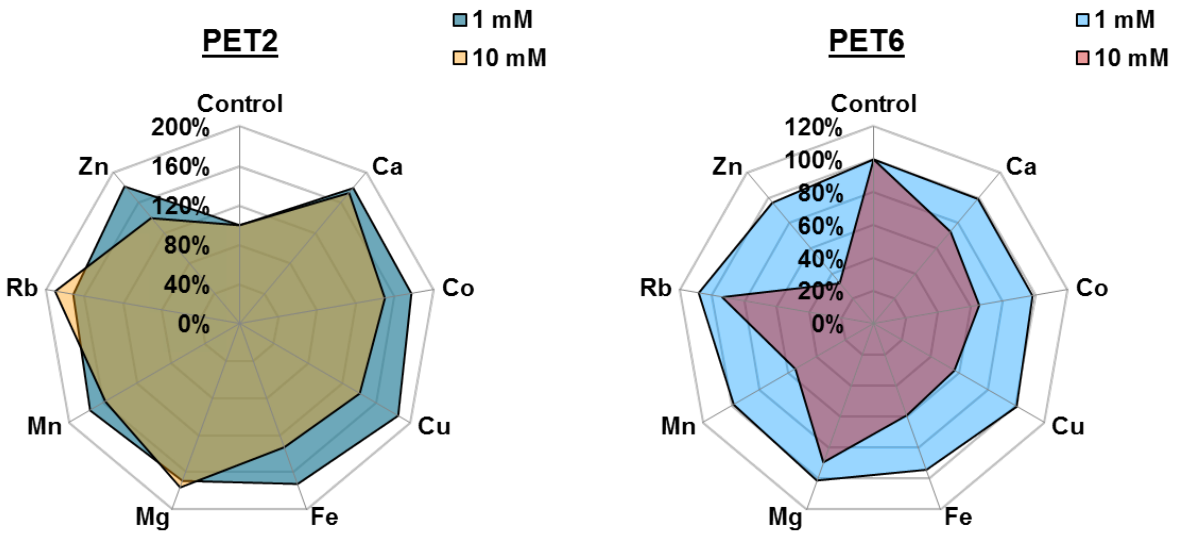

Figure 19: Effect of metal ions on PET2 and PET6 calculated in percentage of relative activity. The effect of different metal ions in concentrations of 1 mM and 10 mM were tested. The standard deviation was below 0.3 and 0.1 for PET2 and PET6 respectively. Data are mean values of three independent measurements.

Inhibitors were tested in two concentrations (1 mM and 10 mM). PET2 showed reduced activity with supplemented inhibitors in both concentrations and for every substance used in this experiment (Figure 20). At a concentration of 1 mM, PMSF had the highest impact on the activity of PET2. The enzyme lost 80% of its activity with 1 mM PMSF and 77% with 10 mM PMSF. EDTA had the smallest effect on the activity. At a concentration of 1 mM, 8% activity was lost and with 10 mM the activity was reduced by just 4%. PET6 showed a little enhancement in activity if the buffer contained 10 mM EDTA.

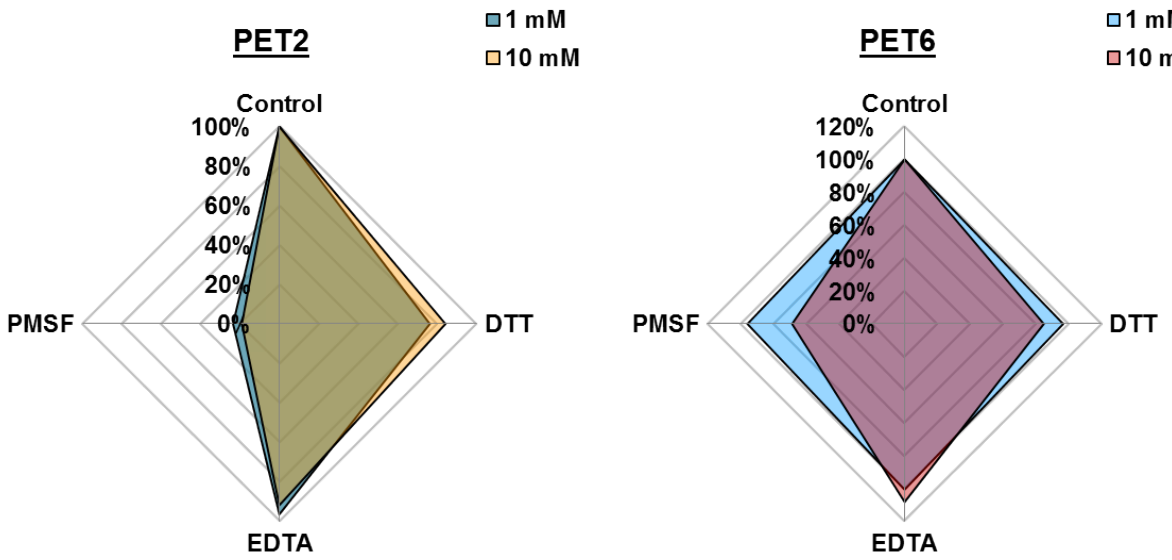

Figure 20: Impact of inhibitors on PET2 and PET6, calculated as percentage of residual relative activity. The assay was performed at concentration of 1 mM and 10 mM. PET2 was less stable against the effect of PMSF than PET6. Both enzymes tolerated EDTA and DTT. The standard deviation was below 0.1 and 0.05 for PET2 and PET6 respectively. Data are mean values of three independent measurements.

The influence of detergents on the activity of PET2 and PET6 (2.7.1.4) showed that Tween 80 improved the activity (Figure 21). A more than 10% higher activity was observed for both enzymes when 1% Tween 80 was added. On the contrary, higher concentrations of Tween 80 reduced the activity. The other tested detergents had a similar effect on PET2 and PET6. All other detergents reduced the relative activity independent of the used concentration.

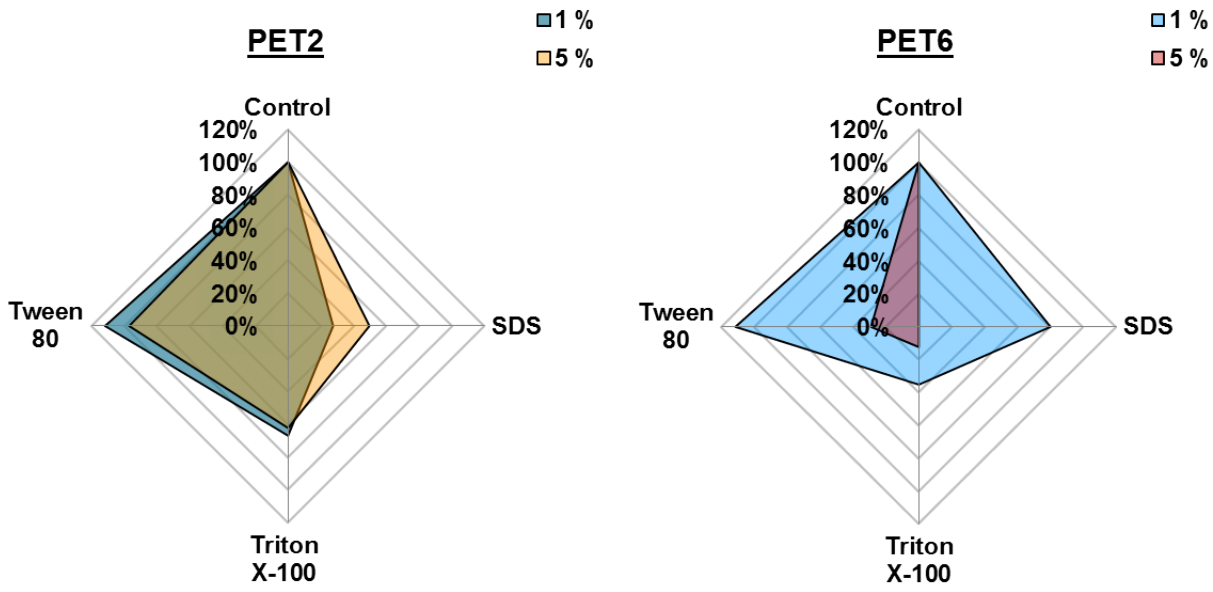

Figure 21: Detergent impact on PET2 and PET6, calculated as percentage of residual relative activity. The activity was tested at concentrations of 1% and 5%. Both enzymes showed improved activity under low concentrations of tween 80. SDS at higher concentrations had the most negative effect on the activity for both PET hydrolases. The standard deviation was below 0.1 for both tested enzymes. Data are mean values of three independent measurements.

PET2 and PET6 were tested for their stability in different organic solvents at different concentrations (2.7.1.4) (Figure 22). PET2 lost 16% of its relative activity at a concentration of 10% isopropanol. All other solvents had no considerable effect on PET2 at 10% concentration. Higher concentrations had a much stronger effect on the activity of PET2. Apart from DMSO, all solvents reduced the activity to at least 70%. PET6 appeared to be much less stable to organic solvents than PET2. Even at a concentration of 10%, the activity of PET6 was lowered by each tested solvent. The highest effect was observed for acetonitrile. Only 1% residual activity was measured at a concentration of 30% acetonitrile.

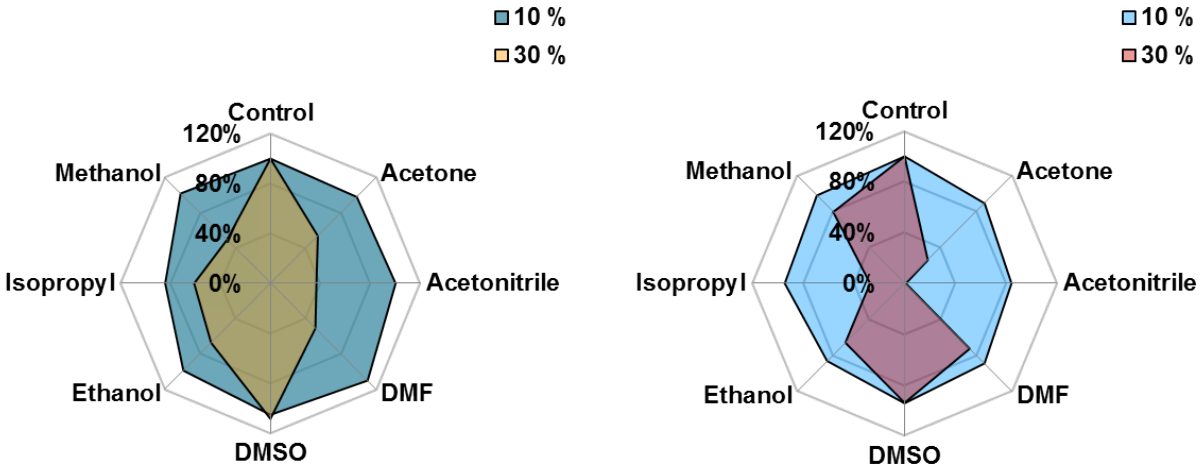

Figure 22: Solvent stability of PET2 and PET6, calculated as percentage of residual relative activity. Concentrations of 10% and 30% were tested. Both PET hydrolases were reduced in activity at higher concentrations. The biggest effect was observed for acetonitrile. The standard deviation was below 0.1 for both tested enzymes. Data are mean values of three independent measurements.

3.4.5 HPLC measurement of PET degradation products

HPLC analysis of PET2 degradation products showed that TPA was the major component, followed by MHET. BHET was almost unobservable (Figure 23 A). Commercially available BHET showed a little impurity which is most likely MHET. Samples taken every hour showed an increasing TPA concentration (Figure 23 B). In experiments using 14 mg of amorphous PET foil as substrate, 900 µM terephthalic acid was measured after 24 h of incubation with 100 µg of PET2. Since no standard was commercially available for MHET, the total amount could not be calculated.

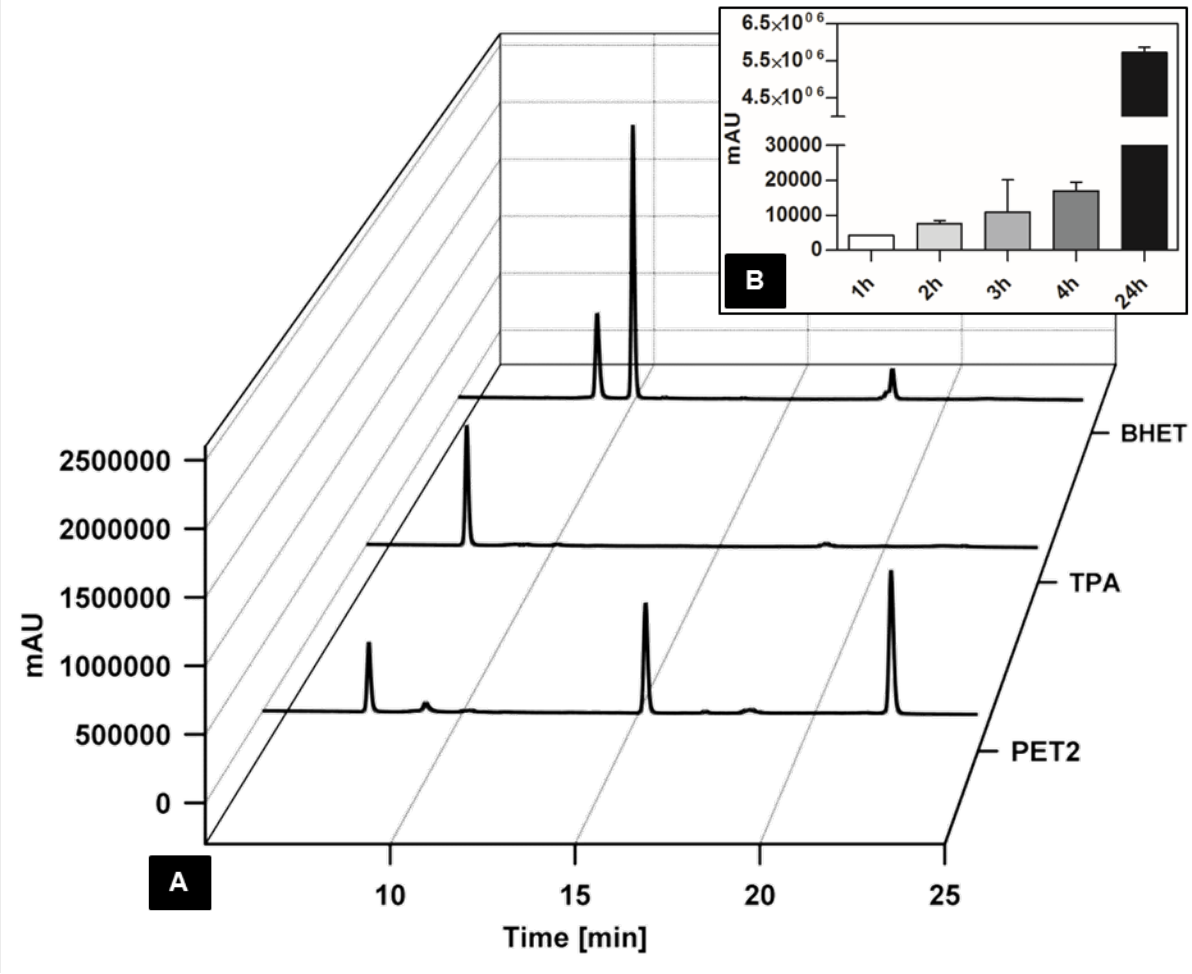

Figure 23: High performance liquid chromatography of PET degradation products. **(A)** PET2 was incubated for 24 h together with 14 mg of amorphous PET foil. The degradation products were compared to commercial available standards. **(B)** Increasing TPA concentration was observed during the experiment. Data in A and B represent an example of a routinely performed experiment. In general assays were done three times.

3.5 *in silico* analysis of functional tested PET hydrolases from this work

The amino acid sequences of PET2, 5, 6 and 12 were compared to the NCBI non-redundant database. The metagenome derived sequence of PET2 showed the highest similarity to a hydrolase from *Marinobacter nanhaiticus* (WP_085988667) with an identity of 65% over the total length of the sequence. All other PET hydrolase sequences were highly similar to sequences from closely related organsims of the same genus. A similarity comparision revealed rather low similarities between the newly identified and the known PET hydrolase sequences (see Table 8). All identified and functionally tested PET hydrolases proofed to have similarity values under 75%. Surprisingly, the sequences showed low similarity also among each other (Figure 24). PET2 and PET6 had the lowest overall similarity values.

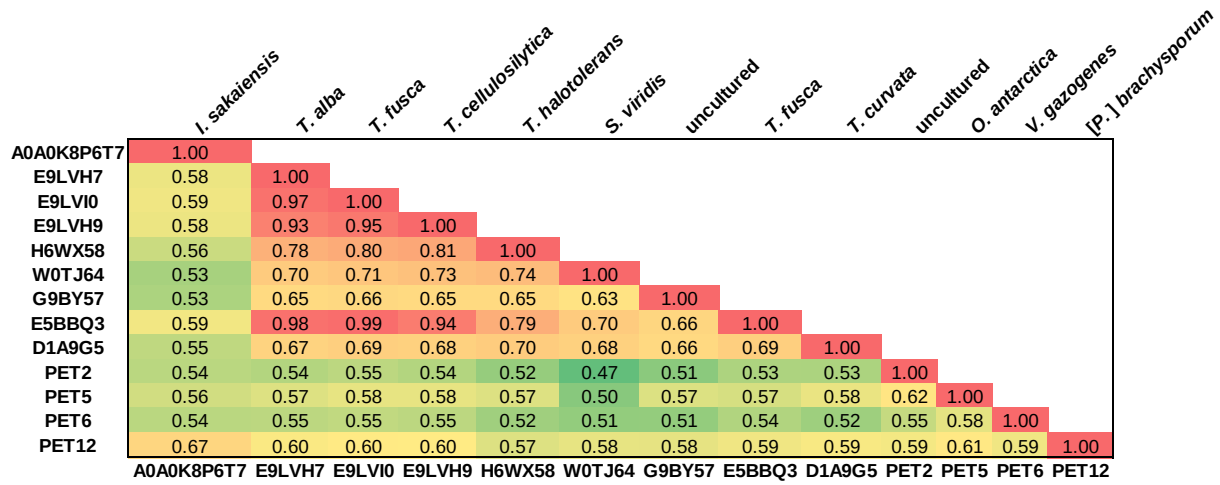

Figure 24: Similarity comparison of known and newly identified PET hydrolase sequences. The comparison was performed with the online tool SiAS (see 2.8). Values range from zero to one. High values are colored red. Low values are colored green, with color transition for values in between the minimum and maximum.

PET 2 performed very well during the biochemical characterization and it had only low similarity to other PET hydrolases and database entries. Therefore, this sequence was chosen for a structure prediction and as most promising PET hydrolase of this work. The protein structure of PET2 was predicted with the Rosetta software on the Robetta server. Afterwards, the model substrate BHET was used for a docking prediction with SwissDock (2.8) (Figure 25). BHET was located in ultimate vicinity to the catalytic triad in a hydrophobic cleft on the surface of the enzyme. Similar protein substrate interaction was observed for other PET hydrolases as well (Han *et al.*, 2017; Joo *et al.*, 2018).

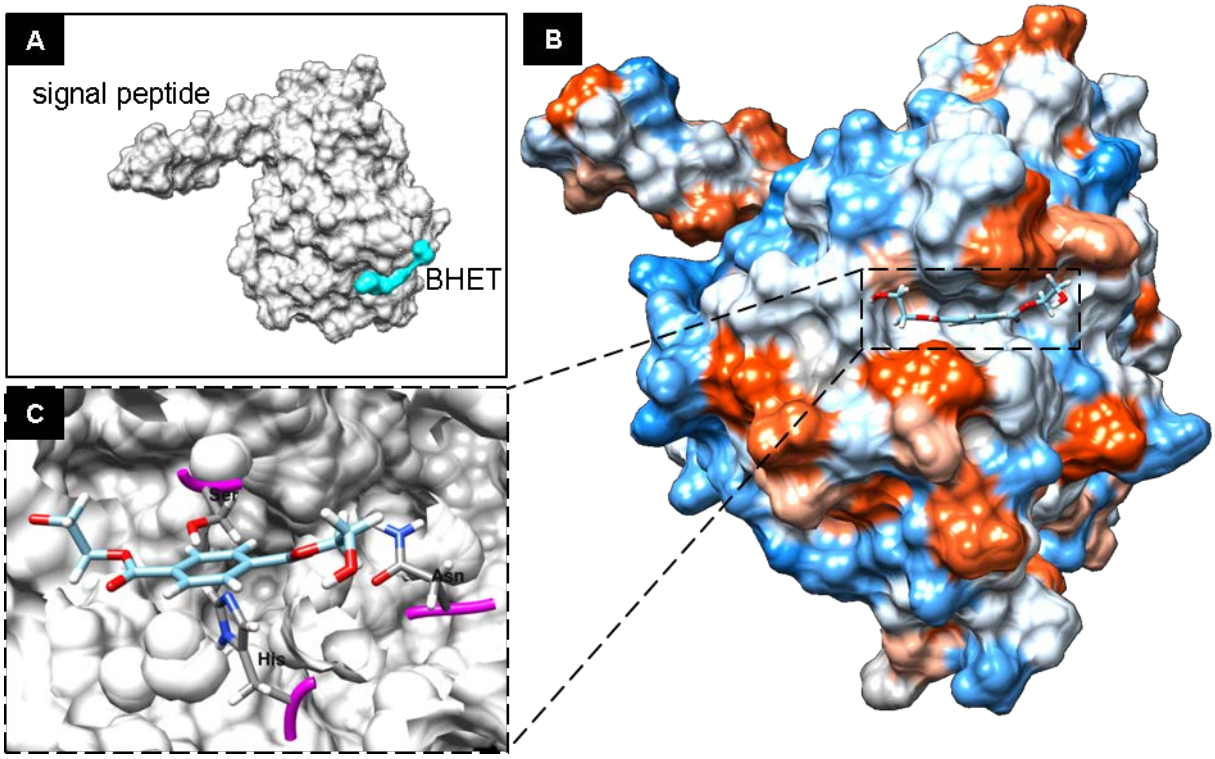

Figure 25: 3D-Structure prediction model of PET2. The native amino acid sequence of PET2 was used for the prediction of this model. BHET was used as substrate due to its similarity to PET and its size. **(A)** The surface view reveals BHET as a fitting substrate for this protein within a cleft on the protein surface. Hydrophobic areas are colored in red, hydrophilic areas are colored in blue. **(B)** Hydrophobicity surface view of PET2 with bound BHET molecule. **(C)** Close up of the active site and the catalytic triad, consisting of a serine (Ser), aspartic acid (Asp) and a histidine (His). BHET was used for the docking experiment.

4 Discussion

Existing studies focused mainly on the mechanism of PET degradation by hydrolases in general or the identification of individual new enzymes and their properties (Wei, Oeser, & Zimmermann, 2014; Yoshida *et al.*, 2016). Traditional enrichment methods are often used for this purpose. Such approaches are useful, but often distorted due to a limited number of culturable organisms under laboratory conditions. Therefore, the resulting novel organisms and enzymes are often taxonomically closely related, as shown by the well-described *Actinobacteria* PET hydrolase examples from *Thermobifida fusca* or *Thermomonospora curvata*.

The aim of this work was to (i) identify new PET-degrading hydrolases, (ii) gain information about the taxonomic diversity of such enzymes, and (iii) combine the results in a global context. To achieve these goals, a combination of enrichment experiments, sequence-based data mining, functional screening of metagenome libraries and functional verification of potential PET hydrolases was used. Based on a HMM search approach, more than 1536 potential PET hydrolases were identified. A subset of four new PET hydrolases were cloned, expressed heterologously in *E. coli* and characterized to verify the HMM for applicability in automated screening procedures. Characterization experiments showed that the enzymes are heat stable or at least heat tolerant. They also tolerate an increased concentration of solvents, detergents and are only slightly inhibited by well-known inhibitors such as EDTA or PMSF, emphasizing their overall stability. The positive results of the functional tests with various model substrates, but also with PET, led to a detailed analysis of the global distribution of PET-degrading hydrolases. An overview of the worldwide occurrence of PET hydrolases could lead to an assessment of the global degradation potential. Therefore, globally distributed metagenome datasets were screened and the resulting output was drawn as hit/MB on a map. The graphic visualization clearly showed that PET hydrolases occur in both terrestrial and marine habitats. The same datasets were used to observe the taxonomic affiliation of PET hydrolases.

4.1 Enrichment of a PET adhering *Comamonas* sp. strain

Enrichment cultures in minimal medium, containing PET yarn as sole carbon source, resulted in a single culture containing a single organism which was highly enriched. The cells formed a biofilm on the PET filaments and couldn’t be found in high yields as planktonic cells within the culture (Figure 7). The observed biofilm had high similarity to the described biofilm formation of *I. sakaiensis* (Yoshida *et al.*, 2016). The as *Comamonas* sp. identified organism showed adhesion to the PET yarn through filamentous structures which were not further elucidated. A

draft genome was sequenced and searched with the HMM for PET hydrolase coding genes. No PET hydrolase was found. A KEGG database comparison revealed that other *Comamonas* species (e.g. *Comamonas* sp. strain E6) are able to take up terephthalic acid via a special tripartite tricarboxylate transporter (TTT) and metabolize it further as sole carbon source (Hosaka *et al.*, 2013; Sasoh *et al.*, 2006). The obtained draft genome might not cover the responsible genes. A PET hydrolase with a different sequence is highly unlikely, but also possible. A metagenomic library of the isolated *Comamonas* sp. strain and a functional screening could provide deeper insights, unfortunately, with PET as the substrate, no high throughput function screening is established. Growth due to tarephthalic acid, which was washed out of the yarn, is the most promising explanation so far.

4.2 Function vs. sequence based screening of novel PET hydrolyzing enzymes from metagenomes and databases

In order to avoid the limitations of traditional cultivation experiments, a functional and sequence-based screening of metagenome libraries was performed. Function-based screening revealed a single fosmid clone that was active on DEP (3.2). DEP was used because of its affiliation to the group of phthalates and its structural similarity with terephthalic acid, which is a major component of PET. Previous studies also showed that DEP can leak into water as it is often an additive for PET bottles (Sax, 2010). DEP is also much easier to use in plate assays and less expensive than PET nanoparticles. Therefore, the first screening run was performed on LB plates containing DEP. After subcloning the protein-encoding gene and heterologous expression in *E. coli*, activity tests using first PCL and further PET showed no activity for this enzyme (3.2). Screening on PCL containing LB plates revealed not a single clone in the existing fosmid libraries (2.3). This poor yield of positive hits was the first indicator for the rather low abundance and probably heterologous expression problems of PET hydrolases. For further screening, a self-constructed HMM was used (2.5). Sequence-based screening provides a convenient and fast way to find sequence homologs in metagenomic datasets and entire sequence databases. The workflow shown in (Figure 5) was used to obtain (i) a PET hydrolase specific HMM and (ii) several novel hydrolases for the depletion of PET. A total of 504 potential PET hydrolase sequences were obtained from databases (3.3.2) and 349 sequences were found in IMG metagenome datasets (Table A1 and Table A2). It is important to mention that the metagenomic sequences could be redundant and may be a subset of the database hits, also different metagenome datasets can have identical sequences. Nonetheless, the sequences found increase the total amount of such enzymes. The found sequences share relatively high similarity to each other due to the search mechanism. It is considered a general problem in

sequence based search approaches that the output is often very similar to the sequences used as reference, thus no completely new sequence variants can be obtained by this method. Still, the sequence based search is a powerful tool to obtain a large amount of new enzymes and enzyme variants for specific tasks (Gullert *et al.*, 2016; Podar *et al.*, 2015). In case of PET hydrolases, the search showed clearly that a multitude of novel sequences can be found in public sequence databases like NCBI and UniProt. Metagenome datasets proofed to be an adequate platform as well.

The finally used 177 terrestrial and 172 marine metagenome hits showed in both cases a maximum similarity within the datasets of >99%. Most genes were annotated as α/β-hydrolases. The minimal similarity was only 70% and 0% respectively, indicating a large diversity within the marine data set. The same results were presented by the calculated mean similarity, which was almost three times smaller for marine metagenomes than for terrestrial metagenomes (0.30 and 0.88, respectively). Comparing the functional and the sequence-based screening method, it is clear that the output of the sequence-based approach is much higher. Restrictions are caused by the fact that only sequences of high similarity are obtained, as observed for sequences obtained from terrestrial metagenomes. In the case of this analysis, sequences of low similarity are biased due to their length. Computing the similarity requires sharing sequence similarities by a sequence length. In this case, the length of the smallest sequence was chosen. Nonetheless, the data obtained suggest that marine PET hydrolase sequences are much more variable than terrestrial ones. A similar effect can be shown at the taxonomic level. Database-derived sequences showed even higher similarity compared to each other when derived from the terrestrial metagenome sequences. Calculated mean similarity was more than 90%, probably caused by sequences from different organisms possessing completly identical PET hydrolases on amino acid level. For example the HMM detected PET hydrolases from *Thermobifida fusca* (G8GER6) and the homologous enzyme in *Burkholderia cepacia* (J7FCD9) share 100% amino acid sequence identity, although these species belong to *Actinobacteria* and *Proteobacteria* respectively. The complete equality of these two PET hydrolases strongly suggests a horizontal gene transfer. At a minimum, these two species must have received the gene from a third species or exchanged it among themselves. In addition, a highly divergent subset was obtained that was less similar to the other sequences and even single sequences that differed greatly from the rest of the data set. Although these enzymes have been found with the PET-hydrolase-specific HMM, it is possible that they have no or lower activity than the others, which share more sequence similarity with each other and in particular with the already described and functionally tested PET-hydrolases. A functional verification of all database and metagenome associated hits

would have been cost intensive and time consuming, therefore the possibility of false positive sequences were considered.

4.2.1 Analysis of database-derived sequence data and unraveling the taxonomic diversity of PET hydrolases

Besides the main goal of this work to obtain novel hydrolases capable to convert PET, it was as well of importance to search for enzymes from different taxa/clades. Taxonomical different enzymes may as well be different in their properties. Therefore, a subsequent taxonomic analysis of the metagenome-derived sequences was performed. The obtained putative PET hydrolase sequences were compared against the NCBI non-redundant database and taxonomy was specified using Megan. The obtained taxonomical distribution showed clearly that most of the enzymes, namely more than 96%, are associated to the domain of Bacteria. The remaining sequences are related to Archaea (2.9%) and only three sequences in total belonged to the domain of *Eukaryota*. Taking into consideration that solely bacterial sequences were used for the HMM construction the described distribution is not surprising. Considering the non-bacterial sequences as negligible within this analysis, efforts were taken to get a deeper look into the bacterial taxonomy. The sequence analysis showed nine bacterial phyla (*Bacteroidetes*, *Proteobacteria*, *Actinobacteria*, *Acidobacteria*, *Firmicutes*, *Cyanobacteria*, *Verrucomicrobia*, *Spirochaetes*, *Lentisphaerae*) to be present in total. Representatives of all nine phyla were found within the marine dataset, but only five phyla (*Bacteroidetes*, *Proteobacteria*, *Actinobacteria*, *Acidobacteria*, *Firmicutes*) were shown to be present in the terrestrial dataset suggesting that the marine PET hydrolase community is more diverse to some extent. The missing representatives make up only a small fraction of the total amount. The main phyla are found in both datasets. The distribution within the two datasets is except for the missing phyla nearly identical. Especially the phyla *Acidobacteria* and *Proteobacteria* are represented nearly equally in both datasets. Exceptions are the phyla *Bacteroidetes* and *Actinobacteria*. While in the marine dataset *Bacteroidetes* sequences are the main representatives, in the terrestrial dataset *Actinobacteria* sequences are the most abundant. Even the percentage with respect to all sequences in the corresponding record is exactly the same in both cases. Therefore the ratio for *Bacteroidetes* to *Actinobacteria* sequences within both datasets is 1:1. The taxonomic analysis showed in this case clearly that PET hydrolase sequences are not exclusively found within the phylum of *Actinobacteria* but also in other phyla like for example *Acidobacteria* or *Proteobacteria*.

Earlier examples of bacteria possessing PET hydrolases were exclusively Gram-positive and mostly, as already mentioned, from *Actinobacteria*. The metagenomic data analysis revealed a multitude of other phyla possessing the potential to hydrolyze PET and not only in terrestrial systems but also in aquatic or at least marine habitats. The land adapted Gram-positive organisms of the phylum *Actinobacteria* (belonging to the group of *Terrabacteria*) are dominating the terrestrial habitats(, whereas the Gram-negative Bacteroidetes (belonging to the group of *Hydrobacteria*) are dominant in marine habitats. The consistent relationship between these two phyla in the analyzed datasets and their habitat preferences may implicate that organisms of one phylum take over the task of the other in their respective habitat.

Deeper looks into the taxonomic diversity of the NCBI non-redundant database derived PET hydrolase sequences (3.3.2) showed that *Actinobacteria* affiliated sequences are by far the most abundant with 1045 representatives. The second most are *Proteobacteria* associated. Although during the time of this study only a single *Proteobacteria* representative was described, which was isolated from PET contaminated soil (Yoshida *et al.*, 2016). These findings are to some extent consistent with the metagenomic analysis of terrestrial habitats. *Bacteroidetes* sequences were much rarer. Only 13 sequences were found in total accounting to slightly more than 1%. Beside seven not assignable sequences there were only representatives of *Cyanobacteria* (1) and *Deinococcus-Thermus* (4) within the database-derived sequences, which is an even more drastic reduction of sequence information. Based on this data it seems like the metagenome and especially the marine taxonomical distribution are not reflected by database derived sequences. Although, far more marine than terrestrial metagenome datasets are available, for example on IMG (as of May 7, 2018), PET hydrolases are mostly described from soil associated *Actinobacteria*. Taking the metagenomic and the database-derived data into account it is questionable if previous efforts failed to identify the marine *Bacteroidetes* representatives. It is also possible that the PET hydrolases show only poor or no activity at all, so that it was not possible to find these enzymes with functional screening methods. Nevertheless, the analysis revealed a much larger taxonomic diversity of these enzymes and a close relationship on amino acid sequence level throughout the phylogenetic tree of PET hydrolase possessing organisms.

4.2.2 Worldwide distribution of PET hydrolases within terrestrial and marine metagenomes

As addition to the taxonomic analysis of metagenome derived PET hydrolases a direct link to the geographical distribution was made, in order to determine hot spots for PET hydrolase activity. Figure 12 represents the obtained geographical metadata for the used metagenome datasets.

PET hydrolases seem to be ubiquitously distributed in a multitude of different marine and terrestrial habitats. Such enzymes are found in tropical regions as well as in regions next to the poles, which may indicate a broad temperature range of such enzymes, which was partially shown with the performed temperature optimum assay (3.4.2). Unfortunately, the obtained metagenome datasets don't give a complete insight into the worldwide occurrence since some regions (e.g. Africa, Asia and South America) were completely left out in this analysis due to a lack of data. Still, the normalized and visualized data is sufficient enough to suggest that PET hydrolases appear at very low frequency whereby it is important to mention that the data gives no information about the expression level of PET hydrolase genes. The overall low gene frequency might suggest that this trait was not able to spread to a higher rate within bacterial life. Considering that PET is mass produced only since the early 1970's it is possible that evolution had not enough time to benefit PET hydrolase expressing bacteria. Compared to the much faster environmental degradation of natural polymers like cellulose or starch (Chen, X. L. *et al.*, 2011; Lopez-Mondejar *et al.*, 2016), PET depletion is limited not only by the very low turnover rates but also by their rareness in natural habitats. In contrast to this observation, one highly contaminated sample site showed to have much more PET hydrolase possessing organisms. A sampling spot in Canada revealed a much higher frequency of such genes than any other included dataset. The sample was collected from a sediment core of a heavy oil reservoir (IMG genome ID: 3300001197). Not only is the frequency the highest among all tested datasets, but also the total number of hits is more than 20 fold higher than the average number of hits for marine datasets and more than 30 fold higher than the average number of hits for terrestrial datasets. The PET monomers ethylenglycole and terephthalic acid are to some extent derived from crude oil, therefore an enrichment of such organisms in crude oil associated habitats is plausible. A direct link between oil contamination and PET hydrolase occurrence could be explained by determining the native substrate of these enzymes.

4.3 Biochemical characterization of PET hydrolases found by sequence mining

After analysis of the obtained sequence information, four promising sequences were chosen for heterologous expression of the respective PET hydrolases in *E. coli*. The PET hydrolases PET2, PET5, PET6 and PET12 were purified and partially tested for their biochemical characteristics. A special focus was set on PET2 and PET6.

4.3.1 Cloning and expression of highly toxic PET hydrolases

During the expression experiments, the partially high toxicity of PET hydrolases was observed. In some cases the induction of the PET hydrolase encoding genes caused reduced growth rates or even death of the culture, with only poor yields of the desired protein. The use of autoinduction medium (2.6.2) proved to be beneficial for expression. Initially, the glucose-containing medium suppressed the production of PET hydrolases which was under control of the lactose promotor of *E. coli*. Therefore, the cells could grow without stress caused by the toxic protein. After complete glucose metabolism, protein production was induced by lactose. This approach resulted in the production of up to 10 mg of PET hydrolase per liter of cell culture. Conventional induction with IPTG has probably failed due to a non-fully adjustable expression system. Without glucose, the PET hydrolase genes were transcribed basally and therefore small amounts of the protein were produced. These small amounts of protein appeared to disrupt the growth of the *E. coli* host used. For even higher yields of active protein, the temperature was reduced during induction (de Groot & Ventura, 2006; Rosano & Ceccarelli, 2014).

4.3.2 Functional analysis revealing highly promiscuous and stable hydrolases

Since PET hydrolases in general are secreted to the environment (Bhardwaj *et al.*, 2013; Lucas *et al.*, 2008; Sivan, 2011) and act there on their substrate, an overall stability of these enzymes would be a clear advantage. Temperature optimum and stability against a variety of chemical compounds were one of the major parameters considered during analysis. The broad temperature range between 17 °C and 70 °C supported the previously observed ubiquitous distribution of such enzymes. Other sequence homologous proteins showed a similar temperature range (Hu *et al.*, 2010; Ronkvist *et al.*, 2009; Wei, Oeser, Then*, et al.*, 2014) Considering PET as substrate it is mandatory for the proteins to withstand higher temperatures, for example during industrial processes of degradation or modification. Due to structural changes within the substrate, at elevated temperatures, PET gets more amorphous and hence provides degradable regions for the protein. Also of major importance is the durability of the protein at such high temperatures. PET2 showed no reduction in activity at 60 °C for up to 6 hours (Figure 17). The combination of stability and increased activity at temperatures above 50 °C may mark PET2 as best suited for industrial applications. Obtained data for substrate specificities showed that at least PET6 is very promiscuous in terms of carbon chain length. The reduction of PET2 activity on longer chain length substrates (>C8) may indicate a higher specifity for a different substrate than PET. The identification of the native substrate is a nearly

non-exploited topic. Most PET hydrolases are assigned as cutinases (EC 3.1.1.74) which might be the original substrate of those enzymes (Acero *et al.*, 2011; Kanelli *et al.*, 2015; Ronkvist *et al.*, 2009; Sulaiman *et al.*, 2012; Thumarat *et al.*, 2012). Other possible targets may are aromatic substances from crude oil (Danso *et al.*, 2018). A detailed description of the regulation mechanism is still missing so far. Since these enzymes are almost exclusively secreted, it is possible that they evolved to be very stable at a multitude of different abiotic conditions. The influence of solvents showed that concentrations up to 10% reduced activity only by up to 16% (Figure 22). Higher concentrations (e.g. 30%) acted much severer on the activity. Since such high concentrations of solvents are not unlikely during an industrial enzymatic process, the solvent stability may have to be improved for both PET hydrolases (PET2 and PET6). Assays with known inhibitors showed clearly that PET2 as well PET6 are significantly influenced by PMSF, since PMSF binds covalently to the serine within the catalytic triad. Other inhibitors showed only minor or no effect at all (Figure 20). As for detergent stability, both enzymes were strongly inhibited by Triton X-100. The inhibiting effect in this particular case was not further investigated, but Triton X-100 has in fact a hydrophilic ether bond connected polyethyleneglycole sidechain which may interfere with the active site of PET hydrolases and therefore minimizes the overall activity (Fernandez-Lorente *et al.*, 2007). Although a connection between Ca^{2+} ions and PET hydrolase activity was clearly shown by others (Then *et al.*, 2015; Then *et al.*, 2016), within this work no clear effect was observed in case of PET2 and PET6. The redundancy of additional metal ions would be an economical advantage considering industrial large scale processes of PET degradation by such enzymes. Beside temperature stability, pH stability was the other major parameter for a stability assessment of the new PET hydrolases PET2 and PET6. The data is consistent for both enzymes, showing that a slightly alkaline milieu of pH 8.0 is preferred. The activity drops to zero at acidic pH below 6.0. Therefore, it is possible that PET2 and PET6 might be inhibited due to acidification by the PET degradation product terephthalic acid at high concentrations. Both enzymes showed good activity and stability under a variety of different conditions making them suitable candidates for an industrial process. A general limitation regarding these proteins is their very low turnover rate when used on PET. Efforts to improve the kinetic properties of PET hydrolases have been attempted by others, leading to promising variants with higher turnover rates (Austin *et al.*, 2018; Fecker *et al.*, 2018). The not further described PET-hydrolases PET5 and PET12 were both badly produced in *E. coli* (Figure 13). A closer look at these enzymes, however, may reveal potential biocatalysts for biotechnological processes. One could consider using the opposite reaction of these enzymes as well. An esterification reaction of monomers to short oligomers or large polymers was already shown for

other plastic degrading enzymes like for example lipases (Hunsen *et al.*, 2007; Kobayashi, 2009).

4.4 *in silico* analysis of the new PET hydrolase PET2

After the successful screening and functional testing of the new PET hydrolases, PET2 became the most interesting and promising enzyme. Its high stability at elevated temperatures and under influence of a multitude of solvents, inhibitors and detergents were considered as desirable characteristics for an industrial application (see Figure 16 – Figure 22). To have even better insight into the function of this protein, the structure prediction and the docking studies were performed. Like other PET hydrolases PET2 showed an α/β-hydrolase like fold without a lid, which is commonly observed for lipases (Kitadokoro *et al.*, 2012). Instead, PET2 showed a cleft on the surface next to the catalytic triad enabling the enzyme to bind such a large polymer like PET. Binding studies with the model substrate BHET showed clearly that aromatic compounds can be bound to the active site of the protein, which is consistent with the functional substrate tests. During the sequence based analysis of PET2 it was found that PET12 has the highest similarity to the PETase from *I. sakaiensis* (Figure 24). Since the PETase is able to hydrolyze PET at room temperature (Yoshida *et al.*, 2016), it might be possible that the newly identified PET hydrolase PET12 from P. *brachysporum* can be used in the same way.

5 Conclusions and outlook

The combination of functional and sequence-based approaches identified novel, to a certain extent stable and promiscuous PET hydrolases from a variety of bacterial organisms. Bioinformatics tools have been used to perform a global analysis of such enzymes that could help to gain a more detailed picture of the actual degradation potential of PET in nature. In addition to new findings, there were also new questions. For example, regulation of PET hydrolases is poorly understood. A closer look at the upstream and downstream regions in multiple organisms has in many cases revealed a conserved gene structure (Figure 26).

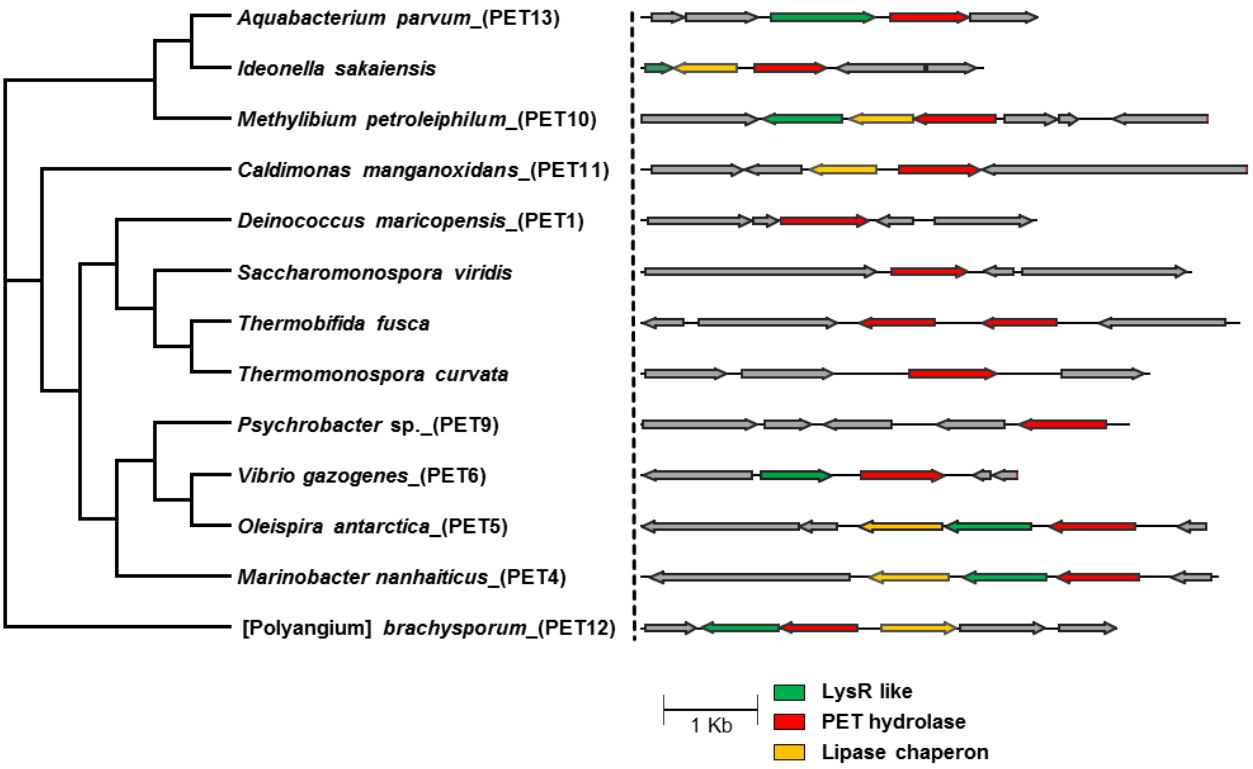

Figure 26: PET-hydrolase flanking genes of PET-degrading bacteria, which were found in this study. In addition to the PET hydrolases, the newly identified PET-degrading bacteria also have a lipase chaperone encoding and a LysR-like encoding gene in the immediate vicinity.

In addition to the PET hydrolases, an ORF assigned as chaperone and a LysR-like coding open reading frame (ORF) can be found. The presence of a transcriptional regulator in the vicinity implies a tight control of the expression of PET hydrolase genes. Similar to other LysR-like transcriptional regulatory proteins, a substrate binding and a DNA binding domain can be found, suggesting the regulation of PET hydrolase gene transcription by a soluble molecule. Since PET itself is not soluble, but its degradation products, one or more of the possible PET hydrolase

products appear to be the best candidates for interaction with the LysR-like protein. For future work on this topic, a detailed analysis of the LysR-like protein and its interaction partners (DNA and substrate) could provide insight into the regulation and even the native substrate of PET hydrolases (Maddocks & Oyston, 2008; Uchiyama *et al.*, 2005). In addition, the mechanism may be used to construct a reporter strain useful in the functional identification of PET degrading enzymes.

Further work may also focus on a better understanding of the interaction between the enzyme and the substrate surface. Only few studies tried to reveal the direct interaction of PET hydrolases with their substrate and the molecular mechanism (Acero *et al.*, 2011; Austin *et al.*, 2018; Joo *et al.*, 2018; Roth *et al.*, 2014). Modelling studies could be the first step to not only understand the polymer enzyme interaction but could also help to identify other enzymes which could have a synergistic effect on the degradation, for example by modifying the substrate surface or enhance activity by additional hydrolyzing steps of intermediate substances and therefore eliminating product inhibition of PET hydrolases (Barth *et al.*, 2015; Carniel *et al.*, 2017; Espino-Rammer *et al.*, 2013; Perz *et al.*, 2015; Ribitsch *et al.*, 2015; Ribitsch *et al.*, 2013). While the two above mentioned points concerns aspects of basic research it will also be of great interest to identify robust PET hydrolases, that can be implemented in biotechnological applications and if possible be used to solve some of the problems affiliated with the large level of PET contamination in nature. As outlined above PET nanoparticles are present in nearly all natural habitats. They represent a large concern with respect to a potential impact on human and animal health . Thus it is highly tempting and at the same time it will be rewarding to identify or develop bacteria and enzymes to address this pressing environmental problem.

6 Appendix

>EU660533.1:828-1754_Uncultured_bacterium_(PET2)

ATGCCAATAACCGCCAGAAACACACTGGCCTCGTTGCTGCTTGCGTCTTCCGCATTGCTGCTCTCCGGAA
CCGCGTTCGCCGCCAACCCGCCGGGGGGCGATCCCGATCCCGGCTGCCAGACCGACTGCAACTATCAGCG
CGGCCCCGATCCGACCGACGCCTACCTCGAAGCCGCGAGCGGACCGTACACCGTGAGCACCATCCGCGTT
TCCAGCCTGGTGCCCGGCTTCGGCGGCGGCACCATCCACTACCCGACCAATGCCGGCGGCGGCAAGATGG
CTGGCATCGTGGTCATTCCGGGCTATCTGTCGTTCGAGTCGTCGATCGAATGGTGGGGACCGCGGCTGGC
TTCGCACGGCTTCGTGGTCATGACCATCGACACCAACACCATCTATGACCAGCCGAGCCAGCGCCGCGAC
CAGATCGAGGCCGCCCTGCAGTACCTGGTCAACCAGAGCAACTCCTCCTCCAGCCCCATCAGCGGCATGG
TCGACAGCTCGCGCCTGGCGGCGGTCGGCTGGTCGATGGGCGGCGGCGGGACGCTGCAACTGGCCGCCGA
CGGCGGAATCAAGGCCGCCATCGCCTTGGCGCCATGGAACTCCAGCATCAACGACTTCAACCGCATCCAG
GTGCCGACCCTGATCTTCGCCTGCCAGCTGGATGCCATCGCGCCAGTGGCCTTGCATGCCTCGCCGTTCT
ATAACCGCATCCCGAACACGACCCCCAAGGCCTTCTTCGAGATGACCGGCGGGGACCACTGGTGCGCCAA
CGGCGGCAACATCTACAGCGCGCTGCTCGGCAAGTACGGCGTGTCTTGGATGAAGCTGCACCTCGACCAG
GATACGCGTTACGCGCCGTTCCTGTGCGGTCCCAACCACGCGGCTCAGACCCTGATCTCCGAATACCGAG
GTAACTGCCCGTACTGA

>ACC95208.1_uncultured_bacterium_(PET2)

MPITARNTLASLLLASSALLLSGTAFAANPPGGDPDPGCQTDCNYQRGPDPTDAYLEAASGPYTVSTIRV
SSLVPGFGGGTIHYPTNAGGGKMAGIVVIPGYLSFESSIEWWGPRLASHGFVVMTIDTNTIYDQPSQRRD
QIEAALQYLVNQSNSSSSPISGMVDSSRLAAVGWSMGGGGTLQLAADGGIKAAIALAPWNSSINDFNRIQ
VPTLIFACQLDAIAPVALHASPFYNRIPNTTPKAFFEMTGGDHWCANGGNIYSALLGKYGVSWMKLHLDQ
DTRYAPFLCGPNHAAQTLISEYRGNCPY

>FO203512.1:861459-862391_Oleispira_antarctica_strain_RB-8_(PET5)

TTAGTAATTACATGTGTCACGGTATTCTGAAACAGAACGATTGGCTGCATGATCAGGACCACATAAGAAC
GCATCGTAGCGTAAGTCATTATCAATAAAGCGCTTCATCCATGAGACACCATATTTGCCTAACAGTCCAC
CGATTGAGCCGCCATCATTTGCACAGAAGTGACTGCCACCATTAATTTCCAAATACGCTTTGTCTGTAGT
GCTAGGGATGCGGTTATAAAAGGGTGAAGCATGGCTGTTGACGGATGCTACAACATCATTTTCACAGGCA
ATAACGAGGGTAGGTGTTTCAATTTGATCAAAGCGGTTGCCACCTTGATTCCAGGGAGCTATTGGAATTG
CTGCGCTTAAACGGTCTCCAGAGGCAAGTTGAAGTGCACCACCTCCTCCCATTGACCAGCCAACAACACC
CAAGCGAGTTGTATCGACCATTCCAGAGATTGGACTATTGCTGCCATTACTTTGATTAATGACATAATCA
AGAGCACGACCTAATTGAGTTGCACGACTAGCAGGCTGGTCAAACCCTGAATTAGCTGAAATCGTAATGA
CTACAAACCCGTGAGAAGCAAGTTTAGGTCCCCAGAAGTCAATAGAGGATTCTTGCAGCAAAAAGCCAGG
AACAACAGCAATCGCGCCCATGGTTCCTGTTGTGTTAGTCGGGTAGTGAAGTGTACCTCCACCAAAGCCA
CTGACTGAACTTGCAACGCTTCGAGTCGCAACGGAATACGGGCCAGTAGAAGCTTCTAAAGAAGAGGGCG
TAGGATTTGGTCCGCGAGTAAAGTCACAATCATCTTGGCATGGTGTTGGGTCAGGTATTGGATCTGGGTT
AGGTGTTGGGCTTGGCGTCCAGCTTAATGCATTCATGGATACGAGTAGTACACTTGTACCGAAAGAGAGT

TTTTTTAGAATAGATTTATTCAC

>CCK74972.1_Oleispira_antarctica_RB-8_(PET5)

MNKSILKKLSFGTSVLLVSMNALSWTPSPTPNPDPIPDPTPCQDDCDFTRGPNPTPSSLEASTGPYSVAT
RSVASSVSGFGGGTLHYPTNTTGTMGAIAVVPGFLLQESSIDFWGPKLASHGFVVITISANSGFDQPASR
ATQLGRALDYVINQSNGSNSPISGMVDTTRLGVVGWSMGGGGALQLASGDRLSAAIPIAPWNQGGNRFDQ
IETPTLVIACENDVVASVNSHASPFYNRIPSTTDKAYLEINGGSHFCANDGGSIGGLLGKYGVSWMKRFI
DNDLRYDAFLCGPDHAANRSVSEYRDTCNY

>CP018835.1:727938-728834_Vibrio_gazogenes_strain_ATCC_43942_(PET6)

ATTATGAATGTTTTAACAAAATGTAAGCTAGCATTGGGAATCATAGCGATATTTTTTTCATTGCCAAGTT
TTGCGGTACCGTGTTCGGACTGTTCAAATGGATTTGAACGGGGCCAAGTTCCTCGGGTTGATCAGTTAGA
ATCCAGTAGAGGACCTTATTCGGTAAAAACTATCAACGTTTCTAGACTTGCCAGAGGGTTTGGTGGTGGC
ACGATCCATTACTCGACGGAATCAGGAGGCCAGCAAGGTATCATTGCGGTCGTTCCGGGCTATGTTTCTT
TAGAAGGTAGTATTAAATGGTGGGGCCCACGTTTGGCATCTTGGGGCTTTACGGTGATTACGATCGATAC
AAATACGATTTATGATCAACCTGACAGTCGAGCAAGTCAGCTCAGTGCAGCAATCGATTATGTCATTGAT
AAAGGTAACGATCGAAGCTCTCCGATTTATGGATTAGTGGATCCAAATCGAGTCGGCGTCATTGGCTGGT
CTATGGGCGGTGGCGGCTCCCTCAAACTGGCAACAGACAGAAAGATTGATGCAGTGATCCCTCAAGCCCC
TTGGTATCTTGGATTGAGTCGTTTTTCTAGTATTACTTCACCAACGATGATCATTGCATGCCAAGCCGAT
GTTGTGGCACCAGTGAGCGTACACGCTTCTCGCTTTTATAATCAAATCCCGGGAACCACCCCAAAAGCAT
ATTTTGAAATTGCGTTAGGCAGTCATTTTTGTGCCAATACAGGGTACCCTAGCGAAGATATCTTGGGGCG
GAATGGTGTTGCATGGATGAAGCGCTTTATTGATAAAGATGAACGCTATACTCAATTTTTATGTGGACAA
AATTTTGATAGCAGTCTTCGCGTCAGTGAATACCGAGACAACTGTAGTTACTATTAG

>ASA57064.1_Vibrio_gazogenes_(PET6)

MMNVLTKCKLALGIIAIFFSLPSFAVPCSDCSNGFERGQVPRVDQLESSRGPYSVKTINVSRLARGFGGG
TIHYSTESGGQQGIIAVVPGYVSLEGSIKWWGPRLASWGFTVITIDTNTIYDQPDSRASQLSAAIDYVID
KGNDRSSPIYGLVDPNRVGVIGWSMGGGGSLKLATDRKIDAVIPQAPWYLGLSRFSSITSPTMIIACQAD
VVAPVSVHASRFYNQIPGTTPKAYFEIALGSHFCANTGYPSEDILGRNGVAWMKRFIDKDERYTQFLCGQ
NFDSSLRVSEYRDNCSYY

>NZ_CP011371.1:2790023-2790919_[Polyangium]_brachysporum_strain_DSM_7029_(PET12)

ATGCCACCGGATTGCGTCCTTCCCCGCCGCCTCGCCGCGGCCGCGCTCCTCGCGAGCGCG
ACACTTGTGCCGCTGTCGGCCGCGGCTCAGACCAACCCCTACCAGCGAGGCCCGGACCCG
ACGACCCGCGACCTGGAAGACAGCCGAGGGCCGTTCCGCTACGCCAGCACCAACGTGCGC
TCGCCCAGCGGTTATGGCGCGGGCACGATCTACTACCCCACCGACGTCAGCGGCAGCGTC
GGCGCGGTGGCGGTGGTGCCCGGCTACCTCGCCCGCCAGTCCAGCATCCGCTGGTGGGGG
CCGCGCCTGGCCTCGCACGGGTTTGTCGTCATCACCCTCGACACCCGCTCGACCTCCGAC

CAGCCCGCCAGCCGTTCCGCTCAACAGATGGCGGCGCTGCGGCAGGTGGTGGCGCTCAGC
GAGACGCGCAGCAGCCCGATCTACGGCAAGGTCGATCCCAACCGTCTCGCGGTGATGGGC
TGGTCGATGGGCGGAGGCGGCACGCTGATCTCTGCGCGCGACAACCCCAGCTTGAAGGCC
GCCGTGCCATTCGCCCCGTGGCACAACACCGCCAACTTCTCGGGCGTGCAAGTGCCGACG
CTCGTGATCGCTTGCGAAAACGACACCGTCGCCCCGATCTCGCGACATGCCTCGTCCTTC
TACAACAGCTTTTCGAGCTCGCTCGCCAAGGCCTATCTCGAGATCAACAACGGCTCGCAC
ACCTGCGCCAACACCGGCAACAGCAACCAGGCCCTGATCGGCAAATACGGTGTGGCGTGG
ATCAAGCGCTTCGTCGACAACGACACACGCTACAGCCCCTTCCTCTGCGGCGCACCACAT
CAGGCCGACCTGCGCAGCAGCCGCCTCTCGGAGTACCGCGAGAGCTGCCCGTACTGA

>WP_047194864.1_[Polyangium]_brachysporum_(PET12)

MPPDCVLPRRLAAAALLASATLVPLSAAAQTNPYQRGPDPTTRDLEDSRGPFRYASTNVRSPSGYGAGTI
YYPTDVSGSVGAVAVVPGYLARQSSIRWWGPRLASHGFVVITLDTRSTSDQPASRSAQQMAALRQVVALS
ETRSSPIYGKVDPNRLAVMGWSMGGGGTLISARDNPSLKAAVPFAPWHNTANFSGVQVPTLVIACENDTV
APISRHASSFYNSFSSSLAKAYLEINNGSHTCANTGNSNQALIGKYGVAWIKRFVDNDTRYSPFLCGAPH
QADLRSSRLSEYRESCPY

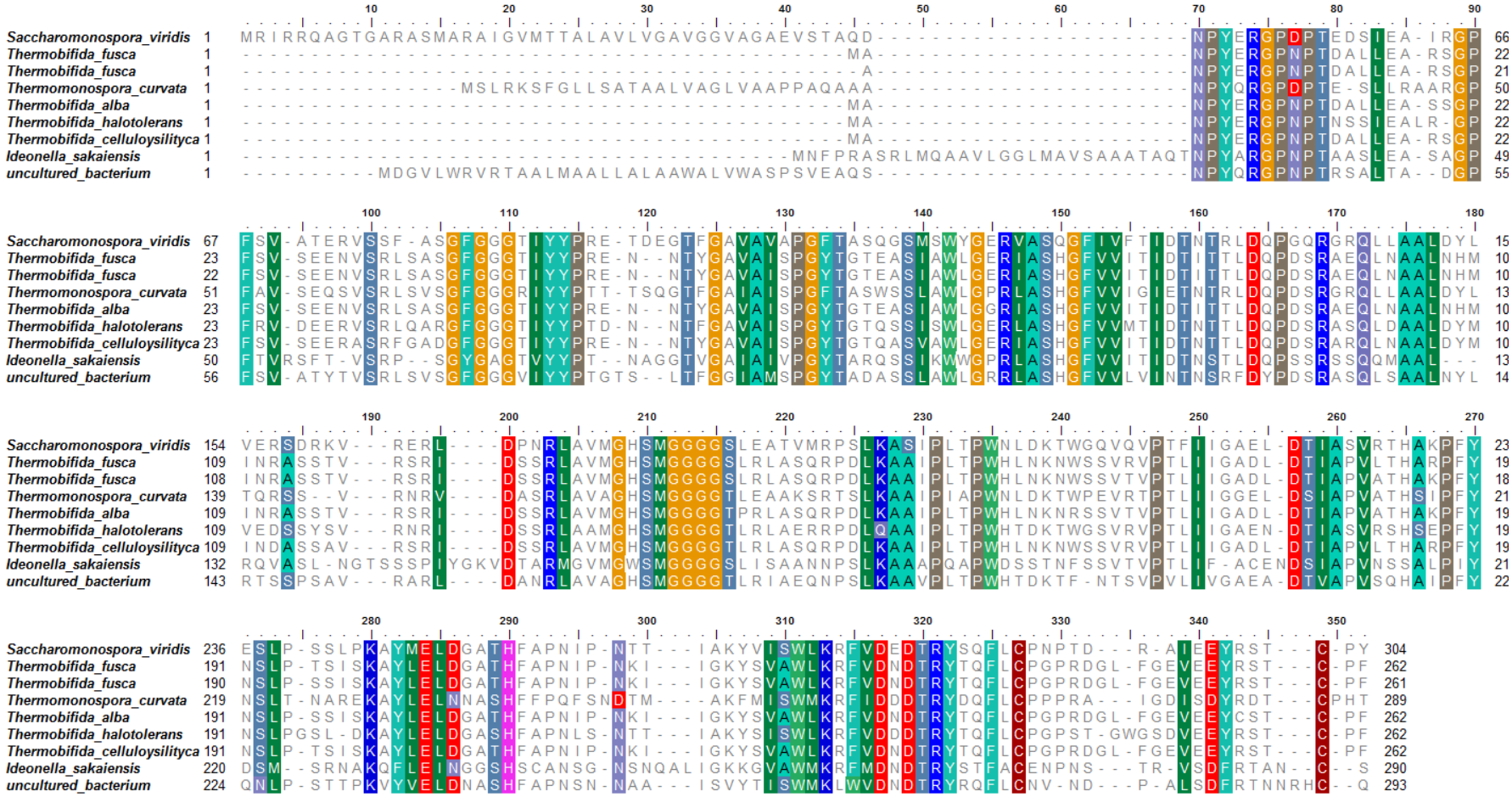

Figure A 1: Alignment for the construction of a PET hydrolase specific HMM. Identical and/or similar residues have a colored background.

Table A 1: Terrestrial metagenomes used for sequence mining.

Terrestrial metagenomes

IMG Genome ID	Sample Name	Latitude	Longitude	Genome Size [bp] (assembled)	Gene Count (assembled)	Hits	Hits/Gb
2070309004	Green-waste compost microbial communities at University of California, Davis, USA, from solid state bioreactor - Luquillo Rain Forest, Puerto Rico	18.311389	-65.8375	281,407,472	781,631	1	0.004
2119805012	Soil microbial communities from sample at FACE Site NTS_067 Nevada Test Site (NTS_067)	36.766667	-115.95	390,792,833	1,036,364	14	0.036
2140918008	Permafrost microbial communities from permafrost in Bonanza Creek, Alaska - Bog_all (Bog_all_CLC)	64.7	-148.3	410,522,816	696,868	5	0.012
2162886008	Soil microbial communities from Puerto Rico rain forest, that decompose switchgrass - Feedstock-adapted consortia SG + Fe (SG + Fe, May 2011 assembly)	18.3724	-65.7166	154,120,208	193,491	2	0.013
2209111000	Soil microbial communities from Colorado Plateau, Greene Butte sample - Dark Crust, Colorado Plateau, Green Butte (Dark Crust, Colorado Plateau, Green Butte June 2011 assem)	38.714972	-109.692944	396,718,301	1,058,313	7	0.018
3300000596	Amended soil microbial communities from Kansas Great Prairies, USA - Total DNA no BrdU F1.4TC (Kansas native prairie Total DNA no BrdU F1.4TC, April 2012 Assem)	39.100992	-96.608258	118,917,357	353,727	1	0.008
3300000956	Soil microbial communities from Great Prairies - Kansas, Native Prairie soil (Kansas, Native Prairie soil, Sept. 2012 Assem JGI Velvet)	39.214012	-96.585283	9,280,824,611	31,876,498	1	0.000
3300001197	Wastewater microbial communities from Syncrude, Ft. McMurray, Alberta - Microbes from Sediment core from a heavy oil reservoir,Alberta Canada Inniskillen 614.3 (Inniskillen 614.3: 454+illumina sequencing assembly)	56.04	-118.13	89,229,498	204,944	135	1.513
3300002468	Deep subsurface microbial communities from Mt. Terri Underground Rock Laboratory, Switzerland - 10_samples_coassembly (concoct_output_2nd_run)	47.379	7.1648	143,773,111	146,707	2	0.014
3300005258	Microbial communities on the surface of bentonite enhanced biochar (D2B)	-33.917926	151.235347	162,951,842	282,734	4	0.025

Terrestrial metagenomes

IMG Genome ID	Sample Name	Latitude	Longitude	Genome Size [bp] (assembled)	Gene Count (assembled)	Hits	Hits/Gb
3300005260	Microbial communities on the surface of kaolinite enhanced biochar from soil with fertiliser in Sydney, Australia (F2B)	-33.917926	151.235347	308,772,089	487,956	5	0.016
2044078003	Miscanthus field bulk soil microbial communities from University of Illinois Energy Farm, Urbana, IL (Bulk soil sample from field growing Miscanthus x giganteus)	40.109	-88.204	58,125,848	165,002	5	0.086
2067725009	Permafrost microbial communities from central Alaska, USA - Permafrost field sample	65.7906	-149.9102	9,656,814	13,277	0	0.000
2189573024	Echo Passage metagenome	31.837801	-110.350292	138,499,921	365,407	0	0.000
3300000044	Arabidopsis rhizosphere microbial communities from the University of North Carolina - sample from Arabidopsis soil old (Arabidopsis soil old, Nov 2011 assem)	35.9	-79.05	92,517,998	280,064	0	0.000
3300000597	Forest soil microbial communities from Amazon forest - 2010 replicate II A1 (Amazon Forest 2010 replicate II A1, April 2012 Assem)	-10.171667	-62.7875	321,952,212	768,309	0	0.000
3300000793	Forest soil microbial communities from Amazon forest - 2010 replicate II A001 (Amazon Forest 2010 replicate II A001, March 2012 Assem)	-10.171667	-62.7875	297,027,237	799,655	0	0.000
3300000825	Subaerial biofilm microbial communities from sulfidic caves, Italy, that are extremely acidic - Acquasanta AS5 (Draft assembly, stringent (mi 99 ml 60 rip))	42.755	13.411118	8,590,020	17,561	0	0.000
3300001490	Fosmid Clones Derived from Amazon Forest Soil Microbial Communities (Amazon Soil Fosmids Plate#2, June 2013 Assem)	-2.5871	-49.041	3,532,881	3,698	0	0.000
3300002157	Saline desert soil microbial communities from Kutch, Gujarat, India - S6	23.940972	70.188444	300,833,452	2,224,072	0	0.000
3300003065	Soil viral communities from Rice paddy at Daejon City	36.3875	127.3392	611,313	1,249	0	0.000
3300003102	Soil and Ice psychrophilic microbial communities from Leh Laddakh, India - psychrophilic sample	34.1453972	77.5676139	3,997,411	11,352	0	0.000
3300005699	Soil microbial communities from Charlotte, North Carolina, that are pyrene degrading (re-annotation)	35.2269444	-80.8433333	5,507,702	9,489	0	0.000

Terrestrial metagenomes

IMG Genome ID	Sample Name	Latitude	Longitude	Genome Size [bp] (assembled)	Gene Count (assembled)	Hits	Hits/Gb
3300006427	Microbial communities of the formation fluids of a supercritical CO2 deposit from McElmo Dome, Colorado	39.7640021	-105.135307	10,773,225	12,085	0	0.000
3300006428	Microbial communities of the formation fluids of a supercritical CO2 deposit from Well 3, McElmo Dome, Colorado	39.7640021	-105.135307	26,625,403	30,536	0	0.000

Table A 2: Marine metagenomes used for sequence mining.

Marine metagenomes

IMG Genome ID	Sample Name	Latitude	Longitude	Genome Size [bp] (assembled)	Gene Count (assembled)	Hits	Hits/Gb
3300000555	Wetland microbial communities from Twitchell Island in the Sacramento Delta, sample from surface sediment Feb2011 Site A1 Bulk (Wetland Surface Sediment Feb2011 Site A1 Bulk-1% Merged Rds,0.27kb Insert)	38.107057	-121.64758	0	0	1	n.d
3300000573	Wetland microbial communities from Twitchell Island in the Sacramento Delta, sample from surface sediment Feb2011 Site A1 Bulk (Wetland Surface Sediment Feb2011 Site A1 Bulk-1% Merged Rds,0.25kb Insert)	38.107057	-121.64758	0	0	1	n.d
3300002359	Wetland sediment microbial communities from Twitchell Island in the Sacramento Delta, sample from surface sediment Aug2011 Site B2 Bulk	38.1072	-121.6485	0	0	44	n.d
3300000444	Marine sediment microbial community from La Parguera, Puerto Rico - PR Tt Sediment 1 (PR Tt Sediment 1 - Bioluminescent Bay in La ParagueraPR, July 2012 Assem)	17.9675	-67.018833	10,852,089	26,822	1	0.0921

Marine metagenomes

IMG Genome ID	Sample Name	Latitude	Longitude	Genome Size [bp] (assembled)	Gene Count (assembled)	Hits	Hits/Gb
3300001939	Marine microbial communities from Block Island, New York, USA - GS009	41.09111	-71.60222	37,747,991	70,271	3	0.0795
3300006559	Marine microbial communities from the Black Sea in Odessa region - Od_3 (Assembly)	46.440968	30.772294	15,417,702	57,359	1	0.0649
3300001935	Marine microbial communities from Northern Gulf of Maine, Canada - GS007	43.63222	-66.84722	32,214,377	54,370	2	0.0621
3300000929	Marine plume microbial communities from the Columbia River - 15 PSU (Columbia River plume metagenome 15 PSU)	46.239	-124.161	448,564,426	1,055,341	26	0.0580
3300003517	Marine microbial communities from Antarctic Ocean - Station_363 -- 3.0 um (Cold Waters_1)	-60.00007	141.23353	86,297,686	253,569	5	0.0579
3300001969	Marine microbial communities from Yucatan Channel, Mexico - GS017	20.5225	-85.41361	174,719,720	322,314	10	0.0572
3300001964	Marine microbial communities from Rosario Bank, Honduras - GS018	18.036667	-83.78472	105,952,902	195,732	6	0.0566
3300001957	Marine microbial communities from Wolf Island, Equador - GS035	1.3891667	-91.81695	88,794,277	150,749	5	0.0563
3300000928	Marine plume microbial communities from the Columbia River - 25 PSU (Columbia River plume metagenome 25 PSU)	46.233	-124.16	465,382,374	999,698	26	0.0559
3300000426	Marine sediment microbial community from Union City, CA, USA - Pond 1C Sediment 3 (Pond 1C Sediment 3 Union City, June 2012 Assem)	37.569083	-122.10327	90,642,744	224,856	5	0.0552
3300001938	Marine microbial communities from Bedford Basin, Nova Scotia, Canada - GS005	44.690277	-63.637222	37,780,382	64,231	2	0.0529
3300000101	Marine microbial communities from Delaware Coast, sample from Delaware MO Early Summer May 2010 (Delaware MO Early Summer May 2010, Feb 2012 assem)	39.004282	-77.101217	647,909,234	1,375,242	31	0.0478
3300001952	Marine microbial communities from Newport Harbor, Rhode Island, USA - GS008	41.485832	-71.35111	85,651,166	142,390	4	0.0467

Marine metagenomes

IMG Genome ID	Sample Name	Latitude	Longitude	Genome Size [bp] (assembled)	Gene Count (assembled)	Hits	Hits/Gb
3300001941	Marine microbial communities from Browns Bank, Gulf of Maine - GS003	42.85278	-66.217224	43,542,027	80,361	2	0.0459
3300001961	Marine microbial communities from Dirty Rock, Cocos Island, Costa Rica - GS025	5.552778	-87.087776	88,355,364	163,872	4	0.0453
3300001958	Marine microbial communities from Gulf of Mexico, USA - GS016	24.174723	-84.344444	88,679,928	152,380	4	0.0451
3300001960	Marine microbial communities from South of Charleston, South Carolina, USA - GS014	32.506943	-79.263885	90,953,628	159,679	4	0.0440
3300000115	Marine microbial communities from Delaware Coast, sample from Delaware MO Summer July 2011 (Delaware MO Summer July 2011, Nov 2011 assem)	38.848917	-75.1076	524,209,194	1,204,138	22	0.0420
3300000864	Marine plume microbial communities from the Columbia River - Metatranscriptome 25 PSU (Columbia River plume metatranscriptome 25 PSU)	46.235	-124.16	24,148,901	56,425	1	0.0414
3300000883	Estuary microbial communities from the Columbia River - 5 PSU (Columbia River estuary metagenome 5 PSU)	46.235	-123.91	339,832,607	816,259	14	0.0412
3300001943	Marine microbial communities from Cape May, New Jersey, USA - GS010	38.94	-74.685	57,052,860	104,685	2	0.0351
3300000093	Wetland microbial communities from Twitchell Island in the Sacramento Delta, sample from surface sediment Feb2011 Site B2 Bulk (Wetland Surface Sediment Feb2011 Site B2 Bulk, Oct 2011 assem)	38.10726	-121.64854	61,310,069	187,990	2	0.0326
3300000312	Wetland microbial communities from Twitchell Island in the Sacramento Delta, sample from surface sediment Feb2011 Site B2 Bulk (Wetland Surface Sediment Feb2011 Site B2 Bulk, Assem Ctgs Oct 2011 assem)	38.10726	-121.64854	61,310,069	187,974	2	0.0326

Marine metagenomes

IMG Genome ID	Sample Name	Latitude	Longitude	Genome Size [bp] (assembled)	Gene Count (assembled)	Hits	Hits/Gb
3300000116	Marine microbial communities from Delaware Coast, sample from Delaware MO Spring March 2010 (Delaware MO Spring March 2010, Nov 2011 assem)	38.848917	-75.1076	590,073,671	1,417,215	19	0.0322
3300001945	Marine microbial communities from Galapagos, Equador - GS026	1.2641667	-90.295	63,080,825	112,133	2	0.0317
3300000094	Wetland microbial communities from Twitchell Island in the Sacramento Delta, sample from surface sediment Feb2011 Site A1 Tule (Wetland Surface Sediment Feb2011 Site A1 Tule, Oct 2011 assem)	38.107057	-121.64758	129,333,279	362,721	4	0.0309
3300000311	Wetland microbial communities from Twitchell Island in the Sacramento Delta, sample from surface sediment Feb2011 Site A1 Bulk (Wetland Surface Sediment Feb2011 Site A1 Bulk, Asm Ctgs IBYY,IIYG 2012 Mar Assem)	38.107057	-121.64758	243,213,055	583,100	7	0.0288
3300000895	Wetland microbial communities from Twitchell Island in the Sacramento Delta, sample from surface sediment Feb2011 Site A1 Bulk (Wetland Surface Sediment Feb2011 Site A1 Bulk, Ctgs/Rds IBYY,IIYG 2012 Mar Assem)	38.107057	-121.64758	243,213,055	583,095	7	0.0288
3300001949	Marine microbial communities from Panama City, Panama - GS022	6.492778	-82.90389	76,322,024	131,400	2	0.0262
3300000310	Wetland microbial communities from Twitchell Island in the Sacramento Delta, sample from surface sediment Feb2011 Site A1 Bulk (Wetland Surface Sediment Feb2011 Site A1 Bulk, Assem Ctgs IIYG 2012 March Assem)	38.107057	-121.64758	120,209,763	331,296	3	0.0250

Marine metagenomes

IMG Genome ID	Sample Name	Latitude	Longitude	Genome Size [bp] (assembled)	Gene Count (assembled)	Hits	Hits/Gb
3300000840	Wetland microbial communities from Twitchell Island in the Sacramento Delta, sample from surface sediment Feb2011 Site A1 Bulk (Wetland Surface Sediment Feb2011 Site A1 Bulk, Ctgs/UnRds IIYG 2012 March Assem)	38.107057	-121.64758	120,209,763	331,295	3	0.0250
3300001940	Marine microbial communities from Bay of Fundy, Nova Scotia, Canada - GS006	45.111668	-64.94666	40,280,931	72,308	1	0.0248
3300001956	Marine microbial communities from Rangirora Atoll, Polynesia Archipelagos - GS051	-15.143611	-147.435	83,225,611	150,651	2	0.0240
3300000786	Wetland microbial communities from Twitchell Island in the Sacramento Delta, sample from surface sediment Feb2011 Site A2 Cattail (Wetland Surface Sediment Feb2011 Site A2 Cattail, Ctgs/UnReads Sept 2011 assem)	38.107057	-121.64758	84,472,288	260,286	2	0.0237
3300001954	Marine microbial communities from Colon, Panama - GS019	10.716389	-80.25445	86,398,015	149,187	2	0.0231
3300001962	Marine microbial communities from Cocos Island, Costa Rica - GS023	5.64	-86.56528	93,108,667	176,476	2	0.0215
3300000117	Marine microbial communities from Delaware Coast, sample from Delaware MO Winter December 2010 (Delaware MO Winter December 2010, Nov 2011 assem)	39.004282	-77.101217	575,155,359	1,341,985	12	0.0209
3300000108	Wetland microbial communities from Twitchell Island in the Sacramento Delta, sample from surface sediment Feb2011 Site A1 Bulk (Wetland Surface Sediment Feb2011 Site A1 Bulk, 2011 Sep Assem)	38.107057	-121.64758	208,340,005	570,282	4	0.0192
3300000309	Wetland microbial communities from Twitchell Island in the Sacramento Delta, sample from surface sediment Feb2011 Site A1 Bulk (Wetland Surface Sediment Feb2011 Site A1 Bulk, Assem Ctgs IBYY 2011 Sep	38.107057	-121.64758	208,340,005	570,282	4	0.0192

Marine metagenomes

IMG Genome ID	Sample Name	Latitude	Longitude	Genome Size [bp] (assembled)	Gene Count (assembled)	Hits	Hits/Gb
	Assem)						
3300000854	Wetland microbial communities from Twitchell Island in the Sacramento Delta, sample from surface sediment Feb2011 Site B2 Tule (Wetland Surface Sediment Feb2011 Site B2 Tule, Ctgs/UnReads Oct 2011 assem)	38.10726	-121.64854	105,539,119	321,832	2	0.0190
3300002053	Marine sediment microbial communities from White Oak River estuary, North Carolina - WOR_SMTZ	34.647811	-77.111208	799,007,063	857,807	12	0.0150
3300000894	Wetland microbial communities from Twitchell Island in the Sacramento Delta, sample from surface sediment Feb2011 Site L1 Bulk (Wetland Surface Sediment Feb2011 Site L1 Bulk, Ctgs/UnReads Jan 2012 assem)	38.106796	-121.64646	70,007,605	199,148	1	0.0143
3300002961	Wetland microbial communities from Twitchell Island in the Sacramento Delta, sample from surface sediment Feb2011 Site A1 Bulk (Wetland microbial communities from Twitchell Island in the Sacramento Delta, sample from surface sediment Feb2011 Site A1 Bulk, ASSEMBLY_DATE=20140701)	38.107057	-121.64758	560,427,570	1,398,510	8	0.0143
3300000917	Wetland microbial communities from Twitchell Island in the Sacramento Delta, sample from surface sediment Feb2011 Site A2 Cattail (Wetland Surface Sediment Feb2011 Site A2 Cattail, Ctgs/UnRds Jan 2012 Assem)	38.107057	-121.64758	146,925,476	408,994	2	0.0136

Marine metagenomes

IMG Genome ID	Sample Name	Latitude	Longitude	Genome Size [bp] (assembled)	Gene Count (assembled)	Hits	Hits/Gb
3300003516	Marine microbial communities from Antarctic ocean - Station_363 -- 0.8 um (Cold Waters_2)	-60.00007	141.23353	77,909,508	239,389	1	0.0128
3300000077	Wetland microbial communities from Twitchell Island in the Sacramento Delta, sample from surface sediment Feb2011 Site B1 Bulk (Wetland Surface Sediment Feb2011 Site B1 Bulk Feb 2012)	38.10726	-121.64854	79,467,785	197,107	1	0.0126
3300001968	Marine microbial communities from Lake Gatun, Panama - GS020	9.164444	-79.83611	166,403,274	296,358	2	0.0120
3300001955	Marine microbial communities from Gulf of Panama, Panama - GS021	8.129167	-79.69111	85,347,158	150,154	1	0.0117
3300003432	Wetland sediment microbial communities from Twitchell Island in the Sacramento Delta, sample from surface sediment Aug2011 Site B2 Bulk (Wetland sediment microbial communities from Twitchell Island in the Sacramento Delta, sample from surface sediment Aug2011 Site B2 Bulk, ASSEMBLY_DATE=20140909)	38.1072	-121.6485	2,212,415,794	5,001,746	23	0.0104
3300003541	Wetland sediment microbial communities from Twitchell Island in the Sacramento Delta, sample from surface sediment Aug2011 Site B2 Bulk (Wetland sediment microbial communities from Twitchell Island in the Sacramento Delta, sample from surface sediment Aug2011 Site B2 Bulk, ASSEMBLY_DATE=20141008)	38.1072	-121.6485	1,842,474,917	3,667,351	18	0.0098
3300000092	Wetland microbial communities from Twitchell Island in the Sacramento Delta, sample from surface sediment Feb2011 Site L2 Tule (Wetland Surface Sediment Feb2011 Site L2 Tule, Sep 2011 assem)	38.106796	-121.64646	111,737,537	334,657	1	0.0089

Marine metagenomes

IMG Genome ID	Sample Name	Latitude	Longitude	Genome Size [bp] (assembled)	Gene Count (assembled)	Hits	Hits/Gb
3300000317	Wetland microbial communities from Twitchell Island in the Sacramento Delta, sample from surface sediment Feb2011 Site L2 Tule (Wetland Surface Sediment Feb2011 Site L2 Tule, Assem Ctgs Sep 2011 assem)	38.106796	-121.64646	111,737,537	334,653	1	0.0089
3300000425	Marine microbial community from Union City, CA, USA - Pond 2C Liquid 2 (Pond 2C Liquid 2 Union City, July 2012 Assem)	37.569017	-122.10243	131,906,500	297,522	1	0.0076
3300001965	Marine microbial communities from Coastal Floreana, Equador - GS028	-1.2169445	-90.319725	135,719,211	247,934	1	0.0074
3300000090	Wetland microbial communities from Twitchell Island in the Sacramento Delta, sample from surface sediment Feb2011 Site B1 Cattail (Wetland Surface Sediment Feb2011 Site B1 Cattail, Sep 2011 assem)	38.10726	-121.64854	291,270,262	747,850	2	0.0069
3300000313	Wetland microbial communities from Twitchell Island in the Sacramento Delta, sample from surface sediment Feb2011 Site B1 Cattail (Wetland Surface Sediment Feb2011 Site B1 Cattail, Assem Ctgs Sep 2011 assem)	38.10726	-121.64854	291,270,262	747,866	2	0.0069
3300000030	Wetland microbial communities from Twitchell Island in the Sacramento Delta, sample from surface sediment Feb2011 Site B1 Bulk (Wetland Surface Sediment Feb2011 Site B1 Bulk, Oct 2011 assem)	38.10726	-121.64854	153,197,283	433,712	1	0.0065
3300001281	Wetland microbial communities from Twitchell Island in the Sacramento Delta, sample from surface sediment Aug2011 Site A1 Bulk (Wetland Surface Sediment Aug2011 Site A1 Bulk Metagenome, ASSEMBLY_DATE=20130408)	38.107	-121.6475	173,204,638	479,751	1	0.0058

Marine metagenomes

IMG Genome ID	Sample Name	Latitude	Longitude	Genome Size [bp] (assembled)	Gene Count (assembled)	Hits	Hits/Gb
3300000418	Marine microbial community from Union City, CA, USA - Pond 2C Liquid 1 (Pond 2C Liquid 1 Union City, June 2012 Assem)	37.569167	-122.1019	194,495,031	424,469	1	0.0051
3300000385	Marine microbial community from Cabo Rojo, Puerto Rico - PR CR 10% Liquid 1 (PR CR 10% Liquid 1 Cabo Rojo PR, June 2012 Assem)	17.951083	-67.193167	210,351,197	409,868	1	0.0048
3300003475	Marine microbial communities from the Indian Ocean - GS112 (warm waters)	-8.505	80.37556	224,938,505	730,224	1	0.0044
3300000091	Wetland microbial communities from Twitchell Island in the Sacramento Delta, sample from surface sediment Feb2011 Site L1 Cattail (Wetland Surface Sediment Feb2011 Site L1 Cattail, Sep 2011 assem)	38.106796	-121.64646	296,519,018	679,298	1	0.0034
3300000318	Wetland microbial communities from Twitchell Island in the Sacramento Delta, sample from surface sediment Feb2011 Site L1 Cattail (Wetland Surface Sediment Feb2011 Site L1 Cattail, Assem Ctgs Sep 2011 assem)	38.106796	-121.64646	296,519,018	679,305	1	0.0034
2199352009	Marine subseafloor sediment microbial communities, sample from White Oak River Estuary, NC, USA 14E (White Oak River Estuary June 2011 assem)	34.690251	-77.106571	22,398,301	51,820	0	0.0000
3300000100	Wetland microbial communities from Twitchell Island in the Sacramento Delta, sample from surface Sediment Feb2011 Site B2 Cattail (Wetland Surface Sediment Feb2011 Site B2 Cattail, Oct 2011 assem)	38.10726	-121.64854	90,589,564	284,067	0	0.0000
3300000316	Wetland microbial communities from Twitchell Island in the Sacramento Delta, sample from surface Sediment Feb2011 Site B2 Cattail (Wetland Surface Sediment Feb2011 Site B2 Cattail, Assem Ctgs Oct	38.10726	-121.64854	90,589,564	284,068	0	0.0000

Marine metagenomes

IMG Genome ID	Sample Name	Latitude	Longitude	Genome Size [bp] (assembled)	Gene Count (assembled)	Hits	Hits/Gb
	2011 assem)						
3300000320	Marine microbial communities from Delaware Coast, sample from Delaware MO Early Summer May 2010	39.004282	-77.101217	3,880,822	10,893	0	0.0000
3300000321	Marine microbial communities from Delaware Coast, sample from Delaware MO Winter December 2010	39.004282	-77.101217	11,256,121	45,129	0	0.0000
3300000369	Marine microbial community from Union City, CA, USA - Pond 2C Liquid 3 (Pond 2C Liquid 3 Union City, June 2012 Assem)	37.568817	-122.10315	17,102,308	18,170	0	0.0000
3300000371	Marine microbial community from Union City, CA, USA - Pond 1C Liquid 3 (Pond 1C Liquid 3 Union City, June 2012 Assem)	37.5693	-122.10252	47,212,914	125,124	0	0.0000
3300000403	Marine sediment microbial community from La Parguera, Puerto Rico - PR Tt Sediment 2 (PR Tt Sediment 2 - Bioluminescent Bay in La ParagueraPR, July 2012 Assem)	17.967317	-67.018833	898,328	2,770	0	0.0000
3300000409	Marine sediment microbial community from Union City, CA, USA - Pond 2C Sediment 1 (Pond 2C Sediment 1 Union City, June 2012 Assem)	37.569167	-122.1019	24,779,847	77,851	0	0.0000
3300000463	Marine sediment microbial community from Union City, CA, USA - Pond 2C Sediment 3 (Pond 2C Sediment 3 Union City, June 2012 Assem)	37.568817	-122.10315	38,427,200	111,900	0	0.0000
3300000517	Marine microbial community from Cabo Rojo, Puerto Rico - PR CR 10% Liquid 3 (PR CR 10% Liquid 3 Cabo Rojo PR, June 2012 Assem)	17.950617	-67.193417	120,470,807	242,233	0	0.0000

Marine metagenomes

IMG Genome ID	Sample Name	Latitude	Longitude	Genome Size [bp] (assembled)	Gene Count (assembled)	Hits	Hits/Gb
3300000853	Marine plume microbial communities from the Columbia River - Metatranscriptome 15 PSU (Columbia River plume metatranscriptome 15 PSU)	46.239	-124.161	7,327,946	18,177	0	0.0000
3300000867	Estuary microbial communities from the Columbia River - metatranscriptome 5 PSU (Columbia River estuary metatranscriptome 5 PSU)	46.235	-123.91	58,151,542	137,277	0	0.0000
3300001279	Wetland sediment microbial communities from Twitchell Island in the Sacramento Delta, sample from surface sediment Aug2011 Site B2 Bulk (Wetland Surface Sediment Aug2011 Site B2 Bulk Metagenome, ASSEMBLY_DATE=20130408)	38.1072	-121.6485	217,817,654	615,872	0	0.0000
3300001922	Marine microbial communities from Polynesia - GS046	-9.571111	-131.49167	336,890	629	0	0.0000
3300001923	Marine microbial communities from the Tropical South Pacific Ocean - GS038	-2.5819445	-97.85139	451,181	795	0	0.0000
3300001924	Marine microbial communities from Tikehau Lagoon, Polynesia Archipelagos - GS050	-15.277778	-148.22444	467,458	887	0	0.0000
3300001925	Marine microbial communities from the Tropical South Pacific Ocean - GS041	-5.93	-108.68694	494,549	909	0	0.0000
3300001926	Marine microbial communities from the Tropical South Pacific Ocean - GS042	-7.1075	-116.11916	542,583	953	0	0.0000
3300001927	Marine microbial communities from Polynesia - GS044	-8.415	-124.23972	562,395	1,015	0	0.0000
3300001928	Marine microbial communities from the Tropical South Pacific Ocean - GS043	-7.661111	-116.11916	612,652	1,073	0	0.0000
3300001930	Marine microbial communities from the Tropical South Pacific Ocean - GS039	-3.3433332	-101.37389	641,644	1,142	0	0.0000
3300001931	Marine microbial communities from Polynesia - GS045	-9.0175	-127.76722	636,185	1,152	0	0.0000

Marine metagenomes

IMG Genome ID	Sample Name	Latitude	Longitude	Genome Size [bp] (assembled)	Gene Count (assembled)	Hits	Hits/Gb
3300001932	Marine microbial communities from Moorea, Cooks Bay, Polynesia Archipelagos - GS049	-17.453056	-149.79889	638,416	1,194	0	0.0000
3300001933	Marine microbial communities from Moorea, Cooks Bay - GS048a	-17.475834	-149.81223	666,592	1,245	0	0.0000
3300001936	Marine microbial communities from Halifax, Nova Scotia, Canada - GS004	44.137222	-63.644444	33,043,040	56,286	0	0.0000
3300001937	Marine microbial communities from the Equatorial Pacific Ocean - GS037	-1.9738889	-95.014725	36,191,958	63,262	0	0.0000
3300001947	Marine microbial communities from the Gulf of Maine, Canada - GS002	42.503056	-67.24	74,202,974	125,810	0	0.0000
3300001950	Marine microbial communities from Delaware Bay, New Jersey, USA - GS011	39.417778	-75.504166	79,799,345	138,086	0	0.0000
3300001953	Marine microbial communities from Key West, Florida, USA - GS015	24.488333	-83.07	84,693,422	146,422	0	0.0000
3300001959	Mangrove swamp microbial communities from Isabella Island, Equador - GS032	-0.5938889	-91.06944	89,398,486	153,552	0	0.0000
3300003066	Marine Synechococcus microbial communities from coastal surface water at La Jolla, California, USA - SRS00536	32.850317	-117.27494	307,378	1,337	0	0.0000
3300003073	Marine surface microbial communities Puget Sound, Washington, USA	47.690656	-122.40441	2,059,964	1,892	0	0.0000
3300003098	Marine microbial communities from surface seawater at Gulf of Maine	43.14	-68.33	5,278,992	8,597	0	0.0000
3300003139	Mat microbial community in the pink berry consortia of the Sippewissett salt marsh	41.575836	-70.63923	33,200,272	56,503	0	0.0000
3300004831	Marine surface microbial communities from the North Atlantic Ocean - filtered matter	41.226956	-8.720528	27,429,400	84,353	0	0.0000
3300005472	Wetland microbial communities from Twitchell Island in the Sacramento Delta, sample from surface sediment Feb2011 Site L1 Cattail (re-annotation)	38.106796	-121.64646	2,364,101	2,630	0	0.0000
3300005698	Marine microbial communities from six Antarctic regions - DNA Fragments (re-	-64.766667	-64.05	43,087	39	0	0.0000

Marine metagenomes

IMG Genome ID	Sample Name	Latitude	Longitude	Genome Size [bp] (assembled)	Gene Count (assembled)	Hits	Hits/Gb
	annotation)						
3300005722	Marine Trichodesmium cyanobacterial communities from the North Pacific Subtropical Gyre outside Oahu, HI, sample from new species B colonies (re-annotation)	22.75	-158	27,189,071	39,432	0	0.0000
3300006561	Marine microbial communities from the Black Sea in Odessa region - Od_1 (Assembly)	46.35116	30.5519	22,413,229	75,560	0	0.0000
3300006562	Marine microbial communities from the Black Sea in Odessa region - Od_2 (Assembly)	46.302094	30.668915	31,351,919	105,939	0	0.0000

AMERICAN SOCIETY FOR MICROBIOLOGY | Applied and Environmental Microbiology®

BIODEGRADATION

New Insights into the Function and Global Distribution of Polyethylene Terephthalate (PET)-Degrading Bacteria and Enzymes in Marine and Terrestrial Metagenomes

Dominik Danso,[a] Christel Schmeisser,[a] Jennifer Chow,[a] Wolfgang Zimmermann,[b] Ren Wei,[b] Christian Leggewie,[c] Xiangzhen Li,[d] Terry Hazen,[e] Wolfgang R. Streit[a]

[a]Department of Microbiology and Biotechnology, Biocenter Klein Flottbek, University of Hamburg, Hamburg, Germany

[b]Institute of Biochemistry, Department of Microbiology and Bioprocess Technology, Leipzig University, Leipzig, Germany

[c]evoxx technologies GmbH, Monheim am Rhein, Germany

[d]Chengdu Institute of Biology, Chengdu, China

[e]The University of Tennessee, Knoxville, Tennessee, USA

ABSTRACT Polyethylene terephthalate (PET) is one of the most important synthetic polymers used today. Unfortunately, the polymers accumulate in nature and to date no highly active enzymes are known that can degrade it at high velocity. Enzymes involved in PET degradation are mainly α- and β-hydrolases, like cutinases and related enzymes (EC 3.1.1). Currently, only a small number of such enzymes are well characterized. In this work, a search algorithm was developed that identified 504 possible PET hydrolase candidate genes from various databases. A further global search that comprised more than 16 Gb of sequence information within 108 marine and 25 terrestrial metagenomes obtained from the Integrated Microbial Genome (IMG) database detected 349 putative PET hydrolases. Heterologous expression of four such candidate enzymes verified the function of these enzymes and confirmed the usefulness of the developed search algorithm. In this way, two novel and thermostable enzymes with high potential for downstream application were partially characterized. Clustering of 504 novel enzyme candidates based on amino acid similarities indicated that PET hydrolases mainly occur in the phyla of *Actinobacteria*, *Proteobacteria*, and *Bacteroidetes*. Within the *Proteobacteria*, the *Betaproteobacteria*, *Deltaproteobacteria*, and *Gammaproteobacteria* were the main hosts. Remarkably enough, in the marine environment, bacteria affiliated with the phylum *Bacteroidetes* appear to be the main hosts of PET hydrolase genes, rather than *Actinobacteria* or *Proteobacteria*, as observed for the terrestrial metagenomes. Our data further imply that PET hydrolases are truly rare enzymes. The highest occurrence of 1.5 hits/Mb was observed in sequences from a sample site containing crude oil.

IMPORTANCE Polyethylene terephthalate (PET) accumulates in our environment without significant microbial conversion. Although a few PET hydrolases are already known, it is still unknown how frequently they appear and with which main bacterial phyla they are affiliated. In this study, deep sequence mining of protein databases and metagenomes demonstrated that PET hydrolases indeed occur at very low frequencies in the environment. Furthermore, it was possible to link them to phyla that were previously not known to harbor such enzymes. This work contributes novel knowledge on the phylogenetic relationships, the recent evolution, and the global distribution of PET hydrolases. Finally, we describe the biochemical traits of four novel PET hydrolases.

KEYWORDS HMM, hydrolases, metagenome, metagenomic screening, PET degradation, polyethylene terephthalate (PET), BHET, TPA, metagenomes

Received 14 December 2017 **Accepted** 29 January 2018

Accepted manuscript posted online 2 February 2018

Citation Danso D, Schmeisser C, Chow J, Zimmermann W, Wei R, Leggewie C, Li X, Hazen T, Streit WR. 2018. New insights into the function and global distribution of polyethylene terephthalate (PET)-degrading bacteria and enzymes in marine and terrestrial metagenomes. Appl Environ Microbiol 84:e02773-17. https://doi.org/10.1128/AEM.02773-17.

Editor Rebecca E. Parales, University of California, Davis

Address correspondence to Wolfgang R. Streit, wolfgang.streit@uni-hamburg.de.

Since its discovery, its first synthesis, and its patenting in 1941, polyethylene terephthalate (PET) has become a widely used material in several industrial branches (1). The worldwide PET resin production amounted to 27.8 million tons in 2015 (https://www.plasticsinsight.com/global-pet-resin-production-capacity).

Due to its massive use, PET is highly enriched in nature. Microplastics and bigger fragments of plastic are found worldwide in oceans and terrestrial environments. The most prominent example is the so-called Pacific garbage patch. PET debris is often eaten by fish and other marine creatures (2, 3). In this way, PET degradation products and additives (i.e., solubilizers) are introduced into the food chain, where they have negative impact on human and animal health (4). Until now, only a few species of bacteria and fungi have been described as capable of partially degrading PET to oligomers or even monomers (5). Within this framework, however, it is noteworthy that all known PET hydrolases have relatively low turnover rates, which makes their use for efficient bioremediation almost impossible (Table 1).

Intriguingly, the trait for PET degradation appears to be limited to a few bacterial phyla, and most bacterial isolates with potential for PET degradation are members of the Gram-positive phylum *Actinobacteria* (12). The best-characterized examples originate from the genera *Thermobifida* and *Thermomonospora* (8, 10–12, 15, 16) (Table 1). The enzymes involved in the degradation (e.g., PET hydrolase and tannase) are typical serine hydrolases, e.g., cutinases (EC 3.1.1.74), lipases (EC 3.1.1.3), and carboxylesterases (EC 3.1.1.1). These enzymes possess a typical α/β-hydrolase fold, and the catalytic triad is composed of a serine, a histidine, and an aspartate residue (17, 18).

More recently, polyethylene (PE)-degrading bacteria were reported in insect guts. In this recent study, *Enterobacteria* and *Bacillus* strains had been isolated and were capable of degrading polyesters (19, 20).

Furthermore, a complete degradation of amorphous PET materials was described for the Gram-negative bacterium *Ideonella sakaiensis* strain 201-F6, which is able to use PET as a major energy and carbon source (13). In addition to the hydrolase, the *I. sakaiensis* genome encodes a second enzyme that appears to be unique, and which is designated as a tannase capable of degrading mono(2-hydroxyethyl) terephthalic acid. In this way, the secreted PET hydrolase produces the intermediate mono(2-hydroxyethyl) terephthalic acid (MHET). MHET is presumably internalized by the cell and hydrolyzed by the MHETase. The resulting monomers are then degraded in a downstream process and used for the bacterial metabolism. *I. sakaiensis* is affiliated with the phylum *Betaproteobacteria* and belongs to the order *Burkholderiales*.

In this work, our intention was to mine metagenomes for the detection of novel genes involved in PET degradation and to establish an overview on their taxonomic distribution within the different bacterial phyla. Therefore, we have developed a hidden Markov model (HMM) to search existing genome and metagenome databases for the presence of potential PET hydrolases (Fig. 1; see also Fig. 3B). Using this approach, we identified >500 potential PET hydrolases in the UniProtKB database. In addition, 349 sequence homologs were obtained from several public metagenome data sets deposited on the IMG server, and four of the identified candidate genes were functionally verified. Together, these results imply that PET hydrolase genes are globally distributed in marine and terrestrial metagenomes. Furthermore, we provide evidence that in marine environments, the PET hydrolases originate mainly from the phylum *Bacteroidetes* and in the terrestrial metagenomes from *Actinobacteria*.

RESULTS

Construction of a hidden Markov model for PET hydrolases. Only a few well-characterized PET hydrolases are currently known. The most prominent examples are PET hydrolases from *T. fusca* and *I. sakaiensis* (see references in Table 1 and Fig. 2). In this study, we set out to increase knowledge of the diversity of this intriguing group of hydrolases. To identify potential novel PET hydrolases, an amino acid sequence alignment of nine already-known examples was constructed using the T-Coffee multiple sequence alignment server. The enzyme sequences used for the model all have verified

TABLE 1 Currently known and partially characterized PET hydrolases

Sequence no.	PDB entry no.[a]	Gene name[b]	Organism	Reference
1	W0TJ64	Cut190	*Saccharomonospora viridis*	6
2	E9LVI0	cut1	*Thermobifida fusca* (*Thermomonospora fusca*)	7
3	E5BBQ3	cut-2	*Thermobifida fusca* (*Thermomonospora fusca*)	8
4	D1A9G5	Tcur_1278	*Thermomonospora curvata*	9
5	E9LVH7	cut1	*Thermobifida alba*	10
6	H6WX58	NA	*Thermobifida halotolerans*	11
7	E9LVH9	cut2	*Thermobifida celluloysilityca*	12
8	A0A0K8P6T7	ISF6_4831	*Ideonella sakaiensis*	13
9	G9BY57	NA	Uncultured bacterium	14

[a]Names and protein database (PDB) entry numbers of currently known PET hydrolases used in this work as references. Sequence data of these examples were used for the initial construction of the HMM.
[b]NA, not applicable.

activity on PET-based substrates (Table 1). Of the proteins used for the model, seven sequences originated from the phylum *Actinobacteria* (i.e., sequences 1 to 7; Table 1) and one from the phylum *Proteobacteria* (*Betaproteobacteria*) (sequence 8; Table 1), and one sequence (AEV21261) encoded a protein with a metagenomic origin not yet assigned to any phylum (sequence 9, Table 1). A comparison with the well-described PET-active cutinase TfCut2 from *Thermobifida fusca* allowed the identification of the location of the catalytic triad and other residues that are commonly involved in binding of the substrates (Fig. 3A). Next to the serine of the catalytic triad in every sequence, a methionine residue was found, which is of importance for forming an oxyanion hole together with an aromatic residue. This aromatic residue is also part of an aromatic clamp, together with similar amino acids like tryptophan, tyrosine, histidine, and phenylalanine (17, 21). Terminal cysteine residues are present in all examples and may be important for the thermostability of these enzymes (14, 22). The alignment was used for the construction of a hidden Markov model. For visualization, an HMM logo was created via the Skylign online tool (http://skylign.org/; Fig. 3B). A subsequent visual analysis and conservation prediction using the JS divergence scoring method revealed at least eight conserved regions (Fig. 3B).

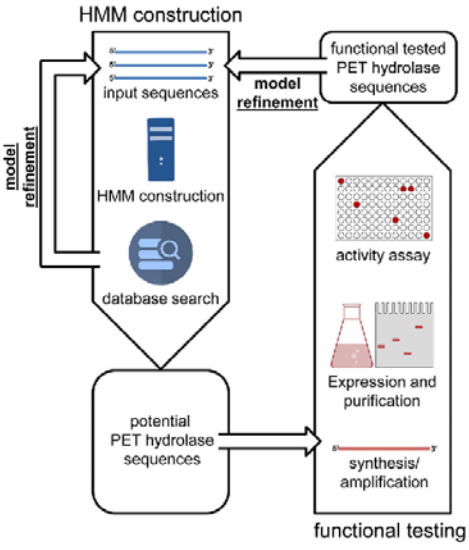

FIG 1 Workflow used in this study to identify and partially characterize novel PET hydrolases from databases and global metagenomes.

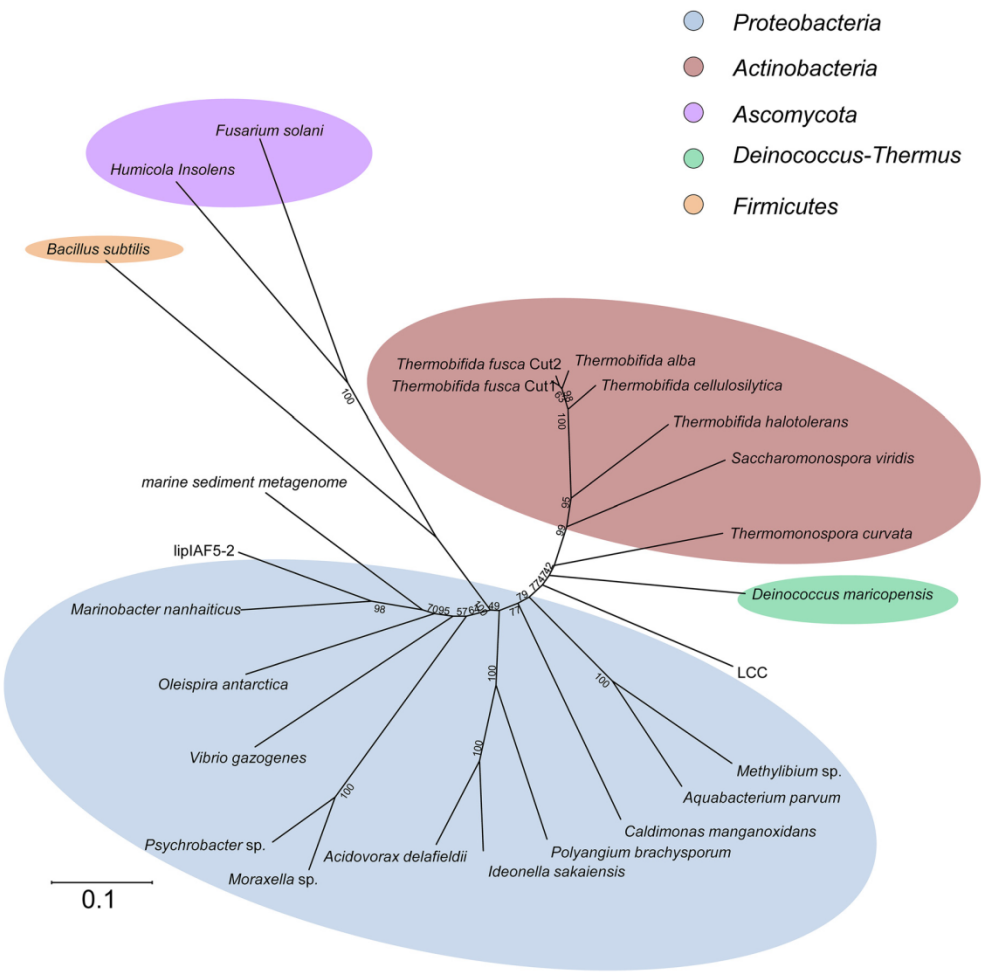

FIG 2 Neighbor-joining tree of manually chosen potential PET hydrolase sequences found in this work. Sequences were obtained from a HMM search in the UniProtKB database and named PET1 to PET13. The tree was calculated using MEGA6. Besides the 13 newly found PET hydrolase sequences (Table S1), 9 already-known PET hydrolases (Table 1) were added to the tree in order to visualize the phylogenetic distribution and similarity of the PET hydrolase sequence homologs.

An initial HMMER online tool database search with the model against the UniProtKB database revealed a total of 10,854 significant query matches, with a highest bit score value of 441.6. Of these, a subset was chosen that showed a bit score value of >180. In addition, a BLAST search was performed using the newly discovered potential PET hydrolases from the HMM search as initial query sequences against the nonredundant and the metagenomic datasets available at the NCBI database in May 2017. This resulted in the detection of 504 potential PET hydrolase candidate genes. From the obtained homologous sequences, 13 potential PET hydrolase homologs (Fig. 2) were manually chosen due to their sequence similarity to known PET hydrolases (PET1 to PET13). These were used for initial verification and further *in silico* and/or biochemical characterization. These novel predicted PET hydrolases are summarized in Table S1 in the supplemental material, together with their UniProt entries and pfam domain similarities. It was of interest to select mainly nonactinobacterial proteins, in order to diversify the HMM. These 13 sequences were added to the alignment and used for a

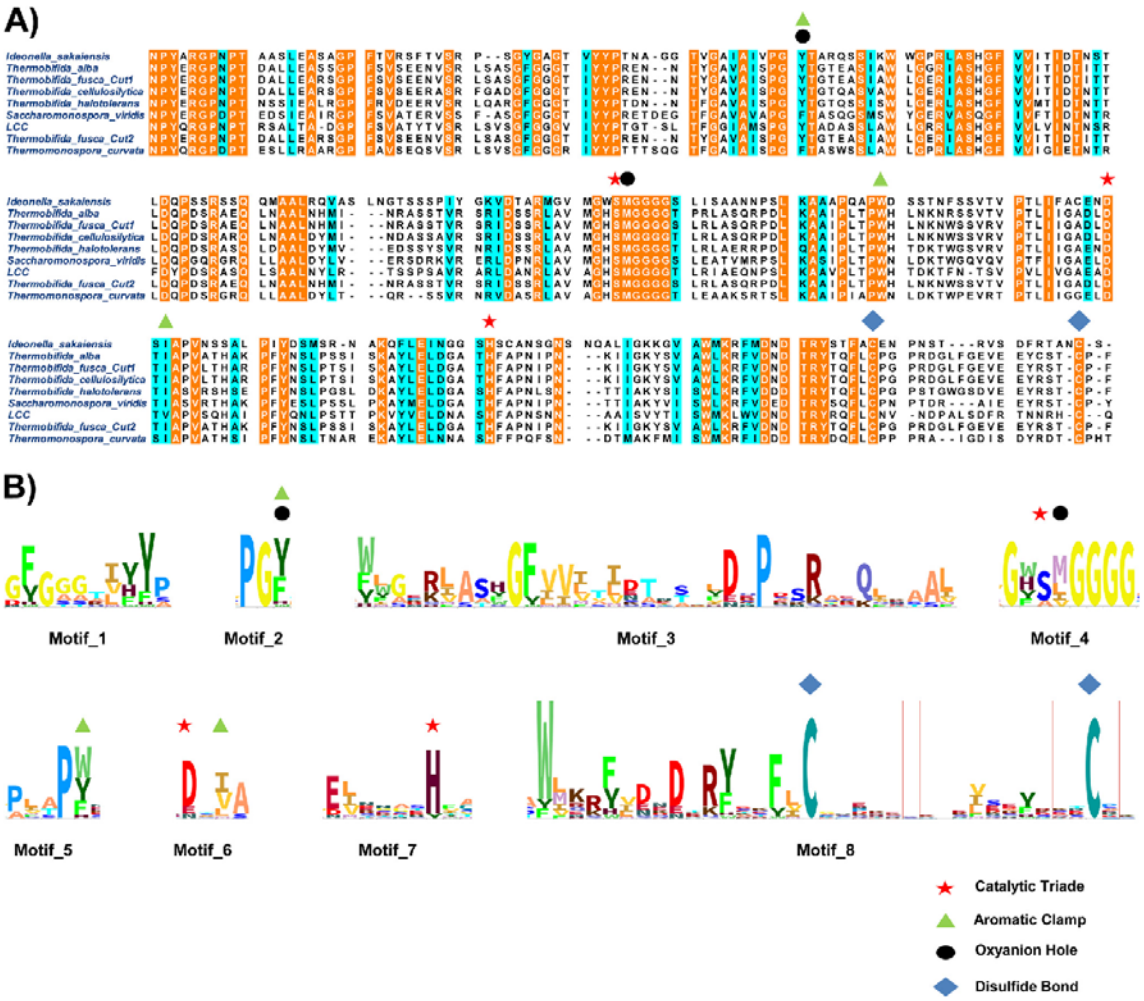

FIG 3 Amino acid sequence alignment of described PET hydrolases. (A) An alignment of the PET hydrolase sequences listed in Table 1 revealed the positions of binding relevant residues and conserved regions. (B) Hidden Markov model (HMM) of PET hydrolase amino acid motifs. The amino acid alignment from panel A was used to calculate a HMM profile. The HMM was consequently visualized as a logo with information content above the background (Skylign; http://skylign.org). Eight sequence motifs are shown in total. Motifs 2 to 8 include amino acids crucial for thermostability, substrate binding, and/or catalytic activity.

modified and refined HMM. The 13 initially identified putative PET hydrolases harbor the above-mentioned residues and motifs (listed in Table 2).

Classification of PET hydrolases and taxonomic assignments. An NCBI conserved domain search in the Conserved Domains Database (CDD) database showed that the nine previously known active PET hydrolases (Table 1), as well as the 13 novel (see Table S1) possible homologs, harbor domains belonging to the superfamily of α/β-hydrolases_5 (pfam12695). Of these, only PET9 and PET12 showed specific hits for the superfamily of acetyl xylan esterases (AXE1, pfam05448). In addition to the specific hits, several nonspecific domain hits were obtained (Table S1).

A further alignment and subsequent tree calculation with the above identified 504 potential PET hydrolases allowed the assignment of all enzymes in 17 subclasses (Fig 4). Interestingly, two sequences (UniProt numbers A0A1N6SMU6 and A0A168EN35) did form individual subclusters and could not be assigned to other clusters. The majority of

TABLE 2 Determined search criteria for the identification of PET hydrolase candidate genes in databases

Sequence no.	Search criterion (criteria)[a]	Function
1	GxS**M**GGGG	Serine of catalytic triade and methionine for oxyanion hole formation
2	F,Y62	Amino acids for oxyanion hole formation and aromatic clamp
3	W,Y157	
4	I,V180	
5	F,W211	Optional aromatic amino acid for aromatic clamp formation
6	C255 C262	C-terminal cysteine residues for thermostability supporting disulfide bond formation
7	DxDxR(Y)xxF(L)**C**	Conserved sequence prior to first thermostability giving cysteine

[a]The letter x indicates a nonconserved position within the sequence pattern. Brackets indicate a less conserved position within the sequence pattern. Numbering of amino acids is according to the HMM (see Fig. 3A).

the subclasses were mainly affiliated with *Actinobacteria*, and only one subclass (XII) was associated with *Proteobacteria*. For the subclass XVII, no clear assignment was possible. The *Thermobifida* PET hydrolase sequences are clustered within subcluster XV, together with the PET hydrolase from *Saccharomonospora*. The leaf compost cutinase (LCC) sequence was found in subcluster XI and the *Thermomonospora curvata* sequence is located in group XII. The PET hydrolase from *Ideonella sakaiensis* is located in subcluster VI.

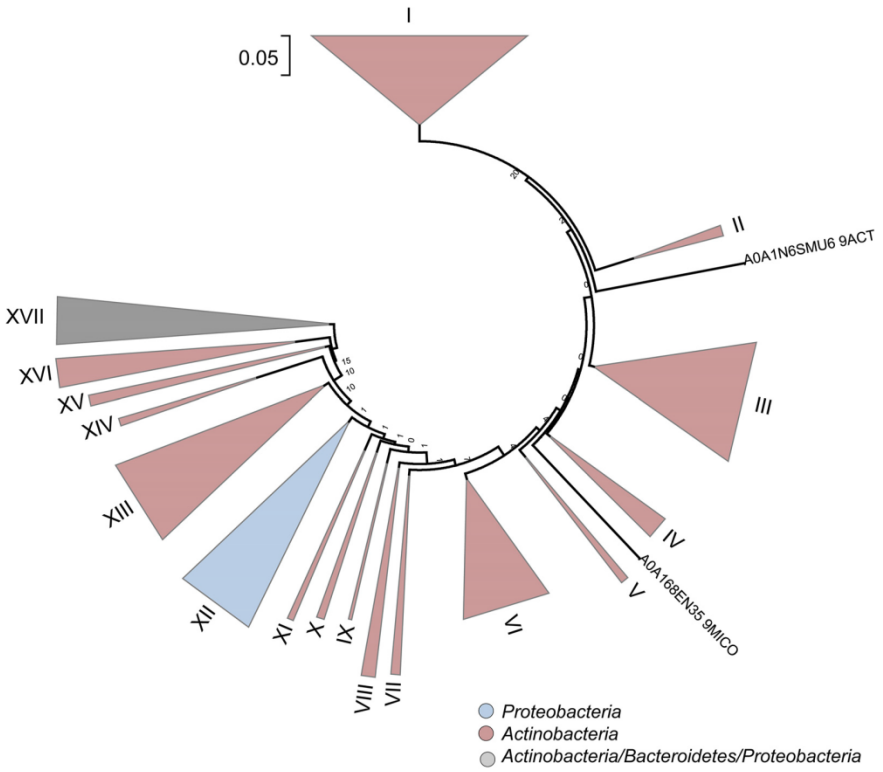

FIG 4 Classification and phylogenetic tree of 504 novel and potential PET hydrolases obtained by HMM searches. Sequences were obtained from the UniProtKB database. A total of 504 sequences identified with the constructed HMM and having a bit score of >180 were visualized, of which the sequences of PET1 to PET13 (Table S1), as well as 9 already described PET hydrolases (Table 1), represent a subset of the newly found potential enzymes.

TABLE 3 Bacterial strains and plasmids used in this work

Strain or plasmid	Property(ies)[a]	Reference or source
Strains		
E. coli DH5α	*supE44 ΔlacU169 (Φ80 lacZ ΔM15) hsdR17 recA1 endA1 gyrA96 thi-1 relA1*	23
E. coli BL21(DE3)	F⁻ *ompT hsdS* B (r_B^- m_B^-) *gal dcm* λDE3	Novagen/Merck (Darmstadt, Germany)
E. coli T7SHuffle Express	*fhuA2 lacZ::T7 gene1 [lon] ompT ahpC gal λatt::pNEB3-r1-cDsbC (SpecR, lacIq) ΔtrxB sulA11 R(mcr-73*::miniTn*10*-TetS*)2 [dcm] R(zgb-210*::Tn*10*-TetS*) endA1 Δgor Δ(mcrC-mrr)114*::IS*10*	NEB (Frankfurt am Main, Germany)
Deinococcus maricopensis DSM-21211	Type strain	DSMZ (Braunschweig, Germany)
Vibrio gazogenes DSM-21264	Type strain	DSMZ (Braunschweig, Germany)
Polyangium brachysporum DSM-7029	Type strain	DSMZ (Braunschweig, Germany)
Plasmids		
pET21a(+)	Expression vector, *lacI*, Amp^r, T7 *lac* promoter, C-terminal His_6-tag coding sequence	Novagen/Merck (Darmstadt, Germany
pET28a(+)	Expression vector, *lacI*, Amp^r, T7 *lac* promoter, C-terminal His_6-tag and N-terminal coding sequence	Novagen/Merck (Darmstadt, Germany
pEX-A2	Cloning vector, Amp^r, P_{lac}*lacZ*, pUC *ori*	Eurofins MWG Operon (Ebersberg, Germany)

[a]Amp^r, ampicillin resistance.

Experimental verification of the HMM and characterization of selected novel PET hydrolases. Since the bioinformatic approach only delivered potential PET hydrolases enzymes, we initiated work to verify a small number of the identified candidate genes with respect to their function. Therefore, we chose the enzymes PET 2, 5, 6, and 12 (Table S1). The respective genes were either synthesized or amplified from genomic DNA using vectors and primers (as outlined in Tables 3 and 4) and were cloned into the expression vectors. Initial tests indicated that all genes coded for active enzymes. On agar plates containing PET nanoparticles or polycaprolactone (PCL) (9, 24), all active clones produced halos after overnight incubation. They were compared to the PET hydrolase from *Thermobifida fusca* as a positive control (see Fig. S1 in the supplemental material). PCL was used as a model substrate, as hydrolysis of this compound indicates possible activities on the more complex PET. From these active enzymes, we chose two enzymes for more detailed biochemical characterization. These were the two enzymes PET 2 and PET 6. PET 2 was derived from a marine metagenomics data set (25), and PET 6 was derived from *Vibrio gazogenes* strain DSM-21264 (26). After successful expression and purification of the two enzymes in sufficient amounts, the obtained enzymes were further characterized using *para*-nitrophenyl esters (*p*NP esters) (Fig. 5; see also Fig. S2 and S3 in the supplemental material). Both enzymes showed best activity against *p*NP esters (C_2 to C_4), but were able to convert long-chain ($>C_{10}$) substrates as well. Their temperature optimums were 55°C and 70°C for PET6 and PET2, respectively. Remarkably, PET2 retained 80% of its relative activity at 90°C after incubation for >5 h. Both enzymes preferred alkaline pH values of 8 to 9. Rubidium at a concentration of 1 mmol/liter had a strong effect in the case of PET2 by increasing the activity by 50%. A similar but significantly smaller effect was observed in the case of PET6. Both enzymes

TABLE 4 Primers used in this work

Primer	Sequence (5′ → 3′)	Length (bp)	T_m (°C)[a]	Source
T7 promoter	TAATACGACTCACTATAGGG	20	53.2	Eurofins MWG (Ebersberg, Germany)
T7 terminator	CTAGTTATTGCTCAGCGGT	19	54.5	Eurofins MWG (Ebersberg, Germany)
PET5_for	CGCCGCCATATGAATAAATCTATTCTAAAAAAACTCTC	38	68	This work
PET5_rev	CGATTCGGCGGCCGCGTAATTACATGTGTCACGG	34	77	This work
PET6_for	CGTAGTCATATGGTACCGTGTTCGGACTG	29	69	This work
PET6_rev	CAGCGGCCGCCTAATAGTAACTACAGTTGTCTCG	34	73	This work
PET12_for	CGCCATATGCAGACCAACCCCTACCAGCGAGGCCC	35	80	This work
PET12_rev	CTTGCGGCCGCTCAGTACGGGCAGCTCTCGCGGTACTCC	39	84	This work

[a]T_m, melting temperature.

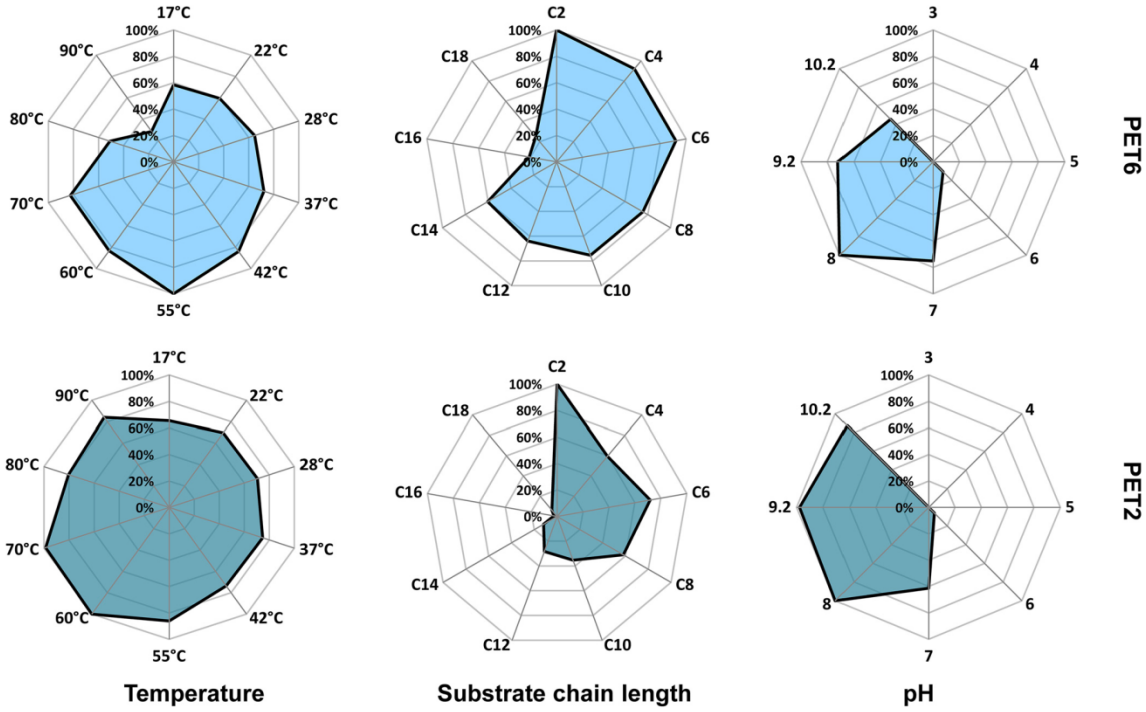

FIG 5 Biochemical characterization of PET2 and PET6 with different *p*NP substrates. Data obtained with a *p*NP assay are shown in net diagrams for PET2 and PET6. Substrate preferences, temperature optimum, and pH optimum were tested. All tests besides substrate preferences were carried out with *p*NP octanoate.

showed reduced hydrolytic activity in the presence of 5% SDS, 10 mM phenylmethylsulfonyl fluoride (PMSF), and 30% acetonitrile. Additional high-performance liquid chromatography (HPLC) analyses confirmed the above findings for PET2. In tests using 14 mg of amorphous PET foil as the substrate, 100 μg of PET 2 was able to release 900 μM terephthalic acid after 24 h of incubation (see Fig. S4 in the supplemental material).

Altogether, the data presented above indicate that the developed search algorithm is useful for the identification of novel and functionally active PET hydrolases from single genomes and metagenomes.

Global distribution of PET hydrolases and their significance in marine and terrestrial environments. After the successful construction of a reliable HMM and the identification, as well as partial characterization, of new PET hydrolases, we asked if these enzymes could be identified on a global level, and if so, to what extent. To evaluate the environmental distribution of sequences encoding PET hydrolases, the data from 108 marine and 25 terrestrial metagenomes were taken into account and downloaded from the IMG database (27) (see Table S2 in the supplemental material). Criteria for the selection of marine data included sample depth (maximum 2 m), assembly status, global distribution of sample locations, and size and availability of the data set. The same criteria, except for the sample depth, were chosen for terrestrial metagenomes. The size of the assembled metagenome data in the case of marine metagenomes ranged from 10.85 Mb up to 7.99 Gb. In the case of terrestrial metagenomes, the number of assembled bases ranged from 58 Mb to 9.2 Gb. The modified HMM was used to find PET hydrolase homologs in the sequence data of those metagenomes on a global scale. The searches identified possible PET hydrolase homologs in 31 marine and 11 terrestrial metagenomes. A total of 349 hits was observed for these 42 samples. The number of hits per sample was normalized, calculated as hits per Mb, and visualized on a global map representing the geographical location as well

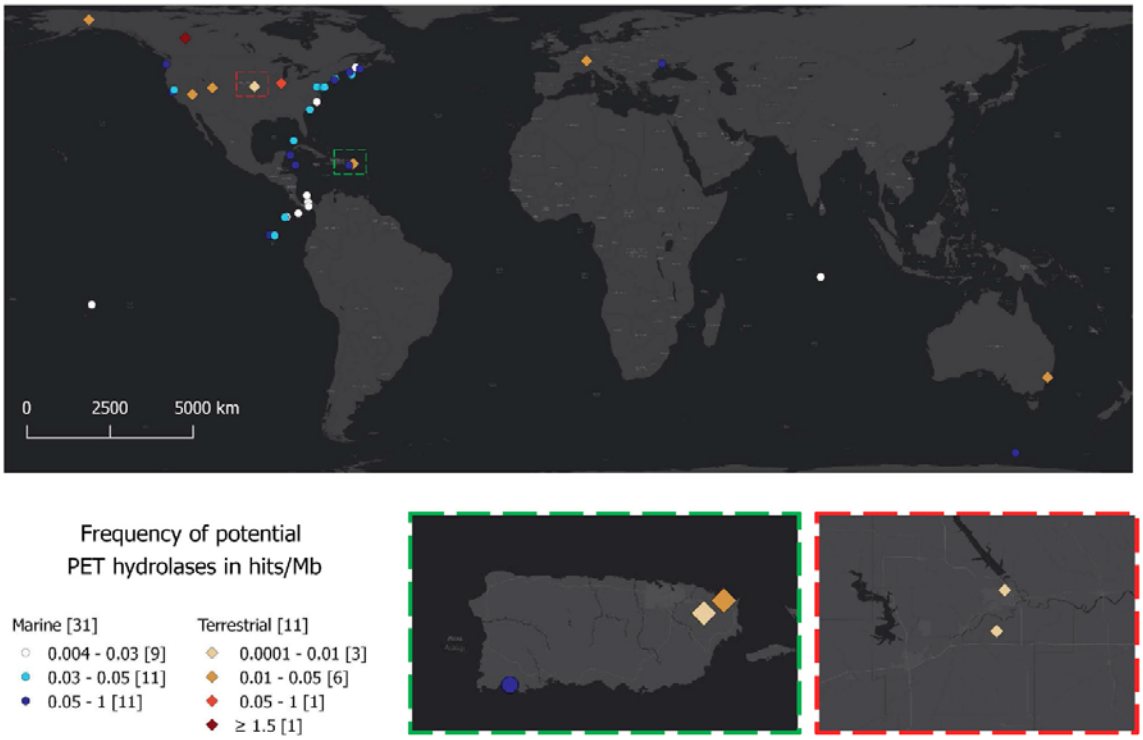

FIG 6 Global distribution of PET hydrolases in available metagenomes. Potential PET hydrolase containing metagenomes were visualized on a world map, using circles for marine and triangles for terrestrial metagenomes. Blue and red color shading indicates the frequency of PET hydrolase genes in hits/Mb for marine and terrestrial metagenomes, respectively. Red and green boxes magnify regions with overlapping spots (sample sites). (Map constructed using qGIS Desktop 2.18.5 [http://www.qgis.org].)

as the frequency of PET hydrolase homologs (Fig. 6). Within the marine and terrestrial metagenomes, PET hydrolase frequencies ranged from 0.004 to 0.92 hits/Mb and 0.0001 to 1.513 hits/Mb, respectively.

The combined genome sizes of terrestrial metagenomes are nearly 2.5-fold higher than those of the marine metagenomes, and they harbor 157 PET hydrolase homologs in average. In contrast, the marine metagenomes harbor an average of 42 PET hydrolases. The terrestrial metagenome with the highest abundance of potential PET hydrolases contains 135 sequence hits and was derived from the sediment core of a heavy oil reservoir in Canada (IMG genome number 3300001197). In the case of the marine metagenomes, the maximum was 31 hits, found within the metagenome data of a sample from the Delaware coast in the United States (Fig. 6).

We further observed that within the terrestrial habitats, the *Actinobacteria* were the main hosts for the terrestrial-derived enzymes. However, in the marine samples, most predicted PET hydrolases originated from the phylum of *Bacteroidetes* (Fig. 7). *Bacteroidetes* sequences in the marine samples were affiliated with 43% of all hits. The phylum *Proteobacteria* was the second-most abundant in both data sets, with 23% of hits in marine and 20% in terrestrial data (Fig. 7).

DISCUSSION

In this work, we developed a search algorithm that allows the *in silico* identification of PET hydrolase gene candidates from genomes and metagenomes. Altogether, we were able to identify 504 novel possible enzyme candidates in the UniProtKB and nonredundant RefSeq databases and the metagenomic database available in the NCBI

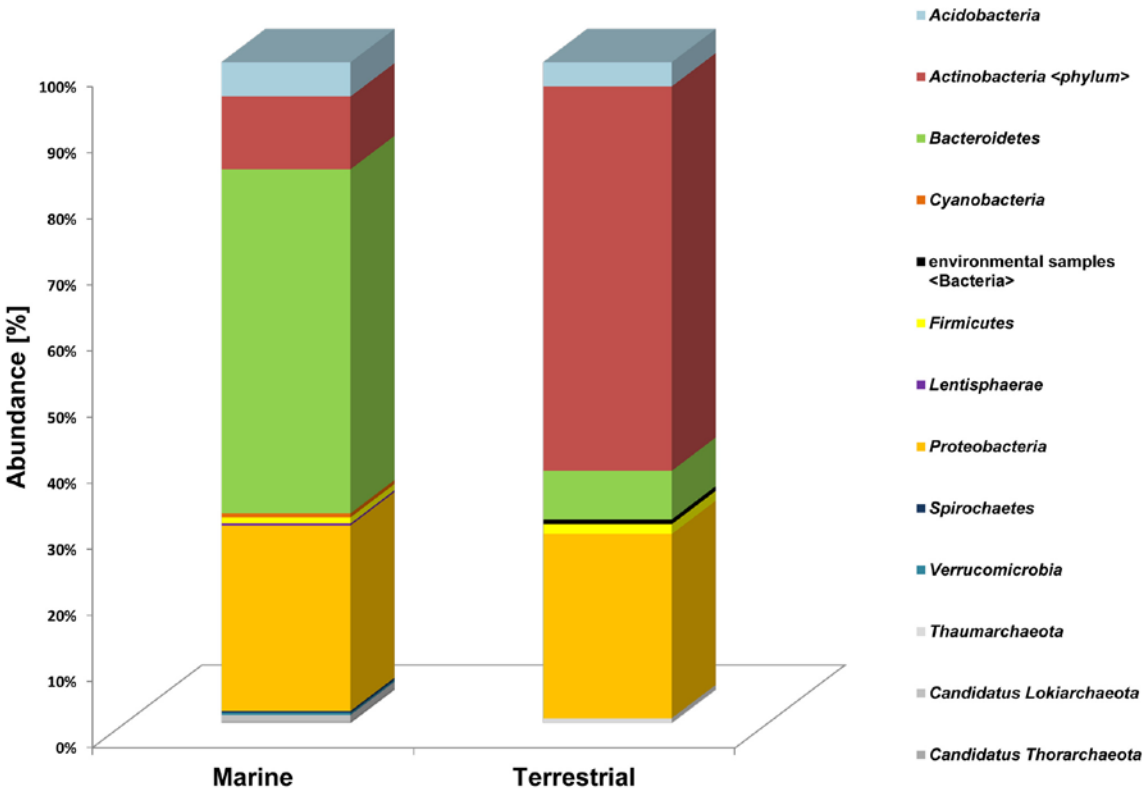

FIG 7 Phylogenetic affiliation of 349 predicted PET hydrolases from 31 marine and 11 terrestrial metagenomes. Colored and stacked bars represent the number of hits per phylum. Data were normalized per Mb of assembled DNA for the analyzed samples.

database. In addition, we identified 349 candidate genes and enzymes from marine and terrestrial metagenomes available at the IMG platform. This is by far the largest collection of PET hydrolase enzyme candidates currently available. A first classification of the PET hydrolases derived from UniProtKB/GenBank due to their protein sequence similarities and occurrence of conserved homologs enabled the formation of 17 enzyme clusters in this work (Fig. 4). An additional search of global metagenomes revealed that the PET hydrolases occur in both marine and terrestrial habitats. However, the frequencies (hits per Mb) are comparably low. The lowest hit rate was observed for a metagenome from a Kansas (United States) prairie soil sample and the highest hit rate was observed for a metagenome derived from a heavy oil reservoir sediment core. Since we included over 100 metagenomes in this analysis, the data give a reliable picture of the overall occurrence of these enzymes, but do not allow estimations on the expression of these genes in the native environment. However, the overall low gene frequencies might suggest that bacterial evolution has not yet allowed the spreading of this trait. This also implies that the overall degradation potential of the oceans is rather low compared to that of other habitats and enzymes involved in the breakdown of natural polymers like starch or cellulose (28, 29).

We verified the usefulness of the developed algorithm by cloning four novel PET hydrolase genes and expressing them heterologously in *E. coli*. All enzyme clones were active and supported the notion that the developed search algorithm is useful. The enzyme properties of the newly characterized PET hydrolases fit well into the overall picture of known PET hydrolyzing enzymes, with an optimum pH at slightly alkaline values and a preference for substrates with a short chain length (14) (30). Overall, these

novel enzymes revealed comparable activities to those already previously characterized (see Table 1 and references therein).

To our surprise, both enzymes characterized in more detail (PET2 and PET6) showed traits of thermostability (Fig. 5) (25). PET2 was stable up to 90°C, with measurable residual activity of more than 50%. This is more stable than the LCC derived from a compost metagenome (14).

All of the newly identified PET hydrolases originated mainly from three bacterial phyla, *Proteobacteria*, *Actinobacteria*, and *Bacteroidetes*. Within this framework, it is notable that the *Bacteroidetes* have so far not been associated with PET degradation, but *Bacteroidetes* species have been described as very potent degraders of other polymers, and they harbor a multitude of hydrolases and binding modules (31–33).

The restriction of PET hydrolases to a few bacterial phyla could indicate that this metabolic capability has only rather recently been evolved and is thus limited to a very few phylogenetic groups. The observation here that in the marine habitat the phylum *Bacteroidetes* is the main host of PET hydrolases is new and intriguing for several reasons. First, using classical searches and biochemical characterization, the *Actinobacteria* and *Proteobacteria* were considered to be the main hosts for these enzymes (Table 1). Second, the searches in UNIProtKB and other databases implemented on the NCBI website underlined the presence of PET hydrolases in the phyla *Actinobacteria* and *Proteobacteria*. Only when we extended our search for metagenomes of mainly noncultivated bacterial phyla did we identify the *Bacteroidetes* as the main hosts for these enzymes in the marine environment.

The recent findings on PET hydrolases described in this publication will significantly extend the knowledge of these enzymes and provide promising candidates for biotechnological applications. In summary, the over 800 enzyme candidates identified in this work will build the basis for a global repository and database of this urgently needed enzyme class.

MATERIALS AND METHODS

Bacterial strains, plasmids, and primers. Bacterial strains, plasmids, and primers used in this study are listed in Tables 3 and 4. If not otherwise mentioned, *Escherichia coli* clones were grown in LB medium (1% tryptone/peptone, 0.5% yeast extract, and 1% NaCl) supplemented with appropriate antibiotics (25 μg/ml kanamycin or 100 μg/ml ampicillin) at 37°C for 18 h.

Databases used in this study and bioinformatic analysis. Nucleotide and amino acid sequences of putative PET hydrolases were acquired from databases integrated into the NCBI (https://www.ncbi.nlm.nih.gov/), UniProt (http://www.uniprot.org/) and the Joint Genome Institute (JGI) IMG (https://img.jgi.doe.gov/) websites (34), (27, 35). Sequences were compared to others deposited in the NCBI databases using BLAST alignment tools (36). Amino acid sequence HMM search was carried out using the HMMER (http://hmmer.org/) webpage or a local version of the software (v3.1b2) with downloaded data sets. Structural information on the enzymes was retrieved from the RCSB-PDB (37) database.

Sequence data were processed using BioEdit and the Clone Manager suite version 9 (Sci-Ed Software, Denver, USA). Neighbor-joining phylogenetic trees based on amino acid sequence alignments were constructed using MEGA6 (38). Nine known and activity-confirmed bacterial PET hydrolase sequences were obtained from NCBI, aligned with T-Coffee (39) and manually revised. Afterwards, the alignment was used to construct a profile HMM with the "hmmbuild" function of the HMMER package (http://hmmer.org/). After the identification of PET hydrolase homologs, the obtained sequences were included in the above-mentioned alignment and the HMM was refined. An HMM logo was visualized using the Skylign online tool (40). Metagenomic data were downloaded from the IMG database using a Globus endpoint and were further analyzed using "hmmsearch" from the HMMER package. Phylogenetic assignment was done via a local diamond-blast search (41) against the nonredundant protein database (36) and subsequent analysis with MEGAN6 (42). The map representing the frequency and geographical distribution of PET hydrolases in metagenomes (Fig. 6) was constructed using qGIS Desktop 2.18.5 (http://www.qgis.org/).

Cloning and heterologous expression of PET2, PET5, PET6, and PET12 in *Escherichia coli* T7-SHuffle. Cloning of PET hydrolase genes into the expression vectors pET21a(+) and pET28a(+) was accomplished after amplification of genomic DNA using specific primer pairs with underlined homolog regions to the vector or restriction sites. The sequence of PET2 was obtained from NCBI (GenBank accession number ACC95208) and synthesized after codon usage optimization for *E. coli* (MWG Eurofins, Germany). Obtained DNAs were cloned into expression vectors, and the constructs were transformed into *E. coli* T7-SHuffle cells. The cultures were grown aerobically in autoinduction medium (ZYM-5052) (43) containing 100 μg/ml ampicillin and 25 μg/ml kanamycin for pET21a(+) and pET28a(+), respectively, at 37°C until they reached an optical density at 600 nm (OD_{600}) of 1.0. The proteins harboring a C- or N-terminal histidine tag were expressed afterwards at 17°C for 16 to 20 h. The cells were harvested

and lysed with pressure using a French press. Afterwards, the proteins were purified with nickel-ion affinity chromatography using nickel-nitrilotriacetic acid (Ni-NTA) agarose (Qiagen, Hilden, Germany) and analyzed by SDS-PAGE. The elution buffer was exchanged against 0.1 mM potassium phosphate buffer pH 8.0 in a 10 kDa Amicon tube (GE Health Care, Solingen, Germany).

Biochemical characterization of PET2 and PET6. For activity tests, both enzymes were assayed using purified recombinant protein. Unless otherwise indicated, the enzymes were added to a substrate solution containing 190 μl of either 0.2 M sodium phosphate buffer or 0.1 M Tris-HCl, with a defined pH between 7 and 8 and 0.5 mM *p*NP substrate dissolved in isopropanol. Incubation time ranged from 15 to 30 min. As substrates, we tested *p*NP esters with chain lengths of C_2, C_4, C_6, C_8, C_{10}, C_{12}, C_{14}, C_{16}, and C_{18}. After incubation at a defined temperature, the color change from colorless to yellow was measured at 405 nm in a plate reader (Biotek, Winooski, USA). All samples were measured in triplicate. For determination of the optimal temperature, samples were incubated between 17°C and 90°C for 15 min. The impact of pH conditions on the activity of each enzyme was measured in citrate phosphate (pH 3.0, 4.0, and 5.0), potassium phosphate (pH 6.0, 7.0, and 8.0), and carbonate bicarbonate buffer (pH 9.2 and 10.2). The influence of possible cofactors, solvents, detergents, and inhibitors was assayed at different concentration levels. After 1 h of incubation in the presence of the substances described below, the residual activity was determined after 15 min incubation at optimal temperature with *p*NP-octanoate and optimal pH. The possible cofactors Ca^{2+}, Co^{2+}, Cu^{2+}, Fe^{3+}, Mg^{2+}, Mn^{2+}, Rb^{2+}, and Zn^{2+}, with final concentrations of 1 and 10 mM, were used. To determine the solvent stability, dimethyl sulfoxide (DMSO), isopropanol, methanol, dimethylformamide (DMF), acetone, acetonitrile, and ethanol, with final concentrations of 10% and 30% (vol/vol) were added to the reaction. Detergent stability was assayed with SDS, Triton X-100, and Tween 80 at 1% and 5% (wt/vol, vol/vol) concentration. The inhibitory effects of EDTA, dithiothreitol (DTT), and PMSF were tested at 1 and 10 mM concentration. Substrate analyses using the HPLC LaChrom Elite system from Hitachi (Tokyo, Japan) with a Lichrospher 100 RP-18e column (VWR International GmbH, Darmstadt, Germany), consisting of 5-μm diameter particles, were done as previously published (14). A 14-mg low-crystallinity PET film (Goodfellow GmbH, Bad Nauheim, Germany) was used as the substrate. For enzymatic hydrolysis, up to 50 μg of protein was incubated at 60°C with continuous shaking at 500 rpm. As a mobile phase, acetonitrile (A) and water with 0.1% trifluoroacetic acid (TFA) (B) were used in a isocratic method with 20% acetonitrile (A). The reaction buffer was 0.1 M Tris-HCl (pH 7.5), with an injection volume of 99 μl. Detection was performed at 241 nm.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at https://doi.org/10.1128/AEM.02773-17.

SUPPLEMENTAL FILE 1, PDF file, 0.5 MB.

ACKNOWLEDGMENTS

This work was supported in part by the BMBF via the programs FuPol and MetaCat and by the European Union project INMARE at the University of Hamburg.

We declare that we have no conflict of interest.

REFERENCES

1. Rex WJ, Tennant DJ. March 1949. Polymeric linear terephthalic esters. US patent 2,465,319.
2. Gregory MR. 2009. Environmental implications of plastic debris in marine settings–entanglement, ingestion, smothering, hangers-on, hitch-hiking and alien invasions. Philos Trans R Soc Lond B Biol Sci 364:2013–2025. https://doi.org/10.1098/rstb.2008.0265.
3. Webb HK, Arnott J, Crawford RJ, Ivanova EP. 2013. Plastic degradation and its environmental implications with special reference to poly(ethylene terephthalate). Polymers 5:1–18. https://doi.org/10.3390/polym5010001.
4. Sax L. 2010. Polyethylene terephthalate may yield endocrine disruptors. Environ Health Perspect 118:445–448. https://doi.org/10.1289/ehp.0901253.
5. Wei R, Zimmermann W. 2017. Microbial enzymes for the recycling of recalcitrant petroleum-based plastics: how far are we? Microb Biotechnol 10:1308–1322. https://doi.org/10.1111/1751-7915.12710.
6. Kawai F, Oda M, Tamashiro T, Waku T, Tanaka N, Yamamoto M, Mizushima H, Miyakawa T, Tanokura M. 2014. A novel Ca2$^+$-activated, thermostabilized polyesterase capable of hydrolyzing polyethylene terephthalate from *Saccharomonospora viridis* AHK190. Appl Microbiol Biotechnol 98:10053–10064. https://doi.org/10.1007/s00253-014-5860-y.
7. Dresler K, van den Heuvel J, Muller RJ, Deckwer WD. 2006. Production of a recombinant polyester-cleaving hydrolase from *Thermobifida fusca* in *Escherichia coli*. Bioprocess Biosyst Eng 29:169–183. https://doi.org/10.1007/s00449-006-0069-9.
8. Chen S, Tong X, Woodard RW, Du GC, Wu J, Chen J. 2008. Identification and characterization of bacterial cutinase. J Biol Chem 283:25854–25862. https://doi.org/10.1074/jbc.M800848200.
9. Wei R, Oeser T, Then J, Kuhn N, Barth M, Schmidt J, Zimmermann W. 2014. Functional characterization and structural modeling of synthetic polyester-degrading hydrolases from *Thermomonospora curvata*. AMB Express 4:44. https://doi.org/10.1186/s13568-014-0044-9.
10. Hu X, Thumarat U, Zhang X, Tang M, Kawai F. 2010. Diversity of polyester-degrading bacteria in compost and molecular analysis of a thermoactive esterase from *Thermobifida alba* AHK119. Appl Microbiol Biotechnol 87:771–779. https://doi.org/10.1007/s00253-010-2555-x.
11. Ribitsch D, Acero EH, Greimel K, Dellacher A, Zitzenbacher S, Marold A, Rodriguez RD, Steinkellner G, Gruber K, Schwab H, Guebitz GM. 2012. A new esterase from *Thermobifida halotolerans* hydrolyses polyethylene terephthalate (PET) and polylactic acid (PLA). Polymers 4:617–629. https://doi.org/10.3390/polym4010617.
12. Acero EH, Ribitsch D, Steinkellner G, Gruber K, Greimel K, Eiteljoerg I, Trotscha E, Wei R, Zimmermann W, Zinn M, Cavaco-Paulo A, Freddi G, Schwab H, Guebitz G. 2011. Enzymatic surface hydrolysis of PET: Effect of structural diversity on kinetic properties of cutinases from *Thermobifida*. Macromolecules 44:4632–4640. https://doi.org/10.1021/ma200949p.
13. Yoshida S, Hiraga K, Takehana T, Taniguchi I, Yamaji H, Maeda Y, Toyohara K, Miyamoto K, Kimura Y, Oda K. 2016. A bacterium that degrades and assimilates poly(ethylene terephthalate). Science 351:1196–1199. https://doi.org/10.1126/science.aad6359.

14. Sulaiman S, Yamato S, Kanaya E, Kim JJ, Koga Y, Takano K, Kanaya S. 2012. Isolation of a novel cutinase homolog with polyethylene terephthalate-degrading activity from leaf-branch compost by using a metagenomic approach. Appl Environ Microbiol 78:1556–1562. https://doi.org/10.1128/AEM.06725-11.
15. Kleeberg I, Hetz C, Kroppenstedt RM, Muller RJ, Deckwer WD. 1998. Biodegradation of aliphatic-aromatic copolyesters by *Thermomonospora fusca* and other thermophilic compost isolates. Appl Environ Microbiol 64:1731–1735.
16. Wei R, Oeser T, Zimmermann W. 2014. Synthetic polyester-hydrolyzing enzymes from thermophilic actinomycetes. Adv Appl Microbiol 89:267–305. https://doi.org/10.1016/B978-0-12-800259-9.00007-X.
17. Roth C, Wei R, Oeser T, Then J, Follner C, Zimmermann W, Strater N. 2014. Structural and functional studies on a thermostable polyethylene terephthalate degrading hydrolase from *Thermobifida fusca*. Appl Microbiol Biotechnol 98:7815–7823. https://doi.org/10.1007/s00253-014-5672-0.
18. Ollis DL, Cheah E, Cygler M, Dijkstra B, Frolow F, Franken SM, Harel M, Remington SJ, Silman I, Schrag J, et al. 1992. The alpha/beta hydrolase fold. Protein Eng 5:197–211. https://doi.org/10.1093/protein/5.3.197.
19. Yang J, Yang Y, Wu WM, Zhao J, Jiang L. 2014. Evidence of polyethylene biodegradation by bacterial strains from the guts of plastic-eating waxworms. Environ Sci Technol 48:13776–13784. https://doi.org/10.1021/es504038a.
20. Yang Y, Yang J, Wu WM, Zhao J, Song YL, Gao LC, Yang RF, Jiang L. 2015. Biodegradation and mineralization of polystyrene by plastic-eating mealworms: Part 1. Chemical and physical characterization and isotopic tests. Environ Sci Technol 49:12080–12086. https://doi.org/10.1021/acs.est.5b02661.
21. Wei R, Oeser T, Schmidt J, Meier R, Barth M, Then J, Zimmermann W. 2016. Engineered bacterial polyester hydrolases efficiently degrade polyethylene terephthalate due to relieved product inhibition. Biotechnol Bioeng 113:1658–1665. https://doi.org/10.1002/bit.25941.
22. Kitadokoro K, Thumarat U, Nakamura R, Nishimura K, Karatani H, Suzuki H, Kawai F. 2012. Crystal structure of cutinase Est119 from *Thermobifida alba* AHK119 that can degrade modified polyethylene terephthalate at 1.76 angstrom resolution. Polym Degradation Stab 97:771–775. https://doi.org/10.1016/j.polymdegradstab.2012.02.003.
23. Hanahan D. 1983. Studies on transformation of *Escherichia coli* with plasmids. J Mol Biol 166:557–580. https://doi.org/10.1016/S0022-2836(83)80284-8.
24. Wei R, Oeser T, Barth M, Weigl N, Luebs A, Schulz-Siegmund M, Hacker MC, Zimmermann W. 2014. Turbidimetric analysis of the enzymatic hydrolysis of polyethylene terephthalate nanoparticles. J Mol Catal B-Enzym 103:72–78. https://doi.org/10.1016/j.molcatb.2013.08.010.
25. Meilleur C, Hupe JF, Juteau P, Shareck F. 2009. Isolation and characterization of a new alkali-thermostable lipase cloned from a metagenomic library. J Ind Microbiol Biotechnol 36:853–861. https://doi.org/10.1007/s10295-009-0562-7.
26. Baumann P, Baumann L, Bang SS, Woolkalis MJ. 1980. Reevaluation of the taxonomy of *Vibrio*, *Beneckea*, and *Photobacterium*: abolition of the genus *Beneckea*. Curr Microbiol 4:127–132. https://doi.org/10.1007/BF02602814.
27. Markowitz VM, Chen IM, Palaniappan K, Chu K, Szeto E, Grechkin Y, Ratner A, Jacob B, Huang J, Williams P, Huntemann M, Anderson I, Mavromatis K, Ivanova NN, Kyrpides NC. 2012. IMG: the Integrated Microbial Genomes database and comparative analysis system. Nucleic Acids Res 40:D115–D122. https://doi.org/10.1093/nar/gkr1044.
28. Morrison M, Pope PB, Denman SE, McSweeney CS. 2009. Plant biomass degradation by gut microbiomes: more of the same or something new? Curr Opin Biotechnol 20:358–363. https://doi.org/10.1016/j.copbio.2009.05.004.
29. Kanokratana P, Uengwetwanit T, Rattanachomsri U, Bunterngsook B, Nimchua T, Tangphatsornruang S, Plengvidhya V, Champreda V, Eurwilaichitr L. 2011. Insights into the phylogeny and metabolic potential of a primary tropical peat swamp forest microbial community by metagenomic analysis. Microb Ecol 61:518–528. https://doi.org/10.1007/s00248-010-9766-7.
30. Thumarat U, Nakamura R, Kawabata T, Suzuki H, Kawai F. 2012. Biochemical and genetic analysis of a cutinase-type polyesterase from a thermophilic *Thermobifida alba* AHK119. Appl Microbiol Biotechnol 95:419–430. https://doi.org/10.1007/s00253-011-3781-6.
31. Dodd D, Mackie RI, Cann IKO. 2011. Xylan degradation, a metabolic property shared by rumen and human colonic Bacteroidetes. Mol Microbiol 79:292–304. https://doi.org/10.1111/j.1365-2958.2010.07473.x.
32. Thomas F, Hehemann J-H, Rebuffet E, Czjzek M, Michel G. 2011. Environmental and gut *Bacteroidetes*: the food connection. Front Microbiol 2:16. https://doi.org/10.3389/fmicb.2011.00093.
33. Foley MH, Cockburn DW, Koropatkin NM. 2016. The Sus operon: a model system for starch uptake by the human gut *Bacteroidetes*. Cell Mol Life Sci 73:2603–2617. https://doi.org/10.1007/s00018-016-2242-x.
34. Coordinators NR. 2017. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res 45:D12–D17. https://doi.org/10.1093/nar/gkw1071.
35. The UniProt Consortium. 2017. UniProt: the universal protein knowledgebase. Nucleic Acids Res 45:D158–D169. https://doi.org/10.1093/nar/gkw1099.
36. Agarwala R, Barrett T, Beck J, Benson DA, Bollin C, Bolton E, Bourexis D, Brister JR, Bryant SH, Lanese K, Charowhas C, Clark K, DiCuccio M, Dondoshansky I, Federhen S, Feolo M, Funk K, Geer LY, Gorelenkov V, Hoeppner M, Holmes B, Johnson M, Khotomlianski V, Kimchi A, Kimelman M, Kitts P, Klimke W, Krasnov S, Kuznetsov A, Landrum MJ, Landsman D, Lee JM, Lipman DJ, Lu ZY, Madden TL, Madcj T, Marchler-Bauer A, Karsch-Mizrachi I, Murphy T, Orris R, Ostell J, O'Sullivan C, Panchenko A, Phan L, Preuss D, Pruitt KD, Rodarmer K, Rubinstein W, Sayers EW, Schneider V, et al. 2016. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res 44:D7–D19. https://doi.org/10.1093/nar/gkv1290.
37. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. 2000. The Protein Data Bank. Nucleic Acids Res 28:235–242. https://doi.org/10.1093/nar/28.1.235.
38. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30:2725–2729. https://doi.org/10.1093/molbev/mst197.
39. Notredame C, Higgins DG, Heringa J. 2000. T-Coffee: a novel method for fast and accurate multiple sequence alignment. J Mol Biol 302:205–217. https://doi.org/10.1006/jmbi.2000.4042.
40. Wheeler TJ, Clements J, Finn RD. 2014. Skylign: a tool for creating informative, interactive logos representing sequence alignments and profile hidden Markov models. BMC Bioinformatics 15:7. https://doi.org/10.1186/1471-2105-15-7.
41. Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. Nat Methods 12:59–60. https://doi.org/10.1038/nmeth.3176.
42. Huson DH, Beier S, Flade I, Górska A, El-Hadidi M, Mitra S, Ruscheweyh H-J, Tappu R. 2016. MEGAN Community Edition - interactive exploration and analysis of large-scale microbiome sequencing data. PLoS Comp Biol 12:e1004957. https://doi.org/10.1371/journal.pcbi.1004957.
43. Studier FW. 2005. Protein production by auto-induction in high-density shaking cultures. Protein Expression Purif 41:207–234. https://doi.org/10.1016/j.pep.2005.01.016.

7 References

Acero, E. H., Ribitsch, D., Steinkellner, G., Gruber, K., Greimel, K., Eiteljoerg, I.*, et al.* (2011). Enzymatic surface hydrolysis of PET: Effect of structural diversity on kinetic properties of cutinases from *Thermobifida*. *Macromolecules, 44*(12), 4632-4640. doi:10.1021/ma200949p

Alves, N. M., Mano, J. F., Balaguer, E., Duenas, J. M. M. & Ribelles, J. L. G. (2002). Glass transition and structural relaxation in semi-crystalline poly(ethylene terephthalate): A dsc study. *Polymer, 43*(15), 4111-4122. doi:10.1016/S0032-3861(02)00236-7

Atthoff, B. & Hilborn, J. (2007). Protein adsorption onto polyester surfaces: Is there a need for surface activation? *Journal of Biomedical Materials Research Part B-Applied Biomaterials, 80b*(1), 121-130. doi:10.1002/jbm.b.30576

Austin, H. P., Allen, M. D., Donohoe, B. S., Rorrer, N. A., Kearns, F. L., Silveira, R. L.*, et al.* (2018). Characterization and engineering of a plastic-degrading aromatic polyesterase. *Proceedings of the National Academy of Sciences of the United States of America, 115*(19), E4350-E4357. doi:10.1073/pnas.1718804115

Awaja, F. & Pavel, D. (2005). Recycling of PET. *European Polymer Journal, 41*(7), 1453-1477. doi:10.1016/j.eurpolymj.2005.02.005

Barnes, D. K. A., Galgani, F., Thompson, R. C. & Barlaz, M. (2009). Accumulation and fragmentation of plastic debris in global environments. *Philosophical Transactions of the Royal Society B-Biological Sciences, 364*(1526), 1985-1998. doi:10.1098/rstb.2008.0205

Barth, M., Honak, A., Oeser, T., Wei, R., Belisario-Ferrari, M. R., Then, J.*, et al.* (2016). A dual enzyme system composed of a polyester hydrolase and a carboxylesterase enhances the biocatalytic degradation of polyethylene terephthalate films. *Biotechnology Journal, 11*(8), 1082-1087. doi:10.1002/biot.201600008

Barth, M., Oeser, T., Wei, R., Then, J., Schmidt, J. & Zimmermann, W. (2015). Effect of hydrolysis products on the enzymatic degradation of polyethylene terephthalate nanoparticles by a polyester hydrolase from *Thermobifida fusca*. *Biochemical Engineering Journal, 93*, 222-228. doi:10.1016/j.bej.2014.10.012

Bertani, G. (1951). Studies on lysogenesis .1. The mode of phage liberation by lysogenic *Escherichia coli*. *Journal of Bacteriology, 62*(3), 293-300.

Bhardwaj, H., Gupta, R. & Tiwari, A. (2013). Communities of microbial enzymes associated with biodegradation of plastics. *Journal of Polymers and the Environment, 21*(2), 575-579. doi:10.1007/s10924-012-0456-z

Billig, S., Oeser, T., Birkemeyer, C. & Zimmermann, W. (2010). Hydrolysis of cyclic poly(ethylene terephthalate) trimers by a carboxylesterase from *Thermobifida fusca* kw3. *Applied Microbiology and Biotechnology, 87*(5), 1753-1764. doi:10.1007/s00253-010-2635-y

Borrelli, G. M. & Trono, D. (2015). Recombinant lipases and phospholipases and their use as biocatalysts for industrial applications. *International Journal of Molecular Sciences, 16*(9), 20774-20840. doi:10.3390/ijms160920774

Buchfink, B., Xie, C. & Huson, D. H. (2015). Fast and sensitive protein alignment using diamond. *Nature Methods, 12*(1), 59-60. doi:10.1038/NMETH.3176

Carniel, A., Valoni, E., Nicomedes, J., Gomes, A. D. & de Castro, A. M. (2017). Lipase from candida antarctica (calb) and cutinase from humicola insolens act synergistically for PET hydrolysis to terephthalic acid. *Process Biochemistry, 59*, 84-90. doi:10.1016/j.procbio.2016.07.023

Charles, T. C., Liles, M. R. & Sessitsch, A. (2017). *Functional metagenomics: Tools and applications*: Springer.

Chen, S., Su, L. Q., Chen, J. & Wu, J. (2013). Cutinase: Characteristics, preparation, and application. *Biotechnology Advances, 31*(8), 1754-1767. doi:10.1016/j.biotechadv.2013.09.005

Chen, S., Tong, X., Woodard, R. W., Du, G. C., Wu, J. & Chen, J. (2008). Identification and characterization of bacterial cutinase. *Journal of Biological Chemistry, 283*(38), 25854-25862. doi:DOI 10.1074/jbc.M800848200

Chen, X. L., Wang, L. M., Shi, J. G., Shi, H. & Liu, Y. L. (2011). Environmental degradation of starch/poly(lactic acid) composite in seawater. *Polymers & Polymer Composites, 19*(7), 559-565.

Chilton, T., Burnley, S. & Nesaratnam, S. (2010). A life cycle assessment of the closed-loop recycling and thermal recovery of post-consumer PET. *Resources Conservation and Recycling, 54*(12), 1241-1249. doi:10.1016/j.resconrec.2010.04.002

Coordinators, N. R. (2017). Database resources of the national center for biotechnology information. *Nucleic Acids Research, 45*(D1), D12-D17. doi:10.1093/nar/gkw1071

Danso, D., Schmeisser, C., Chow, J., Zimmermann, W., Wei, R., Leggewie, C.*, et al.* (2018). New insights into the function and global distribution of polyethylene terephthalate (PET)-degrading bacteria and enzymes in marine and terrestrial metagenomes. *Appl Environ Microbiol, 84*(8). doi:10.1128/aem.02773-17

David, L. O. & Paul, D. C. (2009). A/b-hydrolase fold: An update. *Protein & Peptide Letters, 16*(10), 1137-1148. doi:10.2174/092986609789071298

de Groot, N. S. & Ventura, S. (2006). Effect of temperature on protein quality in bacterial inclusion bodies. *FEBS letters, 580*(27), 6471-6476. doi:10.1016/j.febslet.2006.10.071

Derraik, J. G. (2002). The pollution of the marine environment by plastic debris: A review. *Mar Pollut Bull, 44*(9), 842-852. doi:10.1016/S0025-326X(02)00220-5

Don, R. H., Cox, P. T., Wainwright, B. J., Baker, K. & Mattick, J. S. (1991). 'Touchdown' pcr to circumvent spurious priming during gene amplification. *Nucleic Acids Research, 19*(14), 4008.

Dresler, K., van den Heuvel, J., Muller, R. J. & Deckwer, W. D. (2006). Production of a recombinant polyester-cleaving hydrolase from *Thermobifida fusca* in *Escherichia coli*. *Bioprocess and Biosystems Engineering, 29*(3), 169-183. doi:10.1007/s00449-006-0069-9

Eberl, A., Heumann, S., Bruckner, T., Araujo, R., Cavaco-Paulo, A., Kaufmann, F.*, et al.* (2009). Enzymatic surface hydrolysis of poly(ethylene terephthalate) and bis(benzoyloxyethyl) terephthalate by lipase and cutinase in the presence of surface active molecules. *Journal of Biotechnology, 143*(3), 207-212. doi:10.1016/j.jbiotec.2009.07.008

Eddy, S. R. (1998). Profile hidden markov models. *Bioinformatics, 14*(9), 755-763. doi:10.1093/bioinformatics/14.9.755

Eriksen, M., Lebreton, L. C. M., Carson, H. S., Thiel, M., Moore, C. J., Borerro, J. C.*, et al.* (2014). Plastic pollution in the world's oceans: More than 5 trillion plastic pieces weighing over 250,000 tons afloat at sea. *Plos One, 9*(12). doi:10.1371/journal.pone.0111913

Espino-Rammer, L., Ribitsch, D., Przylucka, A., Marold, A., Greimel, K. J., Acero, E. H.*, et al.* (2013). Two novel class ii hydrophobins from trichoderma spp. Stimulate enzymatic hydrolysis of poly(ethylene terephthalate) when expressed as fusion proteins. *Applied and Environmental Microbiology, 79*(14), 4230-4238. doi:10.1128/Aem.01132-13

Fagerburg, D. R. & Clauberg, H. (2004). Photodegradation of poly(ethylene terephthalate) and poly(ethylene/1,4-cyclohexylenedimethylene terephthalate) *Modern polyesters: Chemistry and technology of polyesters and copolyesters* (pp. 609-641): John Wiley & Sons, Ltd.

Fecker, T., Galaz-Davison, P., Engelberger, F., Narui, Y., Sotomayor, M., Parra, L. P.*, et al.* (2018). Active site flexibility as a hallmark for efficient PET degradation by *i. Sakaiensis* petase. *Biophysical Journal, 114*(6), 1302-1312. doi:10.1016/j.bpj.2018.02.005

Fernandez-Lorente, G., Palomo, J. M., Cabrera, Z., Fernandez-Lafuente, R. & Guisan, J. M. (2007). Improved catalytic properties of immobilized lipases by the presence of very low concentrations of detergents in the reaction medium. *Biotechnology and Bioengineering, 97*(2), 242-250. doi:10.1002/bit.21230

Finn, R. D., Clements, J. & Eddy, S. R. (2011). Hmmer web server: Interactive sequence similarity searching. *Nucleic Acids Research, 39*, W29-W37. doi:10.1093/nar/gkr367

Frank, J. A., Reich, C. I., Sharma, S., Weisbaum, J. S., Wilson, B. A. & Olsen, G. J. (2008). Critical evaluation of two primers commonly used for amplification of bacterial 16s rrna genes. *Applied and Environmental Microbiology, 74*(8), 2461-2470. doi:10.1128/Aem.02272-07

Grosdidier, A., Zoete, V. & Michielin, O. (2011). Swissdock, a protein-small molecule docking web service based on eadock dss. *Nucleic Acids Research, 39*(Web Server issue), W270-277. doi:10.1093/nar/gkr366

Guebitz, G. M. & Cavaco-Paulo, A. (2008). Enzymes go big: Surface hydrolysis and functionalisation of synthetic polymers. *Trends in Biotechnology, 26*(1), 32-38. doi:10.1016/j.tibtech.2007.10.003

Gullert, S., Fischer, M. A., Turaev, D., Noebauer, B., Ilmberger, N., Wemheuer, B.*, et al.* (2016). Deep metagenome and metatranscriptome analyses of microbial communities affiliated with an industrial biogas fermenter, a cow rumen, and elephant feces reveal

major differences in carbohydrate hydrolysis strategies. *Biotechnology for Biofuels, 9*. doi:10.1186/s13068-016-0534-x

GVM. (2016). Aufkommen und verwertung von PET getränkeflaschen in deutschland 2015.

Hall, T. A. (1999). *Bioedit: A user-friendly biological sequence alignment editor and analysis program for windows 95/98/nt.* Paper presented at the Nucleic acids symposium series.

Han, X., Liu, W. D., Huang, J. W., Ma, J. T., Zheng, Y. Y., Ko, T. P.*, et al.* (2017). Structural insight into catalytic mechanism of PET hydrolase. *Nature Communications, 8*. doi:ARTN 2106 10.1038/s41467-017-02255-z

Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *Journal of Molecular Biology, 166*(4), 557-580.

Harwood, C. R. & Cutting, S. M. (1990). *Molecular biological methods for Bacillus*. Chichester; New York: Wiley.

Hasan, F., Shah, A. A. & Hameed, A. (2006). Industrial applications of microbial lipases. *Enzyme and Microbial Technology, 39*(2), 235-251. doi:10.1016/j.enzmictec.2005.10.016

Hosaka, M., Kamimura, N., Toribami, S., Mori, K., Kasai, D., Fukuda, M.*, et al.* (2013). Novel tripartite aromatic acid transporter essential for terephthalate uptake in *Comamonas* sp strain e6. *Applied and Environmental Microbiology, 79*(19), 6148-6155. doi:10.1128/Aem.01600-13

Houde, A., Kademi, A. & Leblanc, D. (2004). Lipases and their industrial applications - an overview. *Applied Biochemistry and Biotechnology, 118*(1-3), 155-170. doi:10.1385/Abab:118:1-3:155

Hu, X., Thumarat, U., Zhang, X., Tang, M. & Kawai, F. (2010). Diversity of polyester-degrading bacteria in compost and molecular analysis of a thermoactive esterase from *Thermobifida alba* ahk119. *Appl Microbiol Biotechnol, 87*(2), 771-779. doi:10.1007/s00253-010-2555-x

Hug, L. A., Baker, B. J., Anantharaman, K., Brown, C. T., Probst, A. J., Castelle, C. J.*, et al.* (2016). A new view of the tree of life. *Nature Microbiology, 1*(5). doi:10.1038/Nmicrobiol.2016.48

Hunsen, M., Azim, A., Mang, H., Wallner, S. R., Ronkvist, A., Xie, W. C.*, et al.* (2007). A cutinase with polyester synthesis activity. *Macromolecules, 40*(2), 148-150. doi:10.1021/ma062095g

Huson, D. H., Beier, S., Flade, I., Górska, A., El-Hadidi, M., Mitra, S.*, et al.* (2016). Megan community edition - interactive exploration and analysis of large-scale microbiome sequencing data. *PLOS Computational Biology, 12*(6), e1004957. doi:10.1371/journal.pcbi.1004957

Ilmberger, N., Güllert, S., Dannenberg, J., Rabausch, U., Torres, J., Wemheuer, B.*, et al.* (2014). A comparative metagenome survey of the fecal microbiota of a breast-and a

plant-fed asian elephant reveals an unexpectedly high diversity of glycoside hydrolase family enzymes. *Plos One, 9*(9), e106707.

Ioakeimidis, C., Fotopoulou, K. N., Karapanagioti, H. K., Geraga, M., Zeri, C., Papathanassiou, E.*, et al.* (2016). The degradation potential of PET bottles in the marine environment: An atr-ftir based approach. *Scientific Reports, 6*. doi:ARTN 23501 10.1038/srep23501

Irwin, J. (2017). Zinc: A free database of commercially available compounds for virtual screening and ligand discovery. *Abstracts of Papers of the American Chemical Society, 254*.

Jambeck, J. R., Geyer, R., Wilcox, C., Siegler, T. R., Perryman, M., Andrady, A.*, et al.* (2015). Plastic waste inputs from land into the ocean. *Science, 347*(6223), 768-771. doi:10.1126/science.1260352

Jensen, W. A., Armstrong, J. M., Degiorgio, J. & Hearn, M. T. W. (1995). Stability studies on maize leaf phosphoenolpyruvate carboxylase - the effect of salts. *Biochemistry, 34*(2), 472-480. doi:DOI 10.1021/bi00002a011

Joo, S., Cho, I. J., Seo, H., Son, H. F., Sagong, H. Y., Shin, T. J.*, et al.* (2018). Structural insight into molecular mechanism of poly (ethylene terephthalate) degradation. *Nature Communications, 9*. doi:ARTN 38210.1038/s41467-018-02881-1

Juretschko, S., Timmermann, G., Schmid, M., Schleifer, K. H., Pommerening-Roser, A., Koops, H. P.*, et al.* (1998). Combined molecular and conventional analyses of nitrifying bacterium diversity in activated sludge: *Nitrosococcus* mobilis and nitrospira-like bacteria as dominant populations. *Applied and Environmental Microbiology, 64*(8), 3042-3051.

Kale, S. K., Deshmukh, A. G., Dudhare, M. S. & Patil, V. B. (2015). Microbial degradation of plastic: A review. *Journal of Biochemical Technology, 6*(2), 952-961.

Kanelli, M., Vasilakos, S., Nikolaivits, E., Ladas, S., Christakopoulos, P. & Topakas, E. (2015). Surface modification of poly(ethylene terephthalate) (PET) fibers by a cutinase from fusarium oxysporum. *Process Biochemistry, 50*(11), 1885-1892. doi:10.1016/j.procbio.2015.08.013

Kawai, F., Oda, M., Tamashiro, T., Waku, T., Tanaka, N., Yamamoto, M.*, et al.* (2014). A novel ca2+-activated, thermostabilized polyesterase capable of hydrolyzing polyethylene terephthalate from saccharomonospora viridis ahk190. *Applied Microbiology and Biotechnology, 98*(24), 10053-10064. doi:10.1007/s00253-014-5860-y

Khan, F. I., Lan, D., Durrani, R., Huan, W., Zhao, Z. & Wang, Y. (2017). The lid domain in lipases: Structural and functional determinant of enzymatic properties. *Front Bioeng Biotechnol, 5*, 16. doi:10.3389/fbioe.2017.00016

Kim, D. E., Chivian, D. & Baker, D. (2004). Protein structure prediction and analysis using the robetta server. *Nucleic Acids Research, 32*, W526-W531. doi:10.1093/nar/gkh468

Kitadokoro, K., Thumarat, U., Nakamura, R., Nishimura, K., Karatani, H., Suzuki, H.*, et al.* (2012). Crystal structure of cutinase est119 from *Thermobifida alba* ahk119 that can degrade modified polyethylene terephthalate at 1.76 angstrom resolution. *Polymer Degradation and Stability, 97*(5), 771-775. doi:10.1016/j.polymdegradstab.2012.02.003

Kobayashi, S. (2009). Recent developments in lipase-catalyzed synthesis of polyesters. *Macromolecular Rapid Communications, 30*(4-5), 237-266. doi:10.1002/marc.200800690

Krohn-Molt, I., Wemheuer, B., Alawi, M., Poehlein, A., Gullert, S., Schmeisser, C.*, et al.* (2013). Metagenome survey of a multispecies and alga-associated biofilm revealed key elements of bacterial-algal interactions in photobioreactors. *Applied and Environmental Microbiology, 79*(20), 6196-6206. doi:10.1128/Aem.01641-13

Laemmli, U. K. (1970). Cleavage of structural proteins during assembly of head of bacteriophage-t4. *Nature, 227*(5259), 680-+. doi:DOI 10.1038/227680a0

Lasken, R. S. & McLean, J. S. (2014). Recent advances in genomic DNA sequencing of microbial species from single cells. *Nat Rev Genet, 15*(9), 577-584. doi:10.1038/nrg3785

Li, S., Yang, X., Yang, S., Zhu, M. & Wang, X. (2012). Technology prospecting on enzymes: Application, marketing and engineering. *Comput Struct Biotechnol J, 2*, e201209017. doi:10.5936/csbj.201209017

Li, W. C., Tse, H. F. & Fok, L. (2016). Plastic waste in the marine environment: A review of sources, occurrence and effects. *Science of the Total Environment, 566*, 333-349. doi:10.1016/j.scitotenv.2016.05.084

Lim, H. C. A. (2017). Chapter 20 - thermoplastic polyesters. In M. Gilbert (Ed.), *Brydson's plastics materials (eighth edition)* (pp. 527-543): Butterworth-Heinemann.

Lopez-Mondejar, R., Zuhlke, D., Becher, D., Riedel, K. & Baldrian, P. (2016). Cellulose and hemicellulose decomposition by forest soil bacteria proceeds by the action of structurally variable enzymatic systems. *Scientific Reports, 6*. doi:10.1038/srep25279

Lucas, N., Bienaime, C., Belloy, C., Queneudec, M., Silvestre, F. & Nava-Saucedo, J. E. (2008). Polymer biodegradation: Mechanisms and estimation techniques. *Chemosphere, 73*(4), 429-442. doi:10.1016/j.chemosphere.2008.06.064

Lund, A. H., Duch, M. & Pedersen, F. S. (1996). Increased cloning efficiency by temperature-cycle ligation. *Nucleic Acids Research, 24*(4), 800-801.

Maddocks, S. E. & Oyston, P. C. (2008). Structure and function of the lysr-type transcriptional regulator (lttr) family proteins. *Microbiology, 154*(Pt 12), 3609-3623. doi:10.1099/mic.0.2008/022772-0

Manikkam, M., Tracey, R., Guerrero-Bosagna, C. & Skinner, M. K. (2013). Plastics derived endocrine disruptors (bpa, dehp and dbp) induce epigenetic transgenerational inheritance of obesity, reproductive disease and sperm epimutations. *Plos One, 8*(1). doi:ARTN e5538710.1371/journal.pone.0055387

Markowitz, V. M., Chen, I. M., Palaniappan, K., Chu, K., Szeto, E., Grechkin, Y.*, et al.* (2012). Img: The integrated microbial genomes database and comparative analysis system. *Nucleic Acids Research, 40*(Database issue), D115-122. doi:10.1093/nar/gkr1044

Methe, B. A., Nelson, K. E., Pop, M., Creasy, H. H., Giglio, M. G., Huttenhower, C.*, et al.* (2012). A framework for human microbiome research. *Nature, 486*(7402), 215-221. doi:10.1038/nature11209

Miyakawa, T., Mizushima, H., Ohtsuka, J., Oda, M., Kawai, F. & Tanokura, M. (2015). Structural basis for the ca2+-enhanced thermostability and activity of PET-degrading cutinase-like enzyme from saccharomonospora viridis ahk190. *Applied Microbiology and Biotechnology, 99*(10), 4297-4307. doi:10.1007/s00253-014-6272-8

Mueller, R. J. (2006). Biological degradation of synthetic polyesters - enzymes as potential catalysts for polyester recycling. *Process Biochemistry, 41*(10), 2124-2128. doi:10.1016/j.procbio.2006.05.018

Muller, R. J., Kleeberg, I. & Deckwer, W. D. (2001). Biodegradation of polyesters containing aromatic constituents. *Journal of Biotechnology, 86*(2), 87-95. doi:Doi 10.1016/S0168-1656(00)00407-7

Notredame, C., Higgins, D. G. & Heringa, J. (2000). T-coffee: A novel method for fast and accurate multiple sequence alignment. *Journal of Molecular Biology, 302*(1), 205-217. doi:10.1006/jmbi.2000.4042

Oeser, T., Wei, R., Baumgarten, T., Billig, S., Follner, C. & Zimmermann, W. (2010). High level expression of a hydrophobic poly(ethylene terephthalate)-hydrolyzing carboxylesterase from *Thermobifida fusca* kw3 in *Escherichia coli* bl21(de3). *Journal of Biotechnology, 146*(3), 100-104. doi:10.1016/j.jbiotec.2010.02.006

Panda, T. & Gowrishankar, B. S. (2005). Production and applications of esterases. *Applied Microbiology and Biotechnology, 67*(2), 160-169. doi:10.1007/s00253-004-1840-y

Parikh, M., Gross, R. A. & Mccarthy, S. P. (1992). The effect of crystalline morphology on enzymatic degradation kinetics. *Abstracts of Papers of the American Chemical Society, 203*, 205-Pmse.

Park, M. T., Lee, M. S., Choi, J. Y., Kim, S. C. & Lee, G. M. (2001). Orthophosphate anion enhances the stability and activity of endoxylanase from *Bacillus* sp. *Biotechnology and Bioengineering, 72*(4), 434-440. doi:10.1002/1097-0290

Perz, V., Zumstein, M. T., Sander, M., Zitzenbacher, S., Ribitsch, D. & Guebitz, G. M. (2015). Biomimetic approach to enhance enzymatic hydrolysis of the synthetic polyester poly(1,4-butylene adipate): Fusing binding modules to esterases. *Biomacromolecules, 16*(12), 3889-3896. doi:10.1021/acs.biomac.5b01219

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., *et al.* (2004). Ucsf chimera--a visualization system for exploratory research and analysis. *J Comput Chem, 25*(13), 1605-1612. doi:10.1002/jcc.20084

Pettersson, E., Lundeberg, J. & Ahmadian, A. (2009). Generations of sequencing technologies. *Genomics, 93*(2), 105-111. doi:10.1016/j.ygeno.2008.10.003

Pio, T. F. & Macedo, G. A. (2009). Cutinases: Properties and industrial applications. *Advances in Applied Microbiology, Vol 66, 66*, 77-95. doi:10.1016/S0065-2164(08)00804-6

PlasticsEurope. (2017). Plastics – the facts 2017. Retrieved from http://www.plasticseurope.org/application/files/5715/1717/4180/Plastics_the_facts_2017_FINAL_for_website_one_page.pdf website:

Podar, M., Gilmour, C. C., Brandt, C. C., Soren, A., Brown, S. D., Crable, B. R.*, et al.* (2015). Global prevalence and distribution of genes and microorganisms involved in mercury methylation. *Science Advances, 1*(9), e1500675. doi:10.1126/sciadv.1500675

R.A., W. & L., A. (2017). Degradation and metabolism of synthetic plastics and associated products by pseudomonas sp.: Capabilities and challenges. *Journal of Applied Microbiology, 123*(3), 582-593. doi:doi:10.1111/jam.13472

Rex, W. J. & Tennant, D. J. (1949). Polymeric linear terephthalic esters: Google Patents.

Ribitsch, D., Acero, E. H., Greimel, K., Dellacher, A., Zitzenbacher, S., Marold, A.*, et al.* (2012). A new esterase from *Thermobifida halotolerans* hydrolyses polyethylene terephthalate (PET) and polylactic acid (pla). *Polymers, 4*(1), 617-629. doi:10.3390/polym4010617

Ribitsch, D., Acero, E. H., Przylucka, A., Zitzenbacher, S., Marold, A., Gamerith, C.*, et al.* (2015). Enhanced cutinase-catalyzed hydrolysis of polyethylene terephthalate by covalent fusion to hydrophobins. *Applied and Environmental Microbiology, 81*(11), 3586-3592. doi:10.1128/Aem.04111-14

Ribitsch, D., Heumann, S., Trotscha, E., Acero, E. H., Greimel, K., Leber, R.*, et al.* (2011). Hydrolysis of polyethyleneterephthalate by p-nitrobenzylesterase from *Bacillus subtilis*. *Biotechnology Progress, 27*(4), 951-960. doi:10.1002/Btpr.610

Ribitsch, D., Yebra, A. O., Zitzenbacher, S., Wu, J., Nowitsch, S., Steinkellner, G.*, et al.* (2013). Fusion of binding domains to *Thermobifida cellulosilytica* cutinase to tune sorption characteristics and enhancing PET hydrolysis. *Biomacromolecules, 14*(6), 1769-1776. doi:10.1021/bm400140u

Rieckmann, T. & Völker, S. (2004). Poly(ethylene terephthalate) polymerization – mechanism, catalysis, kinetics, mass transfer and reactor design *Modern polyesters: Chemistry and technology of polyesters and copolyesters* (pp. 29-115): John Wiley & Sons, Ltd.

Ronkvist, A. M., Xie, W. C., Lu, W. H. & Gross, R. A. (2009). Cutinase-catalyzed hydrolysis of poly(ethylene terephthalate). *Macromolecules, 42*(14), 5128-5138. doi:10.1021/ma9005318

Rosano, G. L. & Ceccarelli, E. A. (2014). Recombinant protein expression in *Escherichia coli*: Advances and challenges. *Frontiers in Microbiology, 5*. doi:10.3389/fmicb.2014.00172

Roth, C., Wei, R., Oeser, T., Then, J., Follner, C., Zimmermann, W.*, et al.* (2014). Structural and functional studies on a thermostable polyethylene terephthalate degrading hydrolase from *Thermobifida fusca*. *Applied Microbiology and Biotechnology, 98*(18), 7815-7823. doi:10.1007/s00253-014-5672-0

Sambrook, J. & Russell, D. (2001). *Molecular cloning: A laboratory manual*: Cold Spring Harbor Laboratory Press.

Sasoh, M., Masai, E., Ishibashi, S., Hara, H., Kamimura, N., Miyauchi, K.*, et al.* (2006). Characterization of the terephthalate degradation genes of *Comamonas* sp strain e6. *Applied and Environmental Microbiology, 72*(3), 1825-1832. doi:10.1128/Aem.72.3.1825-1832.2006

Sax, L. (2010). Polyethylene terephthalate may yield endocrine disruptors. *Environmental Health Perspectives, 118*(4), 445-448. doi:10.1289/ehp.0901253

Schmidt, J., Wei, R., Oeser, T., Belisario-Ferrari, M. R., Barth, M., Then, J.*, et al.* (2016). Effect of tris, mops, and phosphate buffers on the hydrolysis of polyethylene terephthalate films by polyester hydrolases. *Febs Open Bio, 6*(9), 919-927. doi:10.1002/2211-5463.12097

Shaw, D. G. & Day, R. H. (1994). Color-dependent and form-dependent loss of plastic micro-debris from the north pacific-ocean. *Marine Pollution Bulletin, 28*(1), 39-43. doi:Doi 10.1016/0025-326x(94)90184-8

Singh, R., Kumar, M., Mittal, A. & Mehta, P. K. (2016). Microbial enzymes: Industrial progress in 21st century. *3 Biotech, 6*(2), 174. doi:10.1007/s13205-016-0485-8

Sinha, V., Patel, M. R. & Patel, J. V. (2010). Pet waste management by chemical recycling: A review. *Journal of Polymers and the Environment, 18*(1), 8-25. doi:10.1007/s10924-008-0106-7

Sivan, A. (2011). New perspectives in plastic biodegradation. *Current Opinion in Biotechnology, 22*(3), 422-426. doi:10.1016/j.copbio.2011.01.013

Steele, H. L., Jaeger, K.-E., Daniel, R. & Streit, W. R. (2009). Advances in recovery of novel biocatalysts from metagenomes. *Journal of Molecular Microbiology and Biotechnology, 16*(1-2), 25-37.

Studier, F. W. (2005). Protein production by auto-induction in high-density shaking cultures. *Protein Expression and Purification, 41*(1), 207-234. doi:10.1016/j.pep.2005.01.016

Sulaiman, S., Yamato, S., Kanaya, E., Kim, J. J., Koga, Y., Takano, K.*, et al.* (2012). Isolation of a novel cutinase homolog with polyethylene terephthalate-degrading activity from leaf-branch compost by using a metagenomic approach. *Applied and Environmental Microbiology, 78*(5), 1556-1562. doi:Doi 10.1128/Aem.06725-11

Sulaiman, S., You, D. J., Kanaya, E., Koga, Y. & Kanaya, S. (2014). Crystal structure and thermodynamic and kinetic stability of metagenome-derived lc-cutinase. *Biochemistry, 53*(11), 1858-1869. doi:10.1021/bi401561p

Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. (2013). Mega6: Molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution, 30*(12), 2725-2729. doi:10.1093/molbev/mst197

The UniProt, C. (2017). Uniprot: The universal protein knowledgebase. *Nucleic Acids Research, 45*(D1), D158-D169. doi:10.1093/nar/gkw1099

Then, J., Wei, R., Oeser, T., Barth, M., Belisario-Ferrari, M. R., Schmidt, J.*, et al.* (2015). Ca2+ and mg2+ binding site engineering increases the degradation of polyethylene terephthalate films by polyester hydrolases from *Thermobifida fusca*. *Biotechnology Journal, 10*(4), 592-U302. doi:10.1002/biot.201400620

Then, J., Wei, R., Oeser, T., Gerdts, A., Schmidt, J., Barth, M.*, et al.* (2016). A disulfide bridge in the calcium binding site of a polyester hydrolase increases its thermal stability

and activity against polyethylene terephthalate. *Febs Open Bio, 6*(5), 425-432. doi:10.1002/2211-5463.12053

Thompson, L. R., Sanders, J. G., McDonald, D., Amir, A., Ladau, J., Locey, K. J.*, et al.* (2017). A communal catalogue reveals earth's multiscale microbial diversity. *Nature, 551*(7681), 457-+. doi:10.1038/nature24621

Thumarat, U., Kawabata, T., Nakajima, M., Nakajima, H., Sugiyama, A., Yazaki, K.*, et al.* (2015). Comparison of genetic structures and biochemical properties of tandem cutinase-type polyesterases from *Thermobifida alba* ahk119. *J Biosci Bioeng, 120*(5), 491-497. doi:10.1016/j.jbiosc.2015.03.006

Thumarat, U., Nakamura, R., Kawabata, T., Suzuki, H. & Kawai, F. (2012). Biochemical and genetic analysis of a cutinase-type polyesterase from a thermophilic *Thermobifida alba* ahk119. *Appl Microbiol Biotechnol, 95*(2), 419-430. doi:10.1007/s00253-011-3781-6

Towbin, H., Staehelin, T. & Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets - procedure and some applications. *Proceedings of the National Academy of Sciences of the United States of America, 76*(9), 4350-4354. doi:DOI 10.1073/pnas.76.9.4350

Uchiyama, T., Abe, T., Ikemura, T. & Watanabe, K. (2005). Substrate-induced gene-expression screening of environmental metagenome libraries for isolation of catabolic genes. *Nature Biotechnology, 23*(1), 88-93. doi:10.1038/nbt1048

Venter, J. C., Remington, K., Heidelberg, J. F., Halpern, A. L., Rusch, D., Eisen, J. A.*, et al.* (2004). Environmental genome shotgun sequencing of the sargasso sea. *Science, 304*(5667), 66-74. doi:10.1126/science.1093857

Vertommen, M. A. M. E., Nierstrasz, V. A., van der Veer, M. & Warmoeskerken, M. M. C. G. (2005). Enzymatic surface modification of poly(ethylene terephthalate). *Journal of Biotechnology, 120*(4), 376-386. doi:10.1016/j.jbiotec.2005.06.015

Webb, H. K., Arnott, J., Crawford, R. J. & Ivanova, E. P. (2013). Plastic degradation and its environmental implications with special reference to poly(ethylene terephthalate). *Polymers, 5*(1), 1-18. doi:10.3390/Polym5010001

Wei, R., Oeser, T., Barth, M., Weigl, N., Luebs, A., Schulz-Siegmund, M.*, et al.* (2014). Turbidimetric analysis of the enzymatic hydrolysis of polyethylene terephthalate nanoparticles. *Journal of Molecular Catalysis B-Enzymatic, 103*, 72-78. doi:10.1016/j.molcatb.2013.08.010

Wei, R., Oeser, T., Then, J., Kuhn, N., Barth, M., Schmidt, J.*, et al.* (2014). Functional characterization and structural modeling of synthetic polyester-degrading hydrolases from *Thermomonospora curvata*. *AMB Express, 4*, 44. doi:10.1186/s13568-014-0044-9

Wei, R., Oeser, T. & Zimmermann, W. (2014). Synthetic polyester-hydrolyzing enzymes from thermophilic actinomycetes. *Advances in Applied Microbiology, Vol 89, 89*, 267-305. doi:10.1016/B978-0-12-800259-9.00007-X

Welzel, K., Müller, R. J. & Deckwer, W. D. (2002). Enzymatischer abbau von polyester-nanopartikeln. *Chemie Ingenieur Technik, 74*(10), 1496-1500. doi:10.1002/1522-2640(20021015)74:10<1496::AID-CITE1496>3.0.CO;2-P

Wheeler, T. J., Clements, J. & Finn, R. D. (2014). Skylign: A tool for creating informative, interactive logos representing sequence alignments and profile hidden markov models. *Bmc Bioinformatics, 15*. doi:10.1186/1471-2105-15-7

Yoshida, S., Hiraga, K., Takehana, T., Taniguchi, I., Yamaji, H., Maeda, Y.*, et al.* (2016). A bacterium that degrades and assimilates poly(ethylene terephthalate). *Science, 351*(6278), 1196-1199. doi:10.1126/science.aad6359

Zhang, J. F., Wang, X. C., Gong, J. X. & Gu, Z. Y. (2004). A study on the biodegradability of polyethylene terephthalate fiber and diethylene glycol terephthalate. *Journal of Applied Polymer Science, 93*(3), 1089-1096. doi:10.1002/app.20556

8 Acknowledgments

First of all, I would like to thank Prof. Dr. Wolfgang Streit for giving me the opportunity to conduct this dissertation in his group and for his constant support during all stages of the related project. He was always on hand to give advice and guidance when needed, but also showed patience and trust so that I could shape the project and pursue my own ideas.

I further would like to thank Prof. Dr. Bodo Saake and all other members of my examination committee for taking the time to evaluate my work. Beyond that, I would like to show my gratitude to all my co-workers in the department of microbiology and biotechnology. It has always been a nice and constructive working atmosphere in the lab.

My special gratitude goes to my parents, my brother Mel and Sarah for their interest in my work and their constant support. Without their help it would have been impossible for me to get to this point. Finally, I would like to thank my son Colin although he did not actively help, he always reminded me to spare some time for the family and relax after a long day in the lab.