Metagenomic approaches to discover lipolytic enzymes

Dissertation

for the award of the degree

“Doctor rerum naturalium” (Dr. rer. nat.)

of the Georg-August-Universität Göttingen

within the doctoral program in Biology

of the Georg-August University School of Science (GAUSS)

Submitted by

Mingji Lu

from ChongQing, China

Göttingen 2021

Thesis Committee

Prof. Dr. Rolf Daniel, Dept. of Genomic and Applied Microbiology & Göttingen Genomics Laboratory, Georg August University Göttingen

PD Dr. Michael Hoppert, Dept. of General Microbiology, Georg August University Göttingen

Members of the Examination Board

Reviewer: Prof. Dr. Rolf Daniel, Dept. of Genomic and Applied Microbiology & Göttingen Genomics Laboratory, Georg August University Göttingen

Second Reviewer: PD Dr. Michael Hoppert, Dept. of General Microbiology, Georg August University Göttingen

Further members of the Examination Board:

Prof. Dr. Kai Heimel, Dept. of Molecular Microbiology and Genetics, Georg August University Göttingen

Prof. Dr. Stefanie Pöggeler, Dept. Genetics of Eukaryotic Microorganisms, Georg August University Göttingen

Prof. Dr. Gerhard Braus, Dept. of Molecular Microbiology and Genetics, Georg August University Göttingen

Prof. Dr. Dieter Heineke, Assistance of the Dean and Dean of Studies, Georg August University Göttingen

Date of the oral examination: .2021.4.12

Promovierenden-Erklärung
der Georg-August-Universität Göttingen

Name .....
(Name, Vorname)
Anschrift .....
(Straße, PLZ, Wohnort)

Ich beabsichtige, eine Dissertation zum Thema .....
an der Georg-August-Universität Göttingen anzufertigen. Dabei werde ich von ..... betreut.

Ich gebe folgende Erklärung ab:

1. Die Gelegenheit zum vorliegenden Promotionsvorhaben ist mir nicht kommerziell vermittelt worden. Insbesondere habe ich keine Organisation eingeschaltet, die gegen Entgelt Betreuerinnen und Betreuer für die Anfertigung von Dissertationen sucht oder die mir obliegenden Pflichten hinsichtlich der Prüfungsleistungen für mich ganz oder teilweise erledigt.

2. Hilfe Dritter wurde bis jetzt und wird auch künftig nur in wissenschaftlich vertretbarem und prüfungsrechtlich zulässigem Ausmaß in Anspruch genommen. Insbesondere werden alle Teile der Dissertation selbst angefertigt; unzulässige fremde Hilfe habe ich dazu weder unentgeltlich noch entgeltlich entgegengenommen und werde dies auch zukünftig so halten.

3. Die Ordnung zur Sicherung der guten wissenschaftlichen Praxis an der Universität Göttingen wird von mir beachtet.

4. Eine entsprechende Promotion wurde an keiner anderen Hochschule im In- oder Ausland beantragt; die eingereichte Dissertation oder Teile von ihr wurden/werden nicht für ein anderes Promotionsvorhaben verwendet.

Mir ist bekannt, dass unrichtige Angaben die Zulassung zur Promotion ausschließen bzw. später zum Verfahrensabbruch oder zur Rücknahme des erlangten Grades führen können.

(Ort), (Datum)

(Unterschrift)

Contents

Chapter I

General Introduction

1.1 Metagenomics

We are living on a “Planet of microbes”, with microorganisms occupying almost every environmental niche and representing the largest part of the global biodiversity. It is estimated that the total number of microbial cells is more than 10^{30} (Turnbaugh et al. 2009). Prokaryotes represent the largest proportion of individual organisms. The total amount of bacterial and archaeal carbon is estimated to be almost equal to the total carbon of plants (Uritskiy and Di Ruggiero 2019) showing that microorganisms also play an essential role in biogeochemical cycles on this planet. Moreover, the genomes of microorganisms present enormous untapped genetic reservoir of novel enzymes and biomolecules promising for industrial applications (Simon and Daniel 2009; Dukunde et al. 2017).

Despite the obvious importance of microbes, the knowledge of their diversity is largely limited by the traditional culture-based methods. It is widely accepted that less than 1% microorganisms can be cultivated with standard laboratory protocols (Kumar et al. 2015). To circumvent the difficulties and limitations in cultivation techniques, a new discipline, metagenomics, has emerged as a strategic approach for direct exploring the genetic material from so far uncultured organisms (Berini et al. 2017; Laudadio et al. 2019; Almeida et al. 2019).

Metagenomics can be defined as the genomic analysis of the collective microbial assemblage found in an environmental sample (Hugenholtz and Tyson 2008). Through function-based and sequence-based approaches (Simon and Daniel 2011; Madhavan et al. 2017), metagenomics provides new insights into taxonomic and metabolic diversity of microbial communities (Fig. 1-1). According to the bibliographic analysis (Fig. 1-2), an increasing number of studies have applied metagenomic approaches to study the microbial assemblages from as many as approximately 2,192 different sites distributed across the planet (Ferrer et al. 2015).

They include terrestrial habitats (topsoil, forest soil, plant rhizosphere soil, deserts, acid mine site, etc.) (Li et al. 2009; Fang et al. 2015; Dornelas et al. 2017), aquatic environments (deep sea sediments, superficial and deep seawater, river sediment, pond water, lake water, etc.) (Minegishi et al. 2013; Biver and Vandenbol 2013; Rabausch et al. 2013; Lee et al. 2017; Zhang et al. 2017), and eukaryotic-associated microbiomes (marine sponge, termite, rumen, human microbiota, etc.) (Qin et al. 2007; Pehrsson et al. 2016; Pratama et al. 2019; Almeida et al. 2019).

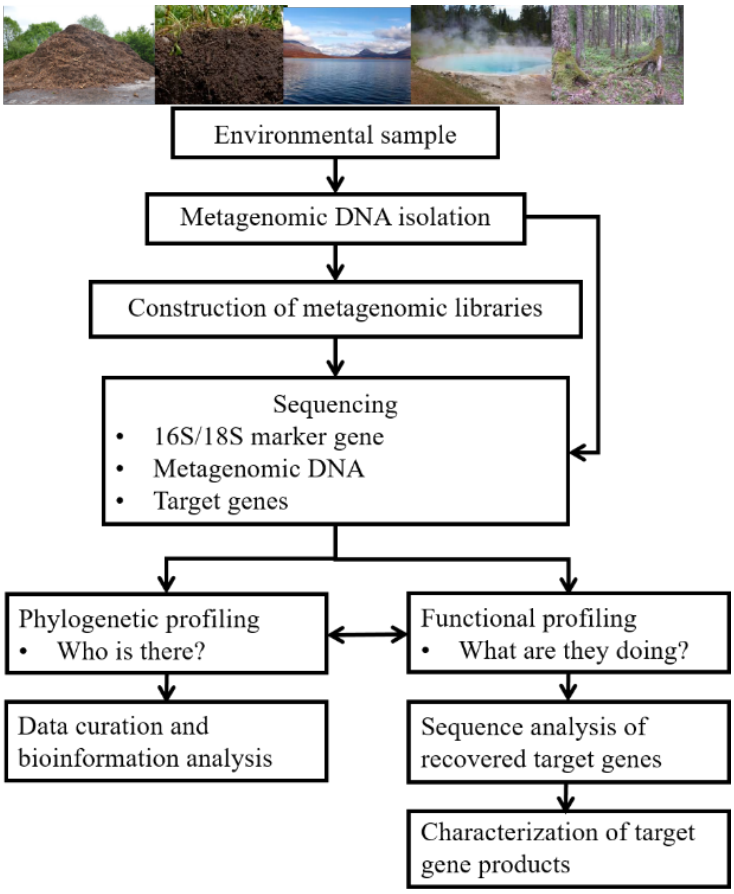


Fig. 1-1: Steps involved in metagenomic approaches (modified from Simon & Daniel, 2009). DNA is first isolated from the habitat of interest. Next, metagenomic DNA can be analyzed through different strategies, including direct sequencing and/or screening of libraries to investigate the phylogenetic and functional diversity of microorganisms.

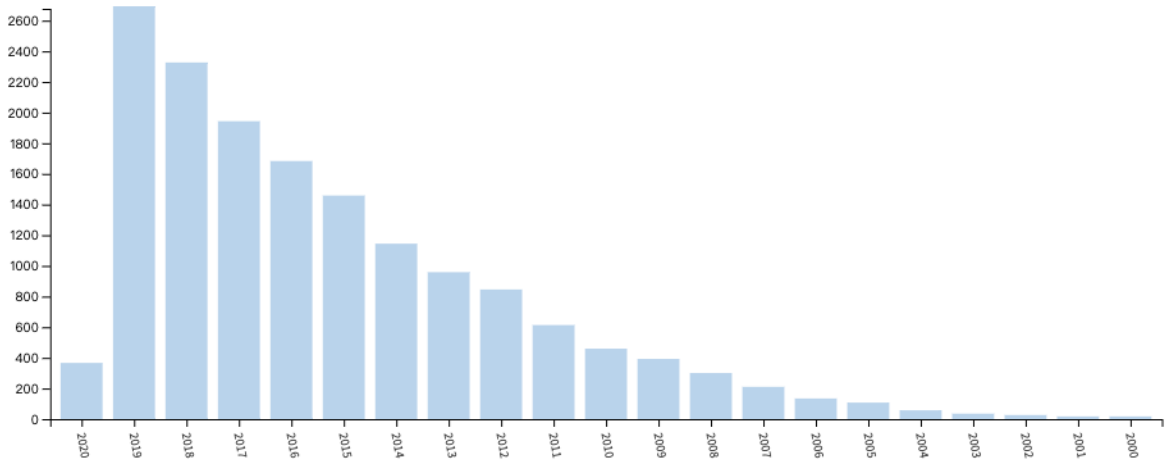


Fig. 1-2: Bibliographic analysis on metagenomics-related studies. "metagenom*" was used as keyword for searching in title and abstract in the Web of knowledge (http://apps.webofknowledge.com/). Time of data collection was 01-03-2020.

1.1.1 Function-driven approach

The function-driven approach involves screening for enzymatic activities expressed from environmental DNA in a surrogate host (Lam et al. 2015). The advantage of this approach is the certainty that the target protein is synthesized in an active form and produced by the host cell (Armour et al. 2019). In principle, it consists of cloning DNA fragments, expressing genes in a surrogate host, and screening for enzymatic activities.

First, environmental DNA is extracted, purified, size-selected and ligated into a vector. The extraction methods are divided into two types: direct and indirect (Miller et al. 1999). In comparison to the indirect isolation, the direct lysis of microbial cells is more commonly used due to the higher yields of nucleic acids (Devi et al. 2015; Satyanarayana et al. 2017). The selection of a vector depends largely on the proteins of interest during heterogeneous expression. For example, plasmids have high copy numbers and strong promotor, hence, usually used in screens in which a single gene

is responsible for the activity. Plasmids are suitable for small-insert library with inserts sizes smaller than 10 kb (Lu et al. 2019). Vectors like fosmids (Fu et al. 2011; De Santi et al. 2016a), cosmids (Craig et al. 2010; Lam et al. 2015), and bacterial artificial chromosomes (BACs) (Ufarté et al. 2015b; Berini et al. 2017) are applied to construct large-insert libraries for identifying complex activities such as multi-gene encoded products, operons and entire biochemical pathways. After ligating the environmental DNA fragments into suitable vectors, the recombinant vectors are transferred into host cells, e.g. various *Escherichia coli* strains, for heterogenous expression. Finally, novel biomolecules are identified by the following strategies: (i) phenotypic trait detection, in which positive clones are identified by the reaction or interaction of an added substance with the expressed gene product (Brady 2007; Tasse et al. 2010; Maruthamuthu et al. 2016); (ii) Heterologous complementation, which relies on the expression of foreign gene product that is vital for host strains or mutants of host strains growing under selective conditions (Donato et al. 2010; Lei et al. 2018); and (iii) induced gene expression, also called substrate-induced gene expression (SIGEX) or metabolite-regulated expression (METREX), which is a high-throughput screening methods particularly suitable for the detection of catabolic genes (Uchiyama and Miyazaki 2009). In Table 1-1, we listed examples of biomolecules identified by the three different methods.

In short, the success of function-based approach significantly depends on the expression of target genes in a foreign host. However, due to the biased expression of foreign genes in host strains such as *E. coli* (Uchiyama and Miyazaki 2009; Reyes-Duarte et al. 2012; Vargas-Albores et al. 2019), the hit rate is relatively low in function-driven screens. McMahon et al. (2012) proved that only about 40% of foreign genes were expressed in the most popular host cell (*E. coli*) in metagenomic studies. To overcome these problems, *E. coli* was improved as a screening host at

the level of both transcription and translation (Lam et al. 2015). In addition, alternative hosts such as *Streptomyces lividans* (McMahon et al. 2012), *Thermus thermophilus* (Angelov et al. 2009), *Sulfolobus solfataricus* (Albers et al. 2006), *Aspergillus oryzae* (Nagamine et al. 2019) and diverse *Proteobacteria* (Craig et al. 2010; Hao et al. 2019) have been developed to allow detection of more and diverse target enzymes (Tripathi and Shrivastava 2019). Other technological bottlenecks such as a low proportion of metagenomic DNA accessible for expression (Guazzaroni et al. 2014), a lack of relevant substrates for screening (Fernández-Arrojo et al. 2010) and a poor performance of enzymatic activities under screening conditions (Fernández-Arrojo et al. 2010) also hamper the identification of novel molecules by function-based screening (Wang et al. 2019).

Nevertheless, function-based approach is still widely used, as it allows for discovery of novel enzymes even having no known homologous structures or sequences. Particularly, new methods such as fluorescence-activated cell sorting (FACS)-driven screening, and microfluidics-driven screening have been developed to improve the sensibility and throughput of function-based screen approaches (Ngara and Zhang 2018). For example, Scanlon and his colleagues (2014) described an ultra-high-throughput screening platform employing microfluidic gel microdroplets for discovery and/or engineering of natural product antibiotics.

1.1.2 Sequence-based approach

The recent advances in high-throughput DNA sequencing technologies have significantly reduced the sequencing costs and paved the way for sequence-based screening by direct sequencing of metagenomes. This approach is often applied in two ways: targeted metagenomics or shotgun metagenomics (Bharagava et al. 2018).

Targeted metagenomics

The targeted metagenomics is employed to explore both the phylogenetic diversity and relative abundance of a particular gene in a sample (Techtmann and Hazen 2016; Bharagava et al. 2018; Awasthi et al. 2020b). Normally, targeted genes are investigated using PCR-based or hybridization-based techniques with primers/probes designed from conserved regions of known genes or gene products (Bender and Bard 2018; Dulanto Chiang and Dekker 2020). To reveal the taxonomic composition of a given community, taxonomic marker gene analysis such as 16S and 18S rRNA gene , has been applied to different types of habitats (Forbes et al. 2017; Varma et al. 2018; Schulz et al. 2019; Meng et al. 2019; Egelkamp et al. 2019). It allows in-depth comparative analysis of microbial community composition in a set of samples, and could reveal, e.g. shifts in microbial diversity before and after a perturbation (Schloss et al. 2003; Turnbaugh and Gordon 2009; Aylward et al. 2012; Garrido-Cardenas and Manzano-Agugliaro 2017). Target genes such as genes encoding lipases (López-López et al. 2015), xylose isomerases (Parachin and Gorwa-Grauslund, 2011), dioxygenases (Iwai et al. 2010; Zaprasis et al. 2010), nitrite reductases (Bartossek et al. 2010), dimethylsulfoniopropionate-degrading gene (Varaljay et al. 2010) and nitrite reductases (Bartossek et al. 2010) were also identified based on PCR amplification. Despite these proven efficiency in amplicon surveys, this approach is limited by the universality of primers (Simon and Daniel, 2009). Moreover, inherited are also the amplification related errors and artifacts, such as biases introduced by the primers and chimeric sequences resulted from co-amplification of homologous genes (Suzuki and Giovannoni, 1996; Wang and Wang 1996).

Whole genome shotgun metagenomics

In contrast to amplicon-based approach, the direct metagenomic shotgun sequencing theoretically allows profiling the taxonomic composition and the function of the entire microbial community.

Studies based on direct sequencing can be divided into two groups: read-based and assembly-based. By classifying single reads with regard to taxonomy and function, the read-based approach is suited to answer questions related to shifts of microbial community composition, and abundances of genes or metabolic pathways (Jünemann et al. 2017). There are well established tools for reference-based classification of short reads. For example, the taxonomic and functional annotation of short reads using the MG-RAST pipeline (Keegan et al. 2016) are performed by mapping representative sequences against a custom M5nr database (Wilke et al. 2012) and a BLAST-like alignment tool (Kent 2002). MEGAN (Huson et al. 2016) uses a lowest common ancestor method to assign the taxonomy/function to each read, after searches with BLAST against a reference database. CARMA3 (Zhang and Lin 2019) implements a reciprocal BLAST search and a HMMER3-based variant (against the Pfam database) for annotation. Taxator-tk (Drö Ge et al. 2015) uses a combined approach of sequence-segmented similarities to a reference dataset and an approximated phylogenetic tree for taxonomic classification.

In assembly-based metagenomics, quality-checked reads are first assembled to long, contiguous sequences (contigs) and/or scaffolds. By now, a range of assemblers were developed specialized for metagenome short reads, such as MetaVelvet (Namiki et al. 2012), Meta-IDBA (Peng et al. 2011), IDBA-UD (Peng et al. 2012), MEGAHIT (Li et al. 2015a) and metaSPAdes (Nurk et al. 2017). The resulting assemblies were clustered into so-called genomic bins for taxonomic classification (Berini et al. 2017; Almeida et al. 2019; Guo et al. 2019; Awasthi et

al. 2020b; Moreno-Indias and Tinahones 2020). To gain insights beyond taxonomic composition, a gene prediction, functional annotation, and metabolic reconstruction are done on assembled contigs. Predicted protein-encoding genes are annotated by similarity/homology searches against: 1) general protein database such as UniProt (Consortium 2018) and NCBI non-redundant database (Consortium 2018); 2) conserved domain databases such as Pfam (Finn et al. 2014) and CATH (Knudsen and Wiuf 2010); 3) metabolic databases such as COG (Clusters of Orthologous Groups of proteins) (Tatusov et al. 2000), eggNOG (evolutionary genealogy of genes: Non-supervised Orthologous Groups) (Huerta-Cepas et al. 2019), and KEGG (Kyoto Encyclopedia of Genes and Genomes) (Kanehisa et al. 2016); 4) motif databases such as MOTIF (Boeva 2016), MEME suite (Bailey et al. 2009) and HOMER (Heinz et al. 2010); and 5) special enzyme sequence resources such as Carbohydrate-Active Enzyme (CAZyme) (Cantarel et al. 2009), PeroxiBase (Fawal et al. 2013), antibiotic resistance genes (McArthur et al. 2013; Alcock et al. 2020) and lactamase (LacED) (Thai et al. 2009). In addition, several analysis platforms like MG-RAST(Keegan et al. 2016), IMG/M(Chen et al. 2017), and CAMERA (Seshadri et al. 2007) are available for the process and deposit of metagenomic data.

1.2 Lipase and esterase

1.2.1 General properties and biological functions

Lipolytic enzymes (LEs) are a diverse group of water-soluble hydrolases that catalyze the cleavage and formation of ester and even non-ester bonds (Rao et al. 2009b; Thakur 2012; Kovacic et al. 2019; Samoylova et al. 2019). They have been studied for more than 150 years since Bernard (1856) first reported the degradation of fats by mammalian pancreatic fluids. LEs can be divided into two groups with respect to the substrate specificity: lipases (EC 3.1.1.1, triacylglycerol hydrolases) and esterases (EC 3.1.1.3, carboxyl ester hydrolases) (Bornscheuer 2002). Briefly, esterases preferentially hydrolyze water soluble esters and triacylglycerols with a fatty acid chain length shorter than C_{10}. Lipases prefer water insoluble substrates, typically triacylglycerols with medium to long chain fatty acids ($\geqslant$C10) (Arpigny and Jaeger 1999; Neves Petersen et al. 2001). In addition, lipases are also distinguished from esterases by the feature of interfacial activation mediated by the hydrophobic domain (lid) covering the active site of the enzyme (Khan et al. 2017; Parapouli et al. 2018).

LEs are ubiquitous in all domains of life (Kovacic et al. 2019), and common in microorganisms. Some microbial strains are industrially important lipase/esterase producers (Rajendran et al. 2009). For example, bacterial species such as *Pseudomonas cepacian* (Secundo and Carrea 2002), *P. aeruginosa* (Ogino et al. 2004), *P. fluorescens* (Kiran et al. 2014), *P. fragi* (Sayali et al. 2013), *Bacillus thermocatenulatus* (Schmidt-Dannert et al. 1996), *B. amyloliquefaciens* (Musa et al. 2018), *Staphylococcus hyicus* (Lee et al. 2012), *S. epidermidis* (Abd Rahman et al. 2010), *S. arlettae* (Chauhan and Garlapati 2013) are used in industrial lipase/esterase production processes. Moreover, due to the catalytic versatility such as esterificatios,

transesterification and inter-esterification in aqueous and non-aqueous media, LEs have been widely applied in a variety of biotechnological applications

The structure of LEs

Winkler et al. (1990) reported the first 3D structure of a human pancreatic lipase determined by X-ray crystallography. Subsequently, a great number of LEs have been evaluated by X-ray crystallography and nuclear magnetic resonance (NMR) (Mandrich et al. 2008; López-López et al. 2015; Kim 2017). These studies have proved that almost all LEs shared a canonical α/β hydrolase fold, despite low sequence similarities (Kim 2017). As shown in Fig. 1-3, the parallel strands display a left-handed super helical twist, β3 to β8 are connected by α helices which pack on either side of the central β sheet (Dukunde et al. 2017). This topology provides a stable scaffold for positioning catalytic sites, of which the spatial positioning of side chains is remarkably well conserved (Glogauer et al. 2011).

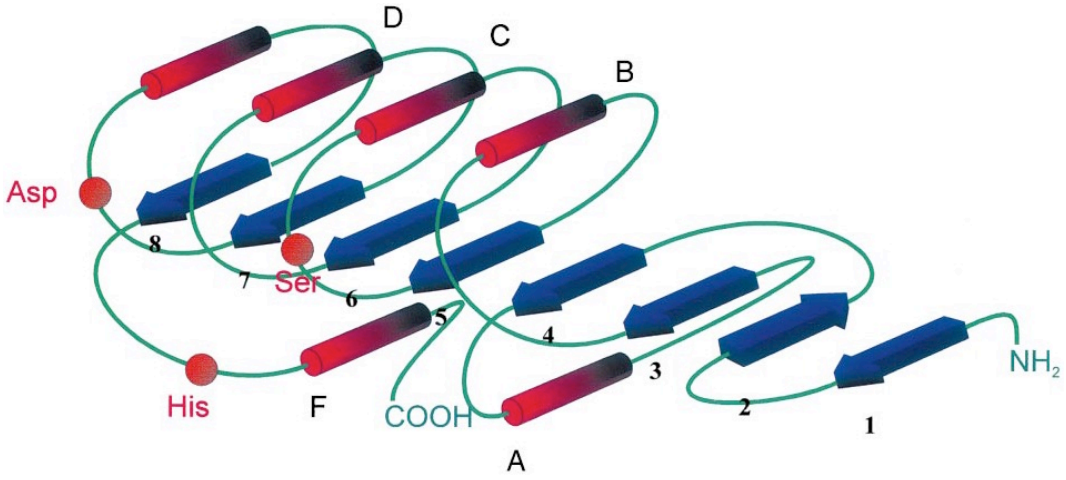


Fig. 1-3: The canonical structure of the α/β hydrolase fold. L-Sheets (1-8) are shown as blue arrows, K-helices (A-F) as red cylinders. The catalytic triad are indicated as red circles. The nucleophile serine residue locates immediately after β5, the aspartic/glutamic acid and histidine residues after β7 and β8, respectively (Ollis et al. 1992; Jaeger et al. 1999). The picture is taken from Bornscheuer (2002).

Additionally, the structural unique feature for lipases is the presence of a flexible domain (referred to as a 'flap' or 'lid') at the protein surface(Ramnath et al. 2016). The lid may consist of a single α helix or two helices flanked by a loop region. In the presence of a minimum substrate concentration, the lid moves apart, making the active site accessible to the substrate (Bornscheuer 2002), and thus, the catalytic activity of the enzyme increases. This phenomenon is designated as interfacial activation (Glogauer et al. 2011; Adlercreutz 2013; Ramnath et al. 2016).

The catalytic triad

The catalytic triad of LEs is composed of three residues: a nucleophilic residue (serine), a catalytic acid residue (aspartate or glutamate acid) and a histidine residue (Fig. 1-3). They are far apart in the primary sequence but spatially close in the folded protein (Brumlik and Buckley 1996). The serine residue is usually embedded in a highly conserved pentapeptide G-X-S-X-G. According to the 3D structure, the serine residue locates in a tight turn between sheet β5 and helix αC, which form a highly conserved β-turn-α motif termed as 'nucleophilic elbow'. Moreover, in the immediate vicinity of the serine residue, there is an oxyanion hole donating backbone amide protons to stabilize the transition state of the substrate in the enzyme (Mandrich et al. 2008). Two main types of oxyanion holes, GX and GGGX, have been found in LEs (Lu et al. 2019).

Catalytic versatility

LEs catalyze a wide range of reactions, such as ester exchange, alcoholysis, acidolysis, aminolysis, hydrolysis, esterification, intramolecular esterification, and synthesis of estolides and other polymers.

- ***Transesterification:***

(a) Ester exchange

R-COO-R' + R"-COO-R* → R-COO-R* + R"-COO-R'

(b) Alcoholysis

R-COO-R' + R"-OH → R-COO-R" + R'-OH

(c) Acidolysis

R-COO-R' + R"-COOH → R"-COO-R' + R-COOH

(d) Aminolysis

R-COO-R' + R"-NH_2 → R-CONH-R" + R'-OH

- ***Hydrolysis:***

R-COO-R' + H_2O → R-COOH + R'-OH

- ***Esterification:***

R-COOH + R'-OH → R-COO-R' + H2O

- ***Intramolecular esterification:***

HO-R-COOH → R-COO-R (*Lactones*)

- ***Synthesis of estolides and other polymers:***

R-COOH + HO-R'-COOH → R-COO-R'-COOH

1.2.2 Classification of LEs

For a long period of time, LEs were grouped simply by their substrate specificities. However, this method is too vague due to the lack of consistency for substrates and methods used in different laboratories. The increasing sequence information in public databases enables the comparison of amino acid sequences, which provides a picture about the similarity and evolutionary relationships among LEs. Thus, LEs are commonly classified based on sequence similarity, i.e. the classification system

of Arpigny and Jaeger (1999). They grouped bacterial LEs into eight families (family I-VIII) according to conserved amino acid sequence motifs and biochemical properties. A recent update to this system was an addition of 11 families (IX to XIX) (Kovacic et al. 2019; Table 1-1). Besides the nineteen families, there are claims of novel families, such as Est22 (Li et al. 2017b), Est9X (Jeon et al. 2009), LipSM54 (Li et al. 2016) and EstDZ2 (Zarafeta et al. 2016).

Table 1-1: Current classification of lipolytic enzymes.

Family	Sub-family	General description	Reference
I	I.1	➢ Similar to *Pseudomonas aeruginosa* lipase; ➢ Have a mass between 30-32kDa; ➢ Require chaperone proteins (Lifs) for expression.	Arpigny and Jaeger, 1999
	I.2	➢ Similar to *Burkholderia glumae* lipase; ➢ Have a mass larger than 32 kDa; ➢ Need Lifs for expression.	Arpigny and Jaeger, 1999
	I.3	➢ Lipases from *Pseudomonas fluorescens* and *Serratia marcescens*; ➢ Have a mass between 50 to 65 kDa.	Arpigny and Jaeger, 1999
	I.4	➢ The smallest lipases, with a molecular mass of less than 20 kDa; ➢ Several with the alternative pentapeptide motif AXSXG.	Arpigny and Jaeger, 1999
	I.5	➢ Commonly from Gram-positive prokaryotes; ➢ Have a mass of approximately 46 kDa.	Arpigny and Jaeger, 1999
	I.6	➢ Start as preproproteins (around 75 kDa) due to an amino acid N-terminal domain of 200 amino acids used as a translocator signal through the cell membrane.	Rosenstein and Götz 2000
	I.7	➢ Have a wide range of substrates, both tri- and mono-glycerides of varying fatty acid chain lengths.	Jaeger and Eggert 2002
	I.8	➢ From *Pseudoalteromonas haloplanktis*; ➢ Lacks the lid structure and Ca^{2+} pockets.	de Pascale et al. 2008
II	-	➢ Modified pentapeptide motif around the active serine: Gly-Asp-Ser-(Leu) [GDS(L)]; ➢ Secreted and membrane-bound esterases.	Mølgaard et al. 2000 Akoh et al. 2004 Li et al. 2019
III	-	➢ Extracellular esterases; ➢ Show sequence similarity (~20%) to human platelet activating factor acetylhydrolase (PAF-AH).	Arpigny and Jaeger, 1999
IV	GTSAG motif	➢ Also called hormone-sensitive lipase (HSL); ➢ Have typical motifs of HGG and GTSAG.	Li et al. 2014
	GDSAG motif	➢ Another subfamily in family IV, but with GDSAG motif.	Li et al. 2014
V	V-1	➢ Have conserved motif HGGG locates upstream of the pentapeptide motif GxSxG; ➢ High sequence similarity with non-lipolytic enzymes: epoxide hydrolases, dehalogenases and haloperoxidases;	Verschueren et al. 1993; Misawa et al. 1998

		➢ Also known as ABHD6 hydrolases.	
	V-2	➢ Known as carboxymethylbutenolide lactonase.	Arpigny and Jaeger, 1999
	V-3	➢ Uncharacterised conserved protein UCP031982, XabL type.	Park et al. 2007
VI		➢ Consists of both phospholipases and carboxylesterases with broad substrate specificity; ➢ Also featured by the small molecular masses (23–26 kDa).	Arpigny and Jaeger, 1999
VII	-	➢ Large bacterial esterases, with a molecular mass of ~ 55 kDa; ➢ Sequence homology with eukaryotic acetylcholine esterases and intestine\liver carboxylesterases.	Arpigny and Jaeger, 1999
VIII	-	➢ Show high sequence similar to class C β-lactamases; ➢ Have a molecular mass of ~42 kDa; ➢ The serine residue in the SXXK tetrapeptide.	Wagner et al. 2002; Hausmann and Jaeger 2010
IX	-	➢ New type of thermoalkalophilic lipase (PhaZ7) from *Paucimonas lemoignei*; ➢ Sshows high specificity for amorphous polyesters.	Handrick et al. 2001
X	X-1	➢ New type of thermostable esterase (EstD) from *Thermotoga maritima*; ➢ Have a molecular mass of 44.5 kDa; ➢ Optimal activity around 95 ˚C and at pH 7.	Levisson et al. 2007
	X-2	➢ Secretory lipase from *Rhodococcus sp*. strain CR-53; ➢ Display an unusual Y-type 0xyanion hole, similar to the *Candida antarctica* lipase clan.	Bassegoda et al. 2012
XI	-	➢ Presence of an Arg-Gly sequence in oxyanion hole instead of His-Gly; ➢ A signature sequence distinctive of filamentous fungal lipases (LipG).	Lee et al. 2006
XII	-	➢ Novel cold-adapted alkaline lipase from an intertidal flat metagenome; ➢ Contain an extra domain in N- or C-terminal: Bacterial Ig-like domain.	Kim et al. 2009
XIII	XIII-1	➢ A Novel Thermostable Carboxylesterase (Est30) from *Geobacillus kaustophilus* HTA426; ➢ Show no more than 17% sequence identity with the closest members in other families.	Montoro-García et al. 2009
	XIII-2	➢ Esterases from *Geobacillus stearothermophilus*, showing high thermostability due to the key salt bridges.	Charbonneau and Beauregard 2013
XIV	-	➢ A thermostable esterase (EstA3) from *Thermoanaerobacter tengcongensis*; ➢ Have a common pentapeptide CHSMG, instead of GXSXG.	Rao et al. 2011

XV	-	➢ Esterases Est10, EstGK1, EstZ3, EstD2, EstGtA2, Est5S and EstWSD belong to this family.	Bayer et al. 2010; Lee et al. 2010; Kim et al. 2012; Charbonneau and Beauregard 2013; Wang et al. 2013; Rodríguez et al. 2015
XVI	-	➢ A novel cold-adapted esterase from an Arctic intertidal metagenomic library.	Fu et al. 2013
XVII	-	➢ A novel thermophilic and halophilic esterase from Janibacter sp. R02.	Castilla et al. 2017
XVIII	-	➢ Esterase estUT1isolated from *Ureibacillus thermosphaericus*; ➢ Have a typical catalytic triad and the active serine is included in a pentapeptide (GGSVG).	Samoylova et al. 2018
XIX	-	➢ The novel, thermostable lipase (LipSm) from *Stenotrophomonas maltophilia*; ➢ Lack the requirement for interfacial activation for small substrates.	Parapouli et al. 2018
EstA	-	➢ Related to family III but different conserved motifs (pentapeptide GHSMG).	Chu et al. 2008
EstF	-	➢ Related to family V but with a modified pentapeptide, GTSXG, and different flanking regions around the HG motif.	Fu et al. 2011
EstY	-	➢ Derived from pathogenic bacteria; ➢ First possible lipolytic virulence factors that do not belong to the GDSL family; ➢ Isolated from surface river water.	Wu and Sun 2009
EM3L4	-	➢ A new esterase derived from a metagenomic library of deep-sea sediment; ➢ show only 33–58% amino acid identities to known proteins.	Jeon et al. 2011
Est9x	-	➢ A new esterase from a marine microbial metagenome of the South China Sea; ➢ Show lower than 27% sequence identities with the characterized lipolytic enzymes.	Fang et al. 2014
Est10	-	➢ A cold-adapted and salt-tolerant esterase from a psychrotrophic bacterium *Psychrobacter pacificensis.*	Wu et al. 2013b
Est12	-	➢ A novel esterase Est12 from a genomic library of a psychrotrophic *Psychrobacter celer* 3Pb1.	Wu et al. 2013b
EstGH	-	➢ A novel esterase EstGH from a metagenomic library of soil sample; ➢ Show low similarity (29%) to known esterases.	Nacke et al. 2011
EstJ	-	➢ A novel alkaliphilic esterase (EstJ) from a soil metagenome of Jeju Island; ➢ Show low similarity (32–45 %) to putative α/β hydrolases, and unique motifs of WMVSGG.	Choi et al. 2013

EstL28	-	➢ A novel cold-active esterase (EstL28) from swamp sediment metagenome, with a molecular mass of 31.3 kDa.	Seo et al. 2014
LipC	-	➢ A halophilic esterolytic enzyme LipC from archeaon *Haloarcula marismortui*.	Rao et al. 2009
LipS	-	➢ A metagenome-derived lipase (LipS) with an optimum temperature at 70 ˚C.	Chow et al. 2012
LipT	-	➢ A metagenome-derived lipase (LipT) with an optimum temperature at 75 ˚C.	Chow et al. 2012
lp_3505	-	➢ A novel esterase from *Lactobacillus plantarum*; ➢ Cold-active and salt-tolerant and show potential application for cheese ripening.	Esteban-Torres et al. 2014
PE10	-	➢ A halotolerant esterase from a marine bacterium *Pelagibacterium halotolerans* B2T.	Jiang et al. 2012
FLS18	-	➢ Two novel esterases FLS18C and FLS18D were derived from a metagenomic library of the South China Sea marine sediment.	Hu et al. 2010
RlipE1	-	➢ A novel esterase from a metagenomic library of China Holstein cow rumen.	Liu et al. 2009

Additionally, several databases are available that group LEs into different families: ESTHER database (Lenfant et al. 2013) (http://bioweb.ensam.inra.fr/ESTHER/general?what=index), microbial esterases and lipases database (MELDB) (Kang et al. 2006), Carbohydrate Active Enzymes database (CAZy) (Cantarel et al. 2009) (www.cazy.org) and Lipase Engineering Database (LED) (Fischer and Pleiss 2003)(http://www.led.uni-stuttgart.de/).

1.2.3 Microbial lipases/esterases potential for industrial applications

Thermostability

Thermostability is always one the most desirable features of LEs for industrial applications. Most of the industrial processes are carried out at temperatures above 45 ºC, generally to reduce the contaminations and accelerate the reaction rates (Vieille and Zeikus 2001; Gotor-Fernández et al. 2006; Ramnath et al. 2016). Thus, the enzymes need to be stable at this temperature and above. Thermostable/thermophilic LEs have successfully been isolated from microbial strains mainly from *Bacillus* and *Pseudomonas*, as well as metagenomes such as compost, hot spring and hydrothermal deep-sea sediment (Table 1-2).

Table 1-2: Examples of thermostable LEs isolated from cultured microorganisms and metagenomic libraries.

Source of LEs	Enzyme properties		Reference
	Optimal temp. (°C)	Optimal pH	
Microorganism			
Bacillus sp. strain L2	70	8.0	Sabri et al. 2009
B. thermoleovorans ID-1	70–75	7.5	Lee et al. 1999
Geobacillus sp.	70	9.0	Abdel-Fattah and Gaballa 2008
Pseudomonas sp.	90	11.0	Rathi et al. 2000
Pyrobaculum calidifontis	90	7.0	Hotta et al. 2002
Pyrococcus furiosus	100	-	Ikeda and Clark 1998
Pyrococcus horikoshii	95	7.0	Feng et al. 2000
Alicyclobacillus acidocaldarius	65	6	Mandrich et al. 2008
Janibacter sp. R02	80	8	Castilla et al. 2017
Sulfolobus tokodaii	80	9	Wei et al. 2013
Fervidobacterium changbaicum	75	8	Cai et al. 2011
Thermus thermophilus	80	8	Leis et al. 2015
Metagenome			
Thermal environmental samples	>95	6.0	Rhee et al. 2005
Hot spring	70	9	Tirawongsaroj et al. 2008
	60	8	Zarafeta et al. 2016
Activated sludge	70	8.5	Shao et al. 2013
Hydrothermal deep-sea sediment	60	8.0	Zhu et al. 2013
Enrichment cultures maintained at 65 to 75 °C	70,75	-	Chow et al. 2012
Red sea brine pool	65	8.5	Mohamed et al. 2013
Composts at thermophilic stage	80	7	Lu et al. 2019
	70	9	

Organic solvent tolerance

Most of the industrial relevant reactions are performed in the presence of organic solvents. It is advantageous to carry out the reactions under a water-restricted environment, due to the shifting of thermodynamic equilibria in favor of synthesis (esterification and transesterification), increasing solubility of substrates and product, improved thermal stability of the enzymes since, simpler removal of solvent (most organic solvents have lower boiling point than water), and reducing chemical waste (Dandavate et al. 2009; Ahmed et al. 2010; Ebrahimpour et al. 2011; Salihu and Alam 2015; Kumar et al. 2016). However, enzymes tend to denature or loose activity in the presence of organic solvents, as a result of removal of water molecules at the enzyme surface and the active sites by organic solvents. Thus, LEs that function in the presence of organic solvents are of particular interest for industrial applications. In recent years, a focus is to find new LEs with excellent activity, selectivity and stability in organic solvents (examples were listed in Table 1-3).

Table 1-3: Examples of organic solvent tolerant LEs isolated from microorganisms and metagenomic libraries.

Source of LEs	Enzyme properties		Reference
	Incubation condition	Stable in[a]	
Microorganism			
Pseudomonas aeruginosa san-ai	30 °C, 48 h	25 % (v/v) chloroform and *n*-hexane	Karadzic et al. 2006
Bacillus sphaericus 205y	37 °C, 30 min	25 % (v/v) *n*-hexane and *p*-xylene.	Hun et al. 2003
Bacillus megaterium	29 °C, 1 h	25–80 % (v/v) ethanol and acetone 100 % (v/v) 2-propanol, 1-butanol, Tol, *n*-hexane and *n*-heptane	Lima et al. 2004
Sulfolobus solfataricus P1	30 °C or 70 °C, 1 h	40 % (v/v) methanol, ethanol and 2-propanol.	Mandrich et al. 2005
Stenotrophomonas maltophilia CGMCC 4254	30 °C, 24 h	20 and 50 % (v/v) benzene, toluene, n-hexane and n-heptane.	Li et al. 2013
Pseudomonas aeruginosa MH38	25 °C, 1 h	30 and 50 % (v/v) benzene and hexane	Jang et al. 2014
Monascus purpureus	40 °C, 24 h	20 % (v/v) methanol, ethanol, acetonitrile, glycerol, acetone, *n*-Hexane, toluene and chloroform	Kang et al. 2017
Chromohalobacter sp.	35 °C, 30 min	20 and 50 % (v/v) benzene and hexane	Ai et al. 2018
Psychrobacter sp. ZY124	37 °C, 5 h	10, 30 and 50 % (v/v) DMSO, methanol, ethanol, acetone, acetonitrile, toluene, pentane, hexane and octane	Zhang et al. 2018b
Metagenome			
Marine mud	30 °C, 12 h	20 % (v/v) ethanol, acetonitrile, DMF and cyclohexane	Gao et al. 2016
Forest soil sample.	37 °C, 2 h	25 % (v/v) DMSO, Benzene and p-xylene	Berlemont et al. 2013
Wastewater treatment plant of a meat packing and dairy industry	4 °C, 48 h	15 and 30 % (v/v) methanol, ethanol, 1-propanol, 2-propanol, glycerol, THF, dioxane, DMSO	Glogauer et al. 2011
Soil	30 °C, 2 h	15 and 30 % (v/v) DMSO, DMF, *p*-xylene, hexane, heptane, and octane	Wang et al. 2013
Compost	Room temp., 26 d	30 % (v/v) methanol, ethanol, isopropanol, DMSO, acetone	Lu et al., 2019
Soil contaminated with petroleum hydrocarbons	No data	30 % (v/v) DMF and DMSO	Pereira et al. 2015

[a] Abbreviations for the organic solvents are: DMF, N,N-dimethylformamide; DMSO, dimethylsulfoxide; THF, tetrahydrofuran.

Halotolerance

Among extremophilic LEs, halophilic/halotolerant LEs are another major group of industrial relevant enzymes. In comparison to the non-halophilic/non-halotolerant

counterparts, halophilic/halotolerant LEs are able to successfully compete with salt ions for hydration and maintain their functional conformation in the presence of high ionic concentration (Karan et al. 2012). Thus, halophilic/halotolerant LEs can generally maintain their function in processes in which water activity is low. Some of them are also thermostable and tolerant to a wide range organic solvents (Delgado-García et al. 2012). These properties make halophilic/halotolerant LEs an option for new enzymatic processes in various products (pharmaceuticals, foods, textiles, chemicals, *etc*). In principle, mechanisms that contribute to halophilic/halotolerant enzyme stability are a high proportion of negatively charged residues (mainly aspartic and glutamic acid) located on the protein surface, conglomeration of slightly hydrophobic groups in the presence of high salt concentrations, and hydration of the protein surface due to amino acid residues with carboxylic groups (Lanyi 1974; Coquelle et al. 2010; Delgado-García et al. 2012; Munawar and Engel 2013). The examples of recently identified halophilic/halotolerant LEs are listed in Table 1-4.

Table 1-4: Examples of organic halophilic/halotolerant LEs isolated from microorganisms and metagenomic libraries.

Source of LEs	Enzyme properties		Reference
	Salt range	Maximum activity (%) [a]	
Microorganism			
Pelagibacterium halotolerans	0-4 M	~ 160 % at 3 M	Jiang et al. 2012
Alcanivorax borkumensis	0-3.5 M	100 % at 0 M	Tchigvintsev et al. 2015
Serratia sp.	0-4 M	100 % at 0 M	Jiang et al. 2016
Alkalibacterium sp.	0-4 M	~ 105 % at 2 M	Wang et al. 2016
Psychrobacter pacificensis	0-5 M	143.2 % at 2 M	Wu et al. 2013a
Lactobacillus plantarum	0-25 %	~ 250 % at 1M	Esteban-Torres et al. 2014
Zunongwangia profunda	0-4.5 M	100 % at 0 M	Rahman et al. 2016
Haloarcula marismortui	0-5 M	800 mU at 3 M	Rao et al. 2009
Thalassospira sp.	0-4 M	283 % at 3 M	De Santi et al. 2016b
Bacillus licheniformis	0-5 M	588 % at 3.5 M	Zhang et al. 2018
Metagenome			
Deep sea sponge	0-24%	100 % at 0 M	Borchert et al. 2017
Marine sponge	0-4 M	234 % at 5 M	Selvin et al. 2012
Soil	0-5 M	155 % at 1 M	Jayanath et al. 2018
Marine arctic sediment	0-4 M	675 % at 3 M	De Santi et al. 2016a
Deep-sea shrimp	0-4 M 0-4 M	~ 250 % at 3.2 M ~ 250 % at 3.6 M	Alcaide et al. 2015b
Marine water	0-3.5 M	~ 250 % at 3.5 M	Tchigvintsev et al. 2015
Desert basin soil	0-5 M	~ 140 % at 1 M	Wang et al. 2013
Marine water	0-4 M	~ 190 % at 4 M	Fang et al. 2014

a The activity measured without salt (NaCl, if not mentioned) was taken as 100 %

1.2.4 Biotechnological application of LEs

Due to the catalytic versatility, LEs are widely used in various biotechnological applications (Table 1-5).

- **Biosensor application**

Qualitative and quantitative determination of lipids and lipid-binding proteins can be used either directly or indirectly in biosensor applications. These biosensors are cheaper and less time-consuming comparing to the chemical methods.

- **Bioremediation**

LEs are used in the treatment of industrial waste and direct bioremediation of contaminated environments.

- **Food and beverage industry**

LEs are used *in situ* to improve the shelf life and rheological properties and produce aromas and emulgents. In addition, LEs are employed *ex situ* to produce flavors and improve the taste and textures of food.

- **Detergent industry**

LEs are one of the most important additives in powder and liquid detergents.

- **Paper industry**

Since early 1990, LEs are used to control the pitch in the large-scale paper-making process (Koseki et al. 2009).

- **Medical and pharmaceutical application**

Products of lipolysis (free fatty acids and diacylglycerols, etc.) can be used for diagnosing diseases, such as development of atherosclerosis, hyperlipidemia, and development of tumors (Verma N et al. 2012; Coughlan et al. 2015; Jensen et al. 2016).

- **Cosmetic industry**

LEs are involved in the synthesis of surfactants and aroma products (Guerrand 2017).

- **Agriculture**

LEs are used to synthesize intermediates for pesticides, insecticides and other agrochemical compounds.

- **Biodiesel production**

Biodiesel is an environmentally friendly alternative fuel to petroleum-based diesel. Generally, biodiesel is produced by digesting the substrates such as vegetable oils,

jatropha oil, animal fat, waste edible oil and industrial acid oil with the cocktail of enzymes (Wang et al. 2017).

- **Textile industry**

Desizing is an important process in textile industry. Traditional desizing uses acidic or oxidizing agents, which damages the cellulose material in the fabric. In comparison, the enzymatic process is more eco-friendly than the traditional method (www.wipo.int).

Table 1-5: Potential applications of LEs

Application fields	Examples	Reference
Biosensor	➢ Erythrocyte cholinesterases for exposure to nerve agents and other chemical weapons ➢ Acetylcholinesterases for detecting insecticides ➢ Amperometric biosensors	Hasan et al. 2006; Camacho et al. 2013
Bioremediation	➢ Breaking down pollutants such as fats generated from generating leather products ➢ Prokaryotic isoesterase breaking down dimethylpthalate ➢ Biofilm deposits, oil contaminated soils, lipid-rich wastewater treatment	Cammarota et al. 2013; Kim 2017; Rigoldi et al. 2018; Laudadio et al. 2019
Food and beverage industry	➢ Production of flavors in cheese ➢ Interesterification of fats and oils ➢ Accelerating the ripening of cheese and lipolysis of butter, fats, and cream ➢ Removal of fat from meat and fish products ➢ Mediating the flavor by releasing different lengths of fatty acids ➢ Emulsifiers ➢ Production of maltose and lactose like sugar fatty acid esters	Jaeger and Reetz 1998; Rajendran et al. 2009; Meng et al. 2015; Coughlan et al. 2015; De Filippis et al. 2017; Khan and Sathya 2017; Kuddus 2018
Detergent industry	➢ Cocktails of enzymes to enhancing the detergents ability to remove stains	Romdhane et al. 2010; Nerurkar et al. 2013; Bora 2014
Paper industry	➢ Eco-friendly degradation of lignin in pulp ➢ Removal of pitch and lipid stains ➢ Avoiding the formation of sticky materials ➢ Modification of raw starch	Koseki et al. 2009; Verma N et al. 2012; Ramnath et al. 2016
Medical and pharmaceutical application	➢ Enantioselective interesterification and transesterification reactions are important for selective acylation and deacylation reactions ➢ Emulsifiers ➢ Synthesizing lovastatin to lowers serum cholesterol level ➢ Diagnosis of heart ailments ➢ Detecting conditions e.g. acute pancreatitis and pancreatic injury	Gotor-Fernández et al. 2006; Gaur and Khare 2011; Coughlan et al. 2015; Luan et al. 2016; Dornelas et al. 2017
Cosmetics industry	➢ Synthesis of pentylferulate ester ➢ Emulsifiers ➢ Lipases from *Pseudomonas cepacia* for resolving the racemic rose oxides produced by the bromomethoxylation of citronellol ➢ Production of flavors	Chandel et al. 2011; Garlapati and Banerjee 2013; Rigoldi et al. 2018

	➢ Esters of cinnamic acid, ellagic acid and ferulic acid for fragrance compound	
Agriculture	➢ Phosphotriesterases from *Brevundimonas diminuta* and *Alteromonas sp.* was used extensively in detoxifying/degrading organophosphorous compounds	Horne et al. 2002; Guerrand 2017
Biodiesel production	➢ Employed as biocatalysts in biodiesel synthesis	Gaur and Khare 2011; Ramnath et al. 2016; Dornelas et al. 2017; Patel et al. 2018
Textile industry	➢ Removing size lubricants ➢ Desizing of the denim and other cotton fabrics at the commercial scale ➢ Improving the ability of polyester fabric to uptake chemical compounds	Andualema and Gessesse 2012; Xiao et al. 2017; Rigoldi et al. 2018

1.3 Thesis overview and research objects

1.3.1 Thermophilic composts for extremophilic LEs isolation

Composting is an aerobic process that accelerates the degradation of organic waste under controlled conditions (Rebollido et al. 2008). The composting process is mainly carried out by a succession of microorganisms that break down complex organic particles into simpler products. According to the temperature change, the composting process can be generally divided into three stages: mesophilic stage, thermophilic stage, and a curing/maturing stage. During the thermophilic phase, heat generated by microbial succession can raise temperatures to above 50 °C (Dougherty et al. 2012). Correspondingly, compost is a potential source for recovery of extremophilic enzymes. Recently, extremophilic enzymes, such as thermophilic, organic solvent tolerant and alkaliphilic lipolytic enzymes, have been successfully identified from compost metagenomes (Lämmle et al. 2007; Kang et al. 2011; Leis et al. 2015; Ufarté et al. 2015a; Wang et al. 2016a; Lu et al. 2019).

In this study, compost samples were collected at a composting company (Göttingen GmbH, Göttingen, Germany, 51° 34′ 25.1″ N 9° 139 54′ 33.0″ E). The sampling piles were the fermentation product of fresh tree branches or household waste. To ensure using mainly thermophilic microorganisms as a source for analysis, compost at the core zone of compost pile was collected.

1.3.2 Aim of this thesis

The aim of this study was to explore LEs in compost microbial consortia through function-driven and sequence-based approaches.

Firstly, to answer the question “who is there”, phylogenetic analyses of the two composts microbial communities were performed based on the pyrotag sequence data of 16S rRNA genes and transcripts.

Then, to identify novel lipolytic genes, functional screening of the constructed metagenomic libraries were performed. A sequence-based screening strategy based on profile HMMs were also developed. Moreover, to explore the distribution of LEs across various ecological niches, comparative analysis of performed based on the screening results of different metagenomes.

Finally, to exhibit the potential of LEs for potentially industrial application, the characterization of three LEs from different lipolytic families were carried out.

1.4 References

Abd Rahman RNZR, Ahmad Kamarudin NH, Yunus J, Salleh AB, Basri M (2010) Expression of an organic solvent stable lipase from *Staphylococcus epidermidis* AT2. Int J Mol Sci 11:3195–3208 . doi: 10.3390/ijms11093195

Abdel-Fattah YR, Gaballa AA (2008) Identification and over-expression of a thermostable lipase from *Geobacillus thermoleovorans Toshki* in *Escherichia coli*. Microbiol Res 163:13–20 . doi: 10.1016/j.micres.2006.02.004

Adlercreutz P (2013) Immobilisation and application of lipases in organic media. Chem Soc Rev 42:6406–6436 . doi: 10.1039/c3cs35446f

Ahmed EH, Raghavendra T, Madamwar D (2010) An alkaline lipase from organic solvent tolerant *Acinetobacter* sp. EH28: Application for ethyl caprylate synthesis. Bioresour Technol 101:3628–3634 . doi: 10.1016/J.BIORTECH.2009.12.107

Ai L, Huang Y, Wang C (2018) Purification and characterization of halophilic lipase of *Chromohalobacter* sp . from ancient salt well. J Basic Microbiol 58:647–657 . doi: 10.1002/jobm.201800116

Akoh CC, Lee GC, Liaw YC, Huang TH, Shaw JF (2004) GDSL family of serine esterases/lipases. Prog Lipid Res 43:534–552 . doi: 10.1016/j.plipres.2004.09.002

Albers S V., Jonuscheit M, Dinkelaker S, Urich T, Kletzin A, Tampé R, Driessen AJM, Schleper C (2006) Production of recombinant and tagged proteins in the hyperthermophilic archaeon *Sulfolobus solfataricus*. Appl Environ Microbiol 72:102–111 . doi: 10.1128/AEM.72.1.102-111.2006

Alcaide M, Tchigvintsev A, Martínez-Martínez M, Popovic A, Reva ON, Lafraya Á, Bargiela R, Nechitaylo TY, Matesanz R, Cambon-Bonavita M-A, Jebbar M, Yakimov MM, Savchenko A, Golyshina O V, Yakunin AF, Golyshin PN, Ferrer M, MAMBA Consortium (2015) Identification and characterization of carboxyl esterases of gill chamber-associated microbiota in the deep-sea shrimp Rimicaris exoculata by using functional metagenomics. Appl Environ Microbiol 81:2125–2136 . doi: 10.1128/AEM.03387-14

Alcock BP, Raphenya AR, Lau TTY, Tsang KK, Bouchard M, Edalatmand A, Huynh W, Nguyen AL V., Cheng AA, Liu S, Min SY, Miroshnichenko A, Tran HK, Werfalli RE, Nasir JA, Oloni M, Speicher DJ, Florescu A, Singh B, Faltyn M, Hernandez-Koutoucheva A, Sharma AN, Bordeleau E, Pawlowski AC, Zubyk HL, Dooley D, Griffiths E, Maguire F, Winsor GL, Beiko RG, Brinkman FSL, Hsiao WWL, Domselaar G V., McArthur AG (2020) CARD 2020: Antibiotic resistome surveillance with the comprehensive antibiotic resistance database. Nucleic Acids Res 48:D517–D525 . doi: 10.1093/nar/gkz935

Almeida A, Mitchell AL, Boland M, Forster SC, Gloor GB, Tarkowska A, Lawley TD, Finn RD (2019) A new genomic blueprint of the human gut microbiota. Nature 568:499–504 . doi: 10.1038/s41586-019-0965-1

Andualema B, Gessesse A (2012) Microbial lipases and their industrial applications: Review. Biotechnology 11:100–118 . doi: 10.3923/biotech.2012.100.118

Angelov A, Mientus M, Liebl S, Liebl W (2009) A two-host fosmid system for functional screening of (meta)genomic libraries from extreme thermophiles. Syst Appl Microbiol 32:177–185 . doi: 10.1016/j.syapm.2008.01.003

Armour CR, Nayfach S, Pollard KS, Sharpton TJ (2019) A metagenomic meta-analysis reveals functional signatures of health and disease in the human gut microbiome. mSystems 4:e00332-18 . doi: 10.1128/msystems.00332-18

Arpigny JL, Jaeger K-E (1999) Bacterial lipolytic enzymes: classification and properties. Biochem J 343:177–183 . doi: 10.1042/0264-6021:3430177

Awasthi MK, Ravindran B, Sarsaiya S, Chen H, Wainaina S, Singh E, Liu T, Kumar S, Pandey A, Singh L, Zhang Z (2020) Metagenomics for taxonomy profiling : tools and approaches. Bioengineered 11:356–374 . doi: 10.1080/21655979.2020.1736238

Aylward FO, Burnum KE, Scott JJ, Suen G, Tringe SG, Adams SM, Barry KW, Nicora CD, Piehowski PD, Purvine SO, Starrett GJ, Goodwin LA, Smith RD, Lipton MS, Currie CR (2012) Metagenomic and metaproteomic insights into bacterial communities in leaf-cutter ant fungus gardens. ISME J 6:1688–1701 . doi: 10.1038/ismej.2012.10

Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS (2009) MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res 37:W202–W208 . doi: 10.1093/nar/gkp335

Bassegoda A, Pastor FIJ, Diaz P (2012) Rhodococcus sp. strain CR-53 lipr, the first member of a new bacterial lipase family (Family X) displaying an unusual Y-type oxyanion hole, similar to the *Candida antarctica* lipase clan. Appl Environ Microbiol 78:1724–1732 . doi: 10.1128/AEM.06332-11

Bayer S, Kunert A, Ballschmiter M, Greiner-Stoeffele T (2010) Indication for a new lipolytic enzyme family: Isolation and characterization of two esterases from a metagenomic library. J Mol Microbiol Biotechnol 18:181–187 . doi: 10.1159/000315459

Bender JM, Bard JDi (2018) Metagenomics in pediatrics: Using a shotgun approach to diagnose infections. Curr Opin Pediatr 30:125–130 . doi: 10.1097/MOP.0000000000000577

Berini F, Casciello C, Marcone GL, Marinelli F (2017) Metagenomics: novel enzymes from non-culturable microbes. FEMS Microbiol Lett 364: . doi: 10.1093/femsle/fnx211

Berlemont R, Spee O, Delsaute M, Lara Y, Schuldes J, Simon C, Power P, Daniel R, Galleni M (2013) Novel organic solvent-tolerant esterase isolated by metagenomics: insights into the lipase/esterase classification. Rev Argent Microbiol 45:3–12

Bharagava RN, Purchase D, Saxena G, Mulla SI (2018) Applications of metagenomics in microbial bioremediation of pollutants: from genomics to environmental cleanup. In: Das S, Dash HR (eds) Microbial Diversity in the Genomic Era. Elsevier Inc., Oxford, UK, pp 459–477

Biver S, Vandenbol M (2013) Characterization of three new carboxylic ester hydrolases isolated by functional screening of a forest soil metagenomic library. J Ind Microbiol Biotechnol 40:191–200 . doi: 10.1007/s10295-012-1217-7

Boeva V (2016) Analysis of genomic sequence motifs for deciphering transcription factor binding and transcriptional regulation in Eukaryotic cells. Front. Genet. 7:24

Bora L (2014) Purification and characterization of highly alkaline lipase from bacillus licheniformis MTCC 2465: And study of its detergent compatibility and applicability. J Surfactants Deterg 17:889–898 . doi: 10.1007/s11743-013-1517-6

Borchert E, Selvin J, Kiran SG, Jackson SA, O'Gara F, Dobson ADW (2017) A novel cold active esterase from a deep sea sponge *Stelletta normani* metagenomic library. Front Mar Sci 4:1–13 . doi: 10.3389/fmars.2017.00287

Bornscheuer UT (2002) Microbial carboxyl esterases: classification, properties and application in biocatalysis. FEMS Microbiol Rev 26:73–81 . doi: 10.1111/j.1574-6976.2002.tb00599.x

Brady SF (2007) Construction of soil environmental DNA cosmid libraries and screening for clones that produce biologically active small molecules. Nat Protoc 2:1297–1305 . doi: 10.1038/nprot.2007.195

Brumlik MJ, Buckley JT (1996) Identification of the catalytic triad of the lipase/acyltransferase from *Aeromonas hydrophila*. J Bacteriol 178:2060–2064 . doi: 10.1128/jb.178.7.2060-2064.1996

Cai J, Xie Y, Song B, Wang Y, Zhang Z, Feng Y (2011) *Fervidobacterium changbaicum* Lip1: Identification, cloning, and characterization of the thermophilic lipase as a new member of bacterial lipase family V. Appl Microbiol Biotechnol 89:1463–1473 . doi: 10.1007/s00253-010-2971-y

Cammarota MC, Rosa DR, Duarte ICS, Saavedra NK, Varesche MBA, Zaiat M, Freire DMG (2013) The effect of enzymatic pre-hydrolysis of dairy wastewater on the granular and immobilized microbial community in anaerobic bioreactors. Environ Technol 34:417–428 . doi: 10.1080/09593330.2012.698649

Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B (2009) The carbohydrate-active enZymes database (CAZy): an expert resource for glycogenomics. Nucleic Acids Res 37:233–238

Castilla A, Panizza P, Rodríguez D, Bonino L, Díaz P, Irazoqui G, Rodríguez Giordano S (2017) A novel thermophilic and halophilic esterase from *Janibacter* sp. R02, the first member of a new lipase family (Family XVII). Enzyme Microb Technol 98:86–95 . doi: 10.1016/J.ENZMICTEC.2016.12.010

Chandel C, Kumar A, Kanwar SS (2011) Enzymatic synthesis of butyl ferulate by silica-immobilized lipase in a non-aqueous medium. J Biomater Nanobiotechnol 2:400–408 . doi: 10.4236/jbnb.2011.24049

Charbonneau DM, Beauregard M (2013) Role of key salt bridges in thermostability of *G. thermodenitrificans* EstGtA2: distinctive patterns within the new bacterial lipolytic enzyme family XV. PLoS One 8:e76675 . doi: 10.1371/journal.pone.0076675

Charoenpanich J, Soongrung T, Chinnasri S, Suebchuea N, Suppoontong M, Thiemsawait S (2018) A novel broad-temperature active and solvent stable esterase from a newly isolated *Bacillus aerophilus*. Biocatal Agric Biotechnol 13:116–122 . doi: 10.1016/j.bcab.2017.12.004

Chauhan M, Garlapati VK (2013) Production and characterization of a halo-, solvent-, thermo-tolerant alkaline lipase by *Staphylococcus arlettae* JPBW-1, isolated from rock salt mine. Appl Biochem Biotechnol 171:1429–1443 . doi: 10.1007/s12010-013-0433-6

Chen IMA, Markowitz VM, Chu K, Palaniappan K, Szeto E, Pillay M, Ratner A, Huang J, Andersen E, Huntemann M, Varghese N, Hadjithomas M, Tennessen K, Nielsen T, Ivanova NN, Kyrpides NC (2017) IMG/M: Integrated genome and metagenome comparative data analysis system. Nucleic Acids Res 45:507–516 . doi: 10.1093/nar/gkw929

Choi JE, Kwon MA, Na HY, Hahm DH, Song JK (2013) Isolation and characterization of a metagenome-derived thermoalkaliphilic esterase with high stability over a broad pH range. Extremophiles 17:1013–1021 . doi: 10.1007/s00792-013-0583-z

Chow J, Kovacic F, Dall Antonia Y, Krauss U, Fersini F, Schmeisser C, Lauinger B, Bongen P, Pietruszka J, Schmidt M, Menyes I, Bornscheuer UT, Eckstein M, Thum O, Liese A, Mueller-Dieckmann J, Jaeger KE, Streit WR (2012) The metagenome-derived enzymes LipS and LipT increase the diversity of known lipases. PLoS One 7:e47665 . doi: 10.1371/journal.pone.0047665

Chu X, He H, Guo C, Sun B (2008) Identification of two novel esterases from a marine metagenomic library derived from South China Sea. Appl Microbiol Biotechnol 80:615–625 . doi: 10.1007/s00253-008-1566-3

The UniProt Consortium (2018) UniProt: The universal protein knowledgebase. Nucleic Acids Res 45:D158–D169 . doi: 10.1093/nar/gkw1099

Coquelle N, Talon R, Juers DH, Girard É, Kahn R, Madern D (2010) Gradual adaptive changes of a protein facing high salt concentrations. J Mol Biol 404:493–505 . doi: 10.1016/j.jmb.2010.09.055

Coughlan LM, Cotter PD, Hill C, Alvarez-Ordóñez A (2015) Biotechnological applications of functional metagenomics in the food and pharmaceutical industries. Front Microbiol 6:672 . doi: 10.3389/fmicb.2015.00672

Craig JW, Chang FY, Kim JH, Obiajulu SC, Brady SF (2010) Expanding small-molecule functional metagenomics through parallel screening of broad-host-range cosmid environmental DNA libraries in diverse *Proteobacteria*. Appl Environ Microbiol 76:1633–1641 . doi: 10.1128/AEM.02169-09

Dandavate V, Jinjala J, Keharia H, Madamwar D (2009) Production, partial purification and characterization of organic solvent tolerant lipase from *Burkholderia multivorans* V2 and its application for ester synthesis. Bioresour Technol 100:3374–3381 . doi: 10.1016/j.biortech.2009.02.011

De Filippis F, Parente E, Ercolini D (2017) Metagenomics insights into food fermentations. Microb Biotechnol 10:91–102 . doi: 10.1111/1751-7915.12421

de Pascale D, Cusano AM, Autore F, Parrilli E, di Prisco G, Marino G, Tutino ML (2008) The cold-active Lip1 lipase from the Antarctic bacterium *Pseudoalteromonas haloplanktis* TAC125 is a member of a new bacterial lipolytic enzyme family. Extremophiles 12:311–323 . doi: 10.1007/s00792-008-0163-9

De Santi C, Altermark B, Pierechod MM, Ambrosino L, de Pascale D, Willassen N-P (2016a) Characterization of a cold-active and salt tolerant esterase identified by functional screening of Arctic metagenomic libraries. BMC Biochem 17: . doi: 10.1186/s12858-016-0057-x

De Santi C, Leiros HKS, Di Scala A, de Pascale D, Altermark B, Willassen NP (2016b) Biochemical characterization and structural analysis of a new cold-active and salt-tolerant esterase from the marine bacterium *Thalassospira* sp. Extremophiles 20:323–336 . doi: 10.1007/s00792-016-0824-z

Delgado-García M, Valdivia-Urdiales B, Aguilar-González CN, Contreras-Esquivel JC, Rodríguez-Herrera R (2012) Halophilic hydrolases as a new tool for the biotechnological industries. J Sci Food Agric 92:2575–2580 . doi: 10.1002/jsfa.5860

Devi SG, Fathima AA, Radha S, Arunraj R, Curtis WR, Ramya M (2015) A rapid and economical method for efficient DNA extraction from diverse soils suitable for metagenomic applications. PLoS One 10:e0132441 . doi: 10.1371/journal.pone.0132441

Donato JJ, Moe LA, Converse BJ, Smart KD, Berklein EC, McManus PS, Handelsman J (2010) Metagenomic analysis of apple orchard soil reveals antibiotic resistance genes encoding predicted bifunctional proteins. Appl Environ Microbiol 76:4396–4401 . doi: 10.1128/AEM.01763-09

Dornelas JCM, Figueiredo JEF, De Abreu CS, Lana UGP, Oliveira CA, Marriel IE (2017) Characterization and phylogenetic affiliation of *Actinobacteria* from tropical soils with potential uses for agro-industrial processes. Genet Mol Res 16:1–16 . doi: 10.4238/gmr16039703

Dougherty MJ, D'haeseleer P, Hazen TC, Simmons BA, Adams PD, Hadi MZ (2012) Glycoside hydrolases from a targeted compost metagenome, activity-screening and functional characterization. BMC Biotechnol 12:38 . doi: 10.1186/1472-6750-12-38

Drö Ge J, Gregor I, Mchardy AC (2015) Taxator-tk: precise taxonomic assignment of metagenomes by fast approximation of evolutionary neighborhoods. Bioinformatics 31:817–824 . doi: 10.1093/bioinformatics/btu745

Dukunde A, Schneider D, Lu M, Brady S, Daniel R (2017) A novel, versatile family IV carboxylesterase exhibits high stability and activity in a broad pH spectrum. Biotechnol Lett 39:577–587 . doi: 10.1007/s10529-016-2282-1

Dulanto Chiang A, Dekker JP (2020) From the pipeline to the bedside: advances and challenges in clinical metagenomics. J Infect Dis 221:S331–S340 . doi: 10.1093/infdis/jiz151

Ebrahimpour A, Rahman RNZRA, Basri M, Salleh AB (2011) High level expression and characterization of a novel thermostable, organic solvent tolerant, 1,3-regioselective lipase from *Geobacillus* sp. strain ARM. Bioresour Technol 102:6972–6981 . doi: 10.1016/j.biortech.2011.03.083

Egelkamp R, Zimmermann T, Schneider D, Hertel R, Daniel R (2019) Impact of nitriles on bacterial communities. Front Environ Sci 7:103 . doi: 10.3389/fenvs.2019.00103

Esteban-Torres M, Santamaría L, de las Rivas B, Muñoz R (2014) Characterisation of a cold-active and salt-tolerant esterase from *Lactobacillus plantarum* with potential application during cheese ripening. Int Dairy J 39:312–315 . doi: 10.1016/j.idairyj.2014.08.004

Fang Y, Wang S, Liu S, Jiao Y (2015) Discovery a novel organic solvent tolerant esterase from *Salinispora arenicola* CNP193 through genome mining. Int J Biol Macromol 80:334–340 . doi: 10.1016/j.ijbiomac.2015.06.045

Fang Z, Li J, Wang Q, Fang W, Peng H, Zhang X, Xiao Y (2014) A novel esterase from a marine metagenomic library exhibiting salt tolerance ability. J Microbiol Biotechnol 24:771–780 . doi: 10.4014/jmb.1311.11071

Fawal N, Li Q, Savelli B, Brette M, Passaia G, Fabre M, Mathé C, Dunand C (2013) PeroxiBase: A database for large-scale evolutionary analysis of peroxidases. Nucleic Acids Res 41:441–444 . doi: 10.1093/nar/gks1083

Feng Y, Joh Y-G, Ishikawa K, Ishida H, Ando S, Yamagaki T, Nakanishi H, Cao S, Matsui I, Kosugi Y (2000) Thermophilic phospholipase A2 in the cytosolic fraction from the archaeon *Pyrococcus horikoshii*. J Am Oil Chem Soc 77:1147–1152 . doi: 10.1007/s11746-000-0180-8

Ferrer M, Martínez-Martínez M, Bargiela R, Streit WR, Golyshina O V, Golyshin PN (2015) Estimating the success of enzyme bioprospecting through metagenomics: current status and future trends. Microb Biotechnol 9:22–34 . doi: 10.1111/1751-7915.12309

Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J, Sonnhammer ELL, Tate J, Punta M (2014) Pfam: the protein families database. Nucleic Acids Res 42:D222–D230 . doi: 10.1093/nar/gkt1223

Fischer M, Pleiss J (2003) The Lipase Engineering Database: a navigation and analysis tool for protein families. Nucleic Acids Res 31:319–321

Forbes JD, Knox NC, Ronholm J, Pagotto F, Reimer A (2017) Metagenomics: The next culture-independent game changer. Front Microbiol 8:1069 . doi: 10.3389/fmicb.2017.01069

Fu C, Hu Y, Xie F, Guo H, Ashforth EJ, Polyak SW, Zhu B, Zhang L (2011) Molecular cloning and characterization of a new cold-active esterase from a deep-sea metagenomic library. Appl Microbiol Biotechnol 90:961–970 . doi: 10.1007/s00253-010-3079-0

Fu J, Leiros HKS, De Pascale D, Johnson KA, Blencke HM, Landfald B (2013) Functional and structural studies of a novel cold-adapted esterase from an Arctic intertidal metagenomic library. Appl Microbiol Biotechnol 97:3965–3978 . doi: 10.1007/s00253-012-4276-9

Gao W, Wu K, Chen L, Fan H, Zhao Z, Gao B, Wang H, Wei D (2016) A novel esterase from a marine mud metagenomic library for biocatalytic synthesis of short-chain flavor esters. Microb Cell Fact 15:41 . doi: 10.1186/s12934-016-0435-5

Garlapati VK, Banerjee R (2013) Solvent-free synthesis of flavour esters through immobilized lipase mediated transesterification. Enzyme Res 2013: . doi: 10.1155/2013/367410

Garrido-Cardenas JA, Manzano-Agugliaro F (2017) The metagenomics worldwide research. Curr Genet 63:819–829 . doi: 10.1007/s00294-017-0693-8

Gaur R, Khare SK (2011) Solvent tolerant *Pseudomonads* as a source of novel lipases for applications in non-aqueous systems. Biocatal Biotransformation 29:161–171 . doi: 10.3109/10242422.2011.609588

Gilbert EJ (1993) Pseudomonas lipases: Biochemical properties and molecular cloning. Enzyme Microb. Technol. 15:634–645

Glogauer A, Martini VP, Faoro H, Couto GH, Müller-Santos M, Monteiro RA, Mitchell DA, de Souza EM, Pedrosa FO, Krieger N (2011) Identification and characterization of a new true lipase isolated through metagenomic approach. Microb Cell Fact 10:54 . doi: 10.1186/1475-2859-10-54

Gotor-Fernández V, Brieva R, Gotor V (2006) Lipases: Useful biocatalysts for the preparation of pharmaceuticals. J Mol Catal B Enzym 40:111–120 . doi: 10.1016/j.molcatb.2006.02.010

Guerrand D (2017) Lipases industrial applications: Focus on food and agroindustries. OCL 24:D403 . doi: 10.1051/ocl/2017031

Guo J, Quensen JF, Sun Y, Wang Q, Brown CT, Cole JR, Tiedje JM (2019) Review, evaluation, and directions for gene-targeted assembly for ecological analyses of metagenomes. Front Genet 10:957 . doi: 10.3389/fgene.2019.00957

Handrick R, Reinhardt S, Focarete ML, Scandola M, Adamus G, Kowalczuk M, Jendrossek D (2001) A new type of thermoalkalophilic hydrolase of *Paucimonas lemoignei* with high

specificity for amorphous polyesters of short chain-length hydroxyalkanoic acids. J Biol Chem 276:36215–36224 . doi: 10.1074/jbc.M101106200

Hao T, Xie Z, Wang M, Liu L, Zhang Y, Wang W, Zhang Z, Zhao X, Li P, Guo Z, Gao S, Lou C, Zhang G, Merritt J, Horsman GP, Chen Y (2019) An anaerobic bacterium host system for heterologous expression of natural product biosynthetic gene clusters. Nat Commun 10:3665 . doi: 10.1038/s41467-019-11673-0

Hasan F, Shah AA, Hameed A (2006) Industrial applications of microbial lipases. Enzyme Microb Technol 39:235–251 . doi: 10.1016/J.ENZMICTEC.2005.10.016

Hausmann S, Jaeger K-E (2010) Lipolytic Enzymes from Bacteria. In: Timmis KN (ed) Handbook of Hydrocarbon and Lipid Microbiology. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 1099–1126

Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell 38:576–589 . doi: 10.1016/j.molcel.2010.05.004

Horne I, Harcourt RL, Sutherland TD, Russell RJ, Oakeshott JG (2002) Isolation of a *Pseudomonas monteilli* strain with a novel phosphotriesterase. FEMS Microbiol Lett 206:51–55 . doi: 10.1111/j.1574-6968.2002.tb10985.x

Hotta Y, Ezaki S, Atomi H, Imanaka T (2002) Extremely stable and versatile carboxylesterase from a hyperthermophilic Archaeon. Appl Environ Microbiol 68:3925–3931 . doi: 10.1128/AEM.68.8.3925-3931.2002

Hu Y, Fu C, Huang Y, Yin Y, Cheng G, Lei F, Lu N, Li J, Ashforth EJ, Zhang L, Zhu B (2010) Novel lipolytic genes from the microbial metagenomic library of the South China Sea marine sediment. FEMS Microbiol Ecol 72:228–37 . doi: 10.1111/j.1574-6941.2010.00851.x

Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, Mende DR, Letunic I, Rattei T, Jensen LJ, Mering C von, Bork P (2019) eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. Nucleic Acids Res 47:D309–D314

Hugenholtz P, Tyson GW (2008) Microbiology: Metagenomics. Nature 455:481–483

Hun CJ, Rahman RNZA, Salleh AB, Basri M (2003) A newly isolated organic solvent tolerant *Bacillus sphaericus* 205y producing organic solvent-stable lipase. Biochem Eng J 15:147–151 . doi: 10.1016/S1369-703X(02)00185-7

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 Comput Biol 12:e1004957 . doi: 10.1371/journal.pcbi.1004957

Ikeda M, Clark DS (1998) Molecular cloning of extremely thermostable esterase gene from hyperthermophilic archaeon *Pyrococcus furiosus* in Escherichia coli. Biotechnol Bioeng 57:624–629 . doi: 10.1002/(SICI)1097-0290(19980305)57:5<624::AID-BIT15>3.0.CO;2-B

Jaeger K-E, Dijkstra BW, Reetz MT (1999) Bacterial biocatalysts: Molecular biology, three-dimensional structures, and biotechnological applications of lipases. Annu Rev Microbiol 53:315–351 . doi: 10.1146/annurev.micro.53.1.315

Jaeger KE, Eggert T (2002) Lipases for biotechnology. Curr Opin Biotechnol 13:390–397 . doi: 10.1016/S0958-1669(02)00341-5

Jaeger KE, Reetz MT (1998) Microbial lipases form versatile tools for biotechnology. Trends Biotechnol 16:396–403 . doi: 10.1016/S0167-7799(98)01195-0

Jang E, Ryu BH, Kim TD (2014) Identification, characterization, and immobilization of an organic solvent-stable alkaline hydrolase (PA27) from *Pseudomonas aeruginosa* MH38. Molecules 19:14396–14405 . doi: 10.3390/molecules190914396

Jayanath G, Mohandas SP, Kachiprath B, Solomon S, Sajeevan TP, Bright Singh IS, Philip R (2018) A novel solvent tolerant esterase of GDSGG motif subfamily from solar saltern through metagenomic approach: Recombinant expression and characterization. Int J Biol Macromol 119:393–401 . doi: 10.1016/j.ijbiomac.2018.06.057

Jensen MB V., Horsfall LE, Wardrope C, Togneri PD, Marles-Wright J, Rosser SJ (2016) Characterisation of a new family of carboxyl esterases with an OsmC domain. PLoS One 11:e0166128 . doi: 10.1371/journal.pone.0166128

Jeon JH, Kim JT, Kim YJ, Kim HK, Lee HS, Kang SG, Kim SJ, Lee JH (2009) Cloning and characterization of a new cold-active lipase from a deep-sea sediment metagenome. Appl Microbiol Biotechnol 81:865–874 . doi: 10.1007/s00253-008-1656-2

Jeon JH, Kim S-J, Lee HS, Cha S-S, Lee JH, Yoon S-H, Koo B-S, Lee C-M, Choi SH, Lee SH, Kang SG, Lee J-H (2011) Novel metagenome-derived carboxylesterase that hydrolyzes β-lactam antibiotics. Appl Environ Microbiol 77:7830–6 . doi: 10.1128/AEM.05363-11

Jiang H, Zhang S, Gao H, Hu N (2016) Characterization of a cold-active esterase from *Serratia* sp. and improvement of thermostability by directed evolution. BMC Biotechnol 16: . doi: 10.1186/s12896-016-0235-3

Jiang X, Huo Y, Cheng H, Zhang X, Zhu X, Wu M (2012) Cloning, expression and characterization of a halotolerant esterase from a marine bacterium *Pelagibacterium halotolerans* B2T. Extremophiles 16:427–435 . doi: 10.1007/s00792-012-0442-3

Jünemann S, Kleinbölting N, Jaenicke S, Henke C, Hassa J, Nelkner J, Stolze Y, Albaum SP, Schlüter A, Goesmann A, Sczyrba A, Stoye J (2017) Bioinformatics for NGS-based metagenomics and the application to biogas research. J Biotechnol 261:10–23 . doi: 10.1016/j.jbiotec.2017.08.012

Kamarudin NHA, Rahman RNZRA, Ali MSM, Leow TC, Basri M, Salleh AB (2014) A new cold-adapted, organic solvent stable lipase from mesophilic *Staphylococcus epidermidis* AT2. Protein J 33:296–307 . doi: 10.1007/s10930-014-9560-3

Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M (2016) KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res 44:D457–D462 . doi: 10.1093/nar/gkv1070

Kang CH, Oh KH, Lee MH, Oh TK, Kim BH, Yoon JH (2011) A novel family VII esterase with industrial potential from compost metagenomic library. Microb Cell Fact 10:41 . doi: 10.1186/1475-2859-10-41

Kang HY, Kim JF, Kim MH, Park SH, Oh TK, Hur CG (2006) MELDB: A database for microbial esterases and lipases. FEBS Lett 580:2736–2740 . doi: 10.1016/j.febslet.2006.04.034

Kang LJ, Meng ZT, Hu C, Zhang Y, Guo HL, Li Q, Li M (2017) Screening, purification, and characterization of a novel organic solvent-tolerant esterase, Lip2, from *Monascus purpureus* strain M7. Extremophiles 21:345–355 . doi: 10.1007/s00792-016-0907-x

Karadzic I, Masui A, Zivkovic LI, Fujiwara N (2006) Purification and characterization of an alkaline lipase from *Pseudomonas aeruginosa* isolated from putrid mineral cutting oil as component of metalworking fluid. J Biosci Bioeng 102:82–89 . doi: 10.1263/jbb.102.82

Karan R, Capes MD, DasSarma S (2012) Function and biotechnology of extremophilic enzymes in low water activity. Aquat Biosyst 8:4 . doi: 10.1186/2046-9063-8-4

Keegan KP, Glass EM, Meyer F (2016) MG-RAST, a metagenomics service for analysis of microbial community structure and function. In: Methods in Molecular Biology. Humana Press Inc., pp 207–233

Kent WJ (2002) BLAT---The BLAST-Like Alignment Tool. Genome Res 12:656–664 . doi: 10.1101/gr.229202

Khan FI, 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

Khan M, Sathya TA (2018) Extremozymes from metagenome: Potential applications in food processing. Crit Rev Food Sci Nutr 58:2017–2025 . doi: 10.1080/10408398.2017.1296408

Kim MK, Kang TH, Kim J, Kim H, Yun HD (2012) Cloning and identification of a new group esterase (Est5S) from noncultured rumen bacterium. J Microbiol Biotechnol 22:1044–1053

Kim S, Joo S, Yoon S, Kim S, Moon J, Ryu Y, Kim KK, Kim TD (2009) Purification, crystallization and preliminary crystallographic analysis of Est-Y29: A novel oligomeric β-lactamase. Acta Crystallogr Sect F Struct Biol Cryst Commun 65:310–312 . doi: 10.1107/S1744309109005442

Kim TD (2017) Bacterial hormone-sensitive lipases (bHSLs): emerging enzymes for biotechnological applications. J Microbiol Biotechnol 27:1907–1915 . doi: 10.4014/jmb.1708.08004

Knudsen M, Wiuf C (2010) The CATH database. Hum Genomics 4:207–212 . doi: 10.1186/1479-7364-4-3-207

Koseki T, Fushinobu S, Ardiansyah, Shirakawa H, Komai M (2009) Occurrence, properties, and applications of feruloyl esterases. Appl Microbiol Biotechnol 84:803–810 . doi: 10.1007/s00253-009-2148-8

Kovacic F, Babic N, Krauss U, Jaeger K (2019) Classification of lipolytic enzymes from bacteria. In: Rojo F (ed) Aerobic Utilization of Hydrocarbons, Oils and Lipids. Handbook of Hydrocarbon and Lipid Microbiology. Springer, Cham, pp 1–35

Kuddus M (2018) Cold-active enzymes in food biotechnology: An updated mini review. J Appl Biol Biotechnol 6:58–63 . doi: 10.7324/JABB.2018.60310

Kumar A, Dhar K, Singh Kanwar S, Arora PK (2016) Lipase catalysis in organic solvents: advantages and applications. Biol Proced Online 18:2 . doi: 10.1186/s12575-016-0033-2

Kumar S, Krishnani KK, Bhushan B, Brahmane MP (2015) Metagenomics: Retrospect and prospects in high throughput age. Biotechnol Res Int 2015:1–13 . doi: 10.1155/2015/121735

Lam KN, Cheng J, Engel K, Neufeld JD, Charles TC (2015) Current and future resources for functional metagenomics. Front Microbiol 6:1196 . doi: 10.3389/fmicb.2015.01196

Lämmle K, Zipper H, Breuer M, Hauer B, Buta C, Brunner H, Rupp S (2007) Identification of novel enzymes with different hydrolytic activities by metagenome expression cloning. J Biotechnol 127:575–592 . doi: 10.1016/j.jbiotec.2006.07.036

Lanyi JK (1974) Salt-dependent properties of proteins from extremely halophilic bacteria. Bacteriol Rev 38:272–290

Laudadio I, Fulci V, Stronati L, Carissimi C (2019) Next-generation metagenomics: methodological challenges and opportunities. Omi A J Integr Biol 23:327–333 . doi: 10.1089/omi.2019.0073

Lee CW, Kwon S, Park S-H, Kim B-Y, Yoo W, Ryu BH, Kim H-W, Shin SC, Kim S, Park H, Kim TD, Lee JH (2017) Crystal structure and functional characterization of an esterase (EaEST) from *Exiguobacterium antarcticum*. PLoS One 12:e0169540 . doi: 10.1371/journal.pone.0169540

Lee D-W, Koh Y-S, Kim K-J, Kim B-C, Choi H-J, Kim D-S, Suhartono MT, Pyun Y-R (1999) Isolation and characterization of a thermophilic lipase from *Bacillus thermoleovorans* ID-1. FEMS Microbiol Lett 179:393–400 . doi: 10.1111/j.1574-6968.1999.tb08754.x

Lee MH, Hong KS, Malhotra S, Park J-H, Hwang EC, Choi HK, Kim YS, Tao W, Lee S-W (2010) A new esterase EstD2 isolated from plant rhizosphere soil metagenome. Appl Microbiol Biotechnol 88:1125–1134 . doi: 10.1007/s00253-010-2729-6

Lee MH, Lee CH, Oh TK, Song JK, Yoon JH (2006) Isolation and characterization of a novel lipase from a metagenomic library of tidal flat sediments: Evidence for a new family of bacterial lipases. Appl Environ Microbiol 72:7406–7409 . doi: 10.1128/AEM.01157-06

Lei J, Miao Y, Lan Y, Han X, Liu H, Gan Y, Niu L, Wang Y, Zheng Z (2018) A novel complementation assay for quick and specific screen of genes encoding glycerol-3-phosphate acyltransferases. Front Plant Sci 9:353 . doi: 10.3389/fpls.2018.00353

Leis B, Angelov A, Mientus M, Li H, Pham VTT, Lauinger B, Bongen P, Pietruszka J, Gonçalves LG, Santos H, Liebl W (2015) Identification of novel esterase-active enzymes from hot environments by use of the host bacterium *Thermus thermophilus*. Front Microbiol 6:275 . doi: 10.3389/fmicb.2015.00275

Lenfant N, Hotelier T, Velluet E, Bourne Y, Marchot P, Chatonnet A (2013) ESTHER, the database of the α/β-hydrolase fold superfamily of proteins: tools to explore diversity of functions. Nucleic Acids Res 41:423–429 . doi: 10.1093/nar/gks1154

Levisson M, Van Der Oost J, Kengen SWM (2007) Characterization and structural modeling of a new type of thermostable esterase from *Thermotoga maritima*. FEBS J 274:2832–2842 . doi: 10.1111/j.1742-4658.2007.05817.x

Li D, Liu CM, Luo R, Sadakane K, Lam TW (2015) MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics 31:1674–1676 . doi: 10.1093/bioinformatics/btv033

Li H, Han X, Qiu W, Xu D, Wang Y, Yu M, Hu X, Zhuo R (2019) Identification and expression analysis of the GDSL esterase/lipase family genes, and the characterization of SaGLIP8 in *Sedum alfredii* Hance under cadmium stress. PeerJ 7:e6741 . doi: 10.7717/peerj.6741

Li LL, McCorkle SR, Monchy S, Taghavi S, van der Lelie D (2009) Bioprospecting metagenomes: Glycosyl hydrolases for converting biomass. Biotechnol Biofuels 2:1–11 . doi: 10.1186/1754-6834-2-10

Li M, Yang LR, Xu G, Wu JP (2016) Cloning and characterization of a novel lipase from *Stenotrophomonas maltophilia* GS11: The first member of a new bacterial lipase family XVI. J Biotechnol 228:30–36 . doi: 10.1016/j.jbiotec.2016.04.034

Li M, Yang LR, Xu G, Wu JP (2013) Screening, purification and characterization of a novel cold-active and organic solvent-tolerant lipase from *Stenotrophomonas maltophilia* CGMCC 4254. Bioresour Technol 148:114–120 . doi: 10.1016/j.biortech.2013.08.101

Li PY, Ji P, Li CY, Zhang Y, Wang GL, Zhang XY, Xie B Bin, Qin QL, Chen XL, Zhou BC, Zhang YZ (2014) Structural basis for dimerization and catalysis of a novel sterase from the GTSAG motif subfamily of the bacterial hormone-sensitive lipase family. J Biol Chem 289:19031–19041 . doi: 10.1074/jbc.M114.574913

Li PY, Yao QQ, Wang P, Zhang Y, Li Y, Zhang YQ, Hao J, Zhou BC, Chen XL, Shi M, Zhang YZ, Zhang XY (2017) A novel subfamily esterase with a homoserine transacetylase-like fold but no transferase activity. Appl Environ Microbiol 83:e00131-17 . doi: 10.1128/AEM.00131-17

Lima VMG, Krieger N, Mitchell DA, Fontana JD (2004) Activity and stability of a crude lipase from *Penicillium aurantiogriseum* in aqueous media and organic solvents. Biochem Eng J 18:65–71 . doi: 10.1016/S1369-703X(03)00165-7

Liu K, Wang J, Bu D, Zhao S, McSweeney C, Yu P, Li D (2009) Isolation and biochemical characterization of two lipases from a metagenomic library of China Holstein cow rumen. Biochem Biophys Res Commun 385:605–611 . doi: 10.1016/j.bbrc.2009.05.110

López-López O, Knapik K, Cerdán ME, González-Siso MI (2015) Metagenomics of an alkaline hot spring in Galicia (Spain): Microbial diversity analysis and screening for novel lipolytic enzymes. Front Microbiol 6:1291 . doi: 10.3389/fmicb.2015.01291

Lu M, Dukunde A, Daniel R (2019) Biochemical profiles of two thermostable and organic solvent–tolerant esterases derived from a compost metagenome. Appl Microbiol Biotechnol 103:3421–3437 . doi: 10.1007/s00253-019-09695-1

Luan Z-J, Yu H-L, Ma B-D, Qi Y-K, Chen Q, Xu J-H (2016) Dramatically improved performance of an esterase for cilastatin synthesis by cap domain engineering. Ind Eng Chem Res 55:12167–12172 . doi: 10.1021/acs.iecr.6b02440

Madhavan A, Sindhu R, Parameswaran B, Sukumaran RK, Pandey A (2017) Metagenome analysis: a powerful tool for enzyme bioprospecting. Appl Biochem Biotechnol 183:636–651 . doi: 10.1007/s12010-017-2568-3

Mandrich L, Menchise V, Alterio V, De Simone G, Pedone C, Rossi M, Manco G (2008) Functional and structural features of the oxyanion hole in a thermophilic esterase from *Alicyclobacillus acidocaldarius*. Proteins Struct Funct Genet 71:1721–1731 . doi: 10.1002/prot.21877

Mandrich L, Merone L, Pezzullo M, Cipolla L, Nicotra F, Rossi M, Manco G (2005) Role of the N terminus in enzyme activity, stability and specificity in thermophilic esterases belonging to the HSL family. J Mol Biol 345:501–512 . doi: 10.1016/j.jmb.2004.10.035

Maruthamuthu M, Jiménez DJ, Stevens P, van Elsas JD (2016) A multi-substrate approach for functional metagenomics-based screening for (hemi)cellulases in two wheat straw-degrading microbial consortia unveils novel thermoalkaliphilic enzymes. BMC Genomics 17:1–16 . doi: 10.1186/s12864-016-2404-0

McArthur AG, Waglechner N, Nizam F, Yan A, Azad MA, Baylay AJ, Bhullar K, Canova MJ, De Pascale G, Ejim L, Kalan L, King AM, Koteva K, Morar M, Mulvey MR, O'Brien JS, Pawlowski AC, Piddock LJV, Spanogiannopoulos P, Sutherland AD, Tang I, Taylor PL, Thaker M, Wang W, Yan M, Yu T, Wright GD (2013) The comprehensive antibiotic resistance database. Antimicrob Agents Chemother 57:3348–3357 . doi: 10.1128/AAC.00419-13

McMahon MD, Guan C, Handelsman J, Thomas MG (2012) Metagenomic analysis of *Streptomyces lividans* reveals host-dependent functional expression. Appl Environ Microbiol 78:3622–3629 . doi: 10.1128/AEM.00044-12

Meng Q, Yang W, Men M, Bello A, Xu X, Xu B, Deng L, Jiang X, Sheng S, Wu X, Han Y, Zhu H (2019) Microbial community succession and response to environmental variables during

cow manure and corn straw composting. Front Microbiol 10:529 . doi: 10.3389/fmicb.2019.00529

Meng Y, Li S, Yuan H, Zou D, Liu Y, Zhu B, Li X (2015) Effect of lipase addition on hydrolysis and biomethane production of Chinese food waste. Bioresour Technol 179:452–459 . doi: 10.1016/j.biortech.2014.12.015

Miller DN, Bryant JE, Madsen EL, Ghiorse WC (1999) Evaluation and optimization of DNA extraction and purification procedures for soil and sediment samples. Appl Environ Microbiol 65:4715–4724 . doi: 10.1128/aem.65.11.4715-4724.1999

Minegishi H, Shimane Y, Echigo A, Ohta Y, Hatada Y, Kamekura M, Maruyama T, Usami R (2013) Thermophilic and halophilic β-agarase from a halophilic archaeon *Halococcus* sp. 197A. Extremophiles 17:931–939 . doi: 10.1007/s00792-013-0575-z

Misawa E, Chan Kwo Chion CKC, Archer I V., Woodland MP, Zhou N-Y, Carter SF, Widdowson DA, Leak DJ (1998) Characterisation of a catabolic epoxide hydrolase from a *Corynebacterium* sp. Eur J Biochem 253:173–183 . doi: 10.1046/j.1432-1327.1998.2530173.x

Mohamed YM, Ghazy MA, Sayed A, Ouf A, El-Dorry H, Siam R (2013) Isolation and characterization of a heavy metal-resistant, thermophilic esterase from a Red Sea Brine Pool. Sci Rep 3:3358 . doi: 10.1038/srep03358

Mølgaard A, Kauppinen S, Larsen S (2000) Rhamnogalacturonan acetylesterase elucidates the structure and function of a new family of hydrolases. Structure 8:373–383 . doi: 10.1016/S0969-2126(00)00118-0

Montoro-García S, Martínez-Martínez I, Navarro-Fernández J, Takami H, García-Carmona F, Sánchez-Ferrer Á (2009) Characterization of a novel thermostable carboxylesterase from *Geobacillus kaustophilus* HTA426 shows the existence of a new carboxylesterase family. J Bacteriol 191:3076–3085 . doi: 10.1128/JB.01060-08

Moreno-Indias I, Tinahones FJ (2020) Metagenomics. Princ Nutr Nutr 81–87 . doi: 10.1016/b978-0-12-804572-5.00011-2

Munawar N, Engel PC (2013) Halophilic enzymes: Characteristics, structural adaptation and potential applications for biocatalysis. Curr Biotechnol 2:334–344 . doi: 10.2174/18722083113076660033

Musa H, Kasim FH, Gunny AAN, Gopinath SCB, Ahmad MA (2018) Biosecretion of higher halophilic lipase by a novel *Bacillus amyloliquefaciens* AIKK2 using agro-waste as supporting substrate. Process Biochem 72:55–62 . doi: 10.1016/j.procbio.2018.06.022

Nacke H, Will C, Herzog S, Nowka B, Engelhaupt M, Daniel R (2011) Identification of novel lipolytic genes and gene families by screening of metagenomic libraries derived from soil samples of the German Biodiversity Exploratories. FEMS Microbiol Ecol 78:188–201 . doi: 10.1111/j.1574-6941.2011.01088.x

Nagamine S, Liu C, Nishishita J, Kozaki T, Sogahata K, Sato Y, Minami A, Ozaki T, Schmidt-Dannert C, Maruyama J ichi, Oikawa H (2019) *Ascomycete Aspergillus oryzae* is an efficient expression host for production of basidiomycete terpenes by using genomic DNA sequences. Appl Environ Microbiol 85:e00409-19 . doi: 10.1128/AEM.00409-19

Namiki T, Hachiya T, Tanaka H, Sakakibara Y (2012) MetaVelvet: An extension of Velvet assembler to de novo metagenome assembly from short sequence reads. Nucleic Acids Res 40:e155 . doi: 10.1093/nar/gks678

Nerurkar M, Joshi M, Pariti S, Adivarekar R (2013) Application of lipase from marine bacteria *Bacillus sonorensis* as an additive in detergent formulation. J Surfactants Deterg 16:435–443 . doi: 10.1007/s11743-012-1434-0

Neves Petersen MT, Fojan P, Petersen SB (2001) How do lipases and esterases work: The electrostatic contribution. J. Biotechnol. 85:115-147

Ngara TR, Zhang H (2018) Recent advances in function-based metagenomic screening. Genomics, Proteomics Bioinforma. 16:405–415

Nurk S, Meleshko D, Korobeynikov A, Pevzner PA (2017) MetaSPAdes: A new versatile metagenomic assembler. Genome Res 27:824–834 . doi: 10.1101/gr.213959.116

Ollis DL, Cheah E, Cygler M, Dijkstra B, Frolow F, Franken SM, Harel M, Remington SJ, Silman I, Schrag J, Sussman JL, Verschueren KHG, Goldman A (1992) The α / β hydrolase fold. Protein Eng Des Sel 5:197–211 . doi: 10.1093/protein/5.3.197

Parapouli M, Foukis A, Stergiou P-Y, Koukouritaki M, Magklaras P, Gkini OA, Papamichael EM, Afendra A-S, Hatziloukas E (2018) Molecular, biochemical and kinetic analysis of a novel, thermostable lipase (LipSm) from *Stenotrophomonas maltophilia* Psi-1, the first member of a new bacterial lipase family (XVIII). J Biol Res 25:4 . doi: 10.1186/s40709-018-0074-6

Park SY, Kim JT, Kang SG, Woo JH, Lee JH, Choi HT, Kim SJ (2007) A new esterase showing similarity to putative dienelactone hydrolase from a strict marine bacterium, *Vibrio* sp. GMD509. Appl Microbiol Biotechnol 77:107–115 . doi: 10.1007/s00253-007-1134-2

Patel U, Chandpura J, Chauhan K, Gupte S (2018) Screening and isolation of an organic solvent tolerant lipase producing bacteria from various oil contaminated sites. Indian J Appl Microbiol 21:22–36

Pehrsson EC, Tsukayama P, Patel S, Mejía-Bautista M, Sosa-Soto G, Navarrete KM, Calderon M, Cabrera L, Hoyos-Arango W, Bertoli MT, Berg DE, Gilman RH, Dantas G (2016) Interconnected microbiomes and resistomes in low-income human habitats. Nature 533:212–216 . doi: 10.1038/nature17672

Peng Y, Leung HCM, Yiu SM, Chin FYL (2011) Meta-IDBA: A de Novo assembler for metagenomic data. Bioinformatics 27:94–101 . doi: 10.1093/bioinformatics/btr216

Peng Y, Leung HCM, Yiu SM, Chin FYL (2012) IDBA-UD: A de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. Bioinformatics 28:1420–1428 . doi: 10.1093/bioinformatics/bts174

Pereira MR, Mercaldi GF, Maester TC, Balan A, de Macedo Lemos EG (2015) Est16, a new esterase isolated from a metagenomic library of a microbial consortium specializing in diesel oil degradation. PLoS One 10:e0133723 . doi: 10.1371/journal.pone.0133723

Pratama R, Schneider D, Böer T, Daniel R (2019) First insights into bacterial gastrointestinal tract communities of the Eurasian beaver (Castor fiber). Front Microbiol 10:1646 . doi: 10.3389/fmicb.2019.01646

Qin Q-L, Zhao D-L, Wang J, Chen X-L, Dang H-Y, Li T-G, Zhang Y-Z, Gao P-J (2007) *Wangia profunda* gen. nov., sp. nov., a novel marine bacterium of the family *Flavobacteriaceae* isolated from southern Okinawa Trough deep-sea sediment. FEMS Microbiol Lett 271:53–58 . doi: 10.1111/j.1574-6968.2007.00694.x

Rabausch U, Juergensen J, Ilmberger N, Böhnke S, Fischer S, Schubach B, Schulte M, Streit W V. (2013) Functional screening of metagenome and genome libraries for detection of novel flavonoid-modifying enzymes. Appl Environ Microbiol 79:4551–4563 . doi: 10.1128/AEM.01077-13

Rahman MA, Culsum U, Tang W, Zhang SW, Wu G, Liu Z (2016) Characterization of a novel cold active and salt tolerant esterase from *Zunongwangia profunda*. Enzyme Microb Technol 85:1–11 . doi: 10.1016/j.enzmictec.2015.12.013

Rajendran A, Palanisamy A, Thangavelu V (2009) Lipase catalyzed ester synthesis for food processing industries. Brazilian Arch Biol Technol 52:207–219 . doi: 10.1590/S1516-89132009000100026

Ramnath L, Sithole B, Govinden R (2016) Classification of lipolytic enzymes and their biotechnological applications in the pulping industry. Can J Microbiol 63:179–192 . doi: 10.1139/cjm-2016-0447

Rao L, Xue Y, Zhou C, Tao J, Li G, Lu JR, Ma Y (2011) A thermostable esterase from *Thermoanaerobacter tengcongensis* opening up a new family of bacterial lipolytic enzymes. Biochim Biophys Acta - Proteins Proteomics 1814:1695–1702 . doi: 10.1016/j.bbapap.2011.08.013

Rao L, Zhao X, Pan F, Li Y, Xue Y, Ma Y, Lu JR (2009) Solution behavior and activity of a halophilic esterase under high salt concentration. PLoS One 4:e6980 . doi: 10.1371/journal.pone.0006980

Rathi P, Bradoo S, Saxena RK, Gupta R (2000) A hyper-thermostable, alkaline lipase from *Pseudomonas* sp. with the property of thermal activation. Biotechnol Lett 22:495–498 . doi: 10.1023/A:1005604617440

Rebollido R, Martínez J, Aguilera Y, Melchor K, Koerner I, Stegmann R (2008) Microbial populations during composting process of organic fraction of municipal solid waste. Appl Ecol Environ Res 6:61–67

Reyes-Duarte D, Ferrer M, García-Arellano H (2012) Functional-based screening methods for lipases, esterases, and phospholipases in metagenomic libraries. Methods Mol Biol 861:101–113 . doi: 10.1007/978-1-61779-600-5_6

Rhee J, Ahn D, Kim Y, Oh J (2005) New thermophilic and thermostable esterase with sequence similarity to the hormone-sensitive lipase family , cloned from a metagenomic library. Appl Environ Microbiol 71:817–825 . doi: 10.1128/AEM.71.2.817

Rigoldi F, Donini S, Redaelli A, Parisini E, Gautieri A (2018) Review: Engineering of thermostable enzymes for industrial applications. APL Bioeng 2:011501 . doi: 10.1063/1.4997367

Rodríguez MC, Loaces I, Amarelle V, Senatore D, Iriarte A, Fabiano E, Noya F (2015) Est10: A novel alkaline esterase isolated from bovine rumen belonging to the new family XV of lipolytic enzymes. PLoS One 10:e0126651 . doi: 10.1371/journal.pone.0126651

Romdhane IB-B, Fendri A, Gargouri Y, Gargouri A, Belghith H (2010) A novel thermoactive and alkaline lipase from *Talaromyces thermophilus* fungus for use in laundry detergents. Biochem Eng J 53:112–120 . doi: 10.1016/J.BEJ.2010.10.002

Rosenstein R, Götz F (2000) Staphylococcal lipases: Biochemical and molecular characterization. Biochimie 82:1005–1014

Sabri S, Rahman RNZRA, Leow TC, Basri M, Salleh AB (2009) Secretory expression and characterization of a highly Ca2+-activated thermostable L2 lipase. Protein Expr Purif 68:161–166 . doi: 10.1016/j.pep.2009.08.002

Salihu A, Alam MZ (2015) Solvent tolerant lipases: A review. Process Biochem 50:86–96

Samoylova Y V., Sorokina KN, Piligaev A V., Parmon VN (2019) Application of bacterial thermostable lipolytic enzymes in the modern biotechnological processes: A review. Catal Ind 11:168–178 . doi: 10.1134/S2070050419020107

Samoylova Y V., Sorokina KN, Romanenko M V., Parmon VN (2018) Cloning, expression and characterization of the esterase estUT1 from *Ureibacillus thermosphaericus* which belongs to a new lipase family XVIII. Extremophiles 22:271–285 . doi: 10.1007/s00792-018-0996-9

Satyanarayana SDV, Krishna MSR, Kumar PP (2017) Optimization of high-yielding protocol for DNA extraction from the forest rhizosphere microbes. 3 Biotech 7:91 . doi: 10.1007/s13205-017-0737-2

Schloss PD, Hay AG, Wilson DB, Walker LP (2003) Tracking temporal changes of bacterial community fingerprints during the initial stages of composting. FEMS Microbiol Ecol 46:1–9 . doi: 10.1016/S0168-6496(03)00153-3

Schulz G, Schneider D, Brinkmann N, Edy N, Daniel R, Polle A, Scheu S, Krashevska V (2019) Changes in trophic groups of protists with conversion of rainforest into rubber and oil palm plantations. Front Microbiol 10:240 . doi: 10.3389/fmicb.2019.00240

Selvin J, Kennedy J, Lejon DPH, Kiran GS, Dobson ADW (2012) Isolation identification and biochemical characterization of a novel halo-tolerant lipase from the metagenome of the marine sponge *Haliclona simulans.* Microb Cell Fact 11:72 . doi: 10.1186/1475-2859-11-72

Seo S, Lee YS, Yoon SH, Kim SJ, Cho JY, Hahn BS, Koo BS, Lee CM (2014) Characterization of a novel cold-active esterase isolated from swamp sediment metagenome. World J Microbiol Biotechnol 30:879–886 . doi: 10.1007/s11274-013-1496-9

Seshadri R, Kravitz SA, Smarr L, Gilna P, Frazier M (2007) CAMERA: A community resource for metagenomics. PLoS Biol 5:e75 . doi: 10.1371/journal.pbio.0050075

Shao H, Xu L, Yan Y (2013) Isolation and characterization of a thermostable esterase from a metagenomic library. J Ind Microbiol Biotechnol 40:1211–1222 . doi: 10.1007/s10295-013-1317-z

Simon C, Daniel R (2009) Achievements and new knowledge unraveled by metagenomic approaches. Appl Microbiol Biotechnol 85:265–276 . doi: 10.1007/s00253-009-2233-z

Simon C, Daniel R (2011) Metagenomic analyses: Past and future trends. Appl Environ Microbiol 77:1153–1161 . doi: 10.1128/AEM.02345-10

Tasse L, Bercovici J, Pizzut-Serin S, Robe P, Tap J, Klopp C, Cantarel BL, Coutinho PM, Henrissat B, Leclerc M, Doré J, Monsan P, Remaud-Simeon M, Potocki-Veronese G (2010) Functional metagenomics to mine the human gut microbiome for dietary fiber catabolic enzymes. Genome Res 20:1605–1612 . doi: 10.1101/gr.108332.110

Tatusov RL, Galperin MY, Natale DA, Koonin E V (2000) The COG database: a tool for genome-scale analysis of protein functions and evolution

Tchigvintsev A, Tran H, Popovic A, Kovacic F, Brown G, Flick R, Hajighasemi M, Egorova O, Somody JC, Tchigvintsev D, Khusnutdinova A, Chernikova TN, Golyshina O V., Yakimov MM, Savchenko A, Golyshin PN, Jaeger KE, Yakunin AF (2015) The environment shapes microbial enzymes: five cold-active and salt-resistant carboxylesterases from marine metagenomes. Appl Microbiol Biotechnol 99:2165–2178 . doi: 10.1007/s00253-014-6038-3

Techtmann SM, Hazen TC (2016) Metagenomic applications in environmental monitoring and bioremediation. J. Ind. Microbiol. Biotechnol. 43:1345–1354

Thai QK, Bös F, Pleiss J (2009) The lactamase engineering database: A critical survey of TEM sequences in public databases. BMC Genomics 10:390 . doi: 10.1186/1471-2164-10-390

Thakur S (2012) Lipases, its sources, properties and applications: A review. Int J Sci Eng Res 3:1–29

Tirawongsaroj P, Sriprang R, Harnpicharnchai P, Thongaram T, Champreda V, Tanapongpipat S, Pootanakit K, Eurwilaichitr L (2008) Novel thermophilic and thermostable lipolytic enzymes from a Thailand hot spring metagenomic library. J Biotechnol 133:42–49 . doi: 10.1016/J.JBIOTEC.2007.08.046

Tripathi NK, Shrivastava A (2019) Recent developments in bioprocessing of recombinant proteins: expression hosts and process development. Front Bioeng Biotechnol 7:420 . doi: 10.3389/fbioe.2019.00420

Turnbaugh PJ, Gordon JI (2009) The core gut microbiome, energy balance and obesity. J Physiol 587:4153–4158 . doi: 10.1113/jphysiol.2009.174136

Turnbaugh PJ, Hamady M, Yatsunenko T, Cantarel BL, Ley RE, Sogin ML, Jones WJ, Roe B a, Jason P, Egholm M, Henrissat B, Heath AC, Knight R, Gordon JI, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, Magris M, Hidalgo G, Baldassano RN, Anokhin AP, Heath AC, Warner B, Reeder J, Kuczynski J, Caporaso JG, Lozupone C a., Lauber C, Clemente JC, Knights D, Knight R, Gordon JI (2009) A core gut microbiome between lean and obesity twins. Nature 457:480–484 . doi: 10.1038/nature07540

Uchiyama T, Miyazaki K (2009) Functional metagenomics for enzyme discovery: challenges to efficient screening. Curr Opin Biotechnol 20:616–622 . doi: 10.1016/j.copbio.2009.09.010

Ufarté L, Laville É, Duquesne S, Potocki-Veronese G (2015a) Metagenomics for the discovery of pollutant degrading enzymes. Biotechnol Adv 33:1845–1854 . doi: 10.1016/j.biotechadv.2015.10.009

Ufarté L, Potocki-Veronese G, Laville É (2015b) Discovery of new protein families and functions: New challenges in functional metagenomics for biotechnologies and microbial ecology. Front. Microbiol. 6:563

Uritskiy G, Di Ruggiero J (2019) Applying genome-resolved metagenomics to deconvolute the halophilic microbiome. Genes (Basel) 10:220 . doi: 10.3390/genes10030220

Vargas-Albores F, Martínez-Córdova LR, Martínez-Porchas M, Calderón K, Lago-Lestón A (2019) Functional metagenomics: a tool to gain knowledge for agronomic and veterinary sciences. Biotechnol Genet Eng Rev 35:69–91 . doi: 10.1080/02648725.2018.1513230

Varma VS, Dhamodharan K, Kalamdhad AS (2018) Characterization of bacterial community structure during in-vessel composting of agricultural waste by 16S rRNA sequencing. 3 Biotech 8:301 . doi: 10.1007/s13205-018-1319-7

Verma N, Thakur S, Bhatt AK (2012) Microbial lipases: Industrial applications and properties (a review). Int Res J Biol Sci Int Res J Biol Sci 1:2278–3202 . doi: 10.4172/2375-4338.1000156

Verschueren KHG, Seljee F, Rozeboom HJ, Kalk KH, Dijkstra BW (1993) Crystallographic analysis of the catalytic mechanism of haloalkane dehalogenase. Nature 363:693–698 . doi: 10.1038/363693a0

Vieille C, Zeikus GJ (2001) Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. Microbiol Mol Biol Rev 65:1–43 . doi: 10.1128/MMBR.65.1.1-43.2001

Wagner UG, Petersen EI, Schwab H, Kratky C (2002) EstB from *Burkholderia gladioli*: A novel esterase with a β-lactamase fold reveals steric factors to discriminate between esterolytic and β-lactam cleaving activity. Protein Sci 11:467–478 . doi: 10.1110/ps.33002

Wang C, Dong D, Wang H, Müller K, Qin Y, Wang H, Wu W (2016a) Metagenomic analysis of microbial consortia enriched from compost : new insights into the role of *Actinobacteria* in lignocellulose decomposition. Biotechnol Biofuels 9:22 . doi: 10.1186/s13068-016-0440-2

Wang G, Wang Q, Lin X, Ng TB, Yan R, Lin J, Ye X (2016b) A novel cold-adapted and highly salt-tolerant esterase from *Alkalibacterium* sp. SL3 from the sediment of a soda lake. Sci Rep 6:19494 . doi: 10.1038/srep19494

Wang H, Hart DJ, An Y (2019) Functional metagenomic technologies for the discovery of novel enzymes for biomass degradation and biofuel production. Bioenergy Res 12:457–470 . doi: 10.1007/s12155-019-10005-w

Wang S, Wang K, Li L, Liu Y (2013) Isolation and characterization of a novel organic solvent-tolerant and halotolerant esterase from a soil metagenomic library. J Mol Catal B Enzym 95:1–8 . doi: 10.1016/j.molcatb.2013.05.015

Wang X, Qin X, Li D, Yang B, Wang Y (2017) One-step synthesis of high-yield biodiesel from waste cooking oils by a novel and highly methanol-tolerant immobilized lipase. Bioresour Technol 235:18–24 . doi: 10.1016/j.biortech.2017.03.086

Wei T, Feng S, Shen Y, He P, Ma G, Yu X, Zhang F, Mao D (2013) Characterization of a novel thermophilic pyrethroid-hydrolyzing carboxylesterase from *Sulfolobus tokodaii* into a new family. J Mol Catal B Enzym 97:225–232 . doi: 10.1016/j.molcatb.2013.07.022

Wilke A, Harrison T, Wilkening J, Field D, Glass EM, Kyrpides N, Mavrommatis K, Meyer F (2012) The M5nr: A novel non-redundant database containing protein sequences and annotations from multiple sources and associated tools. BMC Bioinformatics 13:141 . doi: 10.1186/1471-2105-13-141

Wu C, Sun B (2009) Identification of novel esterase from metagenomic library of Yangtze River. J Microbiol Biotechnol 19:187–193 . doi: 10.4014/jmb.0804.292

Wu G, Wu G, Zhan T, Shao Z, Liu Z (2013a) Characterization of a cold-adapted and salt-tolerant esterase from a psychrotrophic bacterium *Psychrobacter pacificensis*. Extremophiles 17:809–819 . doi: 10.1007/s00792-013-0562-4

Wu G, Zhang S, Zhang H, Zhang S, Liu Z (2013b) A novel esterase from a psychrotrophic bacterium *Psychrobacter celer* 3Pb1 showed cold-adaptation and salt-tolerance. J Mol Catal B Enzym 98:119–126 . doi: 10.1016/j.molcatb.2013.10.012

Xiao F, Li Z, Pan L (2017) Application of microbial lipase and its research progress. Prog Appl Microbiol 8–14

Zarafeta D, Moschidi D, Ladoukakis E, Gavrilov S, Chrysina ED, Chatziioannou A, Kublanov I, Skretas G, Kolisis FN (2016) Metagenomic mining for thermostable esterolytic enzymes uncovers a new family of bacterial esterases. Sci Rep 6:38886 . doi: 10.1038/srep38886

Zhang S, Lin X (2019) CARMA3: Scaffold protein involved in NF-κB signaling. Front Immunol 10:176 . doi: 10.3389/fimmu.2019.00176

Zhang W, Xu H, Wu Y, Zeng J, Guo Z, Wang L, Shen C, Qiao D, Cao Y (2018a) A new cold-adapted, alkali-stable and highly salt-tolerant esterase from *Bacillus licheniformis*. Int J Biol Macromol 111:1183–1193 . doi: 10.1016/J.IJBIOMAC.2018.01.152

Zhang Y, Hao J, Zhang YQ, Chen XL, Xie B Bin, Shi M, Zhou BC, Zhang YZ, Li PY (2017) Identification and characterization of a novel salt-tolerant esterase from the deep-sea sediment of the South China Sea. Front Microbiol 8:441 . doi: 10.3389/fmicb.2017.00441

Zhang Y, Ji F, Wang J, Pu Z, Jiang B, Bao Y (2018b) Purification and characterization of a novel organic solvent-tolerant and cold-adapted lipase from *Psychrobacter* sp. ZY124. Extremophiles 22:287–300 . doi: 10.1007/s00792-018-0997-8

Zhu Y, Li J, Cai H, Ni H, Xiao A, Hou L (2013) Characterization of a new and thermostable esterase from a metagenomic library. Microbiol Res 168:589–597 . doi: 10.1016/J.MICRES.2013.04.004

Chapter II

Metagenomic Screening for lipolytic enzymes reveals an ecology-clustered distribution pattern

Mingji Lu; Dominik Schneider; Rolf Daniel

doi: https://doi.org/10.1101/2021.02.01.429111

https://www.biorxiv.org/content/10.1101/2021.02.01.429111v1

Article

Metagenomic screening for lipolytic genes reveals an ecology-clustered distribution pattern

Mingji Lu[1], Dominik Schneider[1], Rolf Daniel[1*]

[1]Department of Genomic and Applied Microbiology, Institute of Microbiology and Genetics
Georg-August-University of Göttingen, Grisebachstraße 8, 37077, Göttingen, Germany

[*]Correspondence: Rolf Daniel, Genomic and Applied Microbiology & Göttingen Genomics Laboratory, Institute of Microbiology and Genetics, University of Göttingen, Grisebachstr. 8, 37077 Göttingen, Germany. Phone: +49-551-3933827, Fax: +49-551-3912181, Email: rdaniel@gwdg.de

Abstract: Lipolytic enzymes are one of the most important enzyme types for application in various industrial processes. Despite the continuously increasing demand, only a small portion of the so far encountered lipolytic enzymes exhibit adequate stability and activities for biotechnological applications. To explore novel and/or extremophilic lipolytic enzymes, microbial consortia in two composts at thermophilic stage were analyzed using function-driven and sequence-based metagenomic approaches. Analysis of community composition by amplicon-based 16S rRNA genes and transcripts, and direct metagenome sequencing revealed that the communities of the compost samples were dominated by members of the phyla *Actinobacteria*, *Proteobacteria*, *Firmicutes*, *Bacteroidetes* and *Chloroflexi*. Function-driven screening of the metagenomic libraries constructed from the two samples yielded 115 unique genes encoding lipolytic enzymes. The family assignment of these enzymes was conducted by analyzing the phylogenetic relationship and generation of a protein sequence similarity network according to an integral classification system. The sequence-based screening was performed by using a newly developed database, containing a set of profile Hidden Markov models, highly sensitive and specific for detection of lipolytic enzymes. By comparing the lipolytic enzymes identified through both approaches, we demonstrated that the activity-directed complements sequence-based detection, and vice versa. The sequence-based comparative analysis of lipolytic genes regarding diversity, function and taxonomic origin derived from 175 metagenomes indicated significant differences between habitats. Analysis of the prevalent and distinct microbial groups providing the lipolytic genes revealed characteristic patterns and groups driven by ecological factors. The here presented data suggests that the diversity and distribution of lipolytic genes in metagenomes of various habitats are largely constrained by ecological factors.

Keywords: lipolytic enzymes, function-driven metagenomics, sequence-based metagenomics, profile HMM, lipolytic enzyme classification, comparative analysis, compost

Introduction

Lipolytic enzymes (LEs) acting on carboxyl ester bonds in lipids, include esterases (EC 3.1.1.3, carboxylesterases) and true lipases (EC 3.1.1.1, triacylglycerol acyl hydrolases). Due to the catalytic versatility, LEs have remarkable applications in various processes relevant to food, paper, medical, detergent, and pharmaceutical industries (Hita et al. 2009; Romdhane et al. 2010; Ferrer et al. 2015; Sarmah et al. 2018). Nowadays, LEs are considered to be one of the most important biocatalysts for biotechnological applications.

In principle, LEs can be classified on the basis of the substrate preference (Sarmah et al. 2018) and sequence similarity (Chen et al. 2016). The latter provides an easy-to-perform way for classification and indication of the similarity and evolutionary relationship between LEs. Arpigny and Jaeger (1999) have elaborated the most widely accepted classification of bacterial lipases into eight families (I to VIII). The classification system was based on conserved sequence motifs and biological properties of 53 LEs. A recent update of this system resulted in addition of 11 families (IX to XIX) (Kovacic et al. 2019). Besides the nineteen families, there are claims of novel families, such as Est22 (Li et al. 2017), Est9X(Jeon et al. 2009), LipSM54 (Li et al. 2016) and EstDZ2 (Zarafeta et al. 2016). To avoid an artificial inflation of the number of families, these novel families were mostly excluded during classification of newly identified lipolytic enzymes in previous studies (Hitch and Clavel 2019). To the best of our knowledge, we aimed to incorporate all the 'so-called' novel families during classification to avoid false 'novelty'.

LEs are ubiquitous among all aspects of life, with most of them originating from microorganisms (Kovacic et al. 2019). Environmental microbes, including the so far uncultured species, encode a largely untapped reservoir of novel LEs. Metagenomic function-driven and sequence-based approaches provided access to the genetic resources from so far uncultured and uncharacterized microorganisms (Simon and Daniel 2009; Simon and Daniel 2011). LEs are among the most frequent targets in function-based screens of metagenomic libraries derived from diverse habitats, such as compost (Kang et al. 2011; Lu et al. 2019), landfill leachate (Rashamuse et al. 2009),

marine sediment (Peng et al. 2011; Zhang et al. 2017), activated sludges (Liaw et al. 2010) and hot springs (López-López et al. 2015).

In contrast to the easy access to metagenome-derived sequencing data, most published metagenomic screenings for LEs were enzyme activity-driven and not sequence-based (Ferrer et al. 2015; Berini et al. 2017). Only a few studies explored LEs by sequence-driven approaches, including analysis based on regular expression patterns (Masuch et al. 2015), ancestral sequence reconstruction (Verma et al. 2019) and conserved motifs (Zhang et al. 2009; Barriuso and Jesús Martínez 2015; Zarafeta et al. 2016). For various reasons, only a very limited number of LEs were identified by these strategies. Sequence-based approaches primarily rely on the reference database to infer functions of newly-discovered genes and the corresponding enzymes (Hugenholtz and Tyson 2008; Quince et al. 2017; Berini et al. 2017; Ngara and Zhang 2018). With protein-of-interest-specific databases, biomolecules such as antibiotic resistance genes (Gibson et al. 2015; Willmann et al. 2015; Pehrsson et al. 2016) and CAZymes (Wang et al. 2016; Montella et al. 2017; Wang et al. 2019) were successfully profiled across habitats. Recently, with the rapid accumulation of genomic/metagenomic data in public repositories (Keegan et al. 2016; Chen et al. 2017; Eric Sayers et al. 2019), it is likely that current knowledge regarding LEs only reflects the tip of the iceberg and that the full diversity of these enzymes is far from being completely described.

In order to quantitatively analyze LEs distributed in environmental samples, we developed a LE-specific profile Hidden Markov Model (HMM) database. Profile HMMs have been widely adopted for detection of remote homologs (Gibson et al. 2015; Walsh et al. 2017; Berglund et al. 2017) and annotation of general functions in microbial genomes and metagenomes (Skewes-Cox et al. 2014; Reyes et al. 2017; Bzhalava et al. 2018). However, they have not yet been specifically applied to LEs. Once developed and validated, the database was applied to profile the lipolytic genes in metagenomes from various habitats. Profiling the distribution of LEs among various habitats provides researchers a straightforward approach for their downstream analysis. In this study, two composts were sampled and LEs identified through function-based and sequence-based approaches were compared. The distribution of lipolytic genes in 175 metagenomes was also investigated by sequence-based screening.

Material and Methods

Sample collection

Compost samples were collected as described previously (Lu et al. 2019). Briefly, two compost piles fermenting mainly wood chips (Pile_1) or kitchen waste (Pile_2) were sampled. Temperatures at the sampling spots were measured, and the two samples were designated as compst55 (55 °C for Pile_1) and compst76 (76°C for Pile_2). Approximately 50 g compost per sample was collected in sterile plastic tubes and stored at -20 or -80 °C until further use.

Isolation of nucleic acids

Metagenomic DNA of the compost sample was isolated by using the phenol-chloroform method (Zhou et al. 1996) and MoBio Power Soil DNA extraction kit as recommended by the manufacturer (MO BIO Laboratories, Hilden, Germany). DNA obtained from these two methods was pooled per sample and stored at -20 ºC until use.

RNA was extracted by employing the MoBio PowerSoil RNA isolation kit as recommended by the manufacturer (MO BIO Laboratories). Residual DNA was removed by treatment with 2 U Turbo DNase (Applied Biosystems, Darmstadt, Germany) at 37 ºC for 1 h and recovered by using RNeasy MinElute Cleanup kit as recommended by the manufacturer (Qiagen, Hilden, Germany). RNA yields were estimated by employing a Qubit® Fluorometer as recommended by the manufacturer (Thermo Fisher Scientific, Schwerte, Germany). A PCR reaction targeting the 16S rRNA gene was performed to verify the complete removal of DNA as described by Schneider et al. (2015). Subsequently, the DNA-free RNA was converted to cDNA using the SuperScript™ III reverse transcriptase (Thermo Fisher Scientific). Briefly, a mixture (14 µl) containing 100 ng of DNA-free RNA in DEPC-treated water, 2 µM of reverse primer (5' - CCGTCAATTCMTTTGAGT-') and 10 mM dNTP mix was incubated at 65 ºC for 5 min and chilled on ice for at least 1 min. Then, 10 µl of cDNA synthesis mix including reaction buffer, 5 mM $MgCl_2$, 0.01 M DTT, 1 µl 40U RiboLock™ RNase inhibitor (Thermo Fisher Scientific) and 200U SuperScript™ III reverse transcriptase (Thermo Fisher Scientific) was added to each RNA/primer mixture in the previous step, and incubated at 55 ºC for 90 min. The reaction was terminated at 70 ºC for 15 min.

Sequencing of 16S rRNA genes and transcripts

The PCR amplification of the V3-V5 regions of bacterial 16S rRNA genes and transcripts were performed with the following set of primers comprising the Roche 454 pyrosequencing adaptors (underlined), a key (TCAG), a unique 10-bp multiplex identifier (MID), and template-specific sequence per sample: the forward primer V3for_B (5'-CGTATCGCCTCCCTCGCGCCATCAG-MID-TACGGRAGGCAGCAG-3'), (Liu et al. 2007), and reverse primer V5rev_B 5'-CTATGCGCCTTGCCAGCCCGCTCAG-MID-CCGTCAATTCMTTTGAGT-3' (Wang and

Qian 2009). The PCR reaction mixture (50 µl) contained 10 µl of fivefold reaction buffer, 200 µM of each of the four deoxynucleoside triphosphates, 0.2 µM of each primer, 5% DMSO, 1 U of Phusion hot start high-fidelity DNA Polymerase (Finnzymes, Vantaa, Finland) and 50 ng template (DNA or cDNA). The thermal cycling scheme comprised initial denaturation at 98 ºC for 5 min, 25 cycles of denaturation at 98 ºC for 45 s, annealing for 45 s at 60 ºC, and extension at 72 ºC for 30 s, followed by a final extension period at 72 ºC for 5 min. All amplicon PCR reactions were performed in triplicate and pooled in equimolar amounts for sequencing. The Göttingen Genomics Laboratory determined the sequences of the partial 16S rRNA gene and transcript amplicons by using a 454 GS-FLX sequencer and titanium chemistry as recommended by the manufacturer (Roche, Mannheim, Germany).

Quality-filtering and denoising of the recovered 16S rRNA pyrotag reads were performed with the QIIME (1.9.1) software package (Bolyen et al. 2019) by employing the scheme outlined by Schneider et al. (2015). Forward and reverse primer sequences were removed with the *split_libraries.py* script. Pyrosequencing noise and chimeric sequences were removed with UCHIME (Edgar et al. 2011). Operational taxonomic unit (OTU) determination was performed by employing the *pick_open_reference_otus.py* script at genetic divergence level of 3 %. Taxonomic classification of OTUs was performed by *parallel_assign_taxonomy_blast.py* script against the Silva SSU database release 128 (Quast et al. 2013). The *filter_otu_table.py* script was used to remove singletons, chloroplast sequences, extrinsic domain OTUs, and unclassified OTUs. Rarefaction curves was calculated with QIIME software by using *alpha-rarefaction.py*.

Metagenomic sequencing and data processing

The sequencing libraries were constructed and indexed with Nextera DNA Sample Preparation kit and Index kit as recommended by the manufacturer (Illumina, San Diego, CA, USA). Paired-end sequencing was performed using a HiSeq 4000 instrument (2 x 150 bp) as recommended by the manufacturer (Illumina). Raw reads were trimmed with Trimmomatic version 0.36 (Bolger et al. 2014) and verified with FastQC version 0.11.5 (Andrew, 2010). Then, reads were submitted to MG-RAST metagenomics analysis server and processed by the default quality control pipeline (Keegan et al. 2016). Microbial composition analysis was performed using MG-RAST best hit classification tool against the databases of M5RNA (Non-redundant multisource ribosomal RNA annotation) and M5NR (M5 non-redundant protein) with default settings. Functional classification was performed based on clusters of orthologous groups (COGs) and Subsystem categories with default settings. Since we mainly focused on the bacterial community, the baseline for all fractions reported referred to the reads assigned to the bacterial domain.

Construction of metagenomic plasmid libraries and function-based screening for lipolytic activity

Lipolytic genes were screened by constructing small-insert plasmid libraries as described by Lu et al. (2019). Briefly, DNA was sheared by sonication for 3 s at 30% amplitude and cycle 0.5 (UP200S Sonicator, Stuttgart, Germany), and size-separated using a 0.8% low-melting point agarose gel. DNA fragments from 6 to 12 kb were recovered by gel extraction using the peqGold Gel Extraction kit as recommended by the manufacturer (Peqlab Biotechnologie GmbH, Erlangen, Germany). The metagenomic small-insert library was constructed using the vectors pFLD or pCR-XL-TOPO (Thermo Fisher Scientific).

Vector pFLD was digested with *PmlI* at 37ºC for 2 h and dephosphorylated with 5 U Antarctic phosphatase at 37 ºC for 30 min as recommended by the manufacturer (NEB, Ipswich, MA). Subsequently, the ends of DNA fragments were blunt-ended and phosphorylated by employing the Fast DNA End Repair kit (Thermo Fisher Scientific). SureClean was applied to purify DNA or vector between steps as described by the manufacturer (Bioline GmbH, Luckenwalde, Germany). Finally, metagenomic fragments and pFLD vector were ligated using T4 DNA ligase (Thermo Fisher Scientific) at 16 ºC, overnight. Metagenomic DNA fragments were cloned into vector pCR-XL-TOPO following the protocol of the manufacturer recommended in the TOPO-XL-PCR cloning kit (Thermo Fisher Scientific).

To screen for lipolytic activity, *Escherichia coli* TOP10 was used as the host (Dukunde et al. 2017). Library-bearing cells were plated onto LB agar plates (15 g/L) containing 1% (v/v) emulsified tributyrin (Sigma) as the indicator substrate and the appropriate antibiotic (pFLD, 100 µg/ml Ampicillin; pCR-XL-TOPO, 50 µg/ml Kanamycin). The quality of the libraries was controlled by checking the average insert sizes and the percentage of insert-bearing *E. coli* clones (Table 1). Cells were incubated on indicator agar at 37 ºC for 24 h and subsequently for 1 to 7 d at 30 ºC. Lipolytic-positive *E. coli* clones were identified by the formation of clear zones (halos) around individual colonies.

The recombinant plasmid DNA derived from positive clones was isolated by using the QIAGEN plasmid mini kit (QIAGEN) and digested by *PmlI* (vector PFLD) or *EcoRI* (vector pCR-XL-TOPO) at 37 ºC for 2 h. The digestion pattern was analyzed, and phenotype of positive clones was confirmed by transformation of the selected plasmids (from previous step) into the host and rescreening on indicator agar plates. In addition, lipolytic activity towards different triacylglycerides was measured qualitatively by incubating the confirmed lipolytic positive clones on agar plates emulsified with tributyrin (C4), tricaproin (C6), tricaprylin (C8), tricaprin (C10), trilaurin

(C12), trimyristin (C14), or tripalmitin (C16). Formation of clearing zones (halos) on agar plates indicated lipolytic activity.

Analysis of lipolytic genes from function-based screenings

The plasmids recovered from the confirmed positive clones were pooled in equal amounts (50 ng of each clone) for compost55 and compost76. Then, the two plasmid DNA mixtures were sequenced using an Illumina MiSeq instrument with reagent kit version 3 (2x 300 cycles) as recommended by the manufacturer (Illumina)). To remove the vector sequences, raw reads were initially mapped against vector sequences (pFLD or pCR-XL-TOPO) using Bowtie 2 (Langmead and Salzberg 2012). The unmapped reads were quality-filtered by Trimmomatic v0.30 (Bolger et al. 2014) and assembled into contigs by Metavelvet v1.2.01 (Namiki et al. 2012) and MIRA 4 (Chevreux et al. 1999). In addition, both ends of the inserts of each plasmid were sequenced using Sanger technology and the following primers: pFLD504_F (5'-GCCTTACCTGATCGCAATCAGGATTTC-3') and pFLD706_R (5'-CGAGGAGAGGGTTAGGGATAGGCTTAC-3') for vector pFLD, and M13_Forward (5´-GTAAAACGACGGCCAG-3´) and M13_Reverse (5´-CAGGAAACAGCTATGAC-3´) for vector pCR-XL-TOPO . The raw Sanger reads were processed with the Staden package (Staden et al. 2003). Finally, the full insert sequence for each plasmid was reconstructed by mapping the processed Sanger reads on the contigs assembled from the Illumina reads. Open reading frames (ORFs) were predicted by MetaGeneMark (Zhu et al. 2010) using default parameters. Lipolytic genes were annotated by searches against NCBI Non-redundant sequence database (http://www.ncbi.nlm.nih.gov/gorf/gorf.html).

Family classification of lipolytic enzymes revealed from function-based screening

LEs were clustered according to the classification standard defined by Arpigny and Jaeger (1999). In order to classify LEs identified from function-based screening, we have integrated all the so far reported lipolytic families, including families I to XIX, and potential novel families reported in recent studies (Supplementary Table S1). A neighbor-joining tree was constructed with LEs identified from this study and reference proteins (Supplementary Table S2) using MEGA version 7 (Tamura et al. 2013). The robustness of the tree was tested by bootstrap analysis using 500 replications. The phylogenetic tree was depicted by GraPhlAn (Asnicar et al. 2015). To confirm the classification and group proteins in clusters, a protein sequence similarity network was generated. In a protein sequence similarity network, members in a potential isofunctional group consist of nodes (symbol) that share a sequence similarity larger than a selected value and are connected by edges (line). As similarity increases, edges decrease and finally proteins can be separated into defined clusters (Gerlt et al. 2015). In this study, a protein sequence similarity network was generated by submitting the same sequence dataset used in the phylogenetic analysis to the Enzyme Function Initiative-Enzyme Similarity Tool web server (EFI-EST; http://efi.igb.illinois.edu/efi-est/index.php) (Atkinson et al. 2009) with an E-value cutoff of $\leq 1e^{-10}$ and alignment score $\geqslant 16$. The resulting network was visualized in Cytoscape 3.2.1 using the organic layout (Shannon et al. 2003). In addition, multiple-sequence alignments were conducted to explore the presence of catalytic residues, and conservative and distinct motifs in each lipolytic family by employing ClustalW (Larkin et al. 2007).

Building profile Hidden Markov Model (HMM) database for sequence-based screening

A search method based on profile HMMs was developed to identify and annotate putative lipolytic genes in metagenomes (Supplementary Figure 1). In order to target homologous sequences, profile HMMs were built from multiple sequence alignments, which requires relatedness in the input protein sequences. Thus, consistent to the classification of functional-derived LEs, we generally followed the clustering system of Arpigny and Jaeger (1999).

With the exception of LEs belonging to families II and VIII, and patatin-like-proteins, LEs in the other families generally share a conserved α/β-hydrolase fold and a canonical G-x-S-x-G pentapetide around the catalytic serine (Kovacic et al. 2019). ESTHER is a database dedicated to proteins with α/β-hydrolase-fold and their classifications (Lenfant et al. 2013), containing approximately 60,000 α/β hydrolases grouped in 214 clusters so far (as of November 2019). In ESTHER, families I-XIX were integrated into an own classification with corresponding entries (http://bioweb.supagro.inra.fr/ESTHER/Arpigny_Jaeger.table). We thereby designated lipolytic families that were classified and named according to ESTHER database as ELFs (abbreviation of ESTHER Lipolytic Families). For lipolytic families that were not incorporated into the 19 families (I-XIX), their corresponding ELFs were determined by searching LEs against ESTHER database. Generally, a LE was assigned to an ELF if its BLASTp top hit (with lowest e-value) had $\geqslant$60 % amino acid identity and $\geqslant$80% query coverage. Protein sequences in all of the determined ELFs were downloaded from ESTHER database for profile HMM construction.

Firstly, multiple sequence alignments were performed with protein sequences in each ELF, using the following three algorithms and default settings: ClustalW (Thompson et al. 1994), Clustal Omega (Sievers et al. 2011) and Muscle (Edgar 2004). Subsequently, the three alignment sets were run through *hmmbuild* in HMMER3 (Eddy 2018) to create three sets of profile HMMs. Moreover, profile HMMs supplied in the ESTHER database were downloaded. Finally, four profile HMM databases were constructed by concatenating and compressing the

respective set of profile HMMs using *hmmpress*. Thereafter, we designated the four profile HMM databases with respect to the corresponding alignment algorithm (clustalw-pHMMs, omega-pHMMs and muscle-pHMMs) or source (ESTHER-pHMMs).

For families II, VIII and patatin-like-proteins, profile HMMs were retrieved directly from Pfam database (Finn et al. 2014) using the searching keywords of "GDSL", "beta-lactamase" and "patatin", respectively. The profile HMM database was constructed as described above and designated as pfam-pHMMs, specifying for LEs in families II and VIII, and patatin-like-protein.

Validating profile HMM database

The prediction sensitivity and specificity of the profile HMM databases were evaluated using four datasets. Dataset 1, LEs recruited in the UniProtKB database (as of November 2019) using as search strategy the EC numbers 3.1.1.1 or 3.1.1.3, and protein length between 200 to 800 amino acids. Only the prokaryotic LEs were selected for analysis (Supplementary Table S3a). Dataset 2, LEs reported in literature. Most of these enzymes were obtained through metagenomic approaches and biochemically characterized, and with a confirmed lipolytic family assignment by constructing a multiple sequence alignment and/or phylogenetic tree (Supplementary Table S3b). Dataset 3, protein sequences predicted by MetaGeneMark (Zhu et al. 2010) from identified inserts harboring functional lipolytic genes (Supplementary Table S3c). Dataset 4, randomly selected protein sequences (not recruited from ESTHER database) that were annotated in Uniprot or NCBI database as non-lipolytic proteins but with sequence homology to LEs (Supplementary Table S3d). Proteins in the four datasets were screened against the profile HMM databases successively with *hmmscan* using an E-value cutoff of $\leqslant 1e^{-10}$. The sensitivity and specificity of each database were evaluated by the recalls and false positive returns. In addition, we compared our homology-based method (profile HMMs) with the similarity-based pairwise sequence alignment method (BLAST; Altschul et al. 1990). The database for BLAST-based searching was built with the same dataset used for profile HMM construction. BLASTp was performed at an E-value cutoff of $\leqslant 1e^{-10}$.

In order to improve the accuracy for assigning proteins to lipolytic families and distinguishing "true" LEs from the non-lipolytic proteins, protein sequences were annotated by two methods and combined for final assignment. Briefly, putative lipolytic proteins (PLPs) identified by screening against the selected profile HMM database (one from clustalw-pHMMs, omega-pHMMs, muscle-pHMMs and ESTHER-pHMMs) were further searched against the ESTHER database (all entries in the database were included; as of November 2019) by BLASTp using an E-value cutoff of $\leqslant 1e^{-10}$ (Supplementary Figure 1). A PLP was assigned to a lipolytic family only if it was annotated into the same ELF by *hmmscan* and BLASTp. Otherwise, according to the BLAST results, the remaining PLPs were either annotated as "unassigned" PLPs or non-lipolytic proteins (Supplementary Figure 1). In principle, PLPs with the best Blast hits were affiliated to the miscellaneous ESTHER families (functions were not determined, including *5_AlphaBeta_hydrolase*, *6_AlphaBeta_hydrolase*, *Abhydrolase_7* and *AlphaBeta_hydrolase*), or other ESTHER families (with <60 % identity or <70 % query coverage) were classified as unassigned PLPs. The remaining PLPs with the best Blast hits showing $\geqslant$60 % amino acid identity and $\geqslant$70 % query coverage to the non-lipolytic ESTHER families were classified as non-lipolytic proteins.

Family annotation of PLPs obtained by screening against pfam-phmms were confirmed by a further scan against the CATH HMMs database (Knudsen and Wiuf 2010) using the Github repository *cath-tools-genomescan* (https://github.com/UCLOrengoGroup/cath-tools-genomescan). PLPs were assigned to lipolytic families VIII and II, or patatin-like-proteins only if the PLP was assigned to the specific Funfams (functional families) dedicated to lipolytic-related activities, which were inferred from the functionally characterized LEs and gene ontology (GO) annotations (Supplementary Table S4). Additionally, based on our literature search, the LEs in family VIII were generally restricted to PLPs with sequence length between 350 to 450 amino acids. In other cases, the PLP was grouped into non-lipolytic proteins.

For the unassigned PLPs, these sequences show low similarity to any ESTHER family with known function or CATH Funfams, and hence, could contain novel lipolytic or non-lipolytic proteins. Non-lipolytic proteins were excluded from the downstream analysis.

Sequence-based screening for putative lipolytic genes

Sequence-based screening for putative lipolytic genes in the two compost metagenomes were performed as described above. Briefly, the processed metagenomic short reads were assembled into contigs with SPADES version 3.10 (Bankevich et al. 2012). Then, protein sequences were deduced from PROKKA v1.14.5 annotation (Seemann 2014). In order to obtain full-length lipolytic genes, only proteins with amino acid sequence length between 200 and 800 amino acids were retained. Subsequently, the resulting protein sequences were screened against the selected profile HMM databases using *hmmscan* (Eddy 2011) with an E-value cutoff of $\leq 1e^{-10}$. Identified PLPs were further assigned into different lipolytic families as described above (Supplementary Figure 1). Moreover, the lipolytic family classification of assigned PLPs was confirmed by constructing the protein sequence similarity network (Atkinson et al. 2009). The phylogenetic origins of PLP-encoding genes and their

corresponding contigs were determined using Kaiju web server (http://kaiju.binf.ku.dk/server; Menzel et al. 2016). Phylogenetic distributions of assigned PLPs in each lipolytic family were visualized via Circos software (Krzywinski et al. 2009).

Comparative analysis of metagenomic datasets

A total of 175 assembled metagenomes from 15 different habitats were retrieved from the Integrated Microbial Genomes and Microbiomes database (IMG/M). These included metagenomes from anaerobic digestor active sludges (ADAS, n=9), agriculture soils (AS, n=10), composts (COM, n=18), grassland soils (GS, n=11), human gut systems (HG, n=16), hypersaline mats (HM, n=7), hydrocarbon resource environments (HRE, n=6), hot springs (HS, n=14), landfill leachates (LL, n=10), marine sediments (MS, n=12), marine waters (MW, n=10), oil reservoirs (OR, n=13), river waters (RW, n=11), tropical forest soils (TFS, n=14) and wastewater bioreactors (WB, n=13) (Supplementary Table S5). Data processing including open reading frame prediction in assembled contigs and taxonomic assignment of the corresponding deduced protein sequences were conducted by the IMG/M built-in pipelines (Chen et al. 2017). The protein sequences were downloaded from IMG/M database and used in the sequence-based screening as described above (Supplementary Figure 1).

For comparative analysis, the abundance of PLP-encoding genes in each metagenome were normalized according to the method described by Kaminski et al. (2015). The normalized count is in units of LPGM (Lipolytic hits Per Gigabase per Million mapped genes). Unless otherwise stated, LPGM values were used for all calculations. Heatmap was built in R v3.5.2 (R Core Team, 2016) with the function *heatmap.2* using the "Heatplus" package (Ploner et al., 2020). The heatmap hierarchical clustering was performed with "vegan" package (vegdist = "bray", data.dist = "ward.D"). Non-metric multidimensional scaling (NMDS) was also performed with the "vegan" package (Oksanen et al. 2018). The analysis of similarities (ANOSIM) was performed with 9,999 permutations using PAST 4 (Hammer et al. 2001). The phylogenetic annotation of PLPs was retrieved from IMG/M. Association networks between habitats and phylogenetic distribution of PLPs at genus level were generated by mapping significant point biserial correlation values with the "indicspecies" package in R (Cáceres 2013). Only genera with significant correlation coefficients ($P = 0.05$) were included. The resulting bipartite networks were visualized with Cytoscape v3.5 by using the *edge-weighted spring embedded layout* algorithm, whereby the habitats were source nodes, genera target nodes and edges (lines connecting nodes) weighted positive associations between genera and specific habitat or habitats combinations.

In addition, due to the ambiguity of unassigned PLPs, all analyses were performed successively using two datasets: (1) only assigned PLPs, as the consideration of excluding the potential non-lipolytic ones, (2) assigned and unassigned PLPs combined (total PLPs), in order to include all the possible lipolytic ones. This paper mainly focuses on the assigned PLPs for the sake of accuracy, but the comparative analysis of total PLPs was also performed.

Data availability

The short reads and insert sequences were submitted to NCBI databases. Metagenomic short reads are available at in the NCBI sequence read archive (SRA) under accession numbers SRR13115019 (compost55) and SRR13115018 (compost76) and 16S rRNA pyrotag reads under SAMN06859928 (compost55 genes), SAMN06859946 (compost55 transcripts), SAMN06859935 (compost76 genes) and SAMN06859953 (compost76 transcripts). The insert sequences of the plasmids are available in GenBank under accession numbers MW408002--MW408112 (Supplementary Table S9).

Results and Discussion

Phylogenetic and functional profile of microbes in the compost metagenomes

During the heating-up process of composting, the succession of microorganisms plays a key role in degrading organic matter (Dougherty et al. 2012). In this study, the bacterial community compositions in two compost samples with different pile core temperatures of 55 (compost55) and 76 °C (compost76) were revealed by amplicon-based sequencing of 16S rRNA genes (DNA-based, total community) and transcripts (RNA-based, active community) (Supplementary Figures S2 and S3). To extend the taxonomic analysis, the environmental DNA from both metagenomes were also directly sequenced (Supplementary Table S6). Generally, the bacterial community determined by direct sequencing were consistent with that derived from 16S rRNA gene-based analysis. The bacterial phyla *Actinobacteria*, *Proteobacteria*, *Firmicutes*, *Bacteroidetes* and *Chloroflexi* were predominant (relative abundance >5 %) in compost55 and compost76 (Supplementary Figures S3 and S4). This is in agreement with previous studies of bacterial communities in thermophilic composts (Ryckeboer et al., 2003; Antunes et al., 2016; Zhen et al., 2018; Zhou et al., 2018). Differences were detected, which were derived mainly from the different feedstock composition (wood chips vs. kitchen waste) and composting conditions (core temperature 55 vs. 76 °C). *Actinobacteria* was the most abundant phylum (> 25 %) in compost55 (Supplementary Figures S3 and S4), which is accordance with the bacterial communities in composts using mainly plant material

as feedstock (Yu et al. 2007; Zhang et al. 2014). In compost76, members of the *Firmicutes* were most abundant (> 55 %), which was also reported for composts harboring high-nitrogen feedstock, such as animal manure and kitchen waste (Niu et al. 2013; Antunes et al. 2016; Ma et al. 2018; Zhou et al. 2019). The 16S rRNA gene and transcript analysis (Supplementary Table S7) revealed genera that were present (> 1%) in compost55 such as *Brockia, Rhodothermus*, *Thermobispora*, *Longispora*, *Geobacillus*, *Filomicrobium* and *Thermomonospora*, and in compost76 such as *Symbiobacterium*, *Calditerricola* and *Thermaerobacter* were among the typical bacterial taxa previously identified in composting processes (Ryckeboer et al. 2003; Antunes et al. 2016; Yu et al. 2018; Zhou et al. 2018).

Additionally, the metagenomic data were searched against the COG and subsystem databases to assess the functions prominent in compost microbes (Supplementary Figure S5). In principle, compost55 and compost76 share similar metabolic patterns (Supplementary Figure S5). Particularly, the broad diversity and abundance of gene functions in carbohydrate metabolism and transport (COG) and carbohydrates (subsystems) indicated that composts were potential candidates for exploring biocatalysts (Hu et al. 2010a; Leis et al. 2015; Wang et al. 2016; Egelkamp et al. 2019). Notably, the COG category of lipid transport and metabolism as well as the subsystems category of fatty acids, lipids, and isoprenoids were more abundant in the compost55 community than in the compost76 community, suggesting a higher possibility to identify lipolytic genes in the compost55 metagenome.

Function-based screening of LEs in compost metagenomes

In this study, four metagenomic libraries were prepared to probe the diversity of LEs from compost microbes by the function-driven approach using tributyrin-containing indicator agar (Table 1). Overall, approximately 4.89 and 2.56 Gb of cloned compost DNA were screened, yielding 199 and 51 positive clones for compost55 and compost76, respectively. Previous studies have used various vectors such as BACs, fosmids and plasmids for function-based screening of LEs from different bioresources (Lee et al. 2004; Lämmle et al. 2007; Kim et al. 2010; Nacke 2011; Berlemont et al. 2013; Shao et al. 2013; Leis et al. 2015; Jia et al. 2019). The hit rate to recover a lipolytic-positive clone ranged from 0.714 to 208 per Gb of cloned DNA (Table 1). Among the compost metagenomic libraries, the targeting probability towards a LE in our study ranged from 16.1 to 43.6 per Gb and is generally consistent with the values from other studies (Lämmle et al. 2007; Kim et al. 2010; Leis et al. 2015). Also notably, the targeting probabilities in metagenomic libraries from compost and sludge are generally higher than those from other environments, such as grassland, forest soil and river water (Wu and Sun 2009; Nacke 2011; Berlemont et al. 2013). According to Liaw et al. (2010), the targeting probability and/or hit rate for discovering a lipolytic clone is largely attributed to the sample source.

Table 1. Summary of metagenomic libraries used for lipolytic activity screening in this and other studies

Environmental sample[b]	Vector type (average insert size in kb)	No. of library-containing clones (confirmed positive hits, No. of hits per million of clones)	Probability (No. of hits per Gb of DNA screened)	Reference
Compost	Plasmid (5.3)	675,200 (156, 213)	43.6	compost55 (this study)
Compost	Plasmid (5.6)	234,912 (43, 183)	32.7	compost55 (this study)
Compost	Plasmid (6)	281,281 (37, 132)	21.9	compost76 (this study)
Compost	Plasmid (6.2)	140,747 (14, 100)	16.1	compost76 (this study)
Compost	Plasmid (3.2)	21,000 (14, 670)	208	Lämmle et al. 2007
Compost	Fosmid (35)	23,400 (19, 810)	23.2	Kim et al. 2010
Compost	Fosmid (37.5)	1,920 (2, 1040)	27.8	Leis et al. 2015
Compost	Fosmid (-[a])	13,000 (10, 770)	-[a]	Kang et al. 2011
Compost	plasmid (-[a])	66,000 (6, 0.90)	-[a]	Popovic et al. 2017
Grassland soil	Plasmid (5.7)	510,808 (2, 0.4)	0.714	Nacke et al. 2011
Grassland soil	Fosmid (27.8)	50,952(2, 40)	1.41	Nacke et al. 2011
Forest soil	Fosmid (35)	33,700 (8, 240)	6.78	Lee et al. 2004
Forest soil	Plasmid (3.1)	70,000 (3, 42)	13.8	Berlemont et al. 2013
River surface water	BAC (50)	8,000 (1, 120)	2.5	Wu and Sun 2009
Hot spring biofilm	BAC (50)	68,352 (10, 150)	2.93	Yan et al. 2017
Surface sea water	BAC (70)	20,000 (4, 200)	2.86	Chu et al. 2008
Marine sediment	plasmid (4.5)	29,000 (6, 200)	46.0	Ranjan et al. 2018
Marine sediment	Fosmid (36)	40,000 (19, 480)	13.2	Hu et al. 2010
Marine mud	fosmid (40)	40,000 (5, 120)	3.12	Gao et al. 2016
Deep-sea hydrothermal vent	fosmid (35)	18,000 (7, 390)	11.1	Fu et al. 2015
Paper mill sludge	plasmid (5.1)	15,000 (13,870)	170	Jia et al. 2019
Activated sludge	plasmid (5.1)	3,818 (12, 3140)	616	Liaw et al. 2010
Activated sludge	plasmid (2.5)	40,000 (1, 24)	10.0	Shao et al. 2013
Solar saltern	fosmid (35)	51,00 (1, 200)	5.60	Jayanath et al. 2018
Oil field soil	plasmid (3.9)	83,000 (1, 12)	3.09	Fan et al. 2011

[a] This information is not specified in the reference

[b] Except compost metagenomic libraries, only those included the full library information were listed.

Other studies further suggested that samples subjected to specific enrichment processes, such as composting and waste treatment procedures, usually resulted in a high hit rate (Mayumi et al. 2008; Kang et al. 2011; Popovic et al. 2017).

The insert sizes of the recovered plasmids (250 in total) with a confirmed phenotype ranged from 1,038 to 12,587 bp. In all inserts, at least one putative gene showing similarities to known genes encoding lipolytic enzymes was detected. In total, 210 and 60 lipolytic genes were identified from compost55 and compost76 libraries, respectively. To identify unique and full-length LEs, the amino acid sequences deduced from the corresponding lipolytic genes were clustered at 100 % identity. This resulted in 115 (92 for compost55, 23 for compost76, with 7 shared by both samples) unique and full-length LEs (Supplementary Table S8). The length of the unique LEs ranged from 223 to 707 amino acids, with calculated molecular masses from 23.9 to 72.3 kDa (Supplementary Table S9). Forty of these showed the highest similarity to esterases/lipases from uncultured bacteria, and one (EstC55-13) to an enzyme from an uncultured archaeon. Among them, seven LEs showed the highest identity (53 to 65 %) to lipolytic enzymes obtained during function-based screening of metagenomes derived from marine sediment (Hu et al. 2010b), forest topsoil (Lee et al. 2004), mountain soil (Ko et al. 2012), activated sludge (Liaw et al. 2010), wheat field (Stroobants et al. 2015) and compost (Okano et al. 2015). In the remaining 34 cases, the matching esterases/lipases were mainly detected by sequence-based metagenomic surveys of composts (15 LEs), soil (7 LEs), marine sediment (6 LEs) and marine water (3 LEs).

Functionally derived LEs are affiliated with various LE families

The LEs identified through function-based screening were grouped into families based on the classification system reported by Arpigny and Jaeger (1999). With the increasing amount of reports on LEs, claims of new families have been reported (Arpigny and Jaeger 1999; Jeon et al. 2011; Wang et al. 2013; Esteban-Torres et al. 2014; Fang et al. 2014; Rahman et al. 2016; Castilla et al. 2017). In this study, we integrated 29 so-called "novel" families into the classification system for phylogenetic analysis. As shown in the phylogenetic tree (Figure 1), LEs were assigned to 12 families, including families I, II, III, IV, V, VII, VIII, XVII, EM3L4 (Lee et al. 2011), FLS18 (Hu et al. 2010b), EstGS (Nacke et al. 2011), LipT (Chow et al. 2012), patatin-like-proteins and tannases (Supplementary Table S8). The majority of the LEs were affiliated to families V (25 LEs), VIII (21 LEs), IV (15 LEs), I (8 LEs) and patatin-like-proteins (9 LEs). Noteworthy, 7 LEs could not be classified into any known lipolytic family, indicating new branches of LEs. In agreement with previous studies (Arpigny and Jaeger 1999; Glogauer et al. 2011; Akmoussi-Toumi et al. 2018), the "true lipases", which can hydrolyze long-chain substrates ($\geq$ C10) were all affiliated to family I (Figure 1). The remaining LEs exhibiting a preference for short-chain substrates (<C10) were esterases.

To verify the classification result, a protein sequence similarity network was built (Figure 1). The network visualizes relationships among evolutionarily related proteins and is usually considered as an approach complementary to the phylogenetic analysis (Atkinson et al. 2009; Gerlt et al. 2015). At a threshold of $1x10^{-16}$, the network produced clusters that almost matched all the lipolytic families, with the same classification results as obtained by phylogenetic analysis (Figure 1).

Multiple sequence alignments revealed the catalytic residues and conserved motifs in each family (Supplementary Figure S6). For LEs that harbor the canonical α/β-hydrolase fold, the catalytic triad is consistently composed of a nucleophilic serine, an aspartic acid/glutamic acid and a histidine residue (Nardini & Dijkstra, 1999). Most of these LEs contain the conserved motif Gly-x-Ser-x-Gly in which the catalytic serine is embedded (Supplementary Figure S6). Alternatively, three LEs in family I show variations of this conserved motif. The variations were Ala-x-Ser-x-Gly, Thr-x-Ser-x-Gly (Diamond et al. 2019) and Ser-x-Ser-x-Gly (Dalcin Martins et al. 2018) (Supplementary Figure S6).

Family II LEs share a canonical α/β/α-hydrolase fold, which is characterized by a conserved hydrophobic core consisting of five β-strands and at least four α-helices (Akoh et al. 2004). As shown in Supplementary Figure 6a, there are four homology blocks and one conserved residue in each block (serine, glycine, asparagine, and histidine, respectively), which is essential for catalysis (Akoh et al. 2004; Hong et al. 2008). The structures of family VIII enzymes show remarkable sequence similarities to β-lactamases and penicillin-binding proteins (Bornscheuer 2002). Site-directed mutagenesis demonstrated that the catalytic triad is composed of serine and lysine located in a Ser-X-X-Lys motif, and a tyrosine (Supplementary Figure 6) (Biver and Vandenbol 2013; Kovacic et al. 2019). The patatin-like-proteins display an α/β/α-hydrolase fold, in which a central six-stranded beta-sheet is sandwiched between alpha-helices front and back (Banerji and Flieger 2004). Unlike the catalytic triad of Ser-Asp/Glu-His for most lipolytic proteins, the catalytic Ser-Asp dyad is responsible for the catalytic activity of patatin-like-proteins. In addition, they also contained the Gly-x-Ser-x-Gly motif with the catalytic serine embedded (Supplementary Figure 6).

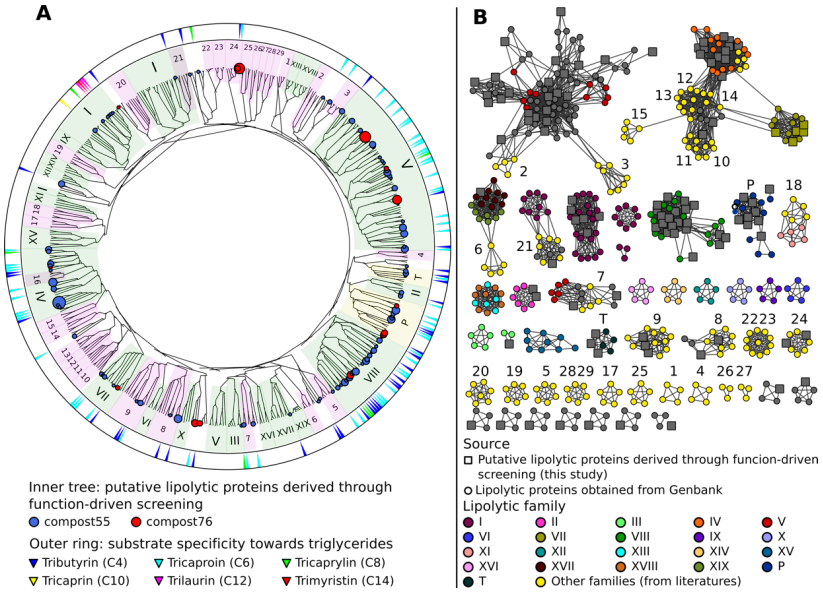


Figure 1. Classification of LEs identified through the function-driven approach. **A,** Unrooted phylogenetic tree was constructed using FA-identified LEs in this study obtained and references retrieved from GenBank (Supplementary Table S2). Phylogenetic tree was constructed using MEGA 7 with neighbor-joining method. The robustness of the tree was tested by bootstrap analysis with 500 replications. Inner tree: the circles represent LEs detected in compost55 (blue) and compost76 (red), sized by abundance (counts of replicates). LEs assigned to families of I-XIX were shaded in green background. Patatin-like-proteins and tannases (designated as P and T, respectively) were shaded in yellow. Other recent reported lipolytic families were shaded in magenta: 1, Est22 (Li et al. 2017); 2, EstL28 (Seo et al. 2014); 3, Rv0045c (Guo et al. 2010); 4, EstGX1 (Jiménez et al. 2012); 5, EstLiu (Rahman et al. 2016); 6, EstY (Wu and Sun 2009); 7; EstGS (Nacke et al. 2011); 8, EM3L4 (Lee et al. 2011); 9, FLS18 (Hu et al. 2010b); 10, Est903 (Jia et al. 2019); 11, EstJ (Choi et al. 2013); 12, PE10 (Jiang et al. 2012); 13, Est12 (Wu et al. 2013); 14, EstDZ2 (Zarafeta et al. 2016); 15, Est9x (Jeon et al. 2009); 16, Lip10 (Guo et al. 2016); 17, EstGH (Nacke et al. 2011); 18, EML1 (Jeon et al. 2009); 19, FnL (Yu et al. 2010); 20, EstP2K (Ouyang et al. 2013); 21, LipA (Couto et al. 2010); 22, LipSM54 (Li et al. 2016); 23, MtEst45 (Lee 2016); 24, LipT (Chow et al. 2012); 25, EstSt7 (Wei et al. 2013); 26, Rlip1 (Liu et al. 2009); 27, EstA (Chu et al. 2008); 28, FLS12 (Hu et al. 2010b); 29, lp_3505 (Esteban-Torres et al. 2014). Outer ring: substrate specificity of corresponding clones towards different carbon chain length (C4 – C14) of triglycerides. **B,** Protein sequence similarity network of LEs belonging to different families. Networks were generated from all-by-all BLAST comparisons of amino acid sequences from the same dataset used for the construction of the phylogenetic tree. Each node represents a sequence. *Larger square nodes* represent LEs derived from function-based screening performed in this study. *Small circle nodes* represent LEs retrieved from GenBank. Nodes were arranged using the *yFiles* organic layout provided in Cytoscape version 3.4.0. Each edge in the network represents a BLAST connection with an E-value cutoff of $\leqslant 1e^{-16}$. At this cut-off, sequences have a mean percent identity and alignment length of 36.3% and 273 amino acids, respectively.

Development of a LE profile HMM database for sequence-based screening

Profile HMMs are statistical models that convert patterns, motifs and other properties from a multiple sequence alignment into a set of position-specific hidden states, i.e. frequencies, insertions, and deletions (Reyes et al. 2017). Profile HMMs are sensitive in detecting remote homologs. Thus, they have been utilized to detect, e.g. viral protein sequences (Skewes-Cox et al. 2014; Bzhalava et al. 2018), antibiotic resistance genes (Gibson et al. 2015), GDSL esterase/lipase family genes (Li et al. 2019) in metagenomes.

In this study, a total of 32 ELFs were determined for profile HMM database construction (Supplementary Table S10). Subsequently, four profile HMM databases (Omega-phmms, Muscle-phmms, Clustalw-phmms, ESTHER-phmms) specific for LEs affiliated to α/β hydrolase superfamily were constructed. Each database consists of 32 profile HMMs (Supplementary Table S11). The prediction sensitivity and specificity of the four databases were evaluated using four datasets (Table 2). All of the four databases obtained high recalls for the datasets 1, 2 and 3 (Table 2), with the highest ones for omega-pHMMs (4,446 in total), followed by muscle-pHMMs (4,444), clustalw-pHMMs (4,425) and ESTHER-pHMMs (4,425). Noteworthy, omega-pHMMs did not identify any false

positive LEs for dataset 3. Thus, omega-pHMMs was chosen for downstream screening. In addition, we compared omega-pHMMs with the pairwise sequence alignment method (BLASTp) for their ability to predict LEs. The omega-pHMM database exhibited improved sensitivity for datasets 1, 2, and 3. In total, 135 more LEs were identified using omega-pHMMs than BLASTp (Table 2).

Table 2 Comparison of profile HMM databases based on different alignment tools to detect LEs

Datasets[a]	Nr. of LEs (α/β hydrolase)	Nr. of LEs (non- α/β hydrolase)	Recall of LEs (α/β hydrolase)					Recall of LEs (non-α/β hydrolase)
			Omega-phmms	Muscle-phmms	Clustalw-phmms	ESTHER-phmms	BLASTp	pfam-phmms
Dataset 1	4382	554	4243	4244	4228	4225	4122	554
Dataset 2	130	32	125	125	121	124	117	32
Dataset 3	80	36	78	75	76	76	70	36
Dataset 4	68	0	56	55	53	53	51	0

[a] Dataset 1, LEs from Uniprot database; Dataset 2, recently reported LEs; Dataset 3, MetaGeneMark-predicted proteins from inserts conferring lipolytic activity; Dataset 4, potential non-lipolytic proteins with homology to LEs.

The accuracy of omega-pHMMs for lipolytic family assignment was also assessed. For datasets 2 and 3, we achieved high precision of annotating LEs to the known lipolytic families, with the exception of LEs from novel families (Supplementary Table S12). Dataset 4 included non-lipolytic proteins, such as epoxide hydrolases, dehalogenases and haloperoxidases and exhibited significant homology with LEs in subfamilies V.1 and V.2 (Arpigny and Jaeger 1999). Our "homology-based" method only differentiated part of these non-lipolytic homologies from "true" LEs (Table 2).

To improve the annotation accuracy, putative lipolytic proteins (PLPs) were further searched against the entire ESTHER database by BLASTp. By combining the annotations from both methods (Supplementary Figure 1), these "novel" LEs in datasets 2 and 3 were correctly identified as "unassigned", in terms of not assigned to any known ELF (Supplementary Table S12). Moreover, almost all of the non-lipolytic proteins (> 92 %) in dataset 4 were distinguished from LEs (Supplementary Table S12).

To identify LEs affiliated to families VIII and II, and patatin-like proteins, enzymes were successively screened against pfam-pHMMs and CATH HMMs database. For the first three datasets, all the LEs in the three families were correctly identified by screening against pfam-pHMMs (Table 2, Supplementary Table S13).

As demonstrated in other sequence-based metagenomic approaches (Liu et al. 2015b; Maimanakos et al. 2016; Azziz et al. 2019), our screening strategy is also vastly dependent on the completeness and accuracy of the reference databases (ESTHER and CATH database in this study). Hence, PLPs exhibiting closest similarity to members affiliated to the miscellaneous ESTHER families or no ESTHER/CATH hits returned, were classified into the "unassigned" group in this study (Supplementary Table S12). This might have resulted in an underestimation of assigned lipolytic proteins (Supplementary Table S13).

Sequence-based screening confirmed compost metagenomes as reservoir for putative lipolytic genes

Initial screening of the assembled metagenomes of compost55 and compost76 resulted in the identification of 4,157 and 2,234 PLPs, respectively. Among them, 1,234 and 759 were further assigned into 28 and 26 families, respectively. The assigned PLPs belonged mainly to family VIII, hormone-sensitive lipase-like proteins, patatin-like proteins, II, A85-Feruloyl-Esterase, Carb_B_Bacteria and homoserine transacetylase (Supplementary Figure S7). The family assignment was also verified by constructing a protein sequence similarity network (supplementary Figure S8). The large number of unassigned PLPs (2,460 for compost55 and 1,208 for compost76) indicated the presence of candidates for novel lipolytic families.

The assigned PLPs were generally of bacterial origin (>95 %), and mainly affiliated to the phyla (> 5 %) *Actinobacteria*, *Proteobacteria*, *Firmicutes* and *Bacteroidetes* (Figure 2). The corresponding contigs were also taxonomically assigned and exhibited a similar phylogeny as seen for the embedded PLP-encoding gene sequences (Figure 2).

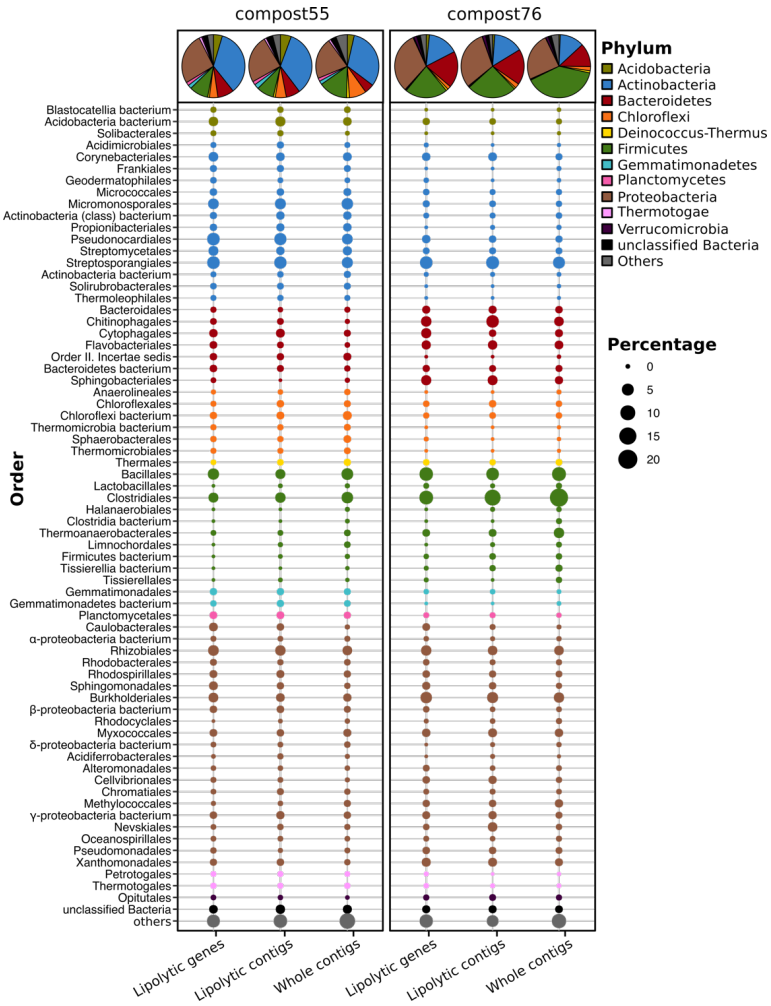


Figure 2. Phylogenetic distribution of assigned PLP-encoding genes identified in compost55 and compost76 metagenomes. The phylogenetic origin of PLP-encoding genes, the contigs harboring these genes, and the whole assembled contigs were annotated by Kaiju (Menzel et al. 2016), and expressed as the proportion of the respective total counts in each sample. The pie charts represent the taxonomic composition at phylum level. Taxa with an abundance of less than 1% were grouped into "others".

Members of the *Actinobacteria* have been reported as important biomass degraders (Ryckeboer et al. 2003; Hubbe et al. 2010; Wang et al. 2016; Lewin et al. 2016). In this study, 34.7 (compost55) and 15.8 % (compost76) of the assigned PLPs originated from *Actinobacteria.* At genus level, the assigned PLPs were affiliated to *Mycobacterium*, *Actinomadura*, *Thermomonospora*, *Streptomyces*, *Micromonospora, Pseudonocardia* and *Thermobifida* (Supplementary Table S14). Members of these genera have been reported as producers for lipases/esterasses (Wei et al. 1998; Alisch et al. 2004; Chahinian et al. 2005; Guo et al. 2010; Hu et al. 2010a; Brault et al. 2012; Mander et al. 2014; Sriyapai et al. 2015). Moreover, some of the corresponding families, such as *Micromonosporaceae, Streptomycetaceae* and *Thermomonosporaceae,* are commonly found in thermophilic composts (Schloss et al. 2003; Blaya et al. 2016; Lima-Junior et al. 2016).

Proteobacteria are also an abundant source for the assigned PLPs in compost55 (26.2 %) and compost76 (31.4 %) (Figure 2). Popovic et al., (2017) identified 80 LEs, of which, 65 % were proteobacterial origin by screening of 16 metagenomic DNA libraries prepared from seawater, soils, compost and wastewater. In our study,

lipolytic genes exhibited high taxonomic diversity at genus level, they were distributed across 97 and 111 genera for compost55 and compot76, respectively (Supplementary Table S14).

The assigned PLPs affiliated to *Firmicutes* originated mainly from *Clostridiales* and *Bacillales* (Figure 2). By analyzing the microbial diversity and metabolic potential of compost metagenomes, members of *Clostridiales* and *Bacillales* were shown to play key roles in degradation of different organic compounds (Martins et al. 2013; Antunes et al. 2016). *Bacteroidetes* is the fourth most abundant phylum for assigned PLPs in compost55 (8.4 %) and compost76 (18.8 %) (Figure 2). At genus level, the assigned PLPs derived mainly from *Rhodothermus* in compost55, and *Sphingobacterium*, *Flavobacterium*, *Niastella* and *Flavihumibacter* in compost76 (Supplementary Table S14). Members of these genera are known as important fermenters during composting (Neher et al. 2013; Antunes et al. 2016; Lapébie et al. 2019).

Strikingly, the phylogenetic distribution of assigned PLPs in each sample, to some extent, corresponded well to the taxonomic composition revealed from the whole contigs (Figure 2) but with minor differences in the rank abundance order. The 16S rRNA amplicon (Supplementary Figure S3) and metagenomic datasets (Supplementary Figure S4) also showed a composition of dominant orders similar to that deduced from lipolytic genes/contigs (Figure 2). Wang et al. (2016) revealed that the phylogenetic distribution of CAZyme genes in the rice straw-adapted compost consortia was in accordance to its microbial composition. By mapping resistance gene dissemination between humans and their environment, Pehrsson et al. (2016) found that resistomes across habitats were generally structured by bacterial phylogeny along ecological gradients.

Comparison between function-driven and sequence-based screening of LEs

Metagenomics allows tapping into the rich genetic resources of so far uncultured microorganisms (Simon and Daniel 2011) through function-driven or sequence-based approaches. The function-driven strategy targets a particular activity of metagenomic library-bearing hosts (Ngara and Zhang 2018). In this way, we identified 13 novel LEs (Supplementary Table S9, Supplementary Figure S6b), which confirmed functional screening as a valuable approach for discovering entirely novel classes of genes and enzymes, particularly when the function could not be predicted based on DNA sequence alone (Reyes-Duarte et al. 2012; Lam et al. 2015; Villamizar et al. 2017).

The sequence-based screening strategy is also frequently used due to the easy access to a wealth of metagenome sequence data and continuous advances in bioinformatics (Chan et al. 2010; Liu et al. 2015c; Maimanakos et al. 2016). Based on the ESTHER and Pfam database, our profile HMM-based search approach efficiently provided an overview of PLP distribution in the two compost metagenomes (Supplementary Figure S7). The hit rate for LEs was higher by sequence-based than by function-based screening, but the sequence-based derived hits need to be functionally verified. In addition, we noticed that only part of the functional screening-derived lipolytic genes were identified during sequence-based screening. By mapping the metagenomic short reads to the functional screening-derived lipolytic genes, 63 genes (out of 115 lipolytic genes in total) had a coverage of 100 % and 88 of $\geqslant$ 99 % (Supplementary Table S15). The BLAST-based comparison between lipolytic genes derived from function-driven and sequence-based approaches indicated that 31 genes from each approach exhibited 100 % sequence identity, and 64 over 99% identity (Supplementary Table S16).

In summary, function-driven and sequence-based strategies have advantages and disadvantages. The function-driven screenings are generally constrained by factors, such as labor-intensive operation, limitations of the employed host systems and low hit rate (Simon and Daniel 2011). However, function-based approaches are activity-directed, and sequence- and database-independent, thus, they bear the potential to discover entirely novel genes for proteins of interest (Rabausch et al. 2013; Lam et al. 2015). Sequencing-based screening, on the other hand, is effective in identifying sequences and potential genes encoding targeted biomolecules in metagenomes. Sequence-based screens largely rely on the used search algorithms, and quality and content of the reference databases to infer the functions of discovered candidate genes (Ngara and Zhang 2018). Thus, the best way to explore novel molecules is to combine the two approaches (Barriuso and Jesús Martínez 2015). Function-driven screens can be employed to complete and verify reference database entries on which sequence-based screening is dependent on. In addition, sequence-based approaches can serve as a pre-selection step for function-driven screens and analysis (Chan et al. 2010; Masuch et al. 2015; Pehrsson et al. 2016; Streit et al. 2018).. The known novel LEs identified by function-based approaches and the functional enzymes identified in this study were employed to expand the LE-specific profile HMM database and annotate the PLPs derived from sequence-based screening.

Assigned PLPs are distributed by ecological factors

In this study, 175 metagenomes representing various ecology niches were selected for sequence-based searching of PLPs. In total, we have screened approx. 1.23 billion genes in 65 Gbp of assembled metagenomes and recovered approx. 0.22 million (absolute counts) PLP-encoding genes. The assigned PLPs (34 % of the total counts) were normalized to LPGM values for comparative analysis. In accordance with the function-based screening, samples subjected to certain enrichment processes, particularly lipid-related, tend to have a higher hit rate (Figure 3). For example, samples with high LPGM values were derived from a hydrocarbon resource environment and oil

reservoir that are enriched with oil-degrading microbes (Liu et al. 2015a; Hu et al. 2016; Vigneron et al. 2017; Liu et al. 2018), and composts and wastewater bioreactors that are reservoir for microbes decomposing organic compounds (Dougherty et al. 2012; Silva et al. 2012; Antunes et al. 2016; Berini et al. 2017). Intriguingly, samples from human gut systems were also candidates for LEs (LPGM values > 7500). The human intestinal microorganisms play an import role in degrading diet components into metabolizable molecules (Wang et al. 2015). The function- and sequence-based study of human gut metagenomes have proved that the human gut microbiome is a rich source for various carbohydrate active enzymes (Li et al. 2009; Turnbaugh et al. 2009; Tasse et al. 2010; Moore et al. 2011).

Overall, the assigned PLPs were classified into 34 lipolytic families (Fig 3). Members of the Hormone-sensitive_lipase_like and patatin-like-protein families were most abundant (average LPGM values across samples > 2000), followed by families of A85-EsteraseD-FGH, VIII and Bacterial_lip_FamI.1 (average LPGM values > 700) (Fig 3). However, no family was shared by all samples. Nevertheless, members from families of Hormone-sensitive-lipase-like, patatin-like-proteins, VIII, homoserine transacetylase, II and A85-Feruloyl-Esterase were detected in more than 90 % of samples (Figure 3). Enzymes belonging to families of PHAZ7_phb_depolymerase, Bact_LipEH166_FamXII and Bacterial_lip_FamI.2 were not or only rarely detected (< 6 % of all samples) and showed a low abundance (LPGM values < 1). The prevalence and abundance of a lipolytic family revealed by the sequence-based screening are dependent on the distribution of corresponding target genes in the microbial consortia (Wang et al. 2016). Taking members from the "abundant" family Hormone-sensitive_lipase_like as example, the corresponding genes are widely distributed in more than 1,200 species as recorded in the ESTHER database so far. This was, somehow, also reflected by the function-based screening, in which a large proportion of the identified LEs belonged to family Hormone-sensitive_lipase_like. In contrast, according to the ESTHER database, only 23, 8 and 6 species harboring LEs were affiliated to the "rare" families like PHAZ7_phb_depolymerase, PC-sterol_acyltransferase and Bact_LipEH166_FamXII, respectively (Supplementary Figure S12).

To investigate the distribution of assigned PLPs that cause the observed lipolytic family profiles across samples and habitats, a matrix with LPGM values representing the abundance of PLPs per lipolytic family identified in each metagenome was generated. The lipolytic family profiles clustered by habitats (Figure 3), which was confirmed by NMDS (stress level 0.2268; Supplementary Figure S9). ANOSIM (Clarke 1993) was used to pairwise compare the multivariate (group) differences of lipolytic family profiles between habitats. A R value-based matrix was generated among habitats (Supplementary Figure S9), a high R value (between 0 to 1) indicated a high group dissimilarity between two habitats. Generally, each habitat exhibited a distinctive pattern of lipolytic family profiles (overall R value = 0.6168; Supplementary Table S17). For example, PLPs detected in agricultural soils were only present in eight lipolytic families with low abundances. In contrast, PLPs in composts were detected in almost all lipolytic families, and with remarkably high abundance in families such as Hormone-sensitive_lipase_like, patatin-like-protein and VIII (Supplementary Figure S10). Notably, the lowest group dissimilarity was observed between the habitats compost and wastewater bioreactor (R=0.1941, P < 0.001, ANISOM; Supplementary Figure S9). The analysis of lipolytic profiles across habitats allows selecting suitable habitats for function-based screening, e.g. targeting LEs of a specific family or with some properties for desired applications. Metagenomes from composts are promising for recovering LEs in families LYsophospholipase_carboxylesterase (family VI), CarbLipBact_2 (family XIII-2/XVIII) and CarbLipBact_1 (family XIII-1) (Supplementary Figure S11).

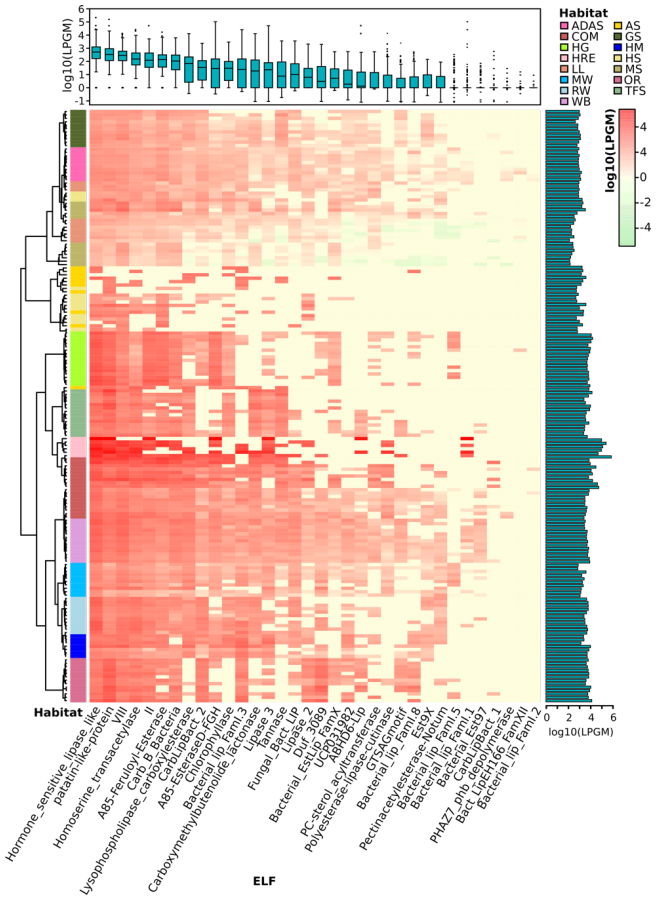


Figure 3. Lipolytic family profile of assigned PLPs across samples. Hierarchical clustering analysis of the lipolytic family profile in each sample was performed using the Ward.D clustering method and Bray-Curtis distance matrices. LPGM values were $\log_{10}$ transformed. The color intensity of the heat map (light green to red) indicates the change of LPGM values (low to high). The habitats are depicted by different colors. The lipolytic family profile in each sample was generally clustered by habitat (overall R value = 0.621, P<0.001, ANOSIM test). The boxplot (top) represents the distribution of the assigned PLPs in each ELF across samples. Mean values (n=175 samples) are given. The bar plot (right) shows the total abundance of assigned PLPs by summing up the abundance in each family of each sample. Abbreviations of habitats: ADAS, anaerobic digestor active sludge; AS, agricultural soil; COM, compost; GS grassland soil; HG, human gut; HM, hypersaline mat; HRE, hydrocarbon resource environment; HS, hot spring; LL, landfill leachate; MS, marine sediment; MW, marine water; OR, oil reservoir; RW, river water; TFS, tropical forest soil; WB, wastewater bioreactor; ELF, ESTHER lipolytic family.

The phylogenetic distribution of assigned PLPs

More than 98 % of the assigned PLPs were encoded by bacterial community members. Although LEs are widely encoded in various microbial genomes (Hausmann and Jaeger 2010; Ramnath et al. 2016; Kovacic et al. 2019a), the assigned PLPs were mainly derived from the bacterial phyla *Proteobacteria* (66.5 %), *Bacteroidetes* (12.5 %), *Actinobacteria* (7.7 %), *Firmicutes* (6.7 %) (Figure 4). This is consistent with the taxonomic origin of reference LEs in ESTHER database (Supplementary Figure S12). Moreover, enzymes from members of *Proteobacteria* were dominant in almost all lipolytic families (Figure 4). At genus level, the phylogenetic origins of assigned PLPs were scattered across approx. 2,000 bacterial genera, with enriched abundance in the genera *Acinetobacter*, *Pseudomonas*, *Bacteroides*, *Bradyrhizobium* and *Mycobacterium* (average LPGM values across samples > 180). Many of the LEs from these genera were described as exoenzymes (Rudek and Haque 1976; Gilbert 1993;

Snellman and Colwell 2004; Guo et al. 2010; Liu et al. 2018). Notably, a similar taxonomic enrichment at genus level was also observed for the reference LEs in ESTHER database as 960 LEs were encoded by *Mycobacterium*, 410 by *Pseudomonas*, 260 by *Bacteroides*, 166 by *Acinetobacter*, and 164 by *Bradyrhizobium* species (Supplementary Table S18).

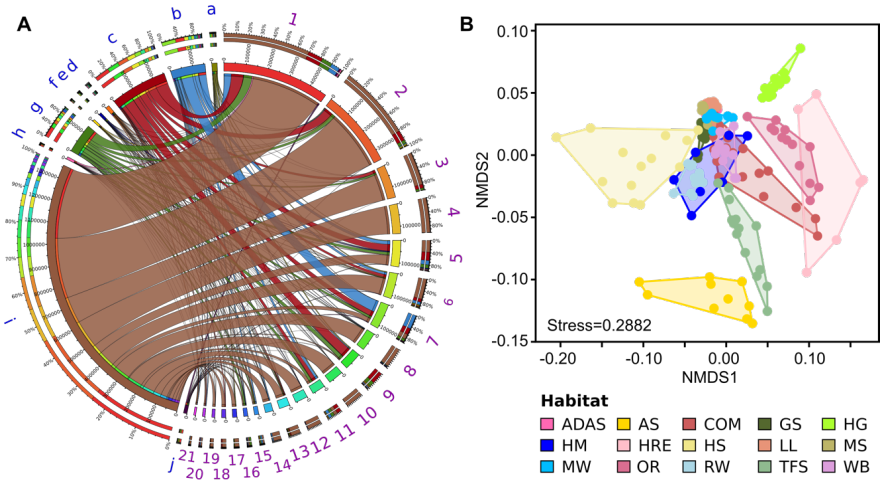


Figure 4. Phylogenetic distribution of assigned PLPs. **A**, phylogenetic distributions of assigned PLPs in abundant bacterial phyla possessing PLP-encoding genes across all the samples. The abundance inferred from LPGM values matrix of assigned PLPs per family identified in each bacterial phylum was generated by summing the corresponding LPGM values across all samples. The width of bars from each bacterial phylum and functional enzyme family indicates their relative abundances across all samples. a-j are bacterial phyla (in blue): a, *Acidobacteria*; b, *Actinobacteria*; c, *Bacteroidetes*; d, *Chloroflexi*; e, *Cyanobacteria*; f, *Deinococcus-Thermus*; g, *Firmicutes*; h, *Planctomycetes*; i, *Proteobacteria*; j, *Verrucomicrobia*. 1-21 are lipolytic families (in purple): 1, Hormone-sensitive_lipase_like; 2, patatin-like-protein; 3, A85-EsteraseD-FGH; 4, Bacterial_lip_FamI.1; 5, VIII; 6, Homoserine_transacetylase; 7, II; 8, Lipase_3; 9, A85-Feruloyl-Esterase; 10, ABHD6-Lip; 11, Carb_B_Bacteria; 12, Bacterial_lip_FamI.3; 13, Lysophospholipase_carboxylesterase; 14, Carboxymethylbutenolide_lactonase; 15, CarbLipBact_2; 16, Chlorophyllase; 17, Tannase; 18, Polyesterase-lipase-cutinase; 19, Duf_3089; 20, Fungal_Bact_LIP; 21, Lipase_2. Only phyla and lipolytic families with a relative abundance > 0.5 % are shown. **B**, Non-metric multidimensional scaling (NMDS) analysis of phylogenetic distribution of assigned PLPs across samples based on Bray-Curtis distances at bacterial genus level. Only genera with a mean LPGM values of ⩾ 0.5 across all the samples were included. Abbreviations of habitats: ADAS, anaerobic digestor active sludge; AS, agricultural soil; COM, compost; GS grassland soil; HG, human gut; HM, hypersaline mat; HRE, hydrocarbon resource environment; HS, hot spring; LL, landfill leachate; MS, marine sediment; MW, marine water; OR, oil reservoir; RW, river water; TFS, tropical forest soil; WB, wastewater bioreactor; ELF, ESTHER lipolytic family.

The taxonomic origin of assigned PLPs at genus level varied significantly across habitats (overall R value = 0.821, P <0.01), especially for the human gut system, oil reservoir and hydrocarbon resource environment (Supplementary Figure S14). The average R value was 0.98, 0.97 and 0.94, respectively (Supplementary Table S19). The lowest dissimilarity was observed between compost and wastewater bioreactor (R value = 0.2317, P <0.001, ANISOM).

Habitats harboring prevalent and distinct microbial clusters are main drivers of PLP distribution

Bipartite association networks have been used to identify microbial taxa responsible for shifts in community structures (Hartmann et al. 2015; Dukunde et al. 2019). In this study, a bipartite association network was constructed to visualize the associations between bacterial members at genus level that harbor lipolytic genes and habitats or habitat combinations (Figure 5). 225 of the total 712 genera, were not significantly separated in abundance and frequency by habitat. These belonged mainly to *Proteobacteria* (82 genera), *Bacteroidetes* (43 genera), *Firmicutes* (33 genera), and *Actinobacteria* (25 genera) (Supplementary Table S20). These non-significant genera were conserved across different habitats, generally represented the "indigenous group" (Hartmann et al. 2015; Wemheuer et al. 2017), and formed the core microbiota harboring lipolytic genes. This core microbiota was also an indication of the prevalence of lipolytic genes across microbes and habitats (Bornscheuer 2002; Hasan et al. 2006; Barriuso and Jesús Martínez 2015; Berini et al. 2017). In contrast, the

significant indicators, with respect to the "characteristic group" (Rime et al. 2016; Dukunde et al. 2019), highlighted the bacterial genera that were responsible for the change of assigned PLPs distribution across habitats (Figure 5). Particularly, the indicators associated with only one habitat defined the distinctiveness of microbiota in each habitat (Hartmann et al. 2015). In this study, the unique-associated indicators accounted for 76% of all significant indicators (Supplementary Table S20). This strongly resembled the ANISOM result, in which the high overall R value (0.8199) suggested a significant distinctiveness of the phylogenetic origins of assigned PLPs across habitats (Supplementary Table S19). With respect to each habitat, a high ratio of unique-associated indicators to the total significant genera in a habitat generally indicated a high R value (Pearson's r correlation = 0.6672, $P < 0.01$, linear regression; Supplementary Figure S15). For example, out of the 75 indicators that were significantly associated to the habitat hydrocarbon resource environment, 65 were unique-associated indicators, with a mean R value of 0.93 (Supplementary Table 19). This is also the case for the habitats oil reservoir (60 out of 75; mean R value = 0.96) and human gut system (35 out of 41; mean R value = 0.97).

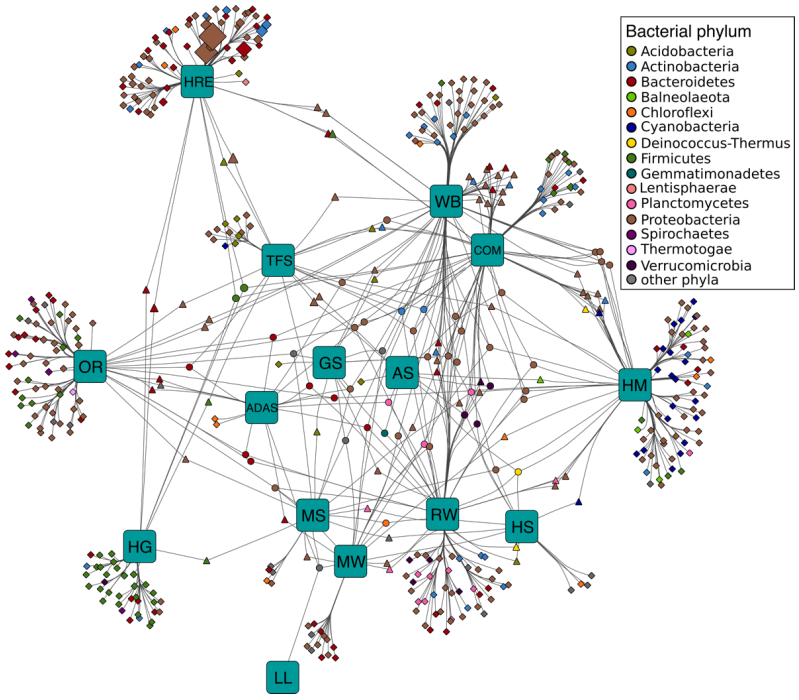


Figure 5. Association networks between bacterial origin of assigned PLPs at genus level and habitats. The abundance of PLPs in each genus per sample was presented by LPGM values, and only genera with mean LPGM values of ⩾ 0.5 across all the samples were used. Source nodes (rounded squares) represent habitats, target node represent bacterial genera (circles, diamonds and triangles), and edges represent associations between habitats and bacterial genera. Target node size represent its mean abundance inferred from LPGM values across habitats. Target node is colored according to its phylogenetic origin at phylum level. The length of edges is weighted according to association strength. Unique clusters, which associate with only one habitat, consist of nodes shaped as diamond. Triangle and circle nodes represent genera with significant cross association between two and more habitats, respectively. Data only represents genera that showed significant positive association with habitats ($P = 0.05$). For ease of visualization, edges were bundled together, with a stress value of 3. Abbreviations of habitats: ADAS, anaerobic digestor active sludge; AS, agricultural soil; COM, compost; GS grassland soil; HG, human gut; HM, hypersaline mat; HRE, hydrocarbon resource environment; HS, hot spring; LL, landfill leachate; MS, marine sediment; MW, marine water; OR, oil reservoir; RW, river water; TFS, tropical forest soil; WB, wastewater bioreactor; ELF, ESTHER lipolytic family.

Only a small fraction of the indicators exhibited cross associations between two (14 % of the total indicators) or more (10 %) habitats. Nevertheless, the 29 cross-associated indicators between habitats compost and wastewater bioreactor explained the low dissimilarity of phylogenetic distributions of assigned PLPs between the two habitats ($R=0.2317$, $P < 0.001$, ANISOM).

Similar to the "indigenous group", the "characteristic group" consisted mainly of genera affiliated to *Proteobacteria* (224 genera), *Bacteroidetes* (72), *Firmicutes* (49) and *Actinobacteria* (36). Among them, proteobacterial genera largely characterized the major habitats, such as tropical forest soil (83 %), wastewater bioreactor (67 %), hypersaline mat (52 %), hydrocarbon resource environment (51 %), oil reservoir (51 %), compost (50 %), marine water (46 %), river water (45 %), and grassland soil (42 %), whereas *Bacteroidetes* and *Firmicutes* characterized the human gut system (68 %) and the active sludge of an anaerobic digestor (53 %) (Supplementary Table S20). Noteworthy, the unique-associated indicators affiliated to *Cyanobacteria* were primarily enriched in the hypersaline mat (95 % indicators), which is also the case for *Planctomycetes* and *Verrucomicrobia* in river water (88 and 80 %, respectively). Pehrsson et al (2015) detected a link between microbial community structure and functional gene repertoire. This link could be extended to the distribution pattern of indicators in our study. For example, various studies have proved that the microbes in human gut systems were dominated by *Firmicutes* (Mahowald et al. 2009; Vital et al. 2014; Rinninella et al. 2019), which in turn leads to the *Firmicutes*-dominated indicators for lipolytic genes (Figure 5). Among all the habitats, only hypersaline mats were featured by the *Cyanobacteria*-dominated oxygenic layer for photosynthesis (Sørensen et al. 2005; Lindemann et al. 2013), which explained that almost all the *Cyanobacteria* indicators were associated with the hypersaline mat (Figure 5).

Conclusions

In this study, two compost samples (compost55 and compost76) were used for metagenomic screening of potential lipolytic genes. Through the function-driven screening, 115 unique LEs were identified and assigned into 12 known lipolytic families. In addition, 7 LEs were not assigned to any known family, indicating new branches of lipolytic families. Our results show that functional screening is a promising approach to discover novel lipolytic genes, particularly for targeted genes, whose function is not predicted based on DNA sequence alone. For sequence-based screening, we have developed a search and annotation strategy specific for putative lipolytic genes in metagenomes (Supplementary Figure 1). Our profile HMM-based searching methods yielded higher sensitivity (recall) for LEs than the BLASTp-derived counterpart. The annotation method also remarkably increased the specificity and accuracy in distinguishing lipolytic from non-lipolytic proteins. With this sequence-based strategy, we identified the putative lipolytic genes within the two compost metagenomes. Analysis of the phylogenetic origin of these genes indicated a potential link between microbial taxa and their functional traits. By comparing the lipolytic hits identified by function-driven and sequence-based screening, we conclude that the best way for exploring and exploiting LEs is to combine both approaches.

In addition, assembled metagenomes from samples of various habitats were used for comparative analysis of the PLP distribution. We profiled the lipolytic family and phylogenetic origin of assigned PLPs for each sample. The two profiles were generally driven by ecological factors, i.e. the habitat. Moreover, the habitat also determined the conserved and distinctive microbial groups harboring the putative lipolytic genes.

PLPs were also mainly enriched in the bacterial phyla *Proteobacteria*, *Bacteroidetes*, *Actinobacteria*, *Firmicutes* (Supplementary Figure S16). The profile of the phylogenetic total PLPdistribution in each sample clustered also by habitats (Supplementary Figures S17, S18 and S19). The bipartite association network identified the conserved and distinctive microbial groups harboring PLP-encoding genes among the habitats (Supplementary Tables S21 and S22). Thus, our study provided a sequence-based strategy for effective identification and annotation of potential lipolytic genes in assembled metagenomes. More importantly, through this strategy, the overview of how the lipolytic genes distributed ecologically (in various habitats), functionally (in different lipolytic enzyme families) and phylogenetically (in diverse microbial groups) is an advantage for novel and/or industrially relevant LE identification.

Supplementary Materials: The supplementary materials are also available online. The Supplementary Figures S1-S19 and the Supplementary Tables S1-S22 are presented in two files (Supplementary Figures S1-S19.pdf and Supplementary Tables S1-S22.xlsx, respectively)

Acknowledgments: We thank Dr. Silja Brady and Mechthild Bömeke for providing technical assistance.

Funding: We acknowledge the support of Mingji Lu by "Erasmus Mundus Action 2 – Lotus I Project". The funders had no role in study design, data collection, and interpretation, or the decision to submit the work for publication.

Conflicts of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Authors Approvals: All authors have seen and approved the manuscript. The manuscript has not been accepted or published elsewhere.

References

Akmoussi-Toumi S, Khemili-Talbi S, Ferioune I, Kebbouche-Gana S (2018) Purification and characterization of an organic solvent-tolerant and detergent-stable lipase from *Haloferax mediterranei* CNCMM 50101. Int J Biol Macromol 116:817–830 . doi: 10.1016/j.ijbiomac.2018.05.087

Akoh CC, Lee GC, Liaw YC, Huang TH, Shaw JF (2004) GDSL family of serine esterases/lipases. Prog Lipid Res 43:534–552 . doi: 10.1016/j.plipres.2004.09.002

Alisch M, Feuerhack A, Müller H, Mensak B, Andreaus J, Zimmermann W (2004) Biocatalytic modification of polyethylene terephthalate fibres by esterases from actinomycete isolates. Biocatal Biotransformation 22:347–351 . doi: 10.1080/10242420400025877

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410 . doi: 10.1016/S0022-2836(05)80360-2

Antunes LP, Martins LF, Pereira RV, Thomas AM, Barbosa D, Lemos LN, Silva GMM, Moura LMS, Epamino GWC, Digiampietri LA, Lombardi KC, Ramos PL, Quaggio RB, De Oliveira JCF, Pascon RC, Da Cruz JB, Da Silva AM, Setubal JC (2016) Microbial community structure and dynamics in thermophilic composting viewed through metagenomics and metatranscriptomics. Sci Rep 6:38915 . doi: 10.1038/srep38915

Arpigny JL, Jaeger K-E (1999) Bacterial lipolytic enzymes: classification and properties. Biochem J 343:177–183 . doi: 10.1042/0264-6021:3430177

Asnicar F, Weingart G, Tickle TL, Huttenhower C, Segata N (2015) Compact graphical representation of phylogenetic data and metadata with GraPhlAn. PeerJ 3:e1029 . doi: 10.7717/peerj.1029

Atkinson HJ, Morris JH, Ferrin TE, Babbitt PC (2009) Using sequence similarity networks for visualization of relationships across diverse protein superfamilies. PLoS One 4:e4345 . doi: 10.1371/journal.pone.0004345

Azziz G, Giménez M, Romero H, Valdespino-Castillo PM, Falcón LI, Ruberto LAM, Mac Cormack WP, Batista S (2019) Detection of presumed genes encoding beta-lactamases by sequence based screening of metagenomes derived from Antarctic microbial mats. Front Environ Sci Eng 13:44 . doi: 10.1007/s11783-019-1128-1

Banerji S, Flieger A (2004) Patatin-like proteins: A new family of lipolytic enzymes present in bacteria? Microbiology 150:522–525 . doi: 10.1099/mic.0.26957-0

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin A V., Sirotkin A V., Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA (2012) SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477 . doi: 10.1089/cmb.2012.0021

Barriuso J, Jesús Martínez M (2015) In silico metagenomes mining to discover novel esterases with industrial application by sequential search strategies. J Microbiol Biotechnol 25:732–737 . doi: 10.4014/jmb.1406.06049

Berglund F, Marathe NP, Österlund T, Bengtsson-Palme J, Kotsakis S, Flach CF, Larsson DGJ, Kristiansson E (2017) Identification of 76 novel B1 metallo-β-lactamases through large-scale screening of genomic and metagenomic data. Microbiome 5:134 . doi: 10.1186/s40168-017-0353-8

Berini F, Casciello C, Marcone GL, Marinelli F (2017) Metagenomics: novel enzymes from non-culturable microbes. FEMS Microbiol Lett 364: . doi: 10.1093/femsle/fnx211

Berlemont R, Spee O, Delsaute M, Lara Y, Schuldes J, Simon C, Power P, Daniel R, Galleni M (2013) Novel organic solvent-tolerant esterase isolated by metagenomics: insights into the lipase/esterase classification. Rev Argent Microbiol 45:3–12

Biver S, Vandenbol M (2013) Characterization of three new carboxylic ester hydrolases isolated by functional screening of a forest soil metagenomic library. J Ind Microbiol Biotechnol 40:191–200 . doi: 10.1007/s10295-012-1217-7

Blaya J, Marhuenda FC, Pascual JA, Ros M (2016) Microbiota characterization of compost using omics approaches opens new perspectives for phytophthora root rot control. PLoS One 11:e0158048 . doi: 10.1371/journal.pone.0158048

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120 . doi: 10.1093/bioinformatics/btu170

Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, Bai Y, Bisanz JE, Bittinger K, Brejnrod A, Brislawn CJ, Brown CT, Callahan BJ, Caraballo-Rodríguez AM, Chase J, Cope EK, Da Silva R, Diener C, Dorrestein PC, Douglas GM, Durall DM, Duvallet C, Edwardson CF, Ernst M, Estaki M, Fouquier J, Gauglitz JM, Gibbons SM, Gibson DL, Gonzalez A, Gorlick K, Guo J, Hillmann B, Holmes S, Holste H, Huttenhower C, Huttley GA, Janssen S, Jarmusch AK, Jiang L, Kaehler BD, Kang K Bin, Keefe CR, Keim P, Kelley ST, Knights D, Koester I, Kosciolek T, Kreps J, Langille MGI, Lee J, Ley R, Liu YX, Loftfield E, Lozupone C, Maher M, Marotz C, Martin BD, McDonald D, McIver LJ, Melnik A V., Metcalf JL, Morgan SC, Morton JT, Naimey AT, Navas-Molina JA, Nothias LF, Orchanian SB, Pearson T, Peoples SL, Petras D, Preuss ML, Pruesse E, Rasmussen LB, Rivers A, Robeson MS, Rosenthal P, Segata N, Shaffer M, Shiffer A, Sinha R, Song SJ, Spear JR, Swafford AD,

Thompson LR, Torres PJ, Trinh P, Tripathi A, Turnbaugh PJ, Ul-Hasan S, van der Hooft JJJ, Vargas F, Vázquez-Baeza Y, Vogtmann E, von Hippel M, Walters W, Wan Y, Wang M, Warren J, Weber KC, Williamson CHD, Willis AD, Xu ZZ, Zaneveld JR, Zhang Y, Zhu Q, Knight R, Caporaso JG (2019) Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat. Biotechnol. 37:852–857

Bornscheuer UT (2002) Microbial carboxyl esterases: classification, properties and application in biocatalysis. FEMS Microbiol Rev 26:73–81 . doi: 10.1111/j.1574-6976.2002.tb00599.x

Brault G, Shareck F, Hurtubise Y, Lépine F, Doucet N (2012) Isolation and characterization of EstC, a new cold-active esterase from Streptomyces coelicolor A3(2). PLoS One 7:e32041 . doi: 10.1371/journal.pone.0032041

Bzhalava Z, Hultin E, Dillner J (2018) Extension of the viral ecology in humans using viral profile hidden Markov models. PLoS One 13:e0190938 . doi: 10.1371/journal.pone.0190938

Cáceres M De (2013) How to use the indicspecies package (ver. 1.7.1). R Proj. 29

Castilla A, Panizza P, Rodríguez D, Bonino L, Díaz P, Irazoqui G, Rodríguez Giordano S (2017) A novel thermophilic and halophilic esterase from Janibacter sp. R02, the first member of a new lipase family (Family XVII). Enzyme Microb Technol 98:86–95 . doi: 10.1016/J.ENZMICTEC.2016.12.010

Chahinian H, Ali Y Ben, Abousalham A, Petry S, Mandrich L, Manco G, Canaan S, Sarda L (2005) Substrate specificity and kinetic properties of enzymes belonging to the hormone-sensitive lipase family: Comparison with non-lipolytic and lipolytic carboxylesterases. Biochim Biophys Acta 1738:29–36 . doi: 10.1016/j.bbalip.2005.11.003

Chan WY, Wong M, Guthrie J, Savchenko A V., Yakunin AF, Pai EF, Edwards EA (2010) Sequence- and activity-based screening of microbial genomes for novel dehalogenases. Microb Biotechnol 3:107–120 . doi: 10.1111/j.1751-7915.2009.00155.x

Chen IMA, Markowitz VM, Chu K, Palaniappan K, Szeto E, Pillay M, Ratner A, Huang J, Andersen E, Huntemann M, Varghese N, Hadjithomas M, Tennessen K, Nielsen T, Ivanova NN, Kyrpides NC (2017) IMG/M: Integrated genome and metagenome comparative data analysis system. Nucleic Acids Res 45:507–516 . doi: 10.1093/nar/gkw929

Chen Y, Black DS, Reilly PJ (2016) Carboxylic ester hydrolases: Classification and database derived from their primary, secondary, and tertiary structures. Protein Sci 25:1942–1953 . doi: 10.1002/pro.3016

Chevreux B, Wetter T, Suhai S (1999) Genome sequence assembly using trace signals and additional sequence information. Comput Sci Biol Proc Ger Conf Bioinforma 99:45–56

Choi JE, Kwon MA, Na HY, Hahm DH, Song JK (2013) Isolation and characterization of a metagenome-derived thermoalkaliphilic esterase with high stability over a broad pH range. Extremophiles 17:1013–1021 . doi: 10.1007/s00792-013-0583-z

Chow J, Kovacic F, Dall Antonia Y, Krauss U, Fersini F, Schmeisser C, Lauinger B, Bongen P, Pietruszka J, Schmidt M, Menyes I, Bornscheuer UT, Eckstein M, Thum O, Liese A, Mueller-Dieckmann J, Jaeger KE, Streit WR (2012) The metagenome-derived enzymes LipS and LipT increase the diversity of known lipases. PLoS One 7:e47665 . doi: 10.1371/journal.pone.0047665

Chu X, He H, Guo C, Sun B (2008) Identification of two novel esterases from a marine metagenomic library derived from South China Sea. Appl Microbiol Biotechnol 80:615–625 . doi: 10.1007/s00253-008-1566-3

Clarke KR (1993) Non‐parametric multivariate analyses of changes in community structure. Aust J Ecol 18:117–143 . doi: 10.1111/j.1442-9993.1993.tb00438.x

Couto GH, Glogauer A, Faoro H, Chubatsu LS, Souza EM, Pedrosa FO (2010) Isolation of a novel lipase from a metagenomic library derived from mangrove sediment from the south Brazilian coast. Genet Mol Res 9:514–523

Dalcin Martins P, Danczak RE, Roux S, Frank J, Borton MA, Wolfe RA, Burris MN, Wilkins MJ (2018) Viral and metabolic controls on high rates of microbial sulfur and carbon cycling in wetland ecosystems. Microbiome 6:138 . doi: 10.1186/s40168-018-0522-4

Diamond S, Andeer PF, Li Z, Crits-Christoph A, Burstein D, Anantharaman K, Lane KR, Thomas BC, Pan C, Northen TR, Banfield JF (2019) Mediterranean grassland soil C–N compound turnover is dependent on rainfall and depth, and is mediated by genomically divergent microorganisms. Nat Microbiol 4:1356–1367 . doi: 10.1038/s41564-019-0449-y

Dougherty MJ, D'haeseleer P, Hazen TC, Simmons BA, Adams PD, Hadi MZ (2012) Glycoside hydrolases from a targeted compost metagenome, activity-screening and functional characterization. BMC Biotechnol 12:38 . doi: 10.1186/1472-6750-12-38

Dukunde A, Schneider D, Lu M, Brady S, Daniel R (2017) A novel, versatile family IV carboxylesterase exhibits high stability and activity in a broad pH spectrum. Biotechnol Lett 39:577–587 . doi: 10.1007/s10529-016-2282-1

Dukunde A, Schneider D, Schmidt M, Veldkamp E, Daniel R (2019) Tree species shape soil bacterial community structure and function in temperate deciduous forests. Front Microbiol 10:1519 . doi: 10.3389/fmicb.2019.01519

Eddy SR (2011) Accelerated Profile HMM Searches. PLoS Comput Biol 7:e1002195 . doi: 10.1371/journal.pcbi.1002195

Eddy SR (2018) HMMER User's Guide Biological sequence analysis using profile hidden Markov models

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797 . doi: 10.1093/nar/gkh340

Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R (2011) UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 27:2194–2200 . doi: 10.1093/bioinformatics/btr381

Egelkamp R, Zimmermann T, Schneider D, Hertel R, Daniel R (2019) Impact of nitriles on bacterial communities. Front Environ Sci 7:103 . doi: 10.3389/fenvs.2019.00103

Eric Sayers GW, Cavanaugh M, Clark K, Ostell J, Pruitt KD, Karsch-Mizrachi I (2019) GenBank. Nucleic Acids Res 47:94–99 . doi: 10.1093/nar/gky989

Esteban-Torres M, Santamaría L, de las Rivas B, Muñoz R (2014) Characterisation of a cold-active and salt-tolerant esterase from Lactobacillus plantarum with potential application during cheese ripening. Int Dairy J 39:312–315 . doi: 10.1016/j.idairyj.2014.08.004

Fan X, Liu X, Wang K, Wang S, Huang R, Liu Y (2011) Highly soluble expression and molecular characterization of an organic solvent-stable and thermotolerant lipase originating from the metagenome. J Mol Catal B Enzym 72:319–325 . doi: 10.1016/j.molcatb.2011.07.009

Fang Z, Li J, Wang Q, Fang W, Peng H, Zhang X, Xiao Y (2014) A novel esterase from a marine metagenomic library exhibiting salt tolerance ability. J Microbiol Biotechnol 24:771–780 . doi: 10.4014/jmb.1311.11071

Ferrer M, Martínez-Martínez M, Bargiela R, Streit WR, Golyshina O V, Golyshin PN (2015) Estimating the success of enzyme bioprospecting through metagenomics: current status and future trends. Microb Biotechnol 9:22–34 . doi: 10.1111/1751-7915.12309

Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J, Sonnhammer ELL, Tate J, Punta M (2014) Pfam: the protein families database. Nucleic Acids Res 42:D222–D230 . doi: 10.1093/nar/gkt1223

Fu L, He Y, Xu F, Ma Q, Wang F, Xu J (2015) Characterization of a novel thermostable patatin-like protein from a Guaymas basin metagenomic library. Extremophiles 19:829–840 . doi: 10.1007/s00792-015-0758-x

Gao W, Wu K, Chen L, Fan H, Zhao Z, Gao B, Wang H, Wei D (2016) A novel esterase from a marine mud metagenomic library for biocatalytic synthesis of short-chain flavor esters. Microb Cell Fact 15:41 . doi: 10.1186/s12934-016-0435-5

Gerlt JA, Bouvier JT, Davidson DB, Imker HJ, Sadkhin B, Slater DR, Whalen KL (2015) Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST): A web tool for generating protein sequence similarity networks. Biochim Biophys Acta 1854:1019–1037 . doi: 10.1002/cncr.27633.Percutaneous

Gibson MK, Forsberg KJ, Dantas G (2015) Improved annotation of antibiotic resistance determinants reveals microbial resistomes cluster by ecology. ISME J 9:207–216 . doi: 10.1038/ismej.2014.106

Gilbert EJ (1993) Pseudomonas lipases: Biochemical properties and molecular cloning. Enzyme Microb. Technol. 15:634–645

Glogauer A, Martini VP, Faoro H, Couto GH, Müller-Santos M, Monteiro RA, Mitchell DA, de Souza EM, Pedrosa FO, Krieger N (2011) Identification and characterization of a new true lipase isolated through metagenomic approach. Microb Cell Fact 10:54 . doi: 10.1186/1475-2859-10-54

Guo H, Zhang Y, Shao Y, Chen W, Chen F, Li M (2016) Cloning, expression and characterization of a novel cold-active and organic solvent-tolerant esterase from Monascus ruber M7. Extremophiles 20:451–459 . doi: 10.1007/s00792-016-0835-9

Guo J, Zheng X, Xu L, Liu Z, Xu K, Li S, Wen T, Liu S, Pang H (2010) Characterization of a novel esterase Rv0045c from Mycobacterium tuberculosis. PLoS One 5:4–10 . doi: 10.1371/journal.pone.0013143

Hammer DAT, Ryan PD, Hammer Ø, Harper DAT (2001) Past: Paleontological Statistics Software Package for Education and Data Analysis

Hartmann M, Frey B, Mayer J, Mäder P, Widmer F (2015) Distinct soil microbial diversity under long-term organic and conventional farming. ISME J 9:1177–1194 . doi: 10.1038/ismej.2014.210

Hasan F, Shah AA, Hameed A (2006) Industrial applications of microbial lipases. Enzyme Microb Technol 39:235–251 . doi: 10.1016/J.ENZMICTEC.2005.10.016

Hausmann S, Jaeger K-E (2010) Lipolytic enzymes from bacteria. In: Timmis KN (ed) Handbook of Hydrocarbon and Lipid Microbiology. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 1099–1126

Hitch TCA, Clavel T (2019) A proposed update for the classification and description of bacterial lipolytic enzymes. PeerJ 7:e7249 . doi: 10.7717/peerj.7249

Hong JK, Choi HW, Hwang IS, Kim DS, Kim NH, Choi DS, Kim YJ, Hwang BK (2008) Function of a novel GDSL-type pepper lipase gene, CaGLIP1, in disease susceptibility and abiotic stress tolerance. Planta 227:539–558 . doi: 10.1007/s00425-007-0637-5

Hu P, Tom L, Singh A, Thomas BC, Baker BJ, Piceno YM, Andersen GL, Banfield JF (2016) Genome-resolved metagenomic analysis reveals roles for candidate phyla and other microbial community members in biogeochemical transformations in oil reservoirs. MBio 7:e01669-15 . doi: 10.1128/mBio.01669-15

Hu X, Thumarat U, Zhang X, Tang M, Kawai F (2010a) Diversity of polyester-degrading bacteria in compost and molecular analysis of a thermoactive esterase from Thermobifida alba AHK119. Appl Microbiol Biotechnol 87:771–779 . doi: 10.1007/s00253-010-2555-x

Hu Y, Fu C, Huang Y, Yin Y, Cheng G, Lei F, Lu N, Li J, Ashforth EJ, Zhang L, Zhu B (2010b) Novel lipolytic genes from the microbial metagenomic library of the South China Sea marine sediment. FEMS Microbiol Ecol 72:228–37 . doi: 10.1111/j.1574-6941.2010.00851.x

Hubbe MA, Nazhad M, Sánchez C (2010) Composting as a way to convert cellulosic biomass and organic waste into high-value soil admendments: a review. BioResources 5:2808–2854

Hugenholtz P, Tyson GW (2008) Microbiology: Metagenomics. Nature 455:481–483

Jayanath G, Mohandas SP, Kachiprath B, Solomon S, Sajeevan TP, Bright Singh IS, Philip R (2018) A novel solvent tolerant esterase of GDSGG motif subfamily from solar saltern through metagenomic approach: Recombinant expression and characterization. Int J Biol Macromol 119:393–401 . doi: 10.1016/j.ijbiomac.2018.06.057

Jeon JH, Kim JT, Kim YJ, Kim HK, Lee HS, Kang SG, Kim SJ, Lee JH (2009) Cloning and characterization of a new cold-active lipase from a deep-sea sediment metagenome. Appl Microbiol Biotechnol 81:865–874 . doi: 10.1007/s00253-008-1656-2

Jeon JH, Kim S-J, Lee HS, Cha S-S, Lee JH, Yoon S-H, Koo B-S, Lee C-M, Choi SH, Lee SH, Kang SG, Lee J-H (2011) Novel metagenome-derived carboxylesterase that hydrolyzes β-lactam antibiotics. Appl Environ Microbiol 77:7830–6 . doi: 10.1128/AEM.05363-11

Jia ML, Zhong XL, Lin ZW, Dong BX, Li G (2019) Expression and characterization of an esterase belonging to a new family via isolation from a metagenomic library of paper mill sludge. Int J Biol Macromol 126:1192–1200 . doi: 10.1016/j.ijbiomac.2019.01.025

Jiang X, Huo Y, Cheng H, Zhang X, Zhu X, Wu M (2012) Cloning, expression and characterization of a halotolerant esterase from a marine bacterium *Pelagibacterium halotolerans* B2T. Extremophiles 16:427–435 . doi: 10.1007/s00792-012-0442-3

Jiménez DJ, Montaña JS, Álvarez D, Baena S (2012) A novel cold active esterase derived from Colombian high Andean forest soil metagenome. World J Microbiol Biotechnol 28:361–370 . doi: 10.1007/s11274-011-0828-x

Kaminski J, Gibson MK, Franzosa EA, Segata N, Dantas G, Huttenhower C (2015) High-specificity targeted functional profiling in microbial communities with ShortBRED. PLOS Comput Biol 11:e1004557 . doi: 10.1371/journal.pcbi.1004557

Kang CH, Oh KH, Lee MH, Oh TK, Kim BH, Yoon JH (2011) A novel family VII esterase with industrial potential from compost metagenomic library. Microb Cell Fact 10:41 . doi: 10.1186/1475-2859-10-41

Keegan KP, Glass EM, Meyer F (2016) MG-RAST, a metagenomics service for analysis of microbial community structure and function. In: Methods in Molecular Biology. Humana Press Inc., pp 207–233

Kim YH, Kwon EJ, Kim SK, Jeong YS, Kim J, Yun HD, Kim H (2010) Molecular cloning and characterization of a novel family VIII alkaline esterase from a compost metagenomic library. Biochem Biophys Res Commun 393:45–49 . doi: 10.1016/j.bbrc.2010.01.070

Knudsen M, Wiuf C (2010) The CATH database. Hum Genomics 4:207–212 . doi: 10.1186/1479-7364-4-3-207

Ko KC, Rim SO, Han Y, Shin BS, Kim GJ, Choi JH, Song JJ (2012) Identification and characterization of a novel cold-adapted esterase from a metagenomic library of mountain soil. J Ind Microbiol Biotechnol 39:681–689 . doi: 10.1007/s10295-011-1080-y

Kovacic F, Babic N, Krauss U, Jaeger K (2019) Classification of Lipolytic Enzymes from Bacteria. In: Rojo F (ed) Aerobic Utilization of Hydrocarbons, Oils and Lipids. Handbook of Hydrocarbon and Lipid Microbiology. Springer, Cham, pp 1–35

Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA (2009) Circos: an information aesthetic for comparative genomics. Genome Res 19:1639–1645 . doi: 10.1101/gr.092759.109

Lam KN, Cheng J, Engel K, Neufeld JD, Charles TC (2015) Current and future resources for functional metagenomics. Front Microbiol 6:1196 . doi: 10.3389/fmicb.2015.01196

Lämmle K, Zipper H, Breuer M, Hauer B, Buta C, Brunner H, Rupp S (2007) Identification of novel enzymes with different hydrolytic activities by metagenome expression cloning. J Biotechnol 127:575–592 . doi: 10.1016/j.jbiotec.2006.07.036

Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359 . doi: 10.1038/nmeth.1923

Lapébie P, Lombard V, Drula E, Terrapon N, Henrissat B (2019) Bacteroidetes use thousands of enzyme combinations to break down glycans. Nat Commun 10:1–7 . doi: 10.1038/s41467-019-10068-5

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23:2947–2948 . doi: 10.1093/bioinformatics/btm404

Lee JH, Jeon JH, Kim JT, Lee HS, Kim SJ, Kang SG, Choi SH (2011) Novel lipolytic enzymes identified from metagenomic library of deep-sea sediment. Evidence-based Complement Altern Med 2011: . doi: 10.1155/2011/271419

Lee SW, Won K, Lim HK, Kim JC, Choi GJ, Cho KY (2004) Screening for novel lipolytic enzymes from uncultured soil microorganisms. Appl Microbiol Biotechnol 65:720–726 . doi: 10.1007/s00253-004-1722-3

Lee Y-S (2016) Isolation and characterization of a novel cold-adapted esterase, MtEst45, from microbulbifer thermotolerans DAU221. Front Microbiol 7:218 . doi: 10.3389/fmicb.2016.00218

Leis B, Angelov A, Mientus M, Li H, Pham VTT, Lauinger B, Bongen P, Pietruszka J, Gonçalves LG, Santos H, Liebl W (2015) Identification of novel esterase-active enzymes from hot environments by use of the host bacterium Thermus thermophilus. Front Microbiol 6:275 . doi: 10.3389/fmicb.2015.00275

Lenfant N, Hotelier T, Velluet E, Bourne Y, Marchot P, Chatonnet A (2013) ESTHER, the database of the α/β-hydrolase fold superfamily of proteins: tools to explore diversity of functions. Nucleic Acids Res 41:423–429 . doi: 10.1093/nar/gks1154

Lewin GR, Carlos C, Chevrette MG, Horn HA, McDonald BR, Stankey RJ, Fox BG, Currie CR (2016) Evolution and ecology of Actinobacteria and their bioenergy applications. Annu Rev Microbiol 70:235–254 . doi: 10.1146/annurev-micro-102215-095748

Li H, Han X, Qiu W, Xu D, Wang Y, Yu M, Hu X, Zhuo R (2019) Identification and expression analysis of the GDSL esterase/lipase family genes, and the characterization of SaGLIP8 in Sedum alfredii Hance under cadmium stress. PeerJ 7:e6741 . doi: 10.7717/peerj.6741

Li LL, McCorkle SR, Monchy S, Taghavi S, van der Lelie D (2009) Bioprospecting metagenomes: Glycosyl hydrolases for converting biomass. Biotechnol Biofuels 2:1–11 . doi: 10.1186/1754-6834-2-10

Li M, Yang LR, Xu G, Wu JP (2016) Cloning and characterization of a novel lipase from *Stenotrophomonas maltophilia* GS11: The first member of a new bacterial lipase family XVI. J Biotechnol 228:30–36 . doi: 10.1016/j.jbiotec.2016.04.034

Li PY, Yao QQ, Wang P, Zhang Y, Li Y, Zhang YQ, Hao J, Zhou BC, Chen XL, Shi M, Zhang YZ, Zhang XY (2017) A novel subfamily esterase with a homoserine transacetylase-like fold but no transferase activity. Appl Environ Microbiol 83:e00131-17 . doi: 10.1128/AEM.00131-17

Liaw R-B, Cheng M-P, Wu M-C, Lee C-Y (2010) Use of metagenomic approaches to isolate lipolytic genes from activated sludge. Bioresour Technol 101:8323–8329 . doi: 10.1016/j.biortech.2010.05.091

Lima-Junior JD, Viana-Niero C, Conde Oliveira D V., Machado GE, Rabello MC da S, Martins-Junior J, Martins LF, Digiampietri LA, da Silva AM, Setubal JC, Russell DA, Jacobs-Sera D, Pope WH, Hatfull GF, Leão SC (2016) Characterization of mycobacteria and mycobacteriophages isolated from compost at the São Paulo Zoo Park Foundation in Brazil and creation of the new mycobacteriophage Cluster U. BMC Microbiol 16:111 . doi: 10.1186/s12866-016-0734-3

Lindemann SR, Moran JJ, Stegen JC, Renslow RS, Hutchison JR, Cole JK, Dohnalkova AC, Tremblay J, Singh K, Malfatti SA, Chen F, Tringe SG, Beyenal H, Fredrickson JK (2013) The epsomitic phototrophic microbial mat of Hot Lake, Washington: Community structural responses to seasonal cycling. Front Microbiol 4:323 . doi: 10.3389/fmicb.2013.00323

Liu JF, Sun XB, Yang GC, Mbadinga SM, Gu JD, Mu BZ (2015a) Analysis of microbial communities in the oil reservoir subjected to CO2-flooding by using functional genes as molecular biomarkers for microbial CO2 sequestration. Front Microbiol 6:236 . doi: 10.3389/fmicb.2015.00236

Liu K, Wang J, Bu D, Zhao S, McSweeney C, Yu P, Li D (2009) Isolation and biochemical characterization of two lipases from a metagenomic library of China Holstein cow rumen. Biochem Biophys Res Commun 385:605–611 . doi: 10.1016/j.bbrc.2009.05.110

Liu SP, Liu RX, Zhang L, Shi GY (2015b) Sequence-based screening and characterization of cytosolic mandelate oxidase using oxygen as electron acceptor. Enzyme Microb Technol 69:24–30 . doi: 10.1016/j.enzmictec.2014.11.001

Liu SP, Liu RX, Zhang L, Shi GY (2015c) Sequence-based screening and characterization of cytosolic mandelate oxidase using oxygen as electron acceptor. Enzyme Microb Technol 69:24–30 . doi: 10.1016/j.enzmictec.2014.11.001

Liu YF, Galzerani DD, Mbadinga SM, Zaramela LS, Gu JD, Mu BZ, Zengler K (2018) Metabolic capability and in situ activity of microorganisms in an oil reservoir. Microbiome 6:5 . doi: 10.1186/s40168-017-0392-1

Liu Z, Lozupone C, Hamady M, Bushman FD, Knight R (2007) Short pyrosequencing reads suffice for accurate microbial community analysis. Nucleic Acids Res 35:e120 . doi: 10.1093/nar/gkm541

López-López O, Knapik K, Cerdán ME, González-Siso MI (2015) Metagenomics of an alkaline hot spring in Galicia (Spain): Microbial diversity analysis and screening for novel lipolytic enzymes. Front Microbiol 6:1291 . doi: 10.3389/fmicb.2015.01291

Lu M, Dukunde A, Daniel R (2019) Biochemical profiles of two thermostable and organic solvent–tolerant esterases derived from a compost metagenome. Appl Microbiol Biotechnol 103:3421–3437 . doi: 10.1007/s00253-019-09695-1

Ma S, Fang C, Sun X, Han L, He X, Huang G (2018) Bacterial community succession during pig manure and wheat straw aerobic composting covered with a semi-permeable membrane under slight positive pressure. Bioresour Technol 259:221–227 . doi: 10.1016/j.biortech.2018.03.054

Mahowald MA, Rey FE, Seedorf H, Turnbaugh PJ, Fulton RS, Wollam A, Shah N, Wang C, Magrini V, Wilson RK, Cantarel BL, Coutinho PM, Henrissat B, Crock LW, Russell A, Verberkmoes NC, Hettich RL, Gordon JI (2009) Characterizing a model human gut microbiota composed of members of its two dominant bacterial phyla. Proc Natl Acad Sci U S A 106:5859–5864 . doi: 10.1073/pnas.0901529106

Maimanakos J, Chow J, Gaßmeyer SK, Güllert S, Busch F, Kourist R, Streit WR (2016) Sequence-based screening for rare enzymes: new insights into the world of AMDases reveal a conserved motif and 58 novel enzymes clustering in eight distinct families. Front Microbiol 7:1332 . doi: 10.3389/fmicb.2016.01332

Mander P, Yoo HY, Kim SW, Choi YH, Cho SS, Yoo JC (2014) Transesterification of waste cooking oil by an organic solvent-tolerant alkaline lipase from *Streptomyces* sp. CS273. Appl Biochem Biotechnol 172:1377–1389 . doi: 10.1007/s12010-013-0610-7

Martins LF, Antunes LP, Pascon RC, de Oliveira JCF, Digiampietri LA, Barbosa D, Peixoto BM, Vallim MA, Viana-Niero C, Ostroski EH, Telles GP, Dias Z, da Cruz JB, Juliano L, Verjovski-Almeida S, da Silva AM, Setubal JC (2013) Metagenomic analysis of a tropical composting operation at the São Paulo zoo park reveals diversity of biomass degradation functions and organisms. PLoS One 8:e61928 . doi: 10.1371/journal.pone.0061928

Masuch T, Kusnezowa A, Nilewski S, Bautista JT, Kourist R, Leichert LI (2015) A combined bioinformatics and functional metagenomics approach to discovering lipolytic biocatalysts. Front Microbiol 6:1110 . doi: 10.3389/fmicb.2015.01110

Mayumi D, Akutsu-Shigeno Y, Uchiyama H, Nomura N, Nakajima-Kambe T (2008) Identification and characterization of novel poly(dl-lactic acid) depolymerases from metagenome. Appl Microbiol Biotechnol 79:743–750 . doi: 10.1007/s00253-008-1477-3

Menzel P, Ng KL, Krogh A (2016) Fast and sensitive taxonomic classification for metagenomics with Kaiju. Nat Commun 7:11257 . doi: 10.1038/ncomms11257

Montella S, Ventorino V, Lombard V, Henrissat B, Pepe O, Faraco V (2017) Discovery of genes coding for carbohydrate-active enzyme by metagenomic analysis of lignocellulosic biomasses. Sci Rep 7:42623 . doi: 10.1038/srep42623

Moore AM, Munck C;, Sommer MOA, Dantas G (2011) Functional metagenomic investigations of the human intestinal microbiota. Front Microbiol 2:188 . doi: 10.3389/fmicb.2011.00188

Nacke H (2011) Identification and characterization of microbial key functions in soils of the German Biodiversity Exploratories representing different land use and management types

Nacke H, Will C, Herzog S, Nowka B, Engelhaupt M, Daniel R (2011) Identification of novel lipolytic genes and gene families by screening of metagenomic libraries derived from soil samples of the German Biodiversity Exploratories. FEMS Microbiol Ecol 78:188–201 . doi: 10.1111/j.1574-6941.2011.01088.x

Namiki T, Hachiya T, Tanaka H, Sakakibara Y (2012) MetaVelvet: An extension of Velvet assembler to de novo metagenome assembly from short sequence reads. Nucleic Acids Res 40:e155 . doi: 10.1093/nar/gks678

Neher DA, Weicht TR, Bates ST, Leff JW, Fierer N (2013) Changes in bacterial and fungal communities across compost recipes, preparation methods, and composting times. PLoS One 8:e79512 . doi: 10.1371/journal.pone.0079512

Ngara TR, Zhang H (2018) Recent advances in function-based metagenomic screening. Genomics, Proteomics Bioinforma. 16:405–415

Niu Q, Qiao W, Qiang H, Li YY (2013) Microbial community shifts and biogas conversion computation during steady, inhibited and recovered stages of thermophilic methane fermentation on chicken manure with a wide variation of ammonia. Bioresour Technol 146:223–233 . doi: 10.1016/j.biortech.2013.07.038

Okano H, Hong X, Kanaya E, Angkawidjaja C, Kanaya S (2015) Structural and biochemical characterization of a metagenome-derived esterase with a long N-terminal extension. Protein Sci 24:93–104 . doi: 10.1002/pro.2591

Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, Mcglinn D, Minchin PR, O'hara RB, Simpson GL, Solymos P, Henry M, Stevens H, Szoecs E, Maintainer HW (2018) vegan: Community Ecology Package. https://cran.r-project.org/web/packages/vegan/vegan.pdf. Accessed 30 Oct 2018

Ouyang LM, Liu JY, Qiao M, Xu JH (2013) Isolation and biochemical characterization of two novel metagenome-derived esterases. Appl Biochem Biotechnol 169:15–28 . doi: 10.1007/s12010-012-9949-4

Pehrsson EC, Tsukayama P, Patel S, Mejía-Bautista M, Sosa-Soto G, Navarrete KM, Calderon M, Cabrera L, Hoyos-Arango W, Bertoli MT, Berg DE, Gilman RH, Dantas G (2016) Interconnected microbiomes and resistomes in low-income human habitats. Nature 533:212–216 . doi: 10.1038/nature17672

Peng Q, Zhang X, Shang M, Wang X, Wang G, Li B, Guan G, Li Y, Wang Y (2011) A novel esterase gene cloned from a metagenomic library from neritic sediments of the South China Sea. Microb Cell Fact 10:95 . doi: 10.1186/1475-2859-10-95

Popovic A, Hai T, Tchigvintsev A, Hajighasemi M, Nocek B, Khusnutdinova AN, Brown G, Glinos J, Flick R, Skarina T, Chernikova TN, Yim V, Brüls T, Paslier D Le, Yakimov MM, Joachimiak A, Ferrer M, Golyshina O V., Savchenko A, Golyshin PN, Yakunin AF (2017) Activity screening of environmental metagenomic libraries reveals novel carboxylesterase families. Sci Rep 7:44103 . doi: 10.1038/srep44103

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO (2013) The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. Nucleic Acids Res 41:D590–D596 . doi: 10.1093/nar/gks1219

Quince C, Walker AW, Simpson JT, Loman NJ, Segata N (2017) Shotgun metagenomics, from sampling to analysis. Nat Biotechnol 35:833–844 . doi: 10.1038/nbt.3935

Rabausch U, Juergensen J, Ilmberger N, Böhnke S, Fischer S, Schubach B, Schulte M, Streit W V. (2013) Functional screening of metagenome and genome libraries for detection of novel flavonoid-modifying enzymes. Appl Environ Microbiol 79:4551–4563 . doi: 10.1128/AEM.01077-13

Rahman MA, Culsum U, Tang W, Zhang SW, Wu G, Liu Z (2016) Characterization of a novel cold active and salt tolerant esterase from *Zunongwangia profunda*. Enzyme Microb Technol 85:1–11 . doi: 10.1016/j.enzmictec.2015.12.013

Ramnath L, Sithole B, Govinden R (2016) Classification of lipolytic enzymes and their biotechnological applications in the pulping industry. Can J Microbiol 63:179–192 . doi: 10.1139/cjm-2016-0447

Ranjan R, Yadav MK, Suneja G, Sharma R (2018) Discovery of a diverse set of esterases from hot spring microbial mat and sea sediment metagenomes. Int J Biol Macromol 119:572–581 . doi: 10.1016/j.ijbiomac.2018.07.170

Rashamuse K, Magomani V, Ronneburg T, Brady D (2009) A novel family VIII carboxylesterase derived from a leachate metagenome library exhibits promiscuous β-lactamase activity on nitrocefin. Appl Microbiol Biotechnol 83:491–500 . doi: 10.1007/s00253-009-1895-x

Reyes-Duarte D, Ferrer M, García-Arellano H (2012) Functional-based screening methods for lipases, esterases, and phospholipases in metagenomic libraries. Methods Mol Biol 861:101–113 . doi: 10.1007/978-1-61779-600-5_6

Reyes A, Alves JMP, Durham AM, Gruber A (2017) Use of profile hidden Markov models in viral discovery: current insights. Adv Genomics Genet 7:29–45 . doi: 10.2147/AGG.S136574

Rime T, Hartmann M, Frey B (2016) Potential sources of microbial colonizers in an initial soil ecosystem after retreat of an alpine glacier. ISME J 10:1625–1641 . doi: 10.1038/ismej.2015.238

Rinninella E, Raoul P, Cintoni M, Franceschi F, Miggiano GAD, Gasbarrini A, Mele MC (2019) What is the healthy gut microbiota composition? A changing ecosystem across age, environment, diet, and diseases. Microorganisms 7:14 . doi: 10.3390/microorganisms7010014

Rudek W, Haque RU (1976) Extracellular enzymes of the genus Bacteroides. J Clin Microbiol 4:458–60

Ryckeboer J, Mergaert J, Vaes K, Klammer S, Clercq D, Coosemans J, Insam H, Swings J (2003) A survey of bacteria and fungi occurring during composting and self-heating processes. Ann Microbiol 53:349–410

Sarmah N, Revathi D, Sheelu G, Yamuna Rani K, Sridhar S, Mehtab V, Sumana C (2018) Recent advances on sources and industrial applications of lipases. Biotechnol Prog 34:5–28 . doi: 10.1002/btpr.2581

Schloss PD, Hay AG, Wilson DB, Walker LP (2003) Tracking temporal changes of bacterial community fingerprints during the initial stages of composting. FEMS Microbiol Ecol 46:1–9 . doi: 10.1016/S0168-6496(03)00153-3

Schneider D, Engelhaupt M, Allen K, Kurniawan S, Krashevska V, Heinemann M, Nacke H, Wijayanti M, Meryandini A, Corre MD, Scheu S, Daniel R (2015) Impact of lowland rainforest transformation on diversity and composition of soil prokaryotic communities in sumatra (Indonesia). Front Microbiol 6:1339 . doi: 10.3389/fmicb.2015.01339

Seemann T (2014) Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069 . doi: 10.1093/bioinformatics/btu153

Seo S, Lee YS, Yoon SH, Kim SJ, Cho JY, Hahn BS, Koo BS, Lee CM (2014) Characterization of a novel cold-active esterase isolated from swamp sediment metagenome. World J Microbiol Biotechnol 30:879–886 . doi: 10.1007/s11274-013-1496-9

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: A software environment for integrated models of biomolecular interaction networks. Genome Res 13:2498–2504 . doi: 10.1101/gr.1239303

Shao H, Xu L, Yan Y (2013) Isolation and characterization of a thermostable esterase from a metagenomic library. J Ind Microbiol Biotechnol 40:1211–1222 . doi: 10.1007/s10295-013-1317-z

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol 7:539 . doi: 10.1038/msb.2011.75

Silva CC, Hayden H, Sawbridge T, Mele P, Kruger RH, Rodrigues MVN, Costa GGL, Vidal RO, Sousa MP, Torres APR, Santiago VMJ, Oliveira VM (2012) Phylogenetic and functional diversity of metagenomic libraries of phenol degrading sludge from petroleum refinery wastewater treatment system. AMB Express 2:18 . doi: 10.1186/2191-0855-2-18

Simon C, Daniel R (2009) Achievements and new knowledge unraveled by metagenomic approaches. Appl Microbiol Biotechnol 85:265–276 . doi: 10.1007/s00253-009-2233-z

Simon C, Daniel R (2011) Metagenomic analyses: Past and future trends. Appl Environ Microbiol 77:1153–1161 . doi: 10.1128/AEM.02345-10

Skewes-Cox P, Sharpton TJ, Pollard KS, DeRisi JL (2014) Profile Hidden Markov Models for the detection of viruses within metagenomic sequence data. PLoS One 9:e105067 . doi: 10.1371/journal.pone.0105067

Snellman EA, Colwell RR (2004) Acinetobacter lipases: Molecular biology, biochemical properties and biotechnological potential. J. Ind. Microbiol. Biotechnol. 31:391–400

Sørensen KB, Canfield DE, Teske AP, Oren A (2005) Community composition of a hypersaline endoevaporitic microbial mat. Appl Environ Microbiol 71:7352–7365 . doi: 10.1128/AEM.71.11.7352-7365.2005

Sriyapai P, Kawai F, Siripoke S, Chansiri K, Sriyapai T (2015) Cloning, expression and characterization of a thermostable esterase HydS14 from *Actinomadura* sp. strain S14 in pichia pastoris. Int J Mol Sci 16:13579–13594 . doi: 10.3390/ijms160613579

Staden R, Judge DP, Bonfield JK (2003) Managing sequencing projects in the GAP4 environment. In: Introduction to Bioinformatics. Humana Press, Totowa, NJ, pp 327–344

Streit W, Chow J, Jaeger K-E, Katzke N, Koch R, Borchert M, Bj'rnvad M, Roggenbuck M, Lippold F, Golyshin P, Golyshina O, Tran H, Ferrer M, Daniel R, Hollensteiner J (2018) A metagenomic collection of novel and highly efficient biocatalysts for industrial biotechnology. Impact 2018:82–84 . doi: 10.21820/23987073.2018.5.82

Stroobants A, Martin R, Roosens L, Portetelle D, Vandenbol M (2015) New lipolytic enzymes identified by screening two metagenomic libraries derived from the soil of a winter wheat field. Biotechnolgie, Agron Société Environ 19:125–131

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 . doi: 10.1093/molbev/mst197

Tasse L, Bercovici J, Pizzut-Serin S, Robe P, Tap J, Klopp C, Cantarel BL, Coutinho PM, Henrissat B, Leclerc M, Doré J, Monsan P, Remaud-Simeon M, Potocki-Veronese G (2010) Functional metagenomics to mine the human gut microbiome for dietary fiber catabolic enzymes. Genome Res 20:1605–1612 . doi: 10.1101/gr.108332.110

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–80

Turnbaugh PJ, Hamady M, Yatsunenko T, Cantarel BL, Duncan A, Ley RE, Sogin ML, Jones WJ, Roe BA, Affourtit JP, Egholm M, Henrissat B, Heath AC, Knight R, Gordon JI (2009) A core gut microbiome in obese and lean twins. Nature 457:480–484 . doi: 10.1038/nature07540

Verma S, Kumar R, Meghwanshi GK (2019) Identification of new members of alkaliphilic lipases in archaea and metagenome database using reconstruction of ancestral sequences. 3 Biotech 9:1–8 . doi: 10.1007/s13205-019-1693-9

Vigneron A, Alsop EB, Lomans BP, Kyrpides NC, Head IM, Tsesmetzis N (2017) Succession in the petroleum reservoir microbiome through an oil field production lifecycle. ISME J 11:2141–2154 . doi: 10.1038/ismej.2017.78

Villamizar GAC, Nacke H, Daniel R (2017) Function-based metagenomic library screening and heterologous expression strategy for genes encoding phosphatase activity. In: Methods in Molecular Biology. Humana Press Inc., pp 249–260

Vital M, Howe AC, Tiedje JM (2014) Revealing the bacterial butyrate synthesis pathways by analyzing (meta)genomic data. MBio 5:e00889-14. . doi: 10.1128/mBio.00889-14

Walsh CJ, Guinane CM, O' Toole PW, Cotter PD (2017) A profile Hidden Markov Model to investigate the distribution and frequency of LanB-encoding lantibiotic modification genes in the human oral and gut microbiome. PeerJ 5:e3254 . doi: 10.7717/peerj.3254

Wang C, Dong D, Wang H, Müller K, Qin Y, Wang H, Wu W (2016) Metagenomic analysis of microbial consortia enriched from compost : new insights into the role of *Actinobacteria* in lignocellulose decomposition. Biotechnol Biofuels 9:22 . doi: 10.1186/s13068-016-0440-2

Wang L, Zhang G, Xu H, Xin H, Zhang Y (2019) Metagenomic analyses of microbial and carbohydrate-active enzymes in the rumen of holstein cows fed different forage-to-concentrate ratios. Front Microbiol 10:649 . doi: 10.3389/fmicb.2019.00649

Wang S, Wang K, Li L, Liu Y (2013) Isolation and characterization of a novel organic solvent-tolerant and halotolerant esterase from a soil metagenomic library. J Mol Catal B Enzym 95:1–8 . doi: 10.1016/j.molcatb.2013.05.015

Wang W-L, Xu S-Y, Ren Z-G, Tao L, Jiang J-W, Zheng S-S (2015) Application of metagenomics in the human gut microbiome. World J Gastroenterol 21:803–814 . doi: 10.3748/wjg.v21.i3.803

Wang Y, Qian P-Y (2009) Conservative fragments in bacterial 16S rRNA genes and primer design for 16S ribosomal DNA amplicons in metagenomic studies. PLoS One 4:e7401 . doi: 10.1371/journal.pone.0007401

Wei T, Feng S, Shen Y, He P, Ma G, Yu X, Zhang F, Mao D (2013) Characterization of a novel thermophilic pyrethroid-hydrolyzing carboxylesterase from *Sulfolobus tokodaii* into a new family. J Mol Catal B Enzym 97:225–232 . doi: 10.1016/j.molcatb.2013.07.022

Wei Y, Swenson L, Castro C, Derewenda U, Minor W, Arai H, Aoki J, Inoue K, Servin-Gonzalez L, Derewenda ZS (1998) Structure of a microbial homologue of mammalian platelet-activating factor acetylhydrolases: Streptomyces exfoliatus lipase at 1.9 Å resolution. Structure 6:511–519 . doi: 10.1016/S0969-2126(98)00052-5

Wemheuer F, Kaiser K, Karlovsky P, Daniel R, Vidal S, Wemheuer B (2017) Bacterial endophyte communities of three agricultural important grass species differ in their response towards management regimes. Sci Rep 7:40914 . doi: 10.1038/srep40914

Willmann M, El-Hadidi M, Huson DH, Schütz M, Weidenmaier C, Autenrieth IB, Peter S (2015) Antibiotic selection pressure determination through sequence-based metagenomics. Antimicrob Agents Chemother 59:7335–7345 . doi: 10.1128/AAC.01504-15

Wu C, Sun B (2009) Identification of novel esterase from metagenomic library of Yangtze River. J Microbiol Biotechnol 19:187–193 . doi: 10.4014/jmb.0804.292

Wu G, Zhang S, Zhang H, Zhang S, Liu Z (2013) A novel esterase from a psychrotrophic bacterium *Psychrobacter celer* 3Pb1 showed cold-adaptation and salt-tolerance. J Mol Catal B Enzym 98:119–126 . doi: 10.1016/j.molcatb.2013.10.012

Yan W, Li F, Wang L, Zhu Y, Dong Z, Bai L (2017) Discovery and characterizaton of a novel lipase with transesterification activity from hot spring metagenomic library. Biotechnol Reports 14:27–33 . doi: 10.1016/j.btre.2016.12.007

Yu H, Zeng G, Huang H, Xi X, Wang R, Huang D, Huang G, Li J (2007) Microbial community succession and lignocellulose degradation during agricultural waste composting. Biodegradation 18:793–802 . doi: 10.1007/s10532-007-9108-8

Yu S, Yu S, Han W, Wang H, Zheng B, Feng Y (2010) A novel thermophilic lipase from *Fervidobacterium nodosum* Rt17-B1 representing a new subfamily of bacterial lipases. J Mol Catal B Enzym 66:81–89 . doi: 10.1016/j.molcatb.2010.03.007

Yu Z, Tang J, Liao H, Liu X, Zhou P, Chen Z, Rensing C, Zhou S (2018) The distinctive microbial community improves composting efficiency in a full-scale hyperthermophilic composting plant. Bioresour Technol 265:146–154 . doi: 10.1016/j.biortech.2018.06.011

Zarafeta D, Moschidi D, Ladoukakis E, Gavrilov S, Chrysina ED, Chatziioannou A, Kublanov I, Skretas G, Kolisis FN (2016) Metagenomic mining for thermostable esterolytic enzymes uncovers a new family of bacterial esterases. Sci Rep 6:38886 . doi: 10.1038/srep38886

Zhang X, Zhong Y, Yang S, Zhang W, Xu M, Ma A, Zhuang G, Chen G, Liu W (2014) Diversity and dynamics of the microbial community on decomposing wheat straw during mushroom compost production. Bioresour Technol 170:183–195 . doi: 10.1016/j.biortech.2014.07.093

Zhang Y, Hao J, Zhang YQ, Chen XL, Xie B Bin, Shi M, Zhou BC, Zhang YZ, Li PY (2017) Identification and characterization of a novel salt-tolerant esterase from the deep-sea sediment of the South China Sea. Front Microbiol 8:441 . doi: 10.3389/fmicb.2017.00441

Zhang Y, Pengjun S, Wanli L, Kun M, Yingguo B, Guozeng W, Zhichun Z, Bin Y (2009) Lipase diversity in glacier soil based on analysis of metagenomic DNA fragments and cell culture. J Microbiol Biotechnol 19:888–897 . doi: 10.4014/jmb.0812.695

Zhou G, Xu X, Qiu X, Zhang J (2019) Biochar influences the succession of microbial communities and the metabolic functions during rice straw composting with pig manure. Bioresour Technol 272:10–18 . doi: 10.1016/j.biortech.2018.09.135

Zhou H, Gu W, Sun W, Hay AG (2018) A microbial community snapshot of windrows from a commercial composting facility. Appl Microbiol Biotechnol 102:8069–8077 . doi: 10.1007/s00253-018-9201-4

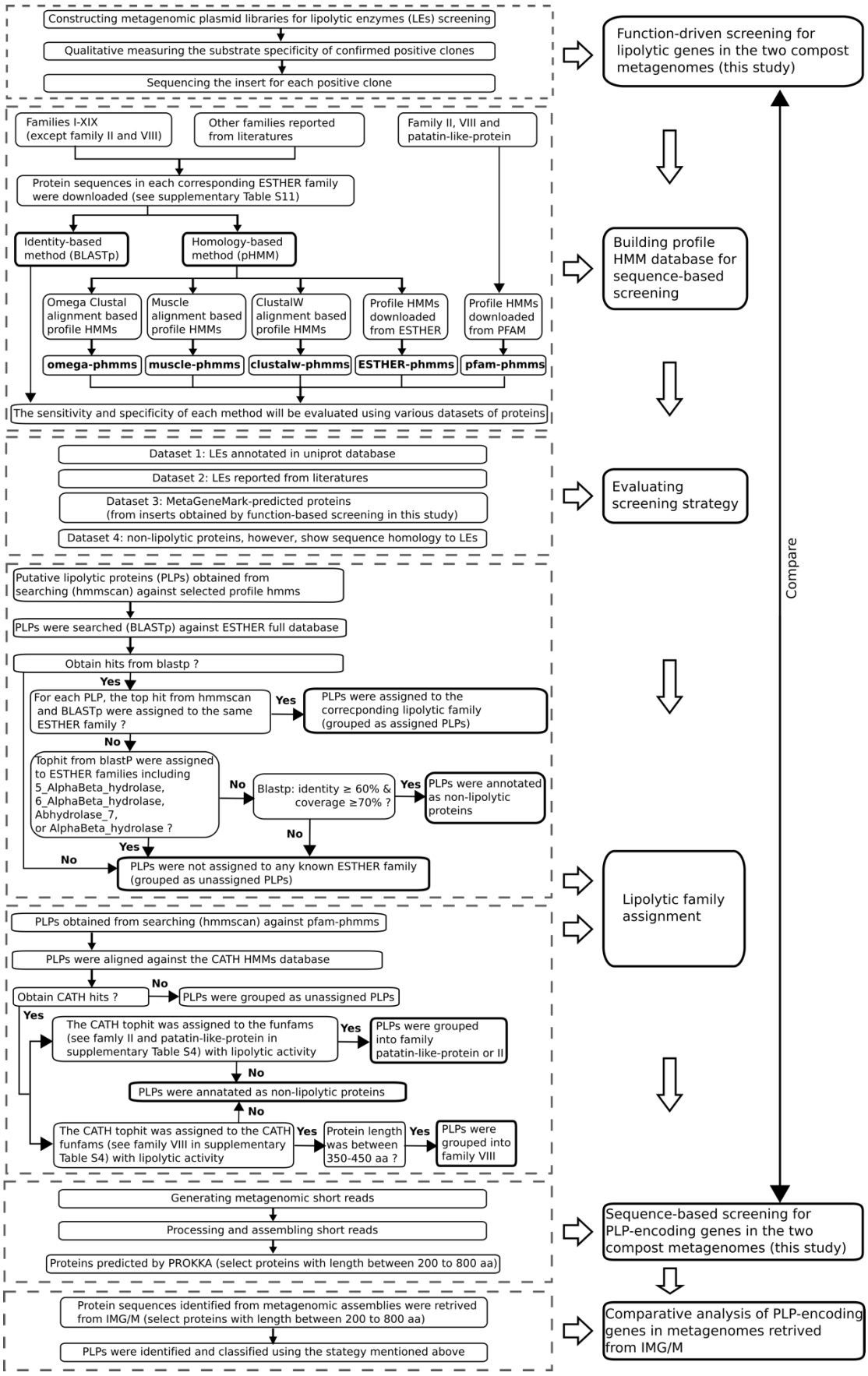
Zhou J, Bruns MA, Tiedje JM (1996) DNA recovery from soils of diverse composition. Appl Environ Microbiol 62:316–322

Zhu W, Lomsadze A, Borodovsky M (2010) Ab initio gene identification in metagenomic sequences. Nucleic Acids Res 38:e132 . doi: 10.1093/nar/gkq275

Supplemental information for chapter II

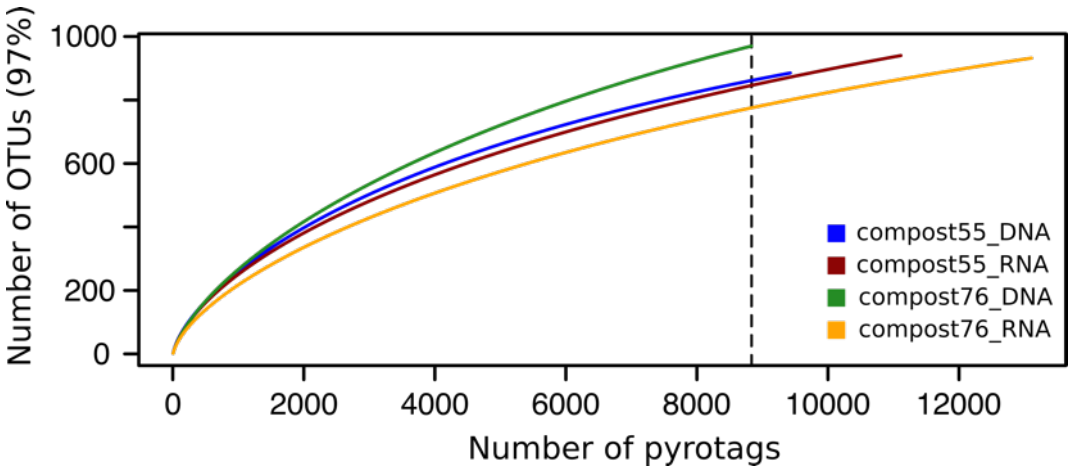
Content

Supplementary Figure S16. Phylogenetic origin of the total PLPs (assigned and unassigned PLPs combined) at (A) domain and (B) phylum level.
Supplementary Figure S17. Heat map of the taxonomic origins at genus level of total PLPs across samples.
Supplementary Figure S18. Analysis of the phylogenetic profile at genus level of total PLPs across samples.
Supplementary Figure S19. Phylogenetic distribution of the total PLPs at phylum level of each habitat.

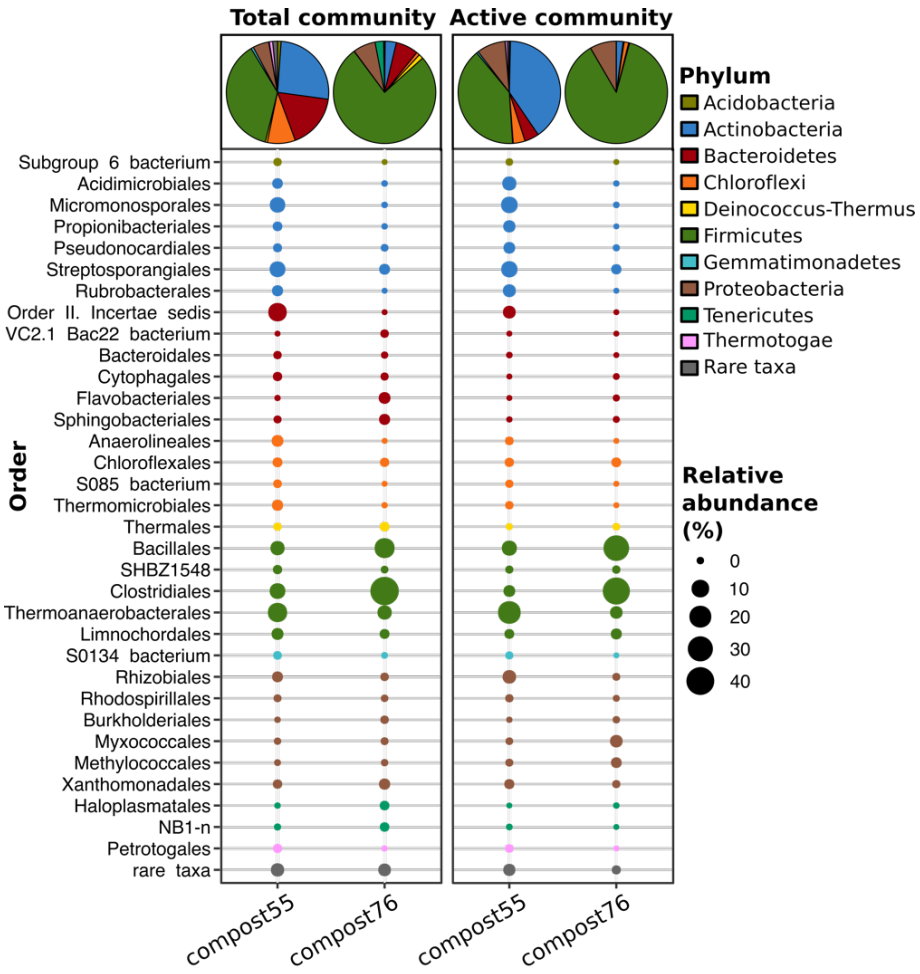

Constructing metagenomic plasmid libraries for lipolytic enzymes (LEs) screening
Qualitative measuring the substrate specificity of confirmed positive clones
Sequencing the insert for each positive clone
Function-driven screening for lipolytic genes in the two compost metagenomes (this study)
Families I-XIX (except family II and VIII)
Other families reported from literatures
Family II, VIII and patatin-like-protein
Protein sequences in each corresponding ESTHER family were downloaded (see supplementary Table S11)
Identity-based method (BLASTp)
Homology-based method (pHMM)
Omega Clustal alignment based profile HMMs
Muscle alignment based profile HMMs
ClustalW alignment based profile HMMs
Profile HMMs downloaded from ESTHER
Profile HMMs downloaded from PFAM
omega-phmms
muscle-phmms
clustalw-phmms
ESTHER-phmms
pfam-phmms
The sensitivity and specificity of each method will be evaluated using various datasets of proteins
Building profile HMM database for sequence-based screening
Dataset 1: LEs annotated in uniprot database
Dataset 2: LEs reported from literatures
Dataset 3: MetaGeneMark-predicted proteins (from inserts obtained by function-based screening in this study)
Dataset 4: non-lipolytic proteins, however, show sequence homology to LEs
Evaluating screening strategy
Putative lipolytic proteins (PLPs) obtained from searching (hmmscan) against selected profile hmms
PLPs were searched (BLASTp) against ESTHER full database
Obtain hits from blastp ?
Yes
For each PLP, the top hit from hmmscan and BLASTp were assigned to the same ESTHER family ?
Yes
PLPs were assigned to the correcponding lipolytic family (grouped as assigned PLPs)
No
Tophit from blastP were assigned to ESTHER families including 5_AlphaBeta_hydrolase, 6_AlphaBeta_hydrolase, Abhydrolase_7, or AlphaBeta_hydrolase ?
No
Blastp: identity ≥ 60% & coverage ≥70% ?
Yes
PLPs were annotated as non-lipolytic proteins
No
Yes
No
PLPs were not assigned to any known ESTHER family (grouped as unassigned PLPs)
Lipolytic family assignment
PLPs obtained from searching (hmmscan) against pfam-phmms
PLPs were aligned against the CATH HMMs database
Obtain CATH hits ?
No
PLPs were grouped as unassigned PLPs
Yes
The CATH tophit was assigned to the funfams (see famly II and patatin-like-protein in supplementary Table S4) with lipolytic activity
Yes
PLPs were grouped into family patatin-like-protein or II
No
PLPs were annatated as non-lipolytic proteins
No
The CATH tophit was assigned to the CATH funfams (see family VIII in supplementary Table S4) with lipolytic activity
Yes
Protein length was between 350-450 aa ?
Yes
PLPs were grouped into family VIII
Compare
Generating metagenomic short reads
Processing and assembling short reads
Proteins predicted by PROKKA (select proteins with length between 200 to 800 aa)
Sequence-based screening for PLP-encoding genes in the two compost metagenomes (this study)
Protein sequences identified from metagenomic assemblies were retrived from IMG/M (select proteins with length between 200 to 800 aa)
PLPs were identified and classified using the stategy mentioned above
Comparative analysis of PLP-encoding genes in metagenomes retrived from IMG/M

Supplementary Figure S1. Overall workflow for identification of lipolytic enzymes (LEs) through function-driven and sequence-based approaches in this study.

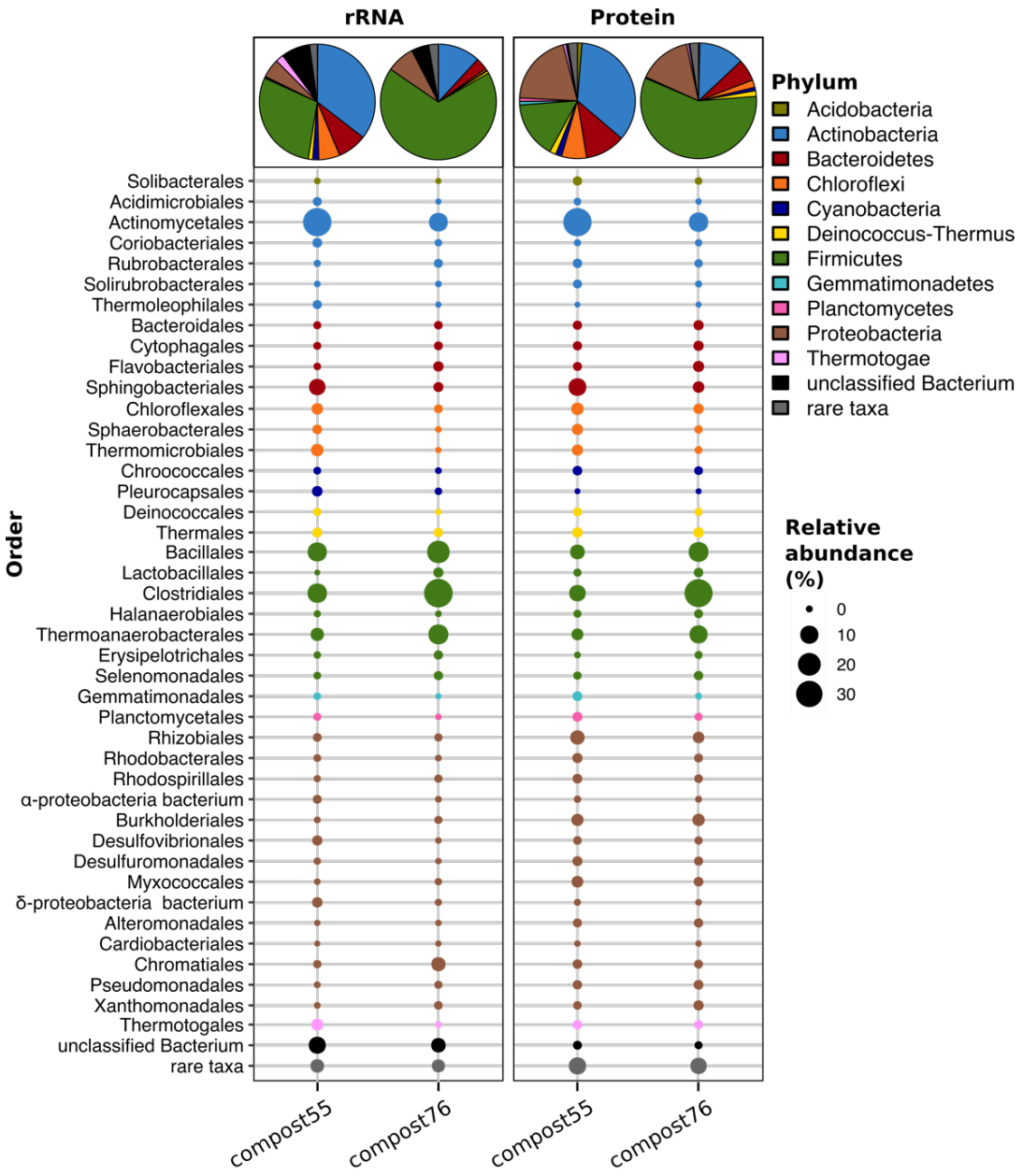
Firstly, LEs were identified by function-based screening of constructed metagenomic libraries. Positive clones were collected and inserts harboring lipolytic genes were sequenced. Lipolytic genes were subsequently revealed and classified. As for sequence-based screening, a search method based on the profile Hidden Markov Models (HMMs) was developed to identify and annotate the putative lipolytic proteins (PLPs) in assembled metagenomes. LEs can be generally divided into two major groups: α/β hydrolase or not α/β hydrolase. For LEs belong to the α/β hydrolase superfamily, four LE-specific profile HMM databases were retrieved (omega-phmms, muscle-phmms, clustalw-phmms and ESTHER-phmms). For LEs that are not α/β hydrolases, profile HMMs were retrieved from the pfam database (pfam-phmms). The prediction sensitivity and specificity of each profile HMM database were evaluated using four datasets, and the best one was selected for subsequent analysis. The lipolytic family assignment of PLPs obtained by screening against the selected profile HMM database (one of omega-phmms, muscle-phmms, clustalw-phmms and ESTHER-phmms) were generally conducted by combining the annotations from *hmmscan* against the profile HMM database and *blastp* against the full ESTHER database. For PLPs obtained by screening against pfam-phmms, the annotation was performed by the subsequent screening against the CATH HMMs database. Based on the strategies for sequence-based screening and lipolytic family assignment, PLPs in the two compost assembled metagenomes were identified and annotated. The results from function-driven and sequence-based screening were also compared. Finally, assembled metagenomes in various habitats were retrieved from the IMG/M database, and PLPs were identified for comparative analysis.



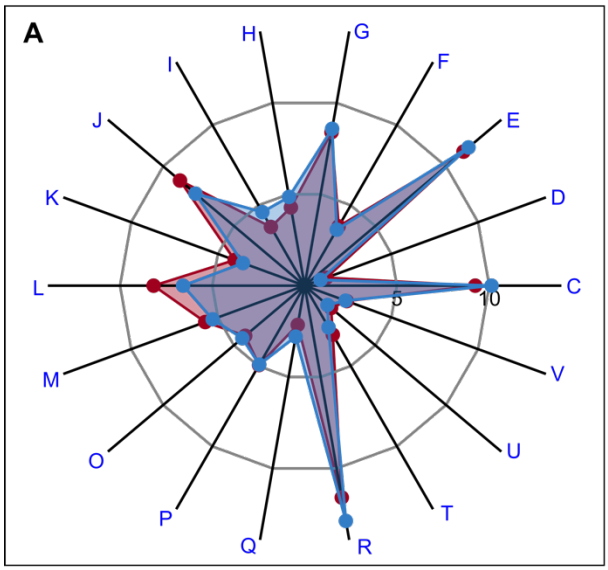
Supplementary Figure S2. Rarefaction curves of subsampled OTUs for 16S rRNA genes (DNA level) and transcripts (RNA level) in compost55 and compost76 at 97% similarity.



Supplementary Figure S3. Phylogenetic composition of bacterial communities in compost55 and compost76, revealed from 16S rRNA genes (DNA-level) and transcripts (RNA-level). Taxonomic specificity ranges from phylum level (pie chart) to order level (point chart) resolution when applicable. Taxonomic classification of DNA- or RNA-derived 16S rRNA gene sequences was performed according to SILVA SSU database 128 (Quast et al. 2013). Low relative abundant groups (<1% at phylum level or <0.5% at order level) were summarized as artificial group “rare taxa”.

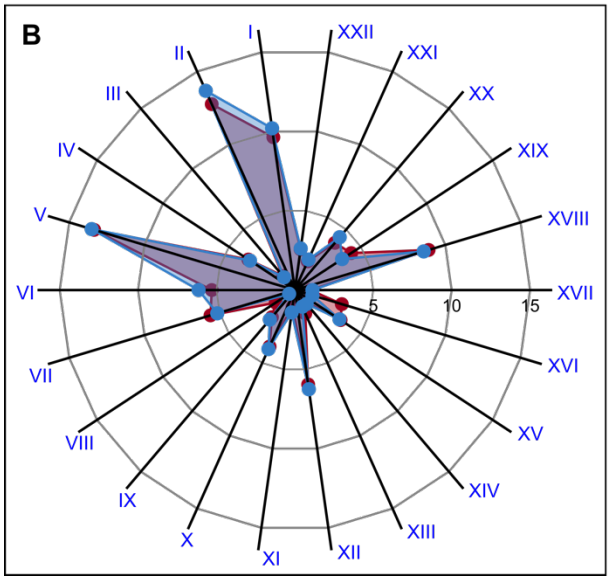


Supplementary Figure S4. Phylogenetic composition of bacterial communities in compost55 and compost76 as annotated by the MG-RAST platform (Keegan et al. 2016). Taxonomic specificity ranges from phylum level (pie chart) to order level (point chart) resolution when applicable. Microbial composition annotation was performed using MG-RAST best hit classification tool against the databases of M5RNA (Non-redundant multisource ribosomal RNA annotation) and M5NR (M5 non-redundant protein) available within MG-RAST with default settings. Low relative abundant groups (<1% at phylum level, <0.5% at order level) were summarized as artificial group "rare taxa".



COG catagory

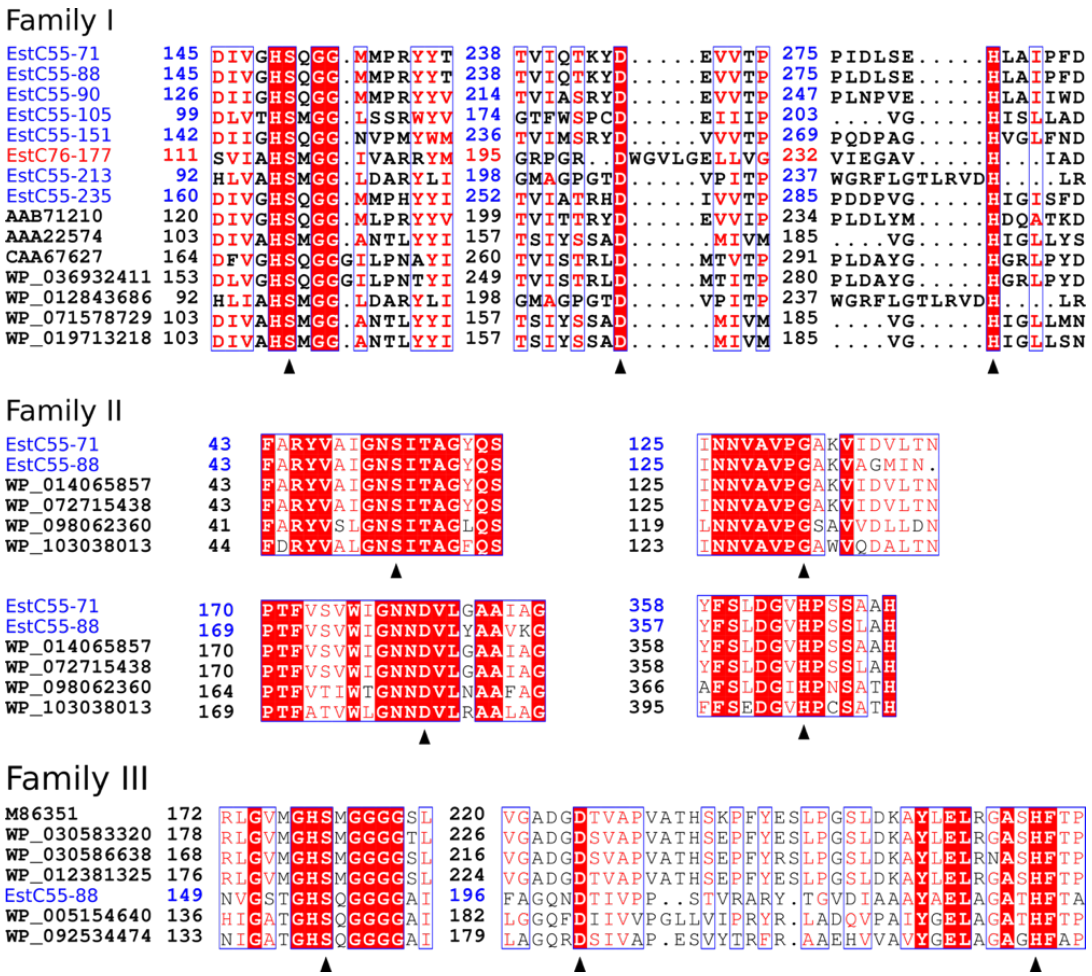
C-Energy production and conversion
D-Cell cycle control, cell division, chromosome partitioning
E-Amino acid transport and metabolism
F-Nucleotide transport and metabolism
G-Carbohydrate transport and metabolism
H-Coenzyme transport and metabolism
I-Lipid transport and metabolism
J-Translation, ribosomal structure and biogenesis
K-Transcription
L-Replication, recombination and repair
M-Cell wall/membrane/envelope biogenesis
O-Posttranslational modification, protein turnover, chaperones
P-Inorganic ion transport and metabolism
Q-Secondary metabolites biosynthesis, transport and catabolism
R-General function prediction only
T-Signal transduction mechanisms
U-Intracellular trafficking, secretion, and vesicular transport
V-Defense mechanisms



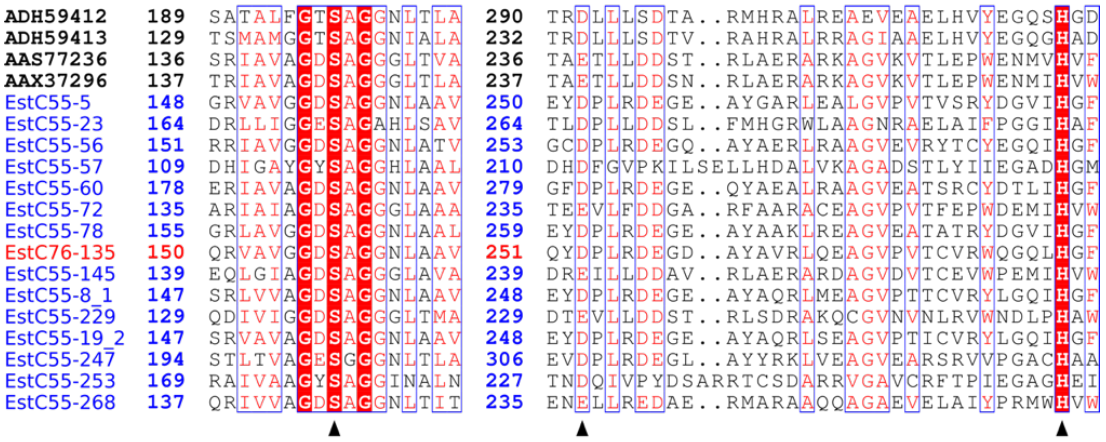
Subsystems catagory

I-Amino Acids and Derivatives
II-Carbohydrates
III-Cell Division and Cell Cycle
IV-Cell Wall and Capsule
V-Clustering-based subsystems
VI-Cofactors, Vitamins, Prosthetic Groups, Pigments
VII-DNA Metabolism
VIII-Dormancy and Sporulation
IX-Fatty Acids, Lipids, and Isoprenoids
X-Membrane Transport
XI-Metabolism of Aromatic Compounds
XII-Miscellaneous
XIII-Motility and Chemotaxis
XIV-Nitrogen Metabolism
XV-Nucleosides and Nucleotides
XVI-Phages, Prophages, Transposable elements, Plasmids
XVII-Phosphorus Metabolism
XVIII-Protein Metabolism
XIX-RNA Metabolism
XX-Respiration
XXI-Stress Response
XXII-Virulence, Disease and Defense

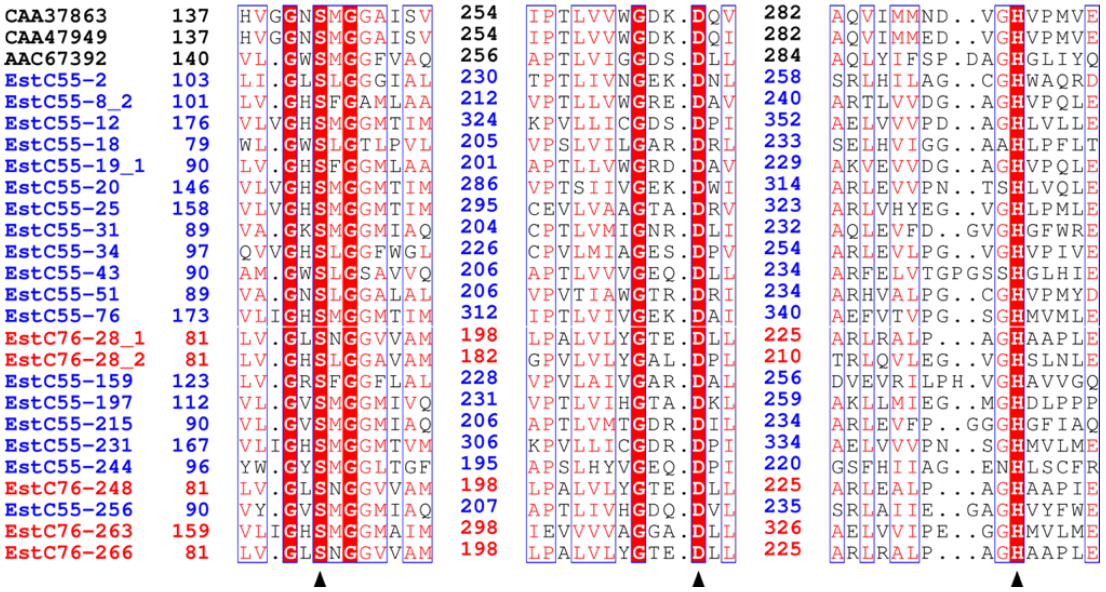
Supplementary Figure S5. Functional distribution pattern in compost55 (**blue** filled circle) and compost76 (**red** filled circle) microbial consortia. MG-RAST annotation against COG (**A**) and (**B**) subsystems database are shown. Only functional categories with relative abundance of more than 1 % were shown. The functional annotation of metagenomic reads was performed by MG-RAST pipeline (Keegan et al. 2016).

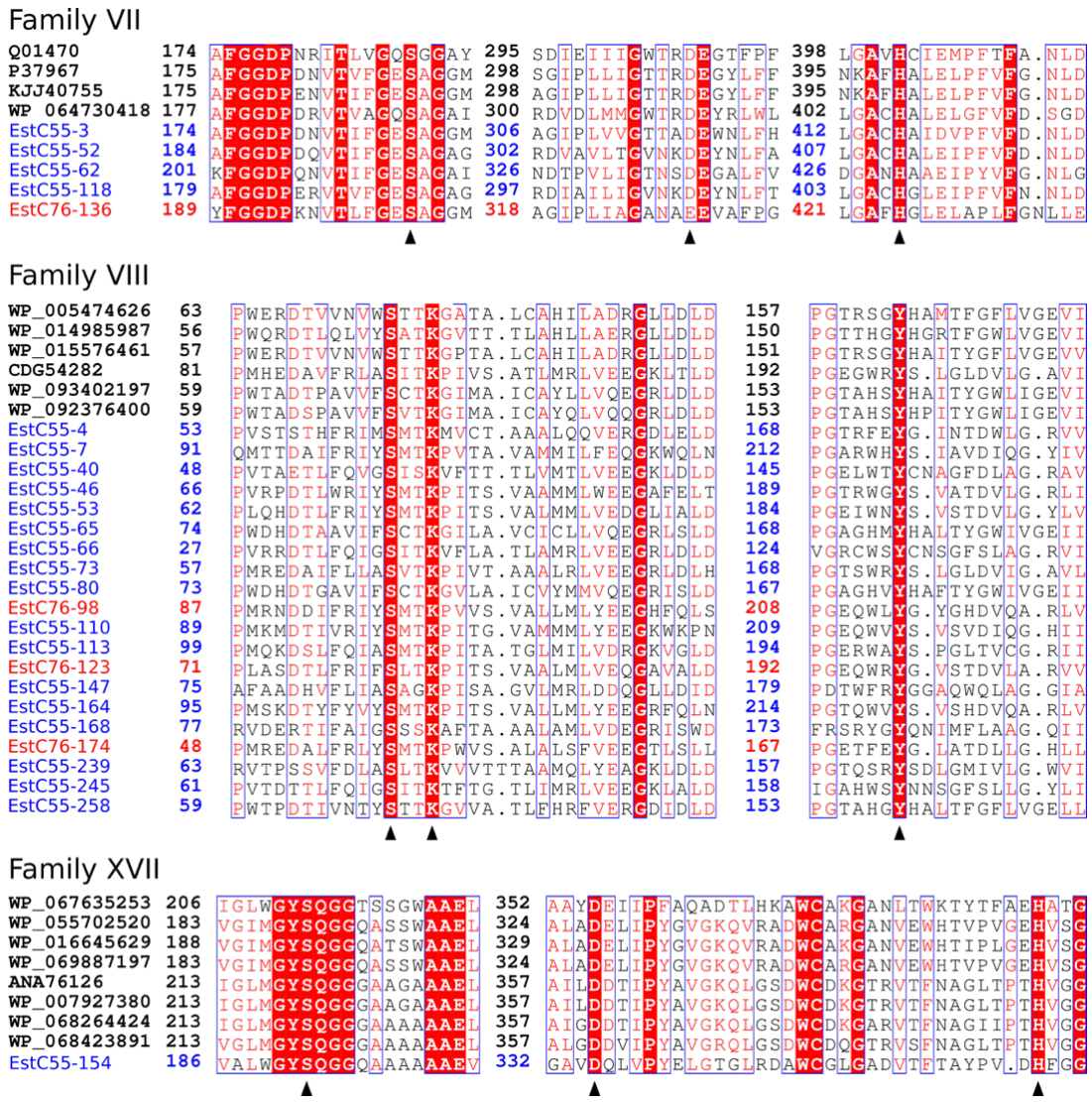

Family I
EstC55-71 145 DIVGHSQGG.MMPRYYT 238 TVIQTKYD.....EVVTP 275 PIDLSE.....HLAIPFD
EstC55-88 145 DIVGHSQGG.MMPRYYT 238 TVIQTKYD.....EVVTP 275 PLDLSE.....HLAIPFD
EstC55-90 126 DIIGHSQGG.MMPRYYV 214 TVIASRYD.....EVVTP 247 PLNPVE.....HLAIIWD
EstC55-105 99 DLVTHSMGG.LSSRWYV 174 GTFWSPCD.....EIIIP 203 ....VG.....HISLLAD
EstC55-151 142 DIIGHSQGG.NVPMYWM 236 TVIMSRYD.....VVVTP 269 PQDPAG.....HVGLFND
EstC76-177 111 SVIAHSMGG.IVARRYM 195 GRPGR..DWGVLGELLVG 232 VIEGAV.....H...IAD
EstC55-213 92 HLVAHSMGG.LDARYLI 198 GMAGPGTD.....VPITP 237 WGRFLGTLRVDH....LR
EstC55-235 160 DIVGHSQGG.MMPHYYI 252 TVIATRHD.....IVVTP 285 PDDPVG.....HIGISFD
AAB71210 120 DIVGHSQGG.MLPRYYV 199 TVITTRYD.....EVVIP 234 PLDLYM.....HDQATKD
AAA22574 103 DIVAHSMGG.ANTLYYI 157 TSIYSSAD.....MIVM 185 ....VG.....HIGLLYS
CAA67627 164 DFVGHSQGGGILPNAYI 260 TVISTRLD.....MTVTP 291 PLDAYG.....HGRLPYD
WP_036932411 153 DLVGHSQGGGILPNTYI 249 TVISTRLD.....MTITP 280 PLDAYG.....HGRLPYD
WP_012843686 92 HLIAHSMGG.LDARYLI 198 GMAGPGTD.....VPITP 237 WGRFLGTLRVDH....LR
WP_071578729 103 DIVAHSMGG.ANTLYYI 157 TSIYSSAD.....MIVM 185 ....VG.....HIGLLMN
WP_019713218 103 DIVAHSMGG.ANTLYYI 157 TSIYSSAD.....MIVM 185 ....VG.....HIGLLSN
Family II
EstC55-71 43 FARYVAIGNSITAGYQS 125 INNVAVPGAKVIDVLTN
EstC55-88 43 FARYVAIGNSITAGYQS 125 INNVAVPGAKVAGMIN.
WP_014065857 43 FARYVAIGNSITAGYQS 125 INNVAVPGAKVIDVLTN
WP_072715438 43 FARYVAIGNSITAGYQS 125 INNVAVPGAKVIDVLTN
WP_098062360 41 FARYVSLGNSITAGLQS 119 LNNVAVPGSAVVDLLDN
WP_103038013 44 FDRYVALGNSITAGFQS 123 INNVAVPGAWVQDALTN
EstC55-71 170 PTFVSVWIGNNDVLGAAIAG 358 YFSLDGVHPSSAAH
EstC55-88 169 PTFVSVWIGNNDVLYAAVKG 357 YFSLDGVHPSSLAH
WP_014065857 170 PTFVSVWIGNNDVLGAAIAG 358 YFSLDGVHPSSAAH
WP_072715438 170 PTFVSVWIGNNDVLGAAIAG 358 YFSLDGVHPSSLAH
WP_098062360 164 PTFVTIWTGNNDVLNAAFAG 366 AFSLDGIHPNSATH
WP_103038013 169 PTFATVWLGNNDVLRAALAG 395 FFSEDGVHPCSATH
Family III
M86351 172 RLGVMGHSMGGGGSL 220 VGADGDTVAPVATHSKPFYESLPGSLDKAYLELRGASHFTP
WP_030583320 178 RLGVMGHSMGGGGTL 226 VGADGDSVAPVATHSEPFYESLPGSLDKAYLELRGASHFTP
WP_030586638 168 RLGVMGHSMGGGGSL 216 VGADGDSVAPVATHSEPFYRSLPGSLDKAYLELRNASHFTP
WP_012381325 176 RLGVMGHSMGGGGSL 224 VGADGDTVAPVATHSEPFYESLPGSLDKAYLELRGASHFTP
EstC55-88 149 NVGSTGHSQGGGGAI 196 FAGQNDTIVPP..STVRARY.TGVDIAAAYAELAGATHFTA
WP_005154640 136 HIGATGHSQGGGGAI 182 LGGQFDIIVVPGLLVIPRYR.LADQVPAIYGELAGATHFTP
WP_092534474 133 NIGATGHSQGGGGAI 179 LAGQRDSIVAP.ESVYTRFR.AAEHVVAVYGELAGAGHFAP

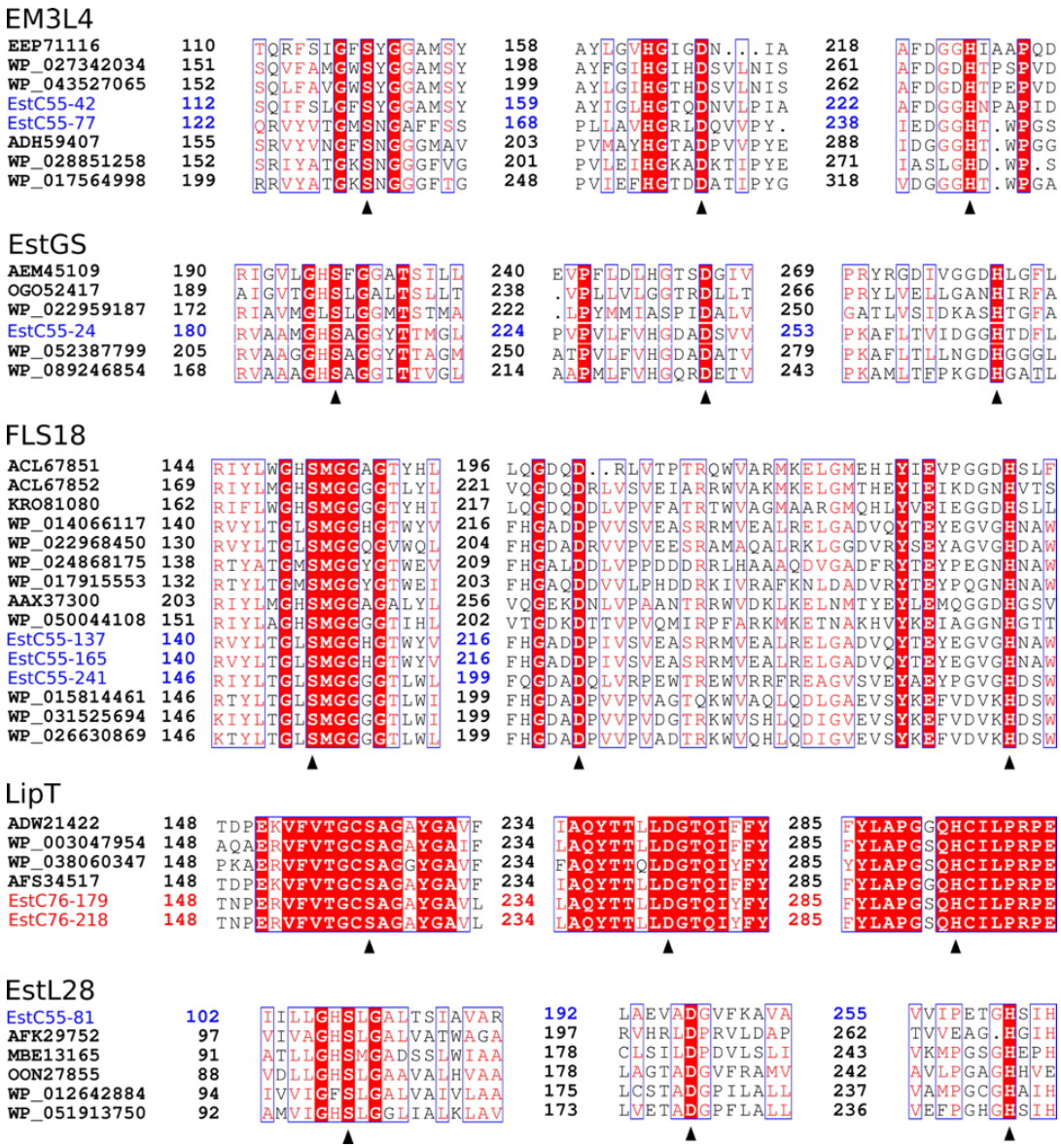
Family IV

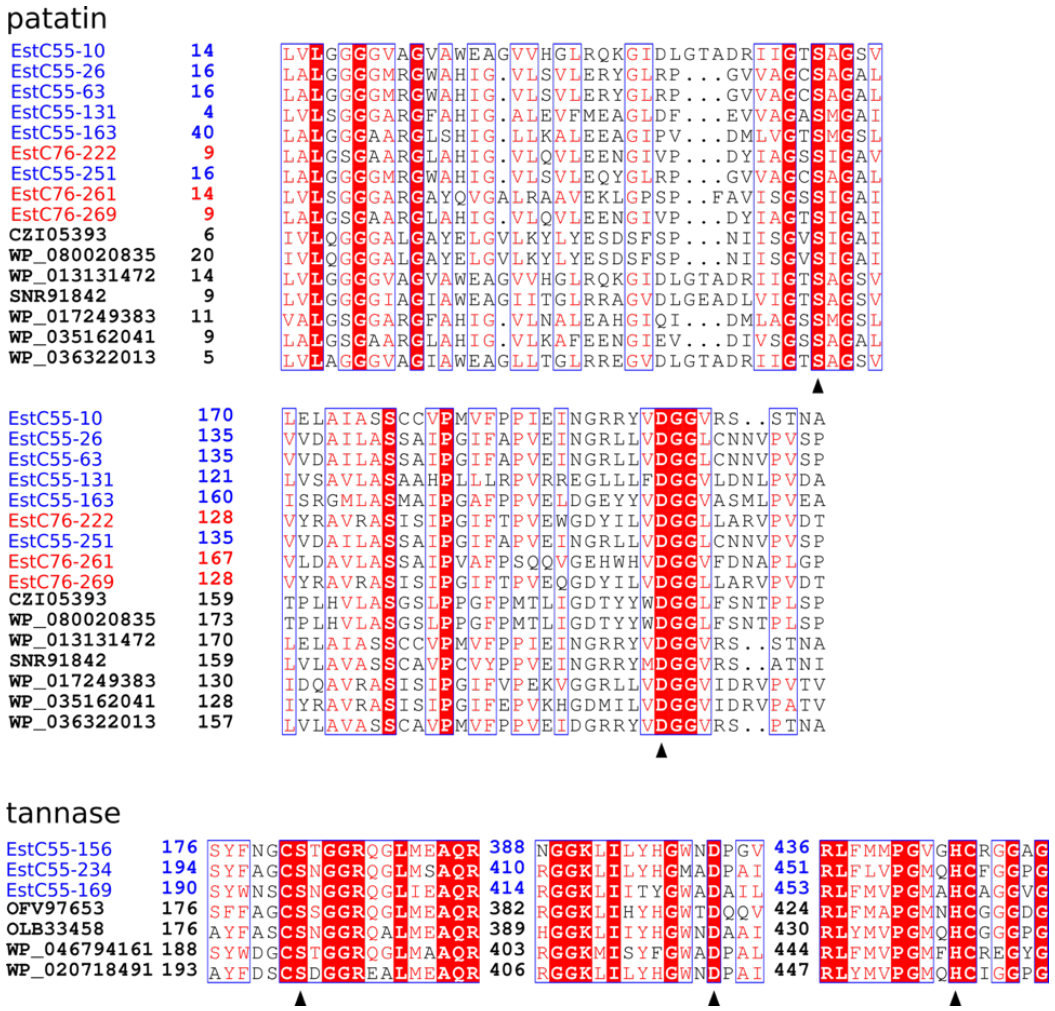


Family V




Family VII
Q01470 174 AFGGDPNRITLVGQSGGAY 295 SDIEIIIGWTRDEGTFPF 398 LGAVHCIEMPFTFA.NLD
P37967 175 AFGGDPDNVTVFGESAGGM 298 SGIPLLIGTTRDEGYLFF 395 NKAFHALELPFVFG.NLD
KJJ40755 175 AFGGDPENVTIFGESAGGM 298 AGIPLLIGTTRDEGYLFF 395 NKAFHALELPFVFG.NLD
WP_064730418 177 AFGGDPDRVTVAGQSAGAI 300 RDVDLMMGWTRDEYRLWL 402 LGACHALELGFVFD.SGD
EstC55-3 174 AFGGDPDNVTIFGESAGGM 306 AGIPLVVGTTADEWNLFH 412 LGACHAIDVPFVFD.NLD
EstC55-52 184 AFGGDPDQVTIFGESAGAG 302 RDVAVLTGVNKDEYNLFA 407 LGACHALEIPFVFD.NLD
EstC55-62 201 KFGGDPQNVTIFGESAGAI 326 NDTPVLIGTNSDEGALFV 426 DGANHAAEIPYVFG.NLG
EstC55-118 179 AFGGDPERVTVFGESAGAG 297 RDIAILIGVNKDEYNLFT 403 LGACHGLEIPFVFN.NLD
EstC76-136 189 YFGGDPKNVTLFGESAGGM 318 AGIPLIAGANAEEVAFPG 421 LGAFHGLELAPLFGNLLE
Family VIII
WP_005474626 63 PWERDTVVNVWSTTKGATA.LCAHILADRGLLDLD 157 PGTRSGYHAMTFGFLVGEVI
WP_014985987 56 PWQRDTLQLVYSATKGVTT.TLAHLLAERGLLDLD 150 PGTTHGYHGRTFGWLVGEVI
WP_015576461 57 PWERDTVVNVWSTTKGPTA.LCAHILADRGLLDLD 151 PGTRSGYHAITYGFLVGEVV
CDG54282 81 PMHEDAVFRLASITKPIVS.ATLMRLVEEGKLTLD 192 PGEGWRYS.LGLDVLG.AVI
WP_093402197 59 PWTADTPAVVFSCTKGIMA.ICAYLLVQEGRLDLD 153 PGTAHSYHAITYGWLIGEVI
WP_092376400 59 PWTADSPAVVFSVTKGIMA.ICAYQLVQQGRLDLD 153 PGTAHSYHPITYGWLIGEVI
EstC55-4 53 PVSTSTHFRIMSMTKMVCT.AAALQQVERGDLELD 168 PGTRFEYG.INTDWLG.RVV
EstC55-7 91 QMTTDAIFRIYSMTKPVTA.VAMMILFEQGKWQLN 212 PGARWHYS.IAVDIQG.YIV
EstC55-40 48 PVTAETLFQVGSISKVFTT.TLVMTLVEEGKLDLD 145 PGELWTYCNAGFDLAG.RAV
EstC55-46 66 PVRPDTLWRIYSMTKPITS.VAAMMLWEEGAFELT 189 PGTRWGYS.VATDVLG.RLI
EstC55-53 62 PLQHDTLFRIYSMTKPITS.VALMMLVEDGLIALD 184 PGEIWNYS.VSTDVLG.YLV
EstC55-65 74 PWDHDTAAVIFSCTKGILA.VCICLLVQEGRLSLD 168 PGAGHMYHALTYGWIVGEII
EstC55-66 27 PVRRDTLFQIGSITKVFLA.TLAMRLVEEGRLDLD 124 VGRCWSYCNSGFSLAG.RVI
EstC55-73 57 PMREDAIFLLASVTKPIVT.AAALRLVEEGRLDLH 168 PGTSWRYS.LGLDVIG.AVL
EstC55-80 73 PWDHDTGAVIFSCTKGVLA.ICVYMMVQEGRISLD 167 PGAGHVYHAFTYGWIVGEII
EstC76-98 87 PMRNDDIFRIYSMTKPVVS.VALLMLYEEGHFQLS 208 PGEQWLYG.YGHDVQA.RLV
EstC55-110 89 PMKMDTIVRIYSMTKPITG.VAMMMLYEEGKWKPN 209 PGEQWVYS.VSVDIQG.HII
EstC55-113 99 PMQKDSLFQIASMTKPITA.TGLMILVDRGKVGLD 194 PGERWAYS.PGLTVCG.RII
EstC76-123 71 PLASDTLFRIFSLTKPITS.VAALMLVEQGAVALD 192 PGEQWRYG.VSTDVLA.RVV
EstC55-147 75 AFAADHVFLIASAGKPISA.GVLMRLDDQGLLDID 179 PDTWFRYGGAQWQLAG.GIA
EstC55-164 95 PMSKDTYFYVYSMTKPITS.VALLMLYEEGRFQLN 214 PGTQWVYS.VSHDVQA.RLV
EstC55-168 77 RVDERTIFAIGSSSKAFTA.AALAMLVDEGRISWD 173 FRSRYGYQNIMFLAAG.QII
EstC76-174 48 PMREDALFRLYSMTKPWVS.ALALSFVEEGTLSLL 167 PGETFEYG.LATDLLG.HLL
EstC55-239 63 RVTPSSVFDLASLTKVVVTTTAAMQLYEAGKLDLD 157 PGTQSRYSDLGMIVLG.WVI
EstC55-245 61 PVTDTTLFQIGSITKTFTG.TLIMRLVEEGKLALD 158 IGAHWSYNNSGFSLLG.YLI
EstC55-258 59 PWTPDTIVNTYSTTKGVVA.TLFHRFVERGDIDLD 153 PGTAHGYHALTFGFLVGELL
Family XVII
WP_067635253 206 IGLWGYSQGGTSSGWAAEL 352 AAYDEIIPFAQADTLHKAWCAKGANLTWKTYTFAEHATG
WP_055702520 183 VGIMGYSQGGQASSWAAEL 324 ALADELIPYGVGKQVRADWCARGANVEWHTVPVGEHVSG
WP_016645629 188 VGIMGYSQGGQATSWAAEL 329 ALADELIPYGVGKQVRADWCAKGANVEWHTIPLGEHVSG
WP_069887197 183 VGIMGYSQGGQASSWAAEL 324 ALADELIPYGVGKQVRADWCARGANVEWHTVPVGEHVSG
ANA76126 213 IGLMGYSQGGGAAGAAAEL 357 AILDDTIPYAVGKQLGSDWCDKGTRVTFNAGLTPTHVGG
WP_007927380 213 IGLMGYSQGGGAAGAAAEL 357 AILDDTIPYAVGKQLGSDWCDKGTRVTFNAGLTPTHVGG
WP_068264424 213 IGLMGYSQGGGAAAAAAEL 357 AIGDDTIPYAVGKQLGSDWCDKGARVTFNAGIIPTHVGG
WP_068423891 213 VGLMGYSQGGGAAAAAAEL 357 ALGDDVIPYAVGRQLGSDWCDQGTRVSFNAGLTPTHVGG
EstC55-154 186 VALWGYSQGGQAAAAAAEV 332 GAVDQLVPYELGTGLRDAWCGLGADVTFTAYPV.DHFGG

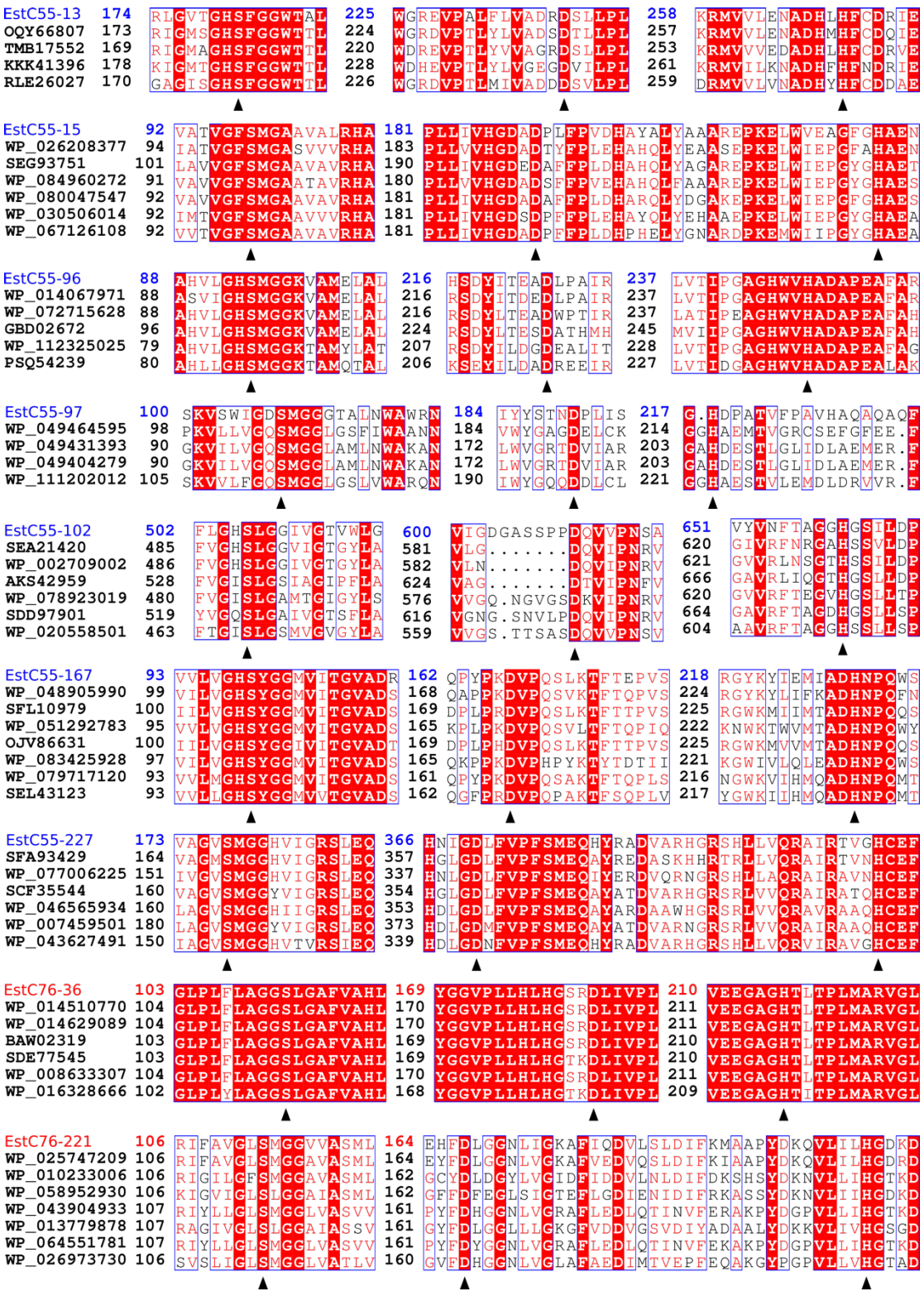

EM3L4
EEP71116 110 TQRFSIGFSYGGAMSY 158 AYLGVHGIGDN..IA 218 AFDGGHIAAPQD
WP_027342034 151 SQVFAMGWSYGGAMSY 198 AYFGIHGIHDSVLNIS 261 AFDGDHTPSPVD
WP_043527065 152 SQLFAVGWSYGGAMSY 199 AYLGIHGTHDSVLNIS 262 AFDGDHTPEPVD
EstC55-42 112 SQIFSLGFSYGGAMSY 159 AYIGLHGTQDNVLPIA 222 AFDGGHNPAPID
EstC55-77 122 QRVYVTGMSNGAFFSS 168 PLLAVHGRLDQVVPY. 238 IEDGGHT.WPGS
ADH59407 155 SRVYVNGFSNGGGMAV 203 PVMAYHGTADPVVPYE 288 IDGGGHT.WPGG
WP_028851258 152 SRIYATGKSNGGGFVG 201 PVLEIHGKADKTIPYE 271 IASLGHD.WP.S
WP_017564998 199 RRVYATGKSNGGGFTG 248 PVIEFHGTDDATIPYG 318 VDGGGHT.WPGA
EstGS
AEM45109 190 RIGVLGHSFGGATSILL 240 EVPFLDLHGTSDGIV 269 PRYRGDIVGGDHLGFL
OGO52417 189 AIGVTGHSLGALTSLLT 238 .VPLLVLGGTRDLLT 266 PRYLVELLGANHIRFA
WP_022959187 172 RIAVMGLSLGGMTSTMA 222 .LPYMMIASPIDALV 250 GATLVSIDKASHTGFA
EstC55-24 180 RVAAMGHSAGGYTTMGL 224 PVPVLFVHGDADSVV 253 PKAFLTVIDGGHTDFL
WP_052387799 205 RVAAGGHSAGGYTTAGM 250 ATPVLFVHGDADATV 279 PKAFLTLLNGDHGGGL
WP_089246854 168 RVAAAGHSAGGITTVGL 214 AAPMLFVHGQRDETV 243 PKAMLTFPKGDHGATL
FLS18
ACL67851 144 RIYLWGHSMGGAGTYHL 196 LQGDQD..RLVTPTRQWVARMKELGMEHIYIEVPGGDHSLF
ACL67852 169 RIYLMGHSMGGGGTLYL 221 VQGDQDRLVSVEIARRWVAKMKELGMTHEYIEIKDGNHVTS
KRO81080 162 RIFLWGHSMGGGGTYHI 217 LQGDQDDLVPVFATRTWVAGMAARGMQHLYVEIEGGDHSLL
WP_014066117 140 RVYLTGLSMGGHGTWYV 216 FHGADDPVVSVEASRRMVEALRELGADVQYTEYEGVGHNAW
WP_022968450 130 RVYLTGLSMGGQGVWQL 204 FHGDADRVVPVEESRAMAQALRKLGGDVRYSEYAGVGHDAW
WP_024868175 138 RTYATGMSMGGYGTWEV 209 FHGALDDLVPPDDDRRLHAAAQDVGADFRYTEYPEGNHNAW
WP_017915553 132 RTYLTGMSMGGYGTWEI 203 FHGAQDDVVLPHDDRKIVRAFKNLDADVRYTEYPQGNHNAW
AAX37300 203 RIYLMGHSMGGAGALYL 256 VQGEKDNLVPAANTRRWVDKLKELNMTYEYLEMQGGDHGSV
WP_050044108 151 RIYLAGHSMGGGGTIHL 202 VTGDKDTTVPVQMIRPFARKMKETNAKHVYKEIAGGNHGTT
EstC55-137 140 RVYLTGLSMGGHGTWYV 216 FHGADDPIVSVEASRRMVEALRELGADVQYTEYEGVGHNAW
EstC55-165 140 RVYLTGLSMGGHGTWYV 216 FHGADDPIVSVEASRRMVEALRELGADVQYTEYEGVGHNAW
EstC55-241 146 RIYLTGLSMGGGGTLWL 199 FQGDADQLVRPEWTREWVRRFREAGVSVEYAEYPGVGHDSW
WP_015814461 146 RTYLTGLSMGGGGTLWI 199 FHGDADPVVPVAGTQKWVAQLQDLGAEVSYKEFVDVKHDSW
WP_031525694 146 KIYLTGLSMGGGGTLWI 199 FHGDADPVVPVDGTRKWVSHLQDIGVEVSYKEFVDVKHDSW
WP_026630869 146 KTYLTGLSMGGGGTLWL 199 FHGDADPVVPVADTRKWVQHLQDIGVEVSYKEFVDVKHDSW
LipT
ADW21422 148 TDPEKVFVTGCSAGAYGAVF 234 IAQYTTLLDGTQIFFY 285 FYLAPGGQHCILPRPE
WP_003047954 148 AQAERVFVTGCSAGAYGAIF 234 LAQYTTLLDGTQIFFY 285 FYLAPGSQHCILPRPE
WP_038060347 148 PKAERVFVTGCSAGGYGAVF 234 FAQYTTQLDGTQIYFY 285 YYLAPGSQHCILPRPE
AFS34517 148 TDPEKVFVTGCSAGAYGAVF 234 IAQYTTLLDGTQIFFY 285 FYLAPGGQHCILPRPE
EstC76-179 148 TNPERVFVTGCSAGAYGAVL 234 LAQYTTLLDGTQIYFY 285 FYLAPGSQHCILPRPE
EstC76-218 148 TNPERVFVTGCSAGAYGAVL 234 LAQYTTLLDGTQIYFY 285 FYLAPGSQHCILPRPE
EstL28
EstC55-81 102 IILLGHSLGALTSIAVAR 192 LAEVADGVFKAVA 255 VVIPETGHSIH
AFK29752 97 VIVAGHSLGALVATWAGA 197 RVHRLDPRVLDAP 262 TVVEAG.HGIH
MBE13165 91 ATLLGHSMGADSSLWIAA 178 CLSILDPDVLSLI 243 VKMPGSGHEPH
OON27855 88 VDLLGHSLGAAVALHVAA 178 LAGTADGVFRAMV 242 AVLPGAGHHVE
WP_012642884 94 IVVIGFSLGALVAIVLAA 175 LCSTADGPILALL 237 VAMPGCGHAIH
WP_051913750 92 AMVIGHSLGGLIALKLAV 173 LVETADGPFLALL 236 VEFPGHGHSIH



Supplementary Figure S6a. Multiple sequence alignments of partial amino acid sequences harboring homologous catalytic regions. Lipolytic enzymes were from reported families. Residues, which are partially consistent, are in frames. Identical residues are shaded in red. Triangles underneath residues indicate the catalytic triad. The functionally identified LEs were assigned to known lipolytic families. **Family I**: EstC55-71, EstC55-88, EstC55-90, EstC55-105, EstC55-151, EstC55-213 and EstC55-235, functionally derived LEs from sample compost55 (this study); EstC76-177, functionally derived LE from sample compost76 (this study); AAB71210, lipase LipA from *Streptomyces cinnamoneus*; AAA22574, lipase from *Bacillus subtilis*; CAA67627, triacylglycerol lipase from *Cutibacterium acnes*; WP_036932411, triacylglycerol lipase from *Cutibacterium avidum*; WP_012843686, alpha/beta fold hydrolase from *Rhodothermus marinus*; WP_071578729, triacylglycerol lipase from *Bacillus* sp. FMQ74; WP_019713218, triacylglycerol lipase from *Bacillus subtilis*. **Family II**: EstC55-111 and EstC55-150, functionally derived LEs from sample compost55 (this study); WP_014065857, SGNH/GDSL hydrolase family protein from *Rhodothermus marinus*; WP_072715438, hypothetical protein from *Rhodothermus profundi*; WP_098062360, hypothetical protein from *Longimonas halophila*; WP_103038013, SGNH/GDSL hydrolase family protein from *Salinivenus iranica*. **Family III**: EstC55-95, functionally derived LE from sample compost55 (this study); M86351, triacylglycerol acylhydrolase from *Streptomyces* sp.; WP_030583320, lipase from *Streptomyces globisporus*; WP_030586638, lipase from *Streptomyces anulatus*; WP_012381325, alpha/beta hydrolase from *Streptomyces*; WP_005154640, lipase from *Amycolatopsis azurea*; WP_092534474, acetylxylan esterase from *Yuhushiella deserti*.

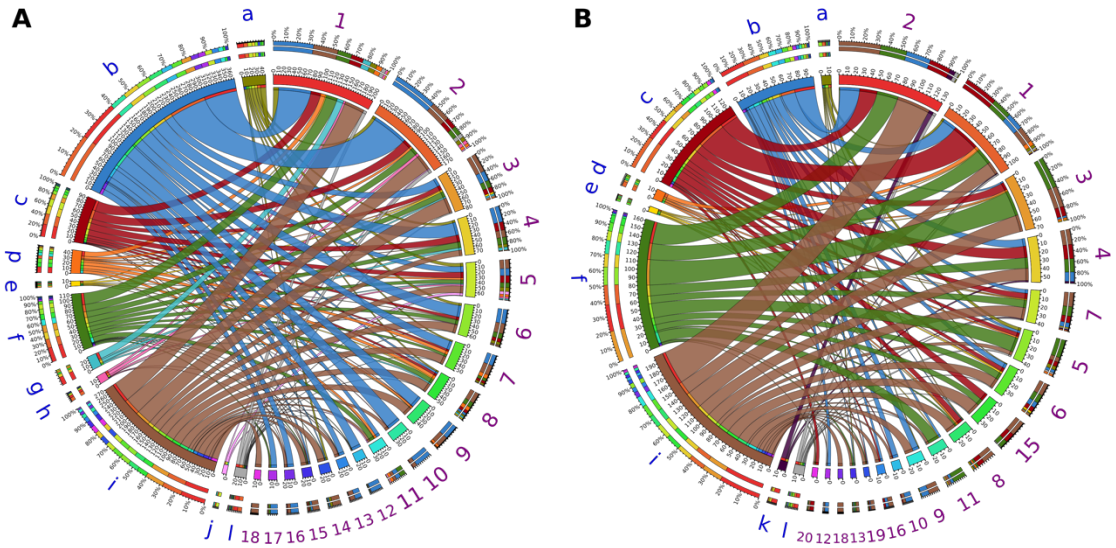
Family IV: EstC55-5, EstC55-23, EstC55-56, EstC55-57, EstC55-60, EstC55-72, EstC55-78, EstC55-145, EstC55-8_1, EstC55-229, EstC55-19_2, EstC55-247, EstC55-253 and EstC55-268, functionally derived LEs from sample compost55 (this study); EstC76-135, F functionally derived LE from sample compost76 (this study); ADH59412, esterase from uncultured bacterium; ADH59413, esterase from uncultured bacterium; AAS77236, lipase/esterase from uncultured bacterium; AAX37296, lipase/esterase from uncultured bacterium. **Family V**: EstC55-2, EstC55-8_2, EstC55-12, EstC55-18, EstC55-19_1, EstC55-20, EstC55-25, EstC55-31, EstC55-34, EstC55-43, EstC55-51, EstC55-76, EstC55-159, EstC55-197, EstC55-215, EstC55-231, EstC55-244 and EstC55-256, functionally derived LEs from sample compost55 (this study); EstC76-28_1, EstC76-28_2, EstC76-248, EstC76-263 and EstC76-266, functionally derived LEs from sample compost76 (this study); CAA37863, triacylglycerol lipase from *Moraxella* sp.; CAA47949, triacylglycerol lipase from *Psychrobacter immobilis*; AAC67392, lipolytic enzyme from *Sulfolobus acidocaldarius*. **Family VII**: EstC55-3, EstC55-52, EstC55-62 and EstC55-118, functionally derived LEs from sample compost55 (this study); EstC76-136, functionally derived LE from sample compost76 (this study); Q01470, serine esterase from *Pseudarthrobacter oxydans*; P37967, para-nitrobenzyl esterase from *Bacillus subtilis*; KJJ40755, para-nitrobenzyl esterase from *Bacillus subtilis*; WP_064730418, carboxylesterase from *Streptomyces parvulus*. **Family VIII**: EstC55-4, EstC55-7, EstC55-40, EstC55-46, EstC55-53, EstC55-65, EstC55-66, EstC55-73, EstC55-80, EstC55-110, EstC55-113, EstC55-147, EstC55-164, EstC55-168, EstC55-239, EstC55-245 and EstC55-258, functionally derived LEs from sample compost55 (this study); EstC76-98, EstC76-123 and EstC76-174, functionally derived LEs from sample compost76 (this study); WP_005474626, esterase from *Streptomyces bottropensis*; WP_014985987, esterase from *Nocardia brasiliensis*; WP_015576461, esterase from *Streptomyces*; CDG54282, esterase EstB from *Halomonas* sp. A3H3; WP_093402197, carboxylesterase from *Verrucosispora sediminis*; WP_092376400, carboxylesterase from *Xiangella phaseoli*. **Family XVII**: EstC55-154, functionally derived LE from sample compost55 (this study); WP_067635253, triacylglycerol lipase from *Actinomadura latina*; WP_055702520, lipase from *Streptomyces silaceus*; WP_016645629, inactive lipase from *Streptomyces aurantiacus*; WP_069887197, lipase from *Streptomyces luteocolor*; ANA76126, secretory lipase LipJ2 from *Janibacter* sp. R02; WP_007927380, secretory lipase from *Janibacter hoylei*; WP_068264424, lipase from *Janibacter limosus*; WP_068423891, lipase from *Janibacter terrae*. **EM3L4**: EstC55-42 and EstC55-77, functionally derived LEs from sample compost55 (this study); EEP71116, ferruloyl esterase fee1B from *Micromonospora* sp. ATCC 39149; WP_027342034, cellulose-binding protein from *Hamadaea tsunoensis*; WP_043527065, cellulose-binding protein from *Actinoplanes utahensis*; ADH59407, esterase/lipase from uncultured bacterium; WP_028851258, hypothetical protein from *Thermocrispum municipal*; WP_017564998, hypothetical protein from *Nocardiopsis synnemataformans*. **EstGS**: EstC55-24, functionally derived LE from sample compost55 (this study); AEM45109, hypothetical protein from uncultured organism; OGO52417, hypothetical protein from *Chloroflexi* bacterium; WP_022959187, alpha/beta hydrolase from *Spongiibacter tropicus*; WP_052387799, alpha/beta hydrolase from *Dactylosporangium aurantiacum*; WP_089246854, chlorophyllase from *Asanoa hainanensis*. **FLS18**: EstC55-137, EstC55-165 and EstC55-241, functionally derived LEs from sample compost55 (this study); ACL67851, esterase/lipase from uncultured bacterium FLS18; ACL67852, esterase/lipase from uncultured bacterium FLS18; KRO81080, hypothetical protein from OM182 bacterium; WP_014066117, phospholipase from *Rhodothermus marinus*; WP_022968450, phospholipase from *Arenimonas oryziterrae*; WP_024868175, phospholipase from *Pseudoxanthomonas suwonensis*; WP_017915553, phospholipase from *Xanthomonas* sp. SHU 308; AAX37300,

lipase/esterase from uncultured bacterium; WP_050044108, alpha/beta hydrolase from *Verrucomicrobia* bacterium SCGC AAA168-F10; WP_015814461, phospholipase/carboxylesterase from *Dyadobacter fermentans*; WP_031525694, phospholipase/carboxylesterase from *Dyadobacter crusticola*; WP_026630869, phospholipase/carboxylesterase from *Dyadobacter alkalitolerans*. **LipT**: EstC76-179 and EstC76-218, functionally derived LEs from sample compost76 (this study); ADW21422, putative esterase from *Thermus scotoductus* SA-01; WP_003047954, esterase from *Thermus aquaticus*; WP_038060347, esterase from *Thermus filiformis*; AFS34517, LipT from uncultured bacterium. EstL28: EstC55-81, functionally derived esterase from sample compost55 (this study); AFK29752, esterase from uncultured bacterium; MBE13165, hypothetical protein from *Chloroflexi* bacterium; OON27855, hypothetical protein from *Micromonospora* sp. Rc5; WP_012642884, alpha/beta hydrolase from *Thermomicrobium roseum*; WP_051913750, alpha/beta hydrolase from *Thermorudis peleae*; **Patatin-like-protein**: EstC55-10, EstC55-26, EstC55-63, EstC55-131, EstC55-163 and EstC55-251, functionally derived LEs from sample compost55 (this study); EstC76-222, EstC76-261 and EstC76-269 functionally derived LEs from sample compost76 (this study); CZI05393, patatin from *Legionella pneumophila*; WP_080020835, patatin-like phospholipase family protein from *Legionella pneumophila*; WP_013131472, patatin from *Thermobispora bispora*; SNR91842, NTE family protein from *Streptosporangium subroseum*; WP_017249383, esterase from *Brevibacillus brevis*; WP_035162041, esterase from *Caloranaerobacter azorensis*; WP_036322013, patatin-like phospholipase family protein from *Microbispora* sp.. **Tannase**: EstC55-156, EstC55-234 and EstC55-269, functionally derived LEs from sample compost55 (this study); OFV97653, hypothetical protein from *Acidobacteria* bacterium; OLB33458, feruloyl esterase from *Acidobacteria bacterium*; WP_046794161, tannase/feruloyl esterase family alpha/beta hydrolase from *Rhizobium* sp.; WP_020718491, tannase/feruloyl esterase family alpha/beta hydrolase from *Acidobacteriaceae* bacterium.

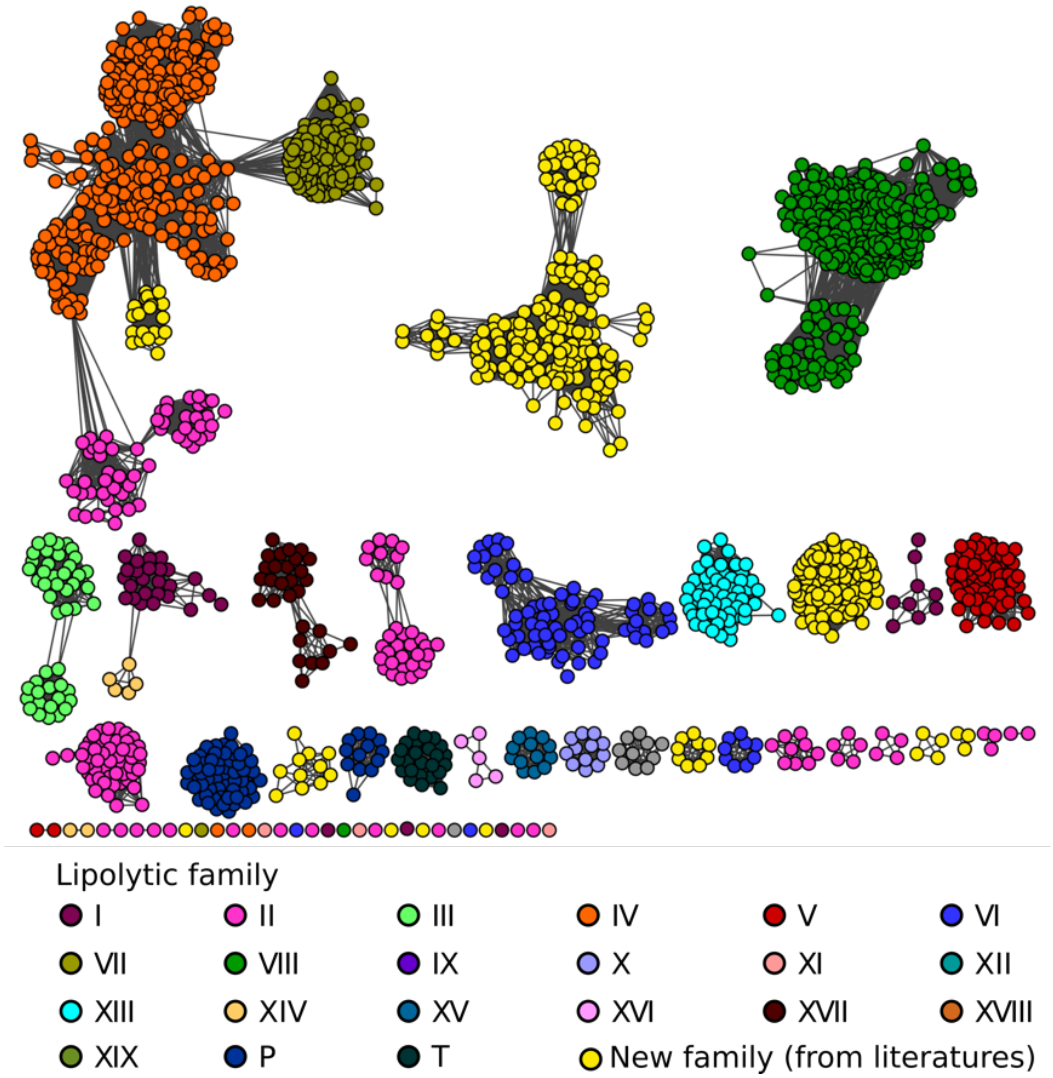


Supplementary Figure S6b. Multiple sequence alignments of partial amino acid sequences harboring homologous catalytic regions. Lipolytic enzymes were from putative novel families identified in this study. Residues, which are partially consistent, are in frames. Identical residues are shaded in red. Triangles underneath residues indicate the catalytic triad. **Putative new family 1**: EstC55-13, functionally derived LE from sample compost55 (this study); OQY66807, hypothetical protein from *Polyangiaceae* bacterium UTPRO1; TMB17552, hypothetical protein from *Deltaproteobacteria*

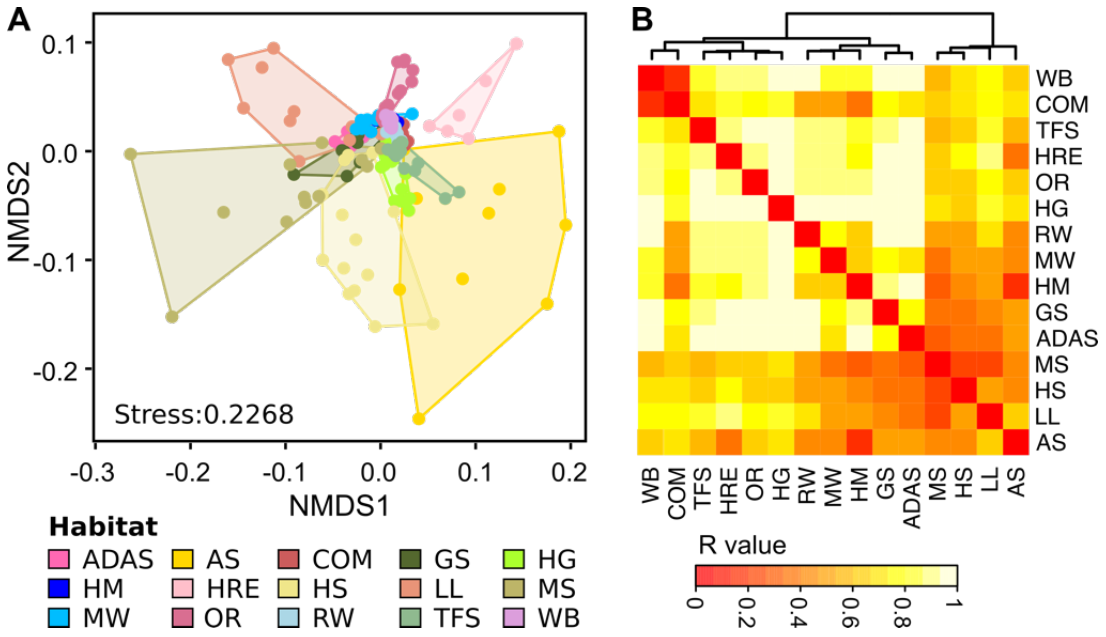
bacterium; KKK41396, Alpha/beta hydrolase family protein from *Lokiarchaeum* sp. GC14_75; RLE26027, hypothetical protein from *Actinobacteria bacterium*. **Putative new family 2**: EstC55-15, functionally derived LE from sample compost55 (this study); WP_026208377, hydrolase from *Catelliglobosispora koreensis*; SEG93751, Alpha/beta hydrolase family protein from *Nonomuraea solani*; WP_084960272,alpha/beta hydrolase from *Thermoactinospora rubra*; WP_080047547, alpha/beta hydrolase from *Nonomuraea* sp. ATCC 55076; WP_030506014, alpha/beta hydrolase from *Microbispora rosea*; WP_067126108, hydrolase from *Microtetraspora malaysiensis*. **Putative new family 3**: EstC55-96, functionally derived LE from sample compost55 (this study); WP_014067971, alpha/beta fold hydrolase from *Rhodothermus marinus*; WP_072715628, alpha/beta fold hydrolase from *Rhodothermus profundi*; GBD02672, Esterase YbfF from *bacterium* HR18; WP_112325025, alpha/beta fold hydrolase from *Rhodothermaceae* bacterium; PSQ54239, alpha/beta hydrolase from *Bacteroidetes* bacterium QH_10_64_37. **Putative new family 4**: EstC55-97, functionally derived LE from sample compost55 (this study); WP_049464595, alpha/beta hydrolase from *Stenotrophomonas maltophilia*; WP_049431393, alpha/beta hydrolase from *Stenotrophomonas maltophilia*; WP_049404279, alpha/beta hydrolase from *Stenotrophomonas maltophilia*; WP_111202012.1, hypothetical protein from *Stenotrophomonas maltophilia*. **Putative new family 5**: EstC55-102, functionally derived LE from sample compost55 (this study); SEA21420, alpha/beta hydrolase family protein from *Thiothrix caldifontis*; WP_002709002, lipase from *Thiothrix nivea*; AKS42959, Extracellular lipase, Pla-1/cef family from *Wenzhouxiangella marina*; WP_078923019, lipase from *Thiothrix eikelboomii*; SDD97901, Alpha/beta hydrolase family protein from *Aquimonas voraii*; WP_020558501, hypothetical protein from *Thiothrix flexilis*. **Putative new family 6**: EstC55-167, functionally derived LE from sample compost55 (this study); WP_048905990, alpha/beta hydrolase from *Pedobacter* sp. V48; SFL10979, alpha/beta hydrolase family protein from *Porphyromonadaceae* bacterium KH3CP3RA; WP_051292783, alpha/beta hydrolase from *Olivibacter sitiensis*; OJV86631, alpha/beta hydrolase from *Bacteroidia* bacterium 44-10; WP_083425928, alpha/beta hydrolase from *Zhouia amylolytica*; WP_079717120, alpha/beta hydrolase from *Parapedobacter luteus*; SEL43123, Pimeloyl-ACP methyl ester carboxylesterase from *Parapedobacter koreensis*. **Putative new family 7**: EstC55-227, functionally derived LE from sample compost55 (this study); SFA93429, hypothetical protein SAMN05216266_102299 from *Amycolatopsis marina*; WP_077006225, hypothetical protein from *Saccharothrix* sp. ALI-22-I; SCF35544, Alpha/beta hydrolase family from *Micromonospora saelicesensis*; WP_046565934, alpha/beta hydrolase from *Micromonospora* sp. HK10; WP_007459501, alpha/beta hydrolase from *Micromonospora lupini*; WP_043627491, alpha/beta hydrolase from *Nonomuraea candida*. **Putative new family 8**: EstC76-36, functionally derived LE from sample compost76 (this study); WP_014510770, alpha/beta hydrolase from *Thermus thermophilus*; WP_014629089, phospholipase from *Thermus thermophilus*; BAW02319, esterase from *Thermus thermophilus*; SDE77545, hypothetical protein SAMN04488243_1105 from *Thermus arciformis*; WP_008633307, phospholipase from *Thermus parvatiensis*; WP_016328666, alpha/beta hydrolase from *Thermus oshimai*. **Putative new family 9**: EstC76-221, functionally derived LE from sample compost76 (this study); WP_025747209, alpha/beta hydrolase from *Caldicoprobacter*; WP_010233006, alpha/beta hydrolase from *Clostridium arbusti*; WP_058952930 alpha/beta fold hydrolase from *Clostridium tyrobutyricum*; WP_043904933 alpha/beta fold hydrolase from *Parageobacillus genomosp.*; WP_013779878 alpha/beta hydrolase from *Mahella australiensis*; WP_064551781 alpha/beta fold hydrolase from *Parageobacillus thermoglucosidasius*; WP_026973730 alpha/beta fold hydrolase from *Alicyclobacillus contaminans*.



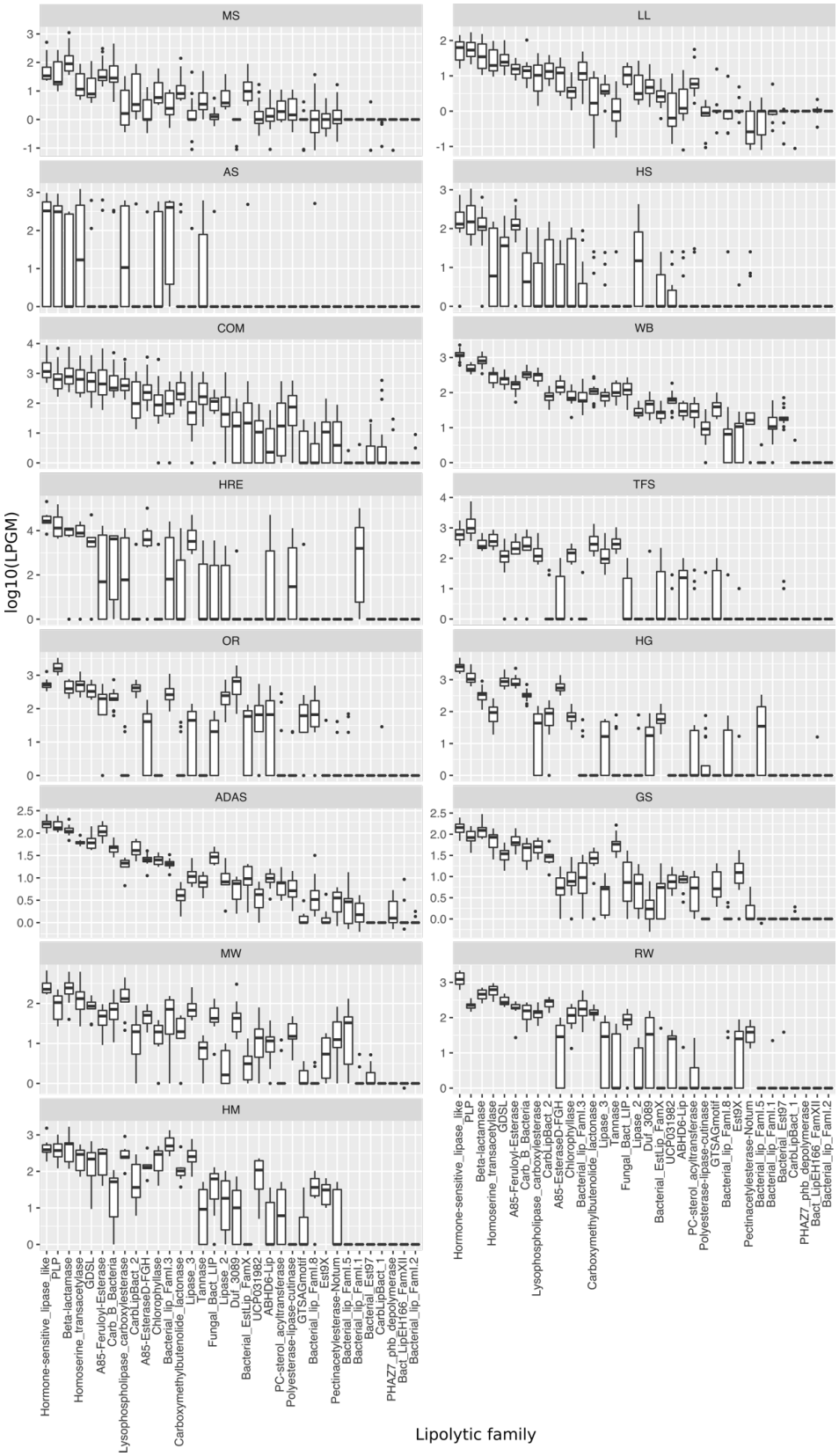
Supplementary Figure S7. Phylogenetic distribution at phylum level of assigned PLPs in the most abundant lipolytic families. Assigned PLPs were identified by screening the assembled metagenome of (**A**) compost55 and (**B**) compost76. Taxonomic information of PLP-encoding genes was annotated by KAIJU (Menzel et al. 2016). The data was visualized via Circos software. The width of bars from each phylum (blue) and ESTHER family (purple) indicates their relative abundance. Bacterial phyla: a, *Acidobacteria*; b, *Actinobacteria*; c, *Bacteroidetes*; d, *Chloroflexi* e, *Deinococcus-Thermus*; f, *Firmicutes*; g, *Gemmatimonadetes*; h, *Planctomycetes*; i, *Proteobacteria*; j, *Thermotogae*; k, *Verrucomicrobia*; l, unclassified Bacteria. Lipolytic families in ESTHER databases: 1, VIII ; 2, Hormone-sensitive_lipase_like; 3, patatin-like-protein; 4, II; 5, A85-Feruloyl-Esterase; 6, Carb_B_Bacteria; 7, Homoserine_transacetylase; 8, Lysophospholipase_carboxylesterase; 9, Carboxymethyl-butenolide_lactonase; 10, Polyesterase-lipase-cutinase; 11, CarbLipBact_2; 12, Lipase_2; 13, Chlorophyllase; 14, Tannase; 15, *A85*-EsteraseD-FGH; 16, Fungal_Bact_LIP; 17, Est9X; 18, Bacterial_lip_FamI.3; 19, PC-sterol_acyltransferase; 20, Lipase_3. Only phyla and protein families with a relative abundance >1% are shown.



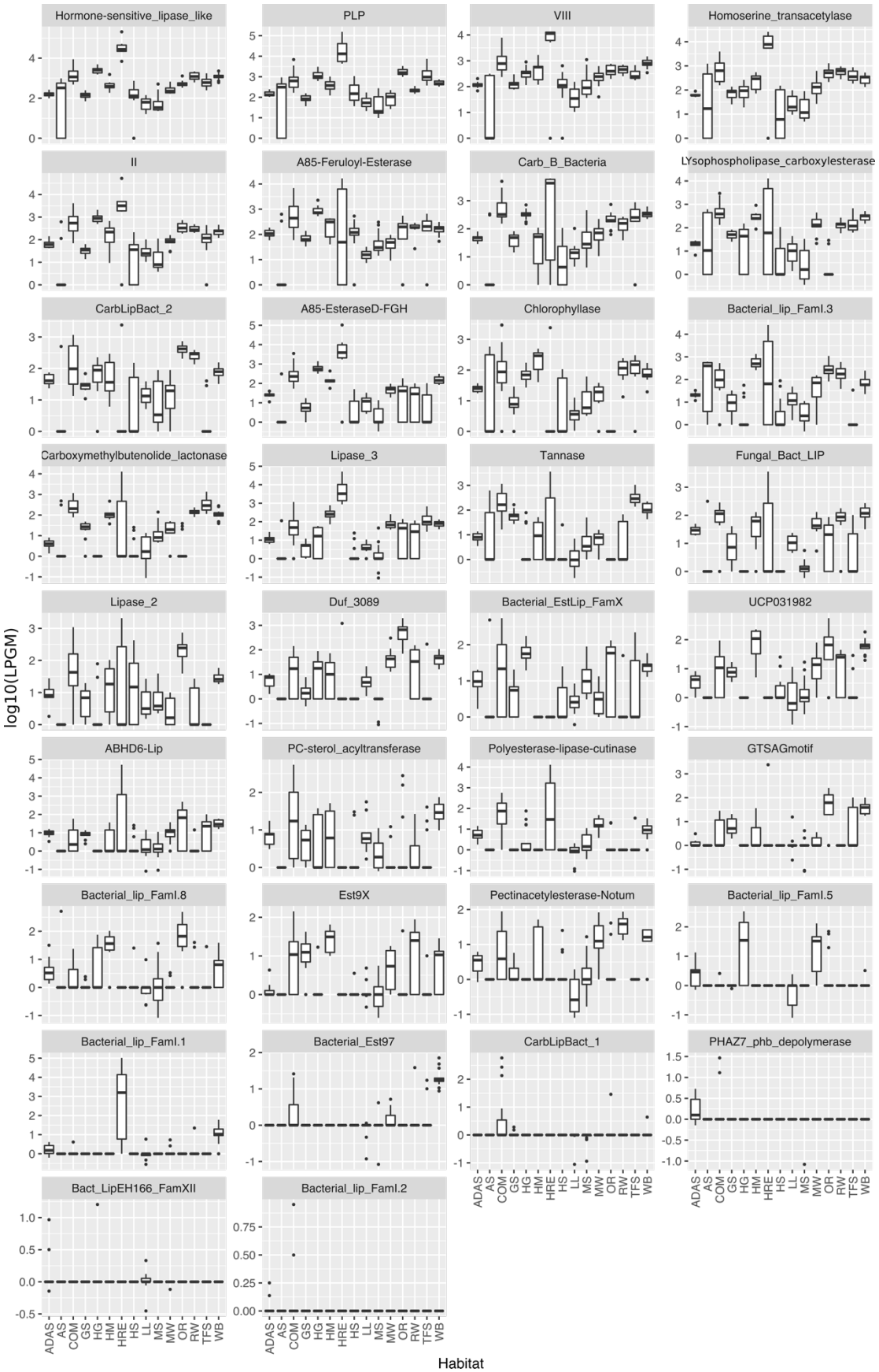
Supplementary Figure S8. Protein Sequence similarity network for classification of assigned PLPs obtained by screening against from compost55 and compost 76 assembled metagenomes. Assigned PLPs were pooled and clustered at 100 % identity using CD-HIT (Huang et al. 2010). Then, the resulting sequences were submitted to the EFI-EST (Gerlt et al. 2015) to generate the network. Each node represents an assigned PLP and is colored according to its lipolytic family. Each edge in the network represents a BLAST connection with an E-value cutoff of $\leq 1e^{-15}$. At this cut-off, sequences have a median percent identity and alignment length of 35% and 291 amino acids, respectively. Lengths of edges are not meaningful except that sequences in tightly clustered groups are relatively more similar to each other than sequences with few connections. Nodes were arranged using the *yFiles* organic layout provided with Cytoscape version 3.4.0 (Shannon et al. 2003).



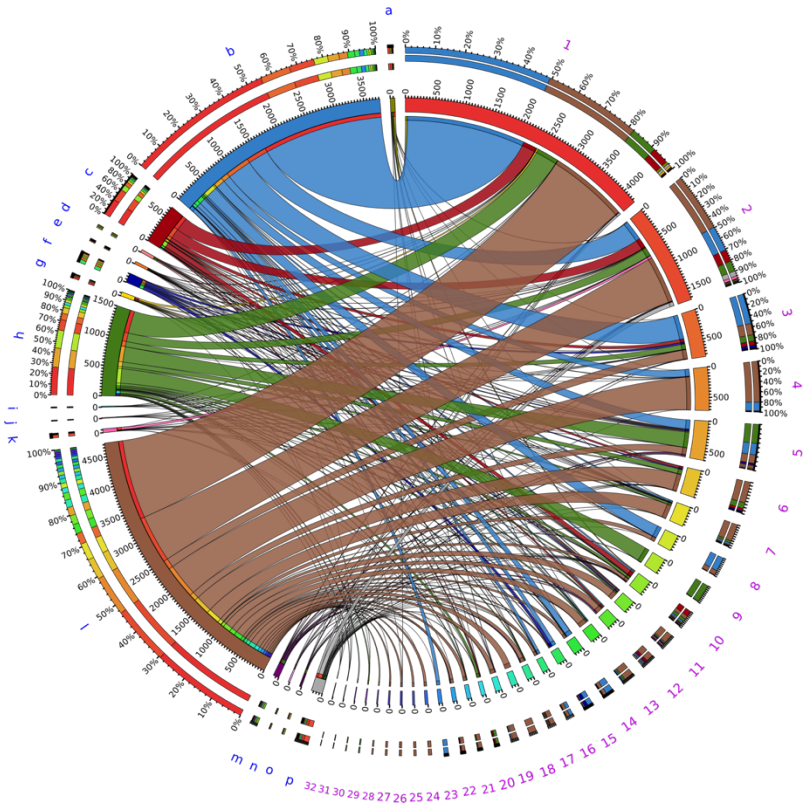
Supplementary Figure S9. Functional lipolytic family profiles of assigned PLPs in different samples. A, Non-metric multidimensional scaling (NMDS) analysis of the lipolytic family profiles across samples was performed. Samples were colored by its habitat source. B, ANOSIM test the group dissimilarity of lipolytic family profiles between habitats (9999 permutations, $p < 0.001$). The resulting R values are shown by the heatmap, and the color intensity (red to light yellow) indicates the change of R values (0 to 1). Hierarchical clustering analysis of R values was performed to generate the cluster dendrogram using the Ward.D clustering method based on Bray-Curtis distance matrices. For all the analysis, LPGM values were $\log_{10}$ transformed. Abbreviations of habitats: ADAS, anaerobic digestor active sludge; AS, agricultural soil; COM, compost; GS grassland soil; HG, human gut; HM, hypersaline mat; HRE, hydrocarbon resource environment; HS, hot spring; LL, landfill leachate; MS, marine sediment; MW, marine water; OR, oil reservoir; RW, river water; TFS, tropical forest soil; WB, wastewater bioreactor.



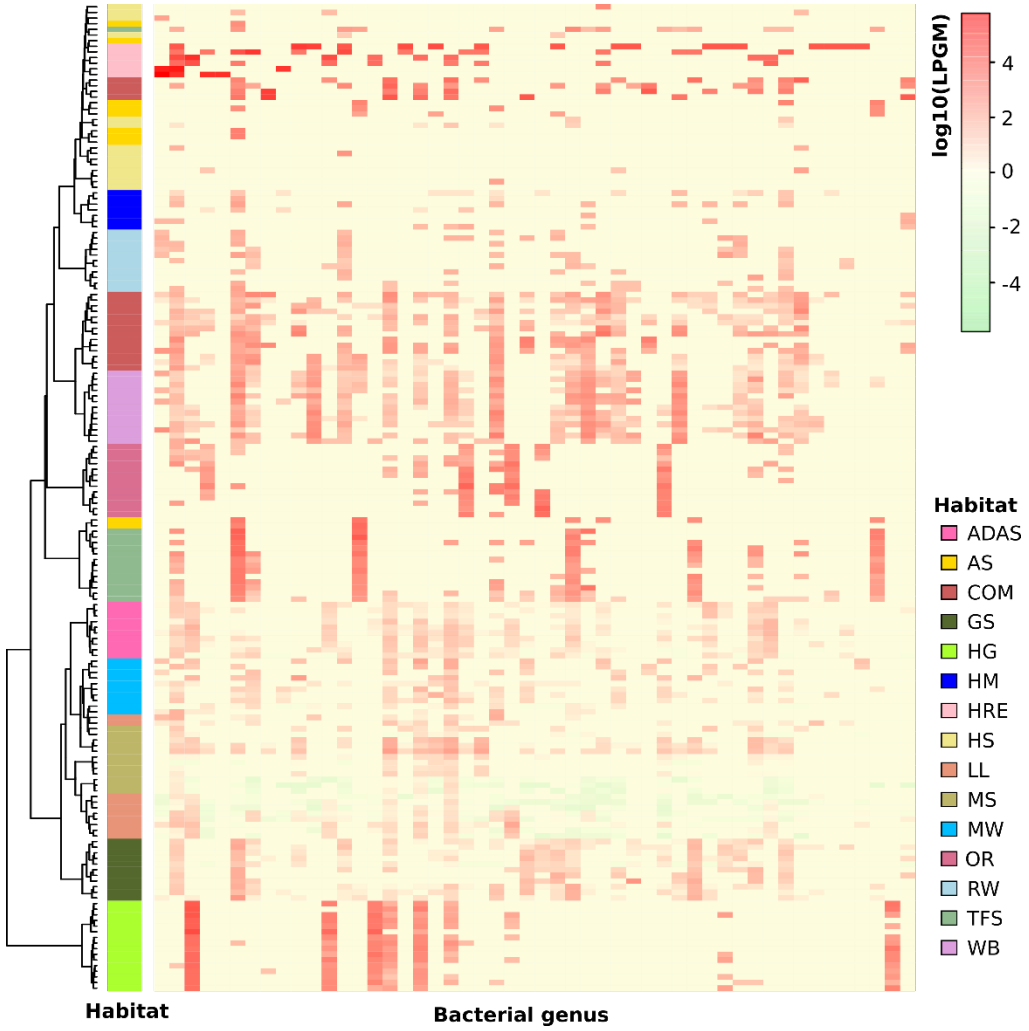
Supplementary Figure S10. Distribution of lipolytic families revealed from assigned PLPs of each habitat. The abundance was inferred from $\log_{10}$ scaled LPGM values.



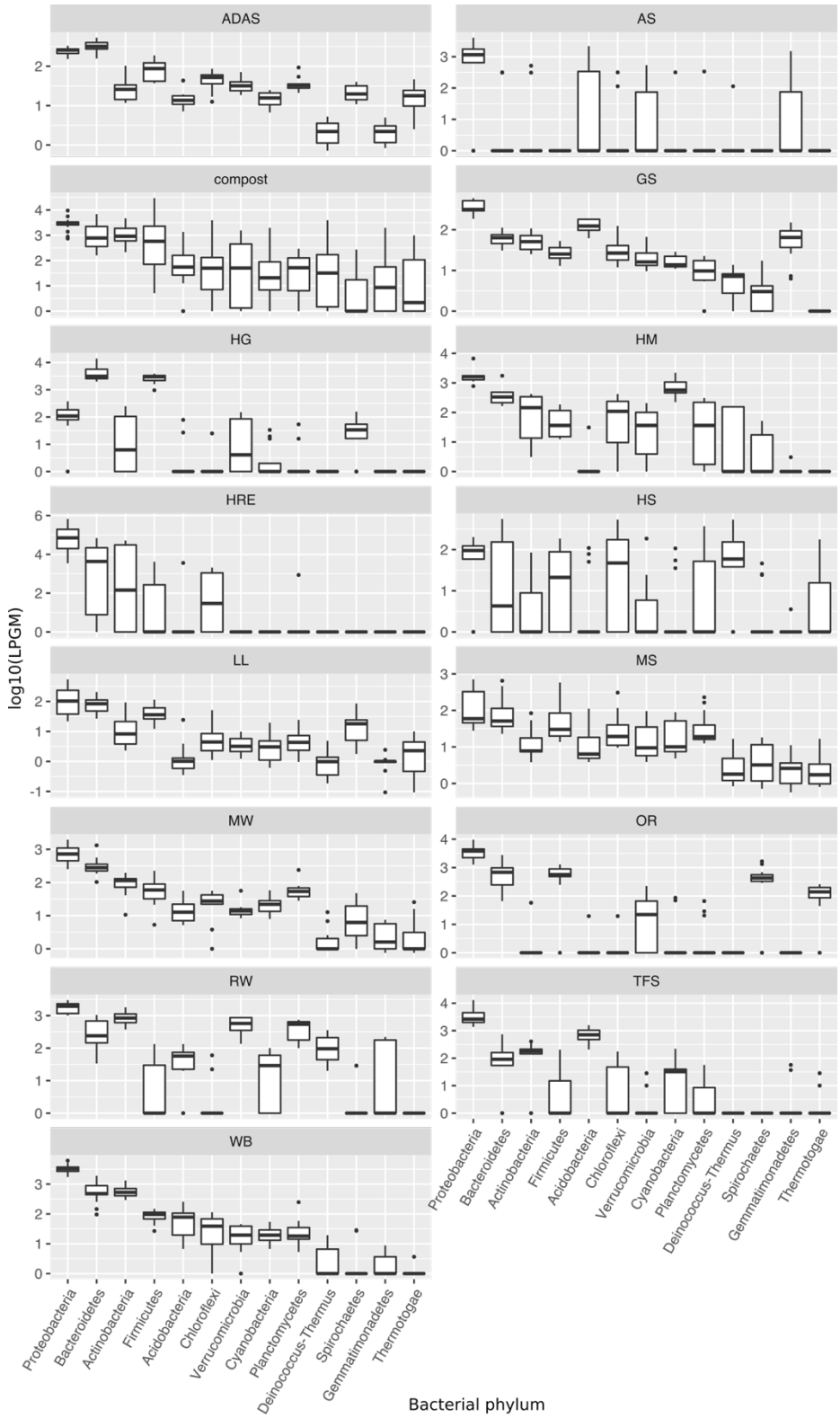
Supplementary Figure S11. Lipolytic families showing significant changes in abundance across different habitats. The abundance was inferred from $\log_{10}$ scaled LPGM values.



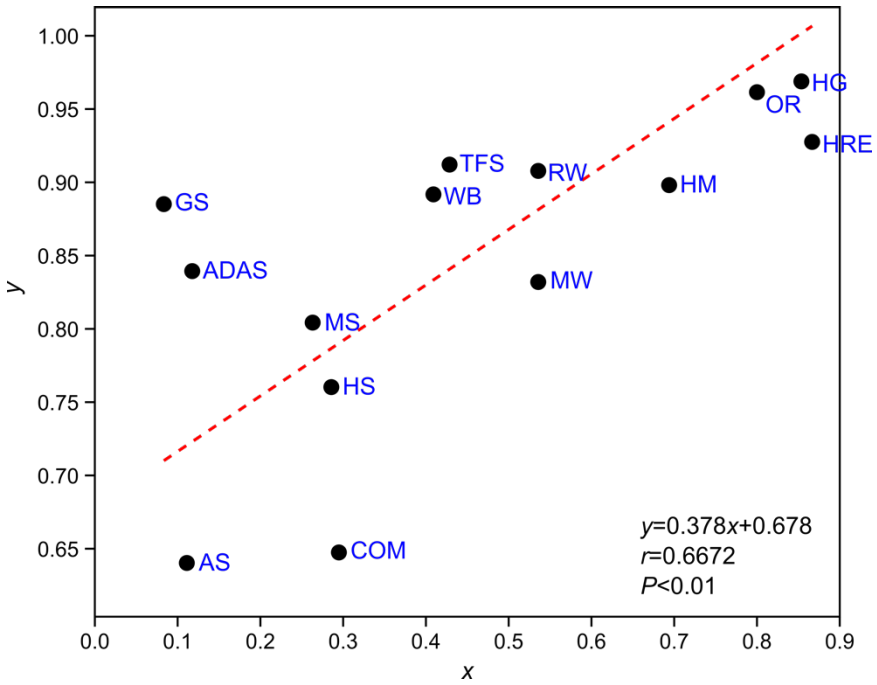
Supplementary Figure S12. Phylogenetic origins of LEs in ESTHER database at phylum level in the most abundant lipolytic families. Taxonomic information of lipolytic genes was retrieved from ESTHER database (Lenfant et al. 2013). Visualization was performed via Circos software. The width of bars from each phylum (blue) and lipolytic family (purple) indicates their relative abundance. Bacterial phyla: a, *Acidobacteria*; b, *Actinobacteria*; c, *Bacteroidetes*; d, Chlorobi; e, *Chloroflexi*; f, Cyanobacteria; g, *Deinococcus-Thermus*; h, *Firmicutes*; i, *Gemmatimonadetes*; j, *Lentisphaerae*; k, *Planctomycetes*; l, *Proteobacteria*; m, *Spirochaetes*; n, *Thermotogae*; o, *Verrucomicrobia*; p, unclassified Bacteria. Lipolytic family based on ESTHER database (Arpingy classification): 1, Carb_B_Bacteria (VII) ; 2, Hormone-sensitive_lipase_like (IV); 3, Chlorophyllase (EstGS); 4, Carboxymethylbutenolide_lactonase (V.2); 5, CarbLipBact_2 (XIII-2); 6, A85-EsteraseD-FGH (lp_3505/FLS12/EstA); 7, Homoserine_transacetylase (Est22); 8, Fungal-Bact_LIP (X-2/XVII/XIX); 9, CarbLipBact_1 (XIII-1/XVIII); 10, A85-Feruloyl-Esterase (Rlip1/EstSt7); 11, LYsophospholipase_carboxylesterase (VI); 12, Duf_3089 (XV); 13, UCP031982 (V.3); 14, Est9X (Est9X); 15, Lipase_2 (I.4/I.7); 16, Polyesterase-lipase-cutinase (III); 17, Pectinacetylesterase-Notum (LipT); 18, Bacterial_lip_FamI.8 (I.8); 19, GTSAGmotif (IV); 20, Bacterial_EstLip_FamX (X); 21, Lipase_3 (XI); 22, ABHD6-Lip (V.1); 23, Bacterial_Est97 (XVI); 24, Bacterial_lip_FamI.1 (I.1); 25, PHAZ7_phb_depolymerase (IX); 26, Tannase (Tannase); 27, Bacterial_lip_FamI.3 (I.3); 28, Bacterial_lip_FamI.2 (I.2); 29, Bacterial_lip_FamI.5 (I.5); 30, Bacterial_lip_FamI.6 (I.6); 31, PC-sterol_acyltransferase (XIV); 32, Bact_LipEH166_FamXII (XII). Only phyla and ESTHER families with a relative abundance >1% are shown..



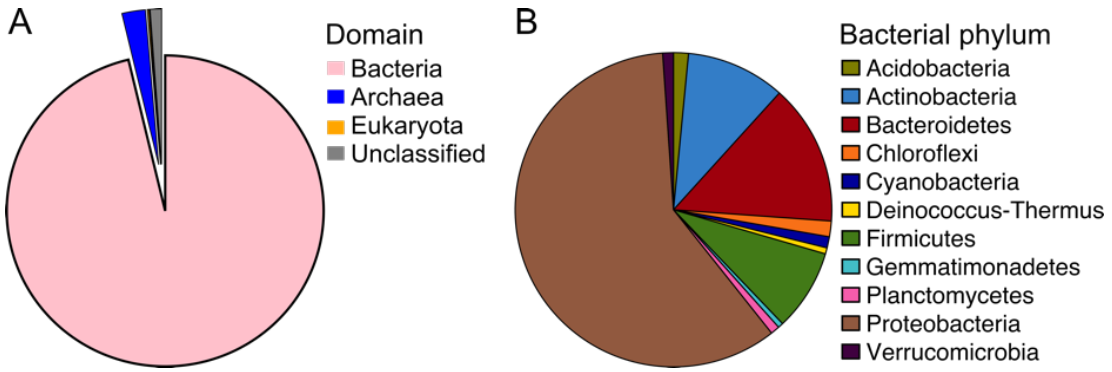
Supplementary Figure S13. Taxonomic origins at genus level of the assigned PLPs across samples. The abundance of assigned PLPs per each genus in each sample was inferred from LPGM values. Only genera with a mean LPGM value of ≥0.5 across all the samples were used for analysis, only the top 50 genera areshown here (ranked by the mean LPGM values across samples). Hierarchical clustering analysis of the phylogenetic distribution profile in each sample was performed using the Ward.D clustering method based on Bray-Curtis distance matrices. The color intensity of the heat map (light green to red) indicates the change of abundance (low to high). The habitats were presented by different colors. The phylogenetic distribution of assigned PLPs in each sample was generally clustered by habitats (overall R value = 0.8199, P <0.001, ANOSIM test). Abbreviations: ADAS, anaerobic digestor active sludge; AS, agriculture soil; COM, compost; GS grassland soil; HG, human gut; HM, hypersaline mat; HRE, hydrocarbon resource environments; HS, hot spring; LL, landfill leachate; MS, marine sediment; MW, marine water; OR, oil reservoir; RW, river water; TFS, tropical forest soil; WB, wastewater bioreactor.



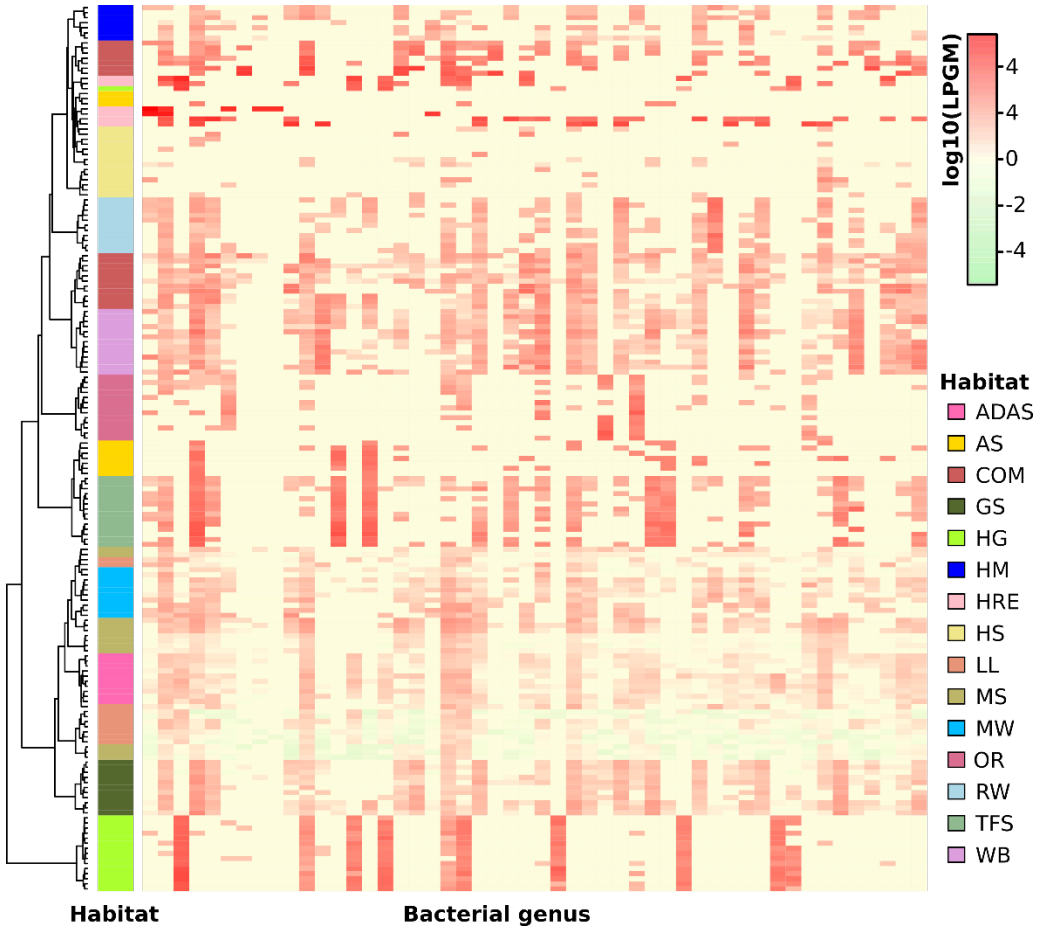
Supplementary Figure S14. Phylogenetic distribution of the assigned PLPs at phylum level in each habitat. The abundance was inferred from $\log_{10}$ scaled LPGM values.



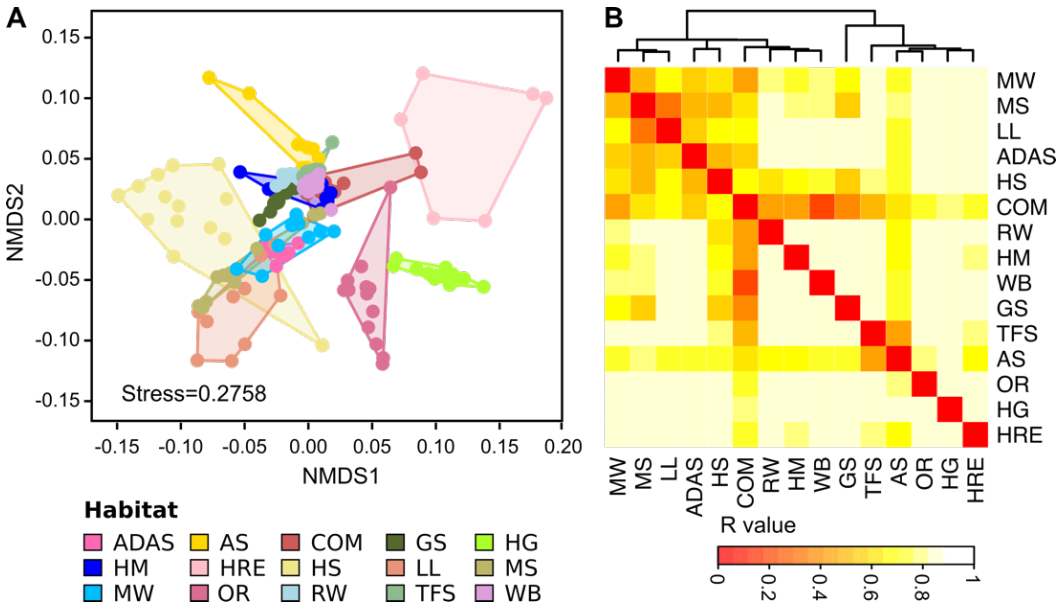
Supplementary Figure S15. Linear regression. The x was the ratio of unique indicators to the total significant indicators in a habitat, as demonstrated by the bipartite association network shown in Figure 5. The corresponding y was the mean dissimilarity of the taxonomic profile of assigned PLPs across habitats, in terms of averaged R values generated by the ANOSIM test (P <0.001; see Supplementary Table S18). The Reduced Major Axis (RMA) algorithm was used for regression. The permutation test on correlation uses 9,999 replicates. Pearson's r correlation r=0.6672, p<0.01.



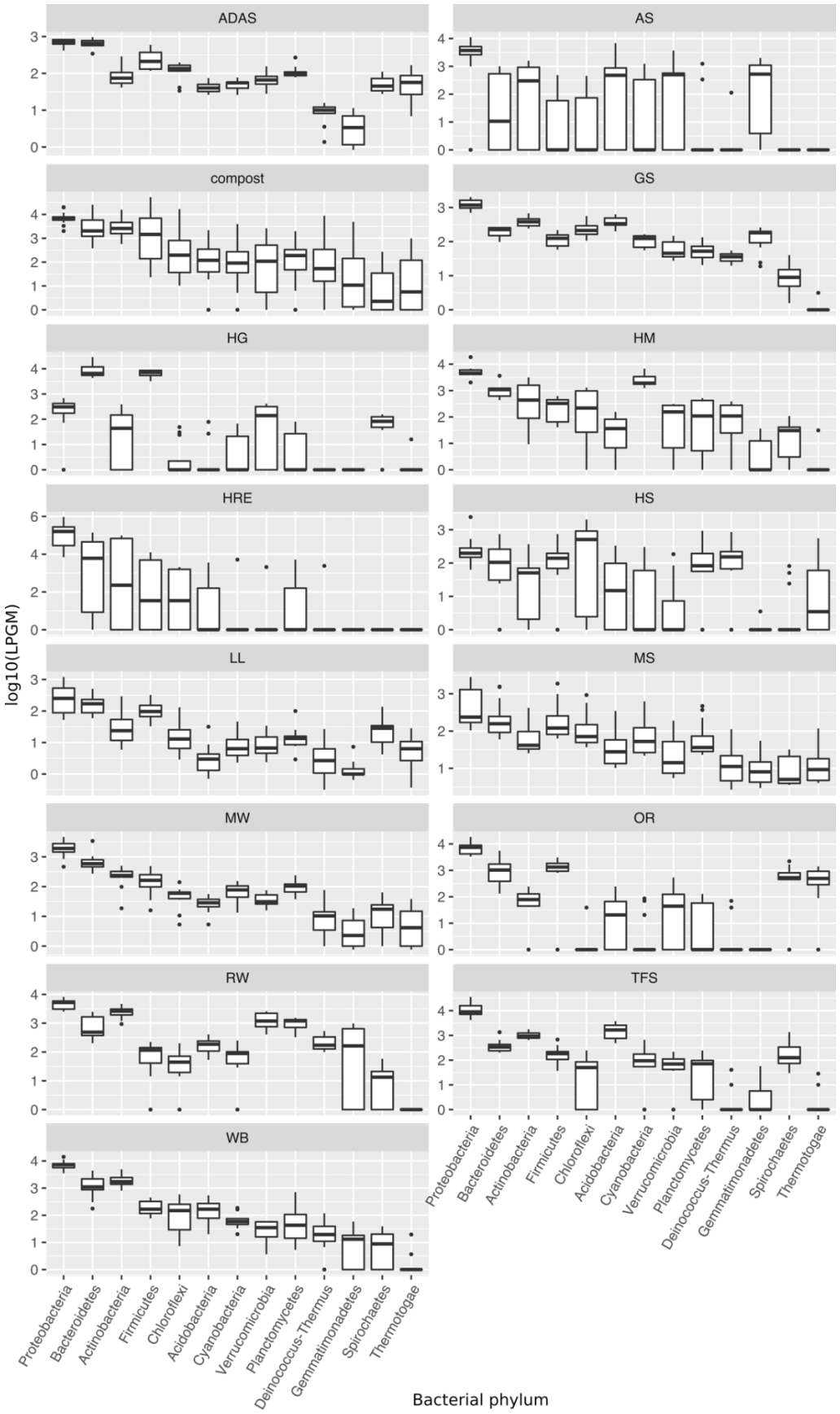
Supplementary Figure S16. Phylogenetic origin of the total PLPs (assigned and unassigned PLPs combined) at (A) domain and (B) phylum level. The abundance of PLPs in each domain or phylum was calculated by summing the LPGMs values across samples.



Supplementary Figure S17. Heat map of the taxonomic origins at genus level of total PLPs across samples. The abundance of total PLPs per genus in each sample was inferred from LPGM values. Only genera with a mean LPGM value of ≥0.5 across all samples were used for analysis and only the top 50 genera are shown here (ranked by the mean LPGM values across samples). The clustering analysis was performed using the Ward.D clustering method based on Bray-Curtis distance matrices. The color intensity of the heat map (light green to red) indicates the change of abundance (low to high). The habitats are presented by different colors. The phylogenetic distribution of assigned PLPs in each sample clustered generally by habitats (overall R value = 0.821, P <0.001, ANOSIM test). Abbreviations: ADAS, anaerobic digestor active sludge; AS, agriculture soil; COM, compost; GS grassland soil; HG, human gut; HM, hypersaline mat; HRE, hydrocarbon resource environments; HS, hot spring; LL, landfill leachate; MS, marine sediment; MW, marine water; OR, oil reservoir; RW, river water; TFS, tropical forest soil; WB, wastewater bioreactor.



Supplementary Figure S18. Analysis of the phylogenetic profile at genus level of total PLPs across samples. A, Non-metric multidimensional scaling (NMDS) analysis of the phylogenetic profile of total PLPs across samples was performed. Samples were colored by its habitat origin. The abundance of PLPs in each genus per sample is presented by the LPGM values. Only genera with mean LPGM values of ≥0.5 across all the samples were used for analysis. B, ANOSIM test the group dissimilarity of these phylogenetic profiles between habitats (9999 putations, P <0.001). The resulting R values are shown by the heatmap, and the color intensity (red to light yellow) indicates the change of R values (0 to 1). Hierarchical clustering analysis of R values was performed to generate the cluster dendrogram using the Ward.D clustering method based on Bray-Curtis distance matrices



Supplementary Figure S19. Phylogenetic distribution of the total PLPs at phylum level of each habitat. The abundance was inferred from $\log_{10}$ scaled LPGM values.

References

Gerlt JA, Bouvier JT, Davidson DB, Imker HJ, Sadkhin B, Slater DR, Whalen KL (2015) Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST): A web tool for generating protein sequence similarity networks. Biochim Biophys Acta 1854:1019–1037 . doi: 10.1002/cncr.27633.Percutaneous

Huang Y, Niu B, Gao Y, Fu L, Li W (2010) CD-HIT Suite: a web server for clustering and comparing biological sequences. Bioinformatics 26:680–692 . doi: 10.1093/bioinformatics/btq003

Keegan KP, Glass EM, Meyer F (2016) MG-RAST, a metagenomics service for analysis of microbial community structure and function. In: Methods in Molecular Biology. Humana Press Inc., pp 207–233

Lenfant N, Hotelier T, Velluet E, Bourne Y, Marchot P, Chatonnet A (2013) ESTHER, the database of the α/β-hydrolase fold superfamily of proteins: tools to explore diversity of functions. Nucleic Acids Res 41:423–429 . doi: 10.1093/nar/gks1154

Menzel P, Ng KL, Krogh A (2016) Fast and sensitive taxonomic classification for metagenomics with Kaiju. Nat Commun 7:11257 . doi: 10.1038/ncomms11257

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO (2013) The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. Nucleic Acids Res 41:D590–D596 . doi: 10.1093/nar/gks1219

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: A software environment for integrated models of biomolecular interaction networks. Genome Res 13:2498–2504 . doi: 10.1101/gr.1239303

Supplementary Table S1 Recent reported novel lipolytic enzymes/families

Lipolytic family	Sequence length (aa)	Function	Reference
Est22	463	Homoserine transacetylase	Li et al., 2017
EstL28	290	Esterase	Seo et al., 2013
Rv0045c	298	Esterase	Guo et al., 2010
EstGX1	201	Lipase	Jiménez et al., 2012
EstLiu	293	Esterase	Rahman et al., 2016
EstY	423	Esterase	Wu & Sun, 2009
EstGS	397	Esterase	Nacke et al., 2011
EM3L4	330	Lipase	Jeon et al., 2011
FLS18	259	Esterase	Hu et al., 2010
Est903	300	Esterase	Jia et al., 2019
EstJ	317	Esterase	Choi et al., 2013
PE10	279	Esterase	Jiang et al., 2012
Est12	329	Esterase	Wu et al., (2013)
EstDZ2	271	Secreted carboxylesterase	Zarafeta et al., 2016
Est9x	294	Lipase	Jeong et al., 2009
Lip10	348	Lipase	Guo et al., 2016
EstGH	448	Esterase	Nacke et al., 2011
EML1	304	Esterase	Jeong et al., 2009
FnL	302	Lipase	Yu et al., 2010
EstP2K	224	Esterase	Ouyang et al., 2013
LipA	277	Lipase	Couto et al., 2010
LipSM54	526	Lipase	Li et al., 2016
MtEst45	516	Acyl esterase	Lee et al., 2016
LipT	329	Esterase	Chow et al., 2012
EstSt7	322	Esterase	Wei et al., 2013
Rlip1	361	Esterase	Liu et al., 2009
EstA	277	Esterase	Chu et al., 2008
FLS12	270	Esterase	Hu et al., 2010
lp_3505	263	Acetyl esterase	Esteban-Torres et al., 2014

Supplementary Table S2 Reference LEs used for constructing phylogentic tree

Accession Nr.	Description	Microbial organism	Lipolytic family
SDT19935	triacylglycerol lipase	*Pseudomonas xinjiangensis*	Family I
AAB53647	extracellular lipase	*Pseudomonas wisconsinensis*	Family I
CAA56780	lipase	*Acinetobacter calcoaceticus*	Family I
AAC15585	lipase	*Pseudomonas fluorescens*	Family I
BAA09135	lipase	*Pseudomonas aeruginosa*	Family I
Q05489	triacylglycerol lipase	*Burkholderia glumae*	Family I
AAC05510	triacylglycerol lipase precursor	*Pseudomonas luteola*	Family I
AAA50466	lipase	*Burkholderia cepacia*	Family I
CAA49812	lipase	*Burkholderia glumae*	Family I
BAA02012	triacylglycerol lipase	*Pseudomonas fluorescens*	Family I
AAD09856	thermostable lipase TliA	*Pseudomonas fluorescens*	Family I
WP_025854682	lipase	*Pseudomonas* sp. CHM02	Family I
AAA22574	lipase	*Bacillus subtilis*	Family I
WP_071578729	triacylglycerol lipase	*Bacillus* sp.	Family I
WP_019713218	triacylglycerol lipase	*Bacillus subtilis*	Family I
WP_059292026	triacylglycerol lipase	*Bacillus malacitensis*	Family I
WP_043021508	triacylglycerol lipase	*Bacillus velezensis*	Family I
AAC12257	lipase	*Geobacillus stearothermophilus*	Family I
AAF40217	lipase	*Geobacillus stearothermophilus*	Family I
CAA64621	triacylglycerol lipase	*Geobacillus thermocatenulatus*	Family I
AAD30278	lipase	*Geobacillus thermoleovorans*	Family I
CAA67627	triacylglycerol lipase	*Propionibacterium acnes*	Family I
AAB71210	lipase LipA	*Streptomyces cinnamoneus*	Family I
WP_036932411	triacylglycerol lipase, partial	*Propionibacterium avidum*	Family I
WP_002519511	triacylglycerol lipase	*Propionibacterium acnes*	Family I
AAA26633	lipase precursor	*Staphylococcus aureus*	Family I
AAF21294	lipase	*Staphylococcus haemolyticus*	Family I
AAC67547	lipase precursor	*Staphylococcus epidermidis*	Family I
AAG35723	lipase precursor	*Staphylococcus warneri*	Family I
WP_042532658	alpha/beta hydrolase	*Oceanobacillus oncorhynchi*	Family II
WP_071394859	alpha/beta hydrolase	*Bacillus tuaregi*	Family II
WP_013842524	alpha/beta hydrolase	*Desulfotomaculum ruminis*	Family II
M86351	triacylglycerol acylhydrolase (lipA)	*Streptomyces exfoliatus*	Family III
WP_030583320	lipase	*Streptomyces globisporus*	Family III
WP_030586638	lipase	*Streptomyces anulatus*	Family III
WP_012381325	alpha/beta hydrolase	*Streptomyces griseus*	Family III
WP_051831627	alpha/beta hydrolase	*Streptomyces violens*	Family III
WP_011196250	alpha/beta hydrolase	*Symbiobacterium thermophilum*	Family IV
ABQ11271	lipase/esterase	uncultured bacterium	Family IV
AAS77236	lipase/esterase	uncultured bacterium	Family IV
AAX37296	lipase/esterase	uncultured bacterium	Family IV
CAA37863	triacylglycerol lipase	*Moraxella* sp.	Family V
WP_002973797	alpha/beta hydrolase	*Leptospira terpstrae*	Family V
WP_002975917	alpha/beta hydrolase	*Leptospira vanthielii*	Family V
CAA47949	triacylglycerol lipase	*Psychrobacter immobilis*	Family V
WP_038681048	3-oxoadipate enol-lactonase	*Rubrobacter radiotolerans*	Family V
WP_007576549	3-oxoadipate enol-lactonase	*Patulibacter medicamentivorans*	Family V
AAC21862	esterase/lipase	*Haemophilus influenzae* Rd KW20	Family V
WP_010945144	alpha/beta fold hydrolase	*Haemophilus ducreyi*	Family V
WP_011147413	esterase	*Photorhabdus luminescens*	Family V

WP_006621073	alpha/beta hydrolase	*Arthrospira platensis*	Family VI
WP_006617627	alpha/beta hydrolase	*Arthrospira platensis*	Family VI
WP_023068329	alpha/beta hydrolase	*Lyngbya aestuarii*	Family VI
OCQ99642	serine esterase	*Oscillatoriales cyanobacterium*	Family VI
WP_015175368	alpha/beta hydrolase	*Oscillatoria nigroviridis*	Family VI
Q01470	Phenmedipham hydrolase	*Pseudarthrobacter oxydans*	Family VII
P37967	Para-nitrobenzyl esterase	*Bacillus subtilis*	Family VII
CAA22794	putative carboxylesterase	*Streptomyces coelicolor*	Family VII
KJJ40755	para-nitrobenzyl esterase	*Bacillus subtilis*	Family VII
WP_063638034	carboxylesterase/lipase family protein	*Bacillus atrophaeus*	Family VII
WP_064730418	carboxylesterase	*Streptomyces parvulus*	Family VII
WP_079426436	carboxylesterase	*Clostridium oryzae*	Family VII
WP_076288838	carboxylesterase	*Paenibacillus pabuli*	Family VII
WP_013412622	carboxylesterase	*Caldicellulosiruptor owensensis*	Family VII
WP_005474626	esterase	*Streptomyces bottropensis*	Family VIII
WP_014985987	esterase	*Nocardia brasiliensis*	Family VIII
WP_015576461	esterase	*Streptomyces* sp.	Family VIII
KXK14623	beta-lactamase	*Chloroflexi bacterium*	Family VIII
ACZ40829	beta-lactamase	*Sphaerobacter thermophilus*	Family VIII
WP_059212061	carboxylesterase	*Streptomyces canus*	Family VIII
WP_040337170	hypothetical protein	*Candidatus Blastococcus massiliensis*	Family VIII
WP_056535681	serine_hydrolase	*Marmoricola* sp.	Family VIII
WP_071051881	serine_hydrolase	*Frankia* sp.	Family VIII
AAK07742	PHB depolymerase PhaZ7 precursor	*Paucimonas lemoignei*	Family IX
WP_056130558	alpha/beta hydrolase	*Lysobacter* sp.	Family IX
WP_052107835	alpha/beta hydrolase	*Lysobacter daejeonensis*	Family IX
WP_012277824	lipase class 2	*Shewanella halifaxensis*	Family IX
WP_012143864	hypothetical protein	*Shewanella sediminis*	Family IX
NP_228147	hypothetical protein TM0336	*Thermotoga maritima*	Family X
WP_041843998	alpha/beta hydrolase	*Thermotoga* sp.	Family X
WP_012057344	DUF3887 domain-containing protein	*Thermosipho melanesiensis*	Family X
WP_052515459	alpha/beta fold hydrolase	*Dethiosulfatarculussandiegensis*	Family X
APW63053	Esterase EstD	*Paludisphaera borealis*	Family X
ABE69172	probable lipase	uncultured bacterium pFosLip	Family XI
GAK59442	probable lipase	*Candidatus Vecturithrix granuli*	Family XI
WP_054029223	lipase family protein	*Desulfatitalea tepidiphila*	Family XI
WP_051945529	lipase family protein	*Verrucomicrobium* sp.	Family XI
GAD03196	lipase	*Agarivorans albus*	Family XI
ACB11220	lipase	uncultured bacterium	Family XII
SFR44003	hypothetical protein SAMN04488073_1267	*Marinobacter gudaonensis*	Family XII
WP_012136589	hypothetical protein	*Marinobacter lipolyticus*	Family XII
WP_007016102	hypothetical protein	*Bermanella marisrubri*	Family XII
WP_046019340	hypothetical protein	*Marinomonas* sp.	Family XII
BAD77330	carboxylesterase	*Geobacillus kaustophilus*	Family XIII
KYD25926	carboxylesterase	*Geobacillus* sp.	Family XIII
WP_081189860	carboxylesterase	*Geobacillus* sp.	Family XIII
WP_044893000	alpha/beta fold hydrolase	*Bacillus alveayuensis*	Family XIII
WP_024030776	alpha/beta fold hydrolase	*Bacillus vireti*	Family XIII
WP_011026365	acetyltransferase	*Caldanaerobacter subterraneus*	Family XIV
WP_049685837	acetyltransferase	*Thermoanaerobacter kivui*	Family XIV
WP_028991973	acetyltransferase	*Thermoanaerobacter thermocopriae*	Family XIV

WP_012995938	acetyltransferase	*Thermoanaerobacter italicus*	Family XIV
WP_041589396	acetyltransferase	*Thermoanaerobacterium thermosaccharolyticum*	Family XIV
ADE28719	EstGK1	uncultured bacterium	Family XV
ADE28720	EstZ3	uncultured bacterium	Family XV
WP_049895542	DUF3089 domain-containing protein	*Oribacterium* sp.	Family XV
AFP50148	esterase	uncultured microorganism	Family XVI
WP_077001818	hypothetical protein	*Variovorax* sp.	Family XVI
WP_030038081	alpha/beta hydrolase	*Streptomyces resistomycificus*	Family XVI
WP_018637498	alpha/beta hydrolase	*Frankia elaeagni*	Family XVI
WP_077040295	alpha/beta hydrolase	*Rhodococcus* sp.	Family XVI
WP_068423891	lipase	*Janibacter terrae*	Family XVII
WP_072625206	lipase	*Janibacter terrae*	Family XVII
WP_068313207	lipase	*Janibacter anophelis*	Family XVII
WP_068264424	lipase	*Janibacter limosus*	Family XVII
WP_007927380	secretory lipase	*Janibacter hoylei*	Family XVII
ANA76126	secretory lipase LipJ2	*Janibacter* sp. R02	Family XVII
WP_066165759	alpha/beta fold hydrolase	*Bacillus* sp. KCTC 13219	Family XVIII
WP_042479182	alpha/beta fold hydrolase	*Bacillus ndiopicus*	Family XVIII
WP_079523184	alpha/beta fold hydrolase	*Solibacillus isronensis*	Family XVIII
ASU50657	esterase	*Ureibacillus thermosphaericus*	Family XVIII
QBK24554	alpha/beta fold hydrolase	*Ureibacillus thermophilus*	Family XVIII
WP_124766500	alpha/beta fold hydrolase	*Lysinibacillus composti*	Family XVIII
WP_126293751	alpha/beta fold hydrolase	*Lysinibacillus telephonicus*	Family XVIII
WP_118877586	alpha/beta fold hydrolase	*Lysinibacillus yapensis*	Family XVIII
ARJ54612	secreted lipase	*Stenotrophomonas maltophilia*	Family XIX
WP_049451625	lipase, partial	*Stenotrophomonas maltophilia*	Family XIX
WP_093999627	lipase	*Stenotrophomonas* sp. CC120222-04	Family XIX
WP_106468945	lipase	*Stenotrophomonas maltophilia*	Family XIX
WP_049430271	lipase	*Stenotrophomonas maltophilia*	Family XIX
WP_057504400	lipase	*Pseudomonas geniculata*	Family XIX
CEO58279	Putative Arylacetamide deacetylase	*Penicillium brasilianum*	Lip10
BAE66557	unnamed protein product	*Aspergillus oryzae*	Lip10
EIT75442	arylacetamide deacetylase	*Aspergillus oryzae*	Lip10
OGE49037	hypothetical protein	*Penicillium arizonense*	Lip10
OJJ29996	hypothetical protein	*Aspergillus wentii*	Lip10
ALN44199	Lip10	*Monascus ruber*	Lip10
ADA70028	lipolytic enzyme precursor	uncultured marine bacterium	EstF
ADH59412	esterase	uncultured bacterium	EstF
ADH59413	esterase	uncultured bacterium	EstF
WP_073290055	alpha/beta hydrolase	*Chryseobacterium polytrichastri*	EstLiu
WP_054509002	alpha/beta hydrolase	*Chryseobacterium sp.*	EstLiu
WP_076503926	esterase	*Chryseobacterium shigense*	EstLiu
WP_084842892	esterase	*Zunongwangia atlantica*	EstLiu
SFC64030	Acetyl esterase/lipase	*Zunongwangia mangrovi*	EstLiu
ADF51938	esterase/lipase-like protein	*Zunongwangia profunda*	EstLiu
WP_022959187	alpha/beta hydrolase	*Spongiibacter tropicus*	EstGS
WP_07599942	hypothetical_protein	*Halioglobus pacificus*	EstGS
OGO52417	hypothetical_protein	*Chloroflexi bacterium*	EstGS
AEM45109	hypothetical protein	uncultured organism	EstGS
WP_011731534	hypothetical protein	*Mycobacterium smegmatis*	EstGH
WP_003898161	hypothetical protein	*Mycobacterium smegmatis*	EstGH
SBS76993	conserved exported hypothetical protein	uncultured *Mycobacterium* sp.	EstGH

WP_071947591	hypothetical protein	*Mycobacterium sp. WY10*	EstGH
AEM45123	hypothetical protein	hypothetical protein	EstGH
AIT16227	putative lipase	uncultured bacterium	EML1
WP_020471493	lipase family protein	*Zavarzinella formosa*	EML1
WP_018291660	lipase family protein	*Verrucomicrobium* sp.	EML1
WP_051926020	lipase family protein	*Leptolyngbya sp.*	EML1
ABB79948	class 3 lipase	uncultured bacterium	EML1
WP_047854544	alpha/beta hydrolase	*Archangium gephyra*	EstL28
EGC22818	alpha/beta domain protein	*Streptococcus sanguinis* SK353	EstL28
AFK29752	esterase	uncultured bacterium	EstL28
WP_082858528	alpha/beta hydrolase	*Planctomyces* sp. SH-PL62	EstJ
WP_015246870	alpha/beta hydrolase	*Singulisphaera acidiphila*	EstJ
OHB78550	esterase partial	*Planctomycetes* bacterium RBG 16 64 10	EstJ
WP_086013385	esterase	*Schlesneria paludicola*	EstJ
WP_076349888	esterase	*Paludisphaera borealis*	EstJ
AFG17170	EstJ	uncultured bacterium	EstJ
WP_060986166	esterase	*Acidovorax delafieldii*	EstY
WP_069103995	esterase	*Acidovorax* sp. RAC01	EstY
OGB09581	esterase	*Burkholderiales bacterium*	EstY
WP_019426758	hypothetical protein	*Limnohabitans sp. Rim28*	EstY
ABY83635	esterase, partial	uncultured bacterium	EstY
OUU14110	poly(3-hydroxybutyrate) depolymerase	*Porticoccaceae bacterium*	FLS18
ACL67852	esterase/lipase	uncultured bacterium FLS18	FLS18
WP_070116203	hypothetical protein	*Pseudohongiella acticola*	FLS18
KRO81080	hypothetical protein	OM182 bacterium BACL3 MAG-120619-bin3	FLS18
ACL67851	esterase/lipase	uncultured bacterium	FLS18
WP_015250293	polyhydroxybutyrate depolymerase	*Singulisphaera acidiphila*	EM3L4
OGO16435	hypothetical protein	*Chloroflexi* bacterium	EM3L4
OGO11912	hypothetical protein	*Chloroflexi* bacterium	EM3L4
OGT27829	hypothetical protein	*Gammaproteobacteria* bacterium	EM3L4
ADH59407	esterase/lipase	uncultured bacterium	EM3L4
WP_019907066	alpha/beta hydrolase	*Methylobacterium* sp. 77	PE10
WP_056144266	alpha/beta hydrolase	*Methylobacterium* sp. Leaf85	PE10
WP_027174325	alpha/beta hydrolase	*Methylobacterium* sp. 10	PE10
WP_046141294	alpha/beta hydrolase	*Devosia soli*	PE10
SDG73374	Acetyl esterase/lipase	*Pelagibacterium luteolum*	PE10
AEV42214	esterase	*Pelagibacterium halotolerans*	PE10
WP_077447883	esterase	*Psychrobacter* sp.	Est12
WP_068405132	esterase	*Psychrobacter* sp.	Est12
WP_011514506	esterase	*Psychrobacter cryohalolentis*	Est12
OEH67893	esterase	*Psychrobacter* sp.	Est12
AGQ21328	esterase	*Psychrobacter* sp.	Est12
WP_033186083	lipase	*Pseudoalteromonas* sp.	Est9x
WP_051275252	alpha/beta hydrolase	*Aestuariibacter salexigens*	Est9x
WP_055012904	lipase	*Pseudoalteromonas* sp.	Est9x
WP_010179535	putative_lipase	*Glaciecola* sp.	Est9x
AFR79233	esterase	Uncultured bacterium	Est9x
YP_765030	esterase	*Rhizobium leguminosarum*	EstA
ZP_02167917	esterase	*Hoeflea phototrophica*	EstA
NP_354481	esterase D	*Agrobacterium fabrum*	EstA
WP_071219778	esterase	*Paenibacillus* sp.	FLS12
ACL67846	esterase	uncultured bacterium FLS12	FLS12
EDV78171	putative esterase	*Geobacillus* sp.	FLS12

WP_024625887	alpha/beta fold hydrolase	*Lactobacillus fabifermentans*	lp_3505
WP_003640069	alpha/beta fold hydrolase	*Lactobacillus pentosus*	lp_3505
YP_004890987	acetyl esterase	*Lactobacillus plantarum*	lp_3505
AFS34518	LipS	uncultured bacterium	LipS
BAD41030	esterase	*Symbiobacterium thermophilum*	LipS
AEN92268	EstGtA2	*Geobacillus thermodenitrificans*	LipS
EAR67363	esterase	*Bacillus* sp.	LipS
EFV74766	esterase	*Bacillus* sp.	LipS
WP_038060347	esterase	*Thermus filiformis*	LipT
WP_003047954	esterase	*Thermus aquaticus*	LipT
WP_038056291	esterase	*Thermus amyloliquefaciens*	LipT
ADW21422	putative esterase	*Thermus scotoductus*	LipT
CZI05393	Patatin	*Legionella pneumophila*	patatin
WP_013131472	Patatin	*Thermobispora bispora*	patatin
WP_036322013	patatin-like phospholipase family protein	*Microbispora* sp.	patatin
WP_071677097	phospholipase	*Thermus brockianus*	patatin
EGK13482.1	patatin family phospholipase	*Desmospora* sp.	patatin
WP_046794161	tannase/feruloyl esterase family alpha/beta hydrolase	*Rhizobium* sp.	tannase
OLB33458	feruloyl esterase	*Acidobacteria* bacterium	tannase
WP_020718491	tannase/feruloyl esterase family alpha/beta hydrolase	*Acidobacteriaceae* bacterium	tannase
WP_043319646	acyl esterase	*Microbulbifer* sp. HZ11	MtEst45
WP_066959386	acyl esterase	*Microbulbifer* sp. Q7	MtEst45
WP_078083231	acyl esterase	*Microbulbifer mangrovi*	MtEst45
WP_073276480	acyl esterase	*Microbulbifer donghaiensis*	MtEst45
WP_074903882	acyl esterase	*Microbulbifer thermotolerans*	MtEst45
AKH15681	acyl esterase	*Microbulbifer thermotolerans*	MtEst45
WP_008099315	esterase	*Verrucomicrobiae bacterium*	Rlip1
WP_079555980	esterase	*Alkalitalea saponilacus*	Rlip1
ACM91105	lipase	uncultured bacterium	Rlip1
WP_067838338	alpha/beta hydrolase	*Mycobacterium mantenii*	Rv0045c
WP_083066055	alpha/beta hydrolase	*Mycobacterium arosiense*	Rv0045c
WP_085163092	alpha/beta hydrolase	*Mycobacterium lacus*	Rv0045c
WP_031683431	alpha/beta hydrolase	*Mycobacterium tuberculosis*	Rv0045c
WP_003400489	alpha/beta hydrolase	*Mycobacterium tuberculosis complex*	Rv0045c
ANQ80463	esterase	uncultured bacterium	Est22
WP_049586806	homoserine O-acetyltransferase	*Alteromonas macleodii*	Est22
WP_118495271	homoserine O-acetyltransferase	*Alteromonas* sp. BL110	Est22
WP_025254722	homoserine O-acetyltransferase	*Alteromonas* sp. ALT199	Est22
WP_005409763	alpha/beta fold hydrolase	*Stenotrophomonas maltophilia*	LipSM54
WP_097047631	alpha/beta fold hydrolase	*Stenotrophomonas* sp. CC120223-11	LipSM54
WP_057499404	alpha/beta fold hydrolase	*Stenotrophomonas maltophilia*	LipSM54
AGF29555	lipase	*Stenotrophomonas maltophilia*	LipSM54
WP_019336523	alpha/beta fold hydrolase	*Stenotrophomonas maltophilia*	LipSM54
WP_010980100	esterase	*Sulfurisphaera tokodaii*	EstSt7
PVU70598	esterase	*Sulfolobus* sp. SCGC AB-777 L09	EstSt7
AAY79544	conserved Archaeal esterase	*Sulfolobus acidocaldarius* DSM 639	EstSt7
WP_069282843	esterase	*Sulfolobus* sp. A20	EstSt7
WP_048099963	esterase	*Candidatus Acidianus copahuensis*	EstSt7
WP_128998162	acetyltransferase	*Fervidobacterium changbaicum*	FnL

ABS61180	conserved hypothetical protein	*Fervidobacterium nodosum* Rt17-B1	FnL
SDG89761	Triacylglycerol esterase/lipase	*Fervidobacterium changbaicum*	FnL
WP_077197788	esterase	*Thermosipho affectus*	FnL
WP_143145293	alpha/beta hydrolase	*Fervidobacterium gondwanense*	FnL
WP_004103732	esterase	*Thermosipho africanus*	FnL
AEH57833	putative lipase	*uncultured microorganism*	EstP2K
WP_092991953	alpha/beta fold hydrolase	*Thiohalomonas denitrificans*	EstP2K
TAJ91530	alpha/beta fold hydrolase	*Gammaproteobacteria* bacterium	EstP2K
RUM94998	alpha/beta hydrolase	*Thiothrix* sp.	EstP2K
RPI52117	alpha/beta fold hydrolase	*Deltaproteobacteria* bacterium	EstP2K
TDJ72546	alpha/beta hydrolase	*Proteobacteria bacterium*	EstP2K
RJQ49102	alpha/beta fold hydrolase	*Gammaproteobacteria* bacterium	EstP2K
ADZ54162	EstGX1	uncultured bacterium	EstGX1
HAM01484	MarR family transcriptional regulator	*Acidimicrobiaceae* bacterium	EstGX1
OYV61387	hypothetical protein B7Z69_00985	*Actinobacteria* bacterium 21-73-9	EstGX1
WP_052604404	MarR family transcriptional regulator	*Acidithrix ferrooxidans*	EstGX1
ANI19854	EstDZ2a	uncultured *Acetothermia* bacterium	EstDZ2a
BAL56305	lipase	uncultured *Acetothermia* bacterium	EstDZ2a
KPJ55944	hypothetical protein AMS16_03300	*Planctomycetes* bacterium DG 58	EstDZ2a
RLS54875	alpha/beta hydrolase	*Planctomycetes* bacterium	EstDZ2a
RLS60676	alpha/beta hydrolase	*Planctomycetes* bacterium	EstDZ2a
HAC89685	alpha/beta hydrolase	*Planctomycetaceae* bacterium	Est903
RLS47763	alpha/beta hydrolase	*Planctomycetes* bacterium	Est903
RLS97429	alpha/beta hydrolase	*Planctomycetes* bacterium	Est903
AXG50964	putative esterase	uncultured bacterium	Est903
AMV32280	Acetylxylan esterase precursor	*Pirellula* sp. SH-Sr6A	Est903

Supplementary Table S3a Lipolytic enzymes downloaded from Uniprot database for validating the profile (part of the table, the full table see the link https://www.biorxiv.org/search/Metagenomic%252Bscreening%252Bfor%252Blipolytic%252Bgenes%252Brevea ls%252Ban%252Becology-clustered%252Bdistribution%252Bpattern)（Dataset 1）

Entry	Protein names	Organism	Super kingdom	Phylum	Order	Length	EC number
A0A009EQT3	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter baumannii 348935	Bacteria	Proteobacteria	Pseudomonadales	323	3.1.1.3
A0A009EUG9	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter baumannii 348935	Bacteria	Proteobacteria	Pseudomonadales	336	3.1.1.3
A0A009FQY3	Esterase TesA (EC 3.1.1.1)	Acinetobacter baumannii 348935	Bacteria	Proteobacteria	Pseudomonadales	209	3.1.1.1
A0A009G8V4	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter baumannii 118362	Bacteria	Proteobacteria	Pseudomonadales	324	3.1.1.3
A0A009GCA4	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter sp. 1289694	Bacteria	Proteobacteria	Pseudomonadales	313	3.1.1.3
A0A009GH23	Esterase TesA (EC 3.1.1.1)	Acinetobacter sp. 1289694	Bacteria	Proteobacteria	Pseudomonadales	207	3.1.1.1
A0A009GYY5	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter sp. 1289694	Bacteria	Proteobacteria	Pseudomonadales	269	3.1.1.3
A0A009HL94	Esterase TesA (EC 3.1.1.1)	Acinetobacter sp. 1295259	Bacteria	Proteobacteria	Pseudomonadales	207	3.1.1.1
A0A009HLH8	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter sp. 1295259	Bacteria	Proteobacteria	Pseudomonadales	313	3.1.1.3
A0A009I9R9	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter baumannii 1295743	Bacteria	Proteobacteria	Pseudomonadales	307	3.1.1.3
A0A009JPY9	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter baumannii 1419130	Bacteria	Proteobacteria	Pseudomonadales	324	3.1.1.3
A0A009L278	Esterase TesA (EC 3.1.1.1)	Acinetobacter baumannii 146457	Bacteria	Proteobacteria	Pseudomonadales	209	3.1.1.1
A0A009LIG6	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter baumannii 146457	Bacteria	Proteobacteria	Pseudomonadales	338	3.1.1.3
A0A009ML24	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter sp. 1475718	Bacteria	Proteobacteria	Pseudomonadales	313	3.1.1.3
A0A009RMH5	Esterase TesA (EC 3.1.1.1)	Acinetobacter sp. 809848	Bacteria	Proteobacteria	Pseudomonadales	207	3.1.1.1
A0A009S415	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter baumannii 951631	Bacteria	Proteobacteria	Pseudomonadales	324	3.1.1.3
A0A009SG56	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter baumannii 99063	Bacteria	Proteobacteria	Pseudomonadales	324	3.1.1.3
A0A009YG53	Esterase TesA (EC 3.1.1.1)	Acinetobacter sp. 479375	Bacteria	Proteobacteria	Pseudomonadales	211	3.1.1.1
A0A009Z6Z7	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter sp. 479375	Bacteria	Proteobacteria	Pseudomonadales	322	3.1.1.3
A0A010IZF7	Esterase TesA (EC 3.1.1.1)	Acinetobacter sp. 1542444	Bacteria	Proteobacteria	Pseudomonadales	207	3.1.1.1
A0A010JDR7	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter sp. 1542444	Bacteria	Proteobacteria	Pseudomonadales	313	3.1.1.3
A0A010JU09	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter sp. 1566109	Bacteria	Proteobacteria	Pseudomonadales	324	3.1.1.3
A0A010L6U0	Esterase TesA (EC 3.1.1.1)	Acinetobacter sp. 1566109	Bacteria	Proteobacteria	Pseudomonadales	207	3.1.1.1
A0A010ULU7	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter sp. 263903-2	Bacteria	Proteobacteria	Pseudomonadales	334	3.1.1.3
A0A010V2A9	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter sp. 263903-2	Bacteria	Proteobacteria	Pseudomonadales	313	3.1.1.3
A0A010WGV4	Esterase TesA (EC 3.1.1.1)	Acinetobacter sp. 263903-2	Bacteria	Proteobacteria	Pseudomonadales	207	3.1.1.1
A0A011IJ62	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter sp. 723929	Bacteria	Proteobacteria	Pseudomonadales	313	3.1.1.3
A0A011JX18	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter sp. 723929	Bacteria	Proteobacteria	Pseudomonadales	269	3.1.1.3
A0A011K3K9	Esterase TesA (EC 3.1.1.1)	Acinetobacter sp. 723929	Bacteria	Proteobacteria	Pseudomonadales	207	3.1.1.1
A0A011M4E2	Esterase TesA (EC 3.1.1.1)	Candidatus Accumulibacter sp. SK-12	Bacteria	Proteobacteria	NA	206	3.1.1.1
A0A011MGM8	Lipase A (EC 3.1.1.3)	Candidatus Accumulibacter sp. SK-12	Bacteria	Proteobacteria	NA	307	3.1.1.3
A0A011MPS8	Esterase TesA (EC 3.1.1.1)	Candidatus Accumulibacter sp. SK-12	Bacteria	Proteobacteria	NA	220	3.1.1.1
A0A011N3B7	Carboxylesterase 2 (EC 3.1.1.1)	Candidatus Accumulibacter sp. BA-92	Bacteria	Proteobacteria	NA	234	3.1.1.1
A0A011NHS5	Esterase EstA (EC 3.1.1.1)	Candidatus Accumulibacter sp. SK-11	Bacteria	Proteobacteria	NA	377	3.1.1.1
A0A011NNZ6	Esterase EstA (EC 3.1.1.1)	Candidatus Accumulibacter sp. SK-12	Bacteria	Proteobacteria	NA	378	3.1.1.1
A0A011NRW2	Esterase TesA (EC 3.1.1.1)	Candidatus Accumulibacter sp. BA-92	Bacteria	Proteobacteria	NA	205	3.1.1.1
A0A011NTI4	Carboxylesterase 2 (EC 3.1.1.1)	Candidatus Accumulibacter sp. SK-12	Bacteria	Proteobacteria	NA	236	3.1.1.1

A0A011NZS4	Esterase EstA (EC 3.1.1.1)	Candidatus Accumulibacter sp. BA-92	Bacteria	Proteobacteria	NA	396	3.1.1.1
A0A011P1E4	Esterase TesA (EC 3.1.1.1)	Candidatus Accumulibacter sp. BA-93	Bacteria	Proteobacteria	NA	205	3.1.1.1
A0A011P216	Esterase TesA (EC 3.1.1.1)	Candidatus Accumulibacter sp. BA-94	Bacteria	Proteobacteria	NA	206	3.1.1.1
A0A011P6G5	Lipase 3 (EC 3.1.1.3)	Candidatus Accumulibacter sp. SK-11	Bacteria	Proteobacteria	NA	295	3.1.1.3
A0A011P911	Esterase TesA (EC 3.1.1.1)	Candidatus Accumulibacter sp. BA-94	Bacteria	Proteobacteria	NA	226	3.1.1.1
A0A011PQC4	Esterase TesA (EC 3.1.1.1)	Candidatus Accumulibacter sp. SK-12	Bacteria	Proteobacteria	NA	206	3.1.1.1
A0A011PQE5	Carboxylesterase 2 (EC 3.1.1.1)	Candidatus Accumulibacter sp. BA-94	Bacteria	Proteobacteria	NA	236	3.1.1.1
A0A011PZA1	Carboxylesterase 2 (EC 3.1.1.1)	Candidatus Accumulibacter sp. BA-94	Bacteria	Proteobacteria	NA	234	3.1.1.1
A0A011Q348	Esterase TesA (EC 3.1.1.1)	Candidatus Accumulibacter sp. SK-11	Bacteria	Proteobacteria	NA	213	3.1.1.1
A0A011Q6T2	Carboxylesterase (EC 3.1.1.1)	Alkalibacterium sp. AK22	Bacteria	Firmicutes	Lactobacillales	207	3.1.1.1
A0A011QDG0	Lipase A (EC 3.1.1.3)	Candidatus Accumulibacter sp. BA-93	Bacteria	Proteobacteria	NA	347	3.1.1.3
A0A011QEA9	Carboxylesterase 2 (EC 3.1.1.1)	Candidatus Accumulibacter sp. BA-93	Bacteria	Proteobacteria	NA	234	3.1.1.1
A0A011QQI8	Lipase (EC 3.1.1.3)	Candidatus Accumulibacter sp. BA-94	Bacteria	Proteobacteria	NA	202	3.1.1.3
A0A011QTT9	Lipase A (EC 3.1.1.3)	Candidatus Accumulibacter sp. BA-92	Bacteria	Proteobacteria	NA	327	3.1.1.3
A0A011RLC2	Carboxylesterase (EC 3.1.1.1)	Alkalibacterium sp. AK22	Bacteria	Firmicutes	Lactobacillales	246	3.1.1.1
A0A013RM51	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter baumannii 573719	Bacteria	Proteobacteria	Pseudomonadales	319	3.1.1.3
A0A013S8D5	Esterase TesA (EC 3.1.1.1)	Acinetobacter sp. 742879	Bacteria	Proteobacteria	Pseudomonadales	207	3.1.1.1
A0A013TE29	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter sp. 742879	Bacteria	Proteobacteria	Pseudomonadales	324	3.1.1.3
A0A013TPS8	Esterase TesA (EC 3.1.1.1)	Acinetobacter sp. 826659	Bacteria	Proteobacteria	Pseudomonadales	207	3.1.1.1
A0A013TXZ3	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter sp. 826659	Bacteria	Proteobacteria	Pseudomonadales	319	3.1.1.3
A0A014B4I3	Esterase TesA (EC 3.1.1.1)	Acinetobacter sp. 25977_8	Bacteria	Proteobacteria	Pseudomonadales	207	3.1.1.1
A0A014BP77	Esterase TesA (EC 3.1.1.1)	Acinetobacter sp. 25977_6	Bacteria	Proteobacteria	Pseudomonadales	207	3.1.1.1
A0A014BQ95	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter sp. 25977_7	Bacteria	Proteobacteria	Pseudomonadales	324	3.1.1.3
A0A014BVI3	Esterase TesA (EC 3.1.1.1)	Acinetobacter sp. 25977_4	Bacteria	Proteobacteria	Pseudomonadales	207	3.1.1.1
A0A014C673	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter sp. 25977_8	Bacteria	Proteobacteria	Pseudomonadales	324	3.1.1.3
A0A014CDI0	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter sp. 25977_3	Bacteria	Proteobacteria	Pseudomonadales	324	3.1.1.3
A0A014CG41	Esterase TesA (EC 3.1.1.1)	Acinetobacter sp. 25977_7	Bacteria	Proteobacteria	Pseudomonadales	207	3.1.1.1
A0A014CNC5	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter sp. 25977_6	Bacteria	Proteobacteria	Pseudomonadales	324	3.1.1.3
A0A014CXJ8	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter sp. 25977_2	Bacteria	Proteobacteria	Pseudomonadales	324	3.1.1.3
A0A014DCG5	Esterase TesA (EC 3.1.1.1)	Acinetobacter sp. 25977_3	Bacteria	Proteobacteria	Pseudomonadales	207	3.1.1.1
A0A014DNG3	Esterase TesA (EC 3.1.1.1)	Acinetobacter sp. 25977_10	Bacteria	Proteobacteria	Pseudomonadales	207	3.1.1.1
A0A014DW77	Esterase TesA (EC 3.1.1.1)	Acinetobacter sp. 25977_1	Bacteria	Proteobacteria	Pseudomonadales	207	3.1.1.1
A0A014E0S6	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter sp. 25977_1	Bacteria	Proteobacteria	Pseudomonadales	324	3.1.1.3
A0A014ELX8	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter sp. 25977_4	Bacteria	Proteobacteria	Pseudomonadales	324	3.1.1.3
A0A014F485	Esterase TesA (EC 3.1.1.1)	Acinetobacter sp. 25977_2	Bacteria	Proteobacteria	Pseudomonadales	207	3.1.1.1
A0A014GFK6	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter sp. 25977_10	Bacteria	Proteobacteria	Pseudomonadales	324	3.1.1.3
A0A017HFW3	Lipase (EC 3.1.1.3)	Limimaricola hongkongensis DSM 17492	Bacteria	Proteobacteria	Rhodobacterales	248	3.1.1.3
A0A021XAE1	Lipolytic enzyme (EC 3.1.1.3)	Shinella sp. DD12	Bacteria	Proteobacteria	Rhizobiales	303	3.1.1.3
A0A021XH55	Carboxylesterase NlhH (EC 3.1.1.1)	Shinella sp. DD12	Bacteria	Proteobacteria	Rhizobiales	311	3.1.1.1
A0A022I9R1	Esterase TesA (EC 3.1.1.1)	Acinetobacter sp. 1000160	Bacteria	Proteobacteria	Pseudomonadales	209	3.1.1.1
A0A022IWK9	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter sp. 1000160	Bacteria	Proteobacteria	Pseudomonadales	338	3.1.1.3
A0A022J7T0	Esterase TesA (EC 3.1.1.1)	Acinetobacter sp. 1564232	Bacteria	Proteobacteria	Pseudomonadales	207	3.1.1.1
A0A022J8K7	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter sp. 1564232	Bacteria	Proteobacteria	Pseudomonadales	319	3.1.1.3
A0A022KIF1	Esterase TesA (EC 3.1.1.1)	Acinetobacter sp. 478810	Bacteria	Proteobacteria	Pseudomonadales	207	3.1.1.1
A0A022KN98	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter sp. 478810	Bacteria	Proteobacteria	Pseudomonadales	269	3.1.1.3
A0A022KNS3	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter sp. 478810	Bacteria	Proteobacteria	Pseudomonadales	313	3.1.1.3
A0A023KQK9	Autotransporter domain-containing protein (Autotransporter outer membrane beta-barrel domain-	Escherichia coli	Bacteria	Proteobacteria	Enterobacterales	232 314 343	3.1.1.3

	containing protein) (Lipase) (EC 3.1.1.3) (Lipase 1) (Outer membrane autotransporter barrel domain protein)					369 441 217	
A0A023M612	EstATII (EC 3.1.1.1)	uncultured bacterium	Bacteria	NA	NA	216	3.1.1.1
A0A023T3X2	Esterase/lipase (EC 3.1.1.1)	Thalassospira sp. GB04J01	Bacteria	Proteobacteria	Rhodospirillales		3.1.1.1
A0A024H3L8	Beta-lactamase family protein (EC 3.1.1.1)	Pseudarthrobacter siccitolerans	Bacteria	Actinobacteria	Micrococcales		3.1.1.1
A0A024H702	Liver carboxylesterase 1 (EC 3.1.1.1)	Pseudarthrobacter siccitolerans	Bacteria	Actinobacteria	Micrococcales		3.1.1.1
A0A024HBB5	Carboxylesterase 1 (EC 3.1.1.1)	Pseudomonas knackmussii (strain DSM 6978 / LMG 23759 / B13)	Bacteria	Proteobacteria	Pseudomonadales		3.1.1.1
A0A024HG51	Chaperone protein clpB (EC 3.1.1.3)	Pseudomonas knackmussii (strain DSM 6978 / LMG 23759 / B13)	Bacteria	Proteobacteria	Pseudomonadales		3.1.1.3
A0A024HK50	Esterase TesA (EC 3.1.1.1)	Pseudomonas knackmussii (strain DSM 6978 / LMG 23759 / B13)	Bacteria	Proteobacteria	Pseudomonadales	201	3.1.1.1
A0A024HQ86	Esterase estA (EC 3.1.1.1)	Pseudomonas knackmussii (strain DSM 6978 / LMG 23759 / B13)	Bacteria	Proteobacteria	Pseudomonadales	638	3.1.1.1
A0A024K9K1	Carboxylesterase (EC 3.1.1.1)	Devosia sp. DBB001	Bacteria	Proteobacteria	Rhizobiales	200	3.1.1.1
A0A031GAI1	Hydrolase, alpha/beta fold family functionally coupled to Phosphoribulokinase (EC 3.1.1.1)	Pseudomonas sp. RIT288	Bacteria	Proteobacteria	Pseudomonadales	332	3.1.1.1
A0A031GCM8	Triacylglycerol lipase (EC 3.1.1.3)	Pseudomonas sp. RIT288	Bacteria	Proteobacteria	Pseudomonadales	562	3.1.1.3
A0A031GJE5	Lipase LipA (EC 3.1.1.3)	Janthinobacterium lividum	Bacteria	Proteobacteria	Burkholderiales	304	3.1.1.3
A0A031ILE0	Carboxylesterase (EC 3.1.1.1)	Pseudomonas sp. RIT357	Bacteria	Proteobacteria	Pseudomonadales		3.1.1.1
A0A031IPG8	Lipase (EC 3.1.1.3)	Pseudomonas sp. RIT357	Bacteria	Proteobacteria	Pseudomonadales		3.1.1.3
A0A031ISI4	Lipase (EC 3.1.1.3)	Pseudomonas sp. RIT357	Bacteria	Proteobacteria	Pseudomonadales	389	3.1.1.3
A0A031IZU6	Carboxylesterase 2 (EC 3.1.1.1)	Pseudomonas sp. RIT357	Bacteria	Proteobacteria	Pseudomonadales	474	3.1.1.1
A0A031J7M0	Hydrolase, alpha/beta fold family functionally coupled to Phosphoribulokinase (EC 3.1.1.1)	Pseudomonas sp. RIT357	Bacteria	Proteobacteria	Pseudomonadales	232	3.1.1.1
A0A037YJP4	Autotransporter outer membrane beta-barrel domain-containing protein (Lipase) (EC 3.1.1.3)	Escherichia coli	Bacteria	Proteobacteria	Enterobacterales	232	3.1.1.3
A0A045H1T3	Carboxylesterase NlhH (Lipase) (EC 3.1.-.-) (EC 3.1.1.3)	Mycobacterium tuberculosis	Bacteria	Actinobacteria	Corynebacteriales	319	3.1.-.-; 3.1.1.3
A0A045IS29	Alpha/beta hydrolase (Lipase) (EC 3.1.-.-; EC 3.1.1.3)	Mycobacterium tuberculosis	Bacteria	Actinobacteria	Corynebacteriales	320	3.1.-.-; 3.1.1.3
A0A045J0I3	Alpha/beta hydrolase (Esterase) (EC 3.1.1.-; EC 3.1.1.3)	Mycobacterium tuberculosis	Bacteria	Actinobacteria	Corynebacteriales	302	3.1.1.-; 3.1.1.3
A0A060L3K4	Lipase (EC 3.1.1.3)	Burkholderia sp. 30(2014)	Bacteria	Proteobacteria	Burkholderiales	364	3.1.1.3
A0A061CR69	Putative 3-oxoadipate enol-lactonase (EC 3.1.1.1)	Pseudomonas pseudoalcaligenes	Bacteria	Proteobacteria	Pseudomonadales	234	3.1.1.1
A0A061CYQ8	Carboxylesterase (EC 3.1.1.1)	Pseudomonas pseudoalcaligenes	Bacteria	Proteobacteria	Pseudomonadales		3.1.1.1
A0A061D1S0	Arylesterase (EC 3.1.2.-) (Lysophospholipase) (EC 3.1.1.1)	Pseudomonas pseudoalcaligenes	Bacteria	Proteobacteria	Pseudomonadales	219 201	3.1.2.-; 3.1.1.1
A0A061D610	Carboxylesterase bioH (EC 3.1.1.1)	Pseudomonas pseudoalcaligenes	Bacteria	Proteobacteria	Pseudomonadales	240	3.1.1.1
A0A061YC10	Alpha/beta hydrolase (Carboxylesterase NlhH) (EC 3.1.1.1)	Escherichia coli	Bacteria	Proteobacteria	Enterobacterales	347 302	3.1.1.1
A0A061YGS5	Alpha/beta hydrolase (Carboxylesterase NlhH) (EC 3.1.1.1) (Lipase)	Escherichia coli	Bacteria	Proteobacteria	Enterobacterales	209 323	3.1.1.1

A0A062BV08	Esterase TesA (EC 3.1.1.1)	Acinetobacter sp. 263903-1	Bacteria	Proteobacteria	Pseudomonadales		3.1.1.1
A0A062BVN1	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter sp. 263903-1	Bacteria	Proteobacteria	Pseudomonadales		3.1.1.3
A0A062C442	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter sp. 263903-1	Bacteria	Proteobacteria	Pseudomonadales		3.1.1.3
A0A062DKE7	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter baumannii 496487	Bacteria	Proteobacteria	Pseudomonadales		3.1.1.3
A0A062FCR0	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter baumannii 754286	Bacteria	Proteobacteria	Pseudomonadales		3.1.1.3
A0A062FNQ7	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter baumannii 940793	Bacteria	Proteobacteria	Pseudomonadales		3.1.1.3
A0A062GEE3	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter sp. 72431	Bacteria	Proteobacteria	Pseudomonadales		3.1.1.3
A0A062GRT9	Esterase TesA (EC 3.1.1.1)	Acinetobacter sp. 72431	Bacteria	Proteobacteria	Pseudomonadales		3.1.1.1
A0A062GV77	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter sp. 72431	Bacteria	Proteobacteria	Pseudomonadales		3.1.1.3
A0A062IBD1	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter baumannii 21072	Bacteria	Proteobacteria	Pseudomonadales		3.1.1.3
A0A062IXY3	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter baumannii 233846	Bacteria	Proteobacteria	Pseudomonadales		3.1.1.3
A0A062K6R8	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter baumannii 1288284	Bacteria	Proteobacteria	Pseudomonadales	239	3.1.1.3
A0A062KD00	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter baumannii 1288284	Bacteria	Proteobacteria	Pseudomonadales	334	3.1.1.3
A0A062KY07	Lactonizing lipase (EC 3.1.1.3)	Acinetobacter baumannii 1571545	Bacteria	Proteobacteria	Pseudomonadales	324	3.1.1.3

Supplementary Table S3b Lipolytic enzymes reported from literatures for validating the profile HMM database (Dataset 2)

Protein name	Length (aa)	Lipolytic family	Accession Nr. to Genebank	Description	Source of microorganis	Reference
499EST	313	IV	AIE44670.1	esterase	*Acidicaldus* sp. USBA-GBX-499	Lopez et al., 2014
7N9	296	IV	AVP72281.1	esterase	uncultured bacterium	Borchert et al., 2017
ABO_1197	318	V	CAL16645.1	carboxylic ester hydrolase	*Alcanivorax borkumensis* SK2	Schneiker et al., 2006
ABO_1251	505	VII	CAL16699.1	carboxylesterase	*Alcanivorax borkumensis* SK2	
ABO0195	319	IV	WP_011587492.1	alpha/beta hydrolase	*Alcanivorax*	
AcXE2	214	II	ACM59679.1	lipolytic protein GDSL family	*Caldicellulosiruptor bescii* DSM 6725	Soni et al., 2017
CaesCCR11	250	XIII	AHL67377.1	carboxylesterase, partial	*Geobacillus thermoleovorans*	Espinosa-Luna et al., 2016
E69	274	IV	AUD08548.1	esterase	*Erythrobacter seohaensis*	Huo et al., 2017
EaEST	272	V	WP_014970431.1	alpha/beta hydrolase	*Exiguobacterium antarcticum*	Lee et zl., 2017
EM3L2	404	XV	ADH59405	esterase	uncultured bacterium	Jeon et al., 2011
EM3L4	330	EM3L4	ADH59407	esterase/lipase	uncultured bacterium	Jeon et al., 2011
EML1	304	XI	ABB79948	class 3 lipase	uncultured bacterium	Jeong et al., 2009
Est_p1	296	V	ACF33459.1	putative lipase/esterase	uncultured bacterium	Peng et al., 2011
Est01	397	VIII	HQ444406	cold-adapted esterase	uncultured bacterium	Cheng et al., 2014
Est1	270	V	KR149567	esterase	uncultured bacterium	Lu et al., 2019
Est10	223	VI	AGD81840.1	esterase	*Psychrobacter pacificensis*	Rodríguez et al., 2015
Est12	329	IV	AGQ21328.1	esterase	*Psychrobacter* sp.	Wu et al., (2013)
Est2	389	VIII	KR149568	esterase	uncultured bacterium	Lu et al., 2019
Est22	478	Est22	ANQ80463.1	esterase	uncultured bacterium	Li et al., 2017
Est22_2	423	VIII	KF052088	Est22	uncultured bacterium	Mokoena et al., 2013
Est25	340	IV	AII23256.1	esterase	uncultured bacterium	Li et al., 2014
Est30	247	XIII	AAN81911.1	thermostable carboxylesterase Est30	*Geobacillus stearothermophilus*	Ewis et al., 2004
Est4	316	V	ALE67003.1	esterase	uncultured bacterium	Gao et al., 2016
Est40	299	IV	5GMR	esterase	uncultured bacterium	Li et al., 2017
Est56	287	IV	KR149569	esterase	uncultured bacterium	our previous research
Est700	208	I	AQX17346.1	esterase	Bacillus licheniformis	Zhang et al., 2018
Est7K	411	VIII	KP756684	family VIII esterase	uncultured bacterium	Lee et al., 2016
Est903	300	IV	AXG50964.1	putative esterase	uncultured bacterium	Jia et al., 2019
Est97	247	XVI	AFP50148	esterase	uncultured microorganism	Fu et al., 2013
Est9x	294	Est9X	AFR79233	esterase	uncultured bacterium	Jeong et al., 2009
EstA	277	EstA	ABY60416.1	esterase	uncultured bacterium	Chu et al., 2008
EstA3	414	XIV	WP_011026365.1	acetyltransferase	*Caldanaerobacter subterraneus*	Lee et al., 2016
EstA3_2	396	VIII	DQ022078	beta-lactamase	uncultured bacterium	Rao et al., 2011
EstAM	314	IV	ACF04196.1	lipase/esterase	uncultured bacterium	Rashamuse et al., 2009
EstATII	314	IV	AHN10469.1	esterase	uncultured bacterium	Mohamed et al., 2013
EstBL	398	VIII	AAX78516.1	EstBL	Burkholderia cepacia	Rashamuse et al., 2007
EstC	427	VIII	ACH88047.1	carboxylesterase	uncultured bacterium	Rashamuse et al., 2009

EstCE1	378	VIII	DQ022079	beta-lactamase	uncultured bacterium	Elend et al., 2006
EstCS2	570	VII	ADB22436.1	carboxylesterase	uncultured bacterium	Kang et al., 2011
EstD	412	X	NP_228147	hypothetical protein TM0336	*Thermotoga maritima*	Nelson et al., 1999
EstD2	397	XV	GQ866023.1	EstD2	uncultured bacterium	Lee et al., 2010
EstDZ2	271	EstDZ2	ANI19854	EstDZ2a	uncultured *Acetothermia* bacterium	Zarafeta et al., 2016
EstEP16	249	XIII	AGM38158.1	esterase	uncultured bacterium	Zhu et al., 2013
EstF27	290	IV	ADU32684.1	esterase	uncultured bacterium	Sang et al., 2011
EstF4K	396	VIII	JN001202	putative beta-lactamase class C	uncultured bacterium	Ouyang et al., 2013
EstGH	448	EstGH	AEM45123	esterase	uncultured bacterium	Nacke et al., 2011
EstGK1	322	XV	ADE28719	esterase	uncultured bacterium	Bayer et al., 2010
EstGS	397	EstGS	AEM45109	esterase	uncultured organism	Nacke et al., 2011
EstGtA2	249	XIII	AEN92268.1	esterase	*Geobacillus thermodenitrificans*	Charbonneau et al., 2010
EstGX1	210	EstGX1	ADZ54162.1	EstGX1	uncultured bacterium	Jime´nez et al., 2012
EstHE1	223	II	BAH03944.1	GDSL family of esterase	uncultured marine bacterium	Okamura et al., 2010
EstJ	317	IV	AFG17170	esterase	uncultured bacterium	Choi et al., 2013
EstKT4	352	IV	ADH59412	esterase	uncultured bacterium	
EstKT7	316	IV	ADH59413	esterase	uncultured bacterium	Jeon et al., 2012
EstKT9	372	IV	ADH59414	esterase	uncultured bacterium	
EstL28	290	EstL28	AFK29752	esterase	uncultured bacterium	Seo et al., 2013
EstLiu	293	IV	ADF51938.1	esterase/lipase-like protein	*Zunongwangia profunda* SM-87	Rahman et al., 2016
EstM-N1	395	VIII	HQ154132	EstM-N1	uncultured bacterium	Yu et al., 2011
EstM-N2	407	VIII	HQ154133	EstM-N2	uncultured bacterium	
EstMY	360	IV	ADM67447.1	esterase MY09-1	uncultured bacterium	Li et al., 2010
EstMY09-2	291	IV	ADM67446.1	esterase MY09-2	uncultured bacterium	Li et al., 2010
EstN7	320	IV	AUV46828.1	lipase	*Bacillus cohnii*	Noby et al., 2018
EstOF4	261	XIII	AGK06467.1	Esterase/lipase	uncultured bacterium	Rao et al., 2013
Est-OKK	269	V	AWD93592.1	putative esterase	uncultured bacterium	Yang et al., 2018
EstP2K	224	I	JN001203.1	putative lipase	uncultured microorganism	Ouyang et al., 2013
EstQE	382	VIII	AOO35454.1	esterase	*Ochrobactrum* sp.	Zhang et al., 2016
EstS	302	IV	AMQ65607.1	esterase	*Serratia* sp. HZ15	Jiang et al., 2016
EstSL3	211	II	AMO51591.1	SGNH esterase	*Alkalibacterium* sp. SL3	Wang et al., 2016
EstSP	354	IV	ARS65737.1	esterase	uncultured bacterium	Jayanath et al., 2018
EstSt7	322	EstSt7	WP_010980100.1	esterase	*Sulfurisphaera tokodaii*	Wei et al., 2013
EstSTR1	390	VIII	AJE68931.1	carboxylesterase	uncultured bacterium	Jeon et al., 2016
EstU1	426	VIII	JF791800	carboxylesterase	uncultured bacterium	Jeon et al., 2011
EstUT1	248	XVIII	ASU50657.1	esterase	*Ureibacillus thermosphaericus*	Samoylova et al., 2018
EstWSD	383	XV	AFY63009	esterase	uncultured bacterium	Wang et al., 2013
Est-XG2	501	VII	AGS38342.1	esterase	uncultured bacterium	Shao et al., 2013
EstY	423	EstY	ABY83635	esterase	uncultured bacterium	Wu & Sun, 2009
EstZ3	318	XV	ADE28720	esterase	uncultured bacterium	Bayer et al., 2010
FLS12	270	FLS12	ACL67846.1	esterase	uncultured bacterium	Hu et al., 2010
FLS18	259	FLS18	ACL67851.1	esterase	uncultured bacterium	Hu et al., 2010
FnL	302	FnL	ABS61180.1	conserved hypothetical protein	*Fervidobacterium nodosum* Rt17-B1	Yu et al., 2010
GDEst-95	498	VII	ANG09062.1	carboxylesterase	*Geobacillus stearothermophilus*	Gudiukaite et al., 2017
H8	305	V	ARH02619.1	esterase	uncultured bacterium	Zhang et al., 2017
H9Est	356	IV	AHK13303.1	esterase	uncultured bacterium	Santi et al., 2015

LAE1	272	VI	SRA059294	lipase	uncultured bacterium	
LAE2	324	II	SRA059294	lipase	uncultured bacterium	
LAE3	300	IV	SRA059294	lipase	uncultured bacterium	
LAE4	309	IV	SRA059294	lipase	uncultured bacterium	Martínez-Martínez et al., 2013
LAE5	217	II	SRA059294	lipase	uncultured bacterium	
LAE6	315	IV	SRA059294	lipase	uncultured bacterium	
LAE7	310	II	SRA059294	lipase	uncultured bacterium	
Lip10	348	IV	ALN44199.1	lipase	*Monascus ruber*	Guo et al., 2016
Lip-1452	483	IV	ADB11056.1	lipase	*Psychrobacter* sp. G	Lin et al., 2010
Lip3	280	XI	AHZ89331.1	class 3 lipase	uncultured bacterium	De Santi et al., 2016
Lip-948	315	V	ADB11055.1	lipase	*Psychrobacter* sp. G	Lin et al., 2010
LipA	277	I	ACJ13070.1	lipase	uncultured bacterium	Couto et al., 2010
LipA9	404	VIII	AYH52116.1	lipase	*Marinobacter lipolyticus*	Park et al., (2018)
Lipab15	272	IV	AID66450.1	lipase	*Halomonas elongata*	
Lipab18	272	IV	AID66448.1	lipase	*Halomonas eurihalina*	
Lipab4	272	IV	AID66447.1	lipase	*Halomonas elongata*	Asyari et al., 2015
Lipab8	272	IV	AID66449.1	lipase	*Chromohalobacter japonicus*	
Lipag18	272	IV	AID66446.1	lipase	*Halomonas meridiana*	
LipBL	404	VIII	CBX87546.1	lipolytic enzyme	*Marinobacter lipolyticus*	Pérez et al., 2012
LipC	327	IV	AAV45166.1	lipase/esterase	*Haloarcula marismortui* ATCC 43049	Baliga et al., 2004
LipC12	293	I	AEK97793.1	lipase	uncultured bacterium	Glogauer et al., 2011
LipEH166	381	XII	ACB11220.1	lipase	uncultured bacterium	Kim et al., 2009
LipG	300	XI	ABE69172.1	probable lipase	uncultured bacterium	Lee et al., 2006
LipJ2	423	XVII	ANA76126.1	secretory lipase	*Janibacter* sp. R02	Castilla et al., (2017)
Lipo1	308	IV	ADC79133.1	lipase/esterase	uncultured sludge bacterium	
Lipo10	432	unclassified	ADC79144.1	lipase/esterase/thioesterase family lipase	uncultured sludge bacterium	
Lipo11	291	IV	ADC79145.1	lipase/esterase	uncultured sludge bacterium	
Lipo12A	291	IV	ADC79146.1	lipase/esterase	uncultured sludge bacterium	
Lipo12B	506	VIII	ADC79147.1	lipase/esterase	uncultured sludge bacterium	
Lipo13	441	II	ADC79148.1	hypothetical protein	uncultured sludge bacterium	
Lipo2	282	I	ADC79134.1	hypothetical protein	uncultured sludge bacterium	
Lipo3	253	V	ADC79135.1	lipase/esterase	uncultured sludge bacterium	
Lipo4A	226	VI	ADC79136.1	lipase/esterase	uncultured sludge bacterium	Liaw et al., (2010)
Lipo4B	293	patatin-like-protein	ADC79137.1	patatin-like phospholipase	uncultured sludge bacterium	
Lipo4C	292	patatin-like-protein	ADC79138.1	patatin-like phospholipase	uncultured sludge bacterium	
Lipo5	348	IV	ADC79139.1	lipase/esterase	uncultured sludge bacterium	
Lipo6	242	VI	ADC79140.1	lipase/esterase	uncultured sludge bacterium	
Lipo7A	471	unclassified	ADC79141.1	lipase/esterase	uncultured sludge bacterium	
Lipo7B	281	unclassified	ADC79142.1	hypothetical protein	uncultured sludge bacterium	
Lipo8	208	II	ADC79143.1	arylesterase	uncultured sludge bacterium	
LipR	449	X	CCC86601.1	triacylglycerol hydrolase	*Rhodococcus* sp.	Bassegoda et al., 2012
LipR_2	467	X	WP_011331609.1	triacylglycerol lipase	*Rhodococcus erythropolis*	Bassegoda et al., 2012
LipR1	417	I	AEL99900.1	lipase	uncultured bacterium	Kumar et al., 2013
LipR3	417	I	AFV46383.1	lipase	uncultured bacterium	Kumar et al., 2013
LipS	280	XIII	AFS34518.1	esterase/lipase	uncultured bacterium	Chow et al., 2012
LipSm	400	XIX	KX353755.1	secreted lipase	*Stenotrophomonas maltophilia*	Parapouli et al., 2018
LipSM54	271	LipSM54	AGF29555	lipase	*Stenotrophomonas maltophilia*	Li et al., 2016

LipT	329	LipT	AFS34517.1	esterase/lipase	uncultured bacterium	Chow et al., 2012
LipYY31	470	I	BAK52029.2	lipase	*Pseudomonas* sp. YY31	Yamashiro et al., 2013
lp_3505	263	lp_3505	YP_004890987.1	acetyl esterase	*Lactobacillus plantarum*	Esteban-Torres et al., 2014
Lpc53E1	387	VIII	AFM09717.1	lipase	uncultured bacterium	Selvin et al., 2012
M37	340	XI	AY527197.1	lipase	*Photobacterium* sp. M37	Ryu et al., 2006
MGS0010	414	VIII	AHG30919.1	esterase	gamma proteobacterium	
MGS0018	277	IV	AGT96416.1	esterase	uncultured bacterium	Tchigvintsev et al., 2015
MGS0105	390	VIII	AGT96414.1	esterase	uncultured bacterium	
MGS0153	318	V	WP_011588480.1	esterase	*Alcanivorax*	
MGS-B1	345	IV	KF831420.1	carboxylesterase	*Tenericutes* bacterium	
MGS-K1	514	VII	KF831421.1	carboxylesterase	uncultured bacterium	
MGS-M1	239	IV	KF831414.1	carboxylesterase	*Firmicutes* bacterium	
MGS-M2	276	V	KF831415.1	carboxylesterase	*Firmicutes* bacterium	
MGS-MG1	261	unclassified	KF831418.1	carboxylesterase	*Geobacillus* sp.	Alcaide et al., 2015
MGS-MT1	348	IV	KF831419.1	carboxylesterase	uncultured gamma proteobacterium	
MGS-RG1	222	VI	KF831416.1	carboxylesterase	uncultured gamma proteobacterium	
MGS-RG2	225	VI	KF831417.1	carboxylesterase	uncultured gamma proteobacterium	
MGS-RG3	277	IV	KC986402.1	esterase	uncultured bacterium	
MtEst45	516	MtEst45	AKH15681.1	acyl esterase	*Microbulbifer thermotolerans*	Lee et al., 2016
PE10	279	IV	AEV42214.1	esterase	*Pelagibacterium halotolerans*	Jiang et al., 2012
PhaZ7	380	IX	AAK07742.1	PHB depolymerase PhaZ7 precursor	*Paucimonas lemoignei*	Handrick et al., 2001
PLP	302	patatin-like-protein	AIT11629.1	patatin-like phospholipase family protein	uncultured bacterium	Fu et al., 2015
PMGL2	343	IV	AMR72657.1	esterase	uncultured bacterium	Petrovskaya et al., 2016
REst1	342	IV	CBN72524.1	esterase protein	*Rheinheimera* sp. Chandigarh	Virk et al., 2011
Rlip1	316	Rlip1	ACM91047.1	esterase	uncultured bacterium	Liu et al., 2009
Rv0045c	298	Rv0045c	NP_214559.1	hydrolase	*Mycobacterium tuberculosis* H37Rv	Guo et al., 2010
SAestA	324	IV	AIY29984.1	esterase	*Salinispora arenicola*	Fang et al., 2015
SBLip1	445	VIII	AFK83589.1	lipolytic enzyme	uncultured bacterium	
SBLip2	346	IV	AFK83595.1	lipolytic enzyme	uncultured bacterium	Biver et al., 2013
SBLip5.1	316	IV	AFK83603.1	lipolytic enzyme	uncultured bacterium	
ThaEst2349	343	IV	AHX83345.1	esterase/lipase	*Thalassospira* sp.	De Santi et al., 2016
Vlip509	338	V	ABS72371.1	alpha/beta hydrolases	*Vibrio* sp. GMD509	Park et al., (2007)

Supplementary Table S3c. Protein sequences predicted from lipolytic inserts identified by function-driven screening (this study) for validating the profile HMM database ((part of the table, the full table see the link https://www.biorxiv.org/search/Metagenomic%252Bscreening%252Bfor%252Blipolytic%252Bgenes%252Breveals%252Ban%252Becology-clustered%252Bdistribution%252Bpattern) (Dataset 3)

Gene name	Protein length (aa)	Strand	Start (bp)	End (bp)	Corresonding plasmid name	Lipolytic enzyme [a]
gene_1	143	+	1	432	pE55-2	N
gene_2	284	+	434	1288	pE55-2	EstC55-2
gene_3	165	-	1320	1817	pE55-2	N
gene_4	57	+	3	176	pE55-3	N
gene_5	513	+	267	1808	pE55-3	EstC55-3
gene_6	232	+	1895	2593	pE55-3	N
gene_7	142	+	2647	3075	pE55-3	N
gene_8	413	-	3098	4339	pE55-3	N
gene_9	351	-	4435	5490	pE55-3	N
gene_10	172	-	5567	6085	pE55-3	N
gene_11	163	-	6195	6686	pE55-3	N
gene_12	355	-	6761	7828	pE55-3	N
gene_13	163	-	7825	8316	pE55-3	N
gene_14	94	-	8313	8597	pE55-3	N
gene_15	271	+	2	817	pE55-4	N
gene_16	271	+	836	1651	pE55-4	EstC55-4_1
gene_17	390	+	1783	2955	pE55-4	EstC55-4_2
gene_18	131	+	3051	3446	pE55-4	N
gene_19	136	+	3560	3970	pE55-4	N
gene_20	153	-	4055	4516	pE55-4	N
gene_21	487	+	4557	6020	pE55-4	N
gene_22	230	+	6017	6709	pE55-4	N
gene_23	325	+	6804	7781	pE55-4	N
gene_24	109	+	7778	8104	pE55-4	N
gene_25	58	+	3	179	pE55-5	N
gene_26	310	+	265	1197	pE55-5	EstC55-5
gene_27	485	-	1267	2724	pE55-5	N
gene_28	208	+	3025	3651	pE55-5	N
gene_29	162	+	3749	4234	pE55-5	N
gene_30	319	+	132	1091	pE55-6	EstC55-6
gene_31	407	-	1102	2325	pE55-6	N
gene_32	160	+	2399	2881	pE55-6	N
gene_33	164	+	2926	3420	pE55-6	N
gene_34	316	+	3515	4465	pE55-6	N
gene_35	258	+	4462	5238	pE55-6	N
gene_36	126	+	5235	5612	pE55-6	N
gene_37	109	-	3	329	pE55-7	N
gene_38	340	-	335	1357	pE55-7	N
gene_39	369	+	1607	2716	pE55-7	N
gene_40	430	+	2947	4239	pE55-7	EstC55-7
gene_41	375	-	4331	5458	pE55-7	N

gene_42	118	+	5907	6263	pE55-7	N
gene_43	368	-	7020	8126	pE55-7	N
gene_44	50	-	8126	8278	pE55-7	N
gene_45	105	+	8659	8976	pE55-7	N
gene_46	80	-	9108	9350	pE55-7	N
gene_47	203	+	44	655	pE55-8	N
gene_48	318	-	669	1625	pE55-8	EstC55-8_1
gene_49	246	-	1646	2386	pE55-8	EstC55-8_2
gene_50	136	+	2	412	pE55-10	N
gene_51	296	+	492	1382	pE55-10	N
gene_52	121	+	1405	1770	pE55-10	N
gene_53	460	+	1880	3262	pE55-10	N
gene_54	379	+	4416	5555	pE55-10	N
gene_55	279	-	5464	6303	pE55-10	EstC55-10
gene_56	648	+	6510	8453	pE55-10	N
gene_57	143	+	3	434	pE55-12	N
gene_58	293	+	545	1426	pE55-12	N
gene_59	212	+	1426	2064	pE55-12	N
gene_60	384	+	2057	3211	pE55-12	EstC55-12
gene_61	157	+	3240	3713	pE55-12	N
gene_62	230	+	3703	4395	pE55-12	N
gene_63	148	+	4392	4838	pE55-12	N
gene_64	362	+	4831	5919	pE55-12	N
gene_65	41	+	5916	6038	pE55-12	N
gene_66	975	+	3	2930	pE55-13	N
gene_67	375	-	2911	4038	pE55-13	N
gene_68	355	-	4049	5116	pE55-13	EstC55-13
gene_69	110	+	5208	5540	pE55-13	N
gene_70	261	+	1	786	pE55-15	EstC55-15
gene_71	197	+	891	1484	pE55-15	N
gene_72	48	+	1543	1686	pE55-15	N
gene_73	267	+	1	804	pE55-18	EstC55-18

[a] N, Nonlipolytic proteins

Supplementary Table S3d Non-lipolytic proteins showing high homology to LEs were retrieved from uniprot or NCBI database for validating the profile HMM databas (Dataset 4)

Entry (accession Nr)	Database	Protein names	Length (aa)
L7VEQ3	uniprot	2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoate hydrolase BphD (EC 3.7.1.8)	295
A0A0U0R506	uniprot	2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoate hydrolase BphD (EC 3.7.1.8)	294
A0A2X1S8R4	uniprot	2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoate hydrolase BphD (EC 3.7.1.8)	290
A0A0T9VGE3	uniprot	2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoate hydrolase BphD (EC 3.7.1.8)	295
F8GQY4	uniprot	2,6-dioxo-6-phenylhexa-3-enoate hydrolase BphD (EC 3.7.1.8)	268
A0A375GEC5	uniprot	2,6-dioxo-6-phenylhexa-3-enoate hydrolase BphD (EC 3.7.1.8)	263
A0A1S7N7P8	uniprot	Non-heme haloperoxidase (EC 1.11.1.10)	326
A0A1S7T7D7	uniprot	Non-heme haloperoxidase (EC 1.11.1.10)	326
A0A1S7T6S9	uniprot	Non-heme haloperoxidase (EC 1.11.1.10)	323
A0A1S7N8N9	uniprot	Non-heme haloperoxidase (EC 1.11.1.10)	323
A0A1S7N8S2	uniprot	Non-heme haloperoxidase (EC 1.11.1.10)	326
A0A1S7S3Y6	uniprot	Non-heme haloperoxidase (EC 1.11.1.10)	333
A0A1S7NDN6	uniprot	Non-heme haloperoxidase (EC 1.11.1.10)	325
A0A2Z5YBG9	uniprot	Non-heme chloroperoxidase (EC 1.11.1.10) (Non-heme haloperoxidase Hpx)	326
A0A2N9APQ2	uniprot	Arylesterase (Alpha/beta hydrolase fold), putative haloperoxidase (EC 1.11.1.10) (EC 3.1.1.2)	273
G4SWJ2	uniprot	Putative non-heme haloperoxidase, alpha/beta hydrolase fold (EC 1.11.1.10)	273
A0A0N7H8D0	uniprot	Alpha/beta hydrolase (Haloperoxidase) (Non-heme haloperoxidase Hpx) (EC 1.11.1.10)	308
A0A102S299	uniprot	Alpha/beta hydrolase (EC 1.11.1.10) (Haloperoxidase)	278
A0A375DNQ7	uniprot	Non-heme haloperoxidase (EC 1.11.1.-) (EC 1.11.1.10)	274
A0A375EC98	uniprot	Non-heme haloperoxidase (EC 1.11.1.-) (EC 1.11.1.10)	274
A0A375FEC6	uniprot	Non-heme haloperoxidase (EC 1.11.1.-) (EC 1.11.1.10)	274
A0A375D6M7	uniprot	Non-heme haloperoxidase (EC 1.11.1.-) (EC 1.11.1.10)	274
A0A375IUU4	uniprot	Non-heme haloperoxidase (EC 1.11.1.-) (EC 1.11.1.10)	269
A0A1S7RCN8	uniprot	Non-heme haloperoxidase (EC 1.11.1.10)	326
A0A1S7R0G9	uniprot	Non-heme haloperoxidase (EC 1.11.1.10)	278
A0A1S7RQ23	uniprot	Non-heme haloperoxidase (EC 1.11.1.10)	278
A0A1S7N1P6	uniprot	Non-heme haloperoxidase (EC 1.11.1.10)	278
A0A1S7R2U2	uniprot	Non-heme haloperoxidase (EC 1.11.1.10)	278
A0A1S7R186	uniprot	Non-heme haloperoxidase (EC 1.11.1.10)	278
A0A1S7S3E0	uniprot	Non-heme haloperoxidase (EC 1.11.1.10)	278
A0A1S7TAY0	uniprot	Non-heme haloperoxidase (EC 1.11.1.10)	278
A0A379M130	uniprot	Non-heme haloperoxidase (EC 1.11.1.10)	299
A0A1S7TZP1	uniprot	Non-heme haloperoxidase (EC 1.11.1.10)	278
A0A375DP21	uniprot	Putative Non-heme haloperoxidase, putative Alpha/beta hydrolase domain (EC 1.11.1.10)	276
A0A1S7RZV4	uniprot	Non-heme haloperoxidase (EC 1.11.1.10)	278
A0A1S7QF97	uniprot	Non-heme haloperoxidase (EC 1.11.1.10)	326
A0A1S7QYM3	uniprot	Non-heme haloperoxidase (EC 1.11.1.10)	280
A0A1S7UBZ3	uniprot	Non-heme haloperoxidase (EC 1.11.1.10)	278

A0A1X2DNB5	uniprot	Non-heme haloperoxidase Hpx (EC 1.11.1.10)	327
A0A1S7R6I3	uniprot	Non-heme haloperoxidase (EC 1.11.1.10)	278
A0A2X4UF21	uniprot	Non-heme haloperoxidase (EC 1.11.1.10)	279
A0A1S7N0Y4	uniprot	Non-heme haloperoxidase (EC 1.11.1.10)	278
A0A1S7RYV9	uniprot	Non-heme haloperoxidase (EC 1.11.1.10)	278
A0A1S7SHR8	uniprot	Non-heme haloperoxidase (EC 1.11.1.10)	325
A0A378W1U8	uniprot	Non-heme haloperoxidase Hpx (EC 1.11.1.10)	308
A0A1S7S3D1	uniprot	Non-heme haloperoxidase (EC 1.11.1.10)	278
A0A1S7MUS0	uniprot	Non-heme haloperoxidase (EC 1.11.1.10)	325
A0A1S7MPM6	uniprot	Non-heme haloperoxidase (EC 1.11.1.10)	278
A0A375EBJ1	uniprot	Putative Non-heme haloperoxidase, putative Alpha/beta hydrolase domain (EC 1.11.1.10)	276
F8JH88	uniprot	Non-heme haloperoxidase (EC 1.11.1.10)	279
F8JA98	uniprot	Non-heme haloperoxidase (EC 1.11.1.10)	324
A0A2N9AMF4	uniprot	Non-heme haloperoxidase (EC 1.11.1.10)	278
SFQ96535.1	NCBI	epoxide hydrolase	358
AWN43677.1	NCBI	epoxide hydrolase	384
ERK71245.1	NCBI	epoxide hydrolase	377
ALO47455.1	NCBI	epoxide hydrolase	410
ALO91291.1	NCBI	epoxide hydrolase	373
ALO98459.1	NCBI	epoxide hydrolase	411
PNG93099.1	NCBI	epoxide hydrolase	374
PNG95674.1	NCBI	epoxide hydrolase	391
PNG97285.1	NCBI	epoxide hydrolase	390
KDN82212.1	NCBI	epoxide hydrolase	372
KEP41268.1	NCBI	epoxide hydrolase	375
AGI60154.1	NCBI	epoxide hydrolase	388
AIT98391.1	NCBI	epoxide hydrolase	383
KIE51135.1	NCBI	epoxide hydrolase	329
KIE52625.1	NCBI	epoxide hydrolase	376
ART90207.1	NCBI	epoxide hydrolase	273

Supplementary Table S4 Functional families in CATH database used to identify LEs affiliate to family II, VIII

Lipolytic family [a]	CATH superfamily	Functional family	Description
II	Autotransporter beta-domain superfamily (2.40.128.130)	FunFam 4295	Autotransporting lipase, GDSL family
	SGNH hydrolase superfamily (3.40.50.1110)	FunFam 15013	Isoamyl acetate-hydrolyzing esterase 1 homolog
		FunFam 15022	Acyl-CoA thioesterase I
		FunFam 15004	GDSL esterase/lipase APG
		FunFam 15024	Putative secreted hydrolase
		FunFam 14962	Phospholipase B1, membrane-associated
		FunFam 15007	Autotransporting lipase, GDSL family
VIII [b]	DD-peptidase/beta-lactamase superfamily (3.40.710.10)	FunFam 22194	Beta-lactamase (Penicillin-binding protein) (Penicillinase)
		FunFam 22219	Extended spectrum beta-lactamase GES-11
		FunFam 4880	Low-affinity penicillin-binding protein 2X
		FunFam 22208	Class C beta-lactamase CMY-10
patatin-like-protein	Cytosolic phospholipase A2 catalytic domain superfamily (3.40.1090.10)	FunFam 5858	YMR313Cp-like protein
		FunFam 6563	Patatin group D-2
		FunFam 6569	Probable inactive patatin-like protein 9
		FunFam 6582	Patatin-like phospholipase domain-containing protein 2
		FunFam 6604	Patatin-like phospholipase domain-containing protein
	Alpha/Beta hydrolase fold, catalytic domain (3.40.50.1820)	FunFam 115466	Serine hydrolase-like protein 2
		FunFam 115488	Triacylglycerol lipase, pancreatic, putative
		FunFam 115552	Hydrolase, alpha/beta fold family
		FunFam 115703	Putative methylesterase 13, chloroplastic

[a] Proteins with no CATH hits were grouped as putative family II, VIII, PLP candidates, respectively.

[b] only proteins with sequence length between 350 and 450 aa were retrieved.

Supplementary Table S5 Description of assembled metagenomes downloaded from Integrated Microbial Genomes database (IMG)

Sample name	Habitat	Accession Nr in IMG	Assembled genome size (bp)	Assembled gene count
ADAS1	Anaerobic digester active sludge	3300009676	896231486	1557165
ADAS2	Anaerobic digester active sludge	3300009776	910918723	1761190
ADAS3	Anaerobic digester active sludge	3300025629	440662511	734719
ADAS4	Anaerobic digester active sludge	3300025867	804659355	1483826
ADAS5	Anaerobic digester active sludge	3300009655	593528173	1062322
ADAS6	Anaerobic digester active sludge	3300025861	767290509	1440611
ADAS7	Anaerobic digester active sludge	3300025784	689484126	1280091
ADAS8	Anaerobic digester active sludge	3300025714	559791615	1003218
ADAS9	Anaerobic digester active sludge	3300025859	779020248	1463790
ADAS10	Anaerobic digester active sludge	3300025748	640699893	1138626
AS1	Agriculture soil	3300026761	24453291	79694
AS2	Agriculture soil	3300026766	25781940	79948
AS3	Agriculture soil	3300027429	24424955	81886
AS4	Agriculture soil	3300027410	31302445	101087
AS5	Agriculture soil	3300026753	23078731	74587
AS6	Agriculture soil	3300027116	23624174	78960
AS7	Agriculture soil	3300026802	34836368	92819
AS8	Agriculture soil	3300026900	54610317	161548
AS9	Agriculture soil	3300027437	33318367	104972
AS10	Agriculture soil	3300027453	31838534	93132
COM1	Compost	3300001258	138073274	247533
COM2	Compost	2199352012	495813128	848933
COM3	Compost	3300017816	351196916	635693
COM4	Compost	3300027271	220508683	349813
COM5	Compost	3300017544	457787885	845994
COM6	Compost	3300017647	448939775	705231
COM7	Compost	3300020591	343503552	564868
COM8	Compost	3300001077	236465129	523041
COM9	Compost	3300003150	104091772	264942
COM10	Compost	3300000145	85065922	132297
COM11	Compost	3300001232	33898870	50489
COM12	Compost	2199352008	256665756	436647
COM13	Compost	2199352035	51405246	71460
COM14	Compost	3300000869	109516588	194772
COM15	Compost	3300005392	29361615	34669
COM16	Compost	3300009949	84449027	140672
COM17	Compost	3300009948	65120098	107383
COM18	Compost	3300009950	115845660	191034
GS1	Grassland soil	3300002916	214104574	539535
GS2	Grassland soil	3300002907	420592052	981120
GS3	Grassland soil	3300002915	165939417	439950
GS4	Grassland soil	3300002561	507792538	1099704
GS5	Grassland soil	3300002917	757346763	1678570
GS6	Grassland soil	3300002562	518915676	1127421
GS7	Grassland soil	3300002557	233865259	622655
GS8	Grassland soil	3300002908	967438113	2077504
GS9	Grassland soil	3300026296	548056717	951020
GS10	Grassland soil	3300026322	503750574	1023488

GS11	Grassland soil	3300026528	618069426	1070598
HG1	Human gut system	3300008571	163086704	220519
HG2	Human gut system	3300007356	126726456	161507
HG3	Human gut system	3300007210	160776829	231021
HG4	Human gut system	3300007361	173118880	229956
HG5	Human gut system	3300008672	200079075	256044
HG6	Human gut system	3300008744	208172635	293086
HG7	Human gut system	3300006463	212850221	293809
HG8	Human gut system	3300007796	173992781	238201
HG9	Human gut system	3300007804	100761432	131805
HG10	Human gut system	3300008622	152622707	215113
HG11	Human gut system	3300007717	164858816	204019
HG12	Human gut system	3300008299	209026199	284195
HG13	Human gut system	3300008479	194179397	270790
HG14	Human gut system	3300007109	141045918	188333
HG15	Human gut system	3300008750	98564025	116189
HG16	Human gut system	3300008100	94952779	133926
HM1	Hypersaline mat	3300005717	315943091	630229
HM2	Hypersaline mat	3300009133	380042747	844906
HM3	Hypersaline mat	3300005642	427721363	765179
HM4	Hypersaline mat	3300005143	124176687	221446
HM5	Hypersaline mat	3300025384	133680257	240521
HM6	Hypersaline mat	3300025377	125674444	244906
HM7	Hypersaline mat	3300025364	103340022	187339
HRE1	Hydrocarbon resource environments (Oil reservoir)	3300001422	23719607	35050
HRE2	Hydrocarbon resource environments (Oil reservoir)	3300001190	3612035	5365
HRE3	Hydrocarbon resource environments (Oil reservoir)	3300001393	11319761	17164
HRE4	Hydrocarbon resource environments (Oil reservoir)	3300001195	7016652	11041
HRE5	Hydrocarbon resource environments (Oil reservoir)	3300000507	28712273	40166
HRE6	Hydrocarbon resource environments (Oil reservoir)	3300000408	18846681	25476
HS1	Hot spring	3300025462	189101311	435914
HS2	Hot spring	3300025374	97476306	232572
HS3	Hot spring	3300025440	152818410	356023
HS4	Hot spring	3300025360	94027272	168331
HS5	Hot spring	3300000345	49094763	110064
HS6	Hot spring	3300031980	267257994	617247
HS7	Hot spring	3300031875	369211308	763354
HS8	Hot spring	3300033894	147463480	268501
HS9	Hot spring	3300007193	76574785	223595
HS10	Hot spring	3300005963	46964575	105967
HS11	Hot spring	3300005856	95778213	171593
HS12	Hot spring	3300007203	73453869	228153
HS13	Hot spring	3300007178	62396224	176450
HS14	Hot spring	3300005966	71045757	131700
LL1	Landfill leachate	3300028028	475545597	864354
LL2	Landfill leachate	3300029288	2066651250	3975154
LL3	Landfill leachate	3300015214	2151718941	3941155
LL4	Landfill leachate	3300028603	1729402186	3057430
LL5	Landfill leachate	3300028601	905611696	1440715
LL6	Landfill leachate	3300014205	2274189263	4713313

LL7	Landfill leachate	3300014204	2297043378	4947135
LL8	Landfill leachate	3300014206	2497637605	5019461
LL9	Landfill leachate	3300028032	746937909	1379627
LL10	Landfill leachate	3300028602	1720483498	3252495
MS1	Marine sediment	3300002053	799007063	857807
MS2	Marine sediment	3300001782	449385852	488203
MS3	Marine sediment	3300001752	546713840	1387307
MS4	Marine sediment	3300002052	406438721	440211
MS5	Marine sediment	3300001751	558305947	1410986
MS6	Marine sediment	3300001855	303084940	795716
MS7	Marine sediment	3300027901	2405913374	4986014
MS8	Marine sediment	3300027888	2093100725	4192278
MS9	Marine sediment	3300027893	2405770886	4638423
MS10	Marine sediment	3300001753	719684507	1773993
MS11	Marine sediment	3300001854	930812755	2122334
MS12	Marine sediment	3300027814	762479598	1922283
MW1	Marine water	3300028189	449905649	856501
MW2	Marine water	3300031687	410396926	914988
MW3	Marine water	3300028177	408634935	735188
MW4	Marine water	3300028535	489316995	1087482
MW5	Marine water	3300031629	806100132	1627096
MW6	Marine water	3300003601	259650553	528258
MW7	Marine water	3300025623	382735651	725551
MW8	Marine water	3300000265	161683159	387725
MW9	Marine water	3300000153	161992569	328376
MW10	Marine water	3300000172	201262566	388542
OR1	Oil reservoir	3300035036	69829225	111598
OR2	Oil reservoir	3300035038	99252794	152636
OR3	Oil reservoir	3300035046	94073907	161691
OR4	Oil reservoir	3300036317	57638688	94156
OR5	Oil reservoir	3300035048	50068740	81530
OR6	Oil reservoir	3300035050	169884070	266430
OR7	Oil reservoir	3300035052	177248887	288315
OR8	Oil reservoir	3300035524	116600525	191865
OR9	Oil reservoir	3300035539	100086178	161409
OR10	Oil reservoir	3300035542	172062358	282985
OR11	Oil reservoir	3300035543	89955684	159016
OR12	Oil reservoir	3300035544	152691170	227435
OR13	Oil reservoir	3300035546	82893933	139497
RW1	River water	3300028071	132265979	279140
RW2	River water	3300027148	154072978	291628
RW3	River water	3300027488	142548013	280720
RW4	River water	3300024499	124449897	277128
RW5	River water	3300024306	183398807	409005
RW6	River water	3300028067	116036125	256146
RW7	River water	3300027160	176666155	395671
RW8	River water	3300026457	166921656	376394
RW9	River water	3300024498	100335601	224516
RW10	River water	3300027596	162454814	307205
RW11	River water	3300027129	109047475	237601
TFS1	Tropical forest soil	3300027313	206272783	481964
TFS2	Tropical forest soil	3300026887	65261237	150920
TFS3	Tropical forest soil	3300026845	48583672	117856
TFS4	Tropical forest soil	3300026852	24305883	60571
TFS5	Tropical forest soil	3300026979	81920775	197536
TFS6	Tropical forest soil	3300026824	43775853	103697

TFS7	Tropical forest soil	3300027014	102013575	241417
TFS8	Tropical forest soil	3300026908	76880979	188041
TFS9	Tropical forest soil	3300026819	44903906	101777
TFS10	Tropical forest soil	3300027063	156820484	368462
TFS11	Tropical forest soil	3300027019	132250044	265708
TFS12	Tropical forest soil	3300027010	104652154	259653
TFS13	Tropical forest soil	3300027049	110067664	266329
TFS14	Tropical forest soil	3300027042	103999555	260245
WB1	Wastewater bioreactor	3300003757	256997933	400666
WB2	Wastewater bioreactor	3300024984	355043206	529140
WB3	Wastewater bioreactor	3300024992	490734981	617734
WB4	Wastewater bioreactor	3300009070	473466628	577450
WB5	Wastewater bioreactor	3300003484	218046762	321182
WB6	Wastewater bioreactor	3300024990	476329968	648818
WB7	Wastewater bioreactor	3300024998	350228177	459430
WB8	Wastewater bioreactor	3300024991	412698899	554457
WB9	Wastewater bioreactor	3300003418	171736237	298355
WB10	Wastewater bioreactor	3300024986	306104122	372134
WB11	Wastewater bioreactor	3300003407	264251597	417208
WB12	Wastewater bioreactor	3300024987	280019008	367005
WB13	Wastewater bioreactor	3300024988	341047186	437514

Supplementary Table S6 Compost55 and compost76 metagenome assembly statistics

	Statistic	compost55	compost76
Raw datasets	Raw reads (counts)	60,423,492	58,519,117
	Total bases (bp)	16398254642	14115487744
	Paired reads (counts)	48,313,538	49,098,115
	Unpaired reads (counts)	2,201,461	2,099,164
Assembly	N50 (bp)	1,393	1,544
	L50 (bp)	104,761	109,695
	Longest contig (bp)	1,089,853	563,504
	Total bases in contigs (bp)	988,194,787	810,383,026
	Total bases in contigs (bp, >= 1000 bp)	630,203,503	499,132,808

Supplementary Table S7 Summary of abundant bacterial genera in total and active bacterial community as revealed from 16 rRAN genes and transcripts, respectively

Phylum	Order	Genus[a]	compost55		compost76	
			DNA	RNA	DNA	RNA
Actinobacteria	*Acidimicrobiales*	(OM1 clade)	0.81%	1.50%	0.00%	0.00%
	Acidimicrobiales	(uncultured)	1.23%	4.15%	0.00%	0.00%
	Micromonosporales	*Longispora*	5.76%	4.51%	0.00%	0.00%
	Micromonosporales	*Salinispora*	1.88%	4.80%	0.01%	0.03%
	Propionibacteriales	*Actinopolymorpha*	1.31%	3.75%	0.00%	0.00%
	Pseudonocardiales	*Pseudonocardia*	0.64%	2.80%	0.08%	0.01%
	Streptosporangiales	*Thermopolyspora*	1.24%	1.08%	1.24%	1.31%
	Streptosporangiales	*Actinomadura*	0.80%	1.58%	0.17%	0.01%
	Streptosporangiales	*Thermobispora*	6.15%	3.25%	0.05%	0.06%
	Streptosporangiales	*Thermomonospora*	0.66%	4.12%	0.76%	0.37%
	Rubrobacterales	*Rubrobacter*	2.54%	4.66%	0.00%	0.00%
Bacteroidetes	*Bacteroidetes* Incertae Sedis Order II	*Rhodothermus*	15.11%	4.49%	0.00%	0.00%
	Flavobacteriales	(NS9 marine group)	0.00%	0.00%	2.21%	0.09%
Chloroflexi	*Chloroflexales*	(FFCH7168)	0.00%	0.00%	0.76%	1.29%
Deinococcus-Thermus	*Thermales*	*Thermus*	0.56%	0.12%	1.39%	0.25%
Firmicutes	*Bacillales*	*Calditerricola*	0.04%	0.00%	15.47%	35.30%
	Bacillales	*Geobacillus*	2.83%	4.94%	1.62%	0.42%
	Bacillales	*Ureibacillus*	0.57%	1.00%	0.15%	0.00%
	Bacillales	*Planifilum*	0.10%	0.21%	0.10%	1.01%
	Clostridiales	*Thermaerobacter*	2.93%	1.76%	7.61%	12.04%
	Clostridiales	*Symbiobacterium*	0.16%	0.04%	31.11%	29.35%
	Clostridiales	*Hydrogenispora*	3.11%	0.89%	2.12%	0.00%
	Clostridiales	*Ruminiclostridium*	0.87%	0.11%	1.38%	0.00%
	Thermoanaerobacterales	*Thermosediminibacter*	0.01%	0.00%	2.05%	0.61%
	Thermoanaerobacterales	(Family III)	0.00%	0.00%	3.56%	2.03%
	Thermoanaerobacterales	*Brockia*	16.98%	26.70%	0.67%	1.29%
	Limnochordales	(*Limnochordaceae*)	0.73%	0.75%	0.75%	1.75%
	Limnochordales	(*Limnochordaceae*)	2.55%	0.44%	0.52%	0.71%
Proteobacteria	*Rhizobiales*	*Filomicrobium*	1.97%	5.17%	0.01%	0.01%
	Myxococcales	*Nannocystis*	0.00%	0.00%	0.23%	3.46%
	Methylococcales	*Methylocaldum*	0.08%	0.33%	0.16%	1.72%

	Xanthomonadales	*Pseudoxanthomonas*	0.13%	0.02%	1.92%	0.10%
	Xanthomonadales	(uncultured bacterium)	0.64%	1.12%	0.06%	0.01%
Tenericutes	*Haloplasmatales*	*Haloplasma*	0.03%	0.00%	1.32%	0.03%

[a] Only genera from abundant orders (greater than 1 %) are given. In the case the order or genus could not be assigned, the taxonomic name at the highest determined taxonomic resolution is given in parenthesis.

Supplementary Table S8 Annotation of functionally derived LEs to NCBI nr database

Name of lipolytic enzyme	Query length	Subject length	Identity %	Closest similar lipolytic enzyme	Organism	Accession Nr.	E-value
EstC55-2	291	291	99	alpha/beta hydrolase	Chloroflexi bacterium	PZN58850	0.00E+00
EstC55-3	513	525	50	carboxylesterase	Acidimicrobiales bacterium	PZS19531	2.00E-128
EstC55-4_1	271	271	83	alpha/beta hydrolase	*Pseudonocardia* sp. CNS-004	WP_075953220	9.00E-161
EstC55-4_2	390	391	81	serine hydrolase	*Frankia* sp. BMG5.36	WP_071051881	0
EstC55-5	310	311	54	lipolytic enzyme	uncultured bacterium	ACL67843	5.00E-102
EstC55-6	345	343	32	patatin-like phospholipase family protein	*Halothermothrix orenii*	WP_015923954	2.00E-44
EstC55-7	430	417	58	beta-lactamase-related serine hydrolase	*Steroidobacter agariperforans*	WP_129641015	1.00E-166
EstC55-8_1	318	317	80	lipase	Bacterium HR24	GBD13261	6.00E-169
EstC55-8_2	270	266	55	alpha/beta fold hydrolase	*Candidatus Entotheonella palauensis*	WP_089934682	2.00E-99
EstC55-10	288	288	100	patatin	*Thermobispora bispora*	WP_013131472	0.00E+00
EstC55-12	384	384	60	alpha/beta hydrolase	*Micromonospora pattaloongensis*	WP_091555337	2.00E-143
EstC55-13	355	360	46	Alpha/beta hydrolase family protein	*Lokiarchaeum* sp. GC14_75	KKK41396	9.00E-111
EstC55-15	253	243	62	alpha/beta fold hydrolase	*Catelliglobosispora koreensis*	WP_026208377	1.00E-100
EstC55-18	267	267	100	pimeloyl- methyl ester esterase	*Methylocaldum szegediense*	WP_026611794	0
EstC55-19_1	270	266	57	alpha/beta fold hydrolase	*Candidatus Entotheonella palauensis*	WP_089934682	1.00E-103
EstC55-19_2	317	317	99	Carboxylesterase NlhH	Bacterium HR24	GBD13261	0.00E+00
EstC55-20	350	350	100	alpha/beta hydrolase	Actinobacteria bacterium	PZN40536	0
EstC55-23	322	327	62	Acetyl esterase	*bacterium HR29*	GBD24385	4.00E-130
EstC55-24	325	342	51	alpha/beta hydrolase	*Dactylosporangium aurantiacum*	WP_052387799	2.00E-79
EstC55-25	367	367	89	alpha/beta hydrolase	*Pseudonocardia* sp. CNS-004	WP_075951996	0
EstC55-26	317	317	97	patatin	*Rhodothermus marinus*	WP_014066731	0
EstC55-31	265	263	76	alpha/beta fold family hydrolase	Bacterium CSP1-2	KRT66319	6.00E-149
EstC55-34	290	290	92	alpha/beta hydrolase	*Rhodothermus marinus*	WP_014067099	0
EstC55-38	329	329	97	esterase	*Thermus thermophilus*	WP_024119707	0
EstC55-40	444	454	53	serine hydrolase	*Sphaerobacter thermophilus*	PZN60880	3.00E-152
EstC55-42	373	363	62	ferruloyl esterase	*Micromonospora* sp. ATCC 39149	EEP71116	6.00E-148
EstC55-46	415	410	73	serine hydrolase	*Micromonospora noduli*	WP_112582259	0
EstC55-51	268	271	67	alpha/beta fold hydrolase	*Micromonospora pattaloongensis*	WP_091559969	1.00E-111
EstC55-52	503	504	87	Carboxylesterase	*Thermaerobacter marianensis*	WP_013495056	0

EstC55-53	406	406	82	beta-lactamase-related serine hydrolase	*Steroidobacter cummioxidans*	WP_116812916	0
EstC55-56	314	313	62	alpha/beta hydrolase	*Variovorax* sp. BK460	WP_130428336	3.00E-124
EstC55-57	271	301	49	lipase	Planctomycetaceae bacterium	MBN00417	1.00E-76
EstC55-60	340	343	59	alpha/beta hydrolase	Sandaracinaceae bacterium	RZO48478	2.00E-120
EstC55-61	388	388	99	beta-lactamase	*Thermus*	WP_008632158	0
EstC55-62	528	528	100	carboxylesterase	Proteobacteria bacterium	PZN79242	0
EstC55-63	317	317	99	patatin	*Rhodothermus marinus*	WP_014066731	0
EstC55-65	409	409	64	beta-lactamase family protein	Chloroflexi bacterium	TME11201	0.00E+00
EstC55-66	429	431	69	beta-lactamase	*Thermomicrobium roseum*	ACM07132	0
EstC55-71	317	317	98	Extracellular esterase EstB	Bacterium HR41	GBD45993	0
EstC55-72	287	296	62	lipase/esterase	uncultured bacterium	AAS77236	8.00E-115
EstC55-73	385	393	71	beta-lactamase class C family	*Paracoccus* sp. J56	WP_085501535	0
EstC55-76	373	357	77	alpha/beta hydrolase	Actinobacteria bacterium	PZM88754	0.00E+00
EstC55-77	278	308	36	Lip3594	uncultured bacterium	QCC19993.1	1.00E-45
EstC55-78	319	311	53	EST1	uncultured microorganism	ADR31550	5.00E-103
EstC55-80	408	375	70	beta-lactamase family protein	*Chloroflexi bacterium*	TMF40125	0.00E+00
EstC55-81	282	269	33	alpha/beta hydrolase	*Thermorudis peleae*	WP_081886744	9.00E-27
EstC55-88	317	317	98	Extracellular esterase EstB	Bacterium HR41	GBD45993	0
EstC55-90	286	286	100	lipase	*Thermomonospora curvata*	WP_012852218	0
EstC55-95	288	283	51	acetylxylan esterase	*Yuhushiella deserti*	WP_092534474	3.00E-85
EstC55-96	267	267	97	alpha/beta hydrolase	*Rhodothermus marinus*	WP_012844996	0
EstC55-97	240	234	28	alpha/beta hydrolase	*Stenotrophomonas maltophilia*	WP_049464595	3.03E-07
EstC55-100	376	363	57	alpha/beta hydrolase	*Dactylosporangium aurantiacum*	WP_052386271	2.00E-131
EstC55-102	707	699	47	lipase	*Candidatus Competibacteraceae bacterium*	RUQ29580	5.00E-160
EstC55-105	223	223	62	triacylglycerol lipase	Chloroflexi bacterium	TMG05867	2.00E-92
EstC55-110	434	436	66	beta-lactamase-related serine hydrolase	*Steroidobacter agariperforans*	WP_129644730	0.00E+00
EstC55-111	392	393	92	SGNH/GDSL hydrolase family protein	*Rhodothermus marinus*	WP_014065857	0
EstC55-113	409	506	56	beta-lactamase	uncultured sludge bacterium	ADC79147	2.00E-148
EstC55-118	497	496	76	carboxylesterase family protein	*Thermaerobacter sp. FW80*	WP_135226007	0
EstC55-131	251	251	99	phospholipase	*Thermus thermophilus*	WP_024119695	2.00E-170
EstC55-145	305	307	56	lipolytic enzyme	uncultured bacterium	ACL67849	1.00E-105

EstC55-147	386	410	50	beta-lactamase-related serine hydrolase	*Acidimicrobiia bacterium*	RTL09767	9.00E-106
EstC55-150	393	393	97	SGNH/GDSL hydrolase family protein	*Rhodothermus marinus*	WP_014065857	0
EstC55-151	311	315	66	lipase (class 2)	*Herbihabitans rhizosphaerae*	RZS45231	8.00E-152
EstC55-154	394	386	47	lipase	*Streptomyces sp. HST28*	WP_127910203	2.00E-96
EstC55-156	517	520	51	tannase/feruloyl esterase family alpha/beta hydrolase	Betaproteobacteria bacterium	TMH79627	1.00E-168
EstC55-159	284	292	59	alpha/beta hydrolase	*Actinomadura amylolytica*	WP_119727475	1.00E-94
EstC55-163	303	267	39	patatin-like phospholipase family protein	*Orenia metallireducens*	WP_068717466	2.00E-55
EstC55-164	430	432	81	serine hydrolase	Proteobacteria bacterium	PZN68398	0.00E+00
EstC55-165	277	277	97	phospholipase	*Rhodothermus marinus*	WP_014066117	0
EstC55-167	248	248	94	alpha/beta hydrolase	Bacteroidetes bacterium	PZN50641	3.00E-162
EstC55-168	513	598	55	serine hydrolase	Blastocatellia bacterium	PWT89381	0
EstC55-169	533	544	51	tannase/feruloyl esterase family alpha/beta hydrolase	Gammaproteobacteria bacterium	TAJ93847	0.00E+00
EstC55-186	515	515	99	carboxylesterase	*Thermus thermophilus*	WP_124105112	0
EstC55-188	237	238	99	alpha/beta hydrolase	*Thermus thermophilus*	WP_014510770	3.00E-158
EstC55-197	291	298	56	alpha/beta hydrolase	Cyanobacterium TDX16	OWY61995	5.00E-100
EstC55-213	286	286	100	esterase	*Rhodothermus marinus*	WP_014067193	0
EstC55-215	263	263	99	alpha/beta hydrolase	Firmicutes bacterium	REJ37793	0.00E+00
EstC55-227	472	455	69	alpha/beta hydrolase	*Micromonospora* sp. HK10	WP_046565934	0
EstC55-229	311	312	58	esterase/lipase AS-Trib30	uncultured bacterium	AIT69759	3.00E-99
EstC55-231	366	384	68	alpha/beta hydrolase	*Micromonospora pattaloongensis*	WP_091555337	7.00E-164
EstC55-234	557	603	58	tannase/feruloyl esterase family alpha/beta hydrolase	*Steroidobacter* sp. JW-3	WP_129775856	0.00E+00
EstC55-235	326	284	60	alpha/beta fold hydrolase	*Streptomyces yeochonensis*	WP_107498754	7.00E-116
EstC55-239	381	968	98	Beta-lactamase	*Rhodothermus marinus*	WP_014067493	0
EstC55-241	567	569	61	phospholipase	Acidobacteria bacterium	RPJ58855.1	0
EstC55-244	255	253	41	esterase	Rhodospirillaceae bacterium	MAG96647	4.00E-49
EstC55-245	462	446	53	penicillin-binding protein	Acidobacteria bacterium	RLE30144	3e-158
EstC55-247	368	363	56	alpha/beta hydrolase	Deltaproteobacteria bacterium	TMA57433	1.00E-143
EstC55-251	317	317	97	patatin	*Rhodothermus marinus*	WP_014066731	0
EstC55-253	304	287	29	alpha/beta hydrolase	*Clostridium felsineum*	WP_077893977	7.00E-32
EstC55-256	270	270	56	alpha/beta hydrolase	Dehalococcoidia bacterium	TET95823.1	2.00E-102
EstC55-258	389	388	82	Esterase EstB	Bacterium HR29	GBD24121.1	0
EstC55-268	303	303	65	esterase LC-Est5	uncultured bacterium	AIT56391	2.00E-121
EstC76-21	288	288	100	patatin	*Thermobispora bispora*	WP_013131472	0
EstC76-28_1	257	257	99	alpha/beta hydrolase	*Thermus thermophilus*	WP_014629868	2.00E-175

EstC76-28_2	243	243	97	alpha/beta hydrolase	*Thermus thermophilus*	WP_014510102	1.83E-168
EstC76-36	237	238	99	alpha/beta hydrolase	*Thermus thermophilus*	WP_014510770	3.00E-158
EstC76-98	425	431	43	beta-lactamase-related serine hydrolase	Acidobacteria bacterium	TDI24993	1.00E-106
EstC76-123	450	412	41	serine hydrolase	*Truepera radiovictrix*	WP_013178639	1.00E-91
EstC76-135	310	308	57	Carboxylesterase NlhH	Bacterium HR25	GBD14217	4.00E-113
EstC76-136	515	515	99	carboxylesterase	*Thermus thermophilus*	WP_014509791	0
EstC76-137	277	277	99	phospholipase	*Rhodothermus marinus*	WP_014066117	0
EstC76-174	388	388	98	beta-lactamase	*Thermus*	WP_008632158	0
EstC76-177	281	281	99	alpha/beta hydrolase	*Thermaerobacter* sp. FW80	WP_135226180	0.00E+00
EstC76-179	329	329	97	esterase	*Thermus thermophilus*	WP_024119707	0
EstC76-202	267	267	100	pimeloyl-[acyl-carrier protein] methyl ester esterase	*Methylocaldum szegediense*	WP_026611794	0
EstC76-218	329	329	99	esterase	*Thermus thermophilus*	WP_014629382	0
EstC76-221	257	257	86	alpha/beta hydrolase	*Caldicoprobacter oshimai*	WP_025747209	1.00E-164
EstC76-222	260	260	91	patatin-like phospholipase family protein	*Caldicoprobacter*	WP_025747046	2.00E-171
EstC76-248	257	257	93	alpha/beta hydrolase	*Thermus thermophilus*	WP_124104731	1.00E-162
EstC76-250	251	251	99	patatin-like phospholipase family protein	*Thermus thermophilus*	WP_024119695	4.00E-172
EstC76-261	345	343	32	patatin-like phospholipase family protein	*Halothermothrix orenii*	WP_015923954	2.00E-44
EstC76-262	415	410	73	serine hydrolase	*Micromonospora noduli*	WP_112582259	0
EstC76-263	376	363	57	alpha/beta hydrolase	*Dactylosporangium aurantiacum*	WP_052386271	2.00E-131
EstC76-266	257	257	94	alpha/beta hydrolase	*Thermus thermophilus*	WP_014510103	2.00E-164
EstC76-269	260	260	98	patatin-like phospholipase family protein	*Caldicoprobacter*	WP_025747046	0

Supplementary Table S9 Description of LEs from functional metagenomic selections

Name of unique putative lipolytic enzymes	Insert size (bp)	Protein length (aa)	Calculate molecular weight (kDa)	Theoretical pI	Substrate specificity towards triacylglycerides with different chain length	Lipolytic family
EstC55-2	1819	291	32.3	6.11	C4, C6	V
EstC55-3	8599	513	54.1	4.9	C4, C6	VII
EstC55-4_1	8105	271	29	4.63	C4, C6	V
EstC55-4_2	8105	390	42.4	4.84	C4, C6	VIII
EstC55-5	4235	310	32.7	4.66	C4, C6	IV
EstC55-6	5612	345	37.2	9.25	C4, C6	patatin-like-protein
EstC55-7	9350	430	47.1	8.4	C4, C6	VIII
EstC55-8_1	2323	318	33.9	5.49	C4, C6	IV
EstC55-8_2	2323	270	29.8	5.89	C4, C6	V
EstC55-10	8453	288	30.7	5.1	C4, C6	Patatin
EstC55-12	6040	384	40.9	7.2	C4, C6, C8	V
EstC55-13	5540	355	38.3	5.63	C4, C6	EstGS
EstC55-15	1688	253	27	10.1	C4, C6, C8	new family (this study)
EstC55-18	1745	267	28.9	7.85	C4, C6	V
EstC55-19_1	5551	259	28.8	5.64	C4, C6	IV
EstC55-19_2	5551	318	33.9	5.26	C4, C6, C8	V
EstC55-20	3965	350	38.1	6.84	C4, C6	V
EstC55-23	8111	322	35.5	5.6	C4, C6	IV
EstC55-24	9857	325	33.7	5.54	C4, C6	EstGS
EstC55-25	1925	367	39.3	7.21	C4, C6	V
EstC55-26	8221	317	35	8.69	C4, C6	patatin-like-protein
EstC55-31	2989	265	29.8	5.76	C4, C6	V
EstC55-34	5249	290	32.2	9.14	C4, C6	V
EstC55-38	2502	329	36	8.08	C4, C6	LipT
EstC55-40	3844	444	48.3	5.5	C4	VIII
EstC55-42	4335	373	39	8.07	C4	EM3L4
EstC55-46	5482	415	45.6	5.16	C4, C6, C8	VIII
EstC55-51	3141	268	28.1	11.23	C4, C6	V
EstC55-52	2703	503	53.9	5.38	C4, C6	VII
EstC55-53	6466	406	44.9	6.46	C4	VIII
EstC55-56	2166	314	33.6	5.85	C4, C6	IV
EstC55-57	4293	271	29.4	5.18	C4, C6	IV
EstC55-60	6444	340	36.7	5.95	C4	IV
EstC55-61	6942	388	42.9	6.07	C4, C6	VIII

EstC55-62	4505	528	56.7	8.2	C4, C6	VII
EstC55-63	3924	317	34.9	7.74	C4	Patatin
EstC55-65	3706	409	44.1	5.07	C4	VIII
EstC55-66	5776	429	47	5.97	C4, C6	VIII
EstC55-71	2814	317	33.2	8.99	C4, C6, C8, C10, C12, C14	I
EstC55-72	6166	287	30.5	4.97	C4, C6	IV
EstC55-73	4617	385	40.9	5.75	C4, C6	VIII
EstC55-76	2896	373	40.9	9.82	C4, C6	V
EstC55-77	3392	278	30.1	5.07	C4	EM3L4
EstC55-78	6416	319	33.5	4.62	C4, C6	IV
EstC55-80	3076	408	43.7	5.24	C4	VIII
EstC55-81	3362	282	30.6	5	C4, C6	V
EstC55-88	8186	317	33.2	9.12	C4, C6, C8, C10, C12	I
EstC55-90	4812	286	30.9	8.68	C4, C6, C8, C10, C12	I
EstC55-95	6415	288	29.5	6.95	C4, C6, C8	III
EstC55-96	5343	267	29.9	5.99	C4, C6	V
EstC55-97	4889	240	25.6	4.32	C4	new family (this study)
EstC55-100	2878	376	40.8	5.23	C4, C6	V
EstC55-102	2936	707	72.3	4.75	C4	new family (this study)
EstC55-105	2714	223	23.9	7	C4, C6, C8, C10	I
EstC55-110	8074	434	48.2	6.85	C4, C6	VIII
EstC55-111	2870	392	41.8	5.69	C4	II
EstC55-113	5055	409	44.4	7.66	C4	VIII
EstC55-118	2603	497	53.9	5.39	C4, C6	VII
EstC55-131	5270	251	27.2	6.01	C4, C6	Patatin
EstC55-145	3366	305	31.5	5.04	C4, C6	IV
EstC55-147	4651	386	41.4	4.61	C4, C6	VIII
EstC55-150	4567	393	41.7	5.23	C4, C6	II
EstC55-151	5771	311	33.5	6.79	C4, C6, C8	I
EstC55-154	5680	394	41.5	4.83	C4	XVII
EstC55-156	4142	517	55.8	4.6	C4	Tannase
EstC55-159	3057	284	30.5	10.87	C4	V
EstC55-163	2207	303	31.6	4.82	C4, C6	patatin-like-protein
EstC55-164	6522	430	46	5.56	C4	VIII
EstC55-165	3926	277	31.1	5.35	C4, C6	FLS18
EstC55-167	4015	248	27.2	6.09	C4	new family (this study)
EstC55-168	5684	513	57.5	9.29	C4	VIII
EstC55-169	5280	533	57.1	4.8	C4	Tannase
EstC55-186	2955	515	56.2	5.83	C4	VII
EstC55-188	6426	237	25.8	8.09	C4	new family (this study)

EstC55-197	6388	291	30.7	4.73	C4, C6	V
EstC55-213	4812	286	31.9	10.12	C4	I (new subfamily)
EstC55-215	5300	263	29.7	10.13	C4	V
EstC55-227	5479	472	50.7	9.45	C4	new family (this study)
EstC55-229	7811	311	33.9	6.92	C4, C6	IV
EstC55-231	6018	366	39.7	5.71	C4, C6	V
EstC55-234	7802	557	59.9	5.86	C4, C6	Tannase
EstC55-235	3084	326	34.4	6.44	C4, C6, C8, C10, C12	I
EstC55-239	6478	381	42.1	8.75	C4	VIII
EstC55-241	3087	567	63	6.79	C4	FLS18
EstC55-244	4612	255	27.8	5.86	C4, C6	V
EstC55-245	9142	462	50	5.92	C4, C6	VIII
EstC55-247	3706	368	39.4	5.12	C4, C6	IV
EstC55-251	8621	317	34.9	8.34	C4	patatin-like-protein
EstC55-253	1398	297	31.7	5.61	C4	IV
EstC55-256	6473	270	29.8	5.63	C4, C6	V
EstC55-258	6696	389	42.1	6.02	C4, C6	VIII
EstC55-268	7750	303	33.4	6.64	C4	IV
EstC76-21	5941	288	30.7	5.1	C4, C6	patatin-like-protein
EstC76-28_1	6905	257	28	6.24	C4, C6	V
EstC76-28_2	6905	243	26	5.98	C4, C6	V
EstC76-36	6361	237	25.8	8.09	C4	new family (this study)
EstC76-98	6234	425	46.4	7.31	C4, C6, C8	VIII
EstC76-123	4787	450	49.2	8.96	C4	VIII
EstC76-135	6286	310	33.8	5.59	C4	IV
EstC76-136	2955	515	56.2	5.83	C4	VII
EstC76-137	1878	277	31.1	5.47	C4	FLS18
EstC76-174	5888	388	42.8	6.29	C4	VIII
EstC76-177	4039	265	29.4	9.53	C4	I (new subfamily)
EstC76-179	3556	329	36	8.08	C4, C6	LipT
EstC76-202	1745	267	28.9	7.85	C4, C6	V
EstC76-218	2449	329	36.1	8.08	C4	LipT
EstC76-221	3944	257	28.9	5.49	C4, C6	new family (this study)
EstC76-222	2289	260	28.6	6.32	C4	patatin-like-protein
EstC76-248	1575	257	28.1	6.25	C4	V
EstC76-250	5270	251	27.2	6.01	C4, C6	patatin-like-protein
EstC76-261	5603	345	37.2	9.25	C4, C6	patatin-like-protein
EstC76-262	5482	415	45.6	5.16	C4, C6, C8	VIII
EstC76-263	2752	376	40.8	5.23	C4, C6	V
EstC76-266	11975	257	28.1	6.54	C4, C6	V

EstC76-269	5057	260	28.6	6.02	C4, C6	patatin-like-protein

Supplementary Table S10 Annotation of recent reported novel LEs (including these from literatures and this study) against ESTHER database

Protein name	Query length	Subject length	Description	Acce. Nr	ESTHER Family	E-value	Identity %	Query coverage %
Est22	463	463	*Alteromonas* sp.; Homoserine O-acetyltransferase	9alte-a0a010qrx3	Homoserine_transacetylase	0	100	100
EstL28	290	322	*Dehalococcoides* sp.; putative hydrolase	dehsc-q3zyd4	6_AlphaBeta_hydrolase	2.00E-29	28	100
Rv0045c	298	298	*Mycobacterium bovis*; hypothetical protein rv0045c	myctu-RV0045C	6_AlphaBeta_hydrolase	e-171	100	100
EstGX1	201	230	*Streptomyces bingchenggensis*; uncharacterized protein	strbb-d7bzy9	DLH-S	0.073	27	40
EstLiu	293	285	*Robiginitalea biformata*; probable lipase	9flao-a4cmr0	Hormone-sensitive_lipase_like	7.00E-48	42	91
EstY	423	423	uncultured bacterium; Esterase	9bact-b0ln78	Fungal-Bact_LIP	0	100	100
EstGS	397	397	uncultured organism; Putative uncharacterized protein	9zzzz-g3crc6	Chlorophyllase	0	100	100
EM3L4	330	304	*Mycobacterium leprae*; putative secreted hydrolase	mycle-LPQC	Esterase_phb	5.00E-36	32	97
FLS18	259	224	uncultured organism EH82; Uncharacterized protein	9zzzz-EH82	5_AlphaBeta_hydrolase	5.00E-47	44	84
Est903	300	300	uncultured bacterium; Est903 Putative esterase	9bact-a0a345g0q2	Hormone-sensitive_lipase_like	e-176	100	100
EstJ	317	310	Planctomycetales bacterium; Esterase	9bact-a0a1u7gdl0	Hormone-sensitive_lipase_like	2.00E-79	52	85
PE10	279	279	Pelagibacterium; halotolerans Phospholipase/carboxylesterase family protein PE10	pelhb-g4rec9	Hormone-sensitive_lipase_like	0	100	100
Est12	329	316	*Psychrobacter cryohalolentis*; alpha/beta hydrolase	psyck-q1q8n0	Hormone-sensitive_lipase_like	e-138	81	91
EstDZ2	271	271	uncultured Acetothermia bacterium; estDZ2a esterase	9bact-a0a1l2dxz2	Hormone-sensitive_lipase_like	e-157	100	100
Est9x	294	294	uncultured bacterium; Esterase	9bact-j9vdv8	Est9X	0	100	100
Lip10	348	361	*Aspergillus oryzae*; arylacetamide deacetylase	aspor-q2tw16	Hormone-sensitive_lipase_like	e-155	71	97
EstGH	448	437	*Xanthomonas campestris*; hypothetical protein xcc3623	xanca-XCC3623	Lipase_3	4.00E-06	29	36
EML1	304	300	uncultured bacterium; probable lipase	9bact-q1paf1	Lipase_3	4.00E-13	34	53
FnL	302	364	*Thermotoga maritima*; esterase	thema-TM0053	AlphaBeta_hydrolase	5.00E-31	31	90
EstP2K	224	218	Uncultured prokaryote; esterase/lipase	9zzzz-KY203033	6_AlphaBeta_hydrolase	2.00E-42	45	91

LipA	277	367	*Magnaporthe grisea*; Triacylglycerol lipase	maggr-q0pnd5	PGAP1	2.00E-24	36	88
LipSM54	526	526	*Xanthomonas gardneri*; hypothetical protein	xanax-XAC0753	Cocaine_esterase	e-161	55	95
MtEst45	516	540	*Hahella chejuensis*; predicted acylesterase	hahch-q2seh8	Cocaine_esterase	e-114	42	95
LipT	329	329	uncultured bacterium; LipT	9bact-k7qe48	Pectinacetylesterase-Notum	0	100	100
EstSt7	322	322	*Sulfolobus tokodaii*; hypothetical protein st2026	sulto-ST2026	A85-Feruloyl-Esterase	0	100	100
Rlip1	361	361	uncultured bacterium; lipase	9bact-c0k075	A85-Feruloyl-Esterase	0	100	100
EstA	277	277	uncultured bacterium; Esterase A	9bact-b0fwn3	A85-EsteraseD-FGH	e-164	100	100
FLS12	270	270	uncultured bacterium; Esterase	9bact-b8y562	A85-EsteraseD-FGH	e-159	100	100
lp_3505	263	263	*Lactobacillus plantarum*; acetylesterase	lacpl-EST2	A85-EsteraseD-FGH	e-157	100	100
EstC55-15	253	278	*Streptomyces coelicolor*; putative hydrolase	strco-SCO4160	AlphaBeta_hydrolase	2.00E-69	56	93
EstC55-97	240	277	*Sphingomonas wittichii*; Alpha/beta hydrolase fold	sphww-a5v750	Carbon-carbon_bond_hydrolase	1.00E-07	35	82
EstC55-102	707	430	*Shewanella oneidensis*; conserved hypothetical protein	sheon-SO2934	Lipase_bact_N_lipase	7.00E-35	29	80
EstC55-167	248	262	Verrucomicrobiae bacterium; DG1235 Putative uncharacterized protein	9bact-b5jkk1	HNLyase_Bact	2.00E-66	52	93
EstC55-227	472	472	*Deinococcus radiodurans*; hypothetical protein	deira-DR0553	6_AlphaBeta_hydrolase	2.00E-57	34	85
EstC76-36	237	238	*Thermus thermophilus*; AlphaBeta_hydrolase protein	theth-TT1662	AlphaBeta_hydrolase	e-131	98	99
EstC76-221	257	265	*Clostridium acetobutylicum*; alpha/beta superfamily hydrolase	cloac-CAC3665	AlphaBeta_hydrolase	8.00E-74	52	100

Supplementary Table S11 Protein families used for building LE-specific profile hidden markov modle (HMM) database

Profile HMM name (ESTHER family)	Database	Lipolytic family	Interpro annotation	Pfam annotation	Reference
Bacterial_lip_FamI.1	ESTHER	I.1	IPR000734 (Lipase)	PF01764 (Lipase_3)	Nardini et al., 2000
Bacterial_lip_FamI.2	ESTHER	I.2	IPR000734 (Lipase)	PF01764 (Lipase_3)	Arpigny & Jaeger, (1999)
Bacterial_lip_FamI.3	ESTHER	I.3	IPR000734 (Lipase)	PF01764 (Lipase_3)	Angkawidjaja et al., 2010
Bacterial_lip_FamI.5	ESTHER	I.5	IPR000734 (Lipase)	PF01764 (Lipase_3)	Tyndall et al., 2002
Bacterial_lip_FamI.6	ESTHER	I.6	IPR000734 (Lipase)	PF01764 (Lipase_3)	Tiesinga et al., 2007
Lipase_2	ESTHER	I.4 I.7	IPR002918 (Lipase EstA/Esterase EstB)	PF01674 (Lipase_2)	Arpigny & Jaeger, (1999)
Bacterial_lip_FamI.8	ESTHER	I.8	-	PF12146 (Hydrolase_4)	de Pascale et al., 2008
Lipase_GDSL	PFAM	II	IPR001087 (GDSL lipase/esterase)	PF00657 (Lipase_GDSL)	
Lipase_GDSL_2	PFAM	II	IPR013830 (SGNH hydrolase-type esterase domain)	PF13472 (Lipase_GDSL_2)	
Lipase_GDSL_3	PFAM	II	IPR013830 (SGNH hydrolase-type esterase domain)	PF14606 (Lipase_GDSL_3)	Molgaard A et al., 2000
Lipase_GDSL_lke	PFAM	II	IPR032588 (Putative GDSL-like Lipase/Acylhydrolase)	PF16255 (Lipase_GDSL_lke)	
Polyesterase-lipase -cutinase	ESTHER	III	IPR041127 (Chlorophyllase enzyme)	PF12695 (Abhydrolase_5), PF12740 (Chlorophyllase2)	Sulaiman et al., (2012)
Hormone-sensitive _lipase_like	ESTHER	IV	IPR013094 (Alphabeta hydrolase fold-3), IPR002168 (Lipolytic enzyme), IPR033140 (Lipase_GDXG_put_SER_AS), IPR002168 (Lipase_GDXG_HIS_AS)	PF07859 (Abhydrolase_3)	Kim (2017)
GTSAGmotif	ESTHER	IV	IPR013094 (Alphabeta hydrolase fold-3)	PF07859 (Abhydrolase_3)	Li et al., 2014
ABHD6-Lip	ESTHER	V.1	-	-	Pribasnig et al., (2015)
Carboxymethylbutenolide _lactonase	ESTHER	V.2	IPR000073 (Alpha/beta hydrolase fold-1), IPR026968 (3-oxoadipate enol-lactonase)	PF12697 (Abhydrolase_6)	Arpigny & Jaeger, (1999)
UCP031982	ESTHER	V.3	IPR016986 (Uncharacterised conserved protein UCP031982, alpha/beta hydrolase, XabL), IPR005065 (Platelet-activating factor acetylhydrolase-like)	-	Park et al., (2007)
LYsophospholipase _carboxylesterase	ESTHER	VI	IPR003140 (Phospholipase/carboxylesterase/thioesterase PLP_Cesterase PLipase/COase/thioEstase)	PF02230 (Abhydrolase_2)	Martinez et al., (2011)
Carb_B_Bacteria	ESTHER	VII	IPR002018 (Carboxylesterase, type B)	PF00135 (COesterase)	Ewis et al., (2004)
Beta-lactamase	PFAM	VIII	IPR001466 (Beta-lactamase-related)	PF00144 (Beta-lactamase)	Hausmann & Jaeger, 2010
PHAZ7_phb_depolymerase	ESTHER	IX	-	-	Handrick et al., (2001)
Bacterial_EstLip_FamX	ESTHER	X.1 X.2	-	PF12695 (Abhydrolase_5)	Levisson et al., (2007) Bassegoda et al., (2012)
Fungal-Bact_LIP	ESTHER	XVII XIX	IPR005152 (Lipase_secreted)	PF03583 (LIP)	Castilla et al., 2017 Parapouli et al., 2018

Lipase_3	ESTHER	XI	IPR000734 (Lipase), IPR002921 (Fungal lipase-like domain), IPR005592 (Mono-/di-acylglycerol lipase, N-terminal Mono/diacylglycerol_lipase_N)	PF03893 (Lipase3_N), PF01764 (Lipase_3)	Lee et al., (2006)
Bact_LipEH166_FamXII	ESTHER	XII	-	-	Kim et al., (2009)
CarbLipBact_1	ESTHER	XIII-1	IPR012354 (Esterase/lipase)	-	Rao et al., (2013)
CarbLipBact_2	ESTHER	XIII-2 XVIII	IPR012354 (Esterase/lipase)	-	Chow et al., (2012) Samoylova et al., 2018
PC-sterol_acyltransferase	ESTHER	XIV	IPR003386 (Lecithin: cholesterol/phospholipid:diacylglycerol acyltransferase)	PF02450 (LACT-Lecithin: cholesterol acyltransferase)	Rao et al., (2011)
Duf_3089	ESTHER	XV	IPR021440 (Protein of unknown function DUF3089)	PF11288 (DUF3089)	Rodriguez et al., (2015); Bayer et al., (2010)
Bacterial_Est97	ESTHER	XVI	-	-	Fu et al., (2013)
Lipae	PFAM	patatin-like-protein	IPR013818 (Lipase/vitellogenin)	PF00151(Lipase)	Tirawongsaroj et al., 2008
Patatin			IPR002641 (Patatin-like phospholipase domain)	PF01734 (Patatin)	
DUF3734			IPR021095 (DUF3734)	PF12536 (Patatin phospholipase)	
Tannase	ESTHER	Tannase	IPR011118 (Tannase and feruloyl esterase)	PF07519 (Tannase)	Banerjee et al., 2012
Est9X	ESTHER	Est9X	-	-	Fang et al., 2014
A85-EsteraseD-FGH	ESTHER	lp_3505 FLS12 EstA	IPR000801 (Esterase_put), IPR014186 (S-formylglutathione_hydrol)	PF00756 (Esterase)	Esteban-Torres et al., 2014 Hu et al., 2010 Chu et al., 2008
Chlorophyllase	ESTHER	EstGS	IPR017395 (Chlorophyllase)	PF07224 (Chlorophyllase)	Nacke et al., 2011
Pectinacetylesterase-Notum	ESTHER	LipT	IPR004963 (Pectinacetylesterase/NOTUM PAE/NOTUM)	PF03283 (PAE)	Chow et al., 2012
A85-Feruloyl-Esterase	ESTHER	Rlip1 EstSt7	IPR000801 (Esterase_put)	PF00756 (Esterase)	Liu et al., 2009 Wei et al., 2013
Homoserine_transacetylase	ESTHER	Est22	IPR008220 (Homoserine/serine acetyltransferase MetX-like HAT_MetX-like), IPR000073 (Alpha/beta hydrolase fold-1)	PF00561 (Abhydrolase_1)	Li et al., 2017

Supplementary Table S12 Validation of the liplolytic family assignment strategy with family/function-known LEs (with α/β-hydrolase fold)

Protein name	Blastp annotation [a]	pHMM-omega annotation (lipolytic family)	Lipolytic family (This study) [b]	Lipolytic family (literatures or database)[c]
Dataset 2				
499EST	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
7N9	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
ABO_1197	ABHD6-Lip	ABHD6-Lip (V)	V	V
ABO_1251	Carb_B_Bacteria	Carb_B_Bacteria (VII)	VII	VII
ABO0195	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
CaesCCR11	CarbLipBact_2	CarbLipBact_2 (XIII-2)	XIII	XIII
E69	GTSAGmotif	GTSAGmotif (IV)	IV	IV
EaEST	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	nonlipolytic protein	V
EM3L2	Duf_3089	Duf_3089 (XV)	XV	XV
EM3L4	Esterase_phb	Lysophospholipase_carboxylesterase (VI)	unassigned	EM3L4 (novel family)
EML1	Lipase_3	Lipase_3 (XI)	XI	XI
Est_p1	Aclacinomycin-methylesterase_RdmC	Carboxymethylbutenolide_lactonase (V)	nonlipolytic protein	V
Est1	6_AlphaBeta_hydrolase	Carboxymethylbutenolide_lactonase (V)	unassigned	V
Est10	LYsophospholipase_carboxylesterase	Lysophospholipase_carboxylesterase (VI)	VI	VI
Est12	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
Est22	Homoserine_transacetylase	Homoserine_transacetylase (Est22)	Est22	Est22
Est25	GTSAGmotif	GTSAGmotif (IV)	IV	IV
Est30	CarbLipBact_1	CarbLipBact_1 (XIII-1/XVIII)	XIII	XIII
Est4	6_AlphaBeta_hydrolase	ABHD6-Lip (V)	unassigned	V
Est40	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
Est56	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
Est700	Lipase_2	Bacterial_lipase (I)	I	I
Est903	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
Est97	Bacterial_Est97	Bacterial_Est97 (XVI)	XVI	XVI
Est9x	Est9X	Est9X (Est9X)	Est9X	Est9X
EstA	A85-EsteraseD-FGH	A85-EsteraseD-FGH (EstA/FLS12/lp_3505)	EstA	EstA
EstA3	PC-sterol_acyltransferase	Pcsterol_acyltransferase (XIV)	XIV	XIV
EstAM	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
EstATII	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
EstCS2	Carb_B_Bacteria	Carb_B_Bacteria (VII)	VII	VII
EstD	Bacterial_EstLip_FamX	Bacterial_EstLip_FamX (X)	X	X
EstD2	Duf_3089	Duf_3089 (XV)	XV	XV
EstDZ2	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	EstDZ2	EstDZ2
EstEP16	CarbLipBact_2	CarbLipBact_2 (XIII-2)	XIII	XIII
EstF27	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
EstGK1	Duf_3089	Duf_3089 (XV)	XV	XV

EstGS	Chlorophyllase	Chlorophyllase (EstGS)	EstGS	EstGS
EstGtA2	CarbLipBact_2	CarbLipBact_2 (XIII-2)	XIII	XIII
EstJ	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
EstKT4	GTSAGmotif	GTSAGmotif (IV)	IV	IV
EstKT7	GTSAGmotif	GTSAGmotif (IV)	IV	IV
EstKT9	GTSAGmotif	GTSAGmotif (IV)	IV	IV
EstL28	6_AlphaBeta_hydrolase	Carboxymethylbutenolide_lactonase (V)	unassigned	EstL28 (novel family)
EstLiu	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
EstMY	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
EstMY09-2	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
EstN7	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
EstOF4	CarbLipBact_1	CarbLipBact_1 (XIII-1/XVIII)	XIII	XIII
Est-OKK	ABHD11-Acetyl_transferase	ABHD6-Lip (V)	nonlipolytic protein	V
EstP2K	6_AlphaBeta_hydrolase	Bacterial_lipase (I)	unassigned	EstP2K (novel family)
EstS	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
EstSP	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
EstSt7	A85-Feruloyl-Esterase	A85-Feruloyl-Esterase (EstSt7/Rlip1)	EstSt7	EstSt7
EstUT1	CarbLipBact_1	CarbLipBact_1 (XIII-1/XVIII)	XVIII	XVIII
EstWSD	Duf_3089	Duf_3089 (XV)	XV	XV
Est-XG2	Carb_B_Bacteria	Carb_B_Bacteria (VII)	VII	VII
EstY	Fungal-Bact_LIP	Fungal_Bact_LIP (X-2/XVII/XIX)	X	X
EstZ3	Duf_3089	Duf_3089 (XV)	XV	XV
FLS12	A85-EsteraseD-FGH	A85-EsteraseD-FGH (EstA/FLS12/lp_3505)	FLS12	FLS12
FLS18	5_AlphaBeta_hydrolase	A85-Feruloyl-Esterase (EstSt7/Rlip1)	unassigned	FLS18 (novel family)
FnL	AlphaBeta_hydrolase	Pcsterol_acyltransferase (XIV)	unassigned	FnL (novel family)
GDEst-95	Carb_B_Bacteria	Carb_B_Bacteria (VII)	VII	VII
H8	Aclacinomycin-methylesterase_RdmC	Carboxymethylbutenolide_lactonase (V)	nonlipolytic protein	V
H9Est	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
LAE3	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
LAE4	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
LAE6	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
Lip10	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
Lip-1452	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
Lip3	Lipase_3	Lipase_3 (XI)	XI	XI
Lip-948	ABHD6-Lip	ABHD6-Lip (V)	V	V
LipA	PGAP1	Bacterial_lipase (I)	unassigned	I
Lipab15	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
Lipab18	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
Lipab4	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
Lipab8	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
Lipag18	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
LipC	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
LipC12	Bacterial_lipase	Bacterial_lipase (I)	I	I
LipEH166	Bact_LipEH166_FamXII	Bact_LipEH166_FamXII (XII)	XII	XII
LipG	Lipase_3	Lipase_3 (XI)	XI	XI

LipJ2	Fungal-Bact_LIP	Fungal_Bact_LIP (X-2/XVII/XIX)	XVII	XVII
Lipo1	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
Lipo11	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
Lipo12A	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
Lipo2	6-AlphaBeta_hydrolase	Lipase_2 (I)	unassigned	I
Lipo3	ABHD11-Acetyl_transferase	ABHD6-Lip (V)	unassigned	V
Lipo4A	LYsophospholipase_carboxylesterase	Lysophospholipase_carboxylesterase (VI)	VI	VI
Lipo5	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
Lipo6	LYsophospholipase_carboxylesterase	Lysophospholipase_carboxylesterase (VI)	VI	VI
LipR	Fungal-Bact_LIP	Fungal_Bact_LIP (X-2/XVII/XIX)	X	X
LipR_2	Fungal-Bact_LIP	Fungal_Bact_LIP (X-2/XVII/XIX)	X	X
LipR1	Bacterial_lipase	Bacterial_lipase (I)	I	I
LipR3	Bacterial_lipase	Bacterial_lipase (I)	I	I
LipS	CarbLipBact_2	CarbLipBact_2 (XIII-2)	XIII	XIII
LipSm	Fungal-Bact_LIP	Fungal_Bact_LIP (X-2/XVII/XIX)	XIX	XIX
LipSM54	Cocaine_esterase	Hormone_sensitive_lipase_like (IV)	unassigned	LipSM54 (novel family)
LipT	Pectinacetylesterase-Notum	Pectinacetylesterase-Notum (LipT)	LipT	LipT
LipYY31	Bacterial_lipase	Bacterial_lipase (I)	I	I
lp_3505	A85-EsteraseD-FGH	A85-EsteraseD-FGH (EstA/FLS12/lp_3505)	lp_3505	lp_3505
M37	Lipase_3	Lipase_3 (XI)	XI	XI
MGS0018	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
MGS0153	ABHD6-Lip	ABHD6-Lip (V)	V	V
MGS-B1	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
MGS-K1	Carb_B_Bacteria	Carb_B_Bacteria (VII)	VII	VII
MGS-M1	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
MGS-M2	6_AlphaBeta_hydrolase	Carboxymethylbutenolide_lactonase (V)	unassigned	V
MGS-MG1	AlphaBeta_hydrolase	Bacterial_EstLip_FamX (X)	unassigned	unclassified
MGS-MT1	GTSAGmotif	GTSAGmotif (IV)	IV	IV
MGS-RG1	LYsophospholipase_carboxylesterase	Lysophospholipase_carboxylesterase (VI)	VI	VI
MGS-RG2	LYsophospholipase_carboxylesterase	Lysophospholipase_carboxylesterase (VI)	VI	VI
MGS-RG3	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
MtEst45	Cocaine_esterase	Chlorophyllase (EstGS)	unassigned	MtEst45 (novel family)
PE10	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
PhaZ7	PHAZ7_phb_depolymerase	PHAZ7_phb_depolymerase (IX)	IX	IX
PMGL2	GTSAGmotif	GTSAGmotif (IV)	IV	IV
REst1	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
Rlip1	A85-Feruloyl-Esterase	A85-Feruloyl-Esterase (EstSt7/Rlip1)	Rlip1	Rlip1
Rv0045c	6_AlphaBeta_hydrolase	ABHD6-Lip (V)	unassigned	Rv0045c (novel family)
SAestA	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
SBLip2	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
SBLip5.1	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
ThaEst2349	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
Vlip509	UCP031982	UCP031982 (V)	V	V
EstGH	ND [d]	ND [d]	ND [d]	EstGH (novel family)
EstGX1	ND [d]	ND [d]	ND [d]	EstGX1 (novel family)

Lipo10	ND [d]	ND [d]	ND [d]	unclassified
Lipo7A	ND [d]	ND [d]	ND [d]	unclassified
Lipo7B	ND [d]	ND [d]	ND [d]	unclassified
Dataset 3				
EstC55-2	Carbon-carbon_bond_hydrolase	ABHD6-Lip (V)	unassigned	V
EstC55-3	Carb_B_Bacteria	Carb_B_Bacteria (VII)	VII	VII
EstC55-4_1	6_AlphaBeta_hydrolase	Carboxymethylbutenolide_lactonase (V)	unassigned	V
EstC55-5	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
EstC55-8_1	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
EstC55-8_2	Epoxide_hydrolase	Carboxymethylbutenolide_lactonase (V)	unassigned	V
EstC55-12	6_AlphaBeta_hydrolase	Carboxymethylbutenolide_lactonase (V)	unassigned	V
EstC55-13	PAF-Acetylhydrolase	Chlorophyllase (EstGS)	unassigned	new subfamily (this study)
EstC55-15	AlphaBeta_hydrolase	CarbLipBact_2 (XIII-2)	unassigned	new family (this study)
EstC55-18	BioH	Carboxymethylbutenolide_lactonase (V)	unassigned	V
EstC55-19_1	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
EstC55-19_2	Epoxide_hydrolase	ABHD6-Lip (V)	unassigned	V
EstC55-20	6_AlphaBeta_hydrolase	Carboxymethylbutenolide_lactonase (V)	unassigned	V
EstC55-23	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
EstC55-24	Chlorophyllase	Chlorophyllase (EstGS)	EstGS	EstGS
EstC55-25	6_AlphaBeta_hydrolase	Carboxymethylbutenolide_lactonase (V)	unassigned	V
EstC55-31	6_AlphaBeta_hydrolase	Carboxymethylbutenolide_lactonase (V)	unassigned	V
EstC55-34	Carboxymethylbutenolide_lactonase	Carboxymethylbutenolide_lactonase (V)	V	V
EstC55-38	Pectinacetylesterase-Notum	Pectinacetylesterase-Notum (LipT)	LipT	LipT
EstC55-42	Esterase_phb	Lysophospholipase_carboxylesterase (VI)	unassigned	EM3L4
EstC55-51	6_AlphaBeta_hydrolase	ABHD6-Lip (V)	unassigned	V
EstC55-52	Carb_B_Bacteria	Carb_B_Bacteria (VII)	VII	VII
EstC55-56	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
EstC55-57	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
EstC55-60	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
EstC55-62	Carb_B_Bacteria	Carb_B_Bacteria (VII)	VII	VII
EstC55-71	Lipase_2	Bacterial_lipase (I)	I	I
EstC55-72	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
EstC55-76	6_AlphaBeta_hydrolase	Carboxymethylbutenolide_lactonase (V)	unassigned	V
EstC55-77	Esterase_phb	Lysophospholipase_carboxylesterase (VI)	unassigned	EM3L4
EstC55-78	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
EstC55-81	6_AlphaBeta_hydrolase	Carboxymethylbutenolide_lactonase (V)	unassigned	V
EstC55-88	Lipase_2	Bacterial_lipase (I)	I	I
EstC55-90	Lipase_2	Bacterial_lipase (I)	I	I
EstC55-95	Polyesterase-lipase-cutinase	Polyesterase-lipase-cutinase (III)	III	III
EstC55-96	ABHD11-Acetyl_transferase	Carboxymethylbutenolide_lactonase (V)	nonlipolytic α/β hydorlase	V
EstC55-100	6_AlphaBeta_hydrolase	Carboxymethylbutenolide_lactonase (V)	unassigned	V
EstC55-102	Lipase_bact_N_lipase	Chlorophyllase (EstGS)	unassigned	new family (this study)
EstC55-105	Lipase_2	Bacterial_lipase (I)	I	I

EstC55-118	Carb_B_Bacteria	Carb_B_Bacteria (VII)	VII	VII
EstC55-145	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
EstC55-151	Lipase_2	Bacterial_lipase (I)	I	I
EstC55-154	Fungal-Bact_LIP	Fungal_Bact_LIP (X-2/XVII/XIX)	XVII	XVII
EstC55-156	Tannase	Tannase (Tannase)	Tannase	Tannase
EstC55-159	6_AlphaBeta_hydrolase	ABHD6-Lip (V)	unassigned	V
EstC55-165	5_AlphaBeta_hydrolase	A85-Feruloyl-Esterase (EstSt7/Rlip1)	unassigned	FLS18
EstC55-169	Tannase	Tannase (Tannase)	Tannase	Tannase
EstC55-186	Carb_B_Bacteria	Carb_B_Bacteria (VII)	VII	VII
EstC55-188	AlphaBeta_hydrolase	Lysophospholipase_carboxylesterase (VI)	unassigned	new family (this study)
EstC55-197	Aclacinomycin-methylesterase_RdmC	Carboxymethylbutenolide_lactonase (V)	unassigned	V
EstC55-213	PGAP1	Bacterial_lipase (I)	unassigned	new subfamily (this study)
EstC55-215	6_AlphaBeta_hydrolase	Carboxymethylbutenolide_lactonase (V)	unassigned	V
EstC55-229	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
EstC55-231	6_AlphaBeta_hydrolase	Carboxymethylbutenolide_lactonase (V)	unassigned	V
EstC55-234	Tannase	Tannase (Tannase)	Tannase	Tannase
EstC55-235	Lipase_2	Bacterial_lipase (I)	I	I
EstC55-241	5_AlphaBeta_hydrolase	A85-Feruloyl-Esterase (EstSt7/Rlip1)	unassigned	FLS18
EstC55-244	6_AlphaBeta_hydrolase	Carboxymethylbutenolide_lactonase (V)	unassigned	V
EstC55-247	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
EstC55-253	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
EstC55-256	6_AlphaBeta_hydrolase	Carboxymethylbutenolide_lactonase (V)	unassigned	V
EstC55-268	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
EstC76-28_1	6_AlphaBeta_hydrolase	Carboxymethylbutenolide_lactonase (V)	unassigned	V
EstC76-28_2	6_AlphaBeta_hydrolase	ABHD6-Lip (V)	unassigned	V
EstC76-36	AlphaBeta_hydrolase	Lysophospholipase_carboxylesterase (VI)	unassigned	new family (this study)
EstC76-135	Hormone-sensitive_lipase_like	Hormone_sensitive_lipase_like (IV)	IV	IV
EstC76-136	Carb_B_Bacteria	Carb_B_Bacteria (VII)	VII	VII
EstC76-137	5_AlphaBeta_hydrolase	A85-Feruloyl-Esterase (EstSt7/Rlip1)	unassigned	FLS18
EstC76-177	6_AlphaBeta_hydrolase	Lipase_2 (I)	unassigned	new subfamily (this study)
EstC76-179	Pectinacetylesterase-Notum	Pectinacetylesterase-Notum (LipT)	LipT	LipT
EstC76-202	BioH	Carboxymethylbutenolide_lactonase (V)	unassigned	V
EstC76-218	Pectinacetylesterase-Notum	Pectinacetylesterase-Notum (LipT)	LipT	LipT
EstC76-221	AlphaBeta_hydrolase	Bacterial_EstLip_FamX (X)	unassigned	new family (this study)
EstC76-248	6_AlphaBeta_hydrolase	Carboxymethylbutenolide_lactonase (V)	unassigned	V
EstC76-263	6_AlphaBeta_hydrolase	Carboxymethylbutenolide_lactonase (V)	unassigned	V
EstC76-266	6_AlphaBeta_hydrolase	Carboxymethylbutenolide_lactonase (V)	unassigned	V
EstC55-227	ND [d]	ND [d]	ND [d]	new family (this study)
EstC55-97	ND [d]	ND [d]	ND [d]	new family (this study)
EstC55-167	ND [d]	ND [d]	ND [d]	new family (this study)
Dataset 4				
L7VEQ3	Carbon-carbon_bond_hydrolase	ABHD6-Lip (V)	Non-lipolytic protein	ND [d]
A0A0U0R506	Carbon-carbon_bond_hydrolase	ABHD6-Lip (V)	Non-lipolytic protein	ND [d]

A0A2X1S8R4	Carbon-carbon_bond_hydrolase	ABHD6-Lip (V)	Non-lipolytic protein	ND [d]
A0A0T9VGE3	Carbon-carbon_bond_hydrolase	ABHD6-Lip (V)	Non-lipolytic protein	ND [d]
A0A1S7N7P8	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
A0A1S7T7D7	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
A0A1S7T6S9	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
A0A1S7N8N9	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
A0A1S7N8S2	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
A0A1S7S3Y6	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
A0A1S7NDN6	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
A0A2Z5YBG9	6_AlphaBeta_hydrolase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
A0A2N9APQ2	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
G4SWJ2	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
A0A0N7H8D0	6_AlphaBeta_hydrolase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
A0A102S299	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
A0A375DNQ7	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
A0A375EC98	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
A0A375FEC6	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
A0A375D6M7	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
A0A375IUU4	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
A0A1S7RCN8	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
A0A1S7R0G9	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
A0A1S7RQ23	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
A0A1S7N1P6	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
A0A1S7R2U2	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
A0A1S7R186	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
A0A1S7S3E0	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
A0A1S7TAY0	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
A0A379M130	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
A0A1S7TZP1	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
A0A375DP21	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
A0A1S7RZV4	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
A0A1S7QF97	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
A0A1S7QYM3	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
A0A1S7UBZ3	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
A0A1X2DNB5	6_AlphaBeta_hydrolase	Carboxymethylbutenolide_lactonase (V)	unassigned	ND [d]
A0A1S7R6I3	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
A0A2X4UF21	Haloperoxidase	ABHD6-Lip (V)	Non-lipolytic protein	ND [d]
A0A1S7N0Y4	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
A0A1S7RYV9	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
A0A1S7SHR8	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
A0A378W1U8	6_AlphaBeta_hydrolase	Carboxymethylbutenolide_lactonase (V)	unassigned	ND [d]
A0A1S7S3D1	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
A0A1S7MUS0	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
A0A1S7MPM6	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
A0A375EBJ1	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]

F8JH88	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
F8JA98	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
A0A2N9AMF4	Haloperoxidase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
ALO47455	Epoxide_hydrolase	ABHD6-Lip (V)	unassigned	ND [d]
PNG95674	Epoxide_hydrolase	Carboxymethylbutenolide_lactonase (V)	Non-lipolytic protein	ND [d]
PNG97285	Epoxide_hydrolase	ABHD6-Lip (V)	Non-lipolytic protein	ND [d]
KDN82212	Epoxide_hydrolase	ABHD6-Lip (V)	unassigned	ND [d]
KIE51135	Epoxide_hydrolase	Carboxymethylbutenolide_lactonase (V)	unassigned	ND [d]
KIE52625	Epoxide_hydrolase	ABHD6-Lip (V)	Non-lipolytic protein	ND [d]

[a] Protein sequences were searched against the whole ESTHER database

[b] Lipolytic family assignment using the strategy developed in this study; unassigned, could not assigned to any lipolytic family

[c] Lipolytic family assignment reported in literatures (or this study) using pylogenetic-related methods

[d] No data

Supplementary Table S13 Validation of the liplolytic family assignment strategy with LEs affiliated to family II, VIII and patatin-like-protein

Protein name	CATH HMMs annotation	pHMM-pfam annotation	Lipolytic family (This study)[a]	Lipolytic family (literatures or database)[b]
Dataset 1				
A0A1B0QVN6	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
A0A248RGG8	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
A0A2C9EHI5	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
B5TWC2	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
D4N4E9	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
F4MYP0	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
F7Q845	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
I7CDN7	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
J2YNH3	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
K4HQE7	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
Q48LQ9	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
Q4KH73	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
Q56CZ4	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
Q9KX30	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
S5Y3D1	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
A0A024H3L8	Class C beta-lactamase CMY-10	Beta-lactamase (VIII)	VIII	VIII
A0A031ILE0	Class C beta-lactamase CMY-10	Beta-lactamase (VIII)	VIII	VIII
A0A0G2RKR9	Class C beta-lactamase CMY-10	Beta-lactamase (VIII)	VIII	VIII
A0A0H3C4U3	Class C beta-lactamase CMY-10	Beta-lactamase (VIII)	VIII	VIII
A0A0N7CSD6	Class C beta-lactamase CMY-10	Beta-lactamase (VIII)	VIII	VIII
A0A0N9R483	Class C beta-lactamase CMY-10	Beta-lactamase (VIII)	VIII	VIII
A1RB78	Class C beta-lactamase CMY-10	Beta-lactamase (VIII)	VIII	VIII
A4F8E6	Class C beta-lactamase CMY-10	Beta-lactamase (VIII)	VIII	VIII
A4FFW1	Class C beta-lactamase CMY-10	Beta-lactamase (VIII)	VIII	VIII
A9WQD8	Class C beta-lactamase CMY-10	Beta-lactamase (VIII)	VIII	VIII
B0M0H4	Class C beta-lactamase CMY-10	Beta-lactamase (VIII)	VIII	VIII
D5P454	Class C beta-lactamase CMY-10	Beta-lactamase (VIII)	VIII	VIII
E1VYY4	Class C beta-lactamase CMY-10	Beta-lactamase (VIII)	VIII	VIII
G8PW82	Class C beta-lactamase CMY-10	Beta-lactamase (VIII)	VIII	VIII
I4KBV5	Class C beta-lactamase CMY-10	Beta-lactamase (VIII)	VIII	VIII
I4KTB2	Class C beta-lactamase CMY-10	Beta-lactamase (VIII)	VIII	VIII
I4KYM0	Class C beta-lactamase CMY-10	Beta-lactamase (VIII)	VIII	VIII
J2F4F4	Class C beta-lactamase CMY-10	Beta-lactamase (VIII)	VIII	VIII
K9NGV2	Class C beta-lactamase CMY-10	Beta-lactamase (VIII)	VIII	VIII
Q8VU79	Class C beta-lactamase CMY-10	Beta-lactamase (VIII)	VIII	VIII

A0A068QRN7	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A068QRU1	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A068R3C8	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A068R3W3	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A068R5Z7	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A077P070	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A077P6A6	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A077PR36	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A077QJI8	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A077QNW8	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A085G4J8	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A0A8NSM3	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A0B7DJ26	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A0D5XSS9	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A0F7Y552	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A0G3SQ37	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A0G5NCG5	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A0H5LZ72	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A0R4FM75	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A0S4IA49	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A0T9KAX4	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A0X8XW81	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A109KKX4	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A126VD94	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A145P5E7	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A168FS98	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A1B8YLL8	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A1C0U1I1	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A1C3HL68	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A1N6MX70	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A1N6MXA7	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A1W5DEF9	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A240A646	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A2H1L984	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A8G7T5	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
B1JE37	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
B4STS4	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
B6VK93	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
D2BTP5	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
D3V7P0	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
D3VEQ2	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
D4DZP7	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
G8Q0A2	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
I4JZT9	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
I4L7G6	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II

J2MF23	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
N1NPB9	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
P40601	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
S2F0I5	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
W1IPR6	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
W1ITZ2	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
W1IUL6	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
W6VSR6	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A011NHS5	GDSL esterase/lipase APG	Lipase_GDSL (II)	II	II
A0A011NNZ6	GDSL esterase/lipase APG	Lipase_GDSL (II)	II	II
A0A177W0K3	GDSL esterase/lipase APG	Lipase_GDSL (II)	II	II
A0A1J5Q6R0	GDSL esterase/lipase APG	Lipase_GDSL (II)	II	II
A0A1J5QST6	GDSL esterase/lipase APG	Lipase_GDSL (II)	II	II
A0A1R4EFW9	GDSL esterase/lipase APG	Lipase_GDSL (II)	II	II
A0A1R4GRG6	GDSL esterase/lipase APG	Lipase_GDSL (II)	II	II
A0A1T0A9C3	GDSL esterase/lipase APG	Lipase_GDSL (II)	II	II
A0A1U6GL33	GDSL esterase/lipase APG	Lipase_GDSL (II)	II	II
A0A1Y5Q328	GDSL esterase/lipase APG	Lipase_GDSL (II)	II	II
A0A2H4UKY2	GDSL esterase/lipase APG	Lipase_GDSL (II)	II	II
A5WHX6	GDSL esterase/lipase APG	Lipase_GDSL (II)	II	II
A7IIP4	GDSL esterase/lipase APG	Lipase_GDSL (II)	II	II
B2I9C9	GDSL esterase/lipase APG	Lipase_GDSL (II)	II	II
D5VAI8	GDSL esterase/lipase APG	Lipase_GDSL (II)	II	II
E6WQ95	GDSL esterase/lipase APG	Lipase_GDSL (II)	II	II
G6FMT2	GDSL esterase/lipase APG	Lipase_GDSL (II)	II	II
K9XFG2	GDSL esterase/lipase APG	Lipase_GDSL (II)	II	II
Q3JVI5	GDSL esterase/lipase APG	Lipase_GDSL (II)	II	II
Q7X4K7	GDSL esterase/lipase APG	Lipase_GDSL (II)	II	II
A0A024HQ86	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A080VLY1	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A0D0T6T3	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A0F6UI63	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A0H3YNS8	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A0K0Q1Y2	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A0P8X538	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A0S4HWF7	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A120G8H3	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A127N102	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A157V5L0	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A193SIR7	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A1Y6JH94	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
E3VST9	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
I4KAN1	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
J2FAM4	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
K9NS61	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II

O33407	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
Q6B6R8	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
Q938A9	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
V6ANT9	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
W0HFT8	Autotransporting lipase, GDSL family	Lipase_GDSL (II)	II	II
A0A0S2KQT2	GDSL esterase/lipase APG	Lipase_GDSL_2 (II)	II	II
A0A1V6JD04	GDSL esterase/lipase APG	Lipase_GDSL_2 (II)	II	II
A0A246KHV9	GDSL esterase/lipase APG	Lipase_GDSL_2 (II)	II	II
G0HD65	GDSL esterase/lipase APG	Lipase_GDSL_2 (II)	II	II
Q3JQU3	GDSL esterase/lipase APG	Lipase_GDSL_2 (II)	II	II
A0A0P0FMI4	Isoamyl acetate-hydrolyzing esterase 1 homolog	Lipase_GDSL_2 (II)	II	II
A0A135YTB5	Isoamyl acetate-hydrolyzing esterase 1 homolog	Lipase_GDSL_2 (II)	II	II
A0A139JRQ6	Isoamyl acetate-hydrolyzing esterase 1 homolog	Lipase_GDSL_2 (II)	II	II
A0A143PQQ6	Isoamyl acetate-hydrolyzing esterase 1 homolog	Lipase_GDSL_2 (II)	II	II
A0A173Z3Y1	Isoamyl acetate-hydrolyzing esterase 1 homolog	Lipase_GDSL_2 (II)	II	II
A0A174CNS8	Isoamyl acetate-hydrolyzing esterase 1 homolog	Lipase_GDSL_2 (II)	II	II
A0A174H821	Isoamyl acetate-hydrolyzing esterase 1 homolog	Lipase_GDSL_2 (II)	II	II
A0A174NAV5	Isoamyl acetate-hydrolyzing esterase 1 homolog	Lipase_GDSL_2 (II)	II	II
A0A174UHE3	Isoamyl acetate-hydrolyzing esterase 1 homolog	Lipase_GDSL_2 (II)	II	II
A0A1E8F0P5	Isoamyl acetate-hydrolyzing esterase 1 homolog	Lipase_GDSL_2 (II)	II	II
A0A1P8WFW9	Isoamyl acetate-hydrolyzing esterase 1 homolog	Lipase_GDSL_2 (II)	II	II
A0A1V5F586	Isoamyl acetate-hydrolyzing esterase 1 homolog	Lipase_GDSL_2 (II)	II	II
A0A1V5G6A1	Isoamyl acetate-hydrolyzing esterase 1 homolog	Lipase_GDSL_2 (II)	II	II
A0A1V5HGJ2	Isoamyl acetate-hydrolyzing esterase 1 homolog	Lipase_GDSL_2 (II)	II	II
A0A1V5UHU4	Isoamyl acetate-hydrolyzing esterase 1 homolog	Lipase_GDSL_2 (II)	II	II
A0A1V5V5J1	Isoamyl acetate-hydrolyzing esterase 1 homolog	Lipase_GDSL_2 (II)	II	II
A0A1V6JUN0	Isoamyl acetate-hydrolyzing esterase 1 homolog	Lipase_GDSL_2 (II)	II	II
A0A2A3N2S7	Isoamyl acetate-hydrolyzing esterase 1 homolog	Lipase_GDSL_2 (II)	II	II
F3Y895	Isoamyl acetate-hydrolyzing esterase 1 homolog	Lipase_GDSL_2 (II)	II	II
S5XQR0	Isoamyl acetate-hydrolyzing esterase 1 homolog	Lipase_GDSL_2 (II)	II	II
A0A009FQY3	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A009GH23	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A009HL94	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A009L278	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A009MNS1	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A009RMH5	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A009YG53	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A010IZF7	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A010L6U0	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A010WGV4	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A011K3K9	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A011M4E2	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A011NRW2	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A011P1E4	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A011P216	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II

A0A011PQC4	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A011Q348	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A013S8D5	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A013TPS8	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A014B4I3	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A014BP77	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A014BVI3	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A014CG41	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A014DCG5	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A014DNG3	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A014DW77	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A014F485	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A022I9R1	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A022J7T0	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A022KIF1	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A024HK50	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A061D1S0	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A062BV08	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A062GRT9	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A062LUI1	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A062N102	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A062SY13	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A078BGB8	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A080LZ83	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A080M659	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A098G8E8	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A0A8RH78	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A0B5FA02	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A0B7DFD2	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A0C6F7I8	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A0D0T8I6	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A0D6H809	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A0F7XZS5	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A0G6AHG2	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A0H4WD71	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A0J6CCW4	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A0K1J6Z2	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A0K2G755	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A0K2ZPA4	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A0K2ZYS9	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A0K2ZZJ3	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A0M2WGX5	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A0M7HJ86	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A0M9IS63	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A0N1JS35	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II

A0A0P0M8K2	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A0P0MG90	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A0P9MUM9	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A0Q0MJ90	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A0Q9YFU6	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A0Q9YN66	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A0S4HZN8	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A0S4I4I4	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A0S4KTX7	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A0S4L7Q0	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A0T7QZN4	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A0T8L2J5	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A0W0S4P5	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A0W0UTZ0	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A0W0XMR7	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A0W0YWH7	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A0W1AA95	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A0X8HBX8	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A0X8X153	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A109KRK3	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A109LCQ0	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A119A1M5	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A127MQK8	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A150HXT3	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A150I107	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A157KKJ6	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A157R8P7	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A157SW77	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A157WLJ5	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A177YYG5	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A193SRH1	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A1A8TGJ1	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A1A8TLQ4	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A1B8NWA4	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A1C3JU89	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A1C9W8B0	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A1D3JUB8	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A1D8AVV5	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A1E2ZHS6	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A1E3GPF1	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A1E7VQY7	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A1E7WA72	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A1E7WRQ8	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A1E7X2W6	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A1G5SHI5	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II

A0A1J5Q6Y2	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A1J5R2D9	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A1J5RJY5	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A1J5S9V8	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A1J5SBR3	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A1J5SZP4	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A1J5TMA6	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A1M9K390	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A1N7SPK8	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A1Q9QUR5	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A1R4H672	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A1R7Q8F7	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A1V5FB52	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A1V5FZ99	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A1V5QVA5	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A1W1I3M1	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A1Y0N3S6	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A1Y6JN14	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A221V155	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A222P1J4	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A238DU51	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A239RYM4	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A239SIU4	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A2H5XLT0	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A2H6A0X2	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
K1LKT4	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
M7MWS1	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
Q9HZY8	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
V4XZF9	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
V5BQM9	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
V6ACN8	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
W5YPF5	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
W7W0N3	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
W7WJF7	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
A0A089X6P7	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A0K2APQ7	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A0K2YGH5	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A0N1FSK5	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A0N1G8S1	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A0N1NB57	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A0N1NV23	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A0T9M857	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A0T9MCT6	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A0U5LEC0	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A0U5LH42	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II

A0A0U5LIQ2	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A0U5LZ22	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A100JAY0	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A100JN53	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A101RXH9	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A117EEV0	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A124C2F1	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A161I105	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A177HFK3	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A177HFS2	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A177HPT5	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A177HYI6	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A178X1Y4	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A178XEM1	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A1A9GGR0	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A1A9GRX0	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A1B2H079	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A1B9EN57	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A1B9EP94	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A1B9EUB5	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A1D2IIH5	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A1D8C1F8	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A1D8FY73	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A1K2FJY5	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A1K2FK87	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A1K2FWE4	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A1K2G0G8	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A1P8Y9C1	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A1Q2ZMI7	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A1Q2ZNA0	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A1U2G8U4	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A1U2UMR2	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A1V2MND5	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A1V2MS28	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A1V2MTK7	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A1V2RA89	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A1V2RFC3	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A1Y2MP54	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A1Y2N1Y0	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A1Y2NGJ5	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A1Y2NN23	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A221W7P5	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A222TGI3	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A222TK45	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A222TQN6	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II

A0A231GU95	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A250VES8	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0ACN1	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A3KIR8	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
D1A8A5	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
D2AXW6	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
D2PMS0	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
D3CSD1	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
D3EZZ4	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
D3FC90	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
D3PZ28	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
D3Q6I6	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
D3Q953	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
D6EML8	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
D6EUD8	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
D9T622	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
D9T737	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
E8W5N2	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
E8WF38	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
F6FXF1	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
G0Q5X5	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
G0Q684	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
G2NLA7	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
G8S919	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
G8SAJ2	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
K0K0J4	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
K0KD57	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
K4RCW1	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
N0CT46	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
N0CWZ8	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
Q93J06	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
Q93MW7	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
Q9S2A5	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
W5W553	Putative secreted hydrolase	Lipase_GDSL_2 (II)	II	II
A0A066WSZ1	YMR313Cp-like protein	Patatin	patatin-like-protein	patatin-like-protein
A0A0L8V990	YMR313Cp-like protein	Patatin	patatin-like-protein	patatin-like-protein
A0A1S5VH73	YMR313Cp-like protein	Patatin	patatin-like-protein	patatin-like-protein
W2UM92	YMR313Cp-like protein	Patatin	patatin-like-protein	patatin-like-protein
A0A0A3W7A6	Patatin-like phospholipase domain-containing protein	Patatin	patatin-like-protein	patatin-like-protein
A0A0X8R679	Patatin-like phospholipase domain-containing protein	Patatin	patatin-like-protein	patatin-like-protein
B5JS88	Patatin-like phospholipase domain-containing protein	Patatin	patatin-like-protein	patatin-like-protein
B8KFQ3	Patatin-like phospholipase domain-containing protein	Patatin	patatin-like-protein	patatin-like-protein
B8KQY9	Patatin-like phospholipase domain-containing protein	Patatin	patatin-like-protein	patatin-like-protein
B8KXA8	Patatin-like phospholipase domain-containing protein	Patatin	patatin-like-protein	patatin-like-protein

Dataset 2

AcXE2	Isoamyl acetate-hydrolyzing esterase 1 homolog	Lipase_GDSL_2 (II)	II	II
Est01	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
Est2	Class C beta-lactamase CMY-10	Beta-lactamase (VIII)	VIII	VIII
Est22_2	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
Est7K	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
EstA3_2	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
EstBL	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
EstC	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
EstCE1	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
EstF4K	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
EstHE1	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
EstM-N1	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
EstM-N2	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
EstQE	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
EstSL3	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
EstSTR1	Class C beta-lactamase CMY-10	Beta-lactamase (VIII)	VIII	VIII
EstU1	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
LAE2	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
LAE5	Isoamyl acetate-hydrolyzing esterase 1 homolog	Lipase_GDSL_2 (II)	II	II
LAE7	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
LipA9	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
LipBL	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
Lipo12B	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
Lipo13	Isoamyl acetate-hydrolyzing esterase 1 homolog	Lipase_GDSL_2 (II)	II	II
Lipo4B	YMR313Cp-like protein	Patatin (patatin-like-protein)	patatin-like-protein	patatin-like-protein
Lipo4C	ND [c]	Patatin (patatin-like-protein)	unassigned	patatin-like-protein
Lipo8	Acyl-CoA thioesterase I	Lipase_GDSL_2 (II)	II	II
Lpc53E1	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
MGS0010	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
MGS0105	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
PLP	YMR313Cp-like protein	Patatin (patatin-like-protein)	patatin-like-protein	patatin-like-protein
SBLip1	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
Dataset 3				
EstC55-4_2	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
EstC55-6	ND [c]	Patatin (patatin-like-protein)	unassigned	patatin-like-protein
EstC55-7	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
EstC55-10	ND [c]	Patatin (patatin-like-protein)	unassigned	patatin-like-protein
EstC55-26	YMR313Cp-like protein	Patatin (patatin-like-protein)	patatin-like-protein	patatin-like-protein
EstC55-40	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII

EstC55-46	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
EstC55-53	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
EstC55-61	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
EstC55-63	YMR313Cp-like protein	Patatin (patatin-like-protein)	patatin-like-protein	patatin-like-protein
EstC55-65	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
EstC55-66	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
EstC55-73	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
EstC55-80	Class C beta-lactamase CMY-10	Beta-lactamase (VIII)	VIII	VIII
EstC55-110	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
EstC55-111	GDSL esterase/lipase APG	Lipase_GDSL (II)	II	II
EstC55-113	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
EstC55-131	YMR313Cp-like protein	Patatin (patatin-like-protein)	patatin-like-protein	patatin-like-protein
EstC55-147	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
EstC55-150	GDSL esterase/lipase APG	Lipase_GDSL_2 (II)	II	II
EstC55-163	ND [c]	Patatin (patatin-like-protein)	unassigned	patatin-like-protein
EstC55-164	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
EstC55-168	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
EstC55-239	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
EstC55-245	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
EstC55-251	YMR313Cp-like protein	Patatin (patatin-like-protein)	patatin-like-protein	patatin-like-protein
EstC55-258	Class C beta-lactamase CMY-10	Beta-lactamase (VIII)	VIII	VIII
EstC76-21	ND [c]	Patatin (patatin-like-protein)	unassigned	patatin-like-protein
EstC76-98	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
EstC76-123	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
EstC76-174	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
EstC76-222	YMR313Cp-like protein	Patatin (patatin-like-protein)	patatin-like-protein	patatin-like-protein
EstC76-250	YMR313Cp-like protein	Patatin (patatin-like-protein)	patatin-like-protein	patatin-like-protein
EstC76-261	ND [c]	Patatin (patatin-like-protein)	unassigned	patatin-like-protein
EstC76-262	Beta-lactamase (Penicillin-binding protein) (Penicillinase)	Beta-lactamase (VIII)	VIII	VIII
EstC76-269	Neuropathy target esterase sws	Patatin (patatin-like-protein)	patatin-like-protein	patatin-like-protein

[a] Lipolytic family assignment using the strategy developed in this study; unassigned, could not assigned to any lipolytic family

[b] Lipolytic family assignment reported in literatures (or this study) using pylogenetic-related methods

[c] No data

Supplementary Table S14 Phylogenetic origin (at genus level) of assigned lipolytic genes and their corresponding contigs for compost55 and compost76

Bacterial genus [a]	Genes		Contigs harboring corresponding genes	
	compost55	compost76	compost55	compost76
P_Actinobacteria|C_Actinobacteria|O_Streptosporangiales|F_Thermomonosporaceae|G_Thermomonospora	22 (1.99%)	21 (3.04%)	30 (2.48%)	44 (5.99%)
P_Actinobacteria|C_Actinobacteria|O_Streptomycetales|F_Streptomycetaceae|G_Streptomyces	40 (3.62%)	7 (1.01%)	24 (1.99%)	3 (0.41%)
P_Acidobacteria|C_unclassified_Acidobacteria	34 (3.08%)	9 (1.3%)	40 (3.31%)	3 (0.41%)
P_Actinobacteria|C_Actinobacteria|O_Micromonosporales|F_Micromonosporaceae|G_Micromonospora	40 (3.62%)	7 (1.01%)	37 (3.06%)	2 (0.27%)
uncultured_bacterium	19 (1.72%)	11 (1.59%)	35 (2.89%)	9 (1.22%)
P_Actinobacteria|C_Actinobacteria|O_Corynebacteriales|F_Mycobacteriaceae|G_Mycobacterium	23 (2.08%)	11 (1.59%)	18 (1.49%)	9 (1.22%)
P_Actinobacteria|C_Actinobacteria|O_Pseudonocardiales|F_Pseudonocardiaceae|G_Pseudonocardia	27 (2.45%)	0	38 (3.14%)	1 (0.14%)
P_Actinobacteria|C_Actinobacteria|O_Streptosporangiales|F_Thermomonosporaceae|G_Actinomadura	22 (1.99%)	9 (1.3%)	10 (0.83%)	5 (0.68%)
P_Firmicutes|C_Bacilli|O_Bacillales|F_Paenibacillaceae|G_Paenibacillus	11 (1%)	15 (2.17%)	3 (0.25%)	6 (0.82%)
P_Firmicutes|C_Clostridia|O_Clostridiales|F_Clostridiales_Family_XVII. Incertae_Sedis|G_Thermaerobacter	15 (1.36%)	13 (1.88%)	16 (1.32%)	13 (1.77%)
P_Chloroflexi|C_unclassified_Chloroflexi	10 (0.91%)	1 (0.14%)	37 (3.06%)	3 (0.41%)
P_Firmicutes|C_Bacilli|O_Bacillales|F_Bacillaceae|G_Bacillus	13 (1.18%)	13 (1.88%)	5 (0.41%)	6 (0.82%)
P_Bacteroidetes|C_unclassified_Bacteroidetes	18 (1.63%)	8 (1.16%)	10 (0.83%)	6 (0.82%)
P_Proteobacteria|C_Deltaproteobacteria|O_Myxococcales|F_Nannocystaceae|G_Nannocystis	0	10 (1.45%)	0	15 (2.04%)
P_Proteobacteria|C_Betaproteobacteria|O_unclassified_Betaproteobacteria	15 (1.36%)	5 (0.72%)	22 (1.82%)	2 (0.27%)
P_Proteobacteria|C_Gammaproteobacteria|O_Methylococcales|F_Methylococcaceae|G_Methylocaldum	4 (0.36%)	5 (0.72%)	5 (0.41%)	8 (1.09%)
P_Chloroflexi|C_Thermomicrobia|O_Sphaerobacterales|F_Sphaerobacteraceae|G_Sphaerobacter	6 (0.54%)	1 (0.14%)	20 (1.65%)	1 (0.14%)
P_Bacteroidetes|C_Sphingobacteriia|O_Sphingobacteriales|F_Sphingobacteriaceae|G_Sphingobacterium	0	13 (1.88%)	0	17 (2.31%)
P_Bacteroidetes|C_Chitinophagia|O_Chitinophagales|F_Chitinophagaceae|G_Niastella	3 (0.27%)	7 (1.01%)	2 (0.17%)	15 (2.04%)
P_Bacteroidetes|C_Chitinophagia|O_Chitinophagales|F_Chitinophagaceae|G_Flavihumibacter	0	8 (1.16%)	0	12 (1.63%)
P_Firmicutes|C_Clostridia|O_Clostridiales|F_Clostridiaceae|G_Clostridium	8 (0.72%)	10 (1.45%)	2 (0.17%)	8 (1.09%)
P_Actinobacteria|C_Actinobacteria|O_Streptosporangiales|F_Streptosporangiaceae|G_Nonomuraea	10 (0.91%)	2 (0.29%)	10 (0.83%)	0
P_Gemmatimonadetes|C_unclassified_Gemmatimonadetes	12 (1.09%)	0	18 (1.49%)	0

P_Actinobacteria|C_Actinobacteria|G_Thermobispora	10 (0.91%)	2 (0.29%)	19 (1.57%)	6 (0.82%)
P_Proteobacteria|C_Gammaproteobacteria|O_unclassified_Gammaproteobacteria	13 (1.18%)	5 (0.72%)	20 (1.65%)	2 (0.27%)
P_Actinobacteria|C_Actinobacteria|O_Propionibacteriales|F_Nocardioidaceae |G_Actinopolymorpha	7 (0.63%)	0	21 (1.74%)	0
P_Actinobacteria|C_Actinobacteria|O_Pseudonocardiales|F_Pseudonocardiaceae |G_Thermocrispum	11 (1%)	0	18 (1.49%)	0
P_Firmicutes|C_Clostridia|O_Clostridiales|F_unclassified_Clostridiales	3 (0.27%)	8 (1.16%)	2 (0.17%)	5 (0.68%)
P_Actinobacteria|C_Actinobacteria|O_Pseudonocardiales|F_Pseudonocardiaceae |G_Amycolatopsis	13 (1.18%)	3 (0.43%)	6 (0.5%)	0
P_Proteobacteria|C_Gammaproteobacteria|O_Pseudomonadales|F_Pseudomonadaceae |G_Pseudomonas	5 (0.45%)	9 (1.3%)	5 (0.41%)	5 (0.68%)
P_Proteobacteria|C_Gammaproteobacteria|O_Xanthomonadales|F_Xanthomonadaceae |G_Pseudoxanthomonas	1 (0.09%)	11 (1.59%)	1 (0.08%)	17 (2.31%)
P_Proteobacteria|C_Alphaproteobacteria|O_Rhizobiales|F_Beijerinckiaceae |G_Chelatococcus	1 (0.09%)	10 (1.45%)	3 (0.25%)	10 (1.36%)
P_Actinobacteria|C_Actinobacteria|O_Streptosporangiales|F_Nocardiopsaceae |G_Thermobifida	0	8 (1.16%)	1 (0.08%)	12 (1.63%)
P_Firmicutes|C_Clostridia|O_Clostridiales|F_Ruminococcaceae|G_Ruminiclostridium	2 (0.18%)	9 (1.3%)	3 (0.25%)	15 (2.04%)
P_Bacteroidetes|C_Flavobacteriia|O_Flavobacteriales|F_Flavobacteriaceae |G_Flavobacterium	3 (0.27%)	8 (1.16%)	2 (0.17%)	7 (0.95%)
P_Firmicutes|C_Bacilli|O_Bacillales|F_Bacillaceae|G_unclassified_Bacillaceae	3 (0.27%)	4 (0.58%)	4 (0.33%)	6 (0.82%)
P_Bacteroidetes|O_Bacteroidetes_Order_II._Incertae_sedis|F_Rhodothermaceae |G_Rhodothermus	13 (1.18%)	0	18 (1.49%)	0
P_Actinobacteria|C_Actinobacteria|O_unclassified_Actinobacteria	3 (0.27%)	1 (0.14%)	9 (0.74%)	1 (0.14%)
P_Proteobacteria|C_Alphaproteobacteria|O_unclassified_Alphaproteobacteria	6 (0.54%)	2 (0.29%)	4 (0.33%)	0
P_Proteobacteria|C_Gammaproteobacteria|O_Nevskiales|F_Sinobacteraceae |G_Steroidobacter	4 (0.36%)	5 (0.72%)	4 (0.33%)	16 (2.18%)
P_Actinobacteria|C_Actinobacteria|O_Micromonosporales|F_Micromonosporaceae |G_Actinoplanes	8 (0.72%)	1 (0.14%)	13 (1.08%)	0
P_Firmicutes|C_Bacilli|O_Bacillales|F_Bacillaceae|G_Calditerricola	4 (0.36%)	7 (1.01%)	0	9 (1.22%)
P_Deinococcus-Thermus|C_Deinococci|O_Thermales|F_Thermaceae|G_Thermus	5 (0.45%)	6 (0.87%)	4 (0.33%)	4 (0.54%)
P_Gemmatimonadetes|C_Gemmatimonadetes|O_Gemmatimonadales |F_Gemmatimonadaceae |G_Gemmatimonas	10 (0.91%)	1 (0.14%)	22 (1.82%)	1 (0.14%)
P_Actinobacteria|C_Actinobacteria|O_Corynebacteriales|F_Nocardiaceae|G_Nocardia	4 (0.36%)	1 (0.14%)	4 (0.33%)	2 (0.27%)
P_Firmicutes|C_Bacilli|O_Bacillales|F_Thermoactinomycetaceae|G_Planifilum	1 (0.09%)	4 (0.58%)	2 (0.17%)	4 (0.54%)
P_Proteobacteria|C_Deltaproteobacteria|O_Myxococcales|F_unclassified_Myxococcales	5 (0.45%)	2 (0.29%)	3 (0.25%)	5 (0.68%)
P_Firmicutes|C_Clostridia|O_Clostridiales|F_Caldicoprobacteraceae|G_Caldicoprobacter	0	7 (1.01%)	0	6 (0.82%)
P_Firmicutes|C_Bacilli|O_Bacillales|F_Alicyclobacillaceae|G_Alicyclobacillus	0	1 (0.14%)	3 (0.25%)	5 (0.68%)
P_Proteobacteria|C_Alphaproteobacteria|O_Rhizobiales|F_Bradyrhizobiaceae |G_Bradyrhizobium	10 (0.91%)	2 (0.29%)	7 (0.58%)	1 (0.14%)

[a] Only genera from abundant orders (greater than 1 %) are given. In the case the order or genus could not be assigned, the taxonomic name at the highest determined taxonomic resolution is given in parenthesis.

Supplementary Table S15 Mapping coverage of FA-identified lipolytic genes by metagenome short reads

Gene name	Gene length (bp)	Nr. of reads mapped on gene (read depth)	Total number of covered bases (with >= 1X coverage depth)	Breadth of coverage (%, at 1X coverage depth)
estC55-2	876	41	876	100
estC55-3	1542	815	1542	100
estC55-4_1	816	102	816	100
estC55-4_2	1173	157	1173	100
estC55-5	933	442	933	100
estC55-6	1038	83	1038	100
estC55-7	1293	35	1068	82.6
estC55-8_1	957	69	954	99.7
estC55-8_2	813	74	813	100
estC55-10	867	579	867	100
estC55-12	1155	1419	1155	100
estC55-13	1068	93	1067	99.9
estC55-15	762	853	762	100
estC55-18	804	9	723	90
estC55-19_1	780	131	780	100
estC55-19_2	957	147	957	100
estC55-20	1053	181	1053	100
estC55-23	969	43	968	99.9
estC55-24	978	63	978	100
estC55-25	1104	170	1104	100
estC55-26	954	843	954	100
estC55-31	798	43	796	99.7
estC55-34	873	939	873	100
estC55-38	990	4	439	44.3
estC55-40	1335	119	1335	100
estC55-42	1122	1449	1122	100
estC55-46	1248	24	638	51.1
estC55-51	807	858	807	100
estC55-52	1512	90	1512	100
estC55-53	1221	11	1106	90.6
estC55-56	945	32	944	99.9
estC55-57	816	36	774	95
estC55-60	1023	26	971	95
estC55-61	1167	28	982	84.1
estC55-62	1587	135	1586	99.9
estC55-63	954	488	954	100
estC55-65	1230	92	1230	100
estC55-66	1290	78	1290	100
estC55-71	954	73	951	99.7
estC55-72	897	20	879	98
estC55-73	1158	47	1154	99.6
estC55-76	1122	90	1117	99.5
estC55-77	837	13	591	70.6
estC55-78	960	104	957	99.7
estC55-80	1227	308	1227	100
estC55-81	849	61	846	99.6
estC55-88	954	56	953	99.9
estC55-90	861	55	860	99.9

estC55-95	867	60	866	99.9
estC55-96	804	1313	804	100
estC55-97	723	8	720	99.6
estC55-100	1131	4	365	32.3
estC55-102	2124	41	1979	93.2
estC55-105	672	730	672	100
estC55-110	1305	32	1304	99.9
estC55-111	1179	2088	1179	100
estC55-113	1230	20	874	71.1
estC55-118	1494	16	1230	82.3
estC55-131	756	32	755	99.9
estC55-145	918	955	918	100
estC55-147	1161	186	1161	100
estC55-150	1182	1469	1182	100
estC55-151	936	22	935	99.9
estC55-154	1185	50	1183	99.9
estC55-156	1554	37	1519	97.7
estC55-159	855	7	634	74.1
estC55-163	912	46	912	100
estC55-164	1293	0	0	0
estC55-165	834	1783	834	100
estC55-167	747	24	747	100
estC55-168	1542	1254	1542	100
estC55-169	1602	96	1581	98.6
estC55-186	1548	84	1540	99.5
estC55-188	714	37	714	100
estC55-197	876	584	876	100
estC55-213	861	1711	861	100
estC55-215	792	123	792	100
estC55-227	1419	1616	1419	100
estC55-229	936	28	929	99.2
estC55-231	1101	61	1071	97.3
estC55-234	1674	17	1380	82.4
estC55-235	981	540	981	100
estC55-239	1146	3222	1146	100
estC55-241	1704	1707	1704	100
estC55-244	768	46	747	97.3
estC55-245	1389	27	1309	94.2
estC55-247	1107	3	499	45.1
estC55-251	954	358	954	100
estC55-253	894	72	894	100
estC55-256	813	166	813	100
estC55-258	1170	23	1153	99
estC55-268	912	489	912	100
estC76-21	867	14	740	85
estC76-28_1	774	9	212	27
estC76-28_2	732	135	732	100
estC76-36	714	155	714	100
estC76-98	1278	8	781	61
estC76-123	1353	284	1353	100
estC76-135	933	71	932	100
estC76-136	1548	360	1548	100
estC76-137	738	512	738	100
estC76-174	1167	197	1167	100
estC76-177	798	699	798	100
estC76-179	990	65	879	89

estC76-202	804	71	804	100
estC76-218	990	48	838	85
estC76-221	774	454	774	100
estC76-222	783	1154	783	100
estC76-248	774	12	614	79
estC76-250	756	190	756	100
estC76-261	1038	58	1038	100
estC76-262	1248	1	61	5
estC76-263	1131	109	1131	100
estC76-266	774	153	774	100
estC76-269	783	418	783	100

Supplementary Table S16 BLAST-based comparison between putative lipolytic genes identified by sequence-based screening against these identified by function-driven screening

Query [a]	query length	Subject [b]	Subject length	Identity (%)	E-value	Bit score	Coverage	Alignment length
EstC55-2	291	LLBHCHNC_108683	291	99.656	0	588	100	291
EstC55-3	513	LLBHCHNC_04031	509	94.303	0	927	99	509
EstC55-4_1	271	LLBHCHNC_60489	271	100	0	539	100	271
EstC55-4_2	390	LLBHCHNC_60488	390	100	0	795	100	390
EstC55-5	310	LLBHCHNC_10894	310	100	0	615	100	310
EstC55-6	345	LLBHCHNC_24927	345	99.71	0	669	100	345
EstC55-7	430	LLBHCHNC_198339	355	99.437	0	732	83	355
EstC55-8_1	318	LLBHCHNC_170023	318	99.686	0	645	100	318
EstC55-8_2	270	LLBHCHNC_170022	270	100	0	539	100	270
EstC55-10	288	LLBHCHNC_150671	279	100	0	551	97	279
EstC55-12	384	LLBHCHNC_52061	384	100	0	760	100	384
EstC55-13	355	LLBHCHNC_01691	355	100	0	709	100	355
EstC55-15	253	LLBHCHNC_23859	265	100	3.67E-178	492	99	252
EstC55-18	267	LLBHCHNC_236409	247	37.402	5.02E-41	144	93	254
EstC55-19_1	317	LLBHCHNC_149694	317	100	0	639	100	317
EstC55-19_2	270	LLBHCHNC_149695	270	99.63	0	539	100	270
EstC55-20	350	LLBHCHNC_54608	350	100	0	703	100	350
EstC55-23	322	LLBHCHNC_76053	322	100	0	650	100	322
EstC55-24	325	LLBHCHNC_164893	325	100	0	637	100	325
EstC55-25	367	LLBHCHNC_31406	358	99.721	0	707	98	358
EstC55-26	317	LLBHCHNC_200149	317	96.845	0	627	100	317
EstC55-31	265	LLBHCHNC_90973	265	100	0	542	100	265
EstC55-34	290	LLBHCHNC_296992	290	98.276	0	568	100	290
EstC55-38	329	LLBHCHNC_417537	329	99.392	0	672	100	329
EstC55-40	444	LLBHCHNC_161203	442	99.774	0	888	99	442
EstC55-42	373	LLBHCHNC_68310	427	56.919	2.95E-140	408	100	383
EstC55-46	415	LLBHCHNC_180099	436	82.339	0	716	100	436
EstC55-51	268	LLBHCHNC_181078	286	61.024	9.53E-98	290	95	254
EstC55-52	503	LLBHCHNC_276493	505	86.558	0	822	97	491
EstC55-53	406	LLBHCHNC_208903	404	52.12	2.26E-135	396	99	401
EstC55-56	314	LLBHCHNC_147899	314	96.166	0	605	99	313
EstC55-57	271	LLBHCHNC_422494	271	100	0	560	100	271
EstC55-60	340	LLBHCHNC_114902	303	100	0	597	89	303
EstC55-61	388	LLBHCHNC_499519	340	98.824	0	654	88	340
EstC55-62	528	LLBHCHNC_23442	528	100	0	1077	100	528
EstC55-63	317	LLBHCHNC_200149	317	97.161	0	629	100	317
EstC55-65	409	LLBHCHNC_225424	408	83.99	0	701	99	406

EstC55-66	429	LLBHCHNC_233808	451	97.203	0	839	100	429
EstC55-71	317	LLBHCHNC_228249	317	97.792	0	634	100	317
EstC55-72	298	LLBHCHNC_523008	297	51.701	7.09E-89	269	98	294
EstC55-73	385	LLBHCHNC_146194	385	99.481	0	764	100	385
EstC55-76	373	LLBHCHNC_58044	373	100	0	736	100	373
EstC55-77	278	LLBHCHNC_517812	267	40.58	1.29E-44	154	95	276
EstC55-78	319	LLBHCHNC_93396	319	100	0	637	100	319
EstC55-80	408	LLBHCHNC_225424	408	99.755	0	828	100	408
EstC55-81	282	LLBHCHNC_53848	312	100	0	561	100	282
EstC55-88	317	LLBHCHNC_228249	317	97.792	0	631	100	317
EstC55-90	286	LLBHCHNC_112431	286	100	0	589	100	286
EstC55-95	288	LLBHCHNC_71403	371	99.653	0	578	100	288
EstC55-96	267	LLBHCHNC_432831	267	99.625	0	536	100	267
EstC55-97	240	LLBHCHNC_354884	753	38.068	1.16E-21	96.3	72	176
EstC55-100	376	LLBHCHNC_195970	378	53.704	1.13E-133	389	95	378
EstC55-102	707	LLBHCHNC_174663	719	60.734	0	790	99	736
EstC55-105	223	LLBHCHNC_309242	223	100	8.96E-166	458	100	223
EstC55-110	434	LLBHCHNC_266798	434	99.77	0	891	100	434
EstC55-111	392	LLBHCHNC_445504	392	100	0	795	100	392
EstC55-113	409	LLBHCHNC_232836	442	39.225	4.19E-85	269	91	413
EstC55-118	497	LLBHCHNC_276493	505	76.578	0	730	99	491
EstC55-131	251	LLBHCHNC_493354	330	97.211	1.24E-168	471	100	251
EstC55-145	305	LLBHCHNC_59760	305	99.672	0	573	100	305
EstC55-147	386	LLBHCHNC_53082	386	99.741	0	775	100	386
EstC55-150	393	LLBHCHNC_445504	392	92.875	0	694	100	393
EstC55-151	311	LLBHCHNC_364509	311	99.678	0	640	100	311
EstC55-154	394	LLBHCHNC_502793	394	100	0	791	100	394
EstC55-156	517	LLBHCHNC_37415	520	79.31	0	828	94	493
EstC55-159	284	LLBHCHNC_338455	395	100	0	544	100	284
EstC55-163	303	LLBHCHNC_311877	558	99.34	0	589	100	303
EstC55-164	430	LLBHCHNC_32909	424	51.741	3.85E-123	366	90	402
EstC55-165	277	LLBHCHNC_112340	276	52.049	1.08E-78	242	87	244
EstC55-167	248	LLBHCHNC_90003	248	100	0	511	100	248
EstC55-168	513	LLBHCHNC_291282	513	99.805	0	1028	100	513
EstC55-169	533	LLBHCHNC_116580	533	100	0	1088	100	533
EstC55-186	515	LLBHCHNC_353867	504	39.6	2.85E-83	270	94	500
EstC55-188	237	LLBHCHNC_239154	97	100	1.08E-62	193	41	97
EstC55-197	292	LLBHCHNC_21317	294	99.656	0	577	99	291
EstC55-213	286	LLBHCHNC_206366	268	97.015	0	524	94	268
EstC55-215	263	LLBHCHNC_65857	263	99.24	0	525	100	263
EstC55-227	459	LLBHCHNC_119846	472	99.564	0	908	100	459
EstC55-229	311	LLBHCHNC_315562	325	100	0	633	100	311
EstC55-231	366	LLBHCHNC_25121	366	100	0	737	100	366
EstC55-234	557	LLBHCHNC_37415	520	40.301	2.73E-120	367	92	531
EstC55-235	326	LLBHCHNC_51521	326	96.599	0	580	90	294

EstC55-239	381	LLBHCHNC_156698	968	99.213	0	773	100	381
EstC55-241	567	LLBHCHNC_129990	567	95.767	0	1071	100	567
EstC55-244	255	LLBHCHNC_100331	255	100	0	526	100	255
EstC55-245	462	LLBHCHNC_507289	231	98.701	9.06E-164	463	50	231
EstC55-247	368	LLBHCHNC_400362	352	53.977	3.80E-121	356	96	352
EstC55-251	317	LLBHCHNC_200149	317	98.423	0	636	100	317
EstC55-253	304	LLBHCHNC_96988	297	100	5.27E-165	462	77	233
EstC55-256	270	LLBHCHNC_152981	270	98.519	0	545	100	270
EstC55-258	389	LLBHCHNC_167603	394	56.633	1.80E-147	426	98	392
EstC55-268	303	LLBHCHNC_244305	303	64.726	2.48E-125	362	96	292
EstC76-21	288	FKLAIAFA_437586	279	97.849	0	539	97	279
EstC76-28_1	257	FKLAIAFA_63857	257	91.829	1.56E-164	457	100	257
EstC76-28_2	243	FKLAIAFA_63856	243	100	5.52E-176	485	100	243
EstC76-36	237	FKLAIAFA_40146	238	98.734	3.29E-162	450	100	237
EstC76-98	425	FKLAIAFA_43605	437	40.92	1.08E-102	314	99	435
EstC76-123	450	FKLAIAFA_91696	450	99.333	0	903	100	450
EstC76-135	310	FKLAIAFA_98475	317	71.613	8.65E-167	468	100	310
EstC76-136	515	FKLAIAFA_43393	515	100	0	1009	100	515
EstC76-137	277	FKLAIAFA_218044	221	43.231	2.10E-54	177	81	229
EstC76-174	388	FKLAIAFA_18366	388	99.485	0	750	100	388
EstC76-177	265	FKLAIAFA_220036	265	100	0	535	100	265
EstC76-179	329	FKLAIAFA_34048	329	97.568	0	664	100	329
EstC76-202	267	FKLAIAFA_18975	267	100	0	538	100	267
EstC76-218	329	FKLAIAFA_34048	329	98.784	0	669	100	329
EstC76-221	257	FKLAIAFA_399351	257	99.611	0	524	100	257
EstC76-222	260	FKLAIAFA_58163	260	99.231	0	523	100	260
EstC76-248	257	FKLAIAFA_63857	257	93.385	8.27E-169	468	100	257
EstC76-250	251	FKLAIAFA_107970	251	98.805	2.16E-174	482	100	251
EstC76-261	345	FKLAIAFA_06843	345	99.71	0	669	100	345
EstC76-262	415	FKLAIAFA_69458	411	57.039	2.80E-151	437	97	412
EstC76-263	376	FKLAIAFA_04547	376	99.734	0	752	100	376
EstC76-266	257	FKLAIAFA_63857	257	100	0	501	100	257
EstC76-269	260	FKLAIAFA_60368	260	98.462	0	513	100	260

[a] Lipolytic genes identifies through function-driven approach in this study

[b] Putative lipolytic genes identifies through sequence-based approach in this study

Supplementary Table S17 Pseudo-value of ANOSIM test for group dissimilarity of lipolytic family profiles across different habitats (LPGM was log10 transformed. The distance is Bray-Curtis, permutation N=9999, overall R value=0.6168, *p*=0.0001)

Habitat	ADAS	COM	AS	GS	HG	HM	HRE	HS	LL	MS	MW	OR	RW	TFS	WB
ADAS	NA[a]	0.705	0.468	0.781	0.999	0.949	0.936	0.358	0.389	0.318	0.67	0.962	0.999	0.937	1
COM	0.705	NA[a]	0.726	0.786	0.852	0.365	0.84	0.733	0.796	0.641	0.52	0.799	0.471	0.686	0.194
AS	0.468	0.726	NA[a]	0.511	0.708	0.175	0.336	0.438	0.627	0.442	0.466	0.631	0.443	0.563	0.62
GS	0.781	0.786	0.511	NA[a]	1	0.885	0.968	0.384	0.46	0.334	0.779	0.995	0.961	0.883	0.972
HG	0.999	0.852	0.708	1	NA[a]	0.995	0.939	0.617	0.844	0.687	0.987	0.982	1	0.952	0.999
HM	0.949	0.365	0.175	0.885	0.995	NA[a]	0.773	0.403	0.495	0.324	0.631	0.875	0.655	0.843	0.865
HRE	0.936	0.84	0.336	0.968	0.939	0.773	NA[a]	0.75	0.899	0.606	0.905	0.929	0.893	0.904	0.9
HS	0.358	0.733	0.438	0.384	0.617	0.403	0.75	NA[a]	0.514	0.252	0.507	0.637	0.516	0.604	0.668
LL	0.389	0.796	0.627	0.46	0.844	0.495	0.899	0.514	NA[a]	0.263	0.484	0.792	0.685	0.838	0.754
MS	0.318	0.641	0.442	0.334	0.687	0.324	0.606	0.252	0.263	NA[a]	0.397	0.65	0.504	0.597	0.584
MW	0.67	0.52	0.466	0.779	0.987	0.631	0.905	0.507	0.484	0.397	NA[a]	0.894	0.783	0.892	0.81
OR	0.962	0.799	0.631	0.995	0.982	0.875	0.929	0.637	0.792	0.65	0.894	NA[a]	0.918	0.977	0.871
RW	0.999	0.471	0.443	0.961	1	0.655	0.893	0.516	0.685	0.504	0.783	0.918	NA[a]	0.875	0.95
TFS	0.937	0.686	0.563	0.883	0.952	0.843	0.904	0.604	0.838	0.597	0.892	0.977	0.875	NA[a]	0.83
WB	1	0.194	0.62	0.972	0.999	0.865	0.9	0.668	0.754	0.584	0.81	0.871	0.95	0.83	NA[a]
Mean[b]	0.748	0.686	0.511	0.764	0.897	0.659	0.827	0.527	0.631	0.471	0.695	0.851	0.761	0.813	0.787

[a] No data

[b] Averaged R values

Supplementary Table S19 Rseudo-value of ANOSIM test for group dissimilarity of phylogenetic distribution (at genus level) of assigned PLPs across habitats (Only genus with a mean LPGM ≥ 0.5 across all the samples were used. The distance is Bray-Curtis, permutation N=9999, overall R value=0.8199, *p*=0.0001)

Habitat	AS	ADAS	COM	GS	HG	HM	HRE	HS	LL	MS	MW	OR	RW	TFS	WB
AS	NA[a]	0.609	0.628	0.596	0.88	0.506	0.567	0.682	0.613	0.683	0.639	0.791	0.646	0.435	0.69
ADAS	0.609	NA[a]	0.588	0.993	1	1	0.991	0.669	0.638	0.573	0.694	0.997	1	1	1
COM	0.628	0.588	NA[a]	0.568	0.868	0.572	0.864	0.752	0.759	0.727	0.53	0.837	0.546	0.595	0.232
GS	0.596	0.993	0.568	NA[a]	1	0.993	0.995	0.738	0.947	0.726	0.866	1	1	0.977	0.994
HG	0.88	1	0.868	1	NA[a]	1	0.937	0.916	1	0.995	1	1	1	1	1
HM	0.506	1	0.572	0.993		NA[a]	0.971	0.736	1	0.943	0.904	1	0.973	0.999	0.978
HRE	0.567	0.991	0.864	0.995	0.937	0.971	NA[a]	0.811	0.998	0.994	0.99	0.98	0.983	0.949	0.956
HS	0.682	0.669	0.752	0.738	0.916	0.736	0.811	NA[a]	0.724	0.704	0.732	0.874	0.651	0.864	0.791
LL	0.613	0.638	0.759	0.947	1	1	0.998	0.724	NA[a]	0.401	0.799	0.996	1	1	0.993
MS	0.683	0.573	0.727	0.726	0.995	0.943	0.994	0.704	0.401	NA[a]	0.622	0.991	0.962	0.992	0.948
MW	0.639	0.694	0.53	0.866	1	0.904	0.99	0.732	0.799	0.622	NA[a]	0.997	0.947	0.992	0.937
OR	0.791	0.997	0.837	1	1	1	0.98	0.874	0.996	0.991	0.997	NA[a]	1	1	0.999
RW	0.646	1	0.546	1	1	0.973	0.983	0.651	1	0.962	0.947	1	NA[a]	1	1
TFS	0.435	1	0.595	0.977	1	0.999	0.949	0.864	1	0.992	0.992	1	1	NA[a]	0.967
WB	0.69	1	0.232	0.994	1	0.978	0.956	0.791	0.993	0.948	0.937	0.999	1	0.967	NA[a]
Mean[b]	0.640	0.839	0.647	0.885	0.969	0.898	0.928	0.760	0.848	0.804	0.832	0.961	0.908	0.912	0.892

[a] No data

[b] Averaged R values

Supplementary Table S21 Pseudo-values of ANOSIM test for group dissimilarity of phylogenetic distributions (at genus level) of total PLPs across habitats (Only genera with a mean LPGM > 0.5 across all the samples were used. The distance is Bray-Curtis, permutation N=9999, overall R value=0.821, *p*=0.0001)

Habitat	ADAS	COM	AS	GS	HG	HM	HRE	HS	LL	MS	MW	OR	RW	TFS	WB
ADAS	NA[a]	0.605	0.862	1	1	0.994	0.989	0.577	0.636	0.537	0.657	0.995	1	1	0.999
COM	0.605	NA[a]	0.677	0.448	0.907	0.47	0.849	0.752	0.791	0.689	0.527	0.853	0.495	0.551	0.218
AS	0.862	0.677	NA[a]	0.804	0.95	0.79	0.791	0.867	0.866	0.879	0.86	0.914	0.798	0.494	0.854
GS	1	0.448	0.804	N [a]	1	0.968	0.978	0.615	0.938	0.607	0.757	1	1	0.962	0.964
HG	1	0.907	0.95	1	NA[a]	1	0.938	0.987	1	0.995	1	1	1	1	1
HM	0.994	0.47	0.79	0.968	1	NA[a]	0.922	0.765	1	0.873	0.859	1	0.952	0.983	0.956
HRE	0.989	0.849	0.791	0.978	0.938	0.922	NA[a]	0.975	0.998	0.986	0.97	0.984	0.967	0.931	0.939
HS	0.577	0.752	0.867	0.615	0.987	0.765	0.975	NA[a]	0.776	0.589	0.682	0.952	0.673	0.908	0.805
LL	0.636	0.791	0.866	0.938	1	1	0.998	0.776	NA[a]	0.369	0.778	0.993	0.999	0.999	0.994
MS	0.537	0.689	0.879	0.607	0.995	0.873	0.986	0.589	0.369	NA[a]	0.541	0.972	0.94	0.977	0.912
MW	0.657	0.527	0.86	0.757	1	0.859	0.97	0.682	0.778	0.541	NA[a]	0.986	0.923	0.972	0.922
OR	0.995	0.853	0.914	1	1	1	0.984	0.952	0.993	0.972	0.986	NA[a]	1	1	1
RW	1	0.495	0.798	1	1	0.952	0.967	0.673	0.999	0.94	0.923	1	NA[a]	0.994	0.998
TFS	1	0.551	0.494	0.962	1	0.983	0.931	0.908	0.999	0.977	0.972	1	0.994	NA[a]	0.95
WB	0.999	0.218	0.854	0.964	1	0.956	0.939	0.805	0.994	0.912	0.922	1	0.998	0.95	NA[a]
Mean[b]	0.846	0.631	0.815	0.860	0.984	0.895	0.944	0.780	0.867	0.776	0.817	0.975	0.910	0.909	0.894

[a] No data

[b] Averaged R values

Chapter III

Biochemical profiles of two thermostable and organic solvent-tolerant esterases derived from a compost metagenome

Mingji Lu, Amélie Dukunde, Rolf Daniel

Doi: https://doi.org/10.1007/s00253-019-09695-1

https://link.springer.com/article/10.1007/s00253-019-09695-1

Applied Microbiology and Biotechnology
https://doi.org/10.1007/s00253-019-09695-1

BIOTECHNOLOGICALLY RELEVANT ENZYMES AND PROTEINS

Biochemical profiles of two thermostable and organic solvent–tolerant esterases derived from a compost metagenome

Mingji Lu[1] · Amélie Dukunde[1] · Rolf Daniel[1]

Received: 1 December 2018 / Revised: 11 February 2019 / Accepted: 12 February 2019

Abstract
Owing to the functional versatility and potential applications in industry, interest in lipolytic enzymes tolerant to organic solvents is increasing. In this study, functional screening of a compost soil metagenome resulted in identification of two lipolytic genes, *est1* and *est2*, encoding 270 and 389 amino acids, respectively. The two genes were heterologously expressed and characterized. Est1 and Est2 are thermostable enzymes with optimal enzyme activities at 80 and 70 °C, respectively. A second-order rotatable design, which allows establishing the relationship between multiple variables with the obtained responses, was used to explore the combined effects of temperature and pH on esterase stability. The response curve indicated that Est1, and particularly Est2, retained high stability within a broad range of temperature and pH values. Furthermore, the effects of organic solvents on Est1 and Est2 activities and stabilities were assessed. Notably, Est2 activity was significantly enhanced (two- to tenfold) in the presence of ethanol, methanol, isopropanol, and 1-propanol over a concentration range between 6 and 30% (*v*/*v*). For the short-term stability (2 h of incubation), Est2 exhibited high tolerance against 60% (*v*/*v*) of ethanol, methanol, isopropanol, DMSO, and acetone, while Est1 activity resisted these solvents only at lower concentrations (below 30%, *v*/*v*). Est2 also displayed high stability towards some water-immiscible organic solvents, such as ethyl acetate, diethyl ether, and toluene. With respect to long-term stability, Est2 retained most of its activity after 26 days of incubation in the presence of 30% (*v*/*v*) ethanol, methanol, isopropanol, DMSO, or acetone. All of these features indicate that Est1 and Est2 possess application potential.

Keywords Carboxylesterases · Metagenomic library · Second-order rotatable design · Thermophilic · Organic solvent tolerance

Introduction

Extreme environments exhibiting elevated temperatures, extreme pH values, and exposure to organic solvents or high salinity are used to recover novel robust bioactive molecules that can be applied under industrial conditions (Antranikian and Egorova 2007). The targeted environments such as hot springs, compost, oil fields, and deep-sea marine sediments are reservoirs for extremophilic microorganisms that could produce potentially relevant industrial enzymes (Auernik et al. 2008). Culture-independent metagenomic approaches are alternatives to conventional culture-based screening methods. Recently, some extremozymes, such as amylases, amidases, proteases, cellulases, and esterases, have been successfully identified through metagenomic approaches (Daniel 2005; Simon and Daniel 2011; González-González et al. 2017; Jayanath et al. 2018; Martínez-Martínez et al. 2018).

Lipolytic enzymes, which catalyze the hydrolysis and synthesis of acylglycerols, are considered as one of the most important groups of biocatalysts. Lipolytic enzymes include esterases (EC 3.1.1.1, carboxylesterases) and true lipases (EC 3.1.1.3, triacylglycerol acyl hydrolases) and are widespread in bacteria, archaea, and eukaryotes (Hasan et al. 2006). Due to their broad substrate, pH, and temperature spectra combined with high regio- and enantioselectivity, lipolytic enzymes are of interest for food, paper, medical, detergent, and pharmaceutical industries (Hita et al. 2009; Romdhane et al. 2010; Ferrer et al. 2015; Sarmah et al. 2018). In particular, lipolytic enzymes that function in non-aqueous solvents have attracted considerable attention, as they offer new possibilities for

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00253-019-09695-1) contains supplementary material, which is available to authorized users.

✉ Rolf Daniel
rdaniel@gwdg.de

[1] Department of Genomic and Applied Microbiology, Göttingen Genomics Laboratory, Institute of Microbiology and Genetics, Georg-August-University of Göttingen, Grisebachstraße 8, 37077 Göttingen, Germany

Published online: 26 February 2019

bioprocesses, such as shifting of thermodynamic equilibrium in favor of synthesis (esterification and transesterification), controlling substrate specificity and solubility by solvent engineering, and suppressing water-dependent side reactions (Secundo and Carrea 2002; Hun et al. 2003; Ahmed et al. 2010). However, the inhibition or inactivation of enzyme activity resulting from organic solvents has restricted the use of many lipolytic enzymes (Klibanov 2001; Jin et al. 2012). To overcome this limitation, some organic solvent–tolerant (OST) lipolytic enzymes have been isolated, including enzymes from *Bacillus licheniformis* S-86 (Torres et al. 2009), *Streptomyces coelicolor* A3(2) (Brault et al. 2012), *Psychrobacter celer* 3Pb1 (Wu et al. 2013), *Alcanivorax dieselolei* B-5(T) (Zhang et al. 2014), and *Acetomicrobium hydrogeniformans* (Kumagai et al. 2018), as well as from metagenomes of seawater (Chu et al. 2008), compost (Kang et al. 2011), lipid-contaminated soil (Glogauer et al. 2011), mountain soil (Jin et al. 2012), swamp sediment (Seo et al. 2014), and deep-sea hydrothermal vents (Yang et al. 2018). However, these enzymes only show tolerance towards specific organic solvents.

Industrially required versatile lipolytic enzymes that exhibit satisfactory activity and stability in both water-miscible and water-immiscible organic solvents are rare (Doukyu and Ogino 2010). It has been shown that the thermostability of an enzyme in water is correlated to its tolerance against organic solvents (Kumar et al. 2016). Thus, it is straightforward to screen naturally evolved OST enzymes from thermostable ones (Lotti and Alberghina 2007; Ahmed et al. 2010). Most industrial processes utilizing lipolytic enzymes are carried out at higher temperatures (above 45 °C); it is required that the enzymes exhibit activity and stability optima around 50 °C (Sharma et al. 2002). Thus, thermostable lipolytic enzymes exhibiting organic solvent tolerance are of high importance with respect to industrial applications.

Composting is the process of biological, aerobic decomposition of organic waste by microorganisms (Ryckeboer et al. 2003). During the thermophilic phase of composting, heat generated by microbial succession can raise temperatures to above 50 °C (Dougherty et al. 2012). Correspondingly, compost is a potential source for recovery of thermostable enzymes. Recently, lipolytic enzymes have been isolated from compost (Lämmle et al. 2007; Tirawongsaroj et al. 2008; Kim et al. 2010; Ohlhoff et al. 2015; Woo Lee et al. 2016), but only one (EstCS2) of the isolated enzymes was moderately thermostable (optimum 55 °C) and showed resistance to certain water-miscible organic solvents (Kang et al. 2011).

In this study, two genes encoding lipolytic enzymes (*est1* and *est2*) were identified from a thermophilic compost metagenome. The corresponding enzymes were purified and characterized. Enzyme characterizations are usually conducted as one-factor-at-a-time for comparison with reported enzyme features from other studies. However, this methodology ignores interacting effects between factors, which may result in misleading conclusions, especially when at least two requirements must be fulfilled simultaneously. An alternative is to employ design of experiments (DOE) methodologies, i.e., second-order rotatable design approach, which use statistical and mathematical approaches to evaluate the combined effect of factors. DOE has been successfully applied in different aspects related to lipolytic enzymes such as the growth condition optimization and enzyme activity or stability measurements (Kamimura et al. 2001; Shieh et al. 2003; Benaiges et al. 2010). In this study, the combined effect of pH and temperature on the stability of Est1and Est2 was evaluated by the second-order rotatable design approach. Analysis of the recovered two metagenome-derived enzymes showed that they are thermophilic and tolerant towards organic solvents. In addition, Est2 was remarkably resistant to both water-miscible and immiscible organic solvents.

Materials and methods

Strains and plasmids

For the construction of metagenomic plasmid libraries, *Escherichia coli* TOP10 and pFLD (Invitrogen GmbH, Karlsruhe, Germany) were used as host and vector, respectively. *Escherichia coli* BL21 Star (DE3) and the pET101/D-TOPO® vector (Invitrogen GmbH) were used for heterologous expression of the recovered lipolytic genes.

Sample collection and DNA extraction

Compost samples were collected at a composting company (Göttingen GmbH, Göttingen, Germany, 51° 34′ 25.1″ N 9° 54′ 33.0″ E). The sampling pile was the fermentation product of household waste and fresh tree branches. To ensure using mainly thermophilic microorganisms as a source for metagenomic library construction, compost at the core zone of compost pile was collected. The temperature at the sampling spot was 55 °C. The compost soil sample (50 g) was collected in sterile plastic bags and stored at − 20 °C until required.

Metagenomic DNA of the compost sample was extracted following a phenol-chloroform method according to Zhou et al. (1996). In addition, DNA was also isolated with MoBio Power Soil DNA extraction kit following the protocol of the manufacturer (MoBio Laboratories, Carlsbad, CA, USA). DNA obtained from these two methods was pooled and stored at − 20 °C until required.

Metagenomic library construction and screening for lipolytic activity

To construct a metagenomic plasmid library, DNA was sheared, and fragments from 6 to 12 kb were recovered by gel extraction with the peqGold gel extraction kit (Peqlab Biotechnologie GmbH, Erlangen, Germany). End-repaired DNA fragments and *PmlI*-digested pFLD vector were ligated by employing T4 DNA ligase at 16 °C, overnight as recommended by the manufacturer (Thermo Scientific, Bremen, Germany). To screen for lipolytic activity, metagenomic library-bearing cells were plated onto LB agar plates containing 100 μg/ml ampicillin and 1% (*v*/*v*) emulsified tributyrin (Sigma, Germany) and subsequently incubated at 30 °C 1 to 7 days. Lipolytic-positive clones were identified by the formation of clear zones (halos) around individual colonies. The phenotype of positive clones was confirmed by the isolation of recombinant plasmids from positive strains, transformation of the isolated plasmids into the host, and rescreening on indicator agar plates.

Sequence analysis and homology modeling

Lipolytic genes (*est1* and *est2*) were initially predicted by using the ORF Finder program (http://www.ncbi.nlm.nih.gov/gorf/gorf.html), and verified by Clone Manager and FramePlot analysis (Ishikawa and Hotta 1999). Similarity searches of the deduced amino acid sequences were performed by BLASTP program against the public GenBank database (Ye et al. 2006). Signal peptides were detected by using the SignalP 4.0 server (Bendtsen et al. 2004). The deduced amino acid sequences of *est1* and *est2* and reference sequences retrieved from GenBank were used to construct a phylogenetic tree with the neighbor-joining method by using MEGA version 6 (Tamura et al. 2013). Bootstrapping of 1000 replicates was used to estimate the confidence level.

Based on the deduced amino acid sequence, secondary structure and tertiary structure predictions were performed with I-TASSER (Zhang 2008). The identified structural analogs were used for multiple-sequence alignment using the Expresso webserver (Notredame et al. 2000). Figures showing secondary structure alignments were exported by ESPript3 (Robert and Gouet 2014). The analog with the highest TM score was also selected for structural superimposition.

Expression and purification of recombinant proteins Est1 and Est2

The primer pairs 5′-CACCATGCCCCTGGCCCGAG TGGA-3′ and 5′-GGCGCCCACCGGC ACCTGAGTC-3′ and 5′-CACCATGACCGAGCTGCCGGTGGGAG-3′ and 5′-GCG TCTTAGCGCGCGGTACAC-3 were used to amplify *est1* and *est2*, respectively. The resulting PCR products were purified and ligated into expression vector pET101/D-TOPO® according to the protocol of the manufacturer (Invitrogen). To produce His_6-tagged Est1 and Est2, recombinant plasmid DNA was transformed into *E. coli* BL21 (DE3) cells and plated on LB agar plates with 100 μg/ml ampicillin. A single colony was picked and grown overnight at 30 °C in 60-ml LB medium containing 100 μg/ml ampicillin. Subsequently, this pre-culture was added to 600-ml LB medium with 100 μg/ml ampicillin and grown with shaking at 30 °C. At an optical density (OD_{600}) of 0.6, isopropyl-beta-D thiogalactopyranoside (IPTG) was added to a final concentration of 0.5 mM. After a 6-h induction at 30 °C, cells were harvested by centrifugation (7000×*g*, 4 °C, 10 min). Cell pellets were washed with 100 ml LEW buffer (50 mM NaH_2PO_4, 300 mM NaCl, pH 8) and stored at − 20 °C until required.

To purify Est1 and Est2, Protino® Ni-TED 2000 packed column (Macherey-Nagel, Germany) was used following the manufacturer's protocol, however, with the modified LEW buffer (50 mM NaH_2PO_4, 1 M NaCl, 10% *v*/*v* glycerol, 0.05% *v*/*v* Triton X-100, pH 8). Protein concentration was measured by the Bradford method (Bradford 1976). Purity and molecular mass of the purified proteins were determined by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) using the procedure of Laemmli (1970). Fractions derived from affinity chromatography showing a single band with the estimated molecular mass of the targeted proteins were pooled, dialyzed with 50 mM sodium phosphate buffer (pH 8), and stored in 50% (*v*/*v*) glycerol at − 20 °C until use.

Standard enzyme assays

Esterase activity was measured by a spectrophotometric method (Jaeger et al. 1999) using *p*-nitrophenyl (*p*-NP) acyl esters (Sigma) as substrates. To minimize substrate auto-hydrolysis at high temperatures, *p*-NP caprylate (C8) was used as a standard substrate. Unless otherwise indicated, Est1 activity was measured at 80 °C in 1 ml containing 50 mM sodium phosphate assay buffer (pH 8),1 mM *p*-NP caprylate (C8), and 1% (*v*/*v*) isopropanol, while Est2 activity was measured at 70 °C in 1 ml containing 50 mM TAPS (3-(2, 4 dinitrostyrl)-(6R,7R-7-(2-thienylacetamido)-ceph-3-em-4-carboxylic acid) assay buffer (pH 9), 1 mM *p*-NP caprylate (C8), and 1% (*v*/*v*) isopropanol. The assay buffer was initially incubated in a screwed-cap test tube for 10 min at assay temperature. Then, the reaction was initiated by adding enzyme and substrate to the buffer. The amount of *p*-nitrophenol released by esterase-catalyzed hydrolysis was continuously monitored at a wavelength of 410 nm against an enzyme-free blank. One unit (U) of enzyme activity was defined as the amount of enzyme that released 1 μmol of *p*-nitrophenol per minute. All experiments

were performed in at least triplicate. Results are shown as mean values ± standard deviation (SD).

Substrate specificities of Est1 and Est2 were checked towards the following *p*-NP acyl esters of different chain lengths: *p*-NP acetate (C2), *p*-NP butyrate (C4), *p*-NP valerate (C5), *p*-NP caproate (C6), *p*-NP caprylate (C8), *p*-NP caprate (C10), *p*-NP laurate (C12), *p*-NP myristate (C14), and *p*-NP palmitate (C16). Considering the instability of short-chain substrates, the assay temperature was decreased to 50 °C. Initial rates of reaction for *p*-NP butyrate and *p*-NP valerate were calculated by estimating Est1 and Est2 activities with different substrate concentrations ranging from 1 to 2000 μM. Values for K_m and V_{max} were determined by employing the Lineweaver-Burk plots (Lineweaver and Burk 1934). Lipolytic activity towards different triacylglycerides was also measured qualitatively by incubating Est1 and Est2 on agar plates emulsified with tributyrin (C4), tricaproin (C6), tricaprylin (C8), tricaprin (C10), trilaurin (C12), trimyristin (C14), or tripalmitin (C16). Formation of clearing zones (halos) on agar plates indicated lipolytic activity. Beta-lactamase activity of Est2 was tested spectrophotometrically at 486 nm, under standard assay conditions with 1 mM nitrocefin (E-isomer) as substrate.

Effect of temperature and pH

The effect of pH on Est1 and Est2 activities was measured at 348 nm (the pH-independent isosbestic wavelength) under standard assay conditions (Glogauer et al. 2011). The following overlapping buffer systems were used: 50 mM acetate buffer (pH 3.0 to 6.0), 50 mM sodium phosphate buffer (pH 6.0 to 8.0), 50 mM TAPS buffer (pH 8.0 to 9.0), and 50 mM CHES (*N*-cyclohexyl-2-aminoethanesulfonic acid) buffer (pH 9.0 to 10.0). Temperature optima for Est1 and Est2 activities were measured in a temperature range of 20 to 100 °C. Thermostability of enzyme activity was determined by incubating Est1 and Est2 in their optimal buffers at various temperatures (50 to 80 °C) for up to 6 days. Subsequently, Est1 and Est2 activities were determined under standard assay conditions.

Combined effect of pH and temperature on the stabilities of Est1 and Est2

Second-order rotatable design was applied to study the combined effect of pH and temperature on the stabilities of Est1and Est2. The design was based on five levels and two variables (Table S2). Experimental data were fitted to the empirical model using Eq. (1):

$$Z\,(\%) = b_0 + b_1X + b_2Y + b_{12}XY + b_{11}X^2 + b_{22}Y^2 \quad (1)$$

in which Z was residual relative activity, presented as the percentage of activity measured before incubation and under standard assay conditions; X and Y were code values of pH and temperature shown in Table S2; b_0, b_1, b_2, b_{12}, b_{11}, and b_{22} were regression coefficients. Significance of regression coefficients was checked by Student's t test ($\alpha = 0.05$). Statistically non-significant coefficients were removed, and best-fit parameters were recalculated (Lazić 2004). The consistency of regression models was checked by Fisher's test ($\alpha = 0.05$). The ratios of the following mean squares were compared with the F-criterion tabular values. Based on the following mean square ratios (Box et al. 2005), models were accepted if:

$$F_1 = \text{Model}/\text{Experimental error} \qquad F_1 \geq F^{\text{num}}_{\text{den}} \quad (2)$$

$$F_2 = \text{Lack of fitting}/\text{Experimental error} \qquad F_2 \leq F^{\text{num}}_{\text{den}} \quad (3)$$

Est1 and Est2 were incubated under conditions described in Table S1 for 2 h, and residual activity was subsequently measured under the respective standard assay conditions.

Effect of miscible and immiscible organic solvents

The following organic solvents with different log *p* values were used in this study: water-miscible organic solvents of DMSO (− 1.3), methanol (− 0.75), ethanol (− 0.24), acetone (− 0.24), isopropanol (0.074), and 1-propanol (0.28), as well as the water-immiscible organic solvents of ethyl acetate (0.68), diethyl ether (0.85), chloroform (2.0), and toluene (2.5). The effects of water-miscible organic solvents on Est1 and Est2 activities were measured by adding each organic solvent into the assay buffer to obtain a final concentration ranging from 6 to 30% (*v*/*v*) under standard assay conditions. Enzyme activity measured in organic solvent-free assay buffer was regarded as 100%. Appropriate controls were also set to eliminate changes in extinction coefficients due to the presence of solvents.

To evaluate short-term stability towards water-miscible and water-immiscible organic, Est1 and Est2 were incubated in 100-μl aliquots with different amounts of water-miscible organic solvents (0 to 75%, *v*/*v* for Est1; 0 to 95%, *v*/*v* for Est2) or water-immiscible organic solvents (15 and 30%, *v*/*v*) at 30 °C for 2 h with vigorous shaking (300 rpm). The long-term stability towards water-miscible solvents was only measured for Est2. In the presence of 30 or 60% (*v*/*v*) organic solvents, Est2 was incubated at 30 °C with constant shaking in a screwed-cap test tube for up to 26 or 13 days, respectively. Enzyme activities in either the aqueous phase (for water-immiscible solvents) or the mixture (for water-miscible solvents) were measured. Each water-miscible organic solvent was equalized to the same final concentration in the assay buffer and the residual activity was measured under standard

assay conditions. A blank reference was prepared by using the same buffer solution without enzyme containing and the same amount and type of organic solvent (Shao et al. 2013). Residual activity was subsequently measured under the respective standard assay conditions. The activity measured at the start of the experiment was taken as 100%.

Effect of additives on Est1 and Est2 activities

The effects of metal ions on Est1 and Est2 activities were examined in 50 mM sodium phosphate buffer (pH 8) at 50 °C, in the presence of 1 mM and 10 mM KCl, $CaCl_2$, $MnCl_2$, $MgCl_2$, $ZnSO_4$, $FeSO_4$, $CuCl_2$, $NiSO_4$, $FeCl_3$, and $AlCl_3$. The inhibitory effect on enzyme activity was measured under standard assay conditions with the following known esterase effectors (each 1 mM and 10 mM): phenylmethylsulfonyl fluoride (PMSF), dithiothreitol (DTT), and ethylenediaminetetraacetic acid (EDTA). In addition, the effect of the following detergents (each 0.1 and 1%, *v*/*v*), Triton X-100, Tween 20, Tween 80, and SDS, was determined. Furthermore, the effect of NaCl and KCl on enzyme activity in a range between 0.5 and 4 M was assessed. Activity measured in additive-free assay buffer was regarded as 100% activity, while reactions that included corresponding additive but no enzyme were used as blanks.

Accession number

The gene sequences are available at the GenBank database under accession numbers KR149567.1 (Est1) and KR149568.1 (Est2).

Results

Metagenomic library screening and analysis of two novel esterase-encoding genes

To isolate novel lipolytic enzymes, a compost sample at the thermophilic stage (55 °C) was used for constructing a metagenomic plasmid library. The library consisted of approximately 675,200 clones with an average insert size of 5.3 kb and comprised a total size of 3.58 GB. Among the 279 lipolytic-positive clones, two *E. coli* clones harboring the plasmids pFLD_Est1 and pFLD_Est2 showed strong lipolytic activity (large halos) on indicator plates and were selected for further characterization.

Sequence analyses of the plasmids pFLD_Est1 and pFLD_Est2 revealed that each contained one putative esterase-encoding gene, *est1* (813 bp) and *est2* (1170 bp), respectively. The deduced proteins comprised 270 (Est1) and 389 (Est2) amino acids. Putative signal peptides indicating extracellular localization were not detected in the deduced protein sequences. Sequence similarity searches showed that Est1 exhibited 49% identity to a hypothetical protein from *Candidatus Entotheonella* (GenBank: ETW96815) and 43% identity to Est28 from a grassland soil metagenomic library (Nacke et al. 2011). Est2 showed 53% sequence identity to a beta-lactamase from *Streptomyces lavenduligriseus* (GenBank: WP_030784121) and 52% sequence identity to a putative esterase from *Streptomyces bottropensis* (GenBank: EMF58012).

Sequence analysis and the subsequently constructed phylogenetic tree revealed that Est1 belonged to family V and Est2 to family VIII of lipolytic enzymes (Fig. S1). The tertiary structure predicted by I-TASSER obtained the *C*-scores of 1.03 for Est1 and 1.27 for Est2, which indicated a significant confidence of good quality.

The two conserved family V motifs G-X-S-X-G-G and P-T-L were present in the Est1 protein sequence at amino acid positions 92 to 97 and 208 to 210, respectively (Fig. S2a). The tertiary structure of Est1 was composed of a cap domain with five α-helices (α4 to α8, Fig. S2a) and an α/β-hydrolase fold core domain (Fig. 1a). The core domain consists of six helices surrounded by eight β-strands that form parallel structures, in which Ser^{94} is located between β5 and α3, Asp^{217} after β7, and His^{268} between β8 and α10. The overall structure of Est1 superimposed on MGS-M2 (TM score 0.96; RMSD 0.94) (Alcaide et al. 2015), with a global amino acid sequence identity of 21.9% (Fig. 1c).

Due to the high sequence similarity between family VIII lipolytic enzymes and class C beta-lactamases/penicillin-binding proteins, amino acid sequences in the two categories were aligned (Fig. S2b). All the aligned sequences shared the same catalytic triad of serine, lysin, and tyrosine. Conserved family VIII motifs including S^{70}-X-X-K^{73}, G^{154}-X-X-X-X-H^{159}, and H/W^{341}-X-G^{343} were detected in Est2 sequences. However, classical β-lactamase motifs such as Y-A-N and L-S/T-G (KTG-box) were absent in the Est2 sequence (Fig. S2b). Similar to Est1, the tertiary structure of Est2 also consists of two domains: a large α/β domain and a small α-helical domain. The small α-helix domain contains four helices and two 3_{10} helices. For the large α/β domain, a central eight-stranded antiparallel β-sheet is flanked by six helices on one face and two on the opposite face (Fig. 1c). The deduced catalytic residues of Ser^{70}, Lys^{73}, and Tyr^{159} are located in the large cavity between the two domains (Fig. 1b). The overall structure of Est2 superimposed well on CcEstA (TM score 0.95; RMSD 0.46) (Woo Lee et al. 2016), with a global amino acid sequence identity of 39.6% (Fig. 1d).

Purification of Est1 and Est2 and substrate specificity towards *p*-NP acyl esters

The esterase-encoding genes *est1* and *est2* were successfully cloned in expression vectors and expressed in *E. coli* BL21(DE3). The produced gene products Est1 and Est2 were

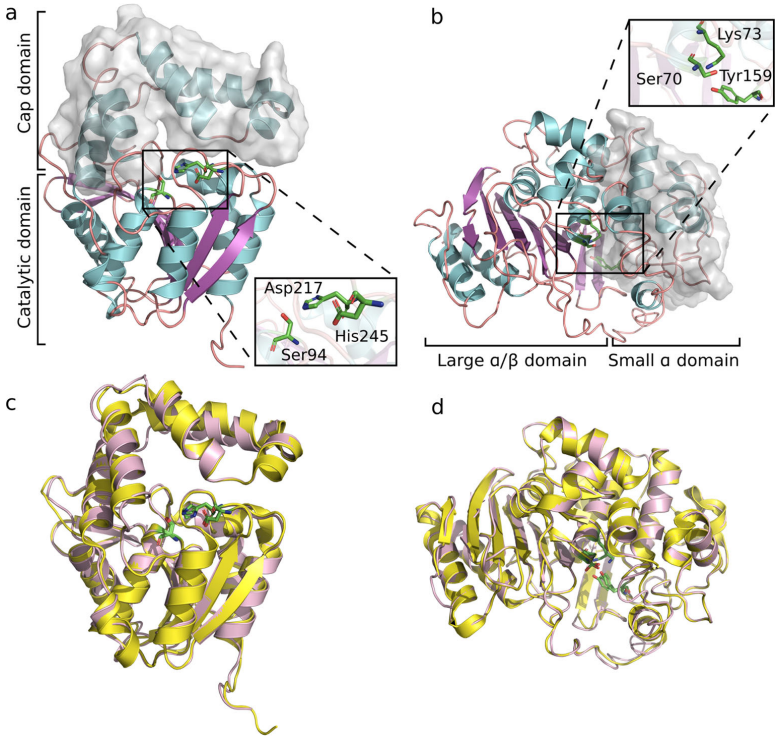


Fig. 1 The modeled three-dimensional structure of Est1 and Est2. **a** A 3D structure model of Est1. The overall structure is composed of two domains: cap domain and catalytic domain. **b** A 3D structure model of Est2. The overall structure harbors two domains: a large α/β domain and a small α-helix domain. α-helices and β-strands are colored in cyan and magenta, respectively. The catalytic triad of Est1 (residues Ser^{94}, Asp^{215}, and His^{245}) and that of Est2 (residues Ser^{70}, Lys^{73}, and Try^{160}) are indicated in stick representation. The cap domain for Est1 and the small α-helix domain for Est2 are presented as transparent surface. **c** Structural superposition of Est1 (pink) onto its structural homolog MGS-M2 (yellow; PDB: 4Q3L). **d** Structural superposition of Est2 (pink) onto its structural homolog CcEstA (yellow; PDB:5GKV)

subsequently purified by Ni-TED affinity chromatography, yielding 28.9-fold and 16.9-fold purification values and specific activities of 22.2 and 7.3 U/mg, respectively (Table S2). SDS-PAGE analysis revealed single bands with molecular masses of approximately 35 kDa (Est1) and 48 kDa (Est2) (Fig. S3), which were in accordance with the calculated masses including during cloning added V5 epitope and His_6-tag.

Est1 and Est2 exhibited a substrate preference for esters with short-chain fatty acids. The maximal activities were detected with *p*-NP butyrate (C4) for Est1 and *p*-NP valerate (C5) for Est2 as substrates (Fig. 2). Correspondingly, *p*-NP butyrate and *p*-NP valerate were used for calculating the K_m and V_{max} values. The K_m and V_{max} values of Est1 were 3.0 μM and 31.2 U/mg, respectively, and that of Est2 2.7 μM and 19.0 U/mg, respectively. With respect to esters with long-chain fatty acids as substrates (C10 to C16), Est1 was active with *p*-NP caprate (C10) retaining 45% activity, but barely with other substrates. In comparison, Est2 hydrolyzes more acyl esters with long-chain fatty acids. It showed 80%, 45%, 12%, and 30% activity towards *p*-NP caprate (C10), laurate (C12), *p*-NP myristate (C14), and *p*-NP palmitate (C16), respectively. Chain length selectivity towards triacylglycerides was recorded for the substrates tributyrin (C_4) and tricaproin (C_6) for both enzymes. The preference for short substrates (< C10) indicates that Est1 and Est2 are esterases, rather than "true" lipases which are commonly found in family I of lipolytic enzymes (Arpigny and Jaeger 1999). Moreover, beta-lactamase activity of Est2 with nitrocefin as substrate was tested, but no significant beta-lactamase activity was detected (data not shown).

Effect of pH and temperature on Est1 and Est2 activities

Est1 and Est2 were active over the entire tested pH range (3 to 10), with the exception of Est1 at pH 3 (Fig. 3a, d). Est1 exhibited maximal activity at pH 7 and retained more than 95% activity between pH 6 and 8. Est2 activity increased with pH and peaked at pH 9. With respect to the temperature dependence of enzyme activity, Est1 and Est2 displayed a similar ascending trend along the temperature gradient, showing a maximal activity at 80 and 70 °C, respectively. At higher temperatures (90 and 100 °C), approximately 80% activity was retained (Fig. 3b, e).

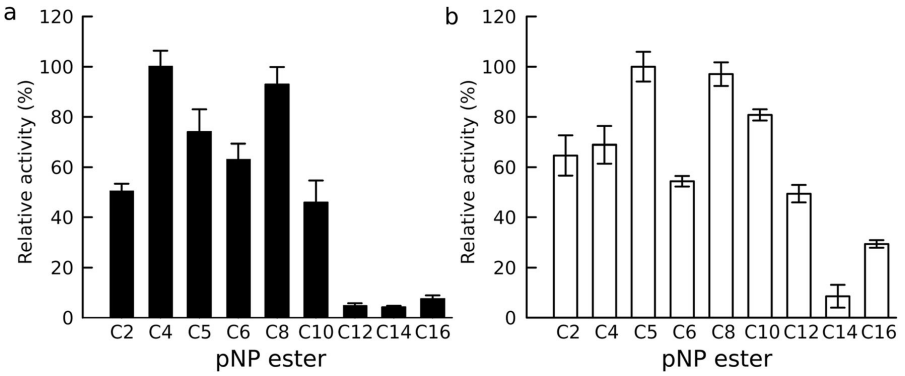


Fig. 2 Substrate specificity towards *p*-NP acyl esters with different chain lengths. **a** Est1, the maximal activity (23.3 U/mg) measured towards *p*-NP butyrate (C4) was taken as 100%. **b** Est2, the maximal activity (10.7 U/mg) measured towards *p*-NP valerate (C5) was taken as 100%

Despite that the activity maximum of Est1 was at 80 °C, stability at 70 °C was low as a rapid drop of activity to 50% was detected after 15 min of incubation. Similar results were obtained for Est2 at 80 °C with 35% residual activity after 30 min. Strikingly, Est2 showed significant stability at 70 °C, with 50% residual activity after 12 h of incubation. At lower temperatures (50 or 60 °C), the activities of both enzymes were remarkably stable for extended incubation times; nonetheless, Est1 displayed a higher stability than Est2 (Fig. 3c). Est1 retained more than 80% activity at 50 °C over the entire incubation period (7 days), whereas 52% residual activity was observed for Est2 after 5 days. When incubated at 60 °C, Est1 exhibited a half-life of 2 days, which is nearly twice as that of Est2 (Fig. 3f).

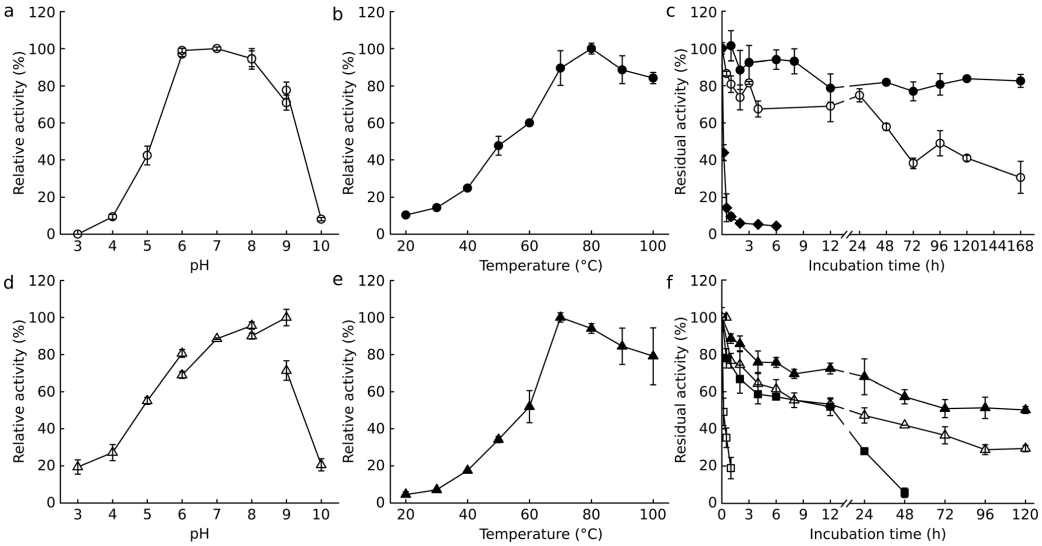


Fig. 3 Effect of pH and temperature on esterase activity. **a** Effect of pH on Est1 activity. Maximal activity at pH 7 (45.8 U/mg) was taken as 100%. **b** Effect of temperature on Est1 activity. Maximal activity at 80 °C (46.1 U/mg) was taken as 100%. **c** Thermostability of Est1 at 50 °C (closed circle), 60 °C (open circle), and 70 °C (closed diamond); activity measured before incubation (40.4 U/mg) was taken as 100%. **d** Effect of pH on Est2 activity. Maximal activity at pH 9 (14.4 U/mg) was taken as 100%. **e** Effect of temperature on Est2 activity. Maximal activity at 70 °C (15.2 U/mg) was taken as 100%. **f** Thermostability of Est2 at 50 °C (closed triangle), 60 °C (open triangle), 70 °C (closed square), and 80 °C (open square); activity measured before incubation (12.0 U/mg) was taken as 100%

Combined effect of pH and temperature on Est1 and Est2 stabilities

The combined effect of temperature and pH on the stability of Est1 and Est2 was visualized by response surfaces (Fig. 4). After removing the insignificant terms and refitting to the empirical model, the following equations for Est1 (4) and Est2 (5) stability under different conditions were obtained:

$$\text{Residual activity}\ (\%) = 76.33 + 13.56\ \text{pH} - 31.89\ T - 11.19\ \text{pH}T - 19.09\ \text{pH}^2 - 17.82T^2 \quad (4)$$

$$\text{Residual activity}\ (\%) = 84.1 + 14.1\ \text{pH} - 19.3\ T - 6.3\ \text{pH}^2 - 23.8\ \text{T}^2 \quad (5)$$

Variance analysis implied that the models for Est1 and Est2 significantly fitted to the experimental data ($F_1 \geq F_{den}^{num}$, $F_2 \leq F_{den}^{num}$, $\alpha = 0.05$; Table S3). In equations (4) and (5), the negative coefficients for T, pH^2 and T^2 indicate the existence of an absolute maximum of residual activity. The calculated optimal conditions were located at pH 8.2 and 39.1 °C (Est1) and pH 8.7 and 48.9 °C (Est2). It was predicted that Est1 and Est2 retain 98.6 and 96.3% residual activity, respectively, under these conditions. Moreover, according to the elliptical contour plot on the pH–T (x–y) dimension, the pH and temperature ranges in which Est1 and Est2 retain more than 80% residual activity were pH 7.2 to 9.2 and 23.6 to 54.6 °C for Est1 and pH 7.0 to 10.4 and 37.4 to 60.4 °C for Est2 (Fig. 4). In conclusion, Est1 and Est2 are stable under thermophilic and alkaline conditions.

Effect of water-miscible organic solvents on Est1 and Est2 activities

Est1 activity increased to approximately 150% in the presence of low concentrations of 1-propanol (below 6%, *v*/*v*), and ethanol and isopropanol (< 12%, *v*/*v*). At higher concentration, Est1 activity decreased rapidly and was not detectable at 18% (*v*/*v*) 1-propanol, 30% (*v*/*v*) ethanol, and 24% (*v*/*v*) isopropanol (Fig. 5a). Addition of DMSO, methanol, and acetone inhibited Est1 activity over the tested concentration range, retaining 50% activity at 30% (*v*/*v*) DMSO, or less than 10% activity at 30% (*v*/*v*) methanol and acetone.

Ethanol, methanol, isopropanol, and 1-propanol had a stimulatory effect on Est2 activity almost at all tested concentrations (Fig. 5b). In the presence of ethanol and methanol, Est2 activity increased continually with raising concentrations. At 30% (*v*/*v*) ethanol and methanol, Est2 activity increased 8.8-fold and 5.4-fold, respectively. For isopropanol and 1-propanol, optimal Est2 activity was recorded at concentrations of 18 and 12% (*v*/*v*), respectively (Fig. 5b). In the presence of DMSO, Est2 retained more than 70% activity at all tested concentrations, with the maximal activity at 6% (*v*/*v*). Acetone caused an activity decrease to 34.1% at 30% (*v*/*v*).

Effect of water-miscible organic solvents on Est1 and Est2 stabilities

Est1 retained almost unchanged residual activity after 2 h of incubation in the presence of 15 and 30% (*v*/*v*) of all tested organic solvents, with the exception of 1-propanol (Fig. 6a). Est1 residual activity rapidly dropped to 17.0% at 30% (*v*/*v*) 1-propanol. Decreased tolerance was also detected after exposure to 45% (*v*/*v*) isopropanol, methanol, acetone, and DMSO for 2 h (Fig. 6a). At 60% (*v*/*v*) solvent concentration, Est1 exhibited moderate tolerance to DMSO (60.5% residual activity), but lost most of its activity (less than 10%

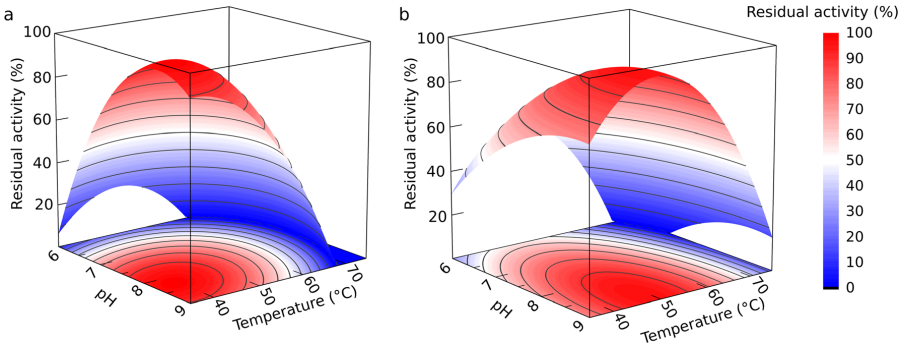


Fig. 4 Response surface corresponding to the combined effect of pH and temperature on Est1 (**a**) and Est2 (**b**) stabilities. Residual activity was expressed as a percentage of the initial activity measured before incubation under standard assay conditions. The elliptical contour plot on the pH–T dimension demonstrates the change of residual activity (%)

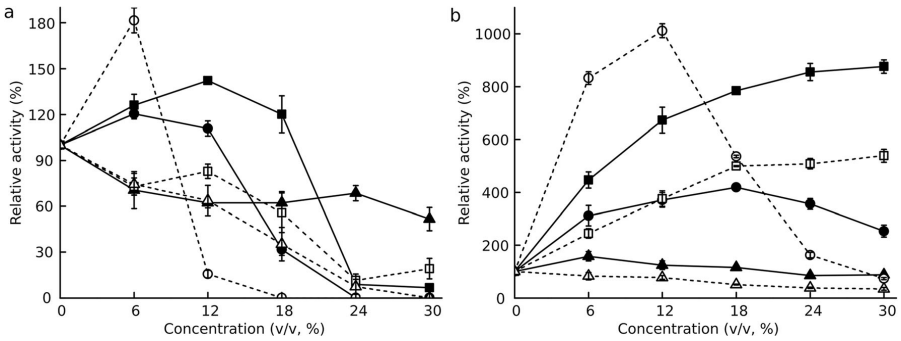


Fig. 5 Effects of water-miscible organic solvents on of Est1 (**a**) and Est2 (**b**) activities. Catalytic activity was measured under standard assay conditions in the presence of different amounts of ethanol (closed square), methanol (open square), isopropanol (closed circle), 1-propanol (open circle), DMSO (closed triangle), and acetone (open triangle). Specific activities corresponding to 100% relative activity were 21.9 U/mg (Est1) and 5.1 U/mg (Est2)

residual activity) in the presence of the other tested organic solvents. Est1 activity was almost inactivated by all the tested organic solvents at 75% (*v*/*v*).

In comparison with Est1, Est2 displayed higher resistance towards water-miscible solvents (Fig. 6b). In the organic solvent concentration range of 0 and 60% (*v*/*v*), Est2 retained its

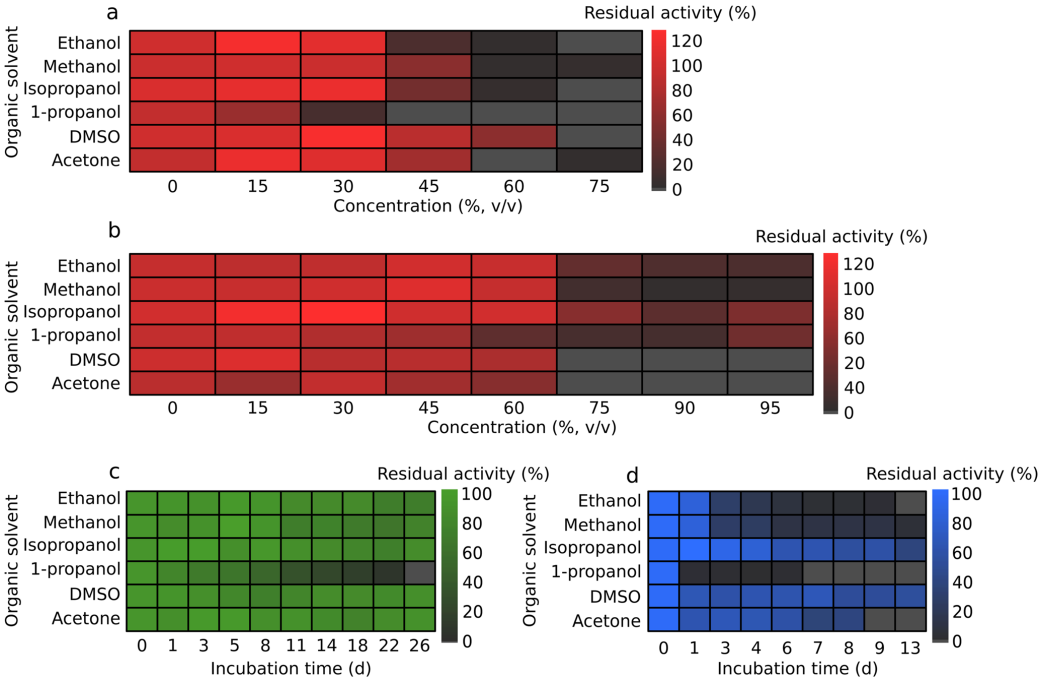


Fig. 6 Heatmap displaying the effect of enzyme stability towards water-miscible organic solvents. Short-term stability of Est1 (**a**) and Est2 (**b**) towards different concentrations of organic solvents at 30 °C for 2 h. The specific activity expressed as percentages of Est1 reference reactions (100%) are 23.3, 17.5, 24.3, 6.6, 15.3, and 13.6 U/mg for reactions in the presence of ethanol, methanol, isopropanol, 1-propanol, DMSO, and acetone, respectively, and that of Est2 are 32.7,14.4, 17.7, 42.7, 4.34, and 3.1 U/mg, respectively. Long-term stability of Est2 towards 30% (*v*/*v*; **c**) and 60% (*v*/*v*; **d**) organic solvents for prolonged time periods at 30 °C. The specific activity values expressed as percentages of Est2 reference reactions (100%) are 14.3, 7.1, 9.2, 25.9, 5.0, and 4.1 U/mg for reactions in the presence of ethanol, methanol, isopropanol, 1-propanol, DMSO, and acetone, respectively

full activity in the presence of isopropanol, methanol, and ethanol and had a slight activity loss by the addition of DMSO and acetone. On the contrary, Est2 stability decreased along the rising concentration of 1-propanol, retaining 32.4% residual activity at 60% (*v*/*v*). Incubating with concentrations above 60% (*v*/*v*) of isopropanol, 1-propanol ethanol, or methanol led to a rapid decline of Est2 residual activity (Fig. 6b). Exposure to 75% (*v*/*v*) DMSO and acetone led to inactivation of Est2 activity.

Est2 exhibited high stability in the presence of 30% (*v*/*v*) ethanol, methanol, isopropanol, DMSO, and acetone, as its residual activity was above 70% over the extended incubation period of up to 26 days (Fig. 6c). This residual activity was even higher than that Est2 exhibited during incubation without additions (Fig. S4). However, a continuous drop in Est2 residual activity was observed when exposed to 1-propanol. Est2 also displayed decreased tolerance towards 60% (*v*/*v*) organic solvents (Fig. 6d). Incubation with 1-propanol, ethanol, and methanol reduced Est2 activity rapidly (< 10% residual activity). Nonetheless, Est2 exhibited substantial tolerance against isopropanol, DMSO, and acetone, retaining approximately 40% residual activity after the 13-day incubation.

Effect of water-immiscible organic solvents on Est1 and Est2 stabilities

Incubation with water-immiscible organic solvents of diethyl ether, chloroform, and toluene resulted in deleterious effects on Est1 enzyme activity, which was not detectable at the tested concentrations (Table 1). However, Est1 activity displayed some resistance towards ethyl acetate, with 46.6 and 21.1% residual activities at the tested concentrations of 15 and 30% (*v*/*v*), respectively. In contrast, Est2 was tolerant to ethyl acetate, diethyl ether, and toluene. Est2 retained its activity after incubation with ethyl acetate and diethyl ether. In the presence of toluene, Est2 retained 87.0 and 68.8% activities at the tested concentrations of 15 and 30% (*v*/*v*), respectively. In addition, Est2 was as Est1 inactivated by chloroform.

Effect of other additives on Est1 and Est2 activities

Metal ions, inhibitors, detergents, and salts were also analyzed for their effects on Est1 and Est2 activities. Est1 and Est2 are generally resistant to various metal ions (Table S4). The addition of tested metal ions at 1 mM concentration had minor effects on Est1 and Est2 activities as approximately 90% of the activity was retained. Moreover, both enzymes exhibited substantial tolerance (above 50% activity) towards some metal ions such as Ca^{2+}, Zn^{2+}, Cu^{2+}, Mn^{2+}, Fe^{3+}, and Al^{3+} at a concentration of 10 mM. Enzyme activity was slightly enhanced in the presence of 10 mM Mg^{2+} (Est1) and Zn^{2+} (Est2). The presence of the chelating agent EDTA did not affect Est1 activity but decreased Est2 activity to less than 70%. The latter results indicated that Est1 activity is independent of metal ions, whereas Est2 might be a metalloenzyme (Mohamed et al. 2013).

With respect to the detergents, Est1 activity was enhanced in the presence of 0.1% (*v*/*v*) Tween 20 and 0.1 and 1% (*v*/*v*) Tween 80. Est2 activity was less resistant than that of Est1 towards the tested detergents. Est2 retained more than 70% activity after addition of 1 mM Triton-100, Tween 20, and Tween 80. The anionic detergent SDS inhibited the activity of both enzymes entirely (Table S4). The activity of Est1 and Est2 was substantially decreased at 10 mM DTT and PMSF. DEPC displayed detrimental effects on Est1 activity, while Est2 activity was almost unaffected even at a concentration of 10 mM. The inhibition of enzyme activity by SDS and PMSF indicates that Est1 and Est2 belong to the serine hydrolases (Peng et al. 2011). In addition, Est1 and Est2 were active in the presence of 0 to 4 M NaCl and KCl. Both enzymes retained more than 50% of activity up to 2.5 M NaCl and KCl, which is indicative of halotolerance (Table S5).

Discussion

Extreme environments such as compost have been used for mining biocatalysts, which are likely to be adapted to harsh industrial reaction conditions (Ryckeboer et al. 2003). In this study, compost derived from the thermophilic core of the pile was used as DNA source to construct a metagenomic library, whereby two putative genes encoding lipolytic enzymes were identified. The low identities of the deduced amino acid sequences to known proteins (Est1 43% and Est2 53%) indicate that Est1and Est2 are novel lipolytic enzymes. Est1 is most related to the characterized esterase EstPS5, which was derived from a screening of a peat-swamp forest soil metagenome (Bunterngsook et al. 2010). Est2 shared the highest sequence identity with a beta-lactamase from *Streptomyces achromogenes*. In addition, most of the enzyme sequences similar to Est2 were derived from members of the *Streptomyces* genus. *Streptomyces* strains are predominantly found in soil and decaying vegetation and well-known as important natural sources for antibiotics (Raja and Prabakarana 2011).

Phylogenetic analyses indicate that Est1 and Est2 belong to family V and family VIII of lipolytic enzymes, respectively (Fig. S1). For most of the known lipolytic enzymes, the catalytic activity generally relies on the typical catalytic triad of Ser, Asp/Gly, and His (Ollis et al. 1992; Jaeger et al. 1999). This triad is also present in the Est1 amino acid sequence (Fig. S2a) and its tertiary structure (Fig. 1a). However, another catalytic triad consisting of Ser^{70}, Lys^{73}, and Tyr^{159} (Sakai et al. 1999) is responsible for Est2 catalytic activity (Fig. 1c, Fig. S2b). This triad is conserved among family VIII lipolytic enzymes, class C β-lactamases, and penicillin-binding proteins (Arpigny and Jaeger 1999; Hausmann and Jaeger 2010; Biver

Table 1 Short-term stabilities of Est1 and Est2 towards water-immiscible organic solvents.

Water-immiscible organic solvent (log *P*)	Residual activity (%)[a]		
	Concentration (*v*/*v*)	Est1	Est2
None[b]	–	100 ± 2.6	100 ± 1.5
Ethyl acetate (0.68)	15	46.6 ± 5.7	95.5 ± 4.2
	30	21.1 ± 3.0	101.5 ± 8.9
Diethyl ether (0.85)	15	ND[c]	116.0 ± 3.8
	30	ND[c]	96.5 ± 4.2
Chloroform (2.0)	15	ND[c]	ND[c]
	30	ND[c]	ND[c]
Toluene (2.5)	15	ND[c]	87.0 ± 6.0
	30	ND[c]	68.8 ± 8.0

[a] Aliquots of enzyme were shortly centrifuged; the aqueous layer was used for activity measurements under standard assay conditions

[b] Specific activity corresponding to 100% residual activities were 21.9 U/mg (Est1) and 11.2 U/mg (Est2)

[c] Not detectable

and Vandenbol 2013; Popovic et al. 2017). Despite the same catalytic triad and high amino acid sequence identity to β-lactamases, some family VIII esterases show promiscuous β-lactamase activity. Similar to Est2, the family VIII esterases EstB (Petersen et al. 2001), Lip8 (Ogino et al. 2004), Est2K (Kim et al. 2010), and Est7K (Woo Lee et al. 2016) showed negligible or no detectable β-lactamase activity, whereas EstC (Rashamuse et al. 2009), EstM-N1 (Yu et al. 2011), and PBS-2 (Boyineni et al. 2014) exhibited moderate or high β-lactamase activity. Yu et al. (2011) suggested that those promiscuous β-lactamase activities could be a result of family VIII esterases evolving from class C β-lactamases or vice versa.

According to the predicted tertiary structure (Fig. 1), the active sites of Est1 and Est2 are protected by an α-helix domain (cap domain). This domain acts as a shield for the catalytic site of many lipolytic enzymes and appears to play a key role in several functional aspects, such as activity, substrate specificity, and thermostability (Gall et al. 2014; Li et al. 2015; Kim 2017). Est2 was capable of utilizing acyl esters with long-chain fatty acids as substrate (C10, C12, and C16) (Fig. 2b). In addition, the Est2 structural homolog Est-Y29 (Ngo et al. 2013) was also reported to hydrolyze a wide variety of hydrophobic compounds, as a result of the deep hydrophobic patch between the large α/β domain in which the small α-helix domain defines a wide active site (Ngo et al. 2014).

Est1 and Est2 were generally thermoalkaline esterases (Fig. 3). In terms of thermophilicity, defined as increased enzymatic activity along a temperature gradient (Georis et al. 2000), Est1 and Est2 have great advantages in comparison with its homologs from other sources. Although esterases from *Fervidobacterium nodosum* (Yu et al. 2010), *Anoxybacillus gonensis* (Faiz et al. 2007), *Sulfolobus tokodaii* (Angkawidjaja et al. 2012), and a compost metagenome (Riedel et al. 2015) displayed optimal activities above 70 °C, Est1and Est2 exhibited higher activities (above 80%) at 90 and 100 °C. Moreover, Est1 and Est2 displayed unprecedented thermostability at high temperatures, with a half-life of more than 7 days at 50 °C and 2 days at 60 °C for Est1 (Fig. 4c) and 5 days at 50 °C, 1 day at 60 °C, and 12 h at 70 °C for Est2 (Fig. 4f). This remarkable feature distinguished the two enzymes from their thermophilic counterparts, such as esterases from the thermophilic microorganisms *Archaeoglobus fulgidus* (40% residual activity after 30 min at 60 °C; D'Auria et al. 2000), *Thermogutta terrifontis* (75% residual activity after 30 min at 70 °C; Sayer et al. 2015a), and *Thermoanaerobacter tengcongensis* (half-life of 2 h at 70 °C; Cook et al. 1996), as well as those from metagenomes of compost (approx. 90% residual activity after 1 h at 60 °C; Kang et al. 2011) and hot spring (half-life of 6 h at 60 °C; Zarafeta et al. 2016).

In general, thermostability is dependent on the structural rigidity, which is an accumulation of various features, including but not limited to amino acid composition, ion pairing, hydrogen bonds, hydrophobic interactions, and sulfide bridges (Sadeghi et al. 2006; Jochens et al. 2010; Ebrahimi et al. 2011; Pezzullo et al. 2013). Interestingly, the two characterized structural homologs of Est1, MGS-M2 (Alcaide et al. 2015) and TtEst (Sayer et al. 2015b), were reported as thermostable esterases (Alcaide et al. 2013; Sayer et al. 2015a). The thermostability of Est1 and Est2 is advantageous for industrial applications, as higher reactivity, stability, and process yields, as well as lower viscosity and fewer contamination problems, can be achieved at an elevated operation temperature (Lima et al. 2004; Panda and Gowrishankar 2005; Doukyu and Ogino 2010; Sood et al. 2016).

Temperature and pH play important roles in enzyme-catalyzed reactions and the two factors have to be considered together for optimization of enzyme reactions. Est1 and Est2 are most stable at conditions which are close to that of the original compost habitat (Fig. 4a, b). Generally, enzymes are stable at conditions similar to their original habitats (Elend et al. 2006; Kovacic et al. 2016), but show maximal activities at higher or lower temperatures (Hardeman and Sjoling 2007; Hu et al. 2010), which is also the case for Est1 and Est2 (Fig. 3). Est1 and Est2 were predicted to retain more than 80% residual activity over a broad temperature and pH range (Fig. 4). Thus, together with the feature of broad substrate specificity (Fig. 2), the two enzymes, particularly Est2, could be potentially utilized in detergents, in which high enzyme stability is required during washing conditions of pH 8 to 11 and temperatures of 30 to 60 °C, as well as versatile substrate specificity (Nerurkar et al. 2013; Bora 2014).

Water-miscible organic solvents are generally detrimental to enzymes. In contrast, Est2 activity was significantly enhanced by the addition of ethanol, isopropanol, methanol, and 1-propanol (Fig. 5b). The stimulated catalytic activity of esterase from *Burkholderia cepacia* was also observed in the presence of DMSO, DMF, methanol, ethanol, 2-propanol, and acetone (Takeda et al. 2006). Similar results were obtained for family VIII esterases Est2K, lpc53E1, and Est7K in the presence of isopropanol and methanol (Kim et al. 2010; Selvin et al. 2012; Woo Lee et al. 2016). The significant activating effect could be attributed to the uniform water phase formed by water-miscible solvents (Ogino and Ishikawa 2001) or the high diffusion rate of substrate in the presence of water-miscible solvents (Metin et al. 2006), which enables substrates quick and easy access to the active site. Moreover, Est2 activity generally showed a bell-shaped dependence on water-miscible organic solvent concentration (Fig. 5b), which indicated an optimal water activity for its hydrolytic activity (Léonard-Nevers et al. 2009; Adlercreutz 2013). Other esterases, including Est1 (Fig. 5a), commonly showed slightly increased (Hotta et al. 2002; Schütte and Fetzner 2007; Faulds et al. 2011; Kang et al. 2017), or decreased (Li and Yu 2013; Monsef Shokri et al. 2014; Dukunde et al. 2017) activity towards water-miscible organic solvents. For these enzymes, water-miscible solvents strip off the crucial water monolayer around the enzyme surface and compete for hydrogen bonds, which at higher solvent concentrations finally leads to denaturation (Ó'Fágáin 2003; Doukyu and Ogino 2010; Monsef Shokri et al. 2014).

In addition, Est1 was stable in the presence of 30% (*v*/*v*) water-miscible organic solvents (Fig. 6a). Similar or even lower solvent tolerance was found for most of the reported OST esterases (Doukyu and Ogino 2010; Kang et al. 2011; Brault et al. 2012; Xing et al. 2012; Zhang et al. 2014; Kang et al. 2017; Kumagai et al. 2018). Thus, in comparison, Est2 exhibited superior stability against higher concentrations of the analyzed water-miscible organic solvents (Fig. 6b). Among the rare counterparts, a hyper-thermophilic archaeal esterase (Hotta et al. 2002) and EstB from *Alcanivorax dieselolei* B-5(T) (Zhang et al. 2014) displayed substantial stability towards certain water-miscible organic solvents at high concentrations. An increase of hydrophobic interactions and hydrogen bonds are essential in enhancing esterase tolerance against water-miscible organic solvents (Song and Rhee 2001; Kawata and Ogino 2009; Park et al. 2012). Est2 also showed considerable tolerance towards 30% (*v*/*v*) ethanol, isopropanol, DMSO, methanol, and acetone for up to 26 days (Fig. 6c) and towards 60% (*v*/*v*) isopropanol, DMSO, and acetone up to 13 days (Fig. 6d). The preservation of high esterase activity over an extended period has been rarely described previously (Sana et al. 2007; Jin et al. 2012; Li and Yu 2013). This feature could allow to apply Est2 as an immobilized biocatalyst in non-aqueous-based continuous bioprocesses (Sana et al. 2007; Yang et al. 2011).

Overall, both enhanced activity and adequate stability towards water-miscible organic solvents are pre-requisite properties for applications of esterases in organic synthesis (Panda and Gowrishankar 2005; Lopez-Lopez et al. 2014). In non-aqueous media, such enzymes could be exploited for ester synthesis and transesterification reactions and a variety of other chemical reactions (Salihu and Alam 2015; Sood et al. 2016; Sarmah et al. 2018). Thus, Est2 could be advantageous for use in biodiesel production, as the acyl acceptors methanol or ethanol are added in esterase/lipase-catalyzed transesterification reactions (Srimhan et al. 2011; Nasaruddin et al. 2014; Wang et al. 2017). In addition, Est1 and Est2 activities responded differently to the presence of polar protic solvents (methanol, ethanol, isopropanol, and 1-propanol) and polar aprotic solvents (acetone and DMSO), which further illustrates that the nature of the solvent influences esterase activity (Torres and Castro 2004). However, the relationship between the corresponding organic solvent logP_{ow} value and esterase activity/stability remains uncertain.

Comparing with water-miscible organic solvents, water-immiscible organic solvents are less deleterious for enzyme-solvent interactions (Rahman et al. 2005). However, most of the reported esterases were similar to Est1 (Table 1), showing a significant decrease of activity or inactivation after the incubation with water-immiscible organic solvents (Schütte and Fetzner 2007; Berlemont et al. 2013; Jin et al. 2012). In contrast, lipolytic enzymes of family I, also known as true lipases, are commonly resistant to water-immiscible organic solvents. In the presence of a water-solvent (hydrophobic) interface, the hydrophobic amino acid residues in the lid/flap region stabilize lipases in a flexible, open conformation (Dandavate et al. 2009; Yang et al. 2011; Kamal et al. 2013). However, the cap domain for most esterases is not intrinsically flexible and does not provide open and closed conformations (Bornscheuer 2002; Gall et al. 2014; Kim 2017). This could be the reason that water-immiscible tolerant esterases are rare. To the best of

our knowledge, only Est2 and the esterases EstC23 (Jin et al. 2012), LipBL (Pérez et al. 2012), Lpc53E1 (Selvin et al. 2012), RBest1 (Berlemont et al. 2013), Pf_Est (Mandelli et al. 2016), EST4 (Gao et al. 2016), and LipA9 (Park et al. 2018) show substantial resistance towards certain water-immiscible organic solvents. This feature further expands the application potential of Est2 to synthetic reactions in the presence of water-immiscible solvents (Gao et al. 2016; Sarmah et al. 2018).

Est1 and Est2 are to some extent resistant to metal ions (Table S4), which is an important feature in the bioremediation of environmental waste (Brault et al. 2012). Est1 and Est2 were also active at a salinity range of up to 4 M (Fig. S5), suggesting halotolerance (Jeon et al. 2012). Halotolerant enzymes are desirable in processes in which water activity is low (Delgado-García et al. 2012). In combination with the tolerance of Est1 and Est2 to organic solvents, it can be further confirmed that halotolerance is somehow positively correlated with organic solvent tolerance (Berlemont et al. 2013).

In conclusion, the characterization of Est1 and Est2 revealed that both, especially Est2, exhibit several application-relevant features, such as a broad substrate range, thermostability, halotolerance, and resistance to various organic solvents. In addition, as revealed by second-order rotatable design, Est1 and Est2 were predicted to be stable at a broad crossed range of temperature and pH. To the best of our knowledge, this is the first time to report an esterase, Est2, which simultaneously exhibits a significantly enhanced activity and unprecedented high stability towards water-miscible organic solvents and substantial tolerance towards water-immiscible organic solvents.

Acknowledgments We thank Dr Silja Brady and Sarah Zachmann for providing assistance. We acknowledge the support of Mingji Lu by "Erasmus Mundus Action 2 – Lotus I Project."

Author contributions The manuscript was written through the contributions of all authors. All authors have given approval to the final version of the manuscript.

Compliance with ethical standards This article does not contain any studies with human participants or animals performed by any of the authors.

Conflict of interest The authors declare that they have no conflict of interest.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

Adlercreutz P (2013) Immobilisation and application of lipases in organic media. Chem Soc Rev 42:6406–6436. https://doi.org/10.1039/c3cs35446f

Ahmed EH, Raghavendra T, Madamwar D (2010) An alkaline lipase from organic solvent–tolerant *Acinetobacter* sp. EH28: application for ethyl caprylate synthesis. Bioresour Technol 101:3628–3634. https://doi.org/10.1016/J.BIORTECH.2009.12.107

Alcaide M, Tornés J, Stogios PJ, Xu X, Gertler C, Di Leo R, Bargiela R, Lafraya Á, Guazzaroni M-E, López-Cortés N, Chernikova TN, Golyshina OV, Nechitaylo TY, Plumeier I, Pieper DH, Yakimov MM, Savchenko A, Golyshin PN, Ferrer M (2013) Single residues dictate the co-evolution of dual esterases: MCP hydrolases from the α/β hydrolase family. Biochem J 454:157–166. https://doi.org/10.1042/BJ20130552

Alcaide M, Stogios PJ, Lafraya Á, Tchigvintsev A, Flick R, Bargiela R, Chernikova TN, Reva ON, Hai T, Leggewie CC, Katzke N, La Cono V, Matesanz R, Jebbar M, Jaeger K-E, Yakimov MM, Yakunin AF, Golyshin PN, Golyshina OV, Savchenko A, Ferrer M, MAMBA Consortium (2015) Pressure adaptation is linked to thermal adaptation in salt-saturated marine habitats. Environ Microbiol 17:332–345. https://doi.org/10.1111/1462-2920.12660

Angkawidjaja C, Koga Y, Takano K, Kanaya S (2012) Structure and stability of a thermostable carboxylesterase from the thermoacidophilic archaeon *Sulfolobus tokodaii*. FEBS J 279:3071–3084. https://doi.org/10.1111/j.1742-4658.2012.08687.x

Antranikian G, Egorova K (2007) Extremophiles, a unique resource of biocatalysts for industrial biotechnology. In: Physiology and biochemistry of extremophiles. American Society of Microbiology, Washington DC, pp 361–406

Arpigny JL, Jaeger K-E (1999) Bacterial lipolytic enzymes: classification and properties. Biochem J 343:177. https://doi.org/10.1042/0264-6021:3430177

Auernik KS, Cooper CR, Kelly RM (2008) Life in hot acid: pathway analyses in extremely thermoacidophilic archaea. Curr Opin Biotechnol 19:445–453. https://doi.org/10.1016/j.copbio.2008.08.001

Benaiges MD, Alarcón M, Fuciños P, Ferrer P, Rua M, Valero F (2010) Recombinant *Candida rugosa* lipase 2 from *Pichia pastoris*: immobilization and use as biocatalyst in a stereoselective reaction. Biotechnol Prog 26:1252–1258. https://doi.org/10.1002/btpr.444

Bendtsen JD, Nielsen H, Von Heijne G, Brunak S (2004) Improved prediction of signal peptides: SignalP 3.0. J Mol Biol 340:783–795. https://doi.org/10.1016/j.jmb.2004.05.028

Berlemont R, Spee O, Delsaute M, Lara Y, Schuldes J, Simon C, Power P, Daniel R, Galleni M (2013) Novel organic solvent–tolerant esterase isolated by metagenomics: insights into the lipase/esterase classification. Rev Argent Microbiol 45:3–12

Biver S, Vandenbol M (2013) Characterization of three new carboxylic ester hydrolases isolated by functional screening of a forest soil metagenomic library. J Ind Microbiol Biotechnol 40:191–200. https://doi.org/10.1007/s10295-012-1217-7

Bora L (2014) Purification and characterization of highly alkaline lipase from *Bacillus licheniformis* MTCC 2465: and study of its detergent compatibility and applicability. J Surfactant Deterg 17:889–898. https://doi.org/10.1007/s11743-013-1517-6

Bornscheuer UT (2002) Microbial carboxyl esterases: classification, properties and application in biocatalysis. FEMS Microbiol Rev 26:73–81. https://doi.org/10.1111/j.1574-6976.2002.tb00599.x

Box GEP, Hunter JS, Hunter WG (2005) Statistics for experimenters: design, innovation and discovery, 2nd edn. Wiley-Interscience, Hoboken

Boyineni J, Kim J, Kang BS, Lee C, Jang SH (2014) Enhanced catalytic site thermal stability of cold-adapted esterase EstK by a W208Y mutation. Biochim Biophys Acta, Proteins Proteomics 1844:1076–1082. https://doi.org/10.1016/j.bbapap.2014.03.009

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72:248–254

Brault G, Shareck F, Hurtubise Y, Lépine F, Doucet N (2012) Isolation and characterization of EstC, a new cold-active esterase from *Streptomyces coelicolor* A3(2). PLoS One 7:e32041. https://doi.org/10.1371/journal.pone.0032041

Bunterngsook B, Kanokratana P, Thongaram T, Tanapongpipat S, Uengwetwanit T, Rachdawong S, Vichitsoonthonkul T, Eurwilaichitr L (2010) Identification and characterization of lipolytic enzymes from a peat-swamp forest soil metagenome. Biosci Biotechnol Biochem 74:1848–1854. https://doi.org/10.1271/bbb.100249

Chu X, He H, Guo C, Sun B (2008) Identification of two novel esterases from a marine metagenomic library derived from South China Sea. Appl Microbiol Biotechnol 80:615–625. https://doi.org/10.1007/s00253-008-1566-3

Cook GM, Rainey FA, Patel BKC, Morgan HW (1996) Characterization of a new obligately anaerobic thermophile, *Thermoanaerobacter wiegelii* sp. nov. Int J Syst Bacteriol 46:123–127. https://doi.org/10.1099/00207713-46-1-123

D'Auria S, Herman P, Lakowicz J, Bertoli E, Tanfani F, Rossi M, Manco G (2000) The thermophilic esterase from *Archaeoglobus fulgidus*: structure and conformational dynamics at high temperature. Proteins Struct Funct Genet 38:351–360

Dandavate V, Jinjala J, Keharia H, Madamwar D (2009) Production, partial purification and characterization of organic solvent–tolerant lipase from *Burkholderia multivorans* V2 and its application for ester synthesis. Bioresour Technol 100:3374–3381. https://doi.org/10.1016/j.biortech.2009.02.011

Daniel R (2005) The metagenomics of soil. Nat Rev Microbiol 3:470–478. https://doi.org/10.1038/nrmicro1160

Delgado-García M, Valdivia-Urdiales B, Aguilar-González CN, Contreras-Esquivel JC, Rodríguez-Herrera R (2012) Halophilic hydrolases as a new tool for the biotechnological industries. J Sci Food Agric 92:2575–2580. https://doi.org/10.1002/jsfa.5860

Dougherty MJ, D'haeseleer P, Hazen TC, Simmons BA, Adams PD, Hadi MZ (2012) Glycoside hydrolases from a targeted compost metagenome, activity screening and functional characterization. BMC Biotechnol 12:38. https://doi.org/10.1186/1472-6750-12-38

Doukyu N, Ogino H (2010) Organic solvent-tolerant enzymes. Biochem Eng J 48:270–282. https://doi.org/10.1016/J.BEJ.2009.09.009

Dukunde A, Schneider D, Lu M, Brady S, Daniel R (2017) A novel, versatile family IV carboxylesterase exhibits high stability and activity in a broad pH spectrum. Biotechnol Lett 39:577–587. https://doi.org/10.1007/s10529-016-2282-1

Ebrahimi M, Lakizadeh A, Agha-Golzadeh P, Ebrahimie E, Ebrahimi M (2011) Prediction of thermostability from amino acid attributes by combination of clustering with attribute weighting: a new vista in engineering enzymes. PLoS One 6:e23146. https://doi.org/10.1371/journal.pone.0023146

Elend C, Schmeisser C, Leggewie C, Babiak P, Steele HL, Reymond J, Jaeger K, Streit R, Carballeira JD, Streit WR (2006) Isolation and biochemical characterization of two novel metagenome-derived esterases. Appl Environ Microbiol 72:3637–3645. https://doi.org/10.1128/AEM.72.5.3637

Faiz O, Colak A, Saglam N, Canakçi S, Beldüz AO (2007) Determination and characterization of thermostable esterolytic activity from a novel thermophilic bacterium *Anoxybacillus gonensis* A4. J Biochem Mol Biol 40:588–594. https://doi.org/10.5483/BMBRep.2007.40.4.588

Faulds CB, Pérez-Boada M, Martínez ÁT (2011) Influence of organic cosolvents on the activity and substrate specificity of feruloyl esterases. Bioresour Technol 102:4962–4967. https://doi.org/10.1016/j.biortech.2011.01.088

Ferrer M, Bargiela R, Martínez-Martínez M, Mir J, Koch R, Golyshina OV, Golyshin PN (2015) Biodiversity for biocatalysis: a review of the α/β-hydrolase fold superfamily of esterases-lipases discovered in metagenomes. Biocatal Biotransformation 33:235–249. https://doi.org/10.3109/10242422.2016.1151416

Gall MG, Nobili A, Pavlidis IV, Bornscheuer UT (2014) Improved thermostability of a *Bacillus subtilis* esterase by domain exchange. Appl Microbiol Biotechnol 98:1719–1726. https://doi.org/10.1007/s00253-013-5053-0

Gao W, Wu K, Chen L, Fan H, Zhao Z, Gao B, Wang H, Wei D (2016) A novel esterase from a marine mud metagenomic library for biocatalytic synthesis of short-chain flavor esters. Microb Cell Factories 15:41. https://doi.org/10.1186/s12934-016-0435-5

Georis J, Esteves FDL, Lamotte-Brasseur J, Bougnet V, Giannotta F, Frère J-M, Devreese B, Granier B (2000) An additional aromatic interaction improves the thermostability and thermophilicity of a mesophilic family 11 xylanase: structural basis and molecular study. Protein Sci 9:466–475. https://doi.org/10.1110/ps.9.3.466

Glogauer A, Martini VP, Faoro H, Couto GH, Müller-Santos M, Monteiro RA, Mitchell DA, de Souza EM, Pedrosa FO, Krieger N (2011) Identification and characterization of a new true lipase isolated through metagenomic approach. Microb Cell Factories 10:54. https://doi.org/10.1186/1475-2859-10-54

González-González R, Fuciños P, Rúa ML (2017) An overview on extremophilic esterases. In: Extremophilic enzymatic processing of lignocellulosic feedstocks to bioenergy. Springer International Publishing, Cham, pp 181–204

Hardeman F, Sjoling S (2007) Metagenomic approach for the isolation of a novel low-temperature-active lipase from uncultured bacteria of marine sediment. FEMS Microbiol Ecol 59:524–534. https://doi.org/10.1111/j.1574-6941.2006.00206.x

Hasan F, Shah AA, Hameed A (2006) Industrial applications of microbial lipases. Enzym Microb Technol 39:235–251. https://doi.org/10.1016/J.ENZMICTEC.2005.10.016

Hausmann S, Jaeger K-E (2010) Lipolytic enzymes from bacteria. In: Handbook of Hydrocarbon and Lipid Microbiology. Springer Berlin Heidelberg, Berlin, pp 1099–1126

Hita E, Robles A, Camacho B, González PA, Esteban L, Jiménez MJ, Muñío MM, Molina E (2009) Production of structured triacylglycerols by acidolysis catalyzed by lipases immobilized in a packed bed reactor. Biochem Eng J 46:257–264. https://doi.org/10.1016/J.BEJ.2009.05.015

Hotta Y, Ezaki S, Atomi H, Imanaka T (2002) Extremely stable and versatile carboxylesterase from a hyperthermophilic archaeon. Appl Environ Microbiol 68:3925–3931. https://doi.org/10.1128/AEM.68.8.3925-3931.2002

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

Hun CJ, Rahman RNZA, Salleh AB, Basri M (2003) A newly isolated organic solvent–tolerant *Bacillus sphaericus* 205y producing organic solvent-stable lipase. Biochem Eng J 15:147–151. https://doi.org/10.1016/S1369-703X(02)00185-7

Ishikawa J, Hotta K (1999) FramePlot: a new implementation of the frame analysis for predicting protein-coding regions in bacterial DNA with a high G + C content. FEMS Microbiol Lett 174:251–263

Jaeger K-E, Dijkstra BW, Reetz MT (1999) Bacterial biocatalysts: molecular biology, three-dimensional structures, and biotechnological applications of lipases. Annu Rev Microbiol 53:315–351. https://doi.org/10.1146/annurev.micro.53.1.315

Jayanath G, Mohandas SP, Kachiprath B, Solomon S, Sajeevan TP, Bright Singh IS, Philip R (2018) A novel solvent–tolerant esterase of GDSGG motif subfamily from solar saltern through metagenomic approach: recombinant expression and characterization. Int J Biol Macromol 119:393–401. https://doi.org/10.1016/j.ijbiomac.2018.06.057

Jeon JH, Lee HS, Kim JT, Kim SJ, Choi SH, Kang SG, Lee JH (2012) Identification of a new subfamily of salt-tolerant esterases from a metagenomic library of tidal flat sediment.

Appl Microbiol Biotechnol 93:623–631. https://doi.org/10.1007/s00253-011-3433-x

Jin P, Pei X, Du P, Yin X, Xiong X, Wu H, Zhou X, Wang Q (2012) Overexpression and characterization of a new organic solvent-tolerant esterase derived from soil metagenomic DNA. Bioresour Technol 116:234–240. https://doi.org/10.1016/j.biortech.2011.10.087

Jochens H, Aerts D, Bornscheuer UT (2010) Thermostabilization of an esterase by alignment-guided focussed directed evolution. Protein Eng Des Sel 23:903–909. https://doi.org/10.1093/protein/gzq071

Kamal Z, Yedavalli P, Deshmukh MV, Rao NM (2013) Lipase in aqueous-polar organic solvents: activity, structure, and stability. Protein Sci 22:904–915. https://doi.org/10.1002/pro.2271

Kamimura ES, Medieta O, Rodrigues MI, Maugeri F (2001) Studies on lipase-affinity adsorption using response surface analysis. Biotechnol Appl Biochem 33:153–159

Kang CH, Oh KH, Lee MH, Oh TK, Kim BH, Yoon JH (2011) A novel family VII esterase with industrial potential from compost metagenomic library. Microb Cell Factories 10:41. https://doi.org/10.1186/1475-2859-10-41

Kang LJ, Meng ZT, Hu C, Zhang Y, Guo HL, Li Q, Li M (2017) Screening, purification, and characterization of a novel organic solvent-tolerant esterase, Lip2, from *Monascus purpureus* strain M7. Extremophiles 21:345–355. https://doi.org/10.1007/s00792-016-0907-x

Kawata T, Ogino H (2009) Enhancement of the organic solvent-stability of the LST-03 lipase by directed evolution. Biotechnol Prog 25:1605–1611. https://doi.org/10.1002/btpr.264

Kim TD (2017) Bacterial hormone-sensitive lipases (bHSLs): Emerging enzymes for biotechnological applications. J Microbiol Biotechnol 27:1907–1915. https://doi.org/10.4014/jmb.1708.08004

Kim YH, Kwon EJ, Kim SK, Jeong YS, Kim J, Yun HD, Kim H (2010) Molecular cloning and characterization of a novel family VIII alkaline esterase from a compost metagenomic library. Biochem Biophys Res Commun 393:45–49. https://doi.org/10.1016/j.bbrc.2010.01.070

Klibanov AM (2001) Improving enzymes by using them in organic solvents. Nature 409:241–246. https://doi.org/10.1038/35051719

Kovacic F, Mandrysch A, Poojari C, Strodel B, Jaeger K-E (2016) Structural features determining thermal adaptation of esterases. Protein Eng Des Sel 29:65–76. https://doi.org/10.1093/protein/gzv061

Kumagai PS, Gutierrez RF, Lopes JLS, Martins JM, Jameson DM, Castro AM, Martins LF, DeMarco R, Bossolan NRS, Wallace BA, Araujo APU (2018) Characterization of esterase activity from an *Acetomicrobium hydrogeniformans* enzyme with high structural stability in extreme conditions. Extremophiles 22:781–793. https://doi.org/10.1007/s00792-018-1038-3

Kumar A, Dhar K, Singh Kanwar S, Arora PK (2016) Lipase catalysis in organic solvents: advantages and applications. Biol Proced Online 18:2. https://doi.org/10.1186/s12575-016-0033-2

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680–685. https://doi.org/10.1038/227680a0

Lämmle K, Zipper H, Breuer M, Hauer B, Buta C, Brunner H, Rupp S (2007) Identification of novel enzymes with different hydrolytic activities by metagenome expression cloning. J Biotechnol 127:575–592. https://doi.org/10.1016/j.jbiotec.2006.07.036

Lazić ŽR (2004) Design of experiments in chemical engineering. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

Léonard-Nevers V, Marton Z, Lamare S, Hult K, Graber M (2009) Understanding water effect on *Candida antarctica* lipase B activity and enantioselectivity towards secondary alcohols. J Mol Catal B Enzym 59:90–95. https://doi.org/10.1016/J.MOLCATB.2009.01.008

Li X, Yu H (2013) Purification and characterization of an extracellular esterase with organic solvent tolerance from a halotolerant isolate, *Salimicrobium* sp. LY19. BMC Biotechnol 13:108

Li P, Chen X, Ji P, Li C, Wang P, Zhang Y, Xie B, Qin Q, Su H, Zhou B, Zhang Y, Zhang X (2015) Interdomain hydrophobic interactions modulate the thermostability of microbial esterases from the hormone-sensitive lipase family. J Biol Chem 290:11188–11,198. https://doi.org/10.1074/jbc.M115.646182

Lima VMG, Krieger N, Mitchell DA, Baratti JC, de FI, Fontana JD (2004) Evaluation of the potential for use in biocatalysis of a lipase from a wild strain of *Bacillus megaterium*. J Mol Catal B Enzym 31:53–61. https://doi.org/10.1016/J.MOLCATB.2004.07.005

Lineweaver H, Burk D (1934) The determination of enzyme dissociation constants. J Am Chem Soc 56:658–666. https://doi.org/10.1021/ja01318a036

Lopez-Lopez O, Cerdan M, Siso M (2014) New extremophilic lipases and esterases from metagenomics. Curr Protein Pept Sci 15:445–455. https://doi.org/10.2174/1389203715666140228153801

Lotti M, Alberghina L (2007) Lipases: molecular structure and function. In: Industrial Enzymes. Springer Netherlands, Dordrecht, pp 263–281

Mandelli F, Goncalves TA, Gandin CA, Oliveira ACP, Oliveira Neto M, Squina FM (2016) Characterization and low-resolution structure of an extremely thermostable esterase of potential biotechnological interest from *Pyrococcus furiosus*. Mol Biotechnol 58:757–766. https://doi.org/10.1007/s12033-016-9975-5

Martínez-Martínez M, Coscolín C, Santiago G, Chow J, Stogios PJ, Bargiela R, Gertler C, Navarro-Fernández J, Bollinger A, Thies S, Méndez-García C, Popovic A, Brown G, Chernikova TN, García-Moyano A, Bjerga GEK, Pérez-García P, Hai T, Del Pozo M V., Stokke R, Steen IH, Cui H, Xu X, Nocek BP, Alcaide M, Distaso M, Mesa V, Peláez AI, Sánchez J, Buchholz PCF, Pleiss J, Fernández-Guerra A, Glöckner FO, Golyshina O V., Yakimov MM, Savchenko A, Jaeger K-E, Yakunin AF, Streit WR, Golyshin PN, Guallar V, Ferrer M, The INMARE Consortium TI (2018) Determinants and prediction of esterase substrate promiscuity patterns. ACS Chem Biol 13:225–234. doi: https://doi.org/10.1021/acschembio.7b00996

Metin K, Burcu Bakir Ateslier Z, Basbulbul G, Halil Biyik H (2006) Characterization of esterase activity in *Geobacillus* sp. HBB-4. J Basic Microbiol 46:400–409. https://doi.org/10.1002/jobm.200510121

Mohamed YM, Ghazy MA, Sayed A, Ouf A, El-Dorry H, Siam R (2013) Isolation and characterization of a heavy metal-resistant, thermophilic esterase from a Red Sea Brine Pool. Sci Rep 3:3358. https://doi.org/10.1038/srep03358

Monsef Shokri M, Ahmadian S, Akbari N, Khajeh K (2014) Hydrophobic substitution of surface residues affects lipase stability in organic solvents. Mol Biotechnol 56:360–368. https://doi.org/10.1007/s12033-013-9716-y

Nacke H, Will C, Herzog S, Nowka B, Engelhaupt M, Daniel R (2011) Identification of novel lipolytic genes and gene families by screening of metagenomic libraries derived from soil samples of the German Biodiversity Exploratories. FEMS Microbiol Ecol 78:188–201. https://doi.org/10.1111/j.1574-6941.2011.01088.x

Nasaruddin RR, Alam MZ, Jami MS (2014) Evaluation of solvent system for the enzymatic synthesis of ethanol-based biodiesel from sludge palm oil (SPO). Bioresour Technol 154:155–161. https://doi.org/10.1016/j.biortech.2013.11.095

Nerurkar M, Joshi M, Pariti S, Adivarekar R (2013) Application of lipase from marine bacteria *Bacillus sonorensis* as an additive in detergent formulation. J Surfactant Deterg 16:435–443. https://doi.org/10.1007/s11743-012-1434-0

Ngo TD, Ryu BH, Ju H, Jang E, Park K, Kim KK, Kim TD (2013) Structural and functional analyses of a bacterial homologue of hormone-sensitive lipase from a metagenomic library. Acta

Crystallogr Sect D Biol Crystallogr 69:1726–1737. https://doi.org/10.1107/S0907444913013425

Ngo TD, Ryu BH, Ju H, Jang EJ, Kim KK, Kim TD (2014) Crystallographic analysis and biochemical applications of a novel penicillin-binding protein/β-lactamase homologue from a metagenomic library. Acta Crystallogr Sect D Biol Crystallogr 70: 2455–2466. https://doi.org/10.1107/S1399004714015272

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

Ó'Fágáin C (2003) Enzyme stabilization—recent experimental progress. Enzym Microb Technol 33:137–149. https://doi.org/10.1016/S0141-0229(03)00160-1

Ogino H, Ishikawa H (2001) Enzymes which are stable in the presence of organic solvents. J Biosci Bioeng 91:109–116

Ogino H, Mimitsuka T, Muto T, Matsumura M, Yasuda M, Ishimi K, Ishikawa H (2004) Cloning, expression, and characterization of a lipolytic enzyme gene (lip8) from *Pseudomonas aeruginosa* LST-03. J Mol Microbiol Biotechnol 7:212–223. https://doi.org/10.1159/000079830

Ohlhoff CW, Kirby BM, Van Zyl L, Mutepfa DLR, Casanueva A, Huddy RJ, Bauer R, Cowan DA, Tuffin M (2015) An unusual feruloyl esterase belonging to family VIII esterases and displaying a broad substrate range. J Mol Catal B Enzym 118:79–88. https://doi.org/10.1016/J.MOLCATB.2015.04.010

Ollis DL, Cheah E, Cygler M, Dijkstra B, Frolow F, Franken SM, Harel M, Remington SJ, Silman I, Schrag J, Sussman JL, Verschueren KHG, Goldman A (1992) The α/β hydrolase fold. Protein Eng Des Sel 5:197–211. https://doi.org/10.1093/protein/5.3.197

Panda T, Gowrishankar BS (2005) Production and applications of esterases. Appl Microbiol Biotechnol 67:160–169. https://doi.org/10.1007/s00253-004-1840-y

Park HJ, Joo JC, Park K, Yoo YJ (2012) Stabilization of *Candida antarctica* lipase B in hydrophilic organic solvent by rational design of hydrogen bond. Biotechnol Bioprocess Eng 17:722–728. https://doi.org/10.1007/s12257-012-0092-4

Park SH, Kim S, Park S, Kim HK (2018) Characterization of organic solvent-tolerant lipolytic enzyme from *Marinobacter lipolyticus* isolated from the Antarctic Ocean. Appl Biochem Biotechnol. https://doi.org/10.1007/s12010-018-2865-5

Peng Q, Zhang X, Shang M, Wang X, Wang G, Li B, Guan G, Li Y, Wang Y (2011) A novel esterase gene cloned from a metagenomic library from neritic sediments of the South China Sea. Microb Cell Factories 10:95. https://doi.org/10.1186/1475-2859-10-95

Pérez D, Kovačic F, Wilhelm S, Jaeger K-E, García MT, Ventosa A, Mellado EN (2012) Identification of amino acids involved in the hydrolytic activity of lipase LipBL from *Marinobacter lipolyticus*. Microbiology 158:2192–2203. doi: https://doi.org/10.1099/mic.0.058792-0

Petersen EI, Valinger G, Sölkner B, Stubenrauch G, Schwab H (2001) A novel esterase from *Burkholderia gladioli* which shows high deacetylation activity on cephalosporins is related to beta-lactamases and DD-peptidases. J Biotechnol 89:11–25

Pezzullo M, Del Vecchio P, Mandrich L, Nucci R, Rossi M, Manco G (2013) Comprehensive analysis of surface charged residues involved in thermal stability in *Alicyclobacillus acidocaldarius* esterase 2. Protein Eng Des Sel 26:47–58. https://doi.org/10.1093/protein/gzs066

Popovic A, Hai T, Tchigvintsev A, Hajighasemi M, Nocek B, Khusnutdinova AN, Brown G, Glinos J, Flick R, Skarina T, Chernikova TN, Yim V, Brüls T, Le PD, Yakimov MM, Joachimiak A, Ferrer M, Golyshina OV, Savchenko A, Golyshin PN, Yakunin AF (2017) Activity screening of environmental metagenomic libraries reveals novel carboxylesterase families. Sci Rep 7:44103. https://doi.org/10.1038/srep44103

Rahman RNZRA, Baharum SN, Basri M, Salleh AB (2005) High-yield purification of an organic solvent-tolerant lipase from *Pseudomonas* sp. strain S5. Anal Biochem 341:267–274. https://doi.org/10.1016/j.ab.2005.03.006

Raja A, Prabakarana P (2011) Actinomycetes and drug-an overview. Am J Drug Discov Dev 1:75–84. https://doi.org/10.3923/ajdd.2011.75.84

Rashamuse K, Magomani V, Ronneburg T, Brady D (2009) A novel family VIII carboxylesterase derived from a leachate metagenome library exhibits promiscuous β-lactamase activity on nitrocefin. Appl Microbiol Biotechnol 83:491–500. https://doi.org/10.1007/s00253-009-1895-x

Riedel K, Sutherland R, Donato JJ, Liebl W, Leis B, Angelov A, Mientus M, Li H, Pham VTT, Lauinger B, Bongen P, Pietruszka J, Gonçalves LG, Santos H (2015) Identification of novel esterase-active enzymes from hot environments by use of the host bacterium *Thermus thermophilus*. Front Microbiol 6:27. https://doi.org/10.3389/fmicb.2015.00275

Robert X, Gouet P (2014) Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res 42:320–324. https://doi.org/10.1093/nar/gku316

Romdhane IB-B, Fendri A, Gargouri Y, Gargouri A, Belghith H (2010) A novel thermoactive and alkaline lipase from *Talaromyces thermophilus* fungus for use in laundry detergents. Biochem Eng J 53:112–120. https://doi.org/10.1016/J.BEJ.2010.10.002

Ryckeboer J, Mergaert J, Vaes K, Klammer S, Clercq D, Coosemans J, Insam H, Swings J (2003) A survey of bacteria and fungi occurring during composting and self-heating processes. Ann Microbiol 53: 349–410

Sadeghi M, Naderi-Manesh H, Zarrabi M, Ranjbar B (2006) Effective factors in thermostability of thermophilic proteins. Biophys Chem 119:256–270. https://doi.org/10.1016/j.bpc.2005.09.018

Sakai Y, Ishikawa J, Fukasaka S, Yurimoto H, Mitsui R, Yanase H, Kato N (1999) A new carboxylesterase from *Brevibacterium linens* IFO 12171 responsible for the conversion of 1,4-butanediol diacrylate to 4-hydroxybutyl acrylate: purification, characterization, gene cloning, and gene expression in *Escherichia coli*. Biosci Biotechnol Biochem 63:688–697. https://doi.org/10.1271/bbb.63.688

Salihu A, Alam MZ (2015) Solvent tolerant lipases: a review. Process Biochem 50:86–96. https://doi.org/10.1016/j.procbio.2014.10.019

Sana B, Ghosh D, Saha M, Mukherjee J (2007) Purification and characterization of an extremely dimethylsulfoxide tolerant esterase from a salt-tolerant Bacillus species isolated from the marine environment of the Sundarbans. Process Biochem 42:1571–1578. https://doi.org/10.1016/j.procbio.2007.05.026

Sarmah N, Revathi D, Sheelu G, Yamuna Rani K, Sridhar S, Mehtab V, Sumana C (2018) Recent advances on sources and industrial applications of lipases. Biotechnol Prog 34:5–28. https://doi.org/10.1002/btpr.2581

Sayer C, Isupov MN, Bonch-Osmolovskaya E, Littlechild JA (2015a) Structural studies of a thermophilic esterase from a new Planctomycetes species, *Thermogutta terrifontis*. FEBS J 282: 2846–2857. https://doi.org/10.1111/febs.13326

Sayer C, Szabo Z, Isupov MN, Ingham C, Littlechild JA (2015b) The structure of a novel thermophilic esterase from the Planctomycetes species, *Thermogutta terrifontis* reveals an open active site due to a minimal "Cap" domain. Front Microbiol 6:1294. https://doi.org/10.3389/fmicb.2015.01294

Schütte M, Fetzner S (2007) EstA from *Arthrobacter nitroguajacolicus* Ru61a, a thermo- and solvent-tolerant carboxylesterase related to class C b-Lactamases. Curr Microbiol 54:230–236. https://doi.org/10.1007/s00284-006-0438-2

Secundo F, Carrea G (2002) Lipase activity and conformation in neat organic solvents. J Mol Catal B Enzym 19:93–102. https://doi.org/10.1016/S1381-1177(02)00155-8

Selvin J, Kennedy J, Lejon DPH, Kiran GS, Dobson ADW (2012) Isolation identification and biochemical characterization of a novel halo-tolerant lipase from the metagenome of the marine sponge *Haliclona simulans*. Microb Cell Factories 11:72. https://doi.org/10.1186/1475-2859-11-72

Seo S, Lee YS, Yoon SH, Kim SJ, Cho JY, Hahn BS, Koo BS, Lee CM (2014) Characterization of a novel cold-active esterase isolated from swamp sediment metagenome. World J Microbiol Biotechnol 30: 879–886. https://doi.org/10.1007/s11274-013-1496-9

Shao H, Xu L, Yan Y (2013) Isolation and characterization of a thermostable esterase from a metagenomic library. J Ind Microbiol Biotechnol 40:1211–1222. https://doi.org/10.1007/s10295-013-1317-z

Sharma R, Soni S, Vohra R, Gupta L, Gupta J (2002) Purification and characterisation of a thermostable alkaline lipase from a new thermophilic *Bacillus* sp. RSJ-1. Process Biochem 37:1075–1084. https://doi.org/10.1016/S0032-9592(01)00316-8

Shieh C-J, Liao H-F, Lee C-C (2003) Optimization of lipase-catalyzed biodiesel by response surface methodology. Bioresour Technol 88: 103–116

Simon C, Daniel R (2011) Metagenomic analyses: past and future trends. Appl Environ Microbiol 77:1153–1161. https://doi.org/10.1128/AEM.02345-10

Song JK, Rhee JS (2001) Enhancement of stability and activity of phospholipase A1 in organic solvents by directed evolution. Biochim Biophys Acta Protein Struct Mol Enzymol 1547:370–378. https://doi.org/10.1016/S0167-4838(01)00204-7

Sood S, Sharma A, Sharma N, Kanwar SS (2016) Carboxylesterases: sources, characterization and broader applications. Insights Enzym Res 1:1. https://doi.org/10.21767/2573-4466.100002

Srimhan P, Kongnum K, Taweerodjanakarn S, Hongpattarakere T (2011) Selection of lipase producing yeasts for methanol-tolerant biocatalyst as whole cell application for palm-oil transesterification. Enzym Microb Technol 48:293–298. https://doi.org/10.1016/j.enzmictec.2010.12.004

Takeda Y, Aono R, Doukyu N (2006) Purification, characterization, and molecular cloning of organic-solvent-tolerant cholesterol esterase from cyclohexane-tolerant *Burkholderia cepacia* strain ST-200. Extremophiles 10:269–277. https://doi.org/10.1007/s00792-005-0494-8

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

Tirawongsaroj P, Sriprang R, Harnpicharnchai P, Thongaram T, Champreda V, Tanapongpipat S, Pootanakit K, Eurwilaichitr L (2008) Novel thermophilic and thermostable lipolytic enzymes from a Thailand hot spring metagenomic library. J Biotechnol 133:42–49. https://doi.org/10.1016/J.JBIOTEC.2007.08.046

Torres S, Castro GR (2004) Non-aqueous biocatalysis in homogeneous solvent systems. Crit Rev Biotechnol 42:271–277

Torres S, Martínez MA, Pandey A, Castro GR (2009) An organic-solvent-tolerant esterase from thermophilic *Bacillus licheniformis* S-86. Bioresour Technol 100:896–902. https://doi.org/10.1016/j.biortech.2008.07.009

Wang X, Qin X, Li D, Yang B, Wang Y (2017) One-step synthesis of high-yield biodiesel from waste cooking oils by a novel and highly methanol-tolerant immobilized lipase. Bioresour Technol 235:18–24. https://doi.org/10.1016/J.BIORTECH.2017.03.086

Woo Lee H, Kyeong Jung W, Ho Kim Y, Han Ryu B, Doohun Kim T, Kim J, Kim H (2016) Characterization of a novel alkaline family VIII esterase with s-enantiomer preference from a compost metagenomic library. J Microbiol Biotechnol 26:315–325. https://doi.org/10.4014/jmb.1509.09081

Wu G, Zhang S, Zhang H, Zhang S, Liu Z (2013) A novel esterase from a psychrotrophic bacterium *Psychrobacter celer* 3Pb1 showed cold-adaptation and salt tolerance. J Mol Catal B Enzym 98:119–126. https://doi.org/10.1016/j.molcatb.2013.10.012

Xing MN, Zhang XZ, Huang H (2012) Application of metagenomic techniques in mining enzymes from microbial communities for biofuel synthesis. Biotechnol Adv 30:920–929. https://doi.org/10.1016/j.biotechadv.2012.01.021

Yang Y, Yu Y, Zhang Y, Liu C, Shi W, Li Q (2011) Lipase/esterase-catalyzed ring-opening polymerization: a green polyester synthesis technique. Process Biochem 46:1900–1908. https://doi.org/10.1016/j.procbio.2011.07.016

Yang X, Wu L, Xu Y, Ke C, Hu F, Xiao X, Huang J (2018) Identification and characterization of a novel alkalistable and salt-tolerant esterase from the deep-sea hydrothermal vent of the East Pacific Rise. Microbiology 7:e00601. https://doi.org/10.1002/mbo3.601

Ye J, McGinnis S, Madden TL (2006) BLAST: improvements for better sequence analysis. Nucleic Acids Res 34:W6–W9. https://doi.org/10.1093/nar/gkl164

Yu S, Zheng B, Zhao X, Feng Y (2010) Gene cloning and characterization of a novel thermophilic esterase from *Fervidobacterium nodosum* Rt17-B1. Acta Biochim Biophys Sin Shanghai 42:288–295. https://doi.org/10.1093/abbs/gmq020

Yu EY, Kwon MA, Lee M, Oh JY, Choi JE, Lee JY, Song BK, Hahm DH, Song JK (2011) Isolation and characterization of cold-active family VIII esterases from an arctic soil metagenome. Appl Microbiol Biotechnol 90:573–581. https://doi.org/10.1007/s00253-011-3132-7

Zarafeta D, Moschidi D, Ladoukakis E, Gavrilov S, Chrysina ED, Chatziioannou A, Kublanov I, Skretas G, Kolisis FN (2016) Metagenomic mining for thermostable esterolytic enzymes uncovers a new family of bacterial esterases. Sci Rep 6:38886. https://doi.org/10.1038/srep38886

Zhang Y (2008) I-TASSER server for protein 3D structure prediction. BMC Bioinformatics 23:40. https://doi.org/10.1186/1471-2105-9-40

Zhang S, Wu G, Liu Z, Shao Z, Liu Z (2014) Characterization of EstB, a novel cold-active and organic solvent-tolerant esterase from marine microorganism *Alcanivorax dieselolei* B-5(T). Extremophiles 18: 251–259. https://doi.org/10.1007/s00792-013-0612-y

Zhou J, Bruns MA, Tiedje JM (1996) DNA recovery from soils of diverse composition. Appl Environ Microbiol 62:316–322

Supplemental information for chapter III

Content

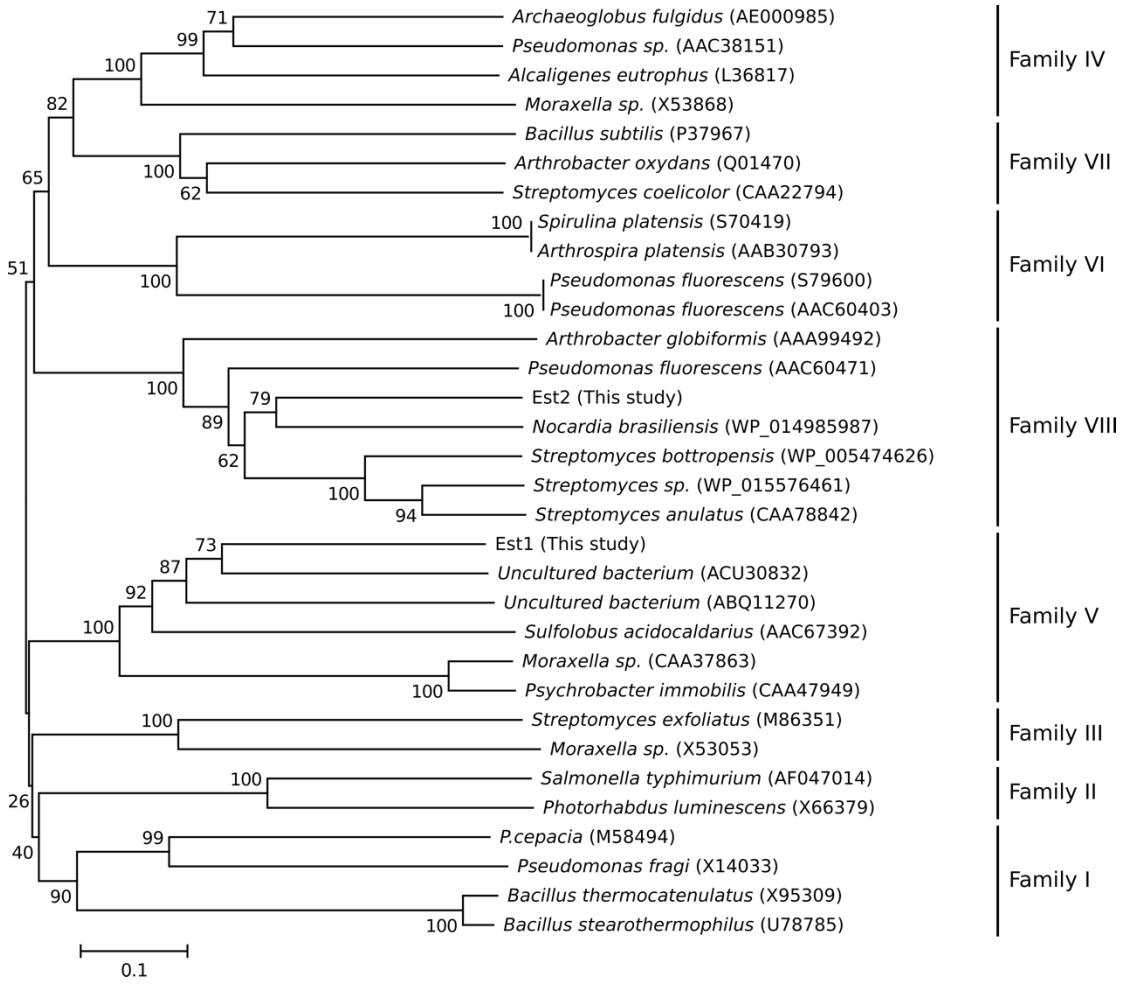
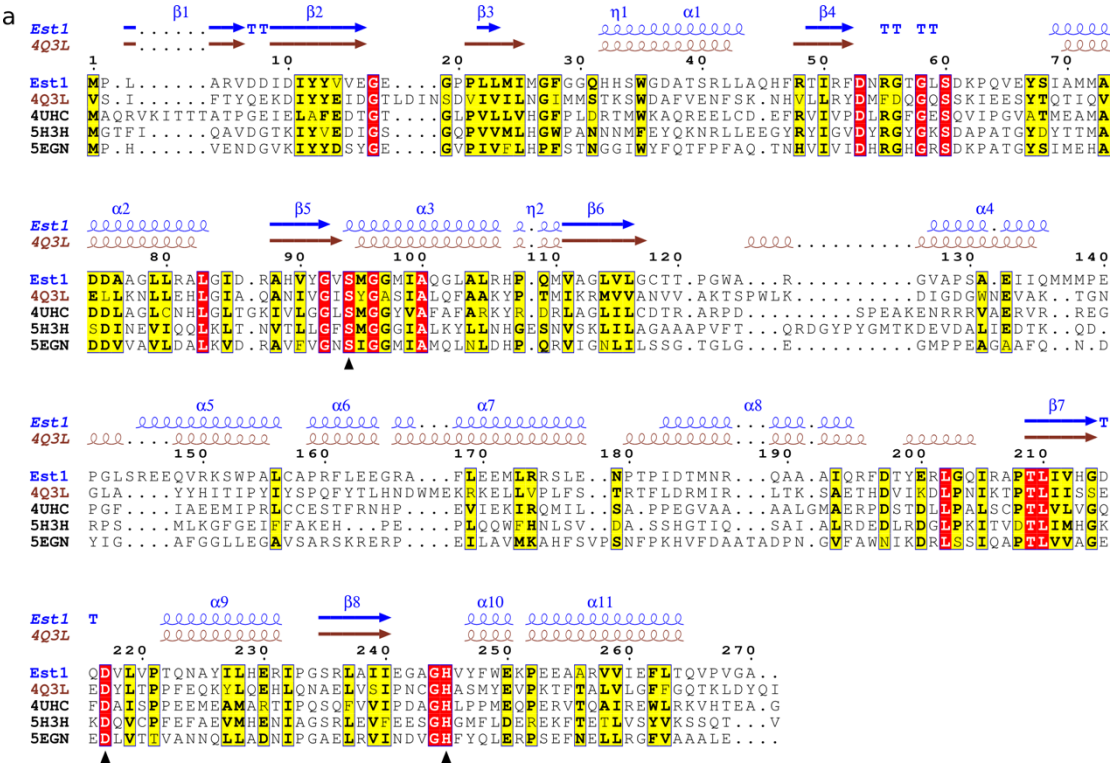


Figure S1 Unrooted phylogenetic tree constructed by the neighbor-joining method with Est1, Est2 and reference lipolytic enzymes from different families. With the exception of Est1 and Est2, other sequences were retrieved from GenBank. The full organism names and accession numbers are given for reference sequences. Only bootstrap values above 50% are shown. Scale represents the number of amino acid substations per site

a


Est1
4Q3L
β1 β2 β3 η1 α1 β4
1 10 20 30 40 50 60 70
Est1 MP.L.....ARVDDIDIYYVVEGE....GPPLLMIMGFGGQHHSWGDATSRLLAQHFRTIRFDNRGTGLSDKPQVEYSIAMMA
4Q3L VS.I.....FTYQEKDIYYEIDGTLDINSDVIVILNGIMMSTKSWDAFVENFSK.NHVLLRYDMFDQGQSSKIEESYTQTIQV
4UHC MAQRVKITTTATPGEIELAFEDTGT....GLPVLLVHGFPLDRTMWKAQREELCD.EFRVIVPDLRGFGESQVIPGVATMEAMA
5H3H MGTFI.....QAVDGTKIYVEDIGS....GQPVVMLHGWPANNNMFEYQKNRLLEEGYRYIGVDYRGYGKSDAPATGYDYTTMA
5EGN MP.H.....VENDGVKIYYDSYGE....GVPIVFLHPFSTNGGIWYFQTFPFAQ.TNHVIVIDHRGHGRSDKPATGYSIMEHA
α2 β5 α3 η2 β6 α4
80 90 100 110 120 130 140
Est1 DDAAGLLRALGID.RAHVYGVSMGGMIAQGLALRHP.QMVAGLVLGCTT.PGWA...R..........GVAPSA.EIIQMMMPE
4Q3L ELLKNLLEHLGIA.QANIVGISYGASIALQFAAKYP.TMIKRMVVANVV.AKTSPWLK..........DIGDGWNEVAK..TGN
4UHC DDLAGLCNHLGLTGKIVLGGLSMGGYVAFAFARKYR.DRLAGLILCDTR.ARPD.........SPEAKENRRRVAERVR..REG
5H3H SDINEVIQQLKLT.NVTLLGFSMGGGIALKYLLNHGESNVSKLILAGAAAPVFT...QRDGYPYGMTKDEVDALIEDTK..QD.
5EGN DDVVAVLDALKVD.RAVFVGNSIGGMIAMQLNLDHP.QRVIGNLILSSG.TGLG...E..........GMPPEAGAAFQ..N.D
α5 α6 α7 α8 β7
150 160 170 180 190 200 210
Est1 PGLSREEQVRKSWPALCAPRFLEEGRA...FLEEMLRRSLE..NPTPIDTMNR...QAA.AIQRFDTYERLGQIRAPTLIVHGD
4Q3L GLA....YYHITIPYIYSPQFYTLHNDWMEKRKELLVPLFS..TRTFLDRMIR...LTK.SAETHDVIKDLPNIKTPTLIISSE
4UHC PGF....IAEEMIPRLCCESTFRNHP....EVIEKIRQMIL..SA.PPEGVAA...AALGMAERPDSTDLLPALSCPTLVLVGQ
5H3H RPS....MLKGFGEIFFAKEH...PE....PLQQWFHNLSV..DA.SSHGTIQ...SAI.ALRDEDLRDGLPKITVDTLIMHGK
5EGN YIG....AFGGLLEGAVSARSKRERP....EILAVMKAHFSVPSNFPKHVFDAATADPN.GVFAWNIKDRLSSIQAPTLVVAGE
α9 β8 α10 α11
220 230 240 250 260 270
Est1 QDVLVPTQNAYILHERIPGSRLAIIEGAGHVYFWEKPEEAARVVIEFLTQVPVGA..
4Q3L EDYLTPPFEQKYLQEHLQNAELVSIPNCGHASMYEVPKTFTALVLGFFGQTKLDYQI
4UHC FDAISPPEEMEAMARTIPQSQFVVIPDAGHLPPMEQPERVTQAIREWLRKVHTEA.G
5H3H KDQVCPFEFAEVMHENIAGSRLEVFEESGHGMFLDERKFTETLVSYVKSSQT...V
5EGN EDLVTTVANNQLLADNIPGAELRVINDVGHFYQLERPSEFNELLRGFVAAALE....

b

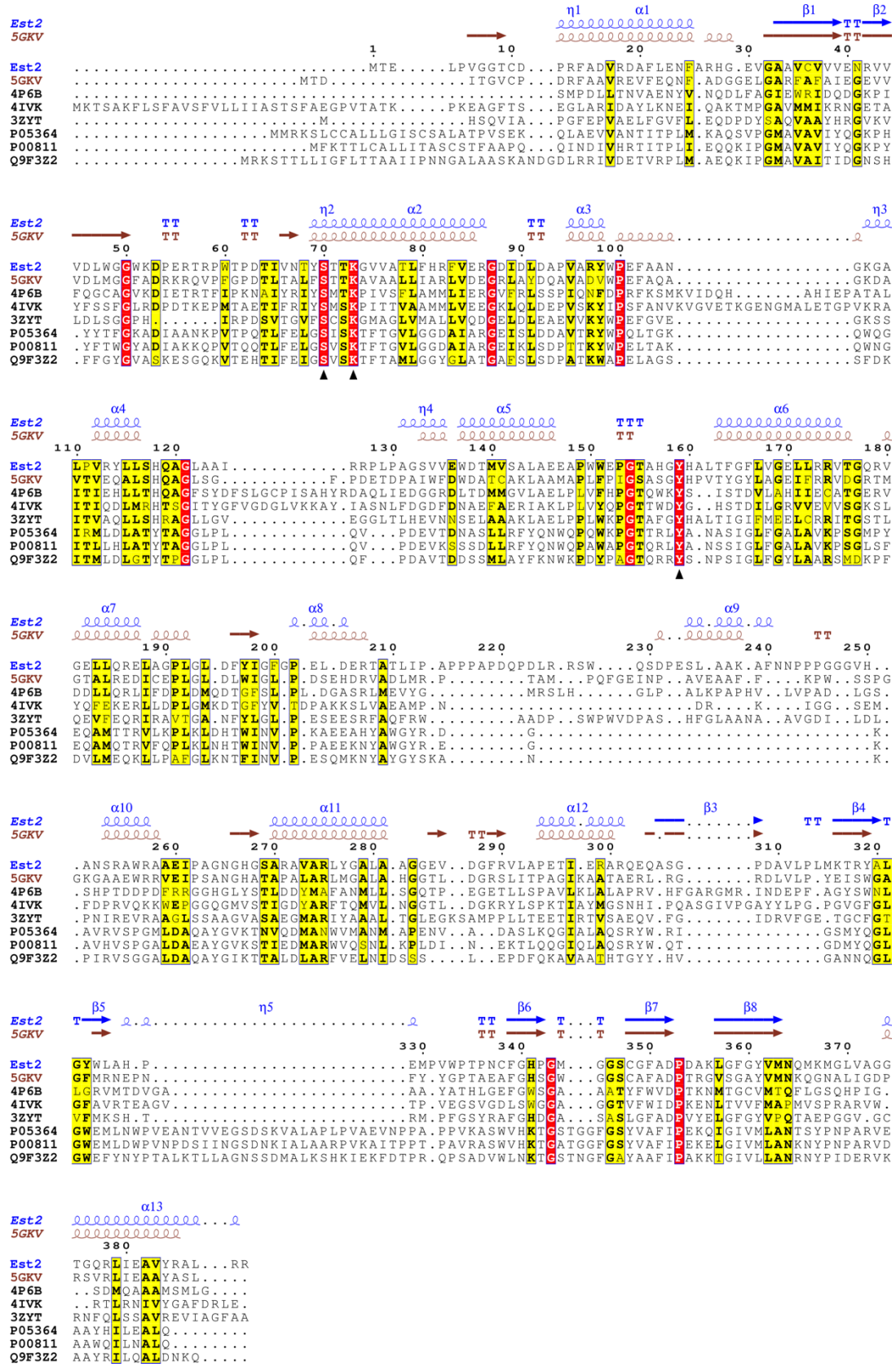


Figure S2 Multiple sequence alignments of Est1 (a) and Est2 (b) with other esterase sequences. Secondary structures of Est1 and Est2 were predicted by I-TASSER. Sequences were aligned using structural information derived from EPRESSO. Identical residues are shaded. Stars underneath residues indicate the catalytic triad. Secondary structures are presented as α-helices and β-strands on top of sequences. a, Alignment of Est1 and its structural analogs. Secondary structures of Est1 and MGS-M2 (PDB: 4Q3L) are presented. b, Alignment of Est2, its esterase structural analogs and class C beta-lactamases (GenBank: P05364, P00811 and Q9F3Z2). Secondary structures of Est2 and CcEstA (PDB: 5GKV) are presented

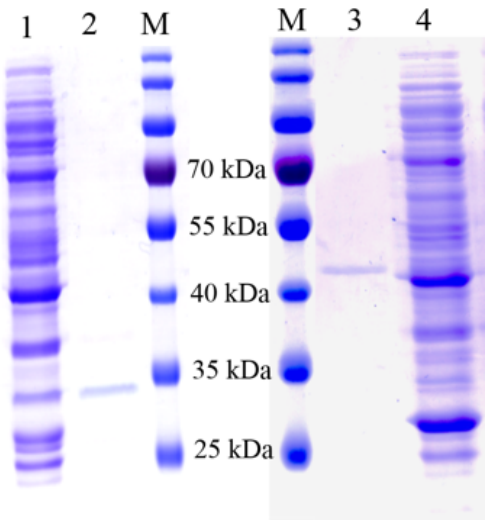


Figure S3 SDS-PAGE of purified Est1 and Est2. Lane M, standard molecular weight marker; Lane 1, soluble fraction of induced Est1 cell lysate; Lane 2, purified Est1; Lane 3, purified Est2; Lane 4, soluble fraction of induced Est2 cell lysate

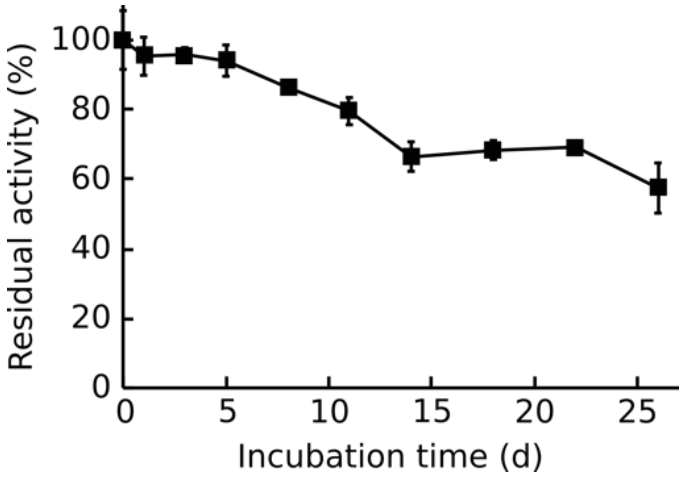


Figure S4 Thermostability of Est2 incubated at 30 ˚C for 26 days. Activity measured before incubation and under standard assay condition (14.8 U/mg) was regarded as 100 %.

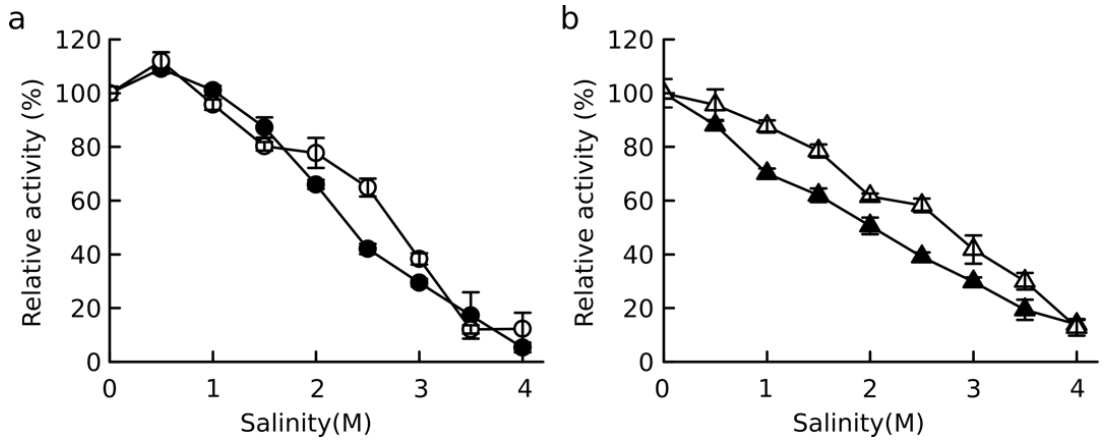


Figure S5 Effect of NaCl and KCl on esterase activity. a, Effect of NaCl (*closed circle*) and KCl (*open circle*) on Est1 activity. b, Effect of NaCl (*closed triangle*) and KCl (*open triangle*) on Est2 activity. Specific activity corresponding to 100 % relative activity is 30.7 U/mg for Est1 and 13.5 U/mg for Est2.

Table S1 Second-order rotable design for the combined effect of temperature and pH on Est1 and Est2 stability

Number of trials (N)	Coded values		Nature values		Residual activity (%)			
	pH	T	pH	T (°C)	Obtained (Est1)	Predicted (Est1)	Obtained (Est2)	Predicted (Est2)
1	1	1	8.5	69.1	1.8	9.8	52.5	48.8
2	-1	1	6.5	69.1	4.3	5.2	19.4	20.6
3	1	-1	8.5	40.9	98.2	96.0	90.4	87.4
4	-1	-1	6.5	40.9	56.0	46.6	53.9	59.2
5	-√2	0	6	55	13.3	19.1	56.1	51.6
6	√2	0	9	55	61.9	57.2	86.8	91.4
7	0	-√2	7.5	35	78.0	85.8	65.6	63.8
8	0	√2	7.5	75	2.3	-4.4	7.6	9.2
9	0	0	7.5	55	69.6	76.3	87.26	84.1
10	0	0	7.5	55	79.5	76.3	82.95	84.1
11	0	0	7.5	55	74.9	76.3	80.21	84.1
12	0	0	7.5	55	81.2	76.3	82.27	84.1
13	0	0	7.5	55	76.4	76.3	87.95	84.1

Table S2 Purification of recombinant Est1 and Est2.

Esterase[a]	Purification step	Total activity (U)	Total protein (mg)	Specific activity (U/mg)	Purification (fold)	Yield (%)
Est1	Crude enzyme	143.0	190.6	0.77	1	100
	Ni-TED	23.1	1.0	22.2	28.9	16.2
Est2	Crude enzyme	52.1	111.3	0.43	1	100
	Ni-TED	5.6	0.78	7.3	16.9	10.8

[a,] Est1 and Est2 activity were measured using 1 mM *p*-NP acyl caprylate (C8) under respective standard assay conditions.

Table S3 Analysis of variance (ANOVA) for the response surface models.

Esterase	Source of variance	DF[a]	SS[b]	MS[c]	Mean square ratios (α=0.05)[d]
Est1 ($\boldsymbol{R^2_{adj} = 0.953}$)[e]	Model	5	14291.3	2858.3	$F_1 = \frac{MSM}{MSEe} = 146.1 > F_4^5 = 6.26$
	Error	7	399.7	57.1	
	Experimental error	4	80.5	20.1	$F_2 = \frac{MSLF}{MSEe} = 5.28 < F_4^3 = 6.59$
	Lack of fit Model	3	319.3	106.4	
Est2 ($\boldsymbol{R^2_{adj} = 0.975}$)[e]	Model	4	8577.7	2144.4	$F_1 = \frac{MSM}{MSEe} = 192.45 > F_4^4 = 6.39$
	Error	8	144.7	18.1	
	Experimental error	4	44.6	11.1	$F_2 = \frac{MSLF}{MSEe} = 2.25 < F_4^4 = 6.39$
	Lack of fit Model	4	100.2	25.0	

[a] Degrees of Freedom;
[b] Sum of Squares;
[c] Mean Squares;
[d] MSM, mean squares of the model; MSMLF, mean squares for the model lack of fit; MSEe, mean squares for experimental error;
[e] $\boldsymbol{R^2_{adj}}$, adjusted R^2, indication of model quality.

Table S4 Effect of metal ions on Est1 and Est2 activity

Variable	Est1 relative activity (%)	Est2 relative activity (%)
None[a]	100± 4.1	100± 2.5
K+ 1mM	121.8±0.8	100.3±1.6
K+ 10mM	111.3±2.9	100.3±1.6
Mg2+ 1mM	98.7±3.3	92.8±5.9
Mg2+ 10mM	114.5±8.7	94.5±1.2
Ca2+ 1mM	90.4±6.0	98.3±7.4
Ca2+ 10mM	80.4±1.3	106.2±7.2
Zn2+ 1mM	84.0±6.0	98.4±2.0
Zn2+ 10mM	60.7±5.1	132.5±9.8
Cu2+ 1mM	83.5±3.3	93.5±4.3
Cu2+ 10mM	56.2±8.6	93.0±9.1
Ni2+ 1mM	87.7±1.6	87.8±0.7
Ni2+ 10mM	27.4±3.7	23.0±4.5
Mn2+ 1mM	94.9±3.5	93.8±3.9
Mn2+ 10mM	53.6±1.0	57.6±2.7
Fe2+ 1mM	86.0±6.2	101.0±6.6
Fe2+ 10mM	34.3±8.8	20.8±0.7
Fe3+ 1mM	86.9±7.8	92.5±5.3
Fe3+ 10mM	58.2±8.7	66.4±6.8
Al3+ 1mM	85.6±0.3	108.5±4.9
Al3+ 10mM	56.4±4.2	54.9±2.3

[a] Assay of metal ion effect for Est1 and Est2 was measured at 50 °C in phosphate buffer (pH=8). Specific activity corresponding to 100 % relative activity is: 23.0 U/mg for Est1, 10.4 U/mg for Est2.

Table S5 Effect of inhibitors and detergent on Est1 and Est2 activity

Variable	Est1 relative activity (%)	Est2 relative activity (%)	Variable	Est1 relative activity (%)	Est2 relative activity (%)
None[a]	100± 3.2	100± 1.6	None[a]	100± 3.2	100± 1.6
Triton X-100 (0.1 %)	67.6±1.0	70.4±2.0	EDTA 1mM	118.7±8.3	77.2±9.5
Triton X-100 (1 %)	30.3±3.8	40.8±5.1	EDTA 10mM	111.3±7.9	64.4±5.9
Tween 20 (0.1 %)	124.1±2.0	77.5±1.0	DTT 1mM	91.7±8.9	96.9±0.2
Tween 20 (1 %)	61.0±4.3	35.1±3.0	DTT 10mM	21.3±0.6	32.4±1.7
Tween 80 (0.1 %)	144.7±4.3	85.2±1.9	PMSF 1mM	84.8±8.2	89.1±13.7
Tween 80 (1 %)	122.9±5.8	36.5±1.5	PMSF 10mM	30.4±11.0	57.7±2.8
SDS (0.1 %)	ND[b]	ND[b]	DEPC 1mM	41.7±4.1	74.9±6.7

[a] Specific activity corresponding to 100 % relative activity is: 32.0 U/mg for Est1, 12.0 U/mg for Est2.
[b] Not detectable

Chapter IV

A novel caboxylesterase derived from a compost metagenome exhibiting high stability and activity towards high salinity

Mingji Lu, Rolf Daniel

Doi: https://doi.org/10.3390/genes12010122

https://www.mdpi.com/2073-4425/12/1/122

genes

MDPI

Article

A Novel Carboxylesterase Derived from a Compost Metagenome Exhibiting High Stability and Activity towards High Salinity

Mingji Lu and Rolf Daniel *

Department of Genomic and Applied Microbiology and Göttingen Genomics Laboratory, Institute of Microbiology and Genetics, Georg-August-University of Göttingen, Grisebachstr. 8, 37077 Göttingen, Germany; mingji.lu@biologie.uni-goettingen.de
* Correspondence: rdaniel@gwdg.de; Tel.: +49-551-393-3827

Abstract: Halotolerant lipolytic enzymes have gained growing interest, due to potential applications under harsh conditions, such as hypersalinity and presence of organic solvents. In this study, a lipolytic gene, *est56*, encoding 287 amino acids was identified by functional screening of a compost metagenome. Subsequently, the gene was heterologously expressed, and the recombinant protein (Est56) was purified and characterized. Est56 is a mesophilic (T_{opt} 50 °C) and moderate alkaliphilic (pH_{opt} 8) enzyme, showing high thermostability at 30 and 40 °C. Strikingly, Est56 is halotolerant as it exhibited high activity and stability in the presence of up to 4 M NaCl or KCl. Est56 also displayed enhanced stability against high temperatures (50 and 60 °C) and urea (2, 4, and 6 M) in the presence of NaCl. In addition, the recently reported halotolerant lipolytic enzymes were summarized. Phylogenetic analysis grouped these enzymes into 13 lipolytic protein families. The majority (45%) including Est56 belonged to family IV. To explore the haloadaptation of halotolerant enzymes, the amino acid composition between halotolerant and halophilic enzymes was statistically compared. The most distinctive feature of halophilic from non-halophilic enzymes are the higher content of acidic residues (Asp and Glu), and a lower content of lysine, aliphatic hydrophobic (Leu, Met and Ile) and polar (Asn) residues. The amino acid composition and 3-D structure analysis suggested that the high content of acidic residues (Asp and Glu, 12.2%) and low content of lysine residues (0.7%), as well as the excess of surface-exposed acidic residues might be responsible for the haloadaptation of Est56.

Keywords: carboxylesterases; metagenome; compost; lipolytic enzymes; halotolerance; halophilic; haloadaptation

Citation: Lu, M.; Daniel, R. A Novel Carboxylesterase Derived from a Compost Metagenome Exhibiting High Stability and Activity towards High Salinity. *Genes* **2021**, *12*, 122. https://doi.org/10.3390/genes12010122

Received: 21 December 2020
Accepted: 13 January 2021
Published: 19 January 2021

1. Introduction

Halophilic and halotolerant enzymes, which show resistance to salinity, are one of the major groups of extremozymes with industrial relevance. Microorganisms growing optimally at high salt concentrations are reservoirs for halophilic and halotolerant enzymes. They are generally divided into halophiles and halotolerant organisms [1,2]. To survive in high salinity, osmotic balance between cell cytoplasm and the external medium has to be maintained [3]. Halophilic archaea of the *Halobacteriales* order primarily adopt the "salt-in" strategy by accumulating equimolar concentrations of inorganic ions such as potassium and chloride ions [4,5]. This mechanism of osmoregulation results in intracellular enzymes, which evolve to halophilic types [6,7], which intrinsically show high activity and stability towards increasing salinity. In contrast, halophilic methanogenic archaea, as well as most halophilic and halotolerant bacteria largely employ the "salt-out" strategy by excluding salt from the cell inside and synthesizing and/or accumulating compatible organic osmolytes [8]. Some show also a combination of "salt-in" and "salt-out" strategies [9,10]. Halotolerant enzymes identified from "salt-out" microorganisms usually show

different levels of salinity resistance [11–13]. Unlike most halophilic enzymes, which are inactive under low salt concentrations, halotolerant enzymes are still active in the absence of NaCl [14].

Haloadaptation of halophilic enzymes has been extensively studied by amino acid sequence and 3D structure comparison [15–17]. The presence of an unusually high proportion of acid residues and a drastic reduction of lysine residues on the surface of proteins play a key role in haloadaptation of enzymes [18,19]. Comparative analyses of halotolerant and halophilic enzymes with respect to amino acid compositions or 3D structures are rare. Given that scrutiny of protein and genome sequences may not unravel small differences during sequence-based comparative analysis [2,20], homologous enzymes sharing relatively high sequence similarity and conserved 3D structure would be promising in unveiling differences potentially related to shared strategies of adaptation to high salt environments.

Lipolytic enzymes, including esterases (EC 3.1.1.1, carboxylesterase) and true lipases (EC 3.1.1.3, triacylglycerol acyl hydrolase), are involved in catalyzing the cleavage and formation of ester bonds. Esterases prefer short-chain substrates with an acyl chain length of less than 10 carbon atoms, while lipases mainly catalyze the hydrolysis of long-chain triacylglycerols ($\geq$10 carbon atoms) [21]. A distinguishing feature of lipolytic enzyme sequences is the conserved catalytic triad composed of a serine residue, which is located in the GXSXG consensus sequence, an aspartate or glutamate, and a histidine residue [22]. Most lipolytic enzymes also exhibit a similar core topology that typically consists of parallel β-pleated strands connected by α-helices [21]. Based on amino acid sequences and biological properties of lipolytic enzymes, 19 families (family I-XIX) have been identified [23]. Subsequently, new families such as EstLiu [24], Em3L4 [25], FLS18 [26], and Est9x [27] were added. Halophilic and halotolerant lipolytic enzymes have been detected in different families but the majority belonged to family IV [11,28–30]. Family IV esterases share high amino acid sequence similarity with mammalian hormone-sensitive lipases (HSL) and hence have also been referred to as the HSL family [21]. Previous studies on members of this family have predominantly explored the thermostability [31–33] and substrate specificity [34,35]. However, studies on salt tolerance of HSL family members are often missing.

Halophilic lipolytic enzymes are rare, among them, only LipC [36], MGS-B1 [37], and LipS2 [38] were characterized. In contrast, an increasing number of halotolerant lipolytic enzymes have been identified, through culture-dependent [28,39] and metagenomic approaches [29,40–42]. Apart from general features, such as broad substrate spectrum, chemo-, regio-, and enantio-selectivity and nonrequirement of cofactors, halophilic and halotolerant enzymes also tend to be resistant to organic solvents [15,31,43]. High salinity or presence of organic solvents both result in reduced water activity. Thus, halophilic and halotolerant lipolytic enzymes are advantageous in applications involving nonaqueous and aqueous/organic media, such as the degradation of organic pollutants in saline wastewater, bioremediation of oil spills, production of inter-esterification substances in food industry and nonaqueous synthesis of non-natural chemical compounds [14,44,45].

Composting is a process of decomposition and humification of organic matter [46]. During composting, enzymes secreted by microorganisms play key roles in degrading various organic substances. Recently, extremophilic enzymes, such as thermophilic, organic solvent tolerant, and alkaliphilic lipolytic enzymes, have been successfully identified from compost metagenomes [47–49]. However, there are no current reports of halophilic or halotolerant lipolytic enzymes from compost. In our previous study, a compost sample at the thermophilic stage (55 °C) was used to construct a metagenomic plasmid library [50] of which the identified esterase-encoding gene *est56* was selected for characterization. Est56 was a novel member of lipolytic family IV and further characterized as halotolerant. The haloadaptation mechanism of Est56 was also explored by comparing its amino acid composition with other halotolerant and halophilic enzymes, and subsequent analysis of homology-modelled 3D structures.

2. Materials and Methods

2.1. Bacterial Strains and Plasmids

The lipolytic recombinant plasmid (pFLD56) harboring the putative lipolytic gene *est56* was derived from the functional screen of the metagenomic compost library as described in our previous study [50]. The plasmid pFLD56 was used as template for amplification of *est56*. *E. coli* strain BL21(DE3) and the plasmid pET101/D-TOPO® (Invitrogen, Karlsruhe, Germany) were used as expression host and vector, respectively.

2.2. Identification and Analysis of Est56 Sequence

Putative open reading frames (ORF) encoding lipolytic enzymes were initially predicted using the ORF Finder program (http://www.ncbi.nlm.nih.gov/gorf/gorf.html), and further verified using FramePlot analysis [51]. Potential signal peptides were detected using SignalP 4.0 [52]. Amino acid sequence similarity searches were performed using the BLASTP program against the public GenBank database [53]. Amino acid sequences of esterases homologous to the deduced gene product of est56 were aligned by employing clustalW. Secondary structure prediction was performed with I-TASSER [54]. Annotation of aligned sequences was performed with ESPript 3.0 [55].

2.3. Cloning and Expression of Est56

The putative lipolytic gene *est56* was amplified using the following primers: 5′-CACCATGCTCGCGCAGTCAC-3′ and 5′-CCCCTGGCGCGGGTAGTGTTCG-3′. The est56 PCR product was cloned into pET101/D-TOPO® vector (Invitrogen, Germany), following the manufacturer's instructions. The resulting recombinant plasmid was transformed into *E. coli* BL21 (DE3) cells. For expression of *est56*, a 6 mL preculture, incubated overnight, was used to inoculate 600 mL LB medium containing 100 µg/mL ampicillin. The culture was incubated overnight with shaking at 30 °C to an OD_{600} of 0.6. Expression was induced by adding IPTG (isopropyl-b-D thiogalactopyranoside) to a final concentration of 0.5 mM. After incubation for 6 h at 30 °C, cells were harvested by centrifugation (7000× *g*, 10 min, 4 °C), suspended in lysis-equilibration-wash buffer (50 mM NaH_2PO_4, 300 mM NaCl, pH 8), and then lysed on ice by sonication using a UPS200S homogenizer (Hielscher Ultrasonics GmbH, Teltow, Germany).

2.4. Purification of Recombinant Est56

To purify His_6-tagged Est56, Protino® Ni-TED 2000 packed columns (Macherey-Nagel, Düren, Germany) were used as recommended by the manufacturer. SDS-PAGE was performed to determine the purity and molecular mass of Est56. Protein concentration was measured by the Bradford method [56]. Finally, fractions containing the purified enzyme were pooled and dialyzed against 50 mM sodium phosphate buffer (pH 8) at 4 °C.

2.5. Esterase Standard Assay

Unless otherwise mentioned, Est56 activity was measured in 1.0 mL assay buffer containing 50 mM sodium phosphate (pH 8) and 1 mM *p*-nitrophenyl (*p*-NP) butyrate (Sigma-Aldrich, Munich, Germany), at 50 °C. The amount of *p*-NP released by enzyme-catalyzed hydrolysis was continuously monitored for at least 2 min at a wavelength of 410 nm against an enzyme-free reference. One unit of enzymatic activity was defined as the amount of Est56 needed to release 1 µmol of *p*-NP per minute under the assay conditions. All experiments were performed in at least triplicate, and extinction coefficients of *p*-NP under every assay condition were determined. Results are shown as mean values ± standard deviation (SD).

2.6. Characterization of Est56

2.6.1. Substrate Specificity

Substrate specificity of Est56 was assessed under standard assay conditions using 1 mM *p*-NP esters (Sigma-Aldrich, Munich, Germany) of different chain lengths as substrates: *p*-NP acetate (C2), *p*-NP butyrate (C4), *p*-NP valerate (C5), *p*-NP caproate (C6), *p*-NP caprylate (C8), *p*-NP caprate (C10), *p*-NP laurate (C12), *p*-NP myristate (C14), and *p*-NP palmitate (C16). All substrates were prepared as a 0.1 M stock solution dissolved in isopropanol. For long chain substrates (>C10), the stock solution was first heated (50 °C) for a short time until the formation of clear transparent solution [40]. Initial reaction rates were calculated by estimating Est56 activity with different substrate concentrations ranging from 5 to 5000 µM. Michaelis–Menten constant (K_m) and the maximal velocity (V_{max}) were determined by employing Lineweaver–Burk plots [57].

2.6.2. Effect of Temperature and pH

Optimum temperature of Est56 activity was determined between 20 and 70 °C. To assess protein thermostability, Est56 was preincubated in assay buffer at 30, 40, 50, and 60 °C for different time periods, and subsequently, residual activity was measured under standard assay condition.

Due to pH-dependent absorption of p-NP in acidic buffers [58], the effect of pH on Est56 activity was determined between pH values 3 and 10 at 348 nm (the pH-independent isosbestic wavelength). The overlapping buffer systems used comprised 50 mM acetate buffer (pH 3.0 to 6.0), 50 mM sodium phosphate buffer (pH 6.0 to 8.0), 50 mM TAPS (3-(2,4 dinitrostyrl)-(6*R*,7*R*-7-(2-thienylacetamido)-ceph-3-em-4-carboxylic acid) buffer (pH 8.0–9.0), and 50 mM CHES (N-cyclohexyl-2-aminoethanesulfonic acid) buffer (pH 9.0 to 10.0). The effect of pH on Est56 stability was examined by preincubating the enzyme at the respective pH value at 10 °C for 24 h. Subsequently, residual activity was measured under standard assay condition.

2.6.3. Effect of Salinity

The effect of salinity on Est56 activity was measured by adding 0.5 to 4 M NaCl or KCl to the standard reaction assay mixture. Est56 stability against salt was determined by incubating the enzyme in assay buffer containing NaCl (0.5 to 4 M) at 10 °C for 24 h. Residual activity was measured under standard assay conditions.

Salt has been reported to protect halophilic proteins against denaturants (such as high temperature and urea) [18]. In this study, the protective effect of NaCl was investigated by adding different amounts of NaCl (0–4 M) to the incubation buffer, in which Est56 was incubated at different temperatures or in the presence of different amounts of urea. Specifically, Est56 was incubated at high temperatures (50 or 60 °C) for 30 min. For urea impact, Est56 was incubated with different amounts of urea (2, 4, and 6 M) at 10 °C for 24 h. Each additive was equalized to the same final concentration in the assay buffer, and the residual activity was measured under standard assay conditions. A blank reference was prepared using the buffer solution without enzyme but containing the same amount and type of additive. Activity measured before incubation was taken as 100%.

2.6.4. Effect of Organic Solvents

The effect of organic solvents on Est56 stability was assayed by incubating Est56 in the presence of 15% and 30% (*v*/*v*) water-miscible organic solvents (DMSO, methanol, ethanol, acetone, isopropanol, and 1-propanol) or water-immiscible organic solvents (ethyl acetate, diethyl ether, chloroform, and toluene) at 10 °C for 24 h under constant shaking. Residual activity was measured under standard assay conditions.

2.6.5. Effect of Other Additives

The effect of metal ions including K^+, Ca^{2+}, Mn^{2+}, Mg^{2+}, Zn^{2+}, Fe^{2+}, Cu^{2+}, Ni^{2+}, Fe^{3+}, and Al^{3+} and inhibitors including phenylmethyl-sulfonyl fluoride (PMSF), dithiothreitol

(DTT), and ethylenediaminetetraacetic acid (EDTA) was measured at concentrations of 1 and 10 mM. In addition, the impact of detergents such as Triton X-100, Tween 20, Tween 80, and SDS at concentrations of 0.1%, 1% and 5% (v/v) on enzyme activity was investigated. The catalytic activity of Est56 was measured under standard reaction conditions by directly adding each additive to the standard assay mixture. Activity measured in additive-free assay buffer was regarded as 100% activity, while reactions that included corresponding additive but no enzyme were used as blanks.

2.7. Sequence Analysis of Halotolerant Enzymes

Phylogenetic trees were constructed with the neighbor-joining method using MEGA version 7 [59]. For this purpose, Est56, other reported halotolerant, and reference lipolytic enzyme sequences retrieved from GenBank were employed (Table S1). A bootstrap value of 500 replicates was used to estimate the confidence level [60]. The phylogenetic tree was subsequently visualized by GraPhlAn [61]. In addition, a tree containing only family IV esterases was constructed.

Characterized halophilic lipolytic enzymes are rare but enzymes from halophilic microorganism that adopt the "salt-in" strategy were evolved to be halophilic [6]. Thus, we retrieved 22 putative halophilic lipolytic enzymes originating from archaea of the Halobacteriales order from GenBank (Table S2). The halophilic feature of each archaeal strain was checked on HaloDom webserver (http://www.halodom.bio.auth.gr) [62]. We summarized experimentally confirmed halophilic proteins as controls (Table S3). Thereafter, we used HT_Lip, HP_Lip, and HP_Enz to refer to the 40 halotolerant lipolytic enzymes, 22 putative halophilic lipolytic enzymes, and 16 experimentally confirmed halophilic proteins, respectively. Amino acid composition, theoretical pI values, and molecular weight of each protein were calculated by the ProtParam tool at Expasy (www.expasy.org) [63].

Differences in amino acid compositions among HT_Lip, HP_Lip, and HP_Enz were statistically compared. Individual amino acids were analyzed as variants for each enzyme. Nonparametric Kruskal–Wallis (KW) and Mann–Whitney (MW) pairwise post hoc tests were used to evaluate median differences among univariate groups. Analysis of similarities (ANOSIM) test was used to pairwise compare the overall differences between multivariate groups based on Bray–Curtis distance, with 9999 permutations. A high R value generated by the ANOSIM test indicates a high dissimilarity between groups. Similarity percentage (SIMPER) tool calculates the average contribution of individual amino acids to the average dissimilarity between groups based on Bray–Curtis similarity. Statistical analyses were performed with R (http://www.r-project.org) using the "vegan" package [64].

2.8. Homology Modeling and Putative Structure Analysis

Based on deduced amino acid sequence, the tertiary structure prediction of Est56 was performed by I-TASSER [54]. PyMOL (PyMOL molecular graphics system, DeLano Scientific, Palo Alto, CA, USA; http://www.pymol.org) was used to visualize the predicted model. The analog with the highest TM score was also selected for structural superimposition. The surface electrostatic potential was calculated by the APBS plugin [65] and visualized by PyMOL.

2.9. Accession Numbers

The amino acid sequence of Est56 is available in the GenBank database under accession number KR149569.1. The compost metagenome sequences are available in the NCBI sequence read archive (SRA) under the accession number SRR13115019.

3. Results

3.1. Identification and Sequence Analysis of a Novel Lipolytic Gene

A metagenomic plasmid library derived from compost using *Escherichia coli* as host was constructed and function-based screened for genes conferring lipolytic activity as previously described [50]. An *E. coli* clone showing strong lipolytic activity (large halos)

on indicator plates was selected for further characterization. Sequence analysis of the plasmid insert (6.1 kb, Figure S1) revealed a putative lipolytic gene (*est56*, 864 bp) encoding 287 amino acids. Putative signal peptides indicating extracytoplasmic localization were not detected in the deduced protein sequence. Similarity searches showed that protein sequences similar to Est56 were mainly identified during metagenome screenings. These comprised ELP45 isolated from a forest topsoil [66], EstC23 from mountain soil [67], Est06 from forest soil [68], and EstMY from activated sludge [69], which showed 62%, 61%, 60%, and 60% amino acid identity to Est56, respectively.

Multiple sequence alignments of Est56 with other esterases revealed that Est56 belongs to the HSL group (family IV) of lipolytic proteins. The conserved family IV motif H-G-G was present in the Est56 protein at amino acid sequence positions from 62 to 64. This motif plays an essential role in the stabilization of the oxyanion hole and catalysis. The catalytic triad composed of Ser^{132}, Glu^{226}, and His^{256}, as well as another conserved motif E-X-L-X-D-D (amino acid residues from 226 to 231), was also detected in Est56 amino acid sequence (Figure S2).

3.2. Purification of Recombinant Est56 and Substrate Specificity

Est56 was heterologously produced in *E. coli* BL21 (DE3). After purification by Ni-TED affinity chromatography, Est56 was purified 73-fold with a specific activity of 90.44 U/mg (Table S4). SDS-PAGE revealed a single band with a molecular mass of approximately 34.0 kDa (Figure S3). This is consistent with the calculated protein mass including the sequences for the V5 epitope and His_6-tag, which were added during cloning of *est56*.

Assays with *p*-NP esters showed that Est56 exhibited a substrate preference for esters with short-chain fatty acids such as *p*-NP acetate (C2), *p*-NP butyrate (C4), *p*-NP valerate (C5), and *p*-NP caproate (C6). The enzyme did show little or no significant activity by employing *p*-NP esters with long-chain fatty acids (C8–C16) as substrates (Figure S4). This indicates that Est56 is an esterase and not a lipase [21]. The maximal specific activity was detected with *p*-NP butyrate (C4). The K_m and V_{max} values with this substrate were 128.0 μM and 102.0 U/mg, respectively.

3.3. Effect of Temperature and pH

Est56 retained high activity (above 40%) over the entire tested temperature range from 20 to 70 °C with maximal activity at 50 °C (Figure 1a). Despite the high activity (above 80%) of Est56 at 50 and 60 °C, thermostability at 50 and 60 °C was low. Est56 retained 50% activity after 30 min of incubation at 50 °C and 12.4% after 10 min at 60 °C. However, Est56 showed high stability for extended incubation times at 30 and 40 °C, with a half-life of 192 and 16 h, respectively (Figure 1c).

Est56 exhibited more than 80% activity between pH 6 and 8, with maximal activity at pH 8 (Figure 1b). Similarly, Est56 was most stable from pH 6 to 8, retaining more than 90% residual activity after 24 h incubation at 10 °C (Figure 1d).

3.4. Effect of Salinity

Addition of salt (NaCl or KCl) produced a stimulatory effect on Est56 activity, with an enhanced activity at NaCl and KCl concentrations of up to 2.5 M. Maximal activities compared to the reference without addition were recorded in the presence of 1.5 M NaCl (130.4%) and 1 M KCl (141.6%). At higher concentrations, Est56 activity decreased gradually with increasing concentration of NaCl and KCl. Notably, Est56 still retained approximately 90% activity at 3 M NaCl and 3.5 M KCl and 40% activity at 4 M NaCl and KCl (Figure 2a).

As for the stability, Est56 was stable over the tested NaCl and KCl concentration ranges (0–4 M), with almost unaltered activity after 24 h incubation at 10 °C (Figure 2b).

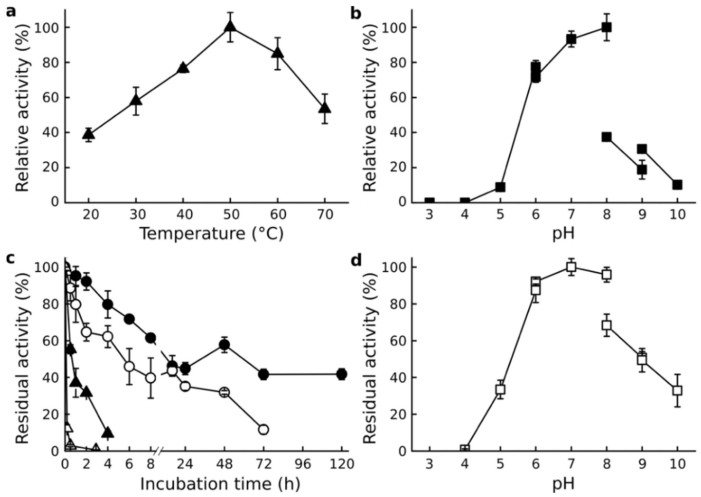


Figure 1. Effect of temperature and pH on Est56 activity. (**a**) Effect of temperature on Est56 activity; the maximal activity (153.9 U/mg) at 50 °C was taken as 100%. (**b**) Effect of pH on Est56 activity; the maximal activity (90.4 U/mg) at pH 8 was taken as 100%. (**c**) Thermostability of Est56 at 30 °C (closed circle), 40 °C (open circle), 50 °C (closed triangle), and 60 °C (open triangle); specific activity (87.2 U/mg) measured at the start of the incubation and under standard assay conditions was taken as 100%. (**d**) Effect of pH on Est56 stability was measured by incubating Est56 at 4 °C for 24 h, the maximal residual activity (33.1 U/mg) at pH 7 was taken as 100%.

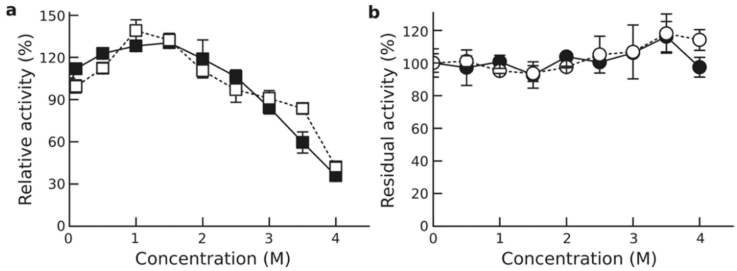


Figure 2. Effect of salinity on Est56 activity and stability. (**a**) Effect of NaCl (closed square) and KCl (open square) on Est56 activity. (**b**) Effect of NaCl (closed circle) and KCl (open circle) on Est56 stability. Specific activities corresponding to 100% activity were 93.1 and 94.1 U/mg for graph (**a**) and (**b**), respectively.

3.5. Protective Effect of NaCl against Denaturants

The presence of NaCl enhanced stability of Est56 to temperature (Table 1). At 50 °C, the addition of 1, 2, and 3 M NaCl significantly ($p < 0.05$) increased Est56 residual activity (above 70%) compared to that incubated in the salt-free assay. The best stabilization effect was detected at 2 M NaCl. However, incubation with 4 M NaCl significantly ($p < 0.05$) decreased Est56 stability. At 60 °C, Est56 was almost inactivated after 30-min without NaCl addition, whereas a significant increase in residual activity was detected by adding 1, 2, 3, or 4 M NaCl.

Table 1. Role of NaCl in protecting Est56 against denaturants.

	Residual Relative Activity (%) [a]				
	0 M	**1 M**	**2 M**	**3 M**	**4 M**
Temperature [b]					
50 °C	54.9 ± 2.4	71.3 ± 8.1 *	82.8 ± 7.4 **	72.9 ± 4.2 *	38.9 ± 3.1 **
60 °C	1.4 ± 0.7	12.7 ± 2.2 **	11.8 ± 1.6 **	8.0 ± 0.7 **	10.6 ± 0.6 **
Urea [c]					
2 M	36.3 ± 4.7	36.2 ± 4.0	43.0 ± 4.7	54.0 ± 2.4 **	41.5 ± 9.5
4 M	19.5 ± 3.5	31.6 ± 2.5 **	33.6 ± 2.2 **	29.3 ± 4.0 *	ND [d]
6 M	17.2 ± 2.4	28.6 ± 2.2 **	ND [d]	ND [d]	ND [d]

[a] Residual activity was measured after incubation with different amounts of NaCl. Two-sample *t* test was performed between residual activities incubated with and without NaCl. * α = 0.05; ** α = 0.01. [b] Est56 was incubated at 50 or 60 °C for 30 min in the presence of different concentrations of NaCl. Specific activity (41.9 U/mg) measured before incubation was taken as 100%. [c] Est56 was incubated at 10 °C for 24 h in the presence of different combinations of NaCl and urea. Specific activity (45.8 U/mg) measured before incubation was taken as 100%. [d] Not detectable.

Urea is another denaturant that can cause inactivation of an enzyme directly and indirectly [70]. Addition of 3 M NaCl significantly ($p < 0.05$) enhanced Est56 tolerance against 2 M urea, with the residual activity increasing from 36.3% without NaCl addition to 54.0% in the presence of 3 M NaCl. A stabilizing effect of NaCl ($p < 0.05$) on Est56 activity was also observed in the presence of 4 or 6 M urea (Table 1).

3.6. Effect of Organic Solvents

Generally, Est56 exhibited enhanced activity after incubating with 15% and 30% (v/v) water-miscible organic solvents (Table 2), with the exception of 30% (v/v) 1-propanol, which caused a considerable loss of Est56 residual activity (23.0%). The highest stimulation of residual activity (286%) was observed after incubation with 30% (v/v) isopropanol. In the presence of water-immiscible organic solvents, Est56 was inhibited by 15% and 30% (v/v) ethyl acetate, chloroform, and toluene (Table 3). Est56 retained its full activity after incubation with 15% and 30% (v/v) diethyl ether.

Table 2. Effect of water-miscible organic solvents on Est56 stability.

Organic Solvent	**Residual Activity (%) [a]**	
	15% (*v*/*v*)	**30% (*v*/*v*)**
DMSO	135.0 ± 11.8	135.4 ± 4.7
Methanol	116.1 ± 6.7	96.6 ± 7.6
Ethanol	116.2 ± 8.7	107.3 ± 4.2
Acetone	138.3 ± 8.9	115.3 ± 14.8
Isopropanol	136.3 ± 10.1	286.6 ± 9.0
1-Propanol	114.0 ± 3.7	23.0 ± 1.0

[a] Specific activity (33.4 U/mg) incubated in the organic solvent-free assay buffer was taken as 100%.

Table 3. Effect of water-immiscible organic solvents on Est56 stability.

Organic Solvent	**Residual Activity (%) [a]**	
	15% (*v*/*v*)	**30% (*v*/*v*)**
Ethyl acetate	55.3 ± 8.1	50.2 ± 3.5
Diethyl ether	110.8 ± 5.7	106.7 ± 10.8
Chloroform	46.5 ± 0.5	6.3 ± 2.5
Toluene	50.0 ± 2.2	54.9 ± 0.7

[a] Specific activity (53.4 U/mg) incubated in the organic solvent-free assay buffer was taken as 100%.

3.7. Effect of Other Additives

Metal ions exhibited different effects on Est56 activity (Table S5). The addition of Al^{3+} and Ca^{2+} at 1 and 10 mM had a stimulatory effect (approximately 130%). Est56 retained its full activity in the presence of 1 and 10 mM Mg^{2+}. In contrast, Est56 activity slightly decreased to approximately 80% in the presence of 1 mM Fe^{2+}, 10 mM Mn^{2+}, and 1 and 10 mM Zn^{2+}. The additives Cu^{2+} and Ni^{2+} at 1 and 10 mM and Fe^{2+} and Fe^{3+} at 10 mM were deleterious to Est56, as enzyme activity dropped to approximately 20%. EDTA did not affect Est56 activity, which indicated that Est56 activity is independent of divalent cations.

The nonionic detergents Triton X-100, Tween 20, and Tween 80 at 0.1% (v/v) significantly enhanced Est56 activity to 156.8%, 154.3%, and 112.7%, respectively. Est56 activity was inhibited or inactivated in the presence of 1% and 5% (v/v) of the other tested detergents (Table S6). The addition of the inhibitors DTT, PMSF, and DEPC at 1 and 10 mM had detrimental effects on Est56 activity (Table S7). The inhibition of enzyme activity by PMSF and DEPC indicated that serine and histidine residues, respectively, are part of the Est56 catalytic triad [71].

3.8. Sequence Analysis of Halotolerant Lipolytic Enzymes

A phylogenetic tree was constructed to group 40 halotolerant lipolytic enzymes (Table S1) into families based on Arpigny and Jaeger [21] classification. As shown in Figure 3, these enzymes covered nine Arpigny and Jaeger families including family I, II, IV, V, VI, VII, VIII, XV, and XVII, as well as four new families including Est9x [27], EstLiu [24], lp_3505 [72], and EM3L4 [25]. Most of the analyzed halotolerant enzymes (18 enzymes) including Est56 belonged to Family IV. Family IV can be further divided into two sub-families based on the conserved GXSXG motif [34]. Est56 and other 10 halotolerant lipolytic enzymes belong to the GDSAG motif subfamily and the remaining 7 enzymes grouped into the GTSAG motif subfamily (Figure S5).

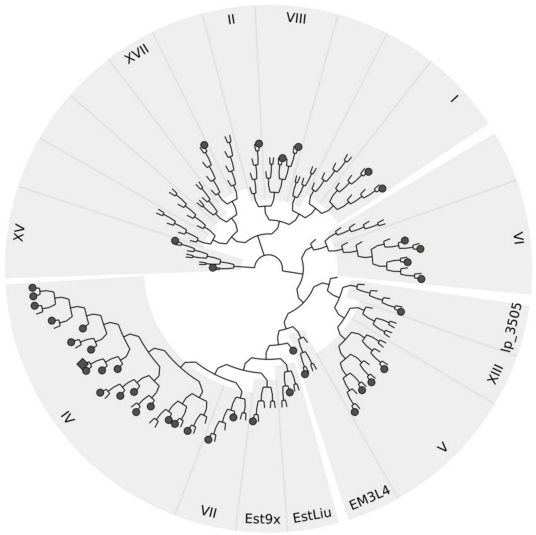


Figure 3. Phylogenetic classification of Est56 (closed diamond) and reported halotolerant (closed circle) lipolytic enzymes by neighbor-joining method. With the exception of Est56, other sequences were retrieved from GenBank.

Generally, the overall amino acid compositions among 3 different groups of selected and characterized halotolerant or halophilic enzymes were analyzed. The halotolerant lipolytic enzymes (HT_Lip; Table S1), the halophilic lipolytic enzymes (HP_Lip; Table S2),

and other characterized halophilic enzymes (HP_Enz; Table S3) are pairwise different (ANOSIM test, $p < 0.01$ in all three cases). The R values between HT_Lip and HP_Lip (0.4254) and HT_Lip and HP_Enz (0.4251) were higher than that of HP_Lip and HP_Enz (0.2065). This result indicated a high separation of amino acid composition between halotolerant (HT_Lip) and halophilic (HP_Lip and HP_Enz) enzymes, rather than within halophilic enzymes. This result was also consistent with the average dissimilarity revealed by SIMPER analysis (Table S8). The KW test identified that 12 residues (Asp, Glu, Lys, Arg, His, Leu, Met, Ile, Ala, Ser, Asn, and Gln) (Figure 4a) and theoretical pI values (Figure 4b) were significantly different ($p < 0.05$) among the three groups. The groups of halophilic enzymes HP_Lip and HP_Enz obtained significantly higher content of aspartic acid, glutamic acid, and arginine residues and lower content of lysine, leucine, methionine, isoleucine, and asparagine residues and theoretical pI values than those in halotolerant lipolytic enzymes HT_Lip (MW post hoc pairwise test, $p < 0.05$). As reported above, the most notable feature for the halophilic adaptation of halophilic enzymes is the excessive number of acidic residues (Asp and Glu) compared to lysine residues. This is reflected by low theoretical pI values [15]. However, halotolerant lipolytic enzymes exhibited broad range of theoretical pI values (4.59 to 9.44) (Figure 4b). Among them, Est56 obtains high content of acidic residues (Asp and Glu, 12.12%) and particularly low content of lysine residues (0.7%), as well as a relatively low theoretical pI value of 4.97.

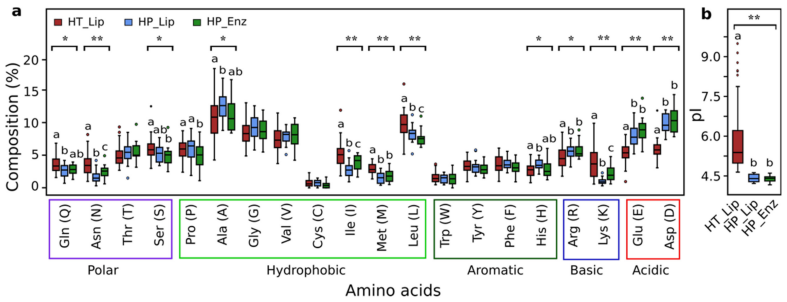


Figure 4. Comparison between halotolerant and halophilic enzymes. Box plots of (**a**) amino acid composition and (**b**) theoretical pI value of enzymes in groups HT_Lip (red), HP_Lip (blue), and HP_Enz (green). Statistical comparison was conducted by Kruskal–Wallis (KW) test (**, $p < 0.01$; *, $p < 0.05$) among three groups and Mann–Whitney (MW) post hoc test between pairwise groups (medians sharing a letter above boxes indicate no significant difference in the pairwise test).

3.9. Structural Modeling of Est56

The tertiary structure of Est56 is composed of a cap domain and an α/β-hydrolase fold core domain. The cap domain of Est56 consists of α-helices at the N-terminal side (α1 to α2) and between β6 and β7 (α6 to α7) (Figure 5a). The core domain comprises six helices surrounded by eight β-strands that form parallel structures. The catalytic triad of Est56 consists of Ser^{132} located between β5 and α5, Glu^{226} after β7, and His^{256} between β8 and α9 (Figure 5a). The overall structure of Est56 superimposed well (TM-score 0.984; RMSD 0.48) on E40 [73], with a global amino acid sequence identity of 53% (Figure 5b). The electrostatic potential of Est56 was calculated and described. The distribution of charges revealed that Est56 had negative charges on the surface (Figure 5c,d).

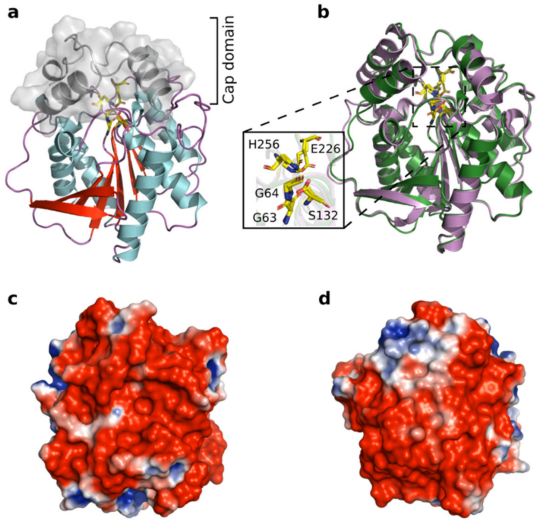


Figure 5. The modelled three-dimensional structure of Est56. (**a**) Ribbon representation of the Est56 monomer colored according to secondary structure elements. The overall structure is composed of two domains: cap domain and catalytic domain. (**b**) Superposition of Est56 (pink) onto its structural homolog E40 (PDB: 4xvc). The residues involved in stabilizing the oxyanion hole (Gly63 and Gly64) and catalytic triad (Ser94, Asp215 and His245) were indicated in stick representation (the sequence number is based on Est56 amino acid sequence). (**c**) Structural electrostatic potential of Est56. The most negative and most positive electrostatic potentials are indicated by red and blue, respectively, from −1 (red) to +1 kT/e (blue). (**d**) The 180° rotated view of (**c**).

4. Discussion

In this study, we identified a halotolerant esterase, Est56, from a compost metagenome and compared it to recently reported halotolerant enzymes (Table S1). In total, the 46 halotolerant lipolytic enzymes were grouped into 13 lipolytic families, which suggested a high diversity of halotolerant lipolytic enzymes (Figure 3). Including Est56, 29 halotolerant lipolytic enzymes were derived from metagenomes and correspondingly from uncultured microorganisms, which confirms the efficiency of metagenomic approaches in exploring novel enzymes [74,75]. The remaining 17 halotolerant enzymes were derived from individual microorganisms of which most thrive in saline environments, i.e., *Zunongwangia profunda* was derived from a surface seawater [24], *Erythrobacter seohaensis* from a tidal flat [76], *Psychrobacter celer* from a deep-sea sediment [77], and *Alkalibacterium* sp. SL3 from a soda lake sediment [78]. Some of the metagenome-derived halotolerant enzymes also originated from saline environments, such as marine water [27,37,41,79], deep sea sponges [42,80], deep-sea sediments [25,81], and deep-sea shrimps [82]. Recently, compost was reported as an habitat for halophilic and halotolerant microorganisms [83,84]. To our knowledge, Est56 is the first halotolerant esterase identified from a compost metagenome.

Est56 activity exhibited a T_{opt} at 50 °C and high stability at 30 °C (Figure 2). This is also the case for the majority of halotolerant esterases, which show a temperature optimum of enzyme activity between 35 and 50 °C and high stability at low temperatures (below 30 °C in most cases) [34,85–87]. Nevertheless, psychrophilic and thermophilic halotolerant lipolytic enzymes were also detected. For example, 7N9 [42], lp_3505 [72], EstS [11], Est10 [88], and EstLiu [24] displayed a temperature optimum ranging from 0 to 30 °C, whereas E69 [76], Est9x [27], BmEST [89], and LipJ2 [39] exhibit their optimal activity between 60 and 80 °C. With respect to pH, most of the halotolerant enzymes including Est56 are moderate alkaliphilic, with the optimum pH between 7.5 and 9 (Table S1).

Generally, halotolerant lipolytic enzymes could be classified into three groups, according to their NaCl-dependent catalytic activity (Table S1). In the first group comprising Lpc53E1 [80], PE10 [86], Lip3 [90], MGS-RG1 [82], LipC12 [40], Est9x [27], and Est700 [91] NaCl stimulated esterase activity was recorded over the tested concentration range (0–4 M in most cases). The second group harbors enzymes for which esterase activity continuously decreases with increasing NaCl concentration. Members of this group are 7N9 [42], ABO_1251 [79], Esth [92], ABO_1197 [79], EstS [11], and EstLiu [24]. For the third group, NaCl serves as an activator at low or moderate concentrations and as inhibitor at higher concentrations. Est56 belonged to this group, as its activity increased in the presence of NaCl up to a concentration of 1.5 M and decreased at higher concentrations (Figure 2). Moreover, almost 50% of the described halotolerant esterases such as E69 [76], ThaEst2349 [85], E25 [34], lp_3505 [72], and Est10 [88] follow this pattern.

Evaluation of enzyme stability towards salts revealed that Est56 retained its full activity after 24 h of incubation with NaCl or KCl concentrations ranging from 0 to 4 M at 10 °C (Figure 2b). Similarly, almost unaltered residual activity was also detected for esterases such as Lpc53E1 [80], EaEST [93], Esth [92], EstSL3 [78], LipC12 [40], Est12 [77], Est-OKK [94], and Est700 [91]. Nevertheless, esterases such as EstSP [30], EstS [11], EstLiu [24], and H8 [81] were inhibited by high salt concentrations. Noteworthy are the stabilities of EM3L4 [25], ThaEst2349 [85], Lip3 [90], and Est10 [88], which were enhanced by addition of NaCl. Although potassium ions are preferable for some halophilic enzymes [15,95], NaCl and KCl exhibited similar effects on Est56 activity and stability. This was tested only for a few other enzymes and similar results were obtained [29].

To shed a light on the mechanism of how halotolerant lipolytic enzymes resist salts, we compared the amino acid compositions among the groups HT_Lip, HP_Lip, and HP_Enz (Figure 4). The amino acid composition between halophilic and halotolerant proteins are significantly different, although enzymes in all groups were salt resistant. Our results were generally consistent with previous comparisons between halophilic and nonhalophilic homologs [9,96–99]. The most crucial feature of halophilic enzymes (HP_Lip and HP_Enz) compared to halotolerant lipolytic enzymes (HT_Lip) is the higher content of acidic residues (Asp and Glu) accompanied by a lower content of lysine residues (Figure 4a). This feature was also reflected by the lower pI values of halophilic enzymes (Figure 4b). Another consistent feature for halophilic enzymes is the low hydrophobicity [4,15], which was presented by a low content of aliphalic hydrophobic residues (Leu, Met and Ile) in this study (Figure 4a). Significant difference in aromatic (Phe, Try, and Trp), small (Gly and Vla), and borderline (Ser and Thr) hydrophobic residues was not detected between HT_Lip and HP_Lip (Figure 4a). On the contrary, different results were reported for certain residues therein by comparing halophilic and non-halophilic homologs [97,99–102]. Moreover, a significantly lower content of polar amino acids (Gln and Asn) was recorded for HP_Lip compared to HT_Lip enzymes (Figure 4a), which has not been observed in other comparative analyses. In general, high acidity and low hydrophobicity reflected by the amino acid composition of halophilic enzymes is a distinctive feature for their "halophilic adaptation," which simultaneously enables them to resist salinity. Nonetheless, we could not find clear patterns in amino acid composition among halotolerant lipolytic enzymes, except the broad range of theoretical pI values (Figure 4b).

Recently, Dassarma and Dassarma [2] proposed a correlation between the halophilic character and acidic nature of proteins by reviewing the proteomes of different halophilic and halotolerant bacteria. Further studies on halophilic protein structure also confirmed that excessive surface-exposed acidic residues are the basis for halophilic adaptation [5,103–106]. Thus, given the low pI values of halophilic enzymes, we assume that halotolerant lipolytic enzymes with relatively low pI values would follow a similar haloadaptation as halophilic enzymes. Est56 exhibits a low pI value (4.97), with predominantly acidic residues located at its surface (Figure 5c,d). This is an indication that Est56 applies a salt resistance strategy similar to that of halophilic enzymes. Moreover, the distinctive feature of enhanced Est56 stability against denaturants (high temperatures and

urea) mediated by NaCl suggests that Est56 possess a halophilic character caused by the high acidic amino acid content. Another family IV esterase, ThaEst2349 (theoretical pI value 4.94), was also reported as halotolerant due to the high ratio of surface-exposed acidic residues [85]. Acidic amino acids were reported to have a greater capacity than other amino acids in keeping proteins hydralated, which is important for the solubility of protein under salt stress [95,107,108]. Interestingly, several reports on the haloadaptation of enzymes with higher pI values than Est56 indicated that instead of acidic residues, basic residues at protein surfaces played a key role in their halotolerance [91,109,110]. By site-directed mutagenesis, Zhang et al. [81] identified two basic residues Arg^{195} and Arg^{236} located on the surface of H8 (theoretical pI value 9.09), which were essential in the salt tolerance. Additional mechanisms were reported to contribute to the haloadaptation of halotolerant enzymes. By introducing hydrophobic residues in the cap and catalytic domain, the halotolerance of E40 (a family IV esterase) was significantly improved [31]. However, this adaptation hardly applies to all of the family IV esterases, since the cap domain is the most variable region [111]. The cap domain was even not observed for the family IV enzyme MGS-M1, which was also reported to resist high salinity [37]. Thus, the underlying mechanism for the haloadaptation of different lipolytic enzymes remains unclear.

To some extent, the hydration characteristics of halotolerant/halophilic enzymes may extend their function to nonaqueous environments [15,43]. In this study, Est56 was stable towards the tested water-miscible organic solvents (Table 2), and moderately tolerant to some water-immiscible organic solvents (Table 3). Some halotolerant lipolytic enzymes, such as LipC12 [40], estHIJ [112], EstSP [30], Est12 [77], and H8 [81] were also reported to resist the presence of organic solvents but to different degrees. The stability towards organic solvents could broaden Est56 application in organic solvent-mediated catalytic processes, such as flavor production in food industry, synthesis of antibiotics and anti-inflammatory compounds in pharmaceutical industry, and production of pesticides for agricultural applications [113–116].

Additionally, other properties, such as the effect of metal ions, inhibitors, and detergents on Est56 activity were also studied. Est56 activity was enhanced by Ca^{2+} and Al^{3+} at 1 and 10 mM. Several esterases and lipases are also reported to be activated by Ca^{2+} ions [117,118]. However, it is rare for esterases to show an increased activity in the presence of Al^{3+} ions. Est56 activity was enhanced at a low concentration (0.1%, v/v) of nonionic detergents such as Triton X-100, Tween 20, and Tween 80, while suppressed at high concentrations (1% and 5%, v/v). This similar concentration-dependent effect of detergents on esterases was also found in other studies [40,88,119].

5. Conclusions

A functional screening of a compost metagenome yielded an esterase, showing activity and stability over a salinity range of 0–4 M. The recently reported halotolerant lipolytic enzymes (40 in total) were also summarized and used for phylogenetic analysis in this study. To explore the haloadaptation of halotolerant lipolytic enzymes, their amino acid compositions were statistically compared with halophilic counterparts. However, no clear pattern was found in the amino acid composition in the halotolerant lipolytic enzymes. For Est56, the excessive content of acidic residues over lysine residues, as well as the predominantly negatively charged surface indicated that it applies a haloadaptation similar to that of halophilic enzymes. In addition, Est56 exhibits a tolerance toward various organic solvents and enhanced activity in the presence of Ca^{2+} and Al^{3+} ions and a low concentration (0.1%, v/v) of nonionic detergents. Thus, Est56 is a novel biocatalyst with application potential, particularly under high salinity and in nonaqueous environments.

Supplementary Materials: The following are available online at https://www.mdpi.com/2073-4425/12/1/122/s1, Figure S1: Genetic organization of the insert harboring *est56* and BLAST search results for the predicted open reading frames (ORFs); Figure S2: Multiple sequence alignment of Est56 and its homologs; Figure S3: SDS-PAGE analysis of the purification of recombinant Est56 (including His_6-tag); Figure S4: Substrate specificity of Est56 towards *p*-NP esters of different chain

length; Figure S5: Unrooted phylogenetic tree of family IV esterases using neighbor-joining method; Table S1: Features of characterized halotolerant lipolytic enzymes; Table S2: Lipolytic enzymes derived from halophilic archaea adapting the "salt in" strategy; Table S3: Characterized halophilic enzymes from other studies; Table S4: Purification of recombinant Est56; Table S5: Effect of metal ions and ethylenediaminetetraacetic acid (EDTA) on Est56 activity; Table S6: Effect of detergents on Est56 activity; Table S7: Effect of inhibitors on Est56 activity; Table S8: Amino acid composition comparison between halotolerant and halophilic enzymes.

Author Contributions: R.D. developed the concept and supervised the work. M.L. performed microbiological and molecular experiments and analysis of the sequences. All authors interpreted the results, contributed to the writing of the final version of the manuscript, and approved submission. All authors have read and agreed to the published version of the manuscript.

Funding: We acknowledge the support of Mingji Lu by "Erasmus Mundus Action 2—Lotus I Project." We acknowledge support by the Open Access Publication Funds of the University of Göttingen.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All datasets are publicly available stored in or derived from NCBI databases (https://www.ncbi.nlm.nih.gov). The amino acid sequence of Est56 is available in the GenBank database under accession number KR149569.1. The compost metagenome sequences are available in the NCBI sequence read archive (SRA) under the accession number SRR13115019.

Acknowledgments: We thank Silja Brady and Mechthild Bömeke for providing technical assistance. We also thank Sarah Zachmann for assistance.

Conflicts of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Yadav, A.N.; Saxena, A.K. Biodiversity and biotechnological applications of halophilic microbes for sustainable agriculture. *J. Appl. Biol. Biotechnol.* **2018**, *6*, 48–55. [CrossRef]
2. Dassarma, S.; Dassarma, P. Halophiles and their enzymes: Negativity put to good use. *Curr. Opin. Microbiol.* **2015**, *25*, 120–126. [CrossRef] [PubMed]
3. Salma, M.; Abdulla, M.K.; Samina, M. Osmoadaptation in halophilic bacteria and archaea. *Res. J. Biotechnol.* **2020**, *15*, 154–161.
4. Siglioccolo, A.; Paiardini, A.; Piscitelli, M.; Pascarella, S. Structural adaptation of extreme halophilic proteins through decrease of conserved hydrophobic contact surface. *BMC Struct. Biol.* **2011**, *11*, 50. [CrossRef] [PubMed]
5. Ishibashi, M.; Hayashi, T.; Yoshida, C.; Tokunaga, M. Increase of salt dependence of halophilic nucleoside diphosphate kinase caused by a single amino acid substitution. *Extremophiles* **2013**, *17*, 585–591. [CrossRef] [PubMed]
6. Moshfegh, M.; Shahverdi, A.R.; Zarrini, G.; Faramarzi, M.A. Biochemical characterization of an extracellular polyextremophilic α-amylase from the halophilic archaeon *Halorubrum xinjiangense*. *Extremophiles* **2013**, *17*, 677–687. [CrossRef] [PubMed]
7. Cycil, L.M.; DasSarma, S.; Pecher, W.; McDonald, R.; AbdulSalam, M.; Hasan, F. Metagenomic insights into the diversity of halophilic microorganisms indigenous to the Karak Salt Mine, Pakistan. *Front. Microbiol.* **2020**, *11*, 1567. [CrossRef]
8. Oren, A. Bioenergetic aspects of halophilism. *Microbiol. Mol. Biol. Rev.* **1999**, *63*, 334–348. [CrossRef]
9. Ventosa, A.; Nieto, J.J.; Oren, A. Biology of moderately halophilic aerobic bacteria. *Microbiol. Mol. Biol. Rev.* **1998**, *62*, 504–544. [CrossRef]
10. Corral, P.; Amoozegar, M.A.; Ventosa, A. Halophiles and their biomolecules: Recent advances and future applications in biomedicine. *Mar. Drugs* **2020**, *18*, 33. [CrossRef]
11. Jiang, H.; Zhang, S.; Gao, H.; Hu, N. Characterization of a cold-active esterase from *Serratia* sp. and improvement of thermostability by directed evolution. *BMC Biotechnol.* **2016**, *16*. [CrossRef] [PubMed]
12. Bai, W.; Xue, Y.; Zhou, C.; Ma, Y. Cloning, expression, and characterization of a novel alkali-tolerant xylanase from alkaliphilic *Bacillus* sp. SN5. *Biotechnol. Lett.* **2012**, *34*, 2093–2099. [CrossRef] [PubMed]
13. Kumar, S.; Paul, D.; Bhushan, B.; Wakchaure, G.C.; Meena, K.K.; Shouche, Y. Traversing the "Omic" landscape of microbial halotolerance for key molecular processes and new insights. *Crit. Rev. Microbiol.* **2020**, *46*, 631–653. [CrossRef] [PubMed]
14. Delgado-García, M.; Valdivia-Urdiales, B.; Aguilar-González, C.N.; Contreras-Esquivel, J.C.; Rodríguez-Herrera, R. Halophilic hydrolases as a new tool for the biotechnological industries. *J. Sci. Food Agric.* **2012**, *92*, 2575–2580. [CrossRef] [PubMed]
15. Munawar, N.; Engel, P.C. Halophilic enzymes: Characteristics, structural adaptation and potential applications for biocatalysis. *Curr. Biotechnol.* **2013**, *2*, 334–344. [CrossRef]
16. Ortega, G.; Laín, A.; Tadeo, X.; López-Méndez, B.; Castaño, D.; Millet, O. Halophilic enzyme activation induced by salts. *Sci. Rep.* **2011**, *1*, 6. [CrossRef]

17. Edbeib, M.F.; Aksoy, H.M.; Kaya, Y.; Wahab, R.A.; Huyop, F. Haloadaptation: Insights from comparative modeling studies between halotolerant and non-halotolerant dehalogenases. *J. Biomol. Struct. Dyn.* **2020**, *38*, 3452–3461. [CrossRef]
18. Sinha, R.; Khare, S.K. Protective role of salt in catalysis and maintaining structure of halophilic proteins against denaturation. *Front. Microbiol.* **2014**, *5*, 165. [CrossRef]
19. Vogler, M.; Karan, R.; Renn, D.; Vancea, A.; Vielberg, M.-T.; Grötzinger, S.W.; DasSarma, P.; DasSarma, S.; Eppinger, J.; Groll, M.; et al. Crystal structure and active site engineering of a halophilic γ-carbonic anhydrase. *Front. Microbiol.* **2020**, *11*, 742. [CrossRef]
20. Bardavid, R.E.; Oren, A. The amino acid composition of proteins from anaerobic halophilic bacteria of the order *Halanaerobiales*. *Extremophiles* **2012**, *16*, 567–572. [CrossRef]
21. Arpigny, J.L.; Jaeger, K.-E. Bacterial lipolytic enzymes: Classification and properties. *Biochem. J.* **1999**, *343*, 177–183. [CrossRef] [PubMed]
22. Bornscheuer, U.T. Microbial carboxyl esterases: Classification, properties and application in biocatalysis. *FEMS Microbiol. Rev.* **2002**, *26*, 73–81. [CrossRef] [PubMed]
23. Kovacic, F.; Babic, N.; Krauss, U.; Jaeger, K. Classification of lipolytic enzymes from bacteria. In *Aerobic Utilization of Hydrocarbons, Oils and Lipids. Handbook of Hydrocarbon and Lipid Microbiology*; Rojo, F., Ed.; Springer: Cham, Switzerland, 2019; pp. 1–35, ISBN 9783319397825.
24. Rahman, M.A.; Culsum, U.; Tang, W.; Zhang, S.W.; Wu, G.; Liu, Z. Characterization of a novel cold active and salt tolerant esterase from *Zunongwangia profunda*. *Enzyme Microb. Technol.* **2016**, *85*, 1–11. [CrossRef] [PubMed]
25. Lee, J.H.; Jeon, J.H.; Kim, J.T.; Lee, H.S.; Kim, S.J.; Kang, S.G.; Choi, S.H. Novel lipolytic enzymes identified from metagenomic library of deep-sea sediment. *Evid. Based Complement. Altern. Med.* **2011**, *2011*. [CrossRef]
26. Hu, Y.; Fu, C.; Huang, Y.; Yin, Y.; Cheng, G.; Lei, F.; Lu, N.; Li, J.; Ashforth, E.J.; Zhang, L.; et al. Novel lipolytic genes from the microbial metagenomic library of the South China Sea marine sediment. *FEMS Microbiol. Ecol.* **2010**, *72*, 228–237. [CrossRef] [PubMed]
27. Fang, Z.; Li, J.; Wang, Q.; Fang, W.; Peng, H.; Zhang, X.; Xiao, Y. A novel esterase from a marine metagenomic library exhibiting salt tolerance ability. *J. Microbiol. Biotechnol.* **2014**, *24*, 771–780. [CrossRef] [PubMed]
28. Asyari, M.; Aditiawati, P.; Akhmaloka; Hertadi, R. Cloning and sequence analysis of lipase gene of halophilic bacteria isolated from mud crater of Bledug Kuwu, Central Java, Indonesia. *Biosci. Biotechnol. Res. Asia* **2015**, *12*, 1903–1912. [CrossRef]
29. Jeon, J.H.; Lee, H.S.; Kim, J.T.; Kim, S.J.; Choi, S.H.; Kang, S.G.; Lee, J.H. Identification of a new subfamily of salt-tolerant esterases from a metagenomic library of tidal flat sediment. *Appl. Microbiol. Biotechnol.* **2012**, *93*, 623–631. [CrossRef]
30. Jayanath, G.; Mohandas, S.P.; Kachiprath, B.; Solomon, S.; Sajeevan, T.P.; Singh, I.S.B.; Philip, R. A novel solvent tolerant esterase of GDSGG motif subfamily from solar saltern through metagenomic approach: Recombinant expression and characterization. *Int. J. Biol. Macromol.* **2018**, *119*, 393–401. [CrossRef]
31. Li, P.-Y.; Zhang, Y.; Xie, B.-B.; Zhang, Y.-Q.; Hao, J.; Wang, Y.; Wang, P.; Li, C.-Y.; Qin, Q.-L.; Zhang, X.-Y.; et al. Structural and mechanistic insights into the improvement of the halotolerance of a marine microbial esterase by increasing intra- and interdomain hydrophobic interactions. *Appl. Environ. Microbiol.* **2017**, *83*, e01286-17. [CrossRef]
32. Mandrich, L.; de Pascale, D. An overview on thermal adaptation of esterases and lipases belonging to the HSL family: New insight on the computational analysis. *Curr. Chem. Biol.* **2011**, *5*, 17–28. [CrossRef]
33. Mandrich, L.; Menchise, V.; Alterio, V.; De Simone, G.; Pedone, C.; Rossi, M.; Manco, G. Functional and structural features of the oxyanion hole in a thermophilic esterase from *Alicyclobacillus acidocaldarius*. *Proteins Struct. Funct. Genet.* **2008**, *71*, 1721–1731. [CrossRef]
34. Li, P.Y.; Ji, P.; Li, C.Y.; Zhang, Y.; Wang, G.L.; Zhang, X.Y.; Xie, B.B.; Qin, Q.L.; Chen, X.L.; Zhou, B.C.; et al. Structural basis for dimerization and catalysis of a novel sterase from the GTSAG motif subfamily of the bacterial hormone-sensitive lipase family. *J. Biol. Chem.* **2014**, *289*, 19031–19041. [CrossRef] [PubMed]
35. Chahinian, H.; Ali, Y.B.; Abousalham, A.; Petry, S.; Mandrich, L.; Manco, G.; Canaan, S.; Sarda, L. Substrate specificity and kinetic properties of enzymes belonging to the hormone-sensitive lipase family: Comparison with non-lipolytic and lipolytic carboxylesterases. *Biochim. Biophys. Acta* **2005**, *1738*, 29–36. [CrossRef]
36. Rao, L.; Zhao, X.; Li, Y.; Xue, Y.; Ma, Y.; Lu, J.R. Solution behavior and activity of a halophilic esterase under high salt concentration. *PLoS ONE* **2009**, *4*, e6980. [CrossRef] [PubMed]
37. Alcaide, M.; Stogios, P.J.; Lafraya, Á.; Tchigvintsev, A.; Flick, R.; Bargiela, R.; Chernikova, T.N.; Reva, O.N.; Hai, T.; Leggewie, C.C.; et al. Pressure adaptation is linked to thermal adaptation in salt-saturated marine habitats. *Environ. Microbiol.* **2015**, *17*, 332–345. [CrossRef] [PubMed]
38. Ai, L.; Huang, Y.; Wang, C. Purification and characterization of halophilic lipase of *Chromohalobacter* sp. from ancient salt well. *J. Basic Microbiol.* **2018**, *58*, 647–657. [CrossRef] [PubMed]
39. Castilla, A.; Panizza, P.; Rodríguez, D.; Bonino, L.; Díaz, P.; Irazoqui, G.; Giordano, S.R. A novel thermophilic and halophilic esterase from *Janibacter* sp. R02, the first member of a new lipase family (Family XVII). *Enzyme Microb. Technol.* **2017**, *98*, 86–95. [CrossRef] [PubMed]
40. Glogauer, A.; Martini, V.P.; Faoro, H.; Couto, G.H.; Müller-Santos, M.; Monteiro, R.A.; Mitchell, D.A.; de Souza, E.M.; Pedrosa, F.O.; Krieger, N. Identification and characterization of a new true lipase isolated through metagenomic approach. *Microb. Cell Fact.* **2011**, *10*, 54. [CrossRef]

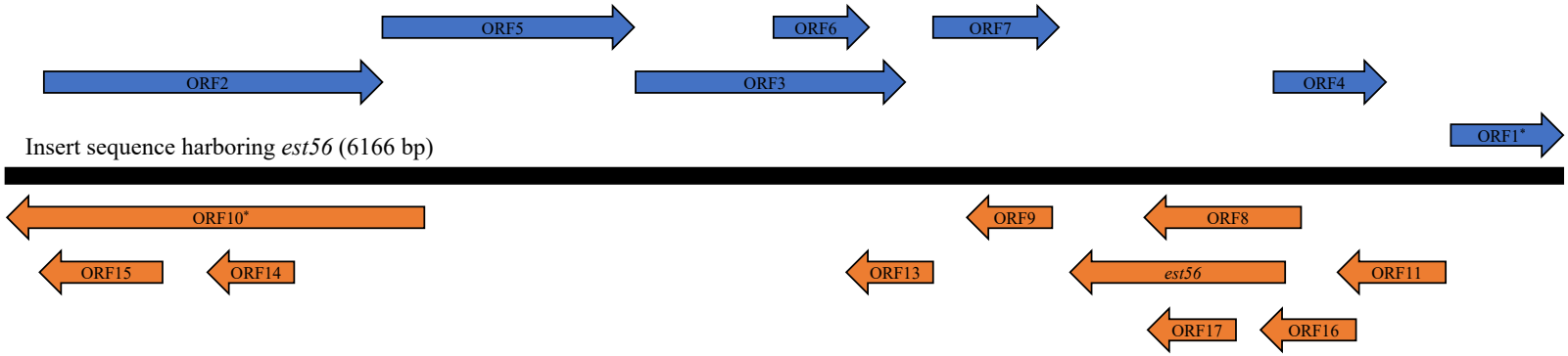
41. Mohamed, Y.M.; Ghazy, M.A.; Sayed, A.; Ouf, A.; El-Dorry, H.; Siam, R. Isolation and characterization of a heavy metal-resistant, thermophilic esterase from a Red Sea Brine Pool. *Sci. Rep.* **2013**, *3*, 3358. [CrossRef]
42. Borchert, E.; Selvin, J.; Kiran, S.G.; Jackson, S.A.; O'Gara, F.; Dobson, A.D.W. A novel cold active esterase from a deep sea sponge *Stelletta normani* metagenomic library. *Front. Mar. Sci.* **2017**, *4*, 1–13. [CrossRef]
43. Marhuenda-Egea, F.C.; Bonete, M.J. Extreme halophilic enzymes in organic solvents. *Curr. Opin. Biotechnol.* **2002**, *13*, 385–389. [CrossRef]
44. Pérez, D.; Martín, S.; Fernández-Lorente, G.; Filice, M.; Guisán, J.M.; Ventosa, A.; García, M.T.; Mellado, E. A novel halophilic lipase, LipBL, showing high efficiency in the production of eicosapentaenoic acid (EPA). *PLoS ONE* **2011**, *6*, e23325. [CrossRef]
45. Oren, A. Industrial and environmental applications of halophilic microorganisms. *Environ. Technol.* **2010**, *31*, 825–834. [CrossRef] [PubMed]
46. Ryckeboer, J.; Mergaert, J.; Vaes, K.; Klammer, S.; Clercq, D.; Coosemans, J.; Insam, H.; Swings, J. A survey of bacteria and fungi occurring during composting and self-heating processes. *Ann. Microbiol.* **2003**, *53*, 349–410.
47. Kang, C.H.; Oh, K.H.; Lee, M.H.; Oh, T.K.; Kim, B.H.; Yoon, J.H. A novel family VII esterase with industrial potential from compost metagenomic library. *Microb. Cell Fact.* **2011**, *10*, 41. [CrossRef]
48. Ilmberger, N.; Meske, D.; Juergensen, J.; Schulte, M.; Barthen, P.; Rabausch, U.; Angelov, A.; Mientus, M.; Liebl, W.; Schmitz, R.A.; et al. Metagenomic cellulases highly tolerant towards the presence of ionic liquids—Linking thermostability and halotolerance. *Appl. Microbiol. Biotechnol.* **2012**, *95*, 135–146. [CrossRef] [PubMed]
49. Dougherty, M.J.; D'haeseleer, P.; Hazen, T.C.; Simmons, B.A.; Adams, P.D.; Hadi, M.Z. Glycoside hydrolases from a targeted compost metagenome, activity-screening and functional characterization. *BMC Biotechnol.* **2012**, *12*, 38. [CrossRef]
50. Lu, M.; Dukunde, A.; Daniel, R. Biochemical profiles of two thermostable and organic solvent–tolerant esterases derived from a compost metagenome. *Appl. Microbiol. Biotechnol.* **2019**, *103*, 3421–3437. [CrossRef]
51. Ishikawa, J.; Hotta, K. FramePlot: A new implementation of the frame analysis for predicting protein-coding regions in bacterial DNA with a high G + C content. *FEMS Microbiol. Lett.* **1999**, *174*, 251–263. [CrossRef]
52. Bendtsen, J.D.; Nielsen, H.; Von Heijne, G.; Brunak, S. Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.* **2004**, *340*, 783–795. [CrossRef] [PubMed]
53. Ye, J.; McGinnis, S.; Madden, T.L. BLAST: Improvements for better sequence analysis. *Nucleic Acids Res.* **2006**, *34*, W6–W9. [CrossRef] [PubMed]
54. Zhang, Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinform.* **2008**, *9*, 40. [CrossRef] [PubMed]
55. Robert, X.; Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **2014**, *42*, 320–324. [CrossRef] [PubMed]
56. Bradford, M.M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **1976**, *72*, 248–254. [CrossRef]
57. Lineweaver, H.; Burk, D. The determination of enzyme dissociation constants. *J. Am. Chem. Soc.* **1934**, *56*, 658–666. [CrossRef]
58. Fitter, J.; Haber-Pohlmeier, S. Structural stability and unfolding properties of thermostable bacterial α-amylases: A comparative study of homologous enzymes. *Biochemistry* **2004**, *43*, 9589–9599. [CrossRef]
59. Tamura, K.; Stecher, G.; Peterson, D.; Filipski, A.; Kumar, S. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* **2013**, *30*, 2725–2729. [CrossRef]
60. Felsenstein, J. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* **1985**, *39*, 783–791. [CrossRef]
61. Asnicar, F.; Weingart, G.; Tickle, T.L.; Huttenhower, C.; Segata, N. Compact graphical representation of phylogenetic data and metadata with GraPhlAn. *PeerJ* **2015**, *3*, e1029. [CrossRef]
62. Loukas, A.; Kappas, I.; Abatzopoulos, T.J. HaloDom: A new database of halophiles across all life domains. *J. Biol. Res.* **2018**, *25*, 2. [CrossRef] [PubMed]
63. Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, M.R.; Appel, R.D.; Bairoch, A. Protein identification and analysis tools on the ExPASy server. In *The Proteomics Protocols Handbook*; Humana Press: Totowa, NJ, USA, 2005; pp. 571–607.
64. Oksanen, J.; Blanchet, F.G.; Friendly, M.; Kindt, R.; Legendre, P.; Mcglinn, D.; Minchin, P.R.; O'hara, R.B.; Simpson, G.L.; Solymos, P.; et al. Vegan: Community Ecology Package. Available online: https://cran.r-project.org/web/packages/vegan/vegan.pdf (accessed on 30 October 2018).
65. Baker, N.A.; Sept, D.; Joseph, S.; Holst, M.J.; McCammon, J.A. Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 10037–10041. [CrossRef]
66. Lee, S.W.; Won, K.; Lim, H.K.; Kim, J.C.; Choi, G.J.; Cho, K.Y. Screening for novel lipolytic enzymes from uncultured soil microorganisms. *Appl. Microbiol. Biotechnol.* **2004**, *65*, 720–726. [CrossRef]
67. Jin, P.; Pei, X.; Du, P.; Yin, X.; Xiong, X.; Wu, H.; Zhou, X.; Wang, Q. Overexpression and characterization of a new organic solvent-tolerant esterase derived from soil metagenomic DNA. *Bioresour. Technol.* **2012**, *116*, 234–240. [CrossRef] [PubMed]
68. Dukunde, A.; Schneider, D.; Lu, M.; Brady, S.; Daniel, R. A novel, versatile family IV carboxylesterase exhibits high stability and activity in a broad pH spectrum. *Biotechnol. Lett.* **2017**, *39*, 577–587. [CrossRef] [PubMed]
69. Li, J.; Zhang, K.; Han, W. Cloning and biochemical characterization of a novel lipolytic gene from activated sludge metagenome, and its gene product. *Microb. Cell Fact.* **2010**, *9*, 83. [CrossRef]
70. Lindgren, M.; Westlund, P.-O. On the stability of chymotrypsin inhibitor 2 in a 10 M urea solution. The role of interaction energies for urea-induced protein denaturation. *Phys. Chem. Chem. Phys.* **2010**, *12*, 9358. [CrossRef]

71. Peng, Q.; Zhang, X.; Shang, M.; Wang, X.; Wang, G.; Li, B.; Guan, G.; Li, Y.; Wang, Y. A novel esterase gene cloned from a metagenomic library from neritic sediments of the South China Sea. *Microb. Cell Fact.* **2011**, *10*, 95. [CrossRef]
72. Esteban-Torres, M.; Santamaría, L.; de las Rivas, B.; Muñoz, R. Characterisation of a cold-active and salt-tolerant esterase from *Lactobacillus plantarum* with potential application during cheese ripening. *Int. Dairy J.* **2014**, *39*, 312–315. [CrossRef]
73. Li, P.; Chen, X.; Ji, P.; Li, C.; Wang, P.; Zhang, Y.; Xie, B.; Qin, Q.; Su, H.; Zhou, B.; et al. Interdomain hydrophobic interactions modulate the thermostability of microbial esterases from the hormone-sensitive lipase family. *J. Biol. Chem.* **2015**, *290*, 11188–11198. [CrossRef]
74. Khan, M.; Sathya, T.A. Extremozymes from metagenome: Potential applications in food processing. *Crit. Rev. Food Sci. Nutr.* **2018**, *58*, 2017–2025. [CrossRef] [PubMed]
75. Simon, C.; Daniel, R. Metagenomic analyses: Past and future trends. *Appl. Environ. Microbiol.* **2011**, *77*, 1153–1161. [CrossRef] [PubMed]
76. Huo, Y.Y.; Rong, Z.; Jian, S.L.; Xu, C.D.; Li, J.; Xu, X.W. A novel halotolerant thermoalkaliphilic esterase from marine bacterium *Erythrobacter seohaensis* SW-135. *Front. Microbiol.* **2017**, *8*, 2315. [CrossRef]
77. Wu, G.; Zhang, S.; Zhang, H.; Zhang, S.; Liu, Z. A novel esterase from a psychrotrophic bacterium *Psychrobacter celer* 3Pb1 showed cold-adaptation and salt-tolerance. *J. Mol. Catal. B Enzym.* **2013**, *98*, 119–126. [CrossRef]
78. Wang, G.; Wang, Q.; Lin, X.; Ng, T.B.; Yan, R.; Lin, J.; Ye, X. A novel cold-adapted and highly salt-tolerant esterase from *Alkalibacterium* sp. SL3 from the sediment of a soda lake. *Sci. Rep.* **2016**, *6*, 19494. [CrossRef] [PubMed]
79. Tchigvintsev, A.; Tran, H.; Popovic, A.; Kovacic, F.; Brown, G.; Flick, R.; Hajighasemi, M.; Egorova, O.; Somody, J.C.; Tchigvintsev, D.; et al. The environment shapes microbial enzymes: Five cold-active and salt-resistant carboxylesterases from marine metagenomes. *Appl. Microbiol. Biotechnol.* **2015**, *99*, 2165–2178. [CrossRef] [PubMed]
80. Selvin, J.; Kennedy, J.; Lejon, D.P.H.; Kiran, G.S.; Dobson, A.D.W. Isolation identification and biochemical characterization of a novel halo-tolerant lipase from the metagenome of the marine sponge *Haliclona simulans*. *Microb. Cell Fact.* **2012**, *11*, 72. [CrossRef]
81. Zhang, Y.; Hao, J.; Zhang, Y.Q.; Chen, X.L.; Xie, B.B.; Shi, M.; Zhou, B.C.; Zhang, Y.Z.; Li, P.Y. Identification and characterization of a novel salt-tolerant esterase from the deep-sea sediment of the South China Sea. *Front. Microbiol.* **2017**, *8*, 441. [CrossRef]
82. Alcaide, M.; Tchigvintsev, A.; Martínez-Martínez, M.; Popovic, A.; Reva, O.N.; Lafraya, Á.; Bargiela, R.; Nechitaylo, T.Y.; Matesanz, R.; Cambon-Bonavita, M.-A.; et al. Identification and characterization of carboxyl esterases of gill chamber-associated microbiota in the deep-sea shrimp *Rimicaris exoculata* by using functional metagenomics. *Appl. Environ. Microbiol.* **2015**, *81*, 2125–2136. [CrossRef]
83. Oliveira, L.C.G.; Ramos, P.L.; Marem, A.; Kondo, M.Y.; Rocha, R.C.S.; Bertolini, T.; Silveira, M.A.; da Cruz, J.B.; de Vasconcellos, S.P.; Juliano, L.; et al. Halotolerant bacteria in the São Paulo Zoo composting process and their hydrolases and bioproducts. *Braz. J. Microbiol.* **2015**, *46*, 347–354. [CrossRef]
84. Chandna, P.; Mayilraj, S.; Kuhad, R.C. *Bacillus pseudoflexus* sp. nov., a moderately halophilic bacterium isolated from compost. *Ann. Microbiol.* **2016**, *66*, 895–905. [CrossRef]
85. De Santi, C.; Leiros, H.K.S.; Di Scala, A.; de Pascale, D.; Altermark, B.; Willassen, N.P. Biochemical characterization and structural analysis of a new cold-active and salt-tolerant esterase from the marine bacterium *Thalassospira* sp. *Extremophiles* **2016**, *20*, 323–336. [CrossRef] [PubMed]
86. Jiang, X.; Huo, Y.; Cheng, H.; Zhang, X.; Zhu, X.; Wu, M. Cloning, expression and characterization of a halotolerant esterase from a marine bacterium *Pelagibacterium halotolerans* B2T. *Extremophiles* **2012**, *16*, 427–435. [CrossRef] [PubMed]
87. Leite, A.E.T.; Briganti, L.; de Araújo, E.A.; Pellegrini, V.D.; Camilo, C.M.; Polikarpov, I. Low-resolution molecular shape, biochemical characterization and emulsification properties of a halotolerant esterase from *Bacillus licheniformis*. *Eur. Biophys. J.* **2020**, *49*, 435–447. [CrossRef]
88. Wu, G.; Wu, G.; Zhan, T.; Shao, Z.; Liu, Z. Characterization of a cold-adapted and salt-tolerant esterase from a psychrotrophic bacterium *Psychrobacter pacificensis*. *Extremophiles* **2013**, *17*, 809–819. [CrossRef]
89. Adıgüzel, A.O. Production and characterization of thermo-, halo- and solvent-stable esterase from *Bacillus mojavensis* TH309. *Biocatal. Biotransform.* **2020**, *38*, 210–226. [CrossRef]
90. De Santi, C.; Altermark, B.; Pierechod, M.M.; Ambrosino, L.; de Pascale, D.; Willassen, N.-P. Characterization of a cold-active and salt tolerant esterase identified by functional screening of Arctic metagenomic libraries. *BMC Biochem.* **2016**, *17*. [CrossRef] [PubMed]
91. Zhang, W.; Xu, H.; Wu, Y.; Zeng, J.; Guo, Z.; Wang, L.; Shen, C.; Qiao, D.; Cao, Y. A new cold-adapted, alkali-stable and highly salt-tolerant esterase from *Bacillus licheniformis*. *Int. J. Biol. Macromol.* **2018**, *111*, 1183–1193. [CrossRef]
92. Hang, Y.; Ran, S.; Wang, X.; Jiao, J.; Wang, S.; Liu, Z. Mutational analysis and stability characterization of a novel esterase of lipolytic enzyme family VI from *Shewanella* sp. *Int. J. Biol. Macromol.* **2016**, *93*, 655–664. [CrossRef]
93. Lee, C.W.; Kwon, S.; Park, S.H.; Kim, B.Y.; Yoo, W.; Ryu, B.H.; Kim, H.W.; Shin, S.C.; Kim, S.; Park, H.; et al. Crystal structure and functional characterization of an esterase (EaEST) from *Exiguobacterium antarcticum*. *PLoS ONE* **2017**, *12*, e0169540. [CrossRef]
94. Yang, X.; Wu, L.; Xu, Y.; Ke, C.; Hu, F.; Xiao, X.; Huang, J. Identification and characterization of a novel alkalistable and salt-tolerant esterase from the deep-sea hydrothermal vent of the East Pacific Rise. *Microbiology* **2018**, *7*, e00601. [CrossRef]
95. Dennis, P.P.; Shimmin, L.C. Evolutionary divergence and salinity-mediated selection in halophilic archaea. *Microbiol. Mol. Biol. Rev.* **1997**, *61*, 90–104. [CrossRef] [PubMed]

96. Margesin, R.; Schinner, F. Potential of halotolerant and halophilic microorganisms for biotechnology. *Extremophiles* **2001**, *5*, 73–83. [CrossRef] [PubMed]
97. Lanyi, J.K. Salt-dependent properties of proteins from extremely halophilic bacteria. *Bacteriol. Rev.* **1974**, *38*, 272–290. [CrossRef]
98. Madern, D.; Ebel, C.; Zaccai, G. Halophilic adaptation of enzymes. *Extremophiles* **2000**, *4*, 91–98. [CrossRef]
99. Hutcheon, G.W.; Vasisht, N.; Bolhuis, A. Characterisation of a highly stable α-amylase from the halophilic archaeon *Haloarcula hispanica*. *Extremophiles* **2005**, *9*, 487–495. [CrossRef] [PubMed]
100. Kastritis, P.L.; Papandreou, N.C.; Hamodrakas, S.J. Haloadaptation: Insights from comparative modeling studies of halophilic archaeal DHFRs. *Int. J. Biol. Macromol.* **2007**, *41*, 447–453. [CrossRef]
101. Madern, D.; Pfister, C.; Zaccai, G. Mutation at a single acidic amino acid enhances the halophilic behaviour of malate dehydrogenase from *Haloarcula marismortui* in physiological salts. *Eur. J. Biochem.* **1995**, *230*, 1088–1095. [CrossRef]
102. Paul, S.; Bag, S.K.; Das, S.; Harvill, E.T.; Dutta, C. Molecular signature of hypersaline adaptation: Insights from genome and proteome composition of halophilic prokaryotes. *Genome Biol.* **2008**, *9*, R70. [CrossRef]
103. Winter, J.A.; Christofi, P.; Morroll, S.; Bunting, K.A. The crystal structure of *Haloferax volcanii* proliferating cell nuclear antigen reveals unique surface charge characteristics due to halophilic adaptation. *BMC Struct. Biol.* **2009**, *9*, 1–15. [CrossRef]
104. Tadeo, X.; Ló Pez-Mé Ndez, B.; Trigueros, T.; Laín, A.; Castañ, D.; Millet, O. Structural basis for the amino acid composition of proteins from halophilic Archea. *PLoS Biol.* **2009**, *7*, e1000257. [CrossRef] [PubMed]
105. Munawar, N.; Engel, P.C. Overexpression in a non-native halophilic host and biotechnological potential of NAD+-dependent glutamate dehydrogenase from *Halobacterium salinarum* strain NRC-36014. *Extremophiles* **2012**, *16*, 463–476. [CrossRef] [PubMed]
106. Coquelle, N.; Talon, R.; Juers, D.H.; Girard, É.; Kahn, R.; Madern, D. Gradual adaptive changes of a protein facing high salt concentrations. *J. Mol. Biol.* **2010**, *404*, 493–505. [CrossRef] [PubMed]
107. Oren, A.; Gavrieli, I.; Gavrieli, J.; Kohen, M.; Lati, J.; Aharoni, M. Microbial communities in the Dead Sea—Past, present and future. In *Adaptation to Life at High Salt Concentrations in Archaea, Bacteria, and Eukarya. Cellular Origin, Life in Extreme Habitats and Astrobiology*; Gunde-Cimerman, N., Oren, A., Plemenitaš, A., Eds.; Springer: Berlin/Heidelberg, Germany, 2005; Volume 9, pp. 27–39, ISBN 1402036337.
108. Bieger, B.; Essen, L.-O.; Oesterhelt, D. Crystal structure of halophilic dodecin: A novel, dodecameric flavin binding protein from *Halobacterium salinarum*. *Structure* **2003**, *11*, 375–385. [CrossRef]
109. Altermark, B.; Helland, R.; Moe, E.; Willassen, N.P.; Smalås, A.O. Structural adaptation of endonuclease I from the cold-adapted and halophilic bacterium *Vibrio salmonicida*. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2008**, *64*, 368–376. [CrossRef] [PubMed]
110. Sivakumar, N.; Li, N.; Tang, J.W.; Patel, B.K.C.; Swaminathan, K. Crystal structure of AmyA lacks acidic surface and provide insights into protein stability at poly-extreme condition. *FEBS Lett.* **2006**, *580*, 2646–2652. [CrossRef]
111. Kim, T.D. Bacterial hormone-sensitive lipases (bHSLs): Emerging enzymes for biotechnological applications. *J. Microbiol. Biotechnol.* **2017**, *27*, 1907–1915. [CrossRef] [PubMed]
112. Noby, N.; Hussein, A.; Saeed, H.; Embaby, A.M. Recombinant cold-adapted halotolerant, organic solvent-stable esterase (estHIJ) from *Bacillus halodurans*. *Anal. Biochem.* **2020**, *591*, 113554. [CrossRef]
113. Panda, T.; Gowrishankar, B.S. Production and applications of esterases. *Appl. Microbiol. Biotechnol.* **2005**, *67*, 160–169. [CrossRef]
114. Hasan, F.; Shah, A.A.; Hameed, A. Industrial applications of microbial lipases. *Enzyme Microb. Technol.* **2006**, *39*, 235–251. [CrossRef]
115. Karan, R.; Capes, M.D.; DasSarma, S. Function and biotechnology of extremophilic enzymes in low water activity. *Aquat. Biosyst.* **2012**, *8*, 4. [CrossRef] [PubMed]
116. Lopez-Lopez, O.; Cerdan, M.; Siso, M. New extremophilic lipases and esterases from metagenomics. *Curr. Protein Pept. Sci.* **2014**, *15*, 445–455. [CrossRef] [PubMed]
117. Rathi, P.; Bradoo, S.; Saxena, R.K.; Gupta, R. A hyper-thermostable, alkaline lipase from *Pseudomonas* sp. with the property of thermal activation. *Biotechnol. Lett.* **2000**, *22*, 495–498. [CrossRef]
118. Castro-Ochoa, L.D.; Rodríguez-Gómez, C.; Valerio-Alfaro, G.; Oliart Ros, R. Screening, purification and characterization of the thermoalkalophilic lipase produced by *Bacillus thermoleovorans* CCR11. *Enzyme Microb. Technol.* **2005**, *37*, 648–654. [CrossRef]
119. Kamarudin, N.H.A.; Rahman, R.N.Z.R.A.; Ali, M.S.M.; Leow, T.C.; Basri, M.; Salleh, A.B. A new cold-adapted, organic solvent stable lipase from mesophilic *Staphylococcus epidermidis* AT2. *Protein J.* **2014**, *33*, 296–307. [CrossRef] [PubMed]

Supplemental information for chapter IV

Content



ORF	Location/Strand (+)	Putative product
ORF1	5725-6165	Hypothetical protein [Chloroflexi bacterium]
ORF2	140-1480	Homoserine dehydrogenase [Tepidiforma bonchosmolovskayae]
ORF3	2483-3559	Threonine synthase [Dehalococcoidia bacterium]
ORF4	5027-5476	No BLAST hit found
ORF5	1482-2486	Hypothetical protein [Chloroflexi bacterium]
ORF6	3036-3422	Pyridoxal-phosphate dependent enzyme [Streptomyces sp. SID7803]
ORF7	3675-4178	Hypothetical protein [Actinobacteria bacterium]

ORF	Location/Strand (-)	Putative product
ORF8	5137-4505	No BLAST hit found
ORF9	4138-3803	No BLAST hit found
ORF10	1654-2	Hypothetical protein [Candidatus Accumulibacter sp. BA-91]
ORF11	5706-5266	No BLAST hit found
est56	5070-4207	Lipase/esterase [uncultured bacterium]
ORF13	3681-3328	No BLAST hit found
ORF14	1152-799	No BLAST hit found
ORF15	615-130	Unknown [Phascolarctobacterium sp. CAG:266]
ORF16	5339-4971	No BLAST hit found
ORF17	4874-4527	No BLAST hit found

Figure S1. Strand, location and BLAST searching results for predicted open reading frames (ORFs) on insert sequence that harboring lipolytic gene *est56*. Only amino acid sequence length ≥ 100 aa were mentioned. The ORF prediction was conducted by ORFfinder (https://www.ncbi.nlm.nih.gov/orffinder/). The closest hit for each ORF was searched using BLASTP. * Partial ORFs.

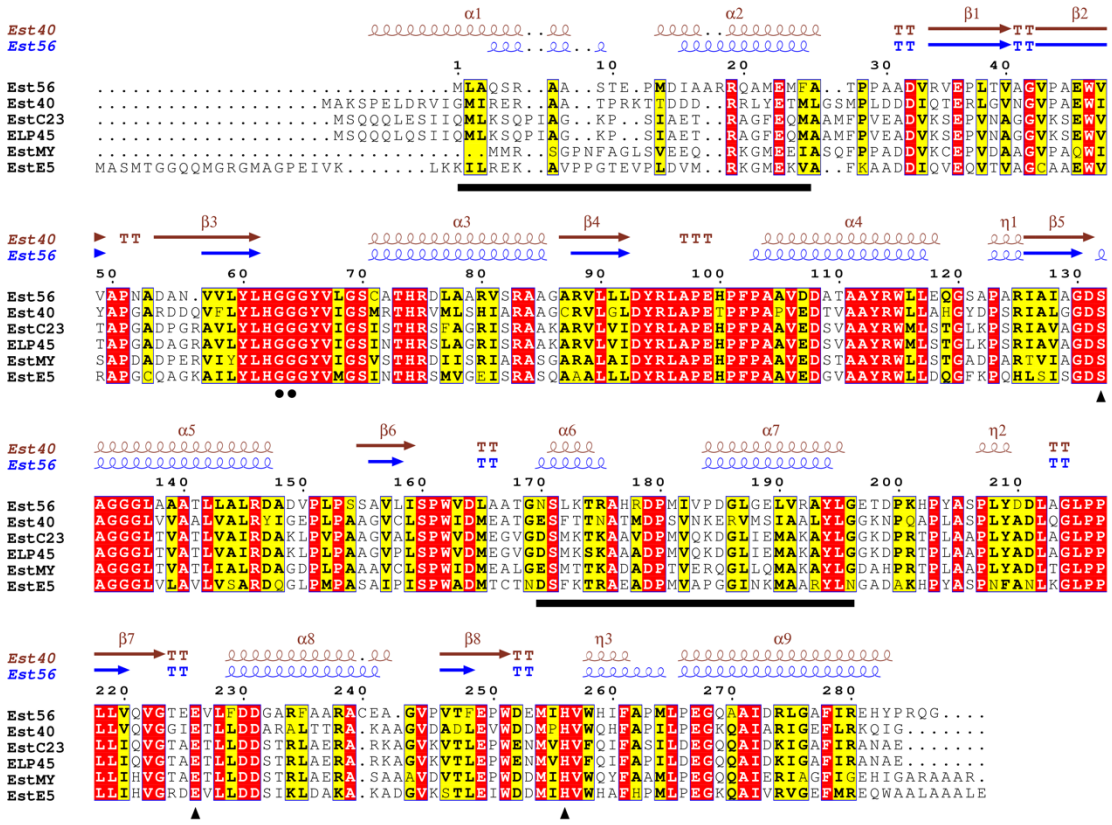


Figure S2. Multiple sequence alignment of Est56 and its homologs. Partially conserved residues are in frames. Identical residues are shaded. Triangles underneath residues indicate the catalytic triad and circles represent residues involved in the oxyanion hole. The secondary structures of Est56 and its structural analog Est40 (Li et al. 2015b) are presented as: squiggles for α helices, arrows for β strands, by TT letters for turns, and η letters for 3_{10}-helices. The square bar represents regions of the cap domain. The reference esterases EstC23 (Jin et al. 2012), ELP45 (Lee et al. 2004), EstMY (Li et al. 2010) and EstE5 (Nam et al. 2009) were derived from GenBank.

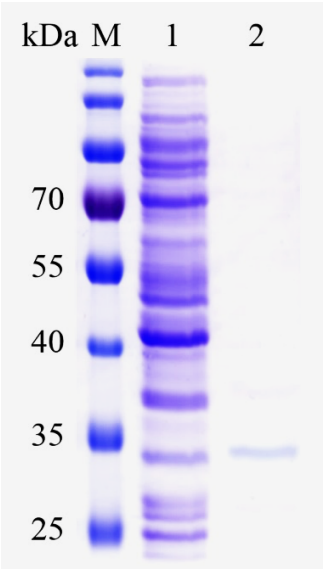


Figure S3. SDS-PAGE analysis of purification of recombinant Est56 (including His_6-tag). Lane M, standard molecular weight marker; Lane 1, cell lysate (21.5 μg); Lane 2, purified Est56 (1.0 μg).

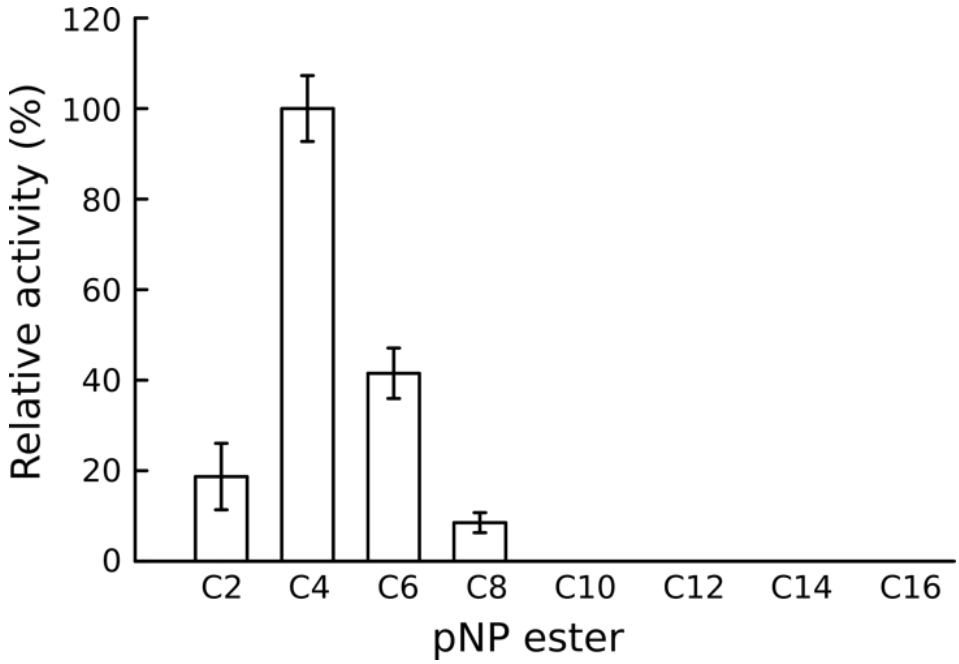


Figure S4. Substrate specificity of Est56 towards p-NP esters of different chain length. The maximal activity (189.5 U/mg) measured with p-NP butyrate (C4) was taken as 100%.

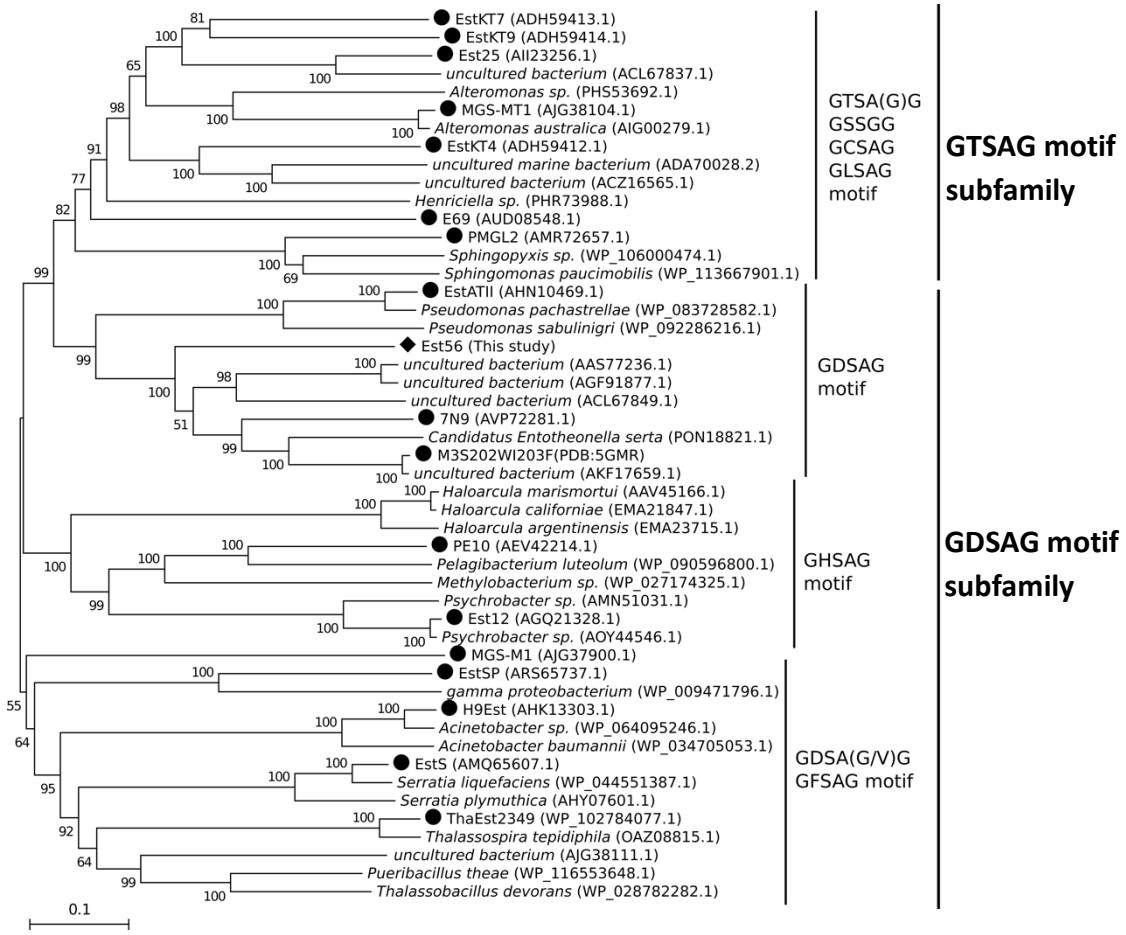


Figure S5. Unrooted phylogenetic tree of family IV esterases using neighbor-joining method. Est56 (closed diamond) and characterized halotolerant (closed circles) are depicted. With the exception of Est56, other sequences were retrieved from GenBank, with accession numbers in parentheses. Only bootstrap values greater than 50% are shown. Scale represents the number of amino acid substitutions per site.

Table S1. Features of characterized halotolerant lipolytic enzymes (HT_LIP)

Lipolytic enzyme	Organism	pI	pH_{opt}	T_{opt} (°C)	Salinity					Reference
					Effect on enzyme activity			Effect on enzyme stability		
					Salt range	Maximum activity (%)[a]	Minimum activity (%)[a]	Incubation condition	Residual activity (%)[b]	
7N9	Uncultured bacterium	4.59	8	0-30	0-24 %	100 % at 0 %	~ 55 % at 24 %	ND[d]	ND[d]	(Borchert et al. 2017)
Lpc53E1	Uncultured bacterium	4.61	7	40	0-4 M	234 % at 5 M	100 % at 0 M	4 °C, 24 h	~ 100 %, over 0-5 M	(Selvin et al. 2012)
EM3L4	Uncultured bacterium	4.61	7.5	35	0-4 M	ND[d]	ND[d]	35 °C, 30 min	> 100 %, over 0-4 M	(Lee et al. 2011)
PE10	*Pelagibacterium halotolerans*	4.65	7.5	45	0-4 M	~ 160 % at 3 M	100 % at 0 M	ND[d]	ND[d]	(Jiang et al. 2012)
EstSP	Uncultured bacterium	4.65	8	40	0-5 M	155 % at 1 M	~ 10 % at 5 M	25 °C, 24 h	> 70 %, over 0-5 M	(Jayanath et al. 2018)
ABO_1251	*Alcanivorax borkumensis*	4.74	ND[d]	35	0-3.5 M	100 % at 0 M	~ 40 % at 3.5 M NaCl/KCl	ND[d]	ND[d]	(Tchigvintsev et al.
E69	*Erythrobacter seohaensis*	4.76	10.5	60	0-3 M	~ 150 % at 0.5 M	~ 40 % at 3 M	ND[d]	ND[d]	(Huo et al. 2017)
MGS-K1	Uncultured bacterium	4.89	7	30	0-4 M	0.8 M NaCl[c]	ND[d]	ND[d]	ND[d]	(Alcaide et al. 2015a)
estHIJ	*Bacillus halodurans*	4.90	7	28	0-4 M	100% at 0 M	~ 70 % at 4 M	25 °C, 6 h	~ 100 % at 4 M	(Noby et al. 2020)
EaEST	*Exiguobacterium antarcticum*	4.91	8	40	0-5 M	ND[d]	ND[d]	25 °C, 1 h	~ 100 %, over 0-5 M	(Lee et al. 2017)
Esth	*Shewanella*	4.93	8	30	0-5 M	100 % at 0 M	~ 50 % at 5 M	4 °C, 24 h	~ 100 %, over 0-5 M	(Hang et al. 2016)
ABO_1197	*Alcanivorax borkumensis*	4.93	ND[d]	30	0-3.5 M	100 % at 0 M	~ 30 % at 3.5 M NaCl/KCl	ND[d]	ND[d]	(Tchigvintsev et al.
ThaEst2349	*Thalassospira* sp.	4.94	8.5	45	0-4 M	283 % at 3 M	40 % at 4 M	4 °C, 24 h	>280 %, over 1-3 M	(De Santi et al. 2016b)
Est56	Uncultured bacterium	4.97	8	50	0-4 M	~ 140 % at 1.5 M NaCl or 1 M KCl	~ 40 % at 4M NaCl/KCl	10 °C, 24 h	~ 100 %, over 0-4 M NaCl/KCl	**This study**
Lip3	Uncultured bacterium	4.98	8	35	0-4 M	675 % at 3 M NaCl	100 % at 0 M	4 °C, 24 h	> 100 %, over 0-3 M	(De Santi et al. 2016a)
EstS	*Serratia* sp.	5.05	8.5	10	0-4 M	100 % at 0 M	94 % at 4 M	4 °C, 24 h	> 80 %, over 0-4 M	(Jiang et al. 2016)
MGS-RG1	Uncultured bacterium	5.07	8	45	0-4 M	~ 250 % at 3.2 M	100 % at 0 M	ND[d]	ND[d]	(Alcaide et al. 2015b)
E25	Uncultured bacterium	5.16	8.5	50	0-4 M	~ 130 % at 1 M	~ 50 % at 4 M	ND[d]	ND[d]	(Li et al. 2014)
M3S202WI 203F	Uncultured bacterium	5.17	ND[d]	55	0-4 M	~ 110 % at 0.5 M	~ 60 % at 4 M	20 °C, 1 h	> 70 %, over 0-4 M	(Li et al. 2017a)
YbfF	*Halomonas elongata*	5.25	8	ND[d]	0-4 M	100 % at 2 M	~15 % at 0 M	ND[d]	ND[d]	(Yoo et al. 2020)
EstSL3	*Alkalibacterium* sp.	5.28	9	30	0-4 M	~ 105 % at 2 M	98 % at 4 M	37 °C, 2 h	~ 100 %, over 0-4 M	(Wang et al. 2016b)
MGS-RG2	Uncultured bacterium	5.31	8	50	0-4 M	~ 250 % at 3.6 M	100 % at 0 M	ND[d]	ND[d]	(Alcaide et al. 2015b)
BlEst1	*Bacillus licheniformis*	5.33	7	40	0-5 M	100 % at 1 M	~ 70 % at 5 M	ND[d]	ND[d]	(Nakamura et al. 2018)
Est10	*Psychrobacter pacificensis*	5.35	7.5	25	0-5 M	143.2 % at 2 M	~ 80 % at 5 M	4 °C, 6.5 h	> 100 %, over 0-5 M	(Wu et al. 2012a)
MGS0010	Uncultured bacterium	5.4	ND[d]	30	0-3.5 M	~ 250 % at 3.5 M	100 % at 0 M	ND[d]	ND[d]	(Tchigvintsev et al.
BlEstA	*Bacillus licheniformis*	5.54	9	30	0-3 M	165 % at 2 M	100 % at 0 M	ND[d]	ND[d]	(Leite et al. 2020)

EstWSD	Uncultured bacterium	5.61	7	50	0-5 M	~ 140 % at 1 M	~ 50 % at 5 M	ND[d]	ND[d]	(Wang et al. 2012)
EstSHJ2	*Chromohalobacter canadensis*	5.71	8	50	0-5 M	100 % at 2.5 M	~10 % at 0 M	ND[d]	ND[d]	(Wang et al. 2020)
PMGL2	Uncultured bacterium	5.72	8.5	45	0-1.75 M	165 % at 0.25 M	84 % at 1.75 M	ND[d]	ND[d]	(Petrovskay et al.
MGS-M1	Uncultured bacterium	5.77	8	25	0-4 M	3.6 M NaCl[c]	ND[d]	ND[d]	ND[d]	(Alcaide et al. 2015a)
EstKT4	Uncultured bacterium	5.81	8.5	40	0-4 M	ND[d]	ND[d]	35 °C, 30 min	> 50%, over 0-3.5 M	(Jeon et al. 2012)
EstKT7	Uncultured bacterium	5.84	8	35	0-4 M	ND[d]	ND[d]	35 °C, 30 min	> 50%, over 0-3 M	(Jeon et al. 2012)
LipC12	Uncultured bacterium	5.98	9	30	0-4 M	1501 % at 1.5 M	100 % at 0 M	4 °C, 24 h	~ 100 %, over 0-3.7 M	(Glogauer et al. 2011)
EstKT9	Uncultured bacterium	6.1	8.5	45	0-4 M	ND[d]	ND[d]	35 °C, 30 min	> 50%, over 0-3.5 M	(Jeon et al. 2012)
lp_3505	*Lactobacillus plantarum*	6.12	6	5	0-25 %	~ 250 % at 5 %	~ 70 % at 25 %	ND[d]	ND[d]	(Esteban-Torres et al.
Est9x	Uncultured bacterium	6.17	8	65	0-4 M	~ 190 % at 4 M	100 % at 0 M	ND[d]	ND[d]	(Fang et al. 2014)
Est12	*Psychrobacter celer*	6.5	7.5	35	0-4.5 M	ND[d]	ND[d]	25 °C, 12 h	~ 100 %, over 0-4.5 M	(Wu et al. 2013b)
EstATII	Uncultured bacterium	7.11	8.5	65	0-4 M	ND[d]	~ 50 % at 4 M	ND[d]	ND[d]	(Mohamed et al. 2013)
Est-OKK	Uncultured bacterium	7.82	9	50	0-3 M	~ 130 % at 1.5 M	~ 100 % at 3 M	RT, 4 h	~ 100 %, over 0-3 M	(Yang et al. 2018)
LipJ2	*Janibacter* sp.	8.25	9	80	1, 10 mM	~ 250 % at 0.1 mM NaCl	100 % at 0 M	ND[d]	ND[d]	(Castilla et al. 2017)
MGS-MT1	Uncultured bacterium	8.4	8.5	50	0-4 M		ND[d]	ND[d]	ND[d]	(Alcaide et al. 2015a)
EstLiu	*Zunongwangia profunda*	8.42	8	30	0-4.5 M	100 % at 0 M	57 % at 4.5 M	4 °C, 12 h	> 80 %, over 0-4.5 M	(Rahman et al. 2016)
H9Est	Uncultured bacterium	8.72	8	40	0-2.5 M	~ 150 % at 1 M	~ 50 % at 2.5 M	ND[d]	ND[d]	(Santi et al. 2015)
H8	Uncultured bacterium	9.09	10	35	0-5 M	~ 105 % at 4 M	~ 10 % at 5 M	0 °C, 1 h	> 80 %, over 0-4.7 M	(Zhang et al. 2017)
Est700	*Bacillus licheniformis*	9.44	8	30	0-5 M	588 % at 3.5 M	100 % at 0 M	4 °C, 1 h	~ 100 %, over 0-5 M	(Zhang et al. 2018)
BmEST	*Bacillus mojavensis*	ND[d]	8	80	0-25 %	~ 300 % at 20 %	100 % at 0 %	80 °C, 1 h	> 60 %, over 0-25 %	(Adıgüzel 2020)

[a] The activity measured without salt (NaCl, if not mentioned) was taken as 100 %

[b] The activity measured after incubating in salt-free (NaCl, unless stated otherwise) buffer was defined as 100%

[c] The activity at $Salt_{opt}$ was set as 100 %

[d] No data

Table S2. Lipolytic enzymes (HP_Lip) derived from halophilic archaea adapting the "salt in" strategy

Lipolytic enzyme	Length (aa)	pI	Microorganism	GeneBank/PDB Accession Nr.	Family	Halophile	Growth condition (NaCl, %)
HP_Lip_1	285	4.45	*Halococcus thailandensis*	EMA51434.1	*Halobacteriaceae*	Extreme	20% - 30%
HP_Lip_2	292	4.55	*Halococcus morrhuae*	EMA45705.1	*Halobacteriaceae*	Extreme	>12%, 23.3%
HP_Lip_3	320	4.16	*Halococcus saccharolyticus*	EMA45019.1	*Halobacteriaceae*	Extreme	15% - satr, optimum 25%
HP_Lip_4	285	4.5	*Halococcus hamelinensis*	EMA39292.1	*Halobacteriaceae*	Moderate	12.5% - 30%, optimum 15%
HP_Lip_5	263	4.4	*Halosimplex carlsbadense*	ELZ28160.1	*Halobacteriaceae*	Extreme	20% - 30% (optimum 25%)
HP_Lip_6	261	4.54	*Halosimplex carlsbadense*	ELZ24957.1	*Halobacteriaceae*	Extreme	20% - 30% (optimum 25%)
HP_Lip_7	330	4.17	*Haladaptatus sp.*	KZN24148.1	*Halobacteriaceae*	Extreme	ND[b]
HP_Lip_8	333	4.32	*Halorhabdus utahensis*	ACV11819.1	*Halobacteriaceae*	Extreme	9% - 30%, optimum 27%
HP_Lip_9	340	4.34	*Halorhabdus utahensis*	ACV10409.1	*Halobacteriaceae*	Extreme	9% - 30%, optimum 27%
HP_Lip_10	258	4.57	*Haloarcula marismortui*	AAV45777.1	*Halobacteriaceae*	Extreme	optimum 20% - 23%
HP_Lip_11	318	4.22	*Haloarcula hispanica*	AHB65276.1	*Halobacteriaceae*	Moderate	>12%
HP_Lip_12	377	4.32	*Haloarcula japonica*	WP_004591147.1	*Halobacteriaceae*	Extreme	15% - 30%, optimum 20%
HP_Lip_13	318	4.18	*Haloarcula japonica*	EMA29911.1	*Halobacteriaceae*	Extreme	15% - 30%, optimum 20%
HP_Lip_14	318	4.19	*Haloarcula vallismortis*	EMA07756.1	*Halobacteriaceae*	Extreme	>15%, optimum 25%
HP_Lip_15	376	4.3	*Haloarcula vallismortis*	WP_004515030.1	*Halobacteriaceae*	Extreme	>15%, optimum 25%
HP_Lip_16	314	4.23	*Haloprofundus marisrubri*	KTG11548.1	*Haloferacaceae*	Extreme	min 7%
HP_Lip_17	260	4.35	*Haloprofundus marisrubri*	KTG08679.1	*Haloferacaceae*	Extreme	min 7%
HP_Lip_18	217	4.44	*Halogeometricum pallidum*	ELZ32922.1	*Haloferacaceae*	Moderate	optimum 18.1%
HP_Lip_19	275	4.48	*Halogeometricum pallidum*	ELZ27142.1	*Haloferacaceae*	Moderate	optimum 18.1%
HP_Lip_20	215	4.53	*Halogeometricum borinquense*	ELY30686.1	*Haloferacaceae*	Extreme	min 8%, optimum 20% - 25%
HP_Lip_21	344	4.27	*Natronolimnobius baerhuensis*	OVE85190.1	*Natrialbaceae*	Extreme	optimum 20%
HP_Lip_22	455	4.54	*Haloterrigena mahii*	OAQ52820.1	*Natrialbaceae*	Extreme	optimum 20.4% - 29.2%

Table S3. Characterized halophilic enzymes (HP_Enz) from other studies

Protein	Length (aa)	pI	Organism	Domain	Family	Halophile	Growth condition (NaCl, %)	Reference
Malate dehydrogenase	303	4.2	*Haloarcula marismortui*	Archaea	*Halobacteriaceae*	Extreme	10% - 30%, optimum 20% - 23%	(Richard et al. 2000)
Malate dehydrogenase	304	4.2	*Haloarcula marismortui*	Archaea	*Halobacteriaceae*	Extreme	10% - 30%, optimum 20% - 23%	(Fioravanti et al. 2007)
Catalase-peroxidase	731	4.32	*Haloarcula marismortui*	Archaea	*Halobacteriaceae*	Extreme	10% - 30%, optimum 20% - 23%	(Yamada et al. 2002)
Esterase	327	4.24	*Haloarcula marismortui*	Archaea	*Halobacteriaceae*	Extreme	10% - 30%, optimum 20% - 23%	(Rao et al. 2009a)
Nucleoside diphosphate kinase	161	4.37	*Halobacterium salinarum*	Archaea	*Halobacteriaceae*	Moderate	>12%, NaCl saturation	(Besir et al. 2005)
Nucleoside diphosphate kinase	164	4.42	*Halobacterium salinarum*	Archaea	*Halobacteriaceae*	Moderate	>12%, NaCl saturation	(Besir et al. 2005)
RNase H1	199	4.36	*Halobacterium salinarum*	Archaea	*Halobacteriaceae*	Moderate	>12%, NaCl saturation	(Tannous and
Phosphatase	431	4.35	*Halobacterium salinarum*	Archaea	*Halobacteriaceae*	Moderate	>12%, NaCl saturation	(Wende et al. 2010)
Dehydrogenase	435	4.39	*Halobacterium salinarum*	Archaea	*Halobacteriaceae*	Moderate	>12%, NaCl saturation	(Munawar and Engel
DNA protecting protein	182	4.3	*Halobacterium salinarum*	Archaea	*Halobacteriaceae*	Moderate	>12%, NaCl saturation	(Zeth et al. 2004)
Beta-galactosidase	663	4.54	*Haloferax lucentense*	Archaea	*Haloferacaceae*	Extreme	10.5% - 29.8% (optimum 25.1%)	(Holmes et al. 1997)
DNA ligase	699	4.34	*Haloferax volcanii*	Archaea	*Haloferacaceae*	Extreme	6% - 29%, optimum 10% - 15%	(Poidevin and
Dihydrofolate reductase	162	4.45	*Haloferax volcanii*	Archaea	*Haloferacaceae*	Extreme	6% - 29%, optimum 10% - 15%	(Pieper et al. 1998)
Glucose dehydrogenase	357	4.55	*Haloferax mediterranei*	Archaea	*Haloferacaceae*	Moderate	>12%	(Britton et al. 2006)
Alpha-amylase	504	4.11	*Natronococcus sp.*	Archaea	*Natrialbaceae*	Moderate	>12%	(Kobayashi et al. 1994)
Esterase	316	4.38	unclutured bacterium	Bacterium	ND[b]	ND[b]	ND[b]	(Alcaide et al. 2015a)

[a] no data

Table S4. Purification of recombinant Est56.

Purification step	Total activity (U)	Total protein (mg)	Specific activity (U/mg)	Purification (fold)	Yield (%)
Crude extract	239.8	193.7	1.24	1.0	100
Ni-TED	108.8	1.2	90.44	73.0	45.4

Table S5. Effect of metal ions and EDTA on Est56 activity

Metal ions	Concentration (mM)	Relative activity (%) [a]
Al^{3+}	1	127.5±3.0
Al^{3+}	10	142.8±6.3
Ca^{2+}	1	126.5±10.3
Ca^{2+}	10	136.9±9.3
Mg^{2+}	1	106.1±0.9
Mg^{2+}	10	105.5±5.6
Mn^{2+}	1	113.2±2.5
Mn^{2+}	10	82.4±3.5
Zn^{2+}	1	75.4±8.2
$Zn2^{+}$	10	83.6±7.3
Fe^{2+}	1	82.4±6.6
Fe^{2+}	10	13.2±1.9
Fe^{3+}	1	60.9±5.0
Fe^{3+}	10	19.5±6.3
Ni^{2+}	1	13.3±4.1
Ni^{2+}	10	22.0±4.4
Cu^{2+}	1	16.7±8.7
Cu^{2+}	10	23.2±2.2
EDTA	1	119.4±2.5
EDTA	10	93.8±10.5

[a] The effects of metal ions on Est56 activity were measured under standard assay conditions. Specific activity (93.1 U/mg) assayed without additive was taken as 100%.

[b] Not detectable.

Table S6. Effect of detergents on Est56 activity

Detergent	Concentration (v/v, %)	Relative activity (%) [a]
Triton X-100	0.1	156.8±3.4
Triton X-100	1	44.6±2.8
Triton X-100	5	18.3±3.6
Tween 20	0.1	154.3±2.3
Tween 20	1	53.6±3.0
Tween 20	5	ND[b]
Tween 80	0.1	112.7±1.7
Tween 80	1	36.6±1.8
Tween 80	5	11.4±1.6

[a] The effects of detergents on Est56 activity were measured under standard assay conditions. Specific activity (93.1 U/mg) assayed without additive was taken as 100%.

[b] Not detectable.

Table S7. Effect of inhibitors on Est56 activity

Inhibitor	Concentration (mM)	Relative activity (%) [a]
DTT	1	102.2±9.3
DTT	10	46.0±1.4
PMSF	1	69.9±0.7
PMSF	10	10.1±1.2
DEPC	1	1.9±1.1
DEPC	10	ND[b]

[a] The effects of inhibitors on Est56 activity were measured under standard assay conditions. Specific activity (93.1 U/mg) assayed without additive was taken as 100%.

[b] Not detectable.

Table S8. Amino acid composition comparison between halotolerant and halophilic enzymes

Amino acid	HT (%)[a]	HP_Lip (%)[a]	HP_Enz (%)[a]	SIMPER analysis[b]		
				HT vs HP_Lip Contribution (%)[c]	HT vs HP_Enz Contribution (%)[d]	HP_Lip vs HP_Enz Contribution (%)[e]
Asp (D)	5.92±1.39	9.87±1.39	10.47±1.99	10.44 (1)	11.58 (1)	6.32 (5)
Ala (A)	10.68±3.19	12.55±2.07	11.36±2.79	8.76 (2)	8.40 (3)	8.73 (1)
Lys (K)	3.74±2.30	1.01±0.48	2.16±1.26	7.34 (3)	5.77 (5)	4.24 (14)
Glu (E)	5.27±1.47	7.86±1.59	8.87±1.98	7.13 (4)	9.26 (2)	6.97 (3)
Ile (I)	5.14±1.97	2.80±1.16	3.86±1.16	6.75 (5)	4.92 (8)	5.20 (9)
Gly (G)	8.42±1.80	9.22±1.89	8.68±1.80	5.59 (6)	5.04 (7)	6.86 (4)
Leu (L)	9.82±1.92	8.30±1.25	7.57±1.19	5.57 (7)	6.59 (4)	4.87 (10)
Asn (N)	3.51±1.56	1.66±0.94	2.49±1.07	5.41 (8)	4.12 (13)	4.34 (12)
Thr (T)	4.93±1.61	5.41±1.61	5.56±1.50	4.71 (9)	4.3 (12)	5.39 (6)
Arg (R)	4.73±1.72	5.60±1.09	5.70±1.31	4.68 (10)	4.67 (11)	4.27 (13)
Pro (P)	5.98±1.50	6.11±1.70	4.83±1.93	4.59 (11)	5.31 (6)	7.393 (2)
Ser (S)	5.93±1.68	5.21±1.33	4.88±1.62	4.39 (12)	4.77 (10)	5.371 (7)
Val (V)	7.52±1.64	7.90±1.13	8.07±1.71	4.2 (13)	4.83 (9)	5.27 (8)
Gln (Q)	3.57±1.29	2.55±0.94	2.79±0.85	3.72 (14)	3.28 (17)	3.34 (17)
Met (M)	2.88±0.83	1.54±0.73	1.86±1.00	3.72 (15)	3.36 (16)	3.26 (18)
His (H)	2.70±1.10	3.58±0.79	2.94±1.45	3.37 (16)	3.51 (14)	4.80 (11)
Phe (F)	3.61±1.21	3.65±0.81	3.13±1.20	3.03 (17)	3.45 (15)	3.94 (15)
Tyr (Y)	3.34±0.98	3.24±1.05	2.91±0.87	2.9 (18)	2.73 (18)	3.47 (16)
Trp (W)	1.54±0.76	1.40±0.63	1.46±0.99	2.02 (19)	2.46 (19)	2.99 (19)
Cys (C)	0.76±0.65	0.72±0.51	0.41±0.47	1.68 (20)	1.65 (20)	1.99 (20)

[a] Data shown are averages with the standard deviation per group.

[b] SIMPER analysis the contribution of each amino acid to the dissimilarity between groups, with permutations 9999. Data shown are contributions (%) with the rank in the brackets.

[c] SIMPER-revealed average dissimilarity was 19.38.

[d] SIMPER-revealed average dissimilarity was 20.03.

[e] SIMPER-revealed average dissimilarity was 15.31.

References

Abd Rahman RNZR, Ahmad Kamarudin NH, Yunus J, Salleh AB, Basri M (2010) Expression of an organic solvent stable lipase from Staphylococcus epidermidis AT2. Int J Mol Sci 11:3195–3208 . doi: 10.3390/ijms11093195

Abdel-Fattah YR, Gaballa AA (2008) Identification and over-expression of a thermostable lipase from Geobacillus thermoleovorans Toshki in Escherichia coli. Microbiol Res 163:13–20 . doi: 10.1016/j.micres.2006.02.004

Adıgüzel AO (2020) Production and characterization of thermo-, halo- and solvent-stable esterase from Bacillus mojavensis TH309. Biocatal Biotransformation 38:210–226 . doi: 10.1080/10242422.2020.1715370

Adlercreutz P (2013) Immobilisation and application of lipases in organic media. Chem Soc Rev 42:6406–6436 . doi: 10.1039/c3cs35446f

Ahmed EH, Raghavendra T, Madamwar D (2010) An alkaline lipase from organic solvent tolerant Acinetobacter sp. EH28: Application for ethyl caprylate synthesis. Bioresour Technol 101:3628–3634 . doi: 10.1016/J.BIORTECH.2009.12.107

Ai L, Huang Y, Wang C (2018) Purification and characterization of halophilic lipase of Chromohalobacter sp . from ancient salt well. J Basic Microbiol 58:647–657 . doi: 10.1002/jobm.201800116

Akoh CC, Lee GC, Liaw YC, Huang TH, Shaw JF (2004) GDSL family of serine esterases/lipases. Prog Lipid Res 43:534–552 . doi: 10.1016/j.plipres.2004.09.002

Albers S V., Jonuscheit M, Dinkelaker S, Urich T, Kletzin A, Tampé R, Driessen AJM, Schleper C (2006) Production of recombinant and tagged proteins in the hyperthermophilic archaeon Sulfolobus solfataricus. Appl Environ Microbiol 72:102–111 . doi: 10.1128/AEM.72.1.102-111.2006

Alcaide M, Stogios PJ, Lafraya Á, Tchigvintsev A, Flick R, Bargiela R, Chernikova TN, Reva ON, Hai T, Leggewie CC, Katzke N, La Cono V, Matesanz R, Jebbar M, Jaeger K-E, Yakimov MM, Yakunin AF, Golyshin PN, Golyshina O V., Savchenko A, Ferrer M, MAMBA Consortium (2015a) Pressure adaptation is linked to thermal adaptation in salt-saturated marine habitats. Environ Microbiol 17:332–345 . doi: 10.1111/1462-2920.12660

Alcaide M, Tchigvintsev A, Martínez-Martínez M, Popovic A, Reva ON, Lafraya Á, Bargiela R, Nechitaylo TY, Matesanz R, Cambon-Bonavita M-A, Jebbar M, Yakimov MM, Savchenko A, Golyshina O V, Yakunin AF, Golyshin PN, Ferrer M, MAMBA Consortium (2015b) Identification and characterization of carboxyl esterases of gill chamber-associated microbiota in the deep-sea shrimp Rimicaris exoculata by using functional metagenomics. Appl Environ Microbiol 81:2125–2136 . doi: 10.1128/AEM.03387-14

Alcaide M, Tornés J, Stogios PJ, Xu X, Gertler C, Di Leo R, Bargiela R, Lafraya Á, Guazzaroni M-E, López-Cortés N, Chernikova TN, Golyshina OV, Nechitaylo TY, Plumeier I, Pieper DH, Yakimov MM, Savchenko A, Golyshin PN, Ferrer M (2013) Single residues dictate the co-evolution of dual esterases: MCP hydrolases from the α/β hydrolase family. Biochem J 454:157–166 . doi: 10.1042/BJ20130552

Alcock BP, Raphenya AR, Lau TTY, Tsang KK, Bouchard M, Edalatmand A, Huynh W, Nguyen AL V., Cheng AA, Liu S, Min SY, Miroshnichenko A, Tran HK, Werfalli RE, Nasir JA, Oloni M, Speicher DJ, Florescu A, Singh B, Faltyn M, Hernandez-Koutoucheva A, Sharma AN, Bordeleau E, Pawlowski AC, Zubyk HL, Dooley D, Griffiths E, Maguire F, Winsor GL, Beiko RG, Brinkman FSL, Hsiao WWL, Domselaar G V., McArthur AG (2020) CARD 2020: Antibiotic resistome surveillance with the comprehensive antibiotic resistance database. Nucleic Acids Res 48:D517–D525 . doi: 10.1093/nar/gkz935

Almeida A, Mitchell AL, Boland M, Forster SC, Gloor GB, Tarkowska A, Lawley TD, Finn RD (2019) A new genomic blueprint of the human gut microbiota. Nature 568:499–504 . doi: 10.1038/s41586-019-0965-1

Altermark B, Helland R, Moe E, Willassen NP, Smalås AO (2008) Structural adaptation of endonuclease I from the cold-adapted and halophilic bacterium *Vibrio salmonicida*. Acta Crystallogr Sect D Biol Crystallogr 64:368–376 . doi: 10.1107/S0907444908000097

Andualema B, Gessesse A (2012) Microbial lipases and their industrial applications: Review. Biotechnology 11:100–118 . doi: 10.3923/biotech.2012.100.118

Angelov A, Mientus M, Liebl S, Liebl W (2009) A two-host fosmid system for functional screening of (meta)genomic libraries from extreme thermophiles. Syst Appl Microbiol 32:177–185 . doi: 10.1016/j.syapm.2008.01.003

Antunes LP, Martins LF, Pereira RV, Thomas AM, Barbosa D, Lemos LN, Silva GMM, Moura LMS, Epamino GWC, Digiampietri LA, Lombardi KC, Ramos PL, Quaggio RB, De Oliveira JCF, Pascon RC, Da Cruz JB, Da Silva AM, Setubal JC (2016) Microbial community structure and dynamics in thermophilic composting viewed through metagenomics and metatranscriptomics. Sci Rep 6:38915 . doi: 10.1038/srep38915

Armour CR, Nayfach S, Pollard KS, Sharpton TJ (2019) A metagenomic meta-analysis reveals functional signatures of health and disease in the human gut microbiome. mSystems 4:e00332-18 . doi: 10.1128/msystems.00332-18

Arpigny JL, Jaeger K-E (1999) Bacterial lipolytic enzymes: classification and properties. Biochem J 343:177–183 . doi: 10.1042/0264-6021:3430177

Awasthi MK, Duan Y, Liu T, Awasthi SK, Zhang Z (2020a) Relevance of biochar to influence the bacterial succession during pig manure composting. Bioresour Technol 304:122962 . doi: 10.1016/j.biortech.2020.122962

Awasthi MK, Ravindran B, Sarsaiya S, Chen H, Wainaina S, Singh E, Liu T, Kumar S, Pandey A, Singh L, Zhang Z (2020b) Metagenomics for taxonomy profiling : tools and approaches. Bioengineered 11:356–374 . doi: 10.1080/21655979.2020.1736238

Aylward FO, Burnum KE, Scott JJ, Suen G, Tringe SG, Adams SM, Barry KW, Nicora CD, Piehowski PD, Purvine SO, Starrett GJ, Goodwin LA, Smith RD, Lipton MS, Currie CR (2012) Metagenomic and metaproteomic insights into bacterial communities in leaf-cutter ant fungus gardens. ISME J 6:1688–1701 . doi: 10.1038/ismej.2012.10

Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS (2009) MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res 37:W202–W208 . doi: 10.1093/nar/gkp335

Barriuso J, Jesús Martínez M (2015) In silico metagenomes mining to discover novel esterases with industrial application by sequential search strategies. J Microbiol Biotechnol 25:732–737 . doi: 10.4014/jmb.1406.06049

Bassegoda A, Pastor FIJ, Diaz P (2012) Rhodococcus sp. strain CR-53 LipR, the first member of a new bacterial lipase family (family X) displaying an unusual Y-type oxyanion hole, similar to the Candida antarctica lipase clan. Appl Environ Microbiol 78:1724–1732 . doi: 10.1128/AEM.06332-11

Bayer S, Kunert A, Ballschmiter M, Greiner-Stoeffele T (2010) Indication for a new lipolytic enzyme family: Isolation and characterization of two esterases from a metagenomic library. J Mol Microbiol Biotechnol 18:181–187 . doi: 10.1159/000315459

Bell PJL, Sunna A, Gibbs MD, Curach NC, Nevalainen H, Bergquist PL (2002) Prospecting for novel lipase genes using PCR. Microbiology 148:2283–2291 . doi: 10.1099/00221287-148-8-2283

Bender JM, Bard JDi (2018) Metagenomics in pediatrics: Using a shotgun approach to diagnose infections. Curr Opin Pediatr 30:125–130 . doi: 10.1097/MOP.0000000000000577

Berini F, Casciello C, Marcone GL, Marinelli F (2017) Metagenomics: novel enzymes from non-culturable microbes. FEMS Microbiol Lett 364: . doi: 10.1093/femsle/fnx211

Berlemont R, Spee O, Delsaute M, Lara Y, Schuldes J, Simon C, Power P, Daniel R, Galleni M (2013) Novel organic solvent-tolerant esterase isolated by metagenomics: insights into the lipase/esterase classification. Rev Argent Microbiol 45:3–12

Besir H, Zeth K, Bracher A, Heider U, Ishibashi M, Tokunaga M, Oesterhelt D (2005) Structure of a halophilic nucleoside diphosphate kinase from Halobacterium salinarum. Febs Lett 579:6595–6600 . doi: 10.2210/PDB2AZ1/PDB

Bharagava RN, Purchase D, Saxena G, Mulla SI (2018) Applications of metagenomics in microbial bioremediation of pollutants: from genomics to environmental cleanup. In: Das S, Dash HR (eds) Microbial Diversity in the Genomic Era. Elsevier Inc., Oxford, UK, pp 459–477

Biver S, Vandenbol M (2013) Characterization of three new carboxylic ester hydrolases isolated by functional screening of a forest soil metagenomic library. J Ind Microbiol Biotechnol 40:191–200 . doi: 10.1007/s10295-012-1217-7

Boeva V (2016) Analysis of genomic sequence motifs for deciphering transcription factor binding and transcriptional regulation in Eukaryotic cells. Front. Genet. 7:24

Boni I V. (2006) Diverse molecular mechanisms of translation initiation in prokaryotes. Mol Biol 40:587–596 . doi: 10.1134/S002689330604011X

Bora L (2014) Purification and characterization of highly alkaline lipase from bacillus licheniformis MTCC 2465: And study of its detergent compatibility and applicability. J Surfactants Deterg 17:889–898 . doi: 10.1007/s11743-013-1517-6

Borchert E, Selvin J, Kiran SG, Jackson SA, O'Gara F, Dobson ADW (2017) A novel cold active esterase from a deep sea sponge Stelletta normani metagenomic library. Front Mar Sci 4:1–13 . doi: 10.3389/fmars.2017.00287

Bornscheuer UT (2002) Microbial carboxyl esterases: classification, properties and application in biocatalysis. FEMS Microbiol Rev 26:73–81 . doi: 10.1111/j.1574-6976.2002.tb00599.x

Brady SF (2007) Construction of soil environmental DNA cosmid libraries and screening for clones that produce biologically active small molecules. Nat Protoc 2:1297–1305 . doi: 10.1038/nprot.2007.195
Britton KL, Baker PJ, Fisher M, Ruzheinikov S, Gilmour DJ, Bonete M-J, Ferrer J, Pire C, Esclapez J, Rice DW (2006) Analysis of protein solvent interactions in glucose dehydrogenase from the extreme halophile Haloferax mediterranei. Proc Natl Acad Sci U S A 103:4846–4851 . doi: 10.1073/pnas.0508854103
Brumlik MJ, Buckley JT (1996) Identification of the catalytic triad of the lipase/acyltransferase from Aeromonas hydrophila. J Bacteriol 178:2060–2064 . doi: 10.1128/jb.178.7.2060-2064.1996
Bzhalava Z, Hultin E, Dillner J (2018) Extension of the viral ecology in humans using viral profile hidden Markov models. PLoS One 13:e0190938 . doi: 10.1371/journal.pone.0190938
Cai J, Xie Y, Song B, Wang Y, Zhang Z, Feng Y (2011) Fervidobacterium changbaicum Lip1: Identification, cloning, and characterization of the thermophilic lipase as a new member of bacterial lipase family V. Appl Microbiol Biotechnol 89:1463–1473 . doi: 10.1007/s00253-010-2971-y
Cammarota MC, Rosa DR, Duarte ICS, Saavedra NK, Varesche MBA, Zaiat M, Freire DMG (2013) The effect of enzymatic pre-hydrolysis of dairy wastewater on the granular and immobilized microbial community in anaerobic bioreactors. Environ Technol 34:417–428 . doi: 10.1080/09593330.2012.698649
Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. Nucleic Acids Res 37:233–238
Castilla A, Panizza P, Rodríguez D, Bonino L, Díaz P, Irazoqui G, Rodríguez Giordano S (2017) A novel thermophilic and halophilic esterase from Janibacter sp. R02, the first member of a new lipase family (Family XVII). Enzyme Microb Technol 98:86–95 . doi: 10.1016/J.ENZMICTEC.2016.12.010
Chandel C, Kumar A, Kanwar SS (2011) Enzymatic synthesis of butyl ferulate by silica-immobilized lipase in a non-aqueous medium. J Biomater Nanobiotechnol 2:400–408 . doi: 10.4236/jbnb.2011.24049
Charbonneau DM, Beauregard M (2013) Role of key salt bridges in thermostability of G. thermodenitrificans EstGtA2: distinctive patterns within the new bacterial lipolytic enzyme family XV. PLoS One 8:e76675 . doi: 10.1371/journal.pone.0076675
Chauhan M, Garlapati VK (2013) Production and characterization of a halo-, solvent-, thermo-tolerant alkaline lipase by staphylococcus arlettae JPBW-1, isolated from rock salt mine. Appl Biochem Biotechnol 171:1429–1443 . doi: 10.1007/s12010-013-0433-6
Chen IMA, Markowitz VM, Chu K, Palaniappan K, Szeto E, Pillay M, Ratner A, Huang J, Andersen E, Huntemann M, Varghese N, Hadjithomas M, Tennessen K, Nielsen T, Ivanova NN, Kyrpides NC (2017) IMG/M: Integrated genome and metagenome comparative data analysis system. Nucleic Acids Res 45:507–516 . doi: 10.1093/nar/gkw929
Choi JE, Kwon MA, Na HY, Hahm DH, Song JK (2013) Isolation and characterization of a metagenome-derived thermoalkaliphilic esterase with high stability over a broad pH range. Extremophiles 17:1013–1021 . doi: 10.1007/s00792-013-0583-z
Chow J, Kovacic F, Dall Antonia Y, Krauss U, Fersini F, Schmeisser C, Lauinger B, Bongen P, Pietruszka J, Schmidt M, Menyes I, Bornscheuer UT, Eckstein M, Thum O, Liese A, Mueller-Dieckmann J, Jaeger KE, Streit WR (2012) The metagenome-derived enzymes LipS and LipT increase the diversity of known lipases. PLoS One 7:e47665 . doi: 10.1371/journal.pone.0047665
Chu X, He H, Guo C, Sun B (2008) Identification of two novel esterases from a marine metagenomic library derived from South China Sea. Appl Microbiol Biotechnol 80:615–625 . doi: 10.1007/s00253-008-1566-3
Consortium TU (2018) UniProt: The universal protein knowledgebase. Nucleic Acids Res 45:D158–D169 . doi: 10.1093/nar/gkw1099
Coquelle N, Talon R, Juers DH, Girard É, Kahn R, Madern D (2010) Gradual adaptive changes of a protein facing high salt concentrations. J Mol Biol 404:493–505 . doi: 10.1016/j.jmb.2010.09.055
Coughlan LM, Cotter PD, Hill C, Alvarez-Ordóñez A (2015) Biotechnological applications of functional metagenomics in the food and pharmaceutical industries. Front Microbiol 6:672 . doi: 10.3389/fmicb.2015.00672
Craig JW, Chang FY, Kim JH, Obiajulu SC, Brady SF (2010) Expanding small-molecule functional metagenomics through parallel screening of broad-host-range cosmid environmental DNA libraries in diverse proteobacteria. Appl Environ Microbiol 76:1633–1641 . doi: 10.1128/AEM.02169-09
Dandavate V, Jinjala J, Keharia H, Madamwar D (2009) Production, partial purification and characterization of organic solvent tolerant lipase from Burkholderia multivorans V2 and its application for ester synthesis. Bioresour Technol 100:3374–3381 . doi: 10.1016/j.biortech.2009.02.011
Daniel R (2005) The metagenomics of soil. Nat Rev Microbiol 3:470–478 . doi: 10.1038/nrmicro1160

De Filippis F, Parente E, Ercolini D (2017) Metagenomics insights into food fermentations. Microb Biotechnol 10:91–102 . doi: 10.1111/1751-7915.12421

de Pascale D, Cusano AM, Autore F, Parrilli E, di Prisco G, Marino G, Tutino ML (2008) The cold-active Lip1 lipase from the Antarctic bacterium Pseudoalteromonas haloplanktis TAC125 is a member of a new bacterial lipolytic enzyme family. Extremophiles 12:311–323 . doi: 10.1007/s00792-008-0163-9

De Santi C, Altermark B, Pierechod MM, Ambrosino L, de Pascale D, Willassen N-P (2016a) Characterization of a cold-active and salt tolerant esterase identified by functional screening of Arctic metagenomic libraries. BMC Biochem 17: . doi: 10.1186/s12858-016-0057-x

De Santi C, Leiros HKS, Di Scala A, de Pascale D, Altermark B, Willassen NP (2016b) Biochemical characterization and structural analysis of a new cold-active and salt-tolerant esterase from the marine bacterium Thalassospira sp. Extremophiles 20:323–336 . doi: 10.1007/s00792-016-0824-z

DeCastro ME, Rodríguez-Belmonte E, González-Siso MI (2016) Metagenomics of thermophiles with a focus on discovery of novel thermozymes. Front Microbiol 7:1521 . doi: 10.3389/fmicb.2016.01521

Delgado-García M, Valdivia-Urdiales B, Aguilar-González CN, Contreras-Esquivel JC, Rodríguez-Herrera R (2012) Halophilic hydrolases as a new tool for the biotechnological industries. J Sci Food Agric 92:2575–2580 . doi: 10.1002/jsfa.5860

Devi SG, Fathima AA, Radha S, Arunraj R, Curtis WR, Ramya M (2015) A rapid and economical method for efficient DNA extraction from diverse soils suitable for metagenomic applications. PLoS One 10:e0132441 . doi: 10.1371/journal.pone.0132441

Donato JJ, Moe LA, Converse BJ, Smart KD, Berklein EC, McManus PS, Handelsman J (2010) Metagenomic analysis of apple orchard soil reveals antibiotic resistance genes encoding predicted bifunctional proteins. Appl Environ Microbiol 76:4396–4401 . doi: 10.1128/AEM.01763-09

Dornelas JCM, Figueiredo JEF, De Abreu CS, Lana UGP, Oliveira CA, Marriel IE (2017) Characterization and phylogenetic affiliation of actinobacteria from tropical soils with potential uses for agro-industrial processes. Genet Mol Res 16:1–16 . doi: 10.4238/gmr16039703

Dougherty MJ, D'haeseleer P, Hazen TC, Simmons BA, Adams PD, Hadi MZ (2012) Glycoside hydrolases from a targeted compost metagenome, activity-screening and functional characterization. BMC Biotechnol 12:38 . doi: 10.1186/1472-6750-12-38

Drö Ge J, Gregor I, Mchardy AC (2015) Taxator-tk: precise taxonomic assignment of metagenomes by fast approximation of evolutionary neighborhoods. Bioinformatics 31:817–824 . doi: 10.1093/bioinformatics/btu745

Dukunde A, Schneider D, Lu M, Brady S, Daniel R (2017) A novel, versatile family IV carboxylesterase exhibits high stability and activity in a broad pH spectrum. Biotechnol Lett 39:577–587 . doi: 10.1007/s10529-016-2282-1

Dulanto Chiang A, Dekker JP (2020) From the pipeline to the bedside: advances and challenges in clinical metagenomics. J Infect Dis 221:S331–S340 . doi: 10.1093/infdis/jiz151

Ebrahimie E, Ebrahimi M, Sarvestani R, Ebrahimi M (2011) Protein attributes contribute to halo-stability, bioinformatics approach. Saline Systems 7:1 . doi: 10.1186/1746-1448-7-1

Ebrahimpour A, Rahman RNZRA, Basri M, Salleh AB (2011) High level expression and characterization of a novel thermostable, organic solvent tolerant, 1,3-regioselective lipase from Geobacillus sp. strain ARM. Bioresour Technol 102:6972–6981 . doi: 10.1016/j.biortech.2011.03.083

Egelkamp R, Zimmermann T, Schneider D, Hertel R, Daniel R (2019) Impact of nitriles on bacterial communities. Front Environ Sci 7:103 . doi: 10.3389/fenvs.2019.00103

Esteban-Torres M, Santamaría L, de las Rivas B, Muñoz R (2014) Characterisation of a cold-active and salt-tolerant esterase from Lactobacillus plantarum with potential application during cheese ripening. Int Dairy J 39:312–315 . doi: 10.1016/j.idairyj.2014.08.004

Fang Y, Wang S, Liu S, Jiao Y (2015) Discovery a novel organic solvent tolerant esterase from Salinispora arenicola CNP193 through genome mining. Int J Biol Macromol 80:334–340 . doi: 10.1016/j.ijbiomac.2015.06.045

Fang Z, Li J, Wang Q, Fang W, Peng H, Zhang X, Xiao Y (2014) A novel esterase from a marine metagenomic library exhibiting salt tolerance ability. J Microbiol Biotechnol 24:771–780 . doi: 10.4014/jmb.1311.11071

Fawal N, Li Q, Savelli B, Brette M, Passaia G, Fabre M, Mathé C, Dunand C (2013) PeroxiBase: A database for large-scale evolutionary analysis of peroxidases. Nucleic Acids Res 41:441–444 . doi: 10.1093/nar/gks1083

Feng Y, Joh Y-G, Ishikawa K, Ishida H, Ando S, Yamagaki T, Nakanishi H, Cao S, Matsui I, Kosugi Y (2000) Thermophilic phospholipase A2 in the cytosolic fraction from the archaeon Pyrococcus horikoshii. J Am Oil Chem Soc 77:1147–1152 . doi: 10.1007/s11746-000-0180-8

Ferrer M, Martínez-Martínez M, Bargiela R, Streit WR, Golyshina O V., Golyshin PN (2016) Estimating the success of enzyme bioprospecting through metagenomics: Current status and future trends. Microb. Biotechnol. 9:22–34

Ferrer M, Martínez-Martínez M, Bargiela R, Streit WR, Golyshina O V, Golyshin PN (2015) Estimating the success of enzyme bioprospecting through metagenomics: current status and future trends. Microb Biotechnol 9:22–34 . doi: 10.1111/1751-7915.12309

Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J, Sonnhammer ELL, Tate J, Punta M (2014) Pfam: the protein families database. Nucleic Acids Res 42:D222–D230 . doi: 10.1093/nar/gkt1223

Fioravanti E, Vellieux FMD, Amara P, Madern D, Weik M (2007) Specific radiation damage to acidic residues and its relation to their chemical and structural environment. J Synchrotron Radiat 14:84–91 . doi: 10.1107/S0909049506038623

Fischer M, Pleiss J (2003) The Lipase Engineering Database: a navigation and analysis tool for protein families. Nucleic Acids Res 31:319–321

Forbes JD, Knox NC, Ronholm J, Pagotto F, Reimer A (2017) Metagenomics: The next culture-independent game changer. Front Microbiol 8:1069 . doi: 10.3389/fmicb.2017.01069

Fu C, Hu Y, Xie F, Guo H, Ashforth EJ, Polyak SW, Zhu B, Zhang L (2011) Molecular cloning and characterization of a new cold-active esterase from a deep-sea metagenomic library. Appl Microbiol Biotechnol 90:961–970 . doi: 10.1007/s00253-010-3079-0

Fu J, Leiros HKS, De Pascale D, Johnson KA, Blencke HM, Landfald B (2013) Functional and structural studies of a novel cold-adapted esterase from an Arctic intertidal metagenomic library. Appl Microbiol Biotechnol 97:3965–3978 . doi: 10.1007/s00253-012-4276-9

Gabor EM, Alkema WBL, Janssen DB (2004) Quantifying the accessibility of the metagenome by random expression cloning techniques. Environ Microbiol 6:879–886 . doi: 10.1111/j.1462-2920.2004.00640.x

Gaida SM, Sandoval NR, Nicolaou SA, Chen Y, Venkataramanan KP, Papoutsakis ET (2015) Expression of heterologous sigma factors enables functional screening of metagenomic and heterologous genomic libraries. Nat Commun 6:1–10 . doi: 10.1038/ncomms8045

Gao W, Wu K, Chen L, Fan H, Zhao Z, Gao B, Wang H, Wei D (2016) A novel esterase from a marine mud metagenomic library for biocatalytic synthesis of short-chain flavor esters. Microb Cell Fact 15:41 . doi: 10.1186/s12934-016-0435-5

Garlapati VK, Banerjee R (2013) Solvent-free synthesis of flavour esters through immobilized lipase mediated transesterification. Enzyme Res 2013: . doi: 10.1155/2013/367410

Garrido-Cardenas JA, Manzano-Agugliaro F (2017) The metagenomics worldwide research. Curr Genet 63:819–829 . doi: 10.1007/s00294-017-0693-8

Gaur R, Khare SK (2011) Solvent tolerant Pseudomonads as a source of novel lipases for applications in non-aqueous systems. Biocatal Biotransformation 29:161–171 . doi: 10.3109/10242422.2011.609588

Gerlt JA, Bouvier JT, Davidson DB, Imker HJ, Sadkhin B, Slater DR, Whalen KL (2015) Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST): A web tool for generating protein sequence similarity networks. Biochim Biophys Acta 1854:1019–1037 . doi: 10.1002/cncr.27633.Percutaneous

Glogauer A, Martini VP, Faoro H, Couto GH, Müller-Santos M, Monteiro RA, Mitchell DA, de Souza EM, Pedrosa FO, Krieger N (2011) Identification and characterization of a new true lipase isolated through metagenomic approach. Microb Cell Fact 10:54 . doi: 10.1186/1475-2859-10-54

Gotor-Fernández V, Brieva R, Gotor V (2006) Lipases: Useful biocatalysts for the preparation of pharmaceuticals. J Mol Catal B Enzym 40:111–120 . doi: 10.1016/j.molcatb.2006.02.010

Guerrand D (2017) Lipases industrial applications: Focus on food and agroindustries. OCL 24:D403 . doi: 10.1051/ocl/2017031

Guo J, Quensen JF, Sun Y, Wang Q, Brown CT, Cole JR, Tiedje JM (2019) Review, evaluation, and directions for gene-targeted assembly for ecological analyses of metagenomes. Front Genet 10:957 . doi: 10.3389/fgene.2019.00957

Handrick R, Reinhardt S, Focarete ML, Scandola M, Adamus G, Kowalczuk M, Jendrossek D (2001) A new type of thermoalkalophilic hydrolase of Paucimonas lemoignei with high specificity for amorphous polyesters of short chain-length hydroxyalkanoic acids. J Biol Chem 276:36215–36224 . doi: 10.1074/jbc.M101106200

Hang Y, Ran S, Wang X, Jiao J, Wang S, Liu Z (2016) Mutational analysis and stability characterization of a novel esterase of lipolytic enzyme family VI from Shewanella sp. Int J Biol Macromol 93:655–664 . doi: 10.1016/j.ijbiomac.2016.09.032

Hao T, Xie Z, Wang M, Liu L, Zhang Y, Wang W, Zhang Z, Zhao X, Li P, Guo Z, Gao S, Lou C, Zhang G, Merritt J, Horsman GP, Chen Y (2019) An anaerobic bacterium host system for heterologous expression of natural product biosynthetic gene clusters. Nat Commun 10:3665 . doi: 10.1038/s41467-019-11673-0

Hasan F, Shah AA, Hameed A (2006) Industrial applications of microbial lipases. Enzyme Microb Technol 39:235–251 . doi: 10.1016/J.ENZMICTEC.2005.10.016

Hausmann S, Jaeger K-E (2010) Lipolytic enzymes from bacteria. In: Timmis KN (ed) Handbook of Hydrocarbon and Lipid Microbiology. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 1099–1126

Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell 38:576–589 . doi: 10.1016/j.molcel.2010.05.004

Hitch TCA, Clavel T (2019) A proposed update for the classification and description of bacterial lipolytic enzymes. PeerJ 7:e7249 . doi: 10.7717/peerj.7249

Holmes ML, Scopes RK, Moritz RL, Simpson RJ, Englert C, Pfeifer F, Dyall-Smith ML (1997) Purification and analysis of an extremely halophilic beta-galactosidase from Haloferax alicantei. Biochim Biophys Acta 1337:276–286

Horne I, Harcourt RL, Sutherland TD, Russell RJ, Oakeshott JG (2002) Isolation of a Pseudomonas monteilli strain with a novel phosphotriesterase. FEMS Microbiol Lett 206:51–55 . doi: 10.1111/j.1574-6968.2002.tb10985.x

Hotta Y, Ezaki S, Atomi H, Imanaka T (2002) Extremely stable and versatile carboxylesterase from a hyperthermophilic Archaeon. Appl Environ Microbiol 68:3925–3931 . doi: 10.1128/AEM.68.8.3925-3931.2002

Hu P, Tom L, Singh A, Thomas BC, Baker BJ, Piceno YM, Andersen GL, Banfield JF (2016) Genome-resolved metagenomic analysis reveals roles for candidate phyla and other microbial community members in biogeochemical transformations in oil reservoirs. MBio 7:e01669-15 . doi: 10.1128/mBio.01669-15

Hu Y, Fu C, Huang Y, Yin Y, Cheng G, Lei F, Lu N, Li J, Ashforth EJ, Zhang L, Zhu B (2010) Novel lipolytic genes from the microbial metagenomic library of the South China Sea marine sediment. FEMS Microbiol Ecol 72:228–37 . doi: 10.1111/j.1574-6941.2010.00851.x

Huang Y, Niu B, Gao Y, Fu L, Li W (2010) CD-HIT Suite: a web server for clustering and comparing biological sequences. Bioinformatics 26:680–692 . doi: 10.1093/bioinformatics/btq003

Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, Mende DR, Letunic I, Rattei T, Jensen LJ, Mering C von, Bork P (2019) eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. Nucleic Acids Res 47:D309–D314

Hugenholtz P, Tyson GW (2008) Microbiology: Metagenomics. Nature 455:481–483

Hun CJ, Rahman RNZA, Salleh AB, Basri M (2003) A newly isolated organic solvent tolerant Bacillus sphaericus 205y producing organic solvent-stable lipase. Biochem Eng J 15:147–151 . doi: 10.1016/S1369-703X(02)00185-7

Huo YY, Rong Z, Jian SL, Xu C Di, Li J xi, Xu XW (2017) A novel halotolerant thermoalkaliphilic esterase from marine bacterium Erythrobacter seohaensis SW-135. Front Microbiol 8:2315 . doi: 10.3389/fmicb.2017.02315

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 Comput Biol 12:e1004957 . doi: 10.1371/journal.pcbi.1004957

Ikeda M, Clark DS (1998) Molecular cloning of extremely thermostable esterase gene from hyperthermophilic archaeon Pyrococcus furiosus in Escherichia coli. Biotechnol Bioeng 57:624–629 . doi: 10.1002/(SICI)1097-0290(19980305)57:5<624::AID-BIT15>3.0.CO;2-B

Ishibashi M, Hayashi T, Yoshida C, Tokunaga M (2013) Increase of salt dependence of halophilic nucleoside diphosphate kinase caused by a single amino acid substitution. Extremophiles 17:585–591 . doi: 10.1007/s00792-013-0541-9

Jaeger K-E, Dijkstra BW, Reetz MT (1999) Bacterial biocatalysts: Molecular biology, three-dimensional structures, and biotechnological applications of lipases. Annu Rev Microbiol 53:315–351 . doi: 10.1146/annurev.micro.53.1.315

Jaeger KE, Eggert T (2002) Lipases for biotechnology. Curr Opin Biotechnol 13:390–397 . doi: 10.1016/S0958-1669(02)00341-5

Jaeger KE, Reetz MT (1998) Microbial lipases form versatile tools for biotechnology. Trends Biotechnol 16:396–403 . doi: 10.1016/S0167-7799(98)01195-0

Jang E, Ryu BH, Kim TD (2014) Identification, characterization, and immobilization of an organic solvent-stable alkaline hydrolase (PA27) from Pseudomonas aeruginosa MH38. Molecules 19:14396–14405 . doi: 10.3390/molecules190914396

Jara-Samaniego J, Pérez-Murcia MD, Bustamante MA, Pérez-Espinosa A, Paredes C, López M, López-Lluch DB, Gavilanes-Terán I, Moral R (2017) Composting as sustainable strategy for municipal solid waste management in the Chimborazo Region, Ecuador: Suitability of the obtained composts for seedling production. J Clean Prod 141:1349–1358 . doi: 10.1016/j.jclepro.2016.09.178

Jayanath G, Mohandas SP, Kachiprath B, Solomon S, Sajeevan TP, Bright Singh IS, Philip R (2018) A novel solvent tolerant esterase of GDSGG motif subfamily from solar saltern through metagenomic approach: Recombinant expression and characterization. Int J Biol Macromol 119:393–401 . doi: 10.1016/j.ijbiomac.2018.06.057

Jensen MB V., Horsfall LE, Wardrope C, Togneri PD, Marles-Wright J, Rosser SJ (2016) Characterisation of a new family of carboxyl esterases with an OsmC domain. PLoS One 11:e0166128 . doi: 10.1371/journal.pone.0166128

Jeon JH, Kim JT, Kim YJ, Kim HK, Lee HS, Kang SG, Kim SJ, Lee JH (2009) Cloning and characterization of a new cold-active lipase from a deep-sea sediment metagenome. Appl Microbiol Biotechnol 81:865–874 . doi: 10.1007/s00253-008-1656-2

Jeon JH, Kim S-J, Lee HS, Cha S-S, Lee JH, Yoon S-H, Koo B-S, Lee C-M, Choi SH, Lee SH, Kang SG, Lee J-H (2011) Novel metagenome-derived carboxylesterase that hydrolyzes β-lactam antibiotics. Appl Environ Microbiol 77:7830–6 . doi: 10.1128/AEM.05363-11

Jeon JH, Lee HS, Kim JT, Kim SJ, Choi SH, Kang SG, Lee JH (2012) Identification of a new subfamily of salt-tolerant esterases from a metagenomic library of tidal flat sediment. Appl Microbiol Biotechnol 93:623–631 . doi: 10.1007/s00253-011-3433-x

Jiang H, Zhang S, Gao H, Hu N (2016) Characterization of a cold-active esterase from Serratia sp. and improvement of thermostability by directed evolution. BMC Biotechnol 16: . doi: 10.1186/s12896-016-0235-3

Jiang X, Huo Y, Cheng H, Zhang X, Zhu X, Wu M (2012) Cloning, expression and characterization of a halotolerant esterase from a marine bacterium Pelagibacterium halotolerans B2T. Extremophiles 16:427–435 . doi: 10.1007/s00792-012-0442-3

Jin P, Pei X, Du P, Yin X, Xiong X, Wu H, Zhou X, Wang Q (2012) Overexpression and characterization of a new organic solvent-tolerant esterase derived from soil metagenomic DNA. Bioresour Technol 116:234–240 . doi: 10.1016/j.biortech.2011.10.087

Jochens H, Aerts D, Bornscheuer UT (2010) Thermostabilization of an esterase by alignment-guided focussed directed evolution. Protein Eng Des Sel 23:903–909 . doi: 10.1093/protein/gzq071

Jünemann S, Kleinbölting N, Jaenicke S, Henke C, Hassa J, Nelkner J, Stolze Y, Albaum SP, Schlüter A, Goesmann A, Sczyrba A, Stoye J (2017) Bioinformatics for NGS-based metagenomics and the application to biogas research. J Biotechnol 261:10–23 . doi: 10.1016/j.jbiotec.2017.08.012

Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M (2016) KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res 44:D457–D462 . doi: 10.1093/nar/gkv1070

Kang CH, Oh KH, Lee MH, Oh TK, Kim BH, Yoon JH (2011) A novel family VII esterase with industrial potential from compost metagenomic library. Microb Cell Fact 10:41 . doi: 10.1186/1475-2859-10-41

Kang HY, Kim JF, Kim MH, Park SH, Oh TK, Hur CG (2006) MELDB: A database for microbial esterases and lipases. FEBS Lett 580:2736–2740 . doi: 10.1016/j.febslet.2006.04.034

Kang LJ, Meng ZT, Hu C, Zhang Y, Guo HL, Li Q, Li M (2017) Screening, purification, and characterization of a novel organic solvent-tolerant esterase, Lip2, from Monascus purpureus strain M7. Extremophiles 21:345–355 . doi: 10.1007/s00792-016-0907-x

Karadzic I, Masui A, Zivkovic LI, Fujiwara N (2006) Purification and characterization of an alkaline lipase from Pseudomonas aeruginosa isolated from putrid mineral cutting oil as component of metalworking fluid. J Biosci Bioeng 102:82–89 . doi: 10.1263/jbb.102.82

Karan R, Capes MD, DasSarma S (2012) Function and biotechnology of extremophilic enzymes in low water activity. Aquat Biosyst 8:4 . doi: 10.1186/2046-9063-8-4

Keegan KP, Glass EM, Meyer F (2016) MG-RAST, a metagenomics service for analysis of microbial community structure and function. In: Methods in Molecular Biology. Humana Press Inc., pp 207–233

Kent WJ (2002) BLAT---The BLAST-Like Alignment Tool. Genome Res 12:656–664 . doi: 10.1101/gr.229202

Khan FI, 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

Khan M, Sathya TA (2018) Extremozymes from metagenome: Potential applications in food processing. Crit Rev Food Sci Nutr 58:2017–2025 . doi: 10.1080/10408398.2017.1296408

Kim MK, Kang TH, Kim J, Kim H, Yun HD (2012) Cloning and identification of a new group esterase (Est5S) from noncultured rumen bacterium. J Microbiol Biotechnol 22:1044–1053

Kim S, Joo S, Yoon S, Kim S, Moon J, Ryu Y, Kim KK, Kim TD (2009) Purification, crystallization and preliminary crystallographic analysis of Est-Y29: A novel oligomeric β-lactamase. Acta Crystallogr Sect F Struct Biol Cryst Commun 65:310–312 . doi: 10.1107/S1744309109005442
Kim TD (2017) Bacterial hormone-sensitive lipases (bHSLs): emerging enzymes for biotechnological applications. J Microbiol Biotechnol 27:1907–1915 . doi: 10.4014/jmb.1708.08004
Kim YH, Kwon EJ, Kim SK, Jeong YS, Kim J, Yun HD, Kim H (2010) Molecular cloning and characterization of a novel family VIII alkaline esterase from a compost metagenomic library. Biochem Biophys Res Commun 393:45–49 . doi: 10.1016/j.bbrc.2010.01.070
Kim YJ, Kim H, Kim SH, Rha E, Choi SL, Yeom SJ, Kim HS, Lee SG (2016) Improved metagenome screening efficiency by random insertion of T7 promoters. J Biotechnol 230:47–53 . doi: 10.1016/j.jbiotec.2016.05.018
Kiran GS, Lipton AN, Kennedy J, Dobson ADW, Selvin J (2014) A halotolerant thermostable lipase from the marine bacterium Oceanobacillus sp. PUMB02 with an ability to disrupt bacterial biofilms. Bioengineered 5:305–318 . doi: 10.4161/bioe.29898
Kixmüller D, Greie JC (2012) Construction and characterization of a gradually inducible expression vector for Halobacterium salinarum, based on the kdp promoter. Appl Environ Microbiol 78:2100–2105 . doi: 10.1128/AEM.07155-11
Knudsen M, Wiuf C (2010) The CATH database. Hum Genomics 4:207–212 . doi: 10.1186/1479-7364-4-3-207
Kobayashi T, Kanai H, Aono R, Horikoshi K, Kudo T (1994) Cloning, expression, and nucleotide sequence of the alpha-amylase gene from the haloalkaliphilic archaeon Natronococcus sp. strain Ah-36. J Bacteriol 176:5131–5144
Koseki T, Fushinobu S, Ardiansyah, Shirakawa H, Komai M (2009) Occurrence, properties, and applications of feruloyl esterases. Appl Microbiol Biotechnol 84:803–810 . doi: 10.1007/s00253-009-2148-8
Kovacic F, Babic N, Krauss U, Jaeger K (2019) Classification of Lipolytic Enzymes from Bacteria. In: Rojo F (ed) Aerobic Utilization of Hydrocarbons, Oils and Lipids. Handbook of Hydrocarbon and Lipid Microbiology. Springer, Cham, pp 1–35
Kuddus M (2018) Cold-active enzymes in food biotechnology: An updated mini review. J Appl Biol Biotechnol 6:58–63 . doi: 10.7324/JABB.2018.60310
Kumar A, Dhar K, Singh Kanwar S, Arora PK (2016) Lipase catalysis in organic solvents: advantages and applications. Biol Proced Online 18:2 . doi: 10.1186/s12575-016-0033-2
Kumar S, Krishnani KK, Bhushan B, Brahmane MP (2015) Metagenomics: Retrospect and prospects in high throughput age. Biotechnol Res Int 2015:1–13 . doi: 10.1155/2015/121735
Lam KN, Cheng J, Engel K, Neufeld JD, Charles TC (2015) Current and future resources for functional metagenomics. Front Microbiol 6:1196 . doi: 10.3389/fmicb.2015.01196
Lämmle K, Zipper H, Breuer M, Hauer B, Buta C, Brunner H, Rupp S (2007) Identification of novel enzymes with different hydrolytic activities by metagenome expression cloning. J Biotechnol 127:575–592 . doi: 10.1016/j.jbiotec.2006.07.036
Lanyi JK (1974) Salt-dependent properties of proteins from extremely halophilic bacteria. Bacteriol Rev 38:272–290
Laudadio I, Fulci V, Stronati L, Carissimi C (2019) Next-generation metagenomics: methodological challenges and opportunities. Omi A J Integr Biol 23:327–333 . doi: 10.1089/omi.2019.0073
Lee CW, Kwon S, Park SH, Kim BY, Yoo W, Ryu BH, Kim HW, Shin SC, Kim S, Park H, Kim TD, Lee JH (2017) Crystal structure and functional characterization of an esterase (EaEST) from Exiguobacterium antarcticum. PLoS One 12:e0169540 . doi: 10.1371/journal.pone.0169540
Lee D-W, Koh Y-S, Kim K-J, Kim B-C, Choi H-J, Kim D-S, Suhartono MT, Pyun Y-R (1999) Isolation and characterization of a thermophilic lipase from Bacillus thermoleovorans ID-1. FEMS Microbiol Lett 179:393–400 . doi: 10.1111/j.1574-6968.1999.tb08754.x
Lee JH, Jeon JH, Kim JT, Lee HS, Kim SJ, Kang SG, Choi SH (2011) Novel lipolytic enzymes identified from metagenomic library of deep-sea sediment. Evidence-based Complement Altern Med 2011: . doi: 10.1155/2011/271419
Lee M-H, Oh K-H, Kang C-H, Kim J-H, Oh T-K, Ryu C-M, Yoon J-H (2012) Novel metagenome-derived, cold-adapted alkaline phospholipase with superior lipase activity as an intermediate between phospholipase and lipase. Appl Environ Microbiol 78:4959–4966 . doi: 10.1128/aem.00260-12
Lee MH, Hong KS, Malhotra S, Park J-H, Hwang EC, Choi HK, Kim YS, Tao W, Lee S-W (2010) A new esterase EstD2 isolated from plant rhizosphere soil metagenome. Appl Microbiol Biotechnol 88:1125–1134 . doi: 10.1007/s00253-010-2729-6

Lee MH, Lee CH, Oh TK, Song JK, Yoon JH (2006) Isolation and characterization of a novel lipase from a metagenomic library of tidal flat sediments: Evidence for a new family of bacterial lipases. Appl Environ Microbiol 72:7406–7409 . doi: 10.1128/AEM.01157-06

Lee SW, Won K, Lim HK, Kim JC, Choi GJ, Cho KY (2004) Screening for novel lipolytic enzymes from uncultured soil microorganisms. Appl Microbiol Biotechnol 65:720–726 . doi: 10.1007/s00253-004-1722-3

Lei J, Miao Y, Lan Y, Han X, Liu H, Gan Y, Niu L, Wang Y, Zheng Z (2018) A novel complementation assay for quick and specific screen of genes encoding glycerol-3-phosphate acyltransferases. Front Plant Sci 9:353 . doi: 10.3389/fpls.2018.00353

Leis B, Angelov A, Mientus M, Li H, Pham VTT, Lauinger B, Bongen P, Pietruszka J, Gonçalves LG, Santos H, Liebl W (2015) Identification of novel esterase-active enzymes from hot environments by use of the host bacterium Thermus thermophilus. Front Microbiol 6:275 . doi: 10.3389/fmicb.2015.00275

Leite AET, Briganti L, de Araújo EA, Pellegrini V de OA, Camilo CM, Polikarpov I (2020) Low-resolution molecular shape, biochemical characterization and emulsification properties of a halotolerant esterase from Bacillus licheniformis. Eur Biophys J 49:435–447 . doi: 10.1007/s00249-020-01448-7

Lenfant N, Hotelier T, Velluet E, Bourne Y, Marchot P, Chatonnet A (2013) ESTHER, the database of the α/β-hydrolase fold superfamily of proteins: tools to explore diversity of functions. Nucleic Acids Res 41:423–429 . doi: 10.1093/nar/gks1154

Levisson M, Van Der Oost J, Kengen SWM (2007) Characterization and structural modeling of a new type of thermostable esterase from Thermotoga maritima. FEBS J 274:2832–2842 . doi: 10.1111/j.1742-4658.2007.05817.x

Li D, Liu CM, Luo R, Sadakane K, Lam TW (2015a) MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics 31:1674–1676 . doi: 10.1093/bioinformatics/btv033

Li H, Han X, Qiu W, Xu D, Wang Y, Yu M, Hu X, Zhuo R (2019) Identification and expression analysis of the GDSL esterase/lipase family genes, and the characterization of SaGLIP8 in Sedum alfredii Hance under cadmium stress. PeerJ 7:e6741 . doi: 10.7717/peerj.6741

Li J, Zhang K, Han W (2010) Cloning and biochemical characterization of a novel lipolytic gene from activated sludge metagenome, and its gene product. Microb Cell Fact 9:83 . doi: 10.1186/1475-2859-9-83

Li LL, McCorkle SR, Monchy S, Taghavi S, van der Lelie D (2009) Bioprospecting metagenomes: Glycosyl hydrolases for converting biomass. Biotechnol Biofuels 2:1–11 . doi: 10.1186/1754-6834-2-10

Li M, Yang LR, Xu G, Wu JP (2016) Cloning and characterization of a novel lipase from Stenotrophomonas maltophilia GS11: The first member of a new bacterial lipase family XVI. J Biotechnol 228:30–36 . doi: 10.1016/j.jbiotec.2016.04.034

Li M, Yang LR, Xu G, Wu JP (2013) Screening, purification and characterization of a novel cold-active and organic solvent-tolerant lipase from Stenotrophomonas maltophilia CGMCC 4254. Bioresour Technol 148:114–120 . doi: 10.1016/j.biortech.2013.08.101

Li P-Y, Zhang Y, Xie B-B, Zhang Y-Q, Hao J, Wang Y, Wang P, Li C-Y, Qin Q-L, Zhang X-Y, Su H-N, Shi M, Zhang Y-Z, Chen X-L (2017a) Structural and mechanistic insights into the improvement of the halotolerance of a marine microbial esterase by increasing intra- and interdomain hydrophobic interactions. Appl Environ Microbiol 83:e01286-17 . doi: 10.1128/AEM.01286-17

Li P, Chen X, Ji P, Li C, Wang P, Zhang Y, Xie B, Qin Q, Su H, Zhou B, Zhang Y, Zhang X (2015b) Interdomain hydrophobic interactions modulate the thermostability of microbial esterases from the hormone-sensitive lipase family. J Biol Chem 290:11188–11198 . doi: 10.1074/jbc.M115.646182

Li PY, Ji P, Li CY, Zhang Y, Wang GL, Zhang XY, Xie B Bin, Qin QL, Chen XL, Zhou BC, Zhang YZ (2014) Structural basis for dimerization and catalysis of a novel sterase from the GTSAG motif subfamily of the bacterial hormone-sensitive lipase family. J Biol Chem 289:19031–19041 . doi: 10.1074/jbc.M114.574913

Li PY, Yao QQ, Wang P, Zhang Y, Li Y, Zhang YQ, Hao J, Zhou BC, Chen XL, Shi M, Zhang YZ, Zhang XY (2017b) A novel subfamily esterase with a homoserine transacetylase-like fold but no transferase activity. Appl Environ Microbiol 83:e00131-17 . doi: 10.1128/AEM.00131-17

Liebl W, Angelov A, Juergensen J, Chow J, Loeschcke A, Drepper T, Classen T, Pietruzska J, Ehrenreich A, Streit WR, Jaeger KE (2014) Alternative hosts for functional (meta)genome analysis. Appl. Microbiol. Biotechnol. 98:8099–8109

Lima VMG, Krieger N, Mitchell DA, Fontana JD (2004) Activity and stability of a crude lipase from Penicillium aurantiogriseum in aqueous media and organic solvents. Biochem Eng J 18:65–71 . doi: 10.1016/S1369-703X(03)00165-7

Liu K, Wang J, Bu D, Zhao S, McSweeney C, Yu P, Li D (2009) Isolation and biochemical characterization of two lipases from a metagenomic library of China Holstein cow rumen. Biochem Biophys Res Commun 385:605–611 . doi: 10.1016/j.bbrc.2009.05.110

Liu SP, Liu RX, Zhang L, Shi GY (2015) Sequence-based screening and characterization of cytosolic mandelate oxidase using oxygen as electron acceptor. Enzyme Microb Technol 69:24–30 . doi: 10.1016/j.enzmictec.2014.11.001

Liu YF, Galzerani DD, Mbadinga SM, Zaramela LS, Gu JD, Mu BZ, Zengler K (2018) Metabolic capability and in situ activity of microorganisms in an oil reservoir. Microbiome 6:5 . doi: 10.1186/s40168-017-0392-1

Lopez-Lopez O, Cerdan M, Siso M (2014) New extremophilic lipases and esterases from metagenomics. Curr Protein Pept Sci 15:445–455 . doi: 10.2174/1389203715666140228153801

López-López O, Knapik K, Cerdán ME, González-Siso MI (2015) Metagenomics of an alkaline hot spring in Galicia (Spain): Microbial diversity analysis and screening for novel lipolytic enzymes. Front Microbiol 6:1291 . doi: 10.3389/fmicb.2015.01291

Lu M, Daniel R (2021) A novel carboxylesterase derived from a compost metagenome exhibiting high stability and activity towards high salinity. Genes (Basel) 12:1–18 . doi: 10.3390/genes12010122

Lu M, Dukunde A, Daniel R (2019) Biochemical profiles of two thermostable and organic solvent–tolerant esterases derived from a compost metagenome. Appl Microbiol Biotechnol 103:3421–3437 . doi: 10.1007/s00253-019-09695-1

Lu M, Schneider D, Daniel R (2021) Metagenomic screening for lipolytic genes reveals an ecology-clustered distribution pattern. bioRxiv 2021.02.01.429111 . doi: 10.1101/2021.02.01.429111

Luan Z-J, Yu H-L, Ma B-D, Qi Y-K, Chen Q, Xu J-H (2016) Dramatically improved performance of an esterase for cilastatin synthesis by cap domain engineering. Ind Eng Chem Res 55:12167–12172 . doi: 10.1021/acs.iecr.6b02440

Ma S, Fang C, Sun X, Han L, He X, Huang G (2018) Bacterial community succession during pig manure and wheat straw aerobic composting covered with a semi-permeable membrane under slight positive pressure. Bioresour Technol 259:221–227 . doi: 10.1016/j.biortech.2018.03.054

Madhavan A, Sindhu R, Parameswaran B, Sukumaran RK, Pandey A (2017) Metagenome analysis: a powerful tool for enzyme bioprospecting. Appl Biochem Biotechnol 183:636–651 . doi: 10.1007/s12010-017-2568-3

Mahapatra GP, Raman S, Nayak S, Gouda S, Das G, Patra JK (2020) Metagenomics approaches in discovery and development of new bioactive compounds from marine Actinomycetes. Curr Microbiol 77:645–656 . doi: 10.1007/s00284-019-01698-5

Mandrich L, Menchise V, Alterio V, De Simone G, Pedone C, Rossi M, Manco G (2008) Functional and structural features of the oxyanion hole in a thermophilic esterase from Alicyclobacillus acidocaldarius. Proteins Struct Funct Genet 71:1721–1731 . doi: 10.1002/prot.21877

Mandrich L, Merone L, Pezzullo M, Cipolla L, Nicotra F, Rossi M, Manco G (2005) Role of the N terminus in enzyme activity, stability and specificity in thermophilic esterases belonging to the HSL family. J Mol Biol 345:501–512 . doi: 10.1016/j.jmb.2004.10.035

Mao H, Lv Z, Sun H, Li R, Zhai B, Wang Z, Awasthi MK, Wang Q, Zhou L (2018) Improvement of biochar and bacterial powder addition on gaseous emission and bacterial community in pig manure compost. Bioresour Technol 258:195–202 . doi: 10.1016/j.biortech.2018.02.082

Markel U, Essani KD, Besirlioglu V, Schiffels J, Streit WR, Schwaneberg U (2020) Advances in ultrahigh-throughput screening for directed enzyme evolution. Chem Soc Rev 49:233–262 . doi: 10.1039/c8cs00981c

Maruthamuthu M, Jiménez DJ, Stevens P, van Elsas JD (2016) A multi-substrate approach for functional metagenomics-based screening for (hemi)cellulases in two wheat straw-degrading microbial consortia unveilsnovel thermoalkaliphilic enzymes. BMC Genomics 17:1–16 . doi: 10.1186/s12864-016-2404-0

McArthur AG, Waglechner N, Nizam F, Yan A, Azad MA, Baylay AJ, Bhullar K, Canova MJ, De Pascale G, Ejim L, Kalan L, King AM, Koteva K, Morar M, Mulvey MR, O'Brien JS, Pawlowski AC, Piddock LJV, Spanogiannopoulos P, Sutherland AD, Tang I, Taylor PL, Thaker M, Wang W, Yan M, Yu T, Wright GD (2013) The comprehensive antibiotic resistance database. Antimicrob Agents Chemother 57:3348–3357 . doi: 10.1128/AAC.00419-13

McMahon MD, Guan C, Handelsman J, Thomas MG (2012) Metagenomic analysis of Streptomyces lividans reveals host-dependent functional expression. Appl Environ Microbiol 78:3622–3629 . doi: 10.1128/AEM.00044-12

Meng Q, Yang W, Men M, Bello A, Xu X, Xu B, Deng L, Jiang X, Sheng S, Wu X, Han Y, Zhu H (2019) Microbial community succession and response to environmental variables during cow manure and corn straw composting. Front Microbiol 10:529 . doi: 10.3389/fmicb.2019.00529

Meng Y, Li S, Yuan H, Zou D, Liu Y, Zhu B, Li X (2015) Effect of lipase addition on hydrolysis and biomethane production of Chinese food waste. Bioresour Technol 179:452–459 . doi: 10.1016/j.biortech.2014.12.015
Menzel P, Ng KL, Krogh A (2016) Fast and sensitive taxonomic classification for metagenomics with Kaiju. Nat Commun 7:11257 . doi: 10.1038/ncomms11257
Metin K, Burcu Bakir Ateslier Z, Basbulbul G, Halil Biyik H (2006) Characterization of esterase activity in Geobacillus sp. HBB-4. J Basic Microbiol 46:400–409 . doi: 10.1002/jobm.200510121
Miller DN, Bryant JE, Madsen EL, Ghiorse WC (1999) Evaluation and optimization of DNA extraction and purification procedures for soil and sediment samples. Appl Environ Microbiol 65:4715–4724 . doi: 10.1128/aem.65.11.4715-4724.1999
Minegishi H, Shimane Y, Echigo A, Ohta Y, Hatada Y, Kamekura M, Maruyama T, Usami R (2013) Thermophilic and halophilic β-agarase from a halophilic archaeon Halococcus sp. 197A. Extremophiles 17:931–939 . doi: 10.1007/s00792-013-0575-z
Mirete S, Morgante V, González-Pastor JE (2016) Functional metagenomics of extreme environments. Curr Opin Biotechnol 38:143–149 . doi: 10.1016/j.copbio.2016.01.017
Misawa E, Chan Kwo Chion CKC, Archer I V., Woodland MP, Zhou N-Y, Carter SF, Widdowson DA, Leak DJ (1998) Characterisation of a catabolic epoxide hydrolase from a Corynebacterium sp. Eur J Biochem 253:173–183 . doi: 10.1046/j.1432-1327.1998.2530173.x
Mohamed YM, Ghazy MA, Sayed A, Ouf A, El-Dorry H, Siam R (2013) Isolation and characterization of a heavy metal-resistant, thermophilic esterase from a Red Sea Brine Pool. Sci Rep 3:3358 . doi: 10.1038/srep03358
Mølgaard A, Kauppinen S, Larsen S (2000) Rhamnogalacturonan acetylesterase elucidates the structure and function of a new family of hydrolases. Structure 8:373–383 . doi: 10.1016/S0969-2126(00)00118-0
Montoro-García S, Martínez-Martínez I, Navarro-Fernández J, Takami H, García-Carmona F, Sánchez-Ferrer Á (2009) Characterization of a novel thermostable carboxylesterase from Geobacillus kaustophilus HTA426 shows the existence of a new carboxylesterase family. J Bacteriol 191:3076–3085 . doi: 10.1128/JB.01060-08
Moreno-Indias I, Tinahones FJ (2020) Metagenomics. Princ Nutr Nutr 81–87 . doi: 10.1016/b978-0-12-804572-5.00011-2
Moriya T, Hikota T, Yumoto I, Ito T, Terui Y, Yamagishi A, Oshima T (2011) Calditerricola satsumensis gen. nov., sp. nov. and Calditerricola yamamurae sp. nov., extreme thermophiles isolated from a high-temperature compost. Int J Syst Evol Microbiol 61:631–636 . doi: 10.1099/ijs.0.018416-0
Munawar N, Engel PC (2013) Halophilic enzymes: Characteristics, structural adaptation and potential applications for biocatalysis. Curr Biotechnol 2:334–344 . doi: 10.2174/18722083113076660033
Munawar N, Engel PC (2012) Overexpression in a non-native halophilic host and biotechnological potential of NAD+-dependent glutamate dehydrogenase from Halobacterium salinarum strain NRC-36014. Extremophiles 16:463–476 . doi: 10.1007/s00792-012-0446-z
Musa H, Kasim FH, Gunny AAN, Gopinath SCB, Ahmad MA (2018) Biosecretion of higher halophilic lipase by a novel Bacillus amyloliquefaciens AIKK2 using agro-waste as supporting substrate. Process Biochem 72:55–62 . doi: 10.1016/j.procbio.2018.06.022
Nacke H, Will C, Herzog S, Nowka B, Engelhaupt M, Daniel R (2011) Identification of novel lipolytic genes and gene families by screening of metagenomic libraries derived from soil samples of the German Biodiversity Exploratories. FEMS Microbiol Ecol 78:188–201 . doi: 10.1111/j.1574-6941.2011.01088.x
Nagamine S, Liu C, Nishishita J, Kozaki T, Sogahata K, Sato Y, Minami A, Ozaki T, Schmidt-Dannert C, Maruyama J ichi, Oikawa H (2019) Ascomycete Aspergillus oryzae is an efficient expression host for production of basidiomycete terpenes by using genomic DNA sequences. Appl Environ Microbiol 85:e00409-19 . doi: 10.1128/AEM.00409-19
Nakamura AM, Kadowaki MAS, Godoy A, Nascimento AS, Polikarpov I (2018) Low-resolution envelope, biophysical analysis and biochemical characterization of a short-chain specific and halotolerant carboxylesterase from Bacillus licheniformis. Int J Biol Macromol 120:1893–1905 . doi: 10.1016/j.ijbiomac.2018.10.003
Nam KH, Kim MY, Kim SJ, Priyadarshi A, Lee WH, Hwang KY (2009) Structural and functional analysis of a novel EstE5 belonging to the subfamily of hormone-sensitive lipase. Biochem Biophys Res Commun 379:553–556 . doi: 10.1016/j.bbrc.2008.12.085
Namiki T, Hachiya T, Tanaka H, Sakakibara Y (2012) MetaVelvet: An extension of Velvet assembler to de novo metagenome assembly from short sequence reads. Nucleic Acids Res 40:e155 . doi: 10.1093/nar/gks678
Neher DA, Weicht TR, Bates ST, Leff JW, Fierer N (2013) Changes in bacterial and fungal communities across compost recipes, preparation methods, and composting times. PLoS One 8:e79512 . doi: 10.1371/journal.pone.0079512

Nerurkar M, Joshi M, Pariti S, Adivarekar R (2013) Application of lipase from marine bacteria Bacillus sonorensis as an additive in detergent formulation. J Surfactants Deterg 16:435–443 . doi: 10.1007/s11743-012-1434-0
Neves Petersen MT, Fojan P, Petersen SB (2001) How do lipases and esterases work: The electrostatic contribution. J Biotechnol 85:115–147 . doi: 10.1016/S0168-1656(00)00360-6
Ngara TR, Zhang H (2018) Recent advances in function-based metagenomic screening. Genomics, Proteomics Bioinforma. 16:405–415
Noby N, Hussein A, Saeed H, Embaby AM (2020) Recombinant cold-adapted halotolerant, organic solvent-stable esterase (estHIJ) from Bacillus halodurans. Anal Biochem 591:113554 . doi: 10.1016/j.ab.2019.113554
Nurk S, Meleshko D, Korobeynikov A, Pevzner PA (2017) MetaSPAdes: A new versatile metagenomic assembler. Genome Res 27:824–834 . doi: 10.1101/gr.213959.116
Ogino H, Ishikawa H (2001) Enzymes which are stable in the presence of organic solvents. J Biosci Bioeng 91:109–116
Ogino H, Mimitsuka T, Muto T, Matsumura M, Yasuda M, Ishimi K, Ishikawa H (2004) Cloning, expression, and characterization of a lipolytic enzyme gene (lip8) from Pseudomonas aeruginosa LST-03. J Mol Microbiol Biotechnol 7:212–223 . doi: 10.1159/000079830
Ollis DL, Cheah E, Cygler M, Dijkstra B, Frolow F, Franken SM, Harel M, Remington SJ, Silman I, Schrag J, Sussman JL, Verschueren KHG, Goldman A (1992) The α / β hydrolase fold. Protein Eng Des Sel 5:197–211 . doi: 10.1093/protein/5.3.197
Palaniveloo K, Amran MA, Norhashim NA, Mohamad-Fauzi N, Peng-Hui F, Hui-Wen L, Kai-Lin Y, Jiale L, Chian-Yee MG, Jing-Yi L, Gunasekaran B, Razak SA (2020) Food waste composting and microbial community structure profiling. Processes 8:723 . doi: 10.3390/pr8060723
Panda T, Gowrishankar BS (2005) Production and applications of esterases. Appl Microbiol Biotechnol 67:160–169 . doi: 10.1007/s00253-004-1840-y
Parapouli M, Foukis A, Stergiou P-Y, Koukouritaki M, Magklaras P, Gkini OA, Papamichael EM, Afendra A-S, Hatziloukas E (2018) Molecular, biochemical and kinetic analysis of a novel, thermostable lipase (LipSm) from Stenotrophomonas maltophilia Psi-1, the first member of a new bacterial lipase family (XVIII). J Biol Res 25:4 . doi: 10.1186/s40709-018-0074-6
Park SY, Kim JT, Kang SG, Woo JH, Lee JH, Choi HT, Kim SJ (2007) A new esterase showing similarity to putative dienelactone hydrolase from a strict marine bacterium, Vibrio sp. GMD509. Appl Microbiol Biotechnol 77:107–115 . doi: 10.1007/s00253-007-1134-2
Patel U, Chandpura J, Chauhan K, Gupte S (2018) Screening and isolation of an organic solvent tolerant lipase producing bacteria from various oil contaminated sites. Indian J Appl Microbiol 21:22–36
Pehrsson EC, Tsukayama P, Patel S, Mejía-Bautista M, Sosa-Soto G, Navarrete KM, Calderon M, Cabrera L, Hoyos-Arango W, Bertoli MT, Berg DE, Gilman RH, Dantas G (2016) Interconnected microbiomes and resistomes in low-income human habitats. Nature 533:212–216 . doi: 10.1038/nature17672
Peng Y, Leung HCM, Yiu SM, Chin FYL (2011) Meta-IDBA: A de Novo assembler for metagenomic data. Bioinformatics 27:94–101 . doi: 10.1093/bioinformatics/btr216
Peng Y, Leung HCM, Yiu SM, Chin FYL (2012) IDBA-UD: A de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. Bioinformatics 28:1420–1428 . doi: 10.1093/bioinformatics/bts174
Pereira MR, Mercaldi GF, Maester TC, Balan A, de Macedo Lemos EG (2015) Est16, a new esterase isolated from a metagenomic library of a microbial consortium specializing in diesel oil degradation. PLoS One 10:e0133723 . doi: 10.1371/journal.pone.0133723
Petrovskaya LE, Novototskaya-Vlasova KA, Spirina E V., Durdenko E V., Lomakina GY, Zavialova MG, Nikolaev EN, Rivkina EM (2016) Expression and characterization of a new esterase with GCSAG motif from a permafrost metagenomic library. FEMS Microbiol Ecol 92:fiw046 . doi: 10.1093/femsec/fiw046
Pezzullo M, Del Vecchio P, Mandrich L, Nucci R, Rossi M, Manco G (2013) Comprehensive analysis of surface charged residues involved in thermal stability in Alicyclobacillus acidocaldarius esterase 2. Protein Eng Des Sel 26:47–58 . doi: 10.1093/protein/gzs066
Pieper U, Kapadia G, Mevarech M, Herzberg O (1998) Structural features of halophilicity derived from the crystal structure of dihydrofolate reductase from the Dead Sea halophilic archaeon, Haloferax volcanii. Structure 6:75–88 . doi: 10.2210/PDB1VDR/PDB
Poidevin L, MacNeill SA (2006) Biochemical characterisation of LigN, an NAD+-dependent DNA ligase from the halophilic euryarchaeon Haloferax volcanii that displays maximal in vitro activity at high salt concentrations. BMC Mol Biol 7:44 . doi: 10.1186/1471-2199-7-44
Pratama R, Schneider D, Böer T, Daniel R (2019) First insights into bacterial gastrointestinal tract communities of the Eurasian beaver (Castor fiber). Front Microbiol 10:1646 . doi: 10.3389/fmicb.2019.01646

Qin Q-L, Zhao D-L, Wang J, Chen X-L, Dang H-Y, Li T-G, Zhang Y-Z, Gao P-J (2007) Wangia profunda gen. nov., sp. nov., a novel marine bacterium of the family Flavobacteriaceae isolated from southern Okinawa Trough deep-sea sediment. FEMS Microbiol Lett 271:53–58 . doi: 10.1111/j.1574-6968.2007.00694.x

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO (2013) The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. Nucleic Acids Res 41:D590–D596 . doi: 10.1093/nar/gks1219

Quince C, Walker AW, Simpson JT, Loman NJ, Segata N (2017) Shotgun metagenomics, from sampling to analysis. Nat Biotechnol 35:833–844 . doi: 10.1038/nbt.3935

Rabausch U, Juergensen J, Ilmberger N, Böhnke S, Fischer S, Schubach B, Schulte M, Streit W V. (2013) Functional screening of metagenome and genome libraries for detection of novel flavonoid-modifying enzymes. Appl Environ Microbiol 79:4551–4563 . doi: 10.1128/AEM.01077-13

Rahman MA, Culsum U, Tang W, Zhang SW, Wu G, Liu Z (2016) Characterization of a novel cold active and salt tolerant esterase from Zunongwangia profunda. Enzyme Microb Technol 85:1–11 . doi: 10.1016/j.enzmictec.2015.12.013

Rajendran A, Palanisamy A, Thangavelu V (2009) Lipase catalyzed ester synthesis for food processing industries. Brazilian Arch Biol Technol 52:207–219 . doi: 10.1590/S1516-89132009000100026

Ramnath L, Sithole B, Govinden R (2016) Classification of lipolytic enzymes and their biotechnological applications in the pulping industry. Can J Microbiol 63:179–192 . doi: 10.1139/cjm-2016-0447

Rao L, Xue Y, Zhou C, Tao J, Li G, Lu JR, Ma Y (2011) A thermostable esterase from Thermoanaerobacter tengcongensis opening up a new family of bacterial lipolytic enzymes. Biochim Biophys Acta - Proteins Proteomics 1814:1695–1702 . doi: 10.1016/j.bbapap.2011.08.013

Rao L, Zhao X, Li Y, Xue Y, Ma Y, Lu JR (2009a) Solution behavior and activity of a halophilic esterase under high salt concentration. PLoS One 4:e6980 . doi: 10.1371/journal.pone.0006980

Rao L, Zhao X, Pan F, Li Y, Xue Y, Ma Y, Lu JR (2009b) Solution behavior and activity of a halophilic esterase under high salt concentration. PLoS One 4:e6980 . doi: 10.1371/journal.pone.0023325

Rao L, Zhao X, Pan F, Li Y, Xue Y, Ma Y, Lu JR (2009c) Solution behavior and activity of a halophilic esterase under high salt concentration. PLoS One 4:e6980 . doi: 10.1371/journal.pone.0006980

Rastogi M, Nandal M, Khosla B (2020) Microbes as vital additives for solid waste composting. Heliyon 6:e03343 . doi: 10.1016/j.heliyon.2020.e03343

Rathi P, Bradoo S, Saxena RK, Gupta R (2000) A hyper-thermostable, alkaline lipase from Pseudomonas sp. with the property of thermal activation. Biotechnol Lett 22:495–498 . doi: 10.1023/A:1005604617440

Rebollido R, Martínez J, Aguilera Y, Melchor K, Koerner I, Stegmann R (2008) Microbial populations during composting process of organic fraction of municipal solid waste. Appl Ecol Environ Res 6:61–67

Reyes-Duarte D, Ferrer M, García-Arellano H (2012) Functional-based screening methods for lipases, esterases, and phospholipases in metagenomic libraries. Methods Mol Biol 861:101–113 . doi: 10.1007/978-1-61779-600-5_6

Reyes A, Alves JMP, Durham AM, Gruber A (2017) Use of profile hidden Markov models in viral discovery: current insights. Adv Genomics Genet 7:29–45 . doi: 10.2147/AGG.S136574

Rhee J, Ahn D, Kim Y, Oh J (2005) New thermophilic and thermostable esterase with sequence similarity to the hormone-sensitive lipase family , cloned from a metagenomic library. Appl Environ Microbiol 71:817–825 . doi: 10.1128/AEM.71.2.817

Richard SB, Madern D, Garcin E, Zaccai G (2000) Halophilic adaptation: novel solvent protein interactions observed in the 2.9 and 2.6 A resolution structures of the wild type and a mutant of malate dehydrogenase from Haloarcula marismortui. Biochemistry 39:992–1000

Rigoldi F, Donini S, Redaelli A, Parisini E, Gautieri A (2018) Review: Engineering of thermostable enzymes for industrial applications. APL Bioeng 2:011501 . doi: 10.1063/1.4997367

Rodríguez MC, Loaces I, Amarelle V, Senatore D, Iriarte A, Fabiano E, Noya F (2015) Est10: A novel alkaline esterase isolated from bovine rumen belonging to the new family XV of lipolytic enzymes. PLoS One 10:e0126651 . doi: 10.1371/journal.pone.0126651

Romdhane IB-B, Fendri A, Gargouri Y, Gargouri A, Belghith H (2010) A novel thermoactive and alkaline lipase from Talaromyces thermophilus fungus for use in laundry detergents. Biochem Eng J 53:112–120 . doi: 10.1016/J.BEJ.2010.10.002

Rong Z, Huo Y-Y, Jian S-L, Wu Y-H, Xu X-W (2018) Characterization of a novel alkaline esterase from Altererythrobacter epoxidivorans CGMCC 1.7731 T. Prep Biochem Biotechnol 48:113–120 . doi: 10.1080/10826068.2017.1387559

Rosenstein R, Götz F (2000) Staphylococcal lipases: Biochemical and molecular characterization. Biochimie 82:1005–1014
Ryckeboer J, Mergaert J, Vaes K, Klammer S, Clercq D, Coosemans J, Insam H, Swings J (2003) A survey of bacteria and fungi occurring during composting and self-heating processes. Ann Microbiol 53:349–410
Sabri S, Rahman RNZRA, Leow TC, Basri M, Salleh AB (2009) Secretory expression and characterization of a highly Ca2+-activated thermostable L2 lipase. Protein Expr Purif 68:161–166 . doi: 10.1016/j.pep.2009.08.002
Sadeghi M, Naderi-Manesh H, Zarrabi M, Ranjbar B (2006) Effective factors in thermostability of thermophilic proteins. Biophys Chem 119:256–270 . doi: 10.1016/j.bpc.2005.09.018
Salihu A, Alam MZ (2015) Solvent tolerant lipases: A review. Process Biochem 50:86–96 . doi: 10.1016/j.procbio.2014.10.019
Samoylova Y V., Sorokina KN, Piligaev A V., Parmon VN (2019) Application of bacterial thermostable lipolytic enzymes in the modern biotechnological processes: A review. Catal Ind 11:168–178 . doi: 10.1134/S2070050419020107
Samoylova Y V., Sorokina KN, Romanenko M V., Parmon VN (2018) Cloning, expression and characterization of the esterase estUT1 from Ureibacillus thermosphaericus which belongs to a new lipase family XVIII. Extremophiles 22:271–285 . doi: 10.1007/s00792-018-0996-9
Santi C De, Zhai L, Ambrosino L, Tedesco P, Zhou C, Xue Y, Ma Y, Pascale D de (2015) Identification and characterization of a novel salt-tolerant esterase from a Tibetan glacier metagenomic library. Biotechnol Prog 31:890–899 . doi: 10.1002/btpr.
Sarmah N, Revathi D, Sheelu G, Yamuna Rani K, Sridhar S, Mehtab V, Sumana C (2018) Recent advances on sources and industrial applications of lipases. Biotechnol Prog 34:5–28 . doi: 10.1002/btpr.2581
Satyanarayana SDV, Krishna MSR, Kumar PP (2017) Optimization of high-yielding protocol for DNA extraction from the forest rhizosphere microbes. 3 Biotech 7:91 . doi: 10.1007/s13205-017-0737-2
Sayali K, Sadichha P, Surekha S (2013) Microbial esterases : An overview. Int J Curr Microbiol Appl Sci 2:135–146
Sayer C, Finnigan W, Isupov MN, Levisson M, Kengen SWM, van der Oost J, Harmer NJ, Littlechild JA (2016) Structural and biochemical characterisation of Archaeoglobus fulgidus esterase reveals a bound CoA molecule in the vicinity of the active site. Sci Rep 6:25542 . doi: 10.1038/srep25542
Schloss PD, Hay AG, Wilson DB, Walker LP (2003) Tracking temporal changes of bacterial community fingerprints during the initial stages of composting. FEMS Microbiol Ecol 46:1–9 . doi: 10.1016/S0168-6496(03)00153-3
Schmidt-Dannert C, Rua ML, Atomi H, Schmid RD (1996) Thermoalkalophilic lipase of Bacillus thermocatenulatus. I. Molecular cloning, nucleotide sequence, purification and some properties. Measurement 1301:105–114
Schulz G, Schneider D, Brinkmann N, Edy N, Daniel R, Polle A, Scheu S, Krashevska V (2019) Changes in trophic groups of protists with conversion of rainforest into rubber and oil palm plantations. Front Microbiol 10:240 . doi: 10.3389/fmicb.2019.00240
Secundo F, Carrea G (2002) Lipase activity and conformation in neat organic solvents. J Mol Catal B Enzym 19:93–102 . doi: 10.1016/S1381-1177(02)00155-8
Selvin J, Kennedy J, Lejon DPH, Kiran GS, Dobson ADW (2012) Isolation identification and biochemical characterization of a novel halo-tolerant lipase from the metagenome of the marine sponge Haliclona simulans. Microb Cell Fact 11:72 . doi: 10.1186/1475-2859-11-72
Seo S, Lee YS, Yoon SH, Kim SJ, Cho JY, Hahn BS, Koo BS, Lee CM (2014) Characterization of a novel cold-active esterase isolated from swamp sediment metagenome. World J Microbiol Biotechnol 30:879–886 . doi: 10.1007/s11274-013-1496-9
Seshadri R, Kravitz SA, Smarr L, Gilna P, Frazier M (2007) CAMERA: A community resource for metagenomics. PLoS Biol 5:e75 . doi: 10.1371/journal.pbio.0050075
Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: A software environment for integrated models of biomolecular interaction networks. Genome Res 13:2498–2504 . doi: 10.1101/gr.1239303
Shao H, Xu L, Yan Y (2013) Isolation and characterization of a thermostable esterase from a metagenomic library. J Ind Microbiol Biotechnol 40:1211–1222 . doi: 10.1007/s10295-013-1317-z
Silva CC, Hayden H, Sawbridge T, Mele P, Kruger RH, Rodrigues MVN, Costa GGL, Vidal RO, Sousa MP, Torres APR, Santiago VMJ, Oliveira VM (2012) Phylogenetic and functional diversity of metagenomic libraries of phenol degrading sludge from petroleum refinery wastewater treatment system. AMB Express 2:18 . doi: 10.1186/2191-0855-2-18

Simon C, Daniel R (2009) Achievements and new knowledge unraveled by metagenomic approaches. Appl Microbiol Biotechnol 85:265–276 . doi: 10.1007/s00253-009-2233-z
Simon C, Daniel R (2011) Metagenomic analyses: Past and future trends. Appl Environ Microbiol 77:1153–1161 . doi: 10.1128/AEM.02345-10
Sivakumar N, Li N, Tang JW, Patel BKC, Swaminathan K (2006) Crystal structure of AmyA lacks acidic surface and provide insights into protein stability at poly-extreme condition. FEBS Lett 580:2646–2652 . doi: 10.1016/J.FEBSLET.2006.04.017
Skewes-Cox P, Sharpton TJ, Pollard KS, DeRisi JL (2014) Profile Hidden Markov Models for the detection of viruses within metagenomic sequence data. PLoS One 9:e105067 . doi: 10.1371/journal.pone.0105067
Sood S, Sharma A, Sharma N, Kanwar SS (2016) Carboxylesterases: Sources, characterization and broader applications. Insights Enzym Res 1:1–11 . doi: 10.21767/2573-4466.100002
Stergiou PY, Foukis A, Filippou M, Koukouritaki M, Parapouli M, Theodorou LG, Hatziloukas E, Afendra A, Pandey A, Papamichael EM, Moh B, Salameh DA, Gupta R, Gupta N, Rathi P, Levisson M, Mrugesh D, Togneri PD, Anobom CD, Pinheiro AS, De-Andrade RA, Aguieiras ECG, Andrade GC, Moura M V, Almeida R V, Freire DM, Nagazawa AA, Nalder TD, Bornscheuer UT, Dheeman DS, Henehan PG, Lopez-Lopez O, Cerdan M, Siso M, Alex D, Udatha G (2014) Biochemical and Structural Analysis of Thermostable Esterases. Biotechnol Adv 31:445–455 . doi: 10.1016/S0168-6445(01)00075-4
Tadeo X, Ló Pez-Mé Ndez B, Trigueros T, Laín A, Castañ D, Millet O (2009) Structural basis for the aminoacid composition of proteins from halophilic Archea. PLoS Biol 7:e1000257 . doi: 10.1371/journal.pbio.1000257
Tannous E, Kanaya S (2014) Divalent metal ion-induced folding mechanism of RNase H1 from extreme halophilic archaeon Halobacterium sp. NRC-1. PLoS One 9:e109016 . doi: 10.1371/journal.pone.0109016
Tasse L, Bercovici J, Pizzut-Serin S, Robe P, Tap J, Klopp C, Cantarel BL, Coutinho PM, Henrissat B, Leclerc M, Doré J, Monsan P, Remaud-Simeon M, Potocki-Veronese G (2010) Functional metagenomics to mine the human gut microbiome for dietary fiber catabolic enzymes. Genome Res 20:1605–1612 . doi: 10.1101/gr.108332.110
Tatusov RL, Galperin MY, Natale DA, Koonin E V (2000) The COG database: a tool for genome-scale analysis of protein functions and evolution
Tchigvintsev A, Tran H, Popovic A, Kovacic F, Brown G, Flick R, Hajighasemi M, Egorova O, Somody JC, Tchigvintsev D, Khusnutdinova A, Chernikova TN, Golyshina O V., Yakimov MM, Savchenko A, Golyshin PN, Jaeger KE, Yakunin AF (2015) The environment shapes microbial enzymes: five cold-active and salt-resistant carboxylesterases from marine metagenomes. Appl Microbiol Biotechnol 99:2165–2178 . doi: 10.1007/s00253-014-6038-3
Techtmann SM, Hazen TC (2016) Metagenomic applications in environmental monitoring and bioremediation. J. Ind. Microbiol. Biotechnol. 43:1345–1354
Thai QK, Bös F, Pleiss J (2009) The lactamase engineering database: A critical survey of TEM sequences in public databases. BMC Genomics 10:390 . doi: 10.1186/1471-2164-10-390
Thakur S (2012) Lipases, its sources, properties and applications: A review. Int J Sci Eng Res 3:1–29
Tirawongsaroj P, Sriprang R, Harnpicharnchai P, Thongaram T, Champreda V, Tanapongpipat S, Pootanakit K, Eurwilaichitr L (2008) Novel thermophilic and thermostable lipolytic enzymes from a Thailand hot spring metagenomic library. J Biotechnol 133:42–49 . doi: 10.1016/J.JBIOTEC.2007.08.046
Tripathi NK, Shrivastava A (2019) Recent developments in bioprocessing of recombinant proteins: expression hosts and process development. Front Bioeng Biotechnol 7:420 . doi: 10.3389/fbioe.2019.00420
Troy SM, Nolan T, Kwapinski W, Leahy JJ, Healy MG, Lawlor PG (2012) Effect of sawdust addition on composting of separated raw and anaerobically digested pig manure. J Environ Manage 111:70–77 . doi: 10.1016/j.jenvman.2012.06.035
Turnbaugh PJ, Gordon JI (2009) The core gut microbiome, energy balance and obesity. J Physiol 587:4153–4158 . doi: 10.1113/jphysiol.2009.174136
Turnbaugh PJ, Hamady M, Yatsunenko T, Cantarel BL, Ley RE, Sogin ML, Jones WJ, Roe B a, Jason P, Egholm M, Henrissat B, Heath AC, Knight R, Gordon JI, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, Magris M, Hidalgo G, Baldassano RN, Anokhin AP, Heath AC, Warner B, Reeder J, Kuczynski J, Caporaso JG, Lozupone C a., Lauber C, Clemente JC, Knights D, Knight R, Gordon JI (2009) A core gut microbiome between lean and obesity twins. Nature 457:480–484 . doi: 10.1038/nature07540
Uchiyama T, Miyazaki K (2009) Functional metagenomics for enzyme discovery: challenges to efficient screening. Curr Opin Biotechnol 20:616–622 . doi: 10.1016/j.copbio.2009.09.010
Ufarté L, Laville É, Duquesne S, Potocki-Veronese G (2015a) Metagenomics for the discovery of pollutant degrading enzymes. Biotechnol Adv 33:1845–1854 . doi: 10.1016/j.biotechadv.2015.10.009

Ufarté L, Potocki-Veronese G, Laville É (2015b) Discovery of new protein families and functions: New challenges in functional metagenomics for biotechnologies and microbial ecology. Front. Microbiol. 6:563
Uritskiy G, Di Ruggiero J (2019) Applying genome-resolved metagenomics to deconvolute the halophilic microbiome. Genes (Basel) 10:220 . doi: 10.3390/genes10030220
Vargas-Albores F, Martínez-Córdova LR, Martínez-Porchas M, Calderón K, Lago-Lestón A (2019) Functional metagenomics: a tool to gain knowledge for agronomic and veterinary sciences. Biotechnol Genet Eng Rev 35:69–91 . doi: 10.1080/02648725.2018.1513230
Varma VS, Dhamodharan K, Kalamdhad AS (2018) Characterization of bacterial community structure during in-vessel composting of agricultural waste by 16S rRNA sequencing. 3 Biotech 8:301 . doi: 10.1007/s13205-018-1319-7
Ventosa A, Nieto JJ, Oren A (1998) Biology of moderately halophilic aerobic bacteria. Microbiol Mol Biol Rev 62:504–544
Verma N, Thakur S, Bhatt AK (2012) Microbial lipases: Industrial applications and properties (a review). Int Res J Biol Sci Int Res J Biol Sci 1:2278–3202 . doi: 10.4172/2375-4338.1000156
Verschueren KHG, Seljee F, Rozeboom HJ, Kalk KH, Dijkstra BW (1993) Crystallographic analysis of the catalytic mechanism of haloalkane dehalogenase. Nature 363:693–698 . doi: 10.1038/363693a0
Vieille C, Zeikus GJ (2001) Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. Microbiol Mol Biol Rev 65:1–43 . doi: 10.1128/MMBR.65.1.1-43.2001
Vigneron A, Alsop EB, Lomans BP, Kyrpides NC, Head IM, Tsesmetzis N (2017) Succession in the petroleum reservoir microbiome through an oil field production lifecycle. ISME J 11:2141–2154 . doi: 10.1038/ismej.2017.78
Wagner UG, Petersen EI, Schwab H, Kratky C (2002) EstB from Burkholderia gladioli: A novel esterase with a β-lactamase fold reveals steric factors to discriminate between esterolytic and β-lactam cleaving activity. Protein Sci 11:467–478 . doi: 10.1110/ps.33002
Wang C, Dong D, Wang H, Müller K, Qin Y, Wang H, Wu W (2016a) Metagenomic analysis of microbial consortia enriched from compost : new insights into the role of Actinobacteria in lignocellulose decomposition. Biotechnol Biofuels 9:22 . doi: 10.1186/s13068-016-0440-2
Wang G, Wang Q, Lin X, Ng TB, Yan R, Lin J, Ye X (2016b) A novel cold-adapted and highly salt-tolerant esterase from Alkalibacterium sp. SL3 from the sediment of a soda lake. Sci Rep 6:19494 . doi: 10.1038/srep19494
Wang H, Hart DJ, An Y (2019) Functional metagenomic technologies for the discovery of novel enzymes for biomass degradation and biofuel production. Bioenergy Res 12:457–470 . doi: 10.1007/s12155-019-10005-w
Wang M, Ai L, Zhang M, Wang F, Wang C (2020) Characterization of a novel halotolerant esterase from Chromohalobacter canadensis isolated from salt well mine. 3 Biotech 10:430 . doi: 10.1007/s13205-020-02420-0
Wang S, Wang K, Li L, Liu Y (2013) Isolation and characterization of a novel organic solvent-tolerant and halotolerant esterase from a soil metagenomic library. J Mol Catal B Enzym 95:1–8 . doi: 10.1016/j.molcatb.2013.05.015
Wang X, Qin X, Li D, Yang B, Wang Y (2017) One-step synthesis of high-yield biodiesel from waste cooking oils by a novel and highly methanol-tolerant immobilized lipase. Bioresour Technol 235:18–24 . doi: 10.1016/j.biortech.2017.03.086
Wei H, Tucker MP, Baker JO, Harris M, Luo Y, Xu Q, Himmel ME, Ding SY (2012) Tracking dynamics of plant biomass composting by changes in substrate structure, microbial community, and enzyme activity. Biotechnol Biofuels 5:1–14 . doi: 10.1186/1754-6834-5-20
Wei H, Wang L, Hassan M, Xie B (2018) Succession of the functional microbial communities and the metabolic functions in maize straw composting process. Bioresour Technol 256:333–341 . doi: 10.1016/j.biortech.2018.02.050
Wei T, Feng S, Shen Y, He P, Ma G, Yu X, Zhang F, Mao D (2013) Characterization of a novel thermophilic pyrethroid-hydrolyzing carboxylesterase from Sulfolobus tokodaii into a new family. J Mol Catal B Enzym 97:225–232 . doi: 10.1016/j.molcatb.2013.07.022
Wende A, Johansson P, Vollrath R, Dyall-Smith M, Oesterhelt D, Grininger M (2010) Structural and biochemical characterization of a halophilic archaeal alkaline phosphatase. J Mol Biol 400:52–62 . doi: 10.1016/J.JMB.2010.04.057
Wilke A, Harrison T, Wilkening J, Field D, Glass EM, Kyrpides N, Mavrommatis K, Meyer F (2012) The M5nr: A novel non-redundant database containing protein sequences and annotations from multiple sources and associated tools. BMC Bioinformatics 13:141 . doi: 10.1186/1471-2105-13-141

Wu C, Sun B (2009) Identification of novel esterase from metagenomic library of Yangtze River. J Microbiol Biotechnol 19:187–193 . doi: 10.4014/jmb.0804.292

Wu G, Wu G, Zhan T, Shao Z, Liu Z (2013a) Characterization of a cold-adapted and salt-tolerant esterase from a psychrotrophic bacterium Psychrobacter pacificensis. Extremophiles 17:809–819 . doi: 10.1007/s00792-013-0562-4

Wu G, Zhang S, Zhang H, Zhang S, Liu Z (2013b) A novel esterase from a psychrotrophic bacterium Psychrobacter celer 3Pb1 showed cold-adaptation and salt-tolerance. J Mol Catal B Enzym 98:119–126 . doi: 10.1016/j.molcatb.2013.10.012

Xiao F, Li Z, Pan L (2017) Application of microbial lipase and its research progress. Prog Appl Microbiol 8–14

Yamada Y, Fujiwara T, Sato T, Igarashi N, Tanaka N (2002) The 2.0 A crystal structure of catalase-peroxidase from Haloarcula marismortui. Nat Struct Biol 9:691–695 . doi: 10.1038/nsb834

Yang X, Wu L, Xu Y, Ke C, Hu F, Xiao X, Huang J (2018) Identification and characterization of a novel alkalistable and salt-tolerant esterase from the deep-sea hydrothermal vent of the East Pacific Rise. Microbiology 7:e00601 . doi: 10.1002/mbo3.601

Yoo W, Kim B, Jeon S, Kim KK, Kim TD (2020) Identification, characterization, and immobilization of a novel YbfF esterase from Halomonas elongata. Int J Biol Macromol 165:1139–1148 . doi: 10.1016/j.ijbiomac.2020.09.247

Zarafeta D, Moschidi D, Ladoukakis E, Gavrilov S, Chrysina ED, Chatziioannou A, Kublanov I, Skretas G, Kolisis FN (2016) Metagenomic mining for thermostable esterolytic enzymes uncovers a new family of bacterial esterases. Sci Rep 6:38886 . doi: 10.1038/srep38886

Zeth K, Offermann S, Essen L-O, Oesterhelt D (2004) Iron-oxo clusters biomineralizing on protein surfaces: structural analysis of Halobacterium salinarum DpsA in its low- and high-iron states. Proc Natl Acad Sci U S A 101:13780–5 . doi: 10.1073/pnas.0401821101

Zhang S, Lin X (2019) CARMA3: Scaffold protein involved in NF-κB signaling. Front Immunol 10:176 . doi: 10.3389/fimmu.2019.00176

Zhang W, Xu H, Wu Y, Zeng J, Guo Z, Wang L, Shen C, Qiao D, Cao Y (2018a) A new cold-adapted, alkali-stable and highly salt-tolerant esterase from Bacillus licheniformis. Int J Biol Macromol 111:1183–1193 . doi: 10.1016/J.IJBIOMAC.2018.01.152

Zhang Y, Hao J, Zhang YQ, Chen XL, Xie B Bin, Shi M, Zhou BC, Zhang YZ, Li PY (2017) Identification and characterization of a novel salt-tolerant esterase from the deep-sea sediment of the South China Sea. Front Microbiol 8:441 . doi: 10.3389/fmicb.2017.00441

Zhang Y, Ji F, Wang J, Pu Z, Jiang B, Bao Y (2018b) Purification and characterization of a novel organic solvent-tolerant and cold-adapted lipase from Psychrobacter sp. ZY124. Extremophiles 22:287–300 . doi: 10.1007/s00792-018-0997-8

Zhang Y, Pengjun S, Wanli L, Kun M, Yingguo B, Guozeng W, Zhichun Z, Bin Y (2009) Lipase diversity in glacier soil based on analysis of metagenomic DNA fragments and cell culture. J Microbiol Biotechnol 19:888–897 . doi: 10.4014/jmb.0812.695

Zhou G, Xu X, Qiu X, Zhang J (2019) Biochar influences the succession of microbial communities and the metabolic functions during rice straw composting with pig manure. Bioresour Technol 272:10–18 . doi: 10.1016/j.biortech.2018.09.135

Zhou H, Gu W, Sun W, Hay AG (2018) A microbial community snapshot of windrows from a commercial composting facility. Appl Microbiol Biotechnol 102:8069–8077 . doi: 10.1007/s00253-018-9201-4

Zhu Y, Li J, Cai H, Ni H, Xiao A, Hou L (2013) Characterization of a new and thermostable esterase from a metagenomic library. Microbiol Res 168:589–597 . doi: 10.1016/J.MICRES.2013.04.004

Chapter V

General discussion

Owing to the functional versatility, LEs are one of the most industrially interested groups of biocatalysts. Particularly, LEs of thermophilic origin gain more attention due to their intrinsic thermal and chemical stability (Stergiou et al. 2014). This thesis describes the results of applying function-driven and sequence-based metagenomic approaches to explore novel LEs in microbial consortia of thermophilic composts. The information obtained in this study illustrated an insight of the genetic diversity of LEs, which have broadened the spectrum of LEs, with the aim to provide more potentials for industrial application.

5.1 Microbial composition in compost consortia

Firstly, this thesis presents the phylogenetic profile of microbial consortia in the two compost samples. Generally, the microbial communities revealed from the pyrotag sequencing data and shotgun whole metagenomic sequencing data were accordance with those in other studies: members in phyla *Actinobacteria*, *Proteobacteria*, *Firmicutes* and *Bacteroidetes* were dominant in compost microbial consortia (Antunes et al. 2016; Ma et al. 2018; Palaniveloo et al. 2020). The incisive taxonomic disparity were found among compost samples (Fig. 5-1), which was the combination of many factors, such as C/N ratio, pH, moisture content, aeration/O2 supply, temperature and particle size (Rastogi et al. 2020). Under aerobic conditions, temperature and the availability of various organic matters are the major selective factors for the microbial population (Ryckeboer et al. 2003). Particularly, the temperature determines the rate metabolic activity (Ma et al. 2018; Rastogi et al. 2020). Antunes et al. (2016) tracked the dynamic change of microbial community throughout the composting process at metagenomic and metatranscriptomic level. Significantly, the rising temperature favors for members in *Firmicutes* that out-competed other organisms, and were dominant over the thermophilic stage. This was found even in different composting systems feed with various organic wastes (Fig.

5-1). During the thermophilic phase, many microbes in *Firmicutes* are able to form endospore-like structures to resist the heat (Wei et al. 2018; Zhou et al. 2019; Awasthi et al. 2020a). They are also reported to hydrolyze sugar, protein and lignocellulose that resulted in a variety of acids, alcohols and lipids (Mao et al. 2018). For members in *Actinobacteria*, they are important degraders for natural polymers and hard to degrade debris, e.g. woody stems, bark or newspaper, and optimally grow at temperature between 45 to 55 °C (Ryckeboer et al. 2003; Wei et al. 2012). Wang et al. (2016) verified that members of the *Actinobacteria* play an essential role in decomposing cellulose and hemicellulose originating from the rice straw. In this study, the excessive heat (76 °C) by the compost pile (Pile_2) with household waste as feedstock, could be associated with high sugar content and water-soluble carbon (Troy et al. 2012; Jara-Samaniego et al. 2017). The other pile (Pile_1) was mainly supplemented with wood chips (i.e. high C/N ratio). A higher C/N ratio (compare to Pile_2) could slowed down composting, with simultaneously reach a lower temperature (55 °C) during the compost process. Therefore, under the same composting operation (i.e. turning regime), the temperature and feedstock composition primarily shape the microbial community in the two compost samples (Fig. 5-1).

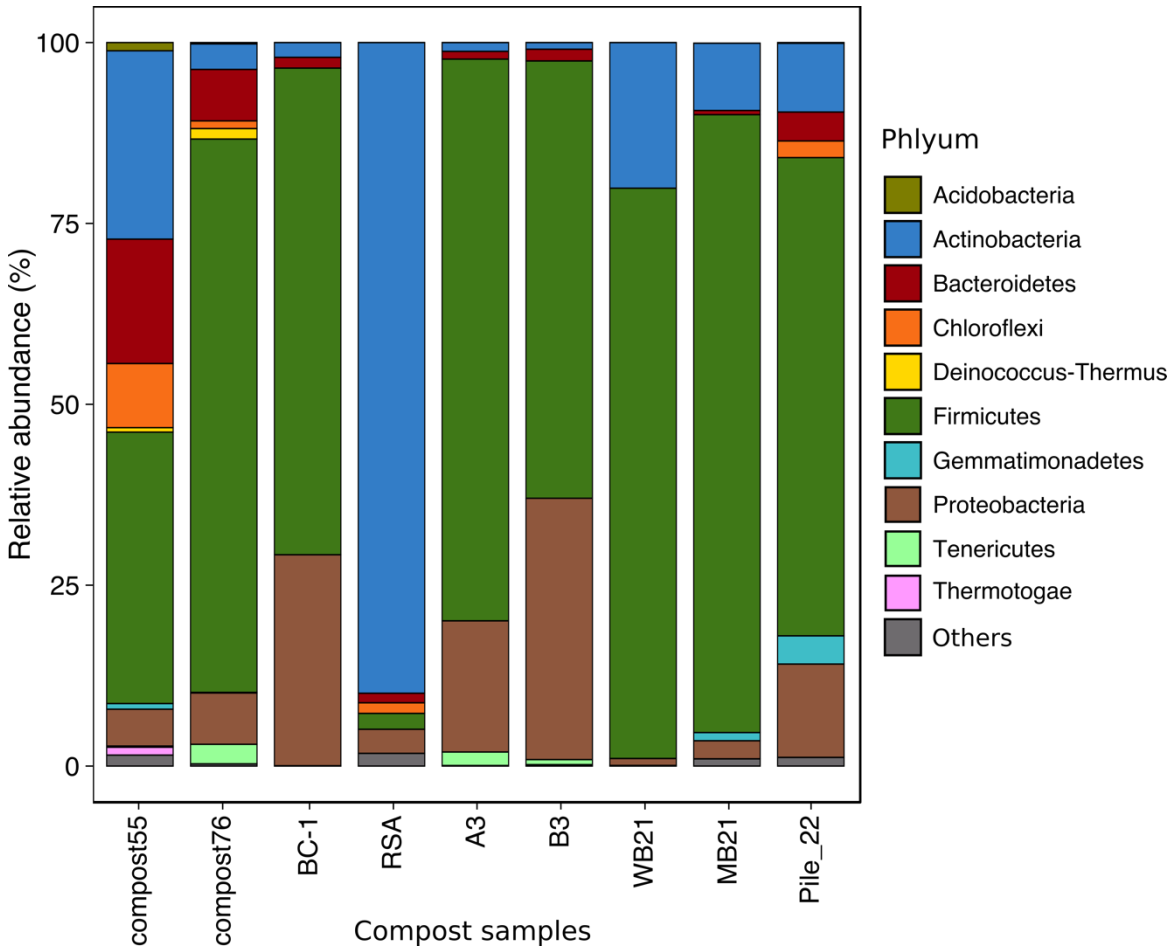


Fig. 5-1. Phylogenetic distribution of dominant bacterial groups (at phylum level) in compost samples investigated in this thesis and other surveys. Compost samples of compost55 (feedstock: green waste) and compost76 (household waste) were from this thesis and at thermophilic stage; BC-1 (manure and sawdust) was at thermophilic stage with temperature at 67.8 °C (Mao et al. 2018); RSA was at maturing stage with temperature at 30 °C (Wang et al. 2016a); A3 (pig manure and wheat straw) and B3 (pig manure and wheat straw covered with a semi-permeable membrane) were at thermophilic stage with temperature at 58 and 57 °C, respectively (Ma et al. 2018); WB21 and MB21 were at cooling/second mesophilic phase stage with temperature at 44 and 45 °C, respectively (Zhou et al. 2019); Pile_22 (manure/silage and hay) was at thermophilic stage (Neher et al. 2013).

Proteobacteria were the major group for decomposing plant straw (Awasthi et al., 2017), and some members of the *Proteobacteria* are known to be involved in sulfur and nitrogen cycling (Zhong et al., 2017; Wei et al., 2018). Many microbes in *Bacteroidetes* were important degraders for organic matter, as they have the ability

to break down cellulose, hemicellulose and lignin (Dodd et al., 2011; Zhong et al., 2017).

At genus level, genera such as *Brockia*, *Thermobispora*, *Thermomonospora*, *Longispora*, *Salinispora*, *Geobacillus*, *Thermaerobacter*, *Filomicrobium*, *Rhodothermus*, *Symbiobacterium*, *Caldicoprobacter* and *Calditerricola* were among major bacterial groups previously identified during composting (Ryckeboer et al. 2003; Antunes et al. 2016; Rong et al. 2018; Zhou et al. 2018). Among them, *Calditerricola*, *Geobacillus, Thermobispora*, *Thermus* and *Thermomonospora* were typical bacterial thermophiles. At temperatures above 60 °C, the degradation of organic compounds such as cellulose, hemicellulose and chitin is performed essentially by these thermophilic bacteria (Moriya et al. 2011; Awasthi et al. 2020a).

5.2 Screening for lipolytic enzymes through metagenomic approaches

Functional metagenomics is activity-based screening of metagenomic libraries that constructed by cloning environmental DNA into expression vectors and propagating them in the appropriate hosts (Daniel 2005). This technique is widely used for novel biocatalyst identification, due to no requiring the cultivation of the native microorganisms or sequence information of the genes *a priori* (DeCastro et al. 2016). Attributing to the functional versatility, lipolytic enzymes are one of the most important groups in industrial applications (Rosenstein and Götz 2000). As summarized by Ferrer et al. (2015) of recent reported 256 metagenomic libraries, 68 % of the 6,038 clones that were active against different tested substrates were lipases and/or esterases. Among them, 4,034 were selected by functional screening of metagenomes from various habitats, such as soil, marine and freshwater environments, microbiota from several hosts, and waste treatments (Lopez-Lopez et al. 2014; Berini et al. 2017). The thermophilic compost samples harbor divergent thermophiles, from which the enzymes are intrinsic thermophilic. Thus, this thesis

used the function-based approach for identifying potential extremophilic lipolytic genes in two thermophilic compost metagenomes. Overall, approximately 4.89 and 2.56 Gb of cloned compost DNA were screened, generating 199 and 51 positive clones for compost55 and compost76, respectively. The targeting probability towards a LE in our study (between 16.1 and 43.6 per Gb) is generally consistent with those from other compost metagenomic libraries (Lämmle et al. 2007; Kim et al. 2010; Leis et al. 2015). The hit rates were affected by many factors, such as the source of the metagenomic DNA, the size of the gene of interest, its abundance in the metagenome, biases in cell lysis during DNA extraction, homegroups expression host and vector, and the expression condition (Uchiyama and Miyazaki 2009).

To date, *E. coli* has been chosen as the cloning host for the vast majority of metagenomic libraries (Uchiyama and Miyazaki 2009), as it possesses a number of desirable attributes, e.g. high transformation efficiency, lacks genes for restriction modification and can deal with diverse translation signals (Boni 2006). Despite these advantages, *E. coli* is unable to express all foreign DNA due to the incompatibility with the heterologous expression system (Lam et al. 2015). As Gabor et al. (2004) estimated, on averagely, only 30%–40% of environmental bacterial genes could be efficiently expressed in *E. coli*, the value even drops to 7% for high G + C DNA from bacterial strains. Moreover, gene product from extremophiles were normally active under its original extreme conditions (Munawar and Engel 2013), whereas, mesophiles, like the expression host *E. coli*, do not survive under these conditions. In the future, to enhance the hit rate, the functional screening can be improved by modifying the *E. coli*-based expression system, such as introducing heterologous sigma factors (Gaida et al. 2015), employing T7 RNA polymerase and forming hybrid ribosomes (Gaida et al. 2015). Alternatively, various hosts (Tripathi and Shrivastava 2019) such as *Haloferax volcanii* (Kixmüller and Greie 2012),

Halobacterium salinarum (Kixmüller and Greie 2012), *Thermus thermophilus* (Leis et al. 2015) and *Bacillus subtilis* (Mahapatra et al. 2020) can be chosen for production of extremophilic gene products .

These functionally derived LEs were further classified into different families according to the gold standard classification system initiated by Arpigny & Jaeger (1999), which has extended to nineteen families nowadays. Until recently, there are still claims of identifying novel families, however, they are not integrated into the classification system. Without comparison of these so-called novel families, the possibility exists that different publications may describe members of the same 'novel' family under different names (Hitch and Clavel 2019). In this study, after extensive literature search, we integrated 29 so-called novel families into the classification system, resulting in a total of 48 lipolytic family (subfamilies in one family were counted as one). According to the phylogenetic tree, the functional-derived LEs were detected in 12 known families and 7 novel families.

Sequence-based screening relies on direct sequencing of environmental DNA, either with or without cloning prior to sequencing, and then subjecting the sequences to bioinformatic analyses. In comparison to the function-based screening, this strategy avoids the cloning-related disadvantages. In addition, sequence-based screening involves less labor-intensive operations than function-based screening (Ngara and Zhang 2018). The sequence-driven approach highly depends on the accuracy and completeness of the reference databases, from which the functions of newly discovered genes are inferred. Thus, if the novel genes have weak sequence similarity to sequences recruited in the database, or the sequence similarity does not correspond to a functional relationship, this approach may return false annotation. Interestingly, in practice, function-based approach was generally more favorable than sequence-based approach. According to Ferrer et al. (2015), 6,100 of the

metagenomic identified bioactive compounds (from 2,192 different sites), only 3.5 % were derived from sequence-based approach.

In the specific case of lipases, only a few studies used sequence-based approaches for LEs identification. For example, Lipases were identified by PCR amplification of metagenomic DNA from a hot spring (Bell et al. 2002) and glacier soil (Zhang et al. 2009), using primers targeting the consensus sequences of the oxyanion hole and regions of active sites. Zarafeta et al. (2016) sequenced the whole-genome of microbial consortia in a water sludge, they found an esterase (EstDZ2) by screening the processed data against the conserved motifs at the active sites, oxyanion hole and signal peptide. Verma et al. (2019) revealed a total 21 unique sequences of new alkaliphilic lipases by searching against the reconstructed ancestral sequence of an alkaliphilic lipase.

In this thesis, a LE-specific profile HMM database was built for sequence-based screening of putative lipolytic genes. Profile HMMs have been widely adopted for improved annotation of general functions in microbial genomes and metagenomes (Skewes-Cox et al. 2014; Reyes et al. 2017; Bzhalava et al. 2018). However, they have not yet been specifically applied to LEs in microbial genomes/metagenomes. In this study, to gain high recall (sensitivity), we compared 4 sets of profile HMMs constructed using different multiple sequence alignment algorithms. The high specificity was achieved by a subsequent BLAST searching after the scan against the selected profile HMM database. This strategy for putative lipolytic gene identification and subsequent family assignment was further validated on different datasets. Using this strategy, we successfully identified 4,157 and 2,234 putative lipolytic proteins (PLPs), among them, 1,234 and 759 were further assigned into 28 and 26 lipolytic families in the assembled metagenomes of compost55 and compost76, respectively. The assigned PLPs generally originated from Bacteria

(>95 %), and mainly affiliated to the phyla *Actinobacteria*, *Proteobacteria*, *Firmicutes* and *Bacteroidetes*. Moreover, we identified a link between microbial taxa and their functional traits (lipolytic genes), which was also reported in studies of profiling CAZyme encoding genes (Wang et al. 2016a) and resistomes (Pehrsson et al. 2016) through sequence-based approaches.

Furthermore, we compared LEs identified in the plasmid library with putative lipolytic proteins derived from the sequence-based screening of the assembled compost55 and compost76 metagenomes. All in all, function-based approach and sequence-based approach has its respective advantages and disadvantages (Table 5-1). The most distinctive advantage of the function-based screening is the certainty that novel compound is correctly biosynthesized by the host cell, and hence, discovery of novel enzymes with no sequence information is required (Ngara and Zhang 2018). However, since this approach is activity-directed, one of the main limitations is the low expression level of heterologous genes in the host (Mirete et al. 2016), which normally led to a low hit rate (Simon and Daniel 2011). Sequencing-based screening is effective in identifying targeted biomolecules in metagenomes, with similarities to known biomolecules. However, sequence-based screening largely relies on the search algorithms, and quality and completeness of reference databases, which are used to infer the functions of newly discovered genes (Lam et al. 2015). Thus, the best way to explore novel molecules is to combine the two approaches (Barriuso and Jesús Martínez 2015). The information obtained from functional metagenomics can broaden the database for future annotation, thus, complements sequence-based metagenomics.

Table 5-1 Comparison of function-driven and sequence-based screening for lipolytic genes

Metagenomic approach	Method	Advantages	Limitations	Future prospects	References
Function-driven	• Hydrolysis of natural triacylglycerols: tributyrin, tricaprylin, triolein, olive oil, formation of hydrolysis halos • Hydrolysis of 1-naphthyl palmitate, *p*-nitrophenyl esters, formation of colored colonies	• Culture-independent • No requirement of sequence information *a priori* • Activity-based, thus, direct detection of bioactive LEs	• Low hit rate • Labor-intensive operations • Only qualitative • Low/no level of gene expression in the most widely used host *E. coli*	• Enrichment of environmental samples under conditions mimicking the application • High-throughput screening strategies, such as fluorescence-activated cell sorting (FACS), substrate induced gene expression (SIGEX), genetic enzyme screening system (GESS) and microfluidics-driven screening • Modification of expression hots *E. coli*, such as random insertion of T7 promoters; introducing heterologous sigma factors; forming hybrid ribosomes • Alternative expression systems, such as *Agrobacterium*, *Bacillus*, *Rhodococcus*, *Streptomyces* and *Pseudomonas*, as well as a few archaea	Gabor et al. 2004; Liebl et al. 2014; Gaida et al. 2015; Kim et al. 2016; Berini et al. 2017; Ngara and Zhang 2018; Mahapatra et al. 2020; Markel et al. 2020
Sequence-based	• Target sequencing: PCR amplification with primers specific for conserved regions of genes of interest • Shotgun sequencing: such as motif- and homology-based searching against database	• Culture-independent • No requirement of cloning *a priori* • Easy access to metagenomic sequencing data nowadays • Plenty of software and platforms for data analyzing	• Highly dependent on the accuracy and completeness of the refence database • Relying on the algorithms available • Genes predicted *in silico* require experimental confirmation • Tend to find proteins with high homologies with known amino acid sequences	• Synthetic metagenomics: identified genes of interest are subjected to high-throughput expression system for further characterization • Constant improvement of database and database annotations	Simon and Daniel 2011; Ufarté et al. 2015a; Ferrer et al. 2016; Quince et al. 2017; Awasthi et al. 2020b

5.3 Comparative analysis of distribution of putative lipolytic genes across various habitats

This thesis retrieved 175 assembled metagenomes representing microbial communities of various habitats from IMG/M database. In total, we have screened approx. 1.23 billion genes in 65 Gbp of assembled metagenomes, with approx. 0.22 million (absolute counts) PLP-encoding genes were identified. Among them, the assigned putative lipolytic proteins were account for 34 % (inferred from the absolute counts). The abundance of putative lipolytic proteins in each sample was normalized to LPGM values for comparative analysis. Generally, putative lipolytic proteins were selectively abundant in samples subjected to certain enrichment processes (Fig. 5-2), such as samples from hydrocarbon resource environments and oil reservoir, which are enriched with oil-degrading microbes (Liu et al. 2015; Hu et al. 2016; Vigneron et al. 2017; Liu et al. 2018); composts and wastewater bioreactor, which are reservoirs of various microbes for decomposing organic compounds (Dougherty et al. 2012; Silva et al. 2012; Antunes et al. 2016; Berini et al. 2017). The unequally distribution of PLPs was also observed across different lipolytic families. Families of Hormone-sensitive_lipase_like and patatin-like-protein were most abundant (average LPGM values across samples > 2000), followed by A85-EsteraseD-FGH, VIII and Bacterial_lip_FamI.1 (average LPGM values > 700). The lipolytic family profile and taxonomic profile revealed from the assigned putative lipolytic proteins were clustered by habitats. Thus, putative lipolytic proteins were distributed ecologically across samples.

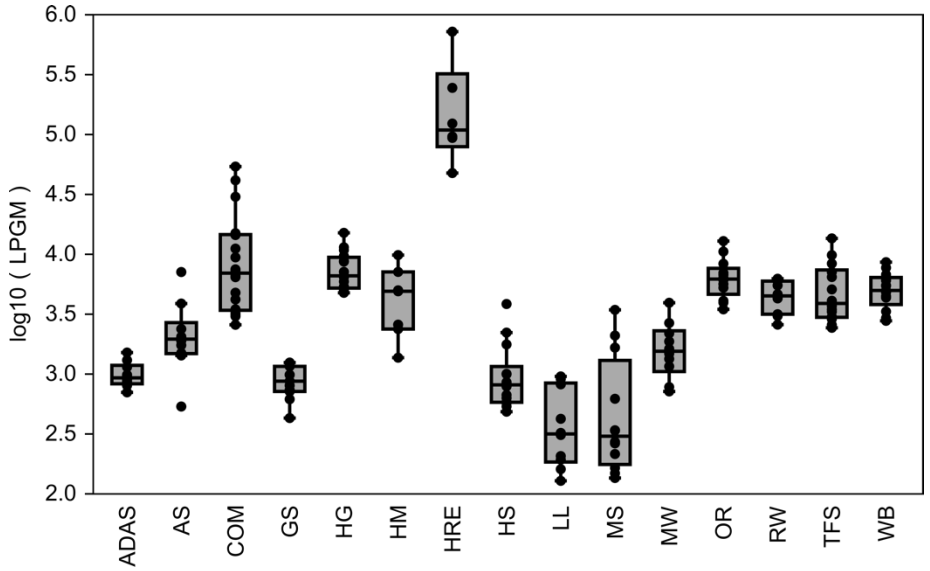


Fig. 5-2: Distribution of assigned putative lipolytic proteins across habitats. LPGM values were $\log_{10}$ transformed. Abbreviations of habitats: ADAS, anaerobic digestor active sludge; AS, agricultural soil; COM, compost; GS grassland soil; HG, human gut; HM, hypersaline mat; HRE, hydrocarbon resource environments; HS, hot spring; LL, landfill leachate; MS, marine sediment; MW, marine water; OR, oil reservoir; RW, river water; TFS, tropical forest soil; WB, wastewater bioreactor.

5.4 Characterization of extremophilic LEs

Finally, three lipolytic genes obtained through functional screening were selected for biochemical characterization. Among them, Est1 and Est2 are thermoalkaline enzymes with optimal enzyme activities at 80 and 70 °C, respectively. Est1 and Est2 also displayed unprecedented thermostability at high temperatures, with a half-life of more than 7 (Est1) and 5 days (Est2) at 50 °C, 2 (Est1) and 1 (Est2) days at 60 °C. Generally, thermostability is dependent on the structural rigidity, which is an accumulation of various features, including but not limited to amino acid composition, ion-pairing, hydrogen bonds, hydrophobic interactions, and sulfide bridges (Sadeghi et al. 2006; Jochens et al. 2010; Ebrahimie et al. 2011; Pezzullo et al. 2013). The modeled three-dimensional structures of Est1 and Est2 were homologous to thermostable esterases, which to some extent, structurally proved their thermostable features (Alcaide et al. 2013; Sayer et al. 2016). As described

before, the thermostability of an enzyme could somehow, extend to its tolerance towards organic-solvent (Khan and Sathya 2018). This is also the case for Est1 and Est2. Particularly for Est2, its activity was significantly enhanced (2- to 10-fold) in the presence of ethanol, methanol, isopropanol and 1-propanol over a concentration range between 6 % and 30 % (v/v). The significant activating effect could be attributed to the uniform water phase formed by water-miscible solvents (Ogino and Ishikawa 2001) or the high diffusion rate of substrate in the presence of water-miscible solvents (Metin et al. 2006), which enables substrates quick and easy access to the active site. With respect to the short-term stability (2 h incubation), Est2 exhibited high tolerance against 60 % (v/v) of ethanol, methanol, isopropanol, DMSO and acetone, while Est1 resisted these solvents only at lower concentrations (below 30 %, v/v). For the long-term stability, Est2 retained most of its activity after 26 days incubation in the presence of 30 % (v/v) ethanol, methanol, isopropanol, DMSO or acetone. Thus, all these features allows for Est1 and Est2, especially Est2, potentially applying in variety of industrial relevant processes, where the reactions are normally carried out at elevated temperature and mediated with different organic solvents (Panda and Gowrishankar 2005; López-López et al. 2015; Sood et al. 2016; Sarmah et al. 2018).

The other esterase (Est56) was tolerant to high salinity. Addition of salt (NaCl or KCl) produced a stimulatory effect on Est56 activity, with an enhanced activity at NaCl and KCl up to 2.5 M. Est56 was stable over the tested NaCl and KCl concentration ranges (0 to 4 M), with almost unaltered activity after 24 h incubation at 10 °C. To shed light on the mechanism of haloadapation of Est56, we summarized the recently reported halotolerant lipolytic enzymes (40 in total), and their amino acid compositions were statistically compared with the halophilic counterparts. However, no clear pattern was found in the amino acid composition of halotolerant

LEs. Actually, LEs with different theoretical pI values apply different mechanisms for haloadapation (Sivakumar et al. 2006; Altermark et al. 2008; De Santi et al. 2016b). Est56 has a low theoretical pI value of 4.97, which is an indication that Est56 apply a salt resistence stategy similar to that of halophilic enzymes (Ishibashi et al. 2013; Munawar and Engel 2013). Similar to other halophilic enzymes, the increased content of acidic residues over lysine residues and a predominantly negatively charged surface were also observed for Est56 (Ventosa et al. 1998; Tadeo et al. 2009; Coquelle et al. 2010; Munawar and Engel 2012; Ishibashi et al. 2013). Moreover, Est56 displayed enhanced stability against high temperatures (50 and 60 ºC) and urea (2, 4, and 6 M) in the presence of NaCl. This distinctive feature of enhanced Est56 stability against denaturants (high temperatures and urea) mediated by NaCl also suggests that Est56 possess a halophilic character caused by the high acidic amino acid content (Coquelle et al. 2010).

5.5 Reference

Alcaide M, Tornés J, Stogios PJ, Xu X, Gertler C, Di Leo R, Bargiela R, Lafraya Á, Guazzaroni M-E, López-Cortés N, Chernikova TN, Golyshina OV, Nechitaylo TY, Plumeier I, Pieper DH, Yakimov MM, Savchenko A, Golyshin PN, Ferrer M (2013) Single residues dictate the co-evolution of dual esterases: MCP hydrolases from the α/β hydrolase family. Biochem J 454:157–166 . doi: 10.1042/BJ20130552

Altermark B, Helland R, Moe E, Willassen NP, Smalås AO (2008) Structural adaptation of endonuclease I from the cold-adapted and halophilic bacterium *Vibrio salmonicida*. Acta Crystallogr Sect D Biol Crystallogr 64:368–376 . doi: 10.1107/S0907444908000097

Antunes LP, Martins LF, Pereira RV, Thomas AM, Barbosa D, Lemos LN, Silva GMM, Moura LMS, Epamino GWC, Digiampietri LA, Lombardi KC, Ramos PL, Quaggio RB, De Oliveira JCF, Pascon RC, Da Cruz JB, Da Silva AM, Setubal JC (2016) Microbial community structure and dynamics in thermophilic composting viewed through metagenomics and metatranscriptomics. Sci Rep 6:38915 . doi: 10.1038/srep38915

Awasthi MK, Duan Y, Liu T, Awasthi SK, Zhang Z (2020a) Relevance of biochar to influence the bacterial succession during pig manure composting. Bioresour Technol 304:122962 . doi: 10.1016/j.biortech.2020.122962

Awasthi MK, Ravindran B, Sarsaiya S, Chen H, Wainaina S, Singh E, Liu T, Kumar S, Pandey A, Singh L, Zhang Z (2020b) Metagenomics for taxonomy profiling : tools and approaches. Bioengineered 11:356–374 . doi: 10.1080/21655979.2020.1736238

Barriuso J, Jesús Martínez M (2015) In silico metagenomes mining to discover novel esterases with industrial application by sequential search strategies. J Microbiol Biotechnol 25:732–737 . doi: 10.4014/jmb.1406.06049

Bell PJL, Sunna A, Gibbs MD, Curach NC, Nevalainen H, Bergquist PL (2002) Prospecting for novel lipase genes using PCR. Microbiology 148:2283–2291 . doi: 10.1099/00221287-148-8-2283

Berini F, Casciello C, Marcone GL, Marinelli F (2017) Metagenomics: novel enzymes from non-culturable microbes. FEMS Microbiol Lett 364. doi: 10.1093/femsle/fnx211

Boni I V. (2006) Diverse molecular mechanisms of translation initiation in prokaryotes. Mol Biol 40:587–596 . doi: 10.1134/S002689330604011X

Bzhalava Z, Hultin E, Dillner J (2018) Extension of the viral ecology in humans using viral profile hidden Markov models. PLoS One 13:e0190938 . doi: 10.1371/journal.pone.0190938

Coquelle N, Talon R, Juers DH, Girard É, Kahn R, Madern D (2010) Gradual adaptive changes of a protein facing high salt concentrations. J Mol Biol 404:493–505 . doi: 10.1016/j.jmb.2010.09.055

Daniel R (2005) The metagenomics of soil. Nat Rev Microbiol 3:470–478 . doi: 10.1038/nrmicro1160

De Santi C, Leiros HKS, Di Scala A, de Pascale D, Altermark B, Willassen NP (2016) Biochemical characterization and structural analysis of a new cold-active and salt-tolerant esterase from the marine bacterium *Thalassospira* sp. Extremophiles 20:323–336 . doi: 10.1007/s00792-016-0824-z

DeCastro ME, Rodríguez-Belmonte E, González-Siso MI (2016) Metagenomics of thermophiles with a focus on discovery of novel thermozymes. Front Microbiol 7:1521 . doi: 10.3389/fmicb.2016.01521

Dougherty MJ, D'haeseleer P, Hazen TC, Simmons BA, Adams PD, Hadi MZ (2012) Glycoside hydrolases from a targeted compost metagenome, activity-screening and functional characterization. BMC Biotechnol 12:38 . doi: 10.1186/1472-6750-12-38

Ebrahimie E, Ebrahimi M, Sarvestani R, Ebrahimi M (2011) Protein attributes contribute to halo-stability, bioinformatics approach. Saline Systems 7:1 . doi: 10.1186/1746-1448-7-1

Ferrer M, Martínez-Martínez M, Bargiela R, Streit WR, Golyshina O V., Golyshin PN (2016) Estimating the success of enzyme bioprospecting through metagenomics: Current status and future trends. Microb. Biotechnol. 9:22–34

Gabor EM, Alkema WBL, Janssen DB (2004) Quantifying the accessibility of the metagenome by random expression cloning techniques. Environ Microbiol 6:879–886 . doi: 10.1111/j.1462-2920.2004.00640.x

Gaida SM, Sandoval NR, Nicolaou SA, Chen Y, Venkataramanan KP, Papoutsakis ET (2015) Expression of heterologous sigma factors enables functional screening of metagenomic and heterologous genomic libraries. Nat Commun 6:1–10 . doi: 10.1038/ncomms8045

Hitch TCA, Clavel T (2019) A proposed update for the classification and description of bacterial lipolytic enzymes. PeerJ 7:e7249 . doi: 10.7717/peerj.7249

Hu P, Tom L, Singh A, Thomas BC, Baker BJ, Piceno YM, Andersen GL, Banfield JF (2016) Genome-resolved metagenomic analysis reveals roles for candidate phyla and other microbial community members in biogeochemical transformations in oil reservoirs. MBio 7:e01669-15 . doi: 10.1128/mBio.01669-15

Ishibashi M, Hayashi T, Yoshida C, Tokunaga M (2013) Increase of salt dependence of halophilic nucleoside diphosphate kinase caused by a single amino acid substitution. Extremophiles 17:585–591 . doi: 10.1007/s00792-013-0541-9

Jara-Samaniego J, Pérez-Murcia MD, Bustamante MA, Pérez-Espinosa A, Paredes C, López M, López-Lluch DB, Gavilanes-Terán I, Moral R (2017) Composting as sustainable strategy for municipal solid waste management in the Chimborazo Region, Ecuador: Suitability of the obtained composts for seedling production. J Clean Prod 141:1349–1358 . doi: 10.1016/j.jclepro.2016.09.178

Jochens H, Aerts D, Bornscheuer UT (2010) Thermostabilization of an esterase by alignment-guided focussed directed evolution. Protein Eng Des Sel 23:903–909 . doi: 10.1093/protein/gzq071

Khan M, Sathya TA (2018) Extremozymes from metagenome: Potential applications in food processing. Crit Rev Food Sci Nutr 58:2017–2025 . doi: 10.1080/10408398.2017.1296408

Kim YH, Kwon EJ, Kim SK, Jeong YS, Kim J, Yun HD, Kim H (2010) Molecular cloning and characterization of a novel family VIII alkaline esterase from a compost metagenomic library. Biochem Biophys Res Commun 393:45–49 . doi: 10.1016/j.bbrc.2010.01.070

Kim YJ, Kim H, Kim SH, Rha E, Choi SL, Yeom SJ, Kim HS, Lee SG (2016) Improved metagenome screening efficiency by random insertion of T7 promoters. J Biotechnol 230:47–53 . doi: 10.1016/j.jbiotec.2016.05.018

Kixmüller D, Greie JC (2012) Construction and characterization of a gradually inducible expression vector for *Halobacterium salinarum*, based on the kdp promoter. Appl Environ Microbiol 78:2100–2105 . doi: 10.1128/AEM.07155-11

Lam KN, Cheng J, Engel K, Neufeld JD, Charles TC (2015) Current and future resources for functional metagenomics. Front Microbiol 6:1196 . doi: 10.3389/fmicb.2015.01196

Lämmle K, Zipper H, Breuer M, Hauer B, Buta C, Brunner H, Rupp S (2007) Identification of novel enzymes with different hydrolytic activities by metagenome expression cloning. J Biotechnol 127:575–592 . doi: 10.1016/j.jbiotec.2006.07.036

Leis B, Angelov A, Mientus M, Li H, Pham VTT, Lauinger B, Bongen P, Pietruszka J, Gonçalves LG, Santos H, Liebl W (2015) Identification of novel esterase-active enzymes from hot environments by use of the host bacterium *Thermus thermophilus*. Front Microbiol 6:275 . doi: 10.3389/fmicb.2015.00275

Liebl W, Angelov A, Juergensen J, Chow J, Loeschcke A, Drepper T, Classen T, Pietruzska J, Ehrenreich A, Streit WR, Jaeger KE (2014) Alternative hosts for functional (meta)genome analysis. Appl. Microbiol. Biotechnol. 98:8099–8109

Liu SP, Liu RX, Zhang L, Shi GY (2015) Sequence-based screening and characterization of cytosolic mandelate oxidase using oxygen as electron acceptor. Enzyme Microb Technol 69:24–30 . doi: 10.1016/j.enzmictec.2014.11.001

Liu YF, Galzerani DD, Mbadinga SM, Zaramela LS, Gu JD, Mu BZ, Zengler K (2018) Metabolic capability and in situ activity of microorganisms in an oil reservoir. Microbiome 6:5 . doi: 10.1186/s40168-017-0392-1

Lopez-Lopez O, Cerdan M, Siso M (2014) New extremophilic lipases and esterases from metagenomics. Curr Protein Pept Sci 15:445–455 . doi: 10.2174/1389203715666140228153801

López-López O, Knapik K, Cerdán ME, González-Siso MI (2015) Metagenomics of an alkaline hot spring in Galicia (Spain): Microbial diversity analysis and screening for novel lipolytic enzymes. Front Microbiol 6:1291 . doi: 10.3389/fmicb.2015.01291

Ma S, Fang C, Sun X, Han L, He X, Huang G (2018) Bacterial community succession during pig manure and wheat straw aerobic composting covered with a semi-permeable membrane under

slight positive pressure. Bioresour Technol 259:221–227 . doi: 10.1016/j.biortech.2018.03.054

Mahapatra GP, Raman S, Nayak S, Gouda S, Das G, Patra JK (2020) Metagenomics approaches in discovery and development of new bioactive compounds from marine Actinomycetes. Curr Microbiol 77:645–656 . doi: 10.1007/s00284-019-01698-5

Mao H, Lv Z, Sun H, Li R, Zhai B, Wang Z, Awasthi MK, Wang Q, Zhou L (2018) Improvement of biochar and bacterial powder addition on gaseous emission and bacterial community in pig manure compost. Bioresour Technol 258:195–202 . doi: 10.1016/j.biortech.2018.02.082

Markel U, Essani KD, Besirlioglu V, Schiffels J, Streit WR, Schwaneberg U (2020) Advances in ultrahigh-throughput screening for directed enzyme evolution. Chem Soc Rev 49:233–262 . doi: 10.1039/c8cs00981c

Metin K, Burcu Bakir Ateslier Z, Basbulbul G, Halil Biyik H (2006) Characterization of esterase activity in Geobacillus sp. HBB-4. J Basic Microbiol 46:400–409 . doi: 10.1002/jobm.200510121

Mirete S, Morgante V, González-Pastor JE (2016) Functional metagenomics of extreme environments. Curr Opin Biotechnol 38:143–149 . doi: 10.1016/j.copbio.2016.01.017

Moriya T, Hikota T, Yumoto I, Ito T, Terui Y, Yamagishi A, Oshima T (2011) *Calditerricola satsumensis* gen. nov., sp. nov. and *Calditerricola yamamurae* sp. nov., extreme thermophiles isolated from a high-temperature compost. Int J Syst Evol Microbiol 61:631–636 . doi: 10.1099/ijs.0.018416-0

Munawar N, Engel PC (2013) Halophilic enzymes: Characteristics, structural adaptation and potential applications for biocatalysis. Curr Biotechnol 2:334–344 . doi: 10.2174/18722083113076660033

Munawar N, Engel PC (2012) Overexpression in a non-native halophilic host and biotechnological potential of NAD+-dependent glutamate dehydrogenase from *Halobacterium salinarum* strain NRC-36014. Extremophiles 16:463–476 . doi: 10.1007/s00792-012-0446-z

Neher DA, Weicht TR, Bates ST, Leff JW, Fierer N (2013) Changes in bacterial and fungal communities across compost recipes, preparation methods, and composting times. PLoS One 8:e79512 . doi: 10.1371/journal.pone.0079512

Ngara TR, Zhang H (2018) Recent advances in function-based metagenomic screening. Genomics, Proteomics Bioinforma. 16:405–415

Ogino H, Ishikawa H (2001) Enzymes which are stable in the presence of organic solvents. J Biosci Bioeng 91:109–116

Palaniveloo K, Amran MA, Norhashim NA, Mohamad-Fauzi N, Peng-Hui F, Hui-Wen L, Kai-Lin Y, Jiale L, Chian-Yee MG, Jing-Yi L, Gunasekaran B, Razak SA (2020) Food waste composting and microbial community structure profiling. Processes 8:723 . doi: 10.3390/pr8060723

Panda T, Gowrishankar BS (2005) Production and applications of esterases. Appl Microbiol Biotechnol 67:160–169 . doi: 10.1007/s00253-004-1840-y

Pehrsson EC, Tsukayama P, Patel S, Mejía-Bautista M, Sosa-Soto G, Navarrete KM, Calderon M, Cabrera L, Hoyos-Arango W, Bertoli MT, Berg DE, Gilman RH, Dantas G (2016) Interconnected microbiomes and resistomes in low-income human habitats. Nature 533:212–216 . doi: 10.1038/nature17672

Pezzullo M, Del Vecchio P, Mandrich L, Nucci R, Rossi M, Manco G (2013) Comprehensive analysis of surface charged residues involved in thermal stability in *Alicyclobacillus acidocaldarius* esterase 2. Protein Eng Des Sel 26:47–58 . doi: 10.1093/protein/gzs066

Quince C, Walker AW, Simpson JT, Loman NJ, Segata N (2017) Shotgun metagenomics, from sampling to analysis. Nat Biotechnol 35:833–844 . doi: 10.1038/nbt.3935

Rastogi M, Nandal M, Khosla B (2020) Microbes as vital additives for solid waste composting. Heliyon 6:e03343 . doi: 10.1016/j.heliyon.2020.e03343

Reyes A, Alves JMP, Durham AM, Gruber A (2017) Use of profile hidden Markov models in viral discovery: current insights. Adv Genomics Genet 7:29–45 . doi: 10.2147/AGG.S136574

Rong Z, Huo YY, Jian SL, Wu YH, Xu XW (2018) Characterization of a novel alkaline esterase from *Altererythrobacter epoxidivorans* CGMCC 1.7731 T. Prep Biochem Biotechnol 48:113–120 . doi: 10.1080/10826068.2017.1387559

Rosenstein R, Götz F (2000) Staphylococcal lipases: Biochemical and molecular characterization. Biochimie 82:1005–1014

Ryckeboer J, Mergaert J, Vaes K, Klammer S, Clercq D, Coosemans J, Insam H, Swings J (2003) A survey of bacteria and fungi occurring during composting and self-heating processes. Ann Microbiol 53:349–410

Sadeghi M, Naderi-Manesh H, Zarrabi M, Ranjbar B (2006) Effective factors in thermostability of thermophilic proteins. Biophys Chem 119:256–270 . doi: 10.1016/j.bpc.2005.09.018

Sarmah N, Revathi D, Sheelu G, Yamuna Rani K, Sridhar S, Mehtab V, Sumana C (2018) Recent advances on sources and industrial applications of lipases. Biotechnol Prog 34:5–28 . doi: 10.1002/btpr.2581

Sayer C, Finnigan W, Isupov MN, Levisson M, Kengen SWM, van der Oost J, Harmer NJ, Littlechild JA (2016) Structural and biochemical characterisation of *Archaeoglobus fulgidus* esterase reveals a bound CoA molecule in the vicinity of the active site. Sci Rep 6:25542 . doi: 10.1038/srep25542

Silva CC, Hayden H, Sawbridge T, Mele P, Kruger RH, Rodrigues MVN, Costa GGL, Vidal RO, Sousa MP, Torres APR, Santiago VMJ, Oliveira VM (2012) Phylogenetic and functional diversity of metagenomic libraries of phenol degrading sludge from petroleum refinery wastewater treatment system. AMB Express 2:18 . doi: 10.1186/2191-0855-2-18

Simon C, Daniel R (2011) Metagenomic analyses: Past and future trends. Appl Environ Microbiol 77:1153–1161 . doi: 10.1128/AEM.02345-10

Sivakumar N, Li N, Tang JW, Patel BKC, Swaminathan K (2006) Crystal structure of AmyA lacks acidic surface and provide insights into protein stability at poly-extreme condition. FEBS Lett 580:2646–2652 . doi: 10.1016/J.FEBSLET.2006.04.017

Skewes-Cox P, Sharpton TJ, Pollard KS, DeRisi JL (2014) Profile Hidden Markov Models for the detection of viruses within metagenomic sequence data. PLoS One 9:e105067 . doi: 10.1371/journal.pone.0105067

Sood S, Sharma A, Sharma N, Kanwar SS (2016) Carboxylesterases: Sources, characterization and broader applications. Insights Enzym Res 1:1–11 . doi: 10.21767/2573-4466.100002

Stergiou PY, Foukis A, Filippou M, Koukouritaki M, Parapouli M, Theodorou LG, Hatziloukas E, Afendra A, Pandey A, Papamichael EM, Moh B, Salameh DA, Gupta R, Gupta N, Rathi P, Levisson M, Mrugesh D, Togneri PD, Anobom CD, Pinheiro AS, De-Andrade RA, Aguieiras

ECG, Andrade GC, Moura M V, Almeida R V, Freire DM, Nagazawa AA, Nalder TD, Bornscheuer UT, Dheeman DS, Henehan PG, Lopez-Lopez O, Cerdan M, Siso M, Alex D, Udatha G (2014) Biochemical and Structural Analysis of Thermostable Esterases. Biotechnol Adv 31:445–455 . doi: 10.1016/S0168-6445(01)00075-4

Tadeo X, Ló Pez-Mé Ndez B, Trigueros T, Laín A, Castañ D, Millet O (2009) Structural basis for the aminoacid composition of proteins from halophilic Archea. PLoS Biol 7:e1000257 . doi: 10.1371/journal.pbio.1000257

Tripathi NK, Shrivastava A (2019) Recent developments in bioprocessing of recombinant proteins: expression hosts and process development. Front Bioeng Biotechnol 7:420 . doi: 10.3389/fbioe.2019.00420

Troy SM, Nolan T, Kwapinski W, Leahy JJ, Healy MG, Lawlor PG (2012) Effect of sawdust addition on composting of separated raw and anaerobically digested pig manure. J Environ Manage 111:70–77 . doi: 10.1016/j.jenvman.2012.06.035

Uchiyama T, Miyazaki K (2009) Functional metagenomics for enzyme discovery: challenges to efficient screening. Curr Opin Biotechnol 20:616–622 . doi: 10.1016/j.copbio.2009.09.010

Ufarté L, Laville É, Duquesne S, Potocki-Veronese G (2015) Metagenomics for the discovery of pollutant degrading enzymes. Biotechnol Adv 33:1845–1854 . doi: 10.1016/j.biotechadv.2015.10.009

Ventosa A, Nieto JJ, Oren A (1998) Biology of moderately halophilic aerobic bacteria. Microbiol Mol Biol Rev 62:504–44

Vigneron A, Alsop EB, Lomans BP, Kyrpides NC, Head IM, Tsesmetzis N (2017) Succession in the petroleum reservoir microbiome through an oil field production lifecycle. ISME J 11:2141–2154 . doi: 10.1038/ismej.2017.78

Wang C, Dong D, Wang H, Müller K, Qin Y, Wang H, Wu W (2016) Metagenomic analysis of microbial consortia enriched from compost : new insights into the role of *Actinobacteria* in lignocellulose decomposition. Biotechnol Biofuels 9:22 . doi: 10.1186/s13068-016-0440-2

Wei H, Tucker MP, Baker JO, Harris M, Luo Y, Xu Q, Himmel ME, Ding SY (2012) Tracking dynamics of plant biomass composting by changes in substrate structure, microbial community, and enzyme activity. Biotechnol Biofuels 5:1–14 . doi: 10.1186/1754-6834-5-20

Wei H, Wang L, Hassan M, Xie B (2018) Succession of the functional microbial communities and the metabolic functions in maize straw composting process. Bioresour Technol 256:333–341 . doi: 10.1016/j.biortech.2018.02.050

Zhang Y, Pengjun S, Wanli L, Kun M, Yingguo B, Guozeng W, Zhichun Z, Bin Y (2009) Lipase diversity in glacier soil based on analysis of metagenomic DNA fragments and cell culture. J Microbiol Biotechnol 19:888–897 . doi: 10.4014/jmb.0812.695

Zhou G, Xu X, Qiu X, Zhang J (2019) Biochar influences the succession of microbial communities and the metabolic functions during rice straw composting with pig manure. Bioresour Technol 272:10–18 . doi: 10.1016/j.biortech.2018.09.135

Zhou H, Gu W, Sun W, Hay AG (2018) A microbial community snapshot of windrows from a commercial composting facility. Appl Microbiol Biotechnol 102:8069–8077 . doi: 10.1007/s00253-018-9201-4

Summary

Lipolytic enzymes, including lipases and esterases, have a wide spectrum of applications in various industrial fields, due to their functional versatility. Despite an increasing number of lipolytic enzymes that have been reported recently, only a small portion were experimentally verified and studied. Until recently, the demand for new lipolytic enzymes, particularly the extremophilic ones that can survive in extreme environments, is still high.

In this thesis, two compost samples (compost55 and compost76) at thermophilic stage were collected for exploring potential extremophilic lipolytic enzymes. Firstly, we investigated the microbial composition in the two composts. The 16S rRNA genes and transcripts data as well as metagenomic data revealed that *Actinobacteria*, *Proteobacteria*, *Firmicutes*, *Bacteroidetes* and *Chloroflexi* were dominant in the microbial consortia of both samples. The taxonomic disparity between compost55 and compost76 was mainly attributed to the different feedstock composition and composting conditions. In general, analysis of the metagenome data of compost55 and compost76 showed that both share similar metabolic pattern. Nevertheless, genes involved in lipid transport and metabolism as well as the categories fatty acids, lipids, and isoprenoids were more abundant in the compost55 than in compost76 community, suggesting a higher possibility to identify lipolytic genes in the compost55 metagenome.

Four metagenomic libraries were prepared to probe the diversity of LEs from compost microbes by function-based approaches. Overall, 199 and 51 positive clones for compost55 and compost76, respectively, were recovered. The inserts of the recovered plasmids with a confirmed phenotype were sequenced, and putative lipolytic genes were identified. Clustering the amino acid sequences deduced from

the corresponding lipolytic genes yielded 115 unique and full-length lipolytic enzymes. Then, the family assignment of these enzymes was conducted by analyzing the phylogenetic relationship and protein sequence similarity network according to an integral classification system. To the best of our knowledge, the system included all the reported lipolytic enzymes so far (46 families in total) to avoid artificial inflation of the number of families during classification. The functional screening-derived lipolytic enzymes were affiliated to 12 lipolytic families. Seven LEs were not assigned to any known families, indicating new branches of lipolytic families. Subsequently, the multiple sequence alignment further showed the catalytic residues and conserved motif in each family.

For sequence-based screening, we have developed a searching and subsequent annotation strategy specific for putative lipolytic genes in metagenomes. The profile hidden Markov models-based searching methods was highly sensitive and accurate for lipolytic enzymes. With this sequence-based screening and annotation strategy, 4,157 and 2,234 putative lipolytic proteins (PLPs) were initially identified in the assembled metagenomes of compost55 and compost76, respectively. Among them, 1,234 (compost55) and 759 (compost76) of these were further assigned into 28 and 26 families, respectively. The enrichments were observed in families, such as VIII, hormone-sensitive lipase-like, patatin-like proteins, II, A85-Feruloyl-Esterase, Carb_B_Bacteria and homoserine transacetylase. Analysis of the phylogenetic origin of the assigned PLPs indicated a potential link between microbial taxa and their functional traits. By comparing the lipolytic hits identified by function-driven and sequence-based screening indicated that the activity-directed selection complements sequence-based selection, and vice versa.

In addition, comparative analysis of the distribution of lipolytic genes in metagenomes from various ecological niches were explored using the sequence-

based approach developed in this thesis. The lipolytic family assignment (functional profile) and phylogenetic origin (taxonomic profile) of assigned PLPs for each sample were driven by the ecological factor (habitat). Moreover, the habitat also determined the conserved and distinctive microbial groups harboring the putative lipolytic genes.

Finally, three lipolytic genes (*est1*, *est2*, and *est56*) belonging to different lipolytic families and showing low sequence identity to known lipolytic enzymes were selected for biochemical characterization. The three genes were heterologously expressed and characterized.

The gene product Est1 (*est1*) and Est2 (*est2*) are thermostable enzymes with optimal enzyme activities at 80 and 70 °C, respectively. The two enzymes, particularly Est2, were also proved as organic solvent tolerant. Est2 activity was significantly enhanced (two- to tenfold) in the presence of ethanol, methanol, isopropanol, and 1-propanol over a concentration range ranging from 6 to 30% (v/v). Moreover, Est2 exhibited short-term (2 h of incubation) and long-term (up to 26 days) stability towards various water-miscible organic solvents at different concentrations. Est2 also displayed high stability towards water-immiscible organic solvents of ethyl acetate, diethyl ether, and toluene. All of these features indicated that Est1 and Est2 possess application potential.

The other lipolytic enzyme, Est56 (gene product of *est56*), was halotolerant. It exhibited high activity and stability towards up to 4 M NaCl and KCl. In the presence of NaCl, Est56 also displayed enhanced stability against denaturants including high temperatures (50 and 60 ºC) and urea (2, 4, and 6 M). The amino acid composition of recently reported halotolerant lipolytic enzymes (40 in total) and halophilic enzymes was statistically compared to reveal the potential salt resistance mechanism for Est56. The results indicate that the haloadaptation of Est56 was mainly attributed

to the high content of acidic residues (Asp and Glu, 12.2 %), the low content of lysine residues (0.7 %), and the excess of surface-exposed acidic residue.

Acknowledgements

First of all, I want to show my deep and sincere appreciation to my doctoral advisor Prof. Rolf Daniel, for giving me the opportunity to join in his excellent team, providing me with this interesting topic, and giving me all the wise supervision, advises and help. He is very kind, knowledgeable, empathetic and brilliant, I appreciated his support throughout the years very much.

I am also grateful to Prof. Michael Hoppert for being my second supervisor and giving me all the insightful suggestions.

It is my honor that Prof. Dr. Kai Heimel, Prof. Dr. Stefanie Pöggeler, Prof. Dr. Gerhard Braus and Prof. Dr. Dieter Heineke agree to join in my examination board.

Many thanks to my friends and also colleagues, Amelie, Martin, Dominik and Silja. Thank you very much for your company, support, encouragement and help. Special thanks to Amelie, I miss and cherish the nice talks, letters, meetings and emails, which helped me in going through the dark times.

I am also thankful to have worked with all these great people: Amelie, Dominik, Silja, Mechthild, Heiko, Martin, Birgit, Bernd, Ines, Sarah, Kristin, Henis, Rhandi, Dirk and Tanja. It was such a wonderful and unforgettable experience.

I am much obliged to my family, in particular to my parents and parents-in-law. They unconditionally support all my decisions, understand my situation, offer me warm hugs and comforts. They are always my solid backing.

Finally, I want to give all my heart to my husband and daughter. Despite all the difficulties, thank you for always holding my hands, support me, encourage me, trust me, and love me. I am so lucky to have such a wonderful and responsible man. For my daughter, she is the most precious and great gift from the god. Thank you for being my daughter and making my life full of sunshine, laughter and joy.