Environmental Technology & Innovation 42 (2026) 104872

Contents lists available at ScienceDirect

Environmental Technology & Innovation

journal homepage: www.elsevier.com/locate/eti

Plastic-degrading clusters of orthologous groups reveal near-universal biodegradation potential in prokaryotes

Shakira Mustari [a,1], Loan Tú Phạm [a,1], Kari Saikkonen [b], Miho Nakamura [c,d,e,g,*,2], Pere Puigbò [a,f,**,2]

[a] Department of Biology, University of Turku, Turku 20500, Finland
[b] Department of Biodiversity Sciences, University of Turku, Turku 20500, Finland
[c] Medicity Research Laboratory, Faculty of Medicine, University of Turku, Tykistökatu 6, Turku 20520, Finland
[d] Institute of Biomaterials and Bioengineering, Tokyo Medical and Dental University, 2-3-10 Kanda-Surugadai, Chiyoda, Tokyo 1010062, Japan
[e] Graduate School of Engineering, Tohoku University, 6-6 Aramaki Aza Aoba, Aoba-ku, Sendai, Miyagi 9808579, Japan
[f] Department of Animal and Food Science, Faculty of Veterinary Medicine, Autonomous University of Barcelona, Bellaterra, Catalonia 08193, Spain
[g] La Salle Campus Barcelona Ramon Llull University, Barcelona, Catalonia 08022, Spain

ARTICLE INFO

Keywords:
Microplastics
Nanoplastics
Biodegradation
Microbial Genomes
Plastic Polymers
Enzymes
Clusters of Orthologous groups

ABSTRACT

Micro- and nanoplastic pollution (MNPP) is an increasing environmental threat due to the persistence, dispersal, and potential toxicity of plastic particles. Although microbial biodegradation offers a sustainable mitigation strategy, a comprehensive understanding of plastic-degrading proteins across microbial taxa is lacking. Here, we present the Plastic-Degrading Clusters of Orthologous Groups (PDCOGs) database (https://phylobone.com/microworld/PDCOG), comprising 625,616 potential plastic-degrading proteins (PPDPs) from free-living prokaryotes organized into 51 orthologous groups. The database/PDCOGs enable systematic analysis of microbial plastic-degrading capacity across ecosystems and phylogenetic lineages. Notably, PPDPs constitute ~3.5% of all prokaryotic proteins, with over 95% of the species having the potential to biodegrade at least one plastic polymer type. This resource provides a genomic tool/framework for exploring the ecological and evolutionary importance of plastic biodegradation and supports future efforts to mitigate the global MNPP crisis.

1. Introduction

Plastic pollution is among the most urgent environmental challenges of the 21st century owing to its small size, persistence in the environment, dispersal, and potential adverse effects on ecosystems and human health (Kakar et al., 2023). Microplastics (1–1000 μm) and nanoplastics (0.001–1 μm) (Bermúdez and Swarzenski, 2021) arise from the fragmentation of larger plastic items or are directly manufactured at small scales (Tong et al., 2022). They have been detected across diverse habitats, including marine, freshwater, terrestrial and atmospheric environments, where they pose serious threats to wildlife and may enter the human food chain (Mamun

* Corresponding author at: Medicity Research Laboratory, Faculty of Medicine, University of Turku, Tykistökatu 6, Turku 20520, Finland.
** Corresponding author at: Department of Biology, University of Turku, Turku 20500, Finland.
E-mail addresses: miho.nakamura@salle.url.edu (M. Nakamura), pere.puigbo@uab.cat (P. Puigbò).
[1] Equal contribution of the authors
[2] Co-senior authors

https://doi.org/10.1016/j.eti.2026.104872
Received 3 November 2025; Received in revised form 4 March 2026; Accepted 9 March 2026
Available online 15 March 2026

et al., 2023; Taipale et al., 2023). In addition to direct toxicity, micro- and nanoplastics (MNPs) can serve as vectors for chemical pollutants and pathogenic microbes, raising concerns about the indirect health impacts on humans when they consume contaminated food and water (Amelia et al., 2021; Shruti et al., 2020). Recent studies have also shown that MNPs may play an important role in the dissemination of antimicrobial-resistant strains because of their capacity to host complex microbial communities and antimicrobial resistance genes through biofilm formation (Gross et al., 2025). In addition, nanoplastics seem to have a more negative impact on ecosystems (Chen et al., 2023) and health (Bridgeman et al., 2025), than microplastics do because of their greater reactivity, abundance, and ability to cross barriers and accumulate in tissues.

Climate change is expected to accelerate MNP accumulation across the Earth. Globally, it is estimated that over 14 million tons of microplastics have accumulated on the seafloor (Barrett et al., 2020). Despite increasing awareness, current strategies to mitigate this MNP pollution problem remain fragmented and largely ineffective in offering long-term solutions. The ubiquity of MNPs, which affect both the most densely populated regions and remote environments such as the Arctic and Antarctic, underscores the global scale of this issue (Bergmann et al., 2022; Obbard et al., 2014). Moreover, the interaction between plastic pollution and climate warming presents

Table 1
List of PDCOGS and their distributions across prokaryotic species, plastic types and environments.

PDCOG	Number	COG	Fun	%Prok.	Plastic types	Environmental distribution
PDCOG001	20,071	COG0154	J	88.81%	N: 0/11; S_He: 1/21; S_Ho: 0/7	Ai (0.56%), Aq (60.27%), Te (38.97%), Un (0.2%)
PDCOG002	21,153	COG0277	C	80.49%	N: 0/11; S_He: 2/21; S_Ho: 1/7	Ai (0.54%), Aq (56.41%), Te (42.84%), Un (0.21%)
PDCOG003	6475	COG0400	R	52.44%	N: 0/11; S_He: 1/21; S_Ho: 0/7	Ai (0.59%), Aq (73.27%), Te (26.01%), Un (0.14%)
PDCOG004	12,730	COG0412	Q	57.01%	N: 2/11; S_He: 9/21; S_Ho: 1/7	Ai (0.46%), Aq (53.94%), Te (45.48%), Un (0.13%)
PDCOG005	16,725	COG0508	C	80.01%	N: 0/11; S_He: 1/21; S_Ho: 1/7	Ai (0.50%), Aq (68.55%), Te (30.75%), Un (0.2%)
PDCOG006	90,850	COG0596	HR	91.29%	N: 4/11; S_He: 8/21; S_Ho: 1/7	Ai (0.56%), Aq (54.26%), Te (45.01%), Un (0.16%)
PDCOG007	4527	COG0627	V	43.07%	N: 1/11; S_He: 5/21; S_Ho: 1/7	Ai (0.46%), Aq (48.13%), Te (51.25%), Un (0.15%)
PDCOG008	15,526	COG0657	I	65.11%	N: 3/11; S_He: 6/21; S_Ho: 2/7	Ai (0.71%), Aq (51.13%), Te (48.04%), Un (0.12%)
PDCOG009	14,723	COG0726	MG	80.66%	N: 0/11; S_He: 0/21; S_Ho: 2/7	Ai (0.60%), Aq (51.69%), Te (47.48%), Un (0.22%)
PDCOG010	43,441	COG0834	TE	76.96%	N: 0/11; S_He: 1/21; S_Ho: 0/7	Ai (0.46%), Aq (54.29%), Te (45.04%), Un (0.22%)
PDCOG011	67,608	COG1012	I	89.46%	N: 0/11; S_He: 1/21; S_Ho: 1/7	Ai (0.44%), Aq (53.08%), Te (46.25%), Un (0.22%)
PDCOG012	139,666	COG1028	I	95.03%	N: 3/11; S_He: 0/21; S_Ho: 2/7	Ai (0.57%), Aq (55.03%), Te (44.21%), Un (0.19%)
PDCOG013	4116	COG1073	T	30.31%	N: 1/11; S_He: 5/21; S_Ho: 1/7	Ai (0.56%), Aq (65.38%), Te (33.79%), Un (0.27%)
PDCOG014	2434	COG1075	I	32.58%	N: 4/11; S_He: 8/21; S_Ho: 1/7	Ai (0.58%), Aq (49.92%), Te (49.30%), Un (0.21%)
PDCOG015	261	COG1382	O	8.89%	N: 0/11; S_He: 1/21; S_Ho: 0/7	Ai (1.53%), Aq (69.35%), Te (29.12%), Un (0.00%)
PDCOG016	4483	COG1398	I	27.57%	N: 0/11; S_He: 0/21; S_Ho: 1/7	Ai (0.27%), Aq (61.25%), Te (38.39%), Un (0.09%)
PDCOG017	6857	COG1404	O	62.37%	N: 1/11; S_He: 4/21; S_Ho: 0/7	Ai (0.61%), Aq (54.53%), Te (44.55%), Un (0.31%)
PDCOG018	9776	COG1506	E	71.65%	N: 4/11; S_He: 4/21; S_Ho: 0/7	Ai (0.30%), Aq (71.69%), Te (27.91%), Un (0.11%)
PDCOG019	2783	COG1647	Q	31.84%	N: 1/11; S_He: 5/21; S_Ho: 1/7	Ai (0.25%), Aq (63.71%), Te (35.79%), Un (0.25%)
PDCOG020	21,699	COG1680	V	66.77%	N: 1/11; S_He: 5/21; S_Ho: 0/7	Ai (0.41%), Aq (59.08%), Te (40.36%), Un (0.15%)
PDCOG021	4746	COG1773	P	42.55%	N: 0/11; S_He: 0/21; S_Ho: 3/7	Ai (0.36%), Aq (62.45%), Te (37.04%), Un (0.15%)
PDCOG022	4733	COG1858	O	35.76%	N: 0/11; S_He: 0/21; S_Ho: 2/7	Ai (0.42%), Aq (58.27%), Te (41.12%), Un (0.19%)
PDCOG023	18,460	COG2010	C	56.58%	N: 0/11; S_He: 0/21; S_Ho: 2/7	Ai (0.64%), Aq (56.21%), Te (43.01%), Un (0.14%)
PDCOG024	6773	COG2132	DMP	51.09%	N: 0/11; S_He: 0/21; S_Ho: 4/7	Ai (0.75%), Aq (56.18%), Te (42.85%), Un (0.22%)
PDCOG025	590	COG2247	M	3.44%	N: 0/11; S_He: 1/21; S_Ho: 0/7	Ai (0.34%), Aq (68.98%), Te (30.68%), Un (0.00%)
PDCOG026	8632	COG2267	I	70.86%	N: 3/11; S_He: 14/21; S_Ho: 0/7	Ai (0.47%), Aq (57.75%), Te (41.52%), Un (0.25%)
PDCOG027	2972	COG2272	I	20.17%	N: 1/11; S_He: 9/21; S_Ho: 0/7	Ai (0.34%), Aq (63.36%), Te (36.07%), Un (0.24%)
PDCOG028	17,646	COG2303	IR	49.96%	N: 0/11; S_He: 1/21; S_Ho: 1/7	Ai (0.49%), Aq (56.58%), Te (42.76%), Un (0.17%)
PDCOG029	1542	COG2703	T	20.82%	N: 0/11; S_He: 0/21; S_Ho: 1/7	Ai (0.71%), Aq (66.67%), Te (32.10%), Un (0.52%)
PDCOG030	10,074	COG2863	C	35.37%	N: 0/11; S_He: 1/21; S_Ho: 1/7	Ai (0.60%), Aq (59.90%), Te (39.40%), Un (0.11%)
PDCOG031	5066	COG2931	Q	28.75%	N: 0/11; S_He: 2/21; S_Ho: 0/7	Ai (0.71%), Aq (61.94%), Te (36.89%), Un (0.45%)
PDCOG032	3397	COG3145	L	33.97%	N: 0/11; S_He: 0/21; S_Ho: 3/7	Ai (0.71%), Aq (59.58%), Te (39.48%), Un (0.24%)
PDCOG033	2814	COG3147	D	28.22%	N: 2/11; S_He: 0/21; S_Ho: 0/7	Ai (0.39%), Aq (66.24%), Te (33.08%), Un (0.28%)
PDCOG034	1515	COG3179	G	11.80%	N: 0/11; S_He: 1/21; S_Ho: 0/7	Ai (0.20%), Aq (43.17%), Te (56.44%), Un (0.20%)
PDCOG035	2522	COG3191	E	29.05%	N: 0/11; S_He: 2/21; S_Ho: 0/7	Ai (0.36%), Aq (42.27%), Te (57.14%), Un (0.24%)
PDCOG036	1,0447	COG3239	I	40.29%	N: 0/11; S_He: 0/21; S_Ho: 4/7	Ai (0.30%), Aq (61.02%), Te (38.57%), Un (0.11%)
PDCOG037	3984	COG3243	I	28.83%	N: 2/11; S_He: 0/21; S_Ho: 0/7	Ai (0.85%), Aq (56.58%), Te (42.42%), Un (0.15%)
PDCOG038	693	COG3266	D	17.77%	N: 2/11; S_He: 0/21; S_Ho: 0/7	Ai (0.43%), Aq (82.40%), Te (16.74%), Un (0.43%)
PDCOG039	1892	COG3325	G	16.33%	N: 0/11; S_He: 4/21; S_Ho: 0/7	Ai (0.16%), Aq (53.91%), Te (45.61%), Un (0.32%)
PDCOG040	278	COG3401	R	11.28%	N: 0/11; S_He: 4/21; S_Ho: 0/7	Ai (0.36%), Aq (57.91%), Te (41.73%), Un (0.00%)
PDCOG041	3327	COG3509	G	25.04%	N: 8/11; S_He: 11/21; S_Ho: 2/7	Ai (0.75%), Aq (53.77%), Te (45.33%), Un (0.15%)
PDCOG042	1500	COG3591	O	21.56%	N: 0/11; S_He: 1/21; S_Ho: 0/7	Ai (0.47%), Aq (49.87%), Te (48.87%), Un (0.80%)
PDCOG043	1294	COG3979	G	11.67%	N: 0/11; S_He: 5/21; S_Ho: 0/7	Ai (0.31%), Aq (69.78%), Te (29.75%), Un (0.15%)
PDCOG044	2311	COG4188	R	21.86%	N: 3/11; S_He: 8/21; S_Ho: 0/7	Ai (0.17%), Aq (65.73%), Te (34.01%), Un (0.09%)
PDCOG045	2504	COG4553	I	14.94%	N: 3/11; S_He: 7/21; S_Ho: 0/7	Ai (0.72%), Aq (38.74%), Te (60.42%), Un (0.12%)
PDCOG046	9970	COG1366	T	35.76%	N: 0/11; S_He: 1/21; S_Ho: 0/7	Ai (0.90%), Aq (57.92%), Te (40.52%), Un (0.65%)
PDCOG047	6948	COG1520	M	53.53%	N: 0/11; S_He: 1/21; S_Ho: 0/7	Ai (0.39%), Aq (70.31%), Te (29.07%), Un (0.23%)
PDCOG048	15,226	COG2755	DI	73.78%	N: 0/11; S_He: 1/21; S_Ho: 0/7	Ai (0.63%), Aq (57.97%), Te (41.11%), Un (0.28%)
PDCOG049	2710	COG2819	R	24.65%	N: 0/11; S_He: 1/21; S_Ho: 0/7	Ai (0.44%), Aq (66.94%), Te (32.55%), Un (0.07%)
PDCOG050	1749	COG3240	I	12.46%	N: 0/11; S_He: 1/21; S_Ho: 0/7	Ai (0.80%), Aq (45.45%), Te (53.57%), Un (0.17%)
PDCOG051	611	COG4099	R	16.16%	N: 0/11; S_He: 1/21; S_Ho: 0/7	Ai (0.33%), Aq (73.65%), Te (26.02%), Un (0.00%)

Fun: Functional category in the COG database; N: Natural polymers (n = 10); S_Ho: Synthetic-homochain polymers (n = 8); S_He: Synthetic-heterochain polymers (n = 21); Ai: Air; Aq: Aquatic; Te: Terrestrial; Un: Unknown

an emerging threat (Sharma et al., 2023). Ocean currents transport plastic debris to polar regions, while melting sea ice releases previously trapped microplastics (Bergmann et al., 2022; Obbard et al., 2014). The Arctic Ocean is predicted to be ice-free before the end of the century (Kelly et al., 2010), whereas ice-free areas in Antarctica may expand from the present 2% to nearly 25% in the same time frame (Lee et al., 2017). These changes will likely exacerbate plastic accumulation in these ecologically vulnerable regions.

Despite increasing awareness, current strategies for managing plastic waste remain fragmented and ineffective in offering long-term solutions at the global scale. Given the persistence of common plastic polymers, such as polyethylene (PE), polystyrene (PS), polypropylene (PP), polyvinyl chloride (PVC), polyethylene terephthalate (PET), and polyurethane (PUR), microbial biodegradation has emerged as a promising, sustainable and environmentally friendly alternative (Yang et al., 2022). Studies of the plastisphere—the microbial biofilms that are formed on plastic surfaces (Zettler et al., 2013)—have revealed a growing diversity of microbial strains capable of plastic degradation (Amaral-Zettler et al., 2020). Although earlier studies focused primarily on mesophilic bacteria (20–45 °C) (Lehmann et al., 2023), a recent study identified plastic-degrading enzymes from thermophiles (i.e., microbial strains with optimal growth temperatures above 45 °C (Lehmann et al., 2023)). Thermostable enzymes are likely more suitable for industrial applications (Hu et al., 2025). Furthermore, microbial strains isolated from the Arctic plastisphere have demonstrated the ability to break down biodegradable plastic films at temperatures below 20 °C (Rüthi et al., 2023), which may offer increased stability and efficiency for industrial and cold-environment applications (Hu et al., 2025). However, despite these advances, the global understanding of plastic-degrading microbial capacities remains limited by the absence of systematic, large-scale comparative resources.

To address this gap, we introduce the Plastic-Degrading Clusters of Orthologous Groups (PDCOGs) database—a comprehensive resource of putative plastic-degrading proteins (PPDPs) from free-living bacteria allocated into 51 orthologous groups. The PDCOGs database enables systematic comparisons of PPDPs across environmental gradients and geographic regions, facilitating a deeper understanding of the global biodegradation landscape. Ultimately, this resource provides a foundation for the development of microbial and enzymatic solutions to mitigate the growing threat of MNP pollution under rapidly changing environmental conditions.

2. Results

2.1. Method summary

A seed of PPDP sequences was obtained from the literature that were mapped onto a database of clusters of orthologous groups in prokaryotes (archaea and bacteria) (Galperin et al., 2025) for the identification of plastic-degrading protein families (PDCOGs). Each PDCOGs was classified based on the 26 biological functions described in the clusters of orthologous groups (COG) database (Galperin et al., 2025; Tatusov et al., 1997), as well as on the annotation of putative plastic-degrading capacities. The COGs method is the most robust approach for genomics classification and phylogenetic inference in prokaryotes. Briefly, the main advantage of the PDCOGs

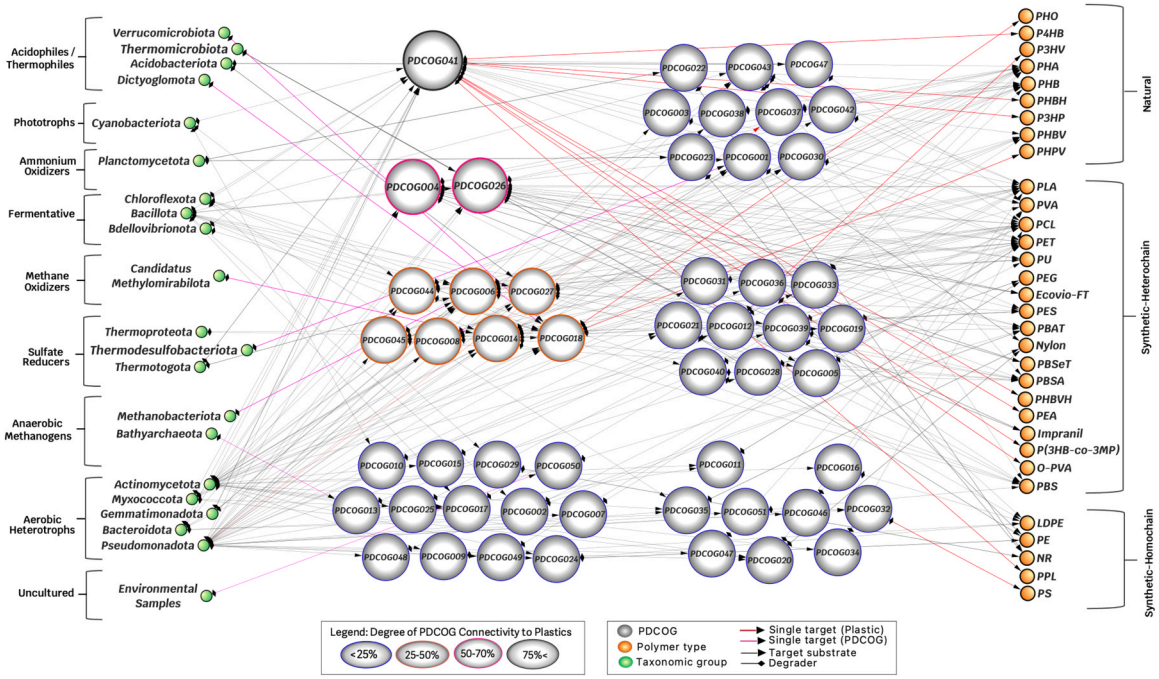

Fig. 1. Bipartite network of associations between the plastic-degrading clusters of orthologous groups (PDCOGs) and microbial species and enzymatic functions. The network was generated from the linkage data of organisms, COGs, and targeted plastics using Cytoscape v3.10.3. In the network, node colors represent (i) the COG functional hierarchy on the basis of connectivity to polymers, (ii) the taxonomic clustering of species, and (iii) the classification of plastic polymers. Edge colors indicate the strength of connections, weighted by interaction confidence.

database is that it provides a standardized manner to compare plastic degrading enzymes across environments and microbial taxonomic groups. The resultant PDCOGs were utilized to annotate proteins from free-living prokaryotes.

2.2. Resource description

2.2.1. A global resource of putative plastic-degrading protein clusters

The resource is organized into six sections and comprises 662,830 PPDP sequences categorized into 51 orthologous clusters (PDCOGs) (Table 1, Supplementary Table S1). The PPDPs span diverse environmental sources and are associated with enzymatic activities involved in the biodegradation of both natural and synthetic polymers. The data capture the microbial degradation potential across 11 natural and 28 synthetic polymers, including 7 homochain and 21 heterochain variants. Notably, polyhydroxyalkanoates such as P(3HB-co-3MP), P(3HV) (Watabe et al., 2023), and P3HP (Meng et al., 2015), although often considered natural owing to their microbial biosynthesis, and classified here as natural (De Angelis and Gobbetti, 1999; García et al., 2011), are engineered for production in *Escherichia coli*, *Clostridium butiricum*, and *Klebsiella pneumoniae* (Yang et al., 2018). All the plastic polymers included in the PDCOGs databases are potentially susceptible to biodegradation by microbes encoding at least one copy of a PDCOG (Supplementary Datasets SD1–2).

2.2.2. Linking PDCOGs with polymers and microbial taxa

A bipartite network analysis linking PDCOGs to polymer types and microbial taxa revealed substantial functional and phylogenetic diversity (Fig. 1 and Supplementary Fig. S1). This modular structure suggests that plastic-degrading potential is broadly distributed

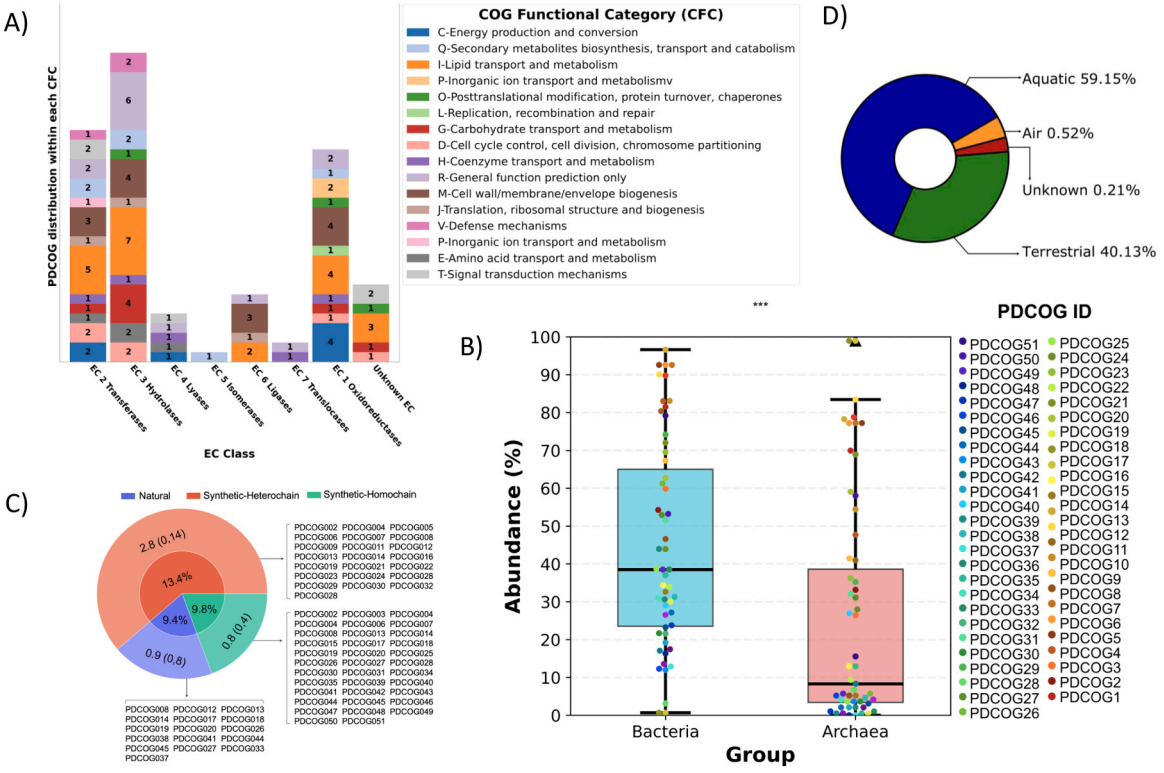

Fig. 2. Functional, genomic and ecological diversity in the plastic-degrading clusters of orthologous groups (PDCOGs). A. Enzymatic roles and genomic functions of plastic biodegradation. The stacked bar plot illustrates the distribution of PDCOGs across COG functional categories (CFCs) within each EC class. The x-axis represents the EC classes (EC1 to EC7), which categorize enzymes on the basis of their catalytic reactions. Each bar is divided into colored segments, with each color representing a distinct CFC (e.g., metabolism, cellular processes, information storage). The count of correlated PDCOGs in each CFC is annotated within the corresponding segment of the bar, the sum of which, on the y-axis, indicates the quantification of COGs associated with each EC class. **B. Relative prokaryotic abundance in the PDCOGs.** Box plots showing the percentages of prokaryotes, bacteria, and Archaea (primary y-axis) across 51 PDCOGs (secondary y-axis), with a swarm plot overlay indicating enzyme-specific contributions (enzymes color-coded). The statistical significance of the dominance of bacteria over Archaea within prokaryotes was determined by a Wilcoxon signed-rank test (***: $p < 0.001$). The outliers are denoted by black triangles. C. Plastic polymer biodegradation profile. The pie chart shows the proportional distributions of natural, synthetic homochain, and synthetic heterochain plastic polymers potentially biodegradable by PDCOGs. The inner segment proportions indicate the relative degradation capacity of different categories of polymers, whereas the outer layer represents the average degradability with minimum and maximum counts. D. Environmental distribution of PDCOGs. Illustration of the mean proportional incidence of plastic degraders in various environmental habitats on the basis of genome counts from the Genome portal.

across bacterial lineages (Supplementary Dataset SD3). The median number of polymer associations per PDCOG was 3 (range: 1–16). PDCOG041 and PDCOG026 display the highest degree of connectivity and are associated with 16 and 12 polymers, respectively. In contrast, 22 PDCOGs are associated exclusively with a single polymer type, suggesting potential specificity: PDCOG003 is linked to nylon, PDCOG005 is linked to PET, PDCOG007 is linked to PET, PDCOG009 is linked to PET, PDCOG010 is linked to low-density polyethylene (LDPE), PDCOG011 is linked to PET, PDCOG015 is linked to polyethylene glycol (PEG), PDCOG016 is linked to polylactic acid (PLA), PDCOG017 is linked to LDPE, PDCOG017 is linked to PLA, PDCOG019 is linked to PET, PDCOG022 is linked to natural rubber (NR), PDCOG023 is linked to NR, PDCOG025 is linked to PLA, PDCOG028 is linked to PEG, PDCOG029 is linked to NR, PDCOG030 is linked to polyvinyl alcohol (PVA), PDCOG031 is linked to polyurethane (PU), PDCOG032 is linked to PS, PDCOG034 is linked to polycaprolactone (PCL), PDCOG035 is linked to nylon, and PDCOG042 is linked to PLA.

The distribution of microbial species in natural environments, based on the best BLAST hits of the seed dataset in the COG database, varies across PDCOGs, ranging from 1 to 56 species. PDCOG041 was associated with the most species (n = 56), followed by PDCOG006 (34), PDCOG014 (33), PDCOG018 (25), PDCOG027 (21), PDCOG045 (19), and PDCOG001 (10). The remaining PDCOGs are each associated with fewer than 10 species (Supplementary Dataset SD3). Moreover, several bacterial groups are associated with a single PDCOG, such as Thermodesulfobacteria with PDCOG001, Dictyoglomota and Verrucomicrobiota with PDCOG018, Methylomirabilota with PDCOG014, and Methanobacteriota with PDCOG006.

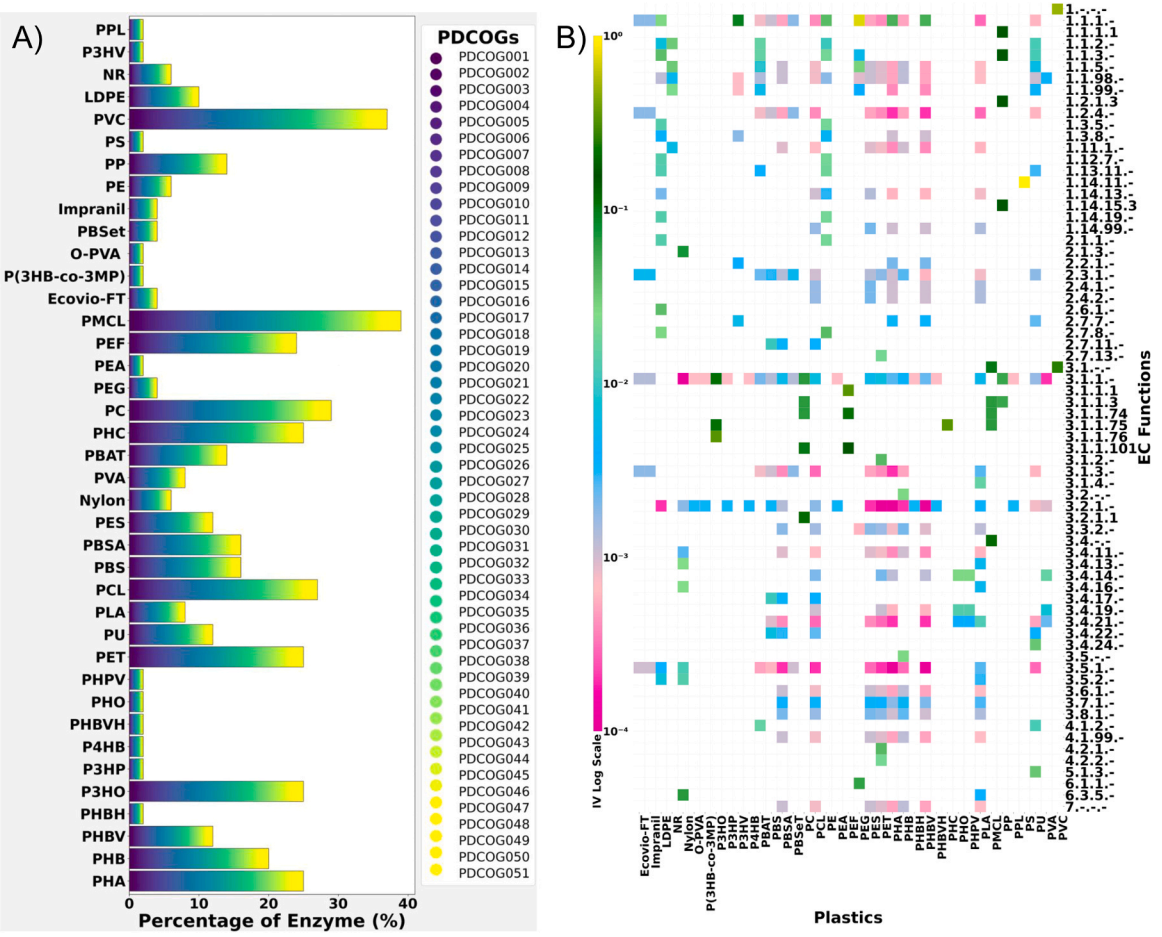

Fig. 3. Enzymatic characterization of plastic-degrading clusters of orthologous groups (PDCOGs) on the basis of known biological enzymatic activities. A. Enzymatic activity of PDCOGs across plastic types. The bars represent the percentage of total enzyme activity (0–100%) among 51 PDCOGs for each plastic type. The x-axis displays plastic types, whereas the y-axis shows enzyme activity (%). Owing to the highest percentage being less than 50%, the y-axis only depicts up to 50% for better visualization. The legend (right) maps PDCOG domains to colors, with a purple-to-yellow gradient (bottom to top) indicating different protein clusters. **B. Heatmap of the indicator values (IVs) for enzymes associated with plastic biodegradation.** This plot shows the IVs for enzymes putatively involved in the biodegradation of 39 plastic types. With plastics on the x-axis and potential enzymes on the y-axis, the IV metric of relative abundance and specificity was calculated on the basis of annotations from the KEGG database (Kanehisa et al., 2025). The color intensity and log-based value (10^{-4} lowest to 1 highest) highlight the strength of their association.

2.2.3. General functional characterization of the PDCOGs

The 51 PDCOGs encompass a wide range of predicted biological functions. A functional characterization of the PDCOGs using the 26 functional categories (Supplementary Table S2) in the COG database revealed that the functional categories of PPDP are I (lipid transport and metabolism; n = 12), R (general function prediction only; n = 6), and G (carbohydrate transport and metabolism; n = 5), which are the most abundant among plastic-degrading enzymes, followed by the other categories in this order: C=D=M=O=T > E = Q>P = V>H=J=L (Fig. 2A). Among the 51 PDCOGs, the enzymes involved in plastic biodegradation include oxidoreductases (EC1s; 33.33%), transferases (EC2s; 33.33%), hydrolases (EC3s; 52.94%), lyases (EC4s; 5.88%), isomerases (EC5s; 1.96%), ligases (EC6s; 9.8%) and translocases (EC7s; 1.96%) (Supplementary Dataset SD4). All PDCOGs were also annotated with KEGG pathway information, gene ontology terms, and Interpro functional domains (Supplementary Dataset SD4), facilitating integrative functional and systems-level analyses.

To assess enzyme—polymer specificity, we computed indicator values (IVs) for 212 EC subclasses across 39 polymers. While most associations had low IVs (75%, < 0.25), enzymes such as carboxylic ester hydrolases and glycosidases were broadly linked to many polymers. Notably, synthetic heterochain polymers were associated with more EC subclasses (n = 52) than were natural (n = 34) or synthetic homochain (n = 26) polymers, indicating greater enzymatic diversity for structurally complex plastics (Fig. 3B, Supplementary Dataset SD4).

2.2.4. Distribution of the PDCOGs across prokaryotes (archaea and bacteria)

Across 2296 prokaryotic species (2103 bacteria and 193 archaea), 13.4% were associated with synthetic heterochain polymers, 9.8% with synthetic homochain polymers, and 9.4% with natural polymers (Fig. 2C; Table 1). Bacteria displayed a significantly broader distribution of PDCOGs than archaea did (the median PDCOGs in bacteria was 38.5%, and the median PDCOGs in archaea was 8.3%, with an overall prevalence of 35.4% in prokaryotes; Wilcoxon signed-rank test, $p < 0.0001$) (Fig. 2B, Supplementary dataset SD1).

On average, the bacterial and archaeal genomes encode 20 and 11 PDCOGs (out of 51), respectively (Supplementary dataset SD1). The largest PDCOGs in archaea were found in Asgardarchaeota (n = 17) and Euryarchaeota (n = 15), whereas Myxococcota (n = 38) and Acidobacteriota (n = 29) were the largest groups in bacteria. Overall, 96.6% of prokaryotes have at least one PDCOG, i.e., a copy of a PPDP. PDCOG015, a chaperonin cofactor present in 99.0% of archaeal species, has protease activity with the potential to breakdown PLA (Gambarini et al., 2022). In bacteria, PDCOG012, which has dehydrogenase activity, has the potential to participate in the intracellular biodegradation of P3HV, PHBV, and PHA (Gambarini et al., 2022). In contrast, PDCOG034 (orthologous group with chitinase activity) in archaea and PDCOG015 in bacteria are poorly represented across the respective lineages.

2.3. Technical validation

2.3.1. Phyletic distribution of the PDCOGs in natural environments

To evaluate the taxonomic overlap in degradation potential, we used BLAST-derived top hits from seed sequences to assess the

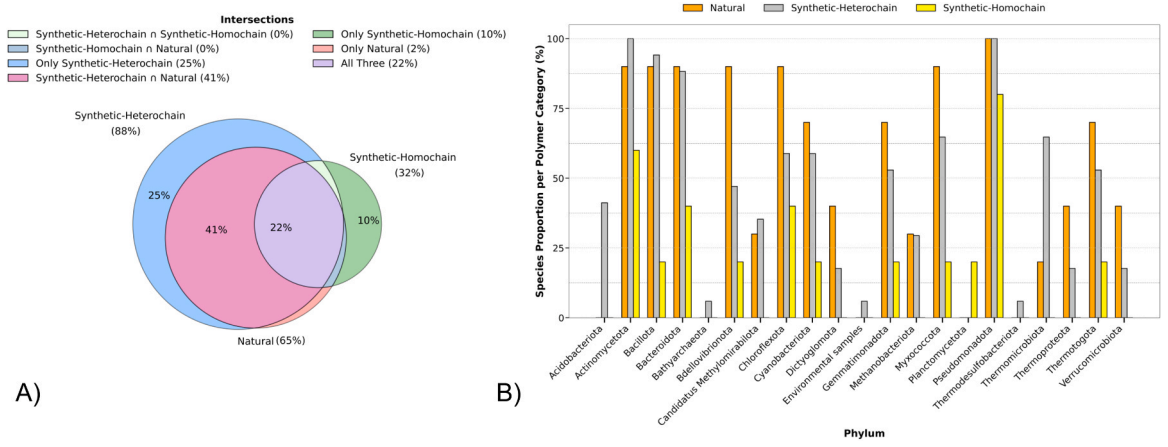

Fig. 4. Taxonomic distribution of plastic-degrading microbes and polymer types. A. Species-level distribution patterns among polymer classes. Venn diagram illustrating the relative distributions of common species among the plastic polymer categories natural, synthetic heterochain, and synthetic homochain for biodegradation. Each circle represents one plastic category with color grading, and the intersections denote shared associations among multiple categories. Percent values are shown for major intersections, whereas low-frequency overlaps (<5%) are denoted in the legend for clarity. Circle positions and sizes vary on the basis of their affinity with species. B. Phylum-specific trends in categorical polymer biodegradation. The grouped bar plot compares the relative associations of phyla within each plastic category. Each bar cluster on the x-axis corresponds to all the unique prokaryotic phyla in the dataset. The height of each bar along the y-axis denotes the phylum-based proportion of unique species that are correlated with the plastics in each category (natural, synthetic heterochain, synthetic homochain), normalized by the total unique plastics within that category.

shared microbial capacities across plastic types (Fig. 4A). We analyzed the relative abundance of the resulting species to determine which plastic types have the most diverse biodegradability potential (synthetic heterochain > natural > synthetic homochain) and which taxa are most frequently associated with specific plastics (Pseudomonadata, Actinomycetota, Bacillota, and Bacteroidota) (Fig. 4B). Synthetic heterochain polymers harbored the highest proportion of unique microbial degraders (88%), followed by natural (65%) and synthetic homochain (32%) polymers. Notably, some strains have the potential capacity to biodegrade more than one polymer type. Only 22% of the species were shared across all three categories, with the greatest overlap between natural and synthetic heterochain polymers (41%); no species were shared between synthetic homochain polymers and the other groups (Fig. 3A).

Most prokaryotic species present enzymes with the capacity to break down plastic polymers (Fig. 4B). PDCOGs are widely distributed across prokaryotic species (Fig. 1, Supplementary Fig. S1, Supplementary Dataset SD3). In nature, Pseudomonadota is the bacterial group that presents the greatest abundance of PPDP of all polymer types (natural and synthetic), followed by Actinomycetota, Bacteroidota and Bacillota (Fig. 4B). In archaea, three phyla, namely, Methanobacteriota, Thermoproteota and Bathyarchaeota, were identified as potential biodegraders of plastic polymers. Our analysis revealed that 65.68% of the cataloged species potentially

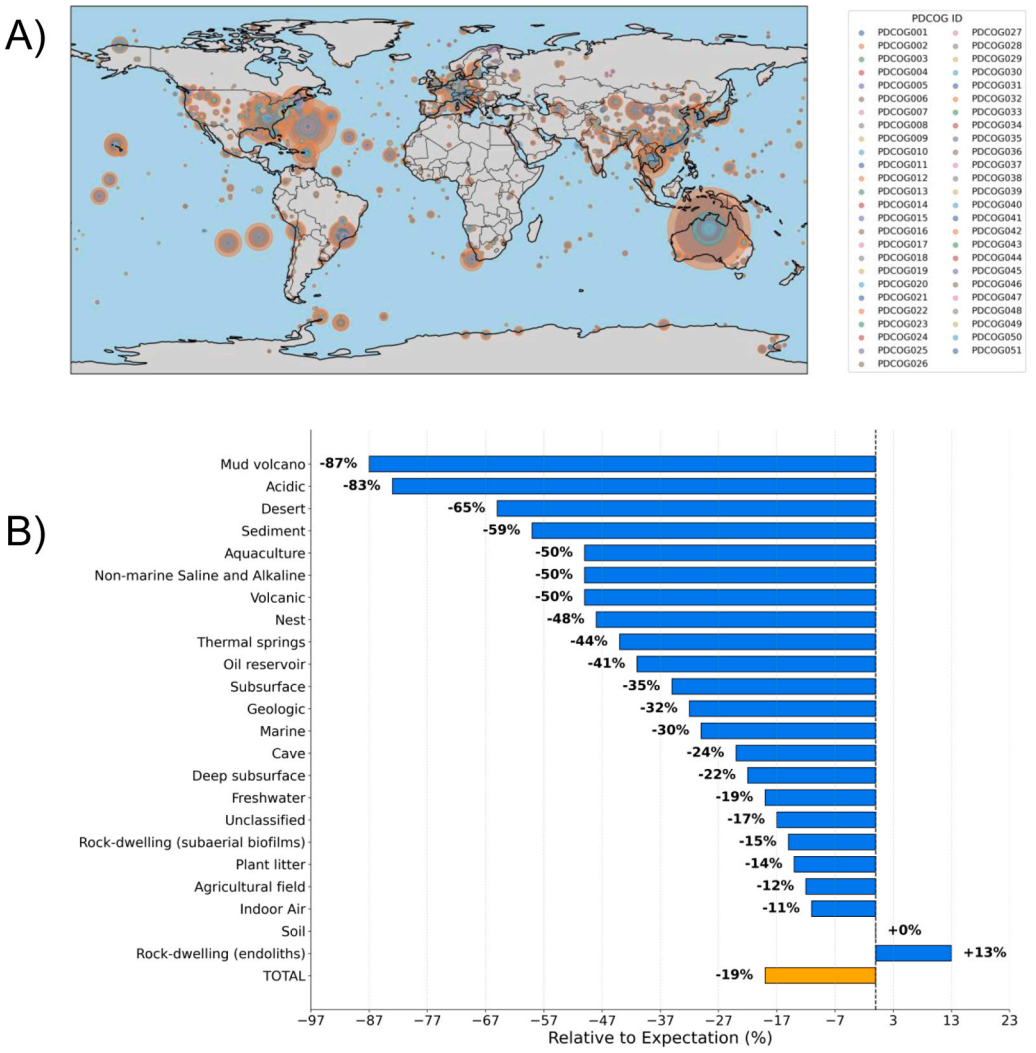

Fig. 5. Global distribution and abundance of plastic-degrading proteins. A. World distribution of plastic-degrading clusters of orthologous groups (PDCOGs) across environments. The geospatial distribution of PDCOGs highlights hotspots of functional diversity and the frequency of plastic-degrading genes across the sampled environments. The color gradients represent functional diversity of PDCOGs, whereas their proportional variation denotes sequence abundance. Hotspots of enzymatic potential correlate with anthropogenic impact zones (e.g., soil, sediment, marine and aquatic) and microbial niches adapted to synthetic polymer hydrolysis. **B. Percentages of observed sequences relative to expectations (O/E) across various environments.** The observed (O) and expected values (E) were calculated on the basis of the data obtained from the JGI Genome Portal and the NCBI COG database, respectively. The values represent the percentage of deviation from the expected baseline (100%). The bars to the left of the central axis indicate environments where the observed values fall below expectations, whereas the bars to the right represent environments exceeding expectations. The "TOTAL" (percentage across environments) is displayed in orange for distinction.

biodegrade natural polymers, 32.47% of the species biodegrade synthetic homochain polymers, and 87.82% biodegrade synthetic heterochain polymers.

2.3.2. Geographic and environmental distributions of the PDCOGs

PPDPs have been detected across a broad spectrum of aquatic and terrestrial environments (Fig. 5; Table 1). Although aquatic environments dominated the dataset due to sampling bias, several PDCOGs (e.g., PDCOG034, PDCOG035, PDCOG045, and PDCOG050) were enriched in continental samples. Geographic data indicate that terrestrial sequences are derived primarily from North America, Western Europe, East Asia, and Australia, whereas marine sequences reflect typical oceanographic sampling routes. Statistical comparison with the expected distributions (chi-square test; Supplementary Table S3) revealed a general underrepresentation of PPDPs in the environmental samples, with the exception of endolithic (rock-dwelling) niches, which were significantly enriched. The soil samples presented the expected representation, but the PDCOG frequencies varied by habitat (Supplementary Dataset SD6), suggesting environment-specific selective pressures on plastic-degrading functions.

2.3.3. Case study: biodegradation potential in plastic-associated microbes

Plastic pollution is a key issue in climate, biodiversity, and resource-use policy (Villarrubia-Gómez et al., 2024). To demonstrate an application of PDCOGs in real-world research, we present a case study assessing the biodegradation potential of microbial communities from beach plastisphere environments (Fig. 6). The biodegradation process involves several key steps: (1) microbial attachment to plastic particles, where species colonize the surface; (2) biofilm formation, providing a stable environment for enzymatic activity; (3) secretion of plastic-degrading enzymes; and (4) assimilation of the resulting monomers as carbon and energy sources (Meng et al., 2024). Within this process, the Plastic-Degrading Clusters of Orthologous Groups (PDCOGs) database contributes to formulating hypotheses regarding steps 4 and 5 of the biodegradation potential. A collection of macroplastic particles from two beaches known for high plastic debris accumulation: La Pineda Beach (Lat: 41.0743152. Long: 1.1782226) and Hanko Beach (59.832394, 22.970695). The collected material included macroplastic particles (Nylon, PC, PP, PET, PP, PS, PU, PVC and PVK) and plastic-associated microbiota (Actinomycetota, Bacillota, Bacteroidota, Pseudomonadota, and Unclassified) identified with Fourier Transform Infrared Spectroscopy (FTIR) and 16S rRNA analysis, respectively.

Linking this data with the PDCOGs database allows a pre-experimental inference of biodegradation potential and relevant enzymatic functions (Fig. 6). This case study validates the PDCOGs dataset in two key ways. First, its robustness is ensured by grounding the classification in the widely accepted COGs framework (Galperin et al., 2025), which provides consistency and comparability across datasets. Second, even a basic annotation of microbial communities enables hypothesis generation on the plastic-degrading enzymes involved. Therefore, pairing the PDCOGs database with more complex methods, e.g. whole-genome shotgun metagenomic sequencing (Guruge et al., 2024), could enable *in situ* identification of plastic-degrading enzymes in plastic-associated microbes.

2.4. Limitations

The PDCOGs contain 51 orthologous groups with potential for plastic degradation, as their seed sequences have been previously identified in the literature. This is based on the idea that orthologous proteins sharing functional domains may possess plastic-degrading potential. Therefore, this resource can serve as a valuable tool for hypothesis testing regarding the plastic-degrading potential of proteins, enabling researchers to explore functional predictions during experimental validation. However, the presence of a plastic-degrading functional domain alone does not confirm the protein's ability to break down plastic polymers to any degree. Thus, enzymatic assays are essential to validate activity against plastic polymers and for determining the optimal environmental conditions required for such activity.

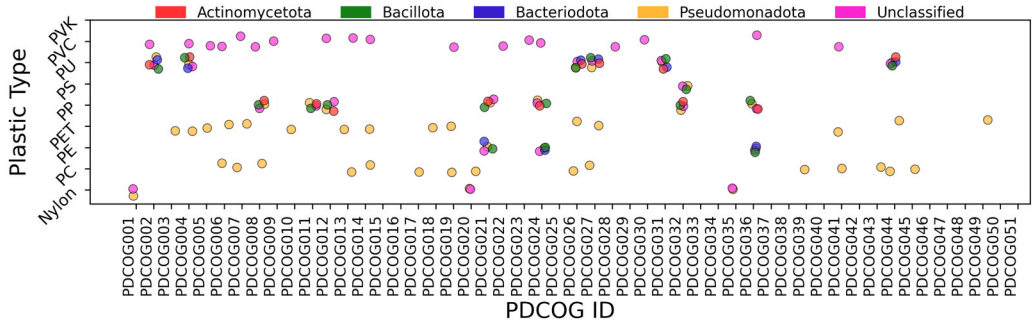

Fig. 6. Plastic-degrading clusters of orthologous groups (PDCOGs) across taxonomic groups and plastic-types. The plot shows the taxonomical groups (actinomycetota: red; bacillota: green; bacteroidota: blue; pseudomonadota: orange; unclassified: pink) identified in the macroplastic particles from beach samples (y axis), and the putative plastic-degrading proteins (n_pdcogs = 51) involved in the degradation process (x axis).

3. Discussion

3.1. Comparison with existing approaches and resources

Individual studies, reviews and web resources report plastic-degrading proteins in microbes (Supplementary Table S1). For instance, the PlasticDB provides a comprehensive list of 228 plastic-degrading enzymes from 824 microorganisms, compiled from the scientific literature (Gambarini et al., 2022). These data were mapped onto metagenomic datasets to identify plastic-degrading potential, leading to the conclusion that such enzymes are widespread regardless of the presence of plastic pollution (Gambarini et al., 2024). Here, we have used these data to identify 662,830 PPDP enzymes from free-living bacteria distributed across the Earth and classified them into 51 orthologous groups. This comprehensive and systematic classification of the PPDP allows hypothesis generation of the plastic-degrading potential in prokaryotes. The previous world, including the PlasticDB compiles, evidence from the scientific literature demonstrating that specific proteins (enzymes) participate in plastic biodegradation. The PDCOGs database builds on this foundation by classifying these enzymes into 51 protein groups. This structured classification enables standardized analysis, facilitates comparisons within and across functional categories and taxa, and—most importantly—supports the generation of new hypotheses for future research. The PDCOGs resource offers a comparative genomics framework for comparative analyses: a) within protein families (e.g., enzymatic activity under different conditions); b) across microbial species (e.g., plastic-degrading activities in archaea vs. bacteria); and c) across environments (e.g., geographical and environmental factors driving local adaptations). In the case study, we showed that mapping PDCOGs onto a 16S rRNA microbial community study can reveal the community's plastic-degrading potential. Therefore, pairing the PDCOGs database with more complex methods, e.g. whole-genome shotgun metagenomic sequencing (Guruge et al., 2024), could enable *in situ* identification of plastic-degrading enzymes in plastic-associated microbes.

3.2. In-depth analysis

The PDCOG database provides a comprehensive and functionally annotated collection of PPDPs, organized into orthologous groups on the basis of sequence homology and enzymatic activity. Most are hydrolases (EC3), followed by oxidoreductases (EC1) and transferases (EC2), spanning diverse COG functional categories. The main categories highlight the biochemical complexity of the process, often involving multistep enzymatic pathways. All PDCOG proteins have known or predicted enzymatic roles. We propose that plastic-contaminated environments may foster the evolution of novel PPDPs through mechanisms such as horizontal gene transfer, gene loss, and rapid sequence diversification, potentially producing enzymes with little similarity to known biodegradative proteins (Tatusov et al., 2000). Polymer structure strongly shapes procaryotic biofilm formation and the distribution of PDCOGs. Overall, synthetic heterochain polymers were linked to more unique PPDP-hosting species (25%) than synthetic homochain (10%) or natural polymers (2%). This analysis is based on seed sequences (Gambarini et al., 2022) derived from prokaryotes known to degrade plastics in lab or natural environments. However, the presence of a PPDP does not guarantee enzymatic activity *in situ*, likely due to: (1) incomplete microbial sampling, (2) enzymatic inactivity under suboptimal environmental conditions, and (3) putative inactivation of the enzyme's main functional domain. Environmental plastic weathering (e.g. via photodegradation, oxidation, or mechanical stress) may also affect microbial colonization and enzymatic access.

Since the term *plastisphere* emerged in 2013, our understanding of plastic-degrading microbial diversity has expanded (Bocci et al., 2024; Zeenat et al., 2021). Over 95% of the surveyed prokaryotic species harbor at least one PDCOG, including $>$ 98% of archaea and $>$ 96% of bacteria, suggesting widespread biodegradation potential. To clarify, these signals should be understood as hypothetical or putative functional expectations, not confirmed activities. The PDCOGs reflect *potential* plastic-degrading proteins inferred from sequence similarity and functional annotation, and thus require experimental validation. Future biochemical and physiological studies will be essential to discriminate which of these candidate enzymes truly participate in plastic biodegradation and under what environmental conditions. We consider this to be one of the main strengths of the PDCOG framework: it provides clear, testable hypotheses for targeted experimental work. Nonetheless, our results highlight that $>$ 95% of microbial genomes contain at least one gene encoding an enzyme with a function that *could potentially* be recruited or co-opted for plastic degradation, reinforcing the notion that the molecular capacity for biodegradation is likely far more widespread than previously recognized.

Among bacteria, *Pseudomonas sp.* has the greatest number of PPDPs in natural environments, aligning with their metabolic versatility (Silby et al., 2011), from degrading medical plastics (*Pseudomonas aeruginosa*) (Howard et al., 2025) toremediating petrochemical waste (Samimi and Shahriari-Moghadam, 2025). We identified three archaeal phyla—Methanobacteriota, Thermoproteota, and Bathyarchaeota—as potential plastic degraders, highlighting a likely underappreciated role for archaea in plastic biodegradation, especially under extreme or anaerobic environments (Gambarini et al., 2021). While PDCOGs are most frequently detected in aquatic systems, terrestrial environments—especially soils and lithic (rock-associated) environments—show disproportionately high PPDP levels. This may reflect selective pressure under nutrient-limited conditions (Kang et al., 2024). In contrast, marine environments, though oligotrophic, support microbial activity through accumulated dissolved organic matter c (Repeta and Boiteau, 2017). Database biases may obscure PDCOG prevalence in some environments, but overall, our findings highlight the role of environmental context in shaping plastic-degrading potential and suggest that microbial adaptation to plastic pollution is both dynamic and environment/habitat specific.

3.3. Real-world usage

A real-world example of pairing PDCOGs with microbial 16S rRNA data illustrates the potential of this resource to generate

hypotheses about plastic-degrading prokaryotes. This type of research is the state-of-the-art for the analysis of environmental and host-associated microbiomes (Mathew et al., 2022), including those from the plastisphere (Magalhães et al., 2024). Furthermore, during the last decade the field of environmental metagenomics has grown exponentially with the use of advanced whole-genome shotgun metagenomic sequencing methods. These methods allow the study of the specific influence of the environment to shape the distribution and variability of prokaryotic (Becsei et al., 2024). Recent studies use metagenomic sequencing to identify the biodiversity in plastic-associated microbial communities and the impact of MNP across environments (Rüthi et al., 2023; Villarrubia-Gómez et al., 2024; Guruge et al., 2024; Becsei et al., 2024). Using the PDCOGs database in microbial metagenomic studies offers a robust tool for detecting plastic-degrading enzymes in environmental and host-associated samples.

3.4. Future applications

The PDCOGs database integrates functional, ecological, and taxonomic information, providing a critical reference for researchers studying microbial plastic degradation. It complements existing resources on plastic-degrading proteins and microbes (Gambarini et al., 2022; Buchholz et al., 2022; Ridley et al., 2024; Wicker et al., 2016; Ellis and Wackett, 2012). Biodegradable polymers, such as polycaprolactone (PCL), polybutylene succinate-coadipate (PBSA), polybutylene succinate (PBS), and polylactic acid (PLA), are more amenable to enzymatic degradation because of their ester linkages (Galperin et al., 2025; Al Hosni et al., 2019). In contrast, high-molecular-weight polymers such as PET require highly efficient and specific enzymes for depolymerization. PDCOGs provide a framework to explore such structure—function relationships and guide the identification of candidate enzymes for biotechnological and material science applications.

Non-specific polymer degradation remains a plausible scenario, and some PPDPs may be involved only in partial or non-targeted breakdown rather than full mineralization into CO_2 and H_2O. Because such incomplete transformations do not necessarily confer environmental benefits, distinguishing between these processes is essential for interpreting their ecological significance. Future research should therefore focus on experimentally validating whether individual PPDPs drive genuine biodegradation or merely partial polymer alteration. Furthermore, in this initial release of the PDCOGs database, the analysis encompasses all bacterial species, irrespective of their exposure to plastic-rich environments, and thus represents a baseline against which future studies can be compared. Microbes that have evolved in plastic-associated niches may harbor enzymes that are more specifically adapted to maximize plastic-degrading capacity.

Addressing the global plastic pollution crisis, particularly the proliferation of MNPs, will require integrative, cross-disciplinary strategies combining genomics, biochemical characterization, synthetic biology, biochemical engineering and environmental microbiology. Given that biodegradation efficacy is strongly influenced by both environmental conditions and polymer structure, PDCOGs provide a foundational resource for advancing our understanding of microbial plastic degradation potential across ecosystems.

4. Methods

4.1. Seeds of putative plastic-degrading proteins

A seed of putative plastic-degrading protein (PPDP) sequences was obtained from the PlasticDB (Gambarini et al., 2022) and articles from the literature to expand the coverage of the dataset (Supplementary Table S4). The initial dataset (available from the PDCOGs webpage) contains a seed of PPDP sequences from bacterial species and potential plastic polymers that can be biodegraded (Supplementary Table S5).

4.2. Orthologous proteins

We collected all proteins from the COGs database (Galperin et al., 2025). This database contains 5640669 protein sequences classified on the basis of orthologous relationships into 5050 COGs. Each protein is functionally annotated, and the COGs are classified into 26 biological functions on the basis of four main functional categories (Supplementary Table S2). As of the 2024 update, the COGDB includes 2296 pieces of genomic information (2103 bacterial and 193 archaeal).

4.2.1. Putative plastic-degrading orthologous groups and phyletic patterns

The PDCOG dataset was created on the basis of our BLASTp results from the NCBI COG database (COGDB) (Galperin et al., 2025) and PlasticDB (Gambarini et al., 2022) (Supplementary Datasets SD1–3). Each sequence from the PPDP dataset was mapped onto the COG dataset with the program BLASP for the identification of potential plastic-degrading orthologous groups. Briefly, a BLASTP search was performed on local databases including NCBI COG database (this database contains 2103 bacteria and 193 archaea complete complete genomes) and the PlasticDB database (this database contains 329 proteins and 857 microbes) to obtain the best COG hits against the PlasticDB (BLASP parameters were outfmt = 6, *e*-value = 0.001, max-target-seqs = 5). These parameters optimize both annotation accuracy and speed. The COG functional category and symbol data were collected from the COGDB.

4.2.2. Functional annotation of the PDCOGs

The functional annotations of the PDCOGs (Supplementary Dataset SD4) were performed with the database Cluster of Orthologous Group Database (COGDB) (Galperin et al., 2025), Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa et al., 2025), InterPro

104.0 (Blum et al., 2025), ExPASy (Bairoch, 2000), and NCBI Taxonomy (Schoch et al., 2020). The KEGG database was used to collect KO codes for each PDCOG, as well as the EC (enzymatic) and reaction codes. The categorical function of each EC class was characterized with ExPASy. The protein functional domain and gene ontology (GO) information of each PDCOG was characterized with the InterPro database.

4.2.3. Reconstruction of the polymers dataset

The systematic workflow for the data curation of the polymers included in the dataset is based on both cheminformatics analyses and literature reviews (Supplementary Table S5 and Supplementary Dataset SD2). We used the PU-REST API of Pubchem DB (Kim et al., 2025) for SMILEs, the Polymer Genome DB, Smiles2Monomers (Dufresne et al., 2015) for monomer mapping, RDKit open source cheminformatics v2025.3.2 in Python for validation of SMILES and monomers, and Beautiful Soup for scraping open access journals from PubMed Central (Sayers et al., 2022). Among the 39 polymers, 12 PSMILEs were obtained from PubChem and the rest were obtained from PI1M DB (Ma and Luo, 2020) for full repeat units vs. fragments, using the filters '−Atomcount $\geq$ 10 to exclude fragments, −Molecular weight 500–5000 Da, Atom count, functional groups, and excluding Non-specific'. Consequently, after the 50 best batches per target polymer were obtained, consensus-based PSMILEs were obtained for the remaining polymers. To evaluate recent advances in the potential biodegradation of additional polymers, including P3HO, PC, PEF, PMCL, PP and PPL, we surveyed 300 peer-reviewed studies published between 2000 and 2025. The selection specifically relied on prokaryotic relevance to these polymers and enzymes used in biodegradation.

4.2.4. Environmental annotation of the PDCOGs

Protein sequences of free-living species with geographical and environmental information were obtained from the Joint Genome Institute (DOE-JGI) (Nordberg et al., 2014). We analyzed 74,651,773 protein sequences with the Python *re* package to identify sequences with geographical and environmental annotations that belong to free-living microbial species. The filtered sequences were mapped on the PDCOGs with BLASTP. Thus, the final dataset included 662,830 protein sequences with environmental and functional annotations (Supplementary Table S3 and Supplementary Datasets SD5–6). These proteins are spread across the Earth in 23 different environments. The distribution map of the PDCOGs was built using pandas, matplotlib, cartopy.crs for handling world map projections, and seaborn for the built-in color palettes. A chi-square test was performed on the PDCOGs dataset to determine significant differences ($p < 0.05$) from an expected distribution of the protein sequences on the basis of protein number in the COGDB.

4.2.5. Plotting and statistical analysis

All data processing, statistical analyses and visualizations were performed with Python (version 3.9.6). Visualizations were generated using Matplotlib (v3.9.4), Plotly (v6.0.0.) and Seaborn (v0.13.2), whereas data processing was performed with the Pandas (v2.2.3), SciPy (v1.13.1), and NumPy (v2.0.2) libraries and the scipy.stats module (wilcoxon). The Wilcoxon signed-rank test (Fig. 2B) was used to assess significant differences (p value < 0.05) in the PDCOG abundance distribution between the bacteria and archaea. The indicator value (IV) in the heatmap (Fig. 3B) was calculated. Briefly, the method combines specificity (Sp) and fidelity (Fd), which are calculated for each enzyme—plastic pair (Wei et al., 2008), and downweighting is applied to reduce the influence of frequently occurring enzymes (Eq. 1). A permutation test (10,000 iterations) was used to assess significance, yielding p values ranging from 0.01 to 0.04 (Fig. 2B).

$$IV = S \cdot F \cdot W$$
$$Sp = \frac{(P \cap e)}{(p \cup e)};\ Fd = \frac{(p \cap e)}{(p)};\ W = \frac{1}{\log(1 + Np)} \quad (1)$$

Where p is the total EC occurrences for the target plastic polymer; e is the total 4th-subclass EC occurrences under the 3rd-subclass EC across plastics; p⋂e is the number of EC occurrences (4th EC subclass) for the target plastic under the 3rd EC subclass; p⋃e is the union of the 3rd EC subclass annotation of the target plastic and across plastics; and W is a downweighting factor.

5. Resource availability

The PDCOGs database is freely available at http://phylobone.com/microworld/PDCOG. The database includes functional and environmental annotation, and protein sequences of the plastic degrading enzymes included in this study. **Lead contacts:** Requests for further information and resources should be directed to and will be fulfilled by the lead contacts, Pere Puigbò (pere.puigbo@uab.cat) and Miho Nakamura (miho.nakamura@utu.fi). **Materials availability:** This study did not generate new unique reagents. **Data availability:** This paper analyzes existing, publicly available data, accessible at (Galperin et al., 2025; Gambarini et al., 2022; Nordberg et al., 2014). **Code availability:** This paper does not report original code. **Additional information:** Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

Author contributions

Conception and design of the study: MN, PP; Funding acquisition: MN, KS, PP; Data collection: SM, LTF, PP; Data analysis: SM, LTF, PP; Manuscript drafting: SM, LTF, PP.; Manuscript revision for critical intellectual content: SM, LTF, KS, MN, PP; Writing the final version of the manuscript: MN, PP. All authors have read and agreed to the published version of the manuscript.

CRediT authorship contribution statement

Shakira Mustari: Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation. **Phạm Loan:** Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Puigbo Pere:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Kari Saikkonen:** Writing – review & editing, Writing – original draft, Validation, Resources, Project administration. **Miho Nakamura:** Writing – review & editing, Supervision, Resources, Project administration, Investigation, Conceptualization.

Funding

Scandinavian-Japan Sasakawa foundation (GA25-FUB-0132). PP is a Serra Húnter Professor in the public university system of Catalonia.

Declaration of Competing Interest

The authors declare that they have no conflict of interest.

Acknowledgements

We thank members of the MicroWorld team, students and collaborators for their helpful discussions.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.eti.2026.104872.

Data availability

All data is publicly available.

References

Al Hosni, A.S., Pittman, J.K., Robson, G.D., 2019. Microbial degradation of four biodegradable polymers in soil and compost demonstrating polycaprolactone as an ideal compostable plastic. Waste Manag 97, 105–114.
Amaral-Zettler, L.A., Zettler, E.R., Mincer, T.J., 2020. Ecology of the plastisphere. Nat. Rev. Microbiol 18, 139–151.
Amelia, T.S.M., Khalik, W.M.A.W.M., Ong, M.C., Shao, Y.T., Pan, H.-J., Bhubalan, K., 2021. Marine microplastics as vectors of major ocean pollutants and its hazards to the marine ecosystem and humans. Prog. Earth Planet. Sci. 8, 12.
Bairoch, A., 2000. The ENZYME database in 2000. Nucleic Acids Res 28, 304–305.
Barrett, J., Chase, Z., Zhang, J., Holl, M.M.B., Willis, K., Williams, A., Hardesty, B.D., Wilcox, C., 2020. Microplastic Pollution in Deep-Sea Sediments From the Great Australian Bight. Front. Mar. Sci. 7.
Becsei, Á., Fuschi, A., Otani, S., Kant, R., Weinstein, I., Alba, P., Stéger, J., Visontai, D., Brinch, C., de Graaf, M., et al., 2024. Time-series sewage metagenomics distinguishes seasonal, human-derived and environmental microbial communities potentially allowing source-attributed surveillance. Nat. Commun. 15, 7551.
Bergmann, M., Collard, F., Fabres, J., Gabrielsen, G.W., Provencher, J.F., Rochman, C.M., van Sebille, E., Tekman, M.B., 2022. Plastic pollution in the Arctic. Nat. Rev. Earth Environ.
Bermúdez, J.R., Swarzenski, P.W., 2021. A microplastic size classification scheme aligned with universal plankton survey methods. MethodsX 8, 101516.
Blum, M., Andreeva, A., Florentino, L.C., Chuguransky, S.R., Grego, T., Hobbs, E., Pinto, B.L., Orr, A., Paysan-Lafosse, T., Ponamareva, I., et al., 2025. InterPro: the protein sequence classification resource in 2025. Nucleic Acids Res 53, D444–D456.
Bocci, V., Galafassi, S., Levantesi, C., Crognale, S., Amalfitano, S., Congestri, R., Matturro, B., Rossetti, S., Di Pippo, F., 2024. Freshwater plastisphere: a review on biodiversity, risks, and biodegradation potential with implications for the aquatic ecosystem health. Front. Microbiol 15, 1395401.
Bridgeman, L., Cimbalo, A., López-Rodríguez, D., Pamies, D., Frangiamone, M., 2025. Exploring toxicological pathways of microplastics and nanoplastics: Insights from animal and cellular models. J. Hazard. Mater. 490, 137795.
Buchholz, P.C.F., Feuerriegel, G., Zhang, H., Perez-Garcia, P., Nover, L.-L., Chow, J., Streit, W.R., Pleiss, J., 2022. Plastics degradation by hydrolytic enzymes: The plastics-active enzymes database-PAZy. Proteins 90, 1443–1456.
Chen, Z., Shi, X., Zhang, J., Wu, L., Wei, W., Ni, B.-J., 2023. Nanoplastics are significantly different from microplastics in urban waters. Water Res. X 19, 100169.
De Angelis, M., Gobbetti, M., 1999. Lactobacillus sanfranciscensis CB1: manganese, oxygen, superoxide dismutase and metabolism. Appl. Microbiol. Biotechnol. 51, 358–363.
Dufresne, Y., Noé, L., Leclère, V., Pupin, M., 2015. Smiles2Monomers: a link between chemical and biological structures for polymers. J. Chemin.. 7, 62.
Ellis, L.B., Wackett, L.P., 2012. Use of the University of Minnesota Biocatalysis/Biodegradation Database for study of microbial degradation. Microb. Inform. Exp. 2 (1).
Galperin, M.Y., Vera Alvarez, R., Karamycheva, S., Makarova, K.S., Wolf, Y.I., Landsman, D., Koonin, E.V., 2025. COG database update 2024. Nucleic Acids Res 53, D356–D363.
Gambarini, V., Pantos, O., Kingsbury, J.M., Weaver, L., Handley, K.M., Lear, G., 2021. Phylogenetic Distribution of Plastic-Degrading Microorganisms. mSystems 6.
Gambarini, V., Pantos, O., Kingsbury, J.M., Weaver, L., Handley, K.M., Lear, G., 2022. PlasticDB: a database of microorganisms and proteins linked to plastic biodegradation. Database (Oxf.) 2022.
Gambarini, V., Drost, C.J., Kingsbury, J.M., Weaver, L., Pantos, O., Handley, K.M., Lear, G., 2024. Uncoupled: investigating the lack of correlation between the transcription of putative plastic-degrading genes in the global ocean microbiome and marine plastic pollution. Environ. Micro 19, 34.

García, B., González-Sabín, J., Menéndez, N., Braña, A.F., Núñez, L.E., Morís, F., Salas, J.A., Méndez, C., 2011. The chromomycin CmmA acetyltransferase: a membrane-bound enzyme as a tool for increasing structural diversity of the antitumour mithramycin. Microb. Biotechnol. 4, 226–238.
Gross, N., Muhvich, J., Ching, C., Gomez, B., Horvath, E., Nahum, Y., Zaman, M.H., 2025. Effects of microplastic concentration, composition, and size on Escherichia coli biofilm-associated antimicrobial resistance. Appl. Environ. Microbiol 91, e0228224.
Guruge, K.S., Goswami, P., Kanda, K., Abeynayaka, A., Kumagai, M., Watanabe, M., Tamamura-Andoh, Y., 2024. Plastiome: Plastisphere-enriched mobile resistome in aquatic environments. J. Hazard. Mater. 471, 134353.
Howard, S.A., de Dios, R., Maslova, E., Myridakis, A., Miller, T.H., McCarthy, R.R., 2025. Pseudomonas aeruginosa clinical isolates can encode plastic-degrading enzymes that allow survival on plastic and augment biofilm formation. Cell Rep. 44, 115650.
Hu, Z., Klupt, K., Zechel, D.L., Jia, Z., Howe, G., 2025. Mining Thermophile Genomes for New PETases with Exceptional Thermostabilities Using Sequence Similarity Networks. Chembiochem, e2500065.
Kakar, F.L., Okoye, F., Onyedibe, V., Hamza, R., Dhar, B.R., Elbeshbishy, E., 2023. Climate change interaction with microplastics and nanoplastics pollution. Current developments in biotechnology and bioengineering. Elsevier, pp. 387–403.
Kanehisa, M., Furumichi, M., Sato, Y., Matsuura, Y., Ishiguro-Watanabe, M., 2025. KEGG: biological systems database as a model of the real world. Nucleic Acids Res 53, D672–D677.
Kang, H., Xue, Y., Cui, Y., Moorhead, D.L., Lambers, H., Wang, D., 2024. Nutrient limitation mediates soil microbial community structure and stability in forest restoration. Sci. Total Environ. 935, 173266.
Kelly, B.P., Whiteley, A., Tallmon, D., 2010. The Arctic melting pot. Nature 468, 891.
Kim, S., Chen, J., Cheng, T., Gindulyte, A., He, J., He, S., Li, Q., Shoemaker, B.A., Thiessen, P.A., Yu, B., et al., 2025. PubChem 2025 update. Nucleic Acids Res 53, D1516–D1525.
Lee, J.R., Raymond, B., Bracegirdle, T.J., Chadès, I., Fuller, R.A., Shaw, J.D., Terauds, A., 2017. Climate change drives expansion of Antarctic ice-free habitat. Nature 547, 49–54.
Lehmann, M., Prohaska, C., Zeldes, B., Poehlein, A., Daniel, R., Basen, M., 2023. Adaptive laboratory evolution of a thermophile toward a reduced growth temperature optimum. Front. Microbiol 14, 1265216.
Ma, R., Luo, T., 2020. PI1M: A benchmark database for polymer informatics. J. Chem. Inf. Model 60, 4684–4690.
Magalhães, E.A., de Jesus, H.E., Pereira, P.H.F., Gomes, A.S., Santos, H.F.D., 2024. Beach sand plastispheres are hotspots for antibiotic resistance genes and potentially pathogenic bacteria even in beaches with good water quality. Environ. Pollut. 344, 123237.
Mamun, A.A., Prasetya, T.A.E., Dewi, I.R., Ahmad, M., 2023. Microplastics in human food chains: Food becoming a threat to health safety. Sci. Total Environ. 858, 159834.
Mathew, S.A., Muola, A., Saikkonen, K., Saloniemi, I., Marjo, H., Puigbò, P., 2022. Quantification of the Potential Impact of Glyphosate-Based Products on Microbiomes. J. Vis. Exp. e *63109*.
Meng, D.-C., Wang, Y., Wu, L.-P., Shen, R., Chen, J.-C., Wu, Q., Chen, G.-Q., 2015. Production of poly(3-hydroxypropionate) and poly(3-hydroxybutyrate-co-3-hydroxypropionate) from glucose by engineering Escherichia coli. Metab. Eng. 29, 189–195.
Meng, L., Liang, L., Shi, Y., Yin, H., Li, L., Xiao, J., Huang, N., Zhao, A., Xia, Y., Hou, J., 2024. Biofilms in plastisphere from freshwater wetlands: Biofilm formation, bacterial community assembly, and biogeochemical cycles. J. Hazard. Mater. 476, 134930.
Nordberg, H., Cantor, M., Dusheyko, S., Hua, S., Poliakov, A., Shabalov, I., Smirnova, T., Grigoriev, I.V., Dubchak, I., 2014. The genome portal of the Department of Energy Joint Genome Institute: 2014 updates. Nucleic Acids Res 42, D26–31.
Obbard, R.W., Sadri, S., Wong, Y.Q., Khitun, A.A., Baker, I., Thompson, R.C., 2014. Global warming releases microplastic legacy frozen in Arctic Sea ice. Earth'S. Future 2, 315–320.
Repeta, D.J., Boiteau, R.M., 2017. Org. Nutr. Chem. Mar. Micro Chem. Micro NCBI Bookshelf.
Ridley, R.S., Conrad, R.E., Lindner, B.G., Woo, S., Konstantinidis, K.T., 2024. Potential routes of plastics biotransformation involving novel plastizymes revealed by global multi-omic analysis of plastic associated microbes. Sci. Rep. 14, 8798.
Rüthi, J., Cerri, M., Brunner, I., Stierli, B., Sander, M., Frey, B., 2023. Discovery of plastic-degrading microbial strains isolated from the alpine and Arctic terrestrial plastisphere. Front. Microbiol 14, 1178474.
Samimi, M., Shahriari-Moghadam, M., 2025. Efficient biodegradation of crude oil by novel *Pseudomonas aeruginosa* strain isolated from petrochemical wastewater: identification and optimization. Pet. Sci. Technol. 43, 1182–1199.
Sayers, E.W., Bolton, E.E., Brister, J.R., Canese, K., Chan, J., Comeau, D.C., Connor, R., Funk, K., Kelly, C., Kim, S., et al., 2022. Database resources of the national center for biotechnology information. Nucleic Acids Res 50, D20–D26.
Schoch, C.L., Ciufo, S., Domrachev, M., Hotton, C.L., Kannan, S., Khovanskaya, R., Leipe, D., Mcveigh, R., O'Neill, K., Robbertse, B., et al., 2020. NCBI Taxonomy: a comprehensive update on curation, resources and tools. Database (Oxf.) 2020, baaa062.
Sharma, S., Sharma, V., Chatterjee, S., 2023. Contribution of plastic and microplastic to global climate change and their conjoining impacts on the environment - A review. Sci. Total Environ. 875, 162627.
Shruti, V.C., Pérez-Guevara, F., Elizalde-Martínez, I., Kutralam-Muniasamy, G., 2020. First study of its kind on the microplastic contamination of soft drinks, cold tea and energy drinks - Future research and environmental considerations. Sci. Total Environ. 726, 138580.
Silby, M.W., Winstanley, C., Godfrey, S.A.C., Levy, S.B., Jackson, R.W., 2011. Pseudomonas genomes: diverse and adaptable. FEMS Microbiol. Rev. 35, 652–680.
Taipale, S.J., Vesamäki, J., Kautonen, P., Kukkonen, J.V.K., Biasi, C., Nissinen, R., Tiirola, M., 2023. Biodegradation of microplastic in freshwaters: A long-lasting process affected by the lake microbiome. Environ. Microbiol 25, 2669–2680.
Tatusov, R.L., Koonin, E.V., Lipman, D.J., 1997. A genomic perspective on protein families. Science 278, 631–637.
Tatusov, R.L., Galperin, M.Y., Natale, D.A., Koonin, E.V., 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res 28, 33–36.
Tong, H., Zhong, X., Duan, Z., Yi, X., Cheng, F., Xu, W., Yang, X., 2022. Micro- and nanoplastics released from biodegradable and conventional plastics during degradation: Formation, aging factors, and toxicity. Sci. Total Environ. 833, 155275.
Villarrubia-Gómez, P., Carney Almroth, B., Eriksen, M., Ryberg, M., Cornell, S.E., 2024. Plastics pollution exacerbates the impacts of all planetary boundaries. One Earth 7, 2119–2138.
Watabe, Y., Shimomura, S., Ono, K., Sasanuma, Y., 2023. Conformational characteristics of poly(3-hydroxyvalerate) (P3HV) and structure-property relationships of P3HV and poly(3-hydroxybutyrate). Phys. Chem. Chem. Phys. 25, 27034–27044.
Wei, H., Therrien, C., Blanchard, A., Guan, S., Zhu, Z., 2008. The Fidelity Index provides a systematic quantitation of star activity of DNA restriction endonucleases. Nucleic Acids Res 36, e50.
Wicker, J., Lorsbach, T., Gütlein, M., Schmid, E., Latino, D., Kramer, S., Fenner, K., 2016. enviPath–The environmental contaminant biotransformation pathway resource. Nucleic Acids Res 44, D502–8.
Yang, B., Liang, S., Liu, H., Liu, J., Cui, Z., Wen, J., 2018. Metabolic engineering of Escherichia coli for 1,3-propanediol biosynthesis from glycerol. Bioresour. Technol. 267, 599–607.
Yang, X.-G., Wen, P.-P., Yang, Y.-F., Jia, P.-P., Li, W.-G., Pei, D.-S., 2022. Plastic biodegradation by in vitro environmental microorganisms and in vivo gut microorganisms of insects. Front. Microbiol 13, 1001750.
Zeenat, Elahi, A., Bukhari, D.A., Shamim, S., Rehman, A., 2021. Plastics degradation by microbes: A sustainable approach. J. King Saud. Univ. Sci. 33, 101538.
Zettler, E.R., Mincer, T.J., Amaral-Zettler, L.A., 2013. Life in the "plastisphere": microbial communities on plastic marine debris. Environ. Sci. Technol. 47, 7137–7146.