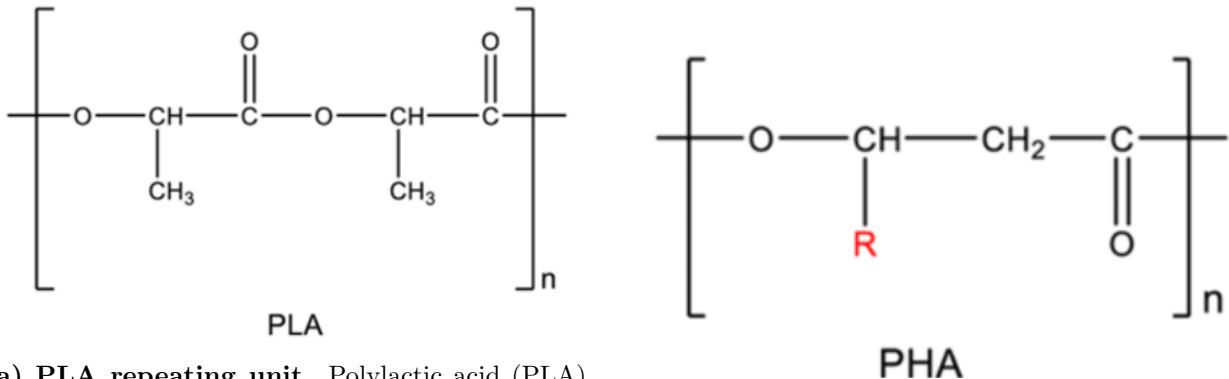


Hybrid Genome Assembly and Multi-Omics of PLA/PHA-Degrading Isolates

Narrative with fully expanded captions

1 Biopolymer context: PLA and PHA chemistry

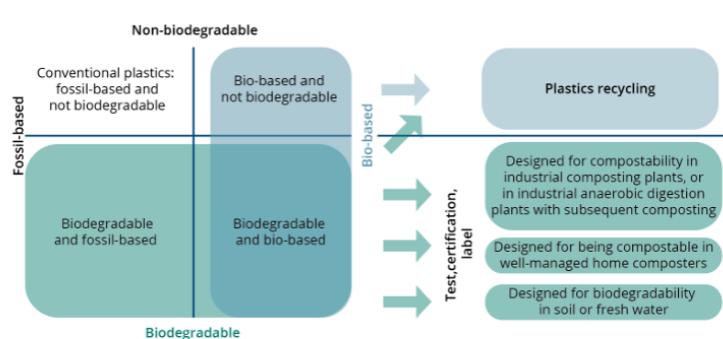


(a) PLA repeating unit. Polylactic acid (PLA) is a thermoplastic polyester obtained from renewable resources (corn starch, sugarcane). Its backbone consists of lactic acid monomers joined by *ester linkages*, highlighted here. These bonds are chemically labile and thus the natural targets of hydrolytic enzymes such as esterases, lipases, and cutinases. However, PLA's crystallinity and stereochemistry (PLLA vs PDLA blends) strongly influence degradation kinetics. Industrial composting reaches the elevated temperatures ($\sim 60^\circ\text{C}$) needed for efficient hydrolysis, whereas ambient soil/compost conditions rarely achieve this, making PLA persistence a real-world issue.

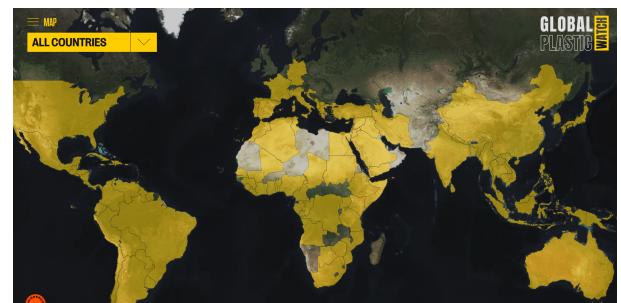
(b) PHA repeating unit. Polyhydroxyalkanoates (PHAs) are microbially synthesized polyesters accumulated as intracellular carbon/energy reserves. The side chain (*R*) varies, producing short-chain-length (scl-PHA, C3–C5) or medium-chain-length (mcl-PHA, C6–C14) polymers. This structural variability affects melting point, crystallinity, and mechanical properties, and dictates the substrate specificity of PHA depolymerases. Unlike PLA, PHAs are truly biodegradable under a wide range of environmental conditions, making them a benchmark against which other “biodegradable plastics” are compared.

Figure 1: PLA and PHA chemistry. These diagrams illustrate why ester bond–hydrolyzing enzymes are focal to this project: they are the molecular “entry points” into breaking down bioplastic polymers.

2 Plastic pollution and bioplastics in context



(a) Conceptual matrix of plastics. This quadrant classifies materials according to (i) origin: fossil- vs bio-based, and (ii) biodegradability: resistant vs degradable. PLA and PHA fall into the bio-based/biodegradable category, yet in practice their environmental degradation varies greatly depending on local temperature, pH, moisture and microbial consortia. This figure is used to frame why even “biodegradable plastics” must be critically studied for real-world breakdown.



(b) Global Plastic Watch planisphere.

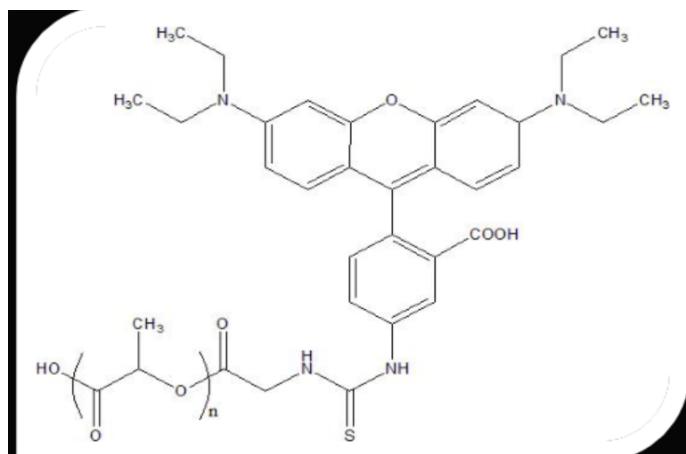
Satellite-based detection of plastic leakage hotspots worldwide. While used here as an *illustrative example*, it highlights the urgency of identifying robust enzymatic solutions. Regions with high leakage may overlap with composting and landfill environments where PLA/PHA could accumulate, reinforcing the relevance of microbial degraders.

Figure 2: Context of plastic pollution. PLA and PHA are marketed as solutions, but global leakage data emphasize the necessity of validating their biodegradability under real-world conditions.

3 Screening of isolates



(a) *Purpureocillium lilacinum* on PLA plate. Growth on PLA-containing agar reveals a degradation halo around the colony. This candidate isolate, listed in the report’s isolate table, was prioritized due to consistent halo formation, fast growth, and known associations with environmental biodegradation.



(b) Rhodamine B assay. Rhodamine B binds PLA particles. Upon ester hydrolysis by extracellular enzymes, local changes in hydrophobicity reduce dye binding, and fluorescence under UV light reveals zones of polymer breakdown. This assay provides a rapid, qualitative proxy for esterase/cutinase activity, suitable for screening large numbers of isolates before omics analyses.

Figure 3: Functional screening of isolates. The combination of PLA agar halos and Rhodamine B fluorescence provides strong preliminary evidence for polyesterase activity.

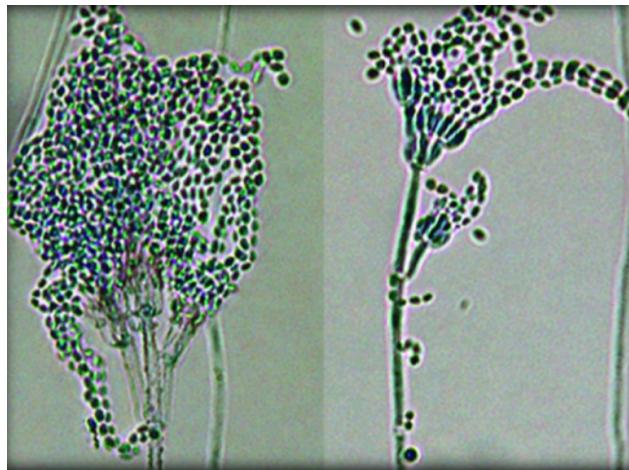
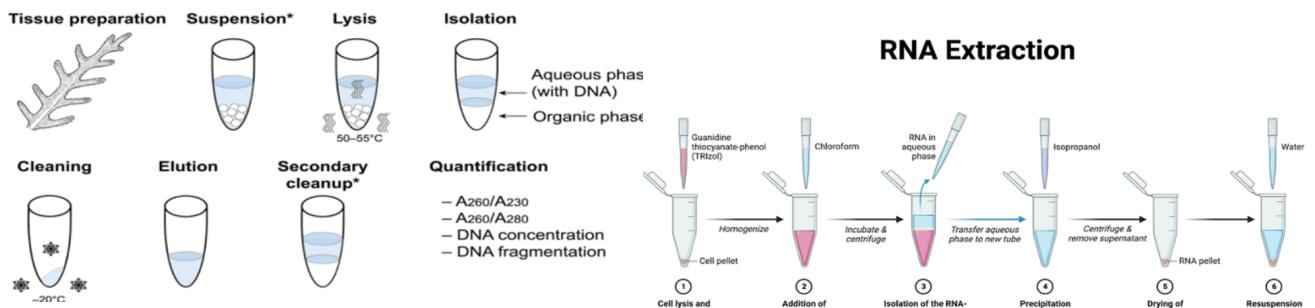


Figure 4: Morphology of *Purpureocillium lilacinum*. Microscopy reveals the diagnostic conidiophores and conidial chains. This morphological evidence, together with ITS sequencing, confirmed isolate identity. Cited in the report as a strong candidate, this species is environmentally relevant and is known for plastic-degrading potential.

4 Nucleic acid extractions

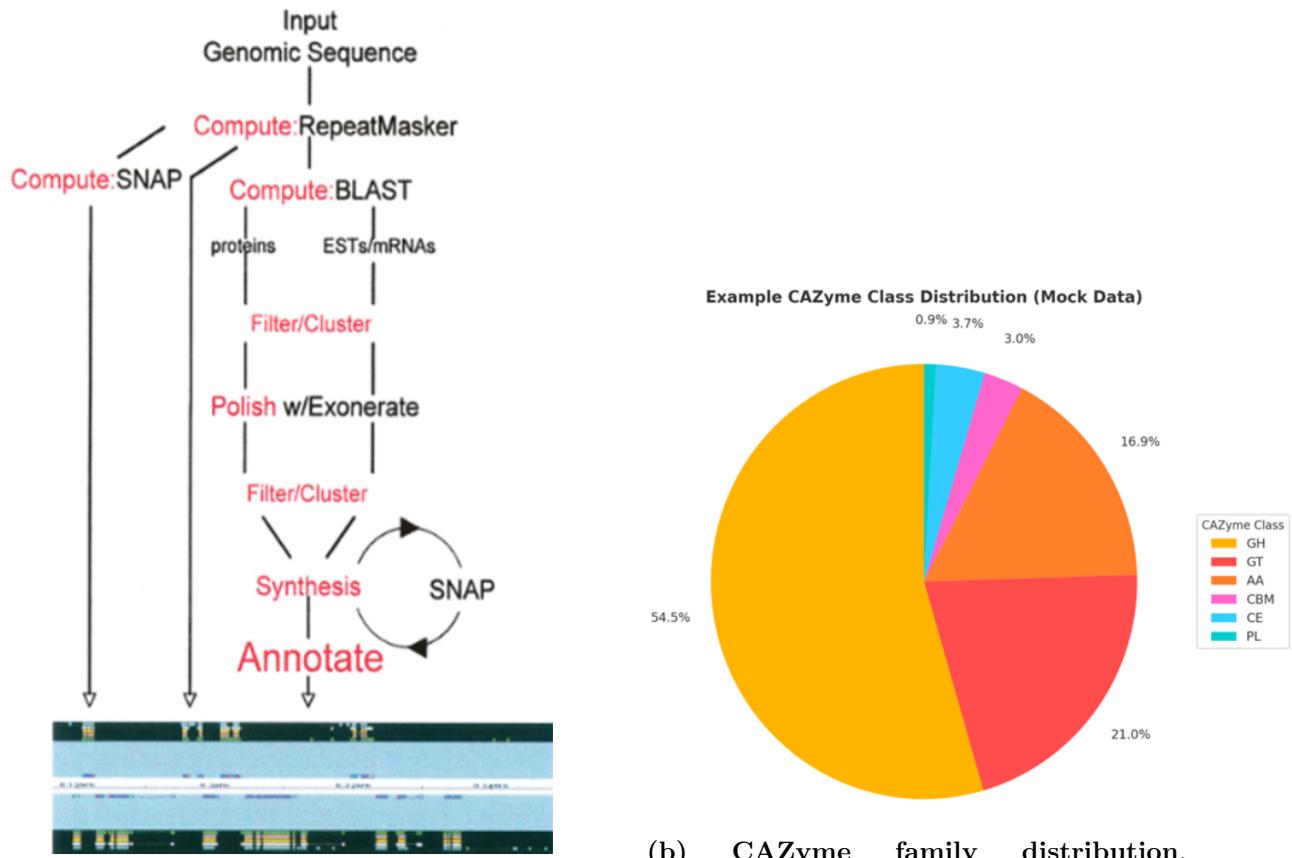


(a) DNA extraction workflow (CTAB-based). High molecular weight genomic DNA is essential for Nanopore sequencing. CTAB effectively removes polysaccharides and phenolic compounds common in fungi. Subsequent phenol-chloroform purification and RNase treatment improve purity, measured by absorbance ratios (A₂₆₀/A₂₈₀, A₂₆₀/A₂₃₀) and verified with Qubit fluorometry.

(b) RNA extraction workflow (TRIzol + silica column). Organic phase separation removes proteins and lipids; silica column purification eliminates polysaccharides and secondary metabolites. High RNA integrity (RIN >8) is critical for RNA-Seq. This workflow balances yield and integrity, ensuring reliable transcriptome data.

Figure 5: DNA and RNA preparation. These protocols were selected for their robustness in fungal samples and compatibility with hybrid genome assembly and RNA-Seq analysis.

5 Genome assembly and annotation

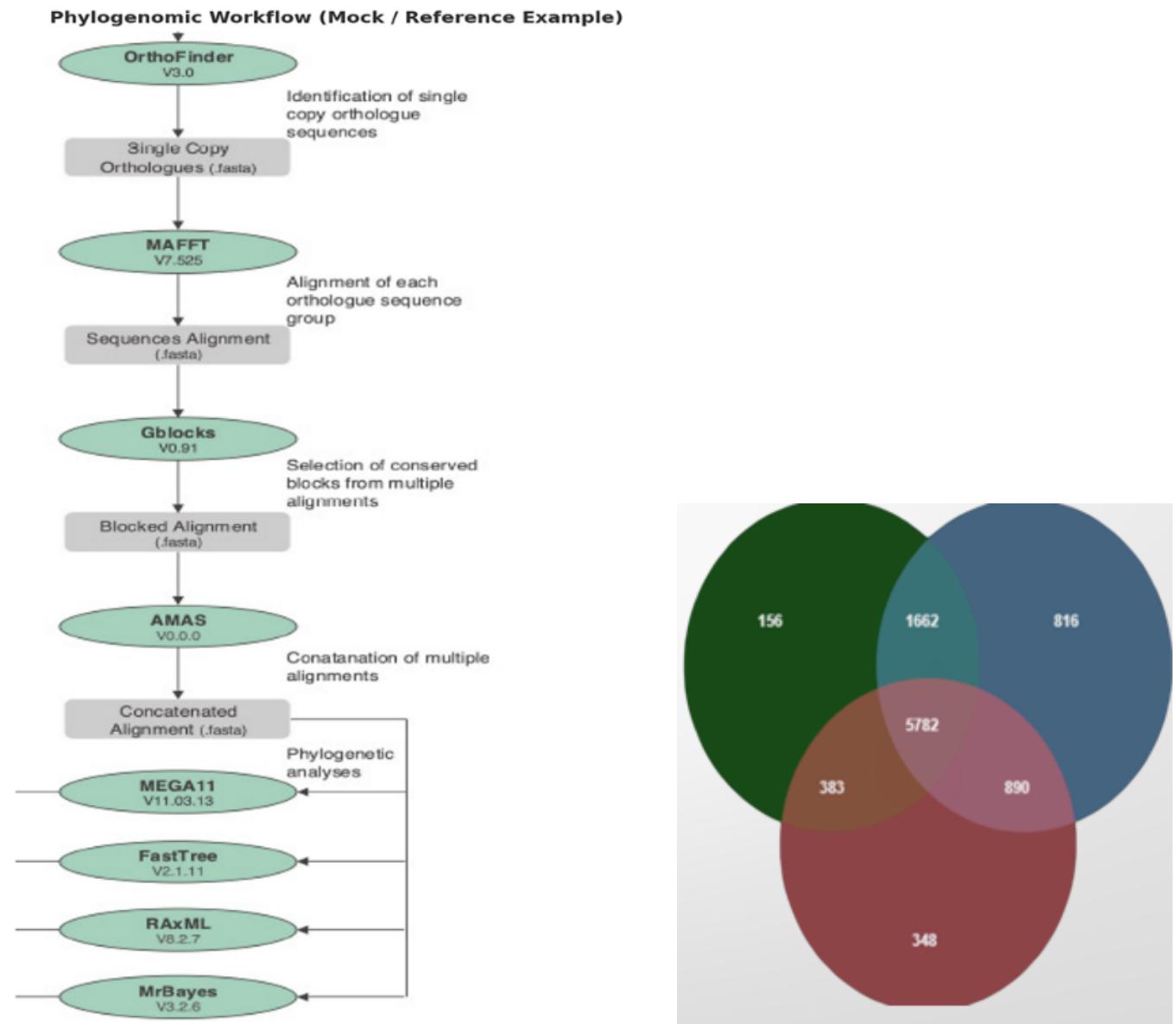


(a) Role of MAKER3 pipeline. MAKER integrates *ab-initio* predictors (SNAP, Augustus) with protein homology and transcript evidence. Iterative retraining improves model accuracy. The output provides structural annotation (gene models, exons, UTRs) and functional annotations when linked to databases. Without this step, the raw genome assembly would remain biologically uninterpretable.

(b) CAZyme family distribution. CAZymes (Carbohydrate-Active enZymes) include glycoside hydrolases, carbohydrate esterases, and auxiliary oxidases. Among them, cutinases and esterases are implicated in polyester degradation. Their relative abundance provides evidence for enzymatic potential relevant to plastic biodegradation.

Figure 6: Genome annotation. MAKER enables gene-level interpretation of the raw genome and links predicted genes to enzymatic functions, highlighting degradative potential.

6 Comparative genomics



(a) OrthoFinder pipeline. OrthoFinder groups genes into orthogroups, extracts single-copy orthologs, aligns them (MAFFT), filters conserved blocks (Gblocks), concatenates alignments (AMAS), and infers phylogenies (RAxML, MrBayes). This workflow provides robust phylogenomic placement of the isolate and highlights lineage-specific genes.

(b) Orthogroup Venn diagram. Shared vs unique orthogroups reveal gene families present only in PLA/PHA-degrading lineages. These unique sets are candidates for encoding novel depolymerases or stress-response enzymes.

Figure 7: Comparative genomics and phylogeny. Evolutionary analysis provides both context and candidate gene families specifically associated with plastic-degrading isolates.

7 Transcriptomics

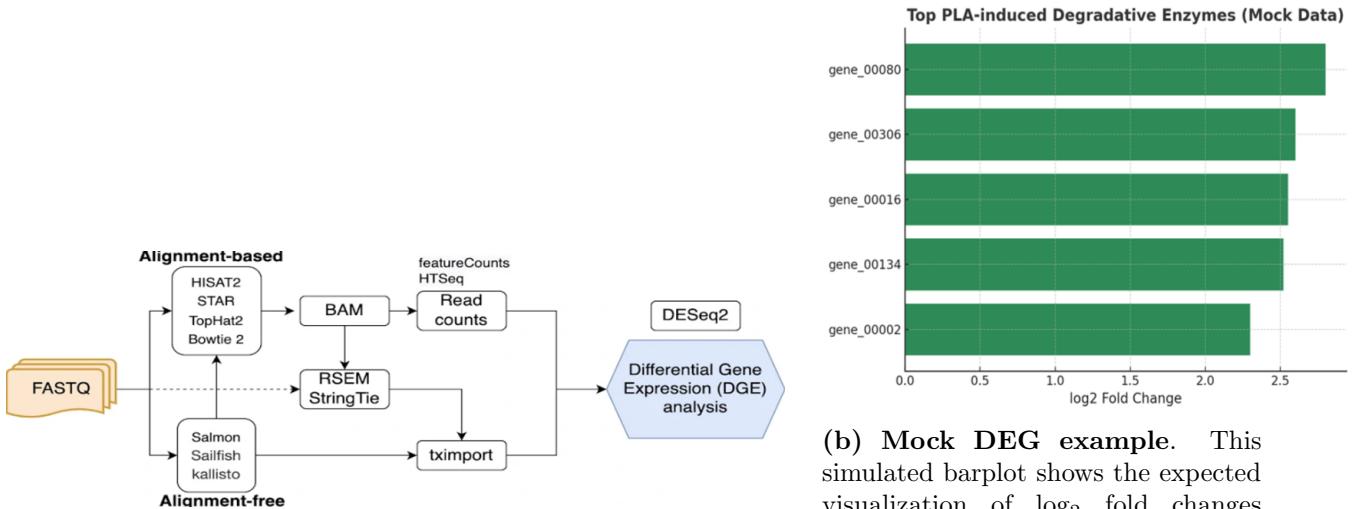


Figure 8: Transcriptome response to PLA. This pipeline identifies which genes are transcriptionally activated in response to PLA, directly linking genomic potential to functional induction.

8 Budget allocation

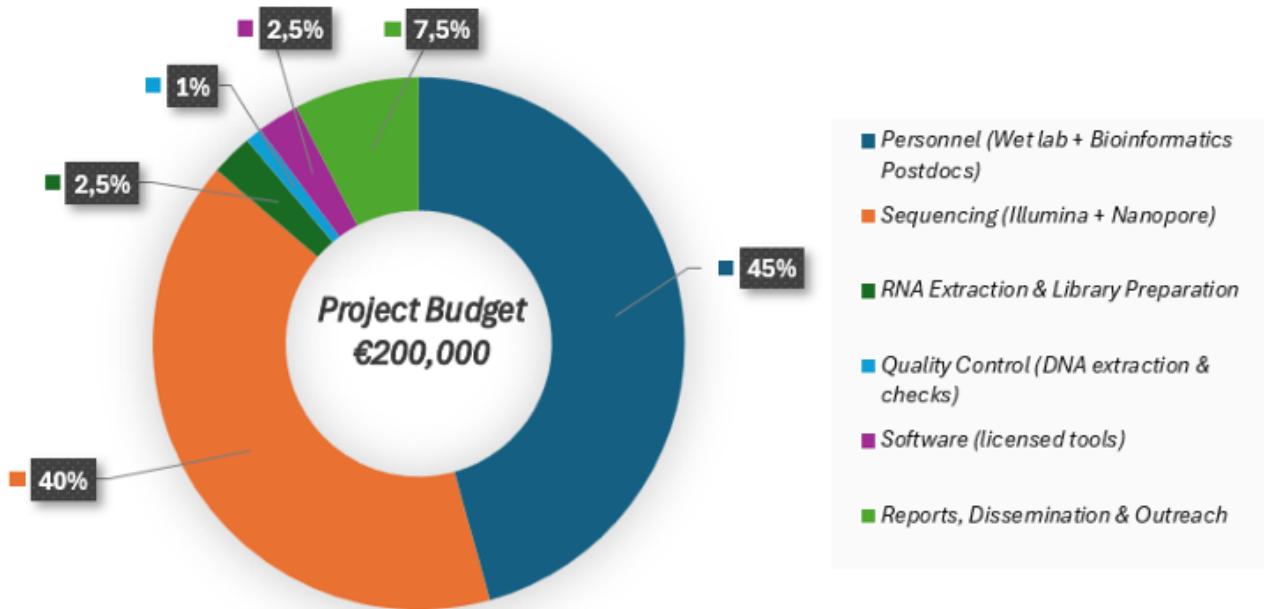


Figure 9: Budget allocation and strategic rationale. The largest fraction (~45%) is dedicated to personnel: wet-lab technicians, computational analysts, and project managers. This guarantees continuity, robust protocols, and reproducibility. Sequencing (~40%) covers NovaSeq Illumina runs and Oxford Nanopore flow cells. The strategy is to conduct an initial pilot run to optimize protocols, followed by scale-up sequencing with multiplexing to reduce cost per sample. RNA extraction and library preparation (~7.5%) ensure replication and statistical power for transcriptomics, using stranded kits for directional information. Quality control (~2.5%) covers consumables such as Qubit kits, Bioanalyzer chips, and ERCC spike-ins to normalize expression data. Software and computational credits (~2.5%) fund licensed tools and cloud-based compute resources. Dissemination (~1%) supports open-access publishing and outreach. The overarching principle is to maximize biological insight per euro: prioritize sequencing depth for genome contiguity, batch RNA-Seq libraries to enhance power, and invest in reproducibility and dissemination to ensure impact beyond the project.

9 Final Integrated Workflow

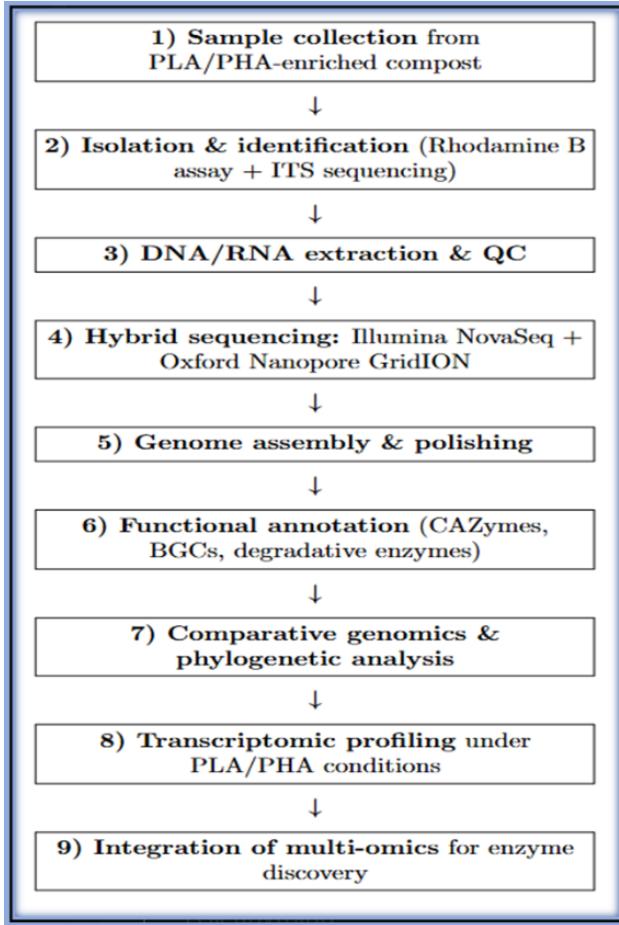


Figure 10: End-to-end workflow of the project. The pipeline integrates all experimental and computational steps: (i) *Sampling* from PLA/PHA-enriched compost (20 isolates obtained, 5 showing clear PLA/PHA degradation activity); (ii) *Functional screening* on PLA plates with Rhodamine B fluorescence to detect esterase/cutinase activity; (iii) *DNA and RNA extraction* (CTAB protocol for HMW DNA; TRIzol+column for RNA with RIN >8); (iv) *Hybrid sequencing strategy* combining Illumina NovaSeq (short reads, high accuracy) with Oxford Nanopore GridION (long reads, high contiguity); (v) *Hybrid assembly & QC* using Flye (Nanopore-first assembly) and Pilon polishing (Illumina), with quality metrics from QUAST (contiguity, misassemblies) and BUSCO (>90% completeness); (vi) *Annotation & functional profiling* through MAKER3 (ab-initio + homology + RNA evidence) and CAZyme annotation with dbCAN3, highlighting esterases, cutinases, lipases; (vii) *Comparative genomics* with OrthoFinder (orthogroups, single-copy orthologs) and synteny analysis (MCScanX, MAUVE) to detect unique gene clusters and structural rearrangements; (viii) *Transcriptomics* with 12 RNA-Seq libraries (Control, PLA, PHA, Blank ×3 replicates, time-points 12h/24h/36h), quantified with Salmon, imported via `tximport`, and tested in DESeq2 (FDR 0.05, $-\log_2\text{FC} \geq -2$); (ix) *Integration & outcome*: linking genome content and expression response, we prioritized candidate esterases, cutinases, lipases, and PHA depolymerases for downstream validation.