Peter_Huang_2024_BINF7001_ Superworm Gut Microbiome Report

Abstract

This paper analyzes microbial community dynamics and functional shifts in superworms (*Zophobas morio*) under three feeding conditions: bran, polystyrene (PS), and no food. Through gene and genome-centric approaches, a metabolically active and diverse microbial community was observed in the bran condition, with key taxa dominating nutrient-rich environments. In contrast, the no food condition exhibited reduced diversity but increased evenness, as microbes adapted by upregulating nutrient-scavenging pathways to survive nutrient deprivation. In the PS condition, moderate microbial activity was driven by genes linked to plastic degradation, revealing the microbial potential for polymer breakdown. Both PS and no food conditions triggered stress responses, reflected in the upregulation of stress-response genes and nutrient acquisition pathways. Comparisons with studies by Sun et al. and Jiang et al. further underscore the adaptive mechanisms of microbial taxa to environmental stressors like PS ingestion. While microbial diversity was reduced in the PS condition, metabolic pathways tied to oxidative stress and carbohydrate transport were upregulated, indicating functional resilience. This research highlights the potential of superworm gut microbiomes in plastic biodegradation and advocates for future multi-omics studies to deepen the understanding of microbial adaptations, with potential biotechnological applications for addressing plastic pollution.

Introduction

Plastic pollution is a global issue with disastrous environmental impacts, and microbial biodegradation has emerged as a potential solution. Research on taxa capable of degrading plastics has expanded our understanding of microbial interactions with synthetic materials. Di Gregorio et al. (2024) highlighted the role of marine microbiomes, in degrading polybutylene succinate-co-adipate (PBSA), identifying previously unreported bacteria like *Kribbella* and *Streptomyces* involved in hydrolytic activities. This study investigates the gut microbiota of *Acartia tonsa*, revealing the association of various bacterial genera with plastic degradation potential, contributing to biotechnological solutions in marine environments. Similarly, Jang et al. (2024) investigates the degradation of low-density polyethylene (LDPE) in the human gut microbiome, isolating pathogens like *Enterococcus casseliflavus* capable of degrading these plastics. This finding raises important questions about the long-term fate of plastics within human systems. Jiang et al. (2024) demonstrates that soil-dwelling *Protaetia brevitarsis* larvae have a remarkable capacity to degrade polystyrene (PS). These findings situate soil microbiomes as key players in the destiny of plastic degradation in terrestrial ecosystems, offering hope for biotechnological waste management. Furthermore, Peng et al. (2024) conducted a study on PVC degradation by *Spodoptera frugiperda* and *Tenebrio molitor*, employing 16S rRNA sequencing and DESeq2 for microbial function prediction. The study revealed the role of *Enterococcus casseliflavus* in PVC degradation, advancing knowledge on microbial plastic biodegradation. Lastly, Satta et al. (2024) explores landfill-derived microbial communities exposed to polyethylene terephthalate (PET), revealing distinct metabolic roles between planktonic and plastic-adhering microbes. Their study emphasizes that these adaptive microbial strategies are pivotal to determining the ultimate fate of plastics in contaminated environments and opens new pathways for harnessing this biotic resilience for plastic upcycling. The goal of this study is to further explore the adaptive responses of microbial communities to plastic contamination, aiming to uncover microbial enzymes and pathways capable of breaking down plastics and providing innovative strategies for biotechnological plastic waste management.

Materials and Methods

In the gene-centric approach, raw reads are pre-processed by removing host sequences to ensure only microbial reads remain. The non-host reads are then processed for 16S rRNA extraction, where small subunit rRNA genes are identified, allowing the microbial taxonomic structure of the community to be profiled [17]. Only a small percentage of reads are expected to encode 16S rRNA genes because 16S rRNA genes make up a small fraction of the total microbial genome. The majority of reads will be protein-coding or non-coding sequences, so assembling multiple reads is necessary to reconstruct the complete 16S rRNA gene for accurate taxonomic profiling [17]. These taxonomic profiles were generated by searching the non-host metagenomic reads for 16S rRNA gene sequences using Hidden Markov Models [3], and taxonomically classifying these against a SILVA reference database [6]. These taxonomic assignments are stored in an operational taxonomic unit (OTU) table, facilitating the exploration of microbial diversity [5].

The community diversity analysis was conducted in RStudio, focusing on quantifying microbial species diversity and relative abundances through alpha diversity metrics, including Simpson indices. The identified taxa counts are analyzed, calculating a range of diversity measures. The Simpson index varies from 0 to 1, and the index increases as the diversity decreases [11]. Principal Component Analysis (PCA) is performed, and a scree plot is generated to show how much variance each principal component explains [12]. The two-dimensional PCA scatter plot helps visualize the clustering of samples by their microbial community profiles, with samples colored by type for clearer differentiation [9]. A heatmap is generated by running RStudio which provides a visual representation of sample-to-sample distances and highlights significant clusters. This hierarchical clustering technique groups similar samples together based on the abundance of taxa, enabling deeper insights into microbial community structure.

DESeq2 is used to identify taxa that show significant changes in abundance between sample groups. By normalizing the taxa counts, DESeq2 handles variability and identifies differentially abundant OTUs, ensuring statistical reliability [18]. Significant taxa are identified by applying a p-value cutoff, with DESeq2 adjusting these p-values to control the false discovery rate using the Bonferroni-Hochberg correction [13]. Finally, the most abundant taxa are identified from the heatmap, providing data on which microbial taxa are most relevant to the conditions being studied. However, while this heatmap only scratches the surface of understanding the complexity within the community. To gain deeper insights into the functional roles of these microbes, a genome-centric analysis is essential. By examining the actual metabolic pathways and genes expressed, we can move beyond mere abundance and begin to unravel the functional annotation of these organisms across environmental pressures.

In this genome-centric analysis, the first step involves performing a quality assessment of the metagenome-assembled genomes (MAGs) using CheckM. By running CheckM, the completeness and contamination levels of each MAG are evaluated to ensure data quality [4]. Afterward, taxonomy is assigned to each bin using GTDB-Tk, which compares the recovered bins to reference genomes and provides lineage classifications [1]. Next, the functional profiling of the MAGs is carried out. Prodigal is used to identify genes [7], while DRAM annotates these genes with metabolic functions and pathways [16]. The annotated results are explored using files like all_bins_dram_distill.xlsx and all_bins_dram_product.tsv, which provide a detailed view of the genes and pathways present in each MAG, including key functions related to plastic degradation. By examining these annotations, I have identified genes involved in methanogenesis or glycolysis and trace them back to specific MAGs using methods in gene-centric analysis. To evaluate the functional profiling of the superworm MAGs, all_bins_dram_distill.xlsx offers an overview of the metabolic functions annotated in each MAG, including gene copy numbers (listed as KEGG, CAZy, etc. IDs) and their associated pathways/modules.

Results

Figure 1:

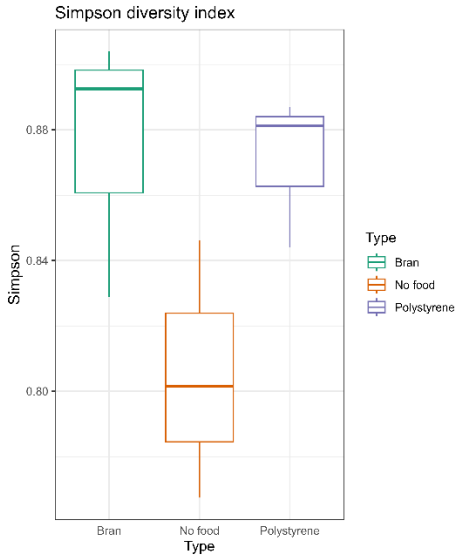


Figure 1 shows the Simpson diversity index boxplot for three different sample types. While the overall diversity is high (per the Shannon index), the dominance of a few species reduces the evenness of the community, which is reflected by the higher Simpson index. Interestingly, the Simpson index is lowest (~0.80) for the "No food" condition, suggesting greater evenness. Although fewer species are present, they are more evenly distributed across the community. For PS, The Simpson index (~0.85) also shows moderate evenness. There is a mix of species with some dominance, but not as pronounced as in the "Bran" condition.

Figure 2:

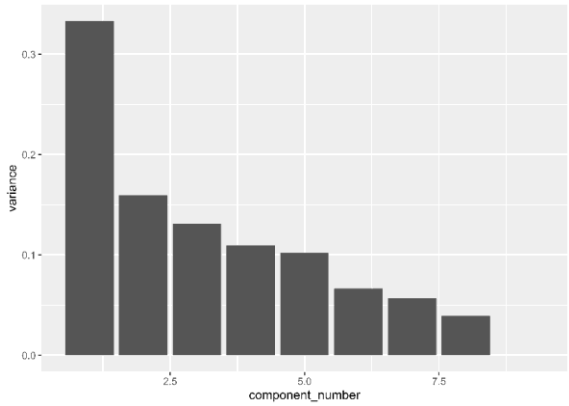


Figure 2 is a scree plot that show the first principal component explains approximately 33% of the variance, while the second explains approximately 16%. Together, the first two components explain about 50% of the total variance. The large difference between the first and second principal components in the scree plot suggests highlights that much of the dataset's complexity can be summarized by the first principal component alone.

Figure 3:

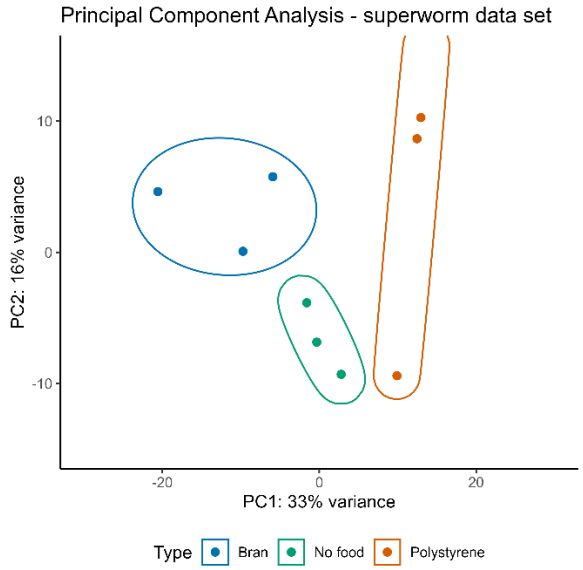


Figure 3 is the PCA plot from the position of each sample in terms of PC1 and PC2. The three groups cluster separately, indicating distinct microbial community compositions between the different feeding conditions. Indeed, one of the samples from the Polystyrene (PS) group is an outlier. This could indicate that the microbial community composition in this particular sample differs significantly from the others in the same group. The outlier might be due to experimental variability, contamination, differences in sequencing depth, or unique biological responses within the superworm gut when exposed to polystyrene.

Figure 4:

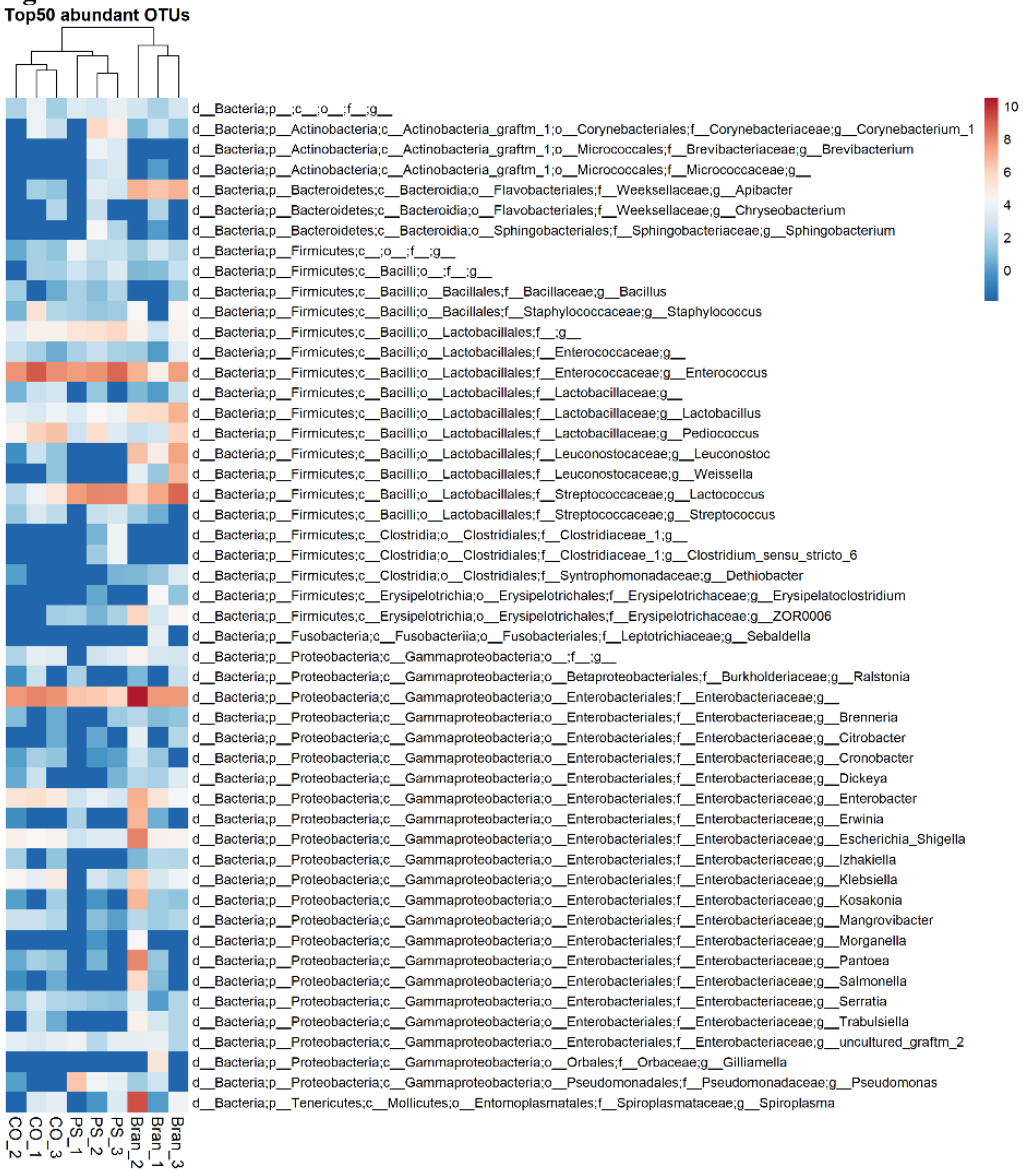


Figure 4 is a heatmap that represents the top 50 most abundant OTUs across Bran, No food, and PS. Each square's color intensity indicates the abundance level of a particular OTU in a sample. Clear patterns of microbial composition differences across the sample types can be observed. Notably, the Bran samples show a high abundance of *Enterobacteriaceae* and *Spiroplasma*, suggesting that nutrient-rich environments strongly influence microbial dominance. Meanwhile, the consistent presence of *Enterococcus* and *Enterobacteriaceae* across all samples suggests these microbes are highly adaptable, thriving from nutrient-rich Bran to synthetic PS and even starvation. Their broad metabolic capabilities enable them to efficiently exploit various carbon sources, whether it's digesting bran or breaking down components of polystyrene. This adaptability showcases their resilience under different feeding conditions, positioning them as key players in the superworm gut ecosystem.

Figure 5: Expression of significant KO terms between prokaryotic communities in the Bran and No food conditions

Figure 5 compares the differential expression of key metabolic genes between prokaryotic communities in the Bran and No food conditions following genome-centric analysis. In the Bran condition, genes involved in chemotaxis, iron transport, and secretion systems are more abundant, reflecting a highly metabolically active community thriving on nutrient availability. Conversely, the No food condition sees an upregulation of genes associated with nutrient scavenging, particularly carbohydrate and amino acid transport systems, indicating a stress response where the community maximizes resource utilization in an environment deprived of nutrients.

Figure 6: Expression of significant KO terms between prokaryotic communities in the Bran and PS conditions

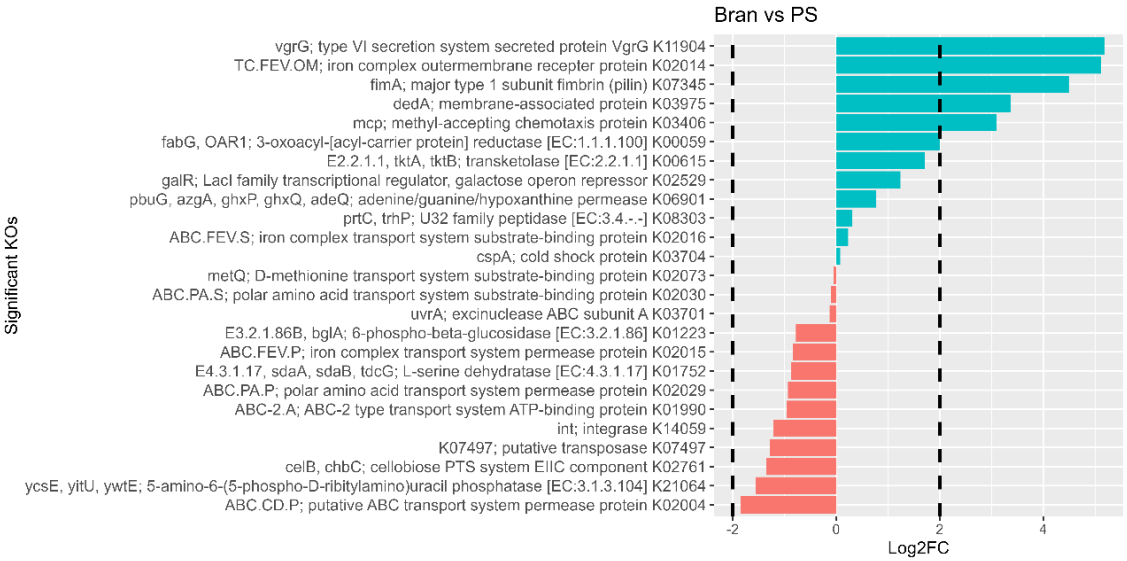
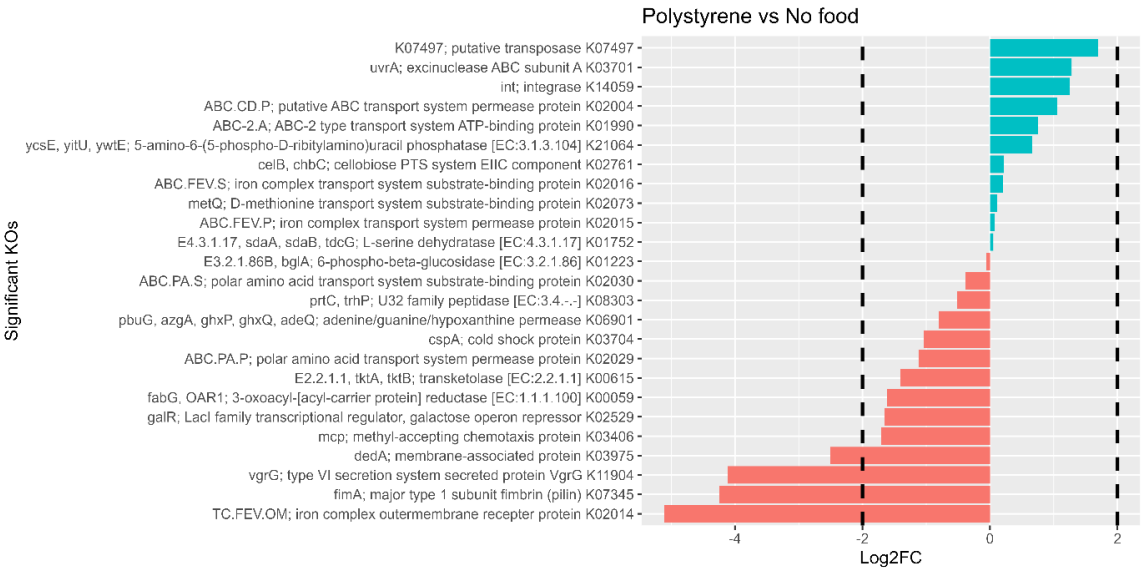


Figure 6 continues the narrative by showcasing the differential expression of significant KEGG Orthology (KO) terms between prokaryotic communities in the Bran and Polystyrene (PS) conditions. The Bran condition fosters an upregulation of genes related to secretion systems, iron transport, and chemotaxis, including vgrG (type VI secretion system), fimA (fimbrial protein), and mcp (chemotaxis protein), all of which play critical roles in nutrient acquisition. In contrast, the PS condition shows a distinct upregulation of genes involved in carbohydrate transport and stress responses, such as celB (cellobiose PTS system) and bglA (beta-glucosidase), suggesting a microbial adaptation to the more limited nutrient availability in the PS environment.

Figure 7: Expression of significant KO terms between prokaryotic communities in the PS and No food conditions.



Building on this, Figure 7 compares the expression of significant KO terms between prokaryotic communities in the Polystyrene (PS) and No food conditions. In the PS condition, genes linked to transposase activity (K07497), excinuclease subunits (urvA), and ABC transport systems are upregulated, aiding in nutrient acquisition and adaptation to the PS substrate. Meanwhile, in the No food condition, the upregulation of genes involved in complex secretion systems (vgrG), chemotaxis (mcp), and fimbrial proteins (fimA) reflects a more aggressive strategy of nutrient-seeking behavior as the community strives for survival in an environment with no external nutrient input.

Discussion

My results revealed the Bran condition is dominated by key taxa (figure 4), which is consistent with the findings of Sun et al. (2022), who showed that nutrient-rich environments foster microbial specialization. On the other hand, the No food condition led to a reduction in diversity, with microbes upregulating scavenging mechanisms to survive nutrient deprivation, showcasing an adaptive stress response (Figures 1 and 5). The Polystyrene (PS) condition promoted moderate microbial activity, favoring genes involved in plastic degradation (Figure 6), reinforcing Sun et al.'s conclusion that microbes play a role in polymer breakdown. Both my study and Sun et al.'s research highlighted a loss of microbial diversity and signs of dysbiosis in the PS condition. Sun et al. reported an increase in opportunistic pathogens like *Enterococcus* and *Klebsiella oxytoca*, indicating gut dysbiosis in superworms. Similarly, I found that the PS condition led to reduced microbial diversity, with an upregulation of stress-response genes (Figure 7). While both studies observed similar stress responses in the No food condition, I did not observe a strong presence of *K. oxytoca* in my findings, suggesting that microbial dynamics under nutrient deprivation might differ between these studies.

The Scree plot in Figure 2 revealed that PC1 captured the majority of the variance in microbial composition. Similarly, Jiang et al.'s study of *P. brevitarsis* grubs showed major metabolic shifts due to PS ingestion, where PC1 explained most of the variance. These findings suggest that both PS and nutrient deprivation drive dominant changes in microbial and metabolic profiles across different species. My Figure 3 shows that the Polystyrene (PS) condition exhibited more dispersed clustering, indicating that PS induced diverse microbial responses. In contrast, Sun et al.'s PCA plot showed a more distinct separation between PS and other diets, suggesting a strong microbial specialization for plastic degradation. The clustering patterns in both studies emphasize that extreme conditions like PS and No food profoundly shape microbial ecology and drive functional differentiation. While Sun et al. noted that PS degradation was accompanied by gut dysbiosis and minimal weight gain in superworms, I found upregulated genes involved in plastic degradation pathways, particularly in carbohydrate transport and breakdown (Figures 6 and 7).

When comparing my findings to Di Gregorio et al.'s study of marine copepods, I observed that both studies revealed how microbial communities adapt to environmental stress. Proteobacteria dominated across conditions in both studies (Figure 4). While Di Gregorio et al. focused on hydrolytic enzymes involved in chitin and polyester degradation, I found that stress-related genes were upregulated in response to PS ingestion (Figures 6 and 7). In both cases, nutrient scarcity or the presence of recalcitrant polymers, like PS, drove microbial shifts toward degradation and nutrient-scavenging pathways. In my study, microbial diversity decreased under the No food condition (Figure 1), reflecting dysbiosis similar to that observed by Jiang et al. They reported enrichment of specific taxa, such as *Promicromonosporaceae* and *Bacillaceae*, in PS-fed grubs, which is similar to the *Enterobacteriaceae* enrichment I observed in the PS samples (Figure 4). Both studies suggest that specific microbial taxa adapt to PS ingestion by activating degradation-related pathways. Key metabolic pathways, such as nicotinate and glutamate metabolism (Jiang et al.), and energy-related genes (Figures 6 and 7 in my study), were upregulated in response to the oxidative stress and metabolic demands induced by PS degradation.

Reflecting on the limitations of my study, I recognize several areas for improvement. Both my study and Sun et al. highlighted similar methodological constraints. Sun et al. called for transcriptomics and click chemistry to identify active microbial members during PS degradation, while I believe the need for metatranscriptomics, or proteomics is clear. These approaches would provide critical insights into dynamic gene expression under environmental stress, which is essential for understanding the real-time activity of microbial communities. Additionally, our reliance on 16S rRNA sequencing restricts the understanding of functional diversity. By employing more comprehensive methodologies, such as deeper metagenomic sequencing with improved assembly tools, I could produce more complete metagenome-assembled genomes (MAGs). This would allow me to better link microbial taxa to specific metabolic functions, especially under stress conditions like PS ingestion. A multi-omics approach combining metabolomics, transcriptomics, and proteomics could provide a more holistic view of microbial adaptations and functional shifts, particularly in complex environmental conditions.

Broader Applications and Conclusion

My research proposes fundamental insights into how microbial communities within the superworm gut adapt to the biodegradation of polystyrene (PS). The identification of microbial taxa that thrive in both nutrient-rich and nutrient-deprived environments, and the upregulation of genes linked to plastic degradation suggests a potential for harnessing these microbes in industrial plastic degradation processes. The resilience and functional versatility of microbial communities observed in the PS conditions offer promising avenues for developing efficient microbial-based recycling technologies. Specifically, the enrichment of genes involved in carbohydrate transport and stress responses in the PS condition highlights the potential for isolating specific enzymes that could break down synthetic polymers more efficiently. This could lead to breakthroughs in reducing plastic waste, one of the most pressing environmental challenges of our time. The limitations identified in my research, particularly the reliance on 16S rRNA sequencing underscore the need for future studies to incorporate metagenomics, metatranscriptomics, and proteomics to capture dynamic gene expression and metabolic activity. In conclusion, my research lays the groundwork for future studies aimed at utilizing microbial communities for biotechnological innovations in plastic waste degradation. My study not only contributes to the growing body of knowledge on microbial adaptation but also opens doors for practical applications that could have a lasting impact on the future of plastic recycling and environmental conservation.

Reference List

1. Chaumeil, P.-A. *et al.* (2019) 'GTDB-TK: A toolkit to classify genomes with the Genome Taxonomy Database', *Bioinformatics*, 36(6), pp. 1925–1927. doi:10.1093/bioinformatics/btz848.

2. Di Gregorio, S. *et al.* (2024) 'Marine copepod culture as a potential source of bioplastic-degrading microbiome: The case of poly(butylene succinate-co-adipate)', *Chemosphere*, 362, p. 142603. doi:10.1016/j.chemosphere.2024.142603.

3. Ghosh, T.S. *et al.* (2012) 'C16s — a hidden Markov model-based algorithm for taxonomic classification of 16S rrna gene sequences', *Genomics*, 99(4), pp. 195–201. doi:10.1016/j.ygeno.2012.01.008.

4. Haryono, M.A. *et al.* (2022) 'Recovery of high-quality metagenome-assembled genomes from full-scale activated sludge microbial communities in a tropical climate using longitudinal metagenome sampling', *Frontiers in Microbiology*, 13. doi:10.3389/fmicb.2022.869135.

5. He, Y. *et al.* (2015) 'Stability of operational taxonomic units: An important but neglected property for analyzing microbial diversity', *Microbiome*, 3(1). doi:10.1186/s40168-015-0081-x.

6. Henderson, G. *et al.* (2019) 'Improved taxonomic assignment of rumen bacterial 16S rrna sequences using a revised Silva taxonomic framework', *PeerJ*, 7. doi:10.7717/peerj.6496.

7. Hyatt, D. *et al.* (2010) 'Prodigal: Prokaryotic gene recognition and translation initiation site identification', *BMC Bioinformatics*, 11(1). doi:10.1186/1471-2105-11-119.

8. Jang, Y. *et al.* (2024) 'Identification of plastic-degrading bacteria in the human gut', *Science of The Total Environment*, 929, p. 172775. doi:10.1016/j.scitotenv.2024.172775.

9. Jian Yang *et al.* (2004) 'Two-dimensional PCA: A new approach to appearance-based face representation and recognition', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(1), pp. 131–137. doi:10.1109/tpami.2004.1261097.

10. Jiang, J. *et al.* (2024) 'Soil-dwelling grub larvae of protaetia brevitarsis biodegrade polystyrene: Responses of gut microbiome and host metabolism', *Science of The Total Environment*, 934, p. 173399. doi:10.1016/j.scitotenv.2024.173399.

11. Kers, J.G. and Saccenti, E. (2022) 'The power of microbiome studies: Some considerations on which alpha and beta metrics to use and how to report results', *Frontiers in Microbiology*, 12. doi:10.3389/fmicb.2021.796025.

12. Ledesma, R.D., Valero-Mora, P. and Macbeth, G. (2015) 'The Scree Test and the number of factors: A Dynamic Graphics Approach', *The Spanish Journal of Psychology*, 18. doi:10.1017/sjp.2015.13.

13. Olejnik, S. *et al.* (1997) 'Multiple testing and statistical power with modified Bonferroni procedures', *Journal of Educational and Behavioral Statistics*, 22(4), p. 389. doi:10.2307/1165229.

14. Peng, H. *et al.* (2024) *Unveiling gut microbiota and metabolic functions contributed to polyvinyl chloride degradation in Spodoptera frugiperda larvae* [Preprint]. doi:10.21203/rs.3.rs-4686713/v1.

15. Satta, A. *et al.* (2024) 'Synergistic functional activity of a landfill microbial consortium in a microplastic-enriched environment', *Science of The Total Environment*, 947, p. 174696. doi:10.1016/j.scitotenv.2024.174696.

16. Shaffer, M. *et al.* (2020) 'Dram for distilling microbial metabolism to automate the curation of microbiome function', *Nucleic Acids Research*, 48(16), pp. 8883–8900. doi:10.1093/nar/gkaa621.

17. Tringe, S.G. and Hugenholtz, P. (2008) 'A Renaissance for the pioneering 16S rrna gene', *Current Opinion in Microbiology*, 11(5), pp. 442–446. doi:10.1016/j.mib.2008.09.011.

18. Weiss, S. *et al.* (2017) 'Normalization and microbial differential abundance strategies depend upon data characteristics', *Microbiome*, 5(1). doi:10.1186/s40168-017-0237-y.