NREL
NATIONAL RENEWABLE ENERGY LABORATORY

U.S. DEPARTMENT OF ENERGY
Office of Science
Early Career Research Program

Cross-kingdom Comparative Genomics of Aromatic Catabolic Pathways in Fungi and Bacteria

Teeratas Kijpornyongpan and **Davinia Salvachúa*** (davinia.salvachua@nrel.gov)
Renewable Resources and Enabling Sciences Center, National Renewable Energy Laboratory, Golden, CO

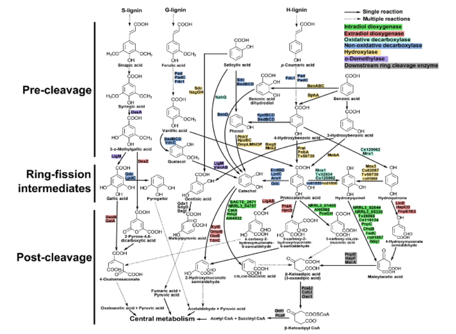


Figure 1 Catabolic pathways of lignin-derived monoaromatic compounds. The catabolic pathways can be divided in three stages: pre-cleavage stage to modify lignin monomers into ring-fission intermediates, ring-cleavage stage to cleave aromatic intermediates, and post-cleavage stage to funnel cleaved products into central metabolism. Previously characterized enzymes from a variety of bacteria and fungi are included next to the arrows in each enzymatic step.

Project Goals

The goal of this project is to test the hypothesis that white-rot fungi can simultaneously depolymerize lignin extracellularly and catabolize depolymerization products intracellularly as carbon and energy sources. The results from this project will lead to improved understanding of lignin utilization by white-rot fungi, and enable identification of promising fungal strains for lignin catabolism and valorization. As part of this effort, we have conducted a comparative genomic study to investigate a diversity of aromatic catabolic pathways in the fungal and bacterial kingdoms. Furthermore, we have combined genomic and phylogenetic approaches to decipher the evolution of certain aromatic catabolic enzymes and discover new classes of enzymes, which represents a foundation for future biochemical and molecular genetic studies.

Background

Although lignin is the second most abundant polymer in plant biomass, its upgrading remains a major hurdle in biorefining due to its recalcitrance and structural heterogeneity. To overcome this challenge, the concept of biological funneling has emerged as an approach to convert heterogeneous mixtures of lignin-derived monomeric aromatic compounds – from various lignin deconstruction processes – to single chemicals by using engineered microbes. For this purpose, knowledge in aromatic catabolic pathways is necessary to design microbes with (1) enhanced substrate diversity utilization, (2) conversion efficiency, and (3) tailored metabolic pathways to produce the desired products (Johnson et al., 2019). Aromatic catabolic pathways have been thoroughly described in several bacteria such as *Pseudomonas putida*, *Burkholderia* sp. SJ98, *Sphingobium* sp. SYK-6, and *Rhodococcus jostii*. However, knowledge of the corresponding aromatic catabolic pathways in white-rot fungi (WRF) is quite limited, even though WRF are known to be the most efficient lignin-degrading organisms in nature (del Cerro et al., 2021).

Results

We examined the protein domains in functionally characterized aromatic catabolic enzymes through InterProScan (**Figure 1**), and we selected protein domains with specific descriptions to the catabolism of aromatic compounds as proxy domains to define a potential capability to modify and/or catabolize aromatic compounds. Then, we performed genome-wide protein domain searches and gene orthology assessment across 255 bacterial genomes and 317 fungal genomes, representing different lineages and nutritional modes across the trees of life in these two kingdoms. Based on the distribution of enzyme domains, we found that fungi have a higher conservation of genes that encode for intradiol dioxygenases, phenol hydroxylases, and 3-oxoacid CoA transferases – compared to bacteria (**Figure 2**). Our association analyses indicated that microbes that are either known to utilize aromatic compounds as a carbon source or classified as plant-associated also have a higher abundance of aromatic catabolic enzymes compared to the ones that are unable to utilize aromatic compounds as a carbon source or known as animal-associated (**Figure 3**). Apart from the genome-wide assessment, we performed phylogenetic analyses to understand the evolution of specific aromatic catabolic enzymes. Our findings highlighted the evolution of intradiol dioxygenases (**Figure 4**): specifically, we found separate origins of catechol dioxygenases and 3,4-protocatechuate dioxygenases in bacteria and fungi. The gene tree phylogeny suggested that hydroxyquinol dioxygenases are the most common type of intradiol dioxygenases found in fungi (**Figure 5**). We also identified a novel class of intradiol dioxygenases, which likely function extracellularly based on the presence of signal peptides.

Based on gene distribution of dioxygenases and phenol hydroxylases (IPR000627 and IPR012941), the ring cleavage and hydroxylation of aromatic compounds is limited to few groups of bacteria, but these reactions are common in fungi

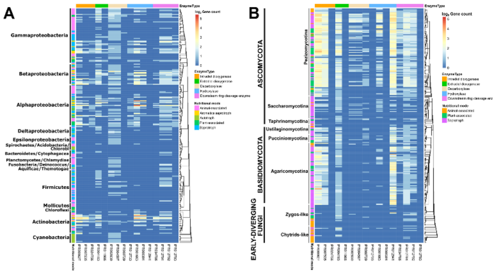


Figure 2 Gene distribution of aromatic catabolic enzymes in (A) Bacteria and (B) Fungi. Each bacterial and fungal genome was subject for genome-wide IPR domain searching. The IPR domains with specific description for modification of aromatic compounds were selected for constructing the heatmap. Color in each cell indicates a log2 abundance of annotated protein-coding genes in each genome. Rows indicate bacterial/fungal genomes. Columns indicate IPR domain identifiers. Color codes for nutritional modes and enzyme types are indicated in the heatmaps. Text in the left side of each graph indicates bacterial and fungal lineages.

Presence and absence patterns of aromatic catabolic enzymes suggest conserved catabolism of aromatic compounds in fungi, including white-rot fungi

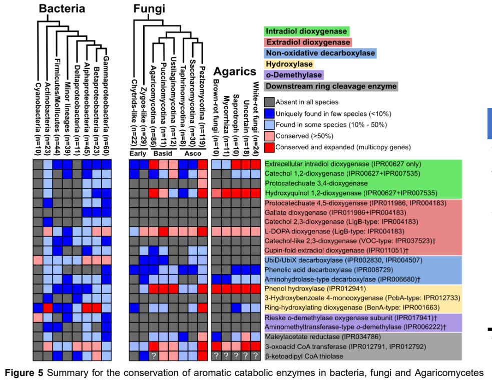
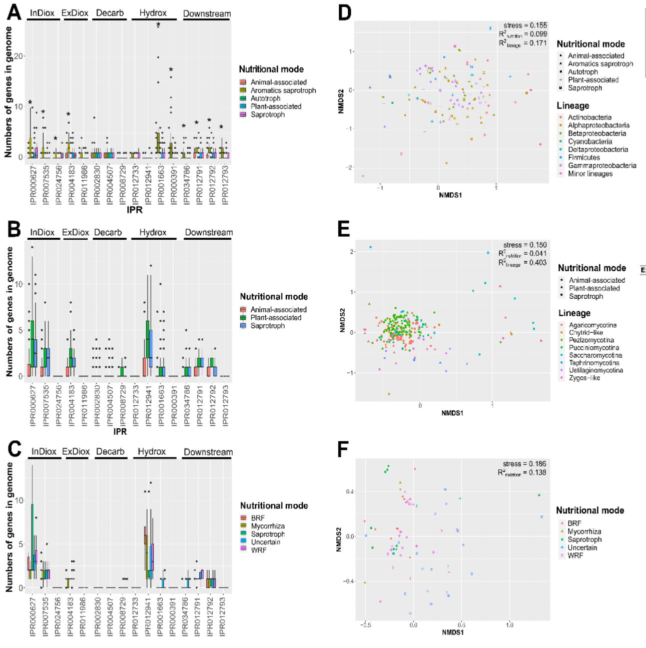
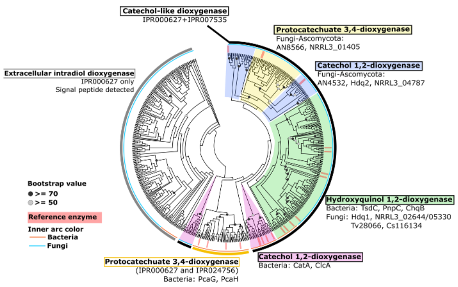


Figure 5 Summary for the conservation of aromatic catabolic enzymes in bacteria, fungi and Agaricomycetes (Agarics) fungi. Colors in cells indicate the levels of conservation. 'n' indicates a number of species/genomes included in the analyses for each lineage.



Aromatic catabolic enzymes are more abundant in plant-associated than animal-associated bacteria or fungi

Figure 3 Association analyses between aromatic catabolic enzymes and nutritional modes/ lineages. (A, B, C) Boxplots showing distribution of gene count for aromatic catabolic enzymes, using IPR domains as proxies, for (A) Bacteria, (B) Fungi and (C) Agaricomycetes classified by nutritional modes. Abbreviations: InDiox, Intradiol dioxygenase; ExDiox, Extradiol dioxygenase; Decarb, Decarboxylase; Hydrox, Hydroxylase; Downstream, Downstream ring cleavage enzymes. (D, E, F) Non-metric multidimensional scaling (NMDS) plots showing data ordination of analyzed species for (D) Bacteria, (E) Fungi and (F) Agaricomycetes. Point shapes indicate lineages.



Gene tree phylogeny reveals divergent substrate specificity of catechol-like dioxygenases in fungi and allows the discovery of novel extracellular intradiol dioxygenases

Figure 4 Phylogenetic tree of intradiol dioxygenases in bacteria and fungi. Representative protein sequences containing IPR000627 (domain for intradiol dioxygenase) were selected for gene tree reconstruction through the maximum-likelihood (RAxML) method. Clades showing different enzyme types and substrate specificity are indicated by outer arcs and shading colors. Substrate specificity is inferred by the presence of reference enzymes in each clade.

Future work and conclusions

Taken together, this work provides new insights related to the catabolism of aromatic compounds in different groups of bacteria and fungi which could enable the discovery of novel microbes, pathways, and enzymes to improve the 'funnel' to convert lignin-derived aromatic compounds to value-added products. We show that fungi are also potential sources of aromatic catabolic enzymes that can be utilized in lignin valorization. Future validation and characterization of these enzymes will lead to better understanding of catabolism of lignin-derived aromatic compounds in the fungal Kingdom and lignin-degrading white-rot fungi.

Reference

del Cerro, C., Erickson, E., Dong, T., Wong, A.R., Eder, E.K., Purvine, S.O., Mitchell, H.D., Weitz, K.K., Markillie, L.M., Burnet, M.C., et al. (2021). Intracellular pathways for lignin catabolism in white-rot fungi. Proceedings of the National Academy of Sciences 118, e2017381118.

Johnson, C.W., Salvachúa, D., Rorrer, N.A., Black, B.A., Vardon, D.R., St. John, P.C., Cleveland, N.S., Dominick, G., Elmore, J.R., Grundl, N., et al. (2019). Innovative Chemicals and Materials from Bacterial Aromatic Catabolic Pathways. Joule 3, 1523-1537.

This research is supported by the U.S. Department of Energy (DOE) Office of Science, Office of Biological and Environmental Research under the Early Career Award Program. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

NREL/PO-2800-81850

Conference Title: Genomic Sciences Program Annual Principal Investigator meeting
Conference Location – Virtual
Conference Dates- 28 February – March 2022