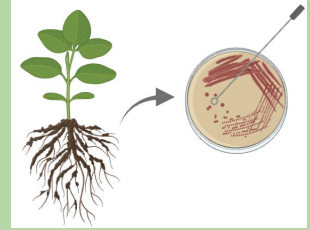# ROOTOMICS DATABASE

ANURADHA BASYAL, ADITHI NATARAAJ, SHIVANI PIMPARKAR

FACULTY ADVISOR : JENNIFER BHATNAGAR
Associate Professor of Biology, Boston university

INTRODUCTION

- Soil microbes drive global biogeochemical cycles

- We know little about their extracellular molecular functions

- RootOmics links gene presence to ecological function

🌱 Root Omics

📘 Introduction   🚀 Explorer   ⬇ Download   📖 Contact   ❓ Help

What is Root Omics?

Root Omics is a functional enzyme database for microbial ecology research. It catalogs soil nutrient-cycling enzymes involved in carbon, nitrogen, and phosphorus transformations, with detailed annotations (EC, KO, GO) and pathway mappings. By integrating genomic and experimental data, it helps researchers assess enzyme distribution, pathway completeness, and functional potential across diverse soil microbial communities.

Developed by Team 13 at Boston University (BF768, Spring 2024) under Dr. G. Benson, Root Omics supports bioinformatics workflows, genomic analyses, and environmental functional gene studies, enabling insights into biogeochemical cycling and microbial ecosystem functions.

MOTIVATION AND BACKGROUND

**KEGG and GO focus mostly on intracellular processes**

Environmental gene function remains under-characterized

No comprehensive tool links genes → ecological role
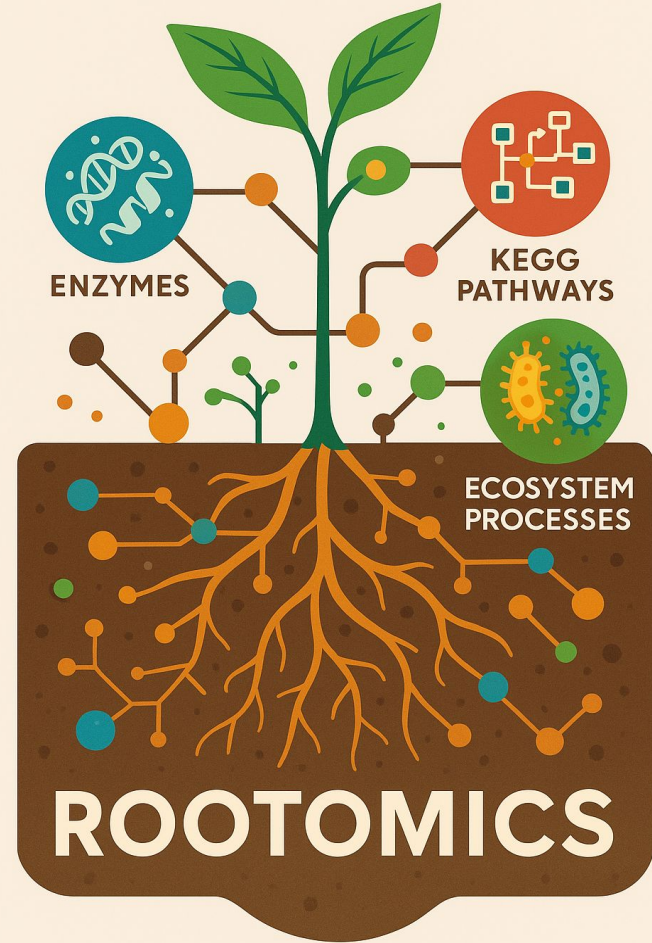
What Can RootOmics Help Us Answer?

1. How do enzymes respond to environmental stressors?

2. What evidence supports a specific enzyme's role in nutrient cycling?

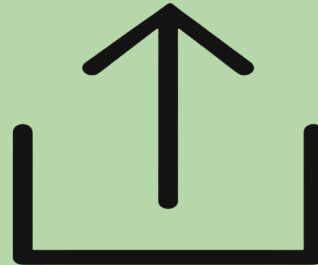3. What are the dominant enzymes driving C, N, P transformations in soil?
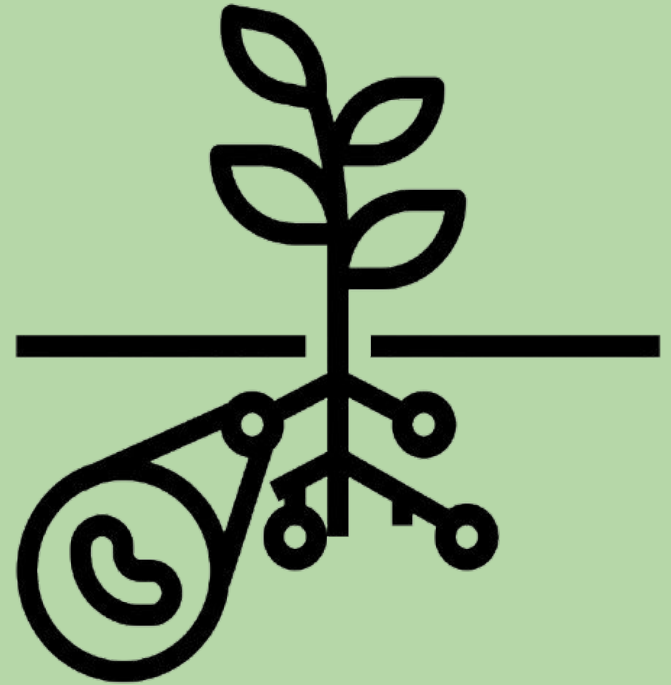
Database Structure & Accessibility

- Open access and searchable—no login required.
- Each entry includes EC number, enzyme name, function, and evidence source.
- Contributor uploads are reviewed to ensure data quality.

Open access

Data Model & ER Diagram

Proteins
PID → Protein_name, Enzyme_Type, EC_numbers, KO_Numbers, GO_terms

Pathways
PAID → KEGG_pathway, CAZyme_class_superfamily

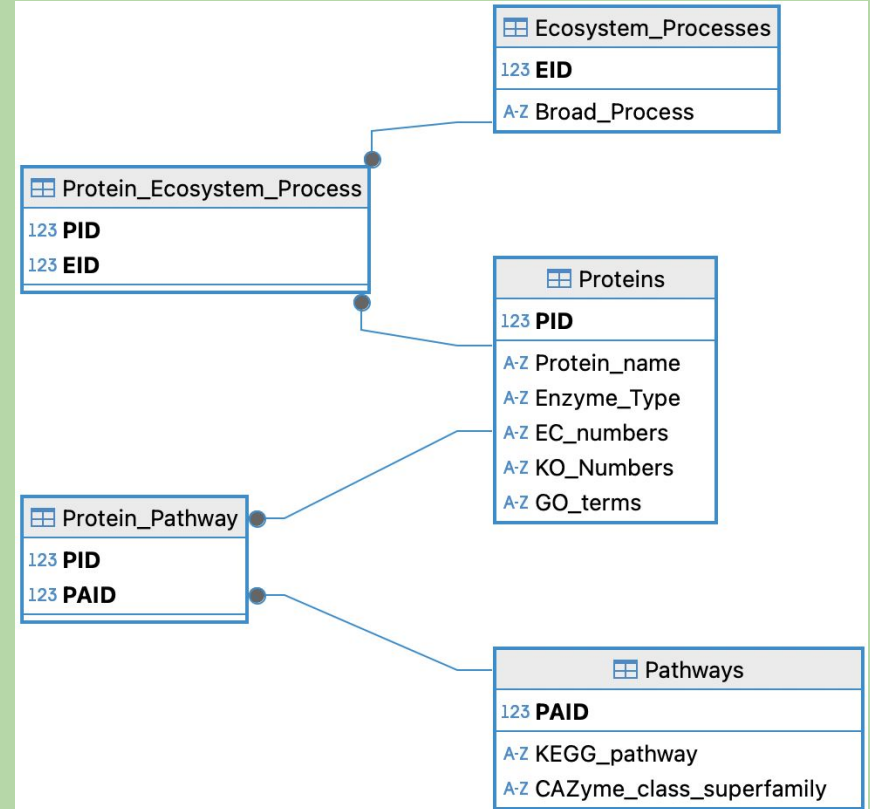**Ecosystem_Processes**
EID → Broad_Process

Links **(many-to-many)**
Protein_Pathway (PID ↔ PAID)
Protein_Ecosystem_Process (PID ↔ EID)

Sample SQL Queries

```sql
-- 1) Search by protein name ("oxidase") and return full annotation
SELECT
  p.Protein_name,
  p.Enzyme_Type,
  p.EC_numbers,
  p.KO_Numbers,
  p.GO_terms,
  GROUP_CONCAT(DISTINCT pw.KEGG_pathway)        AS Pathways,
  GROUP_CONCAT(DISTINCT ep.Broad_Process)       AS Ecosystem_Processes
FROM Proteins p
LEFT JOIN Protein_Pathway pp  ON p.PID = pp.PID
LEFT JOIN Pathways pw         ON pp.PAID = pw.PAID
LEFT JOIN Protein_Ecosystem_Process pep ON p.PID = pep.PID
LEFT JOIN Ecosystem_Processes ep        ON pep.EID = ep.EID
WHERE p.Protein_name LIKE '%oxidase%'
GROUP BY p.PID
LIMIT 100;
```

```sql
-- 3) Combine protein name and KEGG pathway filter
SELECT
  p.Protein_name,
  GROUP_CONCAT(DISTINCT pw.KEGG_pathway) AS Pathways
FROM Proteins p
JOIN Protein_Pathway pp  ON p.PID = pp.PID
JOIN Pathways pw         ON pp.PAID = pw.PAID
WHERE p.Protein_name LIKE '%oxidase%'
  AND pw.KEGG_pathway = '00010'        -- Glycolysis / Gluconeogenesis
GROUP BY p.PID;
```

```sql
-- 2) Filter by EC number (e.g. "3.2.1.14")
SELECT
  p.Protein_name,
  p.Enzyme_Type,
  p.EC_numbers
FROM Proteins p
WHERE p.EC_numbers LIKE '%3.2.1.14%'
LIMIT 100;
```

WEB INTERFACE

🌱 Root Omics

📘 Introduction · 🧬 Explorer · ⬇️ Download · 📖 Contact · ❓ Help

What is Root Omics?

Root Omics is a functional enzyme database for microbial ecology research. It catalogs soil nutrient-cycling enzymes involved in carbon, nitrogen, and phosphorus transformations, with detailed annotations (EC, KO, GO) and pathway mappings. By integrating genomic and experimental data, it helps researchers assess enzyme distribution, pathway completeness, and functional potential across diverse soil microbial communities.

Developed by Team 13 at Boston University (BF768, Spring 2024) under Dr. G. Benson, Root Omics supports bioinformatics workflows, genomic analyses, and environmental functional gene studies, enabling insights into biogeochemical cycling and microbial ecosystem functions.

EXPLORER TAB

🌱 Root Omics

📘 Introduction   🚀 Explorer   ⬇ Download   📫 Contact   ❓ Help

Protein	Enzyme Type	EC #	KO #	GO term	KEGG
e.g. oxidase	oxidase	1.14	e.g. K00844	e.g. GO:0004672	e.g. map00010

Reset Filters

Results

| Protein Name	Enzyme Type	EC Numbers	KO Numbers	GO Terms	KEGG Pathways	Ecosystem Processes
Lignin peroxidase	Peroxidase	1.11.1.14				Soil C cycling; Soil N cycling; Soil N cycling
Cytochrome P450	Polyphenol oxidase	1.14.14.1	K00490; K07408; K07409; K07410; K07411; K07412; K07413; K07414; K07416; K07418; K07420; K07424; K074	https://www.ebi.ac.uk/QuickGO/term/GO:0016682	https://www.kegg.jp/entry/K00490	Soil C cycling; Soil N cycling
Mono-oxygenase	Polyphenol oxidase	1.14.18	K00506		https://www.kegg.jp/entry/K00506	Soil C cycling; Soil N cycling

PATHWAY MAPS
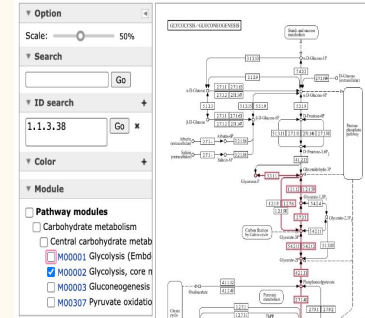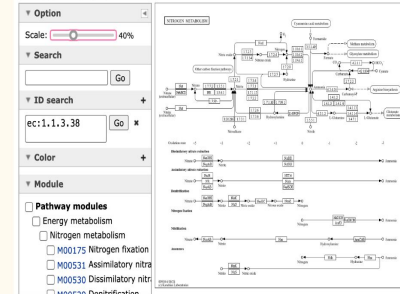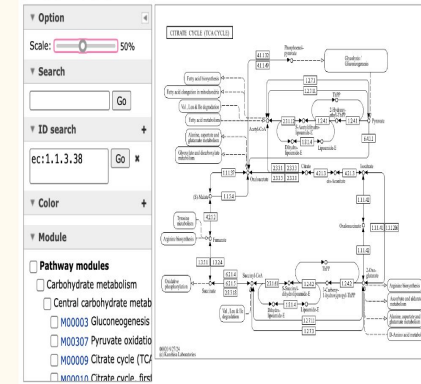
QUERY BASED DATA DOWNLOAD

🌱 Root Omics

📘 Introduction 🧬 Explorer ⬇ Download 📖 Contact ❓ Help

Download Your Data

Export the current view (including any filters) as CSV or Excel:

Download CSV   Download XLSX

CONTACT FORM

HELP CENTER

🌱 Root Omics

📘 Introduction   🧬 Explorer   ⬇ Download   📖 Contact   ❓ Help

❓ Help Center

Quick tips to get the most out of Root Omics:

Filter Data
Enter text in any search field and click `Search`.

Reset Filters
Click `Reset` to clear all fields.

Explore Links
Click any `Pathway` or `GO term` to jump to the external detail page.

Download Data
Head to the `Download` tab to export your filtered results as CSV or Excel.

Need More Help?
Use the `Contact` tab to send us your questions or suggestions.

Thank You!!!