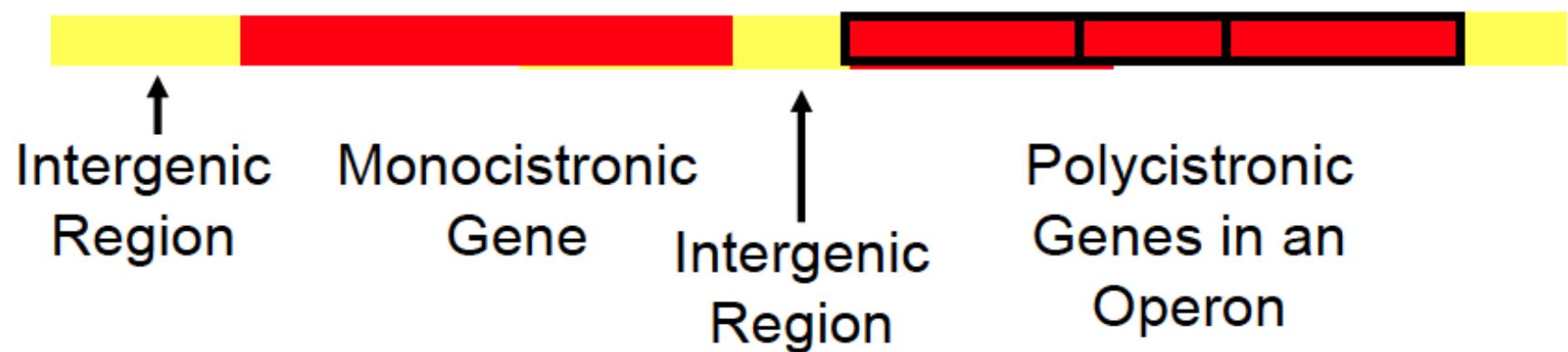


Functional Annotations from Microbial Surveys

Gail Rosen

Genome Structure

Prokaryote

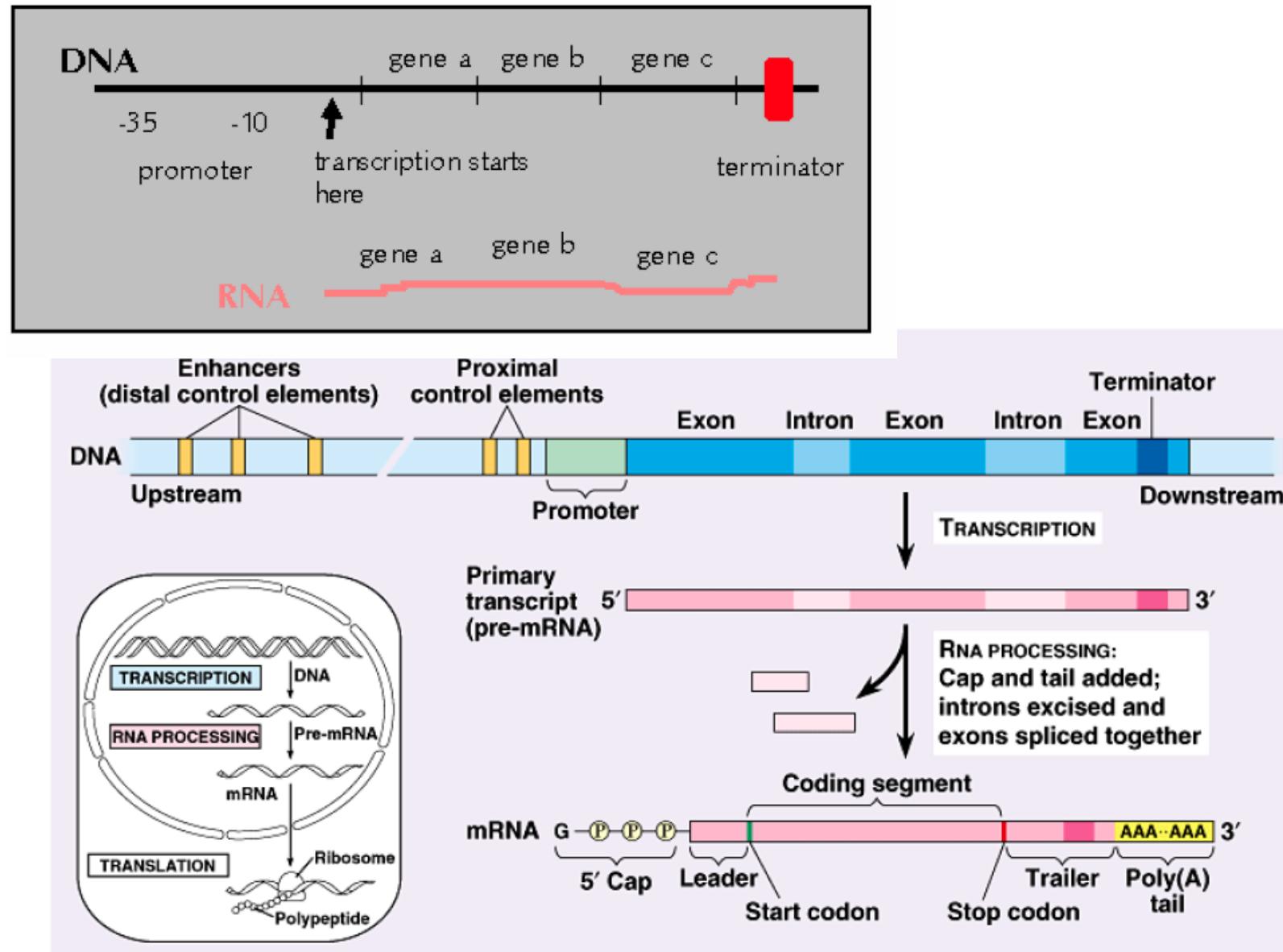


Eukaryote



Transcript “processing”

A 'typical' bacterial operon



Gene Content

- Bacteria/Archaea: ~90% of genome is “coding”
- Eukaryote: ~50% for fungi to 3% for humans!

Functional Assignments

Name

Descriptive common name for the protein, with as much specificity as the evidence supports; gene symbol.

Role

Describe what the protein is doing in the cell and why.

Associated information:

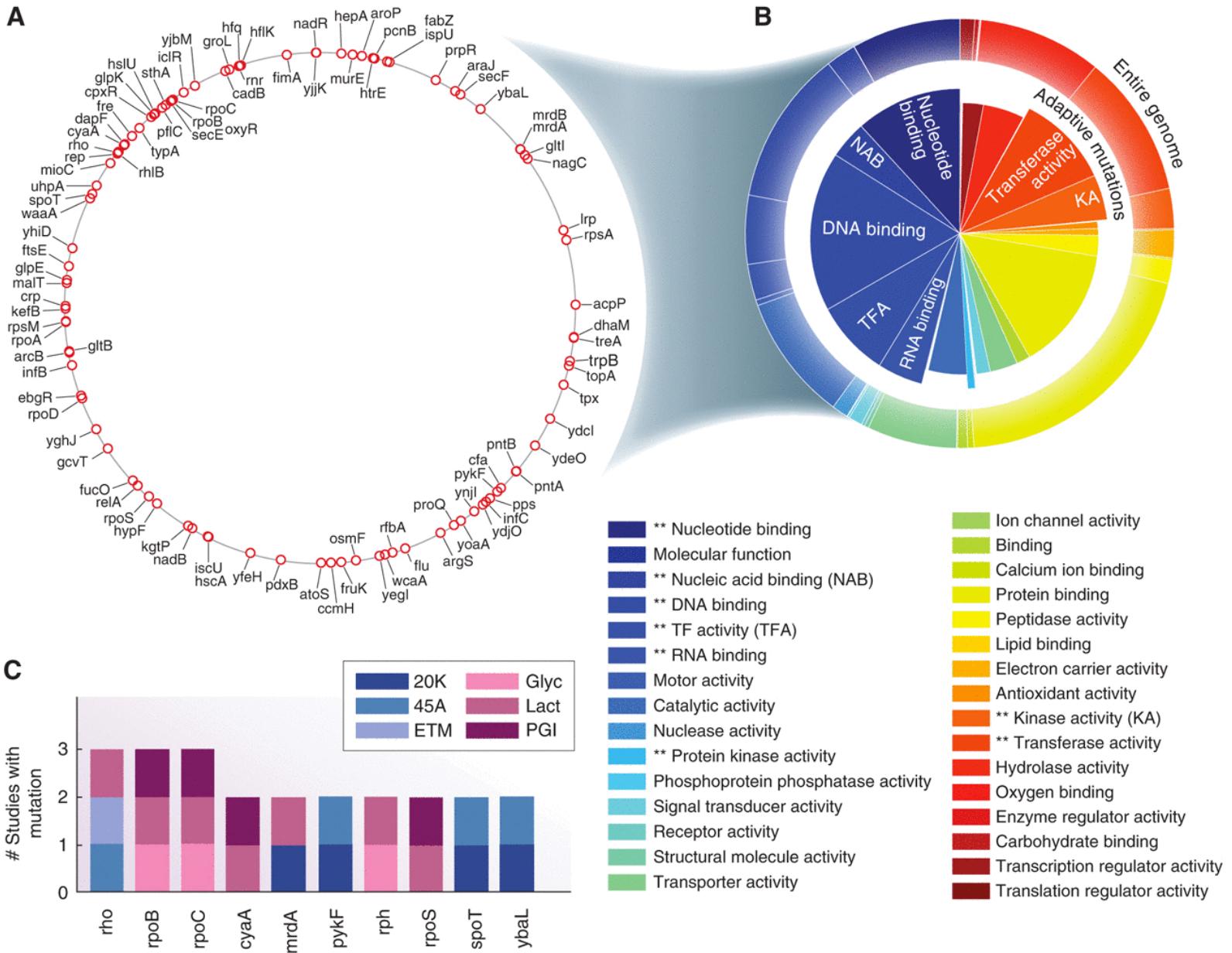
Supporting evidence: Domain and motifs

EC number if protein is an enzyme.

Paralogous family membership.

Metabolic pathways participating in

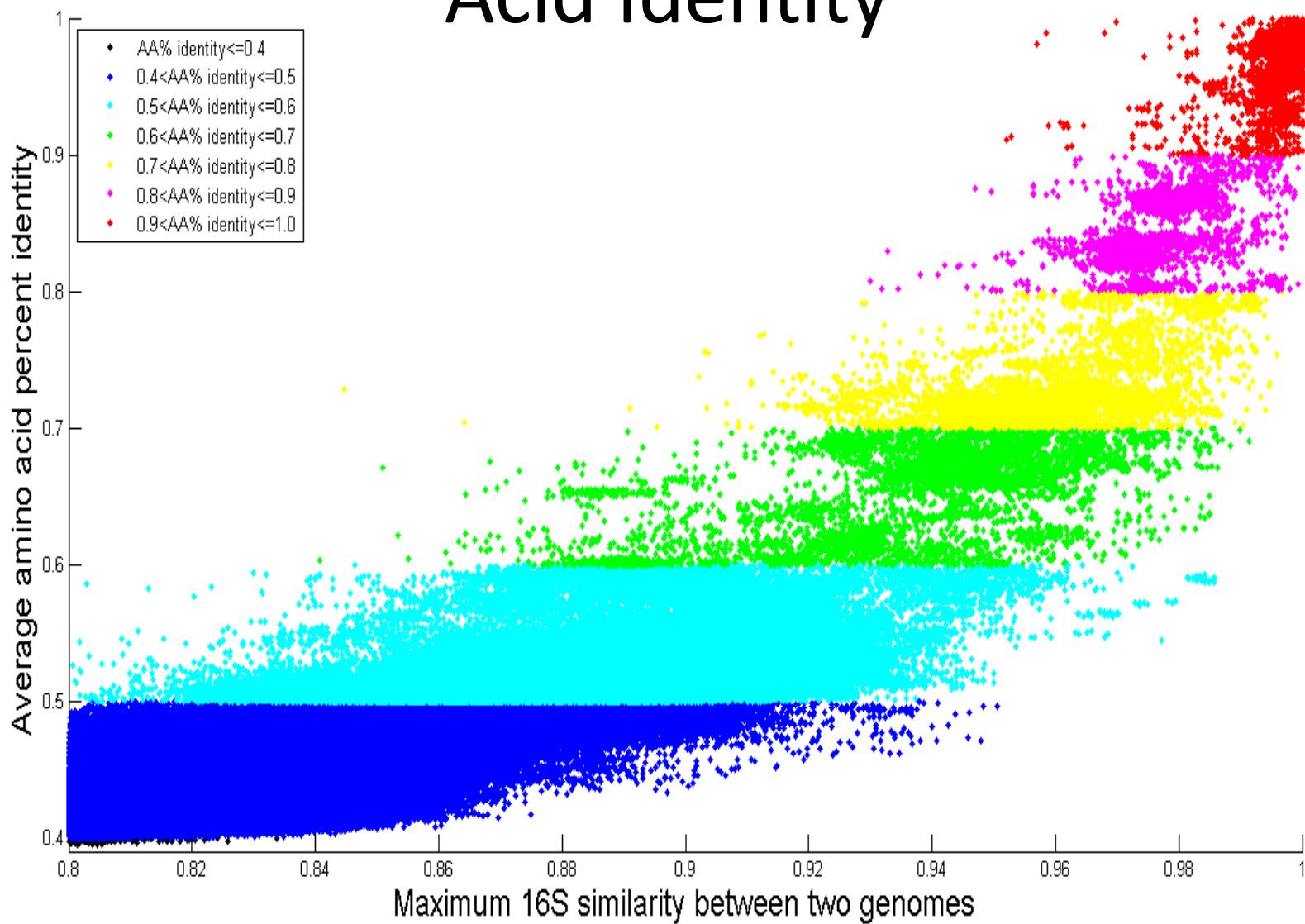
Example E Coli Genome



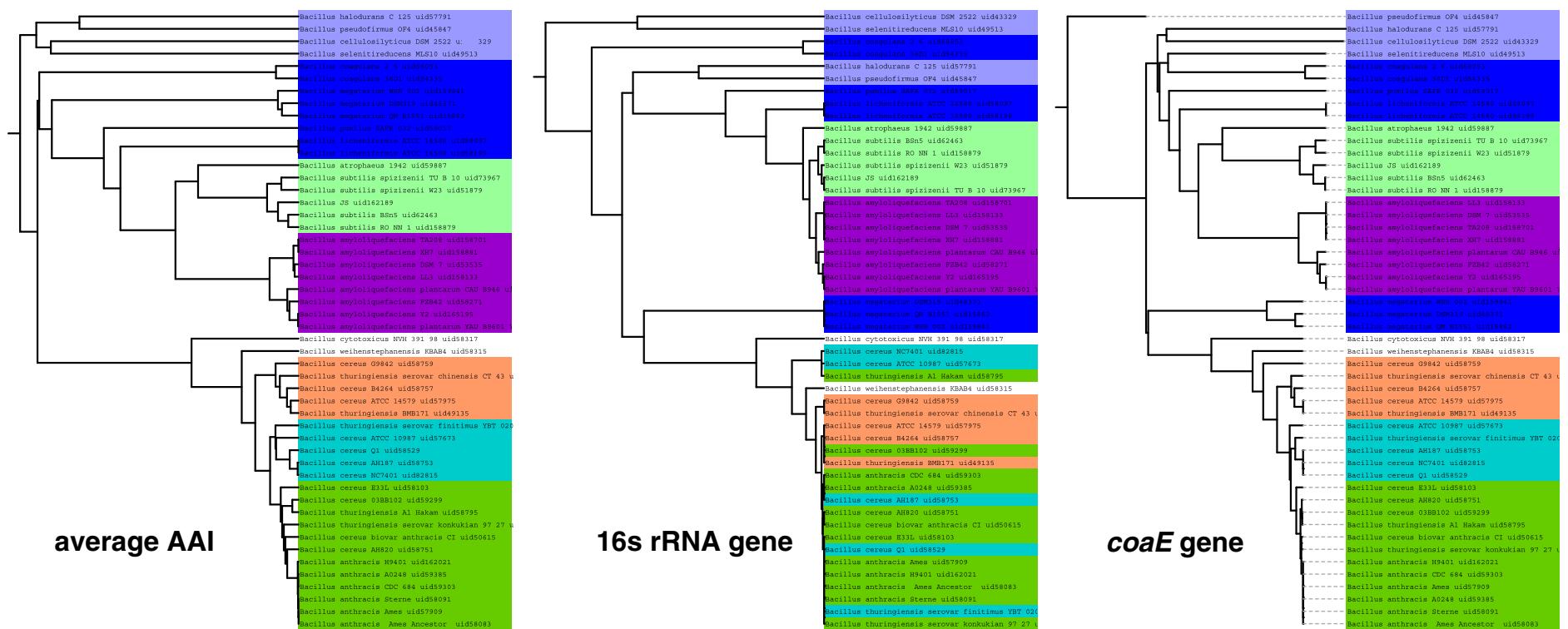
Why so interested in “Who is there”?

- 16S is Cheap
- If know organism, know something about what it does, right?

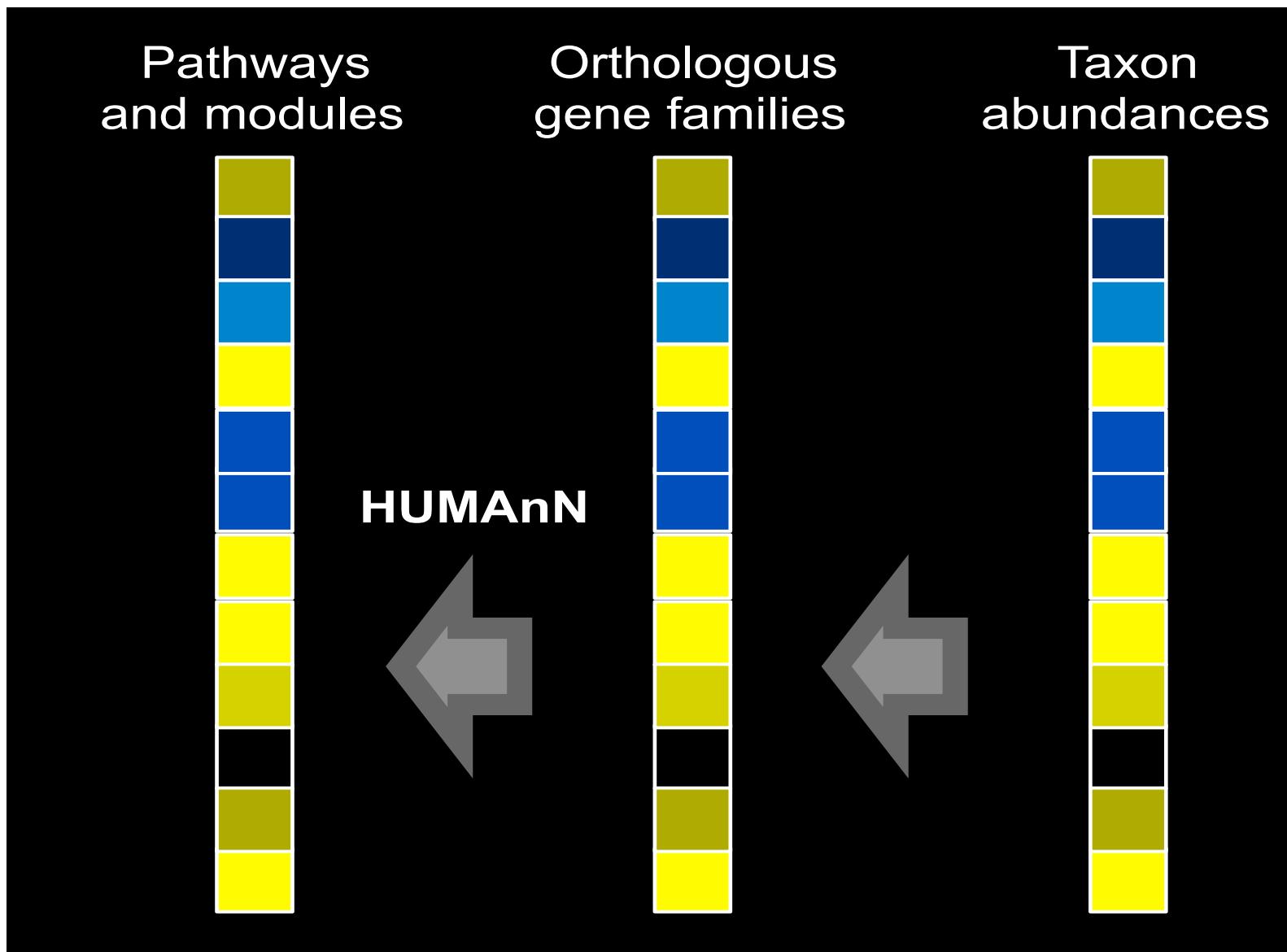
Relationship between 16S and Amino Acid Identity



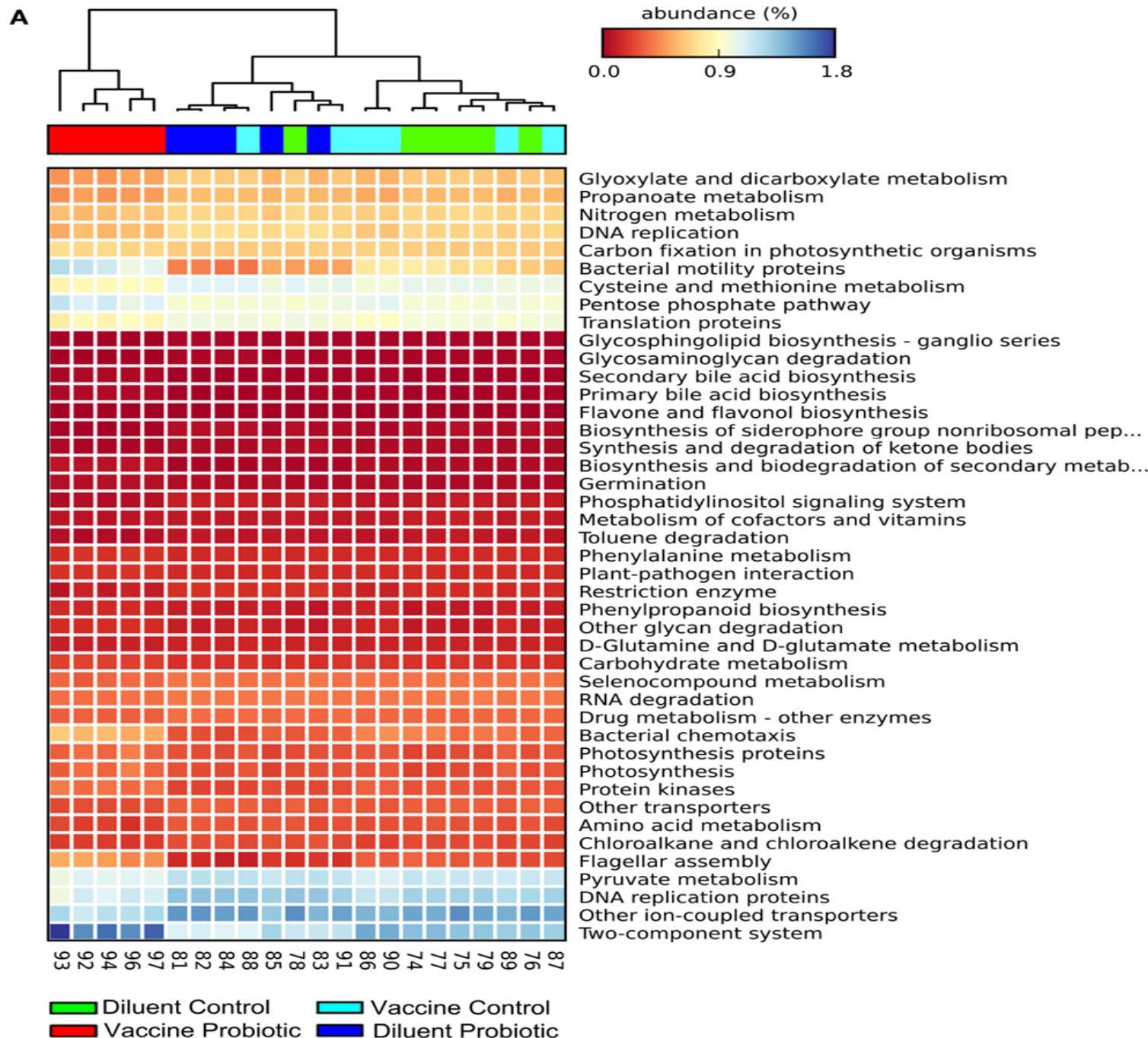
Relationship between 16S and Amino Acid Identity



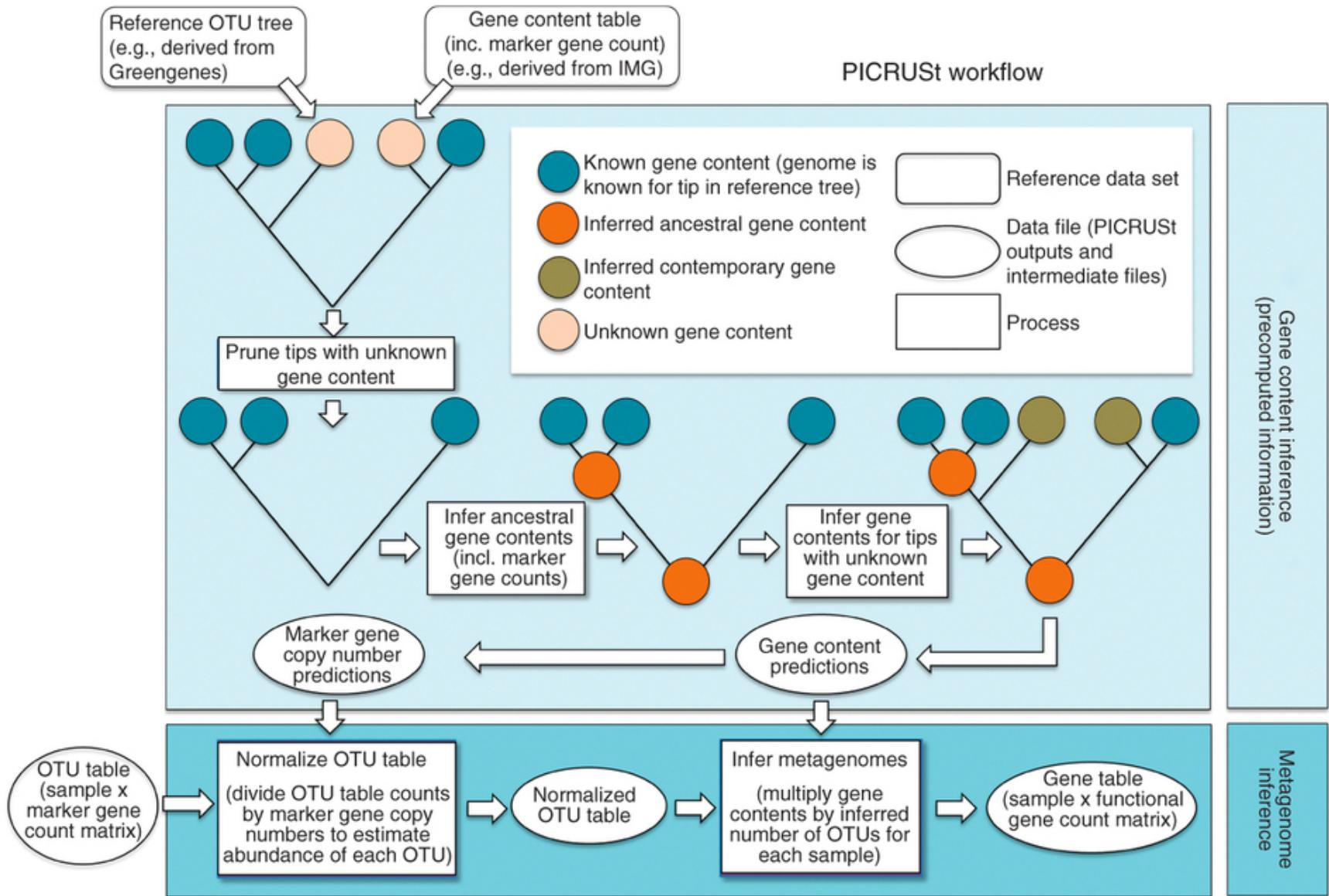
PiCrust



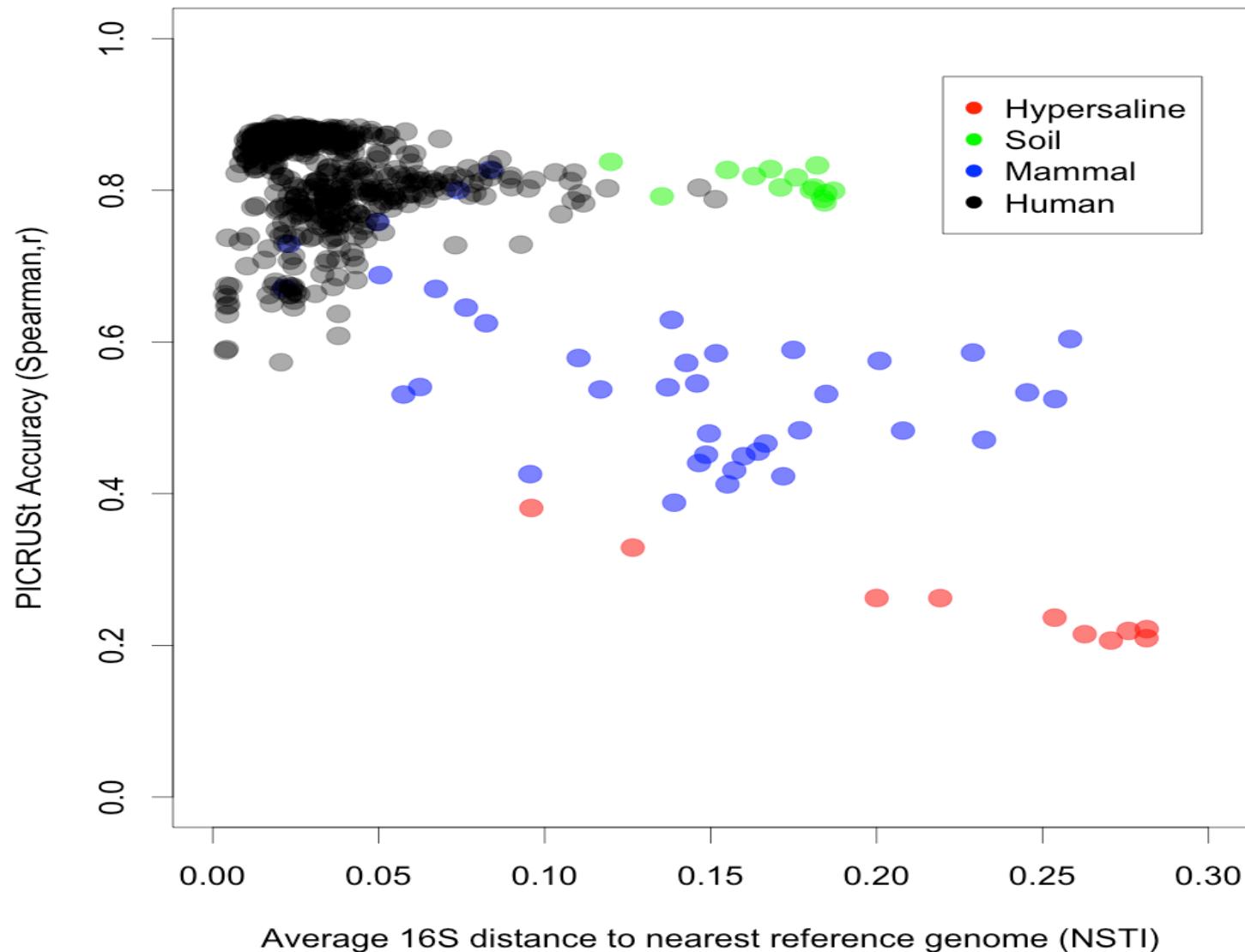
Example results



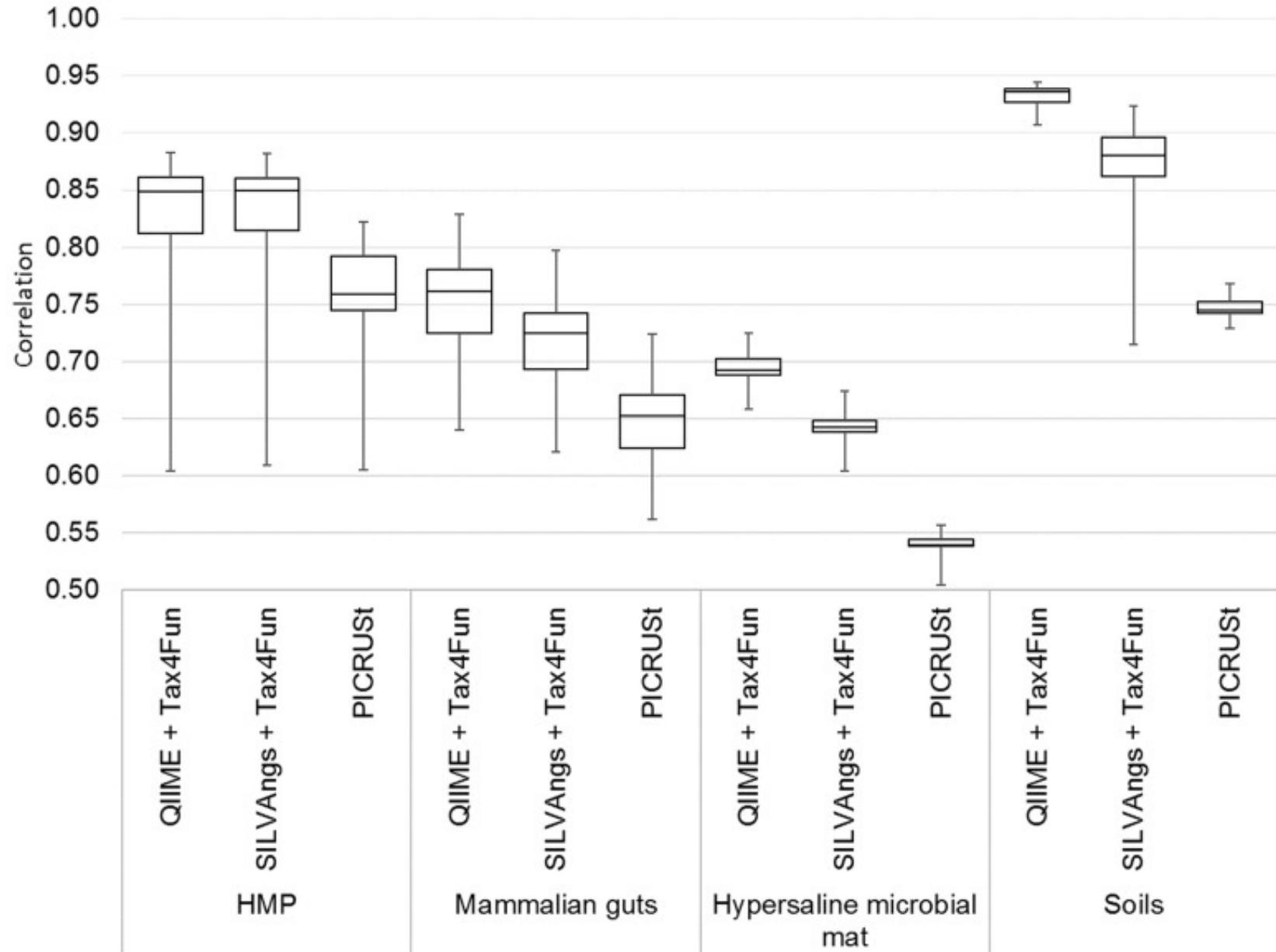
Picrust Workflow



Picrust Accuracy vs. Habitat Microbiome



Tax4Fun performance



Others: e.g. Vikodak

