



# Unlocking Insights from Isolate Genome Collections

Rekha Seshadri, PhD

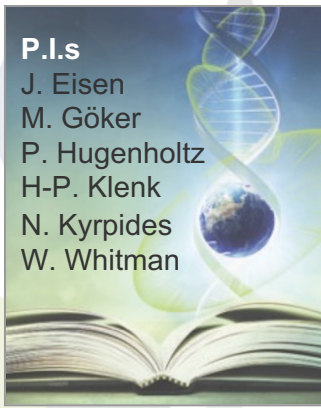
Jun 10, 2024

USCCN Workshop

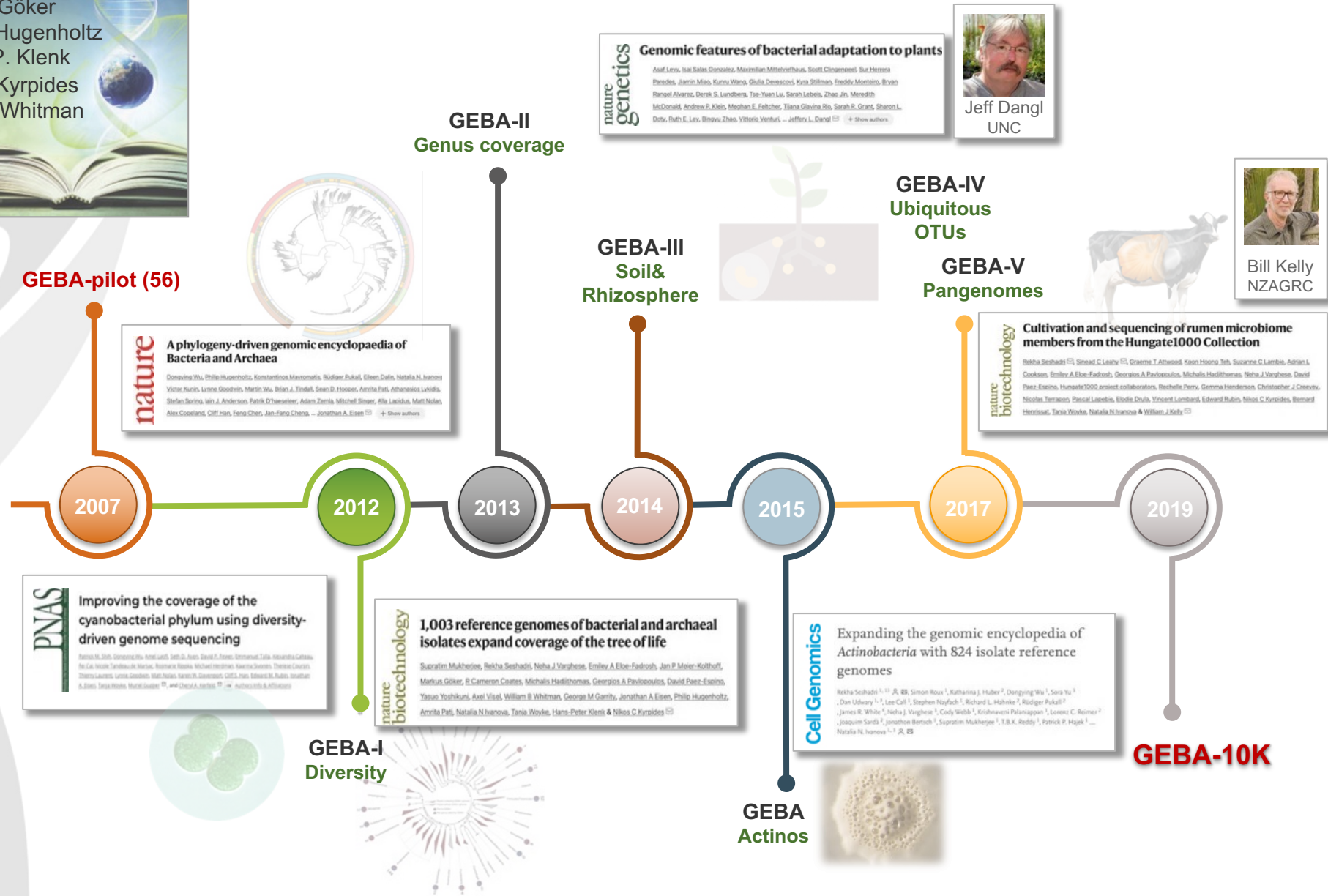
Davis, CA

**P.I.s**

J. Eisen  
M. Göker  
P. Hugenholtz  
H-P. Klenk  
N. Kyrpides  
W. Whitman



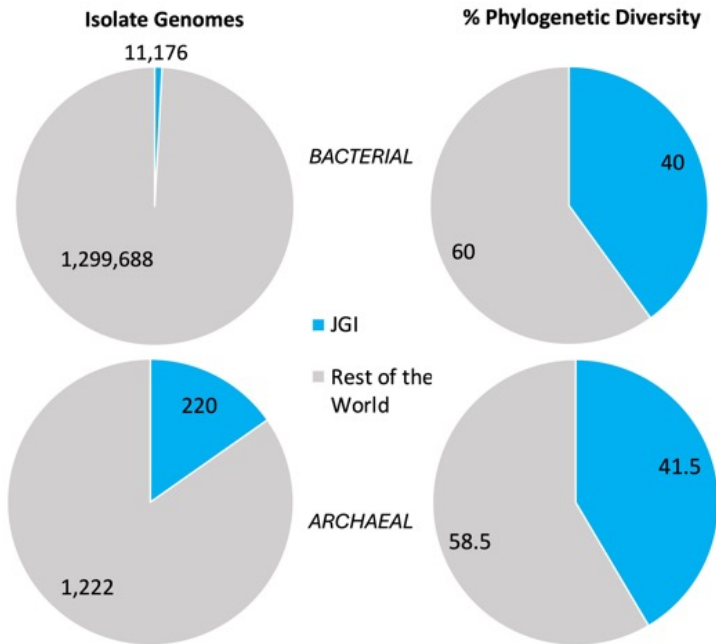
# CSP: Genomic Encyclopedias of Bacteria & Archaea (GEBA)



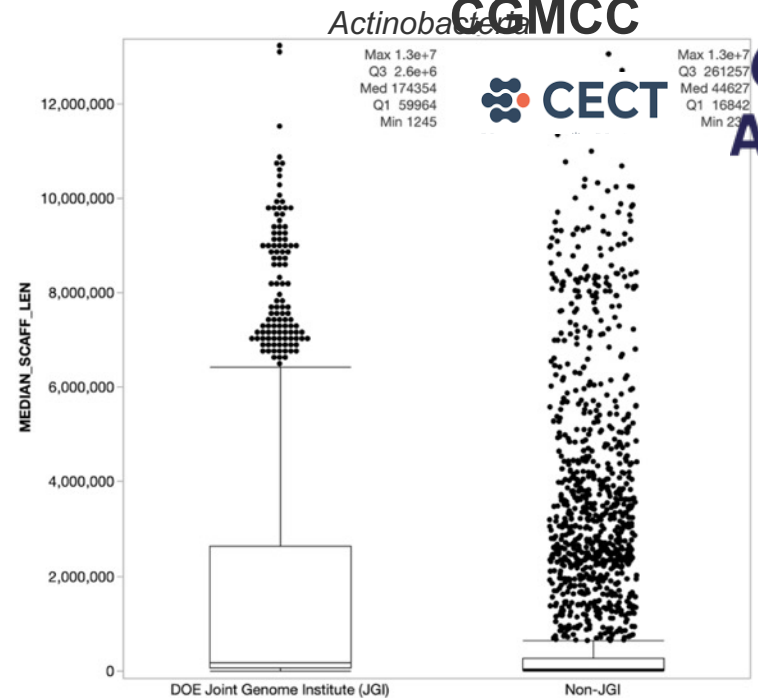
# Expanding Phylogenetic Diversity (w/ Type Strains)



*Result: <10K JGI genomes account for >40% of total diversity of isolate taxa in public databases*



*PD estimated from summed lengths of trees of universally conserved single copy marker gene*



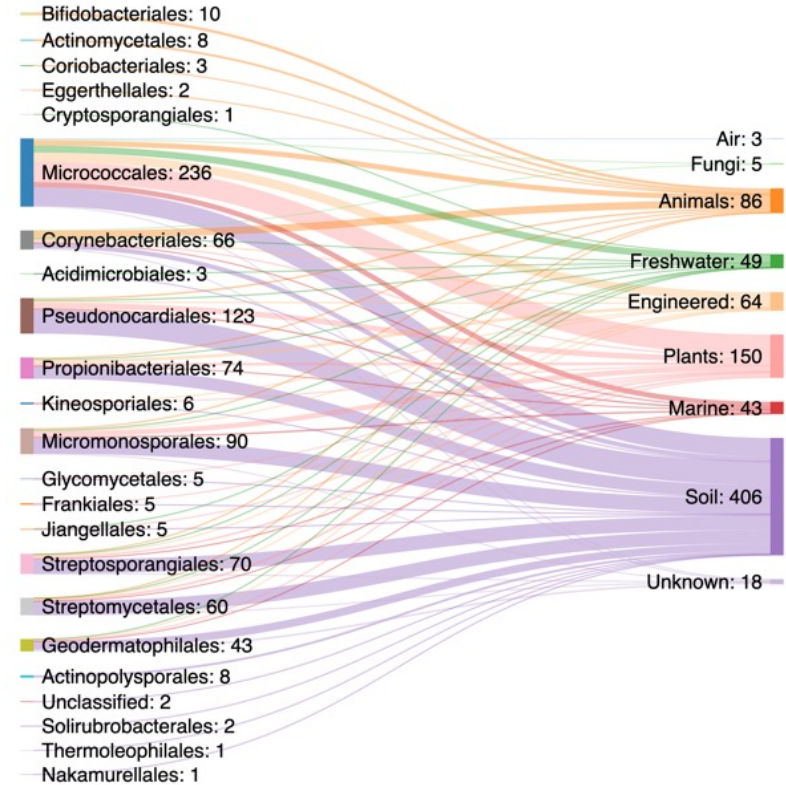
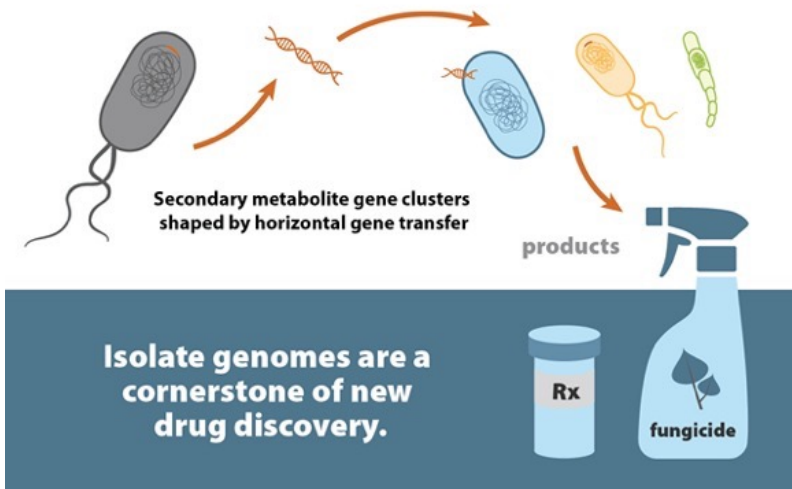
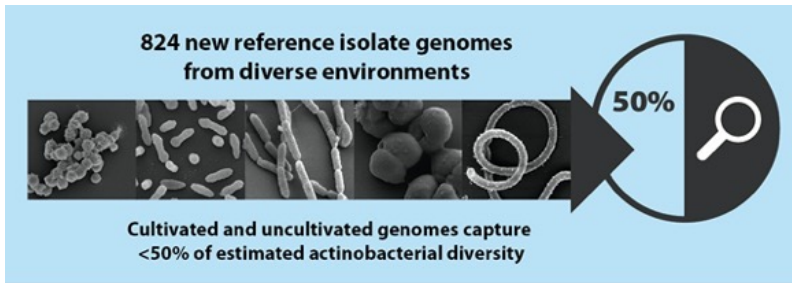
*JGI genomes are less fragmented*

# Vignettes from GEBA-Actinobacteria

## Cell Genomics



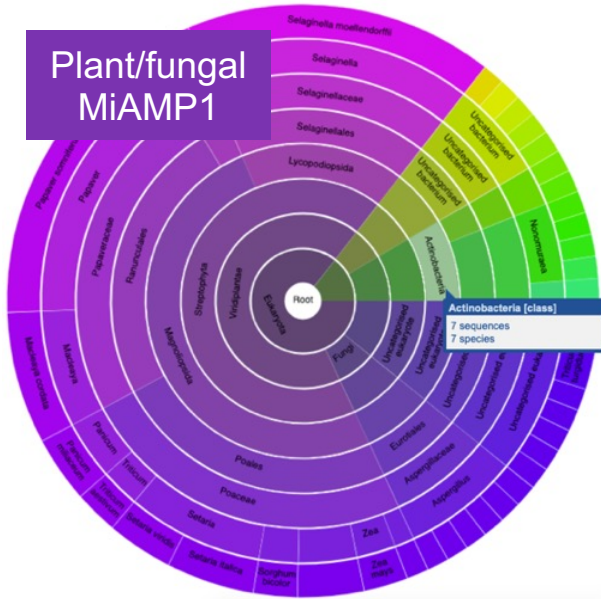
### Expanding the genomic encyclopedia of *Actinobacteria* with 824 isolate reference genomes



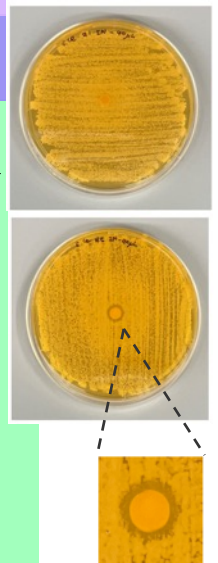
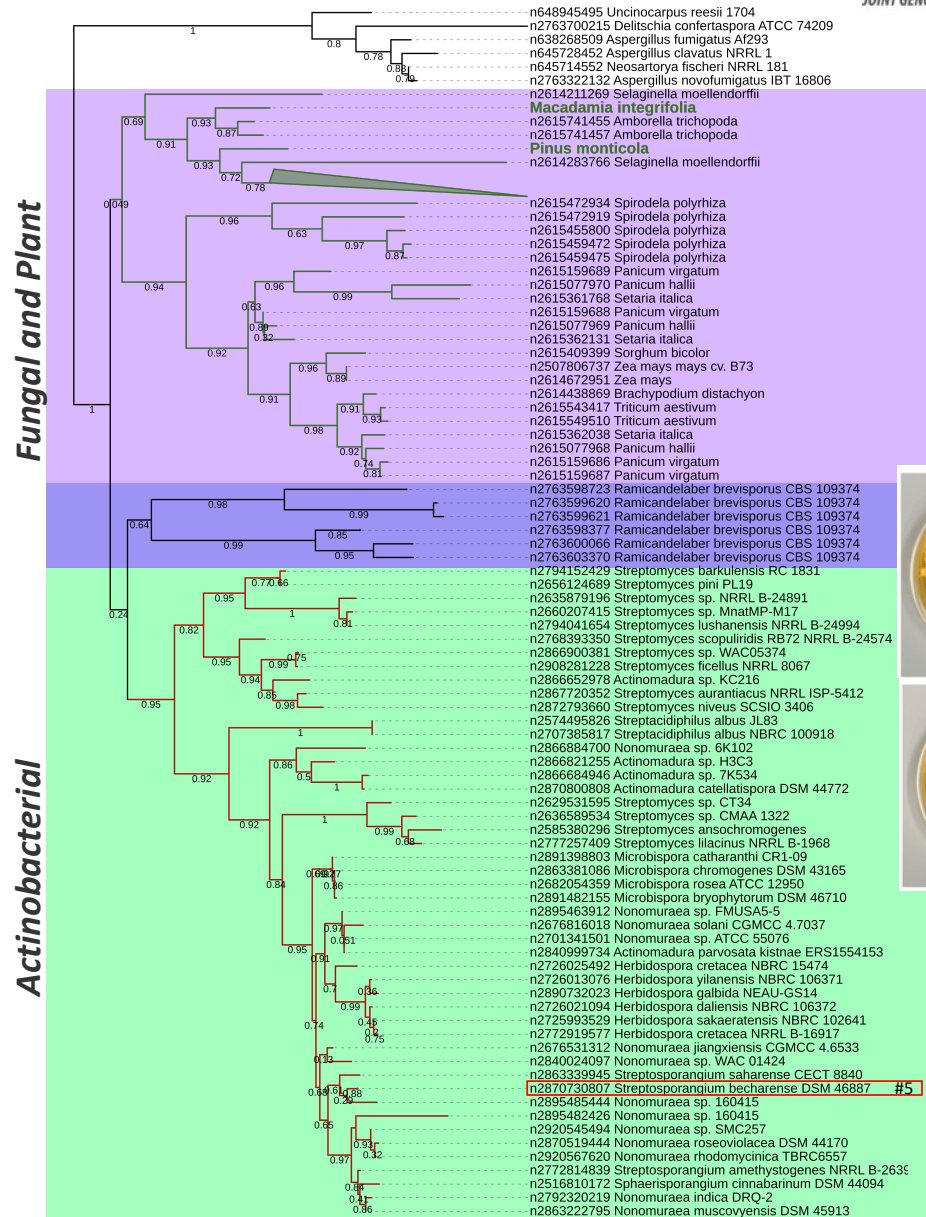
↑ *statistical power from incr. sample size*  
*metadata is rel. poor for isolates*



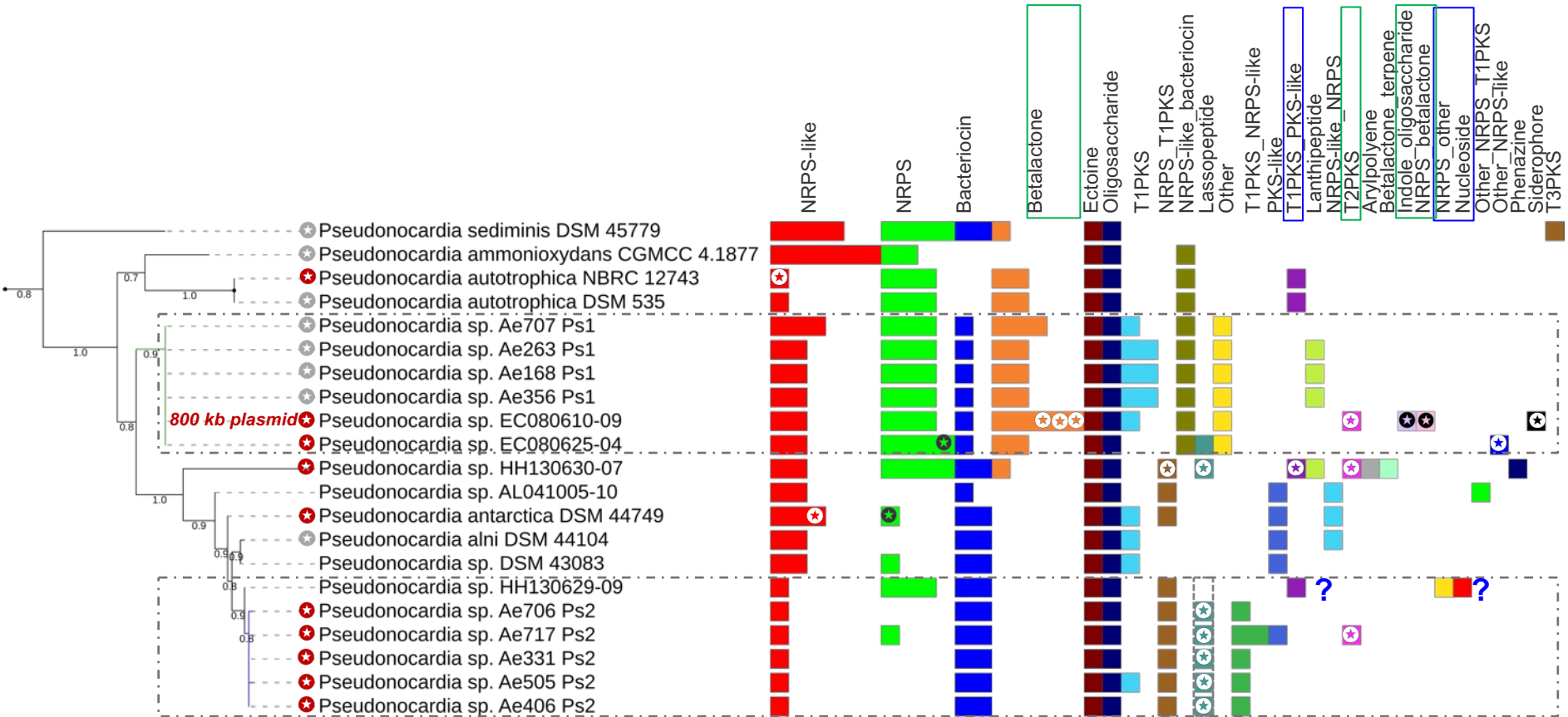
# Antimicrobial Peptide (AMP) in Plant/Soil Actino



Statistical comparison of host vs environ Actinobacteria reveals **enriched AMP** in isolates from plant/rhizosphere



# Sec. metabolites shaped by horizontal gene transfer



# Vignettes from Hungate Genomes

nature biotechnology

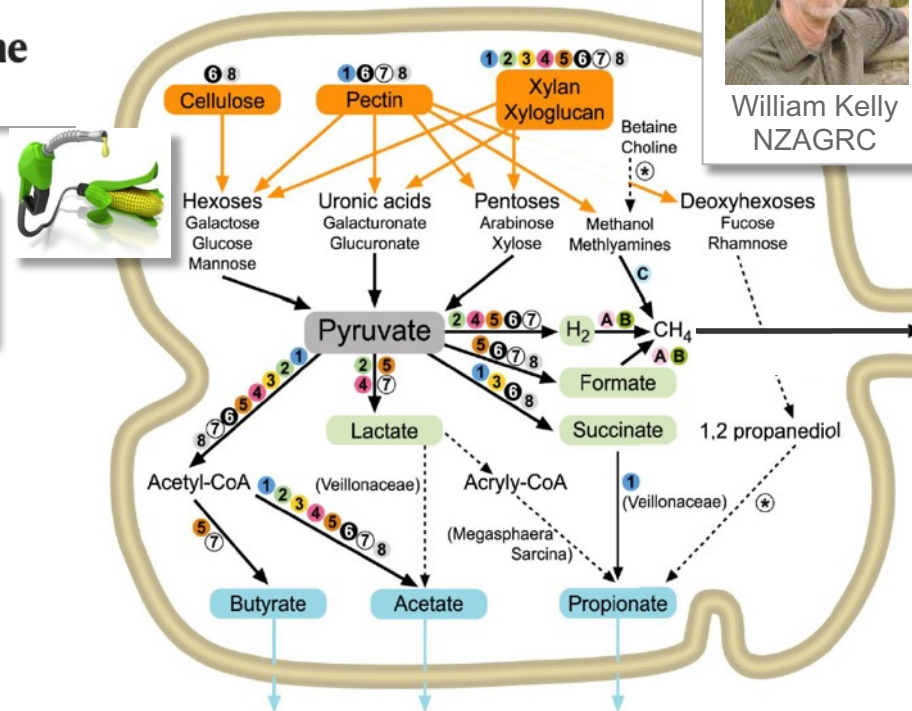
Resource | [Open access](#) | Published: 19 March 2018

## Cultivation and sequencing of rumen microbiome members from the Hungate1000 Collection



William Kelly  
NZAGRC

- Result: 410 new genomes from the rumen
- Previous unknown roles for taxa
- New enzymes for lignocellulose conversion

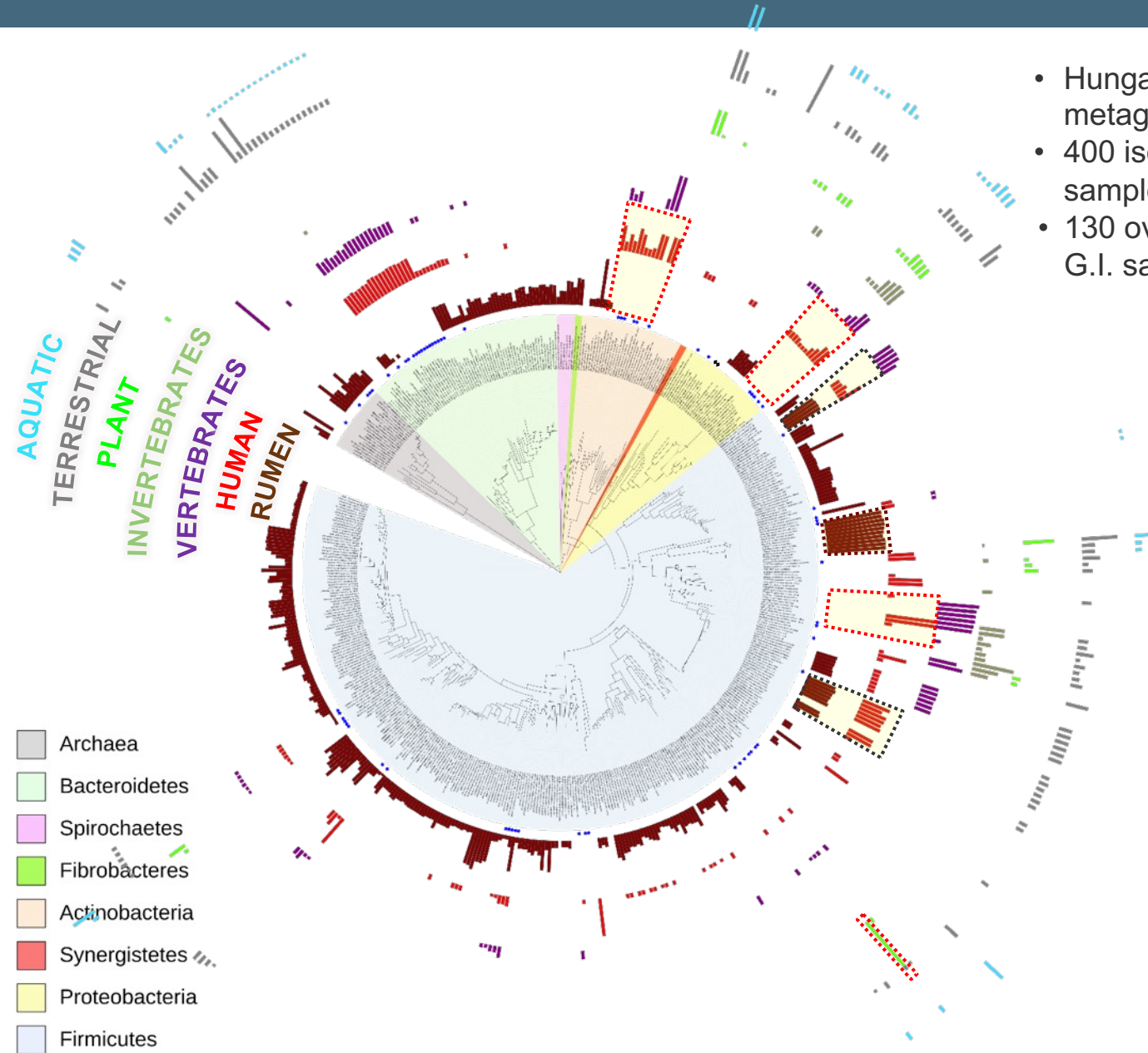


Bacterial Group	Abundance %	Prevalence %
1 <i>Prevotella</i>	22	100
2 Clostridiales	15.3	100
3 Bacteroidales	8.4	100
4 Ruminococcaceae	7.9	100
5 Lachnospiraceae	6.3	100
6 <i>Ruminococcus</i>	3.6	100
7 <i>Butyrivibrio</i>	3.4	100
8 <i>Fibrobacter</i>	2.9	93
Total		69.8
Archaeal Group		
A <i>M. gottschalkii</i>	46.9	100
B <i>M. ruminantium</i>	27.1	99
C <i>Methanomassiliicoccales</i> Group 12 sp.	6.5	87
Total		80.5



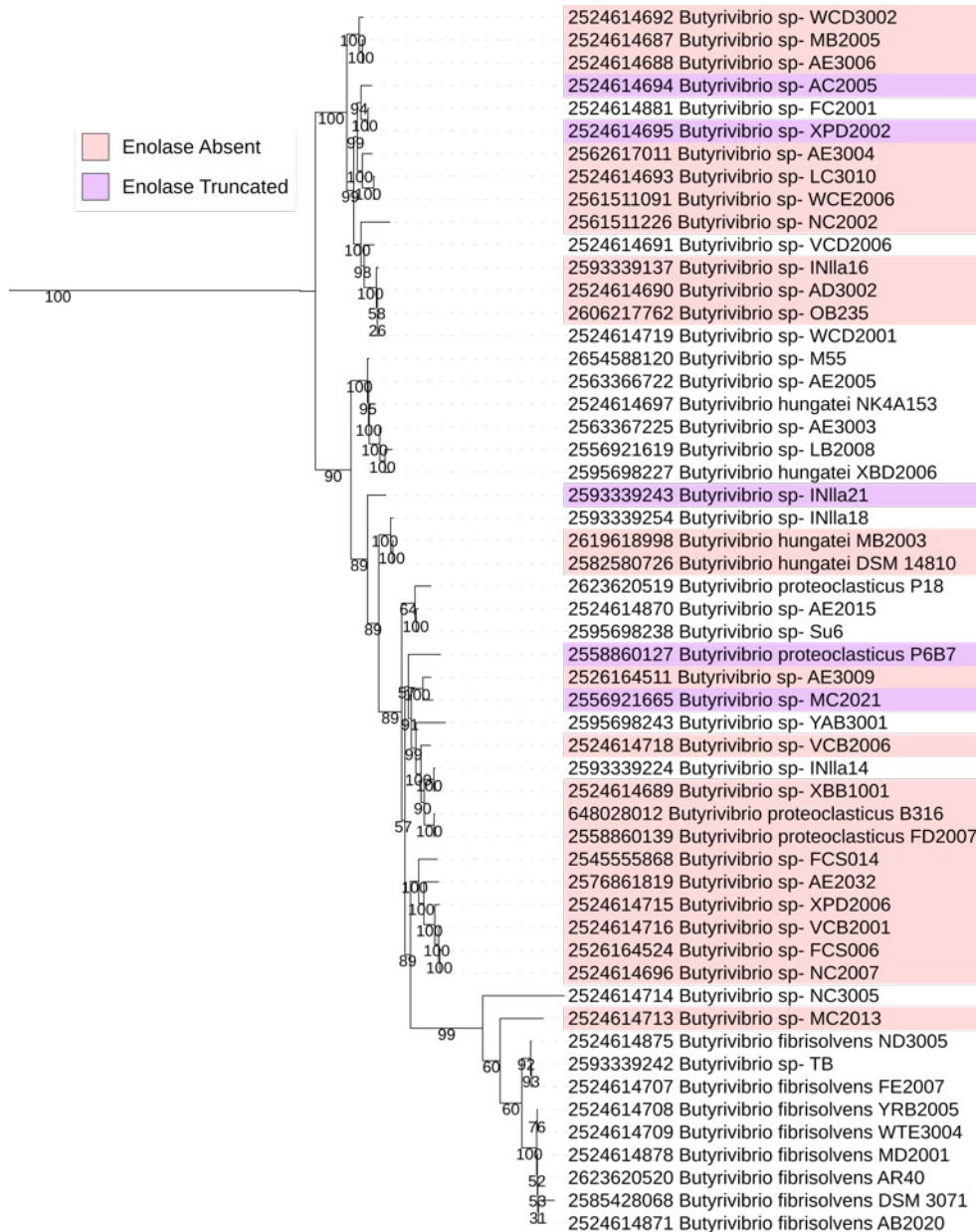
# Environmental relevance

- Hungate ISOLATES against >8,200 metagenomes
- 400 isolates detected in just 40 rumen samples
- 130 overlapping species with human G.I. samples





# Evolution by Gene Loss



- Enolase – essential for growth (ATP, NADP and pyruvate)

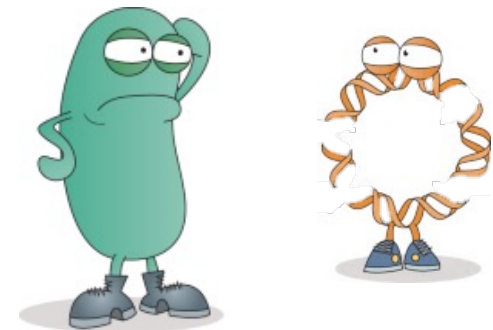
— 25 intact enolases

— 24 absent

— 5 truncated

- mutation, intron insertion

- Multiple loss events



# Vignettes from Root Nodule Bacteria (RNB)

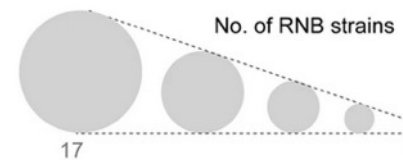
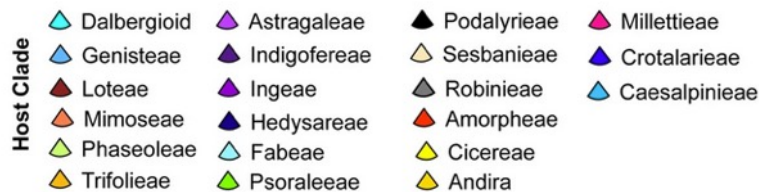
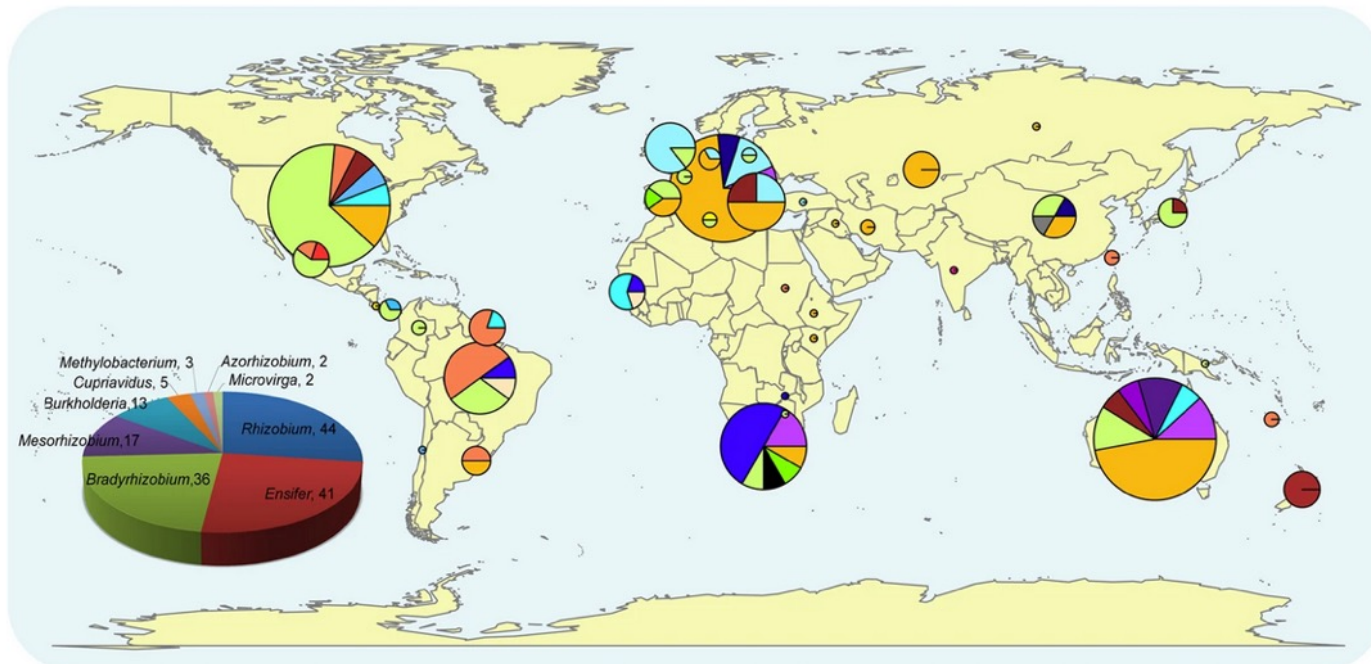
## scientific reports

Article | [Open access](#) | Published: 20 November 2015

### Discovery of Novel Plant Interaction Determinants from the Genomes of 163 Root Nodule Bacteria

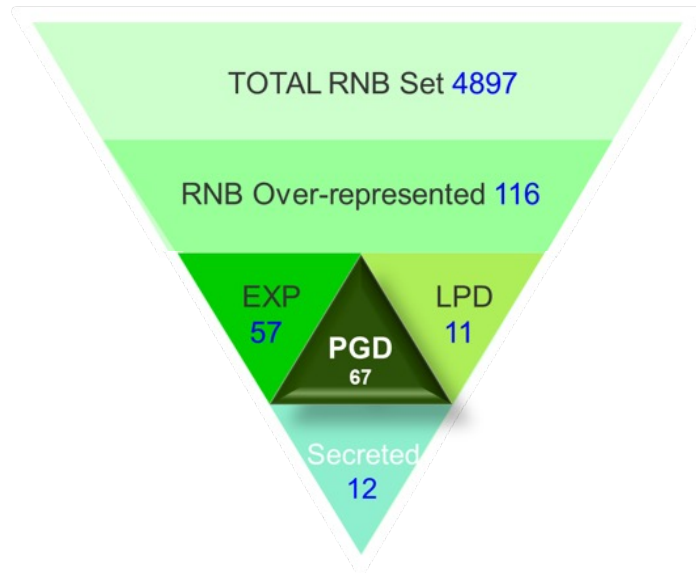


Wayne Reeve

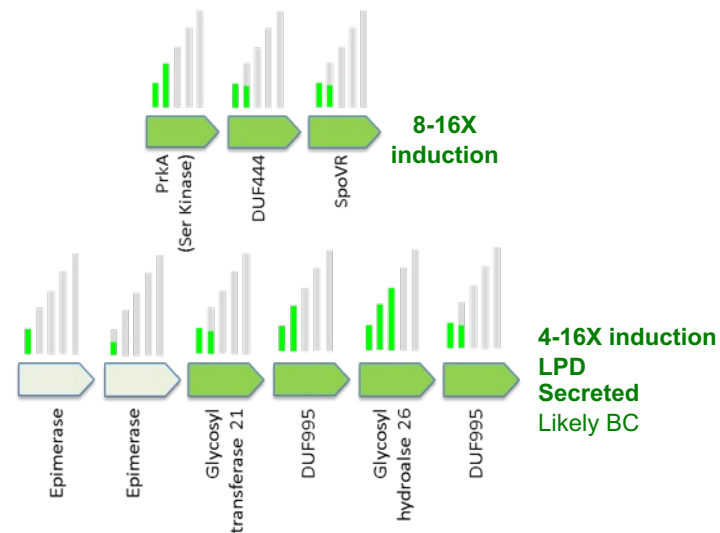


Summary of biogeography and taxonomy of 163 RNB strains analyzed in this study.

# RNB-enriched Functions

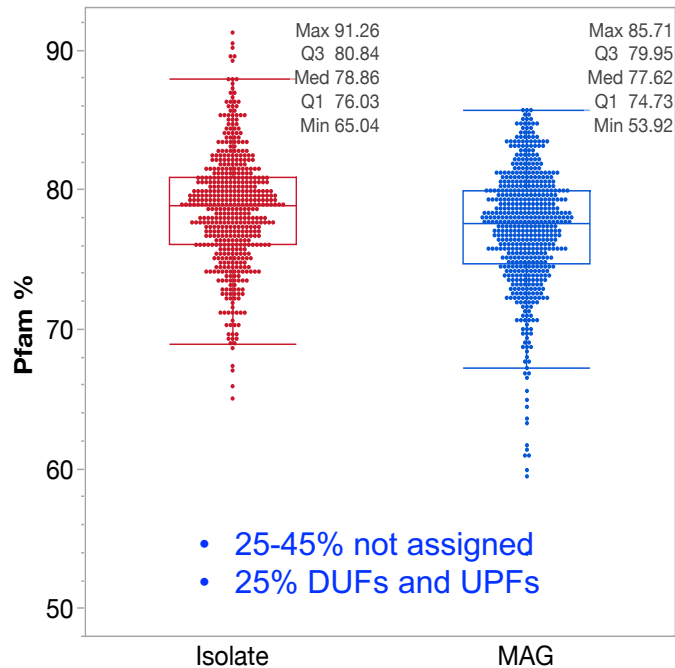


- Novel determinants of microbe-host interaction or plant growth promotion



# Why we need isolates and METADATA

- High quality, more complete (incl. HGT), reliable genome-, multi-omics data
- Confirm or second assertions from uncultivated data
- Support biochemical validation to avoid “homology creep”
  - Only 2.5 - 21% of annotations with experimental evidence for 10 best annotated genomes in 2022, *Cre'cy-lagard et al., Database, Vol 2022, baac062*
- Foundation for an array of experiments, incl. development of new microbial model systems





# JGI Data & Tools



[COVID-19](#) [ABOUT US](#) [CONTACT US](#) Search JGI websites ...

[Our Science](#)

[Our Projects](#)

[Data & Tools](#)

[User Programs](#)

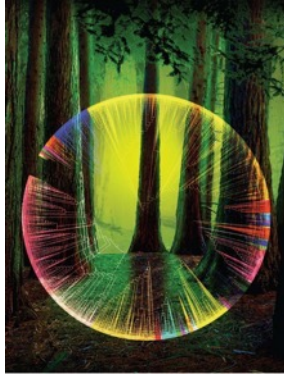
[News & Publications](#)

**IMG/M**

INTEGRATED MICROBIA



**Data Portal**



**MycoCosm**

THE FUNGAL GENOMICS RESOURCE



**PhycoCosm**

THE ALGAL GENOMICS RESOURCE



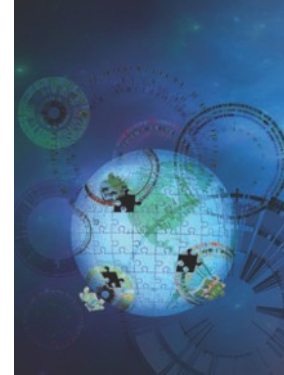
**Phytozome 13**

THE PLANT GENOMICS RESOURCE



**GOLD**

GENOMES ONLINE DATABASE



**IMG/M** Integrated Microbial Genomes & Microbiomes

**Bridging the Gap  
from Sequence  
to Biology**



## JGI IMG/M

### IMG USERS

**25,000+**  
People

**800+**  
Industry

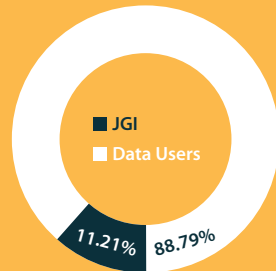
**110+**  
Countries

**70+**  
University  
Courses

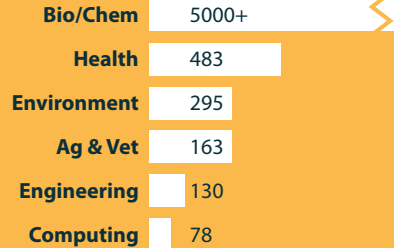
### 8,400+ PUBLICATIONS

**123**  
Filed Patents

#### Impact Beyond JGI



#### Disciplines



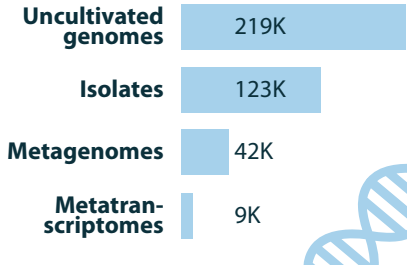
**2,872**  
Organizations

### DATA & TOOLS



#### GENOMES & MICROBIOMES

<b>200,000+</b> Datasets	<b>27T</b> Basepairs	<b>77B</b> Genes
-----------------------------	-------------------------	---------------------



#### OVER 50 TOOLS FOR:

- Metadata-enabled searches
- Genome comparisons
- Refined functional annotation
- Large-scale computation results

**THE MISSION:** The Integrated Microbial Genomes & Microbiomes (IMG/M) system aims to support the annotation, analysis, and distribution of microbial genome and microbiome datasets sequenced at the DOE Joint Genome Institute. It also captures the diversity of public datasets to enable comprehensive comparative analysis and to meet the needs of our diverse community of users.



**IMG/VR**  
INTEGRATED MICROBIAL GENOMES / VIRUS

**IMG/PR**  
INTEGRATED MICROBIAL GENOMES / PLASMID

Viral & Plasmid  
predictions

# IMG Workshops (in person)



Search JGI websites ... **SE**

Email: [rshadri@lbl.gov](mailto:rshadri@lbl.gov)

MGM WORKSHOPS

[Home](#) [Webinar Series](#) [Agenda](#) [Organizers](#) [Logistics](#) [Register](#)

## Microbial Genomics & Metagenomics Workshops



### MGM Workshop Registration

Registration has closed for the MGM Workshop taking place on April 15 – 19, 2024 at the Integrative Genomics Building (Bldg 91), onsite at the Lawrence Berkeley National Lab.

## About this Workshop

The U.S. Department of Energy (DOE) Joint Genome Institute (JGI) is offering a five-day workshop on Microbial Genomics and Metagenomics. This highly hands-on workshop is designed to familiarize users with the **Integrated Microbial Genomes & Microbiomes** (IMG/M) data and workflows for computational analysis and interpretation of sequence data. IMG/M is a web-based platform that provides access to the wealth of public microbiome sequence data and enables the user to analyze these data with existing bioinformatic methods to bridge the gap from sequence to biology.

Users will complete hands-on exercises involving comparative analysis, visualization and interpretation of genomic and metagenomic datasets. They will apply these workflows to individual or group projects and present findings. The workshop also includes brief overviews of sequencing, assembly, metadata curation, annotation, as well as some presentations from the JGI's portfolio of products beyond sequencing such as DNA synthesis, metabolomics, single cell technology, and more.

Learn more about IMG/M's capabilities by watching our [introductory video](#).



**MGM-28**  
February 25-March 1, 2019



**MGM at Valencia, Spain**  
July 2019



**January 2018**  
University of Tennessee



**July 2014**  
Goettingen, Germany