Gall-ID: web-based tools for the rapid identification and characterization of gall-causing phytopathogenic bacteria

Edward W. Davis II[1,2]*, Alexandra J. Weisberg[1]*, Javier F. Tabima[1], Melodie L. Putnam[1], Niklaus J. Grünwald[1,2,3], and Jeff H. Chang[1,2,4]

[1]Department of Botany and Plant Pathology, Oregon State University, Corvallis, OR, 97331, USA, [2]Molecular and Cellular Biology Program, Oregon State University, [3]Horticultural Crops Research Laboratory, USDA-ARS, Corvallis, OR, 97331, USA, [4]Center for Genome Research and Biocomputing, Oregon State University. * contributed equally.

Abstract

Understanding the genetic diversity of plant pathogens and the effect of agricultural practices on pathogen evolution are important for disease management. Advances in DNA sequencing technology have contributed to greater reliance on use of 16S rDNA, multilocus sequence analysis (MLSA), and whole genome sequences to genotype bacteria. Correct analysis and interpretation of sequencing data can be difficult. Therefore we have developed a set of web-based tools, termed Gall-ID, to facilitate the identification and characterization of phytopathogenic bacteria, with a focus on those that cause gall diseases. Users can compare 16S or MLSA gene sequences from an isolate against manually-curated databases, and generate a phylogenetic tree containing their isolate. Gall-ID also includes a tool for uploading and using whole genome sequencing reads to identify homologs of known virulence genes. Finally, Gall-ID provides downloadable software pipelines for core genome analysis (WGS Pipeline), calculation of average nucleotide identity (Auto ANI), and the generation of MLSA gene set databases (Auto MLSA).

Introduction to Gall-ID

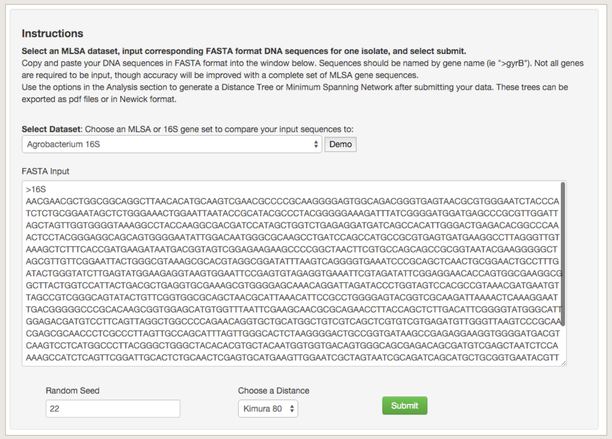
- Web-based tools to easily identify a bacterial isolate using 16S or MLSA gene sequences
- Enables rapid identification without bioinformatics experience
- Downloadable software tools available for core genome and whole genome comparisons
- Available at:

http://gall-id.cgrb.oregonstate.edu

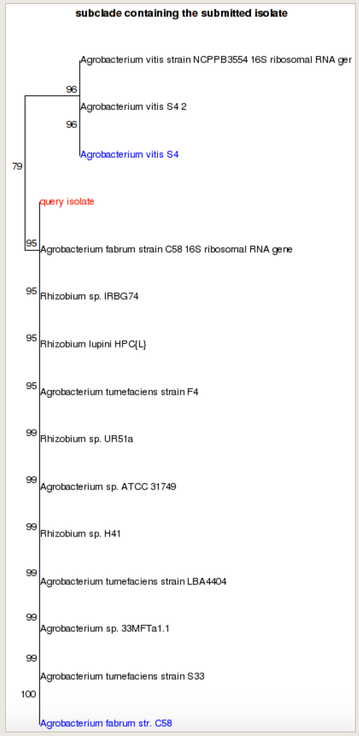
16S rDNA or MLSA gene sequences ← I have… → Whole genome Sequencing reads

(online) (downloadable tools)

Gall Isolate Typing/Phytopath-type Tools



Select an organism, a 16S or MLSA gene dataset, and input your sequences



Gall-ID produces and displays a phylogenetic tree containing your isolate for easy identification

16S database

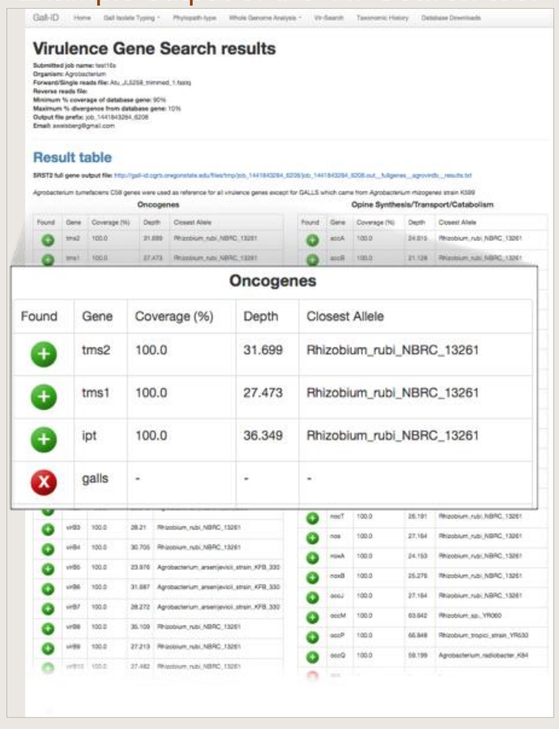
MLSA databases:

- *Agrobacterium*
- *Clavibacter*
- *Dickeya*
- *Pantoea agglomerans*
- *Pectobacterium*
- *Pseudomonas savastanoi*
- *Ralstonia*
- *Rhodococcus fascians*
- *Xanthomonas*
- *Xylella*

Vir-Search

Identify the presence of known virulence genes in raw whole genome sequencing data: Gene coverage and read-mapping depth as well as the most closely related organism allele for each gene are reported.

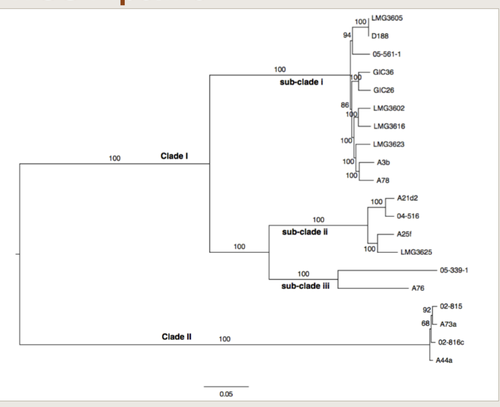
Example output of the **Vir-Search** tool



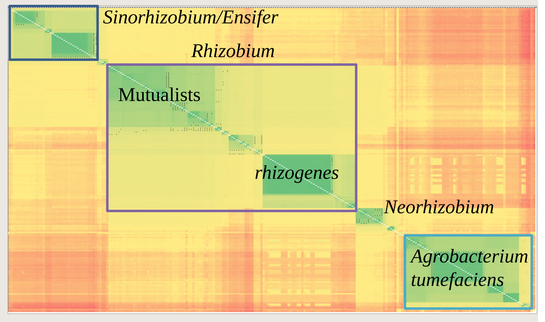
WGS Pipeline/Auto ANI

Simplified whole genome analysis: Use whole genome sequencing reads and reference genome sequences to generate a core genome phylogeny, or calculate pairwise average nucleotide identity (ANI) between genome assemblies, with minimal user input.

Phylogenetic tree of *R. fascians* genomes generated using the **WGS Pipeline**



Average nucleotide identity (ANI) heatmap of *A. tumefaciens* genomes generated using **Auto ANI**



Acknowledgements:

This work was supported by the National Institute of Food and Agriculture, U.S. Department of Agriculture award 2014-51181-22384 (JHC and NJG). Partial support also was provided by the USDA Agricultural Research Service Grant 5358-22000-039-00D (NJG), USDA National Institute of Food and Agriculture Grant 2011-68004-30154 (NJG), and the USDA ARS Floriculture Nursery Research Initiative (NJG). EWD is supported by a Provost's Distinguished Graduate Fellowship awarded by Oregon State University. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1314109 to EDW. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors(s) and do not necessarily reflect the views of the U. S. Department of Agriculture or National Science Foundation.

United States Department of Agriculture | National Institute of Food and Agriculture

WPDN Western Plant Diagnostic Network

Oregon State University