Supporting the bioinformatics community by supporting users of Bactopia



Robert A. Petit III^{1,2}, Taylor Fearing¹, Chayse Rowley¹, Jim Mildenberger¹, and Timothy D. Read³

¹Wyoming Public Health Laboratory, Wyoming Department of Health, Cheyenne, Wyoming, USA ²Theiagen Genomics, Highlands Ranch, CO, USA



Create a Sample Sheet

to create a sample sheet

<u>'bactopia prepare'</u> is used

allowing you to efficiently

process a single sample or

Analysis with Bactopia

150 bioinformatic tools

and is easily adapted to

genomes with more than

Bactopia processes

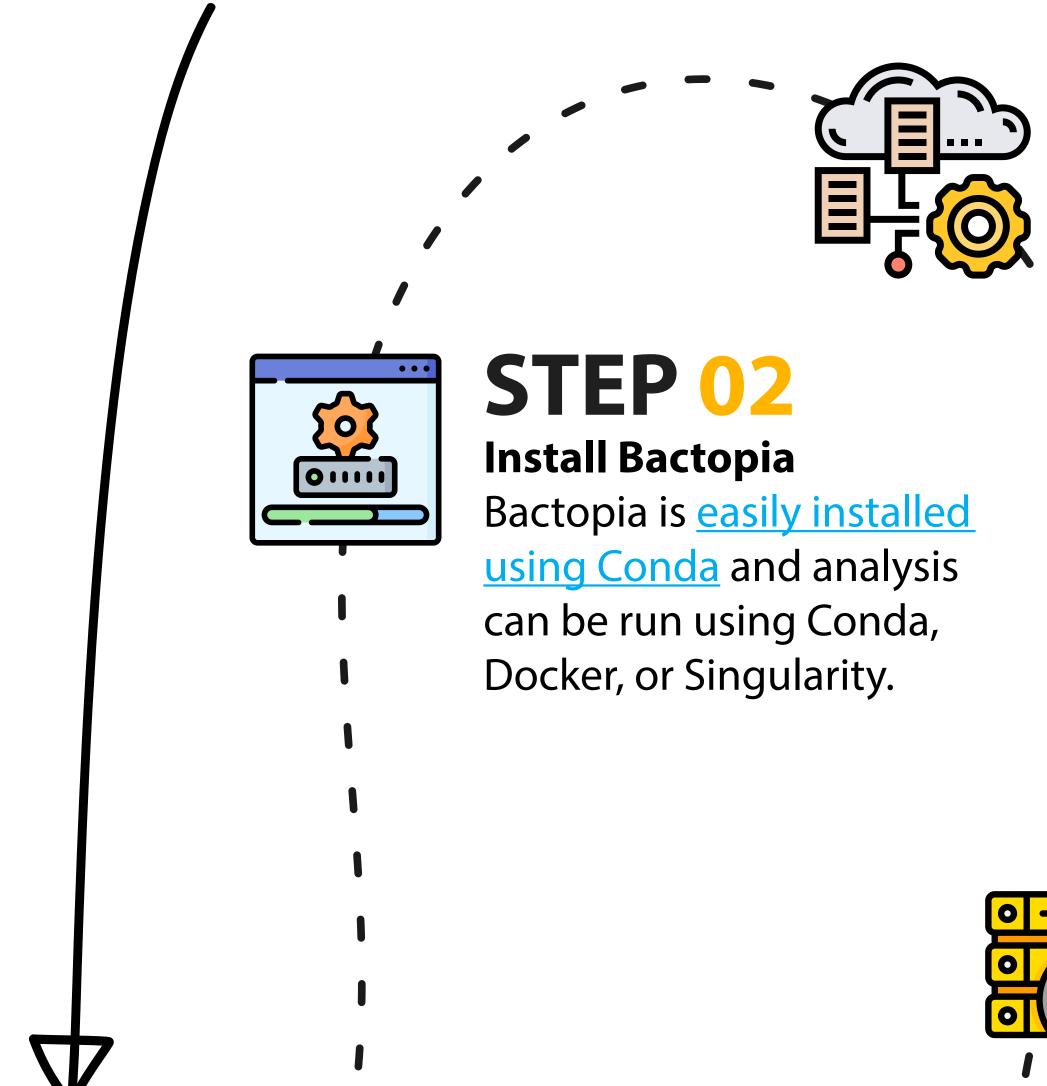
fit your needs.

thousands.



What is Bactopia?

Bactopia is an extensive Nextflow pipeline for the complete analysis of bacterial genomes. To learn more, follow the step-by-step guide below.



STEP 01

Sequence Bacterial Genomes

Bactopia supports sequences

Sequence Read Archive and

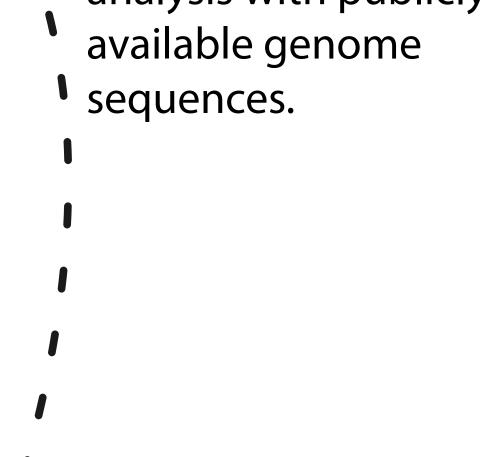
NCBI Assembly accessions.

ACT from Illumina and Nanopore

technologies, as well as

Search Public Databases <u>'bactopia search'</u> can

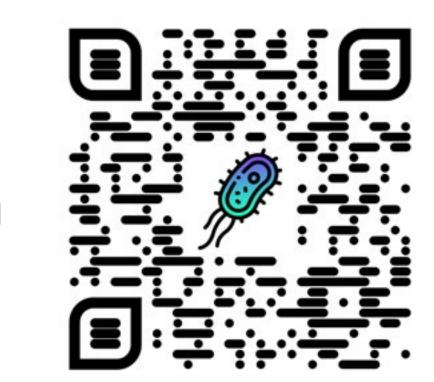
supplement your analysis with publicly available genome sequences.



STEP 04 Build Datasets

'bactopia datasets' downloads publicly available datasets to further supplement your analysis.

Scan to get started with Bactopia



STEP 08 Investigate Results Once completed, numerous output files are available for your /

genomic investigations.

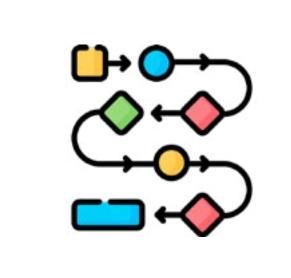
STEP 07 **Further Analysis with Bactopia Tools**

After Bactopia is finished, more than 30 additional workflows, called Bactopia Tools, are available to take your analysis even further.



ACCESSIBLE

Bactopia is open source and available from Conda, Docker and Singularity. Bactopia has been downloaded more than 30,000 times from users around the world.



COMPREHENSIVE

Bactopia is a start-tofinish pipeline which includes <u>numerous</u> tools and workflows commonly used for bacterial genome analysis.



PORTABLE

With a simple profile change you can go from processing genomes on your laptop to an HPC system or any of the major cloud providers (AWS, GCP, Azure).



REPRODUCIBLE

Bactopia was developed following <u>nf-core best</u> **practices** which ensures a robust pipeline with strict version control and an extensive audit trail.



RESILIENT

More than 100 tests, testing 10,000+ variables, assist in identifying potential bugs and downstream changes, before users are affected.



SCALABLE

Bactopia allows you to easily scale from a few genomes to thousands of genomes. For example processing 67,000 genomes in 5 days on AWS.

BACTOPIA

Enhancements to

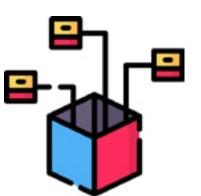
Open Source Science

How supporting Bactopia users supports the **Bioinformatics Community?**

In order to reduce the burden of ongoing maintenance and to better help users, a few design principles were adopted.

- 1. Tools must be free and open source
- 2. Tools must be available from Bioconda or Conda-Forge
- 3. Bactopia Tools must be available from nf-core/modules

By adopting these principles, we have been able to rapidly meet user's needs, while also opening pathways to contribute back to the wider bioinformatics community (described to the right).



Many tools originally developed for Bactopia have been made available as stand-alone tools, incuding dragonflye, pbptyper, fastq-dl, and others. These tools have been downloaded more than 80,000 times from Conda.



By using tools from Conda, it has facilitated contributions to Bioconda and Conda-Forge. To date 28 new recipes have been added, <u>35 recipes updated</u>, and more than <u>1,700</u> Bioconda pull requests reviewed.



Requiring Bactopia Tools be available from nf-core/modules, has also facilitated contributions to nf-core/modules. To date, 59 contributions have been made including 43 new modules and 16 modules updated.



Occasionaly users or CI testing may identify bugs in tools used Bactopia. If a fix is identified, it is submitted upstream to the tool. This has led to 18 contributions to tools including: Ariba, Bowtie 2, Kleborate, Seroba, Shovill, ShigaTyper, and others.



