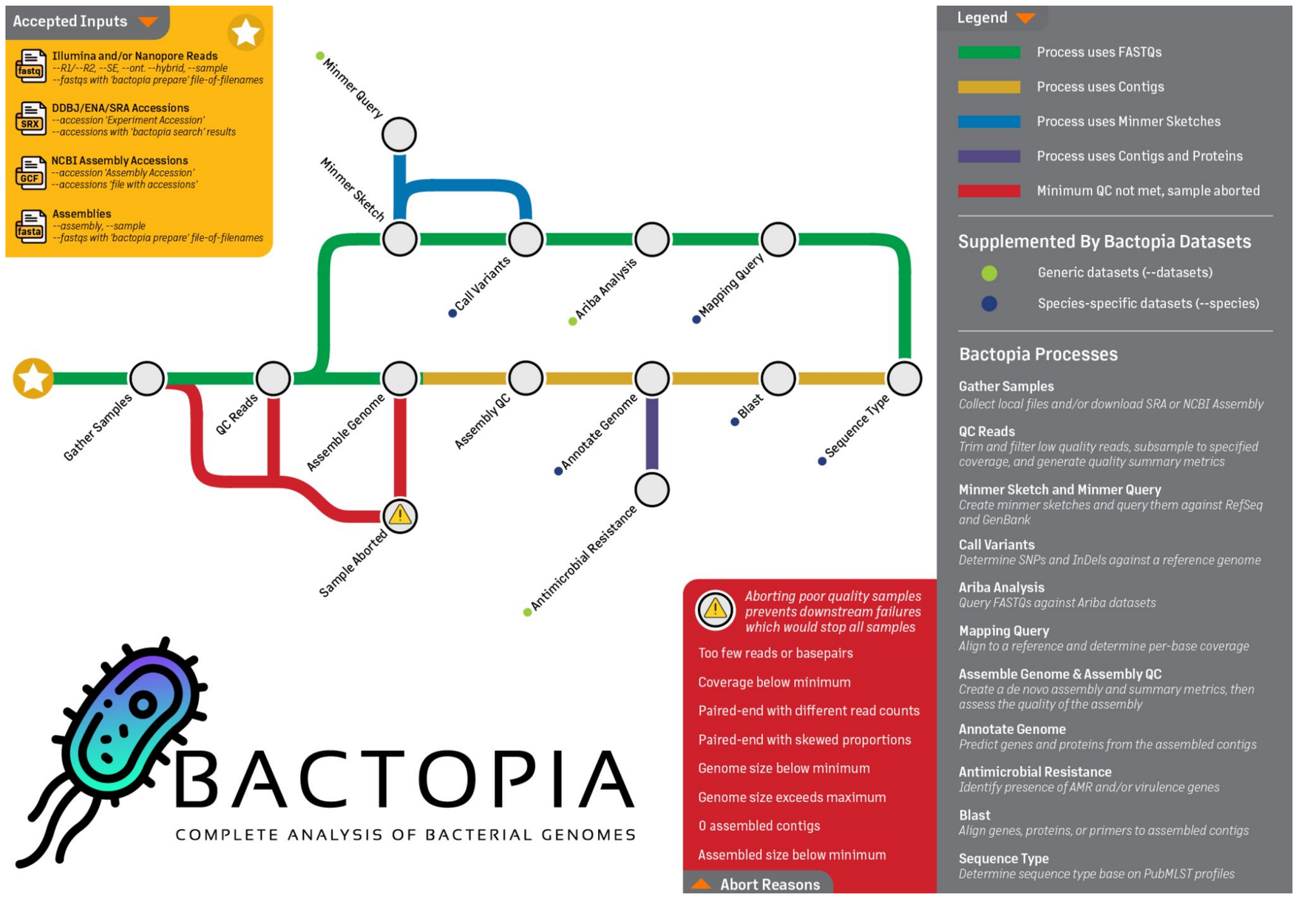
Bactopia: Highly scalable, portable and customizable bacterial genome analyses

Robert A. Petit III[1,2], Davi J. Marcon[3,4], Abhinav Sharma[5], and Timothy D. Read[6]

[1] Theiagen Genomics, Highlands Ranch, CO, USA
[2] Public Health Laboratory, Wyoming Department of Health, Cheyenne, Wyoming, USA
[3] Center of Genomics and Systems Biology, Institute of Biological Sciences, Universidade Federal do Pará, Belém, PA, Brazil
[4] Laboratory of Genetic Engineering, Guamá Science and Technology Park, Belém, PA, Brazil
[5] Faculty of Engineering and Technology. Liverpool John Moores University. Liverpool, United Kingdom.
[6] Division of Infectious Diseases, Department of Medicine, Emory University School of Medicine, Atlanta, Georgia, USA

Correspondence: robert.petit@theiagen.com **Bactopia Documentation**: bactopia.github.io

Learn more about Bactopia!



BACTOPIA
COMPLETE ANALYSIS OF BACTERIAL GENOMES

Highlights

- Complete analysis of bacterial genomes
- Nextflow DSL2, greatly increases Bactopia's ability to fit your needs
- Supports Illumina, Oxford Nanopore technologies as well as NCBI's Sequence Read Archive and Assembly databases
- Includes more than 130 bioinformatics tools
- 33 Bactopia Tools include more workflows for more science
- Extensively tested with 100+ tests for 10,000+ output files
- Available on Bioconda, Docker, and Singularity

Bactopia is platform Independent

By using Nextflow, Bactopia can run on multiple platforms, including:

kubernetes · IBM Spectrum LSF · aws · Google Cloud · GRID ENGINE · slurm workload manager · Terra · Microsoft Azure · nextflow tower

Test Everything

Bactopia features per-module tests using real bacterial sequences available from bactopia-tests. Currently 100+ tests have been created to verify more than 10,000 output files. These tests are all integrated into GitHub Actions to ensure the stability of Bactopia over time.

Bactopia Tools

Bactopia Tools are convenient workflows to do more science with your Bactopia outputs! Currently there are 33 Bactopia Tools for analyses like pan-genome construction, serotyping, and many more. Check out the full set of Bactopia Tools to the right!

nf-core/ influences

nf-core is a community effort to curate analysis pipelines built using Nextflow. This effort has provided a powerful design framework when developing Nextflow pipelines, and Bactopia has followed their lead by implementing:

- Modules from nf-core/modules for Bactopia Tools
- Bactopia has contributed 30+ modules
- Per-module tests modeled after nf-core/modules pytest implementation
- Argument parser based off nf-core/tools
- Single meta variable for general value storage

Ultimately adoption of these practices has made Bactopia a much better pipeline to use and maintain overtime.

(*Thank you nf-core team!*)

Curated Datasets

To facilitate curated datasets, bactopia-datasets was created for community members to curate species- specific datasets,which can then be readily used in Bactopia.

Community Synergy

Without the developers and maintainers of open-source tools, Bactopia would not exist. To provide back to the community, Bactopia has led to 17 Bioconda recipes, 1000+ Bioconda pull-request reviews, and merged pull-requests to popular many popular bioinformatic tools.

@rpetit3 @tdread_emory @abhi18av @rpetit3 @Mxrcon @abhi18av