

Data Analysis Pipelines

Dr. Stan Kwenda SAGESA Online Workshop 4th February 2022









Tools for Bacterial WGS analysis

- Variety of methods and tools are available to analyze bacterial pathogens
- Most bioinformatics tools are implemented in Unix environment
 - Require at least some bioinformatics expertise for usage
- Web-based bioinformatics tools
 - Often free for use
 - Do not require computational power from the user
 - Limited bioinformatics knowledge
 - In some cases demand that that users deposit the analyzed data in public repositories
- Always make an effort to browse through the documentation of web based platforms
 - Helps with choice of parameters and interpretation of results



Web based tools

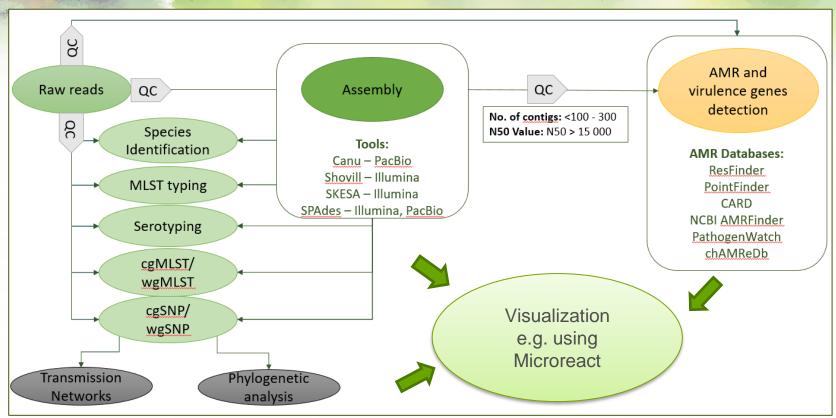
- Web-based tools
 - PathogenWatch (https://pathogen.watch/)
 - CGE tools (https://cge.cbs.dtu.dk/services/)
 - Galaxy (https://usegalaxy.org/)

Locally installed pipelines

- Locally installable:
 - Jekesa (https://github.com/stanikae/jekesa)
 - Tormes (https://github.com/nmquijada/tormes)
 - BacPipe
 https://github.com/wholeGenomeSequencingAnalysisPipeline/BacPipe)
 - Bactopia (https://github.com/bactopia/bactopia)
- Commercial (Licensed) tools
 - Genious
 - CLC Genomics Workbench Server



Bacterial WGS Workflow - Overview





Choice of Tools to Use

- Currently no universally accepted "gold-standard" tools.
- Analysts with WGS expertise may prefer different tools optimized to their needs.
 - Support often needed for individuals with limited expertise in deciding type of analyses and tools to use
- 4 things to consider when deciding on optimal tool for a task:
 - Setting/Resources e.g. type of IT infrastructure you have access to?
 - Type of sequencer used,
 - Genetic characteristics of the sequenced organism
 - Purpose of the analysis
- Results from different laboratories should be comparable/reproducible.



Applications of WGS in AMR surveillance

Transmission networks

Source of

infection or

outbreak

WGS data

Identify highrisk clones

Identify at risk groups

Identify at risk groups

- Phylogenetic Phylogenetic analysis
- analysis
- **Epidemiological** metadata
- Clinical data

Human factors contributing to

AMR evolution

Pathways of

AMR evolution

- Reference database
- AST data
- Epidemiological data
- Data on antibiotic use
- Phylogenetic analysis
- Variant analysis

AMR mechanisms

Develop pointof-care tests

- Reference database
- AST data
- Clinical data
- Data on antibiotic use
- Variant analysis
- Machine learning

- - Strain typing
 - Epidemiological data
 - Clinical Data
 - Reference database



SEQAFRICA

- Supports a consortium of 3 regional sequencing centers in Nigeria, South Africa and Tanzania, including a national centre in Ghana.
 - Provide funding for WGS services for bacterial isolates
- To apply for sequencing you fill in an application form and submit it to the sequencing center in your region
- https://antimicrobialresistance.dk/seqafrica.aspx











Thank you







