rMAP (Rapid Microbial Analysis Pipeline)
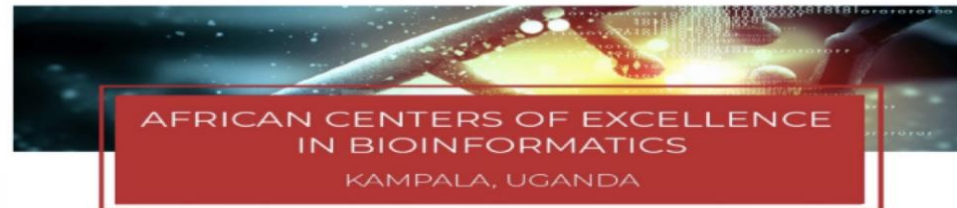
Authors:

Ivan Sserwadda & Gerald Mboowa

MAKERERE UNIVERSITY

AFRICAN CENTERS OF EXCELLENCE IN BIOINFORMATICS
KAMPALA, UGANDA

ACE AFRICAN CENTERS OF EXCELLENCE IN BIOINFORMATICS
Investing in the Future — Impacting Real Lives.

IDI INFECTIOUS DISEASES INSTITUTE

Impacts/Milestones of rMAP (2020 -2022)

1. ASM conference on Rapid Applied Microbial NGS and Bioinformatics pipelines (2020) - *Travel award*

2. African Association for Research and Control of AMR – Validating rMAP (2020) - *Young Investigators Grant*

3. Inaugural International Synthetic Biology & Biosecurity Conference in Africa (14th Oct 2021) – *Travel award*

Impacts/Milestones of rMAP (2020 -2022)

4. Peer reviewed publication (2021) – Microbial Genomics (DOI: https://doi.org/10.1099/mgen.0.000583) - *Publication*

5. African Research Collaboration on Sepsis – Bioinformatics and Molecular Biology training in Transcriptomics (2021) – *Fellowship*

6. PHA4GE - Sub-grants Promoting Sustainable Development in Bioinformatics to Support Public Health (2022) – *Research grant*

Synopsis: rMAP

- Purpose: A resistome-profiling pipeline [amr genes, plasmids, virulence factors, mlsts, insertion sequences, pangenome analysis, amr-SNPs] using Illumina WGS Paired End data inputs

- Architecture:
  - Written in: Shell script, Python, Perl, R
  - Command-line based – requires basic knowledge of Unix terminal
  - OS support: original pre-compilation - unix-based-64-bit platforms [WSL - Ubuntu 20.04.1 LTS (Focal Fossa) & Ubuntu 18.04.4 LTS (Bionic Beaver), MacOS Catalina]: Full compatibility with MacOS Big Sur & MacOS Monterey

Rapid Microbial Analysis Pipeline(rMAP)

GitHub repository: https://github.com/GunzIvan28/rMAP

Target group : ESKAPE WHO high priority bugs
- *Enterococcus faecium*
- *Staphylococcus aureus*
- *Klebsiella pneumoniae,*
- *Acinetobacter baumannii*
- *Pseudomonas aeruginosa*
- *Enterobacter* species

Other bugs:
- *Escherichia coli*
- *Citrobacter species*
- *Salmonella species*
- *Shigella species*
- *Proteus species*

Tool Features:
- QC and sequence trimming
- Assembly
- Annotations
- Variant calling
- Phylogenetic inference
- Pangenome analysis
- AMR profiling
- Virulence factor detection
- Plasmids detection
- Multilocus sequence typing
- Insertion sequence
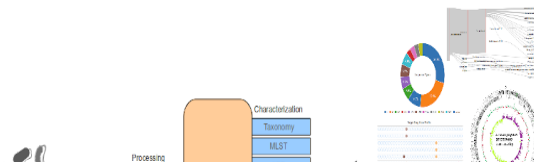
Graphical schema of rMAP workflow and tools
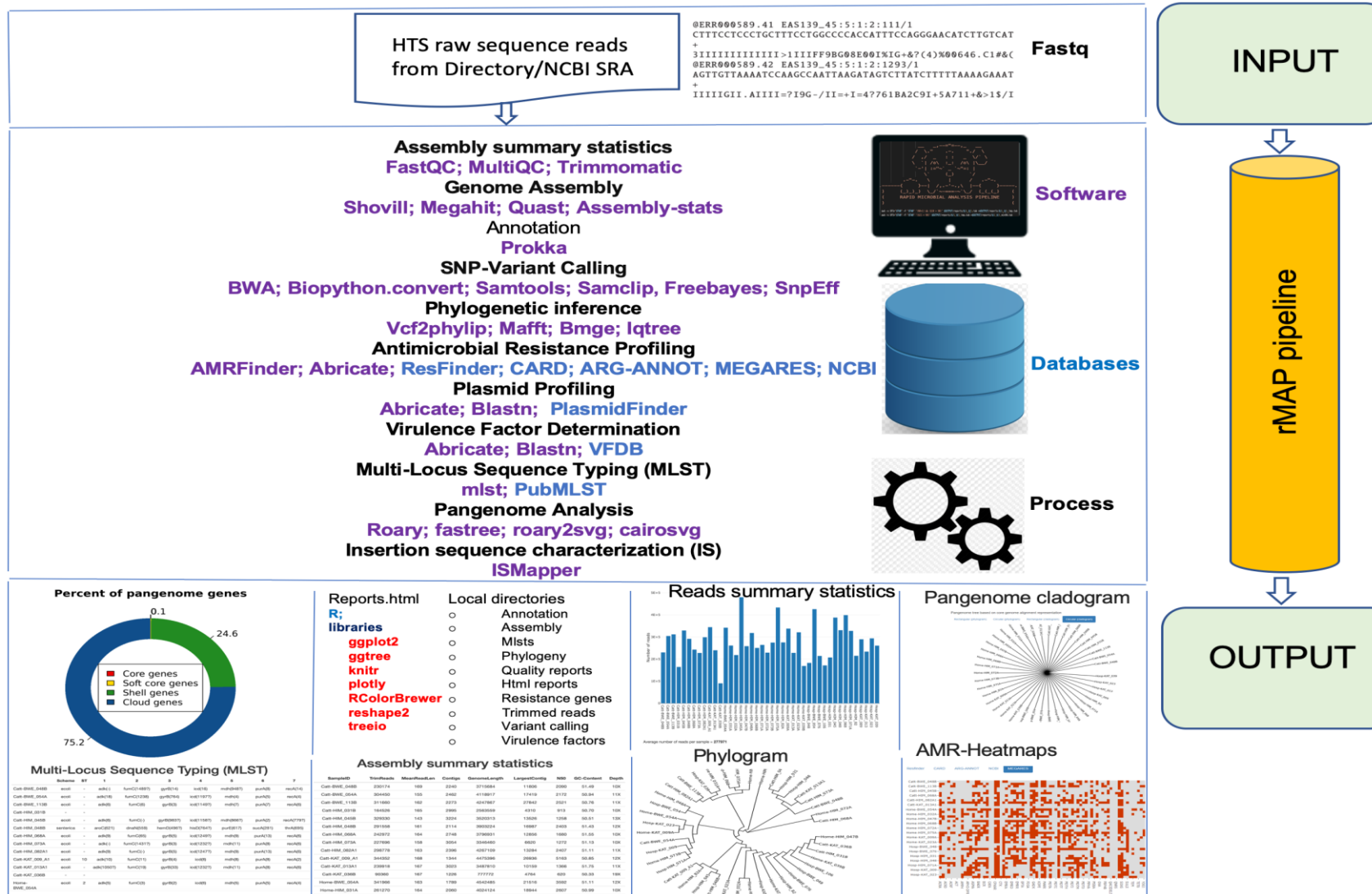
Why use rMAP?

- Highly sensitive to African-derived datasets

- Easy installation and usage

- Application to a vast range of AMR organisms

BACTOPIA
COMPLETE ANALYSIS OF BACTERIAL GENOMES

ASA³P - Automatic Bacterial Isolate Assembly,
Annotation and Analyses Pipeline

License GPL v2   Language Perl 5   Installation difficult

Nullarbor

Pipeline to generate complete public health microbiology reports from sequenced isolates

⚠ This documents the current Nullarbor 2.x version; previous 1.x is here

Anaconda.org 1.3.0   Last updated 08 Jun 2021   Install with conda   downloads 859 total

TORMES

TORMES is An automated and user-friendly pipeline for whole bacterial genome analysis of your genomes (previous assembled or downloaded from any repository) and/or your raw Illumina paired-end sequencing data, regardless th number of bacterial isolates, their origin or taxonomy.

rMAP technical aspects

- Conda installation
  - ➤ Miniconda3 platform [Anaconda as well]
  - ➤ Packages versioned in .yaml file depending on OS

- Code snippets

rMAP -t 8 --reference --input dir_name --output dir_name --quality --assembly megahit --amr --varcall --phylogeny --pangenome --gen-ele
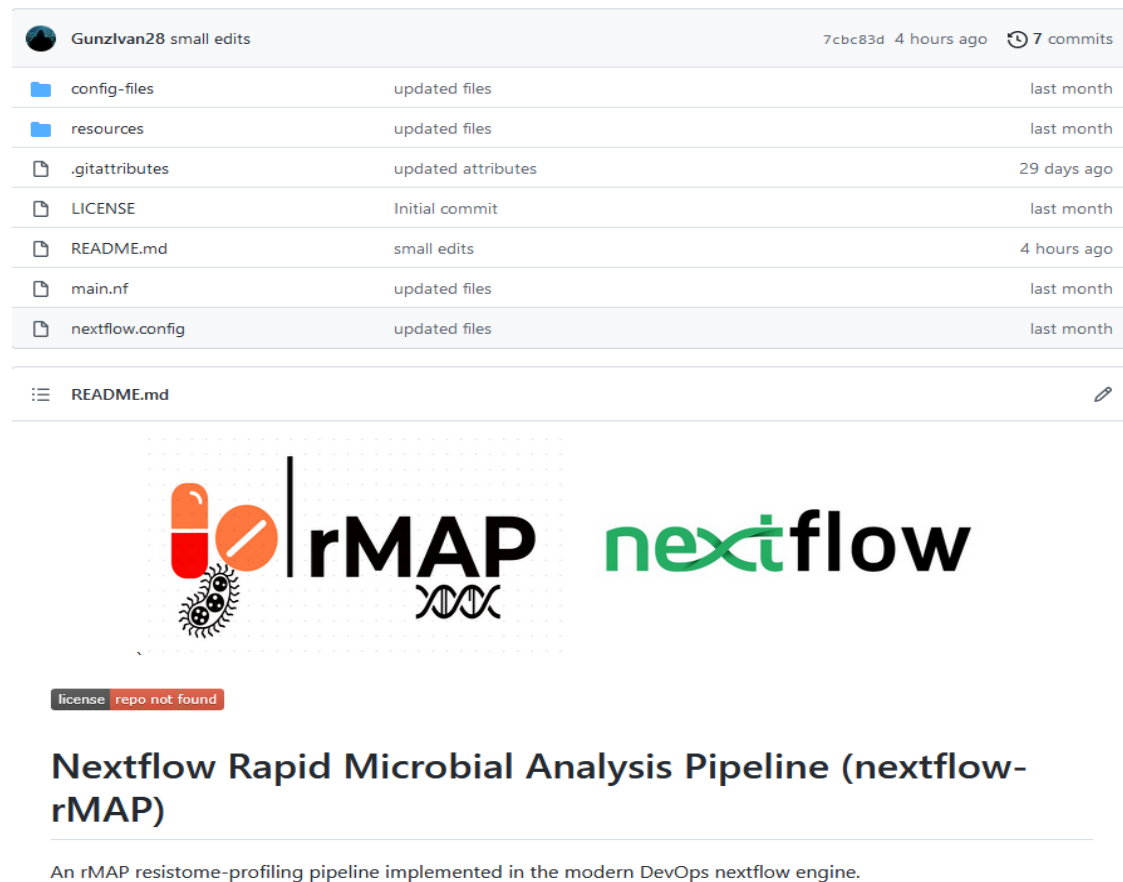
OR:

rMAP -t 8 -r full_genome.gbk -i dir_name -o dir_name -f -a shovill -m -vc -q -p -s -g

- Algorithm parameters

- 24/7 development and user support – https://github.com/GunzIvan28/rMAP/issues

Future prospects

GitHub repository: https://github.com/GunzIvan28/nextflow-rMAP

Implementation facts
- Private – open to only collaborators
- Still under development
- Manuscript – in draft phase
- Opens up the vast possibilities of scalability & reproducibility - overcome the shortcomings of its predecessor
- Designed for - experienced Bioinformatics users continuously interacting with High Performance Compute (HPC) stations, Cloud and Cluster nodes
- High level container integration - Docker, Conda and Singularity
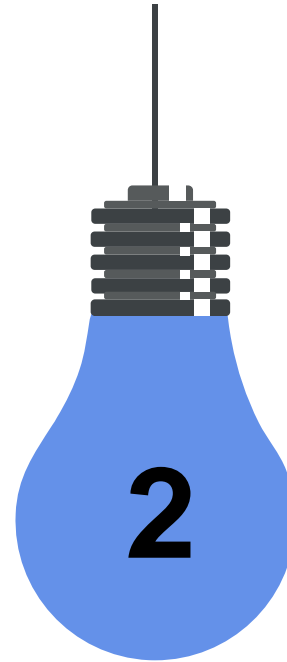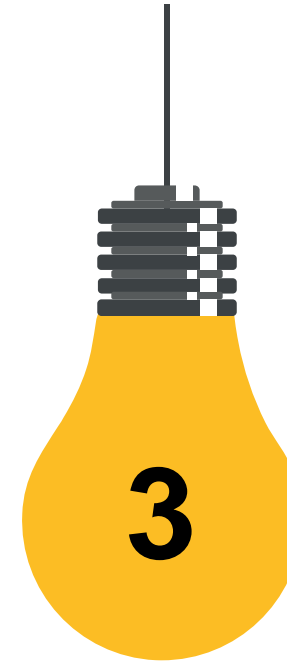
Future prospects
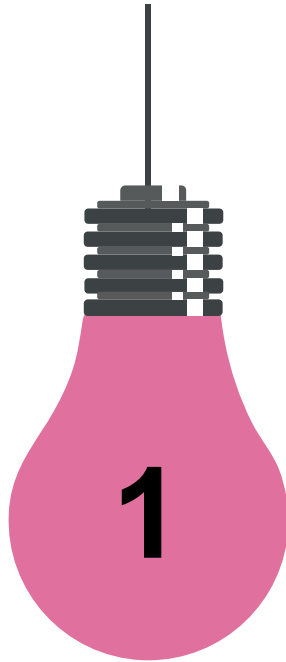
1 — Nationwide AMR surveillance

2 — Validation and implementation of rMAP at Central Public Health Laboratories and National referral hospitals

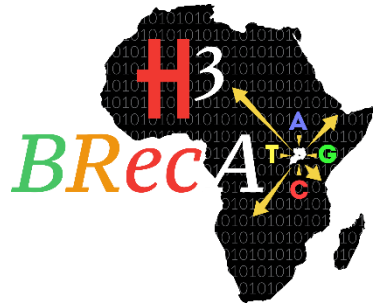3 — Public engagement & AMR policy

Acknowledgements

Special thanks to our funders and collaborators

SAGESA Network