CoP Bioinformatics

Meeting 3 – June 2024

The
Fleming Fund
Regional Grants

6 June 2024

Marco van Zwetselaar

The Problem

Image “Le Petit Prince” is in the public domain in most countries. Exceptions: FR, US.

Bioinformatics CoP

- Mailing list
  - Subscribe: bioinfo-list-join@seqshare.org
- Monthly meet-up
  - Informal, hands-on, brief, “BoF session”
  - 1. Stand-up: who’s doing what (10-15mn)
  - 2. Soapbox: talk, demo, discussion, … (15-20mn)
  - 3. Open mic: questions, tips, ideas (10-15mn)
- More?
  - Face-to-face? “Walk-in” meeting? Topical meetings?

The
Fleming Fund
Regional Grants

Stand-up Round

“Stand-up Meeting”

Original (SCRUM)

- What did you do yesterday?
- What will you do today?
- Is anything blocking you?

Idea: in 1-2mn max(!)

- What are you working on?
- What will you (or would you like to) be working on?
- Is anything blocking you / how can we help?

The Fleming Fund
Regional Grants

Soapbox Time

Some notes & links from
Frontiers Webinar: Pathogen Genomic Surveillance

- Utility of Genomic Surveillance ∝ threat × capacity × response
  → *"A lack in any of these is a lack in overall utility of GS"*
- Success factors: density of sampling & rapidity of sharing
  → *"Surveillance is largely a documentation of past failures"*
  → *"Data sharing fosters trust, integrity and reciprocity"*
- GPAS 'Global Pathogen Surveillance Service': https://gpas.global/
- (TERRA? https://terra.bio)
- Evaluation of ONT-only genomic surveillance (PMCID:9997738)

Nanopore-only assemblies for genomic surveillance of the global priority drug-resistant pathogen, *Klebsiella pneumoniae*

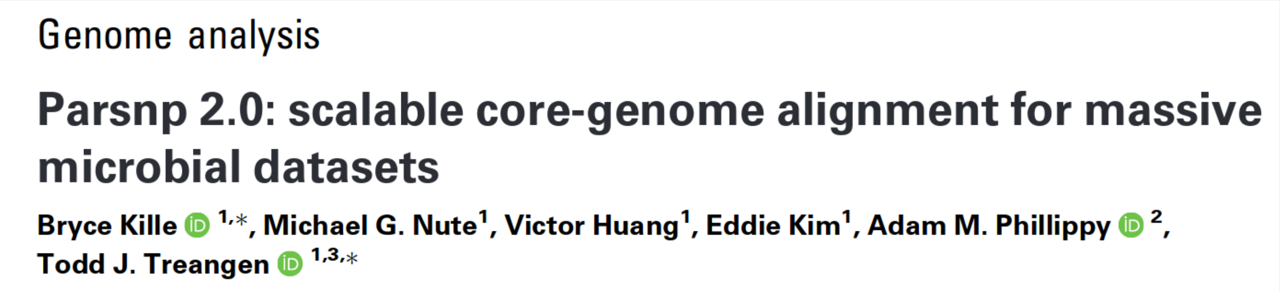
Ebenezer Foster-Nyarko [1], Hugh Cottingham [2], Ryan R Wick [2], Louise M Judd [2], Margaret M C Lam [2], Kelly L Wyres [2], Thomas D Stanton [1], Kara K Tsang [1], Sophia David [3], David M Aanensen [3], Sylvain Brisse [4], Kathryn E Holt [1] [2]

The Final SPAdes (4.0)

- Open Source, moved to GitHub
  - Repo: https://github.com/ablab/spades
  - Docs: https://ablab.github.io/spades/
  - Protocols: https://currentprotocols.onlinelibrary.wiley.com/doi/abs/10.1002/cpbi.102

- 4.0 features
  - Assemble NCBI SRA files
  - BinSPreader: MAG refinement using GFA
  - Run HMMs over a GFA
  - Modes: meta, plasmid, rna, corona, etc

Parsnp 2.0

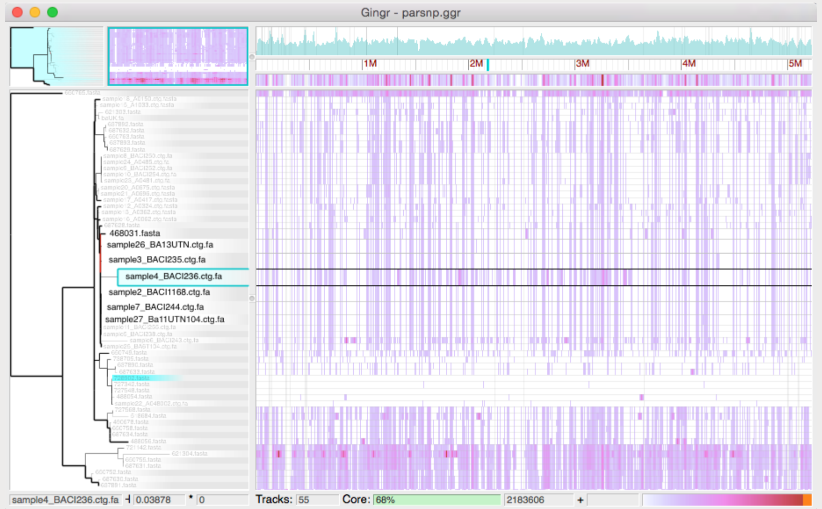


Genome analysis

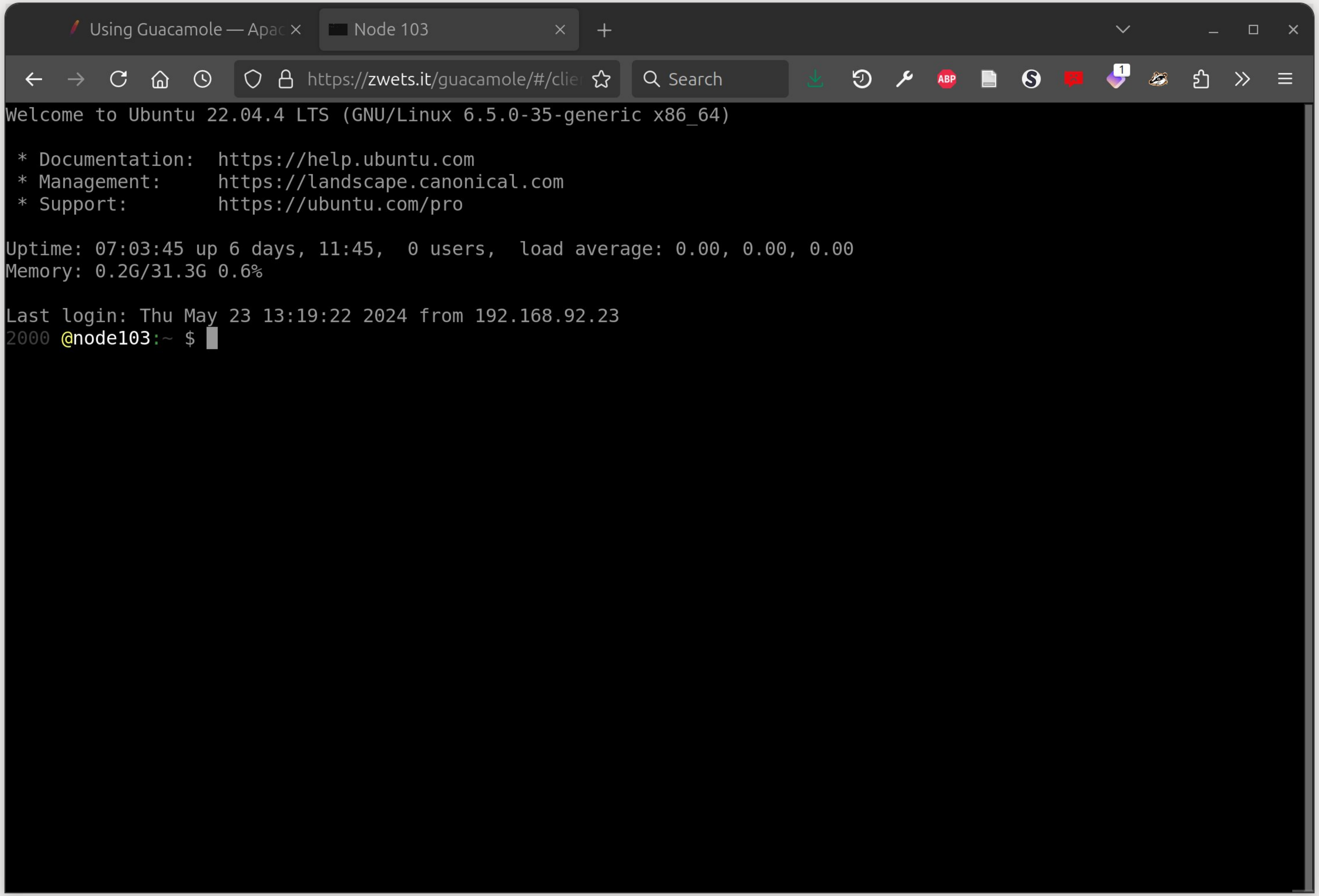
Parsnp 2.0: scalable core-genome alignment for massive microbial datasets

Bryce Kille [1,*], Michael G. Nute[1], Victor Huang[1], Eddie Kim[1], Adam M. Phillippy [2], Todd J. Treangen [1,3,*]

- First update since v1 (2014) (v1)
  - doi: 10.1093/bioinformatics/btae311
  - https://github.com/marbl/parsnp
- Multiple Sequence Alignment of 100s-1000s of genomes
  - Input: assemblies & finished genomes
  - Output: MSA, variant calls, core genome phylogeny
- Gingr for visualisation



Apache Guacamole (1)

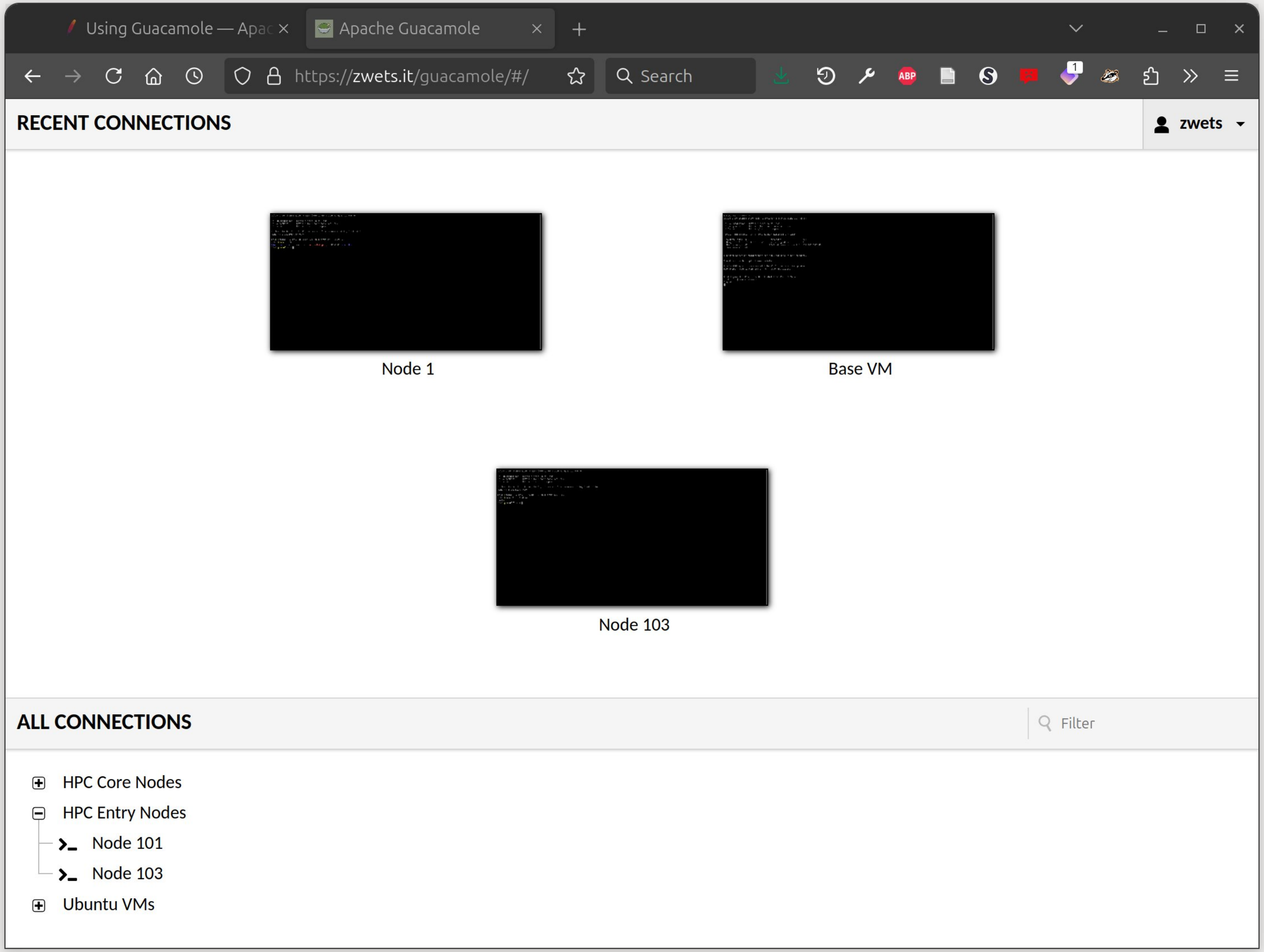


```
Welcome to Ubuntu 22.04.4 LTS (GNU/Linux 6.5.0-35-generic x86_64)

 * Documentation:  https://help.ubuntu.com
 * Management:     https://landscape.canonical.com
 * Support:        https://ubuntu.com/pro

Uptime: 07:03:45 up 6 days, 11:45,  0 users,  load average: 0.00, 0.00, 0.00
Memory: 0.2G/31.3G 0.6%

Last login: Thu May 23 13:19:22 2024 from 192.168.92.23
2000 @node103:~ $ 
```

Client-less Remote Desktop

- In-browser
  - Terminal (SSH)
  - Console (VNC)
  - Desktop
    - Linux (VNC)
    - Windows (RDP)
  - Kubernetes VMs

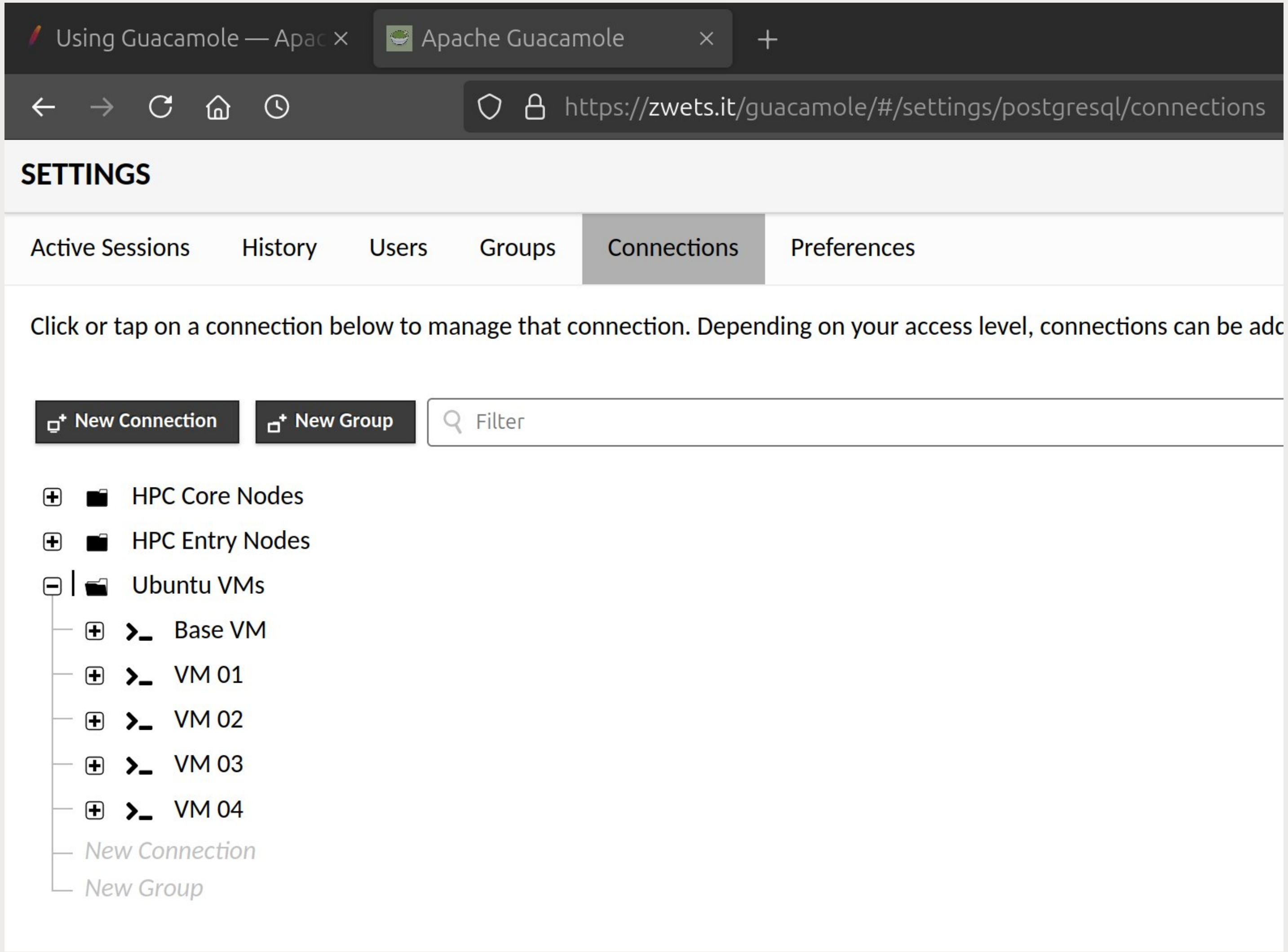
Apache Guacamole (2)



Guacamole

- Tiers
  - Browser
  - Proxy
  - Web-Server
  - Guacamole Server
  - Backends through Plugins
    - SSH, VNC, RDP, K8S

Apache Guacamole (3)



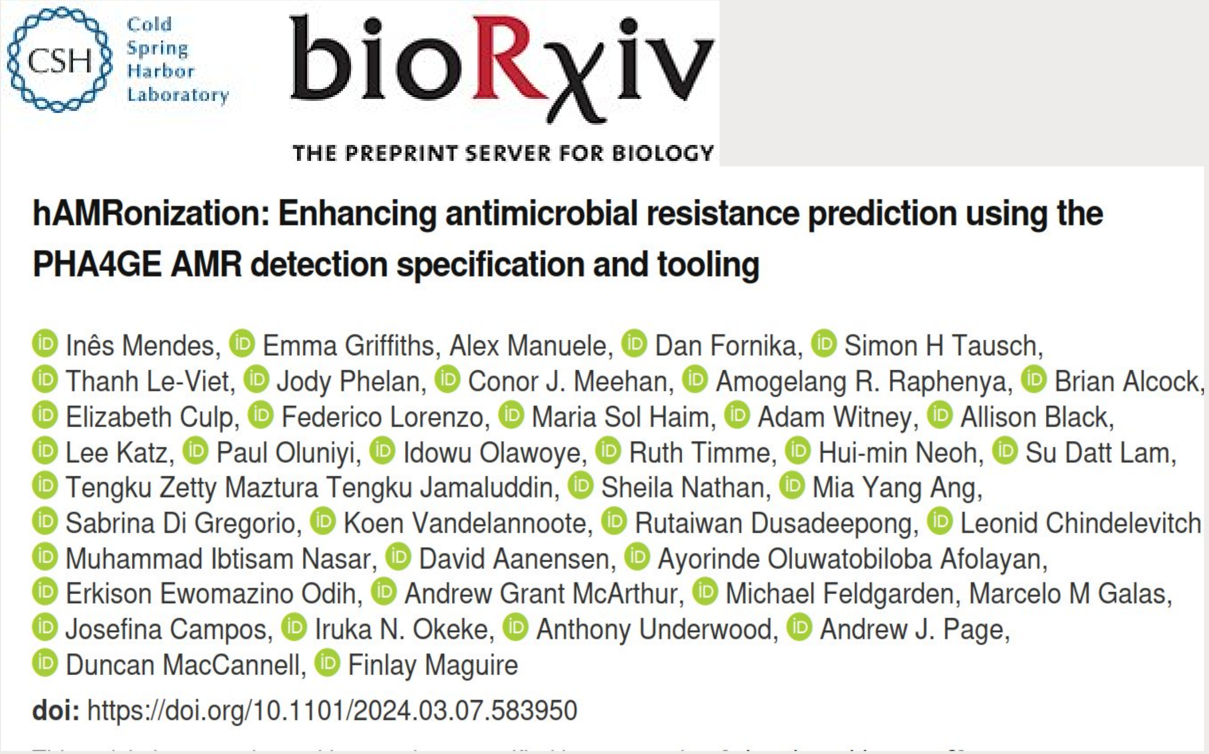
Click or tap on a connection below to manage that connection. Depending on your access level, connections can be ad

Management

- Users
- Groups
- Connections
    - Authorization
    - Connection sharing
    - “On-demand” VMs

hAMRonization

- doi:10.1101/2024.03.07.583950
- PHA4GE = Public Health Alliance for Genomic Epidemiology
- Standardised output specification for bioinformatic AMR detection
- hAMRonization
  - https://github.com/pha4ge/hAMRonization
  - Python package: unifies output from (18) tools to the hAMRonize spec
  - Output: tabular, JSON, interactive HTML

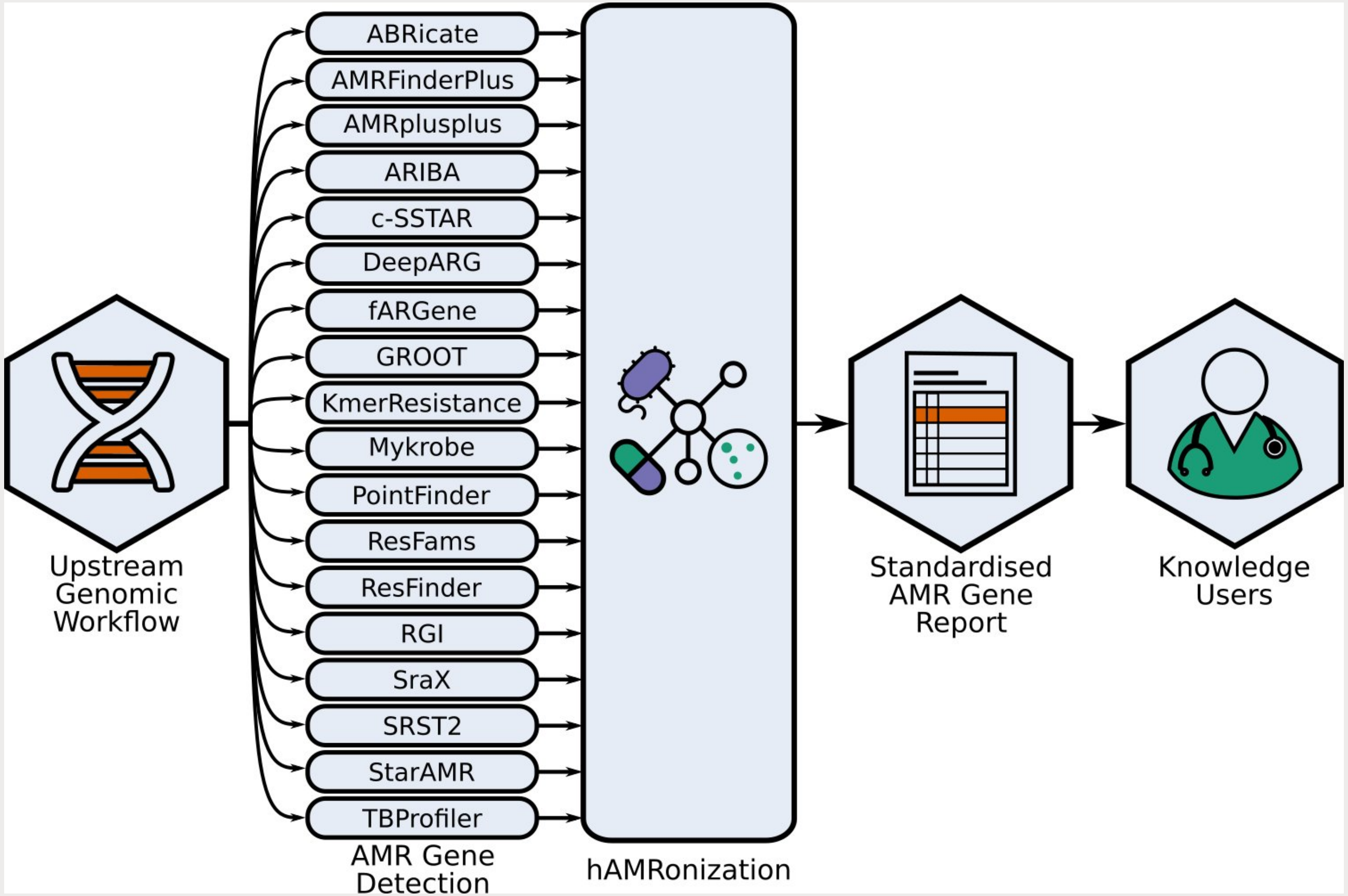


hAMRonization: Enhancing antimicrobial resistance prediction using the PHA4GE AMR detection specification and tooling

Inês Mendes, Emma Griffiths, Alex Manuele, Dan Fornika, Simon H Tausch, Thanh Le-Viet, Jody Phelan, Conor J. Meehan, Amogelang R. Raphenya, Brian Alcock, Elizabeth Culp, Federico Lorenzo, Maria Sol Haim, Adam Witney, Allison Black, Lee Katz, Paul Oluniyi, Idowu Olawoye, Ruth Timme, Hui-min Neoh, Su Datt Lam, Tengku Zetty Maztura Tengku Jamaluddin, Sheila Nathan, Mia Yang Ang, Sabrina Di Gregorio, Koen Vandelannoote, Rutaiwan Dusadeepong, Leonid Chindelevitch Muhammad Ibtisam Nasar, David Aanensen, Ayorinde Oluwatobiloba Afolayan, Erkison Ewomazino Odih, Andrew Grant McArthur, Michael Feldgarden, Marcelo M Galas, Josefina Campos, Iruka N. Okeke, Anthony Underwood, Andrew J. Page, Duncan MacCannell, Finlay Maguire

doi: https://doi.org/10.1101/2024.03.07.583950

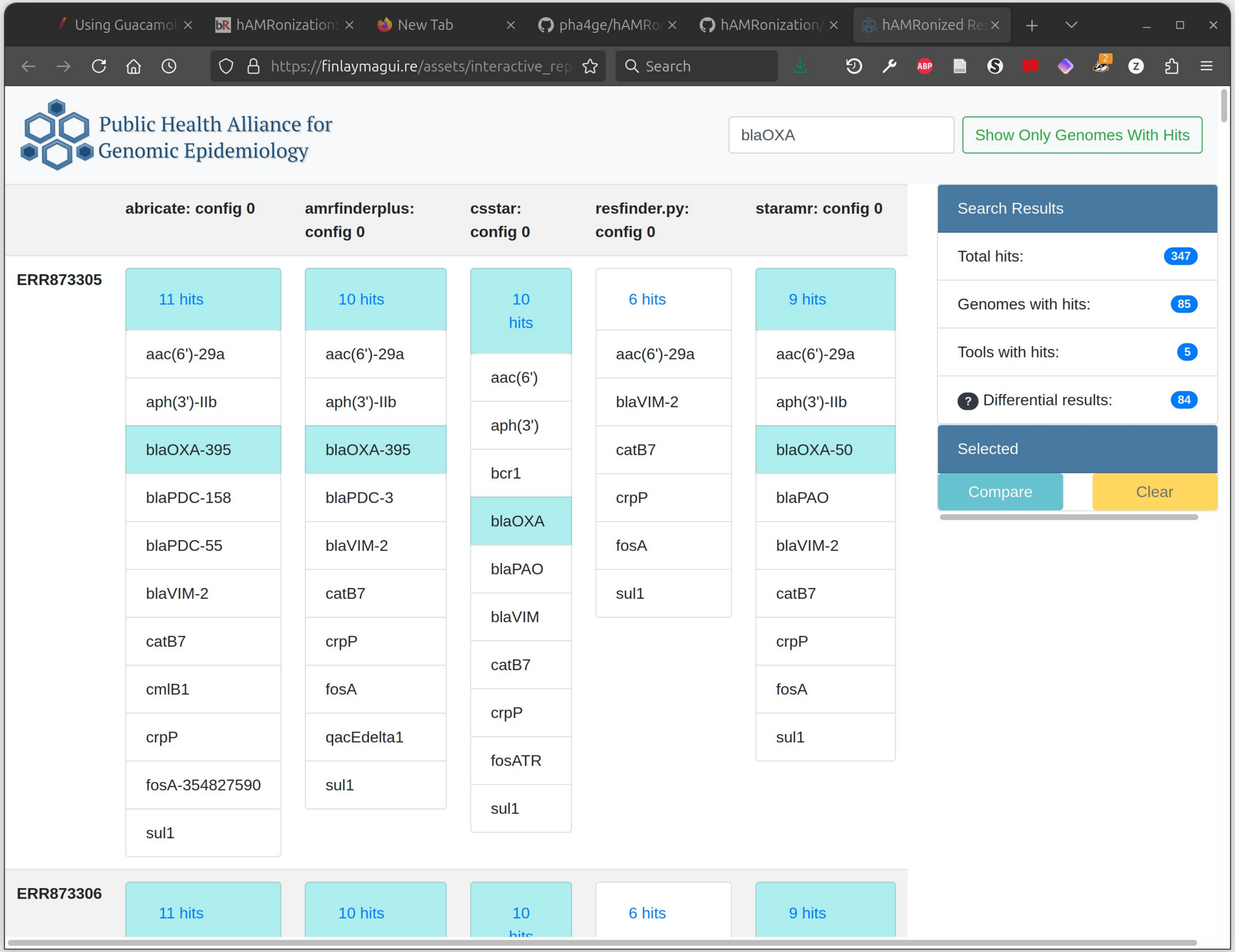
hAMRonization (2)



hAMRonization

- Piloted in 17 PH Labs
- Standardised vocabulary, e.g.
  - “Gene Name”
  - “% Coverage (breadth)”
  - “Drug Class”
  - Provenance (tool, database, versions)
- Harmonised with ontologies
- Conversion from 18 tools that are:
  - In active use
  - Maintained
  - Species-agnostic
  - Open Source
- PoC Workflow (12 tools)

hAMRonization (3)



Public Health Alliance for Genomic Epidemiology

blaOXA

Show Only Genomes With Hits

	abricate: config 0	amrfinderplus: config 0	csstar: config 0	resfinder.py: config 0	staramr: config 0
ERR873305	11 hits	10 hits	10 hits	6 hits	9 hits
	aac(6')-29a	aac(6')-29a	aac(6')	aac(6')-29a	aac(6')-29a
	aph(3')-IIb	aph(3')-IIb	aph(3')	blaVIM-2	aph(3')-IIb
	blaOXA-395	blaOXA-395	bcr1	catB7	blaOXA-50
	blaPDC-158	blaPDC-3	blaOXA	crpP	blaPAO
	blaPDC-55	blaVIM-2	blaPAO	fosA	blaVIM-2
	blaVIM-2	catB7	blaVIM	sul1	catB7
	catB7	crpP	catB7		crpP
	cmlB1	fosA	crpP		fosA
	crpP	qacEdelta1	fosATR		sul1
	fosA-354827590	sul1	sul1		
	sul1				
ERR873306	11 hits	10 hits	10	6 hits	9 hits

Search Results

Total hits: 347

Genomes with hits: 85

Tools with hits: 5

? Differential results: 84

Selected

Compare Clear

- Interactive HTML report
- Samples × tools × genes
- Genes only (tabular report has all)

Open Mic

- Feedback?
- Future themes?
- **Next month:** Niamh

The
Fleming Fund
Regional Grants

Thank you

NOGUCHI MEMORIAL INSTITUTE FOR MEDICAL RESEARCH
UNIVERSITY OF GHANA , LEGON

NATIONAL INSTITUTE FOR COMMUNICABLE DISEASES
Division of the National Health Laboratory Service

INSTITUT PASTEUR
de Dakar

UKaid
from the British people

M M
MOTT MACDONALD

SEQAFRICA

DTU

This programme is being funded by the UK Department of Health and Social Care.
The views expressed do not necessarily reflect the UK Government's official policies.