



Updating and using a Hidden Markov Models-based algorithm to detect Anti-Microbial Resistance sequences in French soils metagenomes

Zéphyrin Enaud, Domitille Jarrige, Olivier Rué, Solène Perrin, Maxime Courcelle, Corinne Cruaud, Patrick Wincker, Claudy Jolivet, Samuel Dequiedt, Antonio Bispo, et al.

► To cite this version:

Zéphyrin Enaud, Domitille Jarrige, Olivier Rué, Solène Perrin, Maxime Courcelle, et al.. Updating and using a Hidden Markov Models-based algorithm to detect Anti-Microbial Resistance sequences in French soils metagenomes. Journées Ouvertes en Biologie, Informatique et Mathématiques 2026 (JOBIM2026), Jun 2026, Strasbourg, France. <10.57745/8Q6R92>. <hal-05681686>

HAL Id: hal-05681686

<https://hal.inrae.fr/hal-05681686v1>

Submitted on 6 Jul 2026

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY-SA 4.0 - Attribution - ShareAlike - International License



Updating a Hidden Markov Model-based algorithm to detect AntiMicrobial Resistance genes in French soil metagenomes

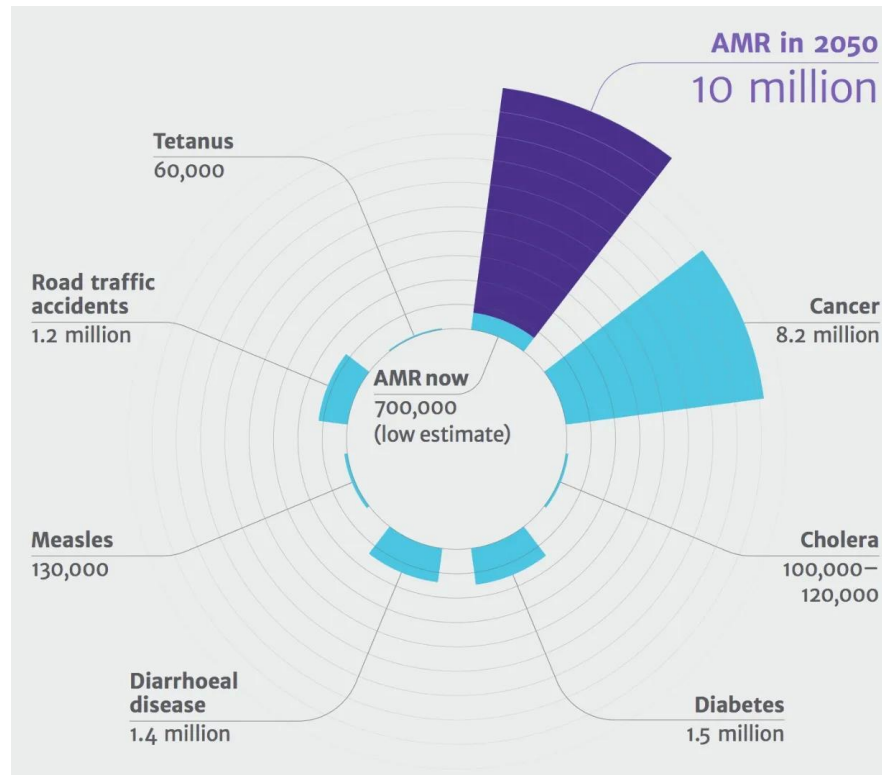
Zéphyrin Enaux, Domitille Jarrige, Olivier Rué, Solène Perrin, Maxime Courcelle, Corinne Cruaud, Patrick Wincker, Claudy Jolivet, Samuel Dequiedt, Antonio Bispo, Lionel Ranjard, Valentin Loux, Sébastien Terrat

JOBIM 2026 Conference

Global context

O'Neil *et al.*, 2016

- Antibiotics (1928) -> one of the greatest medical revolution of the 20th century
- Due to human activity (agriculture, medicine, ...) -> increasing emergence of Antibiotic resistance
- AntiMicrobial resistance (AMR) -> ~ 10 million deaths per year in 2050



Global context

- Antibiotics (1928) -> one of the greatest medical revolution of the 20th century
- Due to human activity (agriculture, medicine, ...) -> increasing emergence of Antibiotic resistance
- AntiMicrobial resistance (AMR) -> ~ 10 million deaths per year in 2050

→ **Need to improve knowledge about AMR diversity and emergence**

Soil : an incredible reservoir of microbial diversity



1g {
10⁹ bacteria, 10⁶ species
10⁶ fungi, 10³ to 10⁴ species



-> Many unknown function in soils microbial communities !

Microflu4AMR project : **anr**[®]

- Identification of new antibiotics and new AMR genes/mechanisms in bacterial communities
- Exploit software specialized in AMR genes detection, to discover potential AMR genes in soils communities

Software for AMR detection

- The team looked for a software :
 - Which could detect AMR in complex environment
 - Adapted to metagenomic data
 - Based on MEGARes, an accurate database, made for large-scale analysis

-> Meta-MARC

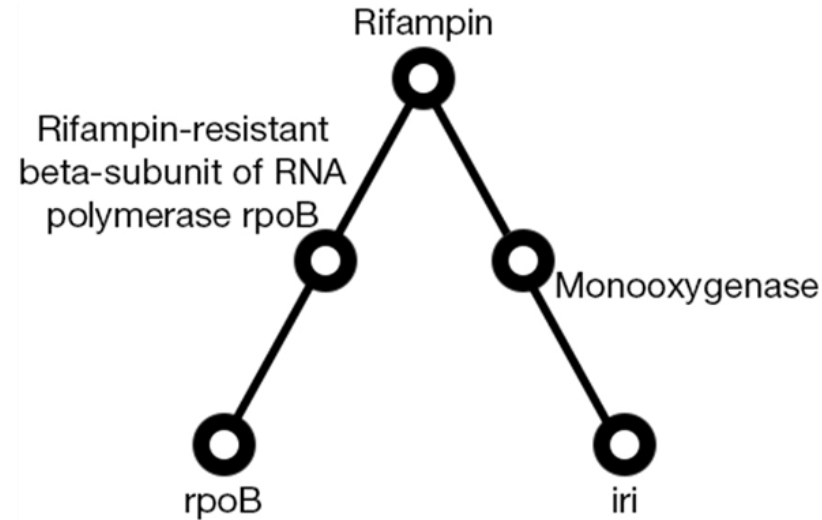
Based on MEGARes database (2016, latest version : 3.0, 2022)

MEGARES

MEGARes database

Lakin *et al.*, 2016

MEGARes annotation structure



Hierarchical and acyclical

Only one parent **per annotation**.

Suited for high-throughput and statistical analysis.

Counts at various rank are straightforward

META-MARC

Meta-MARC

Lakin et al., 2019

- Based on Hidden Markov Models (HMMs) -> Better sensitivity than alignment tools
- Works with contigs
- Initially built on MEGARes v1.0 -> outdated (no detection of metal resistance, ...)

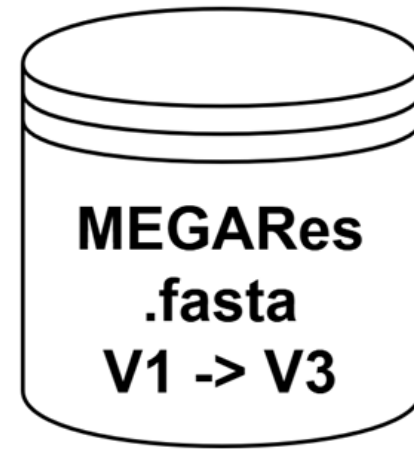
-> Complete Update of Meta-MARC with MEGARes 3.0 sequences

-> Reproduce, as close as possible, Meta-MARC authors methodology using MEGARes 3.0 data (sequence classifications, annotations, ...)

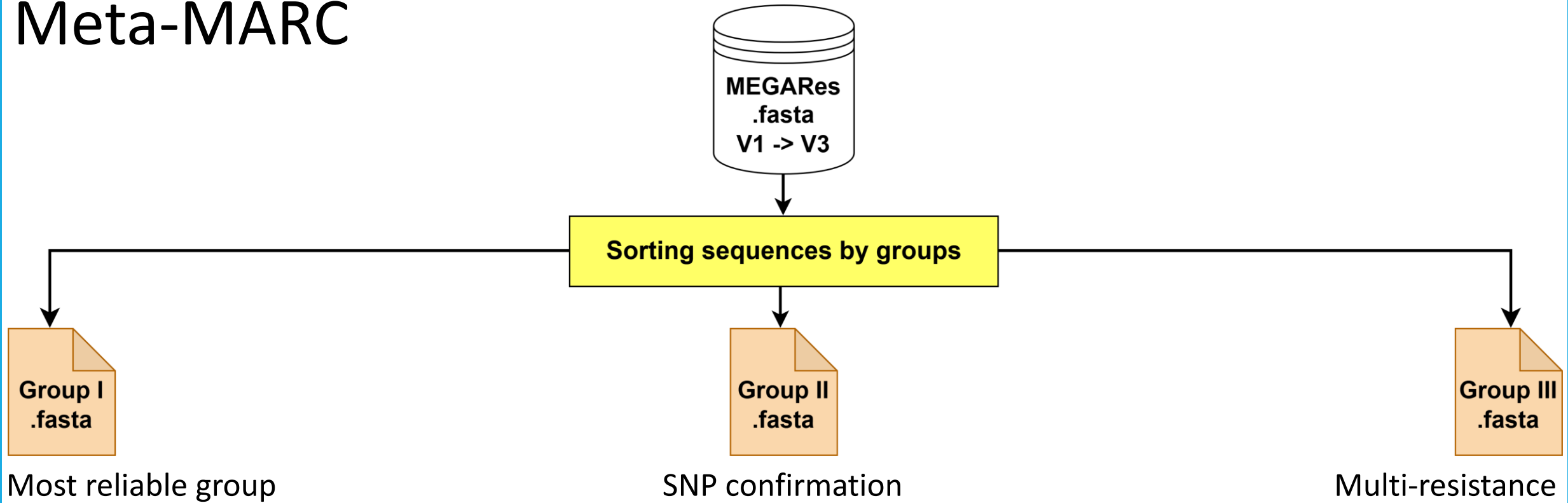
MEGARes version	1.0	3.0
Content detail		
Number of resistance classes	22	49
Total number of sequences	32 888	102 733
Total number of HMMs	1067	1294

WORKFLOW

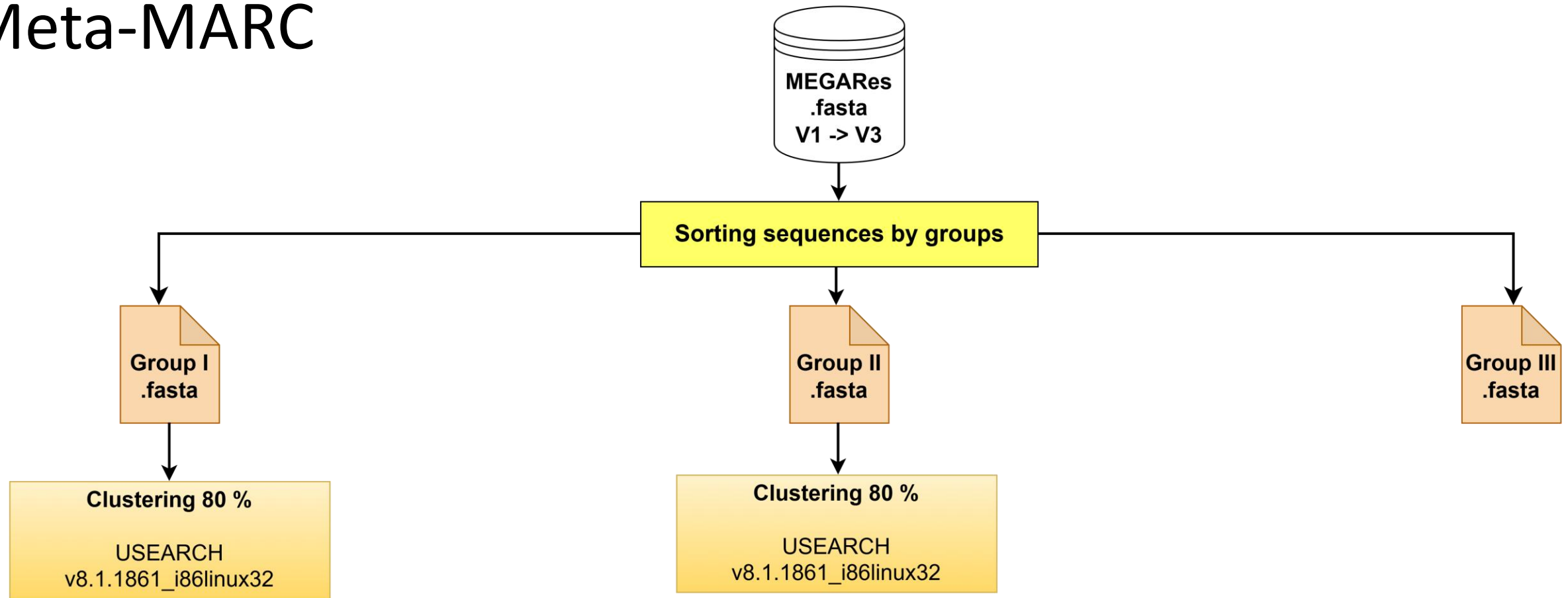
Meta-MARC



Meta-MARC



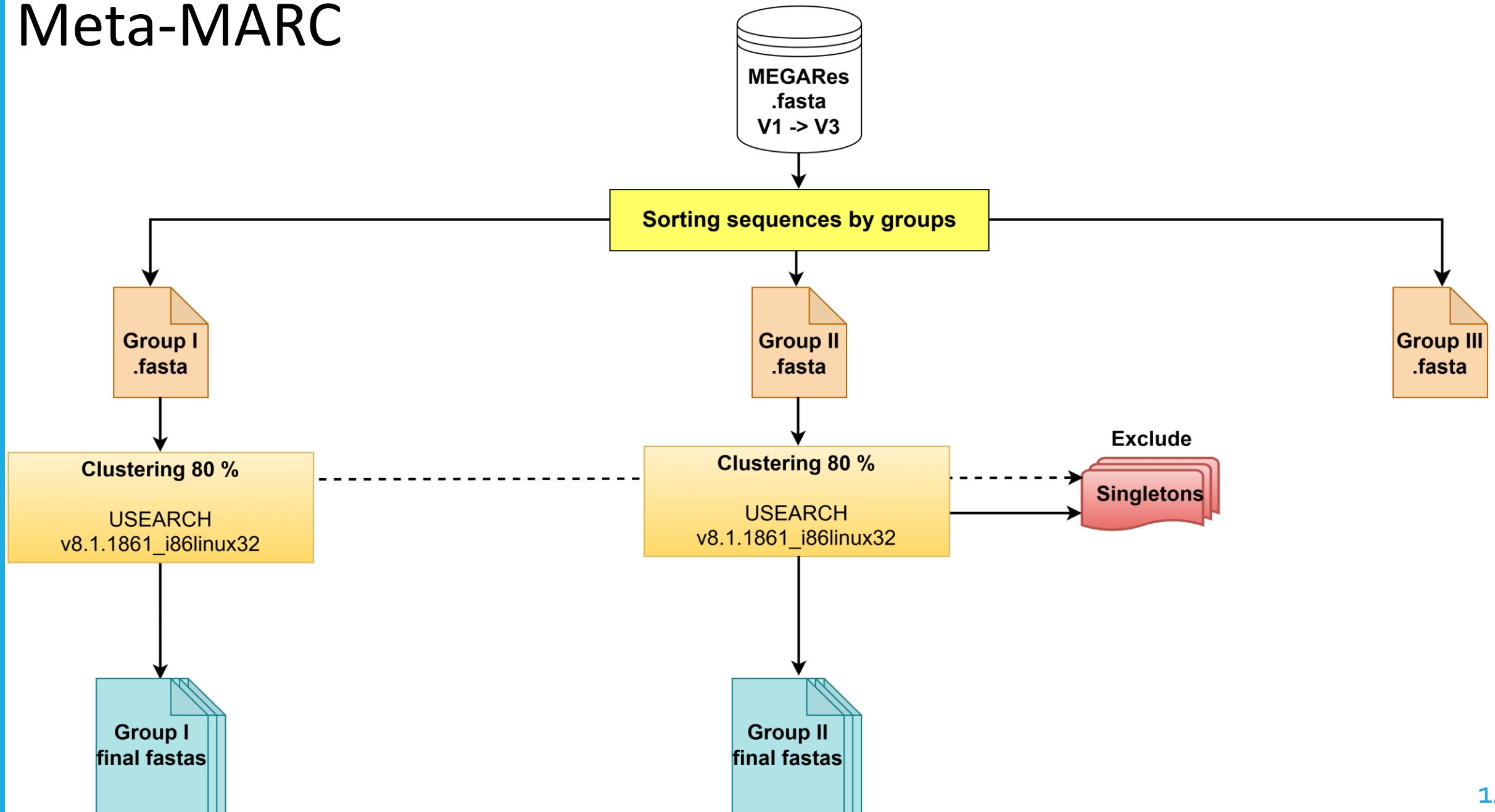
Meta-MARC



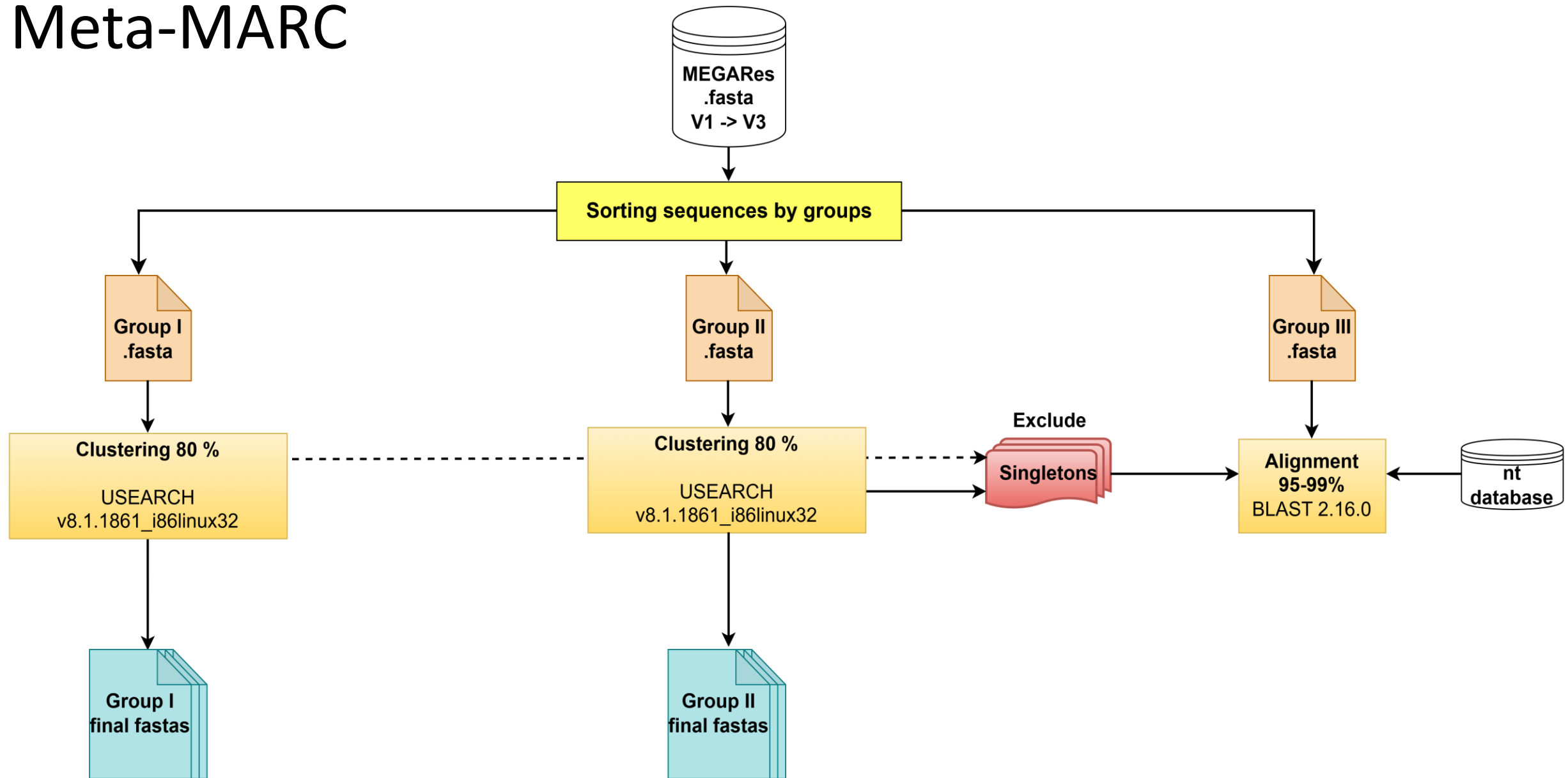
Clustering :

Generation of sequence clusters (alignment files). All sequences with identity $\geq 80\%$ are grouped together.

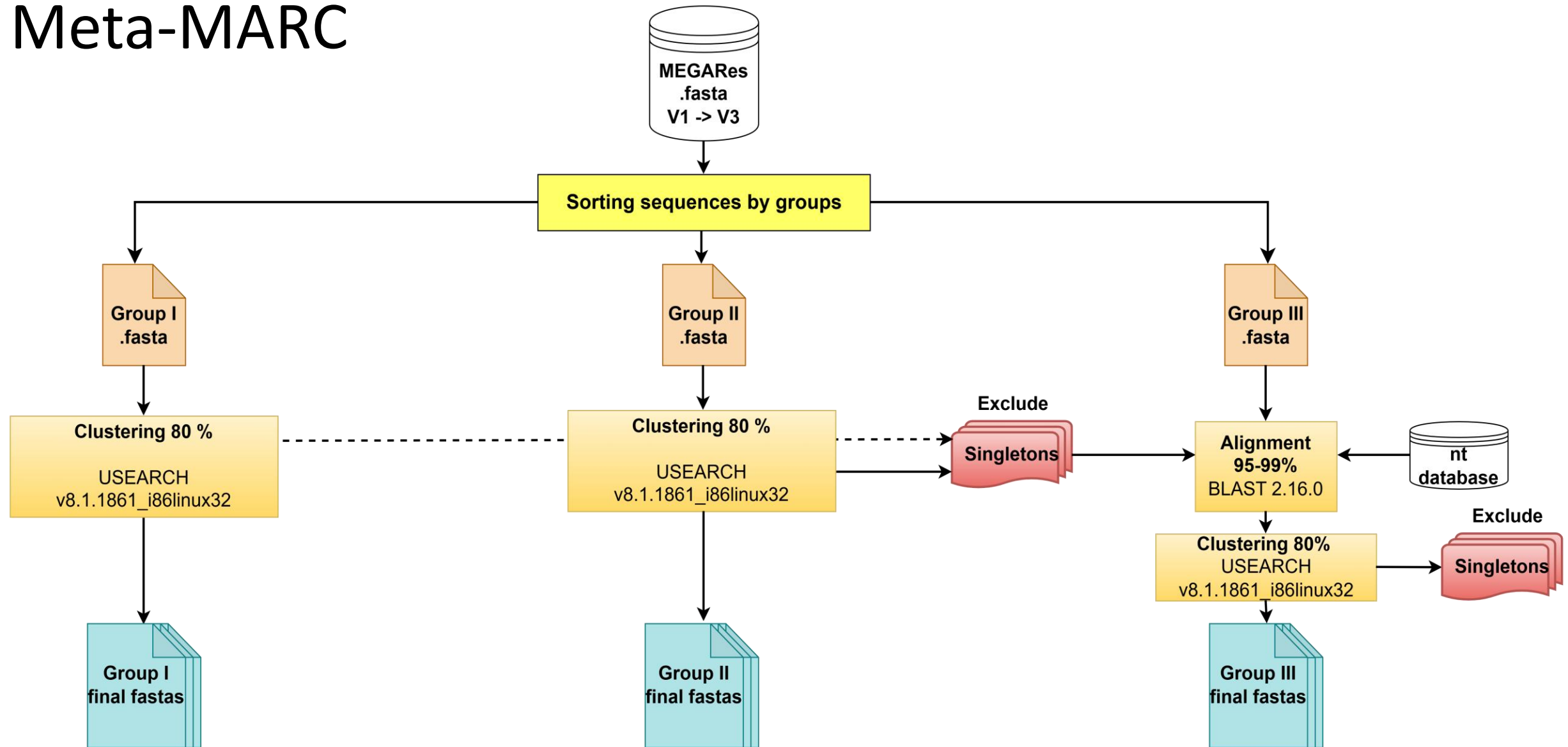
Meta-MARC



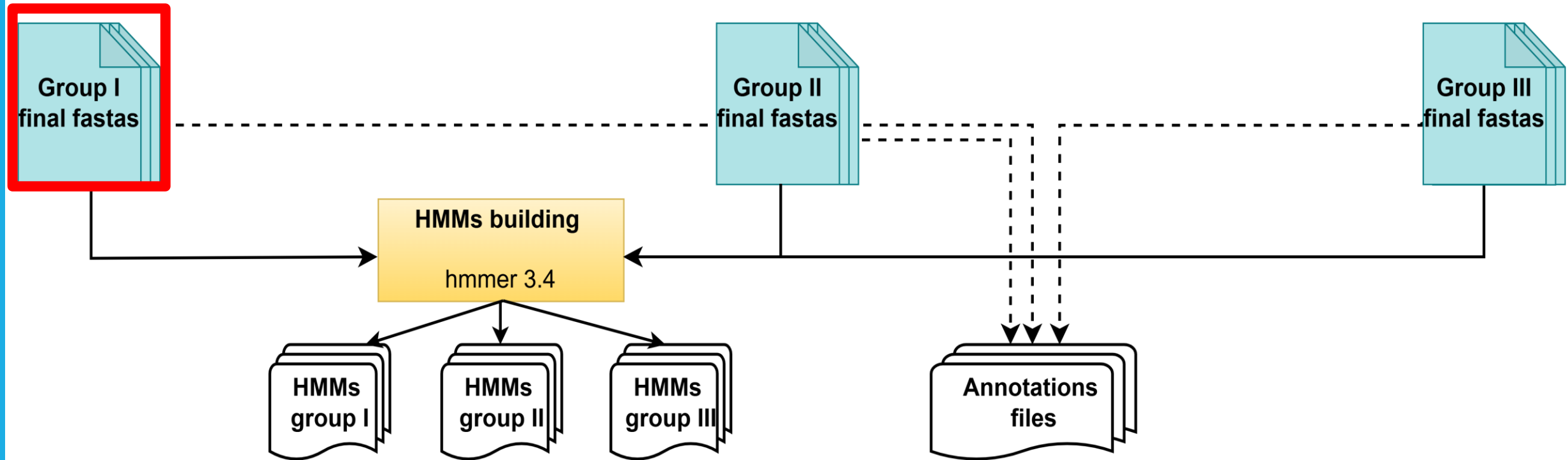
Meta-MARC



Meta-MARC

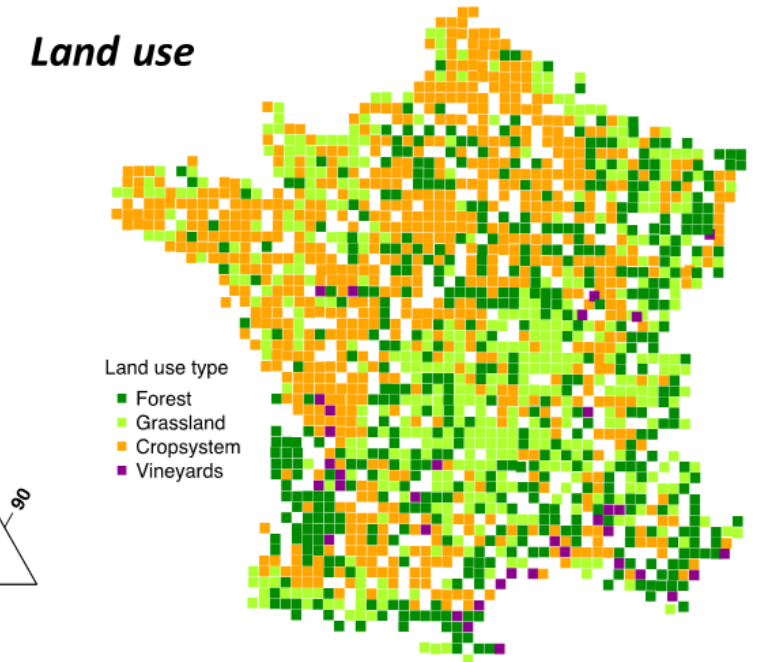
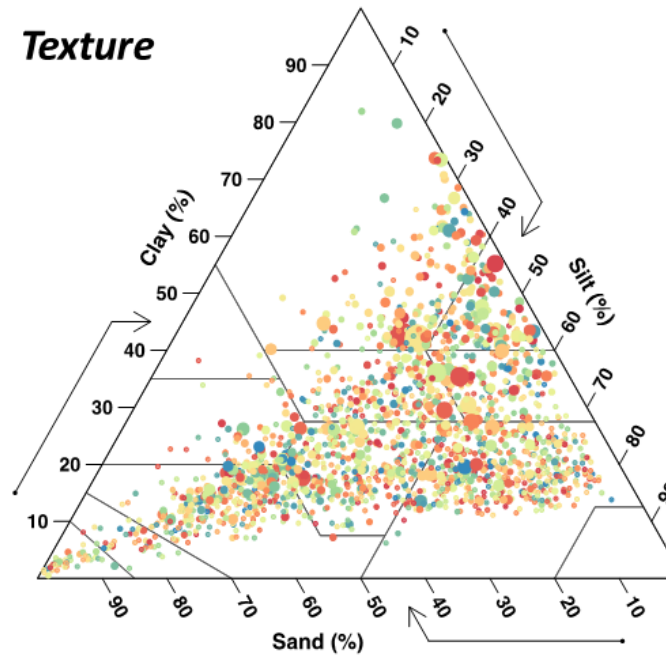
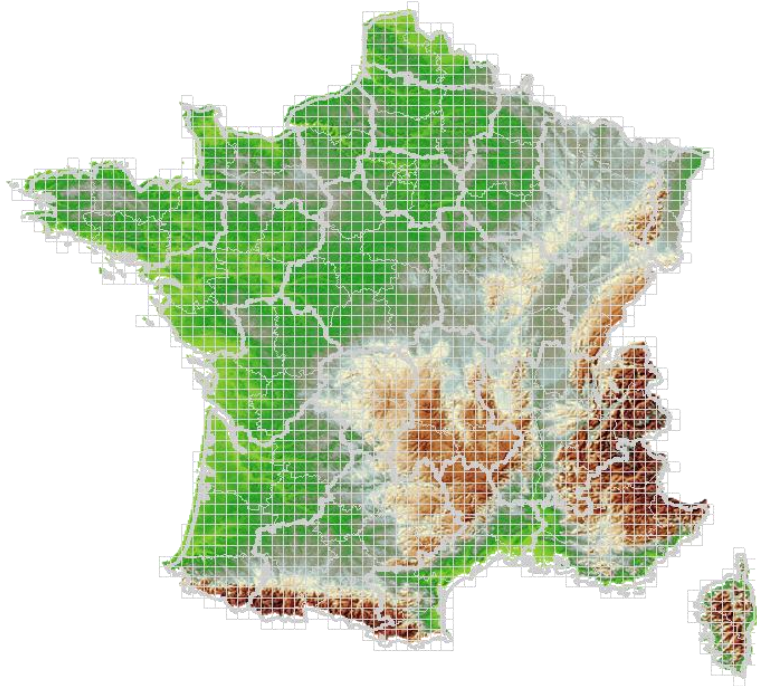


Meta-MARC

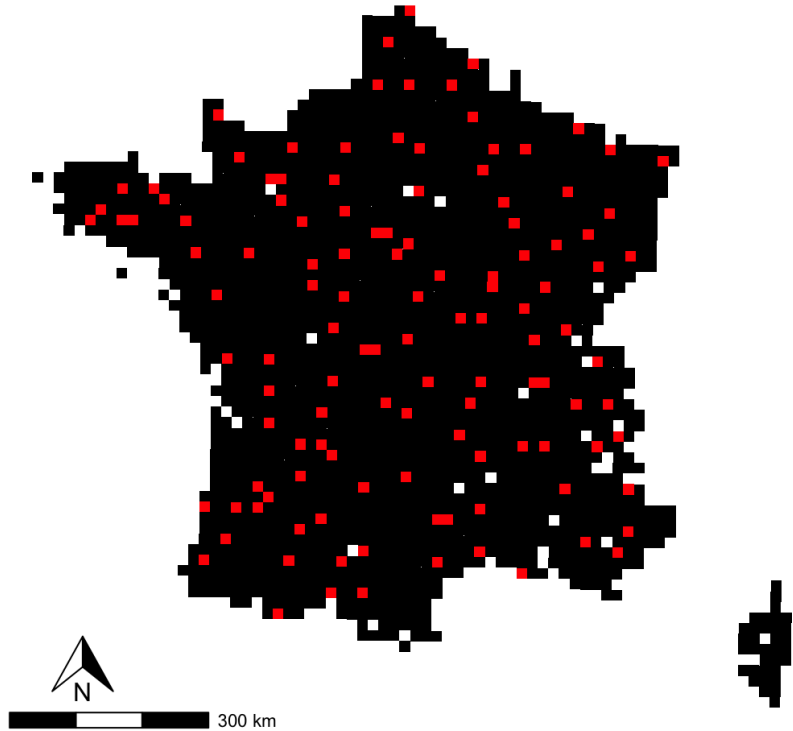


French Soil Quality Monitoring Network (RMQS)

- 2200 sites on a 16 x 16 km grid, representing the diversity of French soils



French Soil Quality Monitoring Network (RMQS)

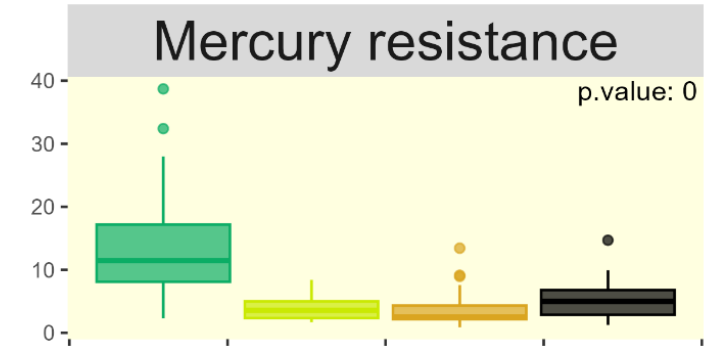
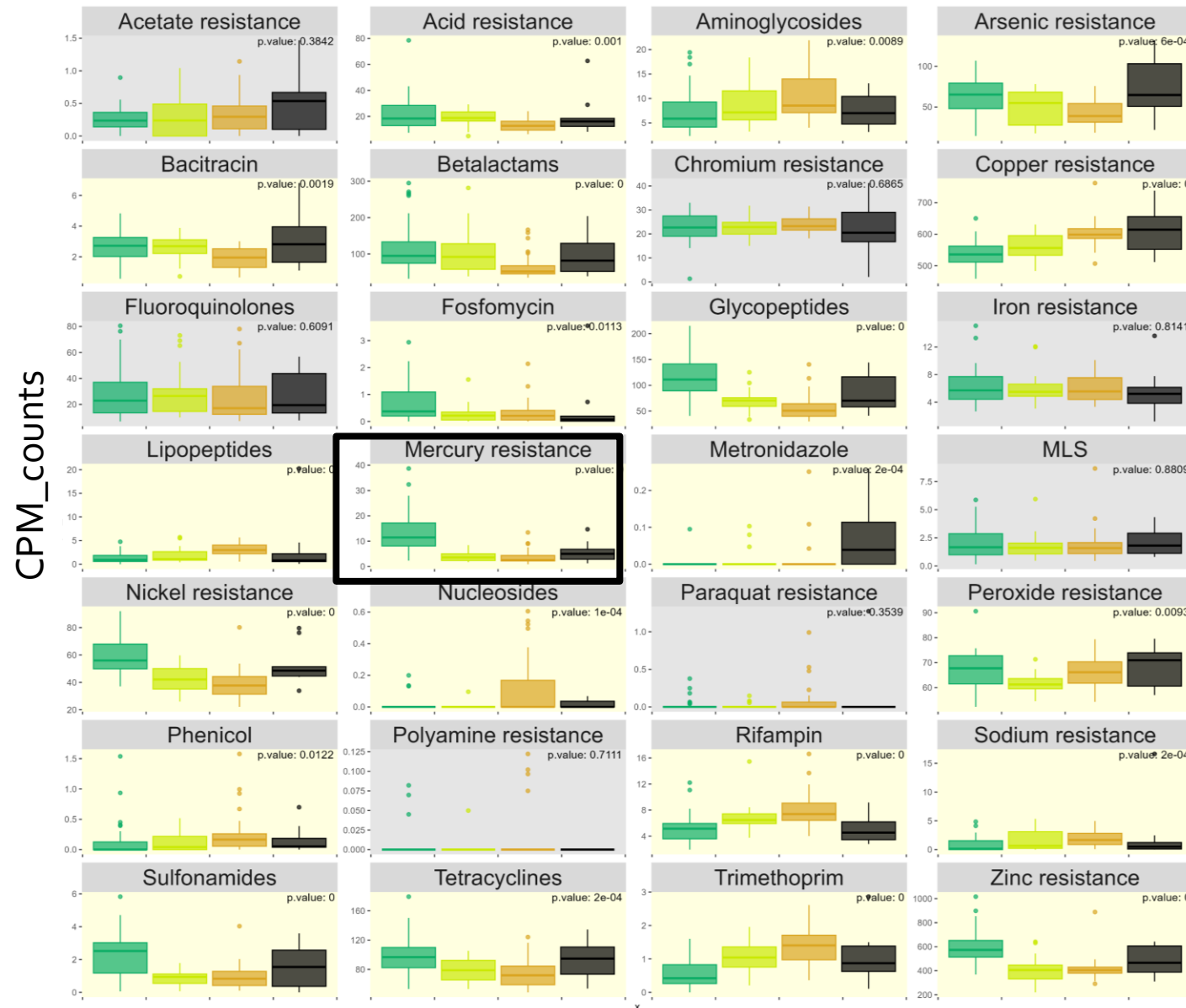


200 samples

Second campaign of the RMQS (2016-2027)

- Associated metadata : more than 60 environmental parameters (pH, C/N ratio, composition, ...)
- Analyzed by global metagenomics (Reads and assembly available for each sample)

AMR classes in land use



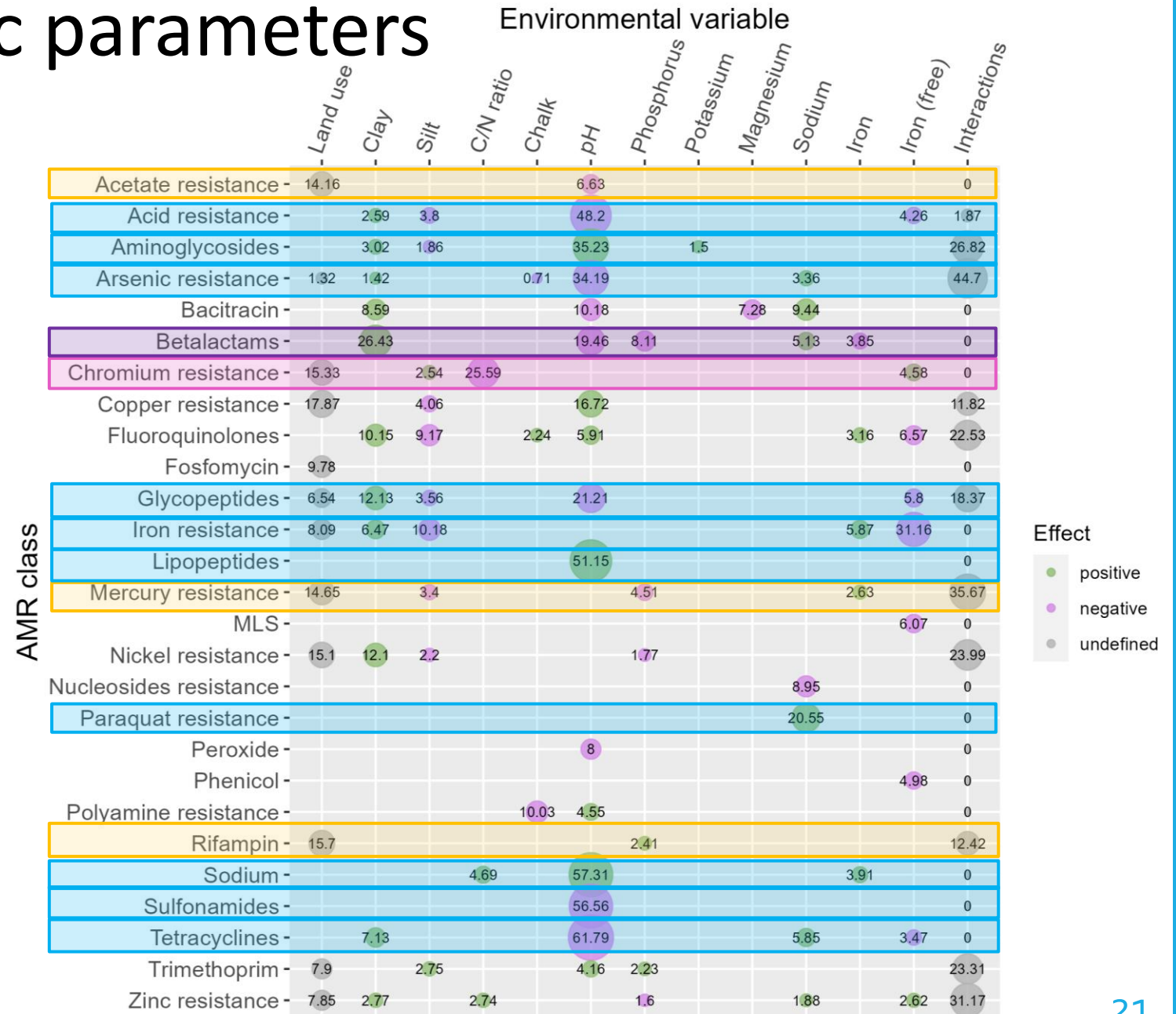
Example:
Mercury is more abundant in forest soil

(because of plant-mediated Hg deposition and vegetation cover that limit re-emission to the atmosphere)

Contribution of abiotic parameters

Environmental parameters influence AMR abundance in French soils.

- AMR class influenced mainly by land use
- AMR class influenced mainly by soil chemistry
- AMR class influenced mainly by soil nutrients
- AMR class influenced mainly by soil texture



First conclusions and perspectives

Software :

- Meta-MARC has been updated and will be used to predict potential AMR genes within more AMR families
- Current update of Meta-MARC with MEGARes v3.0 is already available (DOI : <https://doi.org/10.57745/8Q6R92>)
- Update workflow is publicly available, and could be re-used to upgrade Meta-MARC with any future MEGARes version (DOI : <https://doi.org/10.57745/8Q6R92>)

Perspectives :

- Link observed results with the literature or current knowledge
- Publication of our results
- Explore AMR predictions from group III of Meta-MARC v3.0

Thank you for your attention !



Samuel Dequiedt

Maxime Courcelle

Corinne Cruaud

Alexandra Baron

Domitille Jarrige

Valentin Loux

Patrick Wincker

Aysha Chowdhury

Solène Perrin

Olivier Rué

Andrea Franconi

Lionel Ranjard

Diab-Adams Freyfer

Sébastien Terrat

Jean-Baptiste Gieu

Andrew Griffiths

Paul Rainey

Stéphanie Toetsch



Looking for new opportunities !



`zephyrin.enaux@inrae.fr`

Bioinformatics engineer

Available from 31/10/2026

SUPPLEMENTARIES

Meta-MARC

How Meta-MARC works ?

- User give to Meta-MARC his fasta file, containing contigs of interest,
- In command line, he has to specified against which Meta-MARC group (I, II or III) he wants the AMR prediction to be realized (only one group is available by command line),
- Meta-MARC predicts potential AMR sequences in each contig, using the HMMs of the given group,
- In fact, only AMR classes uses to build the group's HMMs will be used.

mmarc -i contigs.fasta -o output_dir -f output_filename -l group_number -t num_threads

Example : if user write « group I » in command line, he can not predict potential multi-resistance sequences, because all HMMs based on Multi-resistance sequences are in group III.

AMR classes in microbial habitats

Karimi *et al.*, 2020

