



**HAL**  
open science

## Genome assembly and structural variations across geographical isolates of *Meloidogyne enterolobii*

Marine Poulet, Etienne Danchin, Corinne Rancurel, Marine Sallaberry, Céline Lopez-Roques, Joanna Lledo, Hemanth Gopal, Sebastian Kiewnick

### ► To cite this version:

Marine Poulet, Etienne Danchin, Corinne Rancurel, Marine Sallaberry, Céline Lopez-Roques, et al.. Genome assembly and structural variations across geographical isolates of *Meloidogyne enterolobii*. Environmental and Agronomical Genomics Symposium - EAGS 2024, France Génomique; GDR Génomique environnementale, Feb 2024, Toulouse, France. hal-04715094

**HAL Id: hal-04715094**

**<https://hal.science/hal-04715094v1>**

Submitted on 30 Sep 2024

**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 Attribution - NonCommercial - NoDerivatives 4.0 International License

# Genome assembly and structural variations across geographical isolates of *Meloidogyne enterolobii*

---

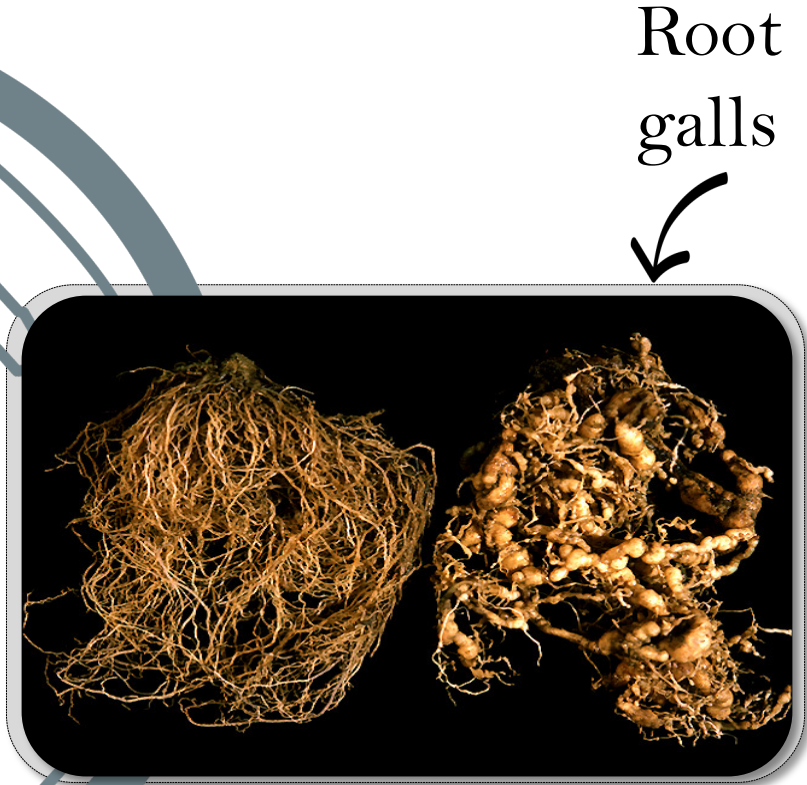
Marine Poulet

Genomic and Adaptive Molecular Evolution team  
Sophia Agrobiotech Institut

# Root-knot nematode genus : *Meloidogyne*

- M. floridensis*
- M. incognita*
- M. arenaria*
- M. javanica*
- M. lopezi*
- M. paranaensis*
- M. konaensis*
- M. luci*
- M. oryzae*
- M. ethiopica*
- M. inornata*
- M. enterolobii*
- M. graminicola*
- M. exigua*
- M. hapla*

- Round worms
- Obligate parasites
- More than 100 described species



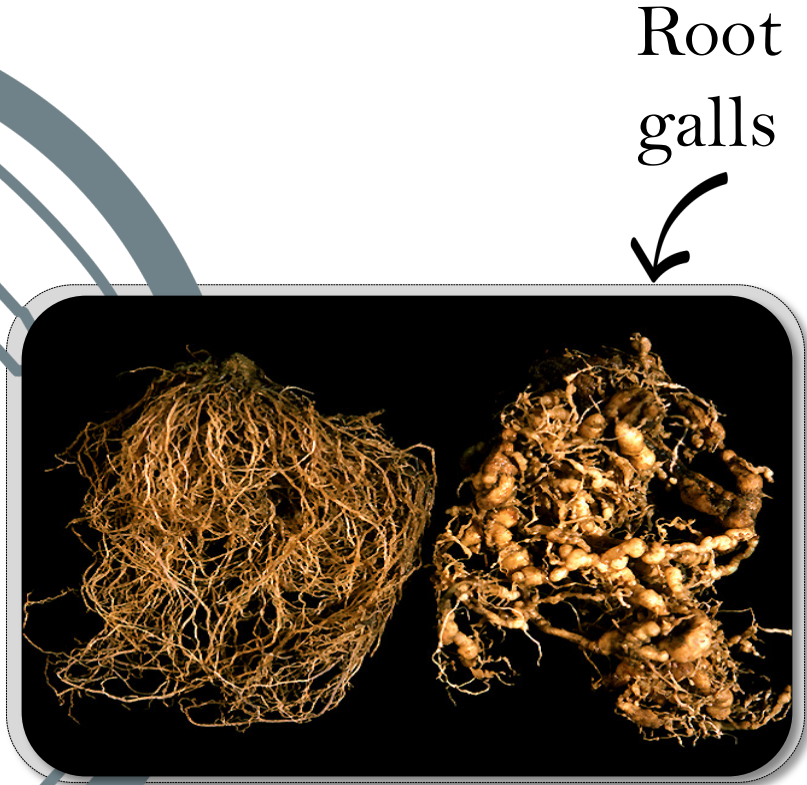
Root galls

# Root-knot nematode genus : *Meloidogyne*

- M. floridensis*
- M. incognita*
- M. arenaria*
- M. javanica*
- M. lopezi*
- M. paranaensis*
- M. konaensis*
- M. luci*
- M. oryzae*
- M. ethiopica*
- M. inornata*
- M. enterolobii*
- M. graminicola*
- M. exigua*
- M. hapla*



- Round worms
- Obligate parasites
- More than 100 described species



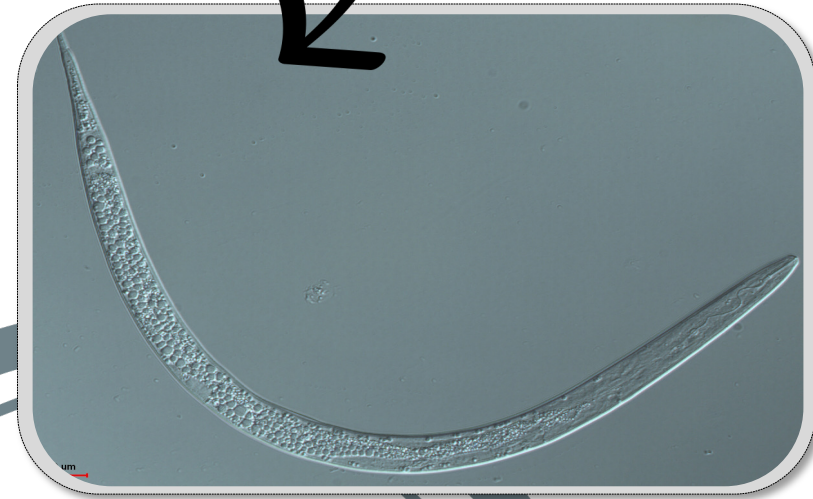


# The nematode *Meloidogyne enterolobii*

- Polyphagous species
- Tropical/Subtropical species
- Insufficient control methods
- Reproduce in plant-resistant genes (containing the Mi-1 gene)

**Impact the global food supply**

 European Quarantine pest

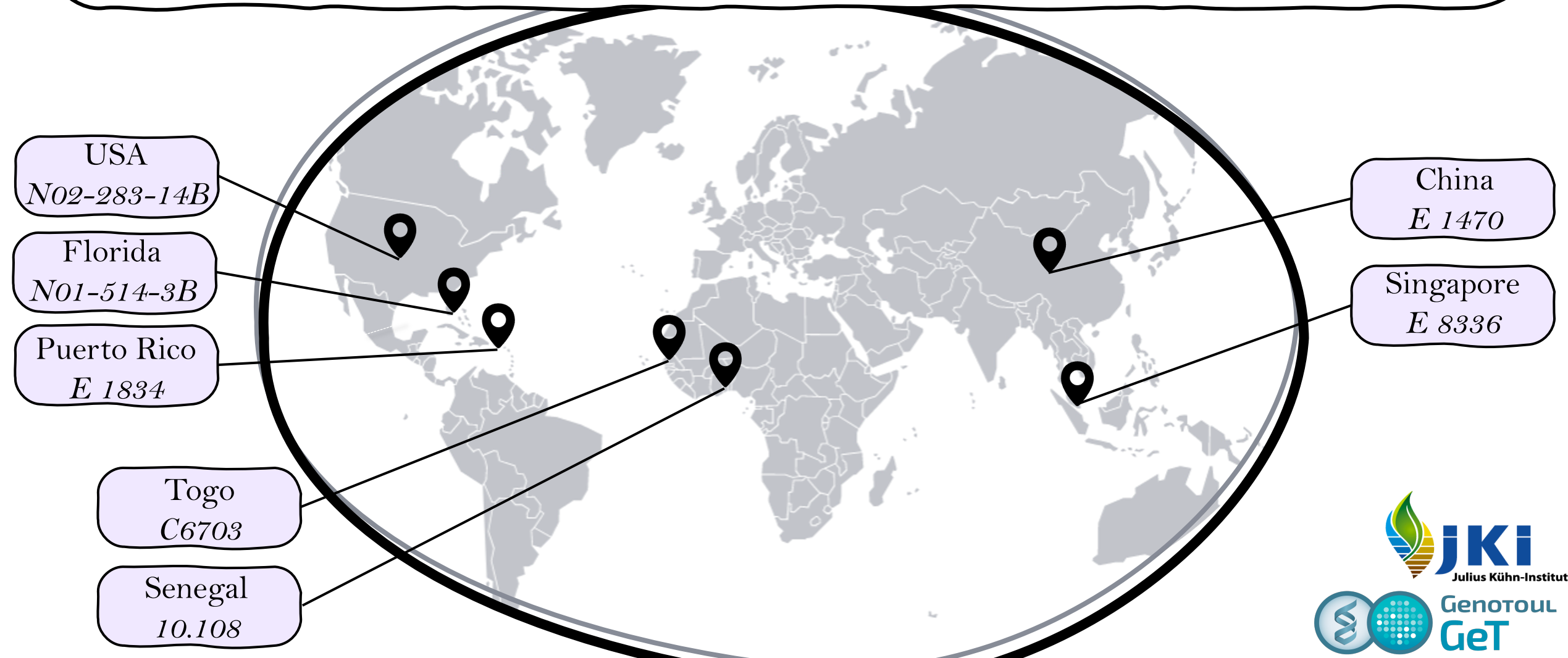


Root galls



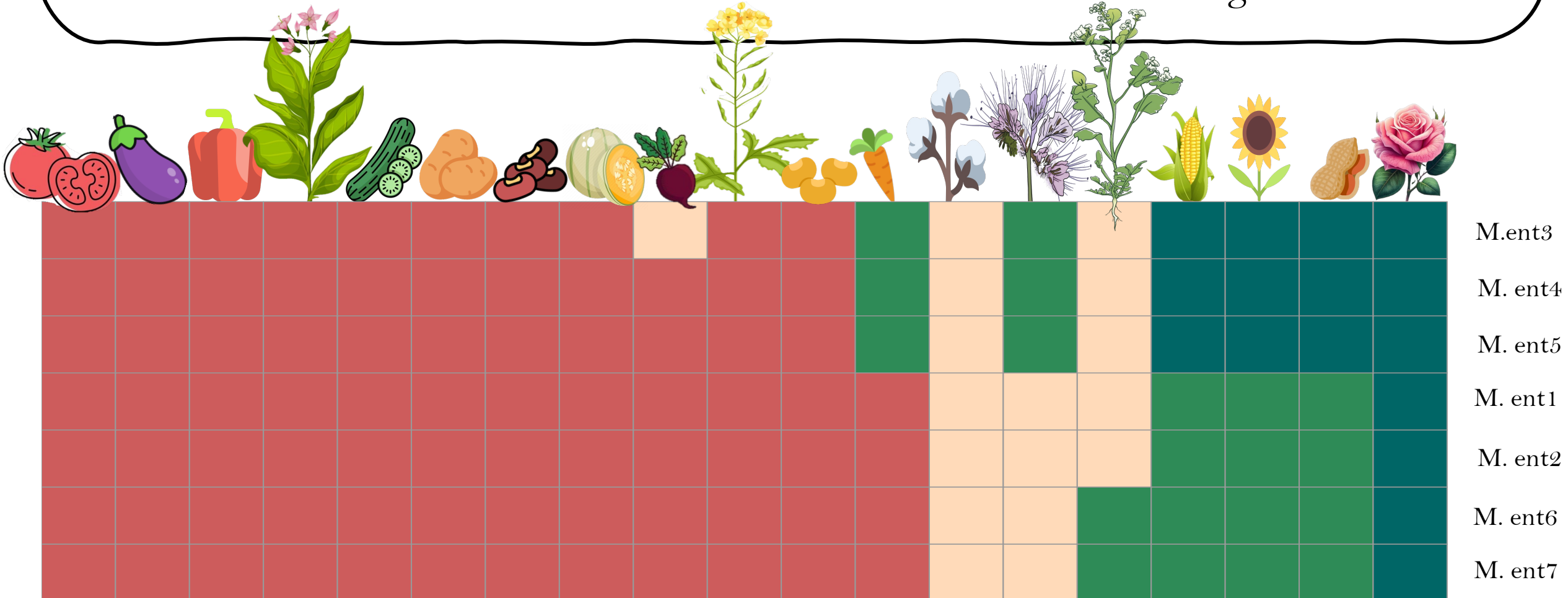
# Exploration of the genomic plasticity of *M. enterolobii*

We have produced high-fidelity genome data (PacBio Hifi) for 7 different geographical isolates of *M. enterolobii*. Each isolate differs in its host range.



# Exploration of the genomic plasticity of *M.enterolobii*

We have produced high-fidelity genome data (PacBio Hifi) for 7 different geographical isolates of *M. enterolobii*. Each isolate differs in its host range.

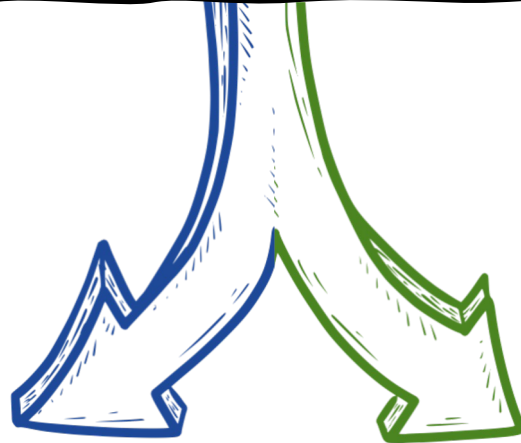
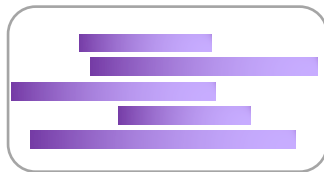


Good host     Non host

# Exploration of the genomic plasticity of *M. enterolobii*

We have produced high-fidelity genome data (PacBio Hifi) for 7 different geographical isolates of *M. enterolobii*. Each isolate differs in its host range.

## Reference genome assembly



## Workflow strategies for variant calling

Compatible host



Geographical distribution

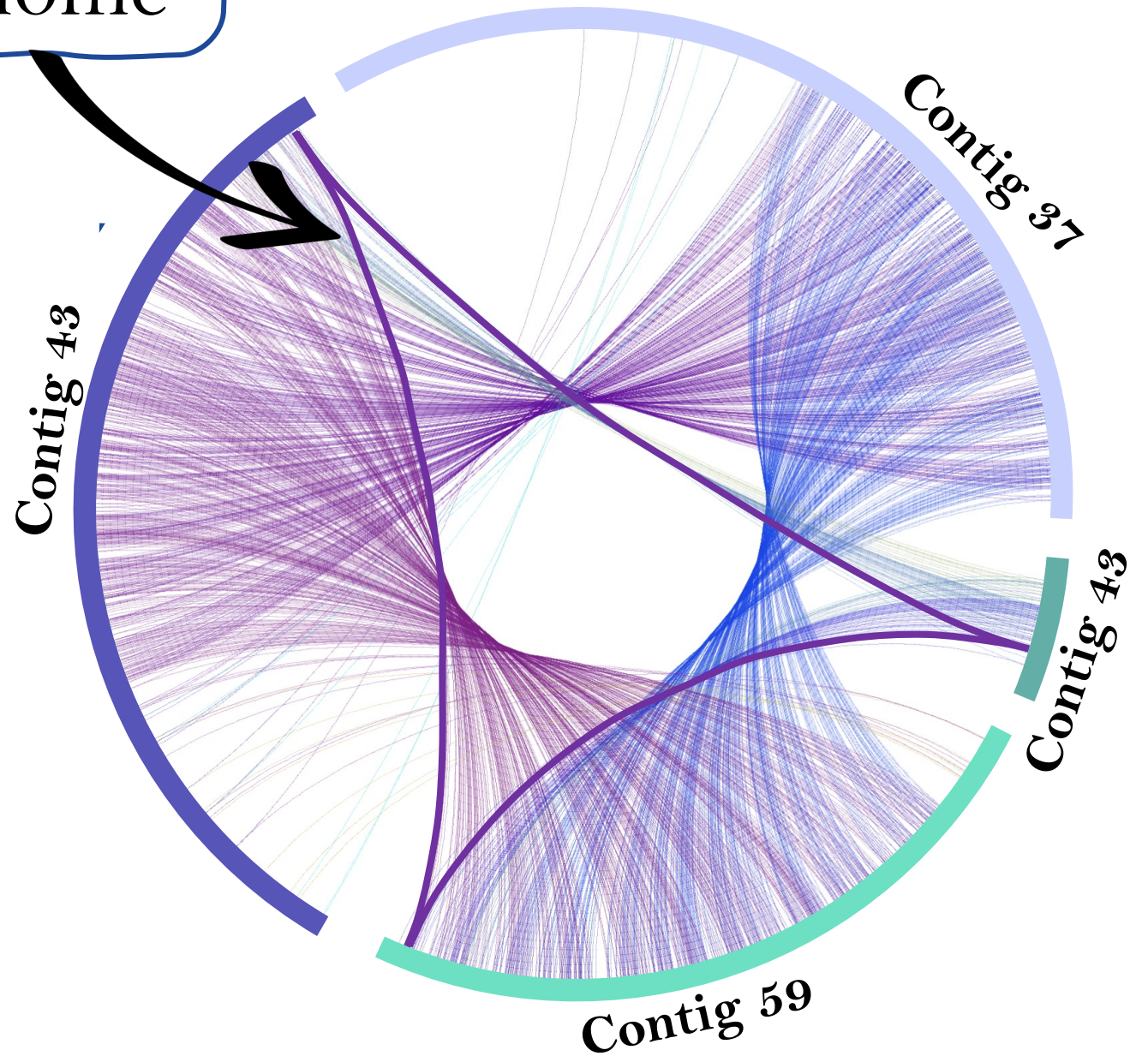


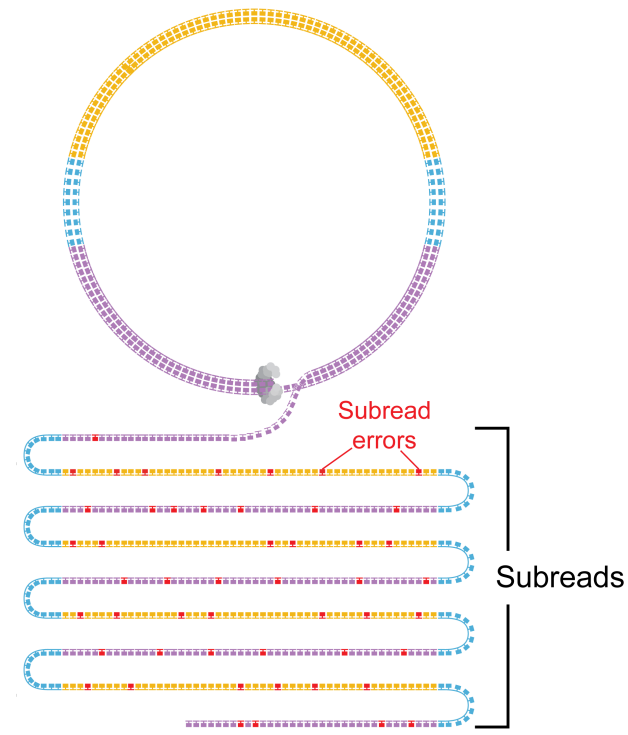


# M.ent has a triploid genome

○ 39, 532 protein coding-genes predicted

○ 75.6% of the protein-coding genes are in three copies



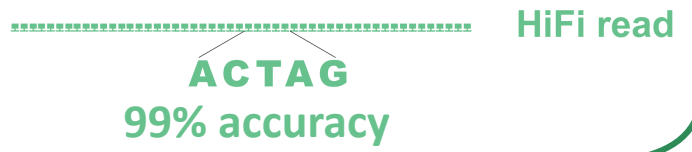


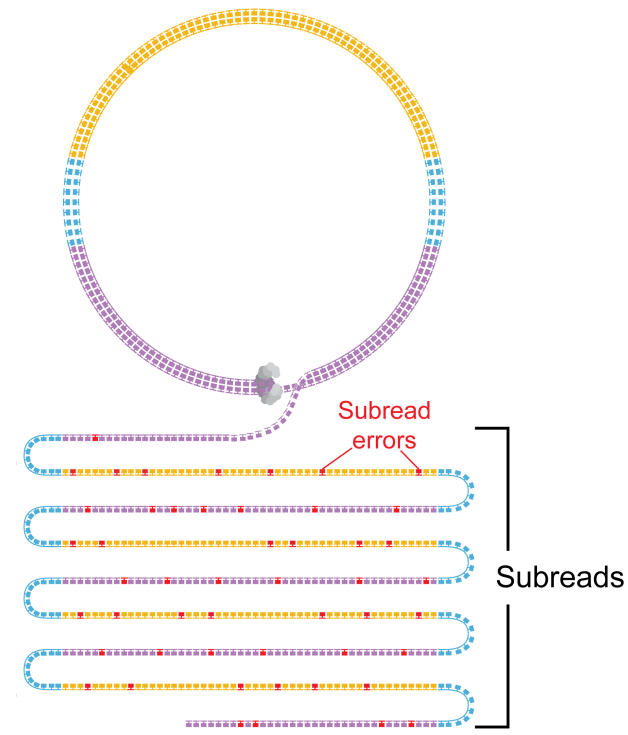
DNA is sequenced multiple times

Trimmed of adapters

Generate consensus read ↓

**pbccs method**



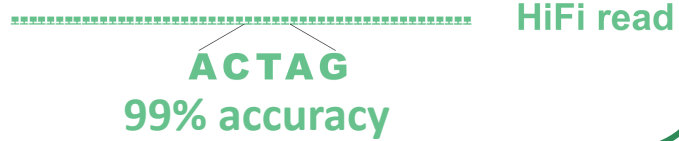


DNA is sequenced multiple times

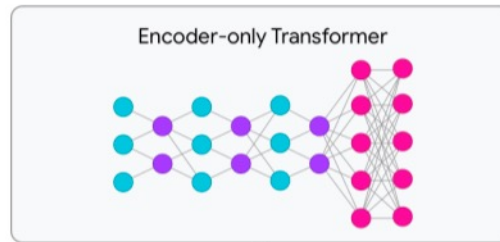
Trimmed of adapters

## pbccs method

Generate consensus read ↓



## DeepConsensus

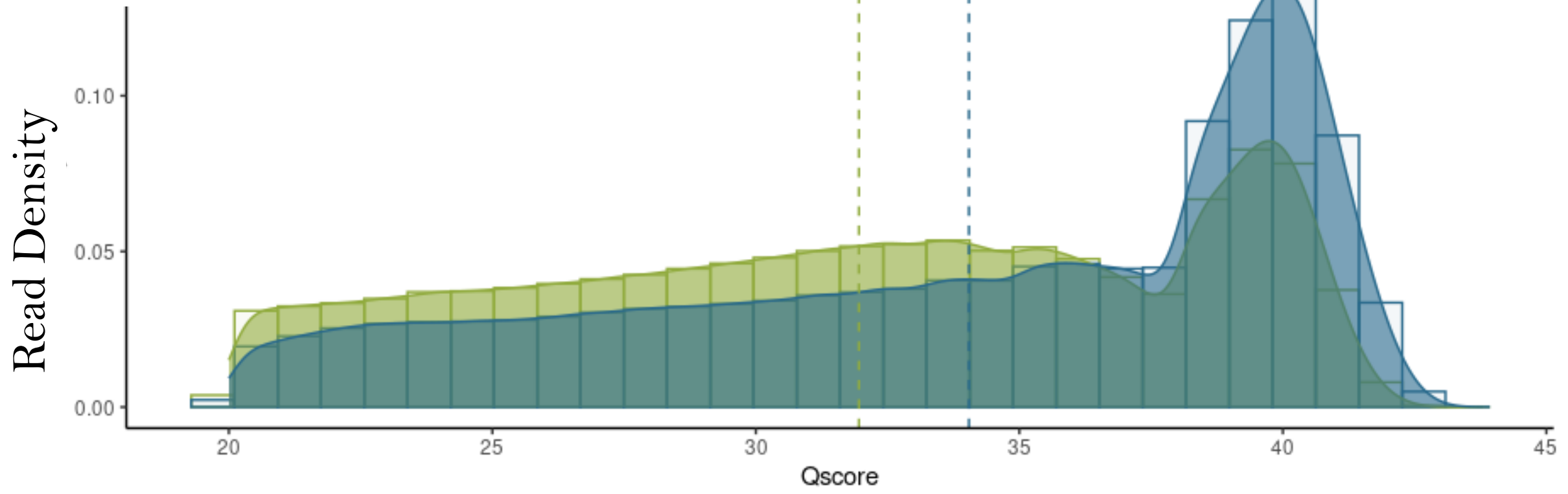


→  New polished read

 PACBIO® Improving Hifi Reads

	No DeepConsensus	DeepConsensus
No. of reads	2 250 199	2 449 079
Mean length	11 194 bp	11 187 bp
Mean Qscore	31.95	34.04

Q20 - 4%  
Q30 + 10%  
**Q40 + 70%**

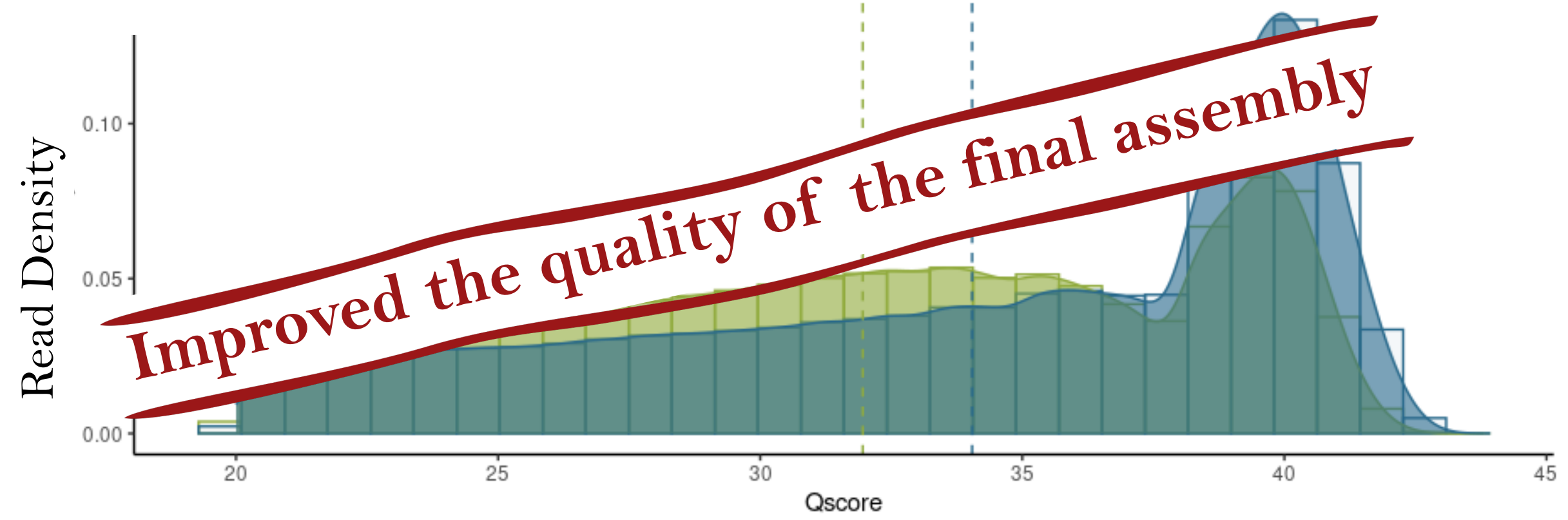


# PACBIO® Improving Hifi Reads

	No DeepConsensus	DeepConsensus
No. of reads	2 250 199	2 449 079
Mean length	11 194 bp	11 187 bp
Mean Qscore	31.95	34.04

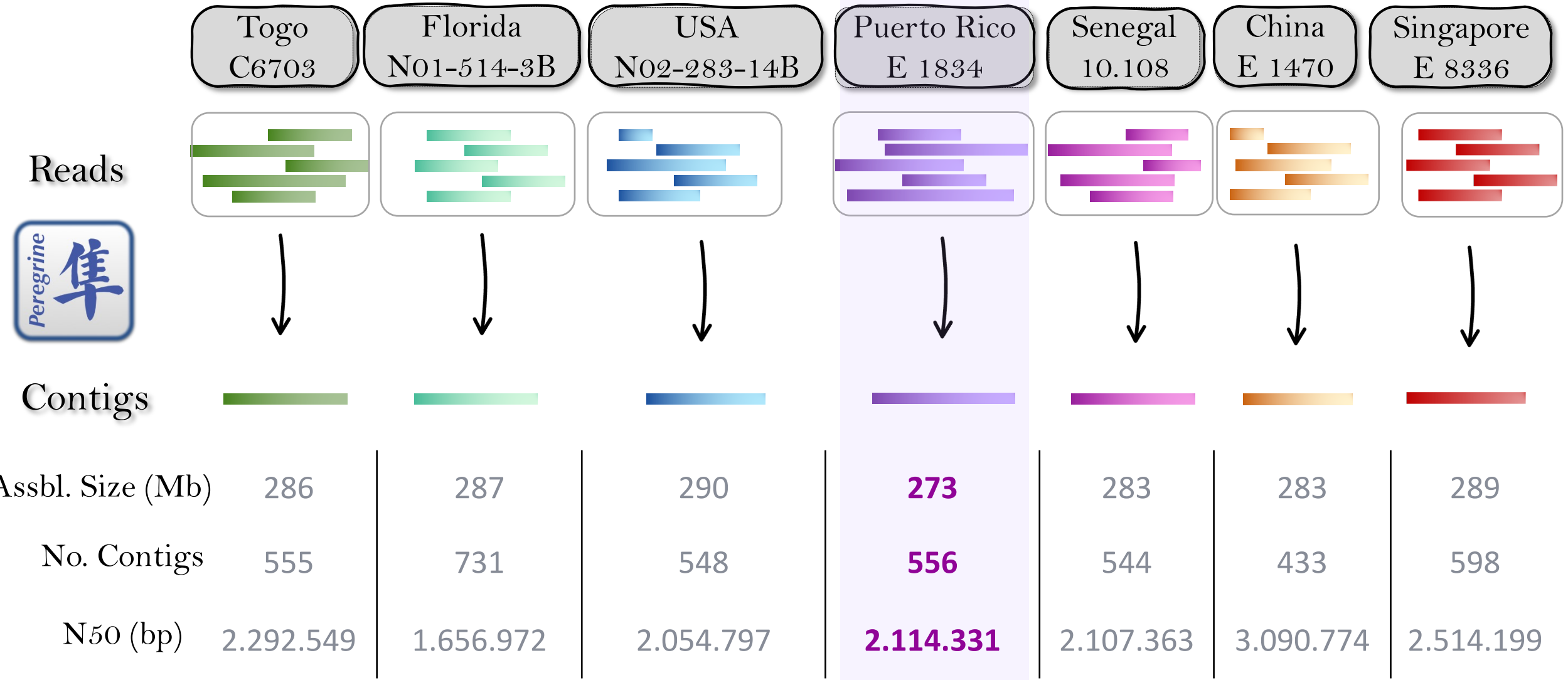
Q20 - 4%  
Q30 + 10%  
**Q40 + 70%**

99.99%  
accuracy





# Assembly statistics



Estimated genome size via flow cytometry : **274.7 Mb ± 18.5** (256.2 – 293.2)

*Koutsovoulos, G.D., et al. (2020) Sci. Data.*

# Exploration of the genomic plasticity of *M. enterolobii*

We have produced high-fidelity genome data (PacBio Hifi) for 7 different geographical isolates of *M. enterolobii*. Each isolate differs in its host range.



## Workflow strategies for variant calling

Compatible host



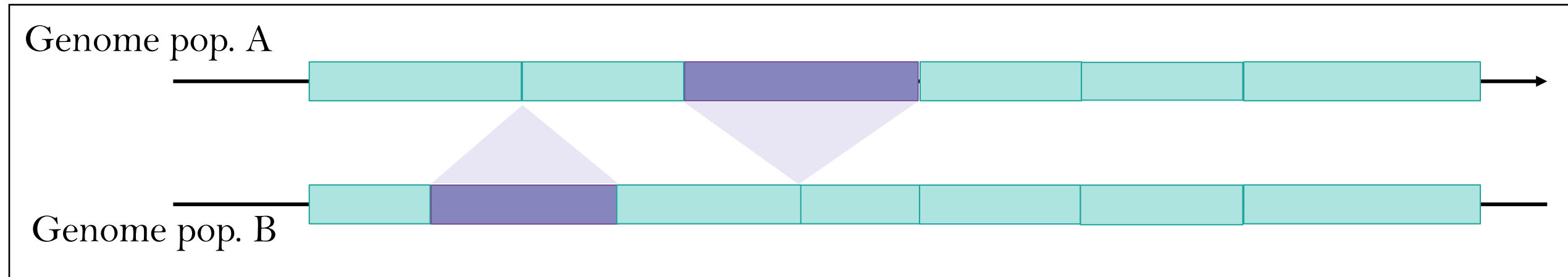
Geographical distribution



# Structural variations in a genome

**Simple definition:** Genome variation of size  $> 50$  bp

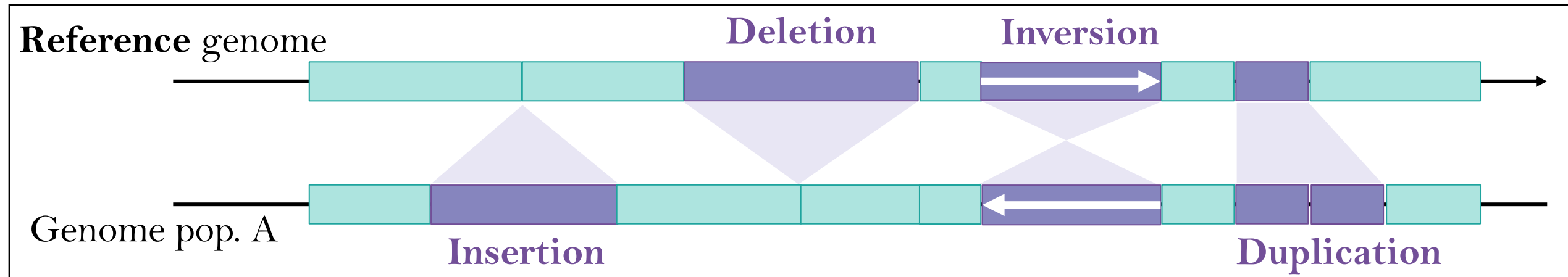
Gathers many different types



# Structural variations in a genome

**Simple definition:** Genome variation of size  $> 50\text{pb}$

Gathers many different types

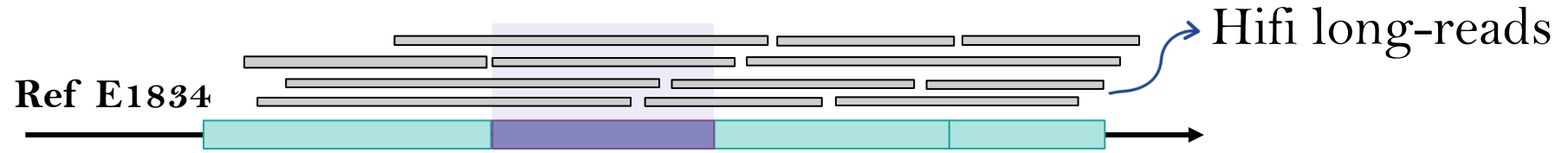
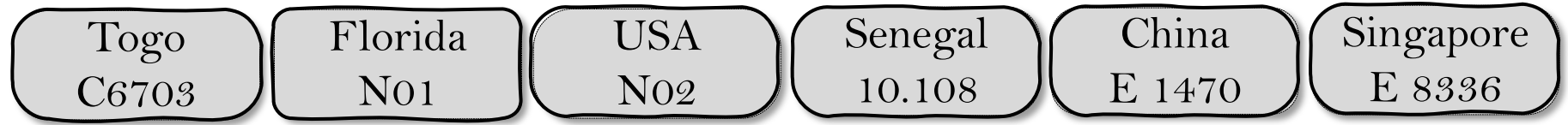


# SVs Workflow

Raw reads (fastq)

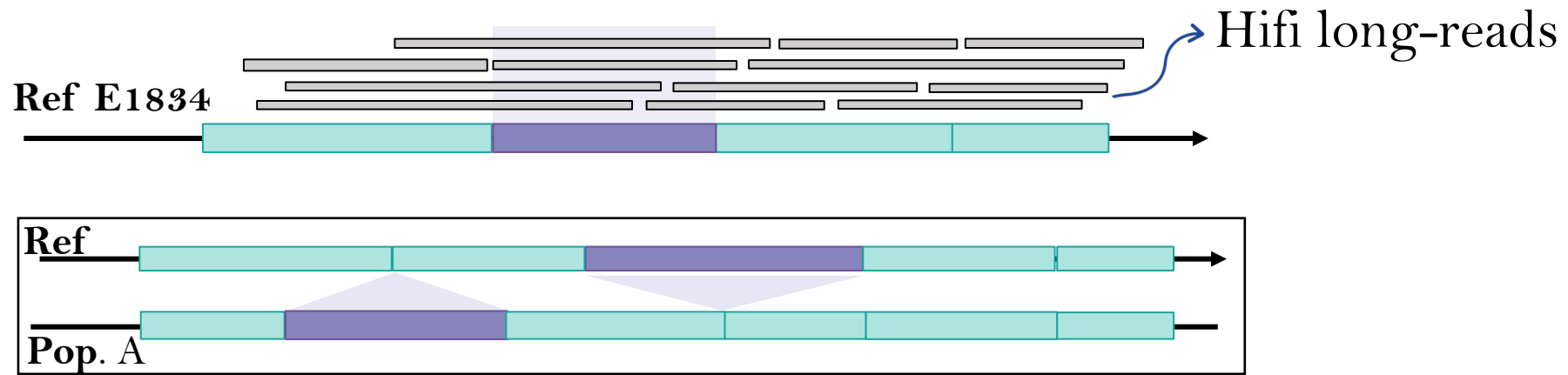
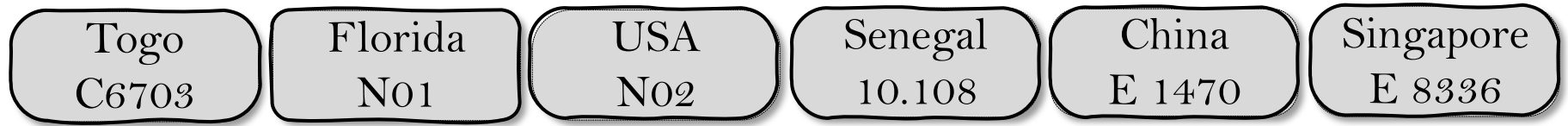
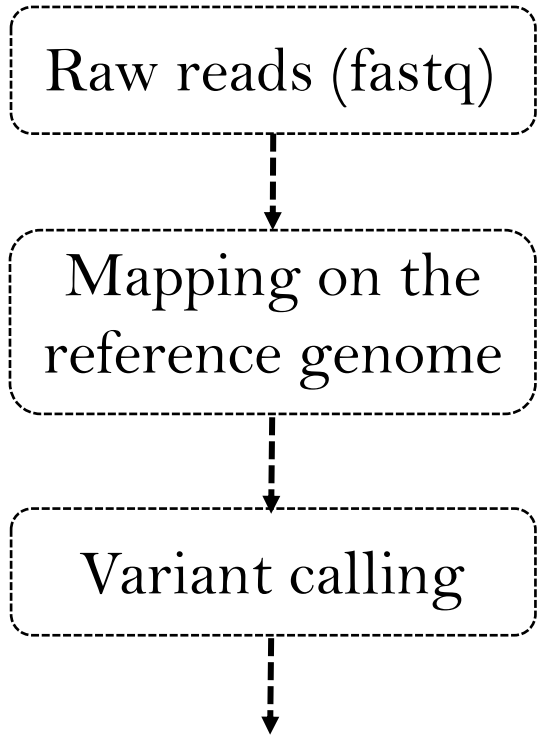
Mapping on the reference genome

Variant calling





# Variant calling



## TOOLS

Average SV found per tool

Sniffles2

11 142

CuteSV

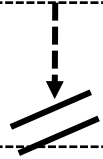
16 232

DeBreak

5 593

# Variant filtering

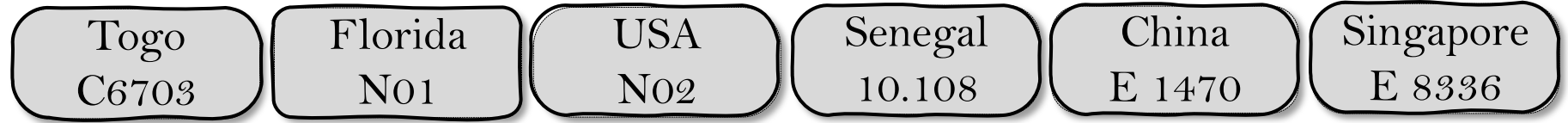
Raw reads (fastq)



Variant calling



Variant filtering



Sniffles2

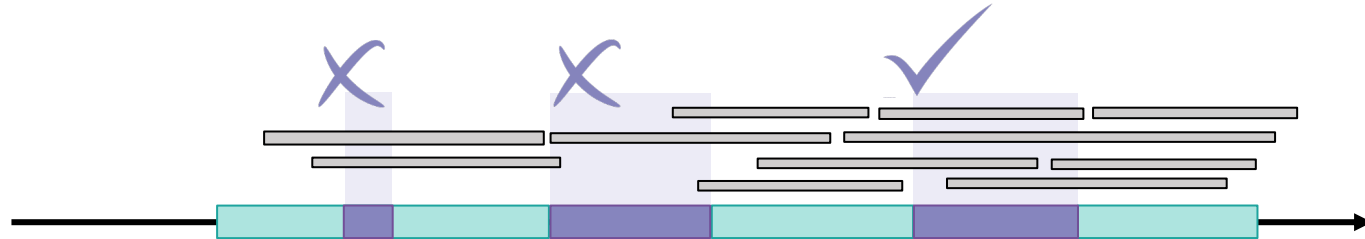
CuteSV

DeBreak

11 142

16 232

5 593 SVs



50bp  $\ll$  SVs size  $\ll$  1Mb

10  $\ll$  SUPP reads  $\ll$  200

Allele Frequency  $\gg$  0.3

MAPQ  $\gg$  10

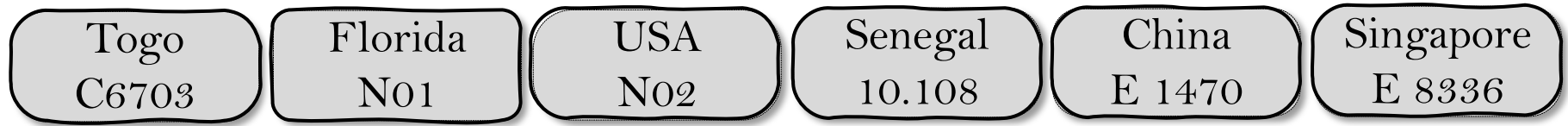
# Variant filtering

Raw reads (fastq)

Variant calling

Variant filtering

50bp  $\ll$  SVs size  $\ll$  1Mb  
10  $\ll$  SUPP reads  $\ll$  200  
Allele Frequency  $\gg$  0.3  
MAPQ  $\gg$  10



Sniffles2

CuteSV

DeBreak

11 142

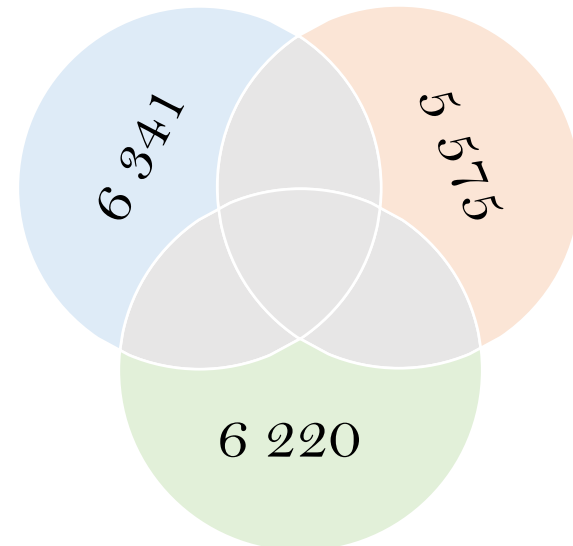
16 232

5 593 SVs

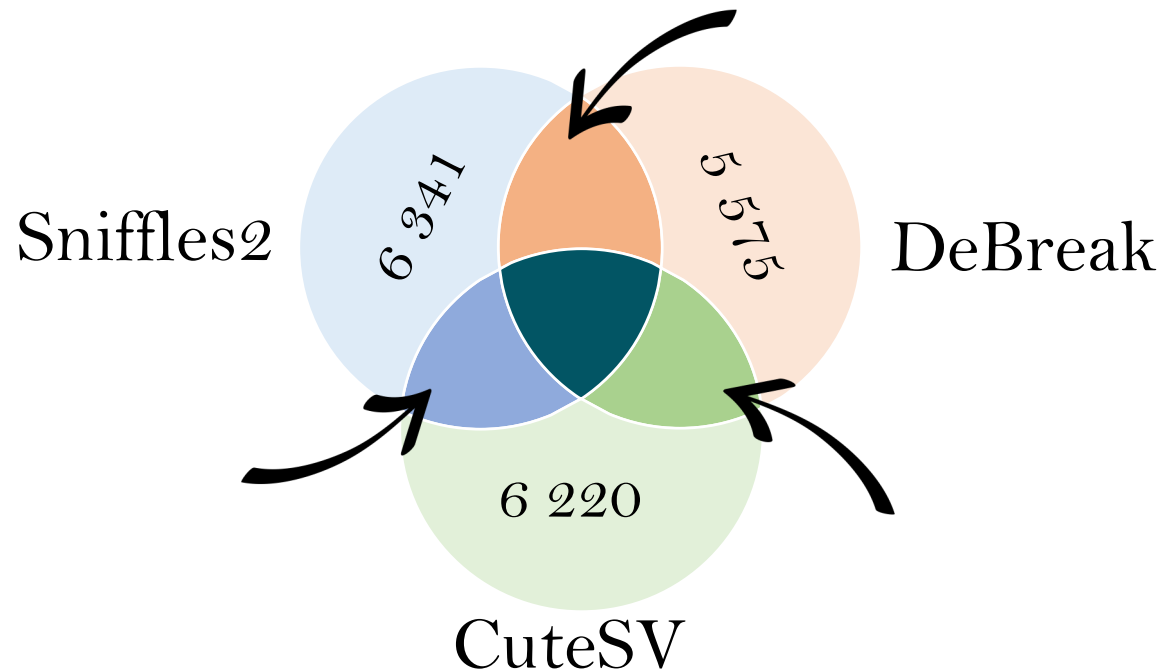
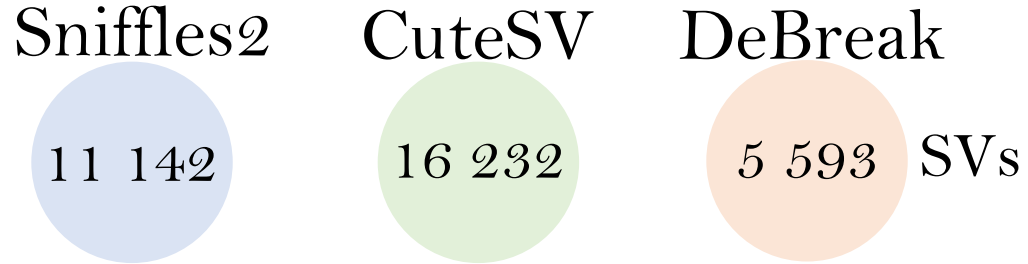
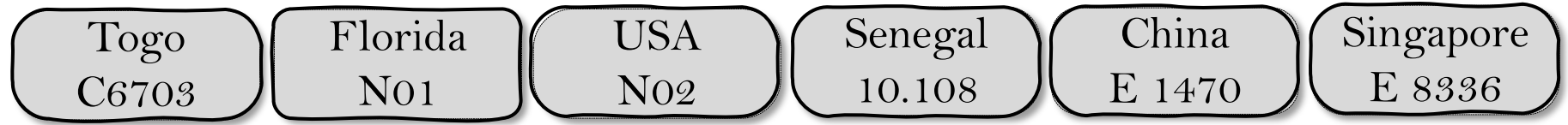
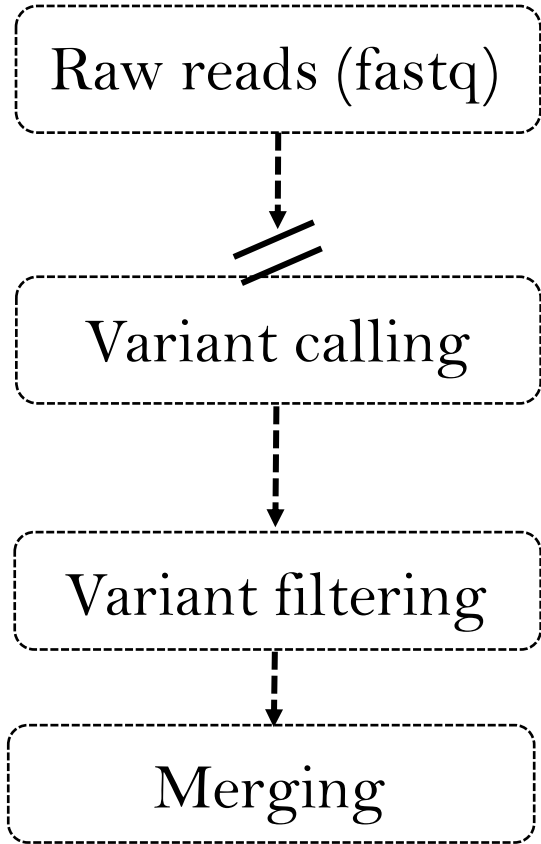
Sniffles2

DeBreak

CuteSV



# Merging all the SVs



## Jasmine

min\_support=2  
min\_overlap=80%

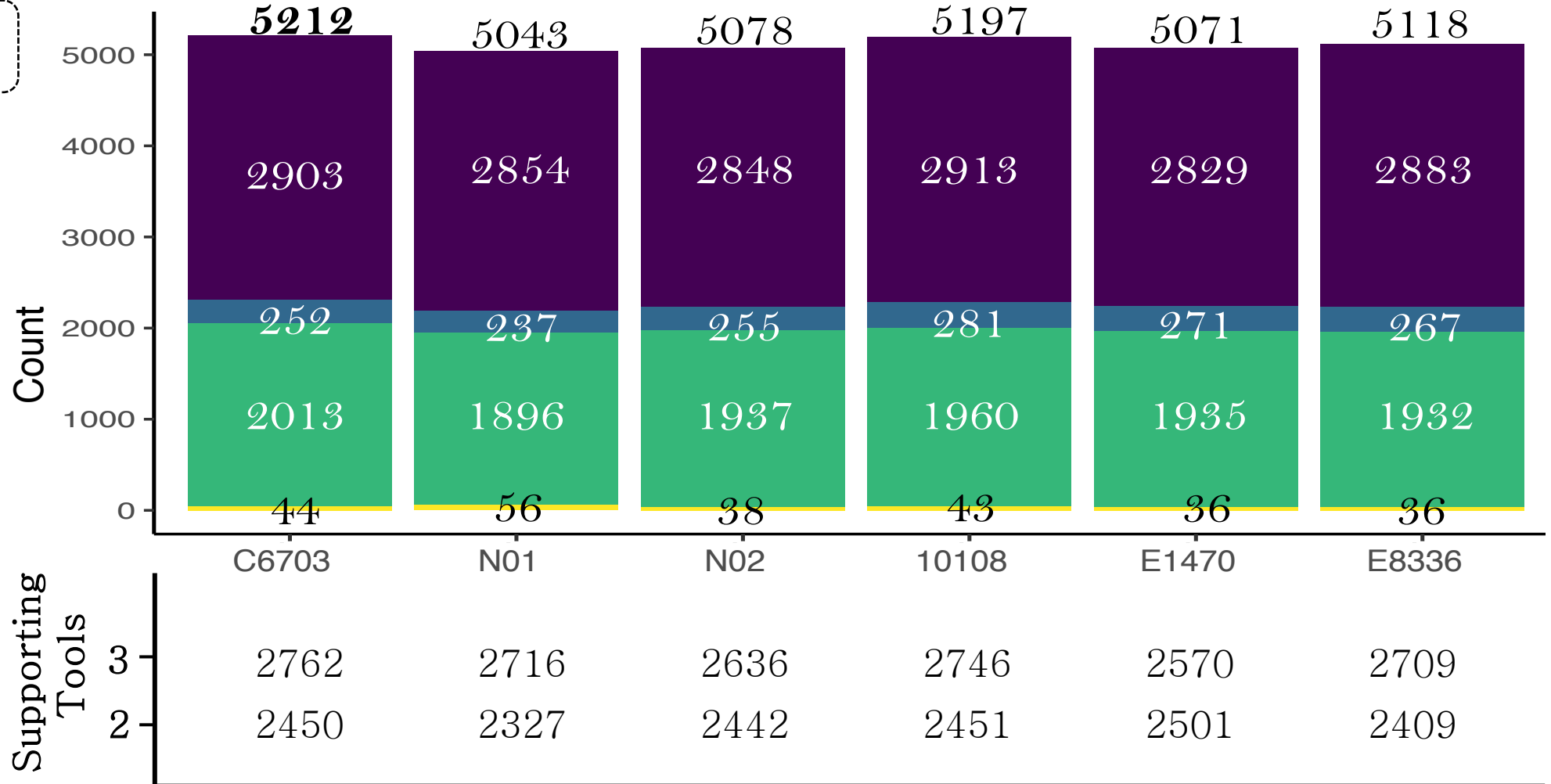
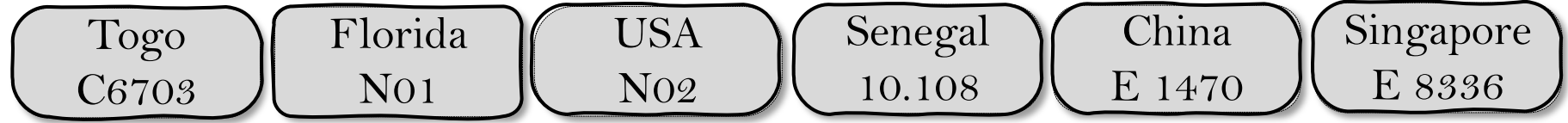
# Structural variants in *M. enterolobii*



Merging

**Jasmine**

min\_support=2  
min\_overlap=80%



SV Type





# Structural variants in *M. enterolobii*

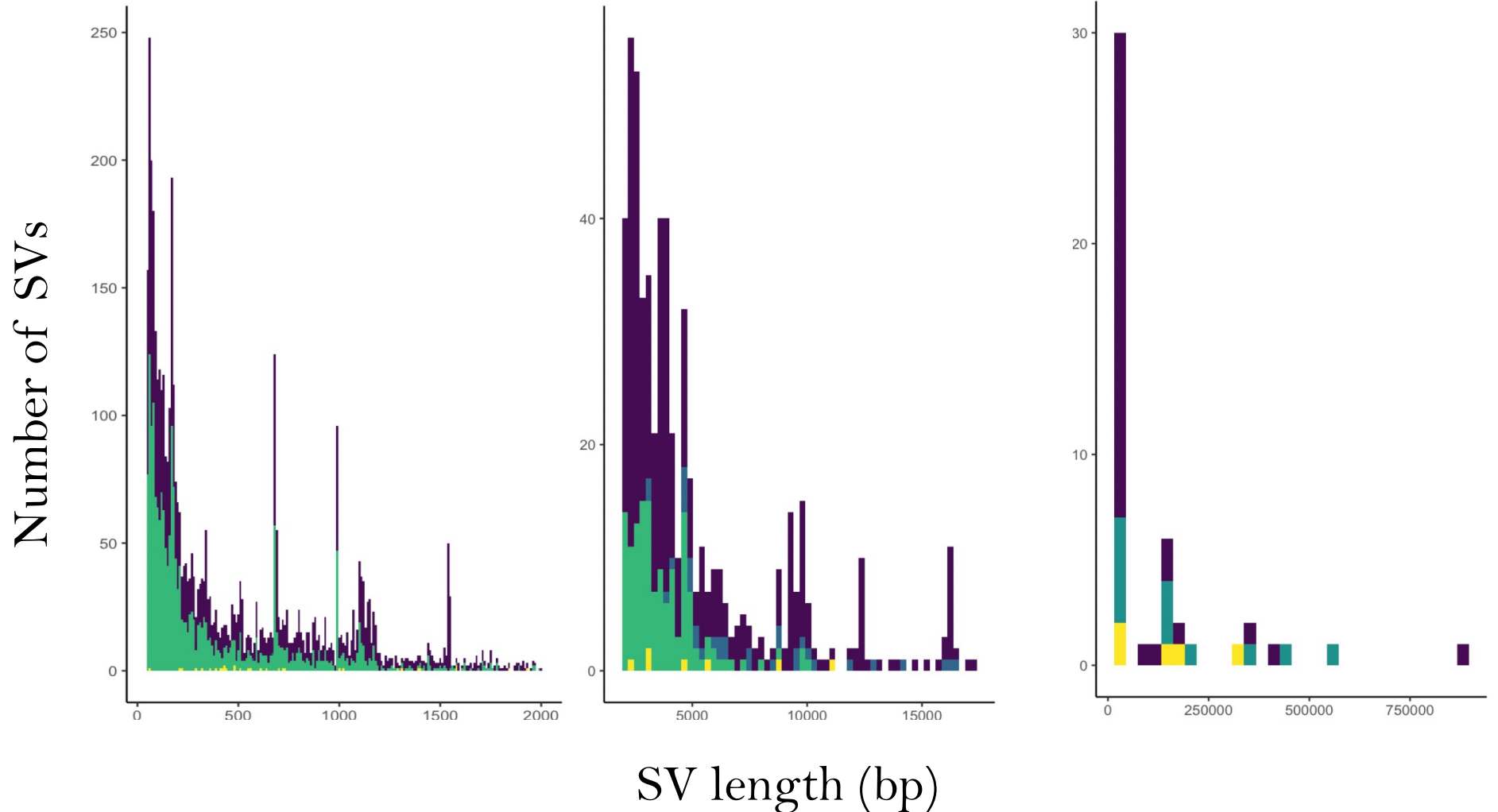
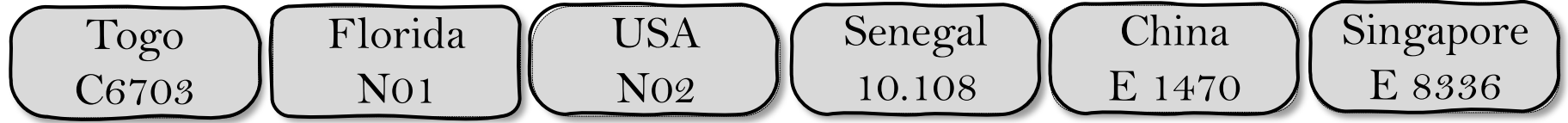


Merging

**Jasmine**

min\_support=2  
min\_overlap=80%

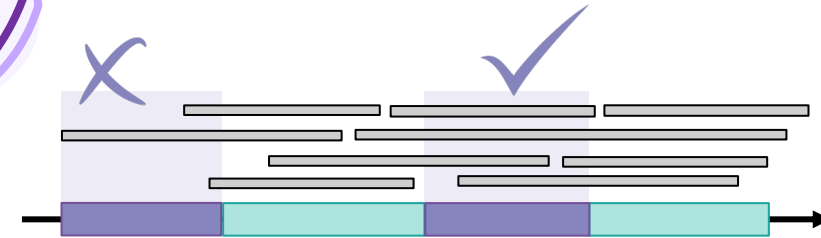
SV Type



# Exploration of the genomic plasticity of *M. enterolobii*

We have produced high-fidelity genome data (PacBio Hifi) for 7 different geographical isolates of *M. enterolobii*. Each isolate differs in its host range.

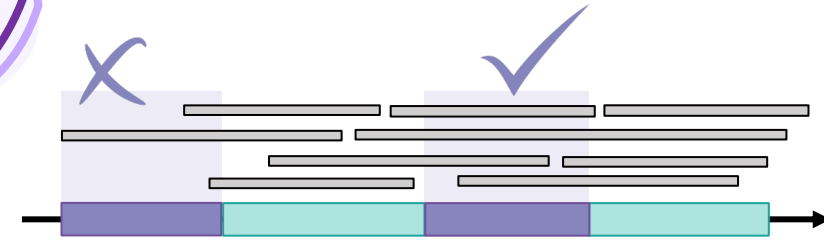
- Reference genome assembly
- Workflow strategies for variant calling



# Exploration of the genomic plasticity of *M. enterolobii*

We have produced high-fidelity genome data (PacBio Hifi) for 7 different geographical isolates of *M. enterolobii*. Each isolate differs in its host range.

- Reference genome assembly
- Workflow strategies for variant calling
- Correlate with phenotypes and snp



Compatible host



Geographical distribution



# Thanks!

EAGS 2024



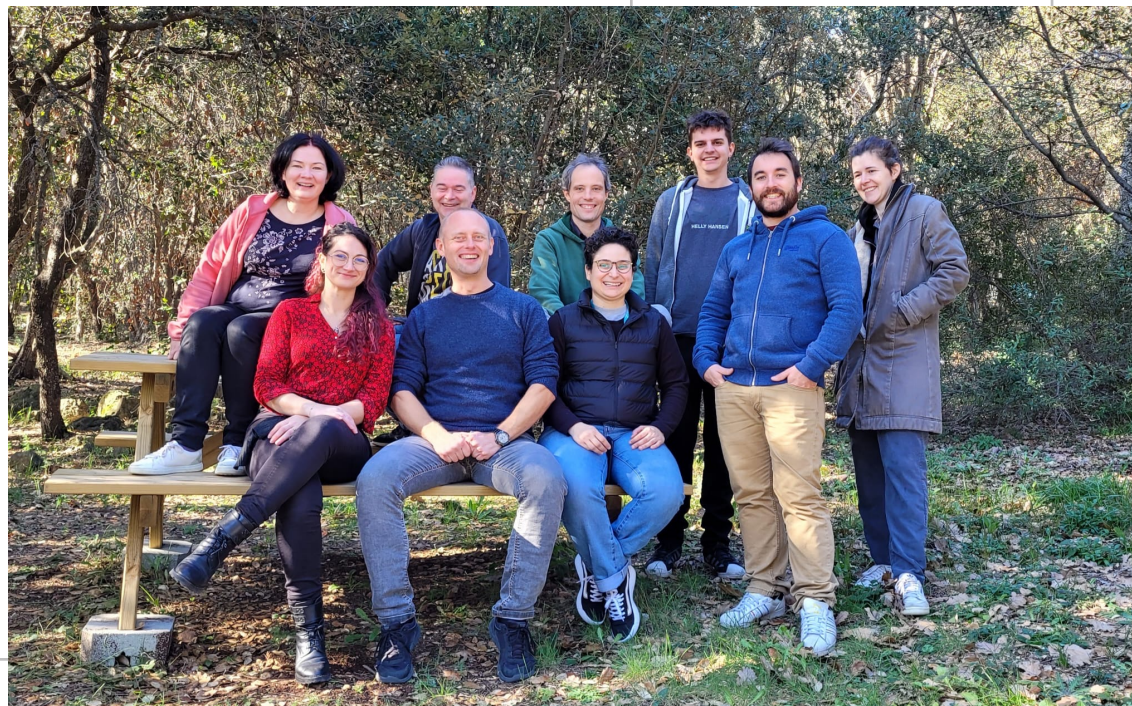
Corinne Rancurel  
Martine Da-Rocha  
Arthur Pere



Marine Sallaberry  
Céline Lopez-Roques  
Joanna Lledo



Sebastian Kiewnick  
Hemanth Gopal



Etienne Danchin  
Karine Robbe-Sermesant  
Marc Bailly-Bechet  
Dominique Colinet  
Carole Belliaro  
Ana Paula Zotta Mota  
Djampa Kozlowski  
Ercan Seckin

