

Unlocking Nanopore sequencing for managing food safety in the food industry

Koen Bossers, PhD, Leratho Biekman & Floyd Wittink, PhD – Bio-informatics department @ Leiden Center for Applied Bioscience & Leiden University of Applied Sciences, the Netherlands

INTRO

- Food pathogens such as *Listeria monocytogenes* are a major cause of food-related illness and product recalls
- Whole genome sequencing (WGS) can link patient-derived food pathogens to food products
- The **goal** of this project is to enable food-producing companies to **use WGS** in monitoring their production and supply chain **to identify the source of the food pathogens**

METHODS

- **700+** *Listeria monocytogenes* isolates from food, raw material and production environment
- Nanopore WGS sequencing
- Custom-built Snakemake pipeline to reconstruct whole genomes [flye], **determine phylogenetic distances** [sourmash], annotate with MLST schemes [mlst] and virulence/resistance genes [abricate]
- Custom-built API and web-application for interactive exploration of results, enabling **in silico food forensics**.



RESULTS

- Interactive web application

ASK ME ABOUT A LIVE DEMO

Precision Food Safety analysis tool
Start typing to perform a search on sequence data

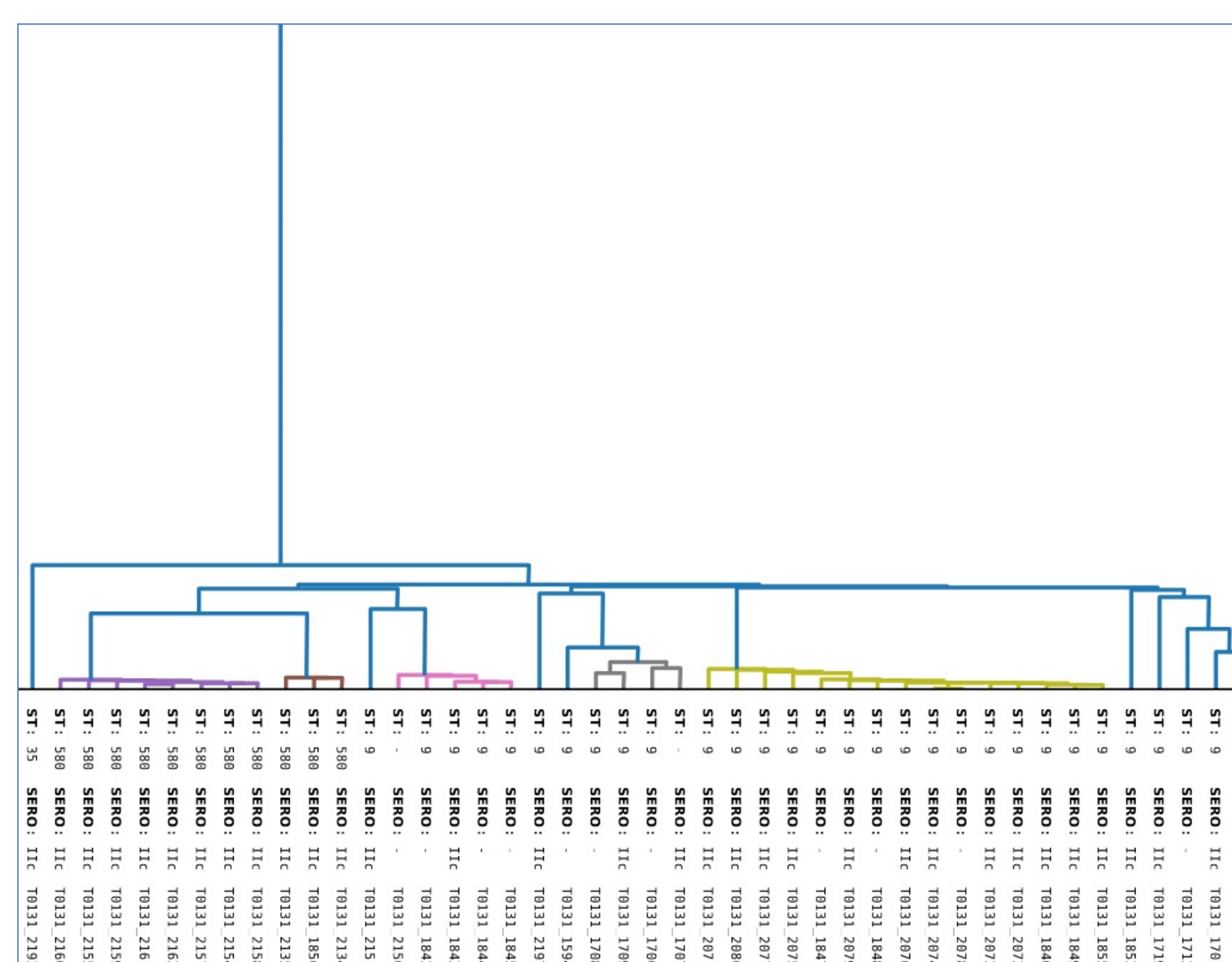
bedrijf:meat_b

0 sequencedata records selected Show Records Reset Selection BUILD TREE

36 sequencedata records found for query bedrijf:meat_b REMOVE ALL ADD ALL

	MEGABASES	PHRED	CONTIGS	GENES	SEQUENCE TYPE	LINEAGE	CLONAL COMPLEX	SEROTYPE
T0131_1201 Meat_B								
SEQUENCE DATA FAP07615_barcode77	144.0	13.4	5	42	8	II	CC8	Ila
T0131_1202 Meat_B								
SEQUENCE DATA FAP07615_barcode85	144.0	13.4	3	42	8	II	CC8	Ila

- **Insight** in related and unrelated **clusters** of *Listeria monocytogenes*



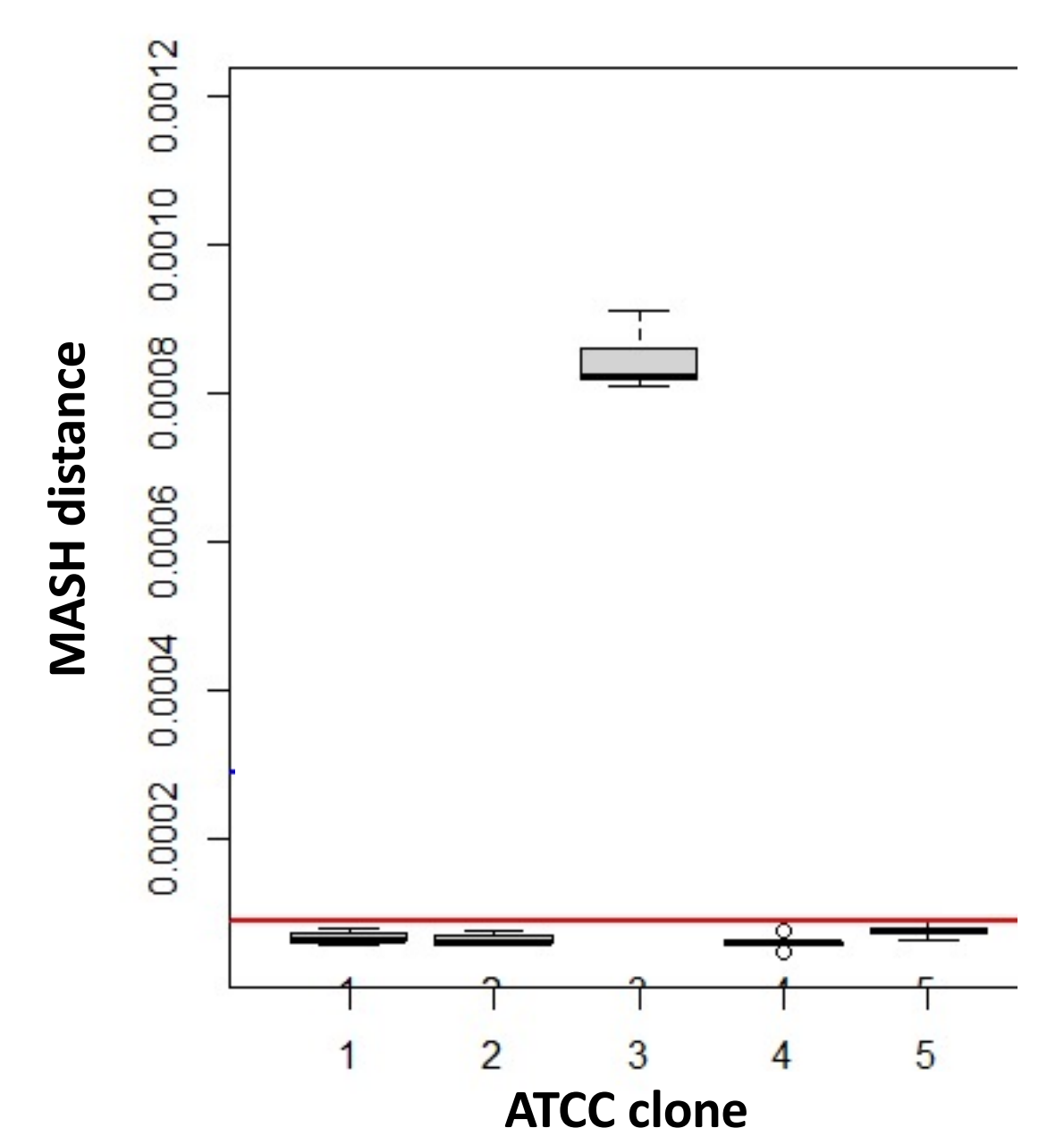
TAKE-HOME MESSAGE

- The Precision Food Safety pipeline and web-app unlocks whole-genome based food forensics without requiring explicit bioinformatics knowledge.

Benchmarking MASH genome distance estimations

- ATCC clones of *Listeria monocytogenes*
- Sequence + reconstruct genome, calculate MASH distance to reference
- Distance threshold: cluster of clones or not?

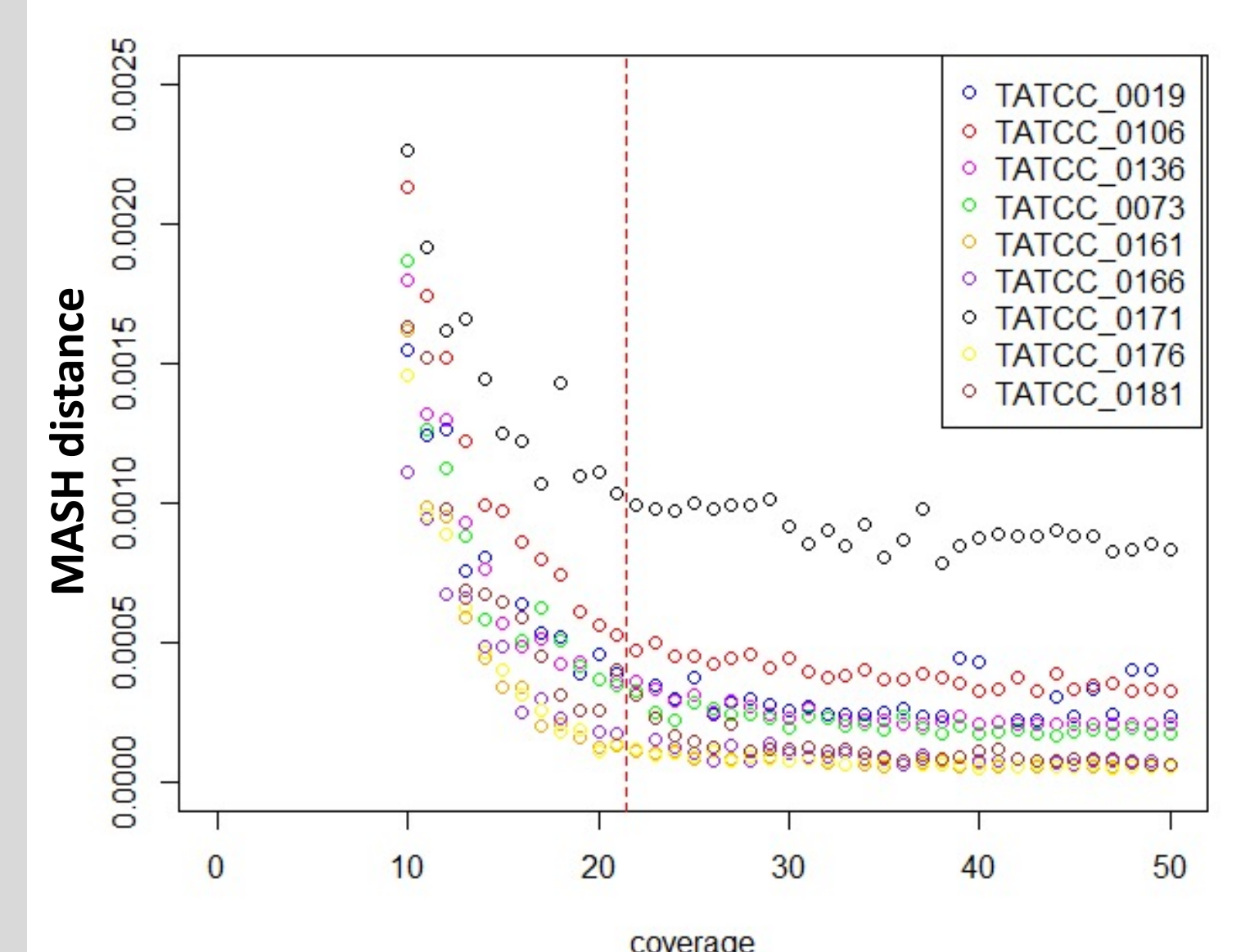
MASH distance to reference genome



Genome coverage versus distance estimation

- Proper genome distance estimation depends on accuracy of genome reconstruction
- Sequence depth is the most important parameter in successful genome reconstruction

MASH distance versus coverage



[flye] : <https://github.com/fenderglass/Flye>
[sourmash] : <https://github.com/sourmash-bio/sourmash>
[mlst] : <https://github.com/tseemann/mlst>
[abricate] : <https://github.com/tseemann/abricate>

Feel free to contact me : bossers.ko@hsleiden.nl