# INRAE
la science pour la vie, l'humain, la terre

# Implementation of a Nanopore metagenomic workflow for time-series of complex microbial communities: MAG catalog and functional annotation.
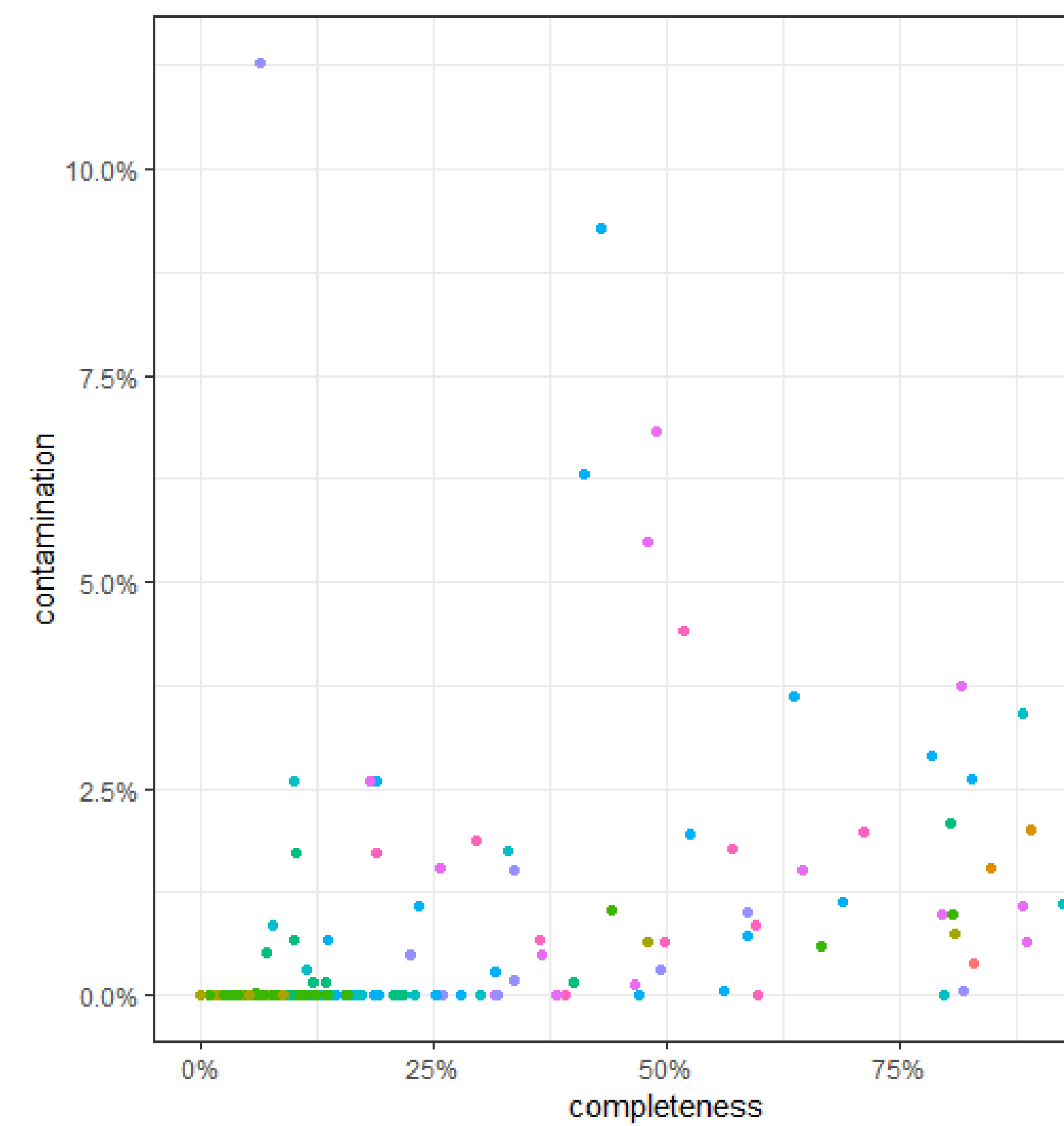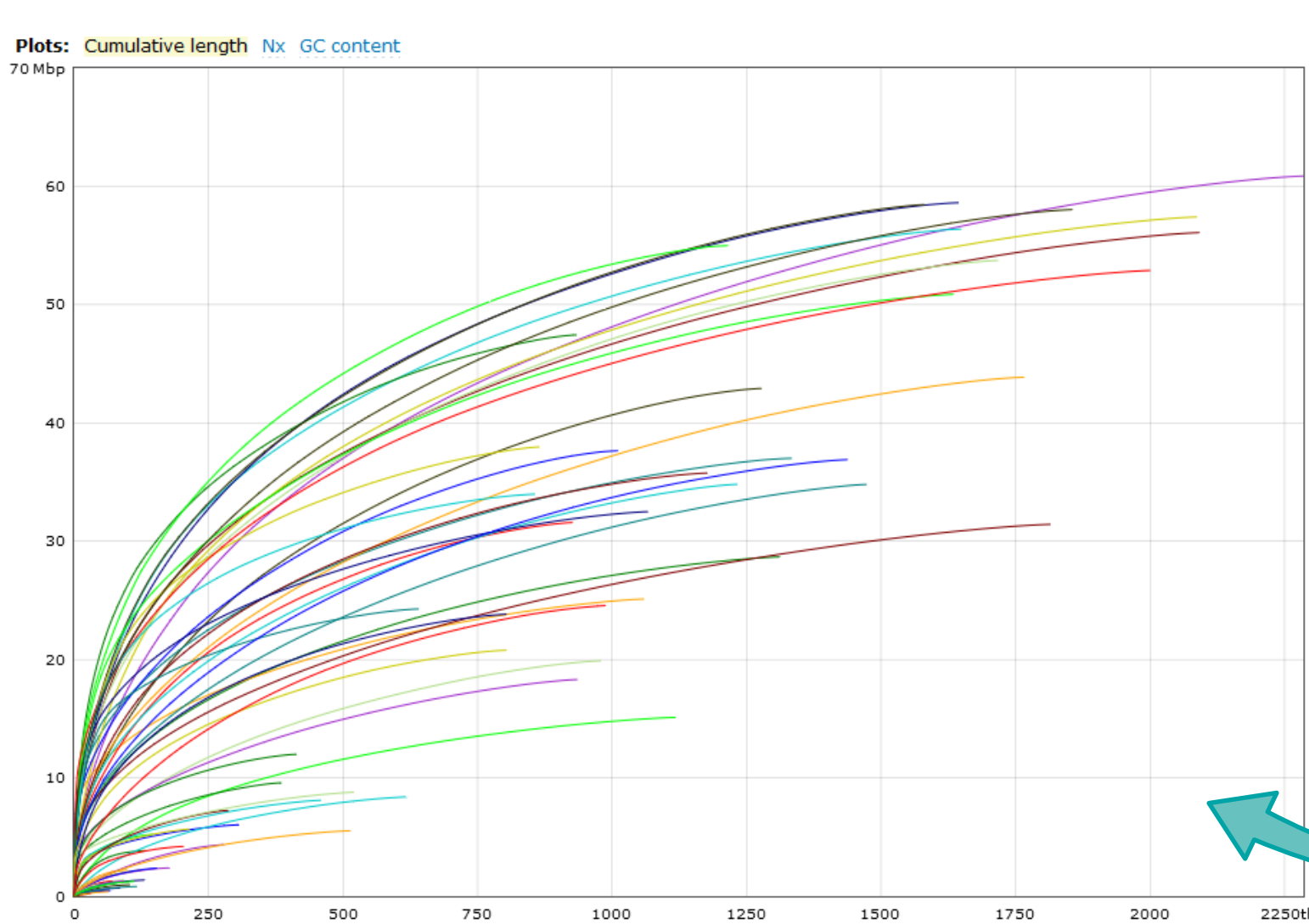
Cédric Midoux[1, 2, 3], Chrystelle Bureau[1], Baptiste Quentin[1] and Olivier Chapleur[1].

1. INRAE, UR 1604 – PROSE
2. INRAE, UR 1461 – MaIAGE
3. BioinfOmics, MIGALE bioinformatics facility

## › Data

- Samples come from 6 anaerobic digesters (3 controls: A, B, C + 3 with perturbation: D, E, F) at 9 different time points (n = 54)
- Lab-made sequencing :
  - MinION Mk1C
  - Flow-cells R9.4.1
  - Rapid Barcoding Kit
  - Multiplexing
- Throughput increases over the runs and our experience

**Raw reads**

**PREPROCESS**
- filtong
- porechop

**Trimmed reads**

**QUALITY CONTROL**
- NanoPlot
- FastQC
- kaiju / nr

**POOL-SAMPLES**
- Pool all samples
- Co-assembly canu
- Polishing racon + 2 × medaka
- Contigs consensus
- Binning MetaBAT2
- MAGs pool-samples

**PER-SAMPLE**
- Assembly per sample canu
- Polishing racon + 2 × medaka
- Contigs consensus
- Binning MetaBAT2
- MAGs per-sample

**POOL-CONSENSUS**
- Pool all contigs consensus
- Binning MetaBAT2
- MAGs pool-consensus

Pool MAGs from all approach

**DE-REPLICATION**
- dRep

**MAGs STATS**
- checkM

**MAGs pool-consensus**

**MAGs ANNOTATION**
- Taxonomy GTDB
- Functional prokka + eggNOG

295 MAGs

**COUNT**
- Alignment minimap2
- Count samtools

**EXPORT**
- pandas
- phyloseq

5767 KEGG_ko

## › Statistical Challenges

- How to normalize count?
- Are reads correctly distributed between MAGs?
- Should we consider the length of the reads for normalization?
- Which effects of sequencing depth?

## › Perspectives

- Statistical analysis and data integration
- Add new data to strengthen catalogs
- Workflow sharing
- Publications …

Centre
**Ile-de-France - Antony-Jouy**

Snakemake workflow avalable
https://forgemia.inra.fr/cedric.midoux/nanosnake