Victor RENAULT
Julien MASLIAH PLANCHON

21 mars 2025

NanoDIAG

Outline:

- Introduction to NanoDIAG
- Various Bioinformatics topics of interest
  - DNA pipelines
  - RNA pipelines
  - Classifiers & ultrafast diagnostics
- Conclusion

ENSEMBLE, PRENONS
LE CANCER DE VITESSE

institut Curie

Victor RENAULT
Julien MASLIAH PLANCHON

21 mars 2025

NanoDIAG

Outline:

Introduction to NanoDIAG

Various Bioinformatics topics of interest

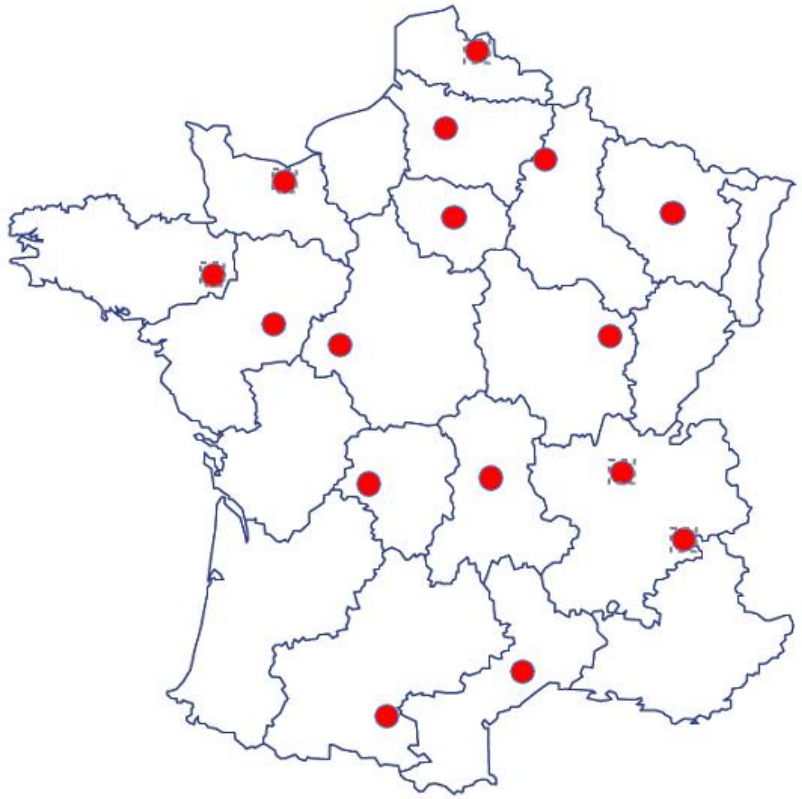
- DNA pipelines
- RNA pipelines
- Classifiers & ultrafast diagnostics

Conclusion

ENSEMBLE, PRENONS
LE CANCER DE VITESSE

institut Curie

Introduction



Martinique

- NanoDIAG = regroupement national des utilisateurs de Nanopore en diagnostique (2023)
- Teams meeting NanoDIAG le premier jeudi chaque mois à 14h
- Intéressé(e) pour **rejoindre NanoDIAG** => julien.masliahplanchon@curie.fr

Victor RENAULT
Julien MASLIAH PLANCHON

21 mars 2025

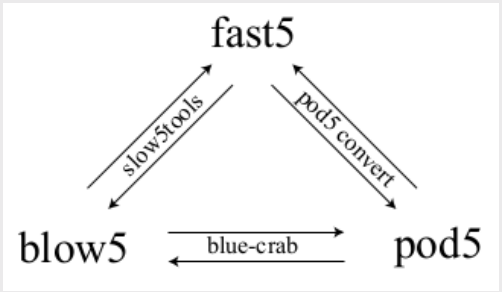
NanoDIAG

Outline:

- Introduction to NanoDIAG
- Various Bioinformatics topics of interest
  - DNA pipelines
  - RNA pipelines
  - Classifiers & ultrafast diagnostics
- Conclusion

DNA pipelines

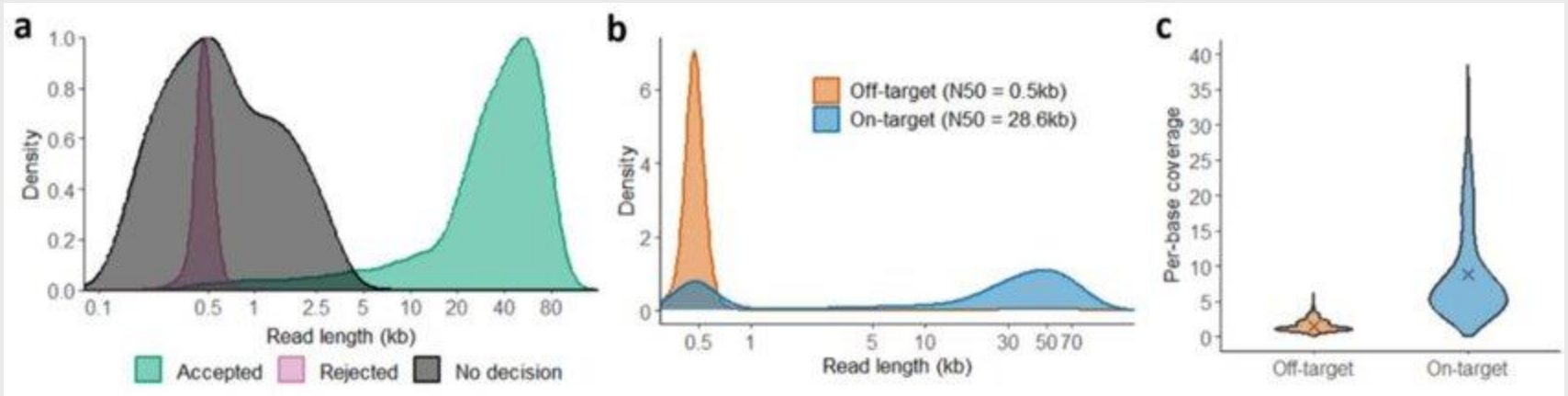
MinKNOW



Various raw Nanopore file formats

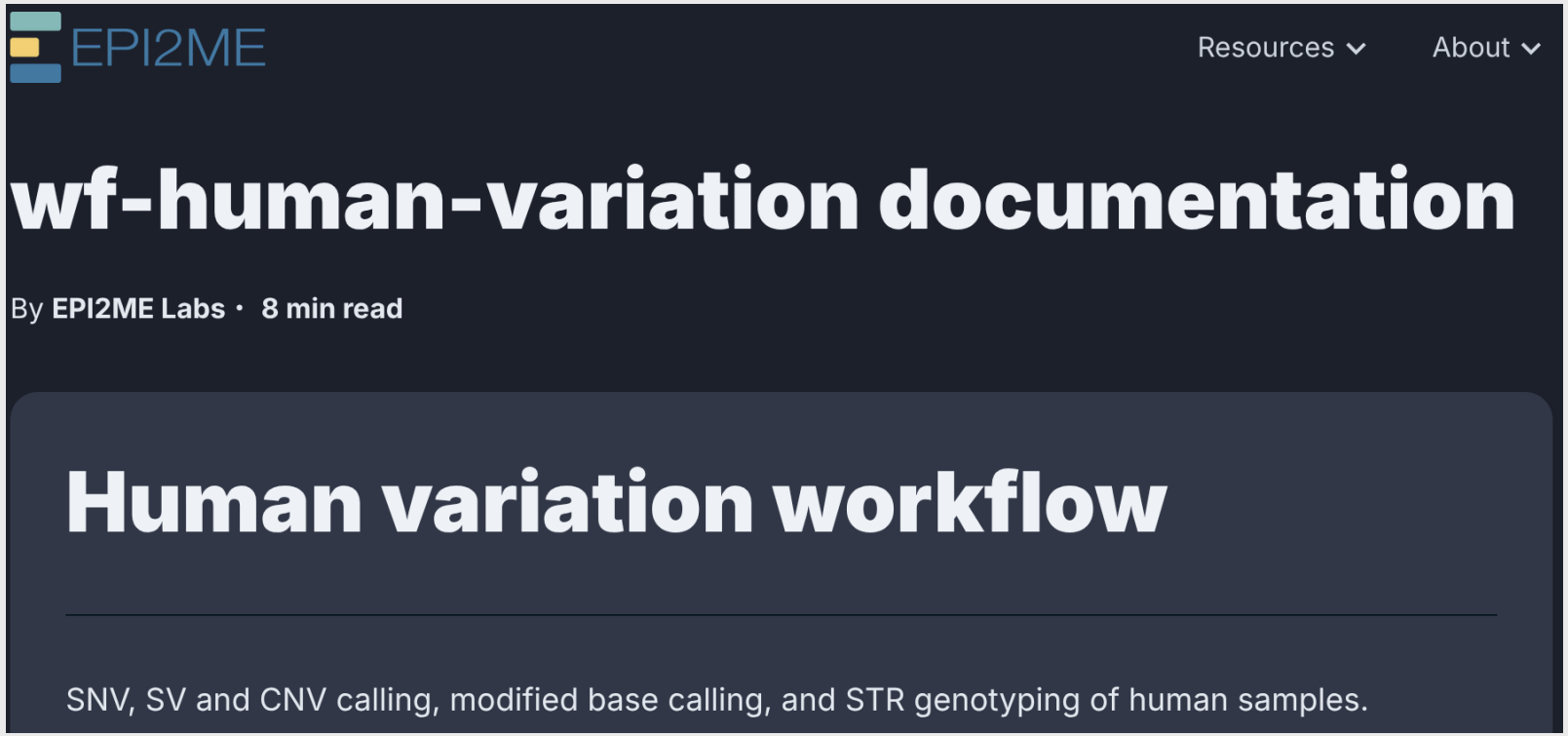
https://nanocem.readthedocs.io/en/latest/preparation/

- Base calling
- QC
- Variant calling
- Complex rearrangements
- CNV
- Phasing (variants / meth.)
- De novo assembly
- Methylation calling
- Others (ctDNA, Telomere seq, etc.)



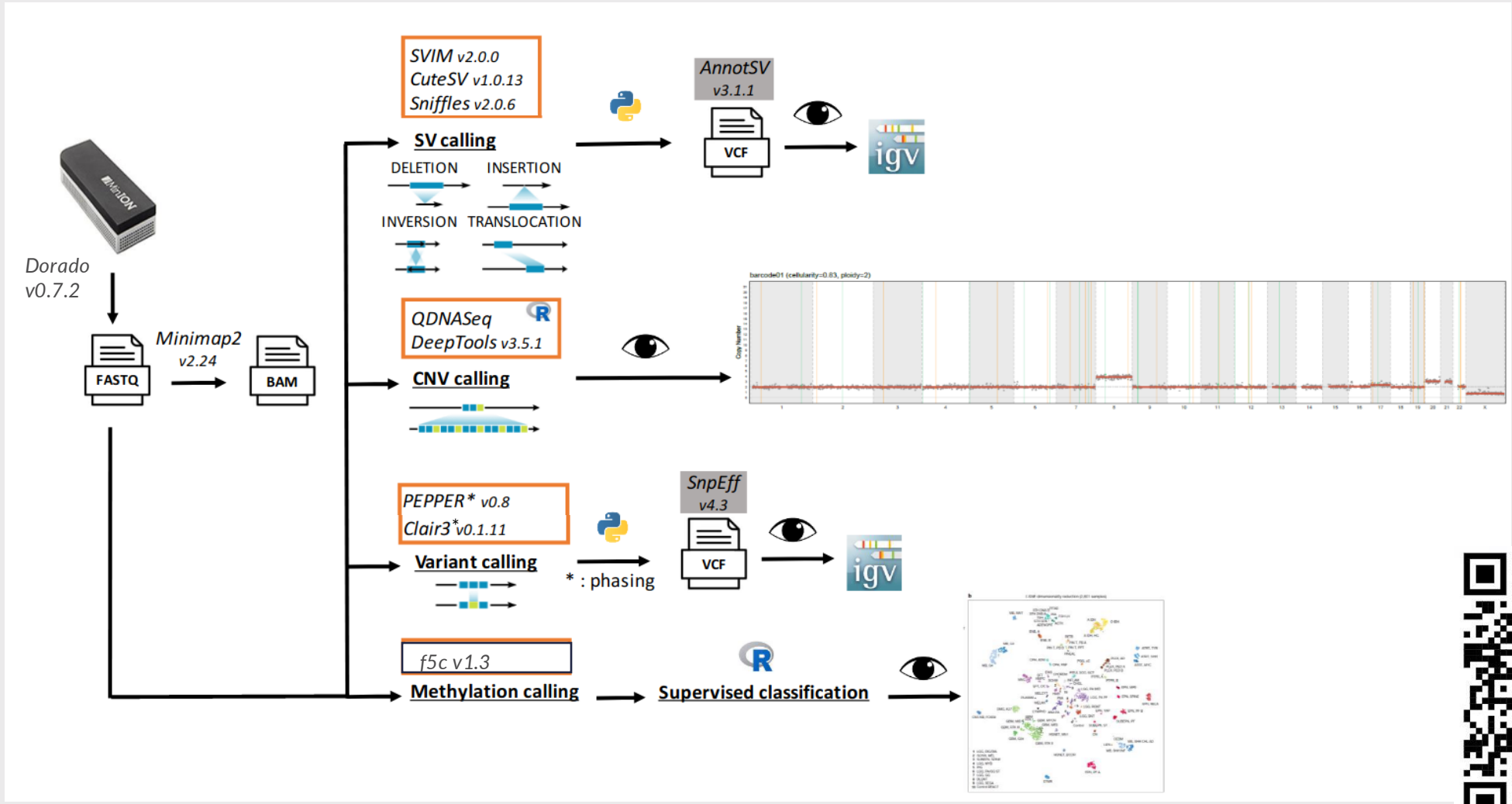
Adaptive sampling (https://www.researchgate.net/figure/Targeted-nanopore-sequencing-with-adaptive-sampling-on-the-reference-DNA-sample-HG002_fig1_376465351)

DNA pipelines - EPI2ME



https://epi2me.nanoporetech.com/wfindex/

DNA pipelines – NanoCliD (**Nano**pore **Cli**nical **D**iagnosis)



GitHub

https://github.com/InstituteCurieClinicalBioinformatics/NanoCliD

DNA pipelines – Variant / methylation phasing

HP1 HP2

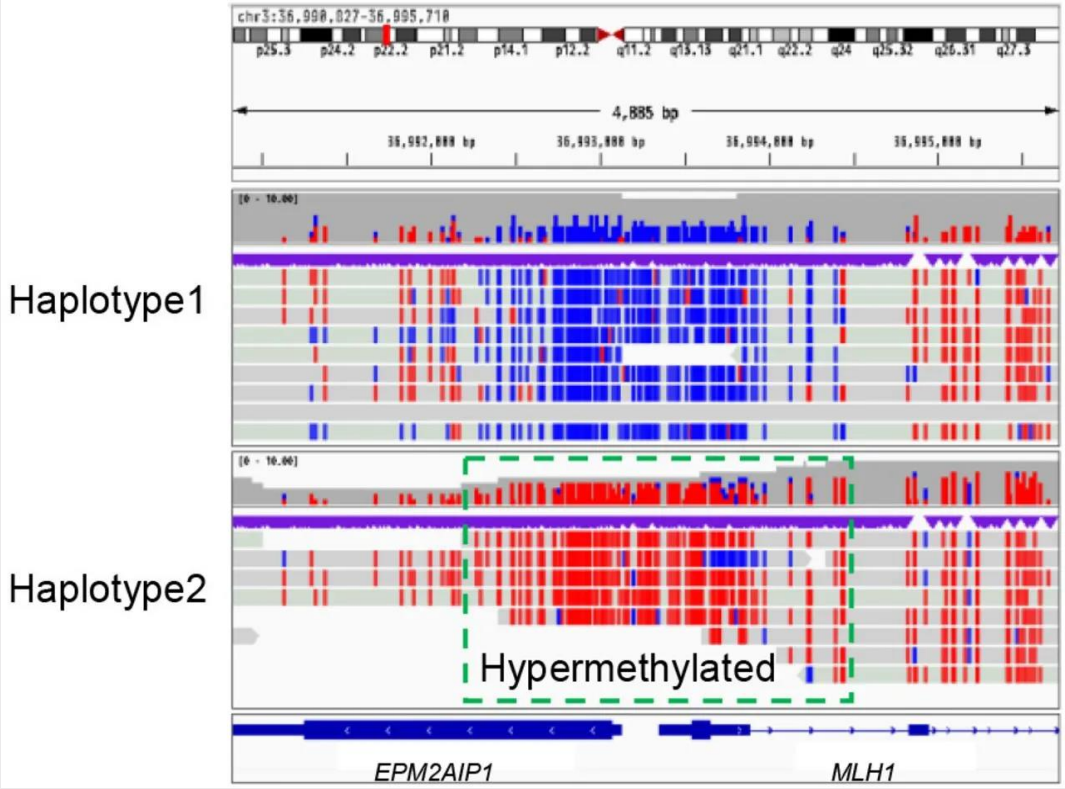
```
chrX    139492497   .   A   G   23.26   PASS   F   GT:GQ:DP:AD:AF:PS   0|1:23:190:71,116:0.6105:139255643
chrX    139492712   .   A   G   24.46   PASS   F   GT:GQ:DP:AD:AF:PS   0|1:24:190:73,116:0.6105:139255643
chrX    139493204   .   T   C   4.98    PASS   F   GT:GQ:DP:AD:AF:PS   1|0:4:190:120,65:0.3421:139255643
```

HP1 : A A C

HP2 : G G T

Genotype

PhaseSet ID



https://nanoporetech.com/ja/resource-centre/faster-analysis-cancer-associated-variants-using-unique-targeted-nanopore

DNA pipelines – methylation calling

modkit

Nombre de CpG	C06	C07	C26
Cov (>=1)	28M	28,6M	28M
Meth Cov (>=1)	26.9M	28M	27,8M
Cov (>= 5)	14,3M	23,1M	21,2M
Meth Cov (>=5)	2.5M + 900k justifies	7.19M + 2,19M justifies	6,3M + 6,1M justifies

DeepMod2

Nombre de CpG	C06	C07	C26
Cov (>=1)	28M	28,6M	28M
Meth Cov (>=1)	27,8M	28,6M	25,3M
Cov (>= 5)	14,3M	23,1M	21,2M
Meth Cov (>=5)	12,2M	21,4M	12,5M

Victor RENAULT
Julien MASLIAH PLANCHON

21 mars 2025

NanoDIAG

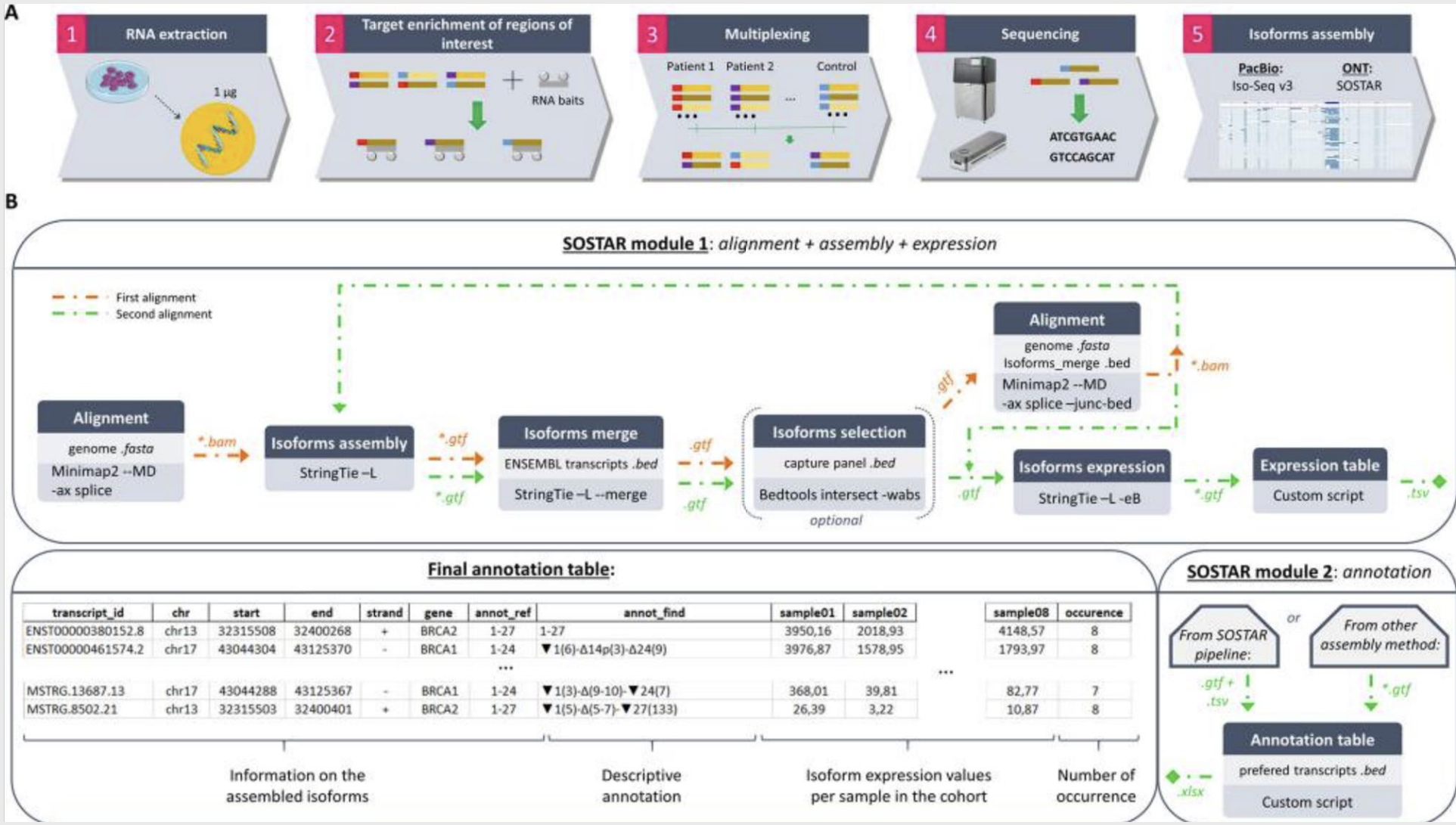
Outline:

- Introduction to NanoDIAG
- Various Bioinformatics topics of interest
  - DNA pipelines
  - RNA pipelines
  - Classifiers & ultrafast diagnostics
- Conclusion

RNA pipelines

- awesome-nanopore : https://github.com/GoekeLab/awesome-nanopore

- https://epi2me.nanoporetech.com/wfindex/ (NF)



C Aucouturier et al. **Fine mapping of RNA isoform diversity using an innovative targeted long-read RNA sequencing protocol with novel dedicated bioinformatics pipeline.** BMC Genomics. 2024

Victor RENAULT
Julien MASLIAH PLANCHON

21 mars 2025

NanoDIAG

Outline:

- Introduction to NanoDIAG
- Various Bioinformatics topics of interest
  - DNA pipelines
  - RNA pipelines
  - Classifiers & ultrafast diagnostics
- Conclusion

Classifiers

- Usual questions :
  - Data (sharing via NanoDIAG)
  - Marker selection
  - Model selection

Classifiers

- Usual questions :
  - Data (sharing via NanoDIAG)
  - Marker selection
  - Model selection

➡

- select most variable CpGs (*Brian Reilly et al.* 2019)
- feature selection with 1st stage classifier (RF [*David Capper et al.* 2018], xgboost)
- differential methylation (*S Peter Wu et al.* 2017)

Classifiers

- Usual questions :
  - Data (sharing via NanoDIAG)
  - Marker selection
  - Model selection

➡

- select most variable CpGs (*Brian Reilly et al.* 2019)
- feature selection with 1st stage classifier (RF [*David Capper et al.* 2018], xgboost)
- differential methylation (*S Peter Wu et al.* 2017)

- Nanopore ➡ sequenced CpGs are sparse, random and poorly covered…

Classifiers

- Usual questions :
  - Data (sharing via NanoDIAG)
  - Marker selection
  - Model selection

➡

- select most variable CpGs (*Brian Reilly et al.* 2019)
- feature selection with 1st stage classifier (RF [*David Capper et al.* 2018], xgboost)
- differential methylation (*S Peter Wu et al.* 2017)

- Nanopore ➡ sequenced CpGs are sparse, random and poorly covered…

One possibility :

- consider the intersection *I* of a large set of markers (eg EPIC) with a current Nanopore
- Train a model using markers *I*
- Predict on current Nanopore

Classifiers

- Usual questions :
  - Data (sharing via NanoDIAG)
  - Marker selection
  - Model selection

➡

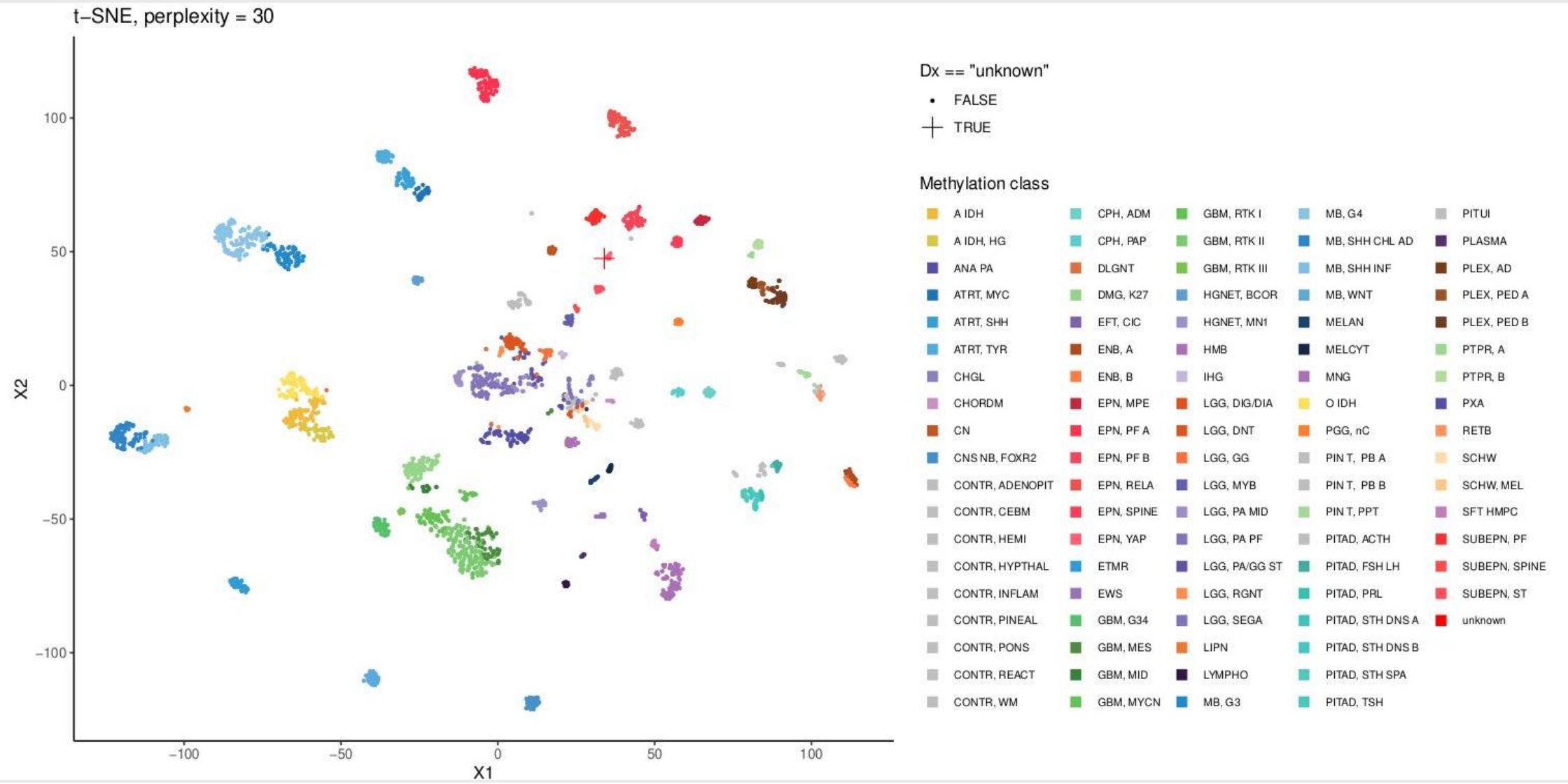
- select most variable CpGs (*Brian Reilly et al.* 2019)
- feature selection with 1st stage classifier (RF [*David Capper et al.* 2018], xgboost)
- differential methylation (*S Peter Wu et al.* 2017)

- Nanopore ➡ sequenced CpGs are sparse, random and poorly covered…

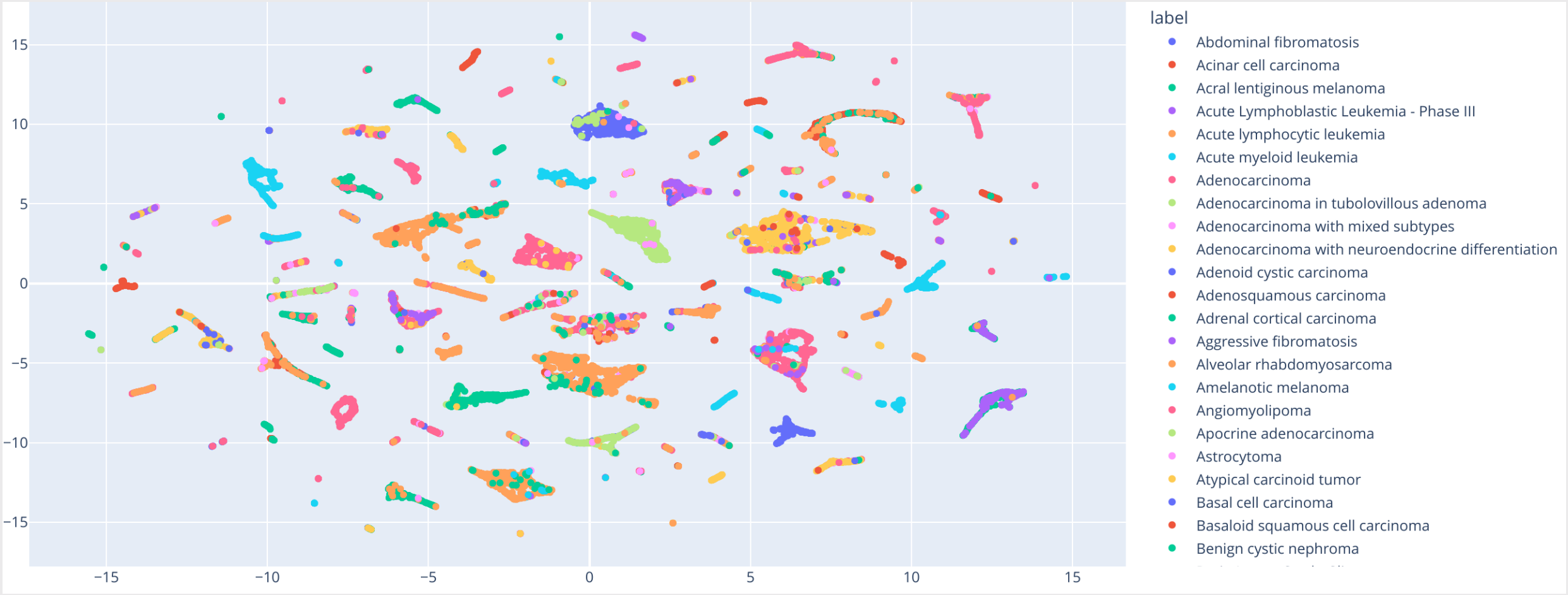
One possibility :
- consider the intersection *I* of a large set of markers (eg EPIC) with a current Nanopore
- Train a model using markers *I*
- Predict on current Nanopore

Luis P Kuschel et al. **Robust methylation-based classification of brain tumours using nanopore sequencing.** Neuropathol Appl Neurobiol. 2023

Classifiers



Classifiers – TCGA

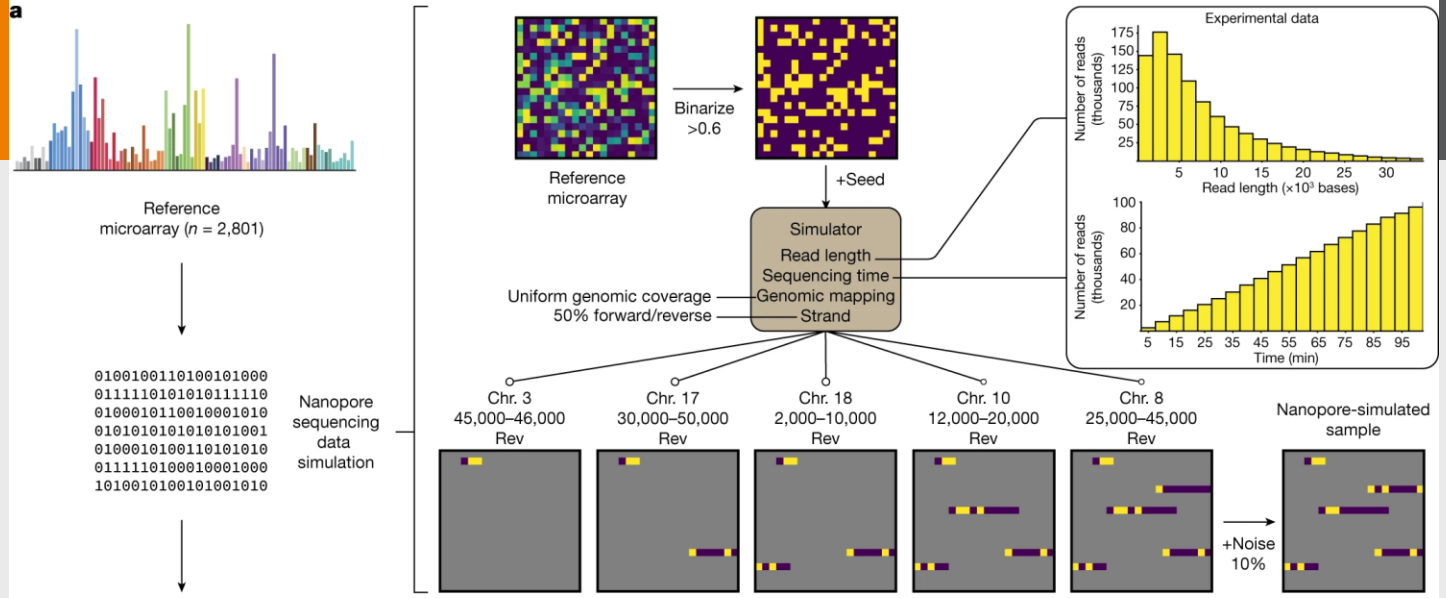


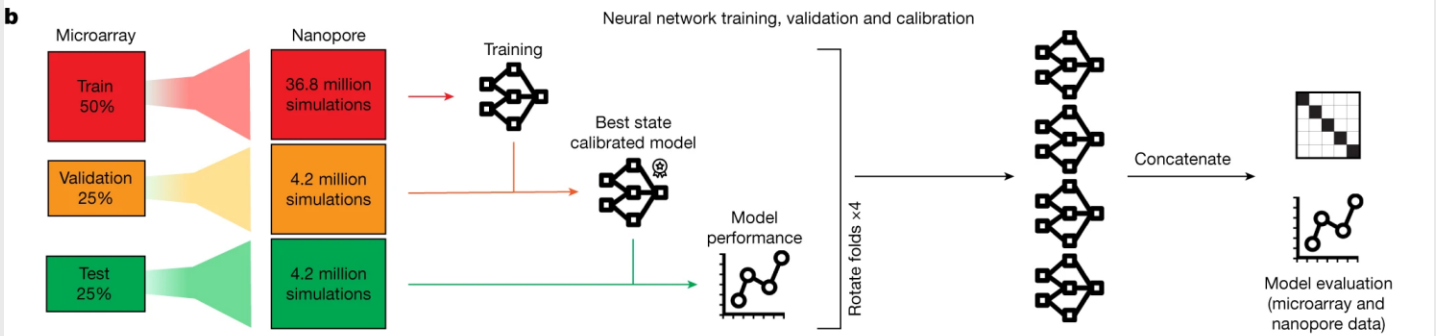
Classifiers – test ctDNA

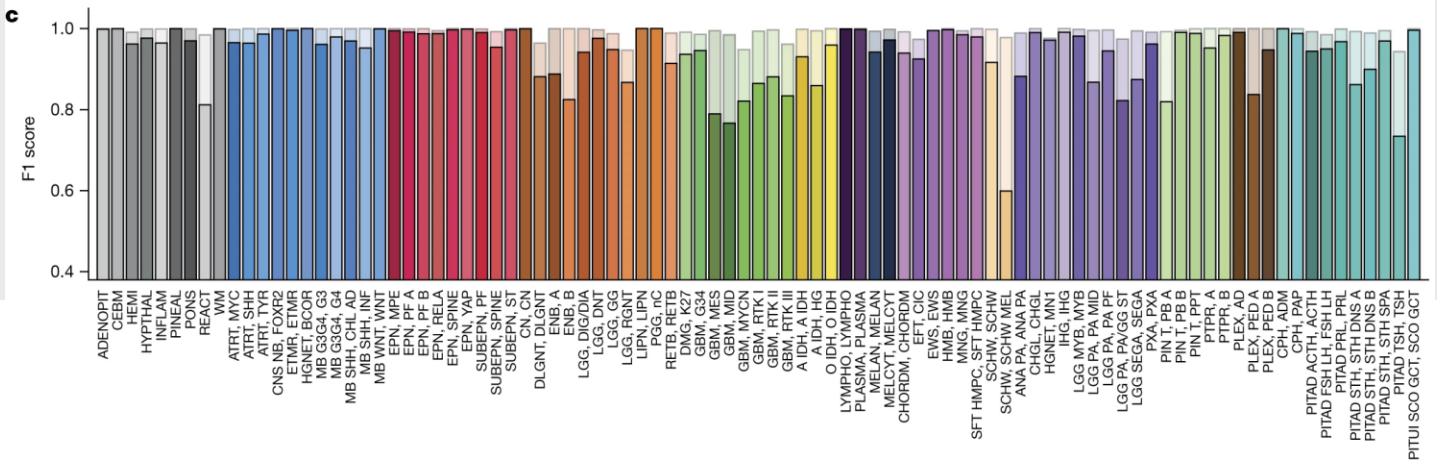
Ultrafast diagnostics - Sturgeon

- *C Vermeulen et al.* **Ultra-fast deep-learned CNS tumour classification during surgery.** Nature 2023

- ROBIN (Rapid nanopOre Brain intraoperatIve classificatioN) : https://github.com/looselab/robin



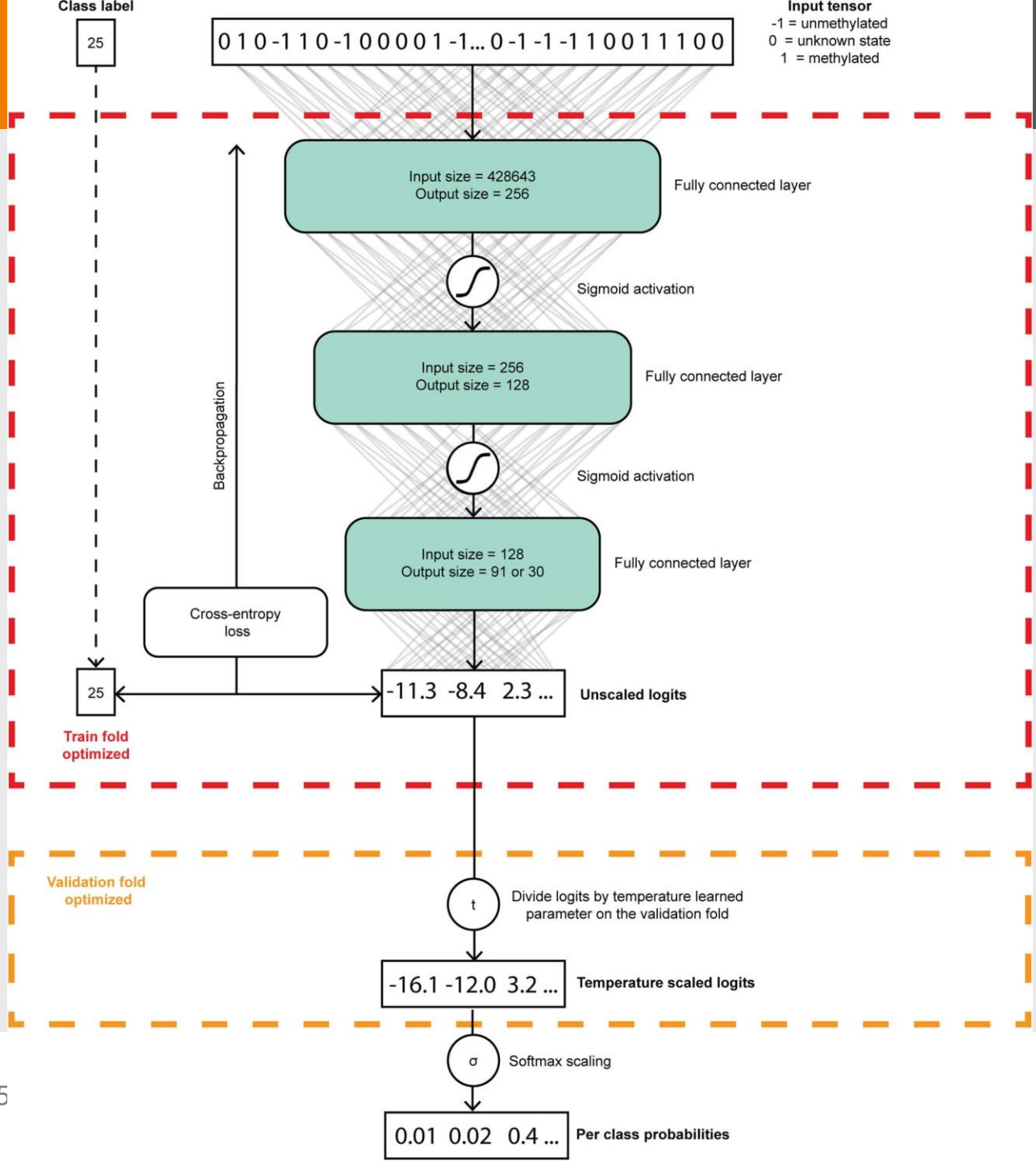




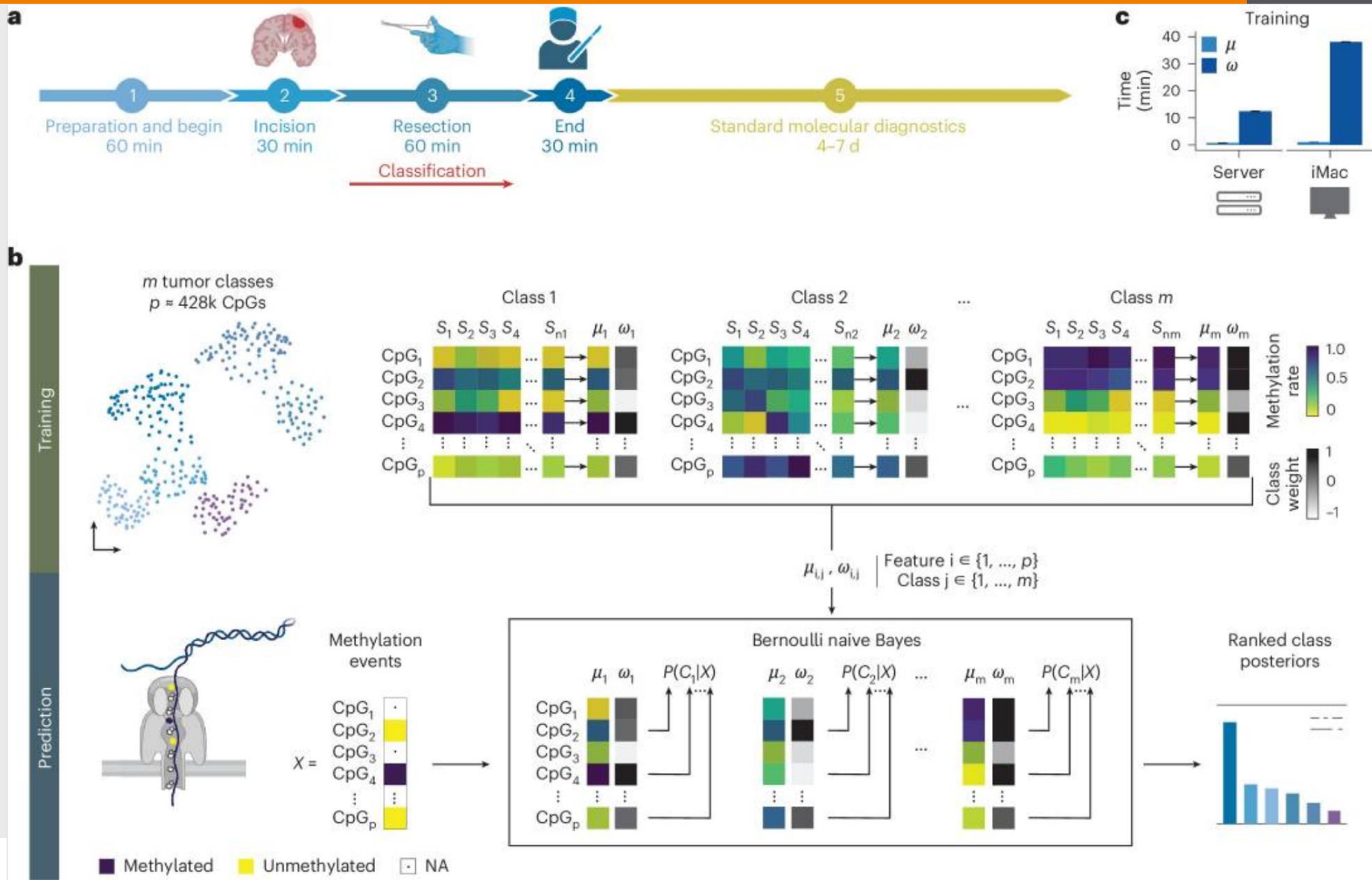
Ultrafast diagnostics - Sturgeon

- *C Vermeulen et al.* **Ultra-fast deep-learned CNS tumour classification during surgery.** Nature 2023

- ROBIN (Rapid nanopOre Brain intraoperatIve classificatioN) : https://github.com/looselab/robin



Ultrafast diagnostics - MethyLYZR



Björn Brändl et al. **Rapid brain tumor classification from sparse epigenomic data.** Nat Med. 2025

Ultrafast diagnostics - Others

- *Francesco E Emiliani et al.* **Nanopore-based random genomic sampling for intraoperative molecular diagnosis.** Genome Med. 2025

- *Maximilian Evers et al.* **Rapid intraoperative amplicon sequencing of CNS tumor markers.** Comput Struct Biotechnol J. 2024

Björn Brändl et al. **Rapid brain tumor classification from sparse epigenomic data.** Nat Med. 2025

Victor RENAULT
Julien MASLIAH PLANCHON

21 mars 2025

NanoDIAG

Outline:

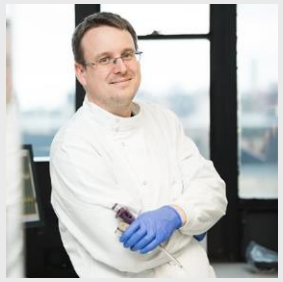
- Introduction to NanoDIAG
- Various Bioinformatics topics of interest
  - DNA pipelines
  - RNA pipelines
  - Classifiers & ultrafast diagnostics
- Conclusion

ENSEMBLE, PRENONS
LE CANCER DE VITESSE

institut Curie

Thank you for your attention !
Contact : victor.renault@curie.fr

- **Merci** à :
  - l'Unité de Génétique Somatique de l'institut Curie :
    - **Julien Masliah Planchon**
    - **Mathilde Filser**
  - l'Unité de bioformatique Clinique de l'institut Curie :
    - **Kévin Merchadou**
    - **Eléonore Frouin**
    - **Siann Chalvin**



- **Journée NanoDIAG** (en Anglais) le **24/06/2025 à Curie** :
  - KeyNote : **Matthew Loose**, University of Nottingham, **pipeline ROBIN** pour la classification des tumeurs cérébrales, y compris en **extemporané**

  - KeyNote : **Eddy de Boer**, UMC Groningen, **panel de 471 gènes** (maladies neurologiques et des épilepsies) avec gènes difficiles à mapper en short read + quelques gènes importants dans ces pathologies pour la méthylation
  - **Appel à contributions + Formulaire d'inscription** : https://shorturl.at/gwe9q