

Séquençage nanopore



Henry Xing
16 février 2021

Plan de la présentation

- Introduction
- Transcriptomique
- Séquençage nanopore
- Étapes bioinformatiques
- Conclusion

C'est quoi le séquençage nanopore

Introduction

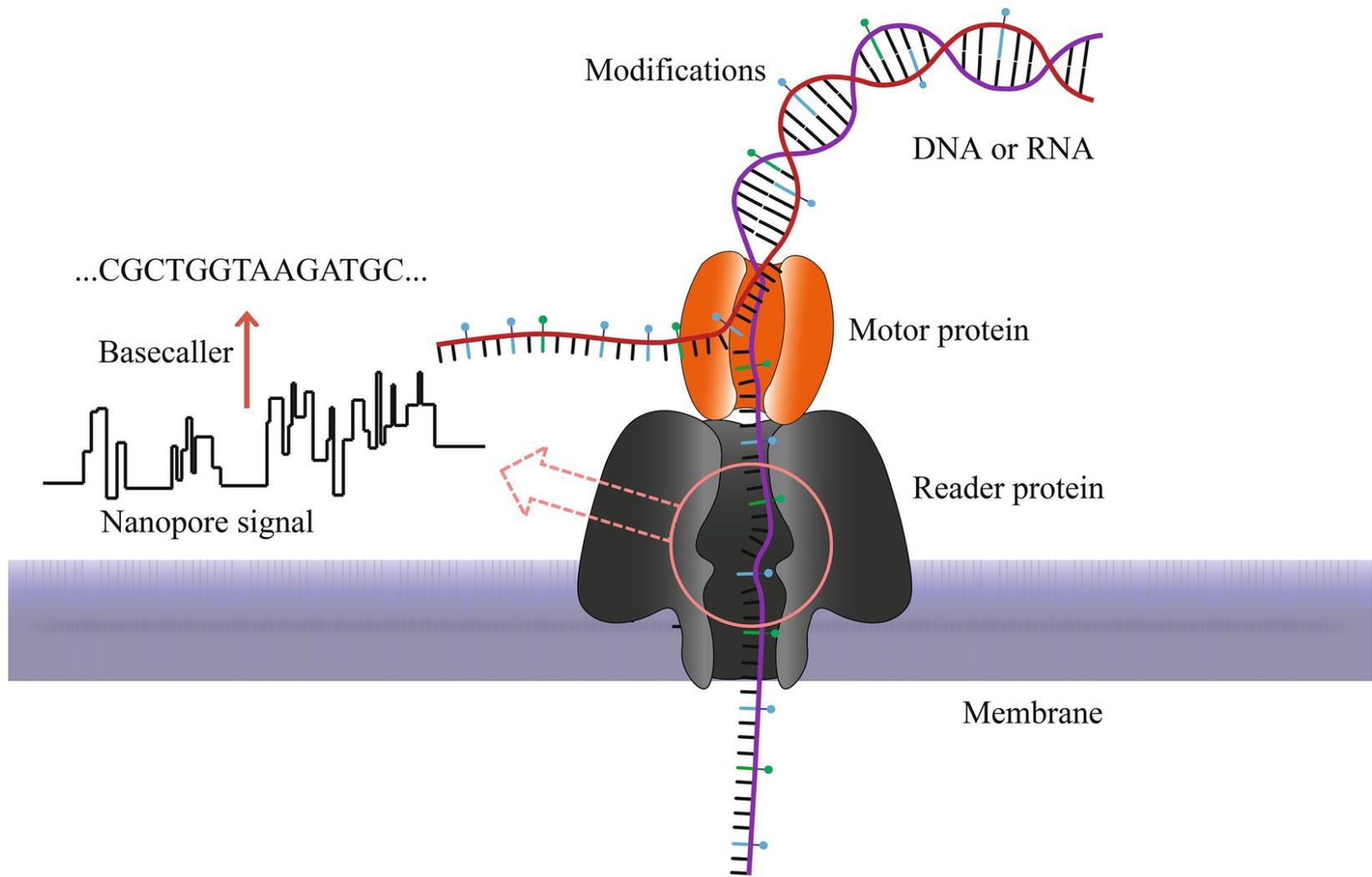
Transcriptomique

Nanopore

Bioinformatique

Conclusion

- Utilisation d'une pore de membrane pour faire passer un brin d'ADN / ARN
- Un courant est induit à travers la membrane
- Les bases passe dans la pore - translocation
- Mesure du courant dans la nanopore



Fonctionnement du séquençage nanopore

Contexte

Introduction

Transcriptomique

Nanopore

Bioinformatique

Conclusion

- L'idée est décrite dans plusieurs laboratoires dans les années 90
- Plusieurs obstacles:
 - Résolution et distinction des nucléotides
 - Vitesse de passage
 - Utilisation à grande échelle

Obstacles majeures

Introduction

Transcriptomique

Nanopore

Bioinformatique

Conclusion

- Dans les premiers papiers, une séquence devait être repassée 100 fois
 - 1-5 μ s par base, pas assez de temps pour obtenir un signal
 - 12-15 bases à la fois dans la région du senseur
 - Différence minimale de courant entre chaque base (A,T,G,C,U)

Contexte

Introduction

Transcriptomique

Nanopore

Bioinformatique

Conclusion

- MinION fait son entrée en 2014
- Premier génome (E. coli) assemblé de novo avec 99.4% de précision en 2015
- Le basecalling (conversion du signal en base AT[U]GC) est basé sur l'apprentissage machine, car beaucoup de bruits
- Encore des améliorations publiées, technologie en progression

Central Dogma of Biology

'Omics

DNA



Genomics

Transcription



RNA



Transcriptomics

Translation



Protein



Proteomics

Transcriptomique

Introduction

Transcriptomique

Nanopore

Bioinformatique

Conclusion

- L'ensemble des ARN à un moment donné
- Expression différentielle
- ARN codant et non-codant
 - 2% séquences codantes
- Différents isoformes du même gène

La complexité biologique au delà du code génétique



Nombre de cellules

La complexité biologique au delà du code génétique



Nombre de gènes

La complexité biologique au delà du code génétique



Codants vs non-codants

Nanopore et transcriptomique

Introduction

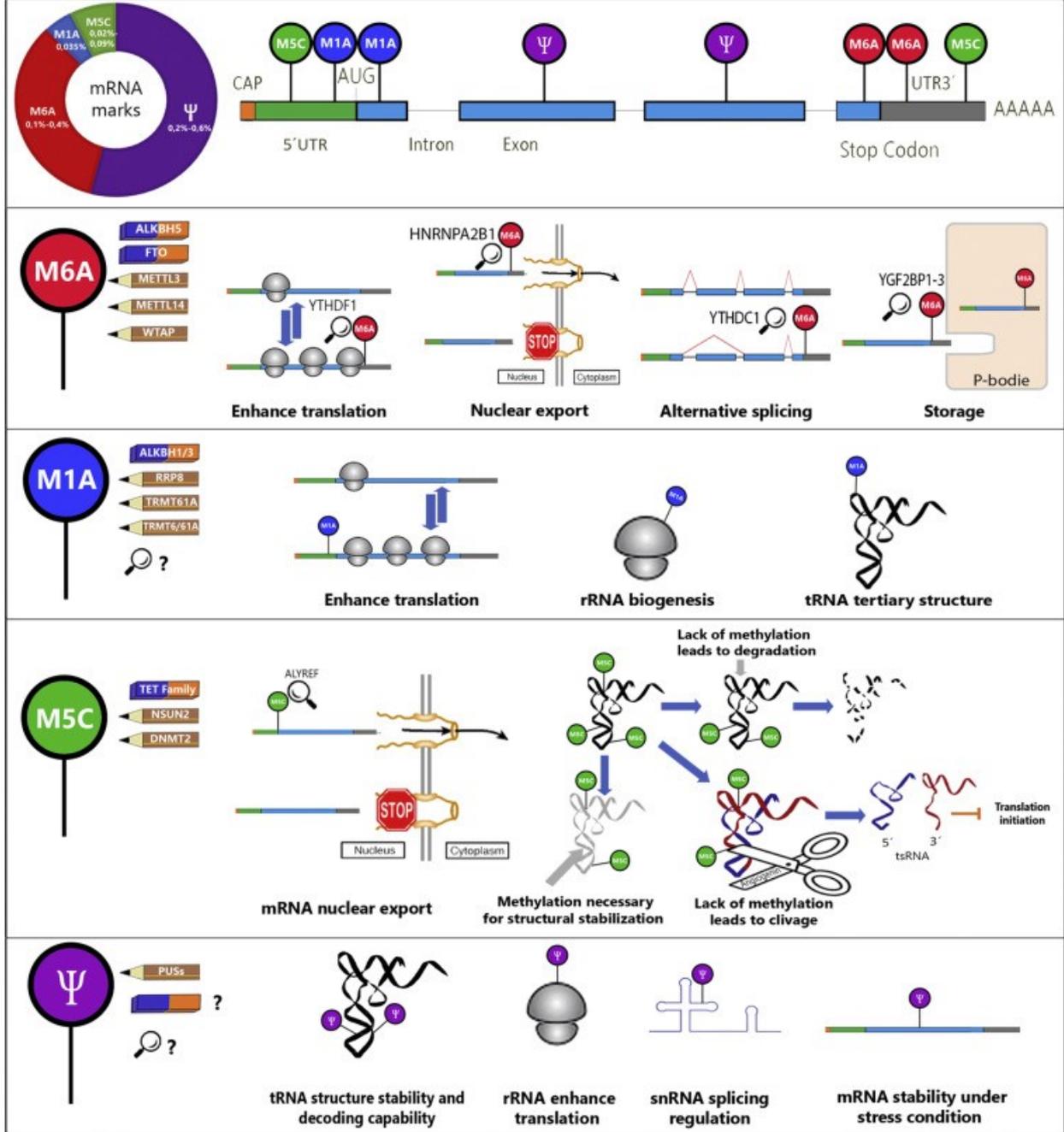
Transcriptomique

Nanopore

Bioinformatique

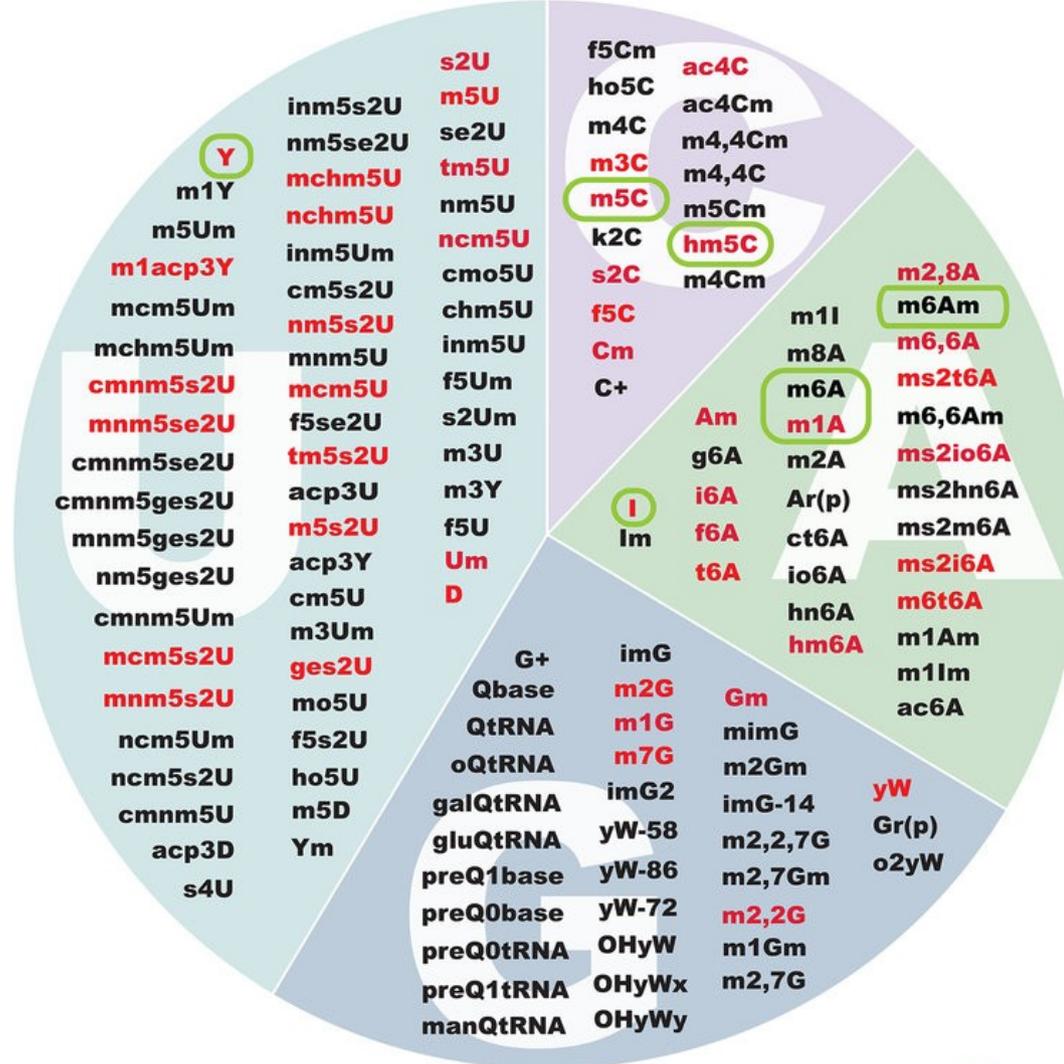
Conclusion

- Pourquoi nanopore?
 - Long et ultra-long reads (record de 2.3Mb)
 - Détection des modifications sur les bases
 - Pas de RT-PCR*
 - Enlève les biais d'amplification
 - Simple à utiliser



Exemples de modifications

<https://www.sciencedirect.com/science/article/pii/S0925477318300054>



Disease related

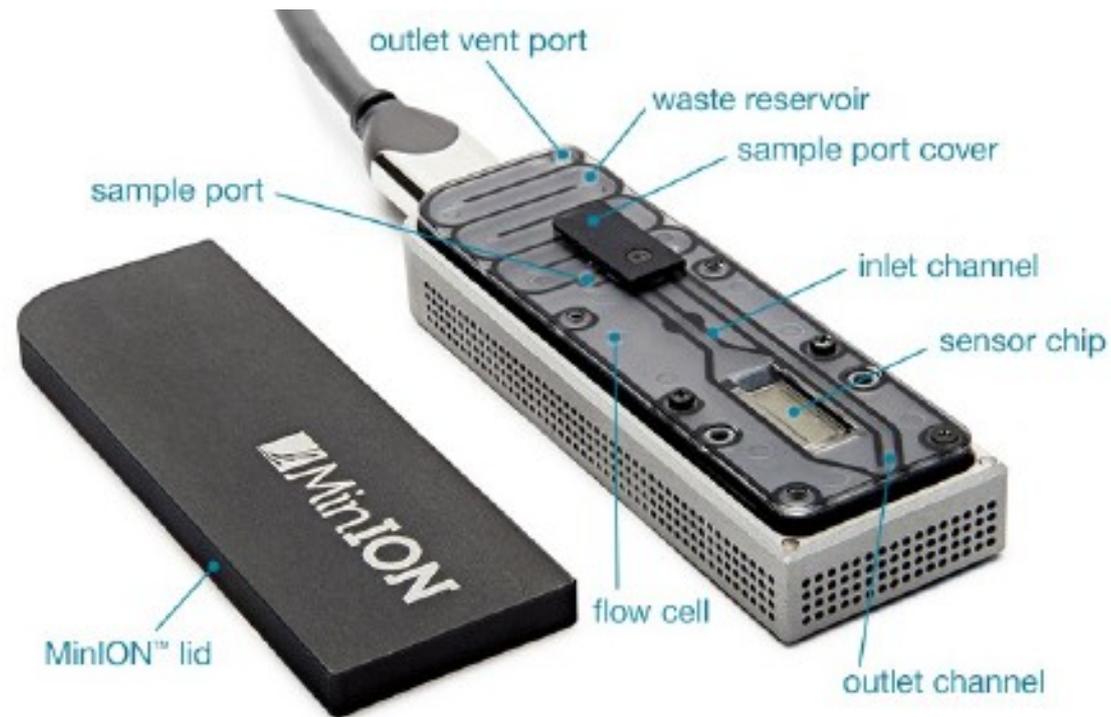
Detection technique available

Exemples de modifications

<https://rnajournal.cshlp.org/content/23/12/1754>

Nanopore

- Flow cell typique pour le séquençage



Flow cell

- 512 canaux dans le ‘sensor chip’
- Chaque canaux = 4 pores
 - Durant le séquençage, les pores peuvent bloquer
- Scan pour décider quelle pore

Nanopore

Introduction

Transcriptomique

Nanopore

Bioinformatique

Conclusion

- Biologique ou synthétique
- Diamètre de quelques nms
- Possible de rajouter des adaptations:
 - Sonde d'ADN pour comparer avec une référence
 - Moteur moléculaire pour l'ADN
 - Autres ligands

Maintenant: Biologiques

Introduction

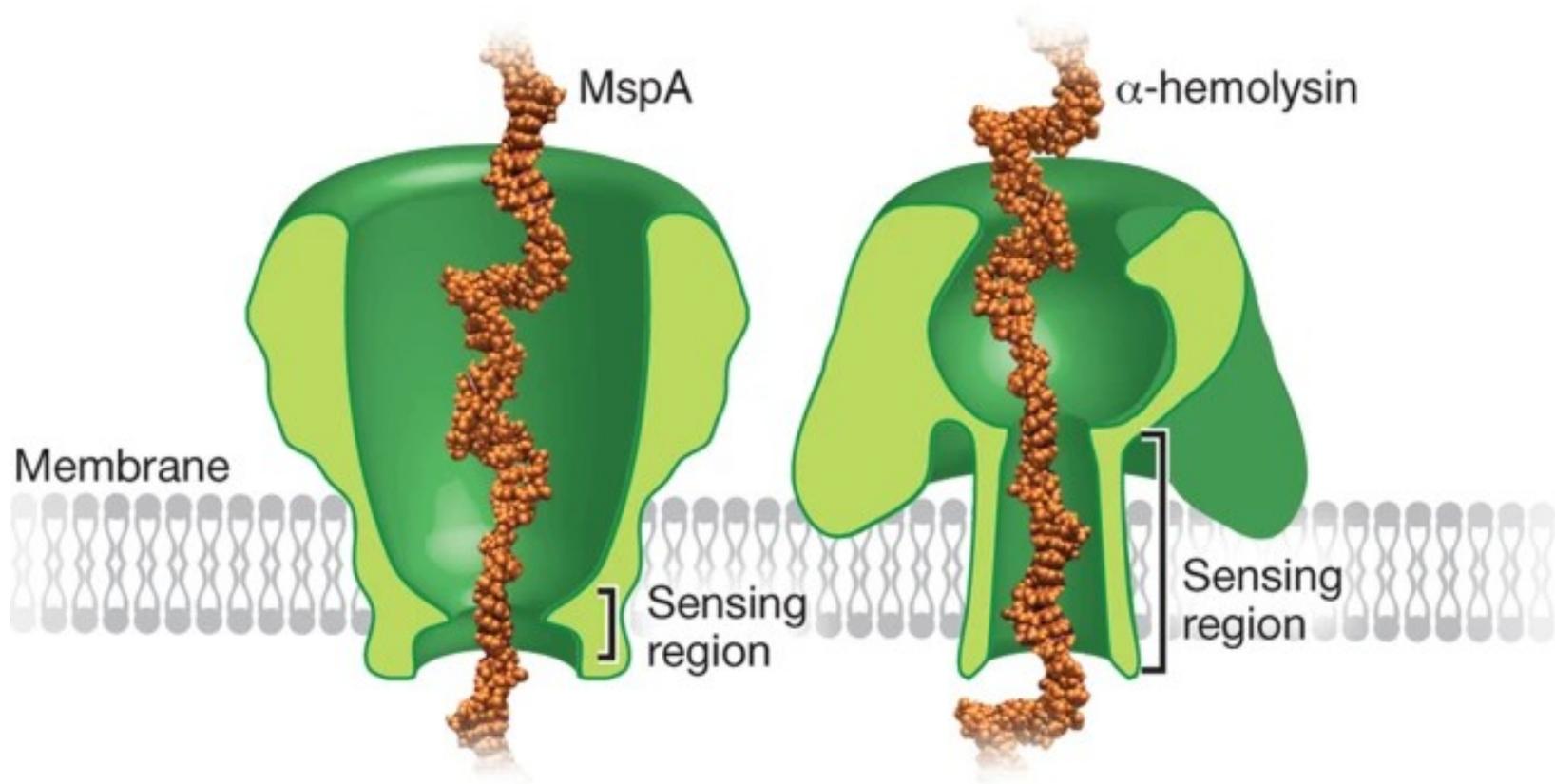
Transcriptomique

Nanopore

Bioinformatique

Conclusion

- Pores (provenant surtout de bactéries) sur une membrane bilipidique
- Avantages:
 - Peu coûteux, reproductible
 - Facile à modifier
- Inconvénients
 - Limitées en conditions
 - Voltage, pH, diamètre
 - Beaucoup de bruits



<https://www.nature.com/articles/nbt.3423>

Pores biologiques

Future: Synthétiques

Introduction

Transcriptomique

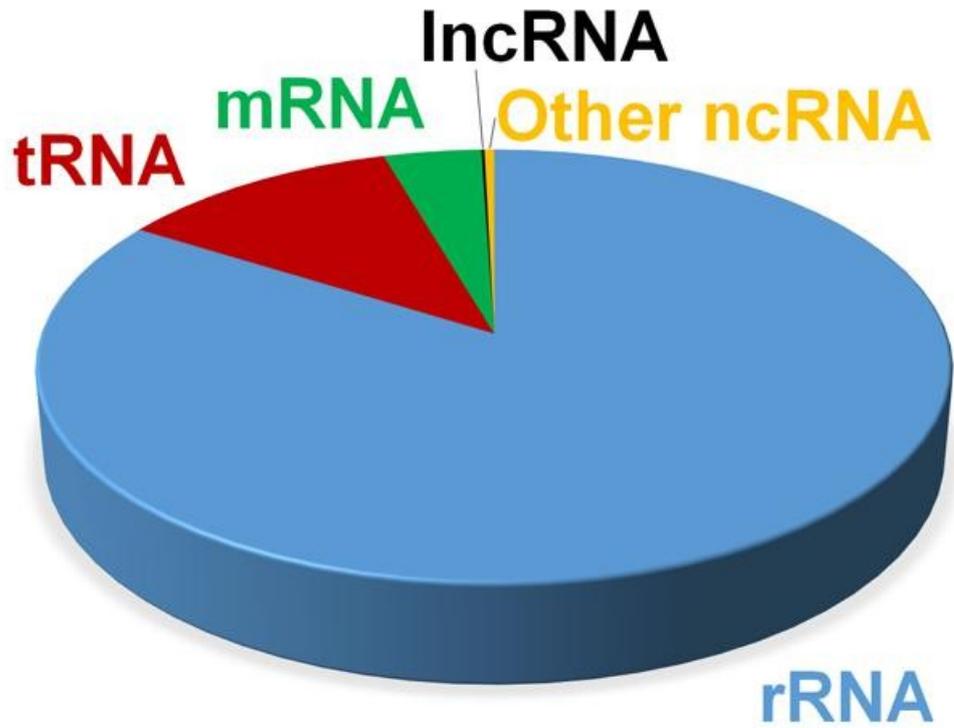
Nanopore

Bioinformatique

Conclusion

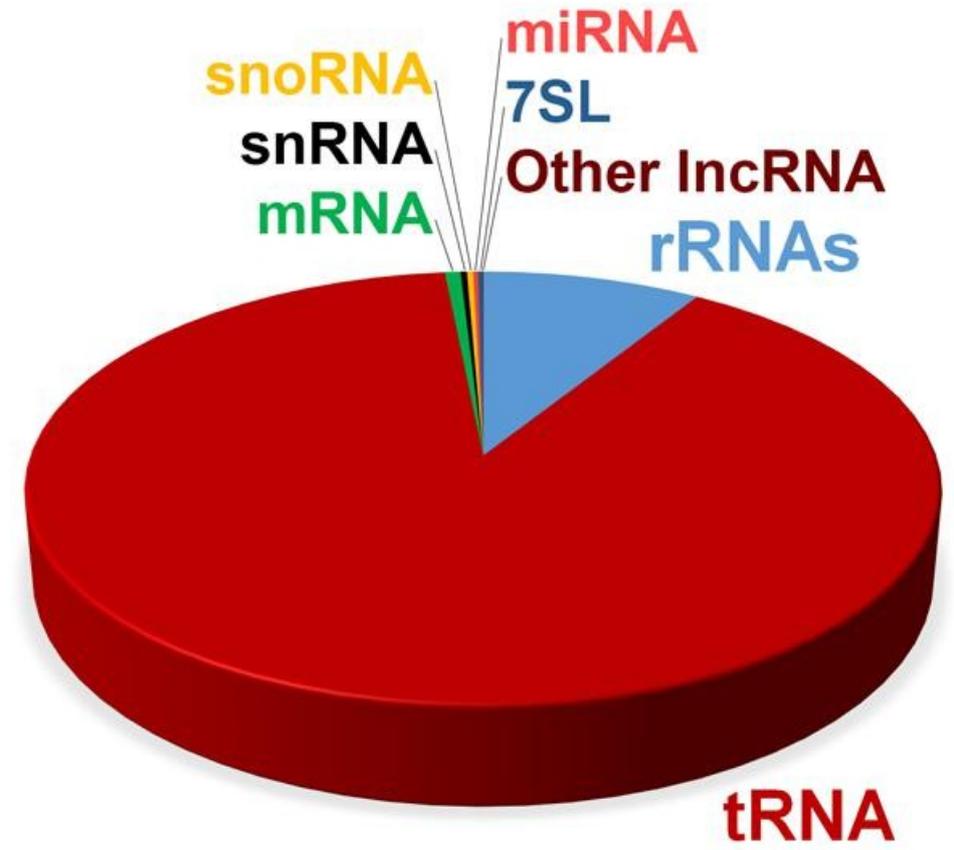
- Pore en graphène sur membrane silicone
- Avantages:
 - Production de masse
 - Compatibilité avec les techniques de mesure
- Inconvénients
 - Difficile de reproduire
 - Conditions biochimiques
- Hybride: membrane synthétique avec pore biologique

A



RNA by mass

B



RNA by number of molecules

<https://www.frontiersin.org/articles/10.3389/fgene.2015.00002/full>

L'abondance relative des ARN

Séquençage

Introduction

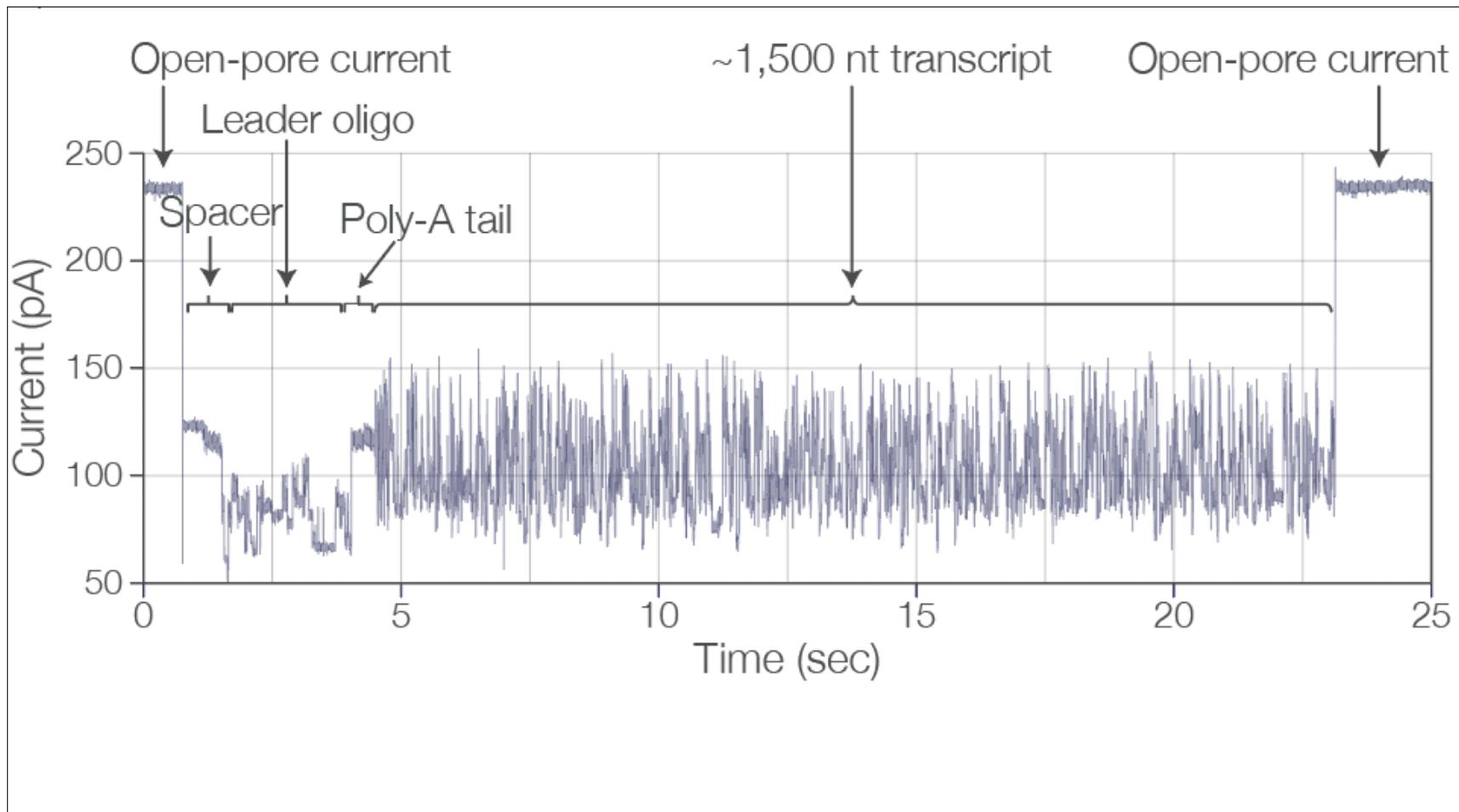
Transcriptomique

Nanopore

Bioinformatique

Conclusion

- Préparation de librairie (relativement court)
 - Primer, adaptateur, barcoding, etc
- Déplétion de l'ARNr et ARNt
- capture polyA
- Charger la flowcell, une run dure autant de temps que voulu
 - Pratique et portable
 - ~100-300 reads/minute
- Produit des fichiers Fast5



Signal nanopore d'ARN

<https://nanoporetech.com/resource-centre/nanopores-allow-direct-sequencing-rna-strands#image1&modal=image1>

Pipeline général

Introduction

Transcriptomique

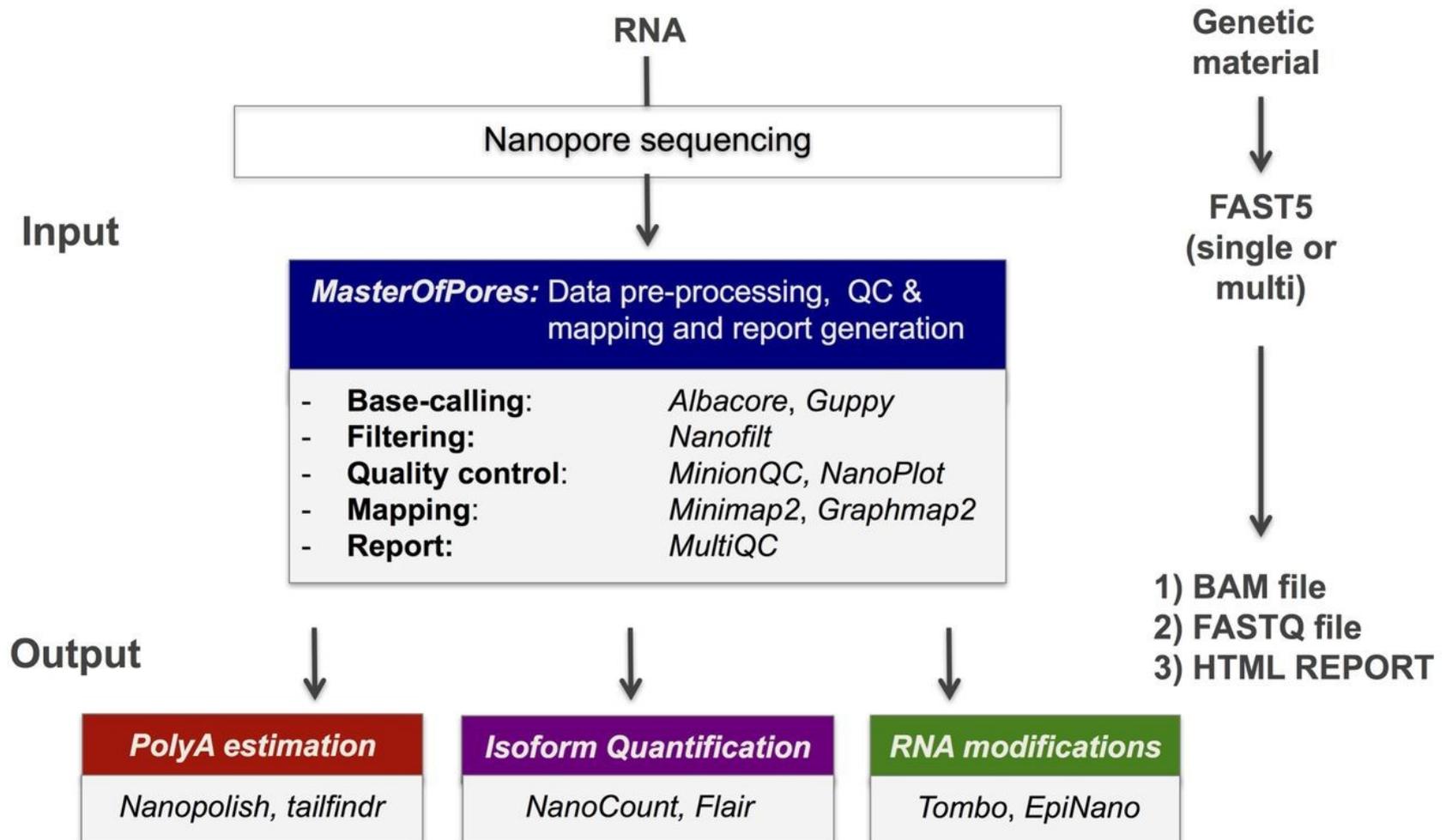
Nanopore

Bioinformatique

Conclusion

- Basecalling
 - Conversion des signaux en bases
- Contrôle qualité
 - Trimming
 - Démultiplexage
 - Chimères, faux positifs
- Mapping, assemblage de génomes
 - Mapping si avec génome de référence
 - Assemblage de novo sans référence
- Polissage

NextFlow workflow: *MasterOfPores*



<https://doi.org/10.1101/818336>

Analyses downstream

Introduction

Transcriptomique

Nanopore

Bioinformatique

Conclusion

- Détection de variants
- Phylogénie
 - T-Rex, MegaX
- Épigénétique / épitranscriptomique
- Quantification
 - FeatureCount, Salmon, Stringtie
- Expression différentielle
 - EdgeR, DESeq2

Ressources

Introduction

Transcriptomique

Nanopore

Bioinformatique

Conclusion

- Beaucoup d'outils bioinformatique sur GitHub
 - Manipulation de fichiers
 - SquiggleKit, Bulkvis, nanoporetech
 - Basecalling
 - Guppy, Albacore, Bonito, Rerio
 - Corrections et polissage
 - Nanopolish, Medaka, NanoReviser
 - Analyses downstream
 - ARTIC, Tombo, Nanomod



Type de séquenceurs

Conclusion

Introduction

Transcriptomique

Nanopore

Bioinformatique

Conclusion

- Beaucoup d'applications possibles
 - Portabilité
 - Diagnostic rapide
 - Omiques
- Opportunités de progression
 - Pores synthétiques
 - Outils bioinformatiques
 - Documentation

Merci!