



**HAL**  
open science

## Optimizing the production of the *Anaplasma phagocytophilum* genome, a crucial step in understanding virulence and host tropism

Harivel Thomas, Clotilde Rouxel, Déborah Merda, Anne-Claire Lagrée, Nadia Haddad, Virginie Chesnais, Pierre Deshuillers

### ► To cite this version:

Harivel Thomas, Clotilde Rouxel, Déborah Merda, Anne-Claire Lagrée, Nadia Haddad, et al.. Optimizing the production of the *Anaplasma phagocytophilum* genome, a crucial step in understanding virulence and host tropism. *Bienvenue chez les Microbes - 19ème congrès national de la SFM*, Oct 2024, Lille, France. hal-04746026

**HAL Id: hal-04746026**

**<https://hal.science/hal-04746026v1>**

Submitted on 21 Oct 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.

# Optimiser la production du génome d'*Anaplasma phagocytophilum*, une étape cruciale pour comprendre la virulence et le tropisme d'hôte

Thomas Harivel<sup>1,2</sup>, Clotilde Rouxel<sup>1</sup>, Déborah Merda<sup>2</sup>, Anne-Claire Lagrée<sup>1</sup>, Nadia Haddad<sup>1</sup>, Virginie Chesnais<sup>2</sup> & Pierre Deshuillers<sup>1</sup>

<sup>1</sup>ANSES, INRAE, Ecole nationale vétérinaire d'Alfort (EnvA), UMR BIPAR, Laboratoire de Santé Animale, 94700 Maisons-Alfort, France.

<sup>2</sup>ANSES, SPAAD, Laboratoire de Sécurité Alimentaire, Laboratoire de Santé Animale, 94700 Maisons-Alfort, France.

## 1. *Anaplasma phagocytophilum*

- Bactérie intracellulaire stricte infectant les granulocytes neutrophiles de l'hôte, culture très complexe → **très faible nombre de génomes disponibles** (33 sur NCBI)
- Transmise par piqûre de tiques (*Ixodes* spp) aux **humains (USA et Europe)** et au **bétail (Europe)**. Responsable de l'**anaplasmose granulocytaire (AG)**
- **AG** = Maladie émergente, variabilité des signes cliniques, manque de connaissances des facteurs bactériens conditionnant la virulence et les barrières d'espèces → **importance de l'analyse de génomes complets**



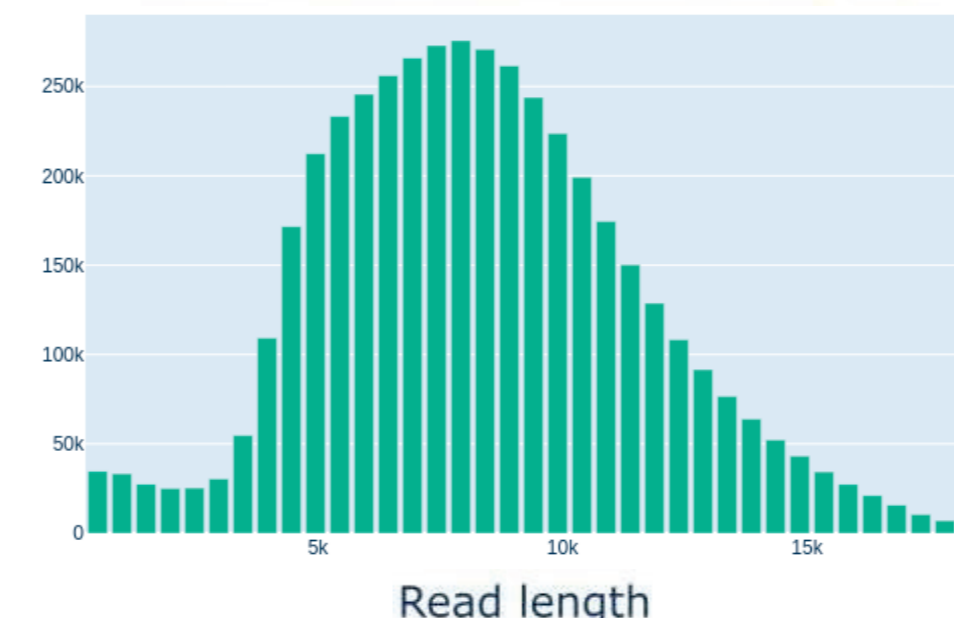
**Objectif :** Réaliser le séquençage des souches au laboratoire par la technologie Nanopore. À partir de ces résultats, mettre au point un pipeline d'analyse bio-informatique pour assembler leur génome

## 2. Séquençage Nanopore



Souche humaine NY18

Souche publiée et cultivable sur cellules HL-60 (humaines)  
Non weighted histogram of read lengths

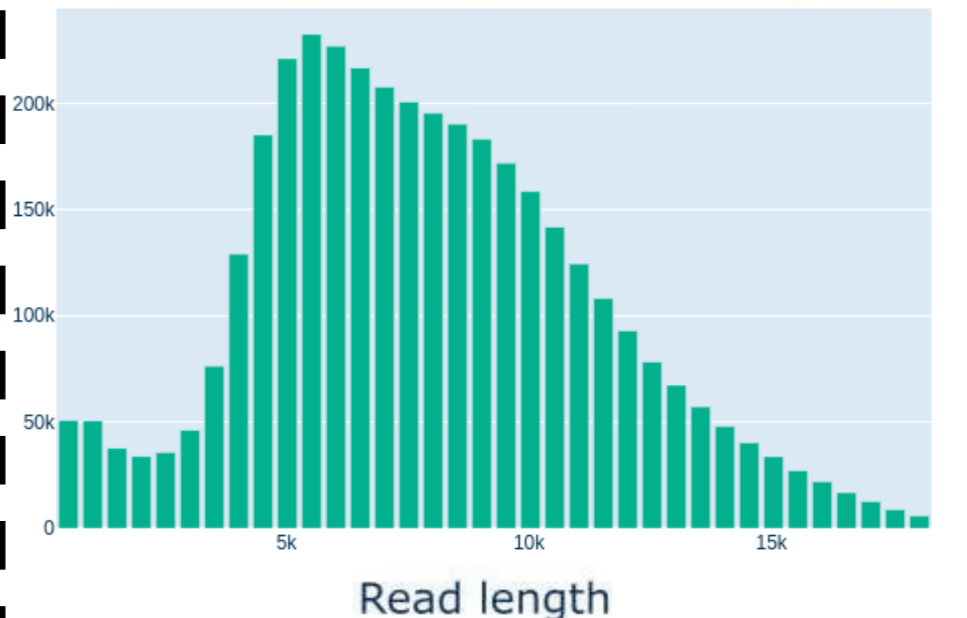


Mean read quality 15.8  
Mean read length 8,112.5  
Number of reads 4,496,201

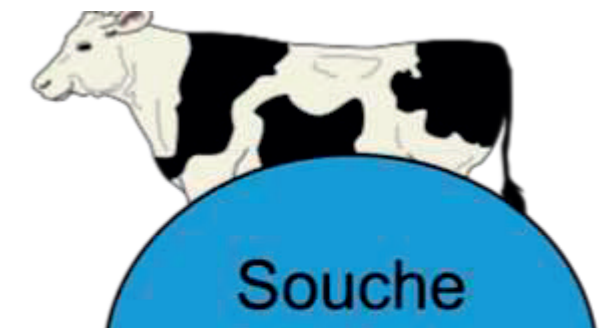


Souche ovine NV20s

Souche publiée et cultivable sur cellules ISE6 (de tiques)  
Non weighted histogram of read lengths

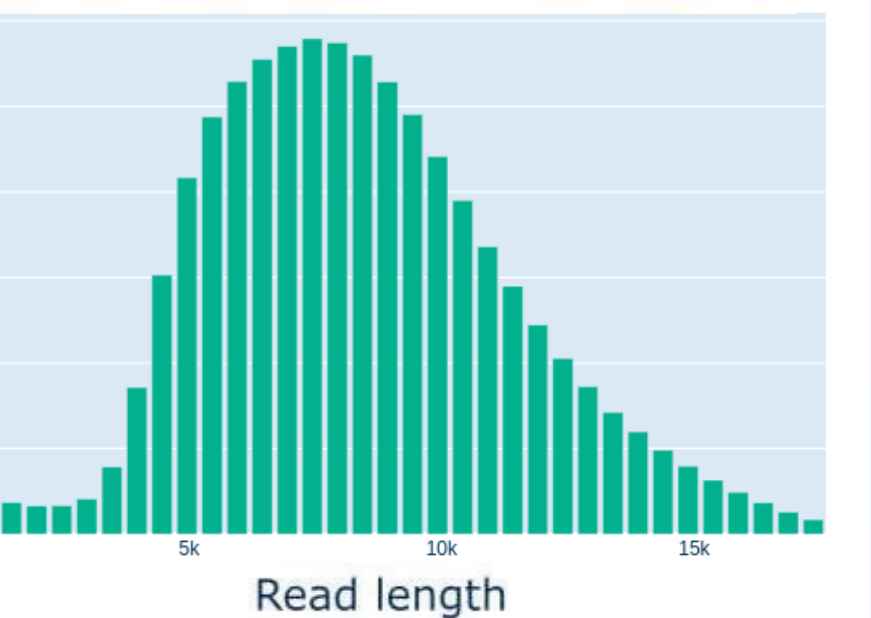


Mean read quality 15.6  
Mean read length 7,687.5  
Number of reads 3,750,906



Souche bovine BV\_10179

Souche publiée et non cultivable (sang infecté)  
Non weighted histogram of read lengths



Mean read quality 16.2  
Mean read length 8,162.4  
Number of reads 4,419,153

**Conclusion :** Reads de bonne qualité et longueur satisfaisante pour les différents échantillons mais **couverture trop importante (> 5000x)** pour NY18 et NV20s. Pour BV\_10179 les reads correspondant à *A. phagocytophilum* sont extrêmement minoritaires et entraînent une **couverture trop faible (≈10x)**

## 3. Analyses Bio-informatiques

NY18 NV20s

**Problématique :**

Comment assembler les génomes dont la couverture est trop importante (> 5000x) ?

**Méthodologie :**

Sous-échantillonnage des échantillons afin d'obtenir des couvertures de 10x à 100x



Assemblage des génomes avec Flye



Analyse des distances phylogénétique par Mash

BV\_10179

**Problématique :**

Comment garder les reads correspondant à *A. phagocytophilum* uniquement ?

**Méthodologie :**

Déplétion des reads du génome de l'hôte « *Bos taurus* »



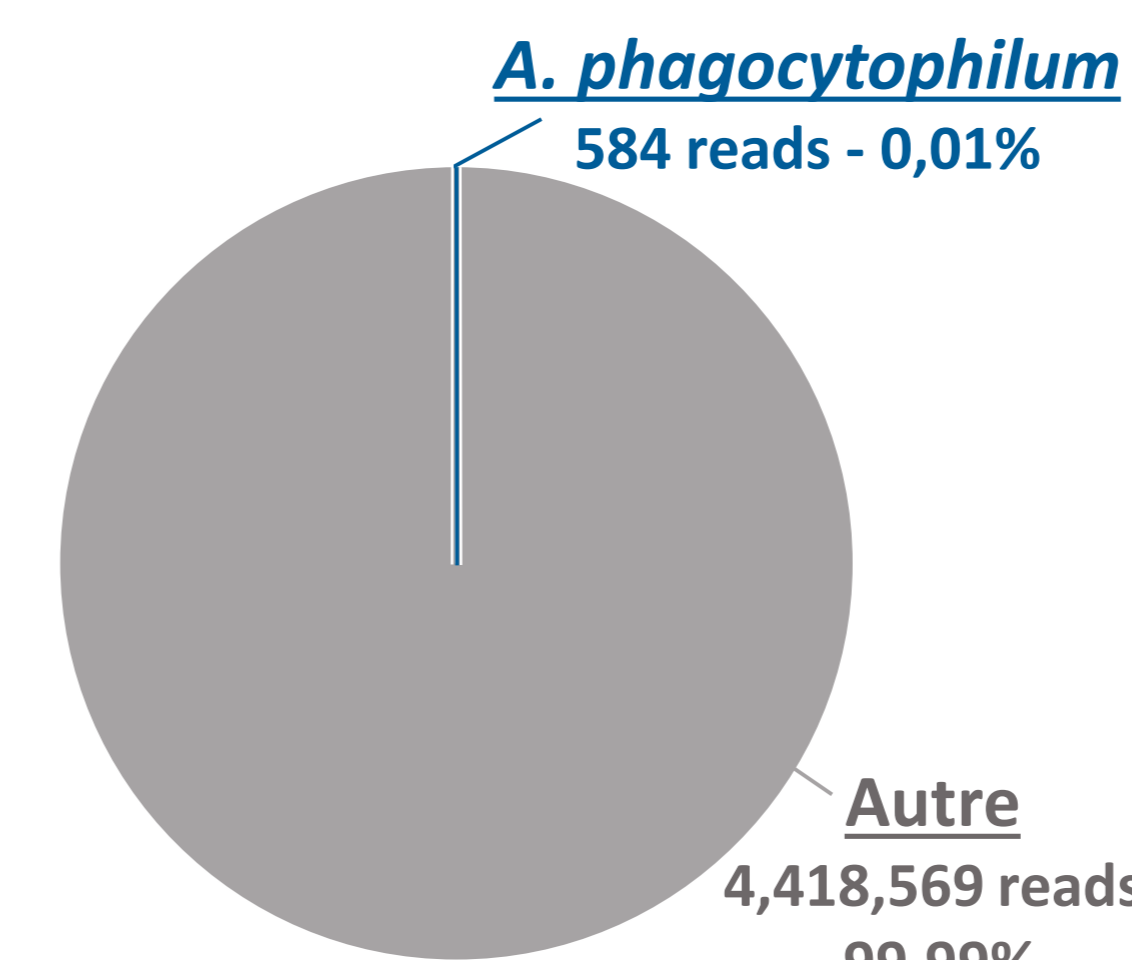
Analyse des fractions de reads correspondants à *A. phagocytophilum*



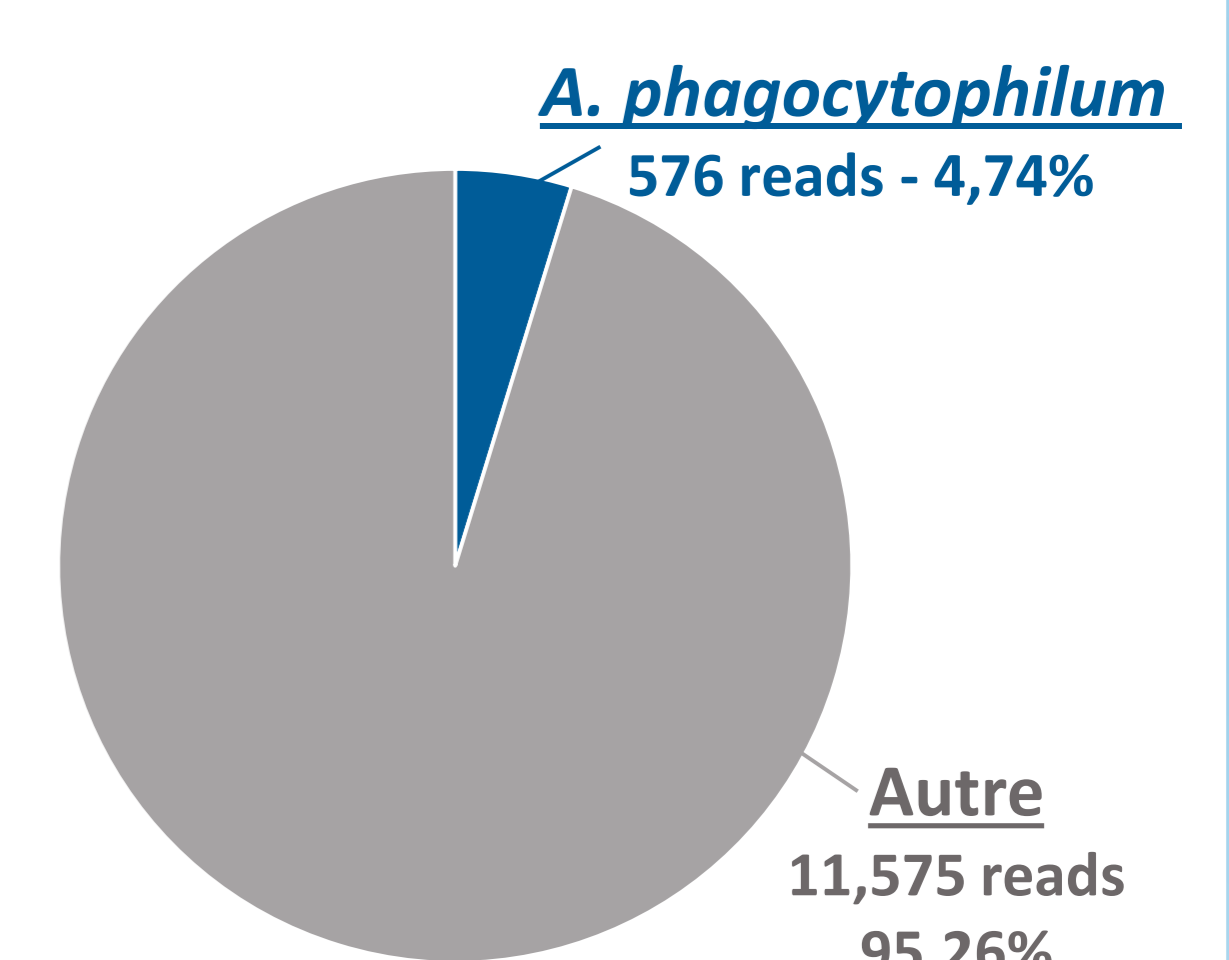
Assemblage du génome avec Flye

## 4.1. Test de déplétion

BV\_10179 brut - Reads



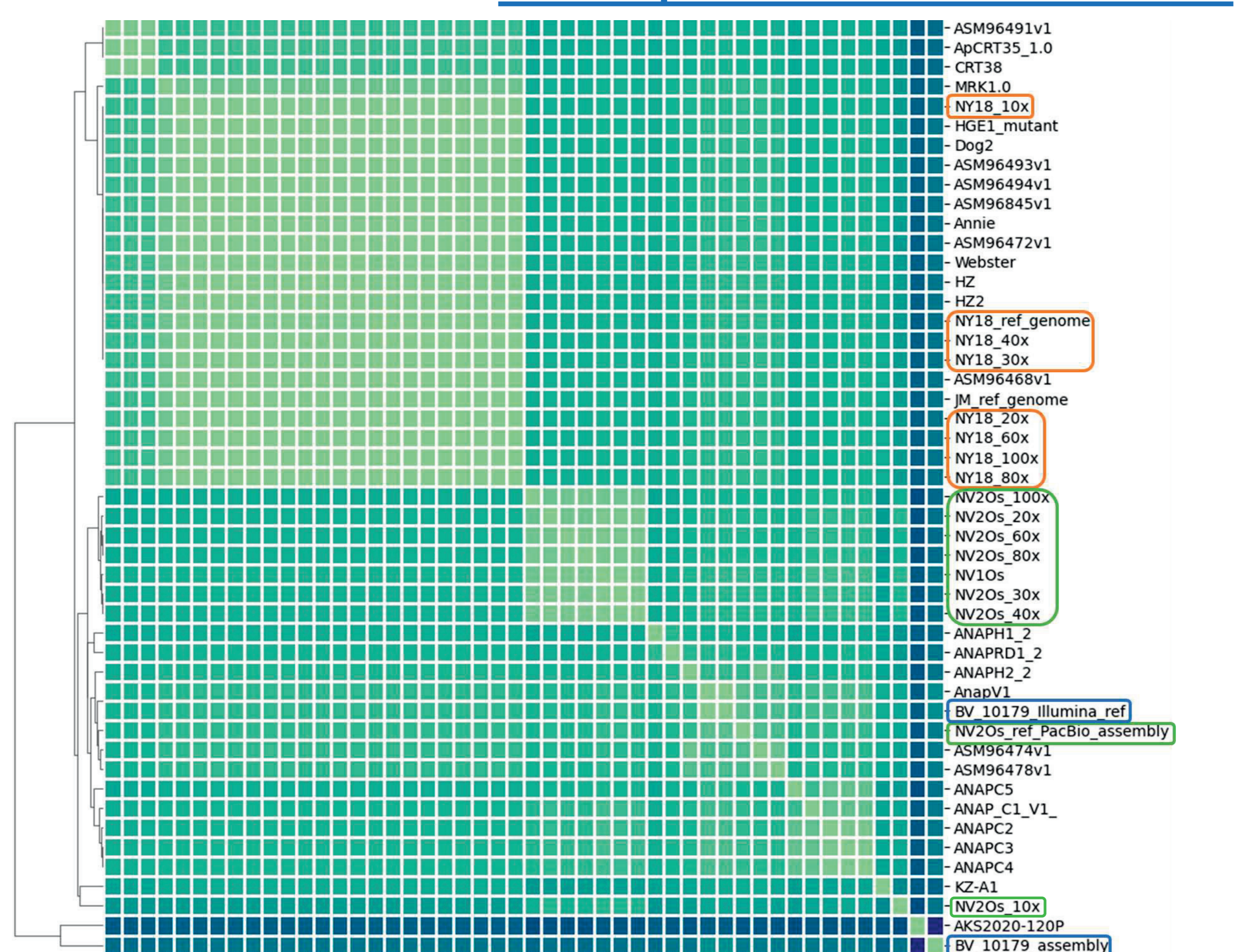
BV\_10179 déplété - Reads



**Conclusion :** La déplétion permet bien d'éliminer une grande quantité de reads ne correspondant pas à *A. phagocytophilum* tout en conservant ceux qui le sont → intérêt pour les échantillons non cultivables

## 4.2. Comparaison des génomes sous-échantillonnés

HeatMap des matrices de distance



NY18 Génomes NY18  
NV20s Génomes NV20s  
BV\_10179 Génomes BV\_10179

Différents  
0.10  
0.08  
0.06  
0.04  
0.02  
0.00  
Identiques

**Distance de Jaccard:**  
Mesure de dissimilarité entre séquences

**Conclusion :** La heatmap montre bien une **clusterisation** des génomes **selon leur origine** et une nécessité d'une **couverture minimale de 20x** pour obtenir un **génome de bonne qualité**

## 5. Conclusions & perspectives

- Obtention de séquences qualitatives (taille et qualité) avec le séquençage Nanopore
- Fonctionnement correct de la déplétion sur les échantillons non cultivables (BV\_10179)
- Nécessité d'une **couverture minimale de 20x** pour obtenir un **génome complet** et de **bonne qualité**
- Application de la méthodologie de déplétion du génome de l'hôte en temps réel (**adaptive sampling**) à d'autres échantillons non cultivables
- Obtention de nouveaux génomes issus de **différentes espèces hôtes** et **zones géographiques** → identification de **marqueurs génétiques du tropisme d'hôte** et de la **virulence** menant au développement d'outils diagnostiques