

TP 7 - Méthodes des données de séquençage

Données issues de séquenceurs :

Séquençage 454 : 400 nucléotides.

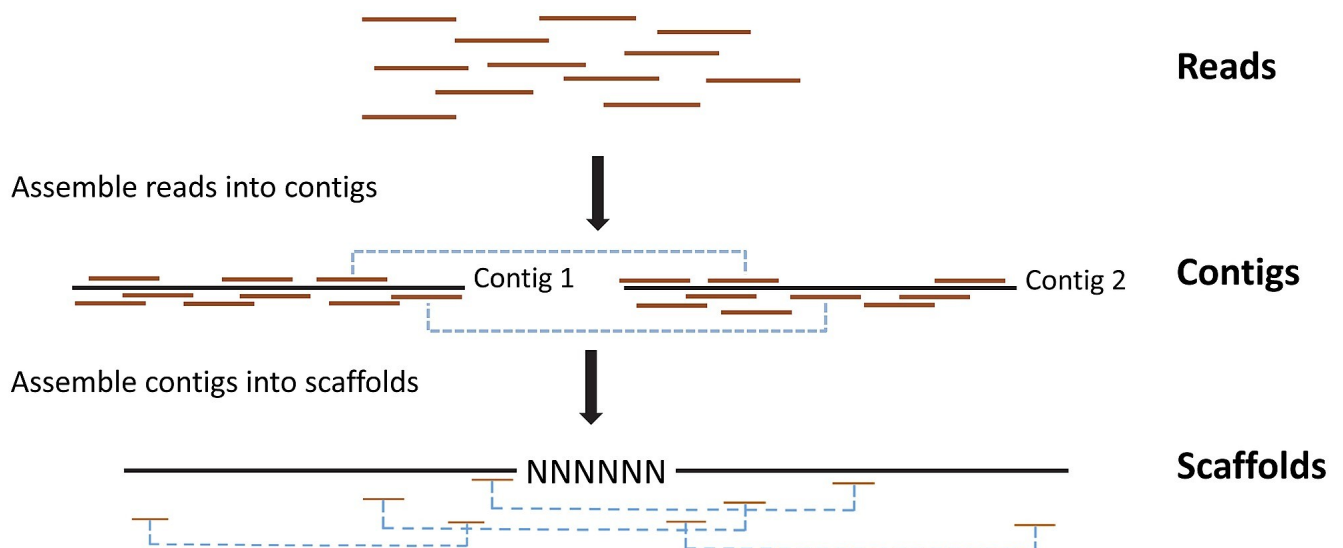
Illumina NGS : séquences de 100 à 300 nucléotides.

Limitations :

Séquençage 454 impossible de déterminer avec certitude la longueur de ces répétitions si elles sont plus longues que 6 à 10 nucléotides.

Illumina NGS pour chaque lecture on ne connaît pas sa position ni son orientation dans le génome dont elle est issue.

L'utilisation des lectures permet la formation de scaffolds qui peuvent être utilisés comme squelettes pour la finition de l'assemblage et à la résolution de répétitions.



Différences entre reads et scaffolds © 1

Enjeu : Alignement des lectures sur le génome efficacement.

