

Présentation Alternance

William Amory
M1 BI-IPFB Université de Paris

24/01/2022



Section 1

CEA - Genoscope



CEA - Genoscope

CEA (Commissariat à l'énergie atomique et aux énergies)

- ???? Salarié
- ??? de labo
- créé en ????

Genoscope (centre national de séquençage)

- ???? salarié
- Créé en ????
 - Séquençage du chromosome 14 humain pour le projet ????
 - Plus grand centre de séquençage français et européen



Organigramme CEA - Genoscope - LBGB



Section 2

Contexte



LBGB (Laboratoire de Bioinformatique pour la Génomique et la Biodiversité)

Arrivé de séquenceurs MGI

- 2 DNBSEQ-G400
- 1 DNBSEQ-T7



Workflow NGS

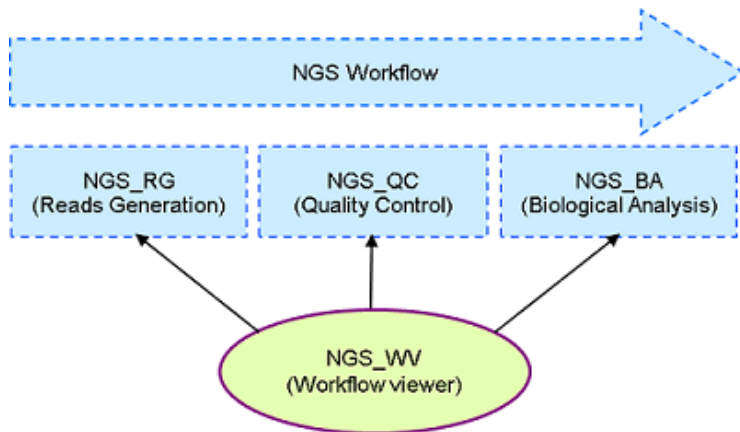


Figure 1: Workflow de génération, de controle qualité et d'analyse biologique des FASTQ

La technologie MGI

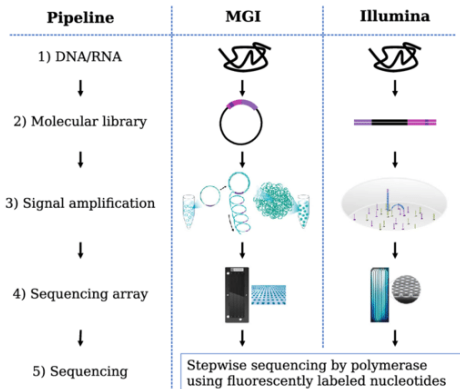


Figure 2: Différences entre Illumina et MGI de technologie NGL

La technologie MGI

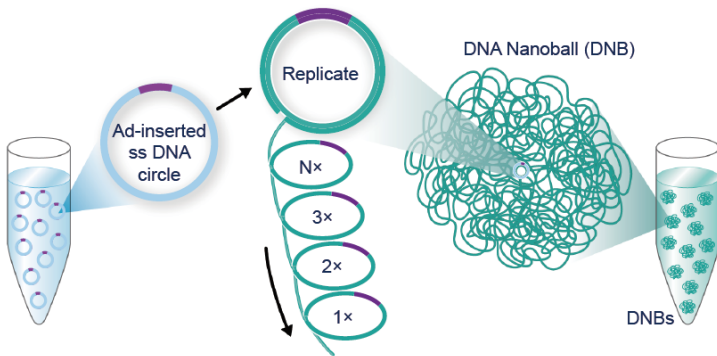


Figure 3: Schéma techno MGI

Section 3

Objectifs



Développement d'un pipeline automatique pour MGI

Détermination des outils bioinformatique

Quels méthodes et outils pour le pipeline



Test de 2 software de génération de FASTQ



Apprentissage du Perl

Pouquoi ?

- Raison historique du laboratoire
- Toutes les librairies et modules utilisés sont en Perl
- Workflow d'Illumina écrit en Perl

Réalisation

- Programme effectuant des analyses statistiques élémentaires
 - compter le taux de GC
 - moyenne de la qualité de chaque read
 - ect ...
- Lecture des modules utilisés dans le workflow d'Illumina



Section 4

Perspective

