

Formation en Bioinformatique

En vue de former une communauté Bioinformatique solide au Sénégal

Dans le cadre du partage d'expériences et de la mise à niveau en bioinformatique, nous souhaitons vous présenter quelques modules que nous envisageons de développer. Nous restons ouvertes à toutes recommandations et suggestions, dans un cadre professionnel favorisant l'épanouissement intellectuel de chacun.

Objectifs pédagogiques

1. Maîtriser l'environnement de travail bioinformatique (Conda, Git)
2. Manipuler les formats de fichiers spécifiques et pipelines d'analyse
3. Acquérir des compétences en analyse de données de séquençage NGS
4. Appliquer les bonnes pratiques de reproductibilité et gestion de code

Module 1 : Environnement technique (15h)

1.1 Gestion des environnements avec Conda

- Création d'environnements isolés
- Gestion des dépendances logicielles
- Partage de configurations via `environment.yml`

1.2 Versioning avec Git

- Workflow basique (`add`, `commit`, `push`)
- Gestion des branches et collaboration
- Intégration avec GitHub/GitLab

1.3 Scripting automatisé

- Bash : piping, job control, scripts batch
- Python : parsing de fichiers biologiques (Biopython)
- Parallélisation basique (GNU Parallel)

Module 2 : Formats de données (10h)

2.1 Formats standards

Format	Structure	Cas d'usage
FASTA/Q	En-tête/Sequences	Séquençage brut
GenBank	Annotations	Banques de données
GFF/GTF	Features génomiques	Annotation
BED	Régions génomiques	Analyse ChIP-seq
SAM/BAM	Alignements	Variants

2.2 Conversion/Validation

- Outils : samtools, bedtools
- Vérification d'intégrité des fichiers

Module 3 : Visualisation (8h)

3.1 Genome Browsers

- UCSC Genome Browser : navigation avancée
- IGV : visualisation d'alignements et variants

3.2 Outils spécialisés

- PyMOL/RasMol (structures 3D)
- Circos (représentations circulaires)

Module 4 : Technologies de séquençage (12h)

4.1 Plateformes comparées

Technologie	Longueur reads	Débit	Applications
Illumina	Court (150bp)	Haut	RNA-seq
PacBio	Long (10kb+)	Moyen	Assemblage
Oxford NanoPore	Long (100kb+)	Variable	Métagénomique

4.2 Contrôle qualité

- FastQC/MultiQC
- Filtrage avec Trimmomatic/Cutadapt

Module 5 : Analyse RNA-seq (20h)

5.1 Alignement

- BWA/Bowtie2 : paramètres optimaux
- STAR : alignement splice-aware

5.3 Bonnes pratiques

- Documentation avec Jupyter/RMarkdown
- Automatisation via Snakemake/Nextflow

Évaluations

1. Projet pratique en binôme: Exemple

- Analyse complète de données RNA-seq
- Versionnage sur dépôt Git
- Rédaction de rapport reproductible

2. Quiz techniques :

- Dépannage de pipeline bash
- Interprétation de sorties d'alignement