



**Ministère de l'Enseignement Supérieur et de la Recherche Scientifique  
République Algérienne Démocratique et Populaire**

*Université des Sciences et de la Technologie d'Oran Mohamed Boudiaf*

*Faculté des Sciences de la Nature et de la Vie*

*Département de Génétique Appliquée Moléculaire*



## « TRAVAIL À FAIRE »

### **TP3 : Git, GitHub et Leur Utilité pour les Biologistes**

#### ❖ Réalisé Par :

- *NOM : BOUTIBA*
- *Prénom : ALIA IMENE*
- *Section 01*
- *Groupe 02*

# Introduction

Dans le domaine des sciences biologiques et expérimentales, les chercheurs manipulent aujourd’hui une grande quantité de données : résultats d’expériences, observations sur le terrain, codes informatiques, scripts d’analyse, tableaux de données, rapports, etc.

La gestion efficace de ces informations ainsi que le travail collaboratif sont devenus des compétences essentielles. Parmi les outils les plus utilisés à l’échelle internationale pour organiser les projets scientifiques, on trouve **Git** et **GitHub**, qui permettent de suivre les modifications, conserver toutes les versions d’un projet et faciliter la collaboration entre chercheurs. L’objectif de ce travail est d’expliquer ce qu’est Git, ce qu’est GitHub, et de montrer leur importance dans le travail scientifique d’un biologiste.

## **1. Qu'est-ce que Git ?**

### A. Définition

Git est un **système de gestion de versions (Version Control System)**.

Il permet d’enregistrer, de suivre et de gérer toutes les modifications apportées à un fichier ou à un projet au fil du temps. Il a été créé en 2005 par **Linus Torvalds**, le créateur de Linux, afin d’assurer un suivi fiable et rapide du développement du noyau Linux.

En termes simples, Git permet : d’enregistrer chaque modification, de revenir à une version précédente à tout moment, de travailler à plusieurs sans conflit, de conserver l’historique complet d’un projet.

### B. Intérêt de Git pour la recherche scientifique

Dans les projets de biologie, il est souvent nécessaire de : suivre l’évolution des données expérimentales, documenter les protocoles, enregistrer les résultats, analyser des fichiers en utilisant des scripts (R, Python...), et travailler en équipe.

Git facilite toutes ces tâches grâce à :

- **La traçabilité des données** : Chaque modification est enregistrée avec la date, l'auteur et une description.
- **La sécurité des informations** : Les anciennes versions sont conservées et récupérables.
- **La reproductibilité scientifique** : Les autres chercheurs peuvent facilement comprendre comment les résultats ont été produits.

- **La collaboration efficace** : Plusieurs personnes peuvent travailler en parallèle sans écraser le travail des autres.

## C. Versioning, historique et commits

Un **commit** représente l'enregistrement d'un ensemble de modifications. Chaque commit contient :

- un message descriptif clair,
- la liste des fichiers modifiés,
- la date et l'auteur.

L'ensemble des commits forme **l'historique du projet**, élément essentiel en recherche pour garder une trace du travail effectué.

## 2. *Qu'est-ce que GitHub ?*

### A. Définition

GitHub est une **plateforme d'hébergement en ligne** qui permet de stocker à distance les projets gérés avec Git. On peut le considérer comme un “cloud” spécialisé pour les chercheurs, étudiants et développeurs. Il facilite : le partage de projets, la sauvegarde en ligne, la collaboration entre plusieurs utilisateurs, l'accès au projet depuis n'importe quel endroit.

### B. Repository

Un dépôt GitHub ou *repository* est un espace où sont stockés : les fichiers du projet, les sous-dossiers, l'historique des commits, les différentes branches, les versions publiées.

Un dépôt peut être **public** (visible par tout le monde) ou **privé** (visible uniquement par le propriétaire et les collaborateurs autorisés).

### C. Branches

Une **branche (branch)** est une version parallèle du projet. Elle permet de tester, modifier ou ajouter de nouvelles fonctionnalités sans toucher au projet principal. Exemple :

- branche principale : **main**
- branche secondaire : *analyse-données*

Les branches servent à : éviter les erreurs dans la version principale, expérimenter librement, travailler en équipe de façon organisée.

## D. Pull Requests

Une **pull request** est une demande de fusion d'une branche secondaire dans la branche principale.

C'est un outil central pour la collaboration, car il permet : de discuter les modifications, de vérifier le travail, de décider ou non d'intégrer les changements.

## E. GitHub et les projets scientifiques

GitHub est utilisé dans : les projets de bio-informatique, l'analyse génomique, les études écologiques, les projets SIG, les analyses de données en R ou Python.

Il apporte organisation, transparence et partage des connaissances.

### *Importance de GitHub pour un biologiste*

#### ***Importance de GitHub pour un biologiste***

##### ***1. Stockage sécurisé des données scientifiques***

GitHub permet d'héberger : des fichiers CSV ou Excel, des données expérimentales, des images de terrain, des scripts d'analyse, et des protocoles de laboratoire.

Les données ne risquent pas de se perdre ou d'être supprimées accidentellement.

##### ***2. Documentation et traçabilité***

Chaque modification, ajout ou suppression est enregistrée. Ceci est essentiel pour : la rigueur scientifique, la compréhension des analyses, la relecture du travail, la préparation d'articles scientifiques.

##### ***3. Travail collaboratif***

GitHub facilite la collaboration entre : étudiants, doctorants, chercheurs. Chacun peut créer une branche, proposer des modifications, commenter le travail des autres, et fusionner les résultats.

## **4. Gestion des projets à long terme**

Les projets de biologie peuvent durer plusieurs mois ou années. GitHub permet de : organiser les fichiers, structurer les analyses, conserver tous les résultats intermédiaires, suivre l'évolution du projet de manière claire.

## **5. Développement de compétences en bio-informatique**

Avec l'augmentation du rôle des données massives (big data) en biologie, un biologiste doit savoir manipuler : Python, outils génomiques, pipelines d'analyse.

### ***GitHub est l'un des outils essentiels dans ces domaines.***

Exemples d'utilisation de GitHub en biologie

#### **1. Analyse génomique (RNA-seq, DNA-seq)**

Les équipes de recherche stockent leurs scripts et pipelines sur GitHub.

#### **- Suivi des expériences de laboratoire**

Les protocoles, fichiers de résultats et analyses sont enregistrés quotidiennement par commits.

#### **- Projets d'écologie**

Les données d'observation (GPS, CSV, photos) sont partagées sur un dépôt commun.

#### **- SIG (Systèmes d'Information Géographique)**

Les fichiers cartographiques sont gérés comme un projet GitHub.

#### **- Rédaction de protocoles scientifiques**

Chaque mise à jour du protocole est sauvegardée et commentée.

## **Conclusion**

Git et GitHub représentent deux outils fondamentaux dans le travail scientifique moderne.

Leur utilisation garantit une meilleure organisation des projets, une collaboration efficace et une traçabilité complète des analyses et des données. Pour un biologiste, maîtriser GitHub devient aujourd'hui une compétence essentielle, car elle permet de gérer correctement les données, d'améliorer la qualité du travail scientifique et de faciliter les échanges au sein d'une équipe de recherche.