

# Git and Sequence Alignment

author: Julin Maloof date: April 02, 2015

## Today’s Lab

We will have two topics for today’s lab

- Version control and collaboration with Git and GitHub
- Sequence Alignment using EMBOSS tools.

## What is Git?

- Tracks changes in projects and documents
- Can restore to an earlier state
- Can collaborate with others
- Can maintain different versions

## What is Git?

- An essential tool for software development
- Should be used whether or not your are collaborating
- Enables transparent development and sharing
- Enables Documenation, Reproducibility, Openness

## Git Intro Videos

- [Video1](#)
- [Video2](#)

## Git: Key concepts

incremental: true \* Repository \* A project that is tracked by Git \* To start a new repository \* `git init` \* To add files to your repository \* `git add` \* To save changes to your repository \* `git commit`

- 
- If you are going to collaborate you will need a
  - **remote** repository
  - If you want to use a remote repository you need to
  - `fork` or `clone` it
  - To download changes from the remote repository
  - `git pull`
  - To upload changes to the remote repsoitory
  - `git push`

## Seq Alignment

A key tool in bioinformatics is sequence alignment.

- Questions:
1. How similar are two sequences?
  2. Given a protein of interest what other proteins are similar in another genome?

Part 2 of today’s lab will have you explore tools for doing alignments and homolgy searches.

You can review principles of Sequence Analysis by looking at the Sequence Analysis Primer (linked to from the lab page).

## Seq Alignment Significance

incremental: true

How do you determine significance?

What does the null distribution look like?

Permutation: Create a randomized data set.

How does the real data compare to randomized data?