

# Introduction à Linux RAQ - Cours 01

Eric Normandeau - 2015-01-29

## Plan de cours

1. Introduction
2. Connection au serveur Linux
3. Le terminal

## 1 - Introduction

### Qu'est-ce que Linux ?

Linux est un système d'exploitation, au même titre que Windows et MacOSX. Bien qu'il soit gratuit, Linux est beaucoup moins fréquemment retrouvé sur les ordinateurs standards (tours et ordinateurs portables). Cependant, dans le monde des ordinateurs très puissants (serveurs de calcul et super-calculateurs), près de 95% des systèmes installés sont dérivés de UNIX, dont la grande majorité sont des versions de Linux. Il existe plusieurs dizaines de versions différentes de Linux, mais il y en a seulement 4-5 qui sont plus fréquentes et elles sont très similaires. Afin d'alléger la discussion, je parlerai de "systèmes UNIX" ou même de "Linux" pour désigner à la fois Linux et MacOSX. Il faut savoir cependant que bien que ces deux systèmes d'exploitation soient similaires, ils sont quand même légèrement différents et certains programmes ou commandes qui fonctionnent avec l'un pourraient ne pas fonctionner avec l'autre.

### Pourquoi utiliser Linux en biologie ?

Voici quelques besoins de la recherche scientifique qui s'appliquent également à la biologie :

Explorer rapidement des données pour explorer des hypothèses  
Reproduire des analyses de façon répétée sur différents jeux de données  
Analyser des données demandant une grande puissance de calcul ou beaucoup de mémoire

Les systèmes dérivés de UNIX, dont Linux, fournissent un environnement et des outils rendent ces tâches plus faciles. En utilisant Linux en combinaison avec un ou plusieurs langages de programmation, on peut plus rapidement explorer ses données, produire des scripts qui permettent de reproduire rapidement des analyses et lancer des programmes gourmands en ressources de calcul.

Dans certains domaines de la bioinformatique, on retrouve presque exclusivement des programmes pour Linux et MacOSX, tous les deux dérivés des systèmes UNIX. Par exemple, la plupart des programmes pour faire des alignements de séquences et découvrir des SNPs sont conçus pour être installés sur Linux. Afin d'avoir accès à ces programmes, il faut donc savoir utiliser des ordinateurs avec des systèmes UNIX.

## **Exemples d'analyses**

- Alignement de séquences
- Assemblage de génomes de novo
- Génotypage par Séquençage
- (... autre chose que génomique)

## 2 - Connection au serveur Linux

Afin de se connecter au serveur, nous devons premièrement ouvrir un terminal. Pour les utilisateurs de MacOSX et Linux, lancez la commande suivante :

```
ssh username@raq.ibis.ulaval.ca  
for i in ${seq 10}; do echo $i; done
```

Vous devez remplacer ‘username’ par le nom d’utilisateur que vous avez reçu. Lorsque demandé, entrez votre mot de passe. Il est normal que vous ne voyez pas d’étoiles ou de cercles apparaître lorsque vous tapez votre mot de passe.

Pour les utilisateurs Windows, double-cliquez sur le logiciel ‘putty.exe’, disponible ici (). Dans

ou lancer putty.exe (Windows).

## 3 - Le terminal

Lorsqu'on fait des analyses sous Linux, il est important d'être familier avec le terminal. Le terminal est une fenêtre où on tape des commandes et où les résultats peuvent être affichés. Les résultats peuvent aussi être envoyés vers des fichiers. Nous allons utiliser le terminal tout au long du cours. Il faut bien comprendre qu'on ne peut pas utiliser Linux pour des analyses si on ne comprend pas bien comment le terminal et le système fonctionnent. Notre but ultime est de faire de la biologie, mais afin de pouvoir faire nos analyses, nous devons bien comprendre les bases de Linux et du terminal.

### Commandes de base

- ls
- cd
- mkdir
- touch