LCG - BEII - Introduction and contents

DUBii 2019

Jacques van Helden

2019-02-05

Scope of the course

This courses provides an overview of bioinformatics methods used to analyse cis-regulation.

1. Detection of cis-regulatory elements and other types of nucleic sequence motifs.
2. Next Generation Sequencing approaches to genomi regulation
   - ChIP-seq for transcription factors
   - ChIP-seq for chromatine marks (histone acetylation, methylation)
   - RNA-seq (differentially expressed genes)
3. Underlying statistics

Schedule

Day	Topics
Day 1	Regulatory sequences: cis-regulatory elements intro + predicting binding sites
Day 2	Regulatory sequences: motif discovery
Day 3	Analysing regulation with NGS data
Day 4	Integrative exercises and evaluations

General goals of the course

- Learn the **concepts and methods** related to bioinformatics analysis of cis-regulation
- Understand the underlying **algorithms** and **statistics**.
- Get a **practical experience** of bioinformatics resources (**databases** and **tools**) used in the domain

Specific goal: skills in statistics

1. Topics
   - Sequence models
   - Statistics of motif detection (scanning, discovery)
   - NGS statistics (read mapping, peak calling, differential expression)
   - Designing control sets to validate the statistical models
2. Approach
   - Minimalist theoretical elements
   - Intuitive approach to the statisics: lots of graphs, simulations, controls
3. Environment
   - R statistical language (https://www.r-project.org/)
   - Bioconductor: (https://www.bioconductor.org/)
   - Developing environment: RStudio (https://www.rstudio.com/)

Specific goal: technical skills

- Writing reports in R markdown (http://rmarkdown.rstudio.com/)
- Version management with git (https://git-scm.com/)
  - see Blischak, Davenport, and Wilson (2016)

Resources for this course

Name	Description + URL
LCG_BEII	The course material for this course on github http://jvanheld.github.io/LCG_BEII
RSAT	Regulatory Sequence Analysis Tools http://rsat.eu/
RSAT course	Teaching material for the analysis of regulatory sequences http://rsa-tools.github.io/course/

Downloading a local copy of the course material

This is optional.

If you wish to download a local copy of all the course material, open a Unix terminal and type.

```
git clone https://github.com/jvanheld/LCG_BEII.git
```

Then open the file *LCG_BEII/index.html* with a Web browser.

Since the course will be periodically improved, you can get the updates with the following command.

```
cd LCG_BEII;
git pull
```

Command-line use of RSAT for LCG students

LCG students can access RSAT on the command line with the following commands

```
## Open an ssh connection to the local server
ssh tepeu

## Load the RSAT configuration file
source /export/apps/rsat/RSAT_config.bashrc

## Check that the paths are correct
supported-organisms
```

References

Blischak, John D, Emily R Davenport, and Greg Wilson. 2016. "A Quick Introduction to Version Control with Git and GitHub." *PLoS Computational Biology* 12 (1): e1004668.