Series 1

Genomics and bioinformatics - Week 1

September 20, 2011

1 Website of this course

Add the course website to your bookmarks.
`http://moodle.epfl.ch/course/view.php?id=11181`

2 Demo of the UCSC genome browser

Go to the address `http://genome.ucsc.edu/`. Click on "Genome browser", select a species and visualize its genomic content; zoom in and out.

3 Programming exercise

To complete this exercise, install R and Python on your computer (see below).

This exercise, which consists in playing with a sample text file, should be completed in both languages. You must write your scripts successively in R and Python. Use as much of the documentation given below as needed to obtain the result we desire here.

Testing file: download `genes_expression_100.txt` on the course Moodle website. It is an output file of a typical bioinformatics program in raw text tab-delimited format - 100 lines of the form:

`Gene name <tab> Expression in condition 1 <tab> Expression in condition 2`

Using Python, then R, do:

1. Read the file and extract gene names and associated numbers;
2. Compute the ratios between the two numerical columns; take log_2 of the result;
3. Compute the geometric means between the two numerical columns; take log_{10} of the result;
4. Write these results in a new text file: `"GeneName <tab>` log_2`(ratio) <tab>` log_{10}`(mean)"`
5. Plot ratios vs means (use `matplotlib` for Python).

4 R

4.1 Installation

Download R 2.13 at `http://stat.ethz.ch/CRAN/`
During the installation, when you are asked what to install, don't forget to include the help files.

If you are using Windows: during the installation, choose where you want to install your folder; then you can run it clicking on the .exe file located in the bin/ sub-folder.

4.2 Tutorial

R tutorial: `http://cran.r-project.org/doc/manuals/R-intro.pdf`
If you need help:

1. Read the tutorial
2. Use the ? or help() R commands
3. Use Google
4. Ask us.

4.3 Reference documentation

`http://cran.r-project.org/doc/manuals/refman.pdf`

5 Python

5.1 Installing

Download the Enthought Python from our USB keys. The packages are also available on `http://www.enthought.com/repo/.hidden_epd_installers` (but the server is very slow).

Open a console and type "ipython" (on Windows, to get a console open the "Start" menu and type "cmd" in the search field).

5.2 Tutorial

Python tutorial: `http://docs.python.org/tutorial/` read 3.1 to 4.6, 5.5 and 7.2.

5.3 Reference documentation

`http://docs.python.org/library/index.html`