

Curriculum vitae

Romain Groux

Bioinformatics software developer /
bioinformatician

32 years old, Swiss

Rue de Lausanne 49G, 1020 Renens

Phone +41(0)76 304 00 20

Email : romain.groux@hotmail.com

Linkedin : <https://www.linkedin.com/in/romain-groux-b15aa4b4/>

GitHub : <https://github.com/romaingroux/>

Strenghts

PhD in bioinformatics

Bioinformatics software developer

NGS data specialist

Naturally curious and always eager to learn

Professional experience:

04.2015 - 09.2019 EPFL, scientific assistant in the laboratory of "Computational Cancer Genomics".

- Development of a software predicting transcription factor binding sites (C). This software predicts which regions of a genome are bound given a specificity model. This software is currently the most efficient available for this type of problem.
- Development of unsupervised classification methods (C++) for genomic region of interest, based on i) the sequence of the regions and ii) diverse NGS sequencing profiles (ChIP-seq, DNase-seq, ATAC-seq, MNase-seq). These methods allow to identify i) functionally different families of regions and ii) their signatures.
- Development, maintenance and optimization of diverse software components and analysis pipelines hosted on the laboratory web server backend.
- Curator for a public database containing NGS data (ChIP-seq, DNase-seq, ATAC-seq, MNase-seq, etc).

Technical skills:

Developing and maintaining of software and analysis pipelines in C++11 (STL, Boost, SFML, SeqAn) Python 2.7/3.6 (STL), R 2.X/3.X (diverse libraries) and bash.

Testing code using dedicated libraries in C++ (Unittest++) and Python (unittest)

Measuring and optimizing the performances of code and analysis pipelines.

Handling and processing of NGS data using dedicated softwares/libraries (samtools, bedtools, pysam), mapping (bowtie, bowtie2) and data quality control, etc.

Genomic data analysis using machine learning methods (unsupervised classification algorithm) and function prediction (transcription factor binding prediction).

Algorithm understanding and ability to implement and optimize them.

Academic projects:

PhD project : setting up and pursuing a research project studying i) transcription factor binding specificity in the human genome (development of methods allowing to model their sequence preferences, training of sequence preference models and performance assessments of the models) and ii) the chromatin structure in the vicinity of transcription factor binding sites in the human genome by the developing dedicated analysis methods. These methods were then applied on NGS datasets (ChIP-seq, DNase-seq, ATAC-seq, etc) in order to characterize some regions of the human genome.

Master project : studying the regulation of gene expression as part of a swiss consortium (CyclIX). I worked on i) the development of a pipeline that performs data processing/formatting and quality control (bash, Python et R), ii) the development of different software modules required for the data analysis and iii) the application of the above for the data analysis (Python et R).

Education:

2016 - 2020 (in progress)

PhD in bioinformatics, EPFL
End planned for February 2020.

2013 - 2014

Master in bioinformatics, UNIL

2010 - 2013

Bachelor in biology, UNIL

2006 - 2009

CFC of biology laboratory technician, EPFL

Languages:

French	mother tongue
English	fluent written and spoken (equivalent to C1)
German	basic scholar knowledge

References:

Dr. Philipp Bucher	Head of the “Computational Cancer Genomics” laboratory of EPFL +41 21 693 09 56 philipp.bucher@epfl.ch
Dr. René Dreos	Senior biostatistician at the Integrative Center for Genomics of UNIL +41 21 692 40 44 rene.dreos@unil.ch
Dr. Giovanna Ambrosini	Senior scientist in the “Computational Cancer Genomics” laboratory of EPFL +41 21 693 09 57 giovanna.ambrosini.epfl.ch
Dr. Vincent Gardeux	Senior scientist in the “Systems Biology and Genetics” laboratory of EPFL vincent.gardeux@epfl.ch