CLAUDIO LORENZI

Bioinformatician, PhD

Contacts:
Basel, CH 4057
claudio.lorenzi@gmail.com
+39 3343048023
www.linkedin.com/in/clorenzi/

OBJECTIVE

Committed and enthusiastic bioinformatician with over six years of experience in developing bioinformatics tools and analysing biological data. Confident in working with different programming languages and frameworks. Fast learner and always curious about new technologies, especially in the machine learning field. Collaborative and creative, promoter of a healthy and productive working environment.

SKILLS

Data Analysis

RNA-seq, ChIP-seq, Genomics, statistical tests, pathway analysis, basics of protein structure analysis.

Machine Learning

Studied at University and used* during PhD:
SVM, Linear and Logistic* regression, Naive Bayes*, Decision tree based models*, k-NN, Deep Learning, Perceptron*, CNN*, Clustering techniques (hierarchical*, k-means*, EM*, DBSCAN*, SOM*), feature reduction (PCA*, NMF, Autoencoders*, t-SNE*), GAN.
Main libraries used: TensorFlow, sk-learn and MLPack.
Additional topics studied on books and online courses:
- Reinforcement Learning and PyTorch (HF certification)
- Transformers and Graph NN (Understanding Deep Learning, S. J.D. Prince, preprint)

Web/GUI Development

Angular, Node.js (Express.js and Sequelize), mySQL, RESTful API, Electron, Bootstrap, Material.

Programming Languages

Python, C++, JavaScript, R, Java (basics)

Other tools

Git, Singularity, Docker, Nextflow, VSCode, Eclipse, several bioinformatics tools and libraries.

LANGUAGES

Italian	Native
English	Proficient
French	Conversational

EXPERIENCE

Bioinformatician - Postdoc

Feb 2022 - Present
Bellinzona, CH

IOR - Institute of Oncology Research
Team: Functional Cancer Genomics

Development of a dedicated nexflow pipeline for the preprocessing of the raw data for large cohorts of patients derived prostate cancer RNA-seq samples (PC atlas).
Design, development and implementation of the web application PCaProfiler, currently in beta at pcaprofilertest.tk, that allows the exploration and analysis of the PC atlas.
Analysis of the PC atlas for the identification of cancer specific isoforms.
RNA-seq and ChIP-seq data analysis for other projects, including splicing analysis and identification of contaminants.

Bioinformatician - PhD candidate

Oct 2018 - Oct 2021
Montpellier, FR

IGH - Institute of Human Genetics
Team: Machine Learning and Gene Regulation

Design, development and implementation of two bioinformatic software: iMOKA and IRFinder-S.
Ideated a fast and compacted data structure for large sparse k-mer matrices, used for the identification of relevant features using Machine Learning models.
Implementation of a GUI for iMOKA and Web Application for IRFinder (IRBase) developed with Electron and Angular/Material the first, Express and Angular/Bootstrap the second.
Machine Learning models generated using MLPack (C++), Tensorflow and scikit-learn (Ptython).
Software benchmarking and other biology-related statistical analysis (R).Software releases on GitHub, DockerHub and Singularity images.

Study and Development Engineer

Oct 2016 - Oct 2018
Montpellier, FR

IGH - Institute of Human Genetics
Team: IMGT - the international ImMunoGeneTics information system

Development of bioinformatics pipelines for the manual curation of immunological sequences on an internal Java based web application.
Improvement of the cross-communication of the internal and external databases and integration of information from different origins.

EDUCATION

PhD in Bioinformatics

Oct 2018 - Oct 2021
University of Montpellier
Montpellier, FR

M.Sc in Bioinformatics

Oct 2014 - Jul 2016
University of Bologna - 110/110 cum laude
Bologna, IT

B.Sc in Medical Biotechnology

Sep 2011 - Sep 2014
University of Milan - 110/110 cum laude
Milan, IT

PUBLICATIONS

- iMOKA: k-mer based software to analyze large collections of sequencing data
  Lorenzi, C., Barriere, S., Villemin, JP. et al. Genome Biol 21, 261 (2020)
  DOI: 10.1186/s13059-020-02165-2
- IRFinder-S: a comprehensive suite to discover and explore intron retention
  Lorenzi, C., Barriere, S., Arnold, K. et al. Genome Biol 22, 307 (2021)
  DOI: 10.1186/s13059-021-02515-8
- A cell-to-patient machine learning transfer approach uncovers novel basal-like breast cancer prognostic markers amongst alternative splice variants
  Villemin, JP., Lorenzi, C., Cabrillac, MS. et al. BMC Biol 19, 70 (2021)
  DOI: 10.1186/s12915-021-01002-7
- GECKO is a genetic algorithm to classify and explore high throughput sequencing data
  Thomas, A., Barriere, S., Broseus, L. et al. Commun Biol 2, 222 (2019)
  DOI: 10.1038/s42003-019-0456-9
- NF90 modulates processing of a subset of human pri-miRNAs
  Grasso, G., Higuchi, T., Mac, V., Barbier, J. et al. Nucleic Acids Research 48, 12 (2020)
  DOI: 10.1093/nar/gkaa386
- Translesion DNA synthesis-driven mutagenesis in very early embryogenesis of fast cleaving embryos
  Lo Furno, E., Busseau, I., Aze, A., Lorenzi, C., et al. Nucleic Acids Research 50, 2 (2022)
  DOI: 10.1093/nar/gkab1223

PUBLIC PROJECTS

A list of public projects is available here: https://cloxd.github.io/CloXD

REFERENCES

- Dr. Theurillat Jean Philippe, Functional Cancer Genomics group Leader, IOR - Bellinzona.
  e-mail: jean-philippe.theurillat@ior.usi.ch
  phone: +41 58 666 7213
- Dr. Ritchie William, Machine Learning and Gene Regulation group Leader, IGH - Montpellier
  e-mail: william.ritchie@igh.cnrs.fr or billthebrute@gmail.com
  phone: +33 7 53 52 92 79