

**Project: Reconstruction of transcriptional circuitries based on epigenetic data of cancer samples**

**Starting date:** January 2019 or later

This project will focus on the development of a Python method to reconstruct transcriptional regulatory circuitries using epigenetic and transcriptional data (ChIP-seq/ATAC-seq + RNA-seq) in cancer samples. The first prototype of the method has been already created in the lab and it combines DNA motif analysis techniques, analysis of high-throughput sequencing data, and graph theory to get candidate transcription factors. The mission of the intern will be to develop the scoring functions and teach the method to predict the experimentally validated sets of master transcriptional regulations in different healthy and malignant cell types.

You will work on this project in the laboratory of Computational Epigenetics of Cancer (**Institut Cochin, Paris**) and in a collaboration with the laboratory of Cancer Systems Biology (PI: Olivier Elemento, **Weill Cornell Medical College, New York**). Our team is the national leader in the analysis of cancer epigenetic data. We work in close collaboration with clinicians and biologists from the Curie Institute, the Cochin Hospital, and Gustave Roussy Institute. In Cochin, we have a dynamic international environment (French, Russian, Algerian and Polish researchers in our lab), a big work-space with a view on the top of the Eiffel tower and excellent expertise in the bioinformatics analysis of high-throughput sequencing data from cancer samples, motif discovery methods, epigenetics and cancer biology.

**Selected publications by the team:**

- Heterogeneity of neuroblastoma cell identity defined by transcriptional circuitries. V. Boeva, et al. *Nature Genetics*, 2017. 49(9):1408-1413. **Key paper.**
- QuantumClone: Clonal assessment of functional mutations in cancer based on a genotype-aware method for clonal reconstruction. P. Deveau, et al. *Bioinformatics*. 2018. 34(11):1808-1816.
- HMCan-diff: a method to detect changes in histone modifications in cells with different genetic characteristics. H. Ashoor, et al. *Nucleic Acids Research*. 2017. 45(8):e58.
- Comparative analyses of super-enhancers reveal conserved elements in vertebrate genomes. Y.A. Pérez-Rico, V. Boeva, et al. *Genome Research*, 2017. 27(2):259-268.
- Analysis of genomic sequence motifs for deciphering transcription factor binding and transcriptional regulation in eukaryotic cells. V. Boeva. *Frontiers in Genetics*. 2016. 7:24.

**Candidate requirements:**

- Experience in data analysis with R
- Good knowledge of Python
- Knowledge of Linux environment (bash, awk, grep,...)
- Good level of spoken English

**Team webpage:** [www.boevalab.com](http://www.boevalab.com)

**Team location:** 5<sup>th</sup> floor of the Faculty building, Institut Cochin, 24 rue du Faubourg Saint-Jacques, 75014 Paris

**Contact:** [valentina.boeva@inserm.fr](mailto:valentina.boeva@inserm.fr) – please send a short motivation letter, your CV and names of two references.