**Master 2 internship project**

**Year 2023-2024**

**Laboratory/Institute:** LPCV **Director:** Eric MARECHAL

**Team:** Floral Regulator (Flo\_Re) **Head of the team:** François PARCY

**Name and status of the scientist in charge of the project:**  **HDR: ~~yes~~ ☐ no ☐**

**Address:** 17 Avenue des Martyrs 38000 Grenoble

**Phone:** 04.38.78.49.78 **e-mail:** romain.blancmathieu@cea.fr

**Program of the Master’s degree in Biology:**

**☐** ~~Microbiology, Infectious Diseases and Immunology~~ **~~☐~~** ~~Structural Biology of Pathogens~~

**☐** Physiology, Epigenetics, Differentiation, Cancer **~~☐~~** ~~Neurosciences and Neurobiology~~

**Title of the project: Deciphering the combinatorial nature of gene expression regulation**

Objectives (up to 3 lines):

Discovering new regulatory protein complexes formed by transcription factors binding to DNA using machine learning

Abstract (up to 10 lines):

Gene expression regulation involves intricate molecular processes, including the binding of transcription factors (TFs) to specific DNA sequences in gene promoters. TFs often associate with other proteins, forming regulatory complexes that modulate the fate of target genes. While the combinatorial aspect of gene regulation is widely recognized, the full scope and diversity of regulatory TF complexes remain unknown. This project aims to develop a machine learning model to comprehensively identify new ternary complexes comprising TFs and DNA in the model plant species Arabidopsis thaliana. By integrating diverse genomic information such as TF binding profiles, co-evolution patterns of interacting amino acids, and temporal-spatial co-expression, we will uncover physical interactions among TFs involved in DNA binding.

Methods (up to 3 lines):

Genomic information will be derived from available TF genome-wide DNA-binding assays, RNA-seq experiments and AlphaFold2 predictions for selected representative TF pairs. The machine learning model will be train using known TFs interaction from the BioGRID database.

Up to 3 relevant publications of the team:

Lai X, et al. 2019. Building Transcription Factor Binding Site Models to Understand Gene Regulation in Plants. Molecular Plant 12:743–763.

Rieu P, et al. 2023. The F-box protein UFO controls flower development by redirecting the master transcription factor LEAFY to new cis-elements. 2. Nat Plants 9:315–329.

Blanc-Mathieu R, et al. 2023. Plant-TFClass: a structural classification for plant transcription factors. Trends in Plant Science, *accepted*

Requested domains of expertise (up to 5 keywords): Bioinformatics, gene expression regulation, strong interest in machine learning and plant science.