Modeling Transcription: an integrated approach to investigate the role of nascent RNAs in the maintenance of healthy gene expression programs

Matthieu Lavigne *[#1], Vaios Theodosiou[1] and Marianna Stagaki[1], Electra Tsaglioti[1], Christos Botos[1]

[1] IMBB

Presenting author * Corresponding author: email: lavigne@imbb.forth.gr.

ABSTRACT

The Gene Control Mechanisms lab is a hybrid group at IMBB-FORTH dedicated to explaining what molecular mechanisms affecting RNA production are disturbed with the onset and progression of various cancers and neurological disorders. We want to understand the contribution of nascent RNA (nRNA) production on global gene expression regulation in human cells and give mechanistic evidence why nRNAs are still overlooked as potential therapeutic targets since they likely interact with other vital chromatin components driving co-transcriptional splicing or genome folding. A hallmark of cancer has been defined as elevated transcriptional outputs and transcription dependency, but we lack thorough molecular-level understanding for the mechanism and the contribution of this process in the development of cancerous phenotype. We hope to identify pathological disturbances related to nRNA pathway. For instance, we focus on showing that growing nRNA molecules feedback transiently and locally on other chromatin components such as H2Bub writer enzyme RNF20/40 that plays a role in RNA polymerase II (RNA pol II) elongation trough nucleosomes (1). Preliminary results characterizing their distribution/structure and local interactome indicate how nRNAs regulate gene expression in cancer processes.

To account for the inter-dependency of the many regulatory steps of the transcription process, like RNA Pol II binding to promoters, Transcription Factors (TFs) association to DNA and the many effects of chromatin modifications on RNA polymerase processing, we also conduct in silico experiments to integrate individual regulatory layers of normal and cancer-specific transcription using mathematical models (2). Synergism between cancer biology, genomics, epigenetics, bioinformatics and mathematics approaches allows us to perform Simulation and Machine learning (ML) experiments to determine the key components of these processes. Our results will strengthen the ground-truth information provided by experimental data and the development of new class of ML model of hybrid type will both help acquiring values of the parameters involved, but also assist in the discovery of hidden mechanisms in the data (Predict gene regulatory networks controling RNA pol II outpout). Fine tuning of the model parameters will allow predictions of molecular and disease phenotypes, which we expect to aid the elucidation of the “principal regulatory code” governing physiological and cancer-specific gene expression programs.

REFERENCES

[1] Fanourgakis S, Synacheri AC, Lavigne MD, Konstantopoulos D, Fousteri M. Histone H2Bub dynamics in the 5' region of active genes are tightly linked to the UV-induced transcriptional response. **Comput Struct Biotechnol Journal**, 21, 614-629.

[2] Siepel A. A Unified Probabilistic Modeling Framework for Eukaryotic Transcription Based on Nascent RNA Sequencing Data. **bioRxiv** 2021.01.12.426408