Master's Internship (M2) - Modelling neuron specific signalling networks using symbolic AI based approach

Supervisors : M.Razzaq, O.Marchand
UMR PRC (INRAE–CNRS–University of Tours), BIOS team

Location: INRAE Val-de-Loire, Nouzilly, France
Duration: 6-month internship
Starting Date: January 2027
Keywords: Neural networks, Artificial Intelligence (AI), RNA sequencing (RNA-seq), Fragile X Syndrome, Phelan-McDermid Syndrome, Social Interaction, Synaptic Plasticity

Scientific Background and Objectives

Phelan-McDermid syndrome (PMDS) and Fragile X syndrome (FXS) are rare genetic disorders that affect brain development. They result in intellectual disabilities, language impairments, and often severe forms of autism. Currently, no effective pharmacological treatment is available for these disorders. Patient care mainly relies on behavioral therapies, which are time-consuming and often produce variable outcomes. Several drug development efforts have failed, likely because the targeted biological pathways are broadly expressed throughout the body, limiting both treatment specificity and efficacy.

To better understand the molecular mechanisms underlying these disorders, our laboratory has established a unique omics database called SOCIALOME, which provides insight into the activity of genes involved in PMDS and FXS. SOCIALOME contains more than 560 ribosome-associated mRNA sequencing (translatomic) profiles obtained from five key neuronal populations in PMDS and FXS mouse models exposed to different social contexts. Previous studies, including our own, suggest a global dysfunction of these neuronal populations, particularly affecting the transport and translation dynamics of messenger RNAs (mRNAs), in both mouse models and patients with FXS and PMDS ([1]-[5]).
Using this unique resource, we aim to precisely identify the biological alterations associated with these disorders. This approach will improve our understanding of how PMDS and FXS disrupt brain function and may ultimately lead to the discovery of novel, more targeted therapeutic strategies.

The objective of this internship is to reconstruct gene regulatory networks and understand how genes interact with one another over time within specific neuronal populations, as well as how these interactions are altered in PMDS and FXS. By identifying causal relationships using artificial intelligence approaches, this project aims to uncover the molecular mechanisms underlying these disorders and highlight novel therapeutic targets for the development of more effective and personalized treatments. No advanced background in biology is required, however, an interest in neuroscience is important for understanding the scientific context of the project.

Internship Project

The bioinformatics preprocessing pipeline, including read mapping, quantification, and differential gene expression analysis using the DESeq2 R package ([6]), has already been completed for all samples. At present, the SOCIALOME database is being analyzed using conventional statistical and functional enrichment approaches. While these methods are effective for identifying dysregulated genes and pathways, they do not provide insights into causal relationships or the regulatory mechanisms underlying the observed phenotypes.

To address this limitation, the project will leverage Answer Set Programming (ASP), a symbolic artificial intelligence framework that represents biological knowledge as logical rules and constraints. By combining these rules with RNA-seq data, ASP enables automated reasoning to identify regulatory mechanisms and infer potential causal relationships between molecular entities that are consistent with the experimental observations.

Using this approach, biological knowledge available through ReactomeFIViz ([7]) within the Cytoscape environment will be integrated with RNA-seq datasets to reconstruct context-specific gene networks. Existing molecular interactions and signaling pathways will be represented as interaction graphs and logical rules, providing a formal framework for automated reasoning on biological regulation and cellular signaling processes.

Graph-based AI methods will be applied to exploit both Reactome-derived interaction networks and transcriptomic data. Network analysis will be used to identify central regulators, key functional modules, and highly connected molecular actors associated with the phenotypes under investigation. Causal inference and logical reasoning approaches will then be employed to explore regulatory cascades and identify the most plausible mechanistic explanations for the observed molecular alterations.

These approaches aim to generate interpretable models of molecular regulation that highlight the biological mechanisms potentially responsible for the phenotypes observed in PMDS and FXS. **Ultimately, this work will result in the reconstruction of causal biological networks and explanatory models that can be visualized and explored in Cytoscape, the identification of key molecular regulators involved in disease-associated processes, and the discovery of novel, syndrome-specific therapeutic targets for future experimental validation.**

Internship Objectives

- Build a comprehensive gene interaction network by integrating biological knowledge from the scientific literature and publicly available pathway databases, particularly through the ReactomeFIViz plugin, together with the transcriptomic data generated within the SOCIALOME project.
- Use logical programming ("Answer Set Programming" ([8])) to model the temporal evolution of gene interactions and identify causal relationships between experimental conditions (mouse models, types of social interactions) and the observed molecular responses.
- Identify the neuronal populations most affected by these disorders and characterize their specific molecular networks.

Methods and Tools

During this internship, the student will receive training in a range of computational and programming approaches, including:

- Artificial intelligence and deep learning methods
- Logic programming
- Omics data analysis
- Use Cytoscape

Candidate Profile

We are seeking a Master's student (M2) in Bioinformatics or Computer Science with an interest in interdisciplinary research at the interface between biology and computer science, as well as data analysis. Candidates should possess the following skills:

- Proficiency in the Linux environment (command line, Bash, virtual environments)
- Proficiency in Python and R programming languages
- Strong interpersonal, writing, and organizational skills, with the ability to work independently
- Good technical English skills

Previous experience with artificial intelligence methods will be considered an asset. Familiarity with Git version control and FAIR data principles is also desirable.

Scientific Working Environment

The student will undertake the internship within the BIOS team of the joint research unit Physiology of Reproduction and Behavior (PRC), affiliated with INRAE UMR85, CNRS UMR7247, and the University of Tours. The interdisciplinary BIOS team is jointly led by Dr. Lucie Pellissier (Neurobiologist) and Dr. Romain Yvinec (Mathematician). The team combines omics approaches, mathematical modeling, and computational methods to investigate complex signaling and gene regulatory networks, from molecular mechanisms to in vivo models. The student will also benefit from the support of the PRC ISLANDE Bioinformatics Platform, which provides access to high-performance computing resources, including a 240-core computing server.

References

[1] Peça J, Feliciano C, Ting JT, Wang W, Wells MF, Venkatraman TN, et al. (2011) Shank3 mutant mice display autistic-like behaviours and striatal dysfunction. *Nature 472: 437–442.*

[2] Tyzio R, Nardou R, Ferrari DC, Tsintsadze T, Shahrokhi A, Eftekhari S, et al. (2014) Oxytocin-mediated GABA inhibition during delivery attenuates autism pathogenesis in rodent offspring. *Science 343: 675–679.*

[3] Jin C, Lee Y, Kang H, Jeong K, Park J, Zhang Y, et al. (2021) Increased ribosomal protein levels and protein synthesis in the striatal synaptosome of Shank3-overexpressing transgenic mice. *Mol Brain 14: 39.*

[4] Niere F, Wilkerson JR, Huber KM (2012) Evidence for a fragile X mental retardation protein-mediated translational switch in metabotropic glutamate receptor-triggered Arc translation and long-term depression. *J Neurosci 32: 5924–5936.*

[5] Parikshak NN, Swarup V, Belgard TG, Irimia M, Ramaswami G, Gandal MJ, et al. (2016) Genome-wide changes in lncRNA, splicing, and regional gene expression patterns in autism. *Nature 540: 423–427.*

[6] Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol 15: 550.*

[7] Wu G, Feng X, Stein L. A human functional protein interaction network and its application to cancer data analysis. *Genome Biol. 2010;11(5):R53.* doi: 10.1186/gb-2010-11-5-r53. Epub 2010 May 19. PMID: 20482850; PMCID: PMC2898064.

[8] Razzaq M, Paulevé L, Siegel A, Saez-Rodriguez J, Bourdon J, Guziolowski C (2018) Computational discovery of dynamic cell line specific Boolean networks from multiplex time-course data ((J. Stelling, editor)). *PLoS Comput Biol 14: e1006538.*