Computational modeling of Human Embryonic Development

Summary of the thesis project

Assisted reproductive technologies (ART), particularly in vitro fertilization (IVF), need novel approaches to improve the pregnancy rate. Current embryo culture systems and embryo quality assessment methods limit the success rate of IVF cycles to only 25%, leading to a social, emotional, and medical burden for the couple and the infertility medical team. In this context, the recent advent of novel technologies, such as transcriptomics, proteomics, and imaging, represents a formidable opportunity to consider in depth each embryo individually and to understand embryo developmental steps from a genetic and metabolic point of view.

One of the outstanding questions of the field is to understand the chain of events regulating human preimplantation development leading to an implantation-competent embryo. To address this question, in (Meistermann *et al.*, 2021), we analyzed single-cell transcriptomic data (scRNAseq) from preimplantation human embryos. scRNAseq data allows following individual cell fate within heterogeneous samples. Our analysis proposed a hierarchy of transcription factors in epiblast, trophectoderm and primitive endoderm lineages, the three-founding cell type of the human embryo. In this project, we aim to generate a computational model of human preimplantation development using single-cell transcriptomic data (scRNAseq).

Objectives

To date we have designed a method using logic programming to interrogate (1) scRNASeq data related to embryonic development, and (2) public gene regulation knowledge databases, such as Pathway Commons, and provide as output *static Boolean models* that explain gene regulation in two stages of embryonic development: medium and late trophectoderm (Bolteau et al., 2023). We want now to move forward and model consecutive steps and different fates of embryo development. The main tasks of this thesis project are:

- **Short term.** To explore other stages of embryonic development. For this, we need to conceive a logic program to extract a set of gene's (pseudo) perturbations associated to discrete gene expression across the 9 developmental stages. This information will be extracted using the scRNASeq datasets for all developmental stages. A first draft of the logic program has been implemented. However, it is not yet applied to a real data setting. If this program succeeds, the prospect becomes very interesting because we could apply a second method, previously developed by us in (Razzaq et al. *PloS Comp Biol* 2018), to obtain dynamic Boolean networks which could explain different evolutionary trajectories of embryonic development.

- **Long term.** To disrupt the computational model, and to alter the dynamics of a given evolution fate. An idea of method to provide this system disruptions (also called perturbations), will be to identify the system dynamical attractors, and then search for sets of perturbations which can guarantee arriving to a fate rather than to another one. This is feasible to implement with logic programs, in an approximation. However other type of verification solvers may be also more adapted to exclude false positive dynamics. This research direction can use the findings and methods of previous studies we have proposed (Videla et al. *Front. Bioeng. Biotechnol.* 2015) (Fitime et al. *Algorithms Mol Bio* 2017), and those proposed recently by (Chevalier et al. LNBI, 2020), all using logic programming.

Context

This thesis is funded by the **I-SITE** (Initiatives for Science, Innovation, Territories and Economy, Excellence label of French Universities) **NExT** (*Nantes Excellence Trajectoire*) program https://next-isite.fr/. Specifically on the context to promote interactions between Engineering sciences and Health. The PhD candidate will integrate the ComBi team of the *Laboratoire de Sciences du Numérique de Nantes* (LS2N, https://www.ls2n.fr/), that brings together Nantes' research expertise in computer science and cybernetics to develop digital sciences, inclusive of other disciplines and taking account of the social challenges involved. The LS2N ComBi (Combinatorics and Bioinformatics) team develops algorithmic and mathematical methods for the study of problems arising from biology. The team's main research themes focus on comparative genomics and systems biology. The PhD contract will be hold within the French Engineering school (*grande école d'ingénieur*), Centrale Nantes (https://www.ec-nantes.fr/). This will offer the PhD candidate the possibility to teach practical sessions within this institution. The starting date of this Phd thesis is programmed to **September 2024**.

Pre-requisite

The PhD candidate will have a Computer Science or Bioinformatic profile (Master degree or equivalent) with knowledge on logic programming or artificial intelligence. Previous experience of analysing (or computationally modelling) massive datasets of biological nature will be helpful.

Application

Please feel free to contact us if you have any question regarding the project, your match with the profile, or the application procedure to:

– carito[dot]guziolowski[at]ec-nantes[dot]fr
– laurent[dot]david[at]univ-nantes[dot]fr

Your application must contain : (1) your CV, (2) your Cover Letter stating your professional project, (3) your transcript from Bac +3 to Bac +5 or equivalent (for the results of Master or equivalent, attach the documents in your possession), and (4) contact information for 2 referees. Please send us your application by email before the 30/06/2024.

References

- (Meistermann *et al.*, 2021) Meistermann, D., Bruneau, A., Loubersac, S., Reignier, A., Firmin, J., et al. : Integrated pseudotime analysis of human pre-implantation embryo single-cell transcriptomes reveals the dynamics of lineage specification. Cell Stem Cell 28(9), 1625–1640.e6 (Sep 2021)
- (Bolteau et al. 2023) Bolteau, M., Bourdon, J., David, L., Guziolowski, C. (2023). Inferring Boolean Networks from Single-Cell Human Embryo Datasets. In: Guo, X., Mangul, S., Patterson, M., Zelikovsky, A. (eds) Bioinformatics Research and Applications. ISBRA 2023. Lecture Notes in Computer Science(), vol 14248. Springer, Singapore. (hal pre print)
- (Razzaq et al. *PloS Comp Biol* 2018) Razzaq M, Paulevé L, Siegel A, Saez-Rodriguez J, Bourdon J, Guziolowski C. Computational discovery of dynamic cell line specific Boolean networks from multiplex time-course data. PLoS Comput Biol. 2018 Oct 29;14(10):e1006538.
- (Videla et al. *Front. Bioeng. Biotechnol.* 2015) S. Videla, I. Konokotina, L. Alexopoulos, J. Saez-Rodriguez, T. Schaub, A. Siegel, C. Guziolowski "Designing experiments to discriminate families of logic models". Frontiers in Bioengineering and Biotechnology, 2015, DOI=10.3389/fbioe.2015.00131
- (Fitime et al. *Algorithms Mol Bio* 2017) L. Fippo Fitime, O. Roux, C. Guziolowski†, L. Paulevé†. Identification of Bifurcation Transitions in Biological Regulatory Networks using Answer-Set Programming, Algorithms Mol Biol. 2017 Jul 20;12:19. doi: 10.1186/s13015-017-0110-3.
- (Chevalier et al. LNBI, 2020) Chevalier, S., Noël, V., Calzone, L., Zinovyev, A., Paulevé, L., Paulevé: Synthesis and simulation of ensembles of boolean networks for cell fate decision pp. 193–209 (2020)