Learning the rules of multicellular self-organisation with interpretable machine learning

Abstract

This project will develop an interpretable AI framework to infer cell fate patterning mechanisms from sequences of microscopy images, and use this to identify evolutionary design principles underlying robust tissue patterning. We hypothesise that a set of motifs in molecular and cellular interaction networks can explain how developmental systems achieve both robustness and evolvability at the same time. Gaining a quantitative understanding of these motifs would enable a breakthrough to rationalise our strategies to control, manipulate, and repair biological systems.

Introduction

Cell fate patterning is a key biological process underpinning tissue formation during development and regeneration. Understanding the rules of multicellular pattern formation could accelerate the field of developmental biology, tissue engineering, and regenerative medicine.

A hallmark of living systems is their ability to self-organise in a manner that is both robust and evolvable. Robustness and evolvability appear to be opposing characteristics, and how biological systems combine these two properties is not understood. Robustness is the ability of a developmental system to yield a reproducible outcome despite genetic and environmental perturbations. Evolvability is the capacity to give rise to viable phenotypic innovations under evolutionary rewiring of genetic and intercellular interaction networks. While robustness provides resistance to intrinsic and extrinsic variability, evolvability allows exploration of phenotype space to maximise adaptability. Resolving this apparent dichotomy between robustness and evolvability in the context of tissue patterning will improve our understanding of living systems and inform biomedical engineering design principles.

We will focus on the specific biological process of axial elongation, whereby early vertebrate embryos break symmetry and undergo unidirectional elongation to lay down the tissues of the posterior axis of the body, including muscles, bones, and the spinal cord. The initiation of axial elongation is an ideal system to study the principles of robust and evolvable cell fate patterning, for which we have recently developed a tractable and standardised experimental system that recapitulates early stages of axial development in vitro.

Research Challenge: The methodological frameworks currently used to discover and validate the rules of tissue patterning is currently resource intensive and relatively limited in scope. Formulating hypotheses based on the current state of the literature is a slow, difficult, and error-prone process. Validating these hypotheses requires complex and costly experimental and computational testing. Furthermore, individual studies typically result in a single candidate mechanism specific to the tissue studied which makes it difficult to tackle big questions such as “Why are there different tissue patterning mechanisms in different tissues/organisms?”, “How are these different mechanisms related?”, and “How do they arise through the process of biological evolution?”.

Learning the rules of multicellular self-organisation with interpretable machine learning

Data & Methodology

Recent advances in AI may help to accelerate biological discovery research but have not been widely leveraged. It may be possible to use AI to generate and analyse a set of mechanistic models for any given patterning process and therefore assist researchers in selecting the most plausible mechanisms and identifying the most informative set of experimental validations. Demonstration of such a framework in practice does not exist yet, in part because many modern AI approaches suffer from a lack of interpretability in terms of biological mechanisms.

We have recently developed a Neural Cellular Automata (NCA) framework for application in biological datasets [Richardson et al. 2024]. As a "hybrid" approach that incorporates constraints similar to mathematical models, NCA have the potential to overcome the interpretability issue. This project will further develop the NCA framework to apply it on quantitative data of in vitro human developmental patterning [Robles Garcia et al. 2023] to generate plausible hypotheses about intercellular mechanisms and inform the biological discovery process. In particular, NCA models have the potential to be mapped to partial differential equations, a well-known class of models in applied mathematics and physics.

Training NCA on imaging data from organoids will yield a potentially large space of patterning mechanisms. Following this we will determine which subset of mechanisms (i.e. trained models) are robust and evolvable. We will computationally screen parameter perturbations in trained models and quantify pattern similarity to predict which experimental perturbations (e.g. biochemical signals or changing tissue geometry) are most informative to distinguish between plausible mechanisms and quantify robustness of patterning.

Understanding evolvability is important to understand how developmental mechanisms channel evolution along non-random trajectories to ensure rapid species adaptation. It can also help us establish engineering guidelines to produce evolvable multicellular systems for biomedical applications and provide causal explanation for the incomplete penetrance of genetic mutations. To evolve interaction rules in silico, we will randomly rewire the robust subset of interaction mechanisms computationally. These evolved interaction rules will then be selected if they produce a similar level of patterning (assessed by entropy measures) to their ancestor rules but at the same time are suitably different in pattern (assessed by information-theoretic distance measures). Interactions rules are deemed more evolvable if they have more selected descendants.

We believe that the systematic analysis of NCA-generated patterning mechanisms has the potential to enable a breakthrough in our understanding of the fundamental principles that underpin robust and evolvable patterning and will in turn help us rationalise our strategies to control, manipulate, and repair biological systems.

RRI/Ethical Considerations

The use of in vitro experiments enables the study of processes resembling human embryonic development, which is of great biomedical importance but otherwise ethically and technically challenging to investigate.

Learning the rules of multicellular self-organisation with interpretable machine learning

Expected Outcome & Impact

We anticipate that the tools we will develop will be broadly applicable well beyond the scope of this study, including many organoid systems in development and disease. Not only will this **new approach accelerate hypothesis generation and testing**, but it will also spearhead a **conceptual advance to understand multicellular self-organisation under the constraints of robustness and evolvability**, and the extended evolutionary synthesis hypothesis more widely.

We expect this study to establish NCA as interpretable AI frameworks to accelerate scientific discovery. There is the further possibility that evolvability is a desirable property of trained AI models more generally, beyond the application to biomedical data.

References

Garcia et al. *bioRxiv:2023.06.01.543323*.
Richardson et al. *arXiv:2310.14809* (accepted at PLOS Computational Biology)