**Worksheet for Developing a Mechanistic Biological Model**

For more information on the process of model definition, please refer to the article “Workflow for Developing Computational Biology-Based Models” in the Hackathon Materials Google Drive: <https://drive.google.com/drive/folders/1ckIYMHSUr9ypQN8dzRClJD5DKsfAYmn7?usp=sharing>

When you have an idea for a project, it’s tempting to immediately start coding a computational model. However, you are more likely to succeed in model-building if you take some time to carefully define what you want to learn from your model, what data are available to build and test it and what components you want to include in it.

**Part 1 Biological Model Definition**

This model, also known as a “conceptual model,” is often created with the help of a domain expert such as a biologist or a clinician. This model is an attempt to explain an observable reality. It should include the parts of the reality that we know about filtered by what we (and the domain expert) believe are relevant to a particular question. This description will include *physical objects* (cells, enzymes, tissues, …) and *processes* (cell proliferation, enzyme reaction, …). In addition, this description may include *spatial* and *temporal* information. Besides this list of objects and processes the biological model should also include a list of measurables and outcomes. For example, which of the physical objects are measurable in terms of count, or volume or concentration? Which of the processes are measurable such as in a time course? Finally, the biological model should identify *outcomes* of interest.

****

**Biological Observations of Your System**

1. **What phenomena are you interested in modeling?**

[For example: Tumor metastasis, or Tumor angiogenesis or Somitogenesis. In other words what is the biological process that you want to understand better]

1. **What data are available about your system?**

[Think about the data in the literature (often semi-quantitative), publicly available data, and data that can be measured with available technologies.]

**Biological Questions & Hypotheses.**

1. **What are the most important biological questions (hypotheses) about your system?**

[For example: What is the relative contribution of diffusion and transcytosis in morphogen gradient patterning during axis formation. Diffusion and regulation by chordin are sufficient to create a BMP morphogen gradient that is consistent with observed experimental data. Just like other biological hypothesis generation – identifying what’s necessary and sufficient are good ways to think about your system]

1. **What are the relevant metrics or outcomes for testing these biological questions?**

[For example: Gradient shape and positional information in bits are two relevant metrics for morphogen gradient patterning. In multi-scale models morphogenesis is often a relevant metric or outcome. (otherwise an ODE model would be good enough]

1. **What do you believe are the key processes involved in the phenomena of interest?**

[For example: cell proliferation. Anoikis, random-walk migration etc]

1. **What are the minimal physical objects needed in the phenomena of interest?**

[For example, if you were modeling a vascularized tumor you might say cancer cells in a tumor, and the endothelial cells of the vasculature are the minimal system. You might decide to leave out the different types of endothelial cells if you are not that interested in how they work.
Alternatively, if you were really interested in the immune reaction – you might choose to ignore the vasculature and just focus on tumor, and immune cells.
Also remember to think beyond cell types. Are there any morphogens or chemical gradients involved? Are there any physical forces that are relevant]

1. **Think carefully about the flow of information in your system. How do signals propagate within the system (both within and between components) and how are these signals mediated (by contact, by diffusible chemicals, by force transduction,..)?**

[]

1. **What spatial and temporal scales are relevant?**

[How long do your phenomena take to happen? How far apart in space and time do different processes happen? Is that relevant? Is there compartmentalization?]

1. **Is noise relevant? Are mechanical cues relevant? Why do you need a multi-scale model?**

[]

1. **Describe your qualitative model. Drawing is often useful.**

[]

**Mathematical Model**

The mathematical model is a mathematical description of the biological model. Here we create mathematical definitions, based on physics, chemical, or other physical models, of the objects and processes. Often the creation of the mathematical model requires making significant assumptions about the underlying mechanisms of the biological system. For example, when modeling transfer of a small molecule into a cell is the process simple diffusion (that can be modelled as a reversible first order ordinary differential equa­tion (ODE)) or is the process trans­porter mediated that may become saturated at high small molecule concentration? This process of converting a biological model into a computational model is the first point where the modeling process adds value to our understanding of the biological system. The mathematical model requires a level of understanding and specification that is rarely present in biological models. The modeler, often with the help of the domain expert, must coerce the available biological knowledge into a numeric framework and in that process decide how the available data can be used to select from a multitude of possible mathematical instantiations. At this stage it often becomes clear that certain aspects of the biological system that are measurable have indeed never been measured. Often, at this stage the modeler also encounters the problem that quantities required in the mathematical model are not directly measurable in the biological assays. This situation may require reformulating the mathematical model to avoid dependence on intrinsically unmeasurable quantities. On the other hand, at this point the mathematical model can also begin to offer new insights into the biological model. Quantities that are not directly *measurable* may instead by *calculable* by the model. This capability is one of the most attractive aspects of a mathematical model. A key aspect of the creation of the mathematical models is that it is the point where the model moves from a biological point of view to a chemical or physics-based description. This allows the modeler to use standard mathematical forms developed in those domains to define the model for the biological domains

1. **Identify relevant models in the literature. There is often at least an ODE model, if not a spatial or CC3D model that is relevant.**

[]

1. **Identify ‘gaps’ in those models that you plan to address.**

[]

1. **Identify components of those models that you can use in your models.**

[]

**Roadmapping a CompuCell3D Simulation**

1. **XML Definition**
	1. **What cell types are in your model?**
	2. **Where are those cell types?**
	3. **What size are the cells or other spatial components?**
	4. **How big is your domain?**
	5. **What are your boundary conditions?**
2. **What processes will your biological components experience?**
	1. **Which plugins will you use?**
	2. **Migration**
	3. **Diffusion**
	4. **Adhesion**
3. **ODE definition**
	1. **What subcellular, intercellular and large-scale processes will you describe with ODE models?**
	2. **For subcellular models (signaling, regulatory and metabolic networks) how do the outputs of these models relate to the parameters controlling cell-level processes (*e.g.* adhesion strength, cell growth and division, cell death, rate of response to external chemicals, secretion and absorption of extracellular components)?**
4. **Model integration**
	1. **What are the feedbacks between the model components? Especially between the ODE systems and the cells and tissue-level components?**