On cellular control networks

Back in graduate school, I became interested in systems biology, but I never quite found exactly what I was looking for. I've stumbled through papers on gene-set enrichment analysis, grouped association tests, critical state transitions, stochastic modeling and related inference methods, large-scale perturbation experiments, and correlation networks. Some of these enable highly detailed modeling motivated by both the biology and the underlying physics. Unfortunately, these do not scale well; they tend to include fewer than 20 interacting genes and proteins, whereas modern sequencing datasets routinely measure tens of thousands of genetic elements. Others grapple head-on with genome-wide inference, but do not explicitly model direct interactions between transcriptional regulators and their targets. So, they can describe cells in lots of detail, but they give limited insight into the mechanisms controlling cell state.

Some time over the last three years, I managed to pin down my questions, and I began finding some of the right papers to satisfy my curiosity. These projects attempt to infer what regulates each gene's activity, and they do it using genome-wide datasets. In the context of my day job, the ultimate goal coming out of that type of model would be a stem cell simulator. Given an RNA dataset, and optionally other inputs such as genes to be overexpressed or small molecules to be dumped on the cells, I want to computationally extend the data weeks into the future. Running many such simulations would help the lab understand how initial conditions affect stem cells' fate.

We are a long way from that ideal, but there is already a daunting amount of work towards it, and that is what I am beginning to explore in this post series. I will try to tackle select examples of existing work in small, coherent batches.

- Quick & dirty biology for people from other fields
- Identifiability of gene regulatory networks
- Data resources: what's out there?
- Evaluating network inference in practice: the DREAM5 competition
- How do people use network inference to study stem cells or developing cells?
- ChIP-seq, ATAC-seq, enhancers, and motif analysis
- Methods of pairing enhancers with their target genes
- Inferring gene regulatory networks from combined RNA and DNA sequencing data