John Vastola

Comments on
"An enriched network motif family regulates multistep cell fate transitions with restricted reversibility"
by Ye et al.

Summary: A cell differentiating (perhaps in development, or in a reprogramming or pathological context) from one cell type to another typically encounters intermediate states; while there has been much work on the mathematical modeling of cell fate transitions controlled by a bistable switch-type architecture, the authors point out that less work has been done studying simple GRNs that generate a cell fate transition with multiple (stable) intermediate steps. The authors identify a network motif (or "design principle") that robustly generates networks with two intermediate states between "initial" and "final" cell types, and use this motif to study T cell development. Their model recapitulates many experimental results (including the order that various transcription factors are turned on and off), and they study its stability, parameter sensitivity, and state transitions in some detail.

In context: To my knowledge, the authors' point about the lack of emphasis on simple GRNs producing multi-step transitions is a good one. They offer concrete examples of when it is important to think about (T cell development, EMT, skin development), and follow up with a detailed study of early T cell development. Their results are carefully validated with many mathematical tools, and important experimental results are recapitulated by their model. Their methods can easily be used to study other problems.

Major comments: Overall, I think the paper is carefully argued, and makes an interesting contribution to the study of GRNs. Also, the authors pointed out many times when their model produced results consistent with experiment, which I found impressive. The structure of the paper is reasonable, and the tables/figures are generally helpful.

It would be helpful if the authors defined "stepwise", which is used many times in the article. Intuitively one has a reasonable idea, but the precise mathematical meaning is only made clear when one sees the bifurcation diagram in Figure 6.

The authors only consider their model's robustness to additive noise, but that is not a huge deal, because that simplification is extraordinarily common, and I doubt their results would be qualitatively any different for state-dependent noise. As a related issue, their use of the Freidlin-Wentzell action to find transition paths is customary in this situation (although it would not be appropriate for non-additive noise).

I was a little mystified regarding what the authors meant by (see lines 260 and 280) 'enhancing' "the ability of the network to produce this system". This is explained in the Methods section (see line 648), but I still don't really understand why they did what they did. As a related issue, it would be nice if they explicitly said that they are quantifying robustness (for example, see line 291) by how large the parameter space region with the desired behavior is.

Minor comments: I have found a bunch of typos/grammar errors.

- Line 84: "Despite **of** the accumulating" (remove the "of")
- Line 131: "**Base** on the complexity atlas" (should be "based")
- Line 260: "ability of the network to produce this **the** system." (remove "the")
- Line 303: "attenuation or **withdraw** of the" (should be "withdrawal")
- Line 304: "the **revert** of the transcription profile" (should be "reversion"; note that revert is used in this way in a few places)
- Line 315: "does not allow the **revert** of lineage progression" (same as previous)
- Line 340: "We next asked how the **duration Notch** signal may control" (should be "duration OF THE Notch signal")
- Lines 422, 435, 440, 457, 458, etc: "**the** T cell development" (should be just "T cell development")