John Vastola

10/18/18

**Journal:** Experimental Biology and Medicine

Thematic issue dedicated to “provocative concepts on the control of stem cell differentiation”

**Title:** From developmental metaphor to quantitative framework: A review of Waddington landscapes and their applications

**Abstract (exactly 150 words):**

Achieving precise control of stem cell differentiation remains one of the most challenging open problems in biology. Some experimental reasons for this will be discussed elsewhere in this thematic issue; however, an important theoretical reason precise control remains elusive is that it is not clear how best to think about one cell changing into another type of cell. The epigenetic landscape view of differentiation, where one imagines a cell as a ball rolling around a rugged pasture full of hills and valleys, offers one way to think; however, it has proven difficult to make this conceptual picture mathematically precise, partly because there are many inequivalent proposals for accomplishing this. In this review, we discuss how to think about the landscape, its different mathematical formulations, and applications to various areas of biology. We conclude with the practical issues currently preventing the landscape from being a useful, predictive tool for understanding differentiation.

**Comments on abstract:** You might note that this abstract is considerably less technical than our previous one: this is on purpose. Whoever reads this thematic issue is probably principally interested in stem cell differentiation, and probably has little mathematical background. For those reasons, I think the abstract should be an invitation to think about the landscape for its promise as a way to think about cells changing state.

It’s probably not necessary to be specific here with the formulations we will discuss, or some of the mathematical subtleties we will emphasize (the different between additive and state-dependent noise, the difference between being interested in global relative stability and local state transitions); they will mean nothing to the experimental biologists reading the abstract, and landscape experts already expect them to be included.

If you gave me more words, I could have included more. But, to be honest, I think this does the job.

**Previous abstract:** Waddington’s epigenetic “landscape” was originally a metaphor for conceptualizing the link between gene regulation dynamics and developmental processes. In recent times however it has become a powerful theoretical framework to map the lineage dynamics of cells in a wide array of contexts, developmental and otherwise. In particular, it has become a powerful method for interpreting the dynamics of model systems describing stem cell dynamics and has the potential to help improve our understanding and control of directed cell differentiation. But how does one quantify the “landscape”? What does the term landscape even mean in a precise quantitative sense? And how can that landscape be used to improve our understanding and control of cells’ states? Here we review different theoretical formulations / interpretations of the “landscape” concept (e.g. Fokker-Planck, quasi-potential, etc. approaches) as well as methods for charting transition paths within them (e.g. least action paths). We also discuss practical issues that must be addressed if these methods are to become predictive, including reconciling the differences between different landscape interpretations, understanding the sensitivity of these landscapes to different types of stochastic noise, the computational burden of landscape construction in high dimensional systems, the need for effective dimensional reduction, and the need to integrate theory with data in these constructions.

**First two paragraphs:**

Changing some skin cells from a patient into heart cells or brain cells—at least, cheaply and in high enough volumes to be clinically useful—has proven to be harder than understanding fundamental properties of nature (like the anomalous magnetic moment of the electron) to ten decimal places, or using astronomical observations to estimate how long ago the universe began. *Why?* Why isn't our encyclopedic knowledge of molecular biology, medicine, mathematics, and computation enough?

There are many, many, many good reasons, some of which will be discussed elsewhere in this thematic issue on iPSC differentiation: difficulties with experimental set-up, culturing, automation, data collection and analysis, and so on. But there are still other reasons, which prevent even the best-designed equipment in the world from being enough to solve the problem on its own. They boil down to the following: how should we *think* about iPSC differentiation? More generally, how should we think about one cell changing into another type of cell?

**Comments on first two paragraphs:**

I really like these two paragraphs, because they frame the problem in an approachable way of interest to someone that works on stem cell differentiation: you need something like the landscape as a way to think about cells changing state, and not because the landscape is just something that you have to learn because lots of people work on it.

They also avoid the usual cliché in landscape articles of starting with: “Waddington’s landscape, first imagined by Waddington in 1957, is a metaphor for gene regulation. Ball rolling down hill bla bla bla…” This is fine, but it’s overdone, and doesn’t tell you why you should care.

I know you will hate them, and will think they are fluffy. But I think academic writing is allowed to be interesting as well as informative.

**OUTLINE**

1. Introduction
2. How to think about the landscape
   1. The standard conceptual picture
   2. Problems with the standard conceptual picture
   3. What factors determine the landscape?
   4. Did the landscape change, or did the ball move?
   5. Is there just one landscape, or are there many?
3. Landscape properties
   1. Generic properties of all landscapes
   2. Global vs local landscapes
   3. Mathematical definition
   4. Intuition for mathematical definition
4. Different landscapes
   1. Simplest model: Markov chain
   2. Steady-state Fokker-Planck landscape
   3. Local and global quasipotential landscapes
   4. Vector decomposition landscapes
5. Applications
   1. Landscapes in reprogramming
   2. Landscapes in developmental biology
   3. Landscapes in cancer biology
6. Pitfalls, issues, and directions for future research
   1. Transition path predictions
   2. Dimensional reduction
   3. Sparseness of available data
   4. Computational burden
7. Conclusion