Available online at www.sciencedirect.com
ScienceDirect
Journal of Theoretical Biology 245 (2007) 610–626
Journal of Theoretical Biology
www.elsevier.com/locate/yjtbi

Analytic formulas for discrete stochastic models of cell populations with both differentiation and de-differentiation

Scott Hotton[a,b,*], Mike E. Colvin[a]

[a]School of Natural Sciences and UC Merced Center for Computational Biology, University of California, Merced, CA 95344, USA
[b]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA

Received 24 November 2005; received in revised form 6 October 2006; accepted 8 November 2006
Available online 15 November 2006

Abstract

Cell differentiation often appears to be a stochastic process particularly in the hemopoietic system. One of the earliest stochastic models for the growth of stem cell populations was proposed by Till et al. in 1964. In this model there are just two cell types: stem cells and specialized cells. At each time step there is a fixed probability that a stem cell differentiates into a specialized cell and a fixed probability that it undergoes mitosis to produce two stem cells. Even though this model is conceptually simple the myriad of possible outcomes has made it difficult to analyse. We present original closed-form expressions for the probability functions and a fast algorithm for computing them. Renewed interest in stem cells has raised questions about the effect de-differentiation has on stem cell populations. We have extended the stochastic model to include de-differentiation and show that even a small amount of de-differentiation can have a large effect on stem cell population growth.

Keywords: Differentiation; De-differentiation; Stochastic population models; Stem cells

1. Introduction

The differentiation and division of cells in multicellular organisms is an exquisitely complex and finely coordinated process. Because of this complexity, mathematical models can play an important role in explaining experimental results and testing new hypotheses. Indeed, it is widely observed that biology is at the threshold of a vast transformation, in which predictive modeling will play an increasing role in advancing biological understanding (Brent, 2000; Ideker et al., 2001). A recent review describes the prospects for understanding differentiated cell populations as "complex adaptive systems" (Theise and d'Inverno, 2004). Specifically, these authors argue that the development of even highly abstract models of stem cell lineages can provide both new understanding of diseases related to cell development as well as new perspectives on major unresolved questions in stem cell biology.

The first and one of the most frequently cited mathematical models for stem cell populations was proposed in 1964 (Till et al., 1964) (hereafter referred to as the "Till model"). Even though this was limited to a single probabilistic decision between proliferation and differentiation this study yielded the significant result that a simple stochastic model could reproduce experimental distributions for the number of proliferative cells in newly formed spleen colonies. This paper presents newly derived closed-form mathematical expressions describing the Till and related models of stem cell populations. In particular, we introduce de-differentiation into the model and determine its effect on population dynamics.

Although the Till model is conceptually simple, the myriad of possible outcomes has made it difficult to analyse and reduced its utility for many stem cell researchers. Most previous studies of this and similar probabilistic models have involved Monte Carlo simulations in which many trials of the cell lineage are generated using random numbers to select each probabilistic choice. Although Monte Carlo methods are easy to implement even for complex models, they yield only specific numerical

*Corresponding author. Tel.: +1 617 496 5945; fax: +1 617 496 5854.
E-mail address: shotton@fas.harvard.edu (S. Hotton).

0022-5193/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jtbi.2006.11.006

outcomes for predefined models and model parameters. Analytical solutions for probabilistic models have the advantage of providing exact mathematical relationships between model parameters and behavior (e.g. an equation relating the extinction probability for an entire stem cell lineage to the stem cell division rate). Moreover, analytical expressions describing the resulting cell populations (such as population size mean and variance) can be used to reduce diverse experimental data sets to a few parameters that characterize the population dynamics. Finally, even with the vast improvements in computer speed, the relatively slow convergence of Monte Carlo results (typically as the square root of the number of trials) limits the usefulness of such methods for predicting the effects of low-probability events, such as de-differentiation described in this paper.

In the Till model time progresses in discrete steps of a fixed length. At each time step an individual stem cell either differentiates into a new cell type or else undergoes mitosis to produce two new stem cells. The probability for an individual stem cell to differentiate is denoted by p_0 and the probability that it undergoes mitosis is denoted by p_2 (Fig. 1). These probabilities are fixed in time and apply to every stem cell in the population. In this simple model the differentiated cells are regarded as performing some function for the organism so they persist unchanged—neither differentiating nor dividing. Every cell acts independently of the others and there is no competition effect when the size of the population becomes large.

In the 40 years since the Till model was first published there has been an explosion of research into stem cells and differentiation. Intriguing data has been obtained suggesting the possibility of trans-differentiation (the transition of a cell from one developmental pathway to another) and even de-differentiation (the transition of a cell to a less differentiated state). For example, Quesenberry et al. (2002) have proposed a "chiaroscuro" nature for stem cells as opposed to a hierarchial nature. Based on their lab work, as well as the work of others (Suda et al., 1984; Pietrzyk et al., 1985; Bradford et al., 1997; Cheshier et al., 1999), they conclude that stem cells are not completely quiescent but that evidence for their mitotic activity requires long-term observation. When cells enter S phase many internal cell structures are dismantled. Quesenberry et al. speculate that the phenotype of a stem cell can be reversibly modulated during the cell cycle, that stem cell identity can be masked in a population of asynchronously dividing stem cells, and that cell fate can depend on an interaction of cell phase with the cell's microenvironment.

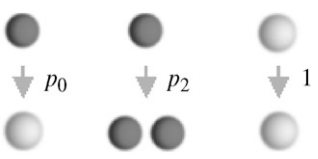

Fig. 1. The possible transitions in the Till model. Stem cells are represented with dark filled circles and differentiated cells are represented with light filled circles. On the left a stem cell differentiates with probability p_0. In the middle a stem cell divides into two stem cells with probability p_2. On the right a differentiated cell persists unchanged with probability 1.

This explosion of research has even raised the question of how to define stem cells. Of particular note is the functional definition of Potten and Loeffler (1990) which has been cited many times. While this definition has proven useful it has been amended by Loeffler and Roeder (2002) in light of recent evidence of trans-differentiation and de-differentiation and their definition for stem cells includes among its criteria flexibility and reversibility. In 2004 Kirkland proposed a phase space model (Kirkland, 2004) in which the stemness of a cell is conceived of as a continuous quantity which can increase or decrease over time. To avoid complications we will only consider a finite number of cell types in this analysis.

It is still not clear how common trans-differentiation or de-differentiation are in animals or even if they happen at all in vivo (Blau et al., 2001; Wagers and Weissman, 2004). Even so, these transitions could greatly increase the complexity of the differentiating cell population. Most of the theoretical studies of stem cell populations have involved Monte Carlo simulations at the level of individual cellular decisions. Although such simulations are easy to implement and computationally feasible, they do not provide generalizable results on how the properties of individual cells (e.g. the values of p_0, p_2 in the Till model) affect the dynamics of entire cell populations. For this reason we have developed closed-form expressions for a generalization of the Till model that includes de-differentiation. This leads to the conclusion that even a small amount of de-differentiation can have a large impact on stem cell population size.

2. Mathematical analysis of the Till model

In this section we analyse the Till model and determine the types of population growth patterns that it exhibits.

2.1. The Till model as a transformation of probability functions

Although the Till model is stochastic it is still possible to make quantitative predictions with the model and to compare those predictions with experiments. The starting generation will be called generation 0. The next generation is generation 1 and so on. We will indicate the generation with superscript notation. We let N^t denote the number of stem cells in the population at generation t. Because the value of N^t depends on random events it is a random variable. We cannot say for sure that N^t will have some particular value for a given population but every possible value that N^t can take will have a precise probability of occurring. In other words at each generation the Till model provides us with a probability function for the random variable N^t. This probability function associates to each