

ADVANCED REVIEW**Kinetic models of hematopoietic differentiation**

Victor Olariu | Carsten Peterson

Department of Computational Biology, Lund University, Lund, Sweden

CorrespondenceCarsten Peterson, Department of Computational Biology, Lund University, Lund, Sweden.
Email: carsten.peterson@thep.lu.se**Funding information**

National Institutes of Health, Grant/Award Number: 18B-1095882; Swedish Research Council, Grant/Award Number: VR 621-2013-4547

As cell and molecular biology is becoming increasingly quantitative, there is an upsurge of interest in mechanistic modeling at different levels of resolution. Such models mostly concern kinetics and include gene and protein interactions as well as cell population dynamics. The final goal of these models is to provide experimental predictions, which is now taking on. However, even without matured predictions, kinetic models serve the purpose of compressing a plurality of experimental results into something that can empower the data interpretation, and importantly, suggesting new experiments by turning “knobs” in silico. Once formulated, kinetic models can be executed in terms of molecular rate equations for concentrations or by stochastic simulations when only a limited number of copies are involved. Developmental processes, in particular those of stem and progenitor cell commitments, are not only topical but also particularly suitable for kinetic modeling due to the finite number of key genes involved in cellular decisions. Stem and progenitor cell commitment processes have been subject to intense experimental studies over the last decade with some emphasis on embryonic and hematopoietic stem cells. Gene and protein interactions governing these processes can be modeled by binary Boolean rules or by continuous-valued models with interactions set by binding strengths. Conceptual insights along with tested predictions have emerged from such kinetic models. Here we review kinetic modeling efforts applied to stem cell developmental systems with focus on hematopoiesis. We highlight the future challenges including multi-scale models integrating cell dynamical and transcriptional models.

This article is categorized under:

Models of Systems Properties and Processes > Mechanistic Models
Developmental Biology > Stem Cell Biology and Regeneration

KEYWORDS

energy landscapes, hematopoiesis, rate equations, reprogramming, stem cells, stochastic simulations, systems biology

1 | INTRODUCTION

Modeling kinetics in hematopoietic processes takes place at two levels; on a cell level with population models and inside the cells with gene/protein dynamics. We confine ourselves in here to the latter. Good reviews exist covering population models in this context (see e.g., MacLean, Celso, & Stumpf, 2016). However, initiatives to fuse the two levels will be touched upon below. Many solid predictive results exist from extracting and analyzing sub-networks for hematopoiesis, which are based entirely upon bioinformatics and statistical analysis. Whereas these techniques are often important for setting the stage for dynamical modeling, they go beyond the scope of this review. For an earlier review containing dynamical approaches for hematopoiesis with focus on diseases, see Whichard, Sarkar, Kimmel, and Corey (2010).

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2018 The Authors. *WIREs Systems Biology and Medicine* published by Wiley Periodicals, Inc.

Kinetic models for determining stem cell fates aim at understanding the role of the interacting genes known to be active in relevant established core network modules. Such models necessarily contain a plurality of parameters and options. One might at first sight think that these necessarily lead to overtraining, with limited data at disposal. However, as it turns out with appropriate questions to ask, this is not necessarily the case. For example, by exploration of different candidate networks with models, fitting to data can single out interesting solutions. Relatedly, different models can guide to what cell states (or attractors) that data exhibits. With models in place for commitment, what-if computer experiments can be performed to develop efficient reprogramming protocols. Furthermore, non-linear dynamics analysis provides insights into whether a cell state transition is irreversible or not. Such mechanistic models appear in two brands, Boolean and continuous-valued models respectively, where the latter either come as deterministic rate equations or their stochastic counterparts. Both approaches have their pros and cons. Whereas continuous-valued models typically contain at least three parameters for each incoming gene interaction vertices, correspondingly a multitude of logical rules (AND, OR, ...) can govern the outcome in Boolean approaches. For mapping out cell states (attractors) and barriers in between for large gene networks, a Boolean approach can be advantageous and likewise when exploring reprogramming trails. This is in particular true when no dynamical data are available. With time series available, continuous-valued models may be more relevant, in particular for establishing switch properties. Also, the real world is far from binary.

In what follows, we will first briefly describe approaches for modeling gene/protein expression dynamics, with deterministic rate equation approaches along with their stochastic implementations as well as with Boolean models. To this end, we will discuss when the different approaches are suitable and how to use the models once established for, for example, model selection and bifurcation analysis. With the understanding of the tools in place, we then proceed with a variety of applications in stem cell developmental systems with focus on hematopoiesis.

Informative overviews covering some of the topics presented here include Papatsenko & Lemischka (2016) and Herberg & Roeder (2015), albeit with more focus on embryonic stem cells.

2 | GENE EXPRESSION MODELS—GENERIC

We first discuss the two different approaches to gene network dynamics, that is, Boolean and continuous-valued models, respectively, followed by brief notes on how to deal with noise and how to explore the corresponding free energy landscapes. Implementation details can be found in the subsection thereafter. Throughout we will use “gene” as representing both incoming proteins and transcribed genes. The translation of mRNA to proteins will be assumed to be absorbed into the formalism.

2.1 | Boolean models

Historically, gene transcriptional models first appeared with binary representations where a gene is either ON or OFF, taking values 0 or 1, respectively, with Boolean interactions for the incoming genes, for example, $A \text{ AND } B \rightarrow C$, where A and B are inputs and C an output (see details in next section). The dynamics arise from updating each gene or node in a network according to Boolean logics until reaching attractors like steady states. The steady states were early on proposed to be identified as cellular states (Kauffman, 1969) whereas other types of attractors may correspond to periodic behavior like cell cycling. These ideas were very likely triggered by the Jacob and Monod experiments on the Lac Operon in *Escherichia coli* suggesting and identifying a gene regulatory network controlling the enzyme levels in cells (Jacob & Monod, 1961). The original Kauffman (1969) model was further developed allowing for multilevel logics, asynchronous updating and more fine-grained kinetics (Thomas & D'Ari, 1990). The early days of Boolean transcriptional models were dominated by issues like what conditions, number of genes and rule choices allow the existence of stable attractors. The number of possible rules for each node increases with K like 2^{2^K} , where K is the number of input genes—16 and 128 for $K = 2$ and 3, respectively. A subset of these, denoted nested canalizing, gives rise to stable solutions (Kauffman, Peterson, Samuelsson, & Troein, 2004). Most combinations of “everyday logics” like AND, OR and NOT belong to this group. Early applications of the Boolean framework confronting real biological data include control of plant flowering (Mendoza, Thieffry, & Alvarez-Buylla, 1999), cell cycle dynamics (Faure, Naldi, Chaouiya, & Thieffry, 2006) and sea urchin transcriptional studies in a full network of 20 genes with verified predictions (Peter, Faure, & Davidson, 2012). Of note, this work also integrates cell–cell signaling interaction geometries in a known cell division framework with gene expressions inside every individual cell synchronously using an absolute time scale. The latter is required to maintain the linkage to cell division. The Boolean framework is currently often employed for fitting rules to established architectures subject to assumed attractor structures, which will be dealt with below in the context of hematopoiesis. Evidently, a potential drawback in Boolean models is the discretization of gene expressions. However, Boolean models can on a crude level, with limited number of rules, capture the kinetics of a transcriptional network often with fewer parameters than in continuous-valued models to be discussed next.

2.2 | Continuous-valued models

In the real world, gene expression values are not binary but rather continuous. Here, one considers transcription as an equilibrium process, that is, assuming that the binding-unbinding of a transcription factor (TF) and RNA polymerase (RNAP) is fast as compared to other reactions involved, for example, translation and decay. Hence, textbook formalisms from equilibrium chemistry can be employed like Michaelis–Menten kinetics or Hill functions when cooperativity is involved, both of which display a non-linear behavior with saturation. With these rate equation approaches one typically has three parameters per input on each interaction node. Another variant is the highly intuitive Shea–Ackers approach (Shea & Ackers, 1985; Ackers, Johnson, & Shea, 1982) with its roots in statistical mechanics. One difference here is that the parameters are not directly kinetic constants. These approaches give rise to similar deterministic rate equations that for a gene network or module display non-linear behavior with solutions defining attractors and whenever present they can pinpoint irreversible behavior. The latter is highly relevant for developmental processes. Early work on continuous-valued gene expression models was performed for *Drosophila* development (Reinitz, Mjolsness, & Sharp, 1995), *E. coli* (Collins, Gardner, & Cantor, 2000) and cell cycle dynamics (Chen et al., 2000), where stability properties and oscillatory solutions were analyzed, respectively. These deterministic approaches assume a large amount of molecular copies, which is the case for bulk gene expression measurements.

2.3 | Noise

It is in general conjectured and to some extent also demonstrated that in developmental processes like stem cell commitment, the transitions are triggered by external or internal noise, where the latter is expected to originate from few molecular copies of key genes. In such cases, the deterministic rate equations provide probabilities for stochastic events depending on the same interaction strengths (Gillespie Monte Carlo algorithms; Gillespie, 1977) as in the deterministic encodings above. In the Boolean modeling context, the noise can be incorporated by implementing Probabilistic Boolean Networks (PBN) where the function for each gene is sampled from a group of acceptable functions, each function having an assigned probability (Shmulevich, Dougherty, Kim, & Zhang, 2002). Another viable option is employing the Stochastic Boolean Networks, which provide an accurate and efficient simulation of a PBN without and with random gene perturbations, where the computational efficiency depends only on the number of genes but not on the number of possible networks like in the PBN case (Liang & Han, 2012).

2.4 | Landscapes

Enforced guiding of developmental processes could benefit from dynamical modeling by tuning parameters for protein concentrations and other factors. However, exhaustive scanning of different parameters is not practical in a rate equation setting. A more profitable approach would be to map out the corresponding free energy landscape. The latter concept goes back to the Waddington landscape metaphor, which is frequently used to qualitatively visualize developmental processes (Ferrell, 2012). The underlying idea is that the dynamics of biochemical equations, governing a specific developmental process, could be represented as movements in an energy landscape such that lineage choices are paths between stable cell states. This notion is based upon a potential correspondence between solving the equations of motion and minimizing the corresponding free energy, a relationship often true in physics but not necessarily in biochemistry. Nevertheless, different approaches to approximate the energy landscape have been explored for small systems (Bhattacharya, Zhang, & Andersen, 2011; Olariu, Manesso, & Peterson, 2017; Wang, Zhang, Xu, & Wang, 2011; Zhou, Aliyu, Aurell, & Huang, 2012). In Olariu, Manesso & Peterson (2017), the Hill functions, which have sigmoid-like behavior, are replaced by true sigmoids, $g(x) = 1/(1 + e^{-2x})$, where x is the incoming signal. In a Boolean framework, the corresponding energy landscapes should be simpler to compute in general, by assigning low energy levels for cases where operator outcomes are true and high energy for false.

2.5 | Are Boolean and continuous-valued approaches related?

Not unexpectedly, the Boolean and rate equation approaches are related since the latter takes binary values in the high gain limit as first discussed in Glass & Kauffman (1973) and later elaborated upon in Thomas & D'Ari (1990). In particular, when using sigmoidal rate equations one has a more formal relation to the Boolean approach (Olariu et al., 2017), which can be formulated as energy functions containing products of binary spin variables. Minimizing these energy functions from sigmoidal rate equations turns out to represent approximate solutions. In a noisy environment (or temperature) the solutions take continuous values between 0 and 1. Along similar lines, detailed physical models were developed in Buchler, Gerland, & Hwa (2003) giving rise to Boolean rules.

3 | DEVELOPING GENE EXPRESSION MODELS

The development of gene expression dynamical models will be illustrated with a single example, that of two positively self-interacting genes A and B that are mutual antagonists (see Figure 1). This is a paradigmatic motif that frequently occur in connection with switch behavior and which is relevant for many stem cell transitions. The work to be discussed later is easily handled with the formalism and techniques illustrated in this example and they mostly follow an implementation strategy described in Figure 2.

3.1 | The Boolean approach

For Boolean modeling, the gene expression can be either high or low corresponding to 1 or 0, respectively.

For example, in Figure 1 considering that only AND and OR logics (most common) are possible, the expression levels of the two genes become:

$$A = A \wedge \neg B \text{ if AND logic is chosen,}$$

$$A = A \vee \neg B \text{ if OR logic is chosen.}$$

Similarly

$$B = B \wedge \neg A \text{ for AND logic,}$$

$$B = B \vee \neg A \text{ for OR logic,}$$

where \wedge is the logical conjunction symbol describing AND Boolean logic, \vee is the logical disjunction symbol describing OR Boolean logic and \neg is the negation symbol.

One updates the genes until a steady state has been reached. This can be either done synchronously as was done originally in Kauffman (1969) or asynchronously. For the latter, where updated values for the input genes are used, there are two

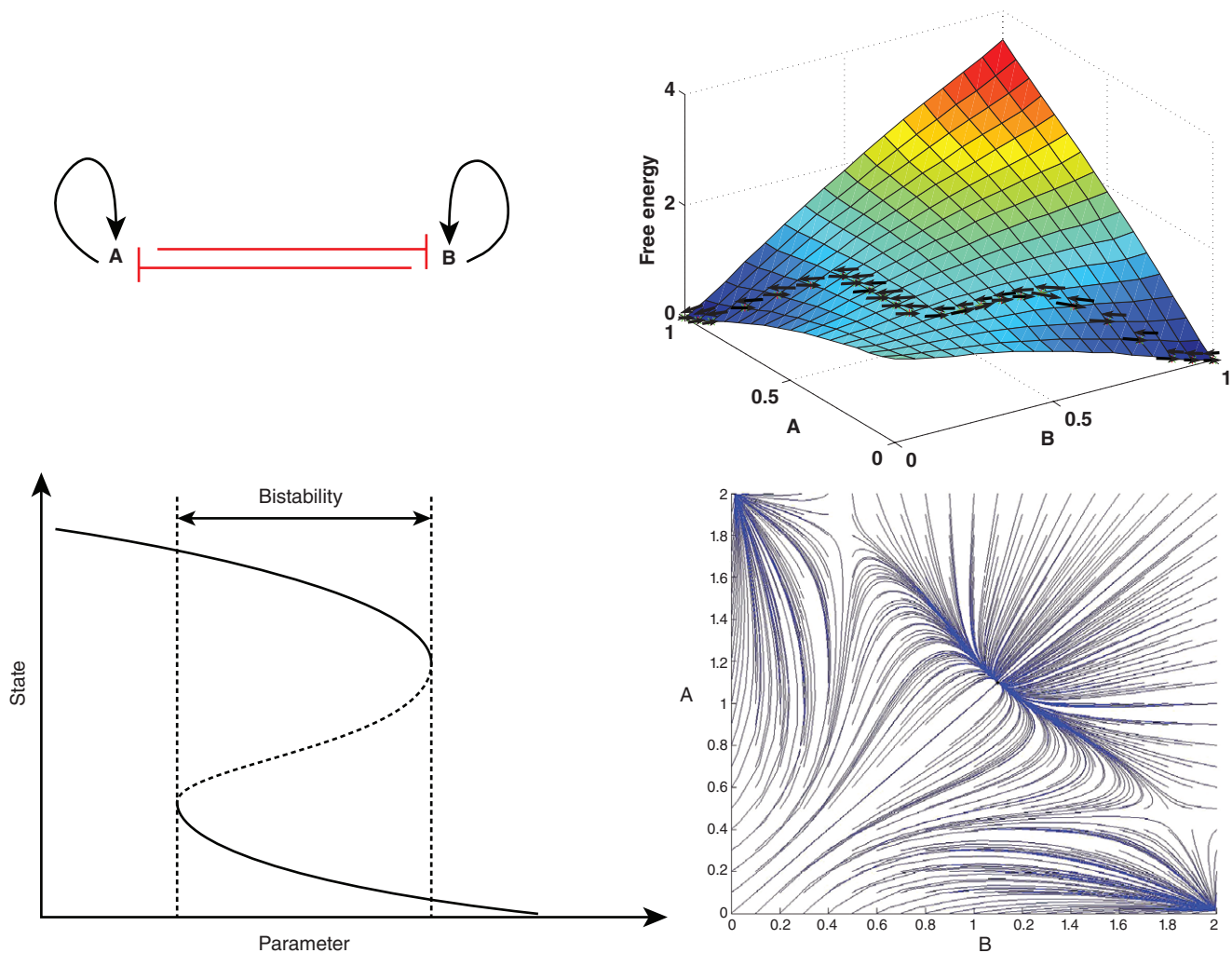


FIGURE 1 Switch architecture with two mutually repressing genes A and B, which are both positively self-interacting along with corresponding energy landscape, bifurcation diagram, and attractor space

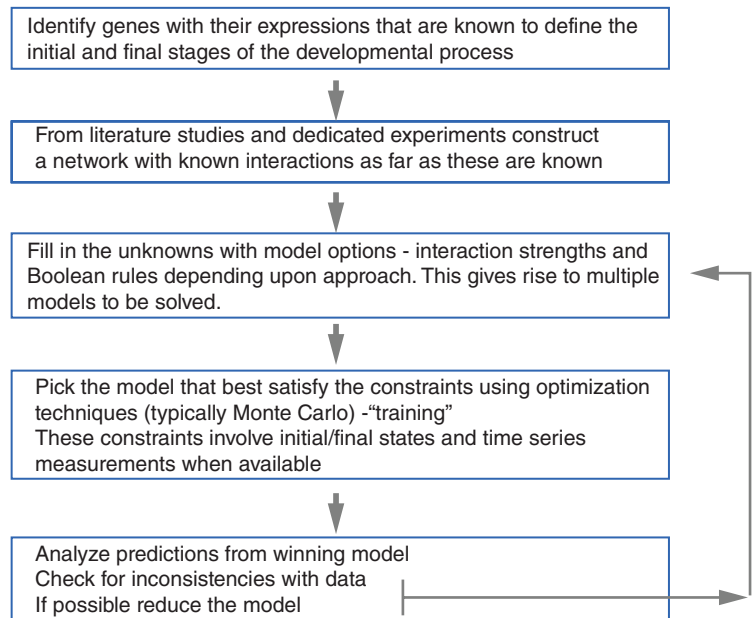


FIGURE 2 Workflow when using computational models to deduce relevant components and interactions

alternatives, serial or random updating. In the serial case, one updates according to a given order. For the other option, one picks a gene randomly to be updated. Issues like convergence properties and reachability of attractor states related to these choices have been studied in Boolean (Greil & Drossel, 2005) and neural networks (Grondin, Porod, Loeffler, & Ferry, 1983). Also, asynchronous updating can be achieved by generating non-deterministic state transition graphs (Thomas & D'Ari, 1990). Needless to say, asynchronous updating is in general closer to a biological reality.

Once the best logical combination meeting the experimental data constraints is found, the resulting Boolean model can be used for making predictions on, for example, the steady state change due to perturbations, which can be further validated by experimental data.

3.2 | The rate equations approach

3.2.1 | Michaelis–Menten formalism

The standard equation for change of concentration in an enzymatic reaction in general and then also for a TF is Michaelis–Menten equation:

$$\frac{d[T]}{dt} = P_T - \frac{[T]}{\tau_T},$$

where $[T]$ is the TF concentration, P_T is the production function and τ_T is the TF half-life. If T is an activator one has

$$\frac{d[T]}{dt} = \alpha \frac{[T]^h}{k^h + [T]^h} - \frac{[T]}{\tau_T}$$

and correspondingly if T is a repressor

$$\frac{d[T]}{dt} = \alpha \frac{k^h}{k^h + [T]^h} - \frac{[T]}{\tau_T},$$

where k is the dissociation constant derived from the law of mass action with the Hill function exponent h that governs the degree of cooperativity and α is the proportionality constant that will be omitted in what follows. For our example, A activates A while B represses A . We have two TFs affecting gene A and their actions, as in the Boolean approach could follow various logical combinations. Assuming that A and B follow an AND logic on their action on A , the equation for A gene expression is written as:

$$\frac{d[A]}{dt} = \frac{[A]^h}{k_A^h + [A]^h} \cdot \frac{k_B^h}{k_B^h + [B]^h} - \frac{[A]}{\tau_A}.$$

If the two TFs activate A and respectively repress A following an OR logic then the equation for A reads as:

$$\frac{d[A]}{dt} = \frac{[A]^h}{k_A^h + [A]^h} + \frac{k_B^h}{k_B^h + [B]^h} - \frac{[A]}{\tau_A}.$$

3.2.2 | Shea–Ackers formalism

In this formalism, transcription is modeled from a statistical mechanics view (Shea & Ackers, 1985; Ackers et al., 1982). Depending on presence or absence of a TF and/or an RNAP, the operator can be in different states s . Each state has a statistical weight $Z(s)$.

The total statistical weight (partition sum) is:

$$Z = \sum_s Z(s) = Z(\text{on}) + Z(\text{off}).$$

The probability of a state being *on* is given by appropriately normalized ratio:

$$P(\text{on}) = \frac{Z(\text{on})}{Z}.$$

For our example, the total statistical weight becomes after some algebra:

$$Z = 1 + k_A[A] + k_B[B],$$

where all k are dissociation constants. The equation describing evolution of expression of A is:

$$\frac{d[A]}{dt} = \alpha \frac{k_A[A]}{1 + k_A[A] + k_B[B]} - \frac{[A]}{\tau_A}.$$

As can be seen, these equations are quite similar to the Michaelis–Menten equations with approximate sigmoidal behavior. In fact, still another option is to formally use sigmoids (Olariu et al., 2017).

Rate equations model simulations are conducted by solving the equations. Most of the time, the solutions are computed using numerical solvers. For simpler models, analytical solutions can also be found. The results consist of time series, which can be compared to experimental time series data for optimizing model parameters. With optimal sets of parameters the model can be further used for: simulating perturbation scenarios, bifurcation analysis, sensitivity analysis, etc.

3.3 | Stochastic simulation approach

The stochastic approach takes into consideration when the next molecular reaction will occur and what type of reaction will take place. It should be noted that in this formalism the variables are number of molecules (A , B) as opposed to concentrations in rate equations approach. Obviously, with many molecular copies the stochastic simulations will reproduce the concentrations. The fundamental premise is that a reaction μ will occur in $[t, t + dt]$ interval given $A(t)$ and $B(t)$, with probability (Gillespie, 1977):

$$\frac{a_\mu(A, B)}{a_0(A, B)} dt = f_\mu c_\mu dt,$$

where

$$a_0(A, B) = \sum_\mu a_\mu(A, B) \tau,$$

where τ is the time increment, c is the constant, f is the combinatorial function of A and B given by one of the model implementations shown in the previous section 3.2.2. For example, in Figure 1 there are four possible reactions ($\mu = 1, \dots, 4$), that is, A is produced with probability a_1/a_0 , A is decayed, B is produced and B is decayed.

The implementation steps for the stochastic simulations of the model for our example are:

1. Initialize $t = t_0$ and $A = A_0$, $B = B_0$.
2. Evaluate a_μ and their sum a_0 given the system is in state A_t , B_t at time t .
3. Generate τ and μ using $\tau = \frac{1}{a_0} \ln\left(\frac{1}{r_1}\right)$, $\mu = \text{smallest } \sum_{i=1}^\mu a_i > r_2 a_0$ where r_1 and r_2 are two samples from $U(0,1)$.
4. Update the system $t = t + \tau$, $A = A + A_\mu$, $B = B + B_\mu$.
5. Record A , B and t . Return to step 2, or else end simulation.

3.4 | When to use Boolean and continuous models respectively?

The answer to this question depends upon the focus on the study and what kind of data are available. There are three major scenarios:

1. One has no access to gene expression time but only arrows in architectures from, for example, binding and under-/over-expression experiments: In this case a preliminary study with a Boolean approach could be productive by exploring different rules for the vertices to map out the attractors. The approach is of particular value for large networks and valuable when it comes finding attractors representing cellular states. Local attractors or complex periodic behaviors could correspond to intermediate cells states trapping the cells during differentiation or reprogramming. Preferably one should use asynchronous updating. With random updating one provides noise to the system and ensembles of such different time courses form continuous time series.
2. With only access to architectures, one can also encode network dynamics using biochemical rate equations allowing for both additive and multiplicative interactions. Investigate attractor structures from iterating rate equations with starting points either from known properties or from point 1 above. With this tool one can also investigate bistability properties. This is suitable for smaller core modules where one suspects that key decisions take place.
3. With access to expression time series, biochemical rate equations are preferable. One fits the dynamical equation parameters to the data, which can be used to discriminate between different architecture hypotheses. Again, bistability analysis is then natural to perform.
4. When statistical experimental data is available, for example, only a percent of cell progenitors commit to a differentiated cell state, one has to conduct a stochastic implementation of the continuous-valued model. Statistics can be obtained from multiple stochastic simulations which should be compared to the statistical experimental data.

How do the number of free parameters compare for the two approaches? In general, the rate equation approach can have more free parameters than the Boolean approach. For very connected networks the existence of “hub” nodes lead to very large number of possible logical combinations. The Boolean approach is mostly affected by these drawbacks as the model training is based solely on finding these logical combinations based on constraints from gene expression experimental data. The rate equation approach can also suffer from this limitation. However, in contrast to Boolean rules rate equation parameters correspond to measurable quantities as binding strengths, decay rates, etc. Hence, these can provide further constraints which facilitate the identification of the best logical combinations of the inputs for each node in the network.

4 | APPLICATIONS—HAS THE APPROACH DELIVERED?

We next briefly describe some applications, where kinetic models, continuous and Boolean, respectively, have delivered in terms of providing new knowledge and predictions being tested.

4.1 | Continuous-valued models for hematopoiesis

Early works to use continuous-valued models for characterizing and quantifying stem cell differentiation processes in terms of switch behavior include Roeder & Glauche (2006) for hematopoiesis and Chickarmane, Troein, Nuber, Sauro, & Peterson (2006) for embryonic stem cells. The former was motivated by experimental observations on the TFs GATA-1 and PU.1, both known to act as key regulators and potential antagonists in the erythroid versus myeloid differentiation processes of hematopoietic progenitor cells as in Figure 1. In Huang, Guo, May, & Enver (2007), a simple model, where GATA-1 and PU.1 are antagonizing each other, is shown to account for how increasing and decreasing the two factors govern the differentiation into the erythroid and myeloid paths. Subsequently, a related model was designed and explored (Chickarmane, Enver, & Peterson, 2009) that not only host a desired Gata1–Pu.1 switch bistability behavior but is also able to keep the progenitor in a “holding pattern.” The latter, which is often denoted as “priming,” has been suggested as a means to keep several options open initially in a symmetrical fashion (see Hu et al., 1997 and references therein). With low molecular numbers for key genes, a stochastic process then moves the system into one of the available directions. Another early computational work on hematopoietic priming concerns the macrophage and neutrophil lineages where the model accounts for the desired bistable and graded behavior and accounts for the observed onset and mixture of differentiation patterns (Laslo et al., 2006). Another study proposed that the Scl–Gata2–Flt1 triad is a network module essential for the development of hematopoietic stem cells, exhibiting robust bistable behavior and irreversible switch between the states under external signals (Narula, Smith, Gottgens, & Igoshin, 2010). So far, the dynamical model results were based upon established gene modules and not fitting parameters to time series data. Rather, the conclusions were drawn from parameter insensitivity with regard to the questions probed.

In May et al. (2013), the Gata1–Pu.1–Gata2 system was subject to time series studies both with regard to gene expression and binding strengths. A putative model was put forward where the known interactions, excitatory and inhibitory, from perturbation experiments were fixed and the 32 different combinations (see Figure 3a) for the remaining ones

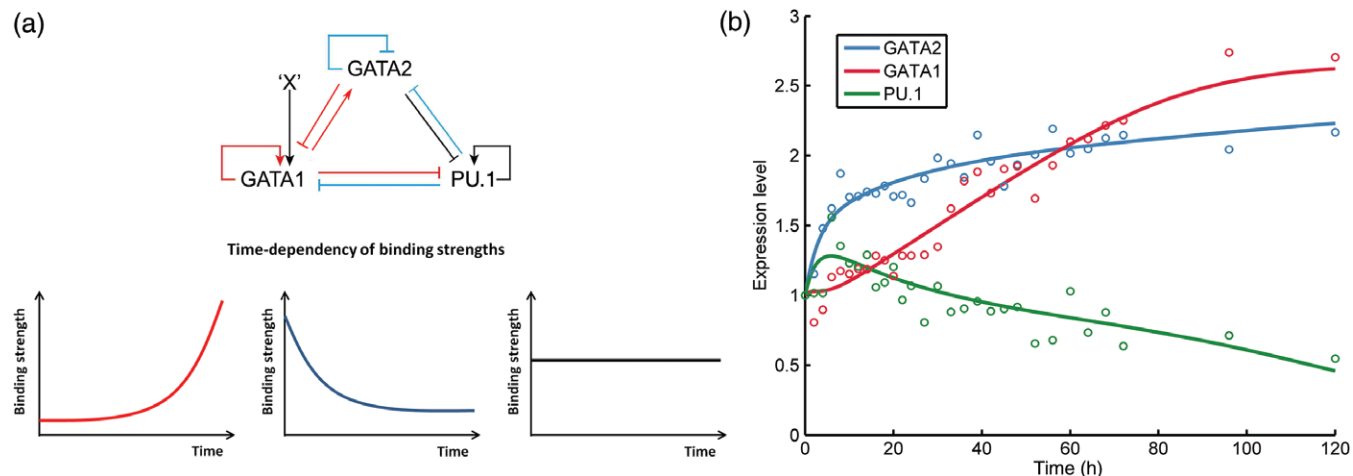


FIGURE 3 The triad interaction circuit depicting the most likely architecture from fitting to gene expression data. (a) The core network for the mutual and self-regulatory interactions between GATA1, PU.1, and GATA2. Red interaction indicates increase in time of the binding strength, blue corresponds to decreasing, and black shows constant binding strengths (as shown in the panel below the network diagram). The external signal Epo is represented in gray. (b) Expression time series of GATA1 (red), PU.1 (green), and GATA2 (blue) from the network model (solid lines) with the best set of estimated parameters. Experimental data are represented by circles. Expression levels are normalized by expression at $t = 0$ (multipotent state). (Reprinted with permission from May et al. (2013). Copyright 2013 Cell Press)

are fitted to the time series. The “winning” solution and its dynamics strongly hint at the Gata2–Pu.1 axis being the main driver for the decision in contrast to the old Gata1–Pu.1 paradigm. This is a prime example of dynamical modeling together with optimization yielding results beyond bistability properties and the like. The Gata1–Pu.1–Gata2 architecture of May et al. (2013) was further probed in Tian & Smith-Miles (2014) with both deterministic and stochastic approaches, where modeling is further constrained by measured degradation constants and disassociation rates for the three key players. Furthermore, a multiple objective method is designed for parameter optimization taking into account stability criteria and robustness. Tri-stability is established consistent with what was suggested in Chickarmane et al. (2009) but with no reference to external specific signals required to retain a primed state. The stochastic implementation of the model predicts, not yet verified, that the synthesis rate of GATA-1 during the decreasing process of GATA-2 determines the probability of the erythroid–myeloid transition.

Along similar lines as in May et al. (2013), the early stages of T-cell commitment were subject to dynamical modeling in Manesso, Kueh, Freedman, Rothenberg, & Peterson (2016), where the key genes Tcf1, Gata3, Pu.1 and Bcl11b, and known interactions responsible for the commitment transition were considered. Again, fitting 32 different model options to time series data one finds that among solutions surviving 95% CI as quality criteria of the fits together with giving rise to reasonable decay rates, one is left with four solutions that all host irreversibility. Furthermore, these solutions constrain what kind of logical operations take place on the Bcl11b locus. Also, these good fits favor a Gata3 dimeric binding on the Bcl11b locus. These results represent real predictions from kinetic modeling.

The B-cell branch of immune system has also been subject to dynamical modeling often with focus on irreversibility properties. In Martinez et al. (2012), a kinetic model of three key gene regulators, BCL6, IRF4, and BLIMP is presented with parameters set by gene expression profiles for the B-cell exit. Two different bistability behaviors that direct B-cells are identified and in silico perturbations elucidate lymphoma-genesis mechanisms and suggest candidate tumorigenic alterations. In Méndez & Mendoza (2016), the dynamical behavior of terminal B-cell differentiation is analyzed under normal and mutant conditions with dynamical models. These recover the B-cell differentiation patterns for a large set of gain- and loss-of-function mutants. Another modeling application for B-cell termination differentiation was pursued in Zhang et al. (2010) with a core involving the repressors Bcl-6, Blimp-1, and Pax5. Even though individual B-cells switch to the plasma cell state in an all-or-none fashion when stimulated appropriately, stochastic fluctuations in gene expression make the switching event probabilistic, leading to heterogeneous differentiation response among individual B-cells. Moreover, stochastic gene expression renders the dose–response behavior of a population of B-cells substantially graded, a result that is consistent with experimental observations. The early B-cell multipotent progenitor transition was modeled in Salerno, Cosentino, Morrone, & Amato (2015). With multi-stability analysis methods, the authors were able to assess the model capability of capturing the experimentally observed switch-like commitment behavior and then confirm the central role of zinc finger protein 521 (ZNF521). Furthermore, a novel putative functional interaction for ZNF521 was identified, which turn out to be essential to realize such characteristic behavior.

An interesting application of clinical relevance is presented in Anguita et al. (2016), where the dynamics of the two alleles of Gfl1b are treated individually to account for the fact that one of them is mutated in acute myeloid leukemia (AML) inhibiting the wild type allele function which leads to abnormal high levels of Pu.1 and Gfl1b. These results demonstrate the Pu.1-Gfl1b transcriptional network to be an important regulatory axis in AML, which is confirmed experimentally.

As mentioned above, a complementary and worth exploring view of gene dynamics can under certain approximations be depicted as movements in an energy landscape. For example, in Olariu et al. (2017), the Gata1–Pu.1–Gata2 system is revisited in this aspect. Its energy landscape is shown in Figure 4. With this landscape information available, one can compute the shortest distance between the two attractors, Gata1/Gata2 ON and Pu.1 OFF, and vice versa. In other words, it offers guiding on how to efficiently go from one attractor to another by varying external factors. This is of high relevance when it comes to reprogramming as was demonstrated (Olariu et al., 2017) in the case of reprogramming to pluripotent state.

As single cell data is becoming increasingly available, one has the opportunity to unravel the stochastic mechanisms likely to be present at the point of lineage commitment. Indeed, by analyzing single cell expression data taken before and after commitment as well as cell growth data (Pina et al., 2012) a Gillespie Monte Carlo time series stochastic model was developed for the hematopoietic triad Gata1–Gata2–Mpo (Teles et al., 2013), where the latter is a proxy for Pu.1. Here, the bursting parameters governing promoter status within the so-called Random Telegraph Model first employed in single-cell gene expression in Mettetal, Muzzey, Pedraza, Ozbudak, & van Oudenaarden (2006) were fitted to experimental static gene expression distributions (see Figure 5). The same holds for the mRNA production and decay

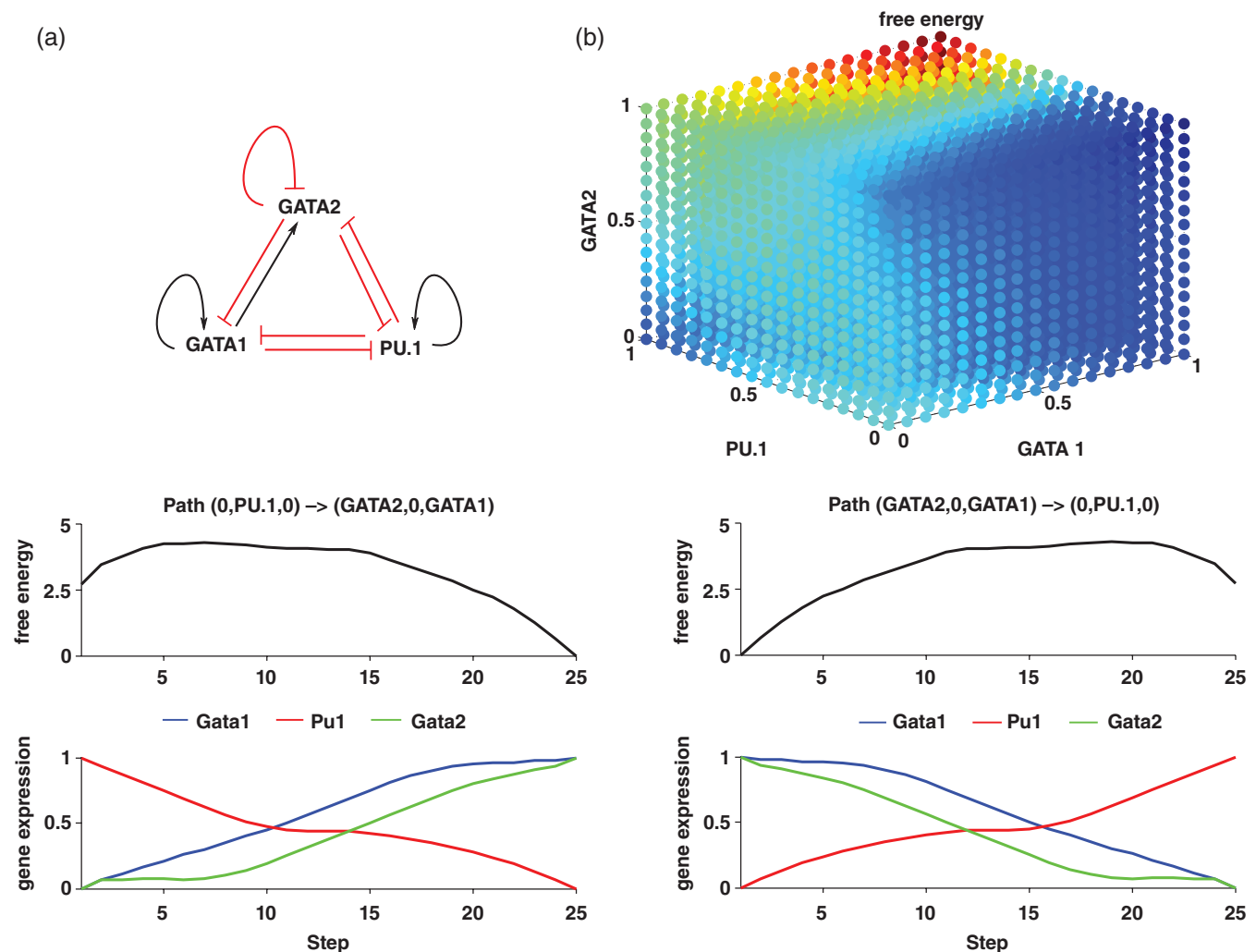


FIGURE 4 The GATA2–GATA1–PU.1 network and the corresponding free energy and the shortest paths between the attractors. (a) The GATA2–GATA1–PU.1 transcription factor (TF) circuit with the well-established GATA1–PU.1 mutual inhibition and positive auto-regulations, GATA1 repressing GATA2 while GATA2 induces GATA1. The circuit includes the interactions of GATA2–PU.1 mutual inhibition and GATA2 negative self-interaction determined by combinatorial searches using rate equations (May et al., 2013). (b) The free energy exhibiting two attractors: (1) GATA2, GATA1-high and PU.1-low; and (2) PU.1-high and GATA2, GATA1-low. (c) Variation of the free energy and gene expressions along the shortest path between (2) and (1). (d) The corresponding variation along the shortest path between (1) and (2). (Reprinted with permission from Olariu et al. (2017). Copyright 2017 Royal Society Publishing)

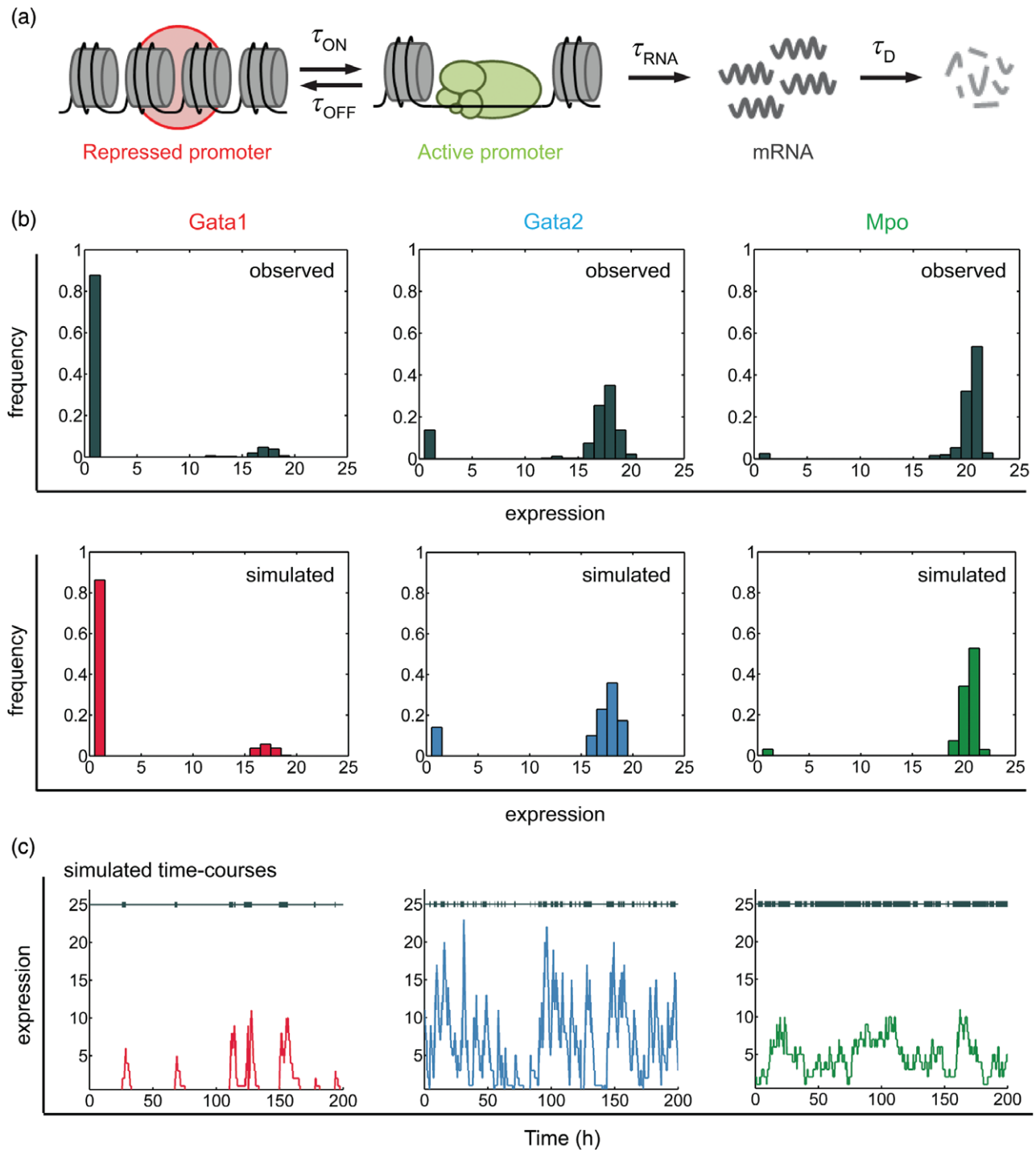


FIGURE 5 (a) Schematic representation of the transcriptional bursting model. For a given gene, the promoter can be in two different states, active or repressed, with the average time spent in each state being controlled by the average times for activation (τ_{ON}) and repression (τ_{OFF}). When in the active promoter state, the gene is transcribed and produces mRNA molecules after an average production time. Finally, mRNA molecules are degraded after an average time, τ_{RNA} , irrespective of promoter states. (b) Best parameter sets for each gene allow for the reconstitution of the experimentally observed distributions (top) within our model simulations (bottom). (c) The parameters suggest different modes of stochastic expression for the different genes, with highly variable burst frequencies and duration (gray bars) as well as mRNA dynamics (colored lines). (Reprinted with permission from Teles et al. (2013). Copyright 2013 PLOS Journals)

rates. Probability of commitment in time is a function of gene expression as defined by a logistic regression model obtained from the same experimental single-cell expression data. The Monte Carlo time scale is converted to physical time using cell culture kinetic data. Robust model solutions are identified with tested predictions. A pipeline procedure around this approach has been developed excluding the commitment probability part (Ezer, Moignard, Göttgens, & Adryan, 2016).

4.2 | Boolean models for hematopoiesis

The procedures and goals with this approach are similar to the ones using continuous-valued models above. Early work here within hematopoiesis was done in Krumsiek, Marr, Schroeder, & Theis (2011) and later followed by Bonzanni et al. (2013), Hamey et al. (2017) and Collombet et al. (2017). These studies were based upon static data—no measured expression time series. Given putative architectures one initially assigns logical rules to interactions from perturbation experiment results and literature searches as far as existing information allows. If all rules are not known from these sources, which is often the case, one needs to explore different rule combinations. For simplicity, one often confines oneself to combinations of the common ones, for example, AND, OR, and NOT. With the requirement that known initial conditions should give rise to desired final states being represented as stable attractors when updating the states, one aims for a solution in terms of a specific set of Boolean rules. In other words, the Boolean updating equations are iterated until convergence for the different rule options is achieved. The winning rule set is then determined by the ability to reproduce the desired final attractors with given initial conditions. This “learning” step corresponds to finding parameters in the continuous case (see May et al., 2013). Similar procedures were followed in the embryonic stem cell case (Dunn, Martello, Yordanov, Emmott, & Smith, 2014). The same group proposed a method based on automated formal reasoning which allows analysis of complete set of logical models consistent with experimental data from myeloid progenitor differentiation among other biological systems applications (Yordanov et al., 2016). With a limited number of available rules, one in general has less number of “parameters” to deal with than in the continuous model case. It is clear that for exploring larger networks, this Boolean approach provides a good starting point. Noise can be imposed upon the dynamics by randomly choosing the updating order. By averaging over different such runs, one obtains continuous time series and can study bifurcations.

In Figure 6 from Collombet et al. (2017), the Boolean state space for progenitor differentiation into B-cells and macrophages are shown as an illustrative example. Here, the initial unstable state and the two final states are the rigid conditions that the putative Boolean models need to satisfy. The inputs to this work were extensive literature studies together with complementing experiments. Through the computational model, data inconsistencies are identified and features detected of relevance for reprogramming.

The increase of single-cell studies in the hematopoietic system, reviewed in Cvejic (2015) and Povinelli, Rodriguez-Meira, & Mead (2018) led to development of computational and statistical approaches, including Boolean models, which

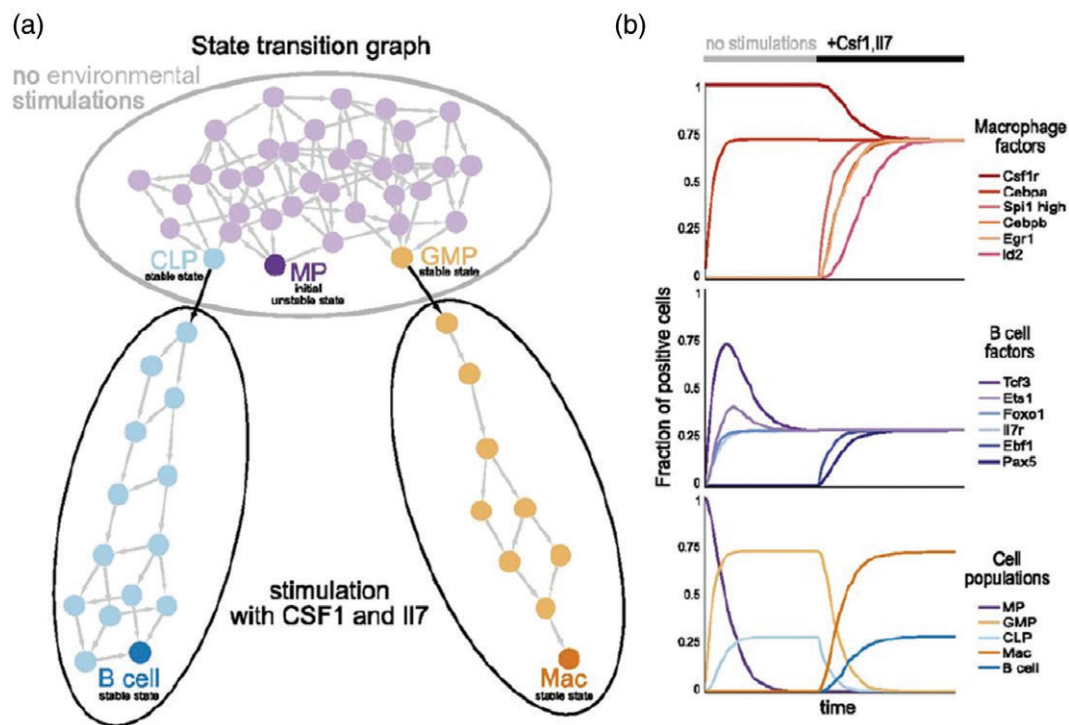


FIGURE 6 (a) Boolean model states for a progenitor (MP) differentiating to B-cells and macrophages, respectively. Model simulations starts from the unstable MP state in the absence of cytokine (upper) and after the addition of CSF1 and IL7 (lower left and lower right). The trajectory from MP to B-cells goes through 9 and 19 states in the two compartments. Similarly, one has 8 and 36 states to pass for the macrophages. (b) Averaged stochastic simulations from random update ordering showing the evolution over time, before and after cytokine exposition, of the fractions of cells expressing specific macrophage factors (top), B-cell factors (middle), and cell-type signatures (bottom). The x and y axes represent time (in arbitrary units) and fractions of positive cells, respectively. (Reprinted with permission from Collombet et al. (2017). Copyright 2017 National Academy of Sciences)

contribute to gaining a better understanding of mechanisms governing lineage commitment and cell fate decisions (Pina et al., 2012; Moignard et al., 2015; Trapnell et al., 2014). When only static measurements are available, one can nevertheless construct pseudo-time trajectories in the case of single cell expression measurements as was done in Hamey et al. (2017). First, one creates diffusion maps (Haghverdi, Buettner, & Theis, 2015; Setty et al., 2016) that follow the trajectories to different cell fates. In the next step one infers a pseudo-time ordering based upon similarities between individual cell expressions (Trapnell et al., 2014)—adjacency is established. In Hamey et al. (2017), networks are reconstructed from correlation measures along the pseudo-time trajectories, which are subsequently subject to Boolean modeling. Predictions for Gata2 regulation of Cbfa2t3h and Nfe2 in two different differentiation trajectories are subsequently verified with chip-seq techniques.

PBN has been inferred from time series data (Martin, Zhang, Martino, & Faulon, 2007) and steady state data (Shmulevich et al., 2002) estimating perturbations and switching probabilities for the gene regulatory networks. In Martin et al. (2007), 901 Boolean networks were extracted from T-cell time series data sets exhibiting three attractors corresponding to states dynamics which were experimentally demonstrated and inferring a gene regulatory network governing the T-cell immune response. Stochastic Boolean Network approach was also shown to be able to recover biologically proven dynamic attractors in the predicted T-cell immune response network (Liang & Han, 2012) and to discover network dynamics when the genes are under perturbation.

5 | CONCLUSIONS

Transcriptional and cell interaction data are rapidly becoming more extensive and accurate, thereby pushing the fields towards a quantitative science like physics. Hence, there is an upsurge in developing theoretical models. In parallel, the access to high performance CPU power is exploding thereby facilitating model simulations. Whole genome and high throughput transcription and binding data based upon different perturbation conditions at different time points are still far from providing entire network reconstructions. Rather one has to focus on finite network modules, where “good-old” down- and up-regulation experiments still play an important role for reconstruction. For developmental processes like those occurring in stem cell fate decisions, relatively few genes are expected to be important. Hence, modular approaches make sense.

Whereas transcription data were initially dominated by “bulk” measurements averaged over many cells, single cell measurements are now becoming more and more standard. This development, which poses new challenges for analysis and modeling, has emphasized the heterogeneity among cells in general and in cells being close to evolution decision points this is very much the case. The latter is often assigned to few molecular copies that are determinants for the transition.

With regard to single cell measurements, real-time imaging techniques are about to revolutionize the field since the cell commitment processes can now be traced in time with appropriate coloring techniques enabling us to see how cells proliferate and key genes develop. Furthermore, single cell gene expression measurements for given surface markers, enable the reconstruction of pseudo-time series.

The computational techniques employed in the examples above for transcription are mature and powerful. The ultimate success is critically dependent upon tight interfaces with the experimental activities within iterative procedure. This is facilitated by the emergence of a new generation researchers with cross-disciplinary training.

Future challenges also include merging transcriptional with cell dynamics—cell division, cell death, and cell–cell interactions. Initial steps in this direction can be found in Roeder & Loeffler (2002) dealing with hematopoietic progenitor colonies. For the case of the mammalian embryo in early stages, informative work has been pursued in this multi-scale direction including mechanical forces (Krupinski, Chickarmane, & Peterson, 2011; Krupinski, Chickarmane, & Peterson, 2012). In another study, the population-based approach was integrated with a single-cell stochastic toggle switch model to study the possibility of inferring mechanistic knowledge of the differentiation of granulocyte–monocyte progenitors to granulocytes or monocytes (Marr, Strasser, Schwarzfischer, Schroeder, & Theis, 2012). A multi-scale model was also proposed for erythropoiesis where a single-cell gene regulatory network governing self-renewal, differentiation, and apoptosis was integrated into a population model which simulated anemia situations (Demin, Crauste, Gandrillon, & Volpert, 2010). Some challenges still remain for transcriptional modeling. For example, in situations where some genes have low copy numbers and others not; one needs to fuse stochastic with deterministic procedures (Puchałka & Kierzek, 2004). Also, one needs to model epigenetic interactions in more detail. So far, these have mostly been accounted for by treating them as effective interactions contained in transcriptional formalisms, see for example, Chickarmane, Olariu, & Peterson (2012) for an embryonic application. One exception, where one goes into more detail (Olariu et al., 2016) incorporating DNA methylation modifications that takes into account the multi-layer nature of regulatory mechanisms governing pluripotency acquisition through reprogramming.

ACKNOWLEDGMENTS

This work was supported by the Swedish Research Council (VR 621-2013-4547) and the National Institutes of Health (18B-1095882).

CONFLICT OF INTEREST

The authors have declared no conflicts of interest for this article.

RELATED WIREs ARTICLES

[Experimentally based sea urchin gene regulatory network and the causal explanation of developmental phenomenology](#)
[Establishing the stem cell state: insights from regulatory network analysis of blood stem cell development](#)

REFERENCES

- Ackers, G. K., Johnson, A. D., & Shea, M. A. (1982). Quantitative model for gene regulation by lambda phage repressor. *Proceedings of the National Academy of Science*, *79*, 1129–1133.
- Anguita, E., Gupta, R., Olariu, V., Valk, P. J., Peterson, C., Delwel, R., & Enver, T. (2016). A somatic mutation of GF11B identified in leukemia alters cell fate via a SPI1 (PU.1.) centered genetic regulatory network. *Developmental Biology*, *411*(2), 277–286.
- Bhattacharya, S., Zhang, Q., & Andersen, M. E. (2011). A deterministic map of Waddington's epigenetic landscape for cell fate specification. *BMC Systems Biology*, *5*(1), 85.
- Bonzanni, N., Garg, A., Feenstra, K. A., Schütte, J., Kinston, S., Miranda-Saavedra, D., ... Göttgens, B. (2013). Hard-wired heterogeneity in blood stem cells revealed using a dynamic regulatory network model. *Bioinformatics*, *29*, i80–i88.
- Buchler, N. E., Gerland, U., & Hwa, T. (2003). On schemes of combinatorial transcription logic. *Proceedings of the National Academy of Sciences*, *100*, 5136–5141.
- Chen, K. C., Csikasz-Nagy, A., Györfy, B., Val, J., Novak, B., & Tyson, J. J. (2000). Kinetic analysis of a molecular model of the budding yeast cell cycle. *Molecular Biology of the Cell*, *11*, 369–391.
- Chickarmane, V., Enver, T., & Peterson, C. (2009). Computational modeling of the hematopoietic erythroid–myeloid switch reveals insights into cooperativity, priming, and irreversibility. *PLoS Computational Biology*, *5*(1), e1000268.
- Chickarmane, V., Olariu, V., & Peterson, C. (2012). Probing the role of stochasticity in a model of the embryonic stem cell—Heterogeneous gene expression and reprogramming efficiency. *BMC Systems Biology*, *6*(1), 98.
- Chickarmane, V., Troein, C., Nuber, U. A., Sauro, H. M., & Peterson, C. (2006). Transcriptional dynamics of the embryonic stem cell switch. *PLoS Computational Biology*, *2*(9), e123.
- Collins, J. J., Gardner, T. S., & Cantor, C. R. (2000). Construction of a genetic toggle switch in *Escherichia coli*. *Nature*, *403*, 339–342.
- Collombet, S., van Oevelen, C., Ortega, J. L. S., Abou-Jaoudé, W., Stefano, B. D., Thomas-Chollier, M., ... Thieffry, D. (2017). Logical modeling of lymphoid and myeloid cell specification and transdifferentiation. *Proceedings of the National Academy of Sciences*, *114*, 5792–5799.
- Cvejic, A. (2015). Mechanisms of fate decision and lineage commitment during haematopoiesis. *Immunology and Cell Biology*, *94*, 230.
- Demin, I., Crauste, F., Gandrillon, O., & Volpert, V. (2010). A multi-scale model of erythropoiesis. *Journal of Biological Dynamics*, *4*, 59–70.
- Dunn, S.-J., Martello, G., Yordanov, B., Emmott, S., & Smith, A. G. (2014). Defining an essential transcription factor program for naive pluripotency. *Science*, *344*, 1156–1160.
- Ezer, D., Moignard, V., Göttgens, B., & Adryan, B. (2016). Determining physical mechanisms of gene expression regulation from single cell gene expression data. *PLoS Computational Biology*, *12*, e1005072.
- Faure, A., Naldi, A., Chaouiya, C., & Thieffry, D. (2006). Dynamical analysis of a generic Boolean model for the control of the mammalian cell cycle. *Bioinformatics*, *22*, e124–e131.
- Ferrell, J. E. (2012). Bistability, bifurcations, and Waddington's epigenetic landscape. *Current Biology*, *22*, R458–R466.
- Gillespie, D. T. (1977). Exact stochastic simulation of coupled chemical reactions. *Journal of Physical Chemistry*, *81*(25), 2340–2361.
- Glass, L., & Kauffman, S. A. (1973). The logical analysis of continuous, non-linear biochemical control networks. *Journal of Theoretical Biology*, *39*, 103–129.
- Greil, F., & Drossel, B. (2005). Dynamics of critical Kauffman networks under asynchronous stochastic update. *Physical Review Letters*, *95*, 048701.
- Grondin, R. O., Porod, W., Loeffler, C. M., & Ferry, D. K. (1983). Synchronous and asynchronous systems of threshold elements. *Biological Cybernetics*, *49*, 1–7.
- Haghverdi, L., Buettner, F., & Theis, F. J. (2015). Diffusion maps for high-dimensional single-cell analysis of differentiation data. *Bioinformatics*, *31*, 2989–2998.
- Hamey, F. K., Nestorowa, S., Kinston, S. J., Kent, D. G., Wilson, N. K., & Göttgens, B. (2017). Reconstructing blood stem cell regulatory network models from single-cell molecular profiles. *Proceedings of the National Academy of Sciences*, *114*, 5822–5829.
- Herberg, M., & Roeder, I. (2015). Computational modelling of embryonic stem-cell fate control. *Development*, *142*, 2250–2260.
- Hu, M., Krause, D., Greaves, M., Sharkis, S., Dexter, M., Heyworth, C., & Enver, T. (1997). Multilineage gene expression precedes commitment in the hemopoietic system. *Genes & Development*, *11*, 774–785.
- Huang, S., Guo, Y.-P., May, G., & Enver, T. (2007). Bifurcation dynamics in lineage-commitment in bipotent progenitor cells. *Developmental Biology*, *305*, 695–713.
- Jacob, F., & Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology*, *3*, 318–356.
- Kauffman, S. (1969). Metabolic stability and epigenesis in randomly constructed genetic nets. *Journal of Theoretical Biology*, *22*, 437–467.
- Kauffman, S., Peterson, C., Samuelsson, B., & Troein, C. (2004). Genetic networks with canalizing boolean rules are always stable. *Proceedings of the National Academy of Sciences*, *101*(49), 17102–17107.
- Krumsiek, J., Marr, C., Schroeder, T., & Theis, F. J. (2011). Hierarchical differentiation of myeloid progenitors is encoded in the transcription factor network. *PLoS One*, *6*, e22649.
- Krupinski, P., Chickarmane, V., & Peterson, C. (2011). Simulating the mammalian blastocyst—molecular and mechanical interactions pattern the embryo. *PLoS Computational Biology*, *7*(5), e1001128.
- Krupinski, P., Chickarmane, V., & Peterson, C. (2012). Computational multiscale modeling of embryo development. *Current Opinion in Genetics & Development*, *22*, 613–618.
- Laslo, P., Spooner, C. J., Warmflash, A., Lancki, D. W., Lee, H.-J., Sciammas, R., ... Singh, H. (2006). Multilineage transcriptional priming and determination of alternate hematopoietic cell fates. *Cell*, *126*, 755–766.

- Liang, J., & Han, J. (2012). Stochastic Boolean networks: An efficient approach to modeling gene regulatory networks. *BMC Systems Biology*, 6, 113.
- MacLean, A. L., Celso, C. L., & Stumpf, M. P. (2016). Concise review: Stem cell population biology—Insights from hematopoiesis. *Stem Cells*, 35, 80–88.
- Manesso, E., Kueh, H. Y., Freedman, G., Rothenberg, E. V., & Peterson, C. (2016). Irreversibility of T-cell specification: Insights from computational modelling of a minimal network architecture. *PLoS One*, 11, e0161260.
- Marr, C., Strasser, M., Schwarzfischer, M., Schroeder, T., & Theis, F. J. (2012). Multi-scale modeling of GMP differentiation based on single-cell genealogies. *FEBS Journal*, 279, 3488–3500.
- Martin, S., Zhang, Z., Martino, A., & Faulon, J.-L. (2007). Boolean dynamics of genetic regulatory networks inferred from microarray time series data. *Bioinformatics*, 23(7), 866–874.
- Martinez, M. R., Corradin, A., Klein, U., Alvarez, M. J., Toffolo, G. M., di Camillo, B., ... Stolovitzky, G. A. (2012). Quantitative modeling of the terminal differentiation of B cells and mechanisms of lymphomagenesis. *Proceedings of the National Academy of Sciences*, 109, 2672–2677.
- May, G., Soneji, S., Tipping, A. J., Teles, J., McGowan, S. J., Wu, M., ... Enver, T. (2013). Dynamic analysis of gene expression and genome-wide transcription factor binding during lineage specification of multipotent progenitors. *Cell Stem Cell*, 13, 754–768.
- Méndez, A., & Mendoza, L. (2016). A network model to describe the terminal differentiation of B cells. *PLoS Computational Biology*, 12, e1004696.
- Mendoza, L., Thieffry, D., & Alvarez-Buylla, E. R. (1999). Genetic control of flower morphogenesis in *Arabidopsis thaliana*: A logical analysis. *Bioinformatics*, 15, 593–606.
- Mettetal, J. T., Muzzey, D., Pedraza, J. M., Ozbudak, E. M., & van Oudenaarden, A. (2006). Predicting stochastic gene expression dynamics in single cells. *Proceedings of the National Academy of Sciences*, 103, 7304–7309.
- Moignard, V., Woodhouse, S., Haghverdi, L., Lilly, A. J., Tanaka, Y., Wilkinson, A. C., ... Göttgens, B. (2015). Decoding the regulatory network of early blood development from single-cell gene expression measurements. *Nature Biotechnology*, 33, 269–276.
- Narula, J., Smith, A. M., Gottgens, B., & Igoshin, O. A. (2010). Modeling reveals bistability and low-pass filtering in the network module determining blood stem cell fate. *PLoS Computational Biology*, 6, e1000771.
- Olariu, V., Lövkvist, C., & Sneppen, K. (2016). Nanog, oct4 and tet1 interplay in establishing pluripotency. *Scientific Reports*, 6, 25438.
- Olariu, V., Manesso, E., & Peterson, C. (2017). A deterministic method for estimating free energy genetic network landscapes with applications to cell commitment and reprogramming paths. *Royal Society Open Science*, 4, 160765.
- Papatsenko, D., & Lemischka, I. R. (2016). Emerging modeling concepts and solutions in stem cell research. *Current Topics in Developmental Biology*, 116, 709–721.
- Peter, I. S., Faure, E., & Davidson, E. H. (2012). Predictive computation of genomic logic processing functions in embryonic development. *Proceedings of the National Academy of Sciences*, 109, 16434–16442.
- Pina, C., Fugazza, C., Tipping, A. J., Brown, J., Soneji, S., Teles, J., ... Enver, T. (2012). Inferring rules of lineage commitment in haematopoiesis. *Nature Cell Biology*, 14(3), 287–294.
- Povinelli, B. J., Rodrigue-Meira, A., & Mead, A. J. (2018). Single cell analysis of normal and leukemic hematopoiesis. *Molecular Aspects of Medicine*, 59, 85–94.
- Puchałka, J., & Kierzek, A. M. (2004). Bridging the gap between stochastic and deterministic regimes in the kinetic simulations of the biochemical reaction networks. *Biophysical Journal*, 86, 1357–1372.
- Reinitz, J., Mjølness, E., & Sharp, D. H. (1995). Model for cooperative control of positional information in *Drosophila* by bicoid and maternal hunchback. *Journal of Experimental Zoology*, 271, 47–56.
- Roeder, I., & Glauche, I. (2006). Towards an understanding of lineage specification in hematopoietic stem cells: A mathematical model for the interaction of transcription factors GATA-1 and PU.1. *Journal of Theoretical Biology*, 241, 852–865.
- Roeder, I., & Loeffler, M. (2002). A novel dynamic model of hematopoietic stem cell organization based on the concept of within-tissue plasticity. *Experimental Hematology*, 30, 853–861.
- Salerno, L., Cosentino, C., Morrone, G., & Amato, F. (2015). Computational modeling of a transcriptional switch underlying b-lymphocyte lineage commitment of hematopoietic multipotent cells. *PLoS One*, 10, e0132208.
- Setty, M., Tadmor, M. D., Reich-Zeliger, S., Angel, O., Salame, T. M., Kathail, P., ... Peer, D. (2016). Wishbone identifies bifurcating developmental trajectories from single-cell data. *Nature Biotechnology*, 34, 637–645.
- Shea, M. A., & Ackers, G. K. (1985). The *or* control system of bacteriophage lambda a physical-chemical model for gene regulation. *Journal of Molecular Biology*, 181, 211–230.
- Shmulevich, I., Dougherty, E. R., Kim, S., & Zhang, W. (2002). Probabilistic Boolean networks: A rule-based uncertainty model for gene regulatory networks. *Bioinformatics*, 18(2), 261–274.
- Teles, J., Pina, C., Edén, P., Ohlsson, M., Enver, T., & Peterson, C. (2013). Transcriptional regulation of lineage commitment—a stochastic model of cell fate decisions. *PLoS Computational Biology*, 9(8), e1003197.
- Thomas, R., & D'Ari, R. (1990). *Biological feedback*. Taylor & Francis: CRC Press, Inc.
- Tian, T., & Smith-Miles, K. (2014). Mathematical modeling of GATA-switching for regulating the differentiation of hematopoietic stem cell. *BMC Systems Biology*, 8(Suppl. 1), S8.
- Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., ... Rinn, J. L. (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nature Biotechnology*, 32, 381–386.
- Wang, J., Zhang, K., Xu, L., & Wang, E. (2011). Quantifying the Waddington landscape and biological paths for development and differentiation. *Proceedings of the National Academy of Sciences*, 108, 8257–8262.
- Whichard, Z. L., Sarkar, C. A., Kimmel, M., & Corey, S. J. (2010). Hematopoiesis and its disorders: A systems biology approach. *Blood*, 115, 2339–2347.
- Yordanov, B., Dunn, S.-J., Kugler, H., Smith, A., Martello, G., & Emmott, S. (2016). A method to identify and analyze biological programs through automated reasoning. *npj Systems Biology and Applications*, 2, 16010.
- Zhang, Q., Bhattacharya, S., Kline, D. E., Crawford, R. B., Conolly, R. B., Thomas, R. S., ... Andersen, M. E. (2010). Stochastic modeling of b lymphocyte terminal differentiation and its suppression by dioxin. *BMC Systems Biology*, 4(1), 40.
- Zhou, J. X., Aliyu, M. D. S., Aurell, E., & Huang, S. (2012). Quasi-potential landscape in complex multi-stable systems. *Journal of the Royal Society Interface*, 9, 3539–3553.