Available online at www.sciencedirect.com
ScienceDirect
Developmental Biology 305 (2007) 695–713
DEVELOPMENTAL BIOLOGY
www.elsevier.com/locate/ydbio

Genomes & Developmental Control

Bifurcation dynamics in lineage-commitment in bipotent progenitor cells

Sui Huang [a,*], Yan-Ping Guo [b], Gillian May [b], Tariq Enver [b]

[a] *Department of Surgery and Vascular Biology Program, Children's Hospital, Harvard Medical School, and Harvard Stem Cell Institute, Children's Hospital, 1 Blackfan Circle, Boston, MA 02115, USA*
[b] *MRC Molecular Haematology Unit, Weatherall Institute of Molecular Medicine, John Radcliffe Hospital, University of Oxford, UK*

Received for publication 21 August 2006; revised 2 February 2007; accepted 26 February 2007
Available online 3 March 2007

Abstract

Lineage specification of multipotent progenitor cells is governed by a balance of lineage-affiliated transcription factors, such as GATA1 and PU.1, which regulate the choice between erythroid and myelomonocytic fates. But how ratios of lineage-determining transcription factors stabilize progenitor cells and resolve their indeterminacy to commit them to discrete, mutually exclusive fates remains unexplained. We used a simple model and experimental measurements to analyze the dynamics of a binary fate decision governed by a gene-circuit containing auto-stimulation and cross-inhibition, as embodied by the GATA1–PU.1 paradigm. This circuit generates stable attractors corresponding to erythroid and myelomonocytic fates, as well as an uncommitted metastable state characterized by coexpression of both regulators, explaining the phenomenon of "multilineage priming". GATA1 and PU.1 mRNA and transcriptome dynamics of differentiating progenitor cells confirm that commitment occurs in two stages, as suggested by the model: first, the progenitor state is destabilized in an almost symmetrical bifurcation event, resulting in a poised state at the boundary between the two lineage-specific attractors; second, the cell is driven to the respective, now accessible attractors. This minimal model captures fundamental features of binary cell fate decisions, uniting the concepts of stochastic (selective) and deterministic (instructive) regulation, and hence, may apply to a wider range of binary fate decision points.

Keywords: Cell fate commitment; Attractor; Bifurcation; Multipotency; Progenitor cell; GATA1; PU.1; Auto-regulation; Stochastic regulation

Introduction

During cell lineage specification, uncommitted multipotent progenitor cells undergo a discrete cell fate decision, selecting between, and committing to, one of a finite number of predefined cell lineages. The latter represent discrete "genetic programs" that are mutually exclusive and intrinsically robust. In molecular terms, lineage commitment ultimately establishes the cell-type-specific gene expression profile (transcriptome) of the mature, differentiated cell. Transcription factors that act in intricate circuits of gene regulation (Davidson et al., 2002; Swiers et al., 2006) specify the stable, lineage-specific transcriptome. This molecular view of lineage commitment is broadly in agreement with the older systems biology concept that terminal cell fates represent stable attractors in gene network dynamics, first proposed by Delbruck, Jacob and Monod and others for small gene circuits, and by Kauffman for genome-wide networks more than 50 years ago (Kauffman, 1969; reviewed in Huang, 2005). More recently, mathematical models, stressing the importance of feedback control in gene regulation, further established the idea that stable attractors of gene circuits represent states of differentiation (Cinquin and Demongeot, 2005; Thomas, 1998). Experiments also suggest that discrete transitions in bistable biochemical systems underlie cell fate decisions or differentiation (Laslo et al., 2006; Xiong and Ferrell, 2003). The concept of attractors captures in formal terms the qualitative metaphors used to described a cell fate, such as an "energy sink" (Graf, 2002) or a "valley" separated by hills in an "epigenetic landscape" (Waddington, 1957) which naturally explain robustness and discreteness of cell fates.

However, there is not much conceptual understanding of the essence of multipotency, a hall mark of stem or progenitor cells, despite recent success in systematic cataloguing of transcripts from various stem cells (Joshi and Enver, 2003; Phillips et al., 2000). The attractor concept, while providing a plausible

* Corresponding author. Fax: +1 617 730 0230.
E-mail address: sui.huang@tch.harvard.edu (S. Huang).

0012-1606/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.ydbio.2007.02.036

explanation for stable cell types, does not address the dynamics of cell fate decision and determination during development. Multipotency may rather be dynamically defined (Mikkers and Frisen, 2005), in that it is associated with the simultaneous expression of lineage-specific genes: the "promiscuous expression" of genes specific for the alternative fates available to the as yet uncommitted cell (Cross and Enver, 1997). Such "multi-lineage priming" has been reported for various progenitor cells (Akashi et al., 2003; Hu et al., 1997; Kim et al., 2005; Miyamoto et al., 2002).

Another challenge in developmental biology is to understand how external regulatory signals control cell fate choice. Two alternative but not mutually exclusive concepts have been discussed (Enver et al., 1998): In the model of intrinsic or stochastic cell fate control, cell fates constitute preexisting programs that can be adopted by the cell in a chance fashion. In this case, external signals lack fate determining capacity and act simply as a survival/growth factor for cells which have already committed. Alternatively, in the extrinsic or deterministic model, external signals actually impose the genetic program to the cell by activating or repressing the "appropriate" set of genes via signal transduction cascades. In developmental biology the stochastic and deterministic models have also been referred to as "selective" or "instructive" regulation, respectively (Wolpert, 1994).

The differentiation of blood cells has been intensively studied as a paradigm for lineage specification, leading to the identification of (*i*) cellular intermediates that act as branch points in the developmental paths (Morrison et al., 1995) and (*ii*) a panel of key transcriptional regulators that control different decision points (Kluger et al., 2004; Orkin, 2000). Specifically, at the level of the common myeloid precursor cell (CMP) which makes the binary decision between the erythroid/megakaryocyte versus myelomonocytic lineages, the zinc-finger transcription factor GATA1 and the Ets family transcription factor PU.1 play key roles. GATA1 is expressed in the erythroid/megakaryocyte and PU.1 in the myelomonocytic lineage. These two lineage-specific and lineage-specifying transcription factors promote the transactivation of large arrays of effector genes that implement the erythroid or the myelomonocytic program, respectively (Ferreira et al., 2005; Rosmarin et al., 2005). They also mutually inhibit each other in that over-expression of either one can suppress expression of the other and thereby influence the respective fate outcome (reviewed in Graf, 2002). Thus, in the classification of gene regulatory network components proposed by Davidson and Erwin (2006), GATA1 and PU.1 not only control directly the "differentiation gene batteries" but also belong to the class of developmental "switch genes". Excess of PU.1 shifts the balance towards myelomonocytic differentiation while excess of GATA1 promotes the erythroid/megakaryocyte fate (Cantor and Orkin, 2001; Graf, 2002). However, this model of a graded "stoichiometric balance" (Cantor and Orkin, 2001) between these two factors does not explain how the progenitor state is normally stabilized and forced to make a discrete "either–or" decision. Importantly, PU.1 and GATA1 also bind to sites in their own promoter region and stimulate their own transcription (Okuno et al., 2005; Tsai et al., 1991). A common (ad hoc) interpretation of this auto-regulation is that it allows cells to maintain the lineage-specific expression of these key regulatory factors. In contrast we have previously proposed a theoretical model in which the auto-regulatory loops of these transcription factors confer (meta)stability to the promiscuous state of the bipotent progenitor cells (Huang, 2005). Similar conclusions for the same as well as other genetic circuits with analogous architectures involved in early stem cells development have recently been presented (Chickarmane et al., 2006; Roeder and Glauche, 2006).

To better understand how the integrated behavior of cross- and auto-regulation can govern the actual cell fate decision, we analyze here a simple mathematical model for predicting the generic behavior of the class of circuits exemplified by the GATA1–PU.1 system and compare it with experimental data from gene expression profiling (Bruno et al., 2004) and cell fate monitoring of differentiating FDCP-mix cells. This bone marrow-derived cell line recapitulates the properties of a bipotential progenitor cell (P) in that it undergoes differentiation in response to specific growth factors into cells with properties of either the erythroid/megakaryocyte (E) or the myelomonocyte lineage (M) (Fig. 1). We propose here that the multipotential state represents a "metastable" state and that during a binary cell fate decision this state is first destabilized upon which it faces the obligatory choice between the two developmental alternatives. This model also unites the "selective" and the "instructive" model of fate determination in a specific fashion. Experimental observations of the trajectories in the GATA1–PU.1 plane, the genome-wide transcriptome changes and cell death kinetics during differentiation are in agreement with the qualitative predictions of the model.

Materials and methods

Experimental data analysis

mRNA profiles in differentiating FDCP-mix cells were derived from microarray data as described by Bruno et al. (2004). The gene set of 3560 genes that were differentially expressed in at least one of the four treatments was used. Normalized signal intensities of gene i, $s_i(t)$ at time t as reported by Bruno et al. (2004) were transformed to $\log_2$ of ratios of relative expression, x_i, for all subsequent computations: $x_i = \log_2[s_i(t)/s_i(t=0)]$.

For GEDI analysis (Eichler et al., 2003), Version 3.0 was used (http://www.childrenshospital.org/research/ingber/GEDI/citation.html). Parameters for the underlying SOM (self-organizing map) were as follows: Linear initialization, Euclidean distance as similarity metric, 20×19 grid. Training parameters were as follows (phase 1, phase 2): train number: 80, 80; neighborhood radius: 1, 3;

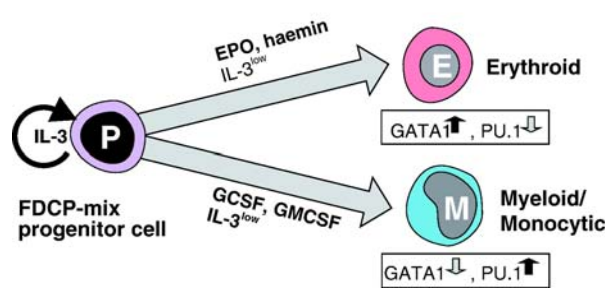

Fig. 1. Experimental model system: FDCP-mix cells.

learning factor: 0.5, 1.0; neighborhood block size: 2, 2; conscience factor: 3.0, 3.0; these settings led to convergence.

State space non-linear dimension reduction with Isomap was performed according to Tenenbaum et al. (2000) (http://isomap.stanford.edu). For calculating the geodesic intertrajectory distances on the manifold the source code of isomap (written in Matlab) was modified to extract the geodesic distances between the sample profiles.

Analysis of cell growth, differentiation, apoptosis and death after induction of differentiation FDCP-mix cells

The cells were routinely cultured in Fischer's medium with 20% horse serum and 2% IL-3-conditioned medium. Differentiation of FDCP-mix cells was performed as previously described by Bruno et al. (2004). Briefly, in the erythroid differentiation for apoptosis analysis, the FDCP-mix cells were cultured in IMDM supplemented with 20% FCS and 100 ng/ml SCF, 10 ng/ml IL-6, 0.2 mM hemin and 10 U/ml Epo. In neutrophil differentiation condition, the growth factor concentrations were 10 ng/ml SCF and 10 000 U/ml G-CSF. For colony assay, cells were plated in MethoCult M3234 (StemCell Technologies, Inc.) containing 0.01 ng/ml IL-3, 100 ng/ml SCF and 10 U/ml Epo, at 500/ml or at adjusted densities to obtain appropriate numbers of colonies for scoring. The annexin V–PE apoptosis detection kit (BD Pharmingen) was used for apoptosis analysis following supplier's instructions.

Modeling

Established regulatory influences

There are many possibilities to formalize the circuit in Fig. 2. One can describe the detailed molecular mechanisms and use a chemical–kinetic formalism. However, the most robust observations, repeated in various experimental systems in different species, concern functional, qualitative relationships between the variables (GATA1 and PU.1) (Galloway et al., 2005; Graf, 2002) whereas molecular processes in vivo, such as post-translational events (phosphorylation, dimerization, regulated proteolysis) remain largely uncharacterized and the published knowledge represent only an idiosyncratic subset of biochemical observations. At the current stage of knowledge they are of qualitative nature and neglect the abundant regulatory inputs from factors outside the modeled circuit. A model describing molecular mechanisms within the circuits may hence suffer more from ignoring inputs from outside the circuit than one that is based on demonstrated functional influences from overexpression or inhibition experiments (Galloway et al., 2005; Graf, 2002; Laslo et al., 2006). Thus,

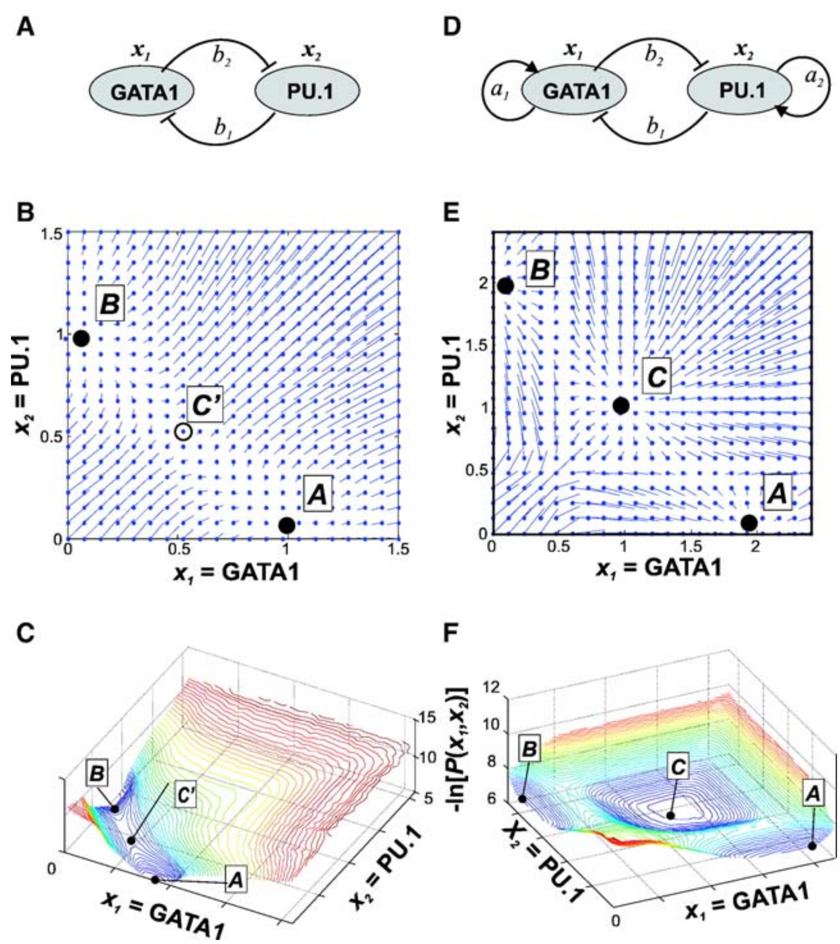

Fig. 2. The GATA1–PU.1 gene regulatory circuit and its dynamic behavior. Didactical schemes of state space. (A, D) Structure of the gene circuit with mutual inhibition between the two transcription factors, GATA1 and PU.1 and without (A) or with (D) auto-stimulation. x_1 and x_2, system variables for activity of GATA1 and PU.1, respectively. Typical dynamic behavior for circuit *A* (bistability) is shown in the state space representations in panels B and C; and for circuit *D* (tristability) is shown in panels E and F. (B, E) Flows of $S(t)$ starting from a grid of initial states in GATA1–PU.1 state space. In this example, $n=4$, $k_1=k_2=1$, $\theta=0.5$, $a_1=a_2=a=0$ for panels B, C and $a=1$ for panels E, F, and $b_1=b_2=b=1$. Black solid circles denote steady states, the attractors *A*, *B* and *C*, and empty circle, the saddle C'. (C, F) "Attractor landscape" with a vertical dimension (visualized by contour curves) representing the negative logarithm of the simulated probability that a state *S* is at position (x_1, x_2). The landscape is for illustration purposes and is not a true potential function. Attractor *C* corresponds to the progenitor state, while attractors *A* and *B* represent the erythroid and the myeloid states, respectively.

Eq. (1) describes the functional relationships of activities of nodes in an influence network and not the actual biochemical interactions.

Meaning of parameters

Since we model here the observed functional interactions in an influence network rather than molecular mechanisms, in Eq. (1), the parameters a_1 and a_2, respectively, represent the relative strength of the "influences" of the self-induction of GATA1 and PU.1, while b_1 and b_2 capture the influence of the cross-inhibition on the basal activation. k_1 and k_2 represent the rate of the first order deactivation. The parameters θ represent the threshold (inflection point) of the explicitly sigmoidal functions, i.e., the strength of the regulatory interaction ($1/\theta$ corresponds to the constant K; n is the Hill coefficient which determines the steepness of the sigmoidal function). Since the detailed biochemistry that is captured by the parameters n and θ is not known we set here $\theta_{a_1}=\theta_{a_2}=\theta_{b_1}=\theta_{b_2}=\theta$ and n equal for all terms of Eq. (1) and to values that can produce bistability ($\theta=0.5$–1.5, $n=4$–8) based on the known separate behavior of parts of the circuit, cross-inhibition or auto-regulation (Gardner et al., 2000; Kaplan and Glass, 1995; Xiong and Ferrell, 2003).

Rationale behind the individual terms of Eq. (1)

The stimulatory and inhibitory influences are explicitly modeled as sigmoidal relationships using Hill functions (Gardner et al., 2000; Kaplan and Glass, 1995; Laslo et al., 2006; Xiong and Ferrell, 2003). The sigmoidal shape of these functions is justified chiefly by the intracellular physicochemical conditions (see Savageau, 1995; Huang, 2002, and references therein), although cooperative kinetics at the molecular level due to multimerization are likely to be involved.

The *first term* in Eq. (1) represents the contribution to activation of transcription factor x_i (PU.1 or GATA1) by auto-stimulation of synthesis/activation and is zero in the absence of x_i. The *second term* represents the cross inhibition, in that a constitutive basal synthesis of PU.1 and GATA1 is cross-inhibited, based on the following rationale: (*i*) the binding by PU.1 and GATA1, respectively, to the complex and not fully understood *cis*-regulatory regions is only a small aspect of gene regulation, which receives many other "background inputs", including chromatin modification processes and (*ii*) experimental observation of basal synthesis of these factors in non-differentiated state in hematopoietic cell lines (S. H., unpublished observation).

Further, to err on the side of over-minimizing the model, inputs from auto-stimulation (first term) and cross-inhibition (second term) are related to each other additively and not multiplicatively (Laslo et al., 2006). This assumes a primary mechanistic independence of these two processes and is justified as follows: Although GATA1 and PU.1 can interact at the protein level, multiple other proteins, some of which are regulated by these factors but are ignored here, bind to them and interfere with this interaction, relieving the direct mutual dependence. In fact multiple, distantly separated GATA1 and PU.1 binding sites exist in the GATA1 and PU.1 genes, respectively, whose relative contribution and regulation modality are not well known (Cantor and Orkin, 2002; Ferreira et al., 2005; Rosmarin et al., 2005). Finally, either factor may promote transcription of itself (auto-regulation) by opening the chromatin structure while being transcribed (Fisher, 2002; Georgopoulos, 2002). In the model the mutual influence between auto-stimulation and cross-inhibition is achieved indirectly via the levels of x, e.g., auto-stimulation is suppressed by cross-inhibition for cases with $a_1<b_1$ when x_1 is sufficiently low (relative to θ_a).

The *third term* is a first-order decay of activity with apparent rate constant k. This simple assumption is made since there is no evidence that GATA1 controls PU.1 degradation and vice versa.

Computation

Numerical solution of Eqs. (1) and (2), state space and parameter space analysis were performed in Matlab (Mathworks, Inc.) using the ode45 solver. Fixed points (in the state space and bifurcation diagrams) were determined numerically as intersects of the $\mathrm{d}x_1/\mathrm{d}t=0$ and $\mathrm{d}x_2/\mathrm{d}t=0$ null-isoclines for Eq. (1) by numerically solving for the roots of the algebraic forms and confirmed by numerically integrating the ODEs using the solver, which also confirmed their stability.

Results

Dynamic system modeling of the GATA1–PU.1 gene regulatory circuit

Circuits containing mutual inhibition, like that between the transcription factors GATA1 and PU.1, are well known and have long been used to describe bistable genetic switches, since they can produce two discretely separated stable states (Gardner et al., 2000; Glass and Kauffman, 1973) (Fig. 2A). However, GATA1 and PU.1 also promote their own transcription (auto-regulation) (Fig. 2D). Since the molecular details of regulation are still largely unknown, a physicochemical model would contain too many uncertain assumptions. Hence, we formulate a minimal model based on the well-documented functional relationships. For simplicity, we begin by assuming an insulated modular system whose system state $S(t)$ is defined by the relative activity levels of two transcription factors, GATA1 and PU.1 at time t: $S(t)=[\text{GATA1, PU.1}]$. The state vector $S(t)$ evolves in time as cellular GATA1 and PU.1 activities change in a mutually dependent manner dictated by the circuit. To determine the potential dynamic behavior of S under these constraints we model the circuit at the same level of coarse-graining as previously used for generic analysis of the same class of circuits (Gardner et al., 2000; Glass and Hill, 1998; Glass and Kauffman, 1973; Laslo et al., 2006). Synthesis and activation of GATA1 and PU.1 (transcription, translation and post-translational modification) were combined together, since these processes that occur at the time scale of minutes to hours can be separated from that of the integrated circuit behavior that drives differentiation at the time scale of multiple days. Then the regulatory influences that constrain the dynamics of $S(t)$ are described by the "activity" of these two proteins, x_1 (for GATA1) and x_2 (PU.1), respectively, leading to the following ordinary differential equations:

$$\frac{\mathrm{d}x_1}{\mathrm{d}t} = a_1\frac{x_1^n}{\theta_{a_1}^n + x_1^n} + b_1\frac{\theta_{b_1}^n}{\theta_{b_1}^n + x_2^n} - k_1x_1$$
$$\frac{\mathrm{d}x_2}{\mathrm{d}t} = a_2\frac{x_2^n}{\theta_{a_2}^n + x_2^n} + b_2\frac{\theta_{b_2}^n}{\theta_{b_2}^n + x_1^n} - k_2x_2 \tag{1}$$

$a_{1/2}$, $b_{1/2}$, $k_{1/2}$, all the θ's and n are non-negative parameters. The first term in each line represents auto-stimulation of activity, the second term is the cross-inhibition (manifest as reduction of basal activity) and the third term is the unregulated decay of activity (for details, see Materials and methods). We focus here on the effects of the parameters a, b and k since we are interested in the relative influence of the various regulatory interactions. This minimal model approach aims not at "fitting" the modeled solution curves for $x_1(t)$ and $x_2(t)$ to observed changes in GATA1 and PU.1 expression but instead, at analyzing the generic behavioral potential of the GATA1 and PU.1 system constrained by the circuit. Its simplicity allows us to exhaustively explore the entire relevant state and parameter space and arrive at a qualitative, testable prediction on general principles that govern cell fate decision.

Behavioral repertoire of the GATA1–PU.1 circuit: tristability

Numerical solutions for Eq. (1) are graphically depicted in the two-dimensional phase plane spanned by x_1 and x_2, showing how the system $S(t)=[x_1(t), x_2(t)]$, which represents a particular GATA1–PU.1 activity configuration, evolves over time. Fig. 2 shows the vector field and contour lines of a schematic "potential function" in the phase plane.

When auto-regulation is not considered ($a_1=a_2=0$) the circuit reduces to the widely studied system of mutual inhibition that exhibits a bistable behavior for a wide range of parameter values (Fig. 2, left panels) (Gardner et al., 2000). This system has three *stationary* states: two stable attractor states, A ($x_1^A \gg x_2^A$) and B ($x_1^B \ll x_2^B$) and (for symmetric parameter values with respect to x_1 and x_2) a central unstable equilibrium state C' ($x_1^{C'} \sim x_2^{C'}$), a saddle state that sits on the separatrix between the two basins of attraction of A and B (Figs. 2B, C). In contrast, presence of auto-stimulation of a similar magnitude as the cross-inhibition ($a_{1,2} \sim b_{1,2}$) (Fig. 2, right panels) converts the unstable saddle C' into a third (meta)stable central attractor C which (for symmetric parameter values) is located on the $x_1=x_2$ diagonal (Figs. 2E, F). The new attractor state has a large basin of attraction, occupying most of the center of the x_1–x_2 phase plane. The three attractors are separated by two nonlinear, stable manifolds.

If cell fates correspond to the attractor states of the gene-regulatory network then the stable attractor state A (GATA1 $\gg$ PU.1) would represent the erythroid fate (E) while the B attractor (PU.1 $\gg$ GATA1) would represent the myeloid (M) fate. Accordingly, in the tristable system (right panels of Fig. 2) the middle attractor C can be interpreted as the bipotential progenitor cell (P), a state with equal access to either lineage. Then the model makes the prediction that the progenitor cell simultaneously expresses low but equal levels of both fate-specific transcription factors, thus, it provides a formal explanation for the promiscuous expression of lineage-specific factors by progenitors (Cross and Enver, 1997). In this model, the undecided progenitor state (P) is "dynamically" defined, namely, as a metastable state in between two neighboring attractors of the prospective differentiated states (M and E).

One can certainly construct alternative explanations for the promiscuous expression of the lineage-specific transcription factors. For instance, the linear stoichiometric balance model (Cantor and Orkin, 2001) would also be compatible with a promiscuous state. However, the non-linear model presented here is the simplest formalism which explains the local stability of the progenitor state and the switch-like dynamics of differentiation as a consequence of the circuit-architecture.

We next asked how the attractor landscape is affected when we change the parameter values of the equations. The systematic exhaustive analysis of the parameter space is possible because of the simplicity of the model. An example of a parameter space analysis is shown in Fig. 3 which represents a 5 × 5 matrix of the state spaces (x_1–x_2 phase planes) of 25 different circuits, each with the trajectories of $S(t)$ starting form a grid of initial states, obtained by combined variations in $a=a_1=a_2$, $b=b_1=b_2$ and $k=k_1=k_2$, i.e., using symmetric variations for x_1 and x_2 (Figs. 3A, B). Symmetric variations of the corresponding parameters for x_1 and x_2 preserve the quality of tristability but just shrink/stretch the attractor landscape. Qualitative changes in the structure of the state space appear around bifurcation points (see below). Asymmetric variations of parameters with respect to x_1 and x_2 result in trivial asymmetries in the position of the attractors with respect to the diagonal in the phase plane until in the extreme case, one of the marginal attractors for the differentiated state overtakes the central progenitor attractor (Fig. 3C). Such asymmetric scenarios were, given the underlying biological rationale, not further analyzed (see below and Discussion). In summary, the regime for existence of tristable dynamics with a central attractor C ($x_1 \sim x_2$) and two marginal attractors A ($x_1 \gg x_2$) and B ($x_2 \gg x_1$) is broad and thus, the tristable behavior appears to be a "generic" ("typical") property of the system described by the Eq. (1).

Bifurcation in the dynamic behavior of the GATA1–PU.1 circuit

The model formalism according to which the progenitor cells sit in a (meta)stable central attractor state, allows us to now ask the question how cells leave this locally stable state when stimulated to differentiate. The parameter space analysis (Fig. 3) revealed that for certain values of parameters the central attractor state C disappeared, i.e., the system undergoes a bifurcation. (This term is *not* related to the biological term of "bifurcating" differentiation paths of a bipotent cell.) We hypothesized that induction of differentiation causes a circuit configuration with a central progenitor state C (with $x_1 \sim x_2$) to change so that C is destabilized and disappears, giving rise to bistability (Fig. 2, left). Then, a state S originally at or near the attractor state C will inevitably be confronted with choosing between stable state A ($x_1 \gg x_2$) or B ($x_1 \ll x_2$). Thus, we asked which modification of parameter values is biologically plausible and can cause such a qualitative change of the "attractor landscape". An associated fundamental question is how the breaking of the symmetry of bipotentiality of the progenitor state is achieved during spontaneous cell fate commitment. We discuss here two theoretical scenarios, which are opposite extremes from a spectrum of possible formalizations of such symmetry-breaking transitions (see Table 1 for a systematic overview):

Scenario (*i*) The differentiation signals change a parameter only in one of the equations in (1). For instance, a myeloid-stimulating agent may increase k_1 or decrease a_1 so as to almost completely deplete the cell of GATA1 activity, collapsing attractor A (Fig. 6C). This model of asymmetric transition recapitulates the purest form of *instructive*, *deterministic* fate regulation by external signals. Such a process would entail that the external signals bear all the burden of the symmetry-breaking and would be manifest in the two differentiation trajectories immediately moving apart in opposite directions, from C towards their respective destination states A, B.

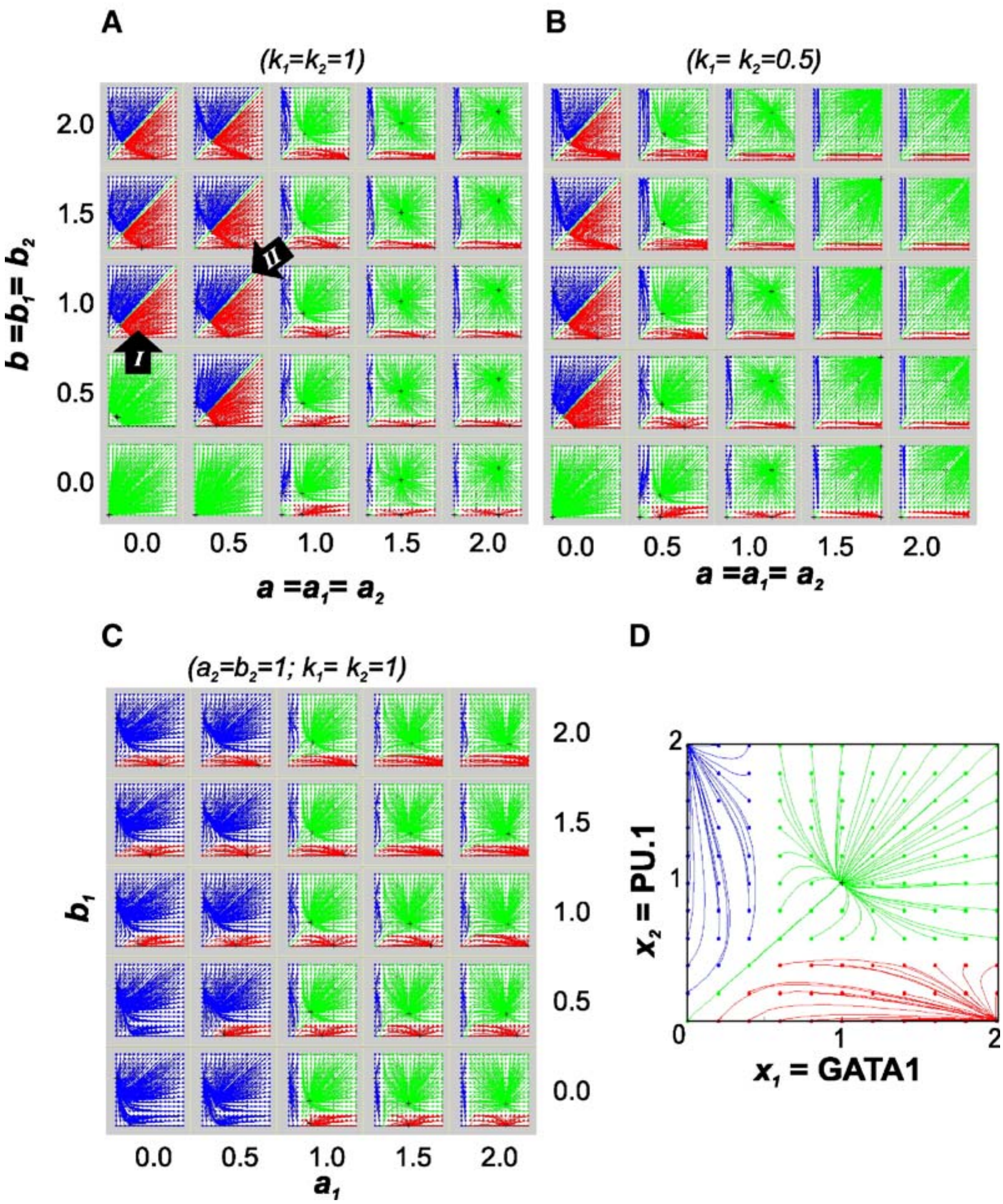

Fig. 3. Global behavior and parameter space. (A–C) Evaluation of the effect of pairwise changes of parameters on state space structure for Eq. (1), examples. Each cell in the 5 × 5 array represents the x_1–x_2 phase plane illustrating the trajectories starting for a grid of initial states $S(t=0)$ in the state space, as shown enlarged in panel D, for parameter values as indicated on the axis of the array. The basic set of parameters that were varied was $a=b=k=1$, $\theta=0.5$, $n=4$. (D) One example phase plane, using the basic parameter set of Fig. 2E. The grid distance between the initial states was 0.2. Blue and red: trajectories ending in the attractors for the committed cell fates *M* and *E*, respectively. Green: trajectories ending in the central attractor for the state of progenitor cells, *P*. The black arrows, *I* and *II*, refer to the type of symmetric bifurcation (Fig. 4).

Scenario (*ii*) The differentiation signals perturb the circuit, affecting x_1 and x_2 concomitantly so as to cause the progenitor cell attractor state to become unstable in a process that is more or less *symmetric* (with respect to x_1 and x_2). This implies initial changes that are similar (but not identical) for both differentiation processes. In a first phase, the “potential well” for *P* (Fig. 2) is converted to a “hill-top” so that a state located in this region of the state space becomes unstable. Then, in a second phase, small deterministic differences in the triggering signals or stochastic fluctuations of regulatory molecules can “tip” the cell into either attractor state on each side of the hill. This concept captures the *stochastic* or *intrinsic* model of cell fate determination, for most of the symmetric-breaking burden is born by the cell. But it also explains increased susceptibility of the destabilized progenitor state to instructive influences.

In the first, instructive model (*i*) the disruption of the symmetry of the *P* state (GATA1 ~ PU.1) is trivial; for instance, an external factor may inhibit GATA1 activation by lowering a_1 to induce myeloid differentiation, represented moving to the left

Table 1
Concepts for modeling quasi-discrete transitions between stable network states (attractors) at a binary fate decision point

(A) Transition via a bifurcation (changing state space structure) caused by regulatory signals that induce a gradual change in control parameters (a, b, k, ...). Two possible scenarios:
i. Asymmetric change leading to state space dominance of one of the two terminal attractors A or B. (*deterministic* regulation)
ii. Symmetric change leading to bistable regime with ~ equally large attractor basins for A and B. (*deterministic*+*stochastic* regulation)
Type I Bifurcation = Super-critical pitchfork bifurcation (Fig. 4A)
Type II Bifurcation = With a subcritical pitchfork bifurcation (Fig. 4B).
(B) Transition as jump in "state space" (static state space structure) caused by perturbation (signal, noise) that modulates values of system variables

on the panels in Fig. 3C. Thus, we asked how the intrinsic symmetry-breaking in the second, two-phase model (scenario *ii*) can occur. To determine the parameter variations that may trigger a (near) symmetrical bifurcation we varied the parameters a, b, θ and k *symmetrically* for both state variables, x_1 and x_2 (i.e., $a=a_1=a_2$, $b=b_1=b_2$, etc.) and analyzed all possible bifurcations at which any central attractor state C, with $x_1 \sim x_2$, disappears and gives rise to a bistable situation with two disparate attractor states A ($x_1 \gg x_2$) and B ($x_1 \ll x_2$). [The symmetric parameter change facilitates calculations. It is clear that slight asymmetries will be present in reality, whereas strong asymmetry would represent the trivial case of instructive "symmetry breaking" discussed above (scenario *i*).] From the systematic variation of parameters and evaluation of the entire relevant state space (not shown), we found for symmetric parameter variations in Eq. (1) that a symmetric bifurcation (scenario *ii*), in which a central attractor state with GATA1 ~ PU.1 is gradually destabilized and converted into a unstable saddle point (hill top), can be classified into two types:

In the *type I* symmetric bifurcation (Figs. 3A and 4A, C) of this two-phase model, the starting central progenitor state C ($x_1 \sim x_2$) is not in the tristable but in a monostable regime and has a basin that occupies the entire phase space. This monostable behavior can undergo a global *supercritical pitchfork bifurcation*, creating the two differentiation attractors A and B of the bistable regime, while the central progenitor attractor becomes an unstable saddle state (Fig. 4C). Such a bifurcation occurs when b ($=b_1=b_2$) is decreased while the other parameters are kept at fixed values. In contrast, the *type II* symmetric bifurcation (Figs. 3A and 4B, D) is a transition starting from a tristable attractor landscape, going into a bistable regime. As parameter values are varied, the progenitor attractor C is locally converted to an unstable saddle by undergoing a *subcritical pitchfork bifurcation* traversed in reversed direction, while the two marginal attractors A and B are not qualitatively affected (Fig. 2D). In our not maximally simplified Eq. (1), this local bifurcation occurs for instance when a ($=a_1=a_2$) is gradually reduced (e.g., for $b=1$, $k=0.5$ or 1, $\theta=0.5$) (Figs. 3A, B). In this case, disappearance of the central attractor is expected since $a=0$ corresponds to the system without auto-regulatory loops which exhibits the bistable behavior (Fig. 2, left). This type of bifurcation from a tristable to a bistable behavior also occurs when k ($=k_1=k_2$) is increased (e.g., from $k=0.5$ to 1.5, for $a=b=1$, $\theta=0.5$), the case shown in Fig. 4D. The qualitatively same type of bifurcation occurs when the universal threshold θ is increased, e.g., from $\theta=0.5$ to 0.75 (for $a=b=k=1$) (not shown).

Trajectory of differentiation in the GATA1–PU.1 phase space

Very little is known about how external factors, such as EPO or G/GM-CSF in the case of FDCP-mix cells, induce differentiation into either one of the two available fates. Since the scenario of symmetric destabilization (*ii*) explains the observed stochastic component in cell fate decision (Graf, 2002; Hume, 2000), we next asked what the trajectories in the GATA1–PU.1 plane would look like if the progenitor cell would undergo a near-symmetric bifurcation. The trajectory of the cellular state $S^d(t)=$[GATA1, PU.1] for cell differentiation process d in the GATA1–PU.1 state space is an experimental observable and can help to evaluate the model. Thus, we computed the possible state trajectories for the two symmetric bifurcations, *type I* and *type II*, and compared them with the trajectory observed from mRNA expression level measurements in differentiating FDCP-mix cells.

In the absence of knowledge on how the cytokine signals impinge on the molecular circuit that governs the fate decision, we hypothesized that inducing the FDCP-mix cells to differentiate into the erythroid (E) or myelomonocytic lineage (M), respectively, initiates a change in one or more parameters, a, b, θ, or/and k to a similar (but not identical) extent for the two variables, x_1 and x_2. The solution trajectories $S^E(t)$ and $S^M(t)$ are obtained by solving the differential equations for parameters that vary in time at the same time scale as the variables x. Thus, the differential equations in Eq. (1) are extended by two algebraic expressions that describe the time dependence of the parameters that will be varied. For instance, for a *type II* transition is achieved by decreasing a in Eq. (1) as follows:

$$\begin{aligned} a_1 &= a_1^0 \exp(-\lambda_1 t) \\ a_2 &= a_2^0 \exp(-\lambda_2 t) \end{aligned} \tag{2}$$

Here $\lambda_{1,2}$ are the rates for the decrease of a_1 and a_2; a_1^0, and a_2^0 represent the values of $a_{1,2}$ at $t=0$. Then, the biology-inspired model assumptions for fate commitment of bipotential progenitor cells were as follows:

(a) Initiation of differentiation triggers a bifurcation in which the central progenitor state $S_0(t_0)=[x_1(t_0),\ x_2(t_0)]$ of attractor C is destabilized.
(b) In reality even a clonal cell population is phenotypically heterogeneous (Kaern et al., 2005; Spudich and Koshland, 1976). Hence, the initial states S_{0j} of the individual progenitor cells j in a cell population are randomly distributed around the central state $S_0(t_0)=$[GATA1 ~ PU.1] with the "promiscuous expression".
(c) The actual determination of cell fate, i.e., the introduction of asymmetry, is hence influenced by two factors:
    (1) Asymmetries in the position of the initial state S_0 of the progenitor cells in attractor C with respect to the prospective separatrix of the bistable regime.

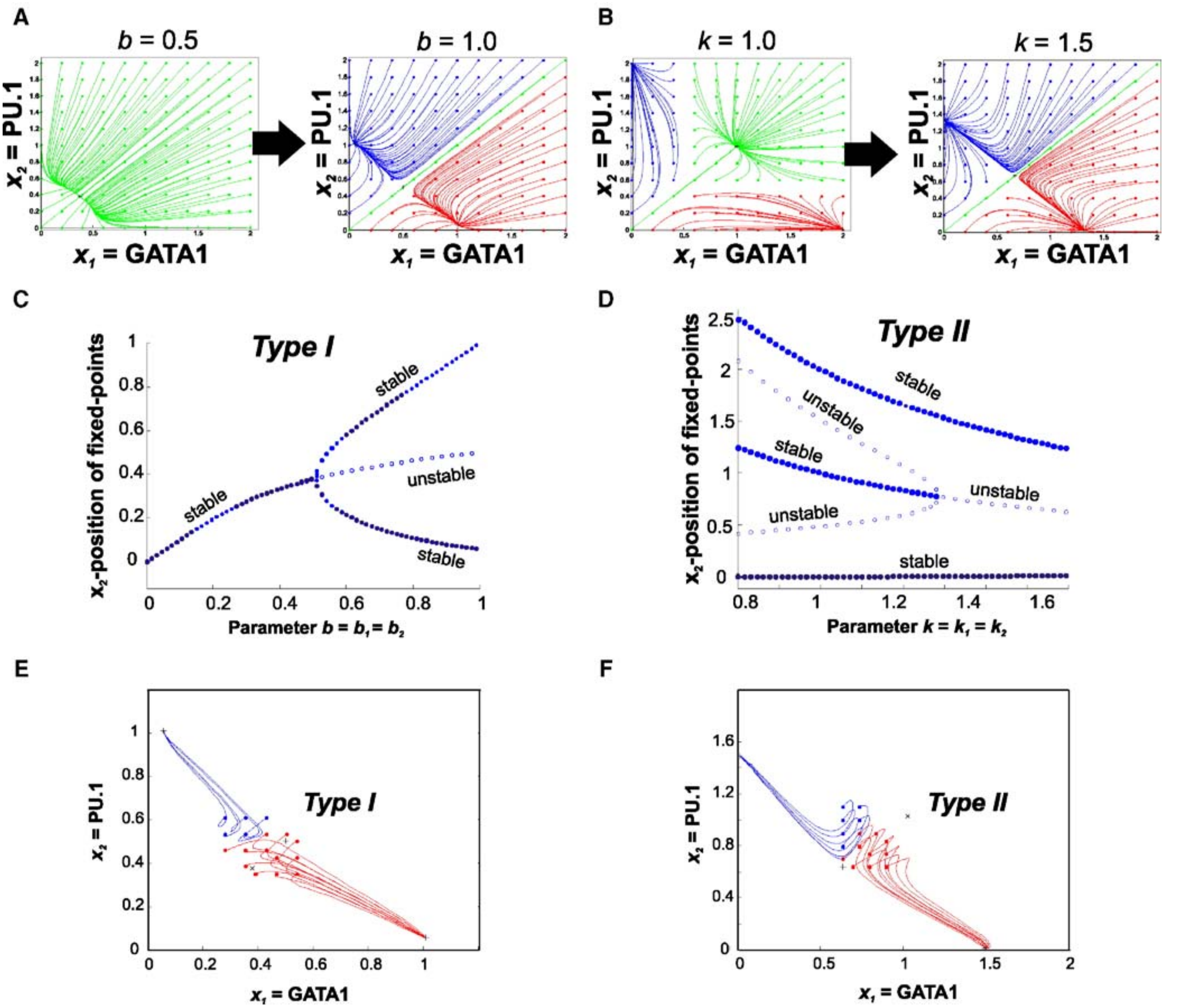

Fig. 4. Two types of symmetric bifurcations that destabilize the central attractor. (A, B) GATA1–PU.1 phase planes with trajectories as in Fig. 3D, before and after the bifurcation point for the parameters (b, k) as indicated. Panel A, type *I* and panel B, type *II* bifurcation. (C, D) Corresponding bifurcation diagrams. The position of the steady states in x_2 dimension (vertical axis) is represented for varying values of the respective parameter (horizontal axis). Solid, closely placed dots: stable steady states (attractors); empty, separated circles: unstable steady states. (E, F) Solution curves of the circuit state as the system undergoes a bifurcation of type *I* and type *II* due to temporal change of the parameters b and a according to Eqs. (1) and (2). Red: trajectories that end in the erythroid state; blue: trajectories that end in the myelomonocyte state. Because a_1 and a_2 in Eq. (2) are treated as system variables the x_1–x_2 plane is not the state space, hence "trajectories" may intersect. For clarity, only relevant initial states in the basin of the original progenitor attractor C with intermediate (x_1, x_2) values are shown. For type *I* bifurcation (*E*), the parameters were $\theta=0.5$, $n=4$, $k_1=k_2=1$, $a_1=a_2=0.01$, whereas b was increased asymptotically from 0.5 to 1.0 in an inverted exponential time course at rates $\lambda_1=0.9$ and $\lambda_2=0.3$. For type *II* bifurcation (F), the basic parameters were $\theta=0.5$, $n=4$, $k_1=k_2=1$, whereas $a_1=a_2$ was exponentially decreased from 1.0 to 0.5 and $b_1=b_2$ was decreased from 1.5 to 1.0; in both cases at rates $\lambda_1=-0.5$ and $\lambda_2=-0.7$.

(2) Asymmetry in the rate of change of the control parameters, e.g. λ_1 vs. λ_2, respectively [see Eq. (2)].

(d) The initial state S moves in state space in accordance to the local vector field by continuously adapting to the changing contours of the attractor landscape in an "*adiabatic*" process (i.e., the parameter change is too fast for the system to reach the steady state for each new parameter value). S moves until it reaches the new attractor state that is established when the change of the parameter itself ceases.

While the model is not intended to predict a particular trajectory course, it divides the potential behavioral repertoire of the system into two qualitatively distinct alternatives for destabilizing the metastable state at C, thus offering the possibility to determine which of the two alternative bifurcation types (*I* or *II*) may be utilized by the cell. To do so we needed to extract distinct observable behaviors associated with either possibility. We first performed a systematic analysis of the parameter space for all trajectories starting in the central region around C. To force cells starting from the very same initial state

P around the center [GATA1 ~ PU.1] to differentiate to either one of the two cell fates (M, E) upon bifurcation, some asymmetry had to be introduced with respect to either S_0 (paracentral position) or λ ($\lambda_1 \neq \lambda_2$) or both. When this enforced asymmetry produced trajectory pairs in which the two trajectories that started from a same paracentral point reached different terminal attractor (either M or E), one of the two trajectories exhibited a characteristic loop as exemplarily shown in Figs. 4E, F. Specifically, numerical analysis revealed that for type *I* bifurcation, given the imposed asymmetry, if a trajectory S^M leading to the myelomonocytic attractor [PU.1 ≫ GATA1] exhibited a loop (not all trajectories do), then it was always in the *counterclockwise* direction (Fig. 4E). In contrast, trajectories undergoing type *II* bifurcation, if they produced loops before reaching the M [PU.1 ≫ GATA1] attractor state, they were always in the *clockwise* direction (Fig. 4F). (Since the model is 'cell fate neutral', scenarios with the inverse trajectory assignment, i.e., with the S^E instead of the S^M trajectory containing the loop, are also generated, but are not relevant for comparison with data.)

Fig. 5A shows the typical predicted trajectories for a type *II* bifurcation that start from the same initial, but not precisely central state S_0 (PU.1 > GATA1), with bias introduced by asymmetric control parameters λ ($\lambda_1 \neq \lambda_2$) tuned so that the two trajectories move towards either one of the two final attractors. Experimentally observed trajectories obtained from three independent microarray measurements of mRNA levels of GATA1 and PU.1 in differentiating FDCP-mix cells displayed a characteristic *clockwise* loop in the myelomonocyte differentiation trajectory (Figs. 5B–D), consistent with a type *II* bifurcation. The characteristic trajectory course for myelomonocyte differentiation in the phase–plane translates into a non-

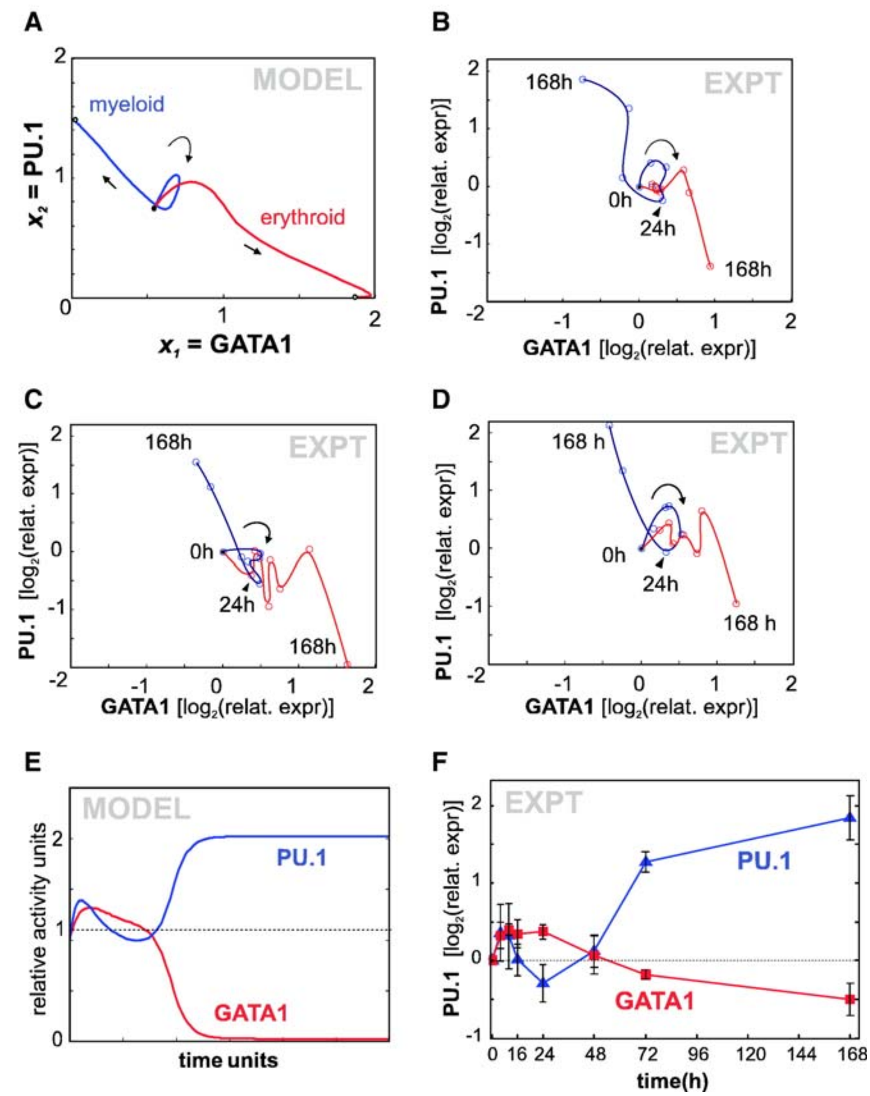

Fig. 5. State space trajectories during *type II* bifurcation: model and experimental data. (A) Typical trajectory for a system starting from one non-central initial state S, differentiating into either the erythroid fate (red) or the myelomonocytic fate (blue) because of asymmetry in λ; $\lambda_1 = -0.3$, $\lambda_2 = -0.5$, with parameters as in Fig. 2E. (B–D) Three independently observed trajectories for mRNA levels of GATA1 and PU.1 (as $\log_2$ of ratio, with 0 h value as reference) measured by microarrays (Bruno et al., 2004) during erythroid and myelomonocytic differentiation of the FDCP-mix cells. Each dot represents a time point for a microarray measurement (0 h, 4 h,..., 168 h – the same time points as in Fig. 6). (E, F) GATA1 (red) and PU.1 (blue) levels during myeloid differentiation for model vs. measurement, displayed as average of time course. Error bars indicate standard error (n=3).

monotonical time course for GATA1 and PU.1 (Figs. 5E, F) which also illustrates qualitative agreement between model prediction and experimental observation.

In conclusion, the exhaustive simulation allowed us to avoid the usual parameter optimization and "curve-fitting" to validate a model. Instead, it presented two alternative, mutually exclusive and qualitatively distinct behaviors within the entire behavioral repertoire of the model system. Hence, comparison with experimental results does not primarily serve to find agreement between observed and predicted behavior but to differentiate between the two possible behaviors. Yet, the close similarity of one predicted behavior with the observed, counterintuitive trajectory courses of differentiating FDCP-mix cells in the GATA1–PU.1 plane lends additional support to the central idea of the model: lineage commitment may be a two-stage process, with a bifurcation event causing a near-symmetric destabilization of the progenitor attractor that precedes fate commitment (scenario *ii*). This makes fate determination asymmetrically imposed by purely "instructive" external signals (scenario *i*) less likely.

Transcriptome change during binary decision

A network topology *motif*, such as the 2-gene circuit in Fig. 2, is not necessarily a functionally independent *module*. The transcription factors GATA1 and PU.1 regulate and are regulated by many other genes and hence, are embedded in an almost genome-wide gene regulatory network ("giant component") (Huang, 2004), which establish and maintain the cell-type-specific transcriptomes. The changes in the transcriptome accompanying differentiation can be formalized as the high-dimensional equivalent of the two-dimensional state $S^d = [x_1, x_2]$ introduced above: in general terms, the state vector $S^d(t)$ for process d, $S^d(t) = [x_1^d(t), x_2^d(t), \ldots, x_N^d(t)]$ describes the dynamics of a complex network of N genes (Huang et al., 2005), where $x_i^d(t)$ is the activity level of gene i at time t in the differentiation process d. As cells differentiate, $S^d(t)$ describes a trajectory in the N-dimensional state space. Each expression profile measurement ("time point") for $d = E$ and $d = M$ is represented by a point S in the N-dimensional state space. Although very little is known about the architecture of genome-spanning regulatory networks in mammalian cells, which precludes genome-scale modeling, we hypothesized that the dynamics of fate decision via an initial destabilization of the progenitor state, as postulated and formalized in the circuit model above, may be manifest in the dynamics of the genome-scale transcriptome because most of the individual genes are interconnected by the giant component of the regulatory network. This would qualitatively predict that because of the near-symmetric initial destabilization of the core regulatory circuit in the progenitor state, the initial transcriptome changes for erythroid and myeloid differentiation would be similar (but not identical) despite the stimulation with disparate external differentiation signals.

Thus, we analyzed the path of change of the transcriptome of over $N = 3560$ genes in FDCP-mix cells stimulated to differentiate either into the erythroid or the myelomonocytic lineage (Bruno et al., 2004). If cell fate decision is initiated by fully asymmetrically acting, instructive signals (alternative *i* in Table 1) then the transcriptome trajectories would rather monotonically move in separate state space directions to adopt the respective cell fate-specific expression patterns. In contrast, if commitment of the progenitor cells to either fate is preceded by an initial, roughly but not absolutely symmetric destabilization of the progenitor state, as postulated above (alternative *ii* in Table 1), and if this is manifest at the transcriptome level, then $S^E(t)$ and $S^M(t)$ would initially move in a similar direction in the N-dimensional state space.

To obtain a simultaneous global view of the time evolution of the transcriptomes of both differentiation paths at the resolution of genes, the time courses of the gene expression profiles of $N = 3560$ genes were displayed with the Gene Expression Dynamics Inspector (GEDI) program. GEDI uses self-organizing maps (SOM) to visually represent parallel time courses of high-dimensional measurements simultaneously (Eichler et al., 2003). Fig. 6 shows the "GEDI maps" each of which represents a gene expression profile at time point t during either differentiation process d, which in turn can be regarded as the N-dimensional state vector $S^d(t) = [x_1^d(t), x_2^d(t), \ldots, x_N^d(t)]$. The GEDI maps confirm that differentiation of the progenitor cells into the erythroid and myeloid lineages was not a progressive transition along the shortest path from the gene expression pattern of the progenitor into separate directions to adopt the respective target expression profiles. Instead, both differentiation paths, S^E and S^M, exhibited strikingly similar transcriptome changes up to ∼24–48 h after induction before separating to develop the expression patterns characteristic of the respective differentiated states that is evident at 168 h (Fig. 6A). This common initial path in N-dimensional state space is consistent with the model of a near-symmetrical destabilization process of the progenitor state. The common initial trajectory was more pronounced when we limited analysis to the previously defined subset of 347 differentiation-specific genes (Bruno et al., 2004) that exhibited significant change in all four analyzed cell fates derived from FDCP-mix progenitors (Fig. 6B).

Genome-scale differentiation trajectories

To more quantitatively characterize the above transcriptome change we next asked how the two N-dimensional trajectories $S^E(t)$ and $S^M(t)$ for $N = 3560$ genes are positioned relative to each other in the N-dimensional state space. We used the "Isomap" technique for non-linear dimension reduction to embed the trajectories into a three dimensional space (Tenenbaum et al., 2000). In contrast to linear dimension reduction, such as principal component analysis, Isomap assumes that all input points [=the measured expression profiles at time points t_i, $S^d(t_i)$] in the N-dimensional space collectively form a non-linear *manifold*, i.e., a subspace of a particular shape and dimension n, where $n \ll N$. The reduction of the maximal N-dimensional space to an n-dimensional manifold on which all measured expression profiles reside reflects the dynamic constraints imposed by the gene regulatory network which collapses the space of "realizable" expression profiles

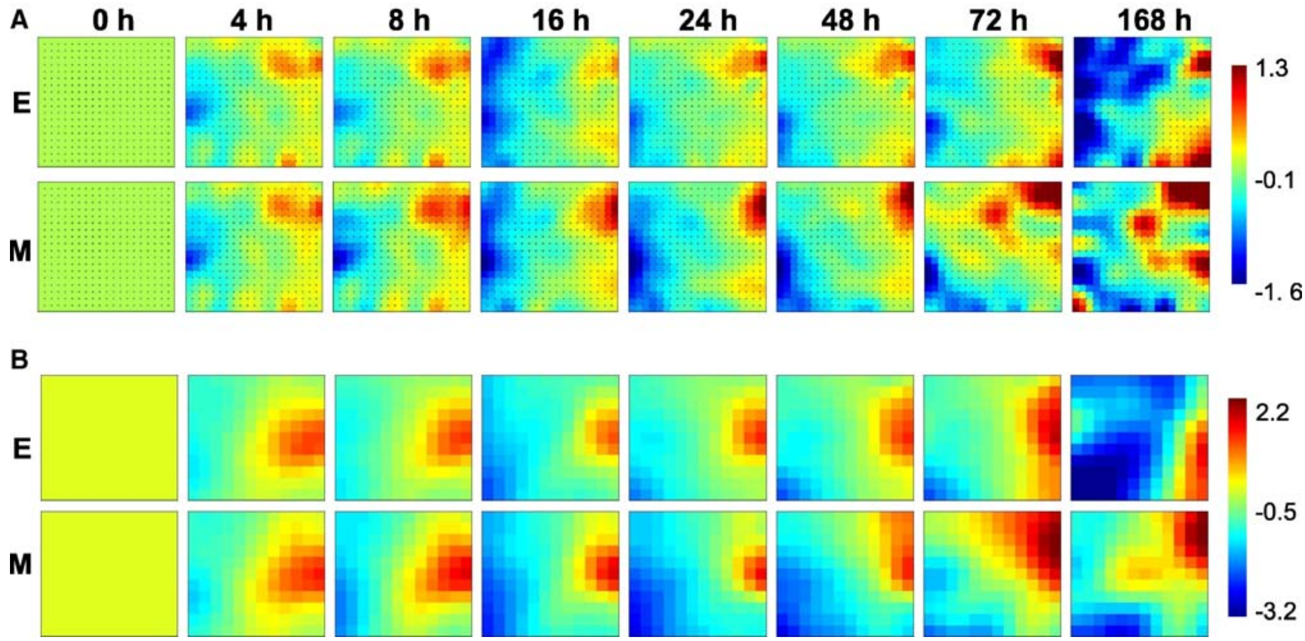

Fig. 6. Time evolution of expression profiles during erythroid and myeloid differentiation of the FDCP-mix cells. Each GEDI map ("mosaic") represents a gene expression profile at a time point during either differentiation process, i.e. $S^d(t)$ (Eichler et al., 2003). For analysis, $\log_2$-transformed, relative expression levels for each gene were used, with expression at 0h as reference, hence the first time point appear in the uniform green color for the value 0. (A) Expression profiles for all $N=3560$ genes expressed in FDCP-mix cell lines (Bruno et al., 2004). Each "tile" in the 19×20 mosaics represents a "minicluster" each containing on average approximately 10 genes that behave highly similarly across all expression profiles. Since the assignment of genes to the miniclusters is universal for the entire set of expression profiles, the GEDI maps can be directly compared to each other. (B) GEDI analysis on a 11×12 grid for the subset of 347 genes that are differentially expressed in all lineages, as determined by Bruno et al. (2004). Color bars indicate $\log_2$(relative expression).

(Huang, 2005). Then, the relevant state space distance between any two states, S_i and S_j on the manifold may not be well represented by the Euclidean distance E_{ij}, but better by the geodesic distance G_{ij} on the manifold. (If the manifold is a "hilly landscape", then the geodesic distance would be the shortest terrestrial path between two points on two different hills, but the Euclidean distance would be the direct aerial distance.) Isomap applies classical multidimensional scaling (MDS) (Cox and Cox, 1994) on the matrix of all the distances between the available profiles S_i (from the two time courses $S^E(t)$ and $S^M(t)$) measured as (estimated) geodesic distances G_{ij} on the manifold instead of Euclidean distance E_{ij}. The true (intrinsic) shape of the manifold is not known but can be approximated by the geodesic structure spanned by all the measured points S_i (Tenenbaum et al., 2000).

Fig. 7A shows how all the measured profiles along the trajectories S^E and S^M can be embedded with Isomap in a three dimensional space. [To achieve the best possible "triangulation" of the manifold, we also included the time points from two other differentiation trajectories for which expression profiles were also measured (Bruno et al., 2004), describing the differentiation of FDCP-mix cells into megakaryocytes (lineage related to the erythroid studied here) and into neutrophils (related to the myeloid–monocyte lineage).] The distribution of a total of 31 time points in the three-dimensional embedding space describes roughly an inverted Y shape, with the progenitor cell (P, $t=0$) at the bottom of the trunk of the Y, on the left. Upon induction of differentiation, the time points of both the trajectories S^E and S^M remained in close proximity within the trunk of the Y for time points up to 48 h after induction, progressing into the same direction along dimension 1. The trajectories separated considerably only at $t=72$ h into different directions along dimensions 2 and 3 (the two arms of the Y) heading towards the terminal states. The residual variance (inset in Fig. 7) indicates that the dimensionality of the manifold as approximated by the 31 available gene expression profiles is three, and that the three dimensional embedding space (Fig. 7A) recovers 95% of the information.

To quantitate the relative course of the two trajectories, we calculated the intertrajectory distances at equivalent time points. For the trajectories in the two-dimensional GATA1–PU.1 plane (Fig. 7B), the intertrajectory distance $D_{\text{int}}(t)$ calculated from the data in Fig. 5 remained minimal until 24 h when it takes off – consistent with the initial pre-bifurcation "wiggling" of GATA1 and PU.1 around the central state C of the progenitor state (Fig. 7B). Indeed, the N-dimensional, genome-scale intertrajectory distance between $S^E(t_i)$ and $S^M(t_i)$ for equivalent time points t_i, measured as estimated geodesic $G_{\text{int}}(t_i)$ distance (Fig. 7C) indicates that the initial divergence remains small and is even flat between 24 h and 48 h but then it increases rapidly. As expected, for short distances up to 16 h, the Euclidean distance $E_{\text{int}}(t)$ (dashed line) was identical to the estimated geodesic distance $G_{\text{int}}(t_i)$, since both measured the nearest neighbors in the manifold, but with increasing scale, $E_{\text{int}}(t)$ appeared to underestimate the intertrajectory distance compared to $G_{\text{int}}(t)$ which follows the complex structure ("curvatures") of the manifold (Fig. 7C). Both measures suggest that there is little to

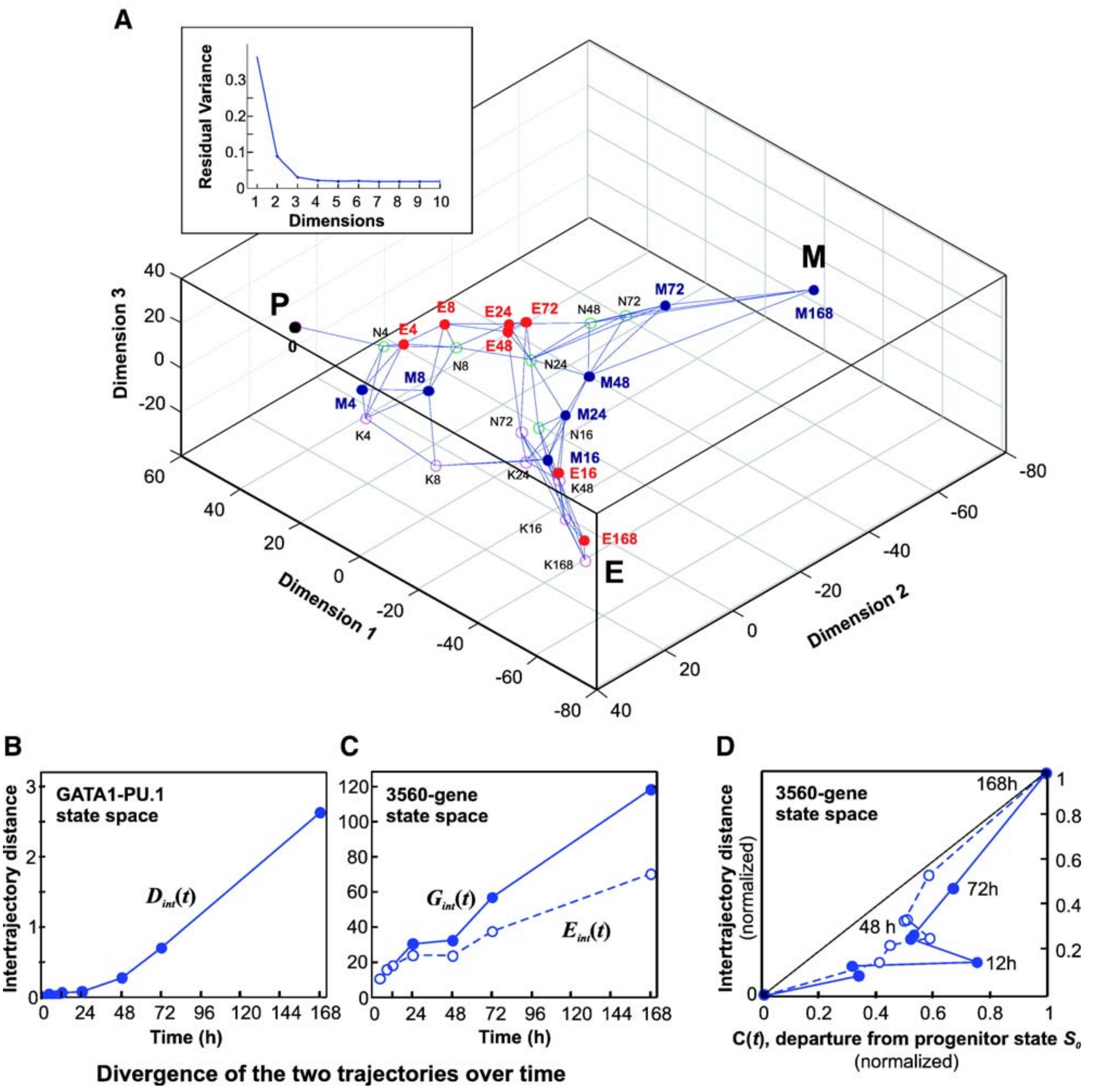

Fig. 7. Isomap representation of state space and the two differentiation trajectories. (A) Each microarray measurement is represented as a point in the three-dimensional embedding space. Position of points is calculated based on Isomap with nearest-neighbor parameter, $k=4$ (Tenenbaum et al., 2000). Solid red dots: points on trajectory S^E; solid blue dots: time points on trajectory S^M. Empty circles denote other trajectories measured by Bruno et al. (2004) and used for Isomap construction. Labels indicate differentiation trajectory and time after induction in hours: E=erythroids, M=myelomonocytes, K=megakaryocyte, N=neutrophils (no 168-h time point available for N). Thin lines connecting dots represent nearest neighbors defining the estimated manifold surface used for Isomap calculation. Units of axes correspond to high-dimensional distances in gene expression state space calculated from the relative expression signals ($\log_2$ of ratio, with 0 h value as reference). Inset: Residual variance of Isomap dimension reduction as in Tenenbaum et al. (2000). (B) Change in the "intertrajectory distance", $D_{\mathrm{int}}(t)$, between the erythroid and myeloid differentiation trajectories in the GATA1–PU.1 state space plane. (C) Intertrajectory distance in the 3560-gene state space, as estimated geodesic distance $G_{\mathrm{int}}(t)$ (solid line) or Euclidean distance $E_{\mathrm{int}}(t)$ (dashed line). (D) Comparison between state space distance traveled away from the progenitor state S_0, averaged for erythroid and myeloid differentiation (horizontal axis) and the intertrajectory distance (vertical axis) as geodesic or Euclidean distances, as in panel C, for each time point. Units in state space distance normalized for 1=distances at final time point of 168 h.

no increase of trajectory separation until 48 h, after which the intertrajectory distance drastically increases as expected from visual inspection of the GEDI maps (Fig. 6).

To account for noise that may inflate the intertrajetcory distance as the cells depart from the progenitor state $S_0(t=0)$ and the relative gene expression levels increase, we also compared the increase of the intertrajetcory distance with the average distance to the progenitor state S_0, $C(t)$, that the differentiating cells have moved away from. Fig. 7D shows that at any time point the normalized intertrajectory distance was smaller than the normalized distance $C(t)$ from the origin (curve stays below the diagonal in Fig. 7C). Initially, the departure from the progenitor state, $C(t)$, drastically exceeded the increase in the intertrajectory distance (flat part of curve, up to 48 h). After 48 h the intertrajectory distance increased more rapidly as cells begin to adopt the lineage-specific transcriptome. Thus the terminal increase of the intertrajetcory distance between the two differentiation processes was not simply due to the increased level of relative gene expression values.

Behavior of individual genes: split genes

To confirm that the predicted but counterintuitive non-monotonical time courses associated with the early changes of GATA1 and PU.1 (Fig. 5F) are a generic property of fate-specific genes in the network that are promiscuously expressed in the progenitor state and not an idiosyncratic behavior of GATA1 and PU.1, we analyzed the time courses of individual genes in the expression profiles. From the set of 347 genes previously reported (Bruno et al., 2004) to exhibit significant change of expression relative to the progenitor cells in all four lineages, we defined a subset using simple but conservative criteria: "split genes" are genes that are expressed in the undifferentiated FDCP-mix cells but whose expression changes in opposite directions ("split") as cells differentiate into the erythroid and myelomonocytic lineages, in that they are down-regulated in one and up-regulated in the other lineage. A gene was defined to be a split gene when the expression levels at the end of the two differentiation paths have changed in the opposite direction (increased and suppressed relative to the starting level) and differed by more than 1.74-fold: $x^M(172\text{ h})/x^E(172\text{ h})$ or $x^E(172\text{ h})/x^M(172\text{ h})=1.74$ ($=2^h$ where $h=0.8$ is the relative expression level as $\log_2$-ratio). With these criteria, 100 of the 347 genes were found to be "split genes". In other words, a considerable number of lineage-specific genes were detectable in progenitor cells at intermediate levels, corroborating the concept of "priming" of multipotent cells by promiscuous expression. Figs. 8A–C show the time courses of three groups of genes among the set of 347 genes: (A) 87 split genes whose expression is increased in myeloid lineage and decreased in the erythroid lineage, (B) 13 split genes with the opposite behavior, and (C) genes that moved in parallel (increased or decreased expression) in the two differentiation processes and whose average expression changed by more than 1.74-fold relative to that of the progenitor cell. Fig. 8 reveals that even as an ensemble, the myeloid-specific split genes displayed an initial transient "dip" of expression before their terminal increase during myeloid differentiation (blue curves), while their suppression in the erythroid lineage also showed a (less characteristic) transient increase early in the process – recapitulating the behavior of PU.1.

If we define a non-monotonical time course as one that exhibits a net change of gene expression during the early period (3 h–48 h) in the opposite direction to that of the late period (72 h–168 h), then the split genes were heavily enriched for genes that displayed non-monotonical behavior: 37% of genes among the split genes meet the above criteria for non-monotonical behavior versus only 3% in the group of non-split genes of group (C). This dramatic difference was not significantly influenced by the particular value h for defining split genes (almost identical results were obtained for range of $h=0.5$–1.0). Thus, the lineage-specific split genes do not simply "move" in opposite directions when cells differentiate into the two separate lineages but tend to initially and transiently follow the course they would take in the alternative differentiation path, as proposed by the GATA1–PU.1 circuit model, thereby contributing to the

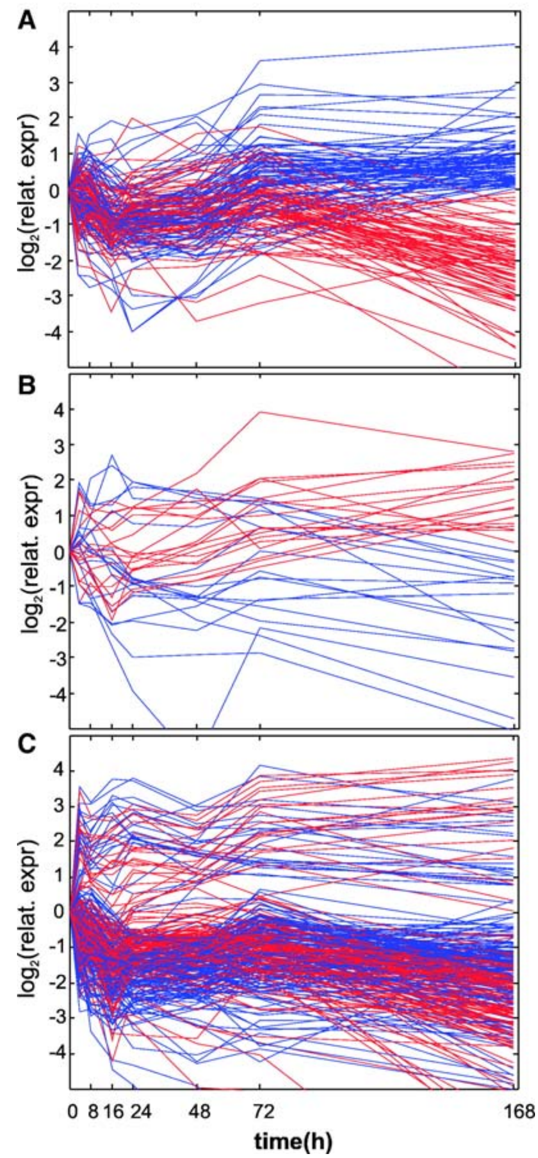

Fig. 8. Time course of expression levels of selected genes and non-monotonical course of "split genes". The set of 347 genes that change their expression significantly as described by Bruno et al. (2004) were non-exhaustively classified based on time course structure into three groups. Expression values are averages from experimental triplicates. For each gene, two curves are displayed, one for its expression behavior in the myelomonocytic (*M*) lineage (blue), and one for behavior when cell choose the erythroid lineage (red). (A) "Split genes" that increased their expression in the myelomonocytic lineage but are suppressed in the erythroid lineage (83 genes). (B) Split gene with the reversed behavior (17 genes). (C) Genes that change their expression during differentiation in both lineages in the same direction (21 genes increased their mRNA expression, 153 genes were suppressed).

apparently common initial stretch in the two high-dimensional differentiation trajectories.

Cell proliferation and cell death during lineage specification and differentiation

Our two-stage model for cell fate choice accounts for both deterministic and stochastic processes in cell fate decision. While the near-symmetric bifurcation and the closely parallel initial transcriptome trajectories make a purely deterministic regulation unlikely, they do not exclude the possibility that the quite sudden separation of the trajectories after 48 h is purely due to a selection in a stochastically heterogeneous cell population by the fate-specifying cytokines. In this scenario, the external factors would promote survival and amplification of the one subpopulation of cells that "happen" to be

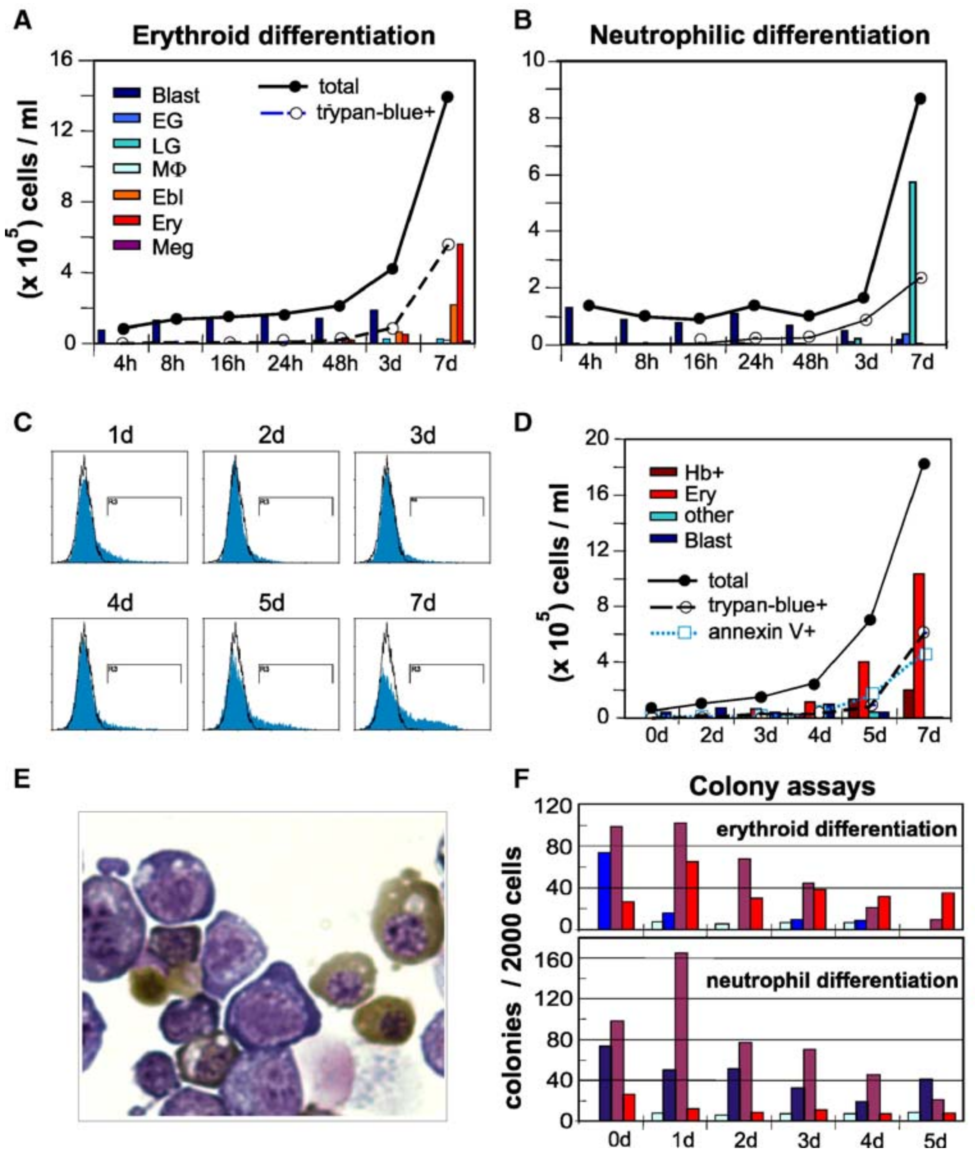

Fig. 9. Analysis of cell death, apoptosis and fates in differentiating of FDCP-mix cells. (A, B) Cell death during erythroid (A) and neutrophil (B) differentiation of FDCP-mix cells. Curves show total cell numbers and numbers of dead cells (assessed by trypan-blue staining) over the time course after induction of differentiation. Bars show distribution of morphologically distinguishable cell types (see legend in graph) at the indicated time points (h = hour, d = days). EG = early granulocytes, *MΦ* = macrophages, Ebl = erythroblasts, Ery = more mature erythroid stages, Meg = megakaryocyte. Cell numbers represent average of the three replicate cultures used for the microarray studies. (C, D) Apoptosis and cell death during erythroid differentiation. Erythroid cultures were sampled at daily intervals and analyzed for apoptosis by annexin V staining by flow cytometry (C) as well as for cell number and cell death (trypan-blue + cells) (D) and changes in morphology (E). (E) Photomicrograph of cells recovered at day 5 of cultures of FDCP-mix cells grown under erythroid differentiation supportive conditions (100 ng/ml SCF, 10 ng/ml IL-6, 10 U/ml Epo, 0.2 mM hemin). Slides were stained with May–Grunwald–Giemsa and counterstained with *o*-dianisidine to detect hemoglobin (Hb^+, hemoglobin-positive by *o*-dianisidine staining). (F) Differentiation of individual colonies. FDCP-mix cells were cultured for varying times as indicated (0–5 days) in suspension culture under erythroid or neutrophil differentiation conditions before plating in semisolid media that sustains all lineages (Methocult 3234 supplemented with 0.01 ng/ml IL-3, 100 ng/ml SCF and 10 U/ml hEpo). Colonies were counted and their morphology was determined (G = granulocyte colony-forming unit, GM = granulocyte–monocyte colony-forming unit, mix = mixed granulocyte–monocyte and erythroid colony-forming unit, Ery = erythroid colony-forming unit). Note that in the cases of no or short (0d, 1d) exposure to differentiating conditions prior to plating, some cells produce pure myeloid and erythroid colonies indicating intrinsic regulation.

"appropriately poised" in the state space [as discussed in Trajectory of differentiation in the GATA1–PU.1 phase space, (see section): (c)], accompanied by the death in the other subpopulation. Microarray-based gene expression profiling of a heterogeneous cell population cannot distinguish a shift in its cellular composition with respect to cell types – each having a distinct gene expression pattern – from a change in the gene expression pattern of every individual cell of the population (Chang et al., 2006).

To exclude the possibility that fate choice is purely due to cell selection, we measured cell proliferation and death during the early phases of cytokine-induced differentiation. A fate choice governed solely by selection would have to result in a substantial portion of the cells dying early after stimulation of differentiation into either one of the two alternative fates. Furthermore, since the two fates are mutually exclusive, the death rates in the two paths of differentiation should be complementary, adding up to approximately 100%. Conversely, if instruction substantially contributes to fate determination, a much lower rate of cell death might be predicted, since "inappropriately poised" cells would change their gene expression pattern in response to the cytokine rather than die. Figs. 9A and B show the total number of cells, their morphology and percentage of dead cells as judged by trypan blue staining at various time points in erythroid and neutrophil differentiation cultures of the FDCP-mix cells used for the gene profiling studies. As expected, there was some evidence of cell death at later stages of cultures when mature cells predominate. However, in both differentiation conditions, cell death was far below 50%. Importantly, at earlier time-points (<3 days) there was little evidence of cell death. In addition, in erythroid differentiation experiments we also monitored cell death through apoptosis by annexin V staining which reveals cells destined to die (Figs. 9C, D). Very few cells stained positive for annexin V and those that did appeared late, around day 3. This finding suggests that cell selection and shift in proportions of cell types is not the major contributor to the observed divergence in the gene expression patterns (Figs. 6 and 7). Success of differentiation in these experiments was confirmed by standard morphological analysis of individual cells and by colony formation assays (Figs. 9E, F). The presence of mixed colonies and the decrease of their number in favour of pure erythroid or myeloid colonies as exposure time to the respective fate-specifying conditions prior to plating was increased (Fig. 9F) are consistent with the combined contribution of intrinsic stochastic processes and the external bias by instructive inputs in cell fate commitment.

Discussion

Biological implications

We studied the integrated behavior of the GATA1–PU.1 mutual inhibition/auto-stimulation circuit that underlies binary cell fate decision using a minimal mathematical model based on well-established interactions. We characterized the generic dynamical constrains intrinsic to this architecture and found that it predicted a behavioral repertoire that in principle could explain a binary cell fate decision of a stable progenitor cell. Comparison of model predictions with observed changes in mRNA expression during differentiation of the bipotent progenitor cell line FDCP-mix into an erythroid lineage and a myelomonocytic lineage helped narrow down the alternative possibilities for how in the first place the circuit can implement a binary fate decision. The qualitative agreement between theoretical prediction and observed behavior supports two main biological principles:

First, the model suggests that the bipotent progenitor state is a metastable attractor state located between the two attractors that represent its prospective differentiation fates. This offers a formal explanation for the idea of "program accessibility" as a hall mark of multipotency. The model also explains the phenomenon of "promiscuous expression" of lineage-affiliated genes in progenitor cells (Cross and Enver, 1997). The mutual inhibition/auto-stimulation circuit (Fig. 2D) that generates the tristable dynamics may represent a general network motif that governs discrete cell fate decisions in multipotent cells. Many key transcription factors involved in cell fate control are in fact also engaged in circuits consisting of mutual regulation and auto-stimulation (Orkin, 2000). Notably, at the level of embryonic stem cells, the transcription factors NANOG, OCT4 and SOX2 which play crucial roles in the earliest fate decisions also mutually regulate each other as well as themselves (Boyer et al., 2005; Ralston and Rossant, 2005).

Second, the process of diversification of a multipotent cell into two distinct fates in response to a differentiation signal is not simply the gradual departure of the transcriptome from that of the progenitor cell to progressively adopt the distinct transcriptome of the respective mature cells. Instead, both model and gene expression data suggest that one can conceptually distinguish between two stages for the commitment process (scenario *ii* in Table 1): despite the distinct signals that induce differentiation into the two different cell types, the two transcriptomes initially change globally in a very similar fashion (Figs. 6 and 7). This may represent a common destabilization process over a period of 24–48 h that allows cells to escape the progenitor attractor and is reflected in the initial changes in GATA1 and PU.1 that remain in a localized state space region (Fig. 7B). Then, in a second stage the separation of the trajectories towards the destination transcriptomes occurs quite abruptly, perhaps reflecting the bifurcation event at which the central progenitor attractor disappears. In fact, attempting to "switch" lineage, e.g. from the erythroid to the myeloid lineage, by changing the medium conditions at varying time points after differentiation into one lineage has been initiated, indicates that the point of commitment is at around 24–48 h (T. E., unpublished observation).

Metaphorically, the destabilization and disappearance of the progenitor attractor can be viewed as S being placed on a "watershed" region in Waddington's epigenetic landscape (Waddington, 1957) where it can easily be "tipped" into *either* side to the now easily accessible attractors of the two prospective lineages by small, deterministic perturbations or by random fluctuations in molecular activities, to reliably produce distinct and specific outcomes. This near-symmetric bifurcation model thus is consistent with the ample evidence for the observed stochasticity in fate determination (Enver et al., 1998). Yet, the separation of the trajectories towards final commitment is not purely due to selection for cells that are (randomly) poised to be responsive to the survival promoting activity of the fate-specifying cytokine, since only a small fraction of cells die at this stage of commitment (Fig. 9). Therefore, cell-intrinsic shifts of gene expression patterns caused by instructive inputs of the cytokines also contribute to the changes of gene expression profile trajectories observed at the cell population level. Such instruction may act by introducing asymmetry in the bifurcation ("tilting the watershed") during the initial destabilization of the progenitor state to harness and bias the stochastic processes. Our model thus naturally unites the stochastic (intrinsic) and deterministic (instructive) mechanisms of cell fate regulation (Enver et al., 1998).

More generally, the "watershed" metaphor explains the observation that many unspecific (hence, non-instructive) signals, such as solvents or mechanical forces, can cause differentiation in many cell systems: they may do so by "tipping" cells into a predefined program (Graf, 2002; Huang and Ingber, 2000; Hume, 2000) but it also allows for the action of specific deterministic signals to steer cells into one of the two "valleys". Moreover, the (hidden) coexistence of the fate-specific attractors with that of the progenitor state as proposed by the tristable behavior and the type II bifurcation (Fig. 4B) may explain why in the absence of culture conditions that prevent differentiation, pluripotent embryonic stem cells or multipotent progenitor cells can spontaneously and apparently stochastically undergo differentiation into a limited set of distinct, predefined cell types within the very same environment (Fairbairn et al., 1993; Keller, 1995). Such intrinsic diversifying behavior cannot be explained by a purely instructive, asymmetric disruption of the progenitor state (scenario *i* in Table 1).

The model describes a progenitor cell that can commit to either the megakaryocyte–erythroid progenitors (MEP; '*E*' in the model), or the granulocyte–monocyte progenitors (GMP; '*M*' in the model) and hence, it corresponds to a common myeloid progenitor cell (CMP) (Akashi et al., 2000). However, this classical scheme of fate paths appears to be rather simplistic. Until recently, this fate decision was thought to occur exclusively at the level of the CMP cell which gives rise to MEP and GMP cells which then produce the mature erythroid and megakaryoctic, or the granulocytic and monocytic cells, respectively. While the existence of the CMP is not in question, recent observations (Adolfsson et al., 2005) suggest that hematopoietic stem cells may directly give rise to MEP progenitors and produce a novel population which possesses lymphoid and myeloid potential but lacks megakaryocytic/erythroid differentiation capacity, thereby producing an alternative road map for hematopoietic differentiation. Future models that incorporate a more complex molecular circuitry may explain these newly described "non-classical" cell fate paths.

Limits of modeling

Although the model presented here is still qualitative, we moved a step beyond the current non-formal, metaphoric concepts, and predicted the existence of a metastable attractor state with a promiscuous expression pattern that represents the uncommitted progenitor cells. This model also captures the idea of "stable suspension" for explaining multipotency (Mikkers and Frisen, 2005). Similar mathematical models for diversifying fate decisions at a more fine-grained (molecular) level of description with similar conclusions have recently been proposed (Chickarmane et al., 2006; Roeder and Glauche, 2006).

Caution is warranted when modeling complex living systems, which have evolved a class of architecture that is inherently robust (Goodwin et al., 1993; Kauffman, 1993) and in which the same or similar system behavior can be produced by a large ensemble of distinct architectures (Bagley and Glass, 1996; Huang, 2005; Tsong et al., 2006). As a consequence, theoretical models of complex living systems, even those that coarsely represent the essential traits, may also harbor this fundamental "robustness" and hence, tolerate inaccuracies in modeling assumptions. In particular, they may contain inherently "soft" parameters. This makes it difficult to falsify model predictions by comparing them to the experimental observation of time courses. To avoid this problem of under-determination and sloppiness of parameters (Brown et al., 2004) as well as the bias associated with the desire to fit model to data, we relied on well-characterized phenotypic influences and only sought to gain general insight on the behavioral repertoire of a minimal system by exploring the entire state space and parameter space. The exhaustive simulation presented two alternative, mutually exclusive, qualitatively distinct behaviors to choose from (Fig. 4) for comparison with experimentally observed state space trajectories. Moreover, the latter involved the entire set of modeled variables, further avoiding fortuitous curve-fitting to a subset of variables due to sloppy parameters (Brown et al., 2004). Nevertheless, although the observed trajectory structure consists of multiple (serial) measurements in both modeled dimensions and was seen in three independent time courses, more measurements in more diverse experimental systems, including cell lines with genetically altered circuits, will be necessary to further evaluate and refine the model.

Our model of fate decision rests on the central idea that transition from one attractor state (P) to another (E or M) can be viewed as a rapid change in the parameters of the system equations that in turn imposes a change in the structure of the attractor landscape, with a bifurcation as a special case. However, an alternative view is that the system state S "jumps" from one attractor to another while the attractor landscape is fixed (Bar-Yam and Epstein, 2004; Huang, 2002; Kauffman, 1993). Noise-driven exit from an attractor can also be modeled in this paradigm (Aurell and Sneppen, 2002; Hasty et al., 2000; Walczak et al., 2005). For regulated transition between high-dimensional attractors the external signal can then be viewed as abruptly causing changes in the expression values of an appropriate set of genes (determined by the inducing cytokine) to "reset" the position of S so that it is displaced across the attractor boundaries into a new basin of attraction. This model requires that the E and M attractors coexist with the central attractor P, and thus, would necessarily favor the tristable attractor landscape that is the basis for the *type II* bifurcation (Figs. 4B, D).

In fact, a central simplification in our model is that it uses deterministic variables that represent macroscopic changes in the averaged activity of the transcription factors in an entire cell population, ignoring cell-to-cell variability in the activity values of PU.1 and GATA1 due to both temporal gene expression noise and stable epigenetic heterogeneity. Thus, the model also discards "noise-driven" state transitions that occur without change in the attractor landscape (Aurell and Sneppen, 2002; Hasty et al., 2000). However, microarrays measure mRNAs levels as cell population averages, thus both model and experiment suffer the equivalent coarsening. Single-cell observations of the differentiation process (Warren et al., 2006) will be needed to examine noise-driven fate determination and discontinuous "jumping" between attractor states (Chang et al., 2006).

Another fundamental limitation of modeling gene circuits is that a circuit motif is not necessarily a functionally independent module. In reality the GATA1–PU.1 is embedded in an almost genome-wide network of regulatory interactions whose inputs onto the GATA1–PU.1 circuit were not considered (Swiers et al., 2006). This shortcoming can be accentuated by attempts to model the dynamics of the system at the level of molecular mechanisms within the circuit while neglecting molecular inputs from outside. In fact, many factors involved in this fate decision also engage in mutual inhibition relationships with GATA1 and PU.1, including FOG, C/EBP, c-Jun, c-Myb, GATA2, etc. (Ferreira et al., 2005; Friedman, 2002; Morceau et al., 2004; Rosmarin et al., 2005). Thus, future models need to describe a higher-dimensional state space. Nevertheless, it appears that the dynamic properties of local circuits of 2–3 genes considered in isolation can have biological relevance (Mangan and Alon, 2003; Prill et al., 2005).

The molecular mechanisms downstream of receptor signaling that embody the parameter changes in a, b, θ and k are not known. The central question remains unanswered as to how lineage-determining cytokines, such as Epo and G/GM-CSF, contribute to the initial symmetric destabilization of the progenitor cell, and how they bend the trajectory towards a particular fate. This question has been difficult to address experimentally perhaps precisely because small differences in a broad intracellular signaling response may suffice to tilt fate commitment in the destabilized cell; in fact, post receptor processes of cytokine signaling are pleiotropic and highly overlapping (Fambrough et al., 1999; Kratchmarova et al., 2005). Concretely, the initial, correlated changes of gene expression patterns in the first 24–48 h of differentiation in both paths may in part be due to the reduction of IL-3 (used to maintain the multipotent state) and in part be caused by (non lineage-specific) overlapping effects of the cytokines Epo and G/GM-CSF. These changes may coincide with general shifts in cell metabolism and physiology, such as exit from cell cycle, which can either be a consequence of differentiation or contribute to it, or both.

Beyond these difficulties in understanding how cytokines and growth factors normally modulate genetic circuits, our model offers the possibility to compare predictions with the trajectory course following a direct (nonphysiological) manipulation of the circuit. For instance, instead of using culture conditions, cell fates can be controlled by GATA1 and PU.1 overexpression (Heyworth et al., 2002; Laslo et al., 2006; Xie et al., 2004). It would be interesting to monitor the transcriptome trajectories for erythroid and myeloid differentiation following overexpression of GATA1 and PU.1, respectively. Our model predicts that such artificial cell fate control may correspond to either the asymmetric case or a "jump" in state space (Table 1 (A) *i*, or (B)), hence, overcoming the initial phase of near-symmetric transcriptome change.

In summary, our model offers a general principle for fate decision that is in agreement with distinct observed qualitative behaviors. Further elucidation of additional proteins involved in the circuitry is necessary to address the question of what are the entry points by which the "instructive" effects of regulators destabilize the progenitor state and break the symmetry to induce commitment. The general principles stimulated by the simple model presented here may provide a conceptual framework that can guide experimental characterization of the molecular basis of cell fate decision.

Acknowledgments

This work was supported by grants from the Air Force Office for Scientific Research (S.H.) and the Medical Research Council and Leukaemia Research Fund of Great Britain (T.E.). We would like to thank Shamit Soneji for help in microarray data analysis and critical discussions and Hannah Chang for discussions and reading the manuscript.

References

Adolfsson, J., Mansson, R., Buza-Vidas, N., Hultquist, A., Liuba, K., Jensen, C.T., Bryder, D., Yang, L., Borge, O.J., Thoren, L.A., Anderson, K., Sitnicka, E., Sasaki, Y., Sigvardsson, M., Jacobsen, S.E., 2005. Identification of Flt3+ lympho-myeloid stem cells lacking erythro-megakaryocytic potential a revised road map for adult blood lineage commitment. Cell 121, 295–306.

Akashi, K., Traver, D., Miyamoto, T., Weissman, I.L., 2000. A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. Nature 404, 193–197.

Akashi, K., He, X., Chen, J., Iwasaki, H., Niu, C., Steenhard, B., Zhang, J., Haug, J., Li, L., 2003. Transcriptional accessibility for genes of multiple tissues and hematopoietic lineages is hierarchically controlled during early hematopoiesis. Blood 101, 383–389.

Aurell, E., Sneppen, K., 2002. Epigenetics as a first exit problem. Phys. Rev. Lett. 88, 048101.

Bagley, R.J., Glass, L., 1996. Counting and classifying attractors in high dimensional dynamical systems. J. Theor. Biol. 183, 269–284.

Bar-Yam, Y., Epstein, I.R., 2004. Response of complex networks to stimuli. Proc. Natl. Acad. Sci. U. S. A. 101, 4341–4345.

Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., Gifford, D.K., Melton, D.A., Jaenisch, R., Young, R.A., 2005. Core transcriptional regulatory circuitry in human embryonic stem cells. Cell 122, 947–956.

Brown, K.S., Hill, C.C., Calero, G.A., Myers, C.R., Lee, K.H., Sethna, J.P., Cerione, R.A., 2004. The statistical mechanics of complex signaling networks: nerve growth factor signaling. Phys. Biol. 1, 184–195.

Bruno, L., Hoffmann, R., McBlane, F., Brown, J., Gupta, R., Joshi, C., Pearson, S., Seidl, T., Heyworth, C., Enver, T., 2004. Molecular signatures of self-renewal, differentiation, and lineage choice in multipotential hemopoietic progenitor cells in vitro. Mol. Cell Biol. 24, 741–756.

Cantor, A.B., Orkin, S.H., 2001. Hematopoietic development: a balancing act. Curr. Opin. Genet. Dev. 11, 513–519.

Cantor, A.B., Orkin, S.H., 2002. Transcriptional regulation of erythropoiesis: an affair involving multiple partners. Oncogene 21, 3368–3376.

Chang, H.H., Oh, P.Y., Ingber, D.E., Huang, S., 2006. Multistable and multistep dynamics in neutrophil differentiation. BMC Cell Biol. 7, 11.

Chickarmane, V., Troein, C., Nuber, U.A., Sauro, H.M., Peterson, C., 2006. Transcriptional dynamics of the embryonic stem cell switch. PLoS Comput. Biol. 2, e123.

Cinquin, O., Demongeot, J., 2005. High-dimensional switches and the modelling of cellular differentiation. J. Theor. Biol. 233, 391–411.

Cox, T.F., Cox, M.A.A., 1994. Multidimensional Scaling. Chapman and Hall, London.

Cross, M.A., Enver, T., 1997. The lineage commitment of haemopoietic progenitor cells. Curr. Opin. Genet. Dev. 7, 609–613.

Davidson, E.H., Erwin, D.H., 2006. Gene regulatory networks and the evolution of animal body plans. Science 311, 796–800.

Davidson, E.H., Rast, J.P., Oliveri, P., Ransick, A., Calestani, C., Yuh, C.H., Minokawa, T., Amore, G., Hinman, V., Arenas-Mena, C., Otim, O., Brown, C.T., Livi, C.B., Lee, P.Y., Revilla, R., Rust, A.G., Pan, Z., Schilstra, M.J., Clarke, P.J., Arnone, M.I., Rowen, L., Cameron, R.A., McClay, D.R., Hood, L., Bolouri, H., 2002. A genomic regulatory network for development. Science 295, 1669–1678.

Eichler, G.S., Huang, S., Ingber, D.E., 2003. Gene Expression Dynamics Inspector (GEDI): for integrative analysis of expression profiles. Bioinformatics 19, 2321–2322.

Enver, T., Heyworth, C.M., Dexter, T.M., 1998. Do stem cells play dice? Blood 92, 348–351 discussion 352.

Fairbairn, L.J., Cowling, G.J., Reipert, B.M., Dexter, T.M., 1993. Suppression of apoptosis allows differentiation and development of a multipotent hemopoietic cell line in the absence of added growth factors. Cell 74, 823–832.

Fambrough, D., McClure, K., Kazlauskas, A., Lander, E.S., 1999. Diverse signaling pathways activated by growth factor receptors induce broadly overlapping, rather than independent, sets of genes. Cell 97, 727–741.

Ferreira, R., Ohneda, K., Yamamoto, M., Philipsen, S., 2005. GATA1 function, a paradigm for transcription factors in hematopoiesis. Mol. Cell Biol. 25, 1215–1227.

Fisher, A.G., 2002. Cellular identity and lineage choice. Nat. Rev., Immunol. 2, 977–982.

Friedman, A.D., 2002. Transcriptional regulation of granulocyte and monocyte development. Oncogene 21, 3377–3390.

Galloway, J.L., Wingert, R.A., Thisse, C., Thisse, B., Zon, L.I., 2005. Loss of *gata1* but not *gata2* converts erythropoiesis to myelopoiesis in zebrafish embryos. Dev. Cell 8, 109–116.

Gardner, T.S., Cantor, C.R., Collins, J.J., 2000. Construction of a genetic toggle switch in *Escherichia coli*. Nature 403, 339–342.

Georgopoulos, K., 2002. Haematopoietic cell-fate decisions, chromatin regulation and ikaros. Nat. Rev., Immunol. 2, 162–174.

Glass, L., Hill, C.C., 1998. Ordered and disordered dynamics in random networks. Europhys. Lett. 41, 599–604.

Glass, L., Kauffman, S.A., 1973. The logical analysis of continuous, non-linear biochemical control networks. J. Theor. Biol. 39, 103–129.

Goodwin, B.C., Kauffman, S., Murray, J.D., 1993. Is morphogenesis an intrinsically robust process? J. Theor. Biol. 163, 135–144.

Graf, T., 2002. Differentiation plasticity of hematopoietic cells. Blood 99, 3089–3101.

Hasty, J., Pradines, J., Dolnik, M., Collins, J.J., 2000. Noise-based switches and amplifiers for gene expression. Proc. Natl. Acad. Sci. U. S. A. 97, 2075–2080.

Heyworth, C., Pearson, S., May, G., Enver, T., 2002. Transcription factor-mediated lineage switching reveals plasticity in primary committed progenitor cells. EMBO J. 21, 3770–3781.

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 Dev. 11, 774–785.

Huang, S., 2002. Regulation of cellular states in mammalian cells from a genome-wide view. Gene Regulation and Metabolism: Post-Genomic Computational Approach. MIT Press, Cambridge, MA, pp. 181–220.

Huang, S., 2004. Back to the biology in systems biology: what can we learn from biomolecular networks. Brief. Funct. Genomics Proteomics 2, 279–297.

Huang, S., 2005. Multistability and multicellularity: cell fates as high-dimensional attractors of gene regulatory networks. In: Kriete, A., Eils, R. (Eds.), Computational Systems Biology. Elsevier Academic Press.

Huang, S., Ingber, D.E., 2000. Shape-dependent control of cell growth, differentiation, and apoptosis: switching between attractors in cell regulatory networks. Exp. Cell Res. 261, 91–103.

Huang, S., Eichler, G., Bar-Yam, Y., Ingber, D.E., 2005. Cell fates as high-dimensional attractor states of a complex gene regulatory network. Phys. Rev. Lett. 94, 128701.

Hume, D.A., 2000. Probability in transcriptional regulation and its implications for leukocyte differentiation and inducible gene expression. Blood 96, 2323–2328.

Joshi, C., Enver, T., 2003. Molecular complexities of stem cells. Curr. Opin. Hematol. 10, 220–228.

Kaern, M., Elston, T.C., Blake, W.J., Collins, J.J., 2005. Stochasticity in gene expression: from theories to phenotypes. Nat. Rev. Genet. 6, 451–464.

Kaplan, D., Glass, L., 1995. Understanding Nonlinear Dynamics. Springer, New York.

Kauffman, S.A., 1969. Metabolic stability and epigenesis in randomly constructed genetic nets. J. Theor. Biol. 22, 437–467.

Kauffman, S.A., 1993. The Origins of Order. Oxford Univ. Press, New York.

Keller, G.M., 1995. In vitro differentiation of embryonic stem cells. Curr. Opin. Cell. Biol. 7, 862–869.

Kim, C.F., Jackson, E.L., Woolfenden, A.E., Lawrence, S., Babar, I., Vogel, S., Crowley, D., Bronson, R.T., Jacks, T., 2005. Identification of bronchioalveolar stem cells in normal lung and lung cancer. Cell 121, 823–835.

Kluger, Y., Lian, Z., Zhang, X., Newburger, P.E., Weissman, S.M., 2004. A panorama of lineage-specific transcription in hematopoiesis. Bioessays 26, 1276–1287.

Kratchmarova, I., Blagoev, B., Haack-Sorensen, M., Kassem, M., Mann, M., 2005. Mechanism of divergent growth factor effects in mesenchymal stem cell differentiation. Science 308, 1472–1477.

Laslo, P., Spooner, C.J., Warmflash, A., Lancki, D.W., Lee, H.J., Sciammas, R., Gantner, B.N., Dinner, A.R., Singh, H., 2006. Multilineage transcriptional priming and determination of alternate hematopoietic cell fates. Cell 126, 755–766.

Mangan, S., Alon, U., 2003. Structure and function of the feed-forward loop network motif. Proc. Natl. Acad. Sci. U. S. A. 100, 11980–11985.

Mikkers, H., Frisen, J., 2005. Deconstructing stemness. EMBO J. 24, 2715–2719.

Miyamoto, T., Iwasaki, H., Reizis, B., Ye, M., Graf, T., Weissman, I.L., Akashi, K., 2002. Myeloid or lymphoid promiscuity as a critical step in hematopoietic lineage commitment. Dev. Cell 3, 137–147.

Morceau, F., Schnekenburger, M., Dicato, M., Diederich, M., 2004. GATA-1: friends, brothers, and coworkers. Ann. N. Y. Acad. Sci. 1030, 537–554.

Morrison, S.J., Uchida, N., Weissman, I.L., 1995. The biology of hematopoietic stem cells. Annu. Rev. Cell Dev. Biol. 11, 35–71.

Okuno, Y., Huang, G., Rosenbauer, F., Evans, E.K., Radomska, H.S., Iwasaki, H., Akashi, K., Moreau-Gachelin, F., Li, Y., Zhang, P., Gottgens, B., Tenen, D.G., 2005. Potential autoregulation of transcription factor PU.1 by an upstream regulatory element. Mol. Cell Biol. 25, 2832–2845.

Orkin, S.H., 2000. Diversification of haematopoietic stem cells to specific lineages. Nat. Rev. Genet. 1, 57–64.

Phillips, R.L., Ernst, R.E., Brunk, B., Ivanova, N., Mahan, M.A., Deanehan, J.K., Moore, K.A., Overton, G.C., Lemischka, I.R., 2000. The genetic program of hematopoietic stem cells. Science 288, 1635–1640.

Prill, R.J., Iglesias, P.A., Levchenko, A., 2005. Dynamic properties of network motifs contribute to biological network organization. PLoS Biol. 3, e343.

Ralston, A., Rossant, J., 2005. Genetic regulation of stem cell origins in the mouse embryo. Clin Genet 68, 106–112.

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. J. Theor. Biol. 241, 852–865.

Rosmarin, A.G., Yang, Z., Resendes, K.K., 2005. Transcriptional regulation in myelopoiesis: hematopoietic fate choice, myeloid differentiation, and leukemogenesis. Exp. Hematol. 33, 131–143.

Savageau, M.A., 1995. Michaelis–Menten mechanism reconsidered: implications of fractal kinetics. J. Theor. Biol. 176, 115–124.

Spudich, J.L., Koshland Jr., D.E., 1976. Non-genetic individuality: chance in the single cell. Nature 262, 467–471.

Swiers, G., Patient, R., Loose, M., 2006. Genetic regulatory networks programming haematopoietic stem cells and erythroid lineage specification. Dev Biol. 294, 525–540.

Tenenbaum, J.B., de Silva, V., Langford, J.C., 2000. A global geometric framework for nonlinear dimensionality reduction. Science 290, 2319–2323.

Thomas, R., 1998. Laws for the dynamics of regulatory networks. Int. J. Dev. Biol. 42, 479–485.

Tsai, S.F., Strauss, E., Orkin, S.H., 1991. Functional analysis and in vivo footprinting implicate the erythroid transcription factor GATA-1 as a positive regulator of its own promoter. Genes Dev. 5, 919–931.

Tsong, A.E., Tuch, B.B., Li, H., Johnson, A.D., 2006. Evolution of alternative transcriptional circuits with identical logic. Nature 443, 415–420.

Waddington, C.H., 1956. Principles of Embryology. Allen and Unwin Ltd., London, p. 351.

Walczak, A.M., Onuchic, J.N., Wolynes, P.G., 2005. Absolute rate theories of epigenetic stability. Proc. Natl. Acad. Sci. U. S. A. 102, 18926–18931.

Warren, L., Bryder, D., Weissman, I.L., Quake, S.R., 2006. Transcription factor profiling in individual hematopoietic progenitors by digital RT–PCR. Proc. Natl. Acad. Sci. U. S. A. 103, 17807–17812.

Wolpert, L., 1994. Do we understand development? Science 266, 571–572.

Xie, H., Ye, M., Feng, R., Graf, T., 2004. Stepwise reprogramming of B cells into macrophages. Cell 117, 663–676.

Xiong, W., Ferrell Jr., J.E., 2003. A positive-feedback-based bistable 'memory module' that governs a cell fate decision. Nature 426, 460–465.