

# SCIENTIFIC REPORTS



OPEN

## Reprogramming, oscillations and transdifferentiation in epigenetic landscapes

Bivash Kaity<sup>1</sup>, Ratan Sarkar<sup>2</sup>, Buddhapriya Chakrabarti<sup>3</sup> & Mithun K. Mitra<sup>1</sup> 

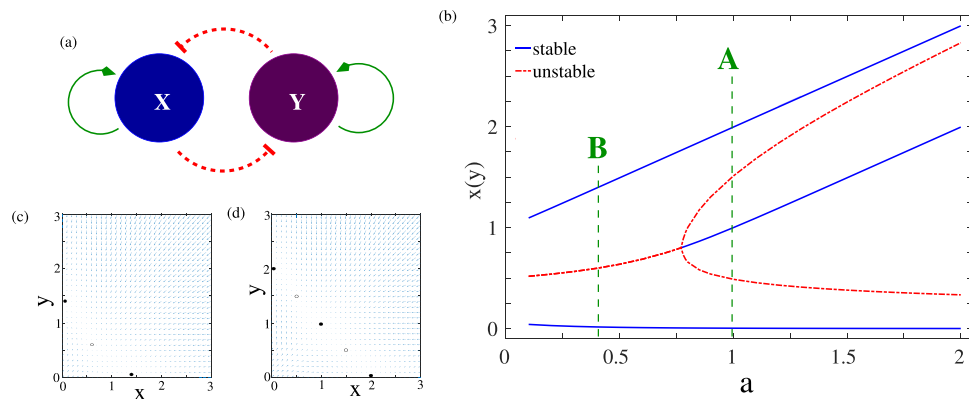
Waddington's epigenetic landscape provides a phenomenological understanding of the cell differentiation pathways from the pluripotent to mature lineage-committed cell lines. In light of recent successes in the reverse programming process there has been significant interest in quantifying the underlying landscape picture through the mathematics of gene regulatory networks. We investigate the role of time delays arising from multi-step chemical reactions and epigenetic rearrangement on the cell differentiation landscape for a realistic two-gene regulatory network, consisting of self-promoting and mutually inhibiting genes. Our work provides the first theoretical basis of the transdifferentiation process in the presence of delays, where one differentiated cell type can transition to another directly without passing through the undifferentiated state. Additionally, the interplay of time-delayed feedback and a time dependent chemical drive leads to long-lived oscillatory states in appropriate parameter regimes. This work emphasizes the important role played by time-delayed feedback loops in gene regulatory circuits and provides a framework for the characterization of epigenetic landscapes.

The rise of transcription factor mediated induced pluripotency<sup>1–6</sup> has given rise to increased efforts in order to understand the epigenetic landscape that underpins the process of cellular differentiation. Yamanaka and co-authors have shown that four transcription factors (TF), Sox2, Oct4, Klf4 and c-Myc are enough to reprogram fully differentiated cells to the pluripotent state, in mouse fibroblast cells<sup>1</sup>, as well as in human fibroblasts<sup>2</sup>. A different program, using a combination of the TFs Sox2, Oct4, NANOG, and LIN28 has also been shown to induce pluripotency in fully differentiated human cells<sup>3–5</sup>.

The process of cellular differentiation has traditionally been thought of in terms of Waddington's epigenetic landscape<sup>7</sup>, using the metaphor of a ball rolling down a hill, with the final valley representing the differentiated states, the top of the hill representing the pluripotent state, and the valley chosen by a cell from among similar ones as cell fate decisions. The discovery of the reprogramming pathway has led to an interest in mathematically modeling the epigenetic landscape through the underlying Gene Regulatory Networks (GRN). Initial studies using a toy single gene regulatory network studied the properties of cellular differentiation using a self-activating gene<sup>8–12</sup>. Later studies have expanded this work to model more realistic two gene networks, that are self-activating and mutually inhibiting<sup>11,13–17</sup>. These models have shown that the cell can follow different pathways during the differentiation and reverse programming pathways, and provided a more nuanced understanding of the epigenetic landscape<sup>11,13–16</sup>.

An experimental attempt to elucidate the reverse programming pathway was made by Nagy and Nagy<sup>18</sup>, in which they attempted to clarify the pathway of a cell undergoing reverse differentiation through the use of a controlled time-dependent chemical drive. Somatic fibroblast cells derived from mouse iPSC have the four pluripotency factors (Sox2, Oct4, Klf4 and c-Myc) incorporated in a latent state. These were activated through a single chemical factor, doxycycline, which was supplied for different durations. They observed that when the doxycycline input was provided for a time of less than  $\sim 7$  days ( $d_{PNR}$ ), the cell stayed in the fully differentiated somatic state. On the other hand, on providing the doxycycline input for a period greater than  $\sim 14$  days ( $d_{CPS}$ ), the cell made a transition to the induced pluripotent state. For a chemical drive between  $d_{PNR}$  and  $d_{CPS}$ , the system was neither in the initial state, nor did it reach the final pluripotent state, and was stuck in a new uncharacterized state, referred to as the "Area 51" state. This provided concrete evidence that the path followed by the cell on the reverse differentiation pathway is not necessarily the same as that during the forward differentiation process.

<sup>1</sup>IIT Bombay, Department of Physics, Mumbai, 400076, India. <sup>2</sup>Indian Institute of Science, Centre for High Energy Physics, Bangalore, 560012, India. <sup>3</sup>University of Sheffield, Department of Physics and Astronomy, Sheffield, S3 7RH, UK. Correspondence and requests for materials should be addressed to M.K.M. (email: [mithun@phy.iitb.ac.in](mailto:mithun@phy.iitb.ac.in))



**Figure 1.** (a) A schematic of gene regulatory network. The two transcription factors X and Y are self-activating, and mutually inhibiting; (b) The bifurcation diagram of the non-delayed system of equations Eq. 1. The dashed green lines correspond to the two values of the feedback parameter  $a$  that were analysed in subsequent section. The other parameters were fixed at  $a_0 = 0.5$ ,  $b = 1.0$ ,  $k = 1.0$ ,  $n = 4$ , and  $S = 0.5$ . The velocity vector plots in the  $(x, y)$  plane for these two points are shown in (c) Regime B and (d) Regime A.

In a previous work<sup>12</sup>, we argued that modeling of the underlying gene regulatory network requires careful consideration of the time delays associated with the feedback regulation of the transcription factors. The reverse differentiation process can take place over a duration of weeks, and is accompanied by a host of changes in the epigenetic markers that characterize the state of the cell, such as changes in histone protein expression levels<sup>19,20</sup>, as well as changes in the chromatin compaction characteristics<sup>21–29</sup>. The timescales associated with these physical changes would impact the feedback regulation, and could have a critical impact in assessing the differentiation pathways. We demonstrated a proof-of-concept of this time-delayed feedback regulation in the context of a single-gene time evolution, where we showed that a competition between the delay timescale and the timescale of the chemical drive can set up long-lived oscillations in the concentration of the transcription factor and interpreted this as a possible explanation of the cells caught in limbo during the differentiation process.

In the current work, we present an analysis of the interplay of time-delayed feedback and a time-dependent chemical drive on a more realistic two-gene regulatory network. The underlying motif of the GRN consists of two self-activating and mutually inhibiting transcription factors, which is an extremely common motif observed in biological GRNs<sup>11,13–17</sup>. Examples include the differentiation pathway of the Common Myeloid Progenitor (CMP) cell which differentiates into the Myeloid and Erythroid cell fates depending on the over-expression of the TF PU.1 or GATA1 respectively. The myeloid and Erythroid cell lines differentiate further to produce the different cell fates in the myeloid lineage<sup>30–34</sup>. While a full analysis of the reverse differentiation pathway involves multiple genes mediated by multiple TFs, we restrict our analysis to this smaller subset of two-gene mediated regulatory networks. We show that starting from a differentiated state, it is possible to follow the reverse pathway to reach the undifferentiated state. Depending on the interplay between the chemical drive and the delayed feedback, we observe long-lived oscillatory states in this two-TF GRN, which might provide an explanation of the uncharacterized states observed in reverse differentiation experiments<sup>18</sup>. We also show that this same framework can provide a theoretical basis for the phenomenon of transdifferentiation, where one differentiated cell can switch to a different differentiated cell fate without passing through the undifferentiated or pluripotent state. To the best of our knowledge, this is the first study that provides a theoretical underpinning to the transdifferentiation process, which is well known in experimental contexts<sup>35–42</sup>.

In subsequent sections, we first define the differential equations that govern the time evolution of the TFs in the presence of a time-dependent chemical drive and time-delayed feedback. We then present results in different parameter regimes, which show oscillatory states and transdifferentiation events. Further, we characterize the phase space of the model in terms of the drive time and the delay timescale as well as characterize the oscillations observed in greater detail. Finally, we discuss the relevance of the current model in the context of generating a more detailed picture of epigenetic landscapes.

## Results

Two-gene networks have been studied in the literature as a model system to investigate the properties of epigenetic landscapes. The most common motif in a two gene network is where the transcription factor corresponding to each gene up-regulates its own production and down-regulates the production of the transcription factor associated with the other gene (Fig. 1a). It is possible to write down differential equations governing the time evolutions of the concentrations of the two transcription factors,

$$\begin{aligned}\frac{dx}{dt} &= a_1 \frac{x^n}{S^n + x^n} + b_1 \frac{S^n}{S^n + y^n} - kx \\ \frac{dy}{dt} &= a_2 \frac{y^n}{S^n + y^n} + b_2 \frac{S^n}{S^n + x^n} - ky\end{aligned}\quad (1)$$

where  $x$  and  $y$  denote the concentrations of the two transcription factors<sup>11,13</sup>. The feedback terms are modeled as Hill functions<sup>11</sup>, with exponent  $n$  controlling the steepness of the switch curve, while the parameter  $S$  controls the concentration at which the half-maximal point is reached. The first term corresponds to a positive feedback mechanism, with each TF up-regulating its own production, while the second term represents the mutual negative feedback between the two components. Finally there is a decay term with the strength of the degradation process controlled by the decay parameter  $k$ . This model has been studied extensively<sup>11,13,16</sup> as a representation of the cell differentiation process and has yielded insights into the epigenetic landscape as a cell differentiates.

In order to model the reverse differentiation process, we incorporate a time dependent chemical drive (along the lines of our earlier work<sup>12</sup>) as in experiments<sup>18</sup>. The feedback regulation depend on the concentrations of the TFs at some previous times, to account for the finite timescales of chromosome reorganization and other epigenetic state markers. We model these finite time processes through the incorporation of a time-delayed feedback. The degradation process is still assumed to be dependent on the instantaneous concentration. The time evolution equations incorporating a simultaneous chemical drive and delayed feedback is given by,

$$\begin{aligned}\frac{dx}{dt} &= a_0\Theta(d-t) + a_1\frac{x^n(t-\tau_1)}{S^n + x^n(t-\tau_1)} + b_1\frac{S^n}{S^n + y^n(t-\tau_2)} - kx(t) \\ \frac{dy}{dt} &= a_0\Theta(d-t) + a_2\frac{y^n(t-\tau_2)}{S^n + y^n(t-\tau_2)} + b_2\frac{S^n}{S^n + x^n(t-\tau_1)} - ky(t)\end{aligned}\quad (2)$$

the first term represents the time dependent drive, which is modeled by the Heaviside theta function, with the parameter  $d$  setting the duration for which the input is provided. Feedback processes mediated by the TF  $x$  are characterised by a delay timescale  $\tau_1$ , while those mediated by the TF  $y$  are characterised by a delay timescale  $\tau_2$ . The parameters  $a_{1/2}$  and  $b_{1/2}$  controls the strength of the positive and negative feedback for the two transcription factors. In order to reduce the dimensionality of the parameter space, we shall assume them to be the same throughout the remainder of this work, *i.e.*  $a_1 = a_2 = a$  and  $b_1 = b_2 = b$ . The decay constant is assumed to be the same for both TFs for similar reasons. A brief discussion on the implications of asymmetric parameters is presented in subsequent sections.

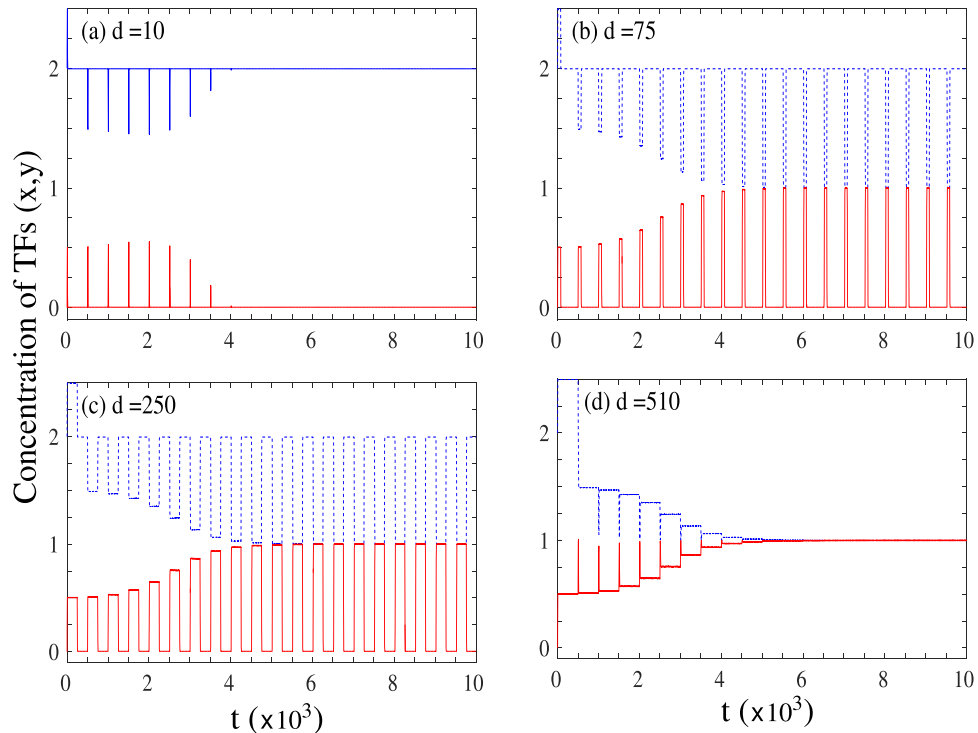
The nonlinear delay differential equations were solved using MATLAB (Mathworks, Inc.) using the dde23 solver. The solutions of the equations yields the trajectories of the system in the  $(x, y)$  concentration space. The fixed points and the bifurcation diagram of the non-delayed model was obtained by plotting the nullclines of the system of equations and subsequently performing a linear stability analysis. The trajectories obtained were analyzed in the long-time limit to classify the phase space of the system.

In the absence of a time-dependent drive and delayed feedback (Eq. 1), the behaviour of the system can be understood in terms of a bifurcation analysis of the phase space. The bifurcation diagram for this system is shown in Fig. 1(b). The phase space analysis shows that this two gene network is a multi-stable system, with a region of bistability for  $a < a_c \sim 0.77$ , and a region of multistability for  $a > a_c$ . For  $a < a_c$ , the system has two stable steady states (solid blue lines), separated by a central unstable state (dashed red line). The phase space velocities in this region are shown in Fig. 1(c), which clearly shows the two stable attractors separated by the central unstable state. For  $a > a_c$ , the system undergoes a pitchfork bifurcation where central unstable state become stable and two new unstable states appear symmetrically on either side of the central steady state. The phase space velocity for this region is shown in Fig. 1(d). The central steady state concentrations of the two TFs are equal *i.e.*,  $x = y$ , which in our model corresponds to the undifferentiated (or progenitor) state. The two terminal attractors, one with  $x = u$ ,  $y \approx 0$  and another with  $x \approx 0$ ,  $y = u$ , correspond to the two differentiated states. Thus, in each of the differentiated states, one TF completely dominates over the other ( $x \ll y$  or  $x \gg y$ ), with the weaker TF being effectively silenced. These then correspond to two distinct valleys of the epigenetic landscape which represent the two differentiated states. A comprehensive phase space and bifurcation analysis for this non-delayed model has been studied previously<sup>13,15</sup>.

We study the reverse differentiation process, starting from one of the differentiated steady state, and follow the time evolution of the system for different values of the drive duration and delay time. Since the bifurcation diagram shows two distinct dynamical behaviour regimes, *i.e.*, for  $a > a_c$  (Regime A), and  $a < a_c$  (Regime B), we study the system at a representative point in both these regimes. The parameters chosen for this paper are marked with a dotted line in the bifurcation diagram Fig. 1(b).

**Long lived oscillatory states.** We first report results for the set of model parameters given by, drive amplitude  $a_0 = 0.5$ , the positive and negative feedback amplitudes  $a = 1.0$ , and  $b = 1.0$ , decay constant  $k = 1.0$ , Hill exponent  $n = 4$ , and  $S = 0.5$ . We shall refer to this representative point in parameter space as Regime A, as illustrated in the bifurcation diagram Fig. 1(b). The central steady state (undifferentiated state) for this choice of parameters occurs at  $(x, y) = (1, 1)$  while the two differentiated stable steady states are at  $(x, y) = \{(2, \varepsilon) \text{ or } (\varepsilon, 2)\}$  with  $\varepsilon \approx 0$ . The time evolution of the system is shown in Fig. 2 for different values of the drive time  $d$  for a given value of the delay time  $\tau_1 = \tau_2 = \tau = 500$  starting from one of the differentiated states.

Figure 2(a) shows the time evolution of the concentrations of the two TFs when the drive time  $d (=10)$  is much less than the delay time  $\tau (=500)$ . In this regime  $d \ll \tau$ , the concentrations show some initial small fluctuations around the initial conditions, which quickly die out leaving the system in the same state as it initially started with. Thus for small drive times, the system remains in the differentiated state it started out in. As we increase the time for which the drive is provided, the initial small fluctuations grow, until beyond a certain critical value, we observe sustained fluctuations in the concentration of the transcription factors. This is shown for two drive times,  $d = 75$  and  $d = 250$  in Fig. 2(b,c). As can be seen from the figures, for these intermediate drive times, the system oscillates



**Figure 2.** Time series of the concentrations of the two transcription factors  $(x, y)$  in regime A for four different drive times for  $\tau = \tau_1 = \tau_2 = 500$ . (a)  $d = 10$ : System stays in initial undifferentiated state; (b)  $d = 75$ : Sustained oscillations with the system spending more time in the vicinity of one state than the other; (c)  $d = 250$ : Sustained oscillations with the system spending approximately equal times in the vicinity of the two states; (d)  $d = 510$ : Successful reprogramming to the undifferentiated state.

between the initial differentiated state and the central progenitor state  $(1, 1)$  for times which are much larger than the characteristic timescales in the system  $(d, \tau)$ .

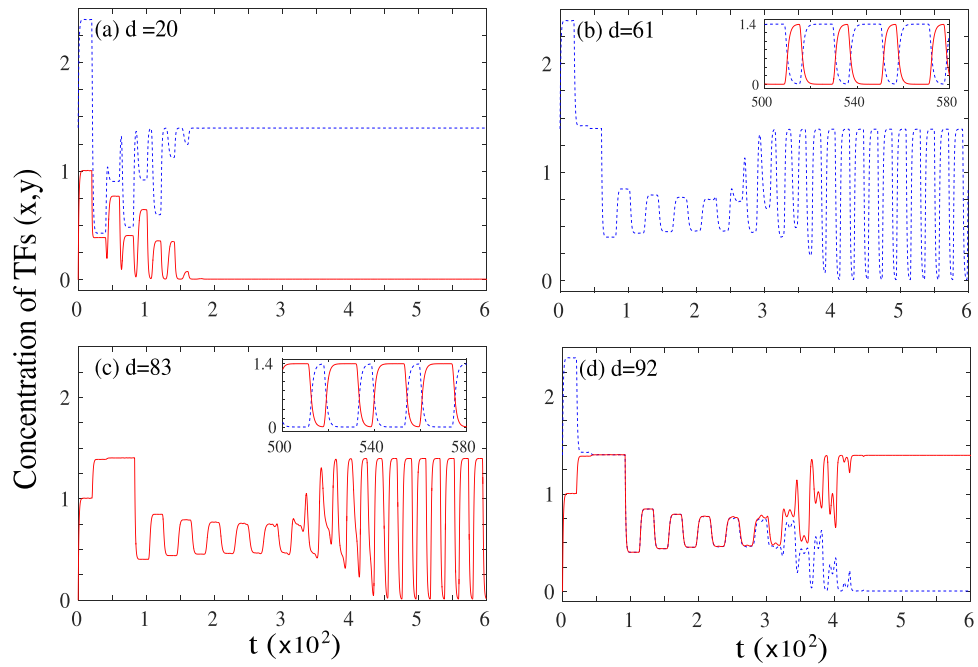
On increasing the drive time even further, the system comes out of this long lived oscillatory state and transitions to central progenitor state beyond a certain critical  $d$ . This is shown in Fig. 2 for  $d = 510$ . The system shows small oscillations as it starts from the initial differentiated state, but it quickly settles into undifferentiated state. This would then correspond to a successful completion of the reverse differentiation process. Thus in regime A, when the input is provided for a small time, the system stays in the differentiated state, while if the input is provided for large enough times, it successfully transitions into the progenitor state. In between, for intermediate drive times, the system shows long lived oscillatory states where it is neither in the differentiated nor in the undifferentiated state, but oscillates between the two. The results of our two-dimensional GRN with this interplay of drive and delayed feedback thus closely mirrors the experimental observations of Nagy and Nagy<sup>18</sup>, with the identification of the undetermined state in the experiments to the oscillatory state predicted by our analysis.

**Transdifferentiation.** While the results described for Regime A are robust for  $a > a_c$  (see bifurcation analysis Fig. 1(b) for details), the system shows a different class of behaviour for  $a < a_c$ . As a representative point in this region of parameter space, we choose  $a_0 = 1.0$ ,  $a = 0.4$ ,  $b = 1.0$ ,  $k = 1.0$ ,  $n = 5$ , and  $S = 0.5$  and follow time trajectories of the two TFs starting from one of the differentiated states, for different drive and delay times. The drive strength  $a_0$  and the Hill coefficient  $n$  are changed from the parameter set of Regime A to illustrate the robustness of our results. We shall refer to this representative point in parameter space as Regime B as shown in Fig. 1(b). The (unstable) progenitor state for this choice of parameters corresponds to  $(x, y) = (0.6, 0.6)$  while the two differentiated states are at  $(x, y) = \{(1.4, \varepsilon) \text{ or } (\varepsilon, 1.4)\}$  with  $\varepsilon \approx 0$ .

The time evolution results in regime B is shown in Fig. 3 for different  $d$  for time delay given by  $\tau_1 = \tau_2 = \tau = 20$ . When the drive time is small,  $d = 20$ , the system shows early time transient fluctuations, before relaxing back to the initial state. This is shown in Fig. 3(a), which is similar to the behaviour observed in regime A for small drive durations, where the system does not undergo reverse differentiation, but continues to remain in its initial differentiated state.

On increasing the drive time, the system shows sustained oscillations, as shown in Fig. 3(b,c) for  $d = 61$  and  $d = 84$  respectively. Crucially in this case, the oscillations are between the two differentiated states and not between the initial differentiated state and the progenitor state, as in regime A described above. The oscillations are again long-lived in comparison to the drive and delay time scales of the system. The insets in panels (b) and (c) of Fig. 3 show a magnified view of the oscillations for both the transcription factors.

Finally Fig. 3(d) shows the time evolution of the system for  $d = 82$ . In this case, after initial transient oscillations, the system stabilizes to a “flipped” state *i.e.* starting from one of the differentiated states  $(1.4, 0)$ , the system



**Figure 3.** Time series of the concentrations of the two transcription factors ( $x/y$ ) in regime B for four different drive times for  $\tau = \tau_1 = \tau_2 = 20$ . **(a)**  $d = 20$ : System stays in initial undifferentiated state; **(b)**  $d = 51$ : Sustained oscillations in concentration of TFs. The main panel shows a single TF ( $x$ ) while the inset shows a magnified view of both the TFs in a narrow window of time; **(c)**  $d = 83$ : Sustained oscillations in concentration of TFs. The main panel shows a single TF ( $y$ ) while the inset shows a magnified view of both the TFs in a narrow window of time; **(d)**  $d = 92$ : Transdifferentiation to the other differentiated state.

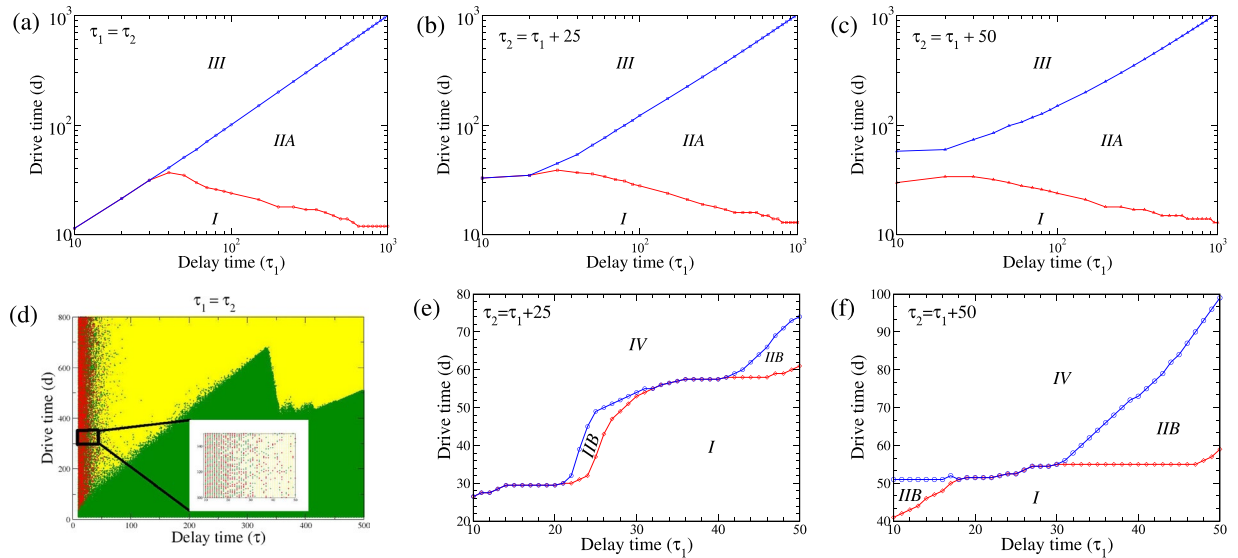
transitions to the other differentiated state (0, 1.4) without reaching the central progenitor state. This dynamics is reminiscent of the phenomenon of transdifferentiation, which has been observed experimentally where cells can directly transition from one differentiated state to another<sup>35–42</sup>.

The time evolution observed in regime B is qualitatively different from the behaviour obtained in regime A. In regime B, the system never reaches the central progenitor state for any choice of the drive and delay times. Similarly, the transdifferentiated state is never observed for regime A. Further the oscillatory state is distinct in the two cases, with oscillations between the progenitor and the differentiated state in regime A (half oscillations) and between the two differentiated states in regime B (full oscillations).

To the best of our knowledge this is the first theoretical model that helps elucidate how delayed feedback processes can give rise to transdifferentiation in gene regulatory networks.

**Phase diagrams.** The different dynamical behaviours in Regimes A and B, characterised by time evolution plots can conveniently be represented as phase diagrams in the drive vs. delay time ( $d$  vs.  $\tau$ ) plane as shown in Fig. 4. Both situations, where the delay timescale for the two TFs are same *i.e.* ( $\tau_1 = \tau_2 = \tau$ ) and different ( $\tau_1 \neq \tau_2$ ) are considered. The long time behaviour can be classified into four phases - (i) the system stays in the initial differentiated state (I); (ii) the system reaches an oscillatory state (IIA or IIB, depending on the nature of oscillations); (iii) the system undergoes reverse differentiation to reach the pluripotent state (III); and (iv) the system reaches a transdifferentiated state (IV).

Figure 4(a), shows the phase behaviour of the system in regime A for  $\tau_1 = \tau_2 = \tau$ . In the region with small delay and drive times  $\tau < \tau_{cr}$  and  $d < d_{cr}$ , the system stays in the initial differentiated state (I). For small delay times  $\tau < \tau_{cr}$ , but beyond a critical drive time  $d > d_{cr}$ , the system transitions directly from the differentiated state (I) to the central progenitor state (III), with the critical drive duration for the reverse differentiation process increasing linearly with the delay time *i.e.* ( $d_{cr} \propto \tau$ ). For  $\tau > \tau_{cr}$  this direct transition from the initial differentiated to the final undifferentiated state is lost. In this region, the system initially passes from the initial state to an oscillatory state (IIA) – as demonstrated in Fig. 2(b,c). In line with the experimental results<sup>18</sup>, we denote this minimal drive time beyond which the initial state is no longer the stable solution as  $d_{PNR}$  (the *Point-of-No-Return*). By keeping  $\tau$  fixed and increasing  $d$ , the oscillatory state is observed to be stable till a second critical value of the drive time  $d_{CPS}$  (*Commitment-to-Pluripotent-State*). Beyond this value of drive time *i.e.*  $d > d_{CPS}$  the undifferentiated state is stable. The threshold for the point-of-no-return  $d_{PNR}$  decreases with increasing time delay, implying that the oscillatory state sets in for lower values of the drive time for higher  $\tau$  values. On the other hand, the threshold for crossing to the pluripotent state increases linearly with delay times ( $d_{CPS} \propto \tau$ ), such that the crossover to the central state requires a chemical drive for longer durations as  $\tau$  is increased. Thus with increasing delay timescales, the regime of stability for sustained oscillations continues to increase widening the gap between the point of no return and commitment to pluripotency.

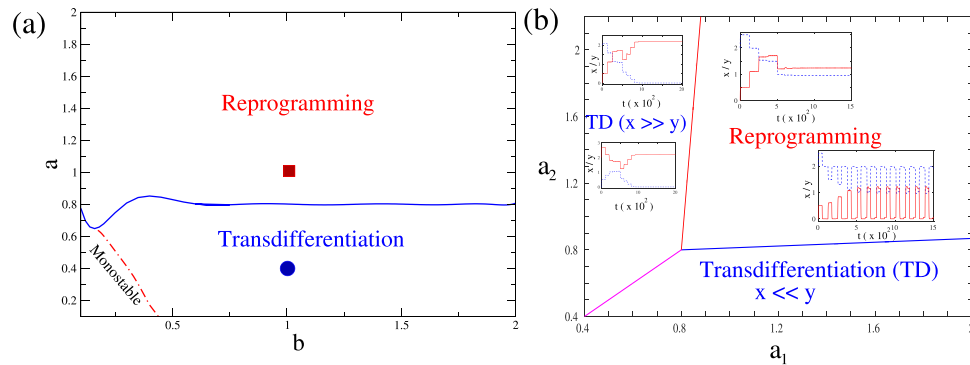


**Figure 4.** Phase space in the drive time ( $d$ ) vs. the delay time ( $\tau_1$ ) plane for (a) Regime A:  $\tau_2 = \tau_1$ ; (b) Regime A:  $\tau_2 = \tau_1 + 25$ ; (c) Regime A:  $\tau_2 = \tau_1 + 50$ ; (d) Regime B:  $\tau_2 = \tau_1$ . The inset shows a magnified view of a small region of this phase space to illustrate the chaotic behaviour in this regime; (e) Regime B:  $\tau_2 = \tau_1 + 25$ ; (f) Regime B:  $\tau_2 = \tau_1 + 50$ . All other parameters for regime A and regime B are as defined in the text.

In Fig. 4(b) we show the phase behaviour in regime A for the situation where the delay times are unequal *i.e.*  $\tau_1 \neq \tau_2$  with a constant difference between the two delay times, *i.e.*  $\Delta\tau = \tau_2 - \tau_1 = 25$ . Similar to the situation when delay times are equal, the system directly transits from the initial state (I) to the final progenitor state (III) for small value of  $\tau_1$  beyond a critical value of drive time  $d_{cr}$ . However unlike equal delay time scenario, the critical drive time at which this transition takes place  $d_{cr}$  remains almost constant as  $\tau_1$  is increased. Beyond a threshold  $\tau_{cr}$  this direct transition gives way to a two-step transition, in which stable state is the initial state for  $d < d_{PNR}$ , the oscillatory state for  $d_{PNR} < d < d_{CPS}$ , and finally the progenitor state for  $d > d_{CPS}$ . This is similar to the equal delay time scenario described earlier. Further, the boundaries  $d_{PNR}$  and  $d_{CPS}$  behave as before, with  $d_{PNR}$  decreasing with increasing  $\tau_1$  and  $d_{CPS}$  increasing linearly with  $\tau_1$ . The critical value of  $\tau_1$  at which this two-step transition take place is lower in this  $\Delta\tau = 25$  case than in the previous  $\Delta\tau = 0$  one. Finally in Fig. 4(c) we show the phase behaviour for the case of  $\Delta\tau = 50$ . Interestingly, in this case, there is no value of the time delay  $\tau_1$  for which the system makes a direct transition from the differentiated (I) to the undifferentiated state (III). In this regime upon increasing the drive time  $d > d_{PNR}$ , the system necessarily passes through a region where the oscillatory state is stable before it can make the transition to the central progenitor state beyond the  $d_{CPS}$  drive time. The width of oscillatory region remains constant with increasing  $\tau_1$  (unlike the  $\Delta\tau = 0$  and 25 cases) for small values of  $\tau_1$  and widens for large values with  $d_{PNR}$  and  $d_{CPS}$  diverging away from each other.

We now turn to the analysis of the phase behaviour in regime B, *i.e.* transdifferentiation regime. This is shown for  $\tau = \tau_1 = \tau_2$  in Fig. 4(d). For small values of the drive time, the system remains in the initial undifferentiated state (I). On increasing the drive time, for small values of the delay time  $\tau$  ( $\lesssim 50$ ), the final long-term behaviour of the system shows extreme sensitivity to parameter values. The inset of Fig. 4(d) shows a magnified view of the phase map for  $\tau < 50$  in the region  $300 < d < 350$ . As can be seen from the inset, the final state shows signatures of chaotic behaviour, where the long time state is either the initial state (I - green dots) or oscillatory state (IIB - yellow dots), or transdifferentiated state (III - red dots) for slight variations in drive time  $d$ . The long time behaviour of the system was obtained by time marching the set of delay differential equations (Eq. 2) for a time  $t$  much greater than the delay and drive timescales of the system ( $t = 10^4 \gg \{d, \tau\}$ ). While it is possible that the final state obtained from our simulations is not the true steady state, biologically relevant timescales would amount to the states observed after times which are of comparable magnitude to the experimental timescales, such as the time  $d$  for which the chemical drive is provided to the system<sup>18</sup>. In these parameter regimes, as shown by the phase map, this deterministic system shows chaotic behavior. For larger delay times ( $\tau \gtrsim 50$ ), the transdifferentiated state is no longer stable (in the range of drive times investigated) and the system transits from the initial state (I) to an oscillatory state (IIB) beyond a certain threshold drive time  $d_{cr}$ . The phase boundary demarcating these two regions shows a non-monotonic behaviour (saw-tooth-like pattern), increasing linearly until  $\tau \sim 340$ , beyond which it drops sharply before continuing to rise with a different slope.

Interestingly, this apparent randomness is lost when the the two delay timescales are no longer equal,  $\Delta\tau = \tau_1 - \tau_2 \neq 0$ . Figure 4(e) shows the phase behaviour in Regime B for  $\Delta\tau = 25$ . For small values of  $\tau_1$  ( $< 20$ ), the system transitions from the initial state to the transdifferentiated state beyond a critical drive time  $d_{cr}$ . For  $20 \lesssim \tau_1 \lesssim 30$ , there exists range of values of  $d$  which show sustained oscillations. Below this range, the system stays in the initial state, while for larger drive times, the transdifferentiated state is stable. For  $30 \lesssim \tau_1 \lesssim 40$ , there is no oscillatory state, and the system again transitions directly from the initial state to the transdifferentiated state. Above this delay timescale range  $\tau_1 > 40$ , the oscillatory state reappears in the phase map, and the region of



**Figure 5.** (a) Parameter phase plane in the positive feedback ( $a$ ) vs negative feedback ( $b$ ) plane for the case of symmetric parameters,  $a = a_1 = a_2$  and  $b = b_1 = b_2$ . The parameter point shown as a red square corresponds to the parameter set for the trajectories shown in Fig. 2, while the point shown as a blue circle corresponds to the parameter set for the trajectories shown in Fig. 3. (b) Parameter phase plane in the  $a_1$  vs.  $a_2$  plane for the case of asymmetric parameters for  $b = 1$ . The insets in the reprogramming region of the phase plane show the time trajectories of the transcription factors for  $a_1 = 1.0$ ,  $a_2 = 1.2$ ,  $d = 150$  for the reprogramming to the central attractor and  $a_1 = 1.0$ ,  $a_2 = 1.2$ ,  $d = 50$  for the long-lived oscillatory state. The insets in the transdifferentiation region of the phase plane show the time trajectories for  $a_1 = 0.6$ ,  $a_2 = 1.2$ ,  $d = 500$  starting from two different initial states, one with  $x \gg y$ , which shows transdifferentiation, and the other with  $x \ll y$ , which does not. All other parameters for both panels are the same as those in Fig. 1. All results are shown for  $\Delta\tau = 25$ .

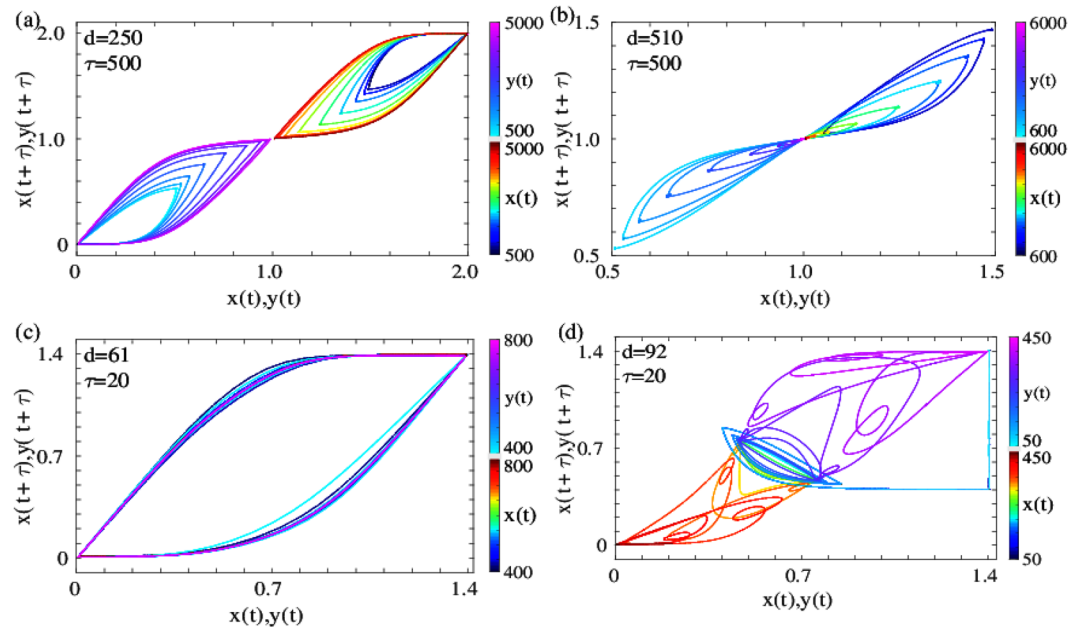
stability of the oscillatory state increases with increasing  $\tau_1$ . While the boundaries between the different phases are sharply defined for regime B in this  $\Delta\tau \neq 0$  condition, the signature of the chaotic behaviour remains in the appearance and disappearance of the stable oscillatory solution for different drive times.

Finally, Fig. 4(f) show the results for regime B with  $\Delta\tau = 50$ . In this case, for low values of  $\tau_1$  ( $\lesssim 18$ ), the oscillatory state is stable for a region of drive times. In an intermediate regime  $18 \lesssim \tau_1 \lesssim 30$ , there is no stable oscillatory state, and the system transitions from the initial state to the transdifferentiated state beyond a threshold drive time. The oscillatory state reappears for  $\tau_1 \gtrsim 30$ , and as in the  $\Delta\tau = 25$  case, the region of stability of the oscillatory solution increases with increasing  $\tau_1$ .

The above analysis of the phase behaviour of the model in regime A and B shows that the mutual interplay of drive and delay timescales can give rise to extremely rich landscapes, with non-trivial dependence of the final steady state on the time for which the input is provided, as well as the delay timescales associated with both the TFs. An interesting question to ask is how these exotic dynamical states are realizable in a biological context.

The observation of a long-lived oscillatory state and a transdifferentiated state is extremely robust that occurs for a range of parameters characterizing the system. In order to show that the introduction of drive and delay generically leads to these novel state, we plot the phase diagram in the parameter space of the positive feedback strength ( $a = a_1 = a_2$ ) and the negative feedback strength ( $b = b_1 = b_2$ ). This is shown in Fig. 5(a). As can be seen from the figure, beyond a certain critical value of the positive feedback strength  $a$ , the system is always in a regime, where an external input provided for a sufficiently long time can drive the system to a reprogrammed state. For very small values of  $a$  and small values of  $b$ , the system has a single stable state and hence this monostable region is not of interest in a biological context. The time trajectories of the concentrations of the transcription factor beyond the critical  $a$  value is qualitatively similar to those shown in Fig. 2. The parameters corresponding to Fig. 2 is shown as a red square in the  $a$ - $b$  plane. Conversely, below this critical  $a$ , an external input provided for a sufficiently long time always drives the system to a transdifferentiated state. The time trajectories in this region are qualitatively similar to those shown in Fig. 3, which point is denoted by a blue circle in the  $a$ - $b$  plane.

The system of delay differential equations characterizing the gene regulatory network were analyzed in the regime where the positive and negative feedback strengths for the two transcription factors were completely symmetrical. It is possible for the cell to drive the differentiation process deterministically to one of the differentiated cell fates through an asymmetric change in the feedback strengths. Various experiments have shown that stochasticity play a role in the cellular differentiation process, which argues that the central attractor corresponding to the undifferentiated state changes its stability through a symmetric variation of the control parameters<sup>13,43,44</sup>. Nevertheless, biochemical control of the feedback strength parameters leading to an asymmetric landscape is a distinct possibility. In order to show, that our predictions of an oscillatory or a transdifferentiated state are independent of whether the control parameters are symmetric or asymmetric, we plot the phase diagram of the system, in the  $a_1$ - $a_2$  plane, where the positive feedback strength of the two transcription factors are allowed to vary independently. This is shown in Fig. 5(b). When  $a_1$  and  $a_2$  are both sufficiently large, the system is in a reprogrammed state, for sufficiently large drive times. The insets in this reprogrammed region of the parameter space shows typical time trajectories for intermediate and large drive times, which leads to oscillations and reprogramming respectively. Note that the central undifferentiated state in this case is not symmetric with respect to the two transcription factors. The asymmetry in the feedback strengths leads to a central state with  $x \neq y$ . However, reprogramming to this central state proceeds as in the case of symmetric parameters. For small values of  $a_1$  and large values of  $a_2$ , the system is in the transdifferentiated regime. However, due to the asymmetry in the feedback



**Figure 6.** Trajectory plots in the  $x/y(t)$  vs.  $x/y(t + \tau)$  plane for (a) Oscillatory state in regime A (state *IIA*). The oscillations are between the differentiated states and the central undifferentiated state, which are represented as half-cycles in this phase plane. (b) Reprogramming to the undifferentiated state in regime A (state *III*). Note the decaying transient oscillations in both  $x$  and  $y$ , which converge to the central steady state. (c) Oscillatory state in regime B (state *IIB*). The oscillations are between the two differentiated states in this regime, which is represented by single limit cycle in this phase plane. (d) Transdifferentiated state in regime B (state *IV*). Note the unstable oscillations around the central state preceding transdifferentiation. In all four cases, the trajectories are coloured by the time values, as shown in the corresponding colorbars.

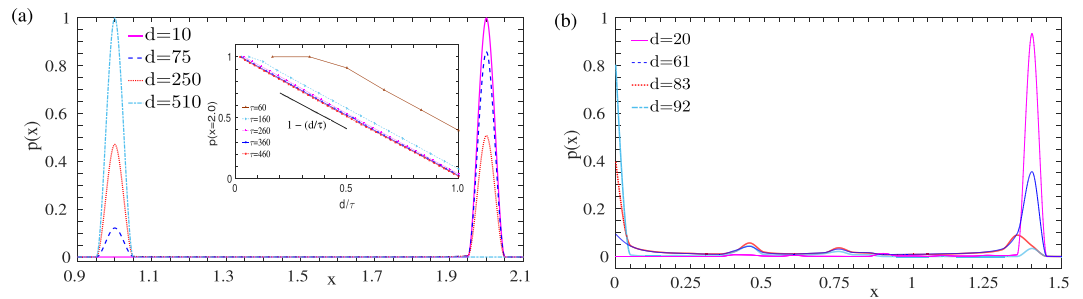
strengths, only one of the differentiated states can be transition to the other one, while the second differentiated state stays stable even for large values of the drive. This is shown in the insets in the transdifferentiated region of the parameter space for initial conditions corresponding to the two differentiated states. If the initial differentiated state has  $y \ll x$ , then the stronger feedback strength of the TF corresponding to  $y(t)$  can cause transdifferentiation to the other stable state in the system. On the other hand, if the initial differentiated state has  $x \ll y$ , then in this regime, it does not undergo any transdifferentiation. The situation is exactly reversed in the region of the parameter space corresponding to large values of  $a_1$  and small values of  $a_2$ . Here the initial state with  $x \ll y$  undergoes transdifferentiation, while the other differentiated state does not.

**Characterizing the oscillations.** We now turn to characterizing the long-lived oscillatory state observed in regimes A and B. In Fig. 6 we show the trajectory of the system in the oscillatory state plotted in the  $x(t)$  vs.  $x(t + \tau)$  plane. While a stable steady state corresponds to the trajectories evolving to a single point in this plane, steady oscillations evolve to a limit cycle.

Figure 6(a) shows the trajectory of the system in this plane in regime A for the concentrations of both the TFs  $x(t)$  and  $y(t)$ . As can be seen clearly from the time evolution of the trajectories, both the concentrations evolve to a steady limit cycle corresponding to the oscillatory region (*IIA*). The oscillations for each TF are between the initial differentiated state and the central undifferentiated state, which is manifested as two distinct limit cycles for the two TFs in the upper right and lower left quadrant. The corresponding trajectories for regime B are shown in Fig. 6(c). In this case, the oscillations are between the two undifferentiated states (*IIB*), and hence the limit cycles for the two TFs completely overlap, unlike the limit cycle in regime A. In contrast, we show the evolution of the trajectories in this plane for the reprogramming in regime A (Fig. 6(b)) and transdifferentiation in regime B (Fig. 6(d)). For successful reprogramming, initial transient oscillations between the initial state and the central state decay to the central stable state for both the TFs. For transdifferentiation, the concentrations of the two TFs perform small oscillations around the undifferentiated state for a short period of time (blue limit cycle), beyond which the limit cycle becomes unstable and the TFs reach the other differentiated state. The oscillations enable the trajectories to mix, ensuring that the history of the initial state is lost, which triggers the switch to the transdifferentiated state.

While the oscillatory state is stable for  $d_{PNR} < d < d_{CPS}$ , the nature of oscillations changes as  $d$  is varied. A quantitative estimate of the difference in the oscillatory state can be obtained by comparing the residence time distributions of the two transcription factors for different drive and delay times. We plot the residence time distribution as a function of the concentration for both regime A (Fig. 7(a)) and regime B (Fig. 7(b)). As can be seen clearly in Fig. 7(a), on increasing drive times, the peak in the probability distribution shifts from a single peak at  $x = 2$  for low values of  $d$  to a bimodal distribution at higher drive times ( $d = 250$ ), and back again to an unimodal distribution, but centered around the undifferentiated state at  $x = 1$  at even higher drive times,  $d = 510$ . In order





**Figure 7.** Probability of the concentration of any one transcription factor for different values of the drive time for (a) regime A with  $\tau = 500$ . The inset shows the probability near one differentiated state ( $x = 2$ ) as a function of the ratio of the drive to delay times  $d/\tau$ ; and (b) regime B with  $\tau = 20$ .

to estimate whether there is a systemic behaviour across a range of delay timescales  $\tau$ , we plot the probability of finding the TF1 at the vicinity of the undifferentiated state,  $x = 2$ , against the non-dimensionalised drive time  $d/\tau$  for a range of  $\tau$  values. Remarkably, this probability appears to have a simple scaling behaviour dictated by a single parameter  $d/\tau$ , viz.  $p(x = 2) = 1 - (d/\tau)$ . While the oscillations in regime B can be characterized by a similar method as shown in Fig. 7(b), the chaotic nature of the oscillatory and transdifferentiated state implies that there is no simple scaling relation with the drive time in this case, unlike in regime A.

## Discussion

Our work illustrates the richness of epigenetic landscapes in the presence of external drive and delayed feedback. Using a generic model of two self-activating and mutually inhibiting genes, we show that the reprogramming to the common progenitor state requires a balance of the drive and delay timescales, as well as appropriate positive and negative feedback strengths. Apart from the reprogrammed state, the system can find itself in long-lived oscillatory states or in a transdifferentiated state in different parameter regimes. An analysis of the phase diagrams in the delay - drive time plane provides a comprehensive picture of the final states of the reverse differentiation process. Additionally, the phase diagrams also provide a signature of chaotic behaviour in appropriate regimes, which have been hypothesized to play an important role in the cell-fate determination process<sup>45</sup>. In the chaotic regime, the model predicts that slight changes in the drive or delay parameters can drive the system from one state to another. The chaotic regime can thus provide a pathway for the control of cell fate determination. The phenomenon of transdifferentiation and oscillatory states is extremely robust and persists in a wide range of parameters as has been shown in Fig. 5. While the chaotic regions provide a regime where small changes in the drive and delay can affect the final state, there also exist regions where the final state is insensitive to wide variations of the drive and delay time.

To the best of our knowledge, this work provides one of the first theoretical basis for understanding the phenomenon of transdifferentiation in the presence of delayed feedback. Transdifferentiation was first shown experimentally in mouse fibroblasts in 1987, where the transcription factor MyoD, belonging to the Myogenic Regulatory Factors (MRF) family was shown to induce transdifferentiation of mouse embryonic fibroblasts to myoblasts<sup>35</sup>. The therapeutic potential of transdifferentiation has led to attempts to direct the process towards functional outcomes. The first proof of such a functional transdifferentiation was carried out in mouse liver cells which transitioned to pancreatic-beta-cell-like cells that helped control the effects of hyperglycemia<sup>46</sup>. Transdifferentiation has been shown to be a valid strategy to alter somatic cell fates in humans as well, which was first demonstrated in the transdifferentiation of a human liver cell to human beta cells through the effect of a single gene PDX-1<sup>47</sup>. The implications of successful transdifferentiation in medical applications has led to a surge of experimental work using a combination of strategies including the forced expression of specific transcription factors and combination of defined factors with microRNAs or small molecules to achieve targeted functional transdifferentiation of one somatic cell type to another<sup>36–42,48–50</sup>. An understanding of the underlying dynamics that characterizes the transdifferentiation process has however been lacking, including the question of whether cells need to necessarily pass through partially reprogrammed states in order to convert to a different cell type.

The mechanism proposed in this paper articulates one possible pathway for transdifferentiation, in the context of somatic cell types of similar lineages<sup>51</sup>. We show that it is possible to achieve transdifferentiation of one somatic cell type to another without passing through the undifferentiated state. Our analysis also provides an explanation for this process in terms of the underlying bifurcation diagram of the model - transdifferentiation arises when the feedback is tuned such that the central progenitor state is unstable. Interestingly, we observe that the transdifferentiation process is preceded by unstable oscillations, as has been seen in a different modeling approach<sup>52</sup>. A different mechanism has been proposed where transdifferentiation proceeds through a distinct stable intermediate state or a series of unstable intermediate states<sup>53</sup>. The incorporation of delays, as in the current work, proposes a new pathway to transdifferentiation, where unstable oscillations around the central state serve as a precursor to the transdifferentiated state. It would be interesting to study the transdifferentiation process in the presence of delayed feedback, in a landscape with multiple lineage branching points in order to characterize the pathway in these more complex situations.

The interplay of delayed feedback and chemical drive also leads to sustained oscillations in the levels of transcription factors. While we predicted the presence of sustained oscillations in the one gene network reported

previously<sup>12</sup>, this study conclusively highlights the importance of these oscillations for a more realistic two-gene network. Additionally, oscillations in the concentrations of transcription factors show different characteristics, (a) oscillations between the differentiated and undifferentiated state in one regime (IIA), and (b) oscillations between the two differentiated states in a different regime (IIB). The two oscillatory states may play different roles in the cellular context. It has been shown experimentally that oscillations in the Notch effector gene *Hes1* regulates the maintenance of neural progenitor cell types<sup>54,55</sup>. It has also been reported that Stella shows heterogeneous expression levels and dynamic equilibrium in embryonic stem cells (ESC), which is responsible for the ESC existing in a metastable state where they can shift between Inner Cell Mass (ICM) and epiblast like phenotypes<sup>56</sup>. Oscillations have been reported to be a possible feature of epigenetic landscapes<sup>52,57</sup>. We show that a possible route to these oscillations is through the incorporation of delayed feedback, and this can potentially explain novel states observed during the reprogramming process<sup>18</sup>. The feedback loops involved in gene regulatory networks are typically composed of multi-step reactions which introduce a characteristic delay time scale. Such delayed feedback loops have been shown to play an important role in both experimental and theoretical contexts<sup>58–61</sup>. While further experiments are needed to investigate whether the novel states seen in the doxycycline induced reprogramming process by Nagy and Nagy<sup>18</sup> are indeed due to oscillations or some other mechanism, this work demonstrates the importance of oscillatory states in general when time delayed feedback circuits are incorporated in the modeling of gene regulatory networks.

In this paper we consider a *deterministic* model of epigenetic landscapes, and consider the effects of time delays and chemical drive. In cellular contexts, it is also important to re-interpret our results in the presence of stochastic fluctuations, which can play a role in cell fate determination<sup>62–69</sup>. Noise induced transitions between different cell types can play an important role in gaining a comprehensive understanding of the differentiation process. The analysis of stochastic fluctuations in these highly nonlinear and non-autonomous delay differential equations present a significant challenge, and the results from an extensive analysis of the role of noise in these systems shall be presented in a forthcoming publication.

The present work provides a framework for the analysis of reprogramming experiments and underscores the importance in accounting for delayed feedback in building a comprehensive picture of the epigenetic landscape. From the perspective of Waddington's landscape the inclusion of delayed feedback translates to an asymmetry in the cell fate choices for the forward and reverse reprogramming processes. Models involving a larger number of transcription factors which incorporate the effects of this time delay can help in understanding the reprogramming pathway as well serve as a guide to targeted experiments in manipulating the cell fate determination process.

## References

1. Takahashi, K. & Yamanaka, S. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell* **126**, 663–676 (2006).
2. Takahashi, K. *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861–872 (2007).
3. Yu, J. *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science* **318**, 1917–1920 (2007).
4. Maherali, N. *et al.* A high-efficiency system for the generation and study of human induced pluripotent stem cells. *Cell Stem Cell* **3**, 340–345 (2008).
5. Ebert, A. D. *et al.* Induced pluripotent stem cells from a spinal muscular atrophy patient. *Nat.* **457**, 277 (2009).
6. Shi, Y. *et al.* Induction of pluripotent stem cells from mouse embryonic fibroblasts by Oct4 and Klf4 with small-molecule compounds. *Cell stem cell* **3**, 568–574 (2008).
7. Waddington, C. H. *The Strategy of Genes* (Allen and Unwin, London, 1957).
8. Ferrell, J. E. & Machleder, E. M. The biochemical basis of an all-or-none cell fate switch in xenopus oocytes. *Sci.* **280**, 895–898 (1998).
9. Ferrell, J. E. Jr. & Xiong, W. Bistability in cell signaling: How to make continuous processes discontinuous, and reversible processes irreversible. *Chaos* **11**, 227–236 (2001).
10. Ferrell, J. E. Jr. *et al.* Simple, realistic models of complex biological processes: Positive feedback and bistability in a cell fate switch and a cell cycle oscillator (2009).
11. Ferrell, J. E. Jr. Bistability, Bifurcations, and Waddington's Epigenetic Landscape. *Curr. Biol.* **22**, R458–R466 (2012).
12. Mitra, M. K., Taylor, P. R., Hutchison, C. J., McLeish, T. C. B. & Chakrabarti, B. Delayed self-regulation and time-dependent chemical drive leads to novel states in epigenetic landscapes. *J. The Royal Soc. Interface* **11**, 20140706 (2014).
13. Huang, S., Guo, Y. P., May, G. & Enver, T. Bifurcation dynamics in lineage-commitment in bipotent progenitor cells. *Dev. Biol.* **305**, 695–713 (2007).
14. Zhou, J. X. & Huang, S. Understanding gene circuits at cell-fate branch points for rational cell reprogramming. *Trends Genet.* **27**, 55–62 (2011).
15. Wang, J., Xu, L., Wang, E. & Huang, S. The potential landscape of genetic circuits imposes the arrow of time in stem cell differentiation. *Biophys. journal* **99**, 29–39 (2010).
16. Wang, J., Zhang, K., Liu, X. & Wang, E. Quantifying the Waddington landscape and biological paths for development and differentiation. *Proc. Natl. Acad. Sci.* **108**, 8257–8262 (2011).
17. Pal, M., Ghosh, S. & Bose, I. Non-genetic heterogeneity, criticality and cell differentiation. *Phys. biology* **12**, 016001 (2014).
18. Nagy, A. & Nagy, K. The mystery of induced pluripotency: where will they lead? *Nat. Methods* **7**, 22–24 (2010).
19. Mellor, J., Dudek, P. & Clynes, D. A glimpse into the epigenetic landscape of gene regulation. *Curr. opinion genetics & development* **18**, 116–122 (2008).
20. Karlič, R., Chung, H.-R., Lasserre, J., Vlahoviček, K. & Vingron, M. Histone modification levels are predictive for gene expression. *Proc. Natl. Acad. Sci.* **107**, 2926–2931 (2010).
21. Di Stefano, M., Paulsen, J., Lien, T. G., Hovig, E. & Micheletti, C. Hi-c-constrained physical models of human chromosomes recover functionally-related properties of genome organization. *Sci. reports* **6**, 35985 (2016).
22. Croft, J. A. *et al.* Differences in the localization and morphology of chromosomes in the human nucleus. *The J. cell biology* **145**, 1119–1131 (1999).
23. Bolzer, A. *et al.* Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. *PLoS biology* **3**, e157 (2005).
24. Meaburn, K. J. & Misteli, T. Cell biology: chromosome territories. *Nat.* **445**, 379–381 (2007).
25. De Carvalho, D. D., You, J. S. & Jones, P. A. Dna methylation and cellular reprogramming. *Trends cell biology* **20**, 609–617 (2010).
26. Kelly, T. K., De Carvalho, D. D. & Jones, P. A. Epigenetic modifications as therapeutic targets. *Nat. biotechnology* **28**, 1069–1078 (2010).
27. Fraser, J. *et al.* Chromatin conformation signatures of cellular differentiation. *Genome biology* **10**, R37 (2009).

28. Groth, A., Rocha, W., Verreault, A. & Almouzni, G. Chromatin challenges during DNA replication and repair. *Cell* **128**, 721–733 (2007).
29. Margueron, R. & Reinberg, D. Chromatin structure and the inheritance of epigenetic information. *Nat. reviews. Genet.* **11**, 285 (2010).
30. Zhang, P. *et al.* Negative cross-talk between hematopoietic regulators: Gata proteins repress pu. 1. *Proc. Natl. Acad. Sci.* **96**, 8705–8710 (1999).
31. Zhang, P. *et al.* Pu. 1 inhibits gata-1 function and erythroid differentiation by blocking gata-1 dna binding. *Blood* **96**, 2641–2648 (2000).
32. Rhodes, J. *et al.* Interplay of pu. 1 and gata1 determines myelo-erythroid progenitor cell fate in zebrafish. *Dev. cell* **8**, 97–108 (2005).
33. Nutt, S. L., Metcalf, D., D'Amico, A., Polli, M. & Wu, L. Dynamic regulation of pu. 1 expression in multipotent hematopoietic progenitors. *J. Exp. Medicine* **201**, 221–231 (2005).
34. Hoppe, P. S. *et al.* Early myeloid lineage choice is not initiated by random pu. 1 to gata1 protein ratios. *Nat.* **535**, 299–302 (2016).
35. Davis, R. L., Weintraub, H. & Lassar, A. B. Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* **51**, 987–1000 (1987).
36. Takeda, Y., Harada, Y., Yoshikawa, T. & Dai, P. Direct conversion of human fibroblasts to brown adipocytes by small chemical compounds. *Sci. Reports* **7** (2017).
37. Dai, P., Harada, Y. & Takamatsu, T. Highly efficient direct conversion of human fibroblasts to neuronal cells by chemical compounds. *J. clinical biochemistry nutrition* **56**, 166–170 (2015).
38. Vierbuchen, T. *et al.* Direct conversion of fibroblasts to functional neurons by defined factors. *Nat.* **463**, 1035 (2010).
39. Ieda, M. *et al.* Direct reprogramming of fibroblasts into functional cardiomyocytes by defined factors. *Cell* **142**, 375–386 (2010).
40. Takasato, M., Vanslambrouck, J. M. & Little, M. H. Reprogramming somatic cells to a kidney fate. *Semin. Nephrol.* **34**, 462–480, <http://www.sciencedirect.com/science/article/pii/S0270929514000850>, Stem Cells and Kidney Regeneration, <https://doi.org/10.1016/j.semnephrol.2014.06.012> (2014).
41. Hendry, C. E. & Little, M. H. Reprogramming the kidney: a novel approach for regeneration. *Kidney Int.* **82**, 138–146, <http://www.sciencedirect.com/science/article/pii/S0085253815555257>, <https://doi.org/10.1038/ki.2012.68> (2012).
42. Zhou, Q., Brown, J., Kanarek, A., Rajagopal, J. & Melton, D. A. *In vivo* reprogramming of adult pancreatic exocrine cells to [beta]-cells. *nature* **455**, 627 (2008).
43. Graf, T. Differentiation plasticity of hematopoietic cells. *Blood* **99**, 3089 (2002).
44. Hume, D. A. Probability in transcriptional regulation and its implications for leukocyte differentiation and inducible gene expression. *Blood* **96**, 2323 (2000).
45. Furusawa, C. & Kaneko, K. Chaotic expression dynamics implies pluripotency: when theory and experiment meet. *Biol. direct* **4**, 17 (2009).
46. Ferber, S. *et al.* Pancreatic and duodenal homeobox gene 1 induces expression of insulin genes in liver and ameliorates streptozotocin-induced hyperglycemia. *Nat. medicine* **6**, 568 (2000).
47. Sapir, T. *et al.* Cell-replacement therapy for diabetes: Generating functional insulin-producing tissue from adult human liver cells. *Proc. Natl. Acad. Sci.* **102**, 7964–7969 (2005).
48. Samavarchi-Tehrani, P. *et al.* Functional genomics reveals a bmp-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. *Cell stem cell* **7**, 64–77 (2010).
49. Treutlein, B. *et al.* Dissecting direct reprogramming from fibroblast to neuron using single-cell rna-seq. *Nat.* **534**, 391–395 (2016).
50. Shen, C.-N., Burke, Z. D. & Tosh, D. Transdifferentiation, metaplasia and tissue regeneration. *Organog.* **1**, 36–44 (2004).
51. Zhou, Q. & Melton, D. A. Extreme makeover: converting one cell into another. *Cell stem cell* **3**, 382–388 (2008).
52. Rabajante, J. F. & Babier, A. L. Branching and oscillations in the epigenetic landscape of cell-fate determination. *Prog. biophysics molecular biology* **117**, 240–249 (2015).
53. Xu, L., Zhang, K. & Wang, J. Exploring the mechanisms of differentiation, dedifferentiation, reprogramming and transdifferentiation. *PLoS one* **9**, e105216 (2014).
54. Shimojo, H., Ohtsuka, T. & Kageyama, R. Oscillations in notch signaling regulate maintenance of neural progenitors. *Neuron* **58**, 52–64 (2008).
55. Bonev, B., Stanley, P. & Papalopulu, N. MicroRNA-9 modulates hes1 ultradian oscillations by forming a double-negative feedback loop. *Cell reports* **2**, 10–18 (2012).
56. Hayashi, K., de Sousa Lopes, S. M. C., Tang, F. & Surani, M. A. Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states. *Cell stem cell* **3**, 391–401 (2008).
57. Suzuki, N., Furusawa, C. & Kaneko, K. Oscillatory protein expression dynamics endows stem cells with robust differentiation potential. *PLoS one* **6**, e27232 (2011).
58. Bratsun, D., Volfson, D., Tsimring, L. S. & Hasty, J. Delay-induced stochastic oscillations in gene regulation. *PNAS* **102**, 14593 (2005).
59. Rao, C. V. & Arkin, A. P. Stochastic chemical kinetics and the quasi-steady-state assumption: Application to the Gillespie algorithm. *J. Chem. Phys.* **118**, 4999 (2003).
60. Talora, C., Franchi, L., Linden, H., Ballario, P. & Macino, G. Role of a white collar-1-white collar-2 complex in blue-light signal transduction. *The EMBO J.* **18**, 4961 (1999).
61. Denault, D. L., Loros, J. J. & Dunlap, J. C. WC-2 mediates WC-1-FRQ interaction within the PAS protein-linked circadian feedback loop of Neurospora. *The EMBO J.* **20**, 109 (2001).
62. Chang, H. H., Hemberg, M., Barahona, M., Ingber, D. E. & Huang, S. Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. *Nat.* **453**, 544 (2008).
63. Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. Stochastic gene expression in a single cell. *Sci.* **297**, 1183–1186 (2002).
64. Kepler, T. B. & Elston, T. C. Stochasticity in transcriptional regulation: origins, consequences, and mathematical representations. *Biophys. journal* **81**, 3116–3136 (2001).
65. Eldar, A. & Elowitz, M. B. Functional roles for noise in genetic circuits. *Nat.* **467**, 167 (2010).
66. Suda, T., Suda, J. & Ogawa, M. Single-cell origin of mouse hemopoietic colonies expressing multiple lineages in variable combinations. *Proc. Natl. Acad. Sci.* **80**, 6689–6693 (1983).
67. Shah, N. M., Groves, A. K. & Anderson, D. J. Alternative neural crest cell fates are instructively promoted by *tgfβ* superfamily members. *Cell* **85**, 331–343 (1996).
68. Yamamaka, Y., Lanner, F. & Rossant, J. Fgf signal-dependent segregation of primitive endoderm and epiblast in the mouse blastocyst. *Dev.* **137**, 715–724 (2010).
69. Krieger, T. & Simons, B. D. Dynamic stem cell heterogeneity. *Dev.* **142**, 1396–1406, <http://dev.biologists.org/content/142/8/1396>, <http://dev.biologists.org/content/142/8/1396.full.pdf>, <https://doi.org/10.1242/dev.101063> (2015).

## Acknowledgements

M.K.M. acknowledges financial support from Ramanujan Fellowship, DST, INDIA (13DST052) and the IIT Bombay Seed Grant (14IRCCSG009). B.C. acknowledges the Biophysical Sciences Institute, and Institute of Advanced Studies, Durham University for financial support at the time the work was done.

### Author Contributions

M.K.M. and B.C. conceived the study, analysed data, and wrote the manuscript. B.K. performed simulations, analysed data, and wrote the manuscript. R.S. performed initial simulations. All authors reviewed the manuscript.

### Additional Information

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018