A mathematical model of cell fate reprogramming: From open loop to closed loop feedback control

Domitilla Del Vecchio
Mechanical Engineering
Synthetic Biology Center
MIT

IPAM EPI2017 Workshop, March 2017
Outline

Background on cell fate reprogramming

Reprogramming GRNs through prefixed overexpression

Reprogramming GRNs through feedback overexpression

A blueprint of a synthetic genetic feedback controller implementing feedback overexpression
Background on cell fate reprogramming

**iPSC reprogramming**
(2006, Yamanaka and colleagues)

**basic rationale:**

- high Oct4, Sox2, Klf-4, c-Myc (OSKM)

overexpress TFs that are found in higher concentration in the target cell type

*(TF-mediated reprogramming)*
Background on cell fate reprogramming

- Today: efficiency remains below 1% across a range of transfection methods
- Majority of resulting cells become partially reprogrammed – overexpressed TFs fail to have the pluripotent concentrations (Morris and Daley, 2013; Chan et al. 2009)
- TF stoichiometry highly affects efficiency (Papapedrou et al. 2009)

Can reprogramming failure be attributed to the structure of the core GRN that is responsible for pluripotent (PL) cell fate determination and maintenance?

If GRN structure can be implicated in these failures, what reprogramming approach can be considered that is highly independent of the endogenous GRN?
Outline

Background on cell fate reprogramming

Reprogramming GRNs through prefixed overexpression

Reprogramming GRNs through feedback overexpression

A blueprint of a synthetic genetic feedback controller implementing feedback overexpression
A mathematical model of cell fate reprogramming

Gene regulatory network (GRN)  

\[ x_i \text{ concentration of TF } x_i \quad x = (x_1, \ldots, x_n) \text{ state of the GRN} \]

\[ u_i \geq 0 \text{ external stimuli (overexpression)} \]

\[ \sum u : \frac{dx_i}{dt} = f_i(x, u_i) = H_i(x) - \gamma_i x_i + u_i \quad \sum_0 : u = 0 \]

Hill function  

\[ \text{decay} \]

\[ S^k \quad \text{stable steady state of} \quad \sum_0 \quad R_0(S^k) \quad \text{region of attraction of } S^k \]

\[ f(S^k, 0) = 0 \]

\[ \sum_0 : \frac{dx}{dt} = f(x, 0) \]

\[ \sum_u : \frac{dx}{dt} = f(x, u) \]

\[ \sum_0 : \frac{dx}{dt} = f(x, 0) \]

The GRN is **strongly reprogrammable to** \( S^0 \) provided there is a \( u \) such that \( \sum_u \) has the attracting set contained in \( R_0(S^0) \)

D. Del Vecchio, MIT
Reprogramming cooperative GRNs

Fact 1: The set of stable steady states in a cooperative GRN has always a maximal element

Theorem: A cooperative GRN is strongly reprogrammable only to its maximal stable steady state

→ A cooperative GRN is not strongly reprogrammable to stable steady states that have intermediate values of some of the TF concentrations $x_i$
Reprogramming cooperative GRNs

Oct4 concentration

\[ \begin{align*}
&\text{low in trophoderm (TR)} \\
&\text{intermediate in pluripotent cells (PL)} \\
&\text{high in primitive endoderm (PE)}
\end{align*} \]

\[ \Rightarrow \text{Core pluripotency GRN is not strongly reprogrammable to PL} \]

Overexpression \( u \) destabilizes TR

But as \( u \) is increased, PL valley may disappear before the TR valley

PE is always an attractor for all \( u \)
Reprogramming cooperative GRNs

A GRN is **weakly reprogrammable to** $S^0$ **from** $x^0$ **if** there is an input $u$ **such** **that** the attracting set of $x^0$ **is** contained in the region of attraction of $S^0$

**Fact:** A cooperative GRN may be weakly reprogrammable to an intermediate state from a state lower than it. **But, whether an input $u$ exists is highly dependent on the GRN**

Example:

\[
H_i(x) = \frac{a_i x_1^2 + b_i x_2 + c_i x_1^2 x_2}{1 + x_1^2 + x_2^2 + dx_1^2 x_2^2}
\]

Preset overexpression is a robust approach to reprogram the GRN to the maximal state but a fragile approach to reprogram the GRN to intermediate states
Reprogramming cooperative GRNs subject to perturbations

When a core GRN is embedded in a larger GRN:

- monotone interactions do not change the reprogramming properties
- non-monotone interactions:

\[
\frac{dx}{dt} = f(x, u) + \epsilon d(x), \quad \|d(x)\| \leq d_M
\]

cooperative vector field non-monotone perturbation

for $\epsilon$ small, the perturbed stable steady states are in $\epsilon$-balls around the nominal ones

**Theorem**: For $\epsilon$ sufficiently small, the perturbed cooperative GRN is strongly reprogrammable only to $S^m_{\epsilon}$ (the perturbation of the maximal stable steady state $S^m$)

$\rightarrow$ A cooperative GRN subject to small non-monotone perturbations is *not strongly reprogrammable* to stable steady states that arise from the perturbation of intermediate ones

D. Del Vecchio, MIT
Reprogramming cooperative GRNs subject to perturbations

Oct4 overexpression: PL disappears before TR cannot reprogram the GRN from TR to PL. GRN strongly reprogrammable to PE only

Preset overexpression of TFs does not have sufficient control over TF concentrations to enforce their target (PL) concentrations within the GRN
Outline

Background on cell fate reprogramming

Reprogramming GRNs through prefixed overexpression

Reprogramming GRNs through feedback overexpression

A blueprint of a synthetic genetic feedback controller implementing feedback overexpression
High-gain feedback globally stabilizes any target state

\[ \sum_u \frac{dx_i}{dt} = H_i(x) - \gamma_i x_i + u_i, \quad u_i = G(x_i^* - x_i) \]

target state

the overexpression level is not preset but it is adjusted based on
the proximity to the target state

High-gain: \( Gx_i^* \gg H_i(x), \quad G \gg \gamma_i \Rightarrow \frac{dx_i}{dt} \approx G(x_i^* - x_i) \)

\( x_i^* \) is unique steady state

Using tools from contraction theory prove that

\[ \lim_{t \to \infty} \|x_i(t) - x_i^*(t)\| \to \frac{(H_M + \gamma_i x_i^*)}{G + \gamma_i} \]

Example:

Pick \( x^* \) in region of attraction of \( S^0 \)

\[ u_1 = G_1(x_1^* - x_1) \]
\[ u_2 = G_2(x_2^* - x_2) \]

D. Del Vecchio, MIT
High-gain feedback globally stabilizes any target state

Qualitatively:

- **preset overexpression**
  \[ u = Gx^* \]

- **enhanced degradation**
  \[ u = -Gx \]

- **combined**
  \[ u = Gx^* - Gx \]
  \[ u = G(x - x^*) \]

We can physically interpret this controller as applying large overexpression combined with similarly large enhance degradation of the transcription factor.

We look for a synthetic genetic circuit design capable of
- Large overexpression of endogenous TF, which is tunable through inducers
- Enhanced degradation of both endogenous and synthetically produced TF

D. Del Vecchio, MIT
Outline

Background on cell fate reprogramming

Reprogramming GRNs through prefixed overexpression

Reprogramming GRNs through feedback overexpression

A blueprint of a synthetic genetic feedback controller implementing feedback overexpression
Synthetic genetic circuit implementing feedback control of TF overexpression

Synthetic circuit

Endogenous system

\[ x_i^s \]

\[ m_i^s \]

Small interfering RNA

\[ S_i \]

\[ D_i(x) \]

\[ x_i^e \]

\[ m_i^e \]

\[ x^* \]

\[ u = Gx^* \]

\[ u = -Gx \]

\[ u = Gx^* - Gx \]

Enzymatic reaction for siRNA [1]

\[ m_i^k + s_i \xrightarrow{a_i} c_i^k \xrightarrow{k_i} s_i, \quad c_i^k \xrightarrow{\beta_i} \emptyset, \quad k \in \{e, s\} \]

Total species concentration (endogenous plus synthetic) assumption [1]

\[ m_i \ll K_i \quad (M-M \text{ constant}) \]

\[
\frac{dm_i}{dt} = \left( \bar{H}_i(x) - \delta_i m_i + G_i(m^*_i - m_i) \right) \left( \frac{1}{1 + \bar{s}_i/K_i} \right) \\
\frac{dx_i}{dt} = \kappa_i m_i - \gamma_i x_i \quad \bar{s}_i = D\alpha_i/\beta_i \]

DNA copy number

Feedback gain

\[ G_i = D \frac{k_i \alpha_i}{K_i \beta_i} \]

Target encoded by inducer

\[ m^*_i = \frac{\gamma_i}{\kappa_i} x^*_i \quad x^*_i = F(I^*_i) \]

Synthetic genetic circuit for feedback control of TF overexpression

**Synthetic circuit**

- $x_s^i$ (gene)
- $m_s^i$ (RNA sequence)
- $D_{h_i}(I_i)$ (production)
- $I_i$ (input)
- $x_i$ (gene)
- $m_i$ (mRNA)
- $s_i$ (RNA sequence)
- $S_i$ (small interfering RNA)
- $D_{\alpha_i}$ (regulator)
- $G_i$ (DNA copy number ~1-5)

**Endogenous system**

- $x_1$ (gene)
- $x_2$ (gene)
- $I_i$ (input)
- $G_i$ (DNA copy number)
- $m_i^*$ (mRNA)
- $m_i$ (mRNA)
- $s_i$ (RNA sequence)
- $S_i$ (small interfering RNA)
- $D_{\alpha_i}$ (regulator)
- $G_i$ (DNA copy number ~1-5)

### Differential Equations

\[
\frac{dx}{dt} = f(x, 0)
\]

### Equations

\[
\begin{align*}
G_i & \gg \delta_i, \quad G_i m_i^* \gg \bar{H}_i(x) \\
\frac{dm_i}{dt} & \approx G_i (m_i^* - m_i) \left( \frac{1}{1 + \bar{s}_i/K_i} \right) \\
\frac{dx_i}{dt} & = \kappa_i m_i - \gamma_i x_i \\
G_i & = D \frac{k_i \alpha_i}{K_i \beta_i} \\
m_i(t) & \rightarrow m_i^* \quad x_i(t) \rightarrow x_i^*
\end{align*}
\]

### Convergence Speed

\[
\text{convergence speed for mRNA } \sim 1/k_i
\]

D. Del Vecchio, MIT
Using the genetic controller for reprogramming

when the controller is shut down, the endogenous species take over

\[
\frac{dm_i}{dt} = (\bar{H}_i(x) - \delta_i m_i)
\]

\[
\frac{dx_i}{dt} = \kappa_i m_i - \gamma_i x_i
\]

\[
\frac{dm_i^s}{dt} = -\delta_i m_i^s \Rightarrow (x^s(t), m^s(t)) \rightarrow 0
\]

\[
\Rightarrow (x^e(t), m^e(t)) \rightarrow S^0
\]

D. Del Vecchio, MIT
Reprogramming the pluripotency network

Nanog → Oct4 → Nanog →...

\( u_1 \) → \( u_2 \)

\([\text{Oct4}] \) → \([\text{Nanog}] \)

stochastic simulations (CLE)

D. Del Vecchio, MIT
Summary

We proposed a mathematical model that implies network structure in reprogramming failure

We proposed a closed loop feedback control approach of overexpression to over-rule network’s dynamics

Del Vecchio et. al. Cell Systems, 2017

On-going and future work

Mathematical framework to determine how intrinsic noise in the low molecule count regime affects reprogramming

Mathematical model of how chromatin dynamics affects reprogramming

Implementation of the genetic feedback controller in reprogrammable cell lines
Thanks to...

Jim Collins  Eduardo Sontag