Irreversibility of T-cell specification: insights from computational modelling of a minimal network architecture

Erica Manesso, Hao Yuan Kueh, George Freedman, Ellen V. Rothenberg, Carsten Peterson

Detailed Materials and Methods

Background

While networks accounting for the subdivision of embryonic tissue fates can be framed in Boolean logic terms\(^1\), GRN accounting for signalling response kinetics are normally modelled with continuous functions. In applying these models to single-cell fate determination, a crucial difference is whether spatial partitioning or temporal features of a response are to be explained. When time-series are available, continuous models for GRN are a good compromise between (a) the necessity to quickly explore different architectures and to perform inference, well-covered by the logical models, and (b) the need to capture stochastic behaviour, accomplished only by single-molecule level models\(^2\). Inspired by topological models proposed in developmental studies\(^3,4,5,6,7,8\), we exploited 32 combinatorial combinations to describe the dynamics of BCL11B as function of Notch signalling, TCF-1, and GATA-3. The results confirmed that a simple sum of these three components (i.e. Notch OR TCF-1 OR GATA-3) in BCL11B production rate is not enough to justify the delay in BCL11B versus the increase of TCF-1 and GATA-3, typical of a feed-forward motif\(^9\), thus suggesting that complex interactions are needed. On the other hand, a simple AND interaction among Notch signalling, TCF-1, and GATA-3 failed to keep high levels BCL11B expressed when Notch signalling disappears.

The initial conditions on which our model was built come from published experimental data. The potential for GATA-3, TCF-1, BCL11B and PU.1 to act on each other is based on gene expression effects measured in acute perturbation experiments\(^10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27\), while genome-wide binding profiles of GATA-3, PU.1, TCF-1 and more limited data for Notch binding indicate candidate sites where direct regulatory interactions may be mediated\(^20,27,28,29\). The proposed minimal model does not include all genes that are known to change in expression during the ETP-DN3a stages\(^29,30,31\). For example, both PU.1 and BCL11B receive additional regulatory inputs from Runx/CBF\(\beta\) complexes\(^14,27,30,32,33\). However, RUNX transcription factors (especially RUNX1) are only slightly affected by Notch signals and change expression very little from ETP to DN3 stage 6 and have therefore been omitted. Also, GATA-3 and PU.1 are thought to interact at the protein level in a mutually antagonistic way\(^34,35,36,37,38\); however, the mechanisms regulating thresholds for this effect are not yet well understood, and it therefore has not been included. Also, the inclusion of additional genes that could act as intermediate links for certain regulatory relationships\(^30\) could add biological detail to the framework of the model. However, this minimal model is robust with ranges of parameter values for which three key regulatory genes can mediate the irreversible change in state from multi-potent progenitors to committed T-cell precursors.
Model equations

The concentration levels of TCF-1, GATA-3, BCL11B, and PU.1 are denoted as \([T], [G], [B],\) and \([P]\) respectively. The Notch signalling is referred as \(\text{Notch}\).

The concentration level of TCF-1 is described by the following dynamical equation:

\[
\frac{1}{f_T} \cdot \frac{dT}{dt} = \frac{\eta_1 \cdot \text{Notch} + \eta_2 \cdot [T]^{n_{T,T}} + \eta_3 \cdot [G]^{n_{T,G}}}{1 + \eta_1 \cdot \text{Notch} + \eta_2 \cdot [T]^{n_{T,T}} + \eta_3 \cdot [G]^{n_{T,G}} + \eta_4 \cdot [P]^{n_{T,P}}} - \gamma_T \cdot [T] \tag{S1}
\]

where: \(\eta_i (i=1, \ldots, 4)\) modulate the production of TCF-1; \(\gamma_T\) is the degradation rate (day\(^{-1}\)) scaled by the factor \(f_T\); \(n_{T,T}, n_{T,G},\) and \(n_{T,P}\) are the Hill coefficients (set to 1). We thus define the vector of parameters describing TCF-1 dynamics as \(\Theta_T=[\eta_1, \eta_2, \eta_3, \eta_4, \gamma_T, n_{T,T}, n_{T,G}, n_{T,P}, f_T]\).

Similarly, the concentration level of GATA-3 is described by:

\[
\frac{1}{f_G} \cdot \frac{dG}{dt} = \frac{\delta_1 \cdot \text{Notch} + \delta_2 \cdot [T]^{n_{G,T}}}{1 + \delta_1 \cdot \text{Notch} + \delta_2 \cdot [T]^{n_{G,T}} + \delta_3 \cdot [P]^{n_{G,P}}} - \gamma_G \cdot [G] \tag{S2}
\]

where: \(\delta_i (i=1, \ldots, 3)\) shape the production of GATA-3; \(\gamma_G\) is the degradation rate (day\(^{-1}\)) scaled by the factor \(f_G\); \(n_{G,T}\) and \(n_{G,P}\) are the Hill coefficients (set to 1). The vector of parameters for GATA-3 dynamics is then \(\Theta_G=[\delta_1, \delta_2, \delta_3, \gamma_G, n_{G,T}, n_{G,P}, f_G]\).

Depending on how the Notch signalling, TCF-1, and GATA-3 collaborate to activate BCL11B, the concentration level of BCL11B is described by one of the following equations which cover all the possible combinatorial interactions among the three players.

1. \textbf{Notch OR TCF-1 OR GATA-3 activate BCL11B}

\[
\frac{1}{f_{B,1}} \cdot \frac{dB}{dt} = \frac{\kappa_{1,1} \cdot \text{Notch} + \kappa_{2,1} \cdot [T]^{n_{B,T,1}} + \kappa_{3,1} \cdot [G]^{n_{B,G,1}}}{1 + \kappa_{1,1} \cdot \text{Notch} + \kappa_{2,1} \cdot [T]^{n_{B,T,1}} + \kappa_{3,1} \cdot [G]^{n_{B,G,1}}} - \gamma_{B,1} \cdot [B] \tag{S3}
\]

2. \textbf{(Notch AND TCF-1) OR GATA-3 turn on BCL11B}

\[
\frac{1}{f_{B,2}} \cdot \frac{dB}{dt} = \frac{\kappa_{1,2} \cdot \text{Notch} \cdot [T]^{n_{B,T,2}} + \kappa_{2,2} \cdot [G]^{n_{B,G,2}}}{1 + \kappa_{1,2} \cdot \text{Notch} \cdot [T]^{n_{B,T,2}} + \kappa_{2,2} \cdot [G]^{n_{B,G,2}}} - \gamma_{B,2} \cdot [B] \tag{S4}
\]

3. \textbf{(Notch AND GATA-3) OR TCF-1 switch on BCL11B}

\[
\frac{1}{f_{B,3}} \cdot \frac{dB}{dt} = \frac{\kappa_{1,3} \cdot \text{Notch} \cdot [G]^{n_{B,G,3}} + \kappa_{2,3} \cdot [T]^{n_{B,T,3}}}{1 + \kappa_{1,3} \cdot \text{Notch} \cdot [G]^{n_{B,G,3}} + \kappa_{2,3} \cdot [T]^{n_{B,T,3}}} - \gamma_{B,3} \cdot [B] \tag{S5}
\]

4. \textbf{Notch OR (TCF-1 AND GATA-3) trigger BCL11B}

\[
\frac{1}{f_{B,4}} \cdot \frac{dB}{dt} = \frac{\kappa_{1,4} \cdot \text{Notch} + \kappa_{2,4} \cdot [T]^{n_{B,T,4}} \cdot [G]^{n_{B,G,4}}}{1 + \kappa_{1,4} \cdot \text{Notch} + \kappa_{2,4} \cdot [T]^{n_{B,T,4}} \cdot [G]^{n_{B,G,4}}} - \gamma_{B,4} \cdot [B] \tag{S6}
\]

5. \textbf{Notch AND (TCF-1 OR GATA-3) activate BCL11B}

\[
\frac{1}{f_{B,5}} \cdot \frac{dB}{dt} = \frac{\text{Notch} \cdot (\kappa_{1,5} \cdot [T]^{n_{B,T,5}} + \kappa_{2,5} \cdot [G]^{n_{B,G,5}})}{1 + \text{Notch} \cdot (\kappa_{1,5} \cdot [T]^{n_{B,T,5}} + \kappa_{2,5} \cdot [G]^{n_{B,G,5}})} - \gamma_{B,5} \cdot [B] \tag{S7}
\]
6. TCF-1 AND (Notch OR GATA-3) turn on BCL11B
\[
\frac{1}{f_{B,6}} \cdot \frac{d[B]}{dt} = \frac{[T]^{n_{B,T,6}} \cdot (\kappa_{1,6} \cdot \text{Notch} + \kappa_{2,6} \cdot [G]^{n_{B,G,6}})}{1 + [T]^{n_{B,T,6}} \cdot (\kappa_{1,6} \cdot \text{Notch} + \kappa_{2,6} \cdot [G]^{n_{B,G,6}})} - \gamma_{B,6} \cdot [B] \quad \text{(S8)}
\]

7. GATA-3 AND (Notch OR TCF-1) switch on BCL11B
\[
\frac{1}{f_{B,7}} \cdot \frac{d[B]}{dt} = \frac{[G]^{n_{B,G,7}} \cdot (\kappa_{1,7} \cdot \text{Notch} + \kappa_{2,7} \cdot [T]^{n_{B,T,7}})}{1 + [G]^{n_{B,G,7}} \cdot (\kappa_{1,7} \cdot \text{Notch} + \kappa_{2,7} \cdot [T]^{n_{B,T,7}})} - \gamma_{B,7} \cdot [B] \quad \text{(S9)}
\]

8. Notch AND TCF-1 AND GATA-3 trigger BCL11B
\[
\frac{1}{f_{B,8}} \cdot \frac{d[B]}{dt} = \frac{\kappa_{1,8} \cdot \text{Notch} \cdot [T]^{n_{B,T,8}} \cdot [G]^{n_{B,G,8}}}{1 + \kappa_{1,8} \cdot \text{Notch} \cdot [T]^{n_{B,T,8}} \cdot [G]^{n_{B,G,8}}} - \gamma_{B,8} \cdot [B] \quad \text{(S10)}
\]

The \(\kappa_{ij} (i=1, \ldots, 3; j=1, \ldots, 8)\) modulate the production of BCL11B; \(\gamma_B\) is the degradation rate (day\(^{-1}\)) scaled by the factor \(f_B\); \(n_{B,T,j}\) and \(n_{B,G,j}\) are the Hill coefficients, assuming the value of 1 in case of a monomer, 2 in case of a dimer. We thus define the vector of parameters describing BCL11B dynamics as \(\Theta_B=\left[\kappa_{i,j}, \gamma_B, n_{B,T,j}, n_{B,G,j}, f_B\right]\). As the Hill coefficients \(n_{B,T,j}\) and \(n_{B,G,j}\) can assume 2 possible values (either 1 or 2), the total number of configurations for the BCL11B production rate are 32 (\(8 \times 2^3\)).

The concentration level of PU.1 is described by:
\[
\frac{1}{f_P} \cdot \frac{d[P]}{dt} = \frac{\alpha_1 \cdot [P]^{n_{P,P}}}{1 + \alpha_1 \cdot [P]^{n_{P,P}} + \alpha_2 \cdot [T]^{n_{P,T}} + \alpha_3 \cdot [G]^{n_{P,G}} + \alpha_4 \cdot [B]^{n_{P,B}}} - \gamma_P \cdot [P] \quad \text{(S11)}
\]

where: the \(\alpha_i (i=1, \ldots, 4)\) shape the production of PU.1; \(\gamma_P\) is the degradation rate (day\(^{-1}\)); \(n_{P,P}\), \(n_{P,T}\), \(n_{P,G}\), and \(n_{P,B}\) are the Hill coefficients (set to 1); \(f_P\) is the uniforming parameter. The vector of parameters for PU.1 dynamics is then \(\Theta_P=\left[\alpha_1, \alpha_2, \alpha_3, \alpha_4, \gamma_P, n_{P,P}, n_{P,T}, n_{P,G}, n_{P,B}, f_P\right]\).

Parameter inference

To determine the parameter space for TCF-1, GATA-3, BCL11B, and PU.1, we first calculated the derivate of each gene expression profile. For this purpose, two independent methods (providing similar results) were applied: (1) a parametric approach; (2) an algorithm based on a stochastic regularization method. The first method consisted in fitting each gene profile with a parametric function and determining the derivative from the best parametric function. The second approach is based on the concept of deconvolution, which allows us to simultaneously perform both data regularization and calculation of the derivative.

We then decided to split the parameter estimation problem in two parts: the first considered PU.1 as known from smoothing procedure (i.e. PU.1 as forcing function) and focused on the estimation of the parameters characterising TCF-1, GATA-3, and BCL11B dynamics; the latter fixed TCF-1, GATA-3, and BCL11B to their respective smoothing profiles (that is, TCF-1, GATA-3, and BCL11B as forcing functions) and aimed to determine PU.1 parameter values.
**TCF-1, GATA-3, and BCL11B parameters.** Once the derivative was calculated, each equation describing the dynamics of TCF-1, GATA-3, and BCL11B (Eqs. S1-S10) was re-written to extrapolate the profile of Notch signalling (generally known except for its multiplying parameter). E.g. for TCF-1

\[ \eta_1 \cdot Notch = -\eta_2 \cdot [T]^{n_T,T} - \eta_3 \cdot [G]^{n_T,G} - \left(1 + \eta_4 \cdot [P]^{n_T,P} \cdot \frac{d[T]}{dT} \cdot \frac{1}{f_T} + \gamma_T \cdot [T]\right) \frac{d[T]}{dT} \cdot \frac{1}{f_T} + \gamma_T \cdot [T] - 1 \] (S12)

The requirement that Notch signalling activity is positive in ETP-DN3a stages\(^{40,41}\) implied a number of inequalities that led to define the lower and upper bounds for most of the unknown parameters. For TCF-1 parameters \(\Theta_T\) the inequalities are

\[ 0 < \eta_2 < \frac{-\eta_3 \cdot [G]^{n_T,G} - a \cdot (1 + \eta_4 \cdot [P]^{n_T,P})}{[T]^{n_T,T}} \] (S13)

\[ 0 < \eta_3 < \frac{-a \cdot (1 + \eta_4 \cdot [P]^{n_T,P})}{[G]^{n_T,G}} \] (S14)

\[ 0 < \gamma_T < \frac{1}{[T]} \] (S15)

\[ f_T > \frac{\frac{d[T]}{dT} \cdot \frac{1}{f_T} + \gamma_T \cdot [T]}{1 - \gamma_T \cdot [T]} \] (S16)

where \(a\) is

\[ a = \frac{\frac{d[T]}{dT} \cdot \frac{1}{f_T} + \gamma_T \cdot [T]}{\frac{d[T]}{dT} \cdot \frac{1}{f_T} + \gamma_T \cdot [T] - 1} \] (S17)

As for some parameters the bounds were not defined by the inequalities (e.g. upper bound for \(f_T\)), we performed a preliminary exploration of the parameter space by the Monte Carlo method to determine the missing bounds. This step was performed separately for the sets \(\Theta_T, \Theta_G,\) and \(\Theta_B\). The measurements of GATA-3 half-life in the range 2-4 hours (Figure S1 - panel B) allowed us to refine the bounds for the degradation rate parameter \(\gamma_G\). To determine \(\Theta_B\) each combinatorial configuration for BCL11B production was analysed. The steps described above were run in parallel for the two regularisation methods and the bounds were then merged so that the largest parameter space was defined.

We then modelled the Notch signalling as a sigmoidal function

\[ Notch = \frac{N}{1 + e^{-\alpha \cdot t}} \] (S18)

where \(N\) is the maximum level of Notch signalling and \(\alpha\) is the increasing rate (day\(^{-1}\)). We denote as \(\Theta_N\) the parameter vector \([N, \alpha]\).

Once defined the bounds for the parameter space, a constrained multi-objective optimisation approach based on the Pareto frontier\(^{42}\) was applied to determine the best parameter configurations
for $\Theta_T, \Theta_G, \Theta_B$, and $\Theta_N$ that fit simultaneously TCF-1, GATA-3, and BCL11B time series (for each of the 32 configurations of BCL11B dynamics). The three different objectives were referred to the sum of the squared residuals of TCF-1, GATA-3, and BCL11B respectively. As in general Notch signalling is generally known except for its multiplying parameters (i.e. $\eta_1, \delta_1, \kappa_{1,j}, j=1, \ldots, 8$), we fixed $N$ to 1 during the optimisation. The constraint included the Notch signalling increase between ETP and DN3a stages by a factor in the range $[1,3.5]$. To allow the parameter estimation, the number of data to fit was increased by interpolation in a uniform grid. A multi-objective optimization approach was preferred to a single global optimization to have better fit of the single gene profiles.

The resulting configurations were then filtered according to the 95% intervals of confidence (assuming an error in the data of 25%). Table S2 sums up the description of the parameters mentioned above.

**PU.1 parameters.** The equation describing the dynamics of PU.1 (Eq. S11 was re-written to extrapolate the profile of BCL11B (known expect for its multiplying parameter)

$$\alpha_4 \cdot [B]^n_{P,B} = -1 - \alpha_2 \cdot [T]^n_{P,T} - \alpha_3 \cdot [G]^n_{P,G} - \frac{\alpha_1 \cdot [P]^n_{P,P} \cdot \left(\frac{d[P]}{dt} \cdot \frac{1}{f_P} + \gamma_P \cdot [P] - 1\right)}{\frac{d[P]}{dt} \cdot \frac{1}{f_P} + \gamma_P \cdot [P]} \quad (S19)$$

The requirement that BCL11B be positive implied a number of inequalities that led to define most the lower and upper bounds for the unknown parameters in $\Theta_P$. Monte Carlo method was applied to determine the missing bounds. $n$ (order $10^2$) runs of simulated annealing under the constraints $\alpha_4 < \alpha_2$ and $\alpha_4 < \alpha_3$ (i.e. decreased inhibition power of PU.1 by BCL11B versus TCF-1 and GATA-3) were performed. Finally the 95% intervals of confidence filtering provided the optimal parameter values for the vector $\Theta_P$.

Data regularisations, Monte Carlo simulations, multi-objective optimisation, simulated annealing, and analysis were performed by MATLAB 2014a software (The Mathworks, Natick, MA).

**Steady-state analysis**

To ensure reaching high levels of TCF-1, GATA-3, and BCL11B, but low levels of PU.1, the entire network was simulated to reach the steady state by taking separately each of the $n$ parameter sets for PU.1 dynamics ($\Theta_P$) resulting from the simulated annealing. The parameters describing TCF-1, GATA-3, BCL11B, and Notch signals, i.e. $\Theta_T, \Theta_G, \Theta_B$, and $\Theta_N$, were fixed to their best values (best in terms of adherence to the data, i.e. parameter values that minimise the mean squared differences between model prediction and data points).

**Bifurcation analysis**

Bifurcation analysis with respect to the maximum value of Notch signalling ($N$, arbitrary unit) was performed by the software XPP-Aut, which contains an interface to AUTO-07p (Concordia University, Montreal, Canada). The kinetic parameters were set to the best configuration values (best in terms of adherence to the data).
References


26. Zarnegar MA, Rothenberg EV. Ikaros represses and activates Pu.1 cell-type-specifically through the multifunctional Sfpi1 URE and a myeloid specific enhancer. *Oncogene* 2012; **31**: 4647-4654.


36. Stopka T, Amanatullah DF, Papetti M, Skoultchi AI. Pu.1 inhibits the erythroid program by binding to Gata-1 on DNA and creating a repressive chromatin structure. *The EMBO journal* 2005; 24: 3712-3723.


Tables

Table S1  Means and Standard Deviations (SD) for the winning parameters of PU.1 dynamics.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Description</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_1$</td>
<td>affinity for PU.1 binding on $Pu.1$ locus</td>
<td>1.5E-03</td>
<td>2.2E-03</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>affinity for TCF-1 binding on $Pu.1$ locus</td>
<td>1.3E-03</td>
<td>1.2E-03</td>
</tr>
<tr>
<td>$\alpha_3$</td>
<td>affinity for GATA-3 binding on $Pu.1$ locus</td>
<td>5.3E-03</td>
<td>6.6E-03</td>
</tr>
<tr>
<td>$\alpha_4$</td>
<td>affinity for BCL11B binding on $Pu.1$ locus</td>
<td>1.2E-03</td>
<td>1.2E-03</td>
</tr>
<tr>
<td>$\gamma_P$</td>
<td>PU.1 degradation rate</td>
<td>1.3E-04</td>
<td>7.3E-05</td>
</tr>
<tr>
<td>$f_P$</td>
<td>PU.1 scaling factor</td>
<td>5.4E+04</td>
<td>6.0E+04</td>
</tr>
<tr>
<td>half-life PU.1</td>
<td>(hour)</td>
<td>4.0E+00</td>
<td>1.9E-01</td>
</tr>
</tbody>
</table>

Table S2  Parameter description for TCF-1, GATA-3 and BCL11B dynamics and Notch signalling profile.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\eta_1$</td>
<td>Notch effect on TCF-1</td>
</tr>
<tr>
<td>$\eta_2$</td>
<td>affinity for TCF-1 binding on $Tcf7$ locus</td>
</tr>
<tr>
<td>$\eta_3$</td>
<td>affinity for GATA-3 binding on $Tcf7$ locus</td>
</tr>
<tr>
<td>$\eta_4$</td>
<td>affinity for PU.1 binding on $Tcf7$ locus</td>
</tr>
<tr>
<td>$\gamma_T$</td>
<td>TCF-1 degradation rate</td>
</tr>
<tr>
<td>$f_T$</td>
<td>TCF-1 scaling factor</td>
</tr>
<tr>
<td>$\delta_1$</td>
<td>Notch effect on GATA-3</td>
</tr>
<tr>
<td>$\delta_2$</td>
<td>affinity for TCF-1 binding on $Gata3$ locus</td>
</tr>
<tr>
<td>$\delta_3$</td>
<td>affinity for PU.1 binding on $Gata3$ locus</td>
</tr>
<tr>
<td>$\gamma_G$</td>
<td>GATA-3 degradation rate</td>
</tr>
<tr>
<td>$f_G$</td>
<td>GATA-3 scaling factor</td>
</tr>
<tr>
<td>$\kappa_{i,j}$</td>
<td>parameters that regulate the production of BCL11B</td>
</tr>
<tr>
<td>$\gamma_B$</td>
<td>BCL11B degradation rate</td>
</tr>
<tr>
<td>$f_B$</td>
<td>BCL11B scaling factor</td>
</tr>
<tr>
<td>$N$</td>
<td>Maximum level of Notch signalling</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>Increasing rate of Notch signalling</td>
</tr>
</tbody>
</table>