Data supplement

Computational modeling of T-cell formation kinetics: output regulated by initial proliferation-linked deferral of developmental competence

Erica Manesso¹, Vijay Chickarmane², Hao Yuan Kueh², Ellen V. Rothenberg², Carsten Peterson¹

¹Computational Biology & Biological Physics, Department of Astronomy and Theoretical Physics, Lund University, Lund, Sweden

²Division of Biology, California Institute of Technology, Pasadena, CA, USA

In vivo data

To investigate the dynamics of DN1 progenitors, data published by Porritt *et al.* [1] were exploited. Briefly, in this study purified bone marrow progenitors (generally 3-5 $\cdot 10^5$ at > 98% purity) were intravenously transplanted into nonirradiated CD45-congenic recipients. At days 7, 8, 9, 10, 11, 12, 14, 15, 19, 20 after transplant, recipient mice were killed (one or more recipients per time point) and the developmental stages of intrathymic progeny derived from transplanted cells were determined. Dots in Figure 2 (reproduced from [1]) show time series of the percentage of donor cells at DN1 (CD4⁻8⁻25⁻44⁺), DN2 (CD4⁻8⁻25⁺44⁺), DN3 (CD4⁻8⁻25⁺44^{lo}), and pDP - equivalent to DN4 - (CD4^{lo}8^{lo}25⁻44^{lo}) differentiation stages. Thymocytes spent a significant period (9-11 days) at the DN1 stage. After this period, DN2 cells began to appear and their percentage among donor cells peaked at around day 12-13. DN3 population started to accumulate at around day 11 and crested at around day 15, while precursors of DP cells (pDP) began to appear at around day 15.

These dynamical data were combined with static information about DN1 cells, e.g. the ~ 10 divisions undergone at this stage [2] and the $2 \cdot 10^4$ fold increase in cell number between DN1 cells at day circa 0 and DN3 cells at day 14 [2].

Model framework for thymocytopoiesis

Assuming each DN1, DN2, DN3 or pDP cell can commit to differentiate to the next stage, die or proliferate (determine to undergo another cell cycle without differentiating) within its cell cycle time, the sum of the probabilities of the three events must equal one at each generation. The cell cycle time characterizes a cell type: T_{DN1} is the cell cycle time for DN1 cells; T_{DN2} is the cell cycle time for DN2 cells; T_{DN3} is the cell cycle time for DN3 cells; T_{pDP} is the cell cycle time for pDP cells. Deterministic population models are used to model all compartments with dynamics defined below.

DN1pre. These cells commit into DN1 cells (generation 0). The dynamic equation for N_{DN1pre} , the number of cells in DN1pre compartment, is then:

$$\frac{dN_{DN1pre}(t)}{dt} = -\frac{N_{DN1pre}(t)}{\tau} + n \cdot \delta(t)$$
(S1)

DN1. In the generation 0 compartment $(N_{DN1,1})$ these cells are formed by commitment of DN1pre cells and are lost either by death with probability $d_{DN1,1}$, commitment to progress to DN2 cells with probability $c_{DN1,1}$, or proliferation in DN1 cells in generation 1 with probability $p_{DN1,1} = 1 - d_{DN1,1} - c_{DN1,1}$. In equation:

$$\frac{dN_{DN1,1}(t)}{dt} = \frac{N_{DN1pre}(t)}{\tau} - (d_{DN1,1} + c_{DN1,1} + 1 - d_{DN1,1} - c_{DN1,1}) \cdot \frac{N_{DN1,1}(t)}{T_{DN1}}$$
$$= \frac{N_{DN1pre}(t)}{\tau} - \frac{N_{DN1,1}(t)}{T_{DN1}}$$
(S2)

DN1 cells in generation i - 1 compartment $(N_{DN1,i})$ are formed by proliferation of DN1 cells in generation i-2 and are lost either by death with probability $d_{DN1,i}$, commitment to become DN2 cells with probability $c_{DN1,i}$, or proliferation in DN1 cells in generation i with probability $p_{DN1,i} = 1 - d_{DN1,i} - c_{DN1,i}$. In equation:

$$\frac{dN_{DN1,i}(t)}{dt} = \frac{1}{T_{DN1}} \cdot \left[2 \cdot p_{DN1,i-1} \cdot N_{DN1,i-1}(t) - N_{DN1,i}(t)\right]$$
(S3)

Eq. S3 is valid for $i=2, \dots, G+1$. The last generation of DN1 cells permitted within a given version of the model can either die or commit, so the sum of these two probabilities must equal 1.

DN2. These cells (N_{DN2}) are formed by proliferation (probability $1-d_{DN2}-c_{DN2}$, with d_{DN2} probability to die and c_{DN2} probability to commit) or commitment of DN1 cells (sum of the contributions from each generation) and are lost either by death or commitment in DN3 cells. The dynamic equation for N_{DN2} is

then:

$$\frac{dN_{DN2}(t)}{dt} = \sum_{i=1}^{G+1} \frac{c_{DN1,i} \cdot N_{DN1,i}(t)}{T_{DN1}} + \frac{1 - 2 \cdot c_{DN2} - 2 \cdot d_{DN2}}{T_{DN2}} \cdot N_{DN2}(t)$$
(S4)

DN3. Similarly, DN3 cells (N_{DN3}) are formed by proliferation (probability $1 - d_{DN3} - c_{DN3}$, with d_{DN3} probability to die and c_{DN3} probability to commit) or commitment of DN2 cells and are lost either by death or commitment in pDP cells. The dynamic equation for N_{DN3} is then:

$$\frac{dN_{DN3}(t)}{dt} = \frac{c_{DN2} \cdot N_{DN2}(t)}{T_{DN2}} + \frac{1 - 2 \cdot c_{DN3} - 2 \cdot d_{DN3}}{T_{DN3}} \cdot N_{DN3}(t)$$
(S5)

pDP. Finally, pDP cells (N_{pDP}) are formed by proliferation (probability $1-d_{pDP} - c_{pDP}$, with d_{pDP} probability to die and c_{pDP} probability to commit) or commitment of DN3 cells and are lost either by death or further differentiation, to DP cells beyond the spectrum of stages analyzed here. In equation:

$$\frac{dN_{pDP}(t)}{dt} = \frac{c_{DN3} \cdot N_{DN3}(t)}{T_{DN3}} + \frac{1 - 2 \cdot c_{pDP} - 2 \cdot d_{pDP}}{T_{pDP}} \cdot N_{pDP}(t)$$
(S6)

Assuming a constant input I of 50 cells per day into DN1pre compartment, in steady state Eqs. S1-S6 become:

$$N_{DN1pre} = \tau \cdot I \tag{S7}$$

$$N_{DN1,1} = \frac{T_{DN1}}{\tau} \cdot N_{DN1pre} \tag{S8}$$

$$N_{DN1,i} = 2 \cdot p_{DN1,i-1} \cdot N_{DN1,i-1}, \text{ for } i = 2, \cdots, G+1$$
(S9)

$$N_{DN2} = -\frac{T_{DN2}}{1 - 2 \cdot c_{DN2} - 2 \cdot d_{DN2}} \cdot \sum_{i=1}^{G+1} \frac{c_{DN1,i} \cdot N_{DN1,i}}{T_{DN1}}$$
(S10)

$$N_{DN3} = -\frac{T_{DN3}}{1 - 2 \cdot c_{DN3} - 2 \cdot d_{DN3}} \cdot \frac{c_{DN2} \cdot N_{DN2}}{T_{DN2}}$$
(S11)

$$N_{pDP} = -\frac{T_{pDP}}{1 - 2 \cdot c_{pDP} - 2 \cdot d_{pDP}} \cdot \frac{c_{DN3} \cdot N_{DN3}}{T_{DN3}}$$
(S12)

From Eqs. S7-S12 it is clear that to obtain plausible values for the number of cells in DN2, DN3, and pDP compartments, the probabilities to proliferate, commit, die characterizing DN2, DN3, and pDP cells must respect the following inequalities:

$$1 - 2 \cdot c_{DN2} - 2 \cdot d_{DN2} = 2 \cdot p_{DN2} - 1 \le 0 \tag{S13}$$

$$1 - 2 \cdot c_{DN3} - 2 \cdot d_{DN3} = 2 \cdot p_{DN3} - 1 \le 0 \tag{S14}$$

$$1 - 2 \cdot c_{pDP} - 2 \cdot d_{pDP} = 2 \cdot p_{pDP} - 1 \le 0$$
(S15)

The constraints represented by Eqs. S13-S15 state that to reach a plausible steady state, the probability to proliferate for DN2, DN3, and pDP cells must be less than 0.5, i.e. these cells do not have stem cell property as expected.

Models for commitment of DN1 cells

Assuming that all DN1 cells have the same probability to die (i.e. $d_{DN1,i} = d_{DN1}$), we explored two model categories for the relationship between the number of generations spent in DN1 stage and their probability to commit to become DN2 cells: DN1 cells commit (A) from all generations or (B) only towards the end of the cascade.

- A1 Constant probability to commit. In this model the probability to commit is independent from the generation (i.e. number of divisions): $c_{DN1,i} = c_{DN1}$. For the last generation the probability to die is not d_{DN1} , but $d_{DN1,G+1} = 1 c_{DN1}$.
- A2 Probability to commit linearly increasing with the generation. For a DN1 cell in the compartment DN1, *i* the probability to commit is:

$$c_{DN1,i} = \frac{1 - d_{DN1}}{G} \cdot i - \frac{1 - d_{DN1}}{G}$$
(S16)

for $i = 1, \dots, G+1$. To obtain Eq. S16 the commitment for cell in generation 0 is assumed to be zero.

A3 Probability to commit semi-quadratically increasing with the generation. For a DN1 cell in the compartment DN1, i the probability to commit increases with the power of q in this fashion:

$$c_{DN1,i} = \frac{1 - d_{DN1}}{(G+1)^q - 1} \cdot (i^q - 1)$$
(S17)

for $i = 1, \dots, G + 1$. Again, to obtain Eq. S17 the commitment for cell in generation 0 is assumed to be zero.

B1 Geometric probability to commit. For a DN1 cell in the compartment DN1, *i* the probability to commit increases with the power of two as follows:

$$c_{DN1,i} = \frac{1 - d_{DN1}}{2^{G+1} - 2} \cdot (2^i - 2)$$
(S18)

for $i = 1, \dots, G + 1$. Again, to obtain Eq. S18 the commitment for cell in generation 0 is assumed to be zero.

B2 Probability to commit only for the last generation. In this model DN1 cells are allowed to commit only after G divisions, that is: $c_{DN1,i} = 0$ for $i = 1, \dots, G$; $c_{DN1,G+1} = 1 - d_{DN1}$.

Implementation details

Parameter estimation

Table S1 contains a detailed list of unknown parameters when the different models for the commitment of DN1 cells are included in the general model for T-cell progenitors (Eqs. S1-S6). The number of generation G has been fixed to integer values varying between 7 and 12 since, as a first approximation, the number of generations should be around 10, considering a mean transit time of 10 days in DN1 stage [1] and a cell cycle of 1 day for DN1 cells. Models A3, B1 and B2 present an exception: values of G up to 14 and 13, respectively, were explored as model predictions generally improved by increasing G in these cases.

For each model and for each value of G, all unknown parameters were identified by nonlinear least squares on data from [1] using the function *lsqnonlin* implemented in MATLAB software (The Mathworks, Natick, MA). We also tried to derive exact solutions to the simplest models, but the time required to achieve them was too long (e.g. the time needed to solve the set of differential equations describing the model where the probability to commit is constant and the number of generation is 7, was around one hour versus the few seconds required to fit the same set of equations to data). As objective function J the sum of the quadratic difference between each datum and its corresponding model prediction, at different time points, for all the differentiation stages was considered. To this sum, the error of prediction for the $2 \cdot 10^4$ fold increase in cell number between DN1 cells at day circa 0 and DN3 cells at day 14 (weighted with 10^4 given its different order of magnitude with respect to the data) was added, since preliminary runs did not bring to reliable profiles for the cell numbers of T-cell progenitors.

Model selection

Selecting the best model is not only a question of finding the one with the lowest objective function J as at the same time overfitting must be avoided. To take into account both features, it is convenient to use the Akaike index (AIC) as a figure-of-merit [3]. This index, which is defined as:

$$AIC = ln(J) + \frac{2 \cdot \text{number of parameters}}{\text{number of data points}}$$
(S19)

was used as criterion to select the best solutions sharing the same model for the commitment of DN1 cells, but different values of G, as well as the best model in absolute terms. Within the same model for the commitment of DN1 cells, solutions for different values of G sharing a low AIC (i.e. lower than -1, Figure

S1) were averaged.

Robustness

The robustness of the best models was tested on the increase in cell number between the first DN1 subcompartment (DN1,1) and DN3 compartment in steady state defined as:

$$fi_{DN1,1-DN3} = \frac{N_{DN1,1}}{N_{DN3}}$$
(S20)

Eq. S20 can be written as function of the unknown parameters using the steady state equations (Eqs. S1-S6):

$$fi_{DN1,1-DN3} = \frac{T_{DN3}}{T_{DN1}} \cdot \frac{2^{G} \cdot c_{DN2}}{(1 - 2 \cdot c_{DN2} - 2 \cdot d_{DN2}) \cdot (1 - 2 \cdot c_{DN3} - 2 \cdot d_{DN3})} \cdot \\ \cdot \sum_{i=1}^{G+1} c_{DN1,i} \cdot \prod_{k=1}^{i-1} p_{DN1,k}$$
(S21)

where the probability to proliferate for DN1 cells $(p_{DN1,k})$ at different generations depends on the probabilities to die and commit. Eq. S21 reveals that the quantity $f_{i_{DN1,1-DN3}}$ is a function of the number of generations G and all, but the following parameters: the time of exit from DN1pre, τ ; the input, I; the cell cycle of DN2 cells, T_{DN2} .

The sensitivity was then calculated according to Savageau formula considering a variation of $\pm 15\%$ in each parameter [4] with the intent to test how small variations in the parameters affect the increase in cell number between the first DN1 sub-compartment and DN3 compartment.

Mean transit times

According to the indicator dilution theory [5], the mean transit time in a system is the first order moment of the distribution function of transit times characterizing that system. In our system, the distribution function of transit times for different stages i of thymocytopoiesis, can be defined as:

$$h_i(t) = \frac{N_i(t)}{\int_0^{+\infty} N_i(t)dt}$$
(S22)

where N_i is the number of cells in the population i.

The mean transit time for DN1pre cells MTT_{DN1pre} is then:

$$MTT_{DN1pre}(t) = \int_0^{+\infty} t \cdot h_{DN1pre}(t)dt$$
(S23)

For the other populations, e.g. DN1, DN2, DN3, and pDP, to the first order moment of the corresponding distribution function of transit times the mean transit times spent in the precedent stages must be subtracted as the donor cells are injected in DN1pre compartment and the time window is referred as days after transplant. The mean transit times for these populations are then:

$$MTT_{DN1}(t) = \int_0^{+\infty} t \cdot h_{DN1}(t)dt - MTT_{DN1pre}$$
(S24)

$$MTT_{DN2}(t) = \int_{0}^{+\infty} t \cdot h_{DN2}(t) dt - MTT_{DN1pre} - MTT_{DN1}$$
(S25)

$$MTT_{DN3}(t) = \int_{0}^{+\infty} t \cdot h_{DN3}(t)dt - MTT_{DN1pre} - MTT_{DN1} - MTT_{DN2}$$
(S26)

$$MTT_{pDP}(t) = \int_{0}^{+\infty} t \cdot h_{pDP}(t) dt - MTT_{DN1pre} - MTT_{DN1} - MTT_{DN2} - MTT_{DN3}$$
(S27)

This theory is valid as long as the time points are enough to see the disappearance of the donor cells in a particular differentiation stage.

In vitro data

Mice. C57BL/6 mice bred and maintained in the Caltech Laboratory Animal Facility were used for these experiments. 4-6 week old male or female mice were used. All animal protocols were reviewed and approved by the Animal Care and Use Committee of the California Institute of Technology.

DN thymocyte purification. Purification of DN thymocytes from mouse thymus was performed as previously described [6]. Briefly, single-cell thymocyte suspensions were prepared, stained for mature cell markers with biotinylated antibodies [CD8a (53-6.7), TCR γ δ (GL3), TCR β (H57-597), Gr1 (R86.8C5), Ter119 (Ter119), CD122 (5H4), NK1.1 (PK136), and CD11c (N418)], incubated with streptavidin-coated magnetic beads and passed through a magnetic column (Miltenyi Biotec, Auburn, CA). Eluted cells were stained [c-Kit-e450 (2B8), CD135 (Flt3)-PE (A2F10), CD25-APC (PC61.5), Streptavidin-PerCPCy5.5, 7-aminoactinmycin D (Invitrogen, Eugene, OR)] and sorted using a FACSAria with Diva software (BD Biosciences, San Jose, CA) for the following populations: Flt3⁺ DN1

(Lin⁻cKit⁺CD25⁻Flt3⁺), Flt3⁻ DN1 (Lin⁻cKit⁺CD25⁻Flt3⁻)

and DN2 (Lin⁻cKit⁺CD25⁻). Sorted cell populations were then stained at 37degC for 5 minutes with Cell-Trace Violet Proliferation Dye (Invitrogen, Eugene, OR). Stained cells were either analyzed immediately using flow cytometry to obtain initial CellTrace Violet fluorescence intensity levels, or cultured on Op9-DL1 stromal cell monolayers.

Cell culture. Culture of DN thymocytes on Op9-DL1 stromal cells was performed as previously described [6]. Briefly, Op9-DL1 cells were plated on 96 well plates one day before the start of the DN culture at a density of 10000 per well. Sorted DN thymocytes were then cultured on Op9-DL1 stromal cell layers at a density of 500/well, and supplemented with 5 ng/mL IL-7 and Flt3L (Peprotech, Rocky Hill, NJ). For flow cytometric analysis, cultured cells were then disaggregated, resuspended, stained [CD44-FITC (CIM7), cKit-PE (2B8), CD45-APC (30-F11), CD25-APC-e780 (PC61.5)], and filtered through a nylon mesh to remove cell clumps. Samples were analyzed using a MacsQuant flow cytometer (Miltenyi Biotec, Auburn, CA), and analyzed after data acquisition using FlowJo flow cytometry analysis software (Tree Star Inc., Ashland, OR).

References

- Porritt HE, Gordon K, Petrie HT (2003) Kinetics of steady-state differentiation and mapping of intrathymic-signaling environments by stem cell transplantation in nonirradiated mice. J Exp Med 198:957-62.
- [2] Petrie HT, Zúñiga-Pflücker JC (2007) Zoned out: functional mapping of stromal signaling microenvironments in the thymus. Annu Rev Immunol 25:649-79.
- [3] Akaike H (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19:71623.
- [4] Savageau MA (1971) Parameter sensitivity as a criterion for evaluating and comparing the performance of biochemical systems. *Nature* 229:542-4.
- [5] Zierler KL (1961) Theory of the use of arterio-venous concentration differences for mea- suring metabolism in steady and nonsteady states. J Clin Invest 40:2111-25.
- [6] Yui MA, Feng N, Rothenberg EV (2010) Fine-scale staging of T cell lineage commitment in adult mouse thymus. J Immunol 185:284-93.

Figure legends

Figure S1 - Akaike index as function of generation number. Akaike index AIC as function of number of generations G in constant (red), linear (yellow), semi-quadratic (green), geometric (blue), and "only last generation" (black) models for commitment of DN1 cells.

Figure S2 - Donor thymocyte cell number. Donor thymocyte cell number at (A) DN1pre, (B) DN1, (C) DN2, (D) DN3, and (E) pDP stages in constant (red), linear (yellow), semi-quadratic (green), geometric (blue), and "only last generation" (black) best models in terms of G, the number of generations.

Figure S3 - Time profiles for different sub-populations of DN1 cells. Time profiles for number of donor cells in each sub-population of DN1 cells in (A) geometric and (B) "only last generation" best models in terms of G, the number of generations.

Figure S4 - Distribution of DN1 cells according to undergone number of divisions. Distribution of DN1 cells in DN1 sub-populations according to generation number in the best models for DN2 progression, i.e. (A) commitment geometrically increasing with the generation and (B) commitment only for the last generation.

Figure S5 - Parameter variability. Parameter change in the configurations closest to the best (*optimum*) in terms of Akaike index in (A) geometric and (B) "only last generation" models.

Figure S6 - Sensitivity. Sensitivity on the increase in cell number between the first DN1 subpopulation and DN3 population in steady state obtained by a variation of $\pm 15\%$ in each parameter in geometric (A) and "only last generation" (B) best models in terms of G, the number of generations.

Figure S7 - Probabilities for DN1 cells in other models. Probabilities to (A) die, (B) commit, and (C) proliferate for DN1 cells as function of the number of generations (i.e. cell divisions) in geometric commitment/decreasing death (red), geometric with short cell cycle time T_{DN1} (yellow), "only last generation" with short cell cycle time T_{DN1} (green), exponential (blue), and Weibull (black) best models in terms of G, the number of generations.

Figure S8 - Data versus model predictions in other models. Model predictions for (A) DN1pre, (B) DN1, (C) DN2, (D) DN3, and (E) pDP cells in in geometric commitment/decreasing death (red), geometric with short cell cycle time T_{DN1} (yellow), "only last generation" with short cell cycle time T_{DN1} (green), exponential (blue), and Weibull (black) best models in terms of G, the number of generations. Data (black dots) reproduced from [1].

Figure S9 - Flow cytometry analysis of thymocyte populations. (A) Flow cytometry plots

showed sort gates for DN1 (Lin⁻cKit^{hi}CD25⁻) and DN2 (Lin⁻cKit^{hi}CD25⁺) thymocytes. (B)-(E) Analysis of sorted DN thymocytes labeled with CellTrace Violet dye. Flow cytometry plots showing CellTrace Violet versus CD25 levels for (A) Flt3⁺DN1 cells, (B) Flt3⁻DN1 cells, and (C) DN2 cells. (E) Histogram showing Flt3 levels for the sorted Flt3⁺DN1 (black), Flt3⁻DN1 (red) and DN2 (green) cell populations.

| Description | Symbol | Unit | Constant | Linear | Semi-quadratic | Geometric | Only last generation |
|---|-----------|-----------------|----------|--------|----------------|-----------|----------------------|
| probability to commit for DN1 cells | c_{DN1} | [dimensionless] | yes | no | no | no | no |
| probability to commit for DN2 cells | c_{DN2} | [dimensionless] | yes | yes | yes | yes | yes |
| probability to commit for DN3 cells | c_{DN3} | [dimensionless] | yes | yes | yes | yes | yes |
| probability to commit for pDP cells | c_{cDP} | [dimensionless] | yes | yes | yes | yes | yes |
| probability to die for DN1 cells | d_{DN1} | [dimensionless] | yes | yes | yes | yes | yes |
| probability to die for DN2 cells | d_{DN2} | [dimensionless] | yes | yes | yes | yes | yes |
| probability to die for DN3 cells | d_{DN3} | [dimensionless] | yes | yes | yes | yes | yes |
| probability to die for pDP cells | d_{dDP} | [dimensionless] | yes | yes | yes | yes | yes |
| cycle time for DN1 cells | T_{DN1} | [day] | yes | yes | yes | yes | yes |
| cycle time for DN2 cells | T_{DN2} | [day] | yes | yes | yes | yes | yes |
| cycle time for DN3 cells | T_{DN3} | [day] | yes | yes | yes | yes | yes |
| cycle time for pDP cells | T_{pDP} | [day] | yes | yes | yes | yes | yes |
| number of cells entering DN1pre as a <i>bolus</i> | n | [cell] | yes | yes | yes | yes | yes |
| time of exit from DN1pre | τ | [day] | yes | yes | yes | yes | yes |
| degree of non linearity for DN1 commitment | q | [dimensionless] | no | no | yes | no | no |

Table S1: List of unknown parameters in the models for commitment of DN1 cells.

Figures



Figure S1



Figure S2



Figure S3



Figure S4



Figure S5



Figure S6



Figure S7



Figure S8



Figure S9