

# Cyton 2

January 25, 2021

## 1 Introduction

1 T and B lymphocytes are crucial components in the adaptive immune system that response to foreign  
2 pathogens by proliferating and differentiating to effector and memory cell types. Over the years, this re-  
3 sponse has been extensively studied not only at a population level by utilising fluorescence-activated cell  
4 sorting (FACS), but also at a single-cell level by an optical microscope (Weston and Parish 1990; Lyons and  
5 Parish 1994). However, analysing such non-linear dynamics due to cell's doubling nature in its blast period  
6 and formation of heterogenous outcomes has been known to be a challenging task. As a result, the impor-  
7 tance in mathematical modelling has been growing, for its practicality and ability to quantitatively extract  
8 inferential information about the system. To approach this problem, one could formalise mathematics with a  
9 deterministic view (Reiner and Adams 2014) and/or typically embrace probabilistic approach (Nordon, Ko,  
10 et al. 2011; Buchholz et al. 2013; Hodgkin, Dowling, and Duffy 2014) in order to capture heterogeneity in  
11 the data.

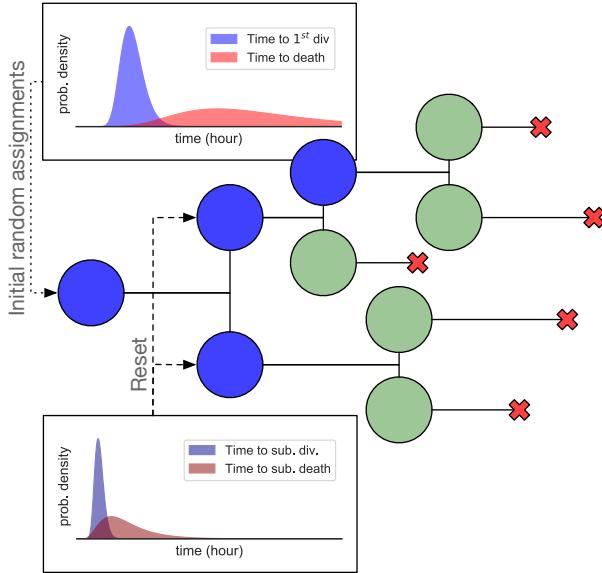
12 On the experimental side, many features of lymphocyte's response can be induced and follow the cells'  
13 division progression as well as survival *in vitro*. Such steady advancement in both experimental and theo-  
14 retical work has enabled huge amount of data to be gathered and inferred about underlying mechanisms of  
15 cell proliferation. For example, in papers Gett and Hodgkin 2000; Boer and Perelson 2005; Ganusov et al.  
16 2005; Asquith et al. 2006; Hawkins, Turner, Dowling, et al. 2007; Luzyanina et al. 2007; Hyrien and Zand  
17 2008; Banks, Sutton, et al. 2011; Miao et al. 2011; Shokhirev and Hoffmann 2013 developed mathematical  
18 frameworks to extract division parameters (e.g. division rate) of lymphocyte population dynamics from  
19 carboxyfluorescein diacetate succinimidyl ester (CFSE) data. Furthermore, studies like in Hawkins, Turner,  
20 Wellard, et al. 2013; Marchingo, Kan, et al. 2014; Heinzel et al. 2016 achieved important deductions with this  
21 approach in answering how the cells process and integrate signals. While these analyses helped identifying  
22 key features of lymphocyte proliferation from population level, how each lymphocyte contributes at family  
23 level still remains largely unknown. To address this question, a series of papers Duffy and Subramanian  
24 2008; Hawkins, Markham, et al. 2009; Markham et al. 2010; Wellard et al. 2010; Duffy, Wellard, et al. 2012;  
25 Shokhirev and Hoffmann 2013; Dowling et al. 2014; Shokhirev, Almaden, et al. 2015; Mitchell et al. 2018

26 studied the system in theoretical and experimental perspectives and reported a highly synchronised pattern  
27 of the response for B cells using time-lapse microscope data. Moreover, recent papers [Marchingo, Prevedello, et al. 2016](#); [Horton et al. 2018](#) concluded similar results for T cells using a novel high-throughput multiplex  
28 assay.

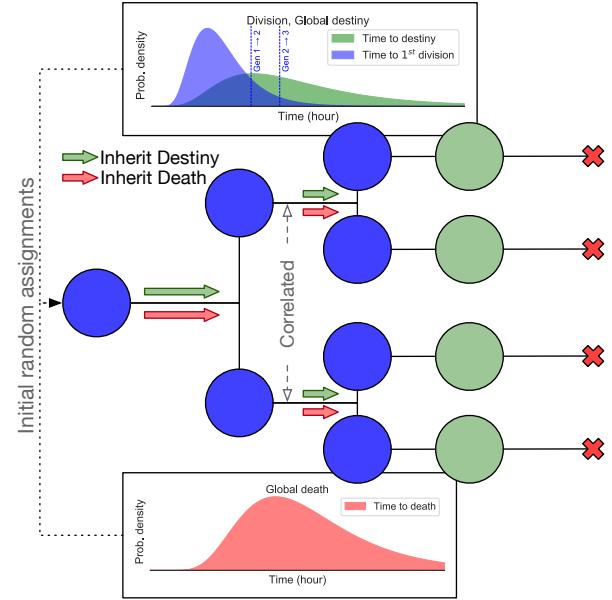
29 [Hawkins, Turner, Dowling, et al. 2007](#) proposed that these features of lymphocyte proliferation can be  
30 encapsulated as a primitive autonomous program response, with activation leading to a number of divisions, a  
31 loss in motivation to divide and eventual death, which they called Cyton model. Since its first introduction,  
32 it has been used to construct efficient analysis pipelines for FACS data ([Hawkins, Hommel, et al. 2007](#);  
33 [Shokhirev and Hoffmann 2013](#)) and to obtain biological insights from both B and T cells ([Hawkins, Turner,](#)  
34 [Wellard, et al. 2013](#); [Marchingo, Kan, et al. 2014](#)). However, a number of experimental investigations have  
35 examined closely the assumptions of the model. While the assumptions of a resetting of division times and  
36 the independence of three controllers of fates, division, death and destiny (time at which a cell returns to  
37 quiescent state) were fit to many situations, two important changes were suggested by analyses from time-  
38 lapse microscope experiments of lymphocytes as well as results published in [Heinzel et al. 2016](#). The first  
39 is that division destiny is not controlled as a "counter", but works instead, as a "timer" by Myc protein  
40 ([Heinzel et al. 2016](#)). An initial burst of Myc expression gradually decreases in a timely manner independent  
41 of number of divisions and acts as a license for cells to divide until they reach critical threshold. Importantly,  
42 the timing of the license for division to continue was found to be passed down through each generation  
43 without being affected. In the second major change compared to the original Cyton model, the time to  
44 death is also found to be programmed early in the stimulated cell and passed to descendants without being  
45 altered in the analogous manner to the transmission of the division destiny time ([Heinzel et al. 2016](#)). As a  
46 result, the fate of whole family members can be highly concordant while allowing significant variation of the  
47 times between families from a homogeneous cell type. Collectively, these findings motivate an alteration to  
48 the model of the cellular processes as well as re-examination of data for revising the model to better capture  
49 cell dynamics.

50 To date with the exception of quorum sensing models ([Weber and Buceta 2013](#); [Kannan and Saini 2018](#)),  
51 essentially all of the lymphocyte proliferation models employed assumption that newly born cell's fate is  
52 independent of its family's history, and, of the lifetimes exponentially distributed ([Smith and Martin 1973](#);  
53 [Nordon, Nakamura, et al. 1999](#); [Revy et al. 2001](#); [Ganusov et al. 2005](#); [Yates, Chan, et al. 2007](#); [Lee et al.](#)  
54 [2009](#); [Banks, Thompson, et al. 2012](#); [Hasenauer, Schittler, and Allgöwer 2012](#); [Mazzocco, Bernard, and Pujo-](#)  
55 [Menjouet 2017](#)), with a few notable exceptions (e.g. Cyton model, [Hyrien, Chen, and Zand 2010](#); [Wellard](#)  
56 [et al. 2010](#); [Zilman, Ganusov, and Perelson 2010](#); [Shokhirev, Almaden, et al. 2015](#); [Yates, Ford, and Mort](#)  
57 [2017](#)). These assumptions are adopted, not because they are consistent with experimental data from, for  
58 example, filming, FACS and multiplex, but for reasons of parsimony, model identifiability and computational

### A Cyton1 (Hawkins et al. 2007)



### B Cyton2: Correlation and Inheritance Features



### C Concept of Cyton2 & Clonal Collapse of a Family Tree

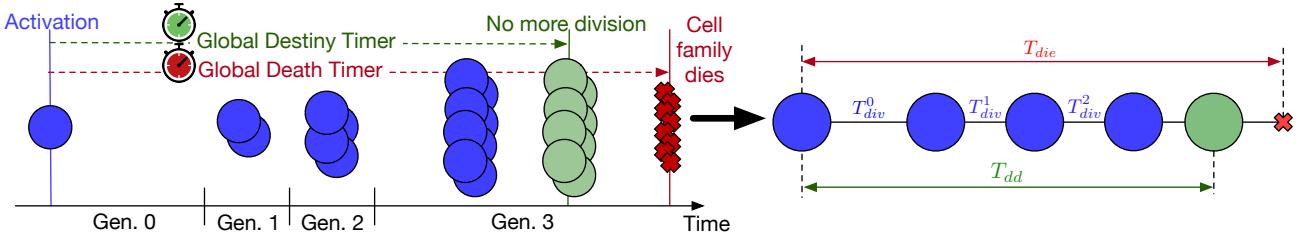


Figure 1: Overview of the two Cyton models. (A) The original Cyton model proposed in Hawkins, Turner, Dowling, et al. 2007. The times to divide and to death get reset every time the cell divides, then draw new random times for the offsprings. The cells cease to divide based on "counter" mechanism. (B) The new Cyton model that includes correlation of division times between siblings, and inheritance feature of time to death and to division destiny (time at which cells return to quiescent state). As opposed to the "counter" mechanism, the cells stop dividing at designated destiny time. (C) As a consequence of the correlation and inheritance, the family tree is highly concordant, depicted by the diagram. By exploiting this property, a family tree is summarised by substituting average values of the times at each generation. An example of clonally collapsed family tree and its key variables is shown.

60 ease of fitting. Here, we introduce a new model, called Cyton 2, where family correlation are included. This  
 61 is done, however, in a way by which identifiability is improved while computational burden is not increased.

62 In this paper, we examine datasets from direct filming of B and T cell families and interrogate these data  
 63 to investigate timed outcomes, measure correlations in each alternative fates and determine suitable para-  
 64 metric distribution classes for collection of those fates in preparation for developing improved mathematical  
 65 descriptions.

## 2 Results

### 2.1 A Stochastic Model of Lymphocyte Response

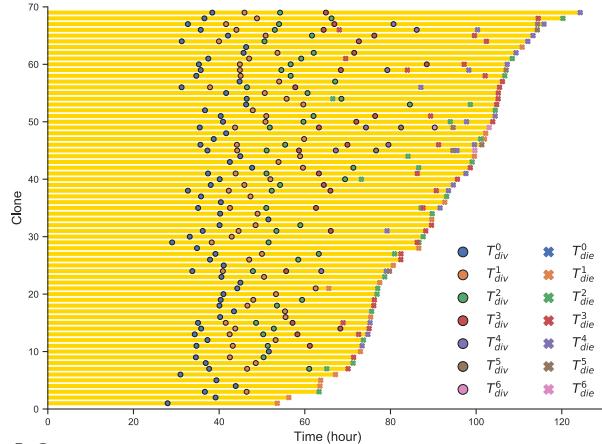
66 The original Cyton model (Hawkins, Turner, Dowling, et al. 2007) postulates that the regulation of kinetic  
 67 profile of expansion, cessation and contraction is a result of stochastic competition between two independent  
 68 cellular timers: time to division and time to death. One of the main features of the model assumes that the

69 division and death times get reset at every round of division and adopt new values randomly drawn from  
70 a lognormal distribution. Additionally, the term division destiny was introduced as the division counting  
71 mechanism to arrest cells that have lost motivation to divide, which was encoded with an extra parameter  
72 called the progressor fraction. As a result, a single founder cell may form various shapes of family trees due  
73 to variations in the timed operation at each generation (Fig. 1A).

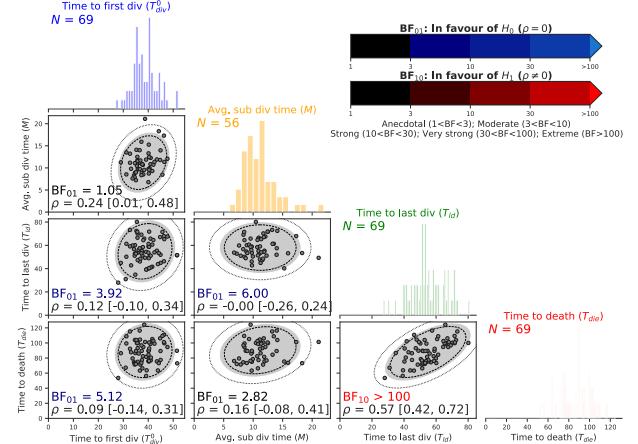
74 While the Cyton model reasonably well captures the cell population dynamics, recent findings in Heinzel  
75 et al. 2016 suggested a modification to the model. In particular, the destiny works as the timed mechanism,  
76 instead of the counting mechanism, similar to the division and death timers. Moreover, the destiny and  
77 death operations were programmed after lymphocyte activation and applied globally within a family, which  
78 means that each of these acts like a timer that is set once at the beginning of the operation and carried  
79 through division round (Heinzel et al. 2016). Consequently, the family tree of an activated lymphocyte is  
80 expected to be regular (Fig. 1B) and, indeed, was observed to carry out series of synchronous operations for  
81 both B and T cells (Hawkins, Markham, et al. 2009; Marchingo, Prevedello, et al. 2016). Based on these  
82 findings, we formulate a new stochastic model of lymphocyte response using sets of random variables (RVs)  
83 that correspond to global death timer, global destiny timer and division machinery (Fig. 1C).

84 The first RV ( $T_{die}$ ) represents the global death time for a family tree, where the founder cell is randomly  
85 assigned a time at which all progeny cells die. The second RV ( $T_{dd}$ ) is the global destiny time after which  
86 no cells are allowed to progress to next generation and awaits only for death if  $T_{dd} < T_{die}$  otherwise destiny  
87 event is censored by death. Lastly, the third RV ( $T_{div}^k$ ) denotes for the time to division of cells at generation  
88  $k$ . We define all three variables to span from the addition of stimulus, which is typically at  $t = 0$  in the  
89 experiment. However, the time to first division ( $T_{div}^0$ ) and the subsequent division time ( $M_k := T_{div}^{k+1} - T_{div}^k$   
90 for  $k > 0$ ) were reported to be two different quantities, for a lymphocyte typically takes longer time to  
91 complete its first division but, having done so, traverses subsequent division rounds at much faster and  
92 consistent rate (Gett and Hodgkin 1998; Gett and Hodgkin 2000). In essence, we call the model comprises  
93 of  $(T_{div}^0, M_k, T_{dd}, T_{die})$  as Cyton 2 model, a variant to the original Cyton model (Hawkins, Turner, Dowling,  
94 et al. 2007) that extends to include not only correlation of division times between progeny cells but also  
95 inheritance of death and destiny times within a family.

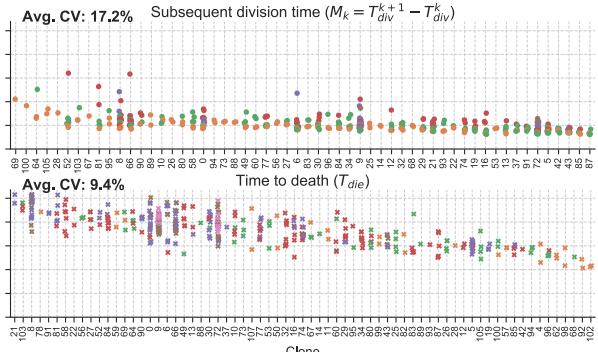
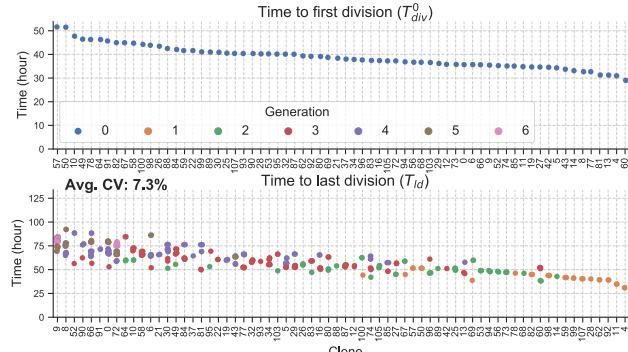
### A1 B cell CpG-stimulated: Clonal Collapse



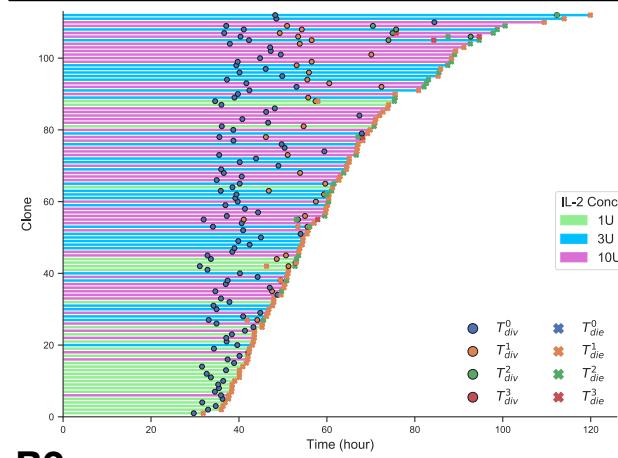
### A3 B cell CpG-stimulated: Correlation



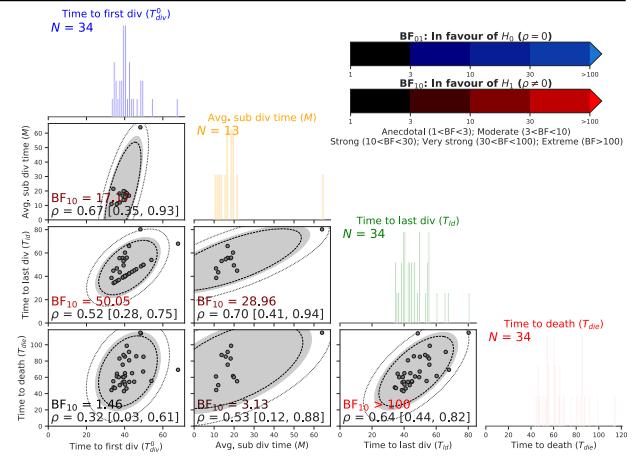
### A2 B cell CpG-stimulated: Event Cascade Plot



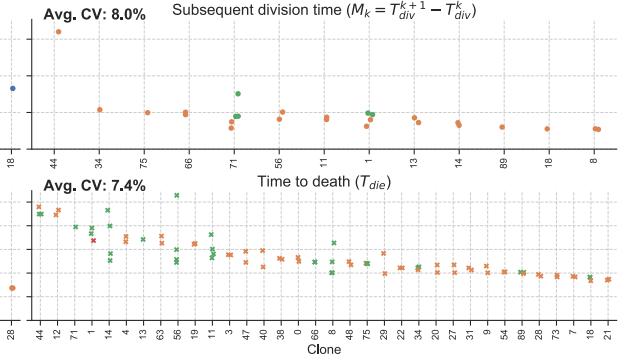
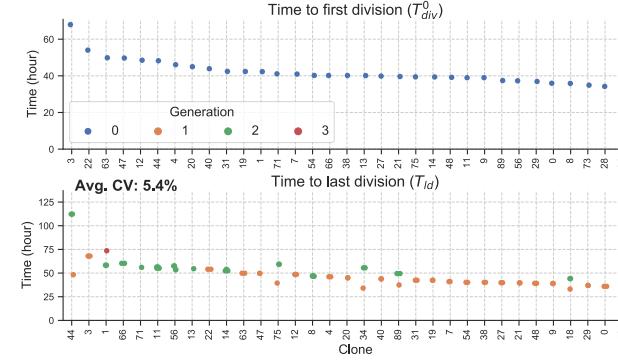
### B1 CD8 T cell [1U, 3U, 10U IL-2]: Clonal Collapse



### B3 CD8 T cell [3U IL-2]: Correlation



### B2 CD8 T cell [3U IL-2]: Event Cascade Plot



**Figure 2: Extracting times to fates from CpG-stimulated B cells and CD8 T cells in the presence of 1U, 3U or 10U IL-2. (A1, B1)** Clonally collapsed family trees from filtered data (lost cells not shown). **(A2, B2)** Measured times to events of the cells are shown and ordered by their average over all generations per clone. **(A3, B3)** Correlation coefficient ( $\rho$ ) from bivariate normal distribution with 95% credible interval is reported for each pair. 90% (---), 95% (■) and 99% (.....) density regions are plotted over the data. Given null,  $H_0: \rho = 0$ , and alternative,  $H_1: \rho \neq 0$ , hypotheses, Bayes factor ( $BF_{01} = 1/BF_{10}$ ) is calculated. If the data is favoured under  $H_0$ , then it is  $BF_{01}$  times more favoured than  $H_1$  (blue-scale), and vice versa (red-scale). Distributions of the times are shown in the diagonal panels collated into 1h time intervals.

## 2.2 Filming Data Supports Independent Operation of the Timers

Cell Type	Stim.	Bayes Factor (01/10) & Correlation Coefficient ( $\rho$ [CI])					
		$(T_{div}^0, M)$	$(T_{div}^0, T_{ld})$	$(T_{div}^0, T_{die})$	$(M, T_{ld})$	$(M, T_{die})$	$(T_{ld}, T_{die})$
B	CpG	1.05 (01) 0.24 [0.01, 0.48]	3.92 (01) 0.12 [-0.10, 0.34]	5.12 (01) 0.09 [-0.14, 0.31]	6.00 (01) 0.0 [-0.26, 0.24]	2.82 (01) 0.16 [-0.08, 0.41]	>100 (10) 0.57 [0.42, 0.72]
CD8 T	1U	1.70 (01) -0.01 [-0.87, 0.85]	39.09 (10) 0.55 [0.29, 0.79]	2.53 (01) 0.19 [-0.16, 0.53]	1.53 (01) -0.15 [-0.97, 0.70]	1.43 (01) 0.20 [-0.67, 0.98]	>100 (10) 0.59 [0.36, 0.81]
(+IL2)	3U	17.17 (10) 0.67 [0.35, 0.93]	50.05 (10) 0.52 [0.28, 0.75]	1.46 (10) 0.32 [0.03, 0.61]	28.96 (10) 0.70 [0.41, 0.94]	3.13 (10) 0.53 [0.12, 0.88]	>100 (10) 0.64 [0.44, 0.82]
	10U	7.55 (10) 0.57 [0.23, 0.87]	>100 (10) 0.58 [0.40, 0.75]	15.39 (10) 0.40 [0.18, 0.62]	>100 (10) 0.76 [0.55, 0.94]	1.56 (10) 0.42 [0.02, 0.79]	>100 (10) 0.63 [0.46, 0.78]

Table 1: **Test correlation of every pair of times to fates.** The correlation coefficient calculated from a bivariate normal distribution with its 95% credible interval is shown for CpG-stimulated B cell and CD8 T cell in the presence of 1U, 3U or 10U of IL-2. The same Bayes factor interpretation criterion is used as shown in Fig. 2.

Having defined the variables that shape the model of the lymphocyte response, we examined the relationship between the variables to verify one of the assumptions embedded for the Cyton model: the law of independence (Gett and Hodgkin 2000; Hodgkin 2005). The bulk culture experiments of B and T cells provided initial evidence for this assumption. It was shown by using Mycophenolic Acid in which it can slow down the division time while preserving the destiny and death timers (Heinzel et al. 2016). Here we sought additional and more direct evidence using the imaging and single cell tracking data. In particular, we revisited two datasets for CpG-stimulated B cells published in Hawkins, Markham, et al. 2009 each consists of 108 clones (CpG3) and 88 clones (CpG4), respectively. Additionally, three CD8 T cell datasets were analysed: (i) with 1U, 3U and 10U of IL-2 seeded with 109, 90 and 163 clones, respectively; (ii) in combination of N4,  $\alpha$ CD28 and IL-2 (Costim1); and, (iii) in combination of N4,  $\alpha$ CD28 and IL-12 (Costim2). As an example, one B cell (CpG3) and T cell (IL-2) datasets are selected in this paper (see Supplementary Fig. S1 for all of raw data).

In Fig. 2A1 and B1, we plotted clonally collapsed family trees of filtered clones as a single representative time line stacked vertically for CpG-stimulated B cells (CpG3: 69 clones) and CD8 T cells in the presence of 3U (34 clones) of IL-2. The average times to division and to death are marked at each generation (see Methods). The key Cyton 2 variables are shown in the cascade plot for all cells in a family with the exception of time to division destiny ( $T_{dd}$ ), which was replaced with time to last division ( $T_{ld}$ ) as a proxy measure (Fig. 2A2, B2; Fig. S3 for Costim1 and Costim2). Given the distributions of these measurements, we asked the spreads of subsequent division time ( $M_k$ ), time to last division and time to death ( $T_{die}$ ) within a clone to see if the cells reached fates synchronously as reported in Hawkins, Markham, et al. 2009; Marchingo, Prevedello, et al. 2016; Mitchell et al. 2018. To quantify this, we calculated coefficient of variation (CV) per clone, and then took the average of calculated CVs for each variables. In the order of  $M_k$ ,  $T_{ld}$  and  $T_{die}$ , we identified 17.2%, 7.3% and 9.4% of average CVs for B cells and 8%, 5.4% and 7.4% for T cells. Similar results were shown for a repeat of B cell (73 clones) and CD8 T cell with 1U (28 clones) and 10U (50 clones) of IL-2

119 (Fig. S2). This signifies low variation around the mean times to fates and is consistent with the reported  
 120 synchronous behaviour of the cells. Given this conclusion, we turned our attention to the question of the  
 121 independence of the variables operating at the clone level. For every pairs of the variables, the correlation  
 122 coefficient ( $\rho$ ) and its 95% credible interval were determined from Bayesian inference. Moreover, Bayes  
 123 factors for two competing hypotheses ( $H_0: \rho = 0$  and  $H_1: \rho \neq 0$ ) were calculated (see Methods) (Fig. 2A3,  
 124 B3). The estimates of these quantities are shown in Table 1 (see Supplementary Table S1 for Costim1 and  
 125 Costim2). With the exception of  $(T_{ld}, T_{die})$  pair for all datasets, CpG-stimulated B cells are consistent with  
 126 the independence assumption of the operation of the times. However, we discovered 10 pairs of variables  
 127 were reported with  $BF_{10} > 10$  for T cells, suggesting that there are strong evidence towards correlation.  
 128 This puzzling results were resolved in the following section in which we generated large number of trees *in*  
 129 *silico* under the independence assumption.

### 2.3 Censorship of the Timers

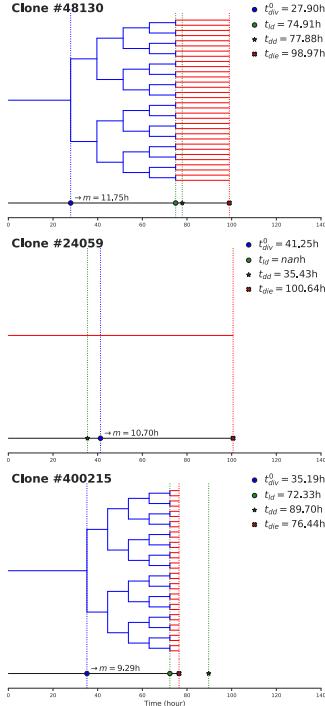
Cell Type	Stim.	Bayes Factor (01/10) of True and Observable values from $10^6$ simulated families											
		$(T_{div}^0, M)$		$(T_{div}^0, T_{ld})$		$(T_{div}^0, T_{die})$		$(M, T_{ld})$		$(M, T_{die})$		$(T_{ld}, T_{die})$	
		True	Obs.	True	Obs.	True	Obs.	True	Obs.	True	Obs.	True	Obs.
B	CpG	>100 (01)	>100 (10)	>100 (01)	>100 (10)	>100 (01)	>100 (01)	>100 (01)	>100 (10)	>100 (01)	4.32 (01)	>100 (01)	>100 (10)
CD8 T (+IL2)	1U	>100 (01)	>100 (10)	>100 (01)	>100 (10)	>100 (01)	>100 (10)	>100 (01)	>100 (10)	>100 (01)	>100 (10)	>100 (01)	>100 (10)
	3U	>100 (01)	>100 (10)	>100 (01)	>100 (10)	>100 (01)	>100 (10)	>100 (01)	>100 (10)	>100 (01)	>100 (10)	>100 (01)	>100 (10)
	10U	>100 (01)	>100 (10)	>100 (01)	>100 (10)	>100 (01)	>100 (10)	>100 (01)	>100 (10)	>100 (01)	>100 (10)	>100 (01)	>100 (10)

Table 2: **Test correlation for the simulated trees.** The Bayes factors were calculated same way as the Table 3. *Blue* indicates more probable under the null hypothesis ( $\rho = 0$ ) and *red* indicates more probably under the alternative hypothesis ( $\rho \neq 0$ ).

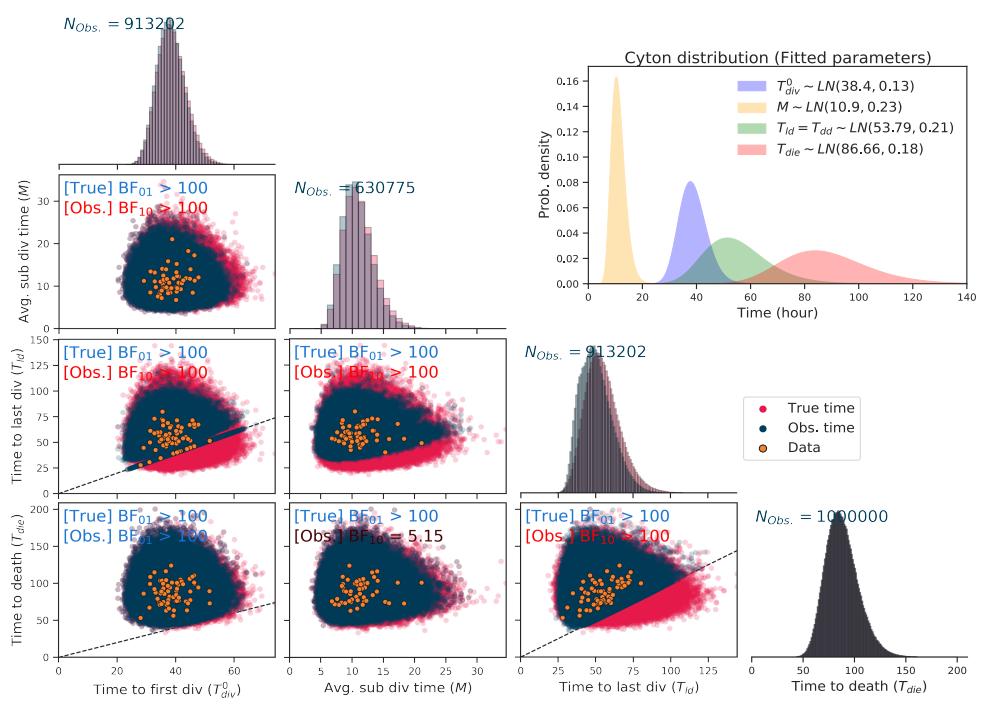
130 In Section 2.2, CpG-stimulated B cells were reported to be more probable under no correlation for most  
 131 of the variable pairs while CD8 T cells had mixed results. The key difference between two datasets is the  
 132 depth of the trees: many of B cells had divided as many as six times whereas T cells had divided at most  
 133 three times (Fig. 2A2, B2). This suggests that many of variables are rendered unobserved or even censored  
 134 for T cells. If the variables are truly independent of each other, any order of times to fates could had been  
 135 assigned to each family. For example, if  $T_{die} < \min(T_{div}, T_{dd})$ , then the time to first division and destiny  
 136 would be censored by the death event. For this reason, the timers may have operated independently, but  
 137 the realisation of those timers are appeared to be necessarily correlated.

138 We simulated Cyton 2-like process via agent-based model (see Methods) to recapitulate the filming  
 139 data. Under the assumption that  $T_{ld} = T_{dd}$ , each of variables was randomly sampled from respective fitted  
 140 lognormal distribution from the data (see Section 2.4). Both observable and true assigned values are plotted  
 141 in Fig. 3A2 for B cells and Fig. 3B1-3 for T cells. It clearly shows that T cells have more censorship than  
 142 that of B cells, which is directly translated from significant overlapping of the distributions. The Bayes

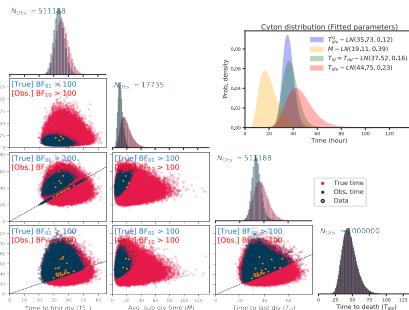
## A1 ABM Realisation



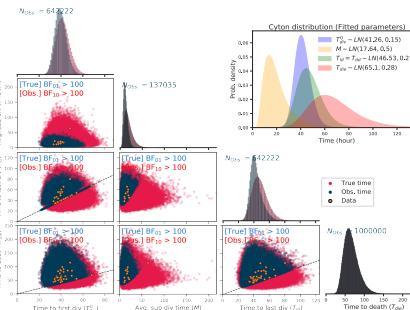
## A2 B cell CpG-stimulated



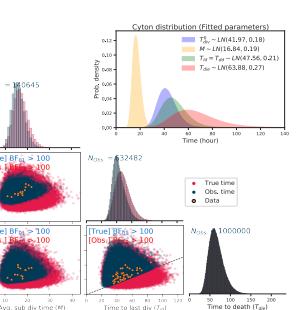
## B1 CD8 T cell [1U IL-2]



## B2 CD8 T cell [3U IL-2]



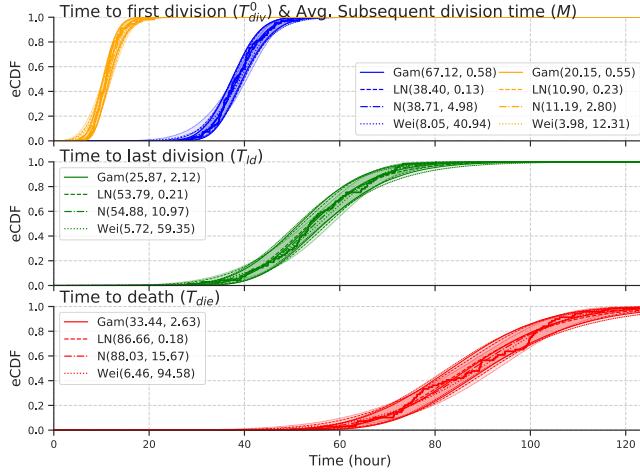
## B3 CD8 T cell [10U IL-2]



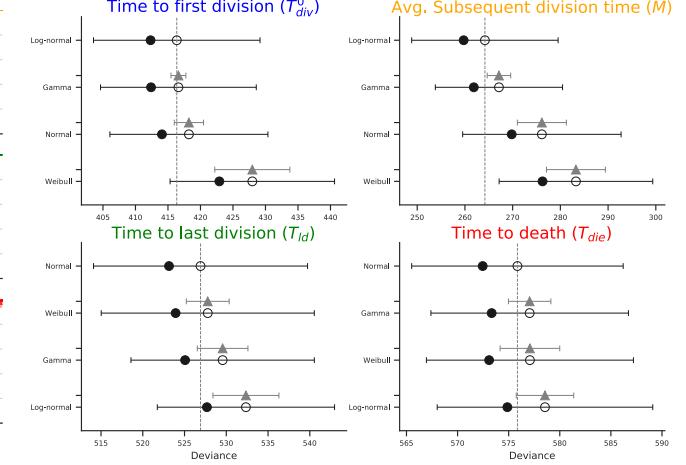
**Figure 3: Simulation under the independence assumption.**  $10^6$  families were simulated via Cyton 2-like ABM given fitted lognormal distributions of  $T_{div}^0$ ,  $M$ ,  $T_{ld}$ ,  $T_{die}$  from respective data. **(A1)** Three example families of ABM realisation for CpG-stimulated B cells: dividing (—) and dying (—) states. The realisations of  $T_{div}^0$  (●),  $M$  ( $m$ ),  $T_{ld}$  (●),  $T_{dd}$  (★) and  $T_{die}$  (●) are annotated on a clonally collapsed line. As features of inheritance and correlation, the cells double in number synchronously whenever division occurs, and likewise, all cells reach destiny and death at the same time. **(A2, B1-3)** Sampled true (●) times to fates for all simulated families and their corresponding observable (●) values are shown in pairs along with the data points (●). The observable and unobservable regions are separated by upper and lower sections of  $y = x$  line (---), respectively. The Bayes factors are reported for the true and observable pairs.

143 factor under the same hypotheses used in Section 2.2 showed that the true values are more probable under  
144 the null hypothesis ( $\rho = 0$ ) as expected while it favours the alternative hypothesis ( $\rho \neq 0$ ) for the observable  
145 values most of times (Table 2). In fact, all of the pairs for the observable values were reported strongly  
146 correlated for T cells as a result of high degree censorship. Interestingly,  $(T_{div}^0, M)$ ,  $(T_{div}^0, T_{ld})$ ,  $(M, T_{ld})$  and  
147  $(T_{ld}, T_{die})$  pairs were reported correlated if we were to measure for larger number of clones from B cells,  
148 otherwise the rest of pairs exhibited lack of correlation as reported directly from the B cell data. In essence,  
149 we illustrated that the timers may well be operating independent to each other but the resulting outcome can  
150 appear correlated. Admittedly this is not a definitive proof of the independence assumption and it remains  
151 a challenging task to prove so, however, we conclude from the simulation that correlation is not from the

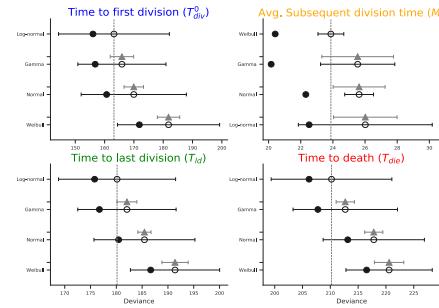
## A1 B cell CpG-stimulated: Compare parametric distribution classes



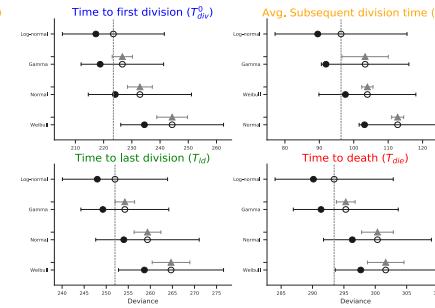
## A2 B cell CpG-stimulated: WAIC Scores



## B1 CD8 T cell [1U IL-2]



## B2 CD8 T cell [3U IL-2]



## B3 CD8 T cell [10U IL-2]

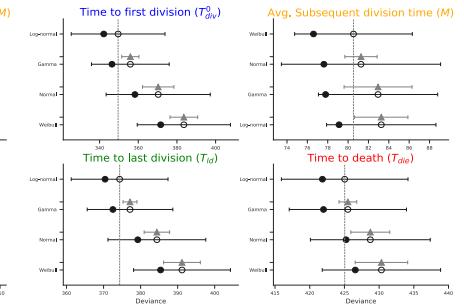


Figure 4: **Best parametric distribution class.** (A1) Empirical CDF of the measured times are overlaid with CDFs of Gamma, Lognormal, Normal and Weibull distributions. 95% confidence bands are plotted by randomly drawing  $10^4$  samples from respective posterior distributions. (A2, B1-3) The in-sample deviance (●), and WAIC scores (○) with 1 standard deviation error bar are shown. The y-axis is sorted from lowest (*top-row*) to highest (*bottom-row*) WAIC score. The lower WAIC score indicates better descriptor of the data. Except the top-ranked one, the value of the difference of WAIC ( $\blacktriangle$ ) between a candidate and the top-ranked are shown with 1 standard error of the difference.

inherit feature of the timers but as a consequence of the censorship. Thus, the data can be interpreted as consistent with the independence assumption.

## 2.4 Using Filming Data to Obtain Best Distribution Class for the Timers

Identifying a parametric distribution class for the variables is another crucial component in the development of the model. In the literature, the probability distributions governing the times to first division and to death were reported to be well approximated by a right-skewed distribution such as Lognormal, Weibull, Gamma or Beta (Hawkins, Markham, et al. 2009). In particular, the time to first division is known to be better described by a Lognormal than other skew distributions whereas Gamma or Weibull distributions can be used to approximate the time to death distribution (Hawkins, Turner, Dowling, et al. 2007). However, the distribution of the time to division destiny is still unknown. Although the "true" distribution of the timers may not be possible to obtain, we can infer a better descriptor by employing Bayesian model selection criteria (see Methods).

In Fig. 4A, we plotted empirical cumulative distribution function (eCDF) of the data overlaid with CDFs

164 of Gamma, Lognormal, Normal and Weibull distributions. Each candidates are parametrised by:  $(\alpha_G, \beta_G)$   
 165 for Gamma;  $(m, s)$  median and shape for lognormal;  $(\mu, \sigma)$  mean and standard deviation for Normal; and,  
 166  $(\alpha_W, \beta_W)$  for Weibull. Qualitatively, all four candidates appear to be excellent descriptors for each of  
 167 measured times (see also Supplementary Fig. S4B1-3). However, we used Widely Applicable Information  
 168 Criterion (Watanabe 2010) to more quantitatively determine the best parametric distribution class (Fig. 4A2,  
 169 B1-3). For both CpG-stimulated B cells (repeat in Fig. S4A1-2) and CD8 T cells, lognormal distribution is  
 170 the best parametric distribution class, or in par with the others, for  $T_{div}^0$  and  $M$ , consistent with the previous  
 171 studies. The result of  $M$  for 1U IL-2 may be unreliable as there were only four observations. Strikingly, while  
 172  $T_{ld}$  and  $T_{die}$  can be well approximated by all four candidates for B cells, lognormal was the best descriptor  
 173 consistently for all T cells and Weibull is notably the worst candidate. In addition to this, we consistently  
 174 observed that the lognormal is a best descriptor for Costim1 and Costim2 datasets (Fig. S5). With strong  
 175 evidence preferred toward lognormal distribution and its strictly positive real number support as an extra  
 176 reason, we will use a lognormal distribution, otherwise mentioned, for the Cyton 2 model presented in this  
 177 paper.

### 3 Model

178 In this section, we describe the derivation of the Cyton 2 model to capture mean population dynamics of  
 179 lymphocytes. Let  $Z_g(t)$  denote the number of cells alive in generation  $g \in \{0, 1, \dots, G\}$  at time  $t \geq 0$ . Then,  
 180  $Z_g(t)$  can be written with the variables shown in Section 2.1 for any chosen probability density functions  
 181 for the random variables. Here, we derived  $\mathbb{E}[Z_g(t)]$  for  $g = 0$  and  $g > 0$  cases separately as lymphocytes  
 182 generally take longer to divide for the first time than at later generation. Also, we assumed that the  
 183 subsequent division time is a constant rather than a random variable as it was shown that the cells divide  
 184 remarkably at consistent rate in later generations (Gett and Hodgkin 1998; Gett and Hodgkin 2000). In  
 185 essence, we begin the derivation with parameters  $\theta = (m, T_{div}^0, T_{dd}, T_{die})$  denoting subsequent division time  
 186 (strictly positive real number) and three random variables for times to first division, to destiny and to death.

#### 3.1 Generation 0

187 For a given family tree, the number of live cells dividing, dying or reaching destiny in generation  $g = 0$  at  
 188 time  $t$  is given by

$$189 \quad Z_0(t) = \mathbb{1}_{\{T_{die} > t\}} \mathbb{1}_{\{\min(t, T_{dd}) < T_{div}^0\}},$$

190 where  $\mathbb{1}$  is indicator function. Assuming that the random variables  $T_{die}$ ,  $T_{dd}$  and  $T_{div}^0$  are independent of  
 191 each other as we established in [Section 2.2](#), the expected value of cell numbers is given by

192 
$$\mathbb{E}[Z_0(t)] = P(T_{die} > t)P(\{t < T_{div}^0 < T_{dd}\} \cup \{t < T_{dd} < T_{div}^0\} \cup \{T_{dd} < T_{div}^0 \leq t\} \cup \{T_{dd} \leq t < T_{div}^0\}). \quad (1)$$

193 The sets in the probability are mutually exclusive events as the intersections of sets in any pairwise combi-  
 194 nation are empty sets. Hence, the probability of union of sets simply becomes a sum. Note that the union  
 195 of the first two sets,  $\{t < T_{div}^0 < T_{dd}\} \cup \{t < T_{dd} < T_{div}^0\}$ , describes a cell that has not yet reached either its  
 196 first division or destiny, while the union of the last two sets,  $\{T_{dd} < T_{div}^0 \leq t\} \cup \{T_{dd} \leq t < T_{div}^0\}$ , represents  
 197 a cell that reached destiny. The probabilities of these events are

198 
$$\begin{aligned} P(\{t < T_{div}^0 < T_{dd}\} \cup \{t < T_{dd} < T_{div}^0\}) &= P(T_{div}^0 > t)P(T_{dd} > t), \\ P(\{T_{dd} < T_{div}^0 \leq t\} \cup \{T_{dd} \leq t < T_{div}^0\}) &= \int_0^t f_{T_{dd}}(\tau)P(T_{div}^0 > \tau)d\tau, \end{aligned} \quad (2)$$

199 where  $f_{T_{dd}}$  is a probability density function of  $T_{dd}$ . Substituting [Eq. \(2\)](#) into [Eq. \(1\)](#), the expected number  
 200 of live cell at  $g = 0$  becomes

201 
$$\mathbb{E}[Z_0(t)] = P(T_{die} > t) \left[ P(T_{div}^0 > t)P(T_{dd} > t) + \int_0^t f_{T_{dd}}(\tau)P(T_{div}^0 > \tau)d\tau \right]. \quad (3)$$

202 This equation can be interpreted as follows: a cell in generation 0 remains alive when  $T_{die} > t$ , and it is  
 203 sorted either in initial state or in destiny state. The cell in the initial state can divide, reach destiny or die  
 204 whichever event comes first. However, the destiny cell can no longer divide but only awaits for the death.

### 3.2 Generation $> 0$

205 To calculate the expected number of live cells for  $g > 0$ , we limit the window of cells being in generation  $g$   
 206 by constraining with  $t \in [T_{div}^0 + (g-1)m, T_{div}^0 + gm]$ , where  $m \in \mathbb{R}_{>0}$  is the subsequent division time.

207 
$$\begin{aligned} Z_g(t) &= 2^g \mathbb{1}_{\{T_{die} > t\}} \mathbb{1}_{\{T_{div}^0 + (g-1)m \leq \min(t, T_{dd}) < T_{div}^0 + gm\}} \\ &= 2^g \mathbb{1}_{\{T_{die} > t\}} \mathbb{1}_{\{T_{div}^0 + (g-1)m \leq \min(t, T_{dd})\} \cap \{\min(t, T_{dd}) < T_{div}^0 + gm\}}. \end{aligned}$$

208 The factor  $2^g$  is required to include effect of clonal expansion of the cells that have divided  $g$  times. Then,  
 209 the expected value is

212 
$$\mathbb{E}[Z_g(t)] = 2^g P(T_{die} > t) P(\{T_{div}^0 + (g-1)m \leq \min(t, T_{dd})\} \cap \{\min(t, T_{dd}) < T_{div}^0 + gm\}). \quad (4)$$

213 Expanding the sets, we readily get a union of four mutually exclusive events that can be categorised into  
 214 two different states similar to  $g = 0$  case: (i) dividing state,  $A = \{T_{div}^0 + (g - 1)m \leq t < T_{dd} < T_{div}^0 + gm\} \cup \{T_{div}^0 + (g - 1)m \leq t < T_{div}^0 + gm < T_{dd}\}$ ; and, (ii) destiny state,  $B = \{T_{div}^0 + (g - 1)m < T_{dd} \leq t < T_{div}^0 + gm\} \cup \{T_{div}^0 + (g - 1)m < T_{dd} < T_{div}^0 + gm \leq t\}$ . With the independence assumption, we obtain the  
 215  
 216 following probability.

$$218 \quad P(A) = P(T_{dd} > t)P(t - gm < T_{div}^0 < t - (g - 1)m) \quad (5)$$

$$P(B) = \int_0^t f_{T_{dd}}(\tau)P(\tau - gm < T_{div}^0 < \tau - (g - 1)m)d\tau$$

219 Thus, substituting Eq. (5) into Eq. (4), we obtain the expected value of live cells in  $g > 0$ ,

220  $\mathbb{E}[Z_g(t)] = 2^g P(T_{die} > t) \times$   
 221  $\left[ P(T_{dd} > t)P(t - gm < T_{div}^0 < t - (g - 1)m) + \int_0^t f_{T_{dd}}(\tau)P(\tau - gm < T_{div}^0 < \tau - (g - 1)m)d\tau \right] \quad (6)$   
 222

223 Together with Eq. (3) and Eq. (6), we can calculate the average number of live cells for a family in generation  
 224  $g$  at time  $t$ . Since the equation is equally applicable for  $N_0$  number of initial founder cells, we generalise the  
 225 equation by multiplying  $N_0$  such that

226  $y_g(t; \boldsymbol{\theta}) = \mathbb{E}[N_0 Z_g(t; \boldsymbol{\theta})] = N_0 \mathbb{E}[Z_g(t; \boldsymbol{\theta})]. \quad (7)$

227 where  $\boldsymbol{\theta} = (m, T_{div}^0, T_{dd}, T_{die})$  is the parameter of Cyton 2 model whose elements denote for subsequent  
 228 division time, time to first division, time to division destiny and time to death, respectively. Typically  
 229 random variables are equipped with lognormal distribution, which has two additional parameters, thus, we  
 230 have total of 7 free parameters to estimate.

## 4 Application to FACS data

### 4.1 B cells: Model assessment

231 In this section, we illustrate the model with a B cell dataset to assess its performance with standard statistical  
 232 tools in terms of the precision of the parameter estimation and the accuracy of the model fit. This data was  
 233 obtained from *in vitro* CpG-stimulated murine Bim<sup>-/-</sup> B cells and recorded cell numbers in each division  
 234 class via flow cytometry. In particular, the experiment was conducted with nine replicates that harvested at  
 235 nine different time points to cover the reasonable range of the cell response. For this dataset, we fitted the  
 236 model for two scenarios: (i) varying the number of replicates while utilising all time points; and (ii) remove  
 237 time points while holding the same number of replicates.

238 In Fig. 5A1-3, the best-fit model and the estimated parameters with 95% confidence intervals are shown

239 (see Methods). Qualitatively, we observed that the confidence bands of the model extrapolation get narrower  
240 around the mean as we increase the number of replicates, suggesting an improvement of the model estimates.  
241 We verified this by performing a principle component analysis (PCA) on a set of estimated parameter vectors  
242 (Fig. 5B1). The PCA result signifies that the first two principle components explain 78% variability in the set  
243 and, furthermore, assures that there is no notable correlation amongst components of the parameter vector  
244 which is consistent with independence of the timers. To assess the precision of the estimates quantitatively,  
245 we computed marginal coefficients of variation as a function of replicate number (Fig. 5B2). Interestingly,  
246 although using all available replicates is certainly a better option, we identified that there is no significant  
247 benefit for the precision of the estimates after three replicates. Therefore, it indicates that three replicates is  
248 sufficient and offers a good balance between obtaining a precise estimate and reducing experimental burdens.

249 In the next scenario, we turned our attention to evaluating the model accuracy as we sequentially removed  
250 time points while maintaining fixed number of replicates (Fig. 5C). Specifically, we considered excluding  ${}^9C_k$   
251 possible positions of time points, where  $k$  is the total number of removed time points. For example, if  $k = 1$ ,  
252 there are only nine possible ways to remove (e.g. 28h). Prior to the model fitting, we imposed the following  
253 rules to avoid any ambiguity and ensure the feasibility of the model fits.

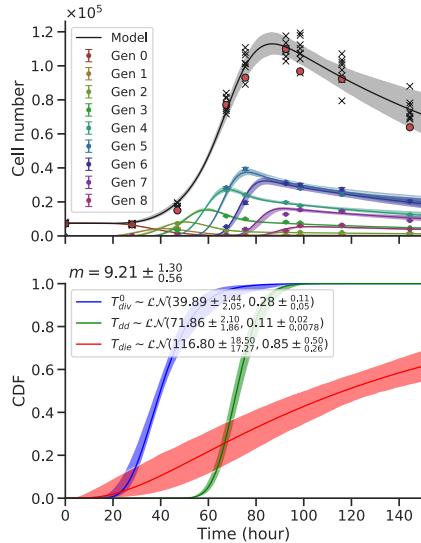
- 254 1. More than two time points must be kept.  
255 2. Either first or second time point must be remained in the set to provide an initial cell number for the  
256 model.  
257 3. Once the position of the time points are determined, all replicates of those time points are removed  
258 from the original dataset.

259 Given these rules, there are 366 cases to consider for all  $k = 1, \dots, 6$ . For each case, we constructed 100  
260 artificial datasets by randomly sampling (with replacement) from the replicates per remained time points and  
261 fitted the model, then calculated the root-mean-square error (RMSE) with the original *unremoved* dataset.

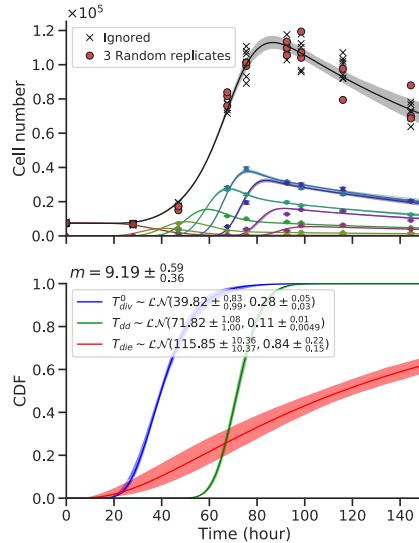
262 In Fig. 5C, we present a rank ordered distributions of RMSE for  $k = 1, 2, 3$  (see Fig. S6 for  $k > 3$ ). For  
263 comparison purposes, the RMSE of a model fit using all time points and replicates is shown as a reference.  
264 The results showed that the times at which cells expanding are the most important information to keep for  
265 the model accuracy. Intuitively, this would represent a regression of a non-linear curve in which data points  
266 around "inflection point" are missing while two ends points are present. Furthermore, we noticed that the  
267 first time point is generally more important than the later ones as RMSEs are higher if the first time point  
268 was removed. Interestingly, we found little to no difference of the RMSE compare to that of the reference  
269 when the positions of the removed time points are sparsely located. As an extreme example ( $k = 6$ ), the  
270 model was capable of accurately fit the data as long as there exists three time points that correspond to early,  
271 expansion and contraction phases (e.g. first, fourth and ninth time points, see Fig. S6B). However, knowing

272 those three time points prior to an experiment is practically impossible, thus, we would interpret this as a  
273 theoretical lower bound of "sampling rate". In summary, we conclude that obtaining the data points for the  
274 expansion phase is crucial, and an absolute minimum number of time points required for an accurate model  
275 fit is three, albeit specific to this particular B cell dataset.

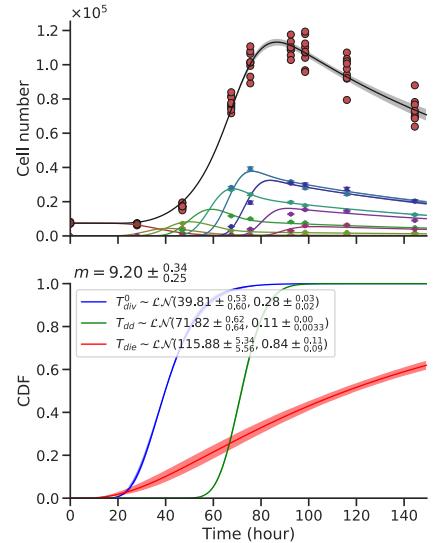
### A1 1 Replicate



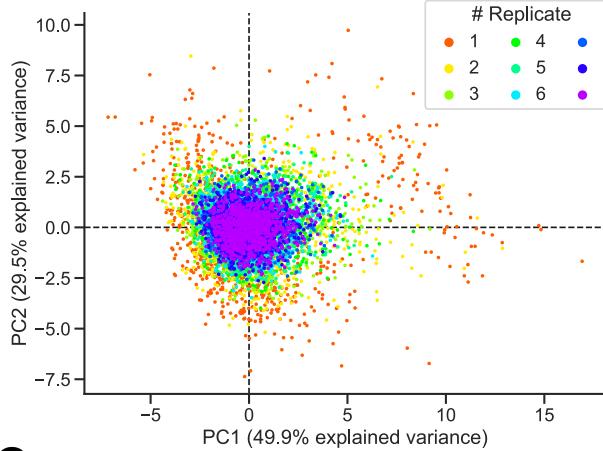
### A2 3 Replicates



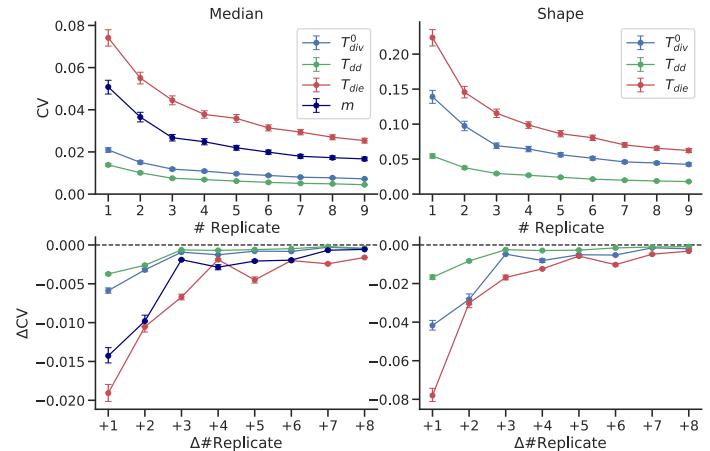
### A3 9 Replicates



### B1 PCA of Estimated Parameters



### B2 Coefficient of Variation of Estimated Parameters



### C Root-mean-square Error of the model fits

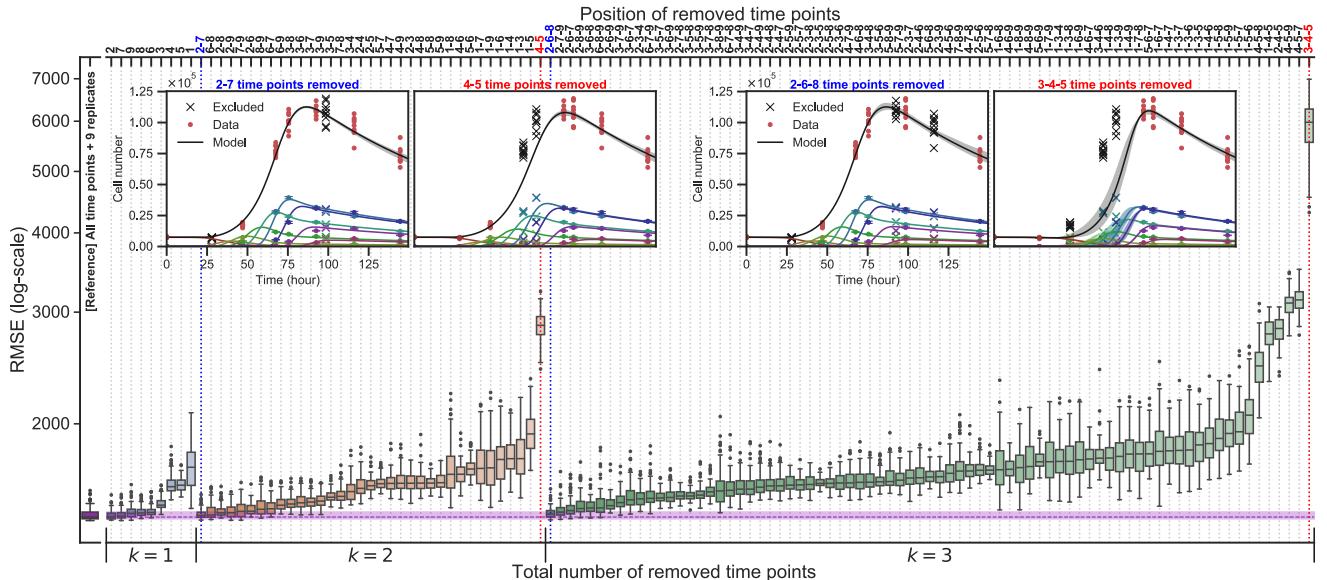
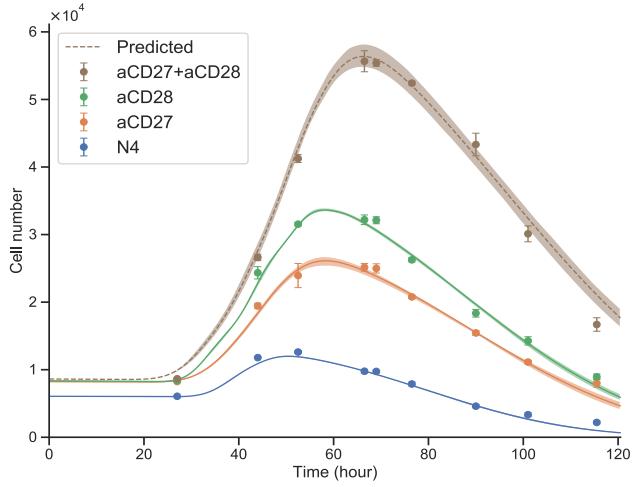
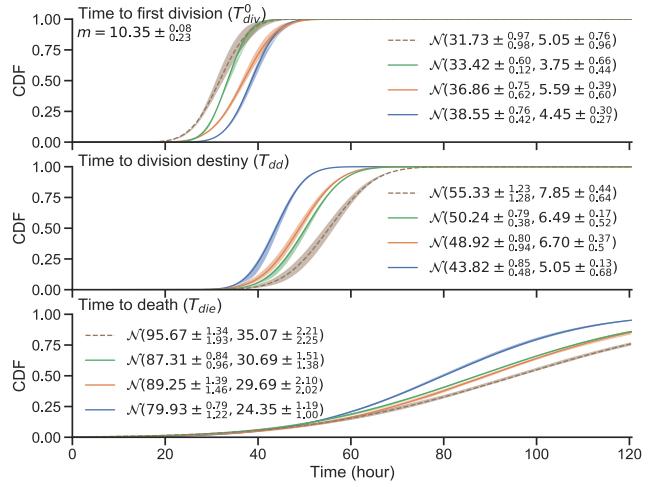


Figure 5: The precision of parameter estimation and accuracy of the model fit with CpG-stimulated  $\text{Bim}^{-/-}$  murine B cell data. (A1-3) The best-fit model extrapolation (top) and its corresponding estimated parameters (bottom). For a given replicate number,  $10^3$  artificial datasets were created by randomly sampling the original dataset with replacement per time point and fitted to calculate 95% confidence bands. (B1) From the sets of estimated parameter vectors, marginal coefficients of variation (CV) was calculated with 95% confidence interval from bootstrapping. (B2) Principle component analysis (PCA) of the estimated parameter vectors. (C) Root-mean-square error (RMSE) evaluated over all time points and replicates after fitting the model to artificially removed datasets. Examples of the best (blue) and worst (red) fits are shown.

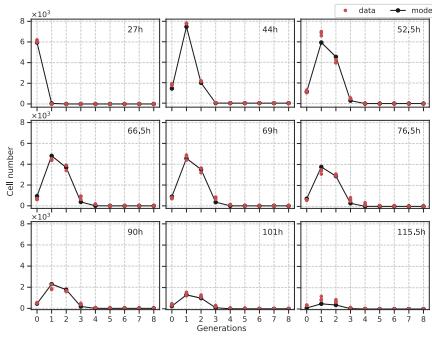
### A OT-I/Bcl2l11<sup>-/-</sup> CD8 T cell: Total cells



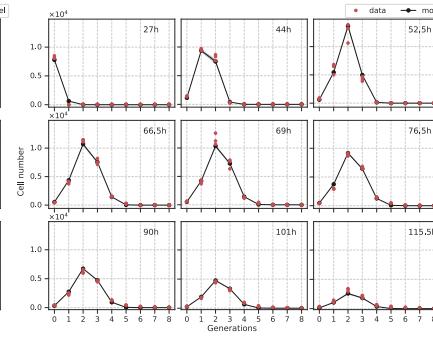
### C Cyton parameters



### B1 N4



### B2 aCD27



### B3 aCD28

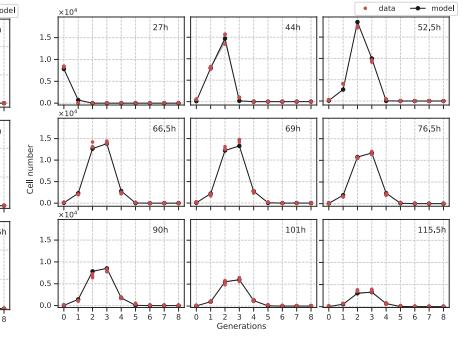


Figure 6: **Cyton 2 fitting to FACS T cell data (Marchingo, Kan, et al. 2014)**. The cells were stimulated with N4 (blue) as basis for all other conditions. **(A)** Harvested total cell numbers (●: mean  $\pm$  SEM) overlaid with the model extrapolation and 95% confidence band from bootstrapping. **(B1-3)** Live cells per generation and the model extrapolation at harvested time points. **(C)** Estimated parameters of  $m$ ,  $T_{div}^0$ ,  $T_{dd}$ ,  $T_{die}$  and their 95% confidence intervals. The fitted and predicted values of mean and standard deviation are labelled in the legend for normally distributed random variables.

## 4.2 T cells: Additive nature of signal integration in time domain

To further illustrate the model, we used it to capture the mean population dynamics of an experiment presented in Marchingo, Kan, et al. 2014. This data was obtained from *in vitro* CTV-labeled OT-I/Bcl2l11<sup>-/-</sup> CD8 T cells stimulated with N4 and cultured with antibodies of CD27 (5 $\mu$ g mL<sup>-1</sup>) and CD28 (2 $\mu$ g mL<sup>-1</sup>) alone and their combination. To accurately capture the effects of the stimuli, 25 $\mu$ g mL<sup>-1</sup> of S4B6, which blocks endogenous IL-2 produced interim, was added to all cultures. CD8 T cells were harvested at 27h, 44h, 52.5h, 66.5h, 69h, 76.5h, 90h, 101h and 115.5h after the stimulation with three replicates at each time point. Refer to the original article for more detail on the experiment setup.

Marchingo, Kan, et al. 2014 showed that each contribution of stimuli was manifested as an increase in mean division number (MDN) from the basis signal (N4). It was concluded that the effect of combination of two or more signals (e.g. aCD27+aCD28) is a simple linear sum of changes in MDN with respect to the basis signal from each component. Here, we asked if the same rule can be applied in the time domain that is, for example,  $T_{dd}^{aCD27+aCD28} = T_{dd}^{N4} + \Delta T_{dd}^{aCD27} + \Delta T_{dd}^{aCD28}$ , where  $\Delta$  denotes for the change in  $T_{dd}$  with respect to N4. To do so, we used *normal* distribution for all variables in the model as there is no closed-form

289 for sum of independent but not necessarily identical two or more lognormal distributions. Thus, assuming  
290 normally distributed random variables, we formulated a simple algebraic sum of the timers by,

$$291 \quad T_{aCD27+aCD28} \sim \mathcal{N} \left( \mu_{N4} + \Delta\mu_{aCD27} + \Delta\mu_{aCD28}, \sqrt{\sigma_{N4}^2 + \Delta\sigma_{aCD27}^2 + \Delta\sigma_{aCD28}^2} \right), \quad (8)$$

292 where  $\Delta\mu_x = \mu_x - \mu_{N4}$  and  $\Delta\sigma_x^2 = \sigma_x^2 - \sigma_{N4}^2$ .

293 In [Fig. 6A](#), we present total cell numbers and the best-fit model with 95% confidence band around the  
294 estimate from the original data (see Methods). The model was simultaneously fitted to N4, aCD27 and  
295 aCD28 datasets with a shared subsequent division time ([Fig. 6B1-3](#)). The estimated  $m$  and cumulative  
296 distribution functions (CDF) of  $T_{div}^0$ ,  $T_{dd}$  and  $T_{die}$  are shown in [Fig. 6C](#). In comparison to N4 alone, the  
297 addition of aCD27 and aCD28 extends both means of  $T_{dd}$  ( $\approx 15\%$ ) and  $T_{die}$  ( $\approx 10\%$ ). Also, we identified  
298 aCD28 reduces mean of  $T_{div}^0$  (13.3%) while aCD27 has minimal impact. Collectively, the compounding effect  
299 of these changes results in larger expansion of cell numbers by letting cells to enter the first division early and  
300 to reach destiny and death at later times. Given the estimates of parameters for N4, aCD27 and aCD28, we  
301 predicted the cell numbers for their combined effect by calculating the  $T_{div}^0$ ,  $T_{dd}$  and  $T_{die}$  according to [Eq. \(8\)](#).  
302 Strikingly, we successfully recreated the expansion kinetics of OT-I/Bcl2l11<sup>-/-</sup> CD8 T cells in the presence  
303 of both aCD27 and aCD28 ([Fig. 6A](#)), supporting the signal integration as a linear sum in a time domain and  
304 the independence of the timers. Additionally, we recapitulated the results in [Marchingo, Kan, et al. 2014](#) by  
305 simulating the family trees using ABM given the fitted and predicted parameter values ([Fig. S7](#)).

## 5 Discussion

306 We have formulated a variant of the original Cyton model that encapsulates feature of inheritance and  
307 correlation structure of cell fates. This was achieved by introducing new parameters called time to division  
308 destiny, which operates independent of the other cellular machinery, and global death time, which describes  
309 a single death time for all members in a family tree. Similar to the Cyton model, this variant offers a  
310 general tool for analysing lymphocyte proliferation and survival, particularly from the data obtained from  
311 CFSE/CTV-labeled division tracking assays. The formula for its computation turned out to be in a relatively  
312 simple form, despite the violation of conventional assumption (e.g. in a standard branching process) that  
313 the progeny are stochastically independent entities. As a result, we believe that this simplicity improves the  
314 model identifiability while computational complexity is not increased.

315 The analysis of the B and T cell filming data provided additional layer of consistency with the inde-  
316 pendence assumption of the cellular machineries, reassuring the very foundation of the Cyton model. We  
317 illustrated this by simulating large number of family trees under the independent operation of Cyton 2 ma-  
318 chineries. It was apparent that the censorship of the events rendered challenges in proving the assumption,

319 however, we recapitulated the data to show its consistency, albeit not a definitive proof. Additionally, we  
320 identified the best parametric distribution class for essential variables of the Cyton 2 model. Although the  
321 true governing distribution may not be possible to obtain, we inferred that lognormal distribution is the best  
322 descriptor of the timers, even for the time to division destiny. However, the normal distribution may be used  
323 to approximate the timers as well.

324 To capture the mean population dynamics, we have derived a formula that parameterised by a constant  
325 and three random variables, where three random variables represent times to first division, to division destiny  
326 and to death, and a constant is for subsequent division time equally applied to all generations after the first.  
327 We provided examples of the model utilisation for both B- and T-cells and illustrated that it equally well  
328 captured the dynamics, which has potential to offer as a general tool for other types of lymphocyte studies.  
329 Our main conclusions through studies of B- and T-cells are that firstly, the design of an experiment has an  
330 impact on the precision of parameter estimates and accuracy of the model. We recommend at least three  
331 replicates per experiment, which then can be further interrogated for confidence intervals via bootstrapping,  
332 and carefully select the harvested time points to ensure that the data contains at least one time point for  
333 each key events (e.g. expansion and/or contraction) in an experiment. Secondly, we further extended the  
334 work of [Marchingo, Kan, et al. 2014](#) to show that the additive nature of T cell stimuli can also be expressed  
335 as a linear sum of normally distributed random variables that resides in time domain. This was verified by  
336 using Cyton 2 model and ABM simulation, and we reached the same conclusion. One of limitations we wish  
337 to raise is that the support of a normal distribution can extend to negative infinity. This poses a challenge  
338 that it is mathematically possible to obtain negative mean values, and our study may had been a special  
339 case for more general rule. For this reason and together with evidence from the filming data, a lognormal  
340 distribution remains as an optimal choice for further investigation. To date, there is no closed-form expression  
341 to calculate sums of independent but not identical lognormal distributions. Moreover, there is no guarantee  
342 that the sum is a lognormal distribution. However, numerical approximations ([Mehta et al. 2007; Lo 2012](#))  
343 could be used as an initial point of attack.

344 One of generalisations of the model can be achieved by replacing the constant subsequent division time to  
345 a random variable. It is evident from the filming data that there exists some degrees of variation in division  
346 times for subsequent generations. However, our current proposed model fails to capture this behaviour.  
347 In this sense, more accurate description of lymphocyte proliferation remains as a challenge that may be  
348 investigated in future work.

## 6 Methods

### 6.1 Time-lapse Microscope

349 [Detailed experiment protocols for the filming data. I need help for this section... For B- and T-cell data, I  
 350 have briefly explained the experiment setup in the main text and referred her original paper for more detail.]

### 6.2 Data selection and tree collapse

Cell Type	Stim.	Number of families for each variables			
		Time to fist division ( $T_{div}^0$ )	Subsequent division time ( $M$ )	Time to last division ( $T_{ld}$ )	Time to death ( $T_{die}$ )
B	CpG	69 (63.9%)	56 (51.9%)	69 (63.9%)	69 (63.9%)
CD8 T (+IL2)	1U	28 (25.7%)	4 (3.7%)	28 (25.7%)	28 (25.7%)
	3U	34 (37.8%)	13 (14.4%)	34 (37.8%)	34 (37.8%)
	10U	50 (30.7%)	16 (9.8%)	50 (30.7%)	60 (30.7%)

Table 3: Number of clones retained after the filtering.

351 For each family tree  $c \in \mathbb{N}_{\geq 0}$ , the times to divide  $\{T_{div}^x\}_c$ , to die  $\{T_{die}^x\}_c$  and to loss  $\{T_{loss}^x\}_c$  of all  
 352 cells were recorded using time-lapse microscope.  $T_{loss}$  is defined as the time at which the cell becomes  
 353 indistinguishable to the nearby cells, or survives until the end of given experiment time frame, thus, were  
 354 lost from the experiment. In order to keep track of the cells' relation, a unique label was given to each cell  
 355 by  $x$ . For example, we denote a founder cell  $x = \langle 0 \rangle$ , and its two daughter cells  $x = \langle 1 \rangle$  and  $x = \langle 2 \rangle$  so that,  
 356 in general,  $x = \langle x_1, x_2, \dots, x_j \rangle$  with  $x_j \in \{1, 2\}$  would represent  $x_j^{\text{th}}$  daughter cell of  $\dots$  of  $x_2^{\text{th}}$  daughter  
 357 cell (see, [Harris 1963](#), Ch.6). Let  $\mathcal{X}_c$  be the collection of all  $x$  for a family  $c$ . Given a unique identifier of  
 358 the cell, its generation  $k$  is extracted by  $g(x) := k$  with  $g(\langle 0 \rangle) = 0$ . With this construct, we define the raw  
 359 measurement of times as a set  $\mathcal{T}_c = \{t(x) \in \{T_{div}^x, T_{die}^x, T_{loss}^x\} : x \in \mathcal{X}_c\}$ .

360 For analyses in [Section 2.2](#) and [Section 2.4](#), we filtered for families that had at least divided once and  
 361 satisfy  $\max(\mathcal{T}_c) = T_{die}^x$  condition. In essence, we eliminated incomplete family trees that contain unusually  
 362 long-surviving cells, but allowed unrecovered cells to be in place as long as the last observed event is death in  
 363 a given family. Indeed there is increasing chance of observing more lost cells as family grows larger. However,  
 364 it was previously shown that the regularity of a family is a result of correlated cell divisions as a biological  
 365 feature inherited within the family even when considering the unrecovered samples ([Marchingo, Prevedello,  
 366 et al. 2016](#)). Therefore, it is highly likely that the lost cells due to indistinguishable circumstance might had  
 367 undergone similar fates with its sibling, thereby maximising the number of data points while reducing any  
 368 potential selection bias, whereas it is difficult to weigh how including the long-surviving cells might affect all  
 369 the other analyses.

370 Given the heritability feature, we summarise a family tree by collapsing it to a single representative  
 371 line ([Fig. 1C](#)). By collapsing we mean substitute average time to divide (and to die) of the cells in a given

372 generation  $k$ , that is

$$373 \quad T_{div}^k := \frac{\sum_{\{x: g(x)=k\}} T_{div}^x}{|\{x : g(x) = k\}|}, \quad T_{die}^k := \frac{\sum_{\{x: g(x)=k\}} T_{die}^x}{|\{x : g(x) = k\}|}.$$

374 We also enumerated all dead cells within a family and calculated mean time to last division ( $T_{ld}$ ) as a proxy to  
 375 the division destiny time. In summary, we represent a single family by  $\mathbf{T}(c) = (T_{div}^0, \dots, T_{div}^k, T_{die}^0, \dots, T_{die}^k)$   
 376 so long as we observed division or death events in each generation  $k$ . [Table 3](#) shows the number of retained  
 377 families used in all analyses presented in this paper given  $\mathbf{T}(c)$  after applying the filtering rule.

### 6.3 Agent-Based Model

378 We developed an agent-based model (ABM) to simulate cells in a single family with the correlated structured  
 379 proposed for the Cyton 2 model. Each realisation of the simulation represents one clonal family. Upon  
 380 initialisation, the founder cell is assigned time to first division, global destiny and global death times, which  
 381 were drawn randomly from three independent lognormal distribution. The time taken to the subsequent  
 382 division after the first is also randomly drawn from a lognormal distribution, but it remains constant and  
 383 equally applied for all subsequent generations. If the founder cell reaches time to first division, it creates  
 384 two daughter cells, which inherit global destiny and death times. If the cell reaches destiny we immediately  
 385 classify it as a destiny cell and prevent from further division. When the cells reach death time they are  
 386 removed from the simulation. The model was implemented in Python (version 3.8.6).

### 6.4 Statistical Analysis: Bayesian Framework

387 The correlations of every pair of time to first division ( $T_{div}^0$ ), average subsequent division time ( $M$ ), time  
 388 to last division ( $T_{ld}$ ) and time to death ( $T_{die}$ ) were estimated using Bayesian inference. For a given pair of  
 389 variables and its observed data, say  $d_i \in \mathcal{D} = \{(x_i, y_i) : i = 1, 2, \dots, n\}$  where  $n$  is the number of observations,  
 390 we used bivariate normal distribution to capture the correlation ( $\rho$ ). This entails  $x_i \sim N(\mu_x, \sigma_x)$  and  
 391  $y_i \sim N(\mu_y, \sigma_y)$ . With uninformative priors on the hyper parameters  $\mu_x, \mu_y \sim U(0, 1000)$ ,  $\sigma_x, \sigma_y \sim U(0, 1000)$   
 392 and  $\rho \sim U(-1, 1)$ , we define the bivariate normal distribution,

$$393 \quad d_i \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}),$$

394 where  $\boldsymbol{\mu} = (\mu_x, \mu_y)$  is a vector of means for  $x_i$  and  $y_i$ , and  $\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{bmatrix}$  is a covariance ma-  
 395 trix. We used an extension of Hamiltonian MCMC algorithm, No-U-Turn Sampler ([Hoffman and Gelman](#)  
 396 [2011](#)), implemented in PyMC3 ([Salvatier, Wiecki, and Fonnesbeck 2016](#)) to infer the posterior distributions  
 397 of  $\rho, \mu_x, \sigma_x, \mu_y, \sigma_y$ . Given these distributions, we calculated 95% credible interval for  $\rho$ , and 90%, 95% and  
 398 99% density regions of  $(x, y)$ . In addition to this, we can formulate bayesian hypothesis testing, where the

399 null hypothesis is  $H_0: \rho = 0$  and alternative hypothesis is  $H_1: \rho \neq 0$  (which translates to  $H_1: \rho \sim U(-1, 1)$ )  
 400 (Jeffreys 1961). This is formally stated as a ratio of likelihoods of hypotheses given the data,

401

$$\frac{P(H_0|\mathcal{D})}{P(H_1|\mathcal{D})} = \frac{P(H_0)}{P(H_1)} \times \frac{P(\mathcal{D}|H_0)}{P(\mathcal{D}|H_1)}.$$

402 In order to grade if the data is more probable under  $H_0$  or  $H_1$ , the Bayes factor  $BF_{01} = P(\mathcal{D}|H_0)/P(\mathcal{D}|H_1)$   
 403 was used given priors of  $P(H_0)$  and  $P(H_1)$ . When  $H_1: \rho \sim U(-1, 1)$ , it can be computed by evaluating the  
 404 following integral (Jeffreys 1961; Wagenmakers, Verhagen, and Ly 2016):

405

$$BF_{01} = 1/BF_{10}, \text{ where } BF_{10} = \frac{1}{2} \int_{-1}^1 \frac{(1 - \rho^2)^{\frac{n-1}{2}}}{(1 - \rho r)^{n-\frac{3}{2}}} d\rho,$$

406 where  $r$  denotes for the sample correlation defined as  $r = \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) / \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}$ .  
 407 Table 4 shows the discrete categories of evidential strength proposed in Jeffreys 1961 for the interpretation  
 of the Bayes factor. The analysis was applied in Section 2.2 and Section 2.3.

Bayes Factor: $BF_{01}$ ( $BF_{10}$ )	Interpretation
$>100$	Extreme evidence for $H_0$ ( $H_1$ )
30 - 100	Very strong evidence for $H_0$ ( $H_1$ )
10 - 30	Strong evidence for $H_0$ ( $H_1$ )
1 - 3	Anecdotal evidence for $H_0$ ( $H_1$ )
1	No evidence

Table 4: Bayes factor interpretation.

408  
 409 In Section 2.4, the candidates of parametric distribution class were assessed for  $T_{div}^0, M, T_{ld}, T_{die}$  variables  
 410 under the Bayesian framework in a similar manner to estimating the correlation coefficient. Table 5 shows  
 the list of candidate distributions and the uninformative priors prescribed for respective parameters. Given

Candidates	Priors	Target Distribution
A	$\alpha_G, \beta_G \sim U(0, 200)$	$T_{div}^0, M, T_{ld}, T_{die} \sim \text{Gamma}(\alpha_G, \beta_G)$
B	$m, s \sim U(0, 200)$	$T_{div}^0, M, T_{ld}, T_{die} \sim \text{LN}(m, s)$
C	$\mu, \sigma \sim U(0, 200)$	$T_{div}^0, M, T_{ld}, T_{die} \sim \text{N}(\mu, \sigma)$
D	$\alpha_W, \beta_W \sim \text{HalfNormal}(0, 200)$	$T_{div}^0, M, T_{ld}, T_{die} \sim \text{Weibull}(\alpha_W, \beta_W)$

Table 5: List of candidate parametric distribution classes.

411  
 412 posterior distributions of the parameters, we adopted WAIC (Watanabe 2010) score to quantitatively assess  
 413 the candidates, which is estimated as follows:

414

$$\text{WAIC}(z, \Theta) = -2 \left( \sum_{i=1}^n \log \left[ \frac{1}{S} \sum_{s=1}^S P(z_i|\Theta_s) \right] - \sum_{i=1}^n \text{Var}_{s=1}^S \log(P(z_i|\Theta_s)) \right),$$

415 where  $z$  is the data with  $n$  independent number of observations,  $\Theta$  is the posterior distribution,  $\Theta_s$  is the  
 416  $s$ -th set of sampled parameter values in the posterior distribution with  $S$  number of samples and  $\text{Var}_{s=1}^S a_s =$

417  $\frac{1}{S-1} \sum_{s=1}^S (a_s - \bar{a})^2$  denotes for the sample variance (Vehtari, Gelman, and Gabry 2017; McElreath 2020).  
 418 The first and the second terms in  $(\cdot)$  are known as the *log-pointwise-predictive-density* (lppd) and the penalty  
 419 term, respectively. For direct comparison of the candidates, we computed the standard error by calculating  
 420 the variance over the individual observations instead of their summation under the assumption of normality  
 421 of WAIC.

422 
$$\text{se}(\text{WAIC}) = \sqrt{n \times \text{Var}_{i=1}^n \left( -2 \log \left[ \frac{1}{S} \sum_{s=1}^S P(z_i | \Theta_s) \right] + 2 \text{Var}_{s=1}^S \log(P(z_i | \Theta_s)) \right)}. \quad (9)$$

423 Let us denote  $\text{WAIC}_i$  to be the term in  $\sqrt{(\cdot)}$  such that  $\text{WAIC} = \sum_{i=1}^n \text{WAIC}_i$ , then the standard error of  
 424 the difference of WAIC between, for instance, candidate A and B can be calculated,

425 
$$\text{se}(\text{WAIC}^A - \text{WAIC}^B) = \sqrt{n \times \text{Var}_{i=1}^n (\text{WAIC}_i^A - \text{WAIC}_i^B)}.$$

## 6.5 Fitting Cyton Model

426 Division structured population datasets obtained from FACS were fitted to the Cyton 2 model (Eq. (7)).  
 427 There are total of 7 parameters to be estimated for each dataset, thus if we have  $N$  number of conditions,  
 428 we have a maximum of  $7N$  free parameters to be fitted. For all conditions, we always used cell numbers at  
 429 the beginning of the stimulus (typically at  $t = 0$ ) as a fixed initial cell number.

430 For each set of cell numbers  $\{n_{g,r}(t_i)\}$  from the data, where  $i \in \{0, 1, \dots, I\}$ ,  $g \in \{0, 1, \dots, G\}$  and  
 431  $r \in \{0, 1, \dots, R\}$  are time, generation and replicate indices, respectively, we obtained point estimates of the  
 432 parameters. To achieve this, we used least-squares method with Levenberg-Marquardt (Marquardt 1963)  
 433 optimisation algorithm implemented in Python library LMFIT (version 1.0.1) (Newville et al. 2014). We  
 434 defined our cost function,

435 
$$C(\boldsymbol{\theta}) = \sum_{i=0}^I \sum_{g=0}^G \sum_{r=0}^R (n_{g,r}(t_i) - y_g(t_i; \boldsymbol{\theta}))^2,$$

436 such that we found an approximate minimum,

437 
$$\{\boldsymbol{\theta}^*\} \in \arg \min_{\boldsymbol{\theta}} C(\boldsymbol{\theta}).$$

438 As each algorithm requires a starting parameter value, we prescribed 100 sets of initial values drawn uniformly  
 439 at random from the appropriate parameter ranges, and recorded residual sum-of-squares for each set to pick  
 440 the best fitted parameters at the end. After identifying the best fit, we performed bootstrap method (Efron  
 441 1979) with an artificial dataset that was resampled with replacement (per time point) from the original  
 442 measured data. We repeated this process 100 times, resulted in 100 additional estimates for each parameter.  
 443 This allowed us to calculate 95% confidence intervals on the best fitted parameter values. Additionally, we

<sup>444</sup> also obtained confidence bands for extrapolated cell numbers by calculating 95% percentile range at each of  
<sup>445</sup> discretised time point from the model.

## AUTHOR CONTRIBUTIONS

## FUNDING

<sup>446</sup> This project has received funding from the European Union’s Horizon 2020 research and innovation pro-  
<sup>447</sup> gramme under the Marie Skłodowska-Curie grant agreement No 764698.

## ACKNOWLEDGEMENTS

## References

1. McElreath, Richard (2020). *Statistical Rethinking*. 2nd ed. CRC Press/Taylor & Francis Group.
2. Horton, Miles B, Giulio Prevedello, Julia M Marchingo, Jie H S Zhou, Ken R Duffy, Susanne Heinzel, and Philip D Hodgkin (2018). “Multiplexed Division Tracking Dyes for Proliferation-Based Clonal Lineage Tracing.” *Journal of immunology (Baltimore, Md. : 1950)* 201, pp. 1097–1103.
3. Kannan, Rajesh E. and Supreet Saini (2018). “Mathematical Modelling of Quorum Sensing in Bacteria”. *INAE Letters* 3, pp. 175–187.
4. Mitchell, Simon, Koushik Roy, Thomas A Zangle, and Alexander Hoffmann (2018). “Nongenetic origins of cell-to-cell variability in B lymphocyte proliferation.” *Proceedings of the National Academy of Sciences of the United States of America* 115, E2888–E2897.
5. Mazzocco, Pauline, Samuel Bernard, and Laurent Pujo-Menjouet (2017). “Estimates and impact of lymphocyte division parameters from CFSE data using mathematical modelling”. *PLOS ONE* 12, e0179768.
6. Vehtari, Aki, Andrew Gelman, and Jonah Gabry (2017). “Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC”. *Statistics and Computing* 27, pp. 1413–1432.
7. Yates, Christian A., Matthew J. Ford, and Richard L. Mort (2017). “A Multi-stage Representation of Cell Proliferation as a Markov Process”. *Bulletin of Mathematical Biology* 79, pp. 2905–2928.
8. Heinzel, Susanne, Tran Binh Giang, Andrey Kan, Julia M Marchingo, Bryan K Lye, Lynn M Corcoran, and Philip D Hodgkin (2016). “A Myc-dependent division timer complements a cell-death timer to regulate T cell and B cell responses.” *Nature immunology* 18, pp. 96–103.
9. Marchingo, J M, G Prevedello, A Kan, S Heinzel, P D Hodgkin, and K R Duffy (2016). “T-cell stimuli independently sum to regulate an inherited clonal division fate.” *Nature communications* 7, p. 13540.
10. Salvatier, John, Thomas V. Wiecki, and Christopher Fonnesbeck (2016). “Probabilistic programming in Python using PyMC3”. *PeerJ Computer Science* 2, e55.
11. Wagenmakers, Eric-Jan, Josine Verhagen, and Alexander Ly (2016). “How to quantify the evidence for the absence of a correlation”. *Behavior Research Methods* 48, pp. 413–426.
12. Shokhirev, Maxim N, Jonathan Almaden, Jeremy Davis-Turak, Harry A Birnbaum, Theresa M Russell, Jesse A D Vargas, and Alexander Hoffmann (2015). “A multi-scale approach reveals that NF- $\kappa$ B cRel enforces a B-cell decision to divide.” *Molecular systems biology* 11, p. 783.

13. Dowling, Mark R, Andrey Kan, Susanne Heinzel, Jie H S Zhou, Julia M Marchingo, Cameron J Wellard, John F Markham, and Philip D Hodgkin (2014). “Stretched cell cycle model for proliferating lymphocytes.” *Proceedings of the National Academy of Sciences of the United States of America* 111, pp. 6377–82.
14. Hodgkin, Philip D, Mark R Dowling, and Ken R Duffy (2014). “Why the immune system takes its chances with randomness.” *Nature reviews. Immunology* 14, p. 711.
15. Marchingo, Julia M, Andrey Kan, et al. (2014). “Antigen affinity, costimulation, and cytokine inputs sum linearly to amplify T cell expansion”. *Science* 346, pp. 1123–1127.
16. Newville, Matthew, Till Stensitzki, Daniel B. Allen, and Antonino Ingargiola (2014). *LMFIT: Non-Linear Least-Square Minimization and Curve-Fitting for Python*. doi: 10.5281/zenodo.11813.
17. Reiner, Steven L and William C Adams (2014). “Lymphocyte fate specification as a deterministic but highly plastic process”. *Nature Reviews Immunology* 14, pp. 699–704.
18. Buchholz, Veit R et al. (2013). “Disparate individual fates compose robust CD8+ T cell immunity.” *Science (New York, N.Y.)* 340, pp. 630–5.
19. Hawkins, E D, M L Turner, C J Wellard, J H S Zhou, M R Dowling, and P D Hodgkin (2013). “Quantal and graded stimulation of B lymphocytes as alternative strategies for regulating adaptive immune responses.” *Nature communications* 4, p. 2406.
20. Shokhirev, Maxim Nikolaievich and Alexander Hoffmann (2013). “FlowMax: A Computational Tool for Maximum Likelihood Deconvolution of CFSE Time Courses”. *PLoS ONE* 8, e67620.
21. Weber, Marc and Javier Buceta (2013). “Dynamics of the quorum sensing switch: stochastic and non-stationary effects”. *BMC Systems Biology* 7, p. 6.
22. Banks, H T, W Clayton Thompson, Cristina Peligero, Sandra Giest, Jordi Argilaguet, and Andreas Meyerhans (2012). “A division-dependent compartmental model for computing cell numbers in CFSE-based lymphocyte proliferation assays”. *Mathematical Biosciences and Engineering* 9, pp. 699–736.
23. Duffy, K R, C J Wellard, J F Markham, J H S Zhou, R Holmberg, E D Hawkins, J Hasbold, M R Dowling, and P D Hodgkin (2012). “Activation-Induced B Cell Fates Are Selected by Intracellular Stochastic Competition”. *Science* 335, pp. 338–341.
24. Hasenauer, J., D. Schittler, and F. Allgöwer (2012). “Analysis and Simulation of Division- and Label-Structured Population Models”. *Bulletin of Mathematical Biology* 74, pp. 2692–2732.
25. Lo, C. F. (2012). “The Sum and Difference of Two Lognormal Random Variables”. *Journal of Applied Mathematics* 2012, pp. 1–13.
26. Banks, H. T., Karyn L. Sutton, W. Clayton Thompson, Gennady Bocharov, Dirk Roose, Tim Schenkel, and Andreas Meyerhans (2011). “Estimation of Cell Proliferation Dynamics Using CFSE Data”. *Bulletin of Mathematical Biology* 73, pp. 116–150.
27. Hoffman, Matthew D and Andrew Gelman (2011). “The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo”. *arXiv*.
28. Miao, Hongyu, Xia Jin, Alan S Perelson, and Hulin Wu (2011). “Evaluation of multitype mathematical models for CFSE-labeling experiment data.” *Bulletin of mathematical biology* 74, pp. 300–26.
29. Nordon, Robert E., Kap-Hyoun Ko, Ross Odell, and Timm Schroeder (2011). “Multi-type branching models to describe cell differentiation programs”. *Journal of Theoretical Biology* 277, pp. 7–18.
30. Hyrien, Ollivier, Rui Chen, and Martin S Zand (2010). “An age-dependent branching process model for the analysis of CFSE-labeling experiments.” *Biology direct* 5, p. 41.

31. Markham, John F, Cameron J Wellard, Edwin D Hawkins, Ken R Duffy, and Philip D Hodgkin (2010). “A minimum of two distinct heritable factors are required to explain correlation structures in proliferating lymphocytes.” *Journal of the Royal Society, Interface / the Royal Society* 7, pp. 1049–59.
32. Watanabe, Sumio (2010). “Asymptotic Equivalence of Bayes Cross Validation and Widely Applicable Information Criterion in Singular Learning Theory”. *arXiv*.
33. Wellard, C, J Markham, E D Hawkins, and P D Hodgkin (2010). “The effect of correlations on the population dynamics of lymphocytes”. *Journal of Theoretical Biology* 264, pp. 443–449.
34. Zilman, Anton, Vitaly V. Ganusov, and Alan S. Perelson (2010). “Stochastic Models of Lymphocyte Proliferation and Death”. *PLoS ONE* 5, e12775.
35. Hawkins, E D, J F Markham, L P McGuinness, and P D Hodgkin (2009). “A single-cell pedigree analysis of alternative stochastic lymphocyte fates”. *Proceedings of the National Academy of Sciences* 106, pp. 13457–13462.
36. Lee, Ha Youn, Edwin Hawkins, Martin S. Zand, Tim Mosmann, Hulin Wu, Philip D. Hodgkin, and Alan S. Perelson (2009). “Interpreting CFSE Obtained Division Histories of B Cells in Vitro with Smith–Martin and Cyton Type Models”. *Bulletin of Mathematical Biology* 71, pp. 1649–1670.
37. Duffy, Ken R and Vijay G Subramanian (2008). “On the impact of correlation between collaterally consanguineous cells on lymphocyte population dynamics”. *Journal of Mathematical Biology* 59, pp. 255–285.
38. Hyrien, Ollivier and Martin S Zand (2008). “A Mixture Model With Dependent Observations for the Analysis of CSFE–Labeling Experiments”. *Journal of the American Statistical Association* 103, pp. 222–239.
39. Hawkins, E D, M L Turner, M R Dowling, C van Gend, and P D Hodgkin (2007). “A model of immune regulation as a consequence of randomized lymphocyte division and death times”. *Proceedings of the National Academy of Sciences* 104, pp. 5032–5037.
40. Hawkins, Edwin D, Mirja Hommel, Marian L Turner, Francis L Battye, John F Markham, and Philip D Hodgkin (2007). “Measuring lymphocyte proliferation, survival and differentiation using CFSE time-series data”. *Nature Protocols* 2, pp. 2057–2067.
41. Luzyanina, Tatyana, Dirk Roose, Tim Schenkel, Martina Sester, Stephan Ehl, Andreas Meyerhans, and Gennady Bocharov (2007). “Numerical modelling of label-structured cell population growth using CFSE distribution data”. *Theoretical Biology and Medical Modelling* 4, p. 26.
42. Mehta, N.B., Jingxian Wu, A.F. Molisch, and Jin Zhang (2007). “Approximating a Sum of Random Variables with a Lognormal”. *IEEE Transactions on Wireless Communications* 6, pp. 2690–2699.
43. Yates, Andrew, Cliburn Chan, Jessica Strid, Simon Moon, Robin Callard, Andrew JT George, and Jaroslav Stark (2007). “Reconstruction of cell population dynamics using CFSE”. *BMC Bioinformatics* 8, p. 196.
44. Asquith, Becca, Christophe Debacq, Arnaud Florins, Nicolas Gillet, Teresa Sanchez-Alcaraz, Angelina Mosley, and Luc Willems (2006). “Quantifying lymphocyte kinetics in vivo using carboxyfluorescein diacetate succinimidyl ester”. *Proceedings of the Royal Society B: Biological Sciences* 273, pp. 1165–1171.
45. Boer, Rob J. De and Alan S. Perelson (2005). “Estimating division and death rates from CFSE data”. *Journal of Computational and Applied Mathematics* 184, pp. 140–164.
46. Ganusov, Vitaly V., Sergei S. Pilyugin, Rob J. de Boer, Kaja Murali-Krishna, Rafi Ahmed, and Rustom Antia (2005). “Quantifying cell turnover using CFSE data”. *Journal of Immunological Methods* 298, pp. 183–200.

47. Hodgkin, P. D. (2005). “Quantitative rules for lymphocyte regulation: the cellular calculus and decisions between tolerance and activation”. *Tissue Antigens* 66, pp. 259–266.
48. Revy, Patrick, Mireia Sospedra, Boris Barbour, and Alain Trautmann (2001). “Functional antigen-independent synapses formed between T cells and dendritic cells”. *Nature Immunology* 2, pp. 925–931.
49. Gett, Amanda V and Philip D Hodgkin (2000). “A cellular calculus for signal integration by T cells”. *Nature Immunology* 1, pp. 239–244.
50. Nordon, Robert E, Masanori Nakamura, Carole Ramirez, and Ross Odell (1999). “Analysis of growth kinetics by division tracking”. *Immunology and Cell Biology* 77, pp. 523–529.
51. Gett, Amanda V and Philip D Hodgkin (1998). “Cell division regulates the T cell cytokine repertoire, revealing a mechanism underlying immune class regulation”. *Proceedings of the National Academy of Sciences* 95, pp. 9488–9493.
52. Lyons, A.Bruce and Christopher R. Parish (1994). “Determination of lymphocyte division by flow cytometry”. *Journal of Immunological Methods* 171, pp. 131–137.
53. Weston, Susan A. and Christopher R. Parish (1990). “New fluorescent dyes for lymphocyte migration studies Analysis by flow cytometry and fluorescence microscopy”. *Journal of Immunological Methods* 133, pp. 87–97.
54. Efron, B. (1979). “Bootstrap Methods: Another Look at the Jackknife”. *The Annals of Statistics* 7, pp. 1–26.
55. Smith, J A and L Martin (1973). “Do Cells Cycle?” *Proceedings of the National Academy of Sciences* 70, pp. 1263–1267.
56. Harris, Theodore E (1963). “The Theory of Branching Processes”.
57. Marquardt, Donald W (1963). “An Algorithm for Least-Squares Estimation of Nonlinear Parameters”. *Journal of the Society for Industrial and Applied Mathematics* 11, pp. 431–441.
58. Jeffreys, Harold (1961). *Theory of Probability*. 3rd ed. Oxford: Oxford University Press.

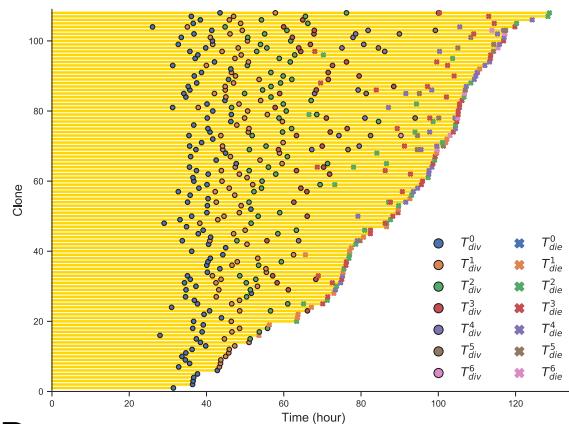


# Supplementary Materials

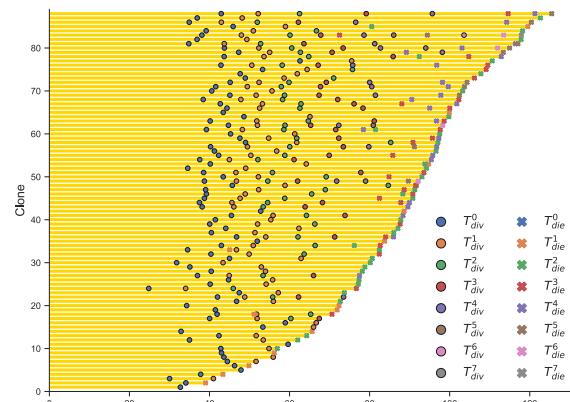
## A Table of data

Data	# clones	Observed cells: # divided cells (# dead cells) [# lost cells]								
		Gen 0	Gen 1	Gen 2	Gen 3	Gen 4	Gen 5	Gen 6	Gen 7	
CpG3	108	108 (0) [0]	169 (30) [17]	204 (67) [67]	153 (147) [108]	82 (134) [90]	37 (95) [32]	1 (58) [15]	0 (0) [2]	
CpG4	88	88 (0) [0]	145 (24) [7]	150 (99) [41]	175 (101) [24]	75 (203) [72]	18 (93) [39]	2 (18) [16]	0 (4) [0]	
1U IL-2	109	29 (79) [1]	6 (50) [2]	0 (12) [0]	-	-	-	-	-	
3U IL-2	90	68 (22) [0]	77 (48) [11]	46 (35) [73]	2 (4) [86]	4 (0) [0]	1 (0) [7]	0 (0) [2]	-	
10U IL-2	163	101 (62) [0]	122 (71) [9]	94 (59) [91]	12 (12) [164]	0 (0) [24]	-	-	-	
N4	37	30 (7) [0]	27 (33) [0]	12 (34) [8]	0 (16) [8]	-	-	-	-	
N4+CD28	38	33 (5) [0]	52 (14) [0]	8 (30) [66]	0 (16) [0]	-	-	-	-	
N4+IL-12	29	22 (7) [0]	33 (11) [0]	12 (18) [36]	5 (17) [2]	0 (6) [4]	-	-	-	
N4+CD28 + IL-12	36	32 (4) [0]	55 (9) [0]	22 (20) [68]	0 (8) [36]	-	-	-	-	
N4	45	20 (25) [0]	0 (40) [0]	-	-	-	-	-	-	
N4+CD28	41	22 (19) [0]	8 (36) [0]	0 (4) [12]	-	-	-	-	-	
N4+IL-2	37	28 (9) [0]	6 (50) [0]	0 (12) [0]	-	-	-	-	-	
N4+CD28 + IL-2	47	36 (11) [0]	31 (41) [0]	20 (38) [4]	0 (19) [21]	-	-	-	-	

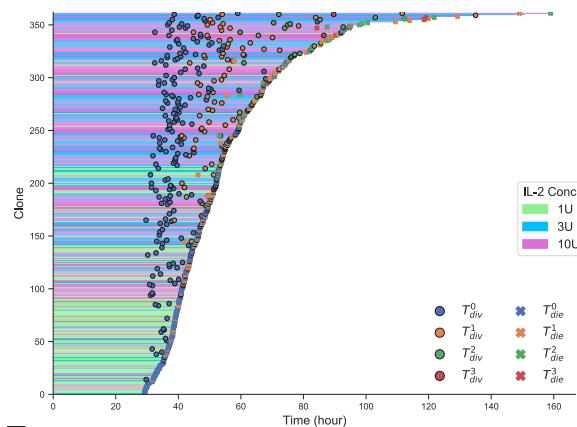
## B B cell (CpG3): All clones



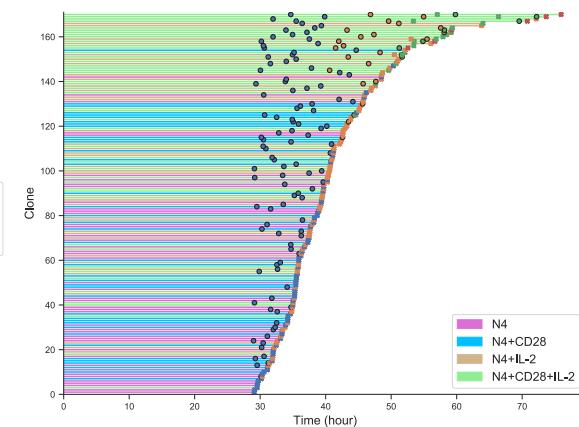
## C Repeat B cell (CpG4): All clones



## D CD8 T cell [1U, 3U, 10U of IL-2]: All clones



## E CD8 T cells [N4, CD28, IL-2] (Costim1): All clones



## F CD8 T cell [N4, CD28, IL-12] (Costim2): All clones

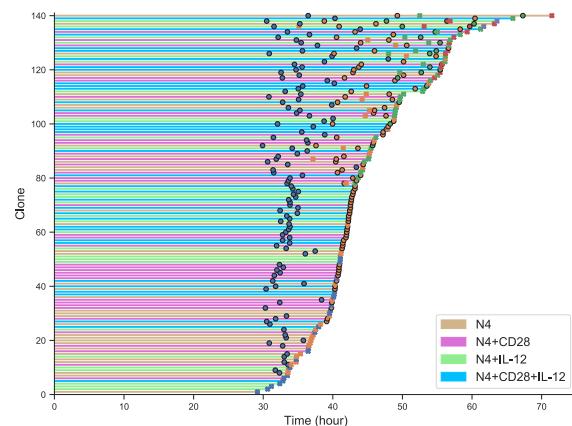
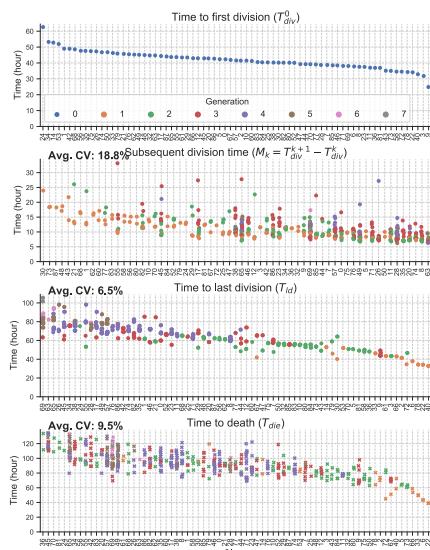


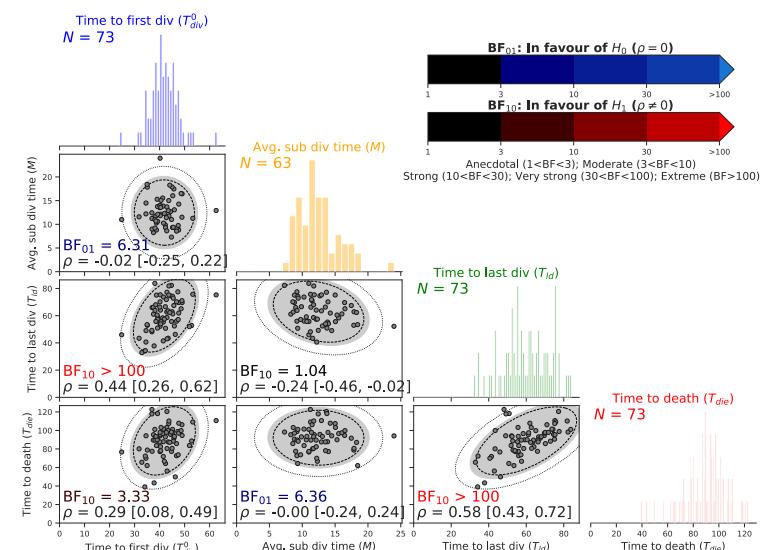
Figure S1: **Raw filming data.** (A) Number of founder cells and raw numbers of observed divided, dead and lost cells at each generation are tabulated. (B-F) Clonally collapsed trees for all families before the filtering. The average division (●) and death (×) times are marked, where the colors represent generation that the event was observed. The lost times are not shown.

**A1**

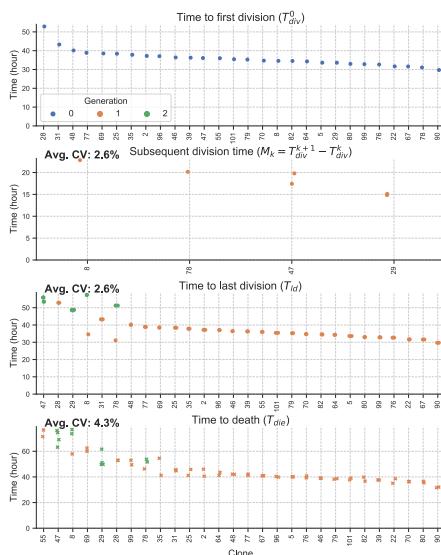
Repeat B cell (CpG4): Event Cascade Plot

**A2**

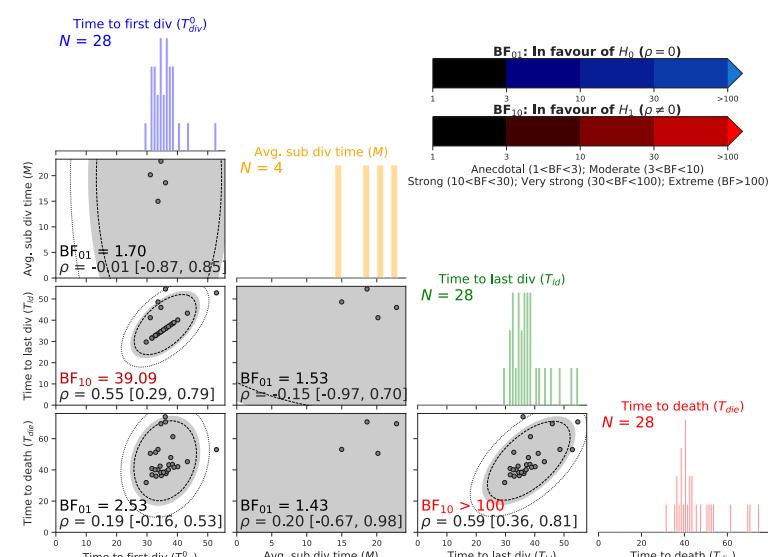
Repeat B cell (CpG4): Correlation

**B1**

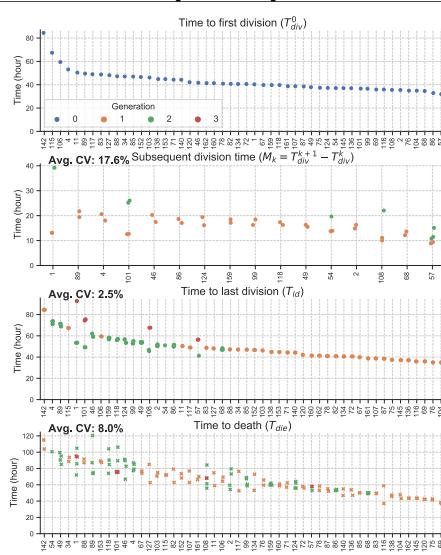
CD8 T cell [1U IL-2]: Event Cascade Plot

**B2**

CD 8 T cell [1U IL-2]: Correlation

**C1**

CD8 T cell [10U IL-2]: Event Cascade Plot

**C2**

CD 8 T cell [10U IL-2]: Correlation

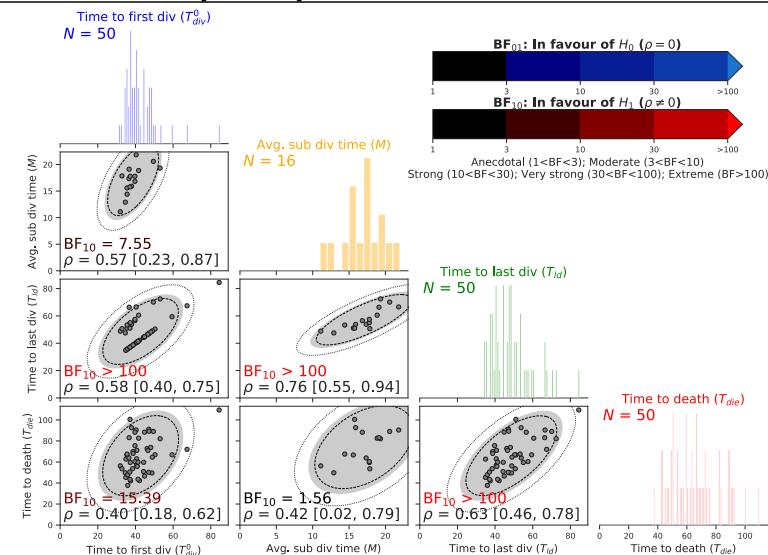
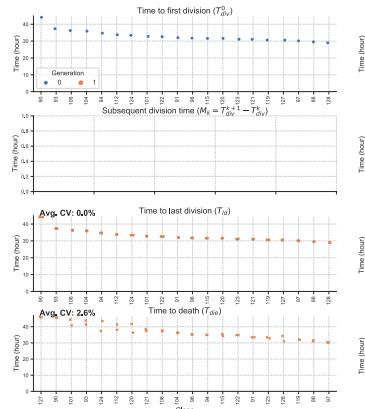


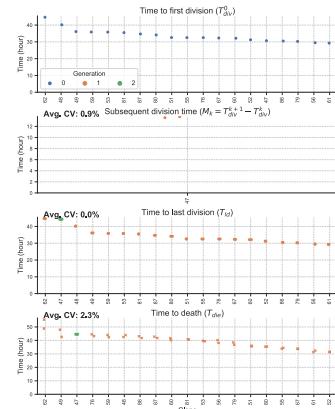
Figure S2: Extracting times to fates and the correlation analysis (A1-2) Repeat of CpG-stimulated B cell (CpG4). (B1-2) CD8 T cells in the presence of 1U of IL-2. (C1-2) CD8 T cells in the presence of 10U of IL-2.

### Costim1

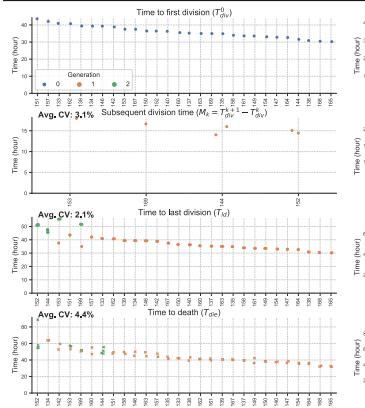
**A1** CD8 T cells [N4]



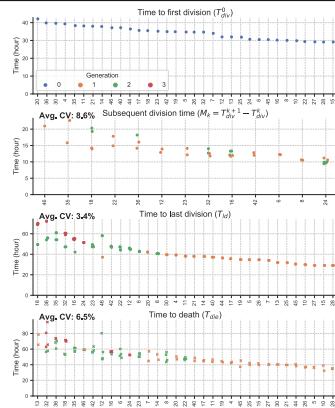
**A2** CD8 T cells [N4+CD28]



**A3** CD8 T cells [N4+IL-2]

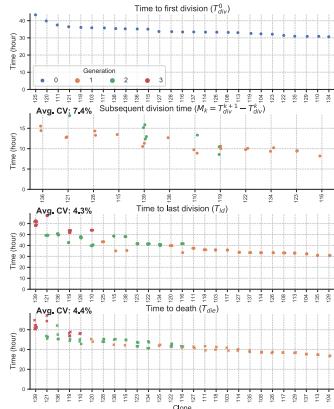


**A4** CD8 T cells [N4+CD28+IL-2]

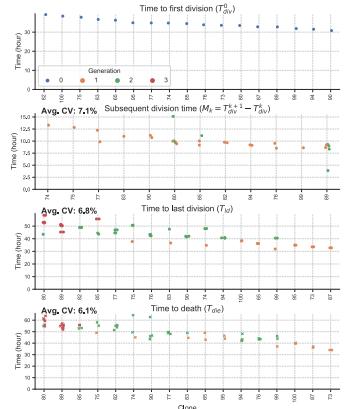


### Costim2

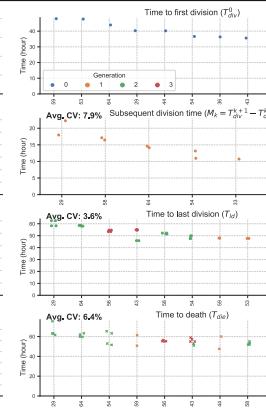
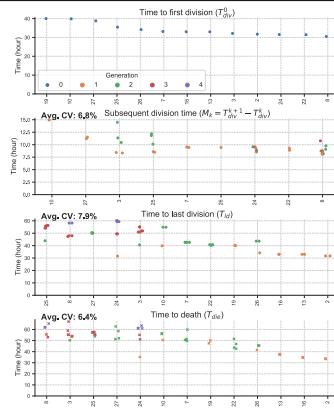
**B1** CD8 T cells [N4]



**B2** CD8 T cells [N4+CD28]



**B3** CD8 T cells [N4+IL-12]

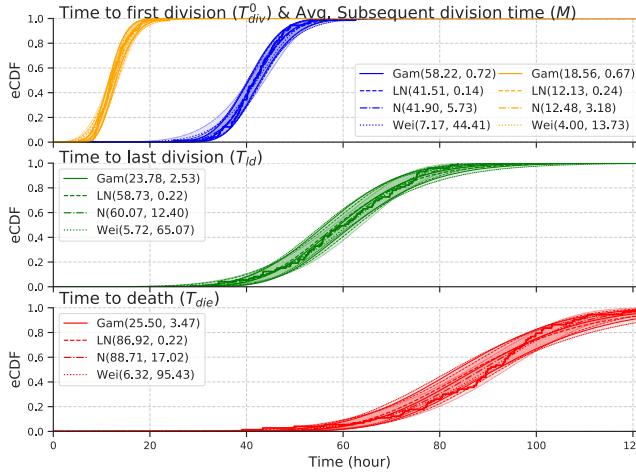


**Figure S3: Extracting times to fates.** **(A1-4)** Experiment Costim1 consists of N4, N4+CD28, N4+IL-2 and N4+CD28+IL-2. **(B1-4)** Experiment Costim2 consists of N4, N4+CD28, N4+IL-12 and N4+CD28+IL-12.

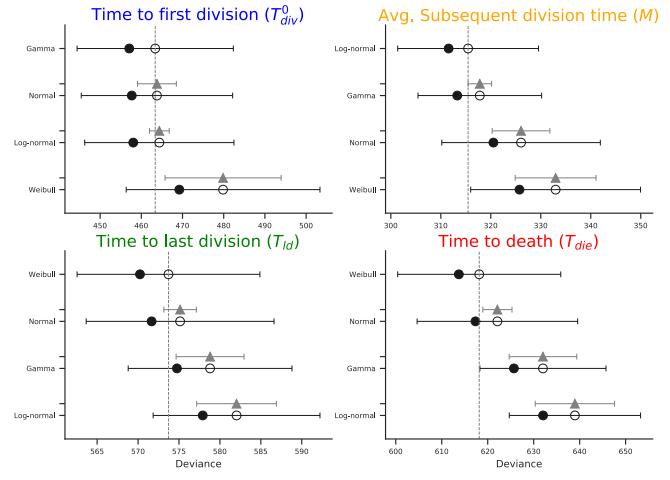
Cell Type	Stim.	Bayes Factor (01/10) & Correlation Coefficient ( $\rho$ [CI])					
		$(T_{div}^0, M)$	$(T_{div}^0, T_{ld})$	$(T_{div}^0, T_{die})$	$(M, T_{ld})$	$(M, T_{die})$	$(T_{ld}, T_{die})$
CD8 T (Costim1)	N4	NA	<b>&gt;100 (10)</b>	<b>8.31 (10)</b>	NA	NA	<b>8.31 (10)</b>
	N4+CD28	NA	1.00 [1.00, 1.00]	0.53 [0.22, 0.82]	NA	NA	0.53 [0.22, 0.82]
	N4+IL-2	NA	<b>85.63 (10)</b>	<b>18.83 (10)</b>	NA	NA	<b>60.52 (10)</b>
	N4+CD28+IL-2	0.67 [0.42, 0.89]	0.59 [0.30, 0.86]	NA	NA	0.66 [0.39, 0.89]	
CD8 T (Costim2)	N4	1.30 (01)	<b>29.98 (10)</b>	<b>2.09 (10)</b>	1.31 (01)	1.58 (01)	<b>&gt;100 (10)</b>
	N4+CD28	0.24 [-0.63, 0.99]	0.54 [0.28, 0.78]	0.38 [0.07, 0.68]	-0.24 [-0.99, 0.64]	-0.12 [-0.96, 0.73]	0.69 [0.49, 0.86]
	N4+IL-12	4.17 (10)	1.03 (10)	<b>3.28 (01)</b>	1.26 (01)	1.77 (01)	<b>&gt;100 (10)</b>
	N4+CD28+IL-12	0.55 [0.17, 0.88]	0.30 [-0.01, 0.59]	0.14 [-0.18, 0.46]	0.33 [-0.14, 0.77]	0.26 [-0.22, 0.73]	0.71 [0.54, 0.86]
	N4	15.12 (10)	2.04 (01)	<b>2.31 (01)</b>	2.69 (10)	<b>3.53 (10)</b>	<b>&gt;100 (10)</b>
	N4+CD28	0.68 [0.36, 0.94]	0.22 [-0.13, 0.56]	0.20 [-0.15, 0.54]	0.52 [0.10, 0.89]	0.55 [0.15, 0.90]	0.95 [0.91, 0.98]
	N4+IL-12	1.34 (01)	2.84 (01)	<b>2.75 (01)</b>	2.82 (01)	1.48 (01)	<b>&gt;100 (10)</b>
	N4+CD28+IL-12	0.33 [-0.17, 0.79]	0.13 [-0.31, 0.57]	0.14 [-0.30, 0.58]	-0.02 [-0.55, 0.52]	0.30 [-0.20, 0.77]	0.83 [0.67, 0.96]
	N4	<b>14.97 (10)</b>	2.58 (01)	<b>2.39 (01)</b>	2.21 (01)	2.24 (01)	<b>&gt;100 (10)</b>
	N4+CD28	0.74 [0.40, 0.98]	0.13 [-0.38, 0.63]	0.17 [-0.33, 0.66]	0.14 [-0.46, 0.74]	0.13 [-0.47, 0.73]	0.95 [0.88, 0.99]
	N4+IL-12	1.33 (01)	<b>2.47 (01)</b>	<b>2.48 (01)</b>	1.00 (01)	1.16 (10)	<b>29.19 (10)</b>
	N4+CD28+IL-12	0.32 [-0.33, 0.91]	0.12 [-0.43, 0.66]	0.12 [-0.44, 0.66]	0.40 [-0.22, 0.94]	0.44 [-0.18, 0.95]	0.74 [0.45, 0.96]

**Table S1: Test correlation of every pair of times to fates.** The Bayes factor and correlation coefficient calculated from a bivariate normal distribution with its 95% credible interval is shown for CD8 T cell in combination of N4, CD28, IL2 (Costim1) and of N4, CD28, IL12 (Costim2). Minimum of two observations per variable pair was required for the calculation.

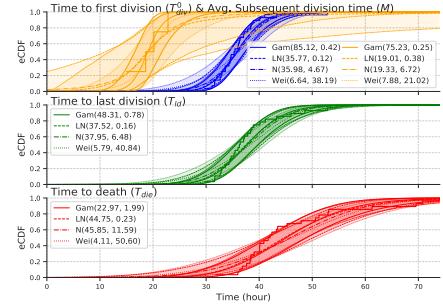
## A1 Repeat B cell (CpG4): Compare parametric distribution classes



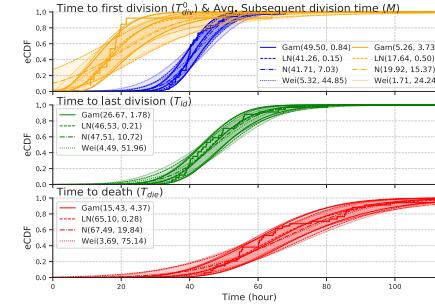
## A2 Repeat B cell (CpG4): WAIC Scores



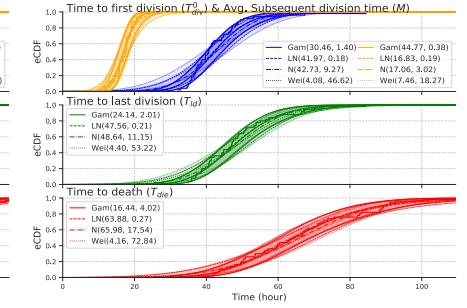
## B1 CD8 T cell [1U IL-2]



## B2 CD8 T cell [3U IL-2]



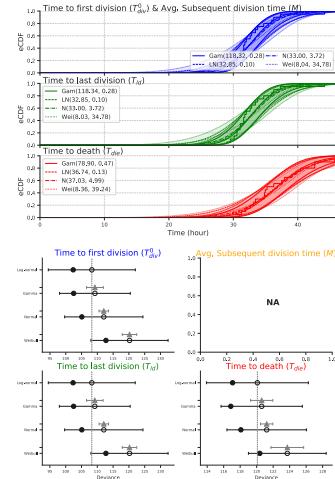
## B3 CD8 T cell [10U IL-2]



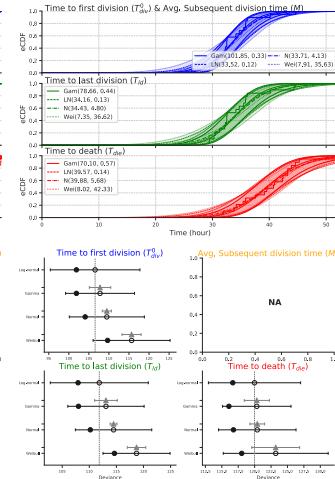
**Figure S4: Best parametric distribution classes.** Four candidate distribution classes parameterised by:  $(\alpha_G, \beta_G)$  for Gamma;  $(m, s)$  median and scale for lognormal;  $(\mu, \sigma)$  mean and standard deviation for Normal; and,  $(\alpha_W, \beta_W)$  for Weibull. **(A1-2)** Repeat of CpG-stimulated B cell (CpG4). The solid line of CDF are plotted by taking the mean of each parameters in respective posterior distribution. **(B1-3)** CD8 T cell in the presence of 1U, 3U and 10U of IL-2. Corresponding distribution fits to the WAIC scores reported Fig. 4B1-3 in the main text.

### Costim1

#### A1 CD8 T cells [N4]

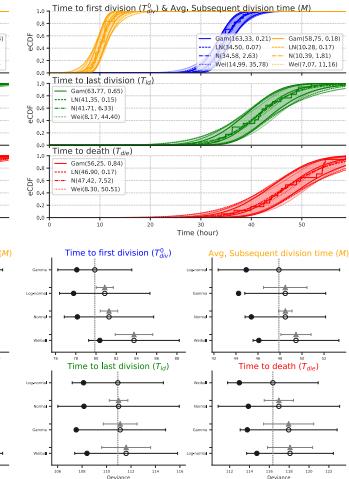


#### A2 CD8 T cells [N4+CD28]

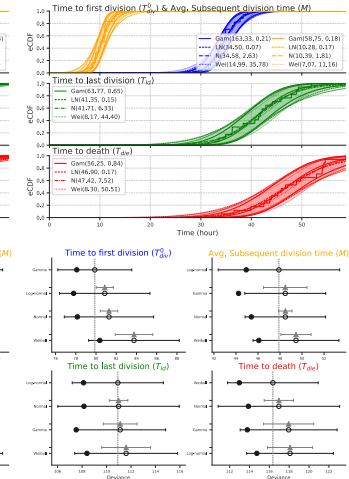


### Costim2

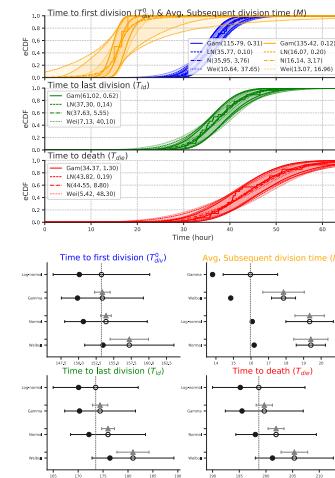
#### B1 CD8 T cells [N4]



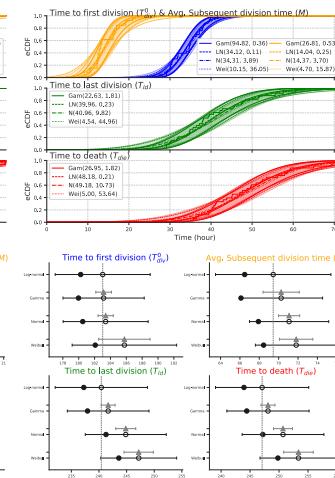
#### B2 CD8 T cells [N4+CD28]



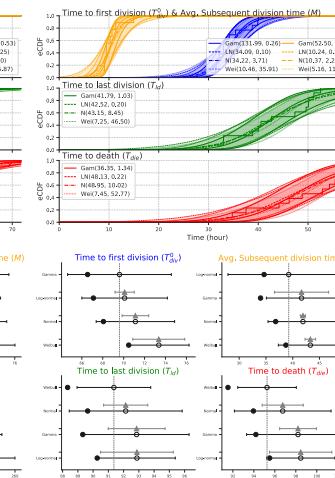
#### A3 CD8 T cells [N4+IL-2]



#### A4 CD8 T cells [N4+CD28+IL-2]



#### B3 CD8 T cells [N4+IL-12]



#### B4 CD8 T cells [N4+CD28+IL-12]

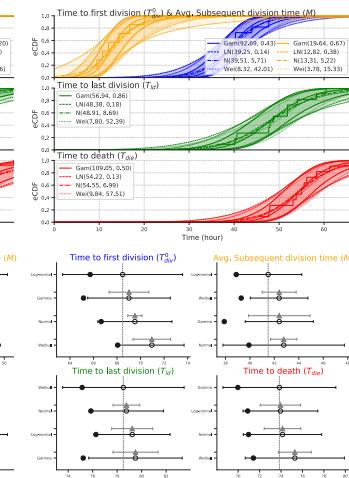
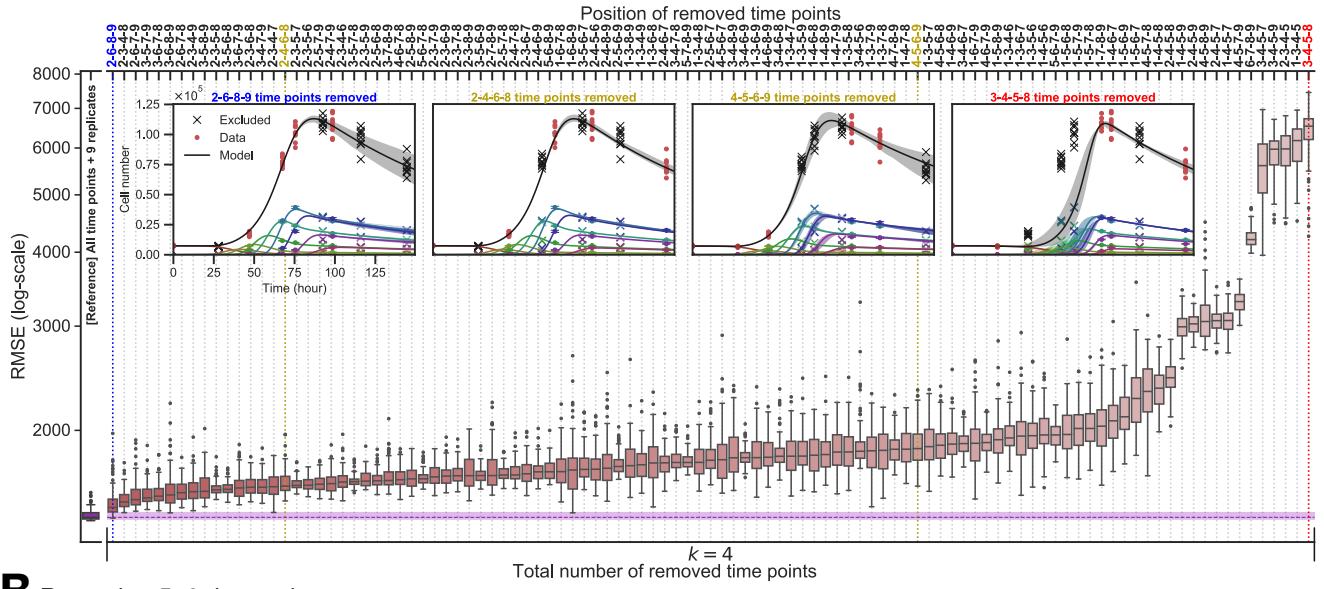


Figure S5: **Best parametric distribution classes.** **(A1-4)** Experiment Costim1 consists of N4, N4+CD28, N4+IL2 and N4+CD28+IL2. **(B1-4)** Experiment Costim2 consists of N4, N4+CD28, N4+IL-12 and N4+CD28+IL-12. Minimum of two observations per variable was required for the calculation.

## A Removing 4 time points



## B Removing 5, 6 time points

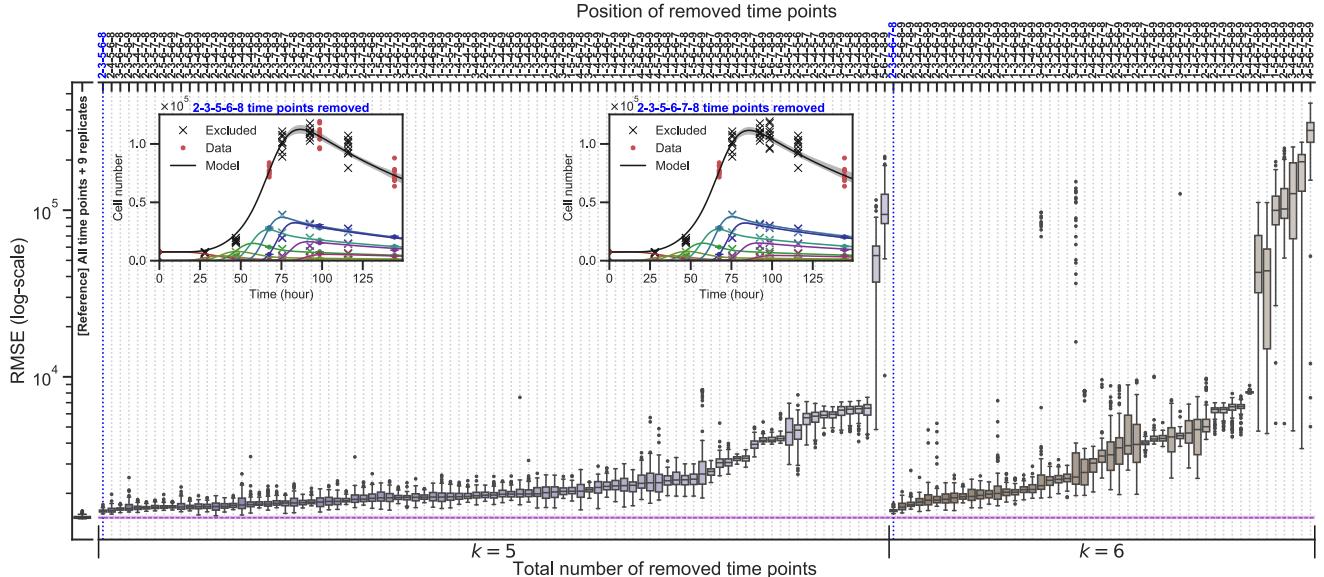


Figure S6: **The accuracy of the model fit with CpG-stimulated  $\text{Bim}^{-/-}$  B cell data.** Root-mean-squared error (RMSE) evaluated over all time points and replicates after fitting the Cyton 2 model to artificially removed dataset. The reference RMSE (purple) was obtained by fitting the model to all available data points. **(A)** All possible combinations of positions of time points for  $k = 4$  case. The best (blue) and worse (red) examples of Cyton fits are shown. **(B)** Best examples for  $k = 5, 6$  are shown. The worst cases failed to provide good fits, resulting in large confidence bands (not shown) and an order of magnitude difference in RMSE.

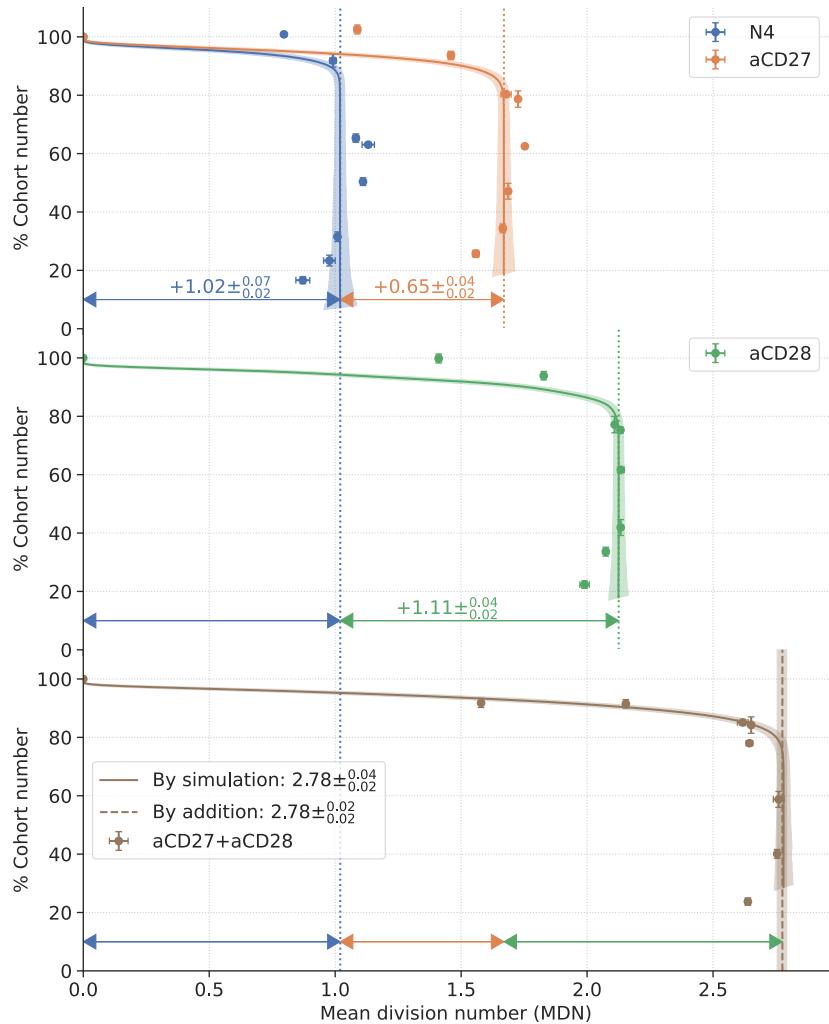


Figure S7: **Linear sum of the signals from simulated trees (Marchingo, Kan, et al. 2014).** The ABM was used to generate family trees. The times to fates were sampled from the estimated parameter values by fitting the Cyton 2 model. To match the data (●: mean  $\pm$  SEM), each round of the simulation was initialised with 6066 (N4), 8252 (aCD27) and 8377 (aCD28) clones and ran for  $t \in [0, 140]$  with  $\Delta t = 0.5$  in hours. This process was repeated  $10^3$  times to obtain 95% confidence bands around the mean. Increase in MDN is labelled in each panel with arrows. The predicted MDNs for aCD27+aCD28 were calculated either by summing the contribution from each individual stimulation (—) or simulating the trees directly with the summed timers (—).