Defining the dynamics of naive CD4 and CD8 T cells across the mouse lifespan

Sanket Rane[1], Thea Hogan[2], Edward Lee[3], Benedict Seddon[2*], Andrew J. Yates[1*]

1 Department of Pathology and Cell Biology, Columbia University Irving Medical Center, 630 West 168th Street, New York, NY 10032, USA

2 Institute of Immunity and Transplantation, Division of Infection and Immunity, UCL, Royal Free Hospital, Rowland Hill Street, London NW3 2PF, United Kingdom

3 Department of Laboratory Medicine, Yale University School of Medicine, 55 Park Street, New Haven, Connecticut 06511, USA

*Address correspondence to either author; benedict.seddon@ucl.ac.uk, andrew.yates@columbia.edu

Abstract

Naive CD4 and CD8 T cells are part of the foundation of adaptive immune responses, but multiple aspects of their behaviour remain elusive. Newly generated T cells continue to develop after they leave the thymus and their dynamics and 'rules of entry' into the mature naive population are challenging to define. The extents to which naive T cells' capacities to survive or self-renew change as they age are also unclear. Further, much of what we know about their behaviour derives from studies in adults, both mouse and human. We know much less about naive T cell dynamics early in life, during which the thymus is highly active and peripheral T cell populations are rapidly established. For example, it has been suggested that neonatal mice are lymphopenic; if so, does this environment impact the behaviour of the earliest thymic emigrants, for example through altered rates of division and loss? In this study we integrate data from multiple experimental systems to construct models of naive CD4 and CD8 T cell population dynamics across the entire mouse lifespan. We infer that both subsets progressively increase their capacity to persist through survival mechanisms rather than through self-renewal, and find that this very simple model of adaptation describes the population dynamics of naive CD4 T cells from birth into old age. In addition, we find that newly generated naive CD8 T cells are lost at an elevated rate for the first 3-4 weeks of life, which may derive from transiently increased recruitment into conventional and virtual memory populations. We find no evidence for elevated rates of division of naive CD4 or CD8 T cells early in life and indeed estimate that these cells divide extremely rarely. Markers of proliferation within peripheral naive T cells are instead inherited from division during thymic development. We also find no evidence for feedback regulation of rates of division or loss of naive T cells at any age in healthy mice, challenging the dogma that their numbers are homeostatically regulated. Our analyses show how confronting an array of mechanistic mathematical models with diverse datasets can move us closer to a complete, and remarkably simple, picture of naive CD4 and CD8 T cell dynamics in mice.

Introduction

To provide a basis for lifelong adaptive immunity, populations of naive CD4 and CD8 T cells with diverse antigen receptors must be generated rapidly from birth and be maintained throughout life. In mice, the number of circulating naive T cells grows from tens of thousands at birth to tens of millions in several weeks, peaking at around 2 months of age[1,2] and waning thereafter. Quantifying the relative contributions of thymic influx, and of loss and self-renewal (processes which boost and preserve diversity, respectively) across the lifespan will therefore help us understand at a mechanistic level how the TCR repertoire is generated and evolves as an individual ages.

Naive T cells in adult mice have a mean lifespan of a few weeks and a mean interdivision time of the order 1-2 years[2,3]. The disparity in these timescales leads to the conclusion that, in mice, most naive T cells never divide and that their numbers are sustained largely by thymic export, which in adult mice contributes 1-2% of the peripheral pool size per day[1–5]. However, the dynamics of the naive pool may be radically different early in life. Thymic output is proportionally larger in young mice; depletion of thymocytes in 2 week-old mice for 7 days early in life drives a rapid and transient 50-70% reduction of peripheral CD4 and CD8 T cell numbers[6]. Despite this influx, it is possible that the neonatal mouse is lymphopenic, a state which, when artificially induced, supports the rapid expansion of newly-introduced T cells through a mechanism referred to as lymphopenia-induced proliferation (LIP)[7–11]. There is some evidence for LIP of naive T cells in healthy neonatal mice[12], although in that study the proliferation was associated with a conversion to a memory phenotype. However, another study, which followed the behaviour of adult-derived T cells transferred into neonatal mice[13], observed low and roughly equal levels of proliferation in both endogenous and donor-derived cells, but no overall expansion of the donor population. The latter results imply that the healthy neonatal environment does not drive increased proliferation, and that naive T cells derived from infant and adult mice divide at similar rates.

The average residence times of naive T cells may also be different in neonates and adults. Loss can occur through death or by activation, which drives naive cells to acquire an effector or memory phenotype. The relative contributions of these two processes are not known, and may not be constant over time. For instance, we have shown that the rate of generation of memory CD4 T cells is elevated early in life, at levels influenced by the antigenic content of the environment[14]. This result suggests that high *per capita* rates of activation upon first exposure to environmental stimuli may increase the apparent rate of loss of naive CD4 T cells in neonatal mice. One might expect a similar process to occur with naive CD8 T cells, with substantial numbers of so-called 'virtual' memory CD8 T cells generated from naive T cells in the periphery soon after birth[15].

The post-thymic age of a cell, rather than the age of the host at which that cell was exported from the thymus, may also impact its behaviour. Following the dynamic period of their establishment, there is evidence that naive T cells do not die or self-renew at constant rates, but continue to respond or adapt to the host environment[16]. Recent thymic emigrants (RTE) are both functionally distinct from mature T cells[17,18], may be lost at a higher rate than mature naive T cells under healthy conditions[10,19–21], and respond differently to mature naive cells under lymphopenia[10]. Phenotypic markers of RTE are poorly defined, however, and so without a strict definition of 'recent' it is difficult to reach a consensus description of their kinetics. It may be more appropriate to view maturation as a continuum of states, and indeed the net loss rates (the balance of loss and self-renewal) of both naive CD4[22] and CD8[22,23] T cells in mice appear to fall smoothly with a cell's post-thymic age, a process we have referred to as adaptation[22]. Such behaviour will lead to increasing heterogeneity in the kinetics of naive T cells over time, as the population's age-distribution broadens, and may also contribute to skewing of the T cell receptor (TCR) repertoire, through a 'first-in, last-out' dynamic in which older naive T cells become progressively fitter than newer immigrants[3].

The inferences above are derived from models in which naive T cells behave independently in the periph-

ery, but the analyses do not preclude the existence of quorum-sensing, in which rates of cell division and/or loss may be influenced by competition for 'space', or shared resources such as cytokines. Density-dependent effects may help to at least partially restore naive T cell numbers in the face of substantially reduced output from the thymus[8], and can explain the decreasing population-average loss rate of naive cells over time in thymectomised mice[2]. However, it is unclear whether density-dependent dynamics are manifest under normal conditions in either neonates (where it may underlie LIP) or in adults.

Taken together, these results indicate that host age, cell age, and cell numbers may all influence naive T cell dynamics to varying degrees. When dealing with cross-sectional observations of cell populations, these effects may be difficult to distinguish. For example, the decreasing population-average loss rate of naive cells over time in thymectomised mice[2] may not derive from reduced competition, as was suggested, but may also be explained by adaptation[22]. It is also possible that elevated loss rates of naive T cells early in life may not be an effect of the neonatal environment *per se*, but just a consequence of the nascent naive T cell pool being comprised almost entirely of RTE with intrinsically shorter residence times than mature cells. These uncertainties invite the use of mathematical models to distinguish different descriptions of naive T cell population dynamics from birth into old age. Here we combine model selection tools with data from three distinct experimental systems, using fits to one dataset to predict another, rather than fitting models to all datasets simultaneously and relying solely on statistical methods to weigh the support for each. We believe this out-of-sample prediction approach provides more intuitive and stringent tests of models' explanatory power. We find that naive CD4 T cell appear to follow simple rules of behaviour throughout the mouse lifespan, dividing very rarely and increasing their survival capacity with cell age, with no evidence for altered behaviour in neonates. Naive CD8 T cells behave similarly, but with an additional, increased rate of loss during the first few weeks of life that may reflect high levels of recruitment into early memory populations. These models are able to explain diverse observations and present a remarkably simple picture in which naive T cells appear to be passively maintained throughout life, with gradually extending lifespans that compensate in part from the decline in thymic output, but no evidence for feedback regulation of cell numbers.

Results

Naive CD4 and CD8 T cells divide very rarely in adult mice and expected lifespans increase with cell age

Recent reports from our group and others[3,22–24] show that the dynamics of naive CD4 and CD8 T cells in adult mice and humans depend on cell age, defined to be time since they (or their ancestor, if they have divided) were released from the thymus. All of these studies found that the net loss rate, which is the balance of their rate of loss through death or differentiation, and self-renewal through homeostatic division, decreases gradually with cell age for both subsets. It is unknown whether these adaptations with age modulate the processes that regulate their residence time, or their ability to self-renew.

To address this question we used data from a well-established system that we have used to quantify lymphocyte dynamics at steady state in healthy mice, over almost the entire lifespan[25]. Hematopoietic stem cells (HSCs) in the bone marrow (BM) are partially and specifically depleted by optimized doses of the transplant conditioning drug busulfan, and reconstituted with T and B cell depleted BM from congenic donor mice. The host's peripheral lymphocyte populations are unperturbed by treatment, and chimerism rapidly stabilises within the thymus and is maintained for the lifetime of the mouse (Figure 1A). As donor T cells develop they progressively replace host T cells in the periphery through natural turnover. This system allows us to estimate the rates of influx into different lymphocyte populations and the net loss rates of cells within them; identify subpopulations with different rates of turnover; and infer whether and how these dynamics

vary with host and/or cell age[3,26,27].

To use this system to model cell dynamics, we generated a cohort of busulfan chimeric mice that underwent bone marrow transplant (BMT) between the ages of 7 and 25 weeks, and at different times post-BMT enumerated host and donor-derived thymocyte subsets and peripheral naive T cells from spleen and lymph nodes (Figure 1B; see Figure S1 for the flow cytometric gating strategy). We begin by normalising the chimerism (fraction donor) within naive CD4 and CD8 T cells to that of DP1 thymocytes to remove the effect of variation across mice in the stable level of bone-marrow chimerism. This normalised donor fraction (f_d) will approach 1 within a population if it turns over completely – that is, if its donor:host composition equilibrates at that of its precursor. Saturation at $f_d < 1$ implies incomplete replacement, which can occur either through waning influx from the precursor population, or if older (host) cells persist longer than new (donor) cells, on average, implying cell-age effects on turnover or self-renewal. Previously we observed incomplete replacement of both naive CD4 and CD8 T cells in adult busulfan chimeric mice[3], and excluded the possibility that this shortfall derived from the natural involution of the thymus, leading us to infer that the net loss rates (the balance of loss and self-renewal through division) of both subsets increase with cell age[3,22]. For the present study we used concurrent measurements of Ki67, a nuclear protein that is expressed following entry into cell cycle and is detectable for approximately 3-4 days afterwards[26,28], and stratified by host and donor cells (Figure 1B, lower panels). We reasoned that this new information would enable us to establish whether these cell-age effects are manifest through survival or self-renewal.

To describe these data we explored variants of a structured population model in which either the rate of division or loss of naive T cells varies exponentially with their post-thymic age. These models are three dimensional linear PDEs that extend those we described previously[3,22], allowing us to simultaneously track the joint distribution of cell age and Ki67 expression within the population. The formulation of these models is detailed in Text S1. We also considered a simpler variant of the age-structured models that explicitly distinguishes RTE from mature naive T cells, with a constant rate of maturation between two, and allowing each to have their own rates of division and loss[21,22]. Finally, we included a class of models of homogeneous cell dynamics; the simplest 'neutral' model with uniform and constant rates of division and loss and density-dependent models that allowed these rates to vary with population size. Each of these models was fitted simultaneously to the measured time-courses of total naive CD4 or CD8 T cell numbers ($N_{\text{donor}}(t) + N_{\text{host}}(t)$, the normalised donor fraction $f_d(t)$, and the proportions of donor and host cells expressing Ki67 ($k_{\text{donor}}(t)$ and $k_{\text{host}}(t)$). To model influx from the thymus we used empirical functions fitted to the numbers and Ki67 expression levels of late stage single-positive CD4 and CD8 thymocytes (Text S2 and Figure S2). Since the rate of export of cells from the thymus is proportional to the number of single-positive thymocytes[19], we used these curves to represent the export of new Ki67$^{\text{hi}}$ and Ki67$^{\text{low}}$ RTE, up to a multiplicative constant which we estimated. We used a Bayesian fitting approach and compared the support for models with the leave-one-out information criterion (LOO-IC)[29,30]. A detailed description of the fitting procedure can be found in ref. 27.

Our analysis confirmed support for the models of cell-age dependent kinetics, with all other candidates, including the RTE model, receiving substantially lower statistical support (Table S1). For naive CD4 T cells, we found strongest support for the age-dependent loss model (relative weight = 84%), which described their dynamics well (Figure 2A) and revealed that their daily loss rate halves roughly every 3 months (Table 1) as they age. For naive CD8 T cells the age-dependent division model was favoured statistically (relative weight = 87%; Figure 2B, dashed lines). However it yielded extremely low division rates, with recently exported cells having estimated mean interdivision times of 18 months (95% CI: 14–25), and the division rate increasing only very slowly with cell age (doubling every 10 months). This model was therefore very similar to a neutral, homogeneous model and predicted that the normalised donor fraction approaches 1 in aged mice. This conclusion contradicts findings from our own and others' studies that demonstrated that models assuming homogene-

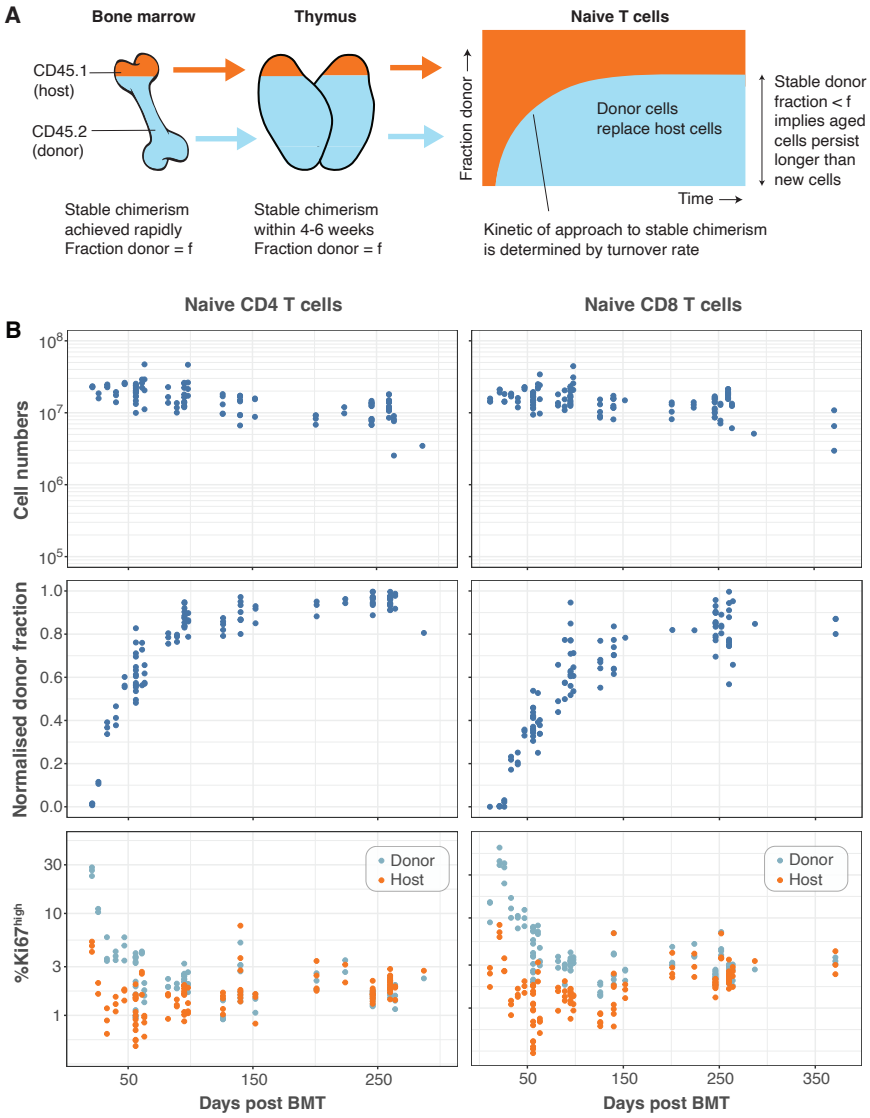

Figure 1: Dynamics of naive CD4 and CD8 T cells in busulfan chimeras. **(A)** The busulfan chimera system for studying lymphocyte population dynamics. **(B)** Observations of total naive T cell numbers derived from spleen and lymph nodes, their normalised donor fraction, and Ki67 expression among host and donor cells, in adult busulfan mice at various times post BMT.

ity in naive CD8 T cells failed to capture their dynamics in adult and aged mice (2-20 months old)[3,22,23]. A prediction of any progressive increase in cell division rates with age is also inconsistent with observations of the frequencies of T cell receptor excision circles (TRECs) in adult mice. TRECs are stable remnants of DNA generated during T cell receptor rearrangement, which are only diluted through mitosis; den Braber *et al.*[2] found that the average TREC content of naive CD4 and CD8 T cells in mice aged between 12-125 weeks closely resembled the average TREC content of SP4 and SP8 thymocytes, suggesting that naive T cells in mice hardly divide. The statistical support for the age-dependent division model may derive from the relatively sparse observations in aged mice in this dataset, which define the asymptotic replacement fraction. For the next phase of analysis we therefore retained the age-dependent loss model, which had the next highest level of support and was similar by visual inspection (Figure 2B, solid lines), as a candidate description of naive CD8 T cell dynamics.

Population	Parameter	Mean	95% confidence interval
Naive CD4	Expected residence time of cells of age 0 (days)	22	(18, 28)
	Time taken for loss rate to halve (days)	92	(71, 130)
	Mean interdivision time (months)	18	(16, 22)
Naive CD8	Expected residence time of cells of age 0 (days)	40	(34, 46)
	Time taken for loss rate to halve (days)	146	(107, 206)
	Mean interdivision time (months)	14	(12, 16)

Table 1: Parameter estimates derived from fitting the age-dependent loss model to data from busulfan chimeras. Residence and interdivision times are defined as the inverses of the instantaneous loss rate ($\delta(a)$) and the division rate (ρ), respectively.

Age-dependent loss models successfully predict the relative persistence of RTE and mature naive CD4 and CD8 T cells

To challenge these models further, we next confronted them with data from a study that compared the ability of RTE and mature naive (MN) CD4 and CD8 T cells to persist following co-transfer to an adult congenic recipient[10]. This study used a reporter mouse strain in which green fluorescent protein (GFP) is linked to the *Rag* gene, which is expressed throughout thymic development and for several days following export into the periphery. GFP expression can then be used as a surrogate marker of RTE status. After transferring RTE and MN cells in equal numbers, the RTE:MN ratio within both CD4 and CD8 populations decreased progressively by approximately 50% by 6 weeks (Figure 3), indicating that both MN CD4 and CD8 cells persist significantly longer than RTE. We simulated this co-transfer using the models fitted to the data from the busulfan chimeric mice, and found that the age-dependent loss model predicted the trend in the RTE:MN ratio (Figure 3A) while the age-dependent division model, which exhibited very weak age effects, predicted that the ratio would remain close to 1 (Figure 3B). We therefore favour models in which the expected residence times of naive CD4 and CD8 T cells increase with their age, while self-renewal occurs at very low, constant levels.

Ki67 expression within naive CD4 and CD8 T cells in adult mice is almost entirely a residual signal of intra-thymic proliferation

Our analyses are consistent with earlier reports that naive T cells divide very rarely[2,3,13]; indeed the disparity between expected lifespans and interdivision times implies that the majority of naive T cells in adult mice will never divide. Our models also explain the apparently contradictory observation that an appreciable fraction (2-3%) of naive T cells in adult mice express Ki67; we can infer that this is almost entirely derived from cells that divided in the thymus and were exported within the previous few days, and not from self-renewal in the periphery. This result also gives an intuitive explanation of the trajectories of Ki67 expression within donor and host cells in the busulfan chimeric mice, which are distinct soon after BMT but converge after

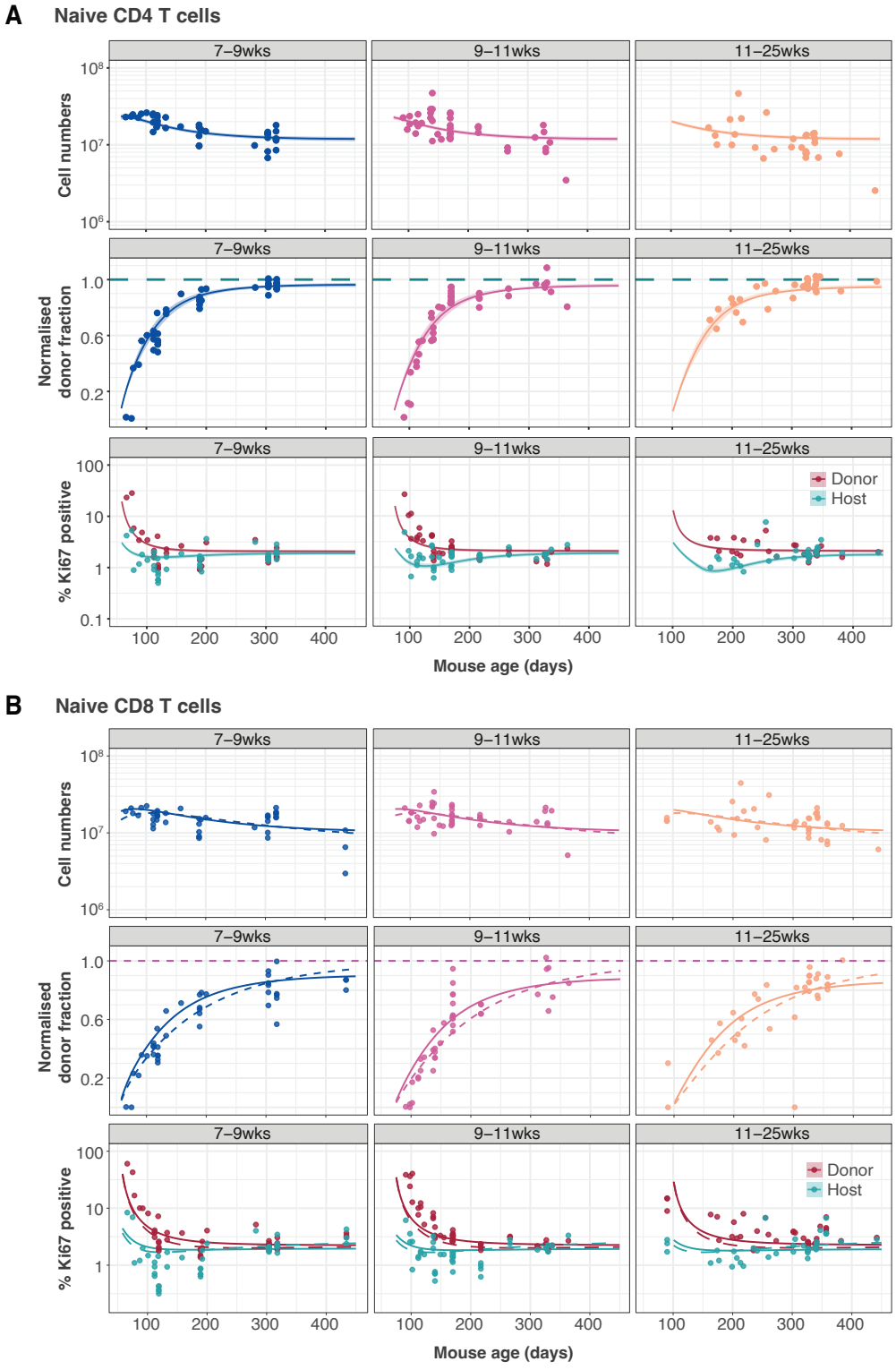

Figure 2: Modelling naive CD4 and CD8 T cell dynamics in adult busulfan chimeric mice. **(A)** The best fitting, age-dependent loss model of naive CD4 T cell dynamics describes the timecourses of cell numbers, chimerism and Ki67 expression within host and donor cells in mice who underwent busulfan treatment and BMT at different ages (indicated in grey bars). **(B)** Naive CD8 T cell dynamics described by the age-dependent division model (dashed lines) and the age-dependent loss model (solid lines).

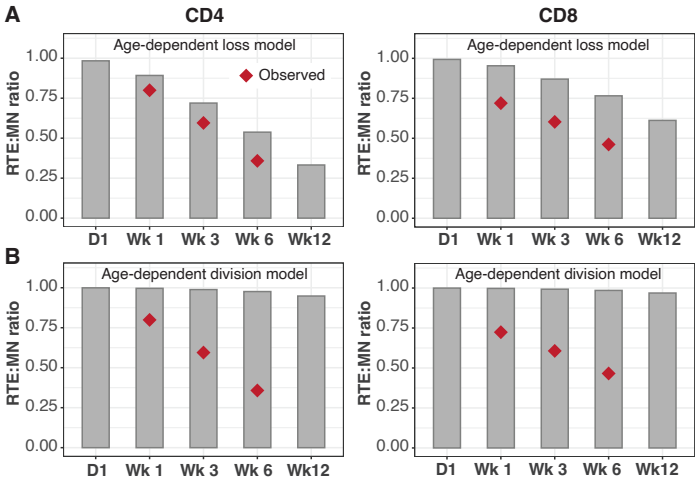

Figure 3: Models of age-dependent division fail to explain the survival kinetics of RTE and mature naive T cells. We simulated the co-transfer experiment described by Houston *et. al.*[10] in which RTE from 5-9 week old Rag^{GFP} reporter mice were co-transferred with equal numbers of mature naive (MN) T cells from mice aged 14 weeks or greater to congenic recipients. Red diamonds represent their observed RTE:MN ratios. Grey bars show the ratios predicted by the best-fitting age-dependent loss models (**A**) and age-dependent division models (**B**).

6-12 months (Figure 1B and Figure 2). This behaviour does not derive from any intrinsic differences between host and donor T cells, but rather from the distinct age profiles of the two populations. Following BMT, the rate of production of host naive T cells declines substantially, as the procedure typically results in 80-90% replacement of host HSC with donor HSC. Since Ki67 is seen almost exclusively within very recent thymic emigrants, the frequency of Ki67-expressing host naive T cells then declines rapidly. Conversely, new donor-derived naive T cells are initially highly enriched for $Ki67^{hi}$ cells. The frequencies of $Ki67^{hi}$ cells within the two populations then gradually converge to pre-transplant levels as aged $Ki67^{low}$ donor cells accumulate, and host-derived naive T cells equilibrate at lower numbers.

Models parameterised using data from adult mice accurately predict the dynamics of naive CD4 T cells in neonates, but not of CD8 T cells

The analyses above were performed in adult mice aged between 2-15 months. Next, we wanted to characterise the dynamics of naive CD4 and CD8 T cells during the first few weeks of life, and connect the two regimes to build unified models of the dynamics of these populations across the lifespan. Due to the difficulty of generating very young chimeric mice, we took the approach of extrapolating the fits from adults back to near birth, and using them to predict observations derived from different experimental systems. First, we drew on data from a cohort of $Rag^{GFP}Ki67^{RFP}$ reporter mice. In these animals a green fluorescent protein (GFP) is linked to the *Rag* gene, as in the study by Houston *et al.*[10], and a red fluorescent protein (RFP) is fused to Ki67, giving a clear readout of its expression without the need for cell fixation and permeabilisation. Tracking GFP and RFP expression simultaneously therefore allows us to study the kinetics and division rates of RTE, which are enriched for GFP+ cells, and of mature naive T cells, which are expected to have largely lost GFP. We could then directly confront the models derived from adult mice with these new data.

Figure 4 shows the numbers and GFP/Ki67 expression of naive CD4 and CD8 T cells in reporter mice aged between 10 and 120 days (black points). The red points show the previously analysed data from the adult busulfan chimeras, with host and donor cells pooled. The blue points are derived from a third independent dataset, comprising the numbers and Ki67 expression of naive T cells derived from wild-type mice aged

between 5 and 300 days. The red curves show the predictions of the age-dependent loss models, extrapolated back to 1 day after birth. To predict the kinetics of GFP+Ki67- and GFP+Ki67+ proportions we needed to estimate only one additional parameter – the average duration of GFP expression. We assume that RTE become GFP-negative with first order kinetics. This timescale is defined both by the intrinsic rate of decay of GFP and the threshold of expression used to define GFP+ cells by flow cytometry. Our estimates of the duration of GFP expression within CD4 and CD8 RTE were similar (11 and 8 days, respectively). Details of our analysis are provided in Text S3.

Strikingly, the model of naive CD4 T cell dynamics in adult chimeric mice captured the dynamics of these cells in neonates remarkably well (Figure 4A). This agreement indicates that the high level of Ki67 expression in naive CD4 T cells early in life does not reflect increased rates of division or LIP, but, as was inferred in adults, is inherited from cells proliferating extensively within the neonatal thymus (Figure S2).

For naive CD8 T cells the age-dependent loss model accurately predicted cell dynamics in both the reporter and wild-type mice back to age $\sim$3 weeks, but underestimated Ki67$^{\text{hi}}$ levels in neonatal mice (Figure 4B), suggesting that naive CD8 T cells exhibit distinct dynamics very early in life. Intuitively this mismatch can be explained in two ways; either CD8 RTE in neonates are lost at a higher rate than in adults, or they divide more rapidly. In the former, a greater proportion of GFP^{+} Ki67^{+} RTE will be lost before they become Ki67$^{\text{low}}$ and so the predicted proportion of cells that are GFP^{+} Ki67$^{\text{low}}$ will be lower (Figure 4D, centre panel). In the latter, the proportion GFP^{+} Ki67$^{\text{hi}}$ will increase (Figure 4D, right panel). Therefore, to explain naive CD8 T cell dynamics in neonates the basic model of cell-age-dependent loss in adults can be extended in two ways, modulating the rates of either division or loss early in life.

Naive CD8 T cells are lost at a higher rate in neonates than in adults

To distinguish between these possibilities we turned to a published study by Reynaldi *et al.*[23], which used an elegant tamoxifen-driven CD4-Cre$^{\text{ERT2}}$-RFP reporter mouse model to track cohorts of CD8 T cells released from the thymus into the peripheral circulation of animals of varying ages. In this model, a pulse of tamoxifen permanently induces RFP in cells expressing CD4, including CD4+CD8+ double positive thymocytes. The cohort of naive CD8 T cells deriving from these precursors will continue to express RFP in the periphery and timecourses of their numbers in individual mice can be estimated with serial sampling of blood[23]. These timecourses showed that the net rate of loss of naive CD8 T cells appears to slow with their post-thymic age (Figure 5A). In addition, the initial rate of loss of these cohorts appears to be greater in neonates than in adults. Without measures of proliferation these survival curves reflect only the balance of survival and self-renewal. Nevertheless we reasoned that confronting our models with these additional data, and triangulating with inferences from other datasets, would allow us to identify a 'universal' model of naive CD8 T cell loss and division across the mouse lifespan.

We re-analysed the data from Reynaldi *et al.* using our models and a Bayesian hierarchical approach (Text S4) to explain the variation in the kinetics of loss of these cohorts of cells across animals and age groups. Since there was no readout of cell division in this system, we simplified the cell-age-dependent loss model to combine division and loss into a net loss rate $\lambda(a)$. We then fitted this model to the timecourses of labelled naive CD8 T cells across the different treatment groups. We tested four possibilities in which either the initial numbers of labelled cells (N_0) and/or the net loss rate of cells of age 0 (λ_0) are varied across groups or animals as normally-distributed hyper-parameters (Table S2). The model in which N_0 was specific to each mouse and λ_0 was specific to each age group gained 100% relative support (Figure 5A, solid lines; Table S2). This model confirmed that CD8 RTE are indeed lost at a significantly higher rate in the younger groups of mice (Figure 5B). We then described this decline in λ_0 with mouse age empirically with a sigmoid (Hill) function, $\lambda_0(t)$ (Figure 5C and Text S4) and used it to replace the discrete group-level variation in λ_0 within the hierar-

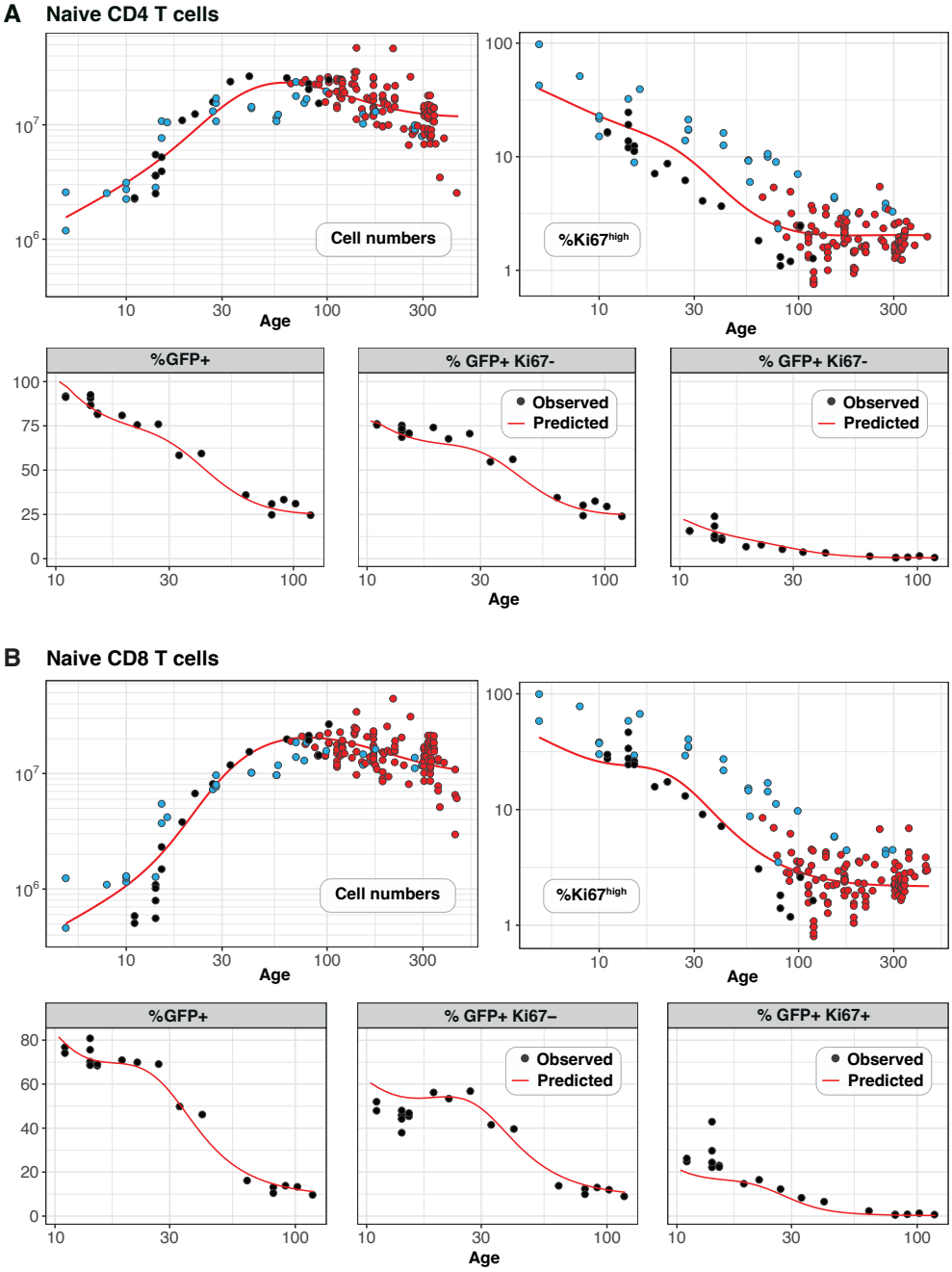

Figure 4: Predicting the kinetics of establishment of naive CD4 and CD8 T cell pools in early life. For naive CD4 (A) and CD8 (B) T cells, we extrapolated the age-dependent loss models (red curves) that were fitted to data from adult busulfan chimeric mice (red points) back to age 1 day. We compared these predicted trajectories with independent observations of naive T cell numbers and Ki67 expression in wild-type mice aged between 5-300 days (blue points), and from $Rag^{GFP}Ki67^{RFP}$ reporter mice (black points). We estimated one additional parameter – the expected duration of GFP expression – by fitting the age-dependent loss model to the timecourse of total numbers of GFP+ cells. We were then able to use the fitted models to predict the timecourses of the percentages of GFP+Ki67+ and GFP+Ki67- naive CD4 and CD8 T cells in these reporter mice.

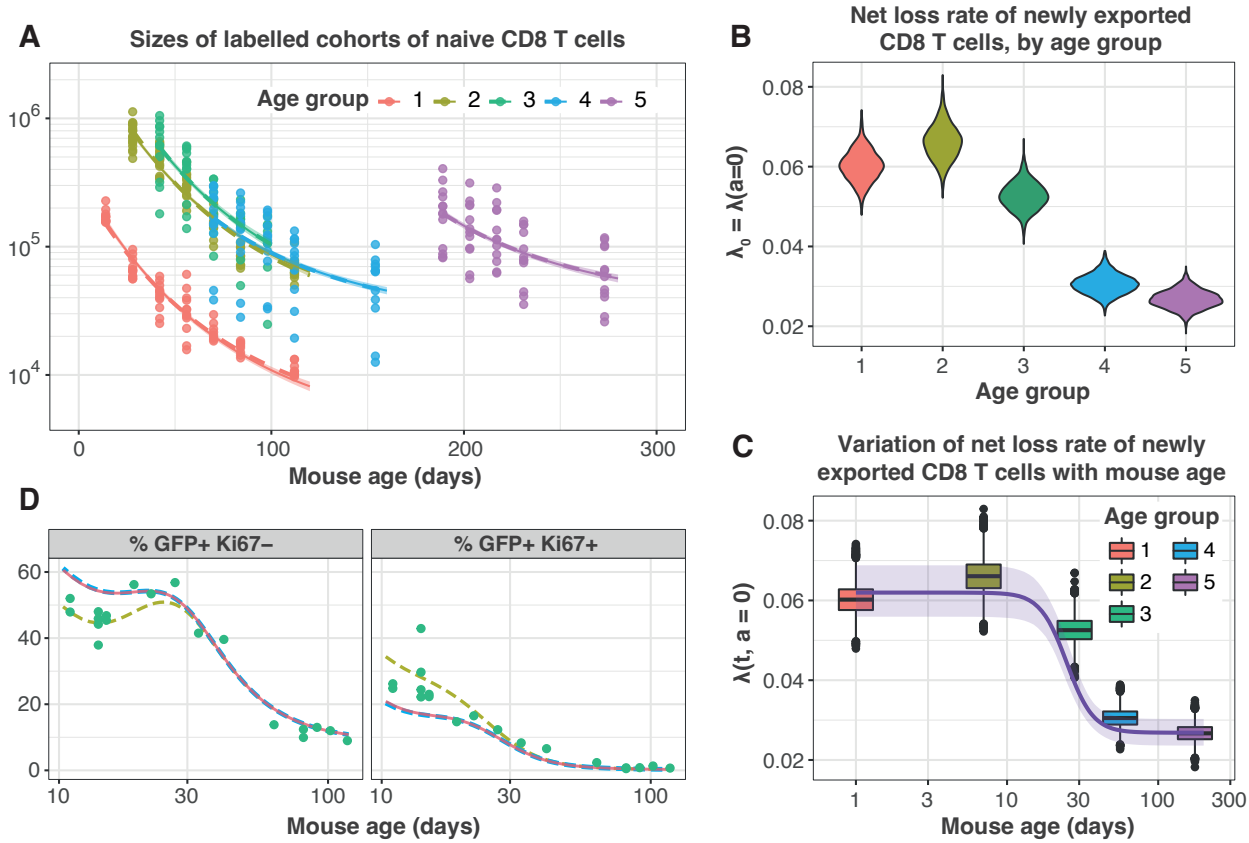

Figure 5: Tracking the persistence of cohorts of naive CD8 T cells *in vivo* – an analysis of data from Reynaldi *et al.*[23]. **(A)** Fitting the age-dependent loss model to numbers of time-stamped naive CD8 T cells in CD4-Cre^{ERT2} reporter mice treated with tamoxifen at different ages. Dashed lines show the model fits from the hierarchical age-structured model, in which the net loss rate λ varies with cell age (a). Group-level effects were considered for the net loss rate of cells of age 0 *i.e.* RTE ($\lambda_0 = \lambda(a = 0)$). **(B)** Estimates of λ_0 for each age-group of mice. **(C)** Fitting an empirical description of the variation in λ_0 with mouse age (see Supporting Information, eqn. S28). **(D)** Predicting the kinetics of the proportions of GFP$^+$ Ki67$^-$ and GFP$^+$ Ki67$^+$ CD8 T cells using the age-dependent loss model, including neonatal age effects in the loss rate (green dashed line) or in the division rate (blue dashed line). The red line (partly concealed by the blue dashed line) shows the original model predictions, with no host age effects.

chical age-structured model. This 'universal' model, in which the loss rate of naive CD8 T cells declines with cell age but begins at higher baseline levels early in life, explained the data equally well (Figure 5A, dashed lines; ΔLOO-IC < 6).

This analysis shows that the baseline net loss rate of CD8 RTE declines from the age of ~3 weeks and stabilises at a level approximately 50% lower by age 9 weeks (Figure 5C). Therefore, newly exported naive CD8 T cells in neonates are either lost at higher rates than in adults, or divide more slowly. Only the former is consistent with our inference from the Rag/Ki67 dual reporter mice. Indeed, we confirmed that simulating the age-dependent loss model from birth with a lower baseline division rate in neonates than in adults failed to improve the description of the early trajectories of the frequencies of GFP$^+$ Ki67$^-$ and GFP$^+$ Ki67$^+$ naive CD8 T cells (Figure 5D, blue dashed line), whereas increasing the baseline loss rate in neonates according to our function we derived from the data in Reynaldi *et al.* captured these dynamics well (Figure 5D, green dashed line).

In summary, we find that naive CD8 T cells rarely divide, increase their capacity to survive with cell age, and those generated within the first few weeks of life are lost at a higher baseline rate than those in adults.

Discussion

Our previous analyses suggest naive T cells operate autonomously and compensate cell-intrinsically for the gradual decline in thymic output with age by increasing their ability to persist with time since they leave the thymus in both adult mice[22] and in humans[24]. This process can also explain the preservation of naive T cells in thymectomized mice[2] without invoking any density-dependent mechanisms[22].

In this study, we take one step further and show through the modeling of a range of datasets that naive T cell adaptation in mice manifests primarily through a progressive decrease in their loss rate, and that they divide very rarely if at all, with mean interdivision times of at least 14 months. This means that throughout the mouse lifetime, newly made CD4 and CD8 RTE are lost at faster rates than their mature counterparts, predicting the preferential retention and accumulation of clones exported early in life. The lack of peripheral expansion combined with high levels of thymic export throughout life implies that the majority of the naive T cell repertoire is made up of long-lived and unique T cell receptor (TCR) clones. This interpretation is consistent with studies showing enormous diversity (low clonality) within naive TCR repertoires[31–34] and supports the idea that any hierarchy within in it is shaped by the generational frequencies of individual clones in the thymus rather than by peripheral expansions[35]. It has been proposed that proliferation driven by low cell densities early in life is responsible for expanded T cell clones in humans[36]. However, it is possible that these are driven instead by early antigen exposure. We also found that levels of intrathymic proliferation are remarkably high in young mice, with close to 100% of late-stage thymocytes expressing Ki67 in neonates, declining to approximately 20% over the first 3 months of life (Figure S2). Presumably this proliferation occurs after TCR rearrangement is complete, implying that clones generated in neonates may be substantially larger on average than those exported from adult thymi. It is possible that a similar mechanism operates in humans.

We showed agreement with the trends demonstrated by Houston *et al.*[10], in which mature naive (MN) CD4 and CD8 T cell persist longer than RTE. However our simulations of their co-transfer experiment yielded somewhat lower levels of enrichment of MN cells (Figure 3A). Aside from differences in experimental conditions, this mismatch may derive in part from differences in the age-distributions of MN T cells in our simulations and their experiments. We considered cells of ages >21 days in 17 week old mice to be MN T cells. Houston *et al.* sourced MN T cells either from mice aged 14 weeks or older that were thymectomized at least 3 weeks prior to transfer, or from >12 week old Rag^{GFP} reporter mice. It is likely that their transferred MN cells were therefore enriched with older cells, giving a more pronounced disparity with RTE in survival.

A study by van Hoeven *et al.*[21] demonstrated that CD4 RTE are lost more rapidly than MN CD4 T cells[21]. They estimated that the loss rate of CD4 RTE was 0.063 day^{-1}, translating to a residence time of 15 days (95%CI: 9-26), and is roughly 4 times shorter than the 66 day (52–90) residence time of MN CD4 T cells. Our results agree closely. We estimate that CD4 RTE (cells of age 0) have an expected residence time of 22 (18-28) days, doubling approximately every 3 months, such that in a 12 week old mouse, the mean residence time of MN CD4 T cells aged 21 days or greater is $\sim$60 days. In contrast, van Hoeven *et al.* concluded that naive CD8 T cells are a kinetically homogeneous population with a mean residence time of 76 (42–83) days. With our favoured age-dependent loss model, we estimate that CD8 RTE initially have an expected residence time of 40 (18–28) days, doubling every $\sim$5 months. However, our predicted average residence time of MN CD8 T cells (aged >21 days) in a 12 week old mouse was approximately 76 days, which agrees with their estimate. We included a similar RTE/MN model in our analysis (illustrated in Figure S3) and found that for CD8 T cells it received statistical support comparable to that of a neutral model of constant division and loss rate, in line with their analysis. Therefore, our different conclusions may stem in part from the specification of our models. It would be instructive to analyse the data from their thymic transplantation and heavy water

labelling studies with the age-structured models we consider here.

The pioneering studies by Berzins *et al.* showed substantial and proportional increases in T cell numbers in mice transplanted with 2, 6 and 9 thymic lobes[19,20]. They concluded that this increase corresponds to the accumulation of RTE exported in the previous 3 weeks. In absence of any homeostatic regulation, the increase in the sizes of the naive CD4 and CD8 pools under hyperthymic conditions is determined by the change in thymic output and by RTE lifespans. Our estimates of these lifespans ($\sim$22 and $\sim$40 days for CD4 and CD8 respectively) are in line with their estimate of 3 weeks[20]. Indeed, simulating the transplantation of 6 thymic lobes using the age-dependent loss models and parameters derived from busulfan chimeric mice recapitulates their observations (Figure S4).

Our analyses indicate that naive CD4 and CD8 T cells divide very rarely in both neonates and adults. However, more than 60% of naive CD4 and CD8 T cells express Ki67 early in life, declining to 2-3% by 3 months of age (Figure 4). This observation suggests that naive T cells undergo more frequent divisions in neonates than in adults, in line with suggestions that neonatal mice are lymphopenic[12,37]. We argue that instead this expression is a shadow of intrathymic division, a conclusion that becomes more apparent in the high degree of correlation between the frequencies of Ki67-positive cells among SP thymocytes and peripheral naive T cells (Figure S5).

Reynaldi *et al.* used a novel fate-mapping system to demonstrate that the net loss rate of naive CD8 T cells (loss minus self-renewal) declines with their age and is higher for CD8 RTE in neonates than in adults[23]. Our addition to this narrative was to reanalyse their data with a more mechanistic modeling approach to isolate the effects of division and loss, and to calculate a functional form for the dependence of the CD8 RTE loss rate on mouse age. In conjunction with our analysis of data from Rag/Ki67 dual reporter mice we inferred that the baseline loss rate naive CD8 T cells immediately following release from the thymus is higher in neonates than in adults, while the rate of division is constant, independent of host and cell age, and close to zero throughout the mouse lifespan. The higher loss rate of CD8 RTE in neonates may derive from high rates of differentiation into memory phenotype cells rather than impaired survival. This idea is consistent with the rapid accumulation of virtual memory CD8 T cells in the periphery during the postnatal period[15]. We found no evidence for a similar process among naive CD4 T cells. We recently showed that increasing the exposure to environmental antigens boosts the generation of memory CD4 T cell subsets early in life[14], but it may be that the flux from naive into the specific pathogen-free mice we consider here is too low to be a detectable drain on naive CD4 T cell numbers.

Our models do not address the mechanisms that prolong the residence time of T cells in the naive pool. Modulation of sensitivity to IL-7 and signaling via Bcl-2 associated molecules have been implicated in increasing naive T cell longevity[10,38] and is consistent with the outcome of co-transfer experiments. However it is possible that increased persistence derives additionally from a progressive or selective decrease in naive T cells' ability to be triggered into effector or memory subsets. Studying the dynamics of naive T cells in busulfan chimeras generated using BM from TCR transgenic donor mice may help us untangle the contributions of survival and differentiation to the increase in their residence time with their age.

Our analyses synthesise the results of multiple studies into a single framework, and conclude that although thymic activity varies with mouse age, the rules of behaviour of naive CD4 T cells – slowly extending their mean residence time – remain constant across the mouse lifespan. Naive CD8 T cells exhibit a similar adaptation, but also experience elevated loss or differentiation early in life. We also find no evidence for feedback regulation of naive T cell numbers through modification of division or loss rates. Comparable studies of naive T cell dynamics are far more challenging, if not impossible, in humans, although adaptation effects in survival also appear to be manifest[24]. However, in contrast to mice, naive T cells in humans appear to undergo self-renewal at appreciable levels and the relative contribution of the thymus to the production of new cells

becomes very small in adulthood[2]. Nevertheless we speculate that the idea of naive T cell homeostasis in the sense of compensatory or quorum sensing behaviour may well be just a theoretical concept in both mice and humans, at least close to healthy conditions. Indeed recently it has been shown that recovery from autologous haematopoietic stem cell transplant results in persistent perturbations of T cell dynamics[39], arguing against the idea of a naturally regulated set-point. Perhaps a better model is one in which the thymus drives the generation of the bulk of the T cell pool in the early life and thereafter naive T cell repertoires 'coast' out into old age in a cell-autonomous manner.

References

1. Scollay RG, Butcher EC, Weissman IL. Thymus cell migration. Quantitative aspects of cellular traffic from the thymus to the periphery in mice. Eur J Immunol. 1980;10(3):210–8.
2. den Braber I, Mugwagwa T, Vrisekoop N, Westera L, Mögling R, de Boer AB, et al. Maintenance of peripheral naive T cells is sustained by thymus output in mice but not humans. Immunity. 2012;36(2):288–97.
3. Hogan T, Gossel G, Yates AJ, Seddon B. Temporal fate mapping reveals age-linked heterogeneity in naive T lymphocytes in mice. Proc Natl Acad Sci U S A. 2015;112(50):E6917–26. doi:10.1073/pnas.1517246112.
4. Egerton M, Scollay R, Shortman K. Kinetics of mature T-cell development in the thymus. Proc Natl Acad Sci U S A. 1990;87(7):2579–82.
5. Graziano M, St-Pierre Y, Beauchemin C, Desrosiers M, Potworowski EF. The fate of thymocytes labeled in vivo with CFSE. Exp Cell Res. 1998;240(1):75–85.
6. Dzierzak E, Daly B, Fraser P, Larsson L, Müller A. Thy-1 tk transgenic mice with a conditional lymphocyte deficiency. Int Immunol. 1993;5(8):975–84. doi:10.1093/intimm/5.8.975.
7. Rocha B, Dautigny N, Pereira P. Peripheral T lymphocytes: expansion potential and homeostatic regulation of pool sizes and CD4/CD8 ratios in vivo. Eur J Immunol. 1989;19(5):905–11. doi:10.1002/eji.1830190518.
8. Almeida AR, Borghans JA, Freitas AA. T cell homeostasis: thymus regeneration and peripheral T cell restoration in mice with a reduced fraction of competent precursors. J Exp Med. 2001;194(5):591–9.
9. Yates A, Saini M, Mathiot A, Seddon B. Mathematical modeling reveals the biological program regulating lymphopenia-induced proliferation. J Immunol. 2008;180(3):1414–1422.
10. Houston EG Jr, Higdon LE, Fink PJ. Recent thymic emigrants are preferentially incorporated only into the depleted T-cell pool. Proc Natl Acad Sci U S A. 2011;108(13):5366–71.
11. Hogan T, Shuvaev A, Commenges D, Yates A, Callard R, Thiebaut R, et al. Clonally diverse T cell homeostasis is maintained by a common program of cell-cycle control. J Immunol. 2013;190(8):3985–93.
12. Min B, McHugh R, Sempowski GD, Mackall C, Foucras G, Paul WE. Neonates support lymphopenia-induced proliferation. Immunity. 2003;18(1):131–40.
13. Modigliani Y, Coutinho G, Burlen-Defranoux O, Coutinho A, Bandeira A. Differential contribution of thymic outputs and peripheral expansion in the development of peripheral T cell pools. Eur J Immunol. 1994;24(5):1223–7. doi:10.1002/eji.1830240533.
14. Hogan T, Nowicka M, Cownden D, Pearson CF, Yates AJ, Seddon B. Differential impact of self and environmental antigens on the ontogeny and maintenance of CD4+ T cell memory. eLife. 2019;8. doi:10.7554/eLife.48901.
15. Akue AD, Lee JY, Jameson SC. Derivation and maintenance of virtual memory CD8 T cells. J Immunol. 2012;188(6):2516–23.
16. Houston EG Jr, Nechanitzky R, Fink PJ. Cutting edge: Contact with secondary lymphoid organs drives postthymic T cell maturation. J Immunol. 2008;181(8):5213–7.
17. Adkins B. T-cell function in newborn mice and humans. Immunol Today. 1999;20(7):330–5.

18. Wang J, Wissink EM, Watson NB, Smith NL, Grimson A, Rudd BD. Fetal and adult progenitors give rise to unique populations of CD8+ T cells. Blood. 2016;128(26):3073–3082.

19. Berzins SP, Boyd RL, Miller JF. The role of the thymus and recent thymic migrants in the maintenance of the adult peripheral lymphocyte pool. J Exp Med. 1998;187(11):1839–48.

20. Berzins SP, Godfrey DI, Miller JF, Boyd RL. A central role for thymic emigrants in peripheral T cell homeostasis. Proc Natl Acad Sci U S A. 1999;96(17):9787–91.

21. van Hoeven V, Drylewicz J, Westera L, den Braber I, Mugwagwa T, Tesselaar K, et al. Dynamics of Recent Thymic Emigrants in Young Adult Mice. Front Immunol. 2017;8.

22. Rane S, Hogan T, Seddon B, Yates AJ. Age is not just a number: Naive T cells increase their ability to persist in the circulation over time. PLoS Biol. 2018;16(4):e2003949.

23. Reynaldi A, Smith NL, Schlub TE, Tabilas C, Venturi V, Rudd BD, et al. Fate mapping reveals the age structure of the peripheral T cell compartment. Proc Natl Acad Sci U S A. 2019;.

24. Mold JE, Réu P, Olin A, Bernard S, Michaëlsson J, Rane S, et al. Cell generation dynamics underlying naive T-cell homeostasis in adult humans. PLOS Biology. 2019;17(10):e3000383. doi:10.1371/journal.pbio.3000383.

25. Hogan T, Yates A, Seddon B. Generation of Busulfan chimeric mice for the analysis of T cell population dynamics. Bio-protocol. 2017;4(24).

26. Gossel G, Hogan T, Cownden D, Seddon B, Yates AJ. Memory CD4 T cell subsets are kinetically heterogeneous and replenished from naive T cells at high levels. eLife. 2017;6. doi:10.7554/eLife.23013.

27. Verheijen M, Rane S, Pearson C, Yates AJ, Seddon B. Fate Mapping Quantifies the Dynamics of B Cell Development and Activation throughout Life. Cell Reports. 2020;33(7):108376. doi:10.1016/j.celrep.2020.108376.

28. Miller I, Min M, Yang C, Tian C, Gookin S, Carter D, et al. Ki67 is a graded rather than a binary marker of proliferation versus quiescence. Cell Rep. 2018;24(5):1105–1112.e5.

29. Vehtari A, Gelman A, Gabry J. Efficient implementation of leave-one-out cross-validation and WAIC for evaluating fitted Bayesian models. Statistics and Computing, arXiv:150704544. 2015;27:1413–1432.

30. Vehtari A, Gelman A, Gabry J. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. Stat Comput. 2017;27:1413–1432.

31. Gonçalves P, Ferrarini M, Molina-Paris C, Lythe G, Vasseur F, Lim A, et al. A new mechanism shapes the naïve CD8 T cell repertoire: the selection for full diversity. Molecular Immunology. 2017;85:66–80. doi:10.1016/j.molimm.2017.01.026.

32. Qi Q, Liu Y, Cheng Y, Glanville J, Zhang D, Lee JY, et al. Diversity and clonal selection in the human T-cell repertoire. PNAS; Proceedings of the National Academy of Sciences. 2014;111(36):13139–13144.

33. de Greef PC, Oakes T, Gerritsen B, Ismail M, Heather JM, Hermsen R, et al. The naive T-cell receptor repertoire has an extremely broad distribution of clone sizes. eLife. 2020;9. doi:10.7554/elife.49900.

34. Mora T, Walczak AM. How many different clonotypes do immune repertoires contain? Current Opinion in Systems Biology. 2019;18:104–110.

35. Quigley M, Greenaway H, Venturi V, Lindsay R, Quinn K, Seder R, et al. Convergent recombination shapes the clonotypic landscape of the naïve T-cell repertoire. PNAS. 2010;107(45):19414–19419.

36. Gaimann MU, Nguyen M, Desponds J, Mayer A. Early life imprints the hierarchy of T cell clone sizes. Elife. 2020;9. doi:10.7554/eLife.61639.

37. Le Campion A, Bourgeois C, Lambolez F, Martin B, Léaument S, Dautigny N, et al. Naive T cells proliferate strongly in neonatal mice in response to self-peptide/self-MHC complexes. PNAS; Proceedings of the National Academy of Sciences. 2002;99(7):4538–4543.

38. Tsukamoto H, Huston GE, Dibble J, Duso DK, Swain SL. Bim dictates naive CD4 T cell lifespan and the development of age-associated functional defects. The Journal of Immunology. 2010;185(8):4535–4544.

39. Baliu-Piqué M, van Hoeven V, Drylewicz J, van der Wagen LE, Janssen A, Otto SA, et al. Cell-density independent increased lymphocyte production and loss rates post-autologous HSCT. Elife. 2021;10. doi:10.7554/eLife.59775.

Supporting Information

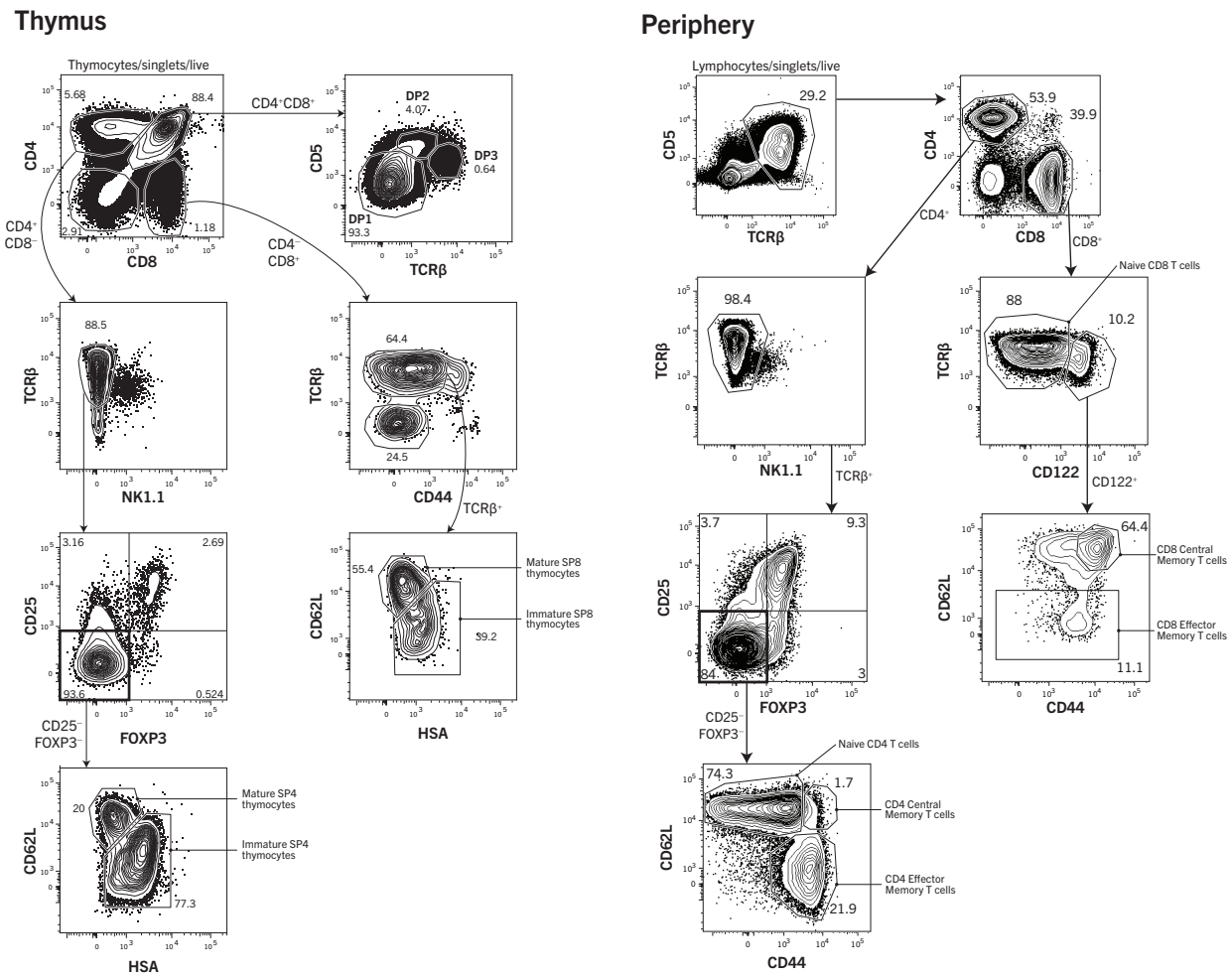

Figure S1: Gating strategies for thymocyte and peripheral naive T cell subsets.

Text S1 Modelling age-structured cell dynamics with continuous Ki67 expression

We aimed to model the size and age structure of the naive T cell pool, and also to track cells' expression of Ki67, which we assume reaches a level $k = 1$ (arbitrarily normalised) immediately after cell division and decays exponentially at rate β. We assume that the rates at which cells divide (ρ) and are lost (δ) can be functions of host age t and also of cell age a, which we define to be the time since that cell, or its ancestor, was exported from the thymus. Here we detail the construction of a structured population PDE model describing the time evolution of the cell population density $u(a, k, t)$.

Initial conditions. We assume that at some host age t_0, the population is of size N_0 and has a cell-age distribution $\gamma(a)$, $0 \leq a \leq t_0$, where $\int_0^{t_0} \gamma(a)da = 1$ (we assume all cells are of age zero when they leave the thymus); and these cells have a distribution of levels of Ki67 expression $\psi(k)$, where $\int_0^1 \psi(k)dk = 1$ and $\psi(k) = 0$ outside of $k \in (0, 1]$. Here for simplicity we are assuming no relation between Ki67 expression and cell age within the cells present at $t = t_0$, but one could easily extend this framework with a more general initial joint distribution $P(a, k)$. At times $t \geq t_0$, we assume that cells of age zero enter the naive pool from the thymus at rate $\theta(t)$ and Ki67 distribution $\phi(k, t)$, where $\int_0^1 \phi(k, t)dk = 1, \ \forall t$.

Breaking the solution into cohorts of cells. Our approach is to track separately the fates of cells that were present at t_0, who will all have age $a > t - t_0$ at some later time t; and the fates of those that were exported from the thymus at time t_0 or later, which will all have age $a < t - t_0$. We then add these to get the full population density $u(a, k, t)$. The master PDE for both populations combined is

$$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial a} - \beta k \frac{\partial u}{\partial k} = -(\rho(a, t) + \delta(a, t))u(a, k, t) \tag{S1}$$

with boundary conditions

$$u(a, k = 1, t) = 2\rho(a, t) \int_{k=0}^{1} u(a, k, t)dk \qquad \text{(Cell division)} \tag{S2}$$

$$u(a, k, t = t_0) = N_0\gamma(a)\psi(k) \qquad \text{(Population present at host age } t_0\text{)} \tag{S3}$$

$$u(a = 0, k, t) = \theta(t)\phi(k, t) \qquad \text{(Influx of new cells from thymus)} \tag{S4}$$

The first condition above derives from cell division; at any host age t, cells of age a at time t divide at rate $\rho(a, t)$ generating two cells of age a with $k = 1$.

Initial cohort

Non-divided cells. First consider those cells present at $t = t_0$ that have yet to divide; this population falls in size at *per capita* rate $\delta(a, t) + \rho(a, t)$, and follows eqn. S1 with the single boundary condition $u(a, k, t = t_0) = N_0\gamma(a)\psi(k)$. We solve this using the method of characteristics by identifying a variable s such that

$$\frac{\mathrm{d}u}{\mathrm{d}s} = -\Upsilon(a, t)u(a, k, t) = \frac{\mathrm{d}t}{\mathrm{d}s}\frac{\partial u}{\partial t} + \frac{\mathrm{d}a}{\mathrm{d}s}\frac{\partial u}{\partial a} + \frac{\mathrm{d}k}{\mathrm{d}s}\frac{\partial u}{\partial k} \tag{S5}$$

where for brevity we define $\Upsilon(a, t) = \rho(a, t) + \delta(a, t)$. Equating terms with eqn. S1 gives

$$dt/ds = 1 \implies s = t - t_0, \quad a = a_0 + t - t_0, \quad dk/ds = -\beta k \implies k = k_0 e^{-\beta(t-t_0)}. \tag{S6}$$

Along the characteristic that starts at (a_0, k_0, t_0) (illustrated in Figure S1.1A), the population density evolves as

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}u(a_0,k_0,t) &= -\Upsilon(a,t)u(a_0,k_0,t) && \text{(S7)}\\ &= -\Upsilon(a,a-a_0+t_0)u(a_0,k_0,t) && \text{(S8)}\\ \implies u(a_0,k_0,t) &= u(a_0,k_0,t_0)\exp\left(-\int_{x=a_0}^{a_0+t-t_0}\Upsilon(x,x-a_0+t_0)dx\right). && \text{(S9)} \end{aligned}$$

But we know from eqn. S6 that $k_0 = ke^{\beta(t-t_0)}$, so

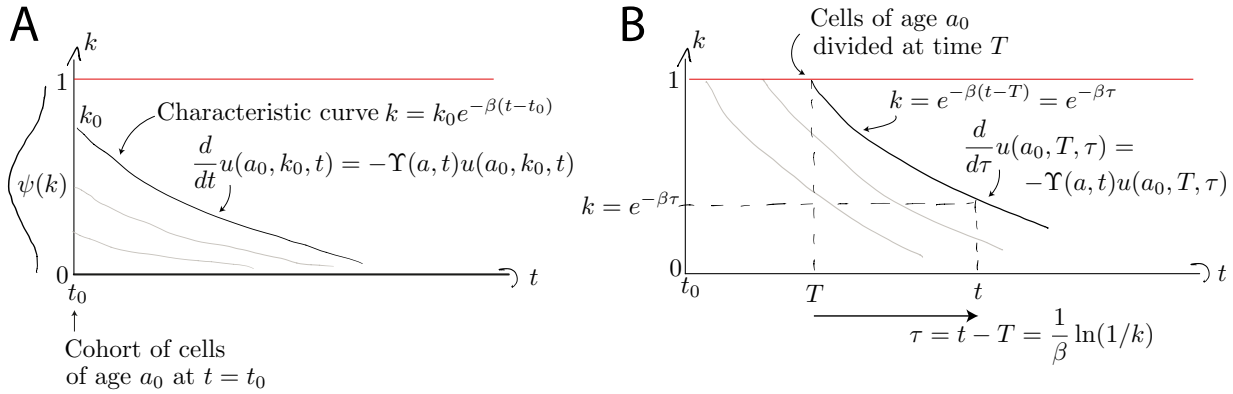

Figure S1.1: Characteristics for the population of age a_0 present at $t = t_0$ (panel A) and the population of age a_0 who divided at time T (panel B).

$$u(a_0, ke^{\beta(t-t_0)}, t) = u(a_0, ke^{\beta(t-t_0)}, t_0)\exp\left(-\int_{x=a_0}^{a_0+t-t_0}\Upsilon(x,x-a_0+t_0)dx\right). \qquad \text{(S10)}$$

The final step here is that $u(.)$ is a population density and so must be transformed with a Jacobian if we want to express it as a density over k rather than over $ke^{\beta(t-t_0)}$. If $y = ke^{\beta(t-t_0)}$, then

$$u(a,k,t) = u(a_0,y,t)|dy/dk| = u(a_0, ke^{\beta(t-t_0)}, t)e^{\beta(t-t_0)}, \qquad \text{(S11)}$$

giving

$$\begin{aligned} u_{i,n}(a,k,t) &= u(a_0, ke^{\beta(t-t_0)}, t_0)\exp\left(\beta(t-t_0) - \int_{x=a_0}^{a_0+t-t_0}\Upsilon(x,x-a_0+t_0)dx\right)\\ &= N_0\gamma(a-t+t_0)\psi(ke^{\beta(t-t_0)})\exp\left(\beta(t-t_0) - \int_{x=a+t_0-t}^{a}\Upsilon(x,x-a+t)\,dx\right),\\ &\quad \text{for } t \geq t_0, a \geq t-t_0,\ k \leq e^{-\beta(t-t_0)}. \qquad \text{(S12)} \end{aligned}$$

where the subscript (i,n) denotes 'initial and non-divided'. Equation S12 holds for $k \leq e^{-\beta(t-t_0)}$ because none of these cells have divided since time t_0 and their Ki67 expression is decaying.

Divided cells. Figure S1.1B illustrates the evolution of cells from the initial cohort who do subsequently divide, each time resetting their Ki67 expression to $k = 1$. To follow these cohorts we solve eqn. S1 with the

boundary condition describing the division of cells of age a at host age t;

$$u(a, k = 1, t) = 2\rho(a, t) \int_{k=0}^{1} u(a, k, t)\, dk. \tag{S13}$$

Characteristic curves are again of the general form $k = k_0 e^{-\beta t}$, but we parameterise them differently to those described above, since these originate at $k = 1$ and not $t = t_0$ (Figure S1.1B). Cells currently of age a at time t with Ki67 expression k must have divided a time $\tau = -(1/\beta)\ln(k)$ in the past. A convenient parameterisation is therefore

$$s = \tau, \quad a = a_0 + \tau, \quad k = e^{-\beta\tau}, \tag{S14}$$

where the cells of age a divided at time $T = t - \tau$ when they were of age $a_0 = a - \tau$. Along these curves,

$$u(a_0, T, \tau = (1/\beta)\ln(1/k)) = 2\rho(a_0, T) \int_{x=0}^{1} u_i(a_0, x, T)\, dx \times \exp\left(- \int_{x=a_0}^{a_0+\tau} \Upsilon(x, x - a_0 + T)\, dx \right) \tag{S15}$$

where the exponential term represents the proportion of cells on this characteristic curve that divide again or die. It integrates a cell's experience from age a_0 at host age T, to age a at host age $t = T + \tau$. Here u_i denotes the entire initial cohort, divided or undivided, integrated over all levels of Ki67 expression, whose cells of age a_0 fed the divided population at time T.

To convert this to a density $u_{i,d}(a, k, t)$, we use the Jacobian $d\tau/dk = 1/(\beta k)$. This gives

$$\begin{aligned} u_{i,d}(a, k, t) &= \frac{2\rho(a - \tau, t - \tau)}{\beta k} \int_{x=0}^{1} u_i(a - \tau, x, t - \tau)\, dx \times \exp\left(- \int_{x=a-\tau}^{a} \Upsilon(x, x - a + t)\, dx \right) \\ &= \frac{2\rho(a - \tau, t - \tau)}{\beta k} U_i(a - \tau, t - \tau) \exp\left(- \int_{x=a-\tau}^{a} \Upsilon(x, x - a + t)\, dx \right), \\ &\text{valid for } \tau = (1/\beta)\ln(1/k),\ a \geq \tau,\ t - t_0 \geq \tau,\ k > e^{-\beta(t-t_0)}, \end{aligned} \tag{S16}$$

where $U_i(a, t)$ is the population density of the initial cohort (divided or undivided) at age a and time t, integrated over all k, which we will return to below.

Cells that enter after $t = t_0$

A similar set of calculations applies for cells that subsequently enter the pool, at which point we define them to be age zero. Again we partition these cells into those that don't divide and those that do. For the former, consider those that are of age $a \leq t - t_0$ at time t; that is, cells that entered later than t_0. We denote this population $u_{\theta,n}(a, k, t)$ where $k \leq e^{-\beta a}$. These cells are what remains of the cohort exported at time $t - a$, at age $a = 0$, of size $\theta(t - a)$ and with Ki67 distribution $\phi(ke^{\beta a}, t - a)$. These are then lost to death or division. By analogy with equation S12, the Jacobian is $e^{\beta a}$ and these non-divided cells evolve as

$$\begin{aligned} u_{\theta,n}(a, k, t) &= \theta(t - a)\phi(ke^{\beta a}, t - a) \exp\left(\beta a - \int_{x=0}^{a} \Upsilon(x, x - a + t)\, dx \right), \\ &\text{for } t > t_0, a < (t - t_0),\ k \leq e^{-\beta a}. \end{aligned} \tag{S17}$$

Here, cells are born with age zero and so the characteristics are $k = k_0 e^{-\beta a}$. Similarly for divided cells; by analogy with equation S16,

$$u_{\theta,d}(a,k,t) = \frac{2\rho(a-\tau,t-\tau)}{\beta k} U_\theta(a-\tau,t-\tau) \times \exp\left(-\int_{x=a-\tau}^{a} \Upsilon(x,x-a+t)\,dx\right) \quad \text{(S18)}$$
$$\text{for } \tau = (1/\beta)\ln(1/k),\ t \geq t_0, a \leq t-t_0,\ 0 \leq \tau \leq a.$$

Here $U_\theta(a,t)$ is the population density of cells of of age a at time t, who entered the pool a time $\tau = t-a$ ago and may have divided or not.

Solving the age-structured PDE only

To complete these solutions we need the population densities $U_i(a,t)$ and $U_\theta(a,t)$, ignoring Ki67 expression. These are straightforward;

$$U_i(a,t) = N_0\gamma(a-t+t_0)\exp\left(-\int_{x=a-t+t_0}^{a} \lambda(x,x-a+t)\,dx\right), \quad \text{for } a \geq t-t_0, \quad \text{(S19)}$$

$$U_\theta(a,t) = \theta(t-a)\exp\left(-\int_{x=0}^{a} \lambda(x,x-a+t)\,dx\right), \quad \text{for } a \leq t-t_0. \quad \text{(S20)}$$

where for brevity $\lambda(a,t)$ is the net loss rate of cells of age a at time t, which is $\delta(a,t)-\rho(a,t)$. The integral in equation S19 follows a cell whose age runs from $a-(t-t_0)$ to a, during which host age runs from $t-a$ to t. The integral in equation S20 follows a cell whose age runs from 0 to a, between host ages of $t-a$ to t. The two solutions join at $a = t-t_0$; the influx at time t_0 must be the density of cells of age zero in the initial cohort; $\theta(t_0) = N_0\gamma(0)$.

Therefore, to obtain the total solution $u(a,k,t) = u_{i,n} + u_{i,d} + u_{\theta,n} + u_{\theta,d}$, we add equations S12, S16, S17, S18, using the solutions for the age-structured model given in equations S19 and S20.

We can connect this solution to gated flow cytometry data by partitioning the popuation into high and low Ki67 expression. We might define Ki67 positive cells to be those which divided no more than a time $1/\beta$ ago, which corresponds to a cut-off of $k = 1/e$. Therefore, the numbers of Ki67 positive and negative cells at time t are

$$N^+(t) = \int_{k=1/e}^{1}\int_{a=0}^{a_{max}} u(a,k,t)\,dk\,da, \qquad N^-(t) = \int_{k=0}^{1/e}\int_{a=0}^{a_{max}} u(a,k,t)\,dk\,da. \quad \text{(S21)}$$

Text S2 Constructing empirical descriptions of the dynamics of mature SP thymocytes over the mouse lifespan

For the analysis of busulfan chimera data, we used phenomenological functions to explain the time-varying rate of export of new naive cells from the thymus over time, assuming that it is proportional to the numbers of single positive (SP4 and SP8) thymocytes[19]. Using data from wild-type mice bred in the same facility as the busulfan chimeras (Figure S2), we fitted the following to the measured numbers of SP4 and SP8 cells;

$$S(t) = S(0) + (A * t^n)\,\frac{(1 - t^q)}{B^q + t^q} \tag{S22}$$

The rate of thymic export is then modelled as $\theta(t) = \phi * S(t)$, with the constant ϕ estimated when fitting models to the busulfan chimera data.

We modelled the proportion Ki67$^{\text{hi}}$ fraction within SP populations using the form

$$\epsilon(t) = \epsilon_0 + e^{-\epsilon_f * (t+C)} \tag{S23}$$

These fitted functions are shown in Figure S2.

When modelling data from the busulfan chimeras, we assumed that the total output from the thymus at any time is identical to that in age-matched wild type mice, but is split between donor and host cells according to the chimerism χ at the DP1 stage of thymic development:

$$\begin{aligned} \theta_{\text{donor}}(t) &= \chi\theta(t) \\ \theta_{\text{host}}(t) &= (1 - \chi)\theta(t) \end{aligned} \tag{S24}$$

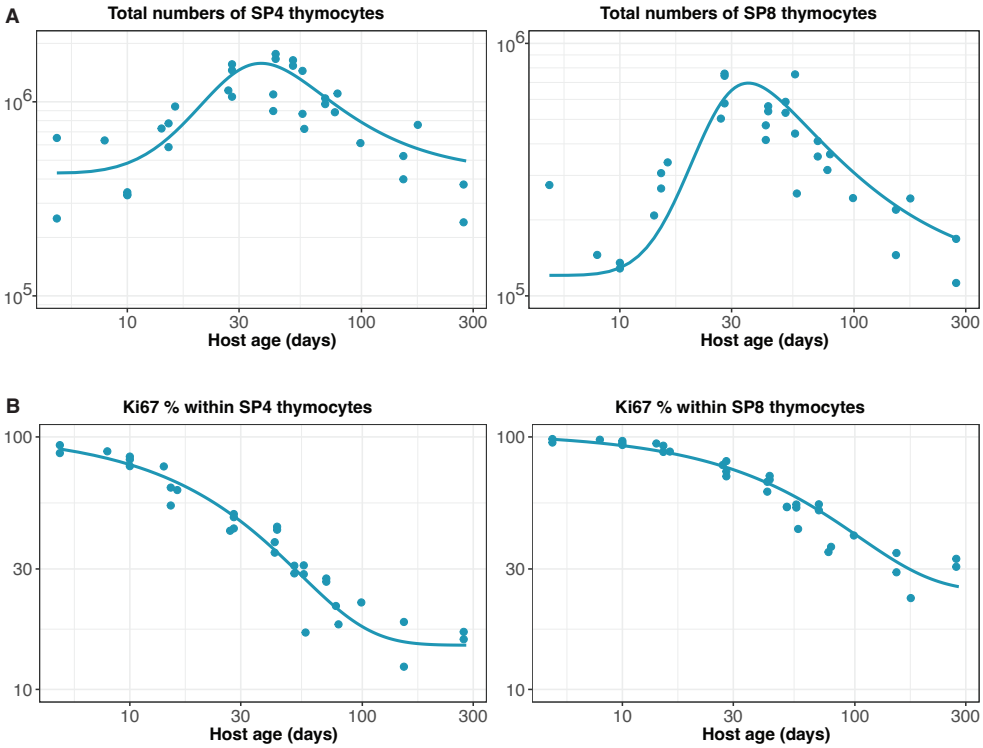

Figure S2: Empirical descriptions of the dynamics of the numbers and Ki67 expression of late-stage thymocytes. These curves (defined in Text S2) were used as inputs to models of the data from adult busulfan chimeric mice.

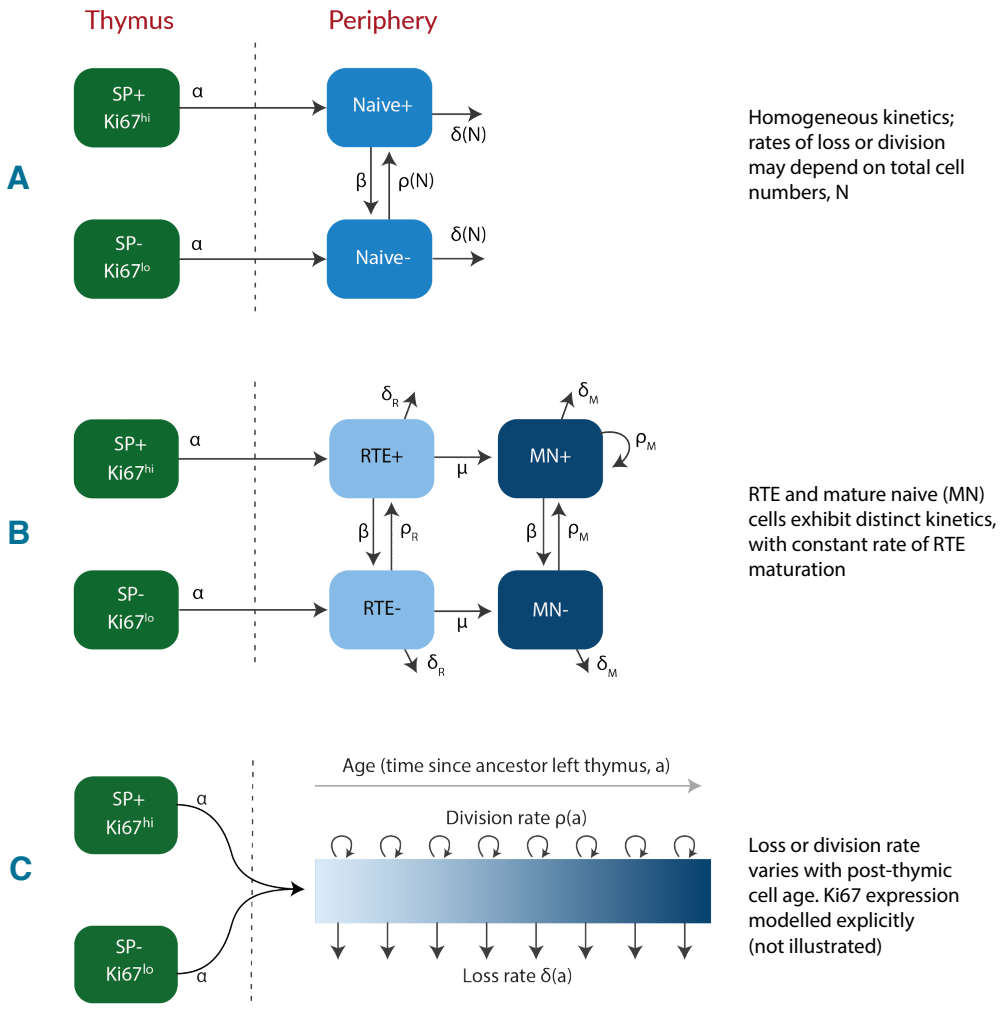

Figure S3: Candidate models of naive T cell dynamics. We considered three classes of model; **(A)** Kinetic homogeneity, in which all cells are lost at the same rate and divide at the same rate. In the simplest 'neutral' case these rates are constant. We also considered extensions in which loss or division rates were allowed to vary with total cell numbers (density-dependent models). **(B)** RTE and mature naive T cells exhibit distinct kinetics, with a constant rate of maturation. **(C)** Loss or division rates vary with post-thymic cell age. In all models we assume $Ki67^{low}$ and $Ki67^{hi}$ cells are exported from the thymus at rates proportional to the numbers of $Ki67^{low}$ and $Ki67^{hi}$ single positive thymocytes.

Population	Model	Δ LOO-IC	Model weight (%)
Naive CD4	C – Loss varying with cell age	0	86.3
	C – Division varying with cell age	21	13
	A – Neutral	38	0.5
	B – Distinct dynamics of RTE mature naive T cells	40	0.2
	A – Density dependent loss	48	0.0
	A – Density dependent division (LIP)	58	0.0
Naive CD8	C – Division varying with cell age	0	85.0
	C – Loss varying with cell age	13	9.0
	A – Density dependent division (LIP)	19	4.5
	A – Density dependent loss	25	1.5
	B – Distinct dynamics of RTE mature naive T cells	26	0.0
	A – Neutral	27	0.0

Table S1: Ranking of models describing naive CD4 and CD8 T cell dynamics in adult busulfan chimeric mice. We fitted different instances of the three classes of model (A, B, C) illustrated in Figure S3, to data from adult busulfan chimeric mice. Measures of support relative to the best fitting model are expressed as differences in the leave-one-out information criterion (LOO-IC) [29,30]. Model weights are calculate on the basis that the relative support for two models is exp(-ΔLOO-IC); so, for instance, a difference in LOO-IC of 6 or more implies at least 20-fold lower support for the model with the larger LOO-IC value.

Text S3 Predicting of RTE dynamics in Rag/Ki67 dual reporter mice

In $\text{Rag}^{\text{GFP}}\text{Ki67}^{\text{RFP}}$ reporter mice, RTE are identified based on the transient expression of GFP, which we assume decays with first order kinetics. Our models estimated very low rates of division among naive CD4 and CD8 T cells, such that any dilution of GFP through division is minimal. There is therefore a simple and direct correlation between GFP expression f and cell age a within naive T cells, which we used to predict the fractions of GFP^+ cells (F) within the naive compartment.

Using the favoured age-structured models described in Text S1, the age distribution of the naive T cell pool at mouse age t is given by equations S19 and S20 as $U(a,t) = U_i(a,t) + U_i(a,t)$. The fraction of cells that are GFP^+ is then the following;

$$F = \frac{\int_0^{\bar{a}} U_i(a,t)}{\int_0^t U_i(a,t)}, \tag{S25}$$

where $\bar{a}$ is the (unknown) time required for a $\text{GFP}^+ \rightarrow \text{GFP}^-$. Implicit in this calculation is the assumption that GFP levels are similar in Ki67^{low} and Ki67^{hi} RTE; mature SP cells are indeed very bright for GFP with minimal differences when stratified by Ki67 expression (data not shown). We then used the parameters derived from busulfan chimera data to generate $U(a,t)$, and estimated $\bar{a}$ by fitting equation S25 to the timecourse of the GFP^+ fraction within naive T cells observed in Rag/Ki67 dual reporter mice. The fits are shown in Figure 4B and D, red lines in the leftmost panels. We then generated the predicted timecourses of GFP^+ Ki67^+ and GFP^+ Ki67^+ fractions by multiplying F with the predictions of Ki67^{hi} and Ki67^{low} fractions that we derived by extrapolating the age-dependent loss model back to birth.

Text S4 Hierarchical modelling of naive CD8 T cell timestamping data

The data from Reynaldi *et al.*[23], shown in Figure 5, comprised longitudinal samples drawn from animals in different age groups. To use these data to characterise how cell loss rates vary as a function of both cell and host age, we took a hierarchical modeling approach. We allowed for animal and/or group-level variation in the initial numbers of RFP labelled cells in each animal (N_0) and in their initial loss rate (that is, the instantaneous net loss rate of cells of age zero, just exported from the thymus). We began by modeling the kinetics of labelled cells with the assumption that their net loss rate λ varies with their post-thymic age a as $\lambda(a) = \lambda_0 e^{-\gamma a}$. The population density of cells of age a at mouse age t, $N(a,t)$, then obeys

$$\frac{\partial N}{\partial t} + \frac{\partial N}{\partial a} = -\lambda(a)\ N(t,a), \tag{S26}$$

with the boundary condition $N(T,a) = N_0\,\delta(a)$, where T is the mouse age at the time of treatment and $\delta(.)$ is the Dirac delta function.

$$\begin{aligned} &y_i \sim \mathcal{N}(\mu_i, \sigma) && \text{(Likelihood)} \\ &\mu_i = f(t_i, N_0, \lambda_0, \gamma) && \text{(Model)} \\ &\textbf{Hyper parameters} \\ &N_0 \sim \mathcal{N}(\mu_N, \sigma_N) && \text{(Initial cell numbers)} \\ &\lambda_0 \sim \mathcal{N}(\mu_\lambda, \sigma_\lambda) && \text{(Loss rate at cell age=0)} \end{aligned} \tag{S27}$$

We define priors for $\mu_N, \sigma_N, \mu_\lambda, \sigma_\lambda$ and γ and fitted permutations of the hierarchical age-structured model to the time courses of labelled CD8 T cell numbers (Table S28). The best-fitting model exhibited group-specific values of the initial RTE loss rate λ_0, and variation in the initial numbers of labelled cells (N_0) across mice, likely deriving from variations in the efficiency of tamoxifen-driven labelling.

We then refitted a model with an explicit, empirical description of the variation in λ_0 with mouse age t, giving the following model of the net loss rate of cells of age a at time t, with estimated parameters λ_h, Q, q and γ;

$$\lambda(t,a) = \lambda_0(t)\, e^{-\gamma\, a} = \lambda_h \left(1 + \frac{Q}{1 + (t/q)^5}\right) e^{-\gamma\, a}. \tag{S28}$$

Model	Δ LOO-IC	Weight %
N_0 varying at animal level	316	0.0
N_0 varying at animal level; λ_0 varying at group level	0.0	100
N_0 varying at animal level; λ_0 varying at animal level	73	0.0
N_0 varying at group level; λ_0 varying at animal level	313	0.0

Table S2: Comparing support for hierarchical age-structured models of the data from ref. 23

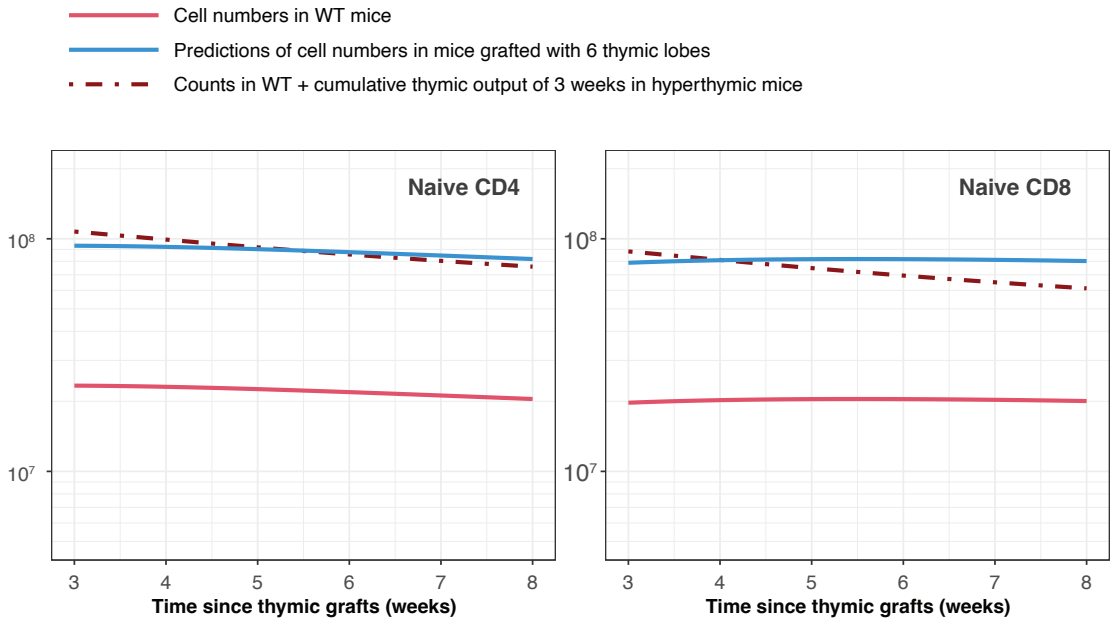

Figure S4: Simulating the outcome of transplanting 6 additional thymi, as described by Berzins *at al.*[20]. The change in numbers of naive CD4 and CD8 T cells is equivalent to 3 weeks of thymic output.

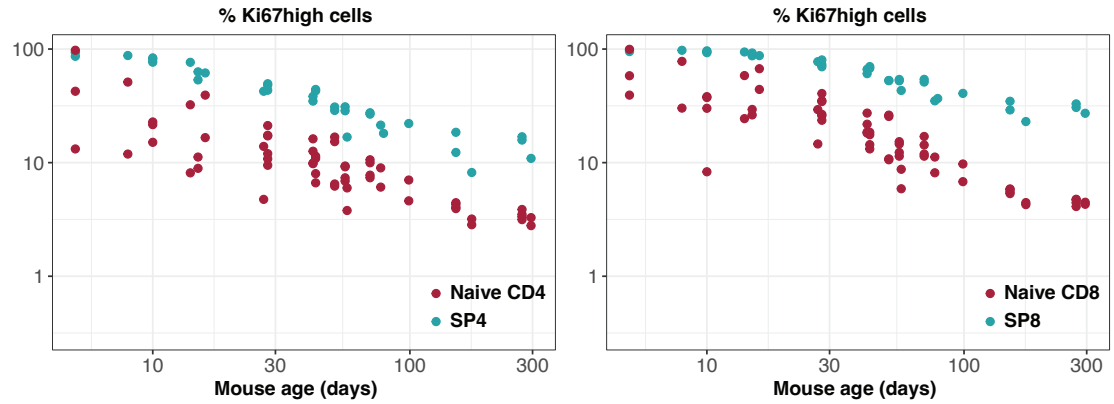

Figure S5: Ki67 levels in late-stage thymocytes and peripheral naive T cells correlate throughout life.