How many TCR clonotypes does a body maintain?

Grant Lythe [a,*], Robin E. Callard [b,c], Rollo L. Hoare [b,c], Carmen Molina-París [a]

[a] *Department of Applied Mathematics, School of Mathematics, University of Leeds, Leeds LS2 9JT, UK*
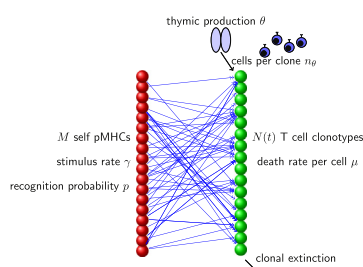[b] *Institute for Child Health, University College London, 30 Guilford Street, London WC1N 1EH, UK*
[c] *Centre for Mathematics and Physics in the Life Sciences and Experimental Biology, University College London, Gower Street, London WC1N 1EH, UK*

HIGHLIGHTS

- The number of T cells of one clonotype is an integer.
- The history of a clonotype starts with release from the thymus, and ends with extinction.
- Competition and cross-reactivity are included in a natural way.
- The average number of cells per clonotype, in a human body, is only of order 10.

GRAPHICAL ABSTRACT

ARTICLE INFO

ABSTRACT

We consider the lifetime of a T cell clonotype, the set of T cells with the same T cell receptor, from its thymic origin to its extinction in a multiclonal repertoire. Using published estimates of total cell numbers and thymic production rates, we calculate the mean number of cells per TCR clonotype, and the total number of clonotypes, in mice and humans. When there is little peripheral division, as in a mouse, the number of cells per clonotype is small and governed by the number of cells with identical TCR that exit the thymus. In humans, peripheral division is important and a clonotype may survive for decades, during which it expands to comprise many cells. We therefore devise and analyse a computational model of homeostasis of a multiclonal population. Each T cell in the model competes for self pMHC stimuli, cells of any one clonotype only recognising a small fraction of the many subsets of stimuli. A constant mean total number of cells is maintained by a balance between cell division and death, and a stable number of clonotypes by a balance between thymic production of new clonotypes and extinction of existing ones. The number of distinct clonotypes in a human body may be smaller than the total number of naive T cells by only one order of magnitude.

1. Introduction

Approximately 4×10^{11} T cells circulate in the adult human body (Jenkins et al., 2009), each with multiple T cell receptors (TCR) (Varma, 2008) on its surface. Each T cell is descended from a T cell that left the thymus after selection by binding to self-peptides expressed in association with major histocompatibility complex molecules (self pMHC) (Bains et al., 2009; Van Laethem et al., 2012), or is itself a thymic emigrant. The immune system can respond to very many different pathogens because very many different TCRs are present in the body. However, the TCRs present on the surface of one T cell are usually all identical. The set of cells with the same TCR defines a T cell clonotype, and the set of T cells in the body can be thought of as a repertoire of clonotypes. How many TCR clonotypes are there in humans, mice and other mammals? (Langman and Cohn, 1987; Blattman et al., 2002; Ciupe et al., 2013; Thomas et al., 2014). Estimates of the number of different TCRs that could, in principle, be produced by VDJ gene rearrangement in the thymus,

are about 10^{15} (Sewell, 2012; Nikolich-Žugich et al., 2004; Zarnit-syna et al., 2013; Murugan et al., 2012). However, the human body cannot contain even one T cell of 10^{15} possible types: 10^{15} T cells would weigh about 500 kg (Mason, 1998).

Extracting and reverse transcribing RNA from pools of T cells, and amplifying, by PCR, the gene sequences that encode the TCR β chain, makes it possible to estimate the TCR diversity in a sample of blood. Initial estimates, based on extrapolation from a small fraction of the repertoire (Arstila et al., 1999), and more recent studies that are able to directly count large number of sequences and perform "missing species" analyses (Robins et al., 2009; Warren et al., 2011; Rempala and Seweryn, 2013), yield estimates of 10^6 to 10^8 (Qi et al., 2014). A new field of immunosequencing has emerged with technologies designed to sequence TCRs (Robins, 2013). Millions of TCR sequences can be amplified in a single multiplex PCR reaction, prepared and then read in parallel from a single sample. The distribution of gene usage can be measured with flow cytometry (Salameire et al., 2009; Ciupe et al., 2013), and used to track the dependence on phenotype, age and variation between individuals (Naylor et al., 2005; Britanova et al., 2014; Elhanati. et al., 2014; Becattini et al., 2015). Depending on the number of TCRα chains that each TCRβ chain combines with, the number of distinct clonotypes in one human may be much higher than estimates based on TCRβ alone (Keşmir et al., 2000). The spleen of a mouse has been estimated to contain 2×10^6 clones of about 10 cells each (Casrouge et al., 2000). In mice, dif-ferent T cell types can be compared and the effects of infections and immunization on the repertoire can be tracked (Bousso et al., 1998; Venturi et al., 2008; Bergot et al., 2015; Thomas et al., 2014).

In an adult, the number of recirculating T cells and the number of distinct clonotypes are believed to be held nearly constant for decades by balancing T cell loss with input from the thymus and homeostatic mechanisms controlling cell division and death in the periphery (Cannon, 1932; Tanchot et al., 1997; De Boer and Per-elson, 1997; Goldrath and Bevan, 1999; Berzins et al., 2002; Murray et al., 2003; Troy and Shen, 2003; Seddon and Zamoyska, 2003; Takada and Jameson, 2009; Rudd et al., 2011; Germain, 2012). The diversity of the T cell repertoire in the periphery is made possible by the enormous variability of their self pMHC ligands (De Boer and Perelson, 1995; Moses et al., 2003; Blanchfield et al., 2013). Division of T cells in the periphery is determined by competition for stimuli from self-peptides, presented in association with MHC class I (for CD8$^+$ T cells) and class II (for CD4$^+$ T cells), found on antigen presenting cells in the lymph nodes, and by soluble factors including IL-7 for naive T cells and IL-15 for memory T cells.

Emerging from the thymus with a pattern of recognition of self pMHC that enabled it to survive positive and negative selection, each TCR clonotype is a species that competes for "space" or "niche" in the periphery (De Boer and Perelson, 1997; Tanchot et al., 1997; Goldrath and Bevan, 1999; Jameson, 2002; Troy and Shen, 2003; Hataye et al., 2006; Moon et al., 2007; Agenes et al., 2008; Leitao et al., 2009). Competition between cells of the same clonotype has been demonstrated by transfer of T cells to TCR transgenic hosts of differing or identical clonotype (Hataye et al., 2006; Min et al., 2004). These and other experiments suggest that, in lymphoreplete conditions, T cells compete for specific survival signals provided by TCR recognition of self pMHC ligands (Hataye et al., 2006; Min et al., 2004). Using pools of 100 T cells each, Singh et al. observed competition between T cell clonotypes for a shared self-ligand, each self pMHC being recognised only by T cells from a small fraction of clonotypes (Singh et al., 2012). Naive T cells in mice have lifetimes measured in months; nearly all are thymic emigrants who have not undergone peripheral division (den Bra-ber et al., 2012; Thomas et al., 2013). Naive T cells in the human body have lifetimes measured in years; most naive peripheral T cells are daughter cells of other naive peripheral T cells (Vrisekoop et al., 2008; Murray et al., 2003; Houston et al., 2011).

The number of distinct TCR clonotypes, N, is equal to the total number of T cells divided by the mean number of cells per clo-notype. Equivalently, N is equal to the product of the rate of release of new clonotypes from the thymus to the periphery, θ, and the mean lifetime of a clonotype in the periphery. We estimate the mean lifetime of a clonotype using analysis of a computational model with simple assumptions. Combined with estimates of the total number of T cells, this calculation yields estimates of the total number of clonotypes. The case of homeostasis in mice is rather simple: because there is little or no cell division in the periphery, the mean number of cells per clonotype is less than the mean number of cells, per clonotype, that emerge from the thymus. To understand homeostasis in the human body, however, we consider rates of thymic output, cell death and division, using published estimates and our model. We are able to find an explicit mathe-matical expression for the mean lifetime of a clonotype, and hence estimate the mean number of clonotypes.

In this paper, we develop a stochastic model of homeostasis of naive T cells, in which TCR signals and competition within the multiclonal T cell pool regulate the size and diversity of the naive T cell pool. We have in mind the homeostasis of naive $\alpha\beta$ CD4$^+$ human T cells, but the structure of the mathematical model is that of an ecological competition process involving many similar spe-cies (Merrill et al., 1994; De Boer and Perelson, 1995, 1997; Tanchot et al., 1997; Freitas and Rocha, 2000; Mahajan et al., 2005; Thomas-Vaslin et al., 2008). The number of individual cells in each species (here, TCR clonotype) changes due to cell death and divi-sion (De Boer and Perelson, 1997; Antia et al., 1998; Ciupe et al., 2009). In our model, the effect of the thymus is to create new clonotypes, not to add new cells to existing clonotypes. We have previously analysed models that focus on the dynamics of one or two clonotypes without explicitly taking the multiclonal nature of the TCR repertoire into account (Stirk et al., 2008, 2010; Molina-París et al., 2011, 2011). In this work, the variables are the integer numbers of cells of each clonotype; a death or division event changes one of the variables by one cell. No clonotype has a pre-established self-limiting size; individual clonotypes compete for resources in a highly cross-reactive pattern of self pMHC recog-nition, which determines whether they die or survive. The fate of individual cells determines the fate of a clonotype and the fate of individual clonotypes determines the homeostasis of the reper-toire. New clonotypes are produced by the thymus but we assume that the thymus does not produce cells of a pre-existing type, and that each individual T cell has only one type of TCR. On the other hand, division of T cells in the periphery can only contribute to the survival or expansion of existing clonotypes (Houston et al., 2011; den Braber et al., 2012; Johnson et al., 2012). Extinction of a clo-notype, when it occurs, is thus irreversible.

1.1. Parameter values for humans and mice

The variables of the model are the numbers of cells of each clonotype; the parameters are the mean cell death rate, μ, the rate of cell division caused by stimulus from one self pMHC subset, γ, the rate at which new clonotypes are produced by the thymus, θ, the number of cells in a newly-produced clonotype, n_θ, the total number of self pMHC subsets, M, and the probability that any given self pMHC is recognised by a randomly-selected T cell clo-notype, p. In order to give a concrete estimate of clonotype numbers, we consider the values of μ, γ, θ, n_θ, M and p that give an appropriate description of the naive CD4$^+$ T cell repertoire in mice and humans.

The parameter μ sets the overall timescale that permits com-parison of computational and physiological dynamics. We take μ

to be constant, based on the simplifications that the T cell population is well mixed and that average availability of the IL-7 resource (Jameson, 2005; Bosco et al., 2005) is constant in time. The mean lifetime of a naive $CD4^+$ T cell in the human body (Michie et al., 1992; Macallan et al., 2004; Borghans and De Boer, 2007) is 1–10 years (De Boer et al., 2012; De Boer and Perelson, 2013; Westera et al., 2013, 2015). Thus, μ should be less than 1 $year^{-1}$. In a mouse, the corresponding lifetime is about 1 month (Labrecque et al., 2001; Bourgeois et al., 2008; Westera et al., 2013); we take $\mu = 1$ $month^{-1}$ when considering mice. Next, consider the total number of naive $CD4^+$ T cells. A human body has about 2×10^{11}; a mouse has about 10^7 (Borghans and De Boer, 2007; Jenkins et al., 2009; Bains et al., 2009; den Braber et al., 2012). As we shall see, this number fixes the value of the combination $\gamma M / \mu$ in humans, and constrains it in mice.

In our model, the number of subsets of pMHC stimuli is large, and individual TCRs are cross-reactive (Perelson and Oster, 1979; Evavold et al., 1995; Borghans et al., 1999; Newell et al., 2011; Calis and De Boer, 2012; van den Berg and Rand, 2007; Hao et al., 2006; Zarnitsyna et al., 2013; Yates, 2014; Birnbaum et al., 2014), recognising many different self pMHC; the number of subsets of self pMHC present is the parameter M. One estimate, based on the number of distinct 9-mers in the human proteome, is 10^7 (Burroughs and de Boer, 2004). Another, based on the observation that peptides up to at least 14 amino acids long can be presented by MHC molecules, is as high as 10^{16} (Mason, 1998; Sewell, 2012). It may be more appropriate to count the different possible antigen presentation patterns on the surface of antigen presenting cells (Van Den Berg et al., 2001). The appropriate value of M is thus difficult to establish. Fortunately, our predictions are not very sensitive to the choice of M.

The total number of cells, and the rate of thymic production, differ from one type of mammal to another. On the other hand, the universe of different epitopes, presented as self pMHC to T cells, is a property of the mammal's environment, lifestyle and physiology. Patterns of cross-reactivity depend on molecular properties of T cell receptors and may depend on HLA class (Košmrlj et al., 2010). We take the value of p, the probability that interaction with a given self pMHC is capable of causing a randomly-selected T cell to divide, to be the same in humans and mice: $p = 10^{-6}$ (Rizzuto et al., 2009; Su et al., 2013; Jenkins et al., 2009). Similarly, even though the body of a mouse is smaller than that of a human by close to a factor of 10^3, we use the estimate $M = 10^{10}$, in both cases. The number of cell divisions per unit time, in the whole body, produced by one self-pMHC subset, γ, is determined by the total amount of the resource available. We expect γ to be proportional to body weight, and thus larger in humans than in mice. The product γM is the total number of cell divisions per unit time in the periphery.

Thymic production rates have been estimated by direct and indirect methods (Scollay et al., 1980; Hazenberg et al., 2003; Thomas-Vaslin et al., 2008; Bains et al., 2009; den Braber et al., 2012). Because we seek to construct a model of homeostasis, we keep the parameter θ constant, although thymic output decreases with age in both mice and humans (Westera et al., 2015). The adult human thymus releases about 10^7 naive $CD4^+$ cells per day (den Braber et al., 2012), or 4×10^9 cells per year. The thymus of a mouse produces about 4×10^5 cells per day (den Braber et al., 2012), or 10^7 per month. The last of our parameters, n_θ, reflects the number of rounds of cell division in the thymus, after formation and selection of the full TCR but before a clonotype is released to the periphery (Hare et al., 1999, 2000; Sinclair et al., 2013; von Boehmer, 2014). For example, the choice $n_\theta = 4$ corresponds to the assumption that a thymocyte undergoes two rounds of cell division, after expression and selection of the TCR but before exit from the thymus.

2. Results

Numerical realisations update the variables

$n_i(t) =$ number of T cells of clonotype i at time t.

A numerical realisation starts with $N(0)$ clonotypes, when the index i that labels a clonotype runs from 1 to $N(0)$. We shall be particularly interested in

$N(t) =$ number of surviving clonotypes at time t.

That is, $N(t)$ is the number of i such that $n_i(t) > 0$. In any short time interval, a T cell either stays alive, dies, or divides into two cells of the same clonotype. The dynamics are governed by the birth (division) and death rates. For the latter, we adopt the simplest hypothesis: that each T cell, regardless of its clonotype, has probability μ per unit time of dying. In the absence of cell division, the number of cells in a population would decrease following an exponential decay curve (Tanchot et al., 1997). The probability per unit time of a T cell undergoing division into two daughter cells of the same clonotype, is proportional to the stimulus that it receives from the set of self pMHCs. Our assumption for cell division is that the stimulus from each subset is divided equally on average between the set of all T cells that recognise it, a fraction of all T cells that comprises T cells of many different clonotypes.

The rates of cell division are calculated from their interactions with M self pMHCs by generating, at the start of the run, the $N(0) \times M$ matrix A with entries (Košmrlj et al., 2010; Molina-París et al., 2011):

$$A_{iq} = \begin{cases} 1 & \text{if self pMHC } q \text{ stimulates T cells of clonotype } i \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

If $A_{iq} = 1$ then we say "self pMHC q stimulates T cells of clonotype i to divide" or, equivalently, "T cells of clonotype i recognise self pMHC q". In graph theory, the matrix A is known as the biadjacency matrix of the bipartite graph consisting of the set of T cell clonotypes, the set of self pMHC and the set of connections (Asratian et al., 1998). A small-scale example is depicted in Fig. 1.

The characteristic signature of a T cell clonotype is its pattern of interaction with the set of M self-pMHC stimuli, illustrated in Fig. 1. Each of the self pMHC is recognised with probability p (Borghans et al., 1999) so that the average number of self pMHCs recognised by a TCR (cross-reactivity) is pM. The number of possible patterns is thus at least as large as the number of ways of choosing pM samples from a total of M possibilities. If $M = 1000$ and $p = 0.1$ then the probability that two clonotypes have the same pattern, given by $((1-p)^2 + p^2)^M$, is already less than 10^{-85}. The recognition pattern of any one clonotype is highly unlikely to ever be repeated and every clonotype, existing in the periphery or emerging from the thymus, may be taken as unique. Similarly, the repertoires of different individual mammals from the same species will be independent samples from the set of possibilities, with little overlap (Bousso et al., 1998).

The number of T cells recognising self pMHC q at time t, $c_q(t)$, and the division rate of clonotype i, $\Lambda_i(t)$, are given by

$$c_q(t) = \sum_{i=1}^{N(t)} A_{iq} n_i(t) \quad \text{and} \quad \Lambda_i(t) = n_i(t) \sum_{q=1}^{M} \gamma_q A_{iq} \frac{1}{c_q(t)}, \quad (2)$$

where γ_q is the stimulus rate of self pMHC q. In this paper, we consider the case where all self pMHC deliver stimulus at the same rate, and the case where the rates are drawn from a log-normal distribution (see Methods). As $\Delta t \to 0$, $\Lambda_i(t)\Delta t$ is the probability that $n_i(t + \Delta t) - n_i(t) = 1$, corresponding to one round of cell division in one of the T cells of clonotype i between t and $t + \Delta t$. The quantity $\frac{n_i(t)}{c_q(t)}$ is the share of the total stimulus from q that is received by clonotype i, because of the assumption that each self pMHC distributes its stimulus, with equal probability, amongst all
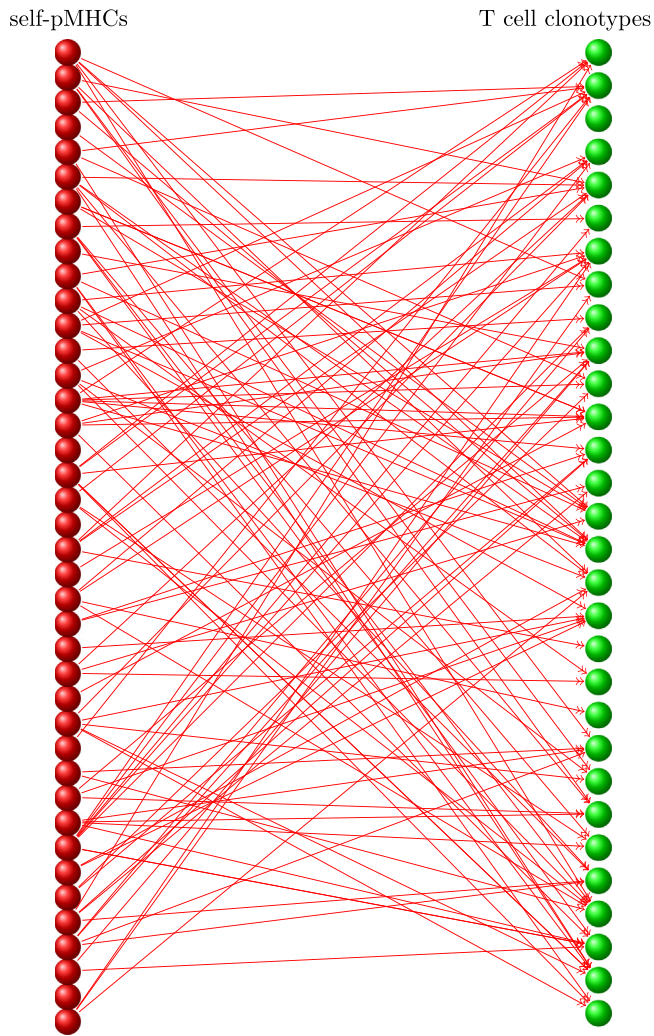
self-pMHCs

T cell clonotypes

Fig. 1. Representation of connections between $N=30$ T cell clonotypes and $M=40$ self pMHC subsets. An arrow connecting a red ball to a green ball indicates that the self pMHC stimulates division of cells in the T cell clonotype. The connections between clonotypes and self pMHCs are assigned randomly with probability $p=0.1$. That is, each entry of the matrix A, independently, is equal to 1 with probability p and equal to 0 with probability $1-p$. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

T cells capable of receiving it. The dynamics of the $n_i(t)$ are coupled because each $\Lambda_i(t)$ depends on $n_j(t)$, for all clonotypes j that share a self pMHC with i.

In our model, the thymus is a source of new T cell specificities, whereas division of T cells in the periphery can only contribute to the survival or expansion of existing clonotypes (den Braber et al., 2012). New clonotypes are produced with rate θ and number of cells n_θ. That is, with probability $\theta\Delta t$ in any small time interval Δt, a new clonotype is defined, by drawing a subset of connections to the M self pMHCs, allocated n_θ cells, and added to the system. Characterised by their self pMHC recognition pattern, recent thymic emigrants (Berzins et al., 1999; Fink, 2013) in our model are not given any special advantage or disadvantage when competing with peripheral cells, although both possibilities have been suggested (Vrisekoop et al., 2008; Thomas-Vaslin et al., 2008; Houston et al., 2011).

2.1. Mean total number of cells

One immediate consequence of the structure of the competition model is that there is a well-defined homeostatic mean

number of cells. Let the total number of T cells at time t be

$$n(t) = \sum_{i=1}^{N(t)} n_i(t)$$

and the mean total number of T cells at time t be $x(t) = \mathbb{E}(n(t))$. Note that $x(t)$, the mean over realisations of the stochastic processes of the model, is not constrained to be an integer and satisfies the ordinary differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}x = -\mu x + \Phi + \theta n_\theta, \tag{3}$$

where $\Phi = \sum_{i=1}^{N(t)} \Lambda_i(t)$ is the total rate of production of new cells by peripheral division. As long as each of the M self pMHCs is recognised by at least one cell,

$$\Phi = \sum_q \gamma_q$$

is constant and

$$x(t) = (\Phi + \theta n_\theta)\mu^{-1}(1 - e^{-\mu t}) + x(0)e^{-\mu t},$$

so that $x(t) \rightarrow \dfrac{\Phi + \theta n_\theta}{\mu}$, after a transient time of duration proportional to μ^{-1}. In the case where γ_q is equal to γ for every q, $\Phi = \gamma M$ and the stationary mean total number of cells is x^*, where

$$x^* = \frac{\gamma M + \theta n_\theta}{\mu}. \tag{4}$$

Note that this stably-maintained mean number of cells (De Boer and Perelson, 1997) arises as a global effect of balancing stochastic influences, but is not a unique state in the multi-dimensional space of clonal sizes. Indeed, as we shall see, individual clonotypes continue to follow their dynamics, reacting to changes in competing clonotypes and chance events, which can lead to extinction or successful establishment in the periphery.

The steady-state fraction, r, of cells that are thymic emigrants that have not divided in the periphery is given by $r = \theta n_\theta/(\gamma M + \theta n_\theta) = \theta n_\theta/(\mu n(t))$ (den Braber et al., 2012). If division of peripheral cells is the dominant contribution, the mean total number of cells is approximately equal to $\gamma M/\mu$. In adult humans, therefore, if $\mu = 0.5$ year^{-1} and the total number of naive CD4$^+$ T cells is 2×10^{11} then $\gamma M = 10^{11}$ year^{-1}. The observation that thymic production, rather than peripheral division, dominates the dynamics of murine T cell homeostasis implies that $\gamma M \ll 10^7$, in a mouse. Assuming that the values of M are similar in humans and mice, it is necessary to assume that the ratio γ/μ is smaller in a mouse than in a human, by a factor which reflects the overall size of the pMHC resource available, proportional to body weight (Table 1).

2.2. How many TCR clonotypes does a body maintain?

We can make estimates of the stable number of clonotypes, N^*, from estimates of the number of cells per clonotype. The dynamics in the "murine" limit is in fact rather simple: clonotypes operate independently, decay stochastically, and comprise few cells. New clonotypes, exported from the thymus, appear in the periphery with rate θ, initially consisting of n_θ cells. Each cell has a constant

Table 1
Estimated parameter values in humans and mice.

Parameter	Human	Mouse
μ	0.5 year^{-1}	1 month^{-1}
γ	10 year^{-1}	10^{-4} month^{-1}
θ	10^9 year^{-1}	2.5×10^6 month^{-1}
M	10^{10}	10^{10}
p	10^{-6}	10^{-6}
n_θ	4	4

probability per unit time of dying but cell division is very unlikely. If $n_\theta = 1$, then nearly every clonotype has only one cell. If not, the mean number of cells per clonotype is less than n_θ. With the simple estimate that the mean number of cells per clonotype, averaged over the lifetime of the clonotype, is $n_\theta/2$, we find that the mean number of clonotypes in the body of a mouse is related to the mean number of naive T cells in the mouse, x, as $t \to \infty$, by

$$N^* \simeq \frac{2}{n_\theta} x. \tag{5}$$

If we choose $n_\theta = 4$ and $x \simeq 10^7$, then we find $N^* \simeq 5 \times 10^6$.

We introduce the parameter

$$\alpha = \frac{n_\theta \theta}{\gamma M}. \tag{6}$$

The relation (5) is valid in the weak peripheral division limit, $\alpha \to 0$, appropriate to describe T cell homeostasis in a mouse. In the human body, on the other hand, many fewer cells result from thymic production than from peripheral division (Borghans and Tesselaar, 2009; den Braber et al., 2012). The stationary mean number of clonotypes in the periphery is found by multiplying the mean lifetime by the rate of production of new clonotypes, θ. In Methods, we compute the mean number of clonotypes by calculating the mean lifetime of a clonotype in the limit $\alpha \ll 1$:

$$N^* \to \frac{\gamma M}{\mu} \alpha (1 - \gamma_E - \log(n_\theta \alpha)), \quad \text{as } \alpha n_\theta \to 0, \tag{7}$$

where $\gamma_E = 0.577\ldots$ is the Euler–Mascheroni constant. Recall that $\frac{\gamma M}{\mu}$ is the mean total number of cells. In Fig. 2, the solid lines use (16) and the dotted line is (7).

Thymic export is estimated to contribute less than 10 percent of the total production of new T cells in adults (den Braber et al., 2012). Let us identify where on the curve in Fig. 2, to locate naive CD4$^+$ T cell homeostasis in adult humans. We use the estimate of annual production of cells by the thymus of 4×10^9, and the estimate of annual number of peripheral divisions of 10^{11}, to obtain $\frac{\theta n_\theta}{\gamma M} \simeq 0.04$. Thus, we find that the number of TCR clonotypes in the body is nine percent of the corresponding total number of T cells in the body, close to 10^{10}. In other words, our model of competition produces a mean clonal size, over the

Fig. 2. Predicted steady-state number of distinct clonotypes. The value of N^* is the product of the rate of production of new clonotypes, θ, and the mean lifetime of a clonotype, (16). The dotted lines are (7), valid in the weak-thymus limit. The parameter $\alpha = \frac{n_\theta \theta}{\gamma M}$ measures the strength of thymic production relative to peripheral division. Three values of n_θ are shown. We use $\gamma M/\mu = 10^{11}$ cells, approximately equal to the number of naive CD4$^+$ T cells in a human.

lifetime of a clone, that is only of order 10, even though some clonotypes expand to thousands of cells.

In Table 1, we collect our estimated values of the parameters. Based on these estimates, the number of distinct clonotypes in the human body is about ten times smaller than the total number of naive T cells ($N \simeq 10^{10}$) and the number of distinct clonotypes in a mouse is about two times smaller than the total number of naive T cells ($N \simeq 10^7$). These estimates are based on considering the lifetime of the population of cells making up one clonotype, starting with emergence from the thymus and ending in extinction in the periphery. Without peripheral division, the number of cells of one type does not exceed its initial value and the mean number of cells, averaged over the lifetime of the clonotype, is less than the number produced by the thymus, n_θ. When there is peripheral division, the number of cells of a given clonotype may be large at some points in the lifetime of the clone, but the average over the lifetime of the clonotype is less than the maximum.

In Methods, we describe the implementation of the Gillespie algorithm to produce numerical realisations of the model. In Figs. 3, 5 and 6, we display results from small-scale numerical realisations, without thymic input, that illustrate the extinction of clonotypes due to competition. Larger systems, maintaining an invariance from the perspective of any one clonotype, are shown in Fig. 7. The effect of thymic production is considered in Fig. 8. In the numerical realisations summarised in Fig. 9, each self pMHC is assigned a strength γ_q that is nor constant but drawn from a log-normal distribution. In Section 4.3, we consider the extinction time of the ith clonotype as a random variable, and derive an expression for its probability distribution using a diffusion approximation.

3. Discussion

Emerging from the thymus with a pattern of recognition of self pMHC, each TCR clonotype is a species that competes in the pool of pre-established ones. We have analysed naive T cell homeostasis by considering the lifetimes of sets of cells making up TCR clonotypes, from generation in the thymus to extinction in the periphery. Our point of view is to consider the stochastic dynamics of the number of T cells of one clonotype as a function of time, which increases or decreases according to the rates of cell division and death. Any one clonotype's time to extinction is a random variable that depends on its cross-reactivity, on the properties of those clonotypes with which it shares self-pMHC stimuli, and on average properties of the whole naive T cell repertoire. The model presented here has both interclonal and intraclonal competition, in which extinction of a clonotype is irreversible. By considering distributions of times to clonal extinction, we give estimates of clonal sizes, and hence diversity, in adult mice and humans. With the estimate that the ratio of thymic production to peripheral division of naive CD4$^+$ T cells is four percent, our model predicts that the number of distinct clonotypes in the human body is nine percent of the total number of naive CD4$^+$ T cells.

Our scheme is consistent with the idea that homeostasis of naive T cell numbers is maintained by availability of cytokines, especially IL-7 (Seddon et al., 2003; Min et al., 2004; Fry and Mackall, 2005; Tan et al., 2001; Pearson et al., 2011; Palmer et al., 2011; Koenen et al., 2013), that are available to T cells of all clonotypes (Ciupe et al., 2009), whereas the diversity of T cells is dependent on the varied self pMHC presented by the body's antigen presenting cells; each self pMHC being recognised only by T cells from a small fraction of clonotypes (Mahajan et al., 2005; Ciupe et al., 2009; Singh et al., 2012). It is possible to produce deterministic models of competition between clonotypes for different resources, with a biadjacency matrix, based on one ODE for
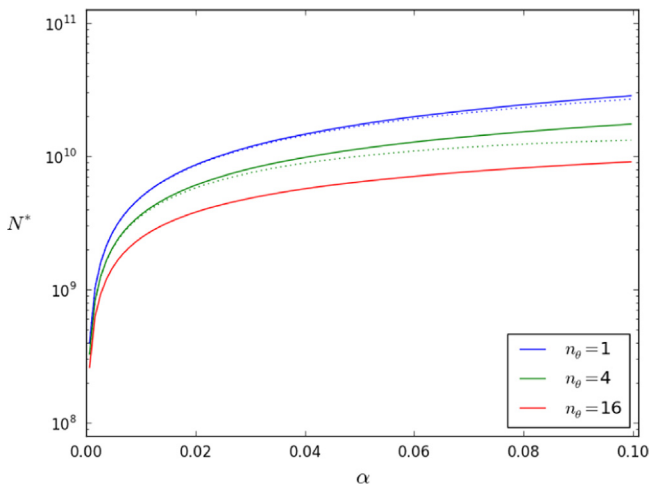
each clonotype (De Boer and Perelson, 1997). Stochastic models have the virtue of being able to include chance, as well as systematic competition effects, of being accurate when the numbers of cells of any type are small, and of describing extinction of clonotypes without extra assumptions.

Producing a large-scale competition model with the simplest assumptions and minimum number of parameters comes, of course, at the cost of simplifying the dynamics of homeostasis of the immune system. There are a number of ways that the simple model used here can be changed to more closely resemble reality. One that we have already explored is to relax the assumption that all self-pMHC complexes provide stimuli that produce cell division with the same rate. Instead, we assigned strengths to each drawn from a log-normal distribution. Even when the variance was larger than the mean, the results were similar to the constant-γ case. We believe that it is desirable to include interclonal and intraclonal heterogeneity (Sinclair et al., 2011; Mandl et al., 2013) to allow exploration of post-thymic maturation (Thomas-Vaslin et al., 2008), of the effects of cell-to-cell differences manifested in variation of CD25, IL-7R and CD5 expression levels, and differentiation into distinct phenotypes at the single-cell level (Zehn et al., 2012; Hataye et al., 2006; Jenkins et al., 2009; Pepper and Jenkins, 2011). Changes over time leading to increase or decrease of division rates in some clonotypes (Johnson et al., 2012) may also occur. Our numerical realisations have used the exact Gillespie algorithm, without resort to approximations based on averaging over populations. In order to undertake direct larger-scale simulations, with numbers of cells closer to physiological values, methods that do not update the Λ_i at every step, or efficient approximations such as "tau-leaping" (Márquez-Lago and Burrage, 2007), will be appropriate.

Although the model is designed to describe the long-term dynamics associated with naive T cell homeostasis, we envisage including strong but short-term stimuli (Hershberg et al., 2003) in the same framework to understand the effect of foreign antigen, the breadth of the immune response to a mutating virus (van Deutekom et al., 2013), and the creation of populations of memory cells, capable of persisting without TCR signals (Seddon and Zamoyska, 2003; Ganusov et al., 2006; Sprent and Surh, 2011). It is also necessary to consider the circulation of T cells around the body (Ganusov and Auerbach, 2014) and the distribution of different T cell subsets in different organs (Thome et al., 2014; Farber et al., 2014).

Dramatic illustrations of the consequences of perturbing homeostatic processes in the peripheral immune system are found in clinical or experimentally-induced lymphopenic environments (Bosco et al., 2005; Hogan et al., 2013; Paul et al., 2013). We can reproduce the reconstitution of the repertoire by thymic output and peripheral division with the model as presented here, but the constant-death-rate assumption will need to be modified to take into account the abundance of trophic factors during recovery from lymphopenia and heterogeneity in TCR and IL-7R expression levels (Palmer et al., 2011; Singh et al., 2012). Another interesting scenario is that of post-thymic transplant dynamics in infants with DiGeorge anomaly, where thymic production is explicitly a function of time (Freitas and Rocha, 2000; Houston et al., 2011; Ciupe et al., 2009).

4. Methods

4.1. The stochastic algorithm

At each step of the Gillespie algorithm (Renshaw, 2011; Wilkinson, 2006), one of the $2N(t)+1$ possible events is chosen. The event is either death of a cell in some clonotype, i, chosen with probability $\mu n_i(t)/S(t)$, birth of a cell in clonotype i, chosen with probability

$\Lambda_i(t)/S(t)$, or thymic production of a new clonotype, chosen with probability $\theta/S(t)$. Here, $S(t)$ is the sum of the rates at time t:

$$S(t) = \sum_{i}^{N(t)} \left(\mu n_i(t) + \Lambda_i(t) + \theta \right).$$

At the end of the step, time is incremented by an amount sampled from the exponential distribution with mean $1/S(t)$ (Wilkinson, 2006).

4.2. Competition

In Fig. 3, we display results from one small-scale numerical realisation, with $\theta = 0$ for simplicity. The effect of competition for stimulus is to drive some clonotypes to extinction, reducing the number of surviving clonotypes (Molina-París et al., 2011). The parameters are $p = 0.1$ and $\gamma_q = 10$ for every q. With the choice $\mu = 1$, one time unit is the mean lifetime of a T cell in the absence of stimulus (Borghans and De Boer, 2007; den Braber et al., 2012; De Boer and Perelson, 2013). In this work, when we consider the cases of T cell homeostasis in humans and mice, we shall use the values $\mu = 1$ month^{-1} (mouse) (Thomas-Vaslin et al., 2008) and $\mu = 1$ year^{-1} (human) (Macallan et al., 2004).

The relationship between the set of clonotypes and the set of self pMHCs is illustrated in Figs. 1 and 4. Consider the following quantities (Stirk et al., 2008, 2010; Molina-París et al., 2011, 2011):

- The set of self pMHCs that are recognised by T cells of clonotype i is denoted by Q_i, and the number of self pMHC subsets in Q_i, by ϕ_i.
- The set of T cells that recognise a self pMHC q is denoted by $C_q(t)$, and the number of surviving clonotypes in C_q by $|C_q|$.
- The average of ϕ_i over all surviving clonotypes is denoted $\overline{\phi}(t)$, and the average of $|C_q|$ over all q is denoted $\overline{C}(t)$ (Stirk et al., 2010).

An important exact relationship is (Mason, 1998; Zarnitsyna et al., 2013):

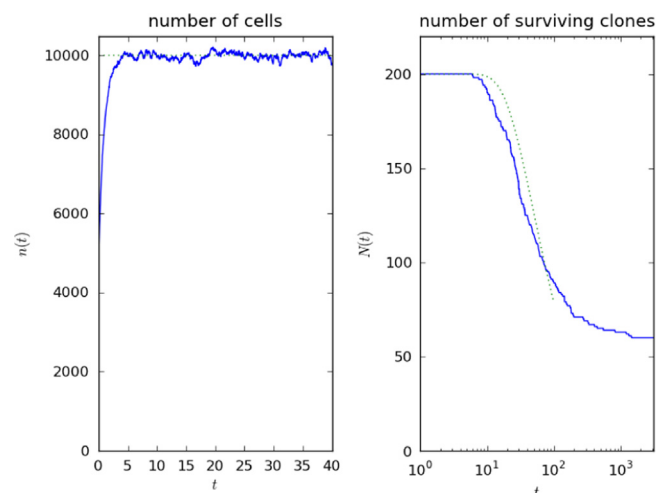$$M\overline{C}(t) = N(t)\overline{\phi}(t), \tag{8}$$

Fig. 3. A numerical solution of the competition model without thymic production. The left panel shows the total number of T cells, as a function of time. (The total number of cells continues to fluctuate about the same mean for times later than shown.) The right panel shows the number of surviving clonotypes, $N(t)$, with a logarithmic time scale. The dotted line is (14). One time unit is the mean lifetime of a T cell in the absence of division stimulus, taken to be one year in a human body. The parameters are $\mu = 1.0$, $\gamma = 10$, $M = 1000$; T cell clonotype-pMHC connections were assigned randomly with probability $p = 0.1$ at the beginning of the run. The initial conditions are $N(0) = 200$ and $n_i(0) = 25$ for each i.

which holds because the LHS and RHS are, in the representation of Fig. 1, different ways of counting the total number of connections between T cell clonotypes and self pMHCs. In our model, M is constant in time; the remaining three quantities in (8) vary in time but the exact equality is maintained.

There are many ways to assign connections between clonotypes and self pMHCs. A simple algorithm, used in this work, is to assign all connections independently with probability p. The values of ϕ_i are thus samples from the binomial distribution with mean pM, and

$$\overline{\phi}(0) = pM.$$

For sufficiently large pM, the distribution of values of ϕ_i is approximately Gaussian with standard deviation $\sqrt{pM}$. Despite its simplicity, the assignment of connections to clonotypes reproduces the basic features of positive and negative thymic selection: all clonotypes recognise some self pMHC, but none are too specific or too promiscuous.

Fig. 5 illustrates the effect of extinction of clonotypes on mean quantities, averaged over the repertoire. Extinction increases $\overline{\phi}(t)$, which is an average over the surviving clonotypes, but the change is necessarily confined within the distribution of values of ϕ_i (mean pM, standard deviation $\sqrt{pM}$). On the other hand, the average number of distinct clonotypes recognising a given self pMHC, $\overline{C}(t)$, is reduced because of extinction of clonotypes. As time increases, in the absence of thymic input, $N(t)$ decreases, $\overline{\phi}(t)$ is nearly constant, and $\overline{C}(t)$ decreases in such a way that the equality (8) is maintained.

4.3. Timescales of clonal extinction

Our point of view is to consider the stochastic dynamics of the number of T cells of one clonotype as a function of time. Using the description of the dynamics of a typical clonotype in the periphery, we can find an expression for the distribution of its time to extinction. Let the extinction time of the ith clonotype be the random variable τ_i. The number of cells, $n_i(t)$, of the clonotype with label i is an integer, increasing or decreasing by one cell at a time, according to the rates for cell division ("birth") and death; thus τ_i, the smallest value of t such that $n_i(t) = 0$. The death rate is $\mu n_i(t)$, independent of the other clonotypes in the system. In contrast, the division rate, $\Lambda_i(t)$, depends on the number of cells of all clonotypes that compete for any of the self pMHC recognised by clonotype i. If we begin with the expression, (2), and replace the sums by their mean values so that $c_q(t) \rightarrow px^*$ and $\sum_{q=1}^{M} \gamma_q A_{iq} \rightarrow \gamma pM$, then

$$\Lambda_i(t) = n_i(t)\frac{\gamma pM}{px^*} = \frac{1}{1+\alpha}\mu n_i(t).\qquad(9)$$

The most remarkable feature of (9) is that the birth rate $\Lambda_i(t)$ takes
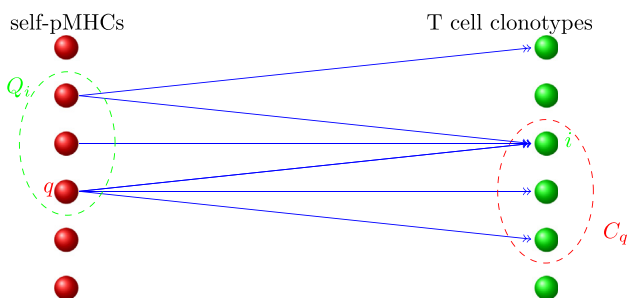
Fig. 4. Diagrammatic representation of the competition model. Each green ball represents a set of T cells; each red ball, a set of self pMHC complexes. An arrow (link) connecting a red ball to a green ball indicates that T cells of the clonotype recognise the self pMHC subset. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)
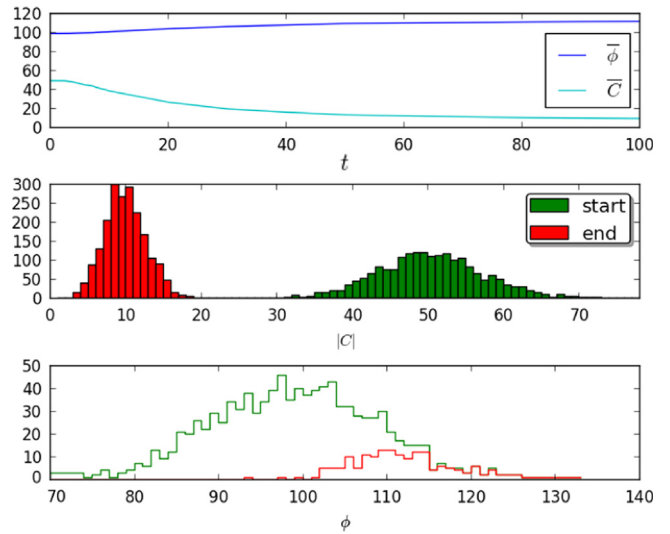
Fig. 5. A numerical solution of the exact clonal competition model without thymic input. In the top panel, the average number of self pMHCs recognised by a clonotype, $\overline{\phi}(t)$, and the average number of clonotypes recognising a pMHC, $\overline{C}(t)$, are shown as a function of time. Due to extinction of clonotypes, $\overline{\phi}(t)$ increases and $\overline{C}(t)$ decreases. The middle panel shows histograms of the 2000 values of $|C_q|$, at the start and end of the realisation. The bottom panel shows the histograms of values of ϕ at the start (1000 clonotypes) and end (168 clonotypes). T cell clonotype-pMHC connections are assigned randomly with probability $p = 0.05$. The remaining parameter values are $\mu = 1.0$, $\gamma = 10$, $M = 2000$. The initial conditions are $N(0) = 1000$ and $n_i(0) = 10 \; \forall i$.

a form that is similar to the death rate, $\mu n_i(t)$. In the absence of thymic production,

$$\Lambda_i(t) \simeq \mu n_i(t).\qquad(10)$$

Although it is not surprising that death rates and birth rates are in balance for a typical clonotype established in the periphery, it is remarkable that they can take identical functional forms.

To derive an expression for the distribution of τ_i, we use the diffusion approximation, where $n_i(t)$ is replaced by a diffusion process on the real line, $\mathbf{X}_t$ (Taylor and Karlin, 1998; Renshaw, 2011). The equation of motion of $\mathbf{X}_t$ is a stochastic differential equation (SDE) that is characterised by the mean and mean-square of its increments over a small time interval Δt (Gardiner, 2004; Wilkinson, 2006). Thus we must consider the mean change in the size of the population n_i between t and $t + \Delta t$:

$$\mathbb{E}(n_i(t + \Delta t)) = n_i(t) + (\Lambda_i(t) - \mu n_i(t))\Delta t.\qquad(11)$$

We require

$$\mathbb{E}(\mathbf{X}_{t+\Delta t} - \mathbf{X}_t) = -\frac{\alpha}{1+\alpha}\mu\mathbf{X}_t\Delta t \quad \text{and} \quad \mathbb{E}((\mathbf{X}_{t+\Delta t} - \mathbf{X}_t)^2) = \frac{2+\alpha}{1+\alpha}\mu\mathbf{X}_t\Delta t.$$

Hence the stochastic differential equation is

$$d\mathbf{X}_t = -\mu\frac{\alpha}{1+\alpha}dt + \sqrt{\frac{2+\alpha}{1+\alpha}}\sqrt{\mu\mathbf{X}_t}d\mathbf{W}_t,\qquad(12)$$

with initial condition $\mathbf{X}_0 = b$.

A simple expression is found when $\alpha = 0$: the distribution of times at which the diffusion process reaches the absorbing state at 0 is Revuz and Yor (2004):

$$\mathcal{P}[\mathbf{X}_t = 0 | \mathbf{X}_0 = b] = \exp\left(-\frac{b}{\mu t}\right).\qquad(13)$$

The appropriate initial condition b is the mean total number of cells, $\gamma M/\mu$, divided by the initial number of clonotypes, $N(0)$. That is, the probability that a randomly-chosen clonotype is extinct

before time t, without thymic production, is

$$\mathcal{P}[\text{extinction}] = \exp\left(-\frac{\gamma M}{N(0)\mu^2 t}\right). \tag{14}$$

The relation (14) is used to plot the dotted lines in Fig. 3.

Now let us consider $0 < \alpha \ll 1$. That is, we consider non-zero thymic production, but where more cells are created by peripheral division than by the thymus. Then

$$d\mathbf{X}_t = -\alpha\mu\mathbf{X}_t d\mathbf{X}_t + \sqrt{2\mu\mathbf{X}_t}\,d\mathbf{W}_t. \tag{15}$$

The mean lifetime of a clonotype is the mean time until $\mathbf{X}_t$ reaches zero, starting at b, $T(\alpha, b)$. It satisfies

$$\mu b \frac{d^2}{db^2} T(\alpha, b) - \alpha\mu b \frac{d}{db} T(\alpha, b) = -1. $$

The solution, with initial number of cells n_θ, is

$$T(\alpha, n_\theta) = \frac{1}{\alpha\mu}(\gamma_E - e^{\alpha n_\theta}\,\text{Ei}(-\alpha n_\theta) + \log(\alpha n_\theta)), \tag{16}$$

where $\gamma_E = 0.577\ldots$ is the Euler–Mascheroni constant. The expression (7) is obtained by using the properties of the exponential integral, Ei, for small values of its argument: as $\alpha n_\theta \to 0$, $\text{Ei}(-\alpha n_\theta) \to \gamma_E + \log(\alpha n_\theta) - \alpha n_\theta$ and so $T(\alpha, n_\theta) \to \frac{n_\theta}{\mu}(1 - \gamma_E - \log(\alpha n_\theta))$.

4.4. Selective or stochastic competition

Is the competition selective or stochastic? That is, do some clonotypes survive and other die because they are better suited to the competitive environment, or is it simply a matter of chance? It is possible to answer this question by carrying out a series of independent realisations, each with the same matrix A, encoding the pattern of connections between self pMHC and T cell clonotypes. The realisations correspond to independent series of stochastic events occurring to identical in silico individuals. The results of ten such realisations are summarised in Fig. 6. About five per cent of clonotypes survive, which could be achieved in two ways. If survival is determined by chance, any one clonotype will survive any one realisation with the same probability (here, about 0.05). If survival is determined by selection, a small subset of clonotypes will survive in most realisations (De Boer and Perelson, 1997). The presence of several near-complete vertical lines, in the subset of clonotypes shown, indicates that the competition is selective. Note, too, the correlation between higher values of ϕ_i and survival frequency.

While it is straightforward to perform computational studies of competition between hundreds or even thousands of T cell clonotypes, it is more challenging to study physiological numbers of clonotypes. It is therefore important to understand how the system can be scaled in such a way that the dynamics of a large system, viewed from any one clonotype, is independent of system size. In the random-connection case, where any one of N T cell clonotypes recognises any one of M self pMHCs with probability p, the number recognised by one T cell clonotype has a binomial distribution with parameter pM. We generate ever-larger systems, maintaining an invariance from the perspective of any one clonotype, by increasing M but keeping the product pM constant. The result, as illustrated in Fig. 7 with a set of three numerical runs, scales in the sense that the dynamics of the quantity $pN(t)$ is preserved. (The product pM is the mean number of self pMHC recognised by a single TCR clonotype; the product pN is the mean number of clonotypes that recognise a single self pMHC.)

4.5. Thymic production

Three numerical realisations are summarised in Fig. 8. As $\theta n_\theta \ll \gamma M$ in all three, thymic production has almost no perceptible effect on the
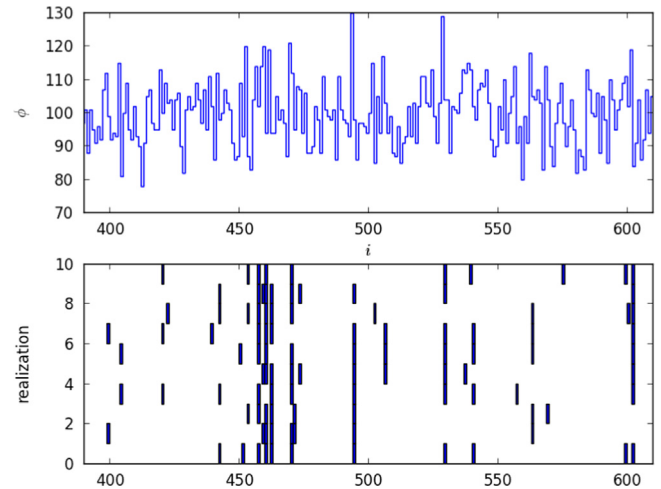

Fig. 6. Competition is more important than chance in determining which clonotypes survive. Of the initial 1000 clonotypes, on average only 50 survive. Ten independent realisations were carried out with the same connection matrix A; the fates of clonotypes $i=390$ to $i=610$ (horizontal axes) are illustrated. The upper panel shows the values of ϕ_i for the subset of clonotypes. (Each i has the same value of ϕ_i in each realisation.) In the lower panel, the vertical axis is the realisation number. A blue rectangle at position i in realisation j indicates that the clonotype has survived to $t = 10^4$ in that realisation. Many clonotypes never survive; some usually do. The parameter values are $\gamma = 10$, $M = 1000$, $p = 0.1$. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)
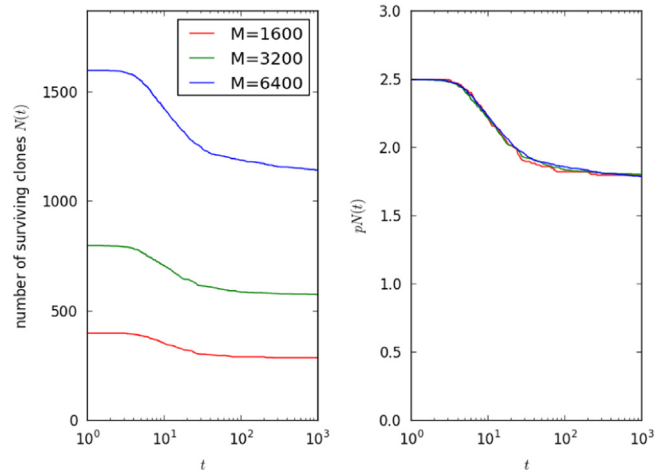

Fig. 7. Scaling of the dynamics with the size of the self pMHC environment. Different numerical runs, with $M = 1600$, $M = 3200$ and $M = 6400$, without thymic input. In each case, $pM = 100$, $\mu = 1$, $N(0) = \frac{1}{4}M$ and $\gamma = 10$. In the vertical axis on the right, the number of surviving clonotypes, as a function of time, is multiplied by p. The product pM is the mean number of self pMHC recognised by a single TCR clonotype; the product pN is the mean number of clonotypes that recognise a single self pMHC.

mean total number of cells, $x(t)$. Even so, it is the strength of thymic production that determines the stationary number of surviving clonotypes. With thymic production included in the system, a bona fide steady state is established, where new clonotypes from the thymus balance those that are driven to extinction. The steady state is established on a long timescale, corresponding to decades in humans. The stationary distribution of clonal sizes (that is, the distribution of numbers of cells in surviving clonotypes) and density of times to extinction, are shown on the right for the case $\theta = 10$ year^{-1}. In this work we have restricted our attention to homeostasis in adults. In future work we plan to extend the model to consider the full human lifetime, taking into account changes in thymic output and body size during childhood (Bains et al., 2009; Johnson et al., 2012).
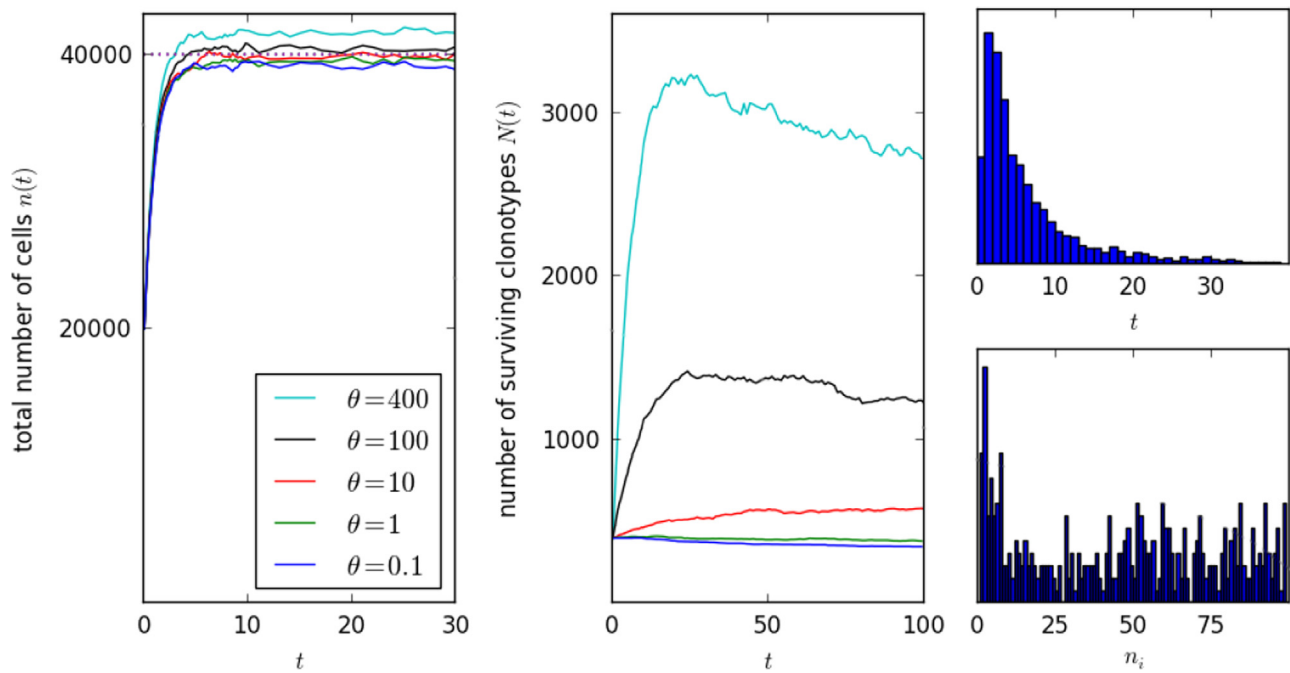
Fig. 8. Input of new clonotypes from the thymus, while barely affecting the total number of cells, determines the late-time number of surviving clonotypes. For comparison with adult human homeostasis, time is measured in years. The thymic output rates are $\theta = 0.1$ year^{-1} (blue), $\theta = 1$ year^{-1} (green), $\theta = 10$ year^{-1} (red), $\theta = 100$ year^{-1} (black) and $\theta = 400$ year^{-1} (cyan). The remaining parameter values are $M=4000$, $pM=100$, $\gamma = 10$ year^{-1}, $\mu = 1$ year^{-1}, $n_\theta = 4$ in all cases. The histograms on the right show, for $\theta = 10$ year^{-1}, the distribution of survival times of clonotypes from the thymus and the distribution of $n_i(t)$ for surviving clonotypes. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)
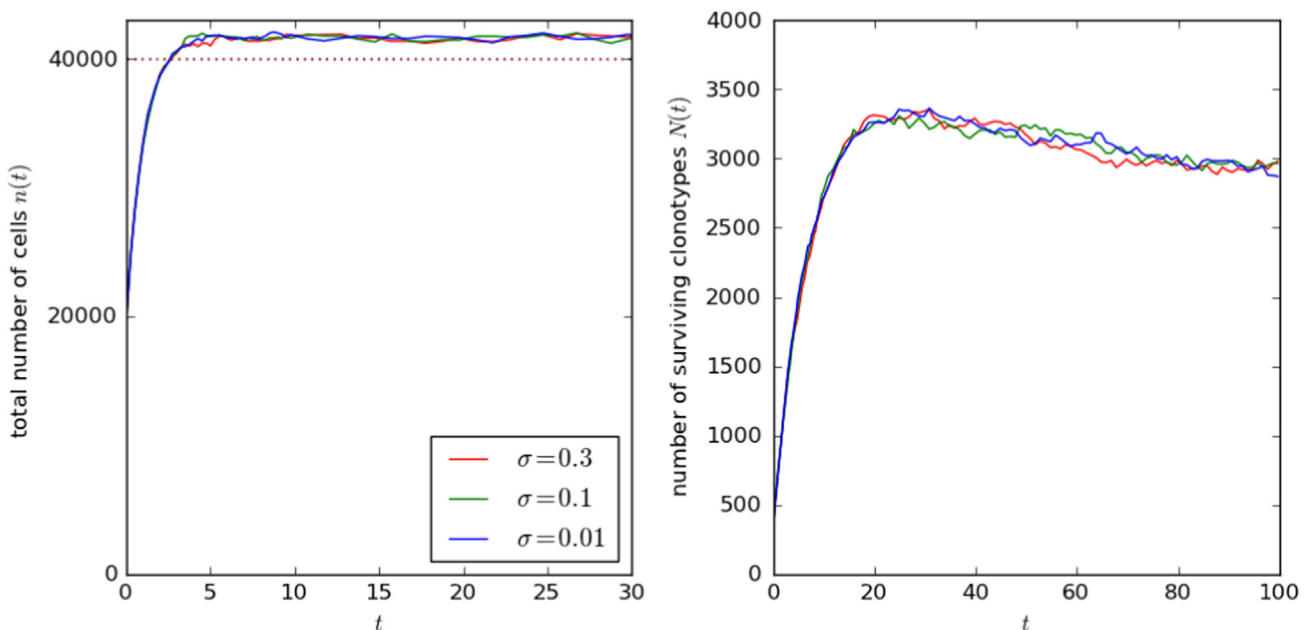

Fig. 9. Total number of cells, and total number of surviving clonotypes, as a function of time. Stimulus rates of the self pMHC, γ_q, are drawn from a log-normal distribution. The values of σ are the ratio of the standard deviation to the mean value of 10 year^{-1}. In the case $\sigma = 1$, the standard deviation is equal to the mean. The remaining parameter values are $M=4000$, $pM=100$, $\theta = 400$ year^{-1}, $\mu = 1$ year^{-1} and $n_\theta = 4$.

In Fig. 9, we display results from a version of the model with an extra element of structural stochasticity. Instead of a constant stimulus rate, each self pMHC is assigned a strength γ_q drawn from a log-normal distribution. The parameter σ is the standard deviation of the distribution, divided by the mean. Even with the widest distribution of rates, there is little discernible effect on the repertoire dynamics.

References

Agenes, F., Dangy, J.P., Kirberg, J., 2008. T cell receptor contact to restricting MHC molecules is a prerequisite for peripheral interclonal T cell competition. J. Exp. Med. 205 (12), 2735.

Antia, R., Pilyugin, S.S., Ahmed, R., 1998. Models of immune memory: on the role of cross-reactive stimulation, competition, and homeostasis in maintaining immune memory. Proc. Natl. Acad. Sci. 95 (25), 14926–14931.

Arstila, T.P., Casrouge, A., Baron, V., Even, J., Kanellopoulos, J., Kourilsky, P., 1999. A direct estimate of the human $\alpha\beta$ T cell receptor diversity. Science 286 (5441), 958.

Asratian, A.S., Denley, T.M.J., Häggkvist, R., 1998. Bipartite Graphs and Their Applications. Cambridge University Press, Cambridge, UK, 131.

Bains, I., Antia, R., Callard, R., Yates, A.J., 2009. Quantifying the development of the peripheral naive CD4+ T-cell pool in humans. Blood 113 (22), 5480.

Bains, I., Thiebaut, R., Yates, A.J., Callard, R., 2009. Quantifying thymic export: combining models of naive T cell proliferation and TCR excision circle dynamics gives an explicit measure of thymic output. J. Immunol. 183 (7), 4329.

Becattini, S., Latorre, D., Mele, F., Foglierini, M., De Gregorio, C., Cassotta, A., et al., 2015. Functional heterogeneity of human memory CD4+ T cell clones primed by pathogens or vaccines. Science 347 (6220), 400–406.

Bergot, A.S., Chaara, W., Ruggiero, E., Mariotti-Ferrandiz, E., Dulauroy, S., Schmidt, M., et al., 2015. TCR sequences and tissue distribution discriminate the subsets of naïve and activated/memory Treg cells in mice. Eur. J. Immunol. 45 (5), 1524–1534.

Berzins, S.P., Godfrey, D., Miller, J., Boyd, R., 1999. A central role for thymic emigrants in peripheral T cell homeostasis. Proc. Natl. Acad. Sci. 96 (17), 9787–9791.

Berzins, S.P., Uldrich, A.P., Sutherland, J.S., Gill, J., Miller, J.F., Godfrey, D.I., et al., 2002. Thymic regeneration: teaching an old immune system new tricks. Trends Mol. Med. 8 (10), 469–476.

Birnbaum, M.E., Mendoza, J.L., Sethi, D.K., Dong, S., Glanville, J., Dobbins, J., et al., 2014. Deconstructing the peptide-MHC specificity of T cell recognition. Cell 157 (5), 1073–1087.

Blanchfield, J.L., Shorter, S.K., Evavold, B.D., 2013. Monitoring the dynamics of T cell clonal diversity using recombinant Peptide: MHC technology. Front. Immunol. 4, 170.

Blattman, J.N., Antia, R., Sourdive, D.J., Wang, X., Kaech, S.M., Murali-Krishna, K., et al., 2002. Estimating the precursor frequency of naive antigen-specific CD8 T cells. J. Exp. Med. 195 (5), 657–664.

Borghans, J.A., De Boer, R.J., 2007. Quantification of T-cell dynamics: from telomeres to DNA labeling. Immunol. Rev. 216 (1), 35–47.

Borghans, J.A., Tesselaar, K., 2009. Be fruitful, multiply, and replenish. Blood 113 (22), 5369–5370.

Borghans, J.A., Noest, A.J., De Boer, R.J., 1999. How specific should immunological memory be? J. Immunol. 163 (2), 569–575.

Bosco, N., Agenès, F., Ceredig, R., 2005. Effects of increasing IL-7 availability on lymphocytes during and after lymphopenia-induced proliferation. J. Immunol. 175 (1), 162–170.

Bourgeois, C., Hao, Z., Rajewsky, K., Potocnik, A.J., Stockinger, B., 2008. Ablation of thymic export causes accelerated decay of naïve CD4 T cells in the periphery because of activation by environmental antigen. Proc. Natl. Acad. Sci. 105 (25), 8691–8696.

Bousso, P., Casrouge, A., Altman, J.D., Haury, M., Kanellopoulos, J., Abastado, J.P., et al., 1998. Individual variations in the murine T cell response to a specific peptide reflect variability in naive repertoires. Immunity 9 (2), 169–178.

Britanova, O.V., Putintseva, E.V., Shugay, M., Merzlyak, E.M., Turchaninova, M.A., Staroverov, D.B., et al., 2014. Age-related decrease in TCR repertoire diversity measured with deep and normalized sequence profiling. J. Immunol. 192 (6), 2689–2698.

Burroughs, N.J., de Boer, R.J., Keşmir, C., 2004. Discriminating self from nonself with short peptides from large proteomes. Immunogenetics 56 (5), 311–320.

Calis, J.J., De Boer, R.J., Keşmir, C., 2012. Degenerate T-cell recognition of peptides on MHC molecules creates large holes in the T-cell repertoire. PLoS Comput. Biol. 8 (3), e1002412.

Cannon, W.B., 1932. The Wisdom of the Body. WW Norton & Co, New York.

Casrouge, A., Beaudoing, E., Dalle, S., Pannetier, C., Kanellopoulos, J., Kourilsky, P., 2000. Size estimate of the $\alpha\beta$ TCR repertoire of naive mouse splenocytes. J. Immunol. 164 (11), 5782–5787.

Ciupe, S.M., Devlin, B.H., Markert, M.L., Kepler, T.B., 2009. The dynamics of T-cell receptor repertoire diversity following thymus transplantation for DiGeorge anomaly. PLoS Comput. Biol. 5 (6), e1000396.

Ciupe, S.M., Devlin, B.H., Markert, M.L., Kepler, T.B., 2013. Quantification of total T-cell receptor diversity by flow cytometry and spectratyping. BMC Immunol. 14 (1), 1–12.

De Boer, R.J., Perelson, A.S., 1995. Towards a general function describing T cell proliferation. J. Theor. Biol. 175 (4), 567–576.

De Boer, R., Perelson, A., 1997. Competitive control of the self-renewing T cell repertoire. Int. Immunol. 9 (5), 779.

De Boer, Rob J, Perelson, Alan S, 2013. Quantifying T lymphocyte turnover. J. Theor. Biol. 327, 45–87.

De Boer, R.J., Perelson, A.S., Ribeiro, R.M., 2012. Modelling deuterium labelling of lymphocytes with temporal and/or kinetic heterogeneity. J. R. Soc. Interface 9 (74), 2191–2200.

den Braber, I., Mugwagwa, T., Vrisekoop, N., Westera, L., Mögling, R., Bregje de Boer, A., et al., 2012. Maintenance of peripheral naive T cells is sustained by thymus output in mice but not humans. Immunity 36 (2), 288–297.

Elhanati, Y., Murugan, A., Callan, Jr CG., Mora, T., Walczak, AM. Quantifying selection in immune receptor repertoires. Arxiv Preprint Arxiv:14044956.

Evavold, B.D., Sloan-Lancastert, J., Wilson, K.J., Rothbard, J.B., Allen, P.M., 1995. Specific T cell recognition of minimally homologous peptides: evidence for multiple endogenous ligands. Immunity 2 (6), 655–663.

Farber, D.L., Yudanin, N.A., Restifo, N.P., 2014. Human memory T cells: generation, compartmentalization and homeostasis. Nat. Rev. Immunol. 14 (1), 24–35.

Fink, P.J., 2013. The biology of recent thymic emigrants. Ann. Rev. Immunol. 31, 31–50.

Freitas, A.A., Rocha, B., 2000. Population biology of lymphocytes: the flight for survival. Ann. Rev. Immunol. 18 (1), 83–111.

Fry, T.J., Mackall, C.L., 2005. The many faces of IL-7: from lymphopoiesis to peripheral T cell maintenance. J. Immunol. 174 (11), 6571–6576.

Ganusov, V.V., Auerbach, J., 2014. Mathematical modeling reveals kinetics of lymphocyte recirculation in the whole organism. PLoS Comput. Biol. 10 (5), e1003586.

Ganusov, V.V., Pilyugin, S.S., Ahmed, R., Antia, R., 2006. How does cross-reactive stimulation affect the longevity of CD8+ T cell memory? PLoS Comput. Biol. 2 (6), e55.

Gardiner, C.W., 2004. Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences, 3rd ed. Springer, Berlin.

Germain, R.N., 2012. Maintaining system homeostasis: the third law of Newtonian immunology. Nat. Immunol. 13 (10), 902–906.

Goldrath, A.W., Bevan, M.J., 1999. Selecting and maintaining a diverse T-cell repertoire. Nature 402 (6759), 255–262.

Hao, Y., Legrand, N., Freitas, A.A., 2006. The clone size of peripheral CD8 T cells is regulated by TCR promiscuity. J. Exp. Med. 203 (7), 1643.

Hare, K.J., Jenkinson, E.J., Anderson, G., 1999. CD69 expression discriminates MHC-dependent and -independent stages of thymocyte positive selection. J. Immunol. 162 (7), 3978–3983.

Hare, K.J., Jenkinson, E.J., Anderson, G., 2000. An essential role for the IL-7 receptor during intrathymic expansion of the positively selected neonatal T cell repertoire. J. Immunol. 165 (5), 2410–2414.

Hataye, J., Moon, J.J., Khoruts, A., Reilly, C., Jenkins, M.K., 2006. Naive and memory CD4+ T cell survival controlled by clonal abundance. Science 312 (5770), 114.

Hazenberg, M.D., Borghans, J.A., de Boer, R.J., Miedema, F., 2003. Thymic output: a bad TREC record. Nat. Immunol. 4 (2), 97–99.

Hershberg, U., Solomon, S., Cohen, I., 2003. What is the basis of the immune systems specificity? Math. Model. Comput. Biol. Med., 377–384.

Hogan, T., Shuvaev, A., Commenges, D., Yates, A., Callard, R., Thiebaut, R., et al., 2013. Clonally diverse T cell homeostasis is maintained by a common program of cell-cycle control. J. Immunol. 190 (8), 3985–3993.

Houston, E.G., Higdon, L.E., Fink, P.J., 2011. Recent thymic emigrants are preferentially incorporated only into the depleted T-cell pool. Proc. Natl. Acad. Sci. 108 (13), 5366.

Jameson, S.C., 2002. Maintaining the norm: T-cell homeostasis. Nat. Rev. Immunol. 2 (8), 547–556.

Jameson, S.C., Elsevier, 2005. T cell homeostasis: keeping useful T cells alive and live T cells useful. Semin. Immunol. 17 (3), 231–237.

Jenkins, M.K., Chu, H.H., McLachlan, J.B., Moon, J.J., 2009. On the composition of the preimmune repertoire of T cells specific for peptide-major histocompatibility complex ligands. Ann. Rev. Immunol. 28, 275–294.

Johnson, P.L., Yates, A.J., Goronzy, J.J., Antia, R., 2012. Peripheral selection rather than thymic involution explains sudden contraction in naive CD4 T-cell diversity with age. Proc. Natl. Acad. Sci. 109 (52), 21432–21437.

Keşmir, C., Borghans, JA., de Boer, RJ., 2000. Diversity of human $\alpha\beta$ T cell receptors. Science 288 (5469), 1135.

Košmrlj, A., Read, E.L., Qi, Y., Allen, T.M., Altfeld, M., Deeks, S.G., et al., 2010. Effects of thymic selection of the T-cell repertoire on HLA-class I associated control of HIV infection. Nature 465 (7296), 350–354.

Koenen, P., Heinzel, S., Carrington, E.M., Happo, L., Alexander, W.S., Zhang, J.G., et al., 2013. Mutually exclusive regulation of T cell survival by IL-7R and antigen receptor-induced signals. Nat. Commun. 4, 1735.

Labrecque, N., Whitfield, L.S., Obst, R., Waltzinger, C., Benoist, C., Mathis, D., 2001. How much TCR does a T cell need? Immunity 15 (1), 71–82.

Langman, R., Cohn, M., 1987. The ET (elephant-tadpole) paradox necessitates the concept of a unit of B-cell function: the protecton. Mol. Immunol. 24 (7), 675–697.

Leitao, C., Freitas, A.A., Garcia, S., 2009. The role of TCR specificity and clonal competition during reconstruction of the peripheral T cell pool. J. Immunol. 182 (9), 5232.

Márquez-Lago, T.T., Burrage, K., 2007. Binomial tau-leap spatial stochastic simulation algorithm for applications in chemical kinetics. J. Chem. Phys. 127, 104101.

Macallan, D.C., Wallace, D., Zhang, Y., de Lara, C., Worth, A.T., Ghattas, H., et al., 2004. Rapid turnover of effector-memory CD4+ T cells in healthy humans. J. Exp. Med. 200 (2), 255–260.

Mahajan, V.S., Leskov, I.B., Chen, J.Z., 2005. Homeostasis of T cell diversity. Cell. Mol. Immunol. 2 (1), 1–10.

Mandl, J.N., Monteiro, J.P., Vrisekoop, N., Germain, R.N., 2013. T cell-positive selection uses self-ligand binding strength to optimize repertoire recognition of foreign antigens. Immunity 38 (2), 263–274.

Mason, D., 1998. A very high level of crossreactivity is an essential feature of the T-cell receptor. Immunol. Today 19 (9), 395–404.

Merrill, S., De Boer, R., Perelson, A.S., 1994. Development of the T-cell repertoire: clone size distribution. Rocky Mt. J. Math. 24 (1), 214–231.

Michie, C.A., McLean, A., Alcock, C., Beverley, P.C., 1992. Lifespan of human lymphocyte subsets defined by CD45 isoforms. Nature 360, 264–265.

Min, B., Foucras, G., Meier-Schellersheim, M., Paul, W.E., 2004. Spontaneous proliferation, a response of naive CD4 T cells determined by the diversity of the memory cell repertoire. Proc. Natl. Acad. Sci. 101 (11), 3874–3879.

Molina-París, C., Stirk, E., Quinn, K., Lythe, G., 2011. Continuous-time birth and death processes: diversity maintenance of Naïve T cells in the periphery. Math. Models Immune Cell Biol., 171–186.

Molina-París, C., Lythe, G., Stirk, E., 2011. Multivariate competition processes: a model for two competing T cell clonotypes. Math. Models Immune Cell Biol., 187–205.

Moon, J.J., Chu, H.H., Pepper, M., McSorley, S.J., Jameson, S.C., Kedl, R.M., et al., 2007. Naive CD4$^+$ T cell frequency varies for different epitopes and predicts repertoire diversity and response magnitude. Immunity 27 (2), 203–213.

Moses, C.T., Thorstenson, K.M., Jameson, S.C., Khoruts, A., 2003. Competition for self ligands restrains homeostatic proliferation of naive CD4 T cells. Proc. Natl. Acad. Sci. 100 (3), 1185–1190.

Murray, J.M., Kaufmann, G.R., Hodgkin, P.D., Lewin, S.R., Kelleher, A.D., Davenport, M.P., et al., 2003. Naive T cells are maintained by thymic output in early ages but by proliferation without phenotypic change after age twenty. Immunol. Cell Biol. 81 (6), 487–495.

Murugan, A., Mora, T., Walczak, A.M., Callan, C.G., 2012. Statistical inference of the generation probability of T-cell receptors from sequence repertoires. Proc. Natl. Acad. Sci. 109 (40), 16161–16166.

Naylor, K., Li, G., Vallejo, A.N., Lee, W.W., Koetz, K., Bryl, E., et al., 2005. The influence of age on T cell generation and TCR diversity. J. Immunol. 174 (11), 7446–7452.

Newell, E.W., Ely, L.K., Kruse, A.C., Reay, P.A., Rodriguez, S.N., Lin, A.E., et al., 2011. Structural basis of specificity and cross-reactivity in t cell receptors specific for cytochrome c-I-E^k. J. Immunol. 186 (10), 5823–5832.

Nikolich-Žugich, J., Slifka, M.K., Messaoudi, I., 2004. The many important facets of T-cell repertoire diversity. Nat. Rev. Immunol. 4 (2), 123–132.

Palmer, M.J., Mahajan, V.S., Chen, J., Irvine, D.J., Lauffenburger, D.A., 2011. Signaling thresholds govern heterogeneity in IL-7-receptor-mediated responses of naïve CD8$^+$ T cells. Immunol. Cell Biol. 89 (5), 581–594.

Paul, W.E., Milner, J.D., Grossman, Z., 2013. Pathogen-sensing, regulatory T cells, and responsiveness-tuning collectively regulate foreign- and self-antigen mediated T-cell responses. Cold Spring Harbor Symposia on Quantitative Biology. Cold Spring Harbor Laboratory Press, New York, p. a020198.

Pearson, C., Silva, A., Saini, M., Seddon, B., 2011. IL-7 determines the homeostatic fitness of T cells by distinct mechanisms at different signalling thresholds in vivo. Eur. J. Immunol. 41, 3656–3666.

Pepper, M., Jenkins, M.K., 2011. Origins of CD4$^+$ effector and central memory T cells. Nat. Immunol. 131 (6), 467–471.

Perelson, A.S., Oster, G.F., 1979. Theoretical studies of clonal selection: minimal antibody repertoire size and reliability of self-non-self discrimination. J. Theor. Biol. 81 (4), 645–670.

Qi, Q., Liu, Y., Cheng, Y., Glanville, J., Zhang, D., Lee, J.Y., et al., 2014. Diversity and clonal selection in the human T-cell repertoire. Proc. Natl. Acad. Sci. 111 (36), 13139–13144.

Rempala, G.A., Seweryn, M., 2013. Methods for diversity and overlap analysis in T-cell receptor populations. J. Math. Biol. 67 (6-7), 1339–1368.

Renshaw, E., 2011. Stochastic Population Processes: Analysis, Approximations, Simulations. Oxford University Press, Oxford, UK.

Revuz, D., Yor, M,. 2004 Continuous Martingales and Brownian Motion, vol. 293. Springer, Berlin.

Rizzuto, G.A., Merghoub, T., Hirschhorn-Cymerman, D., Liu, C., Lesokhin, A.M., Sahawneh, D., et al., 2009. Self-antigen-specific CD8$^+$ T cell precursor frequency determines the quality of the antitumor immune response. J. Exp. Med. 206 (4), 849–866.

Robins, H.S., Campregher, P.V., Srivastava, S.K., Wacher, A., Turtle, C.J., Kahsai, O., et al., 2009. Comprehensive assessment of T-cell receptor β-chain diversity in $\alpha\beta$ T cells. Blood 114 (19), 4099–4107.

Robins, H., 2013. Immunosequencing: applications of immune repertoire deep sequencing. Curr. Opin. Immunol. 25 (5), 646–652.

Rudd, B.D., Venturi, V., Li, G., Samadder, P., Ertelt, J.M., Way, S.S., et al., 2011. Non-random attrition of the naive CD8$^+$ T-cell pool with aging governed by T-cell receptor: pMHC interactions. Proc. Natl. Acad. Sci. 108 (33), 13694–13699.

Salameire, D., Le Bris, Y., Fabre, B., Fauconnier, J., Solly, F., Pernollet, M., et al., 2009. Efficient characterization of the TCR repertoire in lymph nodes by flow cytometry. Cytometry Part A 75 (9), 743–751.

Scollay, R.G., Butcher, E.C., Weissman, I.L., 1980. Thymus cell migration: quantitative aspects of cellular traffic from the thymus to the periphery in mice. Eur. J. Immunol. 10 (3), 210–218.

Seddon, B., Zamoyska, R., 2003. Regulation of peripheral T-cell homeostasis by receptor signalling. Curr. Opin. Immunol. 15 (3), 321–324.

Seddon, B., Tomlinson, P., Zamoyska, R., 2003. Interleukin 7 and T cell receptor signals regulate homeostasis of CD4 memory cells. Nat. Immunol. 4 (7), 680–686.

Sewell, A.K., 2012. Why must T cells be cross-reactive? Nat. Rev. Immunol. 12 (9), 669–677.

Sinclair, C., Saini, M., Schim van der Loeff, I., Sakaguchi, S., Seddon, B., 2011. The long-term survival potential of mature T lymphocytes is programmed during development in the thymus. Sci. Signal. 4 (199), ra77.

Sinclair, C., Bains, I., Yates, A.J., Seddon, B., 2013. Asymmetric thymocyte death underlies the CD4: CD8 T-cell ratio in the adaptive immune system. Proc. Natl. Acad. Sci. 110 (31), E2905–E2914.

Singh, N.J., Bando, J.K., Schwartz, R.H., 2012. Subsets of nonclonal neighboring CD4$^+$ T cells specifically regulate the frequency of individual antigen-reactive T cells. Immunity 37 (4), 735–746.

Sprent, J., Surh, C.D., 2011. Normal T cell homeostasis: the conversion of naive cells into memory-phenotype cells. Nat. Immunol. 131 (6), 478–484.

Stirk, E.R., Molina-París, C., van den Berg, H.A., 2008. Stochastic niche structure and diversity maintenance in the T cell repertoire. J. Theor. Biol. 255 (2), 237–249.

Stirk, E.R., Lythe, G., van den Berg, H.A., Hurst, G.A.D., Molina-París, C., 2010. The limiting conditional probability distribution in a stochastic model of T cell repertoire maintenance. Math. Biosci. 224 (2), 74–86.

Su, L.F., Kidd, B.A., Han, A., Kotzin, J.J., Davis, M.M., 2013. Virus-specific CD4$^+$ memory-phenotype T cells are abundant in unexposed adults. Immunity 38, 373–383.

Takada, K., Jameson, S.C., 2009. Naive T cell homeostasis: from awareness of space to a sense of place. Nat. Rev. Immunol. 9 (12), 823–832.

Tan, J.T., Dudl, E., LeRoy, E., Murray, R, Sprent, J., Weinberg, K.I., et al., 2001. IL-7 is critical for homeostatic proliferation and survival of naive T cells. Proc. Natl. Acad. Sci. 98 (15), 8732–8737.

Tanchot, C., Rosado, M.M., Agenès, F., Freitas, A.A., Rocha, B., 1997. Lymphocyte homeostasis. Semin. Immunol. 9 (6), 331–337.

Taylor, H.M., Karlin, S., 1998. An Introduction to Stochastic Modeling. Academic Press, Boston.

Thomas, P.G., Handel, A., Doherty, P.C., La Gruta, N.L., 2013. Ecological analysis of antigen-specific CTL repertoires defines the relationship between naive and immune T-cell populations. Proc. Natl. Acad. Sci. 110 (5), 1839–1844.

Thomas, N., Best, K., Cinelli, M., Reich-Zeliger, S., Gal, H., Shifrut, E., et al., 2014. Tracking global changes induced in the CD4 T cell receptor repertoire by immunisation with a complex antigen using local sequence features of CDR3 protein sequence, Biorxiv.org.

Thomas-Vaslin, V., Altes, H.K., de Boer, R.J., Klatzmann, D., 2008. Comprehensive assessment and mathematical modeling of T cell population dynamics and homeostasis. J. Immunol. 180 (4), 2240.

Thome, J.J., Yudanin, N., Ohmura, Y., Kubota, M., Grinshpun, B., Sathaliyawala, T., et al., 2014. Spatial map of human T cell compartmentalization and maintenance over decades of life. Cell 159 (4), 814–828.

Troy, A.E., Shen, H., 2003. Cutting edge: homeostatic proliferation of peripheral T lymphocytes is regulated by clonal competition. J. Immunol. 170 (2), 672.

van den Berg, H.A., Rand, D.A., 2007. Quantitative theories of T-cell responsiveness. Immunol. Rev. 216 (1), 81–92.

Van Den Berg, H., Rand, D., Burroughs, N., 2001. A reliable and safe T cell repertoire based on low-affinity T cell receptors. J. Theor. Biol. 209 (4), 465–486.

van Deutekom, H.W., Wijnker, G., de Boer, R.J., 2013. The rate of immune escape vanishes when multiple immune responses control an HIV infection. J. Immunol. 191 (6), 3277–3286.

Van Laethem, F., Tikhonova, A.N., Singer, A., 2012. MHC restriction is imposed on a diverse T cell receptor repertoire by CD4 and CD8 co-receptors during thymic selection. Trends Immunol. 33 (9), 437–441.

Varma, R., 2008. TCR triggering by the pMHC complex: valency, affinity, and dynamics. Sci. Signal. 1 (19), pe21.

Venturi, V., Kedzierska, K., Tanaka, M.M., Turner, S.J., Doherty, P.C., Davenport, M.P., 2008. Method for assessing the similarity between subsets of the T cell receptor repertoire. J. Immunol. Methods 329 (1), 67–80.

von Boehmer, H., 2014. The thymus in immunity and in malignancy. Cancer Immunol. Res. 2 (7), 592–597.

Vrisekoop, N., Den Braber, I., De Boer, A.B., Ruiter, A.F.C., Ackermans, M.T., Van Der Crabben, S.N., et al., 2008. Sparse production but preferential incorporation of recently produced naive T cells in the human peripheral pool. Proc. Natl. Acad. Sci. 105 (16), 6115.

Warren, R.L., Freeman, J.D., Zeng, T., Choe, G., Munro, S., Moore, R., et al., 2011. Exhaustive T-cell repertoire sequencing of human peripheral blood samples reveals signatures of antigen selection and a directly measured repertoire size of at least 1 million clonotypes. Genome Res. 21 (5), 790–797.

Westera, L., Drylewicz, J., den Braber, I., Mugwagwa, T., van der Maas, I., Kwast, L., et al., 2013. Closing the gap between T-cell life span estimates from stable isotope-labeling studies in mice and men. Blood 122, 2205–2212.

Westera, L., Hoeven, V., Drylewicz, J., Spierenburg, G., Velzen, J.F., Boer, R.J., et al., 2015. Lymphocyte maintenance during healthy aging requires no substantial alterations in cellular turnover. Aging Cell 14 (2), 219–227.

Wilkinson, D.J., 2006. Stochastic Modelling for Systems Biology. CRC Press, Boca Raton.

Yates, A., 2014. Theories and quantification of thymic selection. Front. Immunol. 5, 13.

Zarnitsyna, V., Evavold, B., Schoettle, L., Blattman, J., Antia, R., 2013. Estimating the diversity, completeness, and cross-reactivity of the T cell repertoire. Front. Immunol. 4, 485.

Zehn, D., King, C., Bevan, M.J., Palmer, E., 2012. TCR signaling requirements for activating T cells and for generating memory. Cell. Mol. Life Sci. 69 (10), 1565–1575.