Los Alamos NATIONAL LABORATORY EST.1943

LA-UR-15-28321 (Accepted Manuscript)

Mathematics in modern immunology

Lythe, Grant
Castro, Mario
Paris-Molina, Carmen
Ribeiro, Ruy Miguel

Provided by the author(s) and the Los Alamos National Laboratory (2016-12-06).

To be published in: Interface Focus

DOI to publisher's version: 10.1098/rsfs.2015.0093

Permalink to record: http://permalink.lanl.gov/object/view?what=info:lanl-repo/lareport/LA-UR-15-28321

Disclaimer:
Approved for public release. Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by the Los Alamos National Security, LLC for the National Nuclear Security Administration of the U.S. Department of Energy under contract DE-AC52-06NA25396. Los Alamos National Laboratory strongly supports academic freedom and a researcher's right to publish; as an institution, however, the Laboratory does not endorse the viewpoint of a publication or guarantee its technical correctness.

RESEARCH LIBRARY

INTERFACE FOCUS

rsfs.royalsocietypublishing.org

Review

Cite this article: Castro M, Lythe G, Molina-París C, Ribeiro RM. 2016 Mathematics in modern immunology. *Interface Focus* **6**: 20150093.
http://dx.doi.org/10.1098/rsfs.2015.0093

One contribution of 12 to a theme issue 'The Human Physiome: a necessary key to the creative destruction of medicine'.

Subject Areas:
computational biology, biomathematics

Keywords:
mathematical modelling, immunology, T cell, two-photon microscopy, T-cell receptor, diversity

Author for correspondence:
Carmen Molina-París
e-mail: carmen@maths.leeds.ac.uk

THE ROYAL SOCIETY PUBLISHING

Mathematics in modern immunology

Mario Castro[1], Grant Lythe[2], Carmen Molina-París[2] and Ruy M. Ribeiro[3]

[1]Universidad Pontificia Comillas, E28015 Madrid, Spain
[2]Department of Applied Mathematics, School of Mathematics, University of Leeds, Leeds LS2 9JT, UK
[3]Los Alamos National Laboratory, Theoretical Biology and Biophysics, Los Alamos, NM 87545, USA

Mathematical and statistical methods enable multidisciplinary approaches that catalyse discovery. Together with experimental methods, they identify key hypotheses, define measurable observables and reconcile disparate results. We collect a representative sample of studies in T-cell biology that illustrate the benefits of modelling–experimental collaborations and that have proven valuable or even groundbreaking. We conclude that it is possible to find excellent examples of synergy between mathematical modelling and experiment in immunology, which have brought significant insight that would not be available without these collaborations, but that much remains to be discovered.

1. Introduction

Mathematics has a long tradition in biology and medicine, going back at least to Gregor Mendel's work in genetics and Theodor Boveri's work on the nature of the chromosomes [1]. However, as the various subfields have become more specialized, and understanding of biological systems more detailed, modelling has often been dismissed or 'regarded with suspicion, partly because much modelling seems an empty intellectual exercise that fails to deliver biological insight' [2]. In immunology, the situation is not much different [3]. Some commonly cited reasons for the separation between experimental immunology and mathematical modelling are (i) the pace of discovery of new agents and new phenomena in the immune system, accompanied by new jargon, (ii) the rapid advance of technology, producing ever more data, and (iii) the contrast in academic environments, culture and terminology. In our view, these issues (at least the first two) are in fact reasons why mathematical modelling will become increasingly important in immunology (and other fields of biology) [4]. It is precisely because intuition is insufficient beyond a certain level of complexity that analysis of the immune system must become more quantitative [5–7].

As experimental data expose the complexity of the immune system, its nonlinearities [3], feedbacks [8], combinatorial complexity [9] and redundancies [10], it becomes ever more difficult to understand from a purely descriptive and qualitative viewpoint. On the other hand, big data and the so-called information explosion [11] have attracted attention as they promise automatic (computerized) solutions to biological and medical problems. The strategy behind this approach, referred to as *data-driven modelling* [12] or reverse modelling [13], is the use of statistical data analysis with the aim of creating novel knowledge. However, statistical analysis is only one of the many tools that mathematics can provide to immunology and, in general, cannot offer mechanistic explanations. Nor are computational models that include every known agent and interaction sufficiently well developed, yet, to replace animal testing. Current models, using differential equations, the theory of stochastic processes and agent-based simulations [14], are 'often derived from phenomenology and guess work' [13]. Their first benefit is that building the model forces the modeller to lay out all assumptions clearly. Thus, a mathematical model is an accurate description of current, incomplete, thinking, not a complete or accurate description of Nature itself.

Immunology is an excellent field for the application of mathematics, because it has a long tradition of important theories (clonal selection,

© 2016 The Authors. Published by the Royal Society under the terms of the Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/, which permits unrestricted use, provided the original author and source are credited.

immune networks, danger signals, ...) and great thinkers [15]. Still, there are impediments to a marriage of mathematics and immunology [3]. One is that the technical languages of the disciplines are very different. It would be useful for experimental immunologists to have some basic understanding of mathematical modelling, its usefulness and limitations; it is also crucial that modellers learn the required immunology and the experimental systems they are modelling. That is, we have to overcome the third problem identified above. On the other hand, there are examples of successful collaborations between experimental and mathematical immunologists. Often, but not exclusively, these partnerships are more fruitful when collaborations are initiated at an early stage of a given study.

Our aim here is not to review the impact of theoretical ideas in immunology, nor to provide a mathematical tutorial. We instead emphasize the role that mathematical (or computational) modelling has already played in immunology. We present examples from the literature of the contribution of mathematics to T-cell immunology under four headings: (i) generation of hypotheses, (ii) quantification of immunological processes, (iii) definition of observables to measure given an experimental objective, and (iv) reconciling disparate (or even conflicting) experimental results.

With our examples, we seek to illustrate the power that, together, mathematical and experimental immunology have to elucidate the behaviour of one of the most intricate systems that evolution has produced. To Dobzhansky's [16] statement that 'nothing in biology makes sense except in the light of evolution', we may add that nothing can be expressed quantitatively in science except in the light of mathematics.

2. Generating hypotheses

2.1. T-cell movement: directed or random?

T cells interact in lymph nodes (LNs) with antigen presenting cells (APCs) that display peptide–MHC complexes (small fragments of proteins bound to MHC molecules) on their surface. From the point of view of a naive T cell that spends a day or less in an LN, searching for an APC with a few cognate peptide–MHC complexes is a 'needle in a haystack' problem [17,18]. Based on twentieth-century understanding of cell types and the means by which cells communicate and home to different organs [17], derived from *in vitro* studies and snapshot images, the natural hypothesis was that, in the LNs, T cells are guided to their target APCs by chemokine-derived signals.

A new hypothesis emerged from three papers, published together in 2002, where Miller *et al.* [19], Stoll *et al.* [20] and Bousso *et al.* [21] reported *in vivo* imaging studies (using two-photon microscopy) in which individual T and B cells were resolved and tracked: the first direct *in situ* observations of their movement and interaction with dendritic cells (DCs). Movies were produced by stacking two-dimensional slices into three-dimensional images, then into series of three-dimensional images taken at regular time intervals, a minute or less apart. In the movies, 'rather than directed motion in response to sustained chemokine gradients along structural pathways', T-cell motion is seen to be rapid, and described as 'meandering', 'chaotic', 'random' and 'frantic' [22,23]. The data available are the estimated positions of the centres of a set of labelled cells in three-dimensional space at a series of times: $x_i(t)$, $y_i(t)$, $z_i(t)$, where i labels the cell track and t labels time: $t = 0, \Delta t, 2\Delta t, \ldots$, with Δt, the time interval between images. It is well known from modelling of, for example, Brownian motion that, if the movement is indeed random, the means of $x_i(t)-x_i(0)$, $y_i(t)-y_i(0)$ and $z_i(t)-z_i(0)$ are zero, but the mean of the displacement from the initial position,

$$r_i(t) = ((x_i(t) - x_i(0))^2 + (y_i(t) - y_i(0))^2 + (z_i(t) - z_i(0))^2)^{1/2},$$

is not vanishing. When Miller *et al.* plotted the mean displacement against the square root of time, points fell 'on straight lines, consistent with the hypothesis of random movement; the slope is the motility coefficient M' [19,24]. Thus, in an opinion article in 2003, Cahalan *et al.* [25] argued for the new hypothesis 'that lymphocytes migrate randomly within lymphoid organs, and that lymphocyte contact with APCs may be a stochastic process rather than one guided by chemokine gradient'. The 'key point is that vigorous motility in a random walk characterises the behaviour of naive T cells *in vivo*'.

However, the situation is not so simple, because there can be subtle biases in cell tracks pieced together from two-photon imaging data [26], which are limited in both space and time. For example, directed movement in an environment full of physical obstacles (other cells and the endothelial reticular network) could still look approximately random. The best way to analyse these biases is to use mathematical models, which not only provide qualitative clues about the underlying mechanisms but, more importantly, in those cases where the true answer lies somewhere in the middle of two or more competing hypotheses, can describe quantitatively the way in which those different mechanisms are combined in the experiment, and identify additional experiments (for instance, inhibiting one mechanism) to test the validity of the model and gain a deeper biological insight.

Indeed, departures from a straight line in plots of mean displacement versus the square root of time are seen at sufficiently short and long times, indicating possible persistence in T-cell motion and an effect of confinement and finite imaging volumes. To analyse cell tracks in more detail and to compare them with theoretical predictions from random versus directed motion, the dependence on Δt of the distribution of the apparent speed, $v_i(t) = [r_i(t + \Delta t) - r_i(t)]/\Delta t$, and the distribution of turning angles can be used [27–32]. These analyses have been applied to studies of the role of the LN reticular network [33–36], of the effect of localized chemotactic migration [37], and to extrapolate from the high density of labelled cells necessary for two-photon imaging to lower, more physiological, cell densities [38]. For example, careful mixed effects analysis is required to identify a small directed component of the motion of skin-resident T cells [39–41] towards a site of infection.

Not all T cells in LNs crawl vigorously; also noted in the early imaging studies were apparently stationary cells, clusters of cells and swarms of enlarged T cells that repeatedly come into contact with DCs bearing cognate antigen [24]. Bousso & Robey [42] estimated that seven new T cells per minute came into close contact with each individual DC, and extrapolated to conclude that each DC can contact about 500 different T cells in an hour. The description of Mempel *et al.* [43], in which T-cell priming by DCs in LNs occurs in three phases, has been very influential [22]. In the first of these phases, T cells contact antigen-bearing DCs

intermittently and briefly. T-cell–DC contact durations then increase in the second phase and, in the third phase, T cells resume their motility, swarm in the local vicinity, begin the process of cell division and initiate interaction with B cells.

More sophisticated models of cell motion in the LN (such as those based on cellular automata, where cells move only according to stochastic local rules) can include these different phases, and try to use experimental data to calibrate the rates of T-cell interactions with other cells [36]. In yet more advanced computational models [44,45], called *cellular Potts* models [46,47], T cells are simulated as extended entities, that is, cells have shape (which can change over time) and sites that contact other cells, and occupy a physical space determined by their flexible cellular membrane. Each cell is assumed to minimize its surface energy. This gives rise to changes in the cell's configuration and to its movement over the course of time [48]. Extrapolating from experimental data, these models predict that an individual T cell can scan up to a hundred different DCs per hour [44].

These more complicated models allow detailed incorporation of random and directed components of the movement, with different hypotheses and taking into account the complexity of the LN environment. Moreover, they not only simulate the dynamics of cells in the LN, but one can also simulate the experimental measurements, including the effects of limited time and space resolution to clarify our interpretation of the experimental data. For example, it is possible to track single cells in a two-dimensional plane (or stack of two-dimensional planes) for a limited time, and plot their displacement. Thus, in this field of immunology, we find an example of a hierarchy of models, from simple functions to complex computer simulations, that provide more detailed analyses and insights into valuable experimental data and alternative hypotheses to explain T-cell migration in LNs. It is likely that both directed and random components are involved in T-cell movement, and this could depend on the type or state of the T cell. The final verdict is still unknown. Clearly, further rounds of observation–modelling–prediction–observation are needed to advance our understanding of how T cells move and interact with other cells, and how these interactions allow the specific pMHC–TCR interactions that initiate adaptive immune responses [49–51].

2.2. T-cell activation: how are signals integrated?

Adaptive immune responses to infection mediated by $CD8^+$ T cells are characterized by the following time course: a subset of $CD8^+$ T cells, those that are pathogen-specific, undergoes rapid proliferation, expanding the cell numbers of the given subsets, followed by population contraction owing to cell death and a return to homeostasis of the $CD8^+$ T-cell population. T-cell activation and proliferation require a number of receptor-mediated signals. How are $CD8^+$ T-cell stimulatory signals integrated and how do they relate to the strength of the T-cell response? A recent joint experimental and modelling effort by Marchingo *et al.* [52] has provided some answers, whereas previous efforts had been limited to qualitative approaches that do not predict the quantitative effect of altering a combination of stimulatory signals.

The current paradigm is that T cells respond to three signals: signal 1 (or σ_1) is received by the T-cell receptor (TCR) upon ligand binding to peptide–MHC complexes, signal 2 (or σ_2) is co-stimulatory and provided by APCs that express ligands, CD80 and CD86, that can bind co-stimulatory receptors (CD28 and CTLA-4) on the surface of T cells, and signal 3 (or σ_3) is provided by cytokines (such as interleukin (IL)-2, IL-4 or IL-7). As mentioned above, there is an important difference between a qualitative hypothesis (*the three signals are needed*) and a quantitative/mathematical one (*which mathematical function describes the relative weight of these three signals*).

Previous work on B cells by the same authors led them to conclude that B-cell immune responses involve an automated programme of proliferation events, characterized by the number of total divisions and a return to cell quiescence [53–55]. These experimental studies also showed that the number of division events was influenced by the strength of B-cell stimulation. They hypothesized a similar programme for T cells, namely the final number of divisions during a $CD8^+$ T-cell-mediated immune response is a (linear) function of the sum of signals, σ_1, σ_2 and σ_3. The relevance of this hypothesis, if true, is the potential to predict therapeutic strategies of immunomodulation. Furthermore, the hypothesis naturally leads to the development of a quantitative (mathematical and computational) framework to relate the signals received by a given T cell to its behaviour during an immune response as a function of time.

Marchingo *et al.* [52] used the concept of *division destiny* (DD) of a T cell: the number of rounds of division before the cell returns to quiescence or dies. This concept was introduced in a paper describing the 'cyton model' of lymphocyte proliferation [53]. In this model, each cell has autonomous division and death processes in competition, corresponding (mathematically) to countdown clocks for these two fates (division or death). The start time for each clock is an independent random variable. The outcome for the cell is determined by whichever clock reaches zero earlier. DD is the maximum number of divisions allowed in the model [52,53].

A series of experiments allowed them to analyse how DD varies with different combinations of signals and to calibrate the cyton model to compute DDs [53,56,57]. Additionally, an *in vivo* experiment enabled them to test their hypothesis with a set of untrained data. Putting all these together, the authors then concluded that, both *in vitro* and *in vivo*, $CD8^+$ T-cell-mediated immune responses imply two stages of regulation of T-cell DD. During the first stage, σ_1, σ_2 and σ_3 additively programme a heritable number of division rounds before the first cell division. The second stage, which involves external stimuli, such as cytokines, can further increase the DD of a given cell.

The experimental approach of Marchingo *et al.* [52] makes use of transgenic mice with $CD8^+$ T cells that only express one TCR, the OT-I receptor, crossed with fluorescent ubiquitination-based cell-cycle indicator red–green (FucciRG) reporter mice. The cell cycle is tightly regulated by specific proteins that are produced and degraded in a highly controlled fashion. The FucciRG reporter was developed by fusing red and green fluorescent proteins to two cell-cycle regulators, which are expressed at different times [58,59]. (More specifically, the reporter is based on the fact that the cell cycle is not only regulated at the transcriptional and post-translational levels, but also controlled by ubiquitin-mediated proteolysis. The authors harnessed the regulation of cell-cycle-dependent ubiquitination in order to develop two genetically encoded indicators for cell-cycle progression [58], hence the name, fluorescent ubiquitination-based cell-

cycle indicator.) In this way, cells with this reporter (including the $CD8^+$ T cells of the mice used) fluoresce red if they are in the G_0/G_1 phase, or green if they are in the $S/G_2/M$ phase of the cell cycle, and the cells can be classified as quiescent, recently divided or actively dividing. We should emphasize that the development of this reporter construct was a great technical achievement owing to the heroic efforts of Sakaue-Sawano *et al.* [58], opening doors for beautiful and important experiments [60]. Indeed, this is a very good example where new technology allows new insights into the biology, driving hypothesis formulation and model development.

Marchingo *et al.* [52] used the cyton model to fit data from FucciRG mice, allowing them to conclude that DD is around 10 generations (or divisions), yet different T-cell families differ in numbers, with differences up to three orders of magnitude.[1]

A second experimental model is an *in vitro* cell-trace-violet (CTV) division-tracking assay, to allow further exploration of the division progression. With this *in vitro* experimental system, the authors stimulated CTV-labelled OT-I-specific $CD8^+$ T cells with a combination of signals σ_2 and σ_3 (see fig. 2 in [52]), and carried out a titration study for cytokine IL-2 (human). Again, fitting the data to the cyton model enables the authors to estimate the average DD of the population, to obtain a beautiful time course and dose–response curves based on the linearity hypothesis (Marchingo *et al.* [52] fig. 2*c*,*d*). Their *in vitro* CTV data support the hypothesis that summation of DD from multiple co-stimuli geometrically amplifies the T-cell response [52]. With these two sets of experiments, the authors make two key predictions. The first one is the additive nature of the T-cell stimuli *in vivo*, and the second one is that the second stage, which allows the extension of the DD of a given cell, will be more relevant away from the site of priming.

To test these predictions, the authors carried out a final set of experiments that involved co-transfer of T cells to mice infected with two different recombinant influenza viruses. One set of mice was infected with a virus expressing the low-affinity ovalbumin peptide Q_4 (low affinity for the OT-I TCR expressed by the transferred cells), and a second set of mice was infected with an influenza virus expressing the high-affinity peptide N_4 (high affinity for the OT-I TCR). The co-transfers involved two kinds of cells: wild-type (WT) OT-I $CD8^+$ T cells and OT-I $CD8^+$ T cells lacking the high-affinity IL-2 receptor (IL-2Rα or CD25) (mutant), so that they cannot receive IL-2-derived cytokine signals. Thus, in total, the experiments generated four sets of data: (i) high TCR signal (N_4 peptide) and high IL-2R signal (WT cells), (ii) high TCR signal (N_4 peptide) and low IL-2R signal (mutant cells), (iii) low TCR signal (Q_4 peptide) and high IL-2R signal (WT cells), and (iv) low TCR signal (Q_4 peptide) and low IL-2R signal (mutant cells). The authors fitted the cyton model to data from (ii), (iii) and (iv) to predict the effect of each signal (in this case, σ_1 and σ_3) on the distribution of DD for the experimental conditions from (i). Based on that distribution, the authors were able to compute/predict the dynamics of the T-cell population over time, which showed good agreement with the experimental data.

This is one of the few (thus far) examples in which experiments and modelling have gone hand-in-hand and completed the full circle of: (i) experimentation, (ii) hypothesis generation, (iii) further experimentation, (iv) model development, (v) model prediction, and (vi) model validation with a final and independent set of experiments. The relevance of this study is multiple: first, it has managed to resolve long-standing discrepancies between *in vitro* and *in vivo* systems; second, it has provided a quantitative and predictive framework that underpins the time course of $CD8^+$ T-cell responses for a combination of stimuli. Finally, this quantitative framework will allow further *in silico* testing for combinations of stimuli that might not be experimentally feasible now, yet relevant to the development of future immunotherapeutic strategies.

3. Defining what to measure

3.1. T-cell receptor excision circles to measure thymic output

To understand healthy homeostasis and disease dysregulation of the peripheral T-cell compartment, it is important to quantify thymic output [61]. However, there is no simple way to perform this quantification experimentally in humans (or mice). Ideally, recent thymic emigrants (RTEs) would express some surface protein with a short half-life that could be measured. This knowledge, together with the levels of cells expressing that protein in the periphery, would allow the calculation of thymic export, assuming RTEs did not proliferate, at least while expressing the identifier protein. This situation is the case in chicken, where the chT1 antigen on thymic T-cell precursors seems to fulfil this role [62]. In mammals, there is no such convenient marker, but it was proposed that one could use T-cell receptor excision circles (TRECs) to quantify thymic output [63,64].

TRECs are episomal DNA circles that are a by-product of TCR rearrangement. During T-cell development in the thymus, the α chain of the TCR is rearranged by recombination of its subcomponents (the variable and joining regions) on chromosome 14, which involves excision of the δ locus that resides between the V and J segment genes of the α chain on that chromosome. The excised DNA circularizes and contains specific sequences (so-called signal-joint and coding-joint) that are common on 70% of $\alpha\beta$ T cells [63,64]. These TRECs have properties that make them good candidates for markers of RTEs with which one could quantify thymic output. In particular, they are stable [64,65], they do not divide [63], they are of exclusive thymic origin [63,64,66], they are identical in 70% of $\alpha\beta$ T cells [67], and they mark RTEs in chicken [64].

Several groups used TRECs to quantify thymic output, especially in the context of HIV infection, because this infection leads to declining T-cell numbers and it was not clear what the contribution of a lower thymic output was [63,68,69]. In these studies, TRECs were reported as frequency of $TREC^+$ cells among all $CD4^+$ T cells (or, sometimes, naive $CD4^+$ T cells or peripheral blood mononuclear cells (PBMCs)). It was found that TREC frequency decreases with age [63,68], that it is lower in HIV-infected individuals [63,68], although for many this frequency was in the normal range [68], and that it increases significantly in antiretroviral-treated HIV^+ individuals, who have a lower than normal TREC frequency [63,68]. Another early study indicated that TREC frequency was a predictor, beyond viral load and $CD4^+$ T-cell count, of the rate of

HIV-1 disease progression [69]. These studies led to somewhat different conclusions, with some stating that RTEs and thymic output, as measured by TREC frequency, had an important role in HIV disease [63,69], and others concluding that 'it is difficult to invoke thymic regenerative failure as a generalised mechanism for the depletion of $CD4^+$ lymphocytes' and 'we therefore caution against the use of such phenotypic markers as a direct indicator of thymic output' [68]. This initial research was followed by a flurry of studies using TRECs to quantify thymic output [61,70–72].

One problem with the interpretation of TREC frequency is that it can vary owing to changes in both cells with TRECs (the numerator) and cells without TRECs (in the denominator). Thus, other dynamical processes besides thymic input into the periphery, such as T-cell proliferation or death, could influence the observed changes in TREC frequency. For example, T-cell proliferation of non-TREC-containing cells would lead to a decrease in TREC frequency, and it is well known that T-cell proliferation is increased in the context of HIV infection (see below). This problem was recognized and explicitly mentioned by the initial studies [63,64,68], and the assumptions about the importance of this issue affected the conclusions drawn by those authors.

This is a case where modelling can help clarify the interpretation of the data by identifying what to measure, and it can be argued that one must use a model to properly analyse the data. Exactly this was proposed in a seminal study, where a simple mathematical model was used to interpret TREC frequency data [73]. The model equations for TREC-containing cells (C) and total T cells (T) are [73]

$$\frac{dC}{dt} = \alpha\sigma - dC \tag{3.1}$$

and

$$\frac{dT}{dt} = \sigma + (p - d)T, \tag{3.2}$$

where σ is the output of the thymus, α is the fraction of cells exiting the thymus that contain a TREC and p and d are the rates of proliferation and death of T cells, respectively. The ordinary differential equations (3.1) and (3.2) are linear and, thus, most simple, yet allow one to explain the experimental data. In particular, with these assumptions for the dynamics of T cells, TREC frequency ($F = C/T$) is given by the division rule of differentiation and one obtains

$$\frac{dF}{dt} = \alpha\frac{\sigma}{T} - \left(p - \frac{\sigma}{T}\right)F. \tag{3.3}$$

Equation (3.3) shows that the dynamics of TREC frequency depends explicitly on thymic output and also on the proliferation rate p, and cannot simply be interpreted directly as thymic output [74]. Hazenberg *et al.* [73] went on to analyse the predicted dynamics in the context of ageing, HIV infection and HIV treatment, and concluded that 'loss of TREC in HIV-1 infection may primarily be caused by continuous hyper-activation of the immune system as reflected by increased cell division' and 'measurements of TREC content (…) fail to provide experimental evidence for thymic impairment'.

The contribution from the mathematical model was to clarify the assumptions made when thinking of TREC frequency as a direct measurement of thymic output, demonstrating that such interpretation of the data was abusive. However, the mathematical model also suggests that proper assessment of thymic output may be possible using TREC concentration (i.e. $TREC^+$ cells per microlitre of blood), which is described by equation (3.1). This is clear from inspection of the equation, but was not realized until models of the system were developed. Measuring TREC per microlitre became an important variable in studies trying to understand the effects of thymic output [75–79], and also a relevant issue for model development [74,80–82]. In particular, equation (3.1) states that total abrogation of thymic output should result in an exponential decay in TREC concentration, and analyses of such data would allow estimation of the parameter d (the slope of that decay). Moreover, if we assume that TREC concentration is at an approximate steady state (C_0) before abrogation of thymic output, then one can estimate thymic output as $\alpha\sigma = dC_0$. Thus, the model suggests an experiment of thymectomy (removal of the thymus). This experiment was performed on a macaque model of HIV infection, allowing direct quantification of the number of $CD4^+$ and $CD8^+$ T cells exported from the thymus (of rhesus macaques) per day [83].

The study of TRECs shows how development of novel experimental and technical capabilities can be leveraged even with simple mathematical models. It is clear that the model presented in (3.1)–(3.3) is a simplified description of complicated biological phenomena, and yet this approach was highly valuable to properly interpret new datasets.

4. Quantifying immunological processes

4.1. Estimating T-cell receptor diversity

The number of different TCRs that could be produced by the processes of DNA rearrangement that occur in the thymus is enormous, far greater than the total number of cells in a human body [84]. A first question is: how many different TCRs are found at any one time in one mouse or human? Let us denote this quantity by $N(t)$. With the slight simplification that all the TCRs on the surface of any one T cell are identical, the T-cell repertoire can be divided into clonotypes, each clonotype characterized by a unique TCR. If $n_i(t)$, $i = 1, \ldots, N(t)$ is the number of T cells expressing TCR of type i at time t, then we can pose a second question: what is the distribution of values of $n_i(t)$?

The TCR is formed in the thymus by pairing a larger β chain and a smaller α chain. The DNA subset that encodes the β chain, TRB, is formed by rearranging V, D and J gene segments, plus additional nucleotide rearrangements at V–D and D–J junctions. TCRA, which encodes the α chain, rearranges with V and J only. Assuming that the TCR nucleotide sequence provides a unique molecular tag for each clonotype, the total number of distinct TCR sequences in a body is equal to $N(t)$. A new field of immunosequencing has emerged with technologies designed to sequence TCRs [85]. Millions of TCR sequences can be amplified in a single multiplex polymerase chain reaction (PCR), prepared and then read in parallel from a single sample. The distribution of VJ gene segment usage can be measured with flow cytometry [86,87], and used to track changes with age and between individuals [88,89].

In 1999, Arstila *et al.* [90] published a first estimate of the number of distinct TCRβ chains in a sample of 10^8 T cells from a healthy donor. By sequencing only DNA from cells that use VB18 (estimated to be 0.8% of T cells), VJ1.4 (estimated to be 3% of T cells) and with a CDR3 length of 12 amino acids (estimated to be 9.3% of T cells), they were able to identify 17

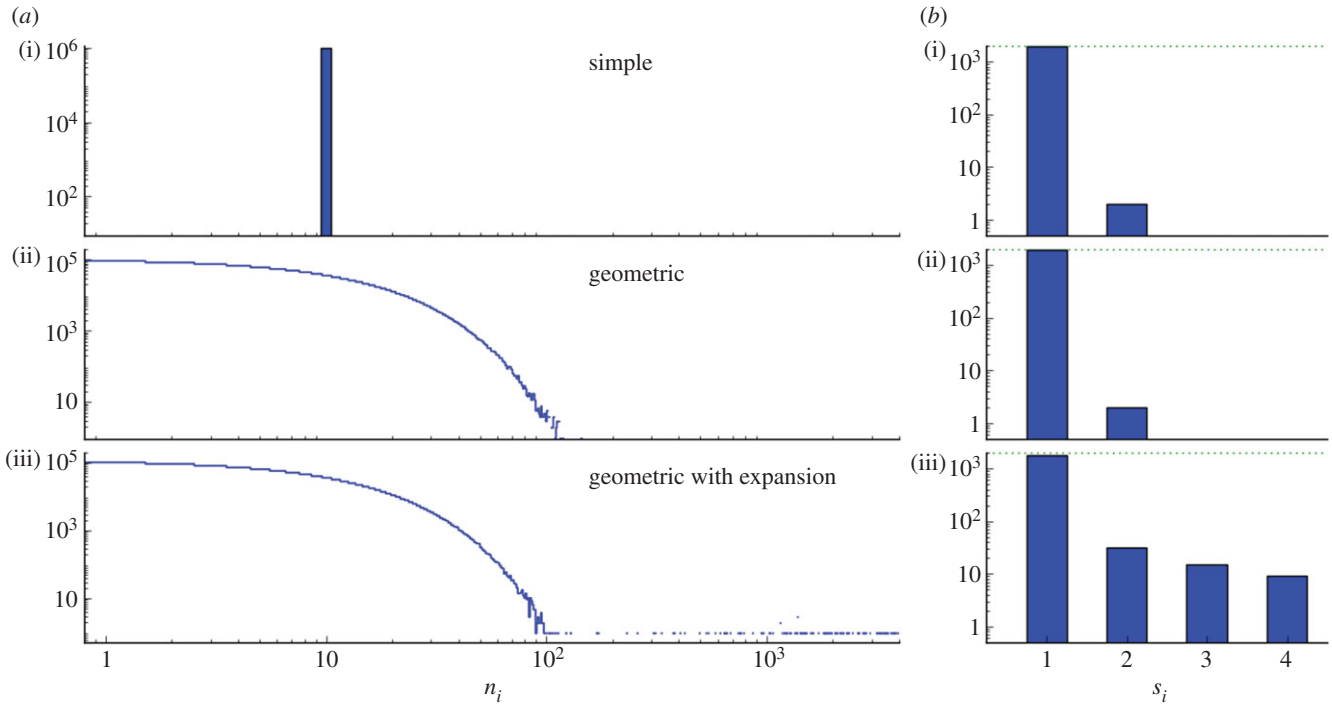

Figure 1. Simulated sampling from repertoires. In each of the three simulated examples shown, a sample of 2000 cells is selected from a repertoire that contains 10 million cells divided into 1 million different clonotypes. (*a*) The three examples have different distributions of clonal sizes, n_i. In the first, every clonotype consists of 10 cells (*a*(i)). In the second, the number of cells per clone is drawn from a geometric distribution with mean 10 cells (*a*(ii)). The third example is similar to the second, but 200 'expanded' clonotypes have sizes drawn from a geometric distribution with mean 1000 cells (*a*(iii)). (*b*) Histograms of the samples (of 2000 cells) from each repertoire. The number of cells, s_i, of a clonotype, *i*, that is present in the sample is most often equal to 1. Only in (*b*(iii)), where there are expanded clonotypes in the repertoire, are there more than two cells of some clonotypes in the sample. (Online version in colour.)

different TCRβ chains. Extrapolating, they concluded that the blood sample contained $17/(0.008 \times 0.03 \times 0.093) = 0.8 \times 10^6$ distinct TCRβ chains. Fifteen sequences from the same sample, restricted to VB16, VJ2.2 and CDR3 length of 13 amino acids, gave a similar estimate of TCRβ diversity, close to 10^6. Subsequent studies have tended to restrict themselves to analysing the β chain, believed to contain the majority of TCR diversity. The estimate of 10^6 TCRβ chains in human blood, first obtained by audacious extrapolation, has been backed up by more recent studies [91].

In 2000, Casrouge *et al.* [92] analysed mouse spleens [93]. The same mathematical extrapolation as that just described, from the subset of VB10, VJ1.2 cells with 10-amino-acid-long CDR3, yielded an estimate of half a million distinct TCRβ chains. They also estimated that the combination with the TCRα chain contributes a factor of more than two to the diversity, thus concluding that a mouse spleen contains about 10^6 clones of about 10 cells each. In mice, different T-cell subsets can be compared, and the effects of infections and immunization on the repertoire can be tracked [94–97]. Current single-cell approaches [98] are able to provide information about the promiscuity of the $\alpha\beta$ pairing, which is needed to establish the true extent of TCR clonal diversity.

Freeman *et al.* [99] identified 33 664 distinct TCRβ sequences in a pool of cells from 550 people. Recent studies, which were able to directly count large numbers of sequences, have identified 10^6 distinct TCRβ chains in a sample of blood from one person. It is possible to use ecological 'missing species' analyses [100–102] to estimate the total number of distinct TCRβ chains in the blood. Estimates range from a few million to as high as 100 million [100,101,103]. Li *et al.* [104] isolated PBMCs and extracted genomic DNA from 10 ml of blood from two rhesus monkeys. They identified nearly a quarter of a million unique CDR3 amino acid sequences in the sample.

Counting experimentally the total number of distinct TCRs in a body may never be possible, but even if it becomes possible, it will not fully characterize the T-cell repertoire, i.e. the number of cells of each clone. In fact, (deep) sequencing does not provide a direct measurement of the distribution of values of n_i, because the number of times a sequence is found depends on mRNA copy numbers inside cells and on biases in PCR amplification [105]. Sequencing one cell at a time is an attractive alternative, but is currently limited in practice to a few hundred cells per experiment. Computational models, on the other hand, are rapid and cheap and may provide the only alternative to properly quantify the distribution of clonotype sizes. Here, at least two approaches are possible: (i) using statistical models to compare distributions and define observables and (ii) developing mechanistic models of cell division and death to establish what parameters influence the steady-state distribution of clonotype sizes. For example, in the first approach, we could define three ways of allocating 10^7 cells (of a mouse) into the estimated 10^6 different clonotypes [92], as illustrated in figure 1. We could then take samples of 2000 cells from each distribution, to replicate an experimental protocol, and compare the resulting histograms from experimental data and the computational model. In the case shown, most of the TCRs that are present in a sample are found only once. The presence of two or more cells of the same TCR clonotype in such a sample is evidence that there are clonotypes with large numbers of cells in the repertoire (e.g. expanded in response to antigen).

Whatever statistically stationary set of $n_i(t)$ is compatible with human (or mouse) experiments results from a balance between T-cell generation (in the thymus), division and

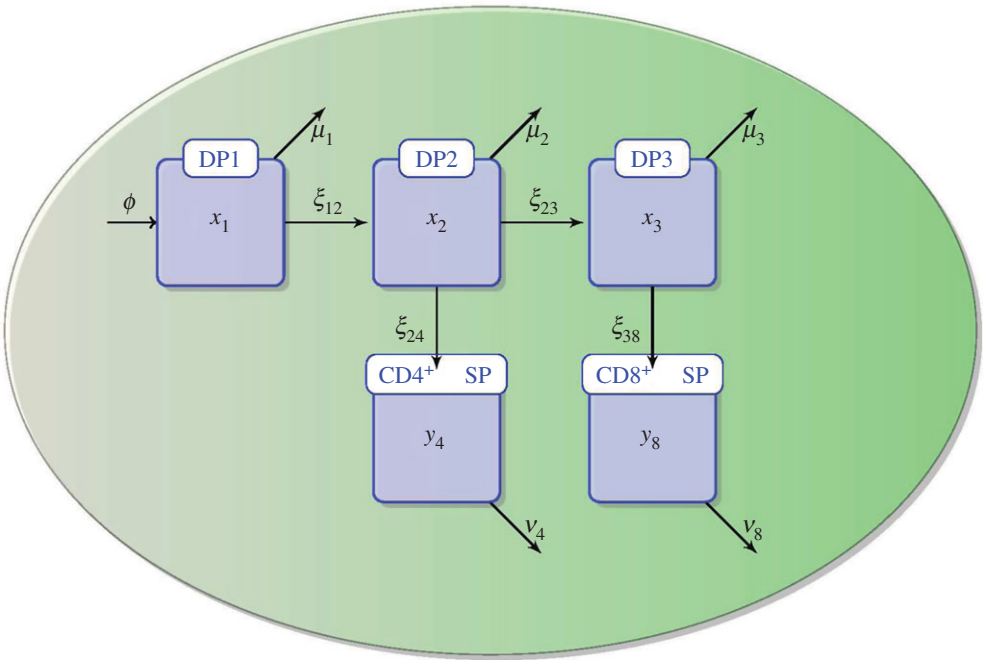

Figure 2. Mathematical model of thymocyte development as introduced and developed in Sinclair *et al.* [111]. Thymocytes enter the DP1 stage from DN cells with influx ϕ. At the DP1 stage, thymocytes can either die with rate μ_1 or differentiate to the DP2 stage with rate ξ_{12}. At the DP2 stage, thymocytes can either die with rate μ_2, or differentiate to the DP3 stage with rate ξ_{23}, or to the SP CD4$^+$ stage with rate ξ_{24}. DP3 thymocytes can either die with rate μ_3 or differentiate to the SP CD8$^+$ stage with rate ξ_{38}. CD4$^+$ SP thymocytes die or exit to the periphery with a total loss rate ν_4, and CD8$^+$ SP do so with rate ν_8. (Online version in colour.)

death. Thus, mathematical models of the dynamics of the T-cell population, in which the $n_i(t)$ are functions of time, governed by thymic production of new T-cell clonotypes, and division and death of existing cells in the periphery have been proposed [106–110]. Often these models simulate ecological competition within and among clones for TCR signals and other resources (e.g. IL7) to study what factors are more important in determining the distribution of T-cell clones. Moreover, using estimates for the basic processes (of thymic generation, division and death) from human and mice, the models allow comparison of predictions of the distribution of clonotypes in those two cases. One such model predicts a much larger number of clonotypes than estimated experimentally so far [110]. The model could be used to produce figures similar to figure 1 to start comparing the predictions with the data, and if needed the model could be refined to better describe experimental data. In summary, as well as requiring statistical models to deduce the size of the T-cell repertoire from experimental measurements, mathematical modelling is needed to understand how homeostasis of that repertoire is attained and maintained.

4.2. Quantifying thymic development of T cells

A beautiful example of how experimental data together with a mathematical model allows us to quantify immunological processes is provided in [111]. This paper provides an excellent lesson on the role that modelling can play in quantifying immunological process; in doing so, new light is shed into the mechanisms that are at the heart of the given immunological process.

Cells of the mature $\alpha\beta$ T-cell repertoire arise from the development in the thymus of bone marrow precursors (or thymocytes). $\alpha\beta$ T-cell maturation is characterized by the expression of thousands of copies of identical $\alpha\beta$ TCRs and the CD4 and/or CD8 co-receptors on the surface of thymocytes. The overall maturation stages of a thymocyte are (i) double negative (DN) (TCR$^-$, CD4$^-$ and CD8$^-$), (ii) double positive (DP) (TCR$^+$, CD4$^+$ and CD8$^+$), and (iii) single positive (SP) (TCR$^+$, CD4$^+$ or CD8$^+$). Thymic APCs provide the appropriate microarchitecture for the maturation of thymocytes, which 'sense' the signalling environment via their randomly generated TCRs. Naive CD4$^+$ and CD8$^+$ T cells that populate the immune system derive, thus, from the SP CD4$^+$ and CD8$^+$ thymocytes, respectively, that were successfully selected and exited the thymus into the periphery. The observed ratio of naive CD4$^+$ to CD8$^+$ T cells is about 2 : 1; that is, there are more CD4$^+$ helper T cells than CD8$^+$ cytotoxic T cells in the peripheral pool [81,112]. A long-standing immunological question is the origin of this asymmetry in CD4$^+$ and CD8$^+$ T-cell numbers and, in this paper, the authors set out to study thymic development in search for clues to this bias [111].

Their approach combines a carefully designed experimental study and a mathematical model that allows them to quantify differentiation and death rates of different subsets of thymocytes during their development (figure 2). Their results are not only the estimation of these parameters, but also the ability to discriminate between two hypotheses for the observed asymmetry in CD4$^+$ and CD8$^+$ T-cell numbers. The authors set out to answer the question: what is the source of the asymmetry between the CD4$^+$ and CD8$^+$ T-cell lineages during thymic development? The experimental study is based on their ability to synchronize cohorts of thymocytes *in vivo* to follow their kinetics as they develop and are being selected. The mathematical model is based on a set of deterministic ordinary differential equations that describe the dynamics of the thymocyte populations from the DP stage (figure 2). The model includes five populations (DP1, DP2, DP3, CD4$^+$ SP and CD8$^+$ SP), where the DP population is subdivided into three distinct temporal stages of development (DP1, DP2 and DP3), which had previously been experimentally identified based on the surface

expression of TCRs and CD5 (see fig. 1 in [111,113]). The model also includes the rates of maturation from one stage to the next (ξ_{12}, ξ_{23}, ξ_{24} and ξ_{38}), and loss (ν_4 and ν_8) or death (μ_1, μ_2 and μ_3). Loss is used to jointly describe death or migration in the SP thymocyte compartments. A migration term is included, ϕ, to represent the influx of cells into the DP1 stage from DN thymocytes. Finally, the mathematical model does not include cell proliferation as recent estimates suggest that there is no proliferation during the DP stage, and at most one round of proliferation during the SP stage [111,114].

Quantifying and deciphering the mechanisms/sources of asymmetry is a big challenge! At least two different hypotheses can explain the thymic origin of the bias:

— at the DP1 stage, there is an asymmetry in the number of DP thymocytes with a TCR that can bind to MHC class I versus MHC class II, or
— during the DP or SP stage, there exists some asymmetry in either maturation or loss processes (migration to the periphery or death).

Three sets of experiments are carried out to identify the source of the CD4/CD8 bias [111]. The first (or control) experimental assay arrests thymic maturation at the DP stage to then switch it on, and thus T-cell development can be followed during 10 days in a synchronized way. Cell numbers were analysed with flow cytometry at days 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 and 10. The second experimental assay is exactly as the first, but with MHC class-I-deficient mice, so that only $CD4^+$ T cells can be selected. Finally, the third experimental assay is performed, with MHC class-II-deficient mice, so that only $CD8^+$ T cells can be selected. In this way, the authors can simultaneously and independently study the development of $CD4^+$ and $CD8^+$ T cells. That is, the experiments with deficient mice allow them to understand lineage-specific maturation and death rates from the DP2 stage under conditions of isolated development.

Given the experimental data and the mathematical model, the authors obtain estimates for the parameters of the model. The most striking differences appear in the DP2 compartment. For example, for the control experiments, they estimate $\xi_{24} = 0.14$ per day and $\xi_{23} = 0.07$ per day, whereas for the MHC class-I-deficient case $\mu_2^{(4)} = 0.28$ per day and $\xi_{24}^{(4)} = 0.21$ per day, and for the MHC class-II-deficient case $\mu_2^{(8)} = 0.88$ per day and $\xi_{23}^{(8)} = 0.20$.[2] This indicates that a source of asymmetry is the difference in death for the $CD4^+$ and $CD8^+$ thymocytes at the DP2 stage. It is at the DP stage when the commitment to either $CD4^+$ or $CD8^+$ lineage is decided. Furthermore, they introduce the dimensionless parameter s, or selection efficiency, for any of the DP compartments as the ratio between the maturation rate and the sum of the maturation and death rates. For example, for DP2 thymocytes in the case of MHC class-I-deficient T-cell development, the selection efficiency is given by

$$s_4 = \frac{\xi_{24}^{(4)}}{\xi_{24}^{(4)} + \mu_2^{(4)}} = \frac{0.21}{0.21 + 0.28} = 0.43. \tag{4.1}$$

In the case of MHC class-II-deficient T-cell development, the selection efficiency is given by

$$s_8 = \frac{\xi_{23}^{(8)}}{\xi_{23}^{(8)} + \mu_2^{(8)}} = \frac{0.20}{0.20 + 0.88} = 0.19. \tag{4.2}$$

Thus, not only is the $CD8^+$ lineage at a disadvantage in terms of death rates during the DP2 stage, but also the selection efficiency is lower than that of the $CD4^+$ lineage. On the other hand, at the DP1 stage, no significant differences are observed between the death μ_1 and maturation rates ξ_{12} estimated for the three different experimental conditions: control, MHC class-I- and MHC class-II-deficient mice.

Taken together, experimental observations and parameter estimates for MHC-deficient mice, the authors can identify that the source of CD4 enrichment takes place at the DP2 stage, given the differences in both death rates and selection efficiencies. Thus, one can now discriminate between the two hypotheses, as the data and the quantification do not support a greater abundance of MHC class-II-restricted thymocytes than MHC class I at the DP1 stage (first hypothesis).

5. Reconciling experimental results

5.1. Quantifying T-cell proliferation

HIV-1 infection results in depletion of $CD4^+$ T cells, but it is not clear what dynamical processes are responsible for this cell loss. One hypothesis is that reduced production from the thymus (see §3.1) causes a decline in cell numbers. Other possibilities are changes in proliferation and death rates of $CD4^+$ T cells. Several clever experimental assays have been developed to analyse proliferation rates of T cells in humans, allowing comparison of healthy and HIV^+ individuals.

One of the first measurements of $CD4^+$ and $CD8^+$ T-cell proliferation was based on the length of their respective telomeres. These are segments of repeated nucleotide sequences at the end of chromosomes, which protect them during replication, because copying of the DNA cannot proceed to the very end of the molecule, since the copying enzymes fall off. Thus, at every cell division, chromosomes lose some nucleotides at the end and, without telomeres, information would be lost (i.e. genes could be truncated). Proliferation of T cells (and others) leads to shortening of the telomeres. Wolthers *et al*. [115] used this idea and measured telomere length over time (up to 10 years) in healthy and HIV-infected individuals, in both their $CD4^+$ and $CD8^+$ T cells. They found stable telomeres in both subsets of healthy people, whereas in HIV-infected people there was no change in the length of telomeres in $CD4^+$ T cells, but a large decrease in this length in $CD8^+$ T cells. This was interpreted as providing 'evidence that turnover in the course of HIV-1 infection can be increased considerably in $CD8^+$ T cells, but not in $CD4^+$ T cells' [115].

A second method to assess T-cell proliferation is based on the expression of the Ki-67 antigen, which is a protein expressed only during cell division. Thus measuring the fraction of T cells bearing this antigen gives a snapshot of the proliferative state of these cells. Analyses of this marker showed that both populations of $CD4^+$ and $CD8^+$ T cells have increased expression of Ki-67 in HIV-1 infection compared with healthy age-matched controls [116]. Although there was large person-to-person variability, there was a significant correlation between the fractions of $Ki\text{-}67^+$ $CD4^+$ T cells and $Ki\text{-}67^+$ $CD8^+$ T cells, as well as between $Ki\text{-}67^+$ $CD4^+$ T cells and HIV-1 viral load in infected people, indicating that both $CD4^+$ and $CD8^+$ T-cell subsets have increased turnover in the setting of HIV-1 infection [116]. These results were confirmed and extended in another study that analysed Ki-67 expression before and after antiretroviral

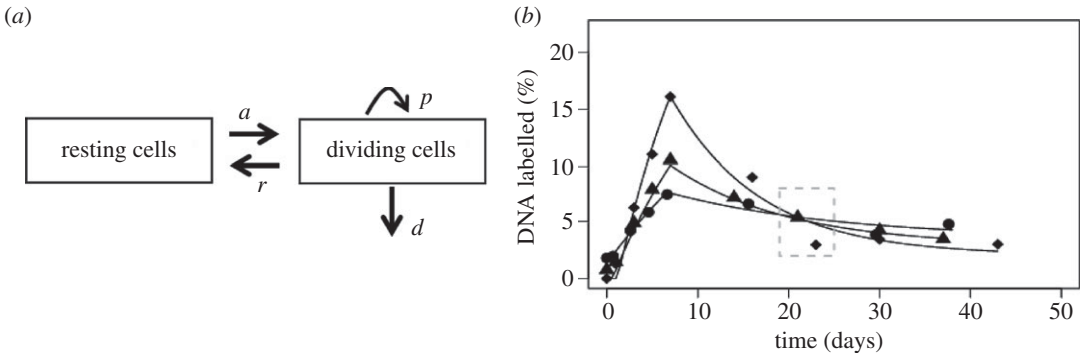

Figure 3. Model used to analyse deuterated glucose labelling of dividing T cells. (*a*) Schematic of the model showing the populations of resting and dividing cells, their interconversion (activation, *a*, and reversion, *r*), proliferation (*p*) and death (*d*) of dividing cells. (*b*) Example of model fit to data for the labelling of cells in one HIV-1-infected person before therapy (diamonds), after short-term therapy (triangles) and long-term therapy (circles). Therapy leads to lower label incorporation, indicating slower turnover; however, if one measures incorporation only after 20–25 days (dashed grey box), there are higher levels of labelling after long-term treatment, which would suggest elevated turnover. The model clearly shows that it is the slope of incorporation or loss of label that contains the relevant information about turnover, not the labelling level at any one time.

treatment for HIV-1 infection [117], which showed that not only was proliferation increased in relation to healthy controls, but this proliferation also declined rapidly with the initiation of treatment [117].

In part, because the methods described are indirect and, in part, because there were relevant discrepancies among the results, a direct measurement of T-cell turnover was applied in the context of HIV-1 infection [118]. This method is based on infusion of deuterated glucose (i.e. glucose where hydrogen has been partly replaced by deuterium) to volunteers. This deuterated glucose will be incorporated in newly synthesized DNA of proliferating cells, in effect labelling them. Sampling T cells from blood after infusion and measuring the fraction of cells labelled with deuterated glucose allows the quantification of T-cell turnover. However, even this 'direct measurement' needs some mathematical model to correctly interpret the labelling data. In the first study of HIV-1 infection using this technique, with a 2-day infusion of deuterated glucose, the increase in proliferation of both $CD4^+$ and $CD8^+$ T cells during HIV-1 infection was confirmed, but, surprisingly, it was found that both proliferation rates were even more increased during successful therapy (i.e. suppressing viral load below detection) [118].

Altogether, these experimental studies resulted in contradictory conclusions, which is a good indication that a modelling approach can be very useful. Initial modelling results suggested that a longer infusion of deuterated glucose, leading to higher T-cell-labelled fractions, together with time courses of labelling, would be more informative. A study was devised with 7-day continuous infusion of deuterated glucose, followed by sampling up to six weeks post-infusion [119]. These data were analysed with a new model allowing for cell activation, proliferation and death (represented in figure 3) [120]. The equations for the resting cells, *R*, and activated cells, *A*, are

$$\frac{dR}{dt} = -aR + rA \tag{5.1}$$

and

$$\frac{dA}{dt} = (p - d)A + a'R - rA, \tag{5.2}$$

where *a* is the activation rate from a compartment of truly resting cells (with an allowance for clonal activation, hence the *a*′), *r* is the rate of return to rest from the activated (dividing) state, and *p* and *d* are the proliferation and death rates of activated (dividing) cells, respectively (a full analysis of the model is provided in [121]). With the help of this model, it was established that proliferation of both $CD4^+$ and $CD8^+$ T cells was increased during HIV-1 infection, and it quickly reduced after the initiation of antiretroviral treatment [119,120]. The estimates of proliferation from the model were highly correlated with $Ki\text{-}67^+$ fractions in both $CD4^+$ and $CD8^+$ T cells. Moreover, the model indicated that activation, proliferation and death of $CD4^+$ T cells was increased, thus leading to stable or lower numbers of proliferating cells over time. On the other hand, in the $CD8^+$ T-cell pool, the increased proliferation leads to larger numbers of proliferating cells [120]. These modelling results explained the telomere data: although $CD4^+$ T cells have higher proliferation rates, these proliferating cells also die quickly leaving behind cells that have proliferated less, with closer to normal telomere lengths; on the other hand, proliferating $CD8^+$ T cells accumulate resulting in shorter telomeres. (This possibility had already been raised by modelling of proliferation and telomere shortening [122].) In addition these models of deuterated glucose showed that, to estimate proliferation, it is important to analyse the slopes of label uptake and loss [120], rather than just the level of labelled cells at one time point [118] (figure 3).

Modelling, thus, contributed to reconcile disparate, even contradictory, results of different techniques to measure T-cell proliferation. At the same time, these initial models also suggested that analyses of this type of data may be more complex than originally thought. For example, one assumes that all cells are proliferating at the same average rate, although it is more likely that different subpopulations have different proliferation rates and/or that these change over time. This has led to numerous studies, of both modelling [79,123–128] and experimental [129–132] aspects, related to T-cell proliferation (an excellent recent review on the subject is provided in [133]).

6. Looking forward

Here, we have focused on examples based on T-cell immunology at the level of populations of cells to maintain a unifying theme. However, many studies from other areas of immunology would also illustrate the richness of mathematical contributions to developing insight into physiological processes. Mathematical and computational models have also been successful in humoral [134,135] and innate immunity [136,137]. The recent discovery of *de facto* innate immune memory [138,139], and of properties of tissue-resident memory T cells, reminiscent of those of innate immune cells [140], are provoking a rethink of the divisions between subfields of immunology, more in tune with the mathematician's natural urge to unite descriptions of the symphonies of molecular and cellular interactions by identifying what is essential.

These examples are often driven by technical advances in genomics, proteomics, mRNA quantification, single-cell capabilities and others, demonstrating how novel experimental approaches push model development. Indeed, new modelling approaches are also needed to make use of the complexity of the data generated [141–144]. It is important to note that these models span multiple biological scales, including intracellular and molecular levels [8,145,146]. Obviously, good mathematical models need good quantitative data, and this is often harder to obtain at finer scales, such as the molecular level. Future technological developments will bring new challenges for modellers.

In spite of these beautiful studies, collaborations between mathematical and experimental immunologists are still infrequent. These collaborations are a fertile ground for new ideas and primed for great impact; however, there are challenges that need to be taken into consideration as the field moves forward. One shared feature of emergent fields is the heterogeneity of approaches, methods and available data. In this context, a certain level of standardization could be helpful to facilitate reproducibility and model sharing and to reduce publication errors by means, for example, of language-independent model definition, with repositories of models contributed by the community. There are several such repositories, such as CellML [147], which is a free, open-source, extensible markup language-based standard for defining mathematical models of cellular function. It currently incorporates more than 50 models in the broad area of mathematical immunology. Other examples are systems biology markup language [148] and JSim [149], which allows storage of data and model analysis.

Recent technical advances permit single-cell measurements [150–152], revealing the inherent heterogeneity in immunological processes. Single-cell data interpretation is almost impossible without mathematics. To realize the full promise of mathematical immunology, more and stronger collaborations between modellers and experimental scientists are needed. One way to stimulate these is to provide venues for enhanced interactions among scientists, at international meetings such as Systems Approaches in Immunology, q-Bio conferences, the KITP Quantitative Immunology programmes, HIV Dynamics and Evolution Conference and collaborative networks such as QuanTI [153], ENLIGHT-TEN [154], INDOEUROPEAN-MATHDS [155] and INTI [156].

In conclusion, it is possible to find excellent examples of synergy between modelling and experiment in immunology, but still much remains to be done.

Author's contribution. All authors contributed equally to the ideas and writing of this article.

Competing interests. We declare we have no competing interests.

Funding. This work has been partially supported by grants BBSRC BB/F003811/1 (G.L., C.M.P.), BBSRC BB/G023395/1 (C.M.P.), FIS2013–47949-C2-2-P (M.C., G.L., C.M.P.), PIRSES-GA-2012-317893 (M.C., G.L., C.M.P.) and National Institutes of Health grant no. R01-AI104373 (R.M.R.). This work has been possible with travel support from the EU, via the FP7 IRSES Network INDOEUROPEAN-MATHDS: Mathematics for health and disease.

Acknowledgements. M.C., G.L. and C.M.P. thank Los Alamos National Laboratory and the University of New Mexico for their hospitality during the drafting and writing of this manuscript. The authors thank Stig Omholt and Peter Hunter for feedback on preliminary versions of this manuscript, as well as two anonymous reviewers, whose comments helped to substantially improve the manuscript.

Disclaimer. The content is solely the responsibility of the authors and does not necessarily represent the official views of funding organizations.

Endnotes

[1]This is in agreement and consistent with other single-cell tracking studies [52].

[2]The superscript, (4) or (8), in a given parameter indicates that it has been estimated making use of MHC class-I- or MHC class-II-deficient mice data, respectively.

References

1. Baltzer F. 1964 Theodor Boveri. *Science* **144**, 809–815. (doi:10.1126/science.144.3620.809)
2. BMC Biology. 2014 Beyond Mendel: modeling in biology. See http://www.biomedcentral.com/bmcbiol/series/beyondmendel.
3. Callard RE, Yates AJ. 2005 Immunology and mathematics: crossing the divide. *Immunology* **115**, 21–33. (doi:10.1111/j.1365-2567.2005.02142.x)
4. Lander AD. 2014 Making sense in biology: an appreciation of Julian Lewis. *BMC Biol*. **12**, 57. (doi:10.1186/s12915-014-0057-5)
5. Noble D. 2002 The rise of computational biology. *Nat. Rev. Mol. Cell Biol*. **3**, 459–463. (doi:10.1038/nrm810)
6. Germain RN. 2004 An innately interesting decade of research in immunology. *Nat. Med*. **10**, 1307–1320. (doi:10.1038/nm1159)
7. Chakraborty AK, Das J. 2010 Pairing computation with experimentation: a powerful coupling for understanding T cell signalling. *Nat. Rev. Immunol*. **10**, 59–71. (doi:10.1038/nri2688)
8. Altan-Bonnet G, Germain RN. 2005 Modeling T cell antigen discrimination based on feedback control of digital ERK responses. *PLoS Biol*. **3**, e356. (doi:10.1371/journal.pbio.0030356)
9. Germain RN, Meier-Schellersheim M, Nita-Lazar A, Fraser ID. 2011 Systems biology in immunology—a computational modeling perspective. *Annu. Rev. Immunol*. **29**, 527–585. (doi:10.1146/annurev-immunol-030409-101317)
10. Rochman Y, Spolski R, Leonard WJ. 2009 New insights into the regulation of T cells by c family cytokines. *Nat. Rev. Immunol*. **9**, 480–490. (doi:10.1038/nri2580)
11. Hunter PJ, Borg TK. 2003 Integration from proteins to organs: the Physiome project. *Nat. Rev. Mol. Cell Biol*. **4**, 237–243. (doi:10.1038/nrm1054)
12. Arazi A, Pendergraft WF, Ribeiro RM, Perelson AS, Hacohen N. 2013 Human systems immunology: hypothesis-based modeling and unbiased data-

driven approaches. *Semin. Immunol*. **25**, 193–200. (doi:10.1016/j.smim.2012.11.003)

13. Gunawardena J. 2014 Models in biology: 'accurate descriptions of our pathetic thinking'. *BMC Biol*. **12**, 29. (doi:10.1186/1741-7007-12-29)
14. Molina-Paŕis C, Lythe G. 2011 *Mathematical models and immune cell biology*. New York, NY: Springer Science and Business Media.
15. Tauber AI, Podolsky SH. 2000 *The generation of diversity: clonal selection theory and the rise of molecular immunology*. Cambridge, MA: Harvard University Press.
16. Dobzhansky T. 1964 Biology, molecular and organismic. *Am. Zool*. **4**, 443–452. (doi:10.1093/icb/4.4.443)
17. Cyster JG. 1999 Chemokines and cell migration in secondary lymphoid organs. *Science* **286**, 2098–2102. (doi:10.1126/science.286.5447.2098)
18. Bajenoff M, Egen J, Qi H, Huang A, Castellino F, Germain R. 2007 Highways, byways and breadcrumbs: directing lymphocyte traffic in the lymph node. *Trends Immunol*. **28**, 346–352. (doi:10.1016/j.it.2007.06.005)
19. Miller MJ, Wei SH, Parker I, Cahalan MD. 2002 Two-photon imaging of lymphocyte motility and antigen response in intact lymph node. *Science* **296**, 1869–1873. (doi:10.1126/science.1070051)
20. Stoll S, Delon J, Brotz TM, Germain RN. 2002 Dynamic imaging of T cell-dendritic cell interactions in lymph nodes. *Science* **296**, 1873–1876. (doi:10.1126/science.1071065)
21. Bousso P, Bhakta NR, Lewis RS, Robey E. 2002 Dynamics of thymocyte-stromal cell interactions visualized by two photon microscopy. *Science* **296**, 1876–1880. (doi:10.1126/science.1070945)
22. Germain RN, Robey EA, Cahalan MD. 2012 A decade of imaging cellular motility and interaction dynamics in the immune system. *Science* **336**, 1676–1681. (doi:10.1126/science.1221063)
23. Moreau HD, Bousso P. 2014 Visualizing how T cells collect activation signals *in vivo*. *Curr. Opin. Immunol*. **26**, 56–62. (doi:10.1016/j.coi.2013.10.013)
24. Zinselmeyer BH *et al.* 2005 *In situ* characterization of $CD4^+$ T cell behavior in mucosal and systemic lymphoid tissues during the induction of oral priming and tolerance. *J. Exp. Med*. **201**, 1815–1823. (doi:10.1084/jem.20050203)
25. Cahalan MD, Parker I, Wei SH, Miller MJ. 2003 Real-time imaging of lymphocytes *in vivo*. *Curr. Opin. Immunol*. **15**, 372–377. (doi:10.1016/S0952-7915(03)00079-7)
26. Beltman JB, Maŕee AF, de Boer RJ. 2009 Analysing immune cell migration. *Nat. Rev. Immunol*. **9**, 789–798. (doi:10.1038/nri2638)
27. Preston S, Waters S, Jensen O, Heaton P, Pritchard D. 2006 T-cell motility in the early stages of the immune response modeled as a random walk amongst targets. *Phys. Rev. E* **74**, 11910. (doi:10.1103/PhysRevE.74.011910)
28. Meyer-Hermann ME, Maini PK. 2005 Interpreting two-photon imaging data of lymphocyte motility. *Phys. Rev. E* **71**, 61912. (doi:10.1103/PhysRevE.71.061912)
29. Beauchemin C, Dixit NM, Perelson AS. 2007 Characterizing T cell movement within lymph nodes in the absence of antigen. *J. Immunol*. **178**, 5505–5512. (doi:10.4049/jimmunol.178.9.5505)
30. Textor J, Henrickson SE, Mandl JN, von Andrian UH, Westermann J, de Boer RJ, Beltman JB. 2014 Random migration and signal integration promote rapid and robust T cell recruitment. *PLoS Comput. Biol*. **10**, e1003752. (doi:10.1371/journal.pcbi.1003752)
31. Mandl JN, Liou R, Klauschen F, Vrisekoop N, Monteiro JP, Yates AJ, Huang AY, Germain RN. 2012 Quantification of lymph node transit times reveals differences in antigen surveillance strategies of naive $CD4^+$ and $CD8^+$ T cells. *Proc. Natl Acad. Sci. USA* **109**, 18 036–18 041. (doi:10.1073/pnas.1211717109)
32. Letendre K, Donnadieu E, Moses ME, Cannon JL. 2015 Bringing statistics up to speed with data in analysis of lymphocyte motility. *PLoS ONE* **10**, e1026333. (doi:10.1371/journal.pone.0126333)
33. Bajenoff M, Egen JG, Koo LY, Laugier JP, Brau F, Glaichenhaus N, Germain RN. 2006 Stromal cell networks regulate lymphocyte entry, migration, and territoriality in lymph nodes. *Immunity* **25**, 989–1001. (doi:10.1016/j.immuni.2006.10.011)
34. Mueller SN, Germain RN. 2009 Stromal cell contributions to the homeostasis and functionality of the immune system. *Nat. Rev. Immunol*. **9**, 618–629. (doi:10.1038/nri2588)
35. Donovan GM, Lythe G. 2012 T-cell movement on the reticular network. *J. Theor. Biol*. **295**, 59–67. (doi:10.1016/j.jtbi.2011.11.001)
36. Graw F, Regoes RR. 2012 Influence of the fibroblastic reticular network on cell-cell interactions in lymphoid organs. *PLoS Comput. Biol*. **8**, e1002436. (doi:10.1371/journal.pcbi.1002436)
37. Vroomans RM, Maŕee AF, de Boer RJ, Beltman JB. 2012 Chemotactic migration of T cells towards dendritic cells promotes the detection of rare antigens. *PLoS Comput. Biol*. **8**, e1002763. (doi:10.1371/journal.pcbi.1002763)
38. Celli S, Day M, Müller AJ, Molina-Paris C, Lythe G, Bousso P. 2012 How many dendritic cells are required to initiate a T-cell response? *Blood* **120**, 3945–3948. (doi:10.1182/blood-2012-01-408260)
39. Ariotti S, Hogenbirk MA, Dijkgraaf FE, Visser LL, Hoekstra ME, Song JY, Jacobs H, Haanen JB, Schumacher TN. 2014 Skin-resident memory $CD8^+$ T cells trigger a state of tissue-wide pathogen alert. *Science* **346**, 101–105. (doi:10.1126/science.1254803)
40. Ariotti S *et al.* 2012 Tissue-resident memory $CD8^+$ T cells continuously patrol skin epithelia to quickly recognize local antigen. *Proc. Natl Acad. Sci. USA* **109**, 19 739–19 744. (doi:10.1073/pnas.1208927109)
41. Ariotti S, Beltman JB, Borsje R, Hoekstra ME, Halford WP, Haanen JB, de Boer RJ, Schumacher TNM. 2015 Subtle CXCR3-dependent chemotaxis of CTLs within infected tissue allows efficient target localization. *J. Immunol*. **195**, 5285–5295. (doi:10.4049/jimmunol.1500853)
42. Bousso P, Robey EA. 2004 Dynamic behavior of T cells and thymocytes in lymphoid organs as revealed by two-photon microscopy. *Immunity* **21**, 349–355. (doi:10.1016/j.immuni.2004.08.005)
43. Mempel TR, Henrickson SE, Von Andrian UH. 2004 T-cell priming by dendritic cells in lymph nodes occurs in three distinct phases. *Nature* **427**, 154–159. (doi:10.1038/nature02238)
44. Beltman JB, Maŕee AF, Lynch JN, Miller MJ, de Boer RJ. 2007 Lymph node topology dictates T cell migration behavior. *J. Exp. Med*. **204**, 771–780. (doi:10.1084/jem.20061278)
45. Bogle G, Dunbar PR. 2009 Agent-based simulation of T-cell activation and proliferation within a lymph node. *Immunol. Cell Biol*. **88**, 172–179. (doi:10.1038/icb.2009.78)
46. Graner F, Glazier JA. 1992 Simulation of biological cell sorting using a two-dimensional extended Potts model. *Phys. Rev. Lett*. **69**, 2013. (doi:10.1103/PhysRevLett.69.2013)
47. Glazier JA, Graner F. 1993 Simulation of the differential adhesion driven rearrangement of biological cells. *Phys. Rev. E* **47**, 2128. (doi:10.1103/PhysRevE.47.2128)
48. Beltman JB, Maŕee AF, De Boer RJ. 2007 Spatial modelling of brief and long interactions between T cells and dendritic cells. *Immunol. Cell Biol*. **85**, 306–314. (doi:10.1038/sj.icb.7100054)
49. Coombs D, Kalergis AM, Nathenson SG, Wofsy C, Goldstein B. 2002 Activated TCRs remain marked for internalization after dissociation from pMHC. *Nat. Immunol*. **3**, 926–931. (doi:10.1038/ni838)
50. Dushek O, Das R, Coombs D. 2009 A role for rebinding in rapid and reliable T cell responses to antigen. *PLoS Comput. Biol*. **5**, e1000578. (doi:10.1371/journal.pcbi.1000578)
51. Aleksic M, Dushek O, Zhang H, Shenderov E, Chen JL, Cerundolo V, Coombs D, van der Merwe PA. 2010 Dependence of T cell antigen recognition on T cell receptor-peptide MHC confinement time. *Immunity* **32**, 163–174. (doi:10.1016/j.immuni.2009.11.013)
52. Marchingo JM *et al.* 2014 Antigen affinity, costimulation, and cytokine inputs sum linearly to amplify T cell expansion. *Science* **346**, 1123–1127. (doi:10.1126/science.1260044)
53. Hawkins ED, Turner ML, Dowling MR, Van Gend C, Hodgkin PD. 2007 A model of immune regulation as a consequence of randomized lymphocyte division and death times. *Proc. Natl Acad. Sci. USA* **104**, 5032–5037. (doi:10.1073/pnas.0700026104)
54. Turner ML, Hawkins ED, Hodgkin PD. 2008 Quantitative regulation of B cell division destiny by signal strength. *J. Immunol*. **181**, 374–382. (doi:10.4049/jimmunol.181.1.374)
55. Hawkins E, Markham J, McGuinness L, Hodgkin P. 2009 A single-cell pedigree analysis of alternative stochastic lymphocyte fates. *Proc. Natl Acad. Sci. USA* **106**, 13 457–13 462. (doi:10.1073/pnas.0905629106)
56. Subramanian VG, Duffy KR, Turner ML, Hodgkin PD. 2008 Determining the expected variability of

immune responses using the cyton model. *J. Math. Biol*. **56**, 861–892. (doi:10.1007/s00285-007-0142-2)

57. Hawkins ED, Turner ML, Wellard CJ, Zhou J, Dowling MR, Hodgkin P. 2013 Quantal and graded stimulation of B lymphocytes as alternative strategies for regulating adaptive immune responses. *Nat. Commun*. **4**, 2406. (doi:10.1038/ncomms3406)
58. Sakaue-Sawano A *et al.* 2008 Visualizing spatiotemporal dynamics of multicellular cell-cycle progression. *Cell* **132**, 487–498. (doi:10.1016/j.cell.2007.12.033)
59. Mechali M, Lutzmann M. 2008 The cell cycle: now live and in color. *Cell* **132**, 341–343. (doi:10.1016/j.cell.2008.01.031)
60. Sakaue-Sawano A, Miyawaki A. 2014 Visualizing spatiotemporal dynamics of multicellular cell-cycle progressions with fucci technology. *Cold Spring Harbour Protoc.* **2014**, pii: pdb.prot080408. (doi:10.1101/pdb.prot080408)
61. Ribeiro RM, Perelson AS. 2007 Determining thymic output quantitatively: using models to interpret experimental T-cell receptor excision circle (TREC) data. *Immunol. Rev*. **216**, 21–34. (doi:10.1111/j.1600-065X.2006.00493.x)
62. Kong F, Chen CH, Cooper MD. 1998 Thymic function can be accurately monitored by the level of recent T cell emigrants in the circulation. *Immunity* **8**, 97–104. (doi:10.1016/S1074-7613(00)80462-8)
63. Douek DC *et al.* 1998 Changes in thymic function with age and during the treatment of HIV infection. *Nature* **396**, 690–695. (doi:10.1038/25374)
64. Kong FK, Chen CL, Six A, Hockett RD, Cooper MD. 1999 T cell receptor gene deletion circles identify recent thymic emigrants in the peripheral T cell pool. *Proc. Natl Acad. Sci. USA* **96**, 1536–1540. (doi:10.1073/pnas.96.4.1536)
65. Livak F, Schatz DG. 1996 T-cell receptor alpha locus V(D)J recombination by-products are abundant in thymocytes and mature T cells. *Mol. Cell Biol*. **16**, 609–618. (doi:10.1128/MCB.16.2.609)
66. Guy-Grand D, Azogui O, Celli S, Darche S, Nussenzweig MC, Kourilsky P, Vassalli P. 2003 Extrathymic T cell lymphopoiesis ontogeny and contribution to gut intraepithelial lymphocytes in athymic and euthymic mice. *J. Exp. Med*. **197**, 333–341. (doi:10.1084/jem.20021639)
67. Verschuren MC, Wolvers-Tettero IL, Breit TM, Noordzij J, van Wering ER, van Dongen JJ. 1997 Preferential rearrangements of the T cell receptor-delta-deleting elements in human T cells. *J. Immunol*. **158**, 1208–1216.
68. Zhang L *et al.* 1999 Measuring recent thymic emigrants in blood of normal and HIV-1-infected individuals before and after effective therapy. *J. Exp. Med*. **190**, 725–732. (doi:10.1084/jem.190.5.725)
69. Hatzakis A *et al.* 2000 Effect of recent thymic emigrants on progression of HIV-1 disease. *Lancet* **355**, 599–604. (doi:10.1016/S0140-6736(99)10311-8)
70. Ye P, Kirschner DE. 2002 Measuring emigration of human thymocytes by T-cell receptor excision circles. *Crit. Rev. Immunol*. **22**, 483–497. (doi:10.1615/critrevimmunol.v22.i5-6.80)
71. Hazenberg MD, Borghans JA, de Boer RJ, Miedema F. 2003 Thymic output: a bad TREC record. *Nat. Immunol*. **4**, 97–99. (doi:10.1038/ni0203-97)
72. Dion ML, Sekaly RP, Cheynier R. 2007 Estimating thymic function through quantification of T-cell receptor excision circles. *Methods Mol. Biol*. **380**, 197–213. (doi:10.1007/978-1-59745-395-0_12)
73. Hazenberg MD *et al.* 2000 Increased cell division but not thymic dysfunction rapidly affects the T-cell receptor excision circle content of the naive T cell population in HIV-1 infection. *Nat. Med*. **6**, 1036–1042. (doi:10.1038/79549)
74. Dutilh BE, de Boer RJ. 2003 Decline in excision circles requires homeostatic renewal or homeostatic death of naive T cells. *J. Theor. Biol*. **224**, 351–358. (doi:10.1016/S0022-5193(03)00172-3)
75. Lewin SR *et al.* 2002 Dynamics of T cells and TCR excision circles differ after treatment of acute and chronic HIV infection. *J. Immunol*. **169**, 4657–4666. (doi:10.4049/jimmunol.169.8.4657)
76. Carcelain G *et al.* 2003 IL-2 therapy and thymic production of naive CD4T cells in HIV-infected patients with severe CD4 lymphopenia. *AIDS* **17**, 841–850. (doi:10.1097/00002030-200304110-00009)
77. Di Mascio M *et al.* 2006 Naive T-cell dynamics in human immunodeficiency virus type 1 infection: effects of highly active antiretroviral therapy provide insights into the mechanisms of naive T-cell depletion. *J. Virol*. **80**, 2665–2674. (doi:10.1128/JVI.80.6.2665-2674.2006)
78. Ribeiro RM, de Boer RJ. 2008 The contribution of the thymus to the recovery of peripheral naive T-cell numbers during antiretroviral treatment for HIV infection. *J. Acquir. Immune Defic. Syndr*. **49**, 1–8. (doi:10.1097/QAI.0b013e318184fb28)
79. Bains I, Thiebaut R, Yates AJ, 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**, 4329–4336. (doi:10.4049/jimmunol.0900743)
80. Ye P, Kirschner DE. 2002 Reevaluation of T cell receptor excision circles as a measure of human recent thymic emigrants. *J. Immunol*. **168**, 4968–4979. (doi:10.4049/jimmunol.168.10.4968)
81. Murray JM, Kaufmann GR, Hodgkin PD, Lewin SR, Kelleher AD, Davenport MP, Zaunders JJ. 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**, 487–495. (doi:10.1046/j.1440-1711.2003.01191.x)
82. van den Dool C, de Boer RJ. 2006 The effects of age, thymectomy, and HIV infection on alpha and beta TCR excision circles in naive T cells. *J. Immunol*. **177**, 4391–4401. (doi:10.4049/jimmunol.177.7.4391)
83. Arron ST *et al.* 2005 Impact of thymectomy on the peripheral T cell pool in rhesus macaques before and after infection with simian immunodeficiency virus. *Eur. J. Immunol*. **35**, 46–55. (doi:10.1002/eji.200424996)
84. Zarnitsyna V, Evavold B, Schoettle L, Blattman J, Antia R. 2013 Estimating the diversity, completeness, and crossreactivity of the T cell repertoire. *Front. Immunol*. **4**, 485. (doi:10.3389/fimmu.2013.00485)
85. Robins H. 2013 Immunosequencing: applications of immune repertoire deep sequencing. *Curr. Opin. Immunol*. **25**, 646–652. (doi:10.1016/j.coi.2013.09.017)
86. Salameire D *et al.* 2009 Efficient characterization of the TCR repertoire in lymph nodes by flow cytometry. *Cytometry Part A*. **75**, 743–751. (doi:10.1002/cyto.a.20767)
87. Ciupe SM, Devlin BH, Markert ML, Kepler TB. 2013 Quantification of total T-cell receptor diversity by flow cytometry and spectratyping. *BMC Immunol*. **14**, 35. (doi:10.1186/1471-2172-14-35)
88. Britanova OV *et al.* 2014 Age-related decrease in TCR repertoire diversity measured with deep and normalized sequence profiling. *J. Immunol*. **192**, 2689–2698. (doi:10.4049/jimmunol.1302064)
89. Elhanati Y, Murugan A, Callan CG, Mora T, Walczak AM. 2014 Quantifying selection in immune receptor repertoires. *Proc. Natl Acad. Sci. USA* **111**, 9875–9880. (doi:10.1073/pnas.1409572111)
90. Arstila TP, 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**, 958–961. (doi:10.1126/science.286.5441.958)
91. Baum PD, Young JJ, McCune JM. 2011 Measurement of absolute T cell receptor rearrangement diversity. *J. Immunol. Methods* **368**, 45–53. (doi:10.1016/j.jim.2011.03.001)
92. 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**, 5782–5787. (doi:10.4049/jimmunol.164.11.5782)
93. Bousso P, Levraud JP, Kourilsky P, Abastado JP. 1999 The composition of a primary T cell response is largely determined by the timing of recruitment of individual T cell clones. *J. Exp. Med*. **189**, 1591–1600. (doi:10.1084/jem.189.10.1591)
94. Bousso P, Casrouge A, Altman JD, Haury M, Kanellopoulos J, Abastado JP, Kourilsky P. 1998 Individual variations in the murine T cell response to a specific peptide reflect variability in naive repertoires. *Immunity* **9**, 169–178. (doi:10.1016/S1074-7613(00)80599-3)
95. Venturi V, Kedzierska K, Tanaka MM, Turner SJ, Doherty PC, Davenport MP. 2008 Method for assessing the similarity between subsets of the T cell receptor repertoire. *J. Immunol. Methods* **329**, 67–80. (doi:10.1016/j.jim.2007.09.016)
96. Bergot AS, Chaara W, Ruggiero E, Mariotti-Ferrandiz E, Dulauroy S, Schmidt M, Six A, Klatzmann D. 2015 TCR sequences and tissue distribution discriminate the subsets of naive and activated/memory Treg cells in mice. *Eur. J. Immunol*. **45**, 1524–1534. (doi:10.1002/eji.201445269)
97. Thomas N *et al.* 2014 Tracking global changes induced in the CD4T cell receptor repertoire by

immunization with a complex antigen using short stretches of CDR3 protein sequence. *Bioinformatics*. **30**, 3181–3188. (doi:10.1093/bioinformatics/btu523)

98. Cukalac T, Kan WT, Dash P, Guan J, Quinn KM, Gras S, Thomas PG, La Gruta NL. 2015 Paired TCR$\alpha\beta$ analysis of virus-specific $CD8^+$ T cells exposes diversity in a previously defined narrow repertoire. *Immunol. Cell Biol*. **93**, 804. (doi:10.1038/icb.2015.44)
99. Freeman JD, Warren RL, Webb JR, Nelson BH, Holt RA. 2009 Profiling the T-cell receptor beta-chain repertoire by massively parallel sequencing. *Genome Res*. **19**, 1817–1824. (doi:10.1101/gr.092924.109)
100. Robins HS, Campregher PV, Srivastava SK, Wacher A, Turtle CJ, Kahsai O, Riddell SR, Warren EH, Carlson CS. 2009 Comprehensive assessment of T-cell receptor β-chain diversity in $\alpha\beta$ T cells. *Blood* **114**, 4099–4107. (doi:10.1182/blood-2009-04-217604)
101. Warren RL, Freeman JD, Zeng T, Choe G, Munro S, Moore R, Webb JR, Holt RA. 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**, 790–797. (doi:10.1101/gr.115428.110)
102. Rempala GA, Seweryn M. 2013 Methods for diversity and overlap analysis in T-cell receptor populations. *J. Math. Biol*. **67**, 1339–1368. (doi:10.1007/s00285-012-0589-7)
103. Qi Q *et al.* 2014 Diversity and clonal selection in the human T-cell repertoire. *Proc. Natl Acad. Sci. USA* **111**, 13 139–13 144. (doi:10.1073/pnas.1409155111)
104. Li Z *et al.* 2015 Comprehensive analysis of the T-cell receptor beta chain gene in rhesus monkey by high throughput sequencing. *Sci. Rep*. **5**, 10092. (doi:10.1038/srep10092)
105. Best K, Oakes T, Heather JM, Shawe-Taylor J, Chain B. 2015 Computational analysis of stochastic heterogeneity in PCR amplification efficiency revealed by single molecule barcoding. *Sci. Rep*. **5**. (doi:10.1038/srep14629)
106. Antia R, Pilyugin SS, 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. USA* **95**, 14 926–14 931. (doi:10.1073/pnas.95.25.14926)
107. Stirk ER, Molina-París C, van den Berg HA. 2008 Stochastic niche structure and diversity maintenance in the T cell repertoire. *J. Theor. Biol*. **255**, 237–249. (doi:10.1016/j.jtbi.2008.07.017)
108. Zilman A, Ganusov VV, Perelson AS. 2010 Stochastic models of lymphocyte proliferation and death. *PLoS ONE* **5**, e12775. (doi:10.1371/journal.pone.0012775)
109. Johnson PL, Yates AJ, Goronzy JJ, Antia R. 2012 Peripheral selection rather than thymic involution explains sudden contraction in naive CD4 T-cell diversity with age. *Proc. Natl Acad. Sci. USA* **109**, 21 432–21 437. (doi:10.1073/pnas.1209283110)
110. Lythe G, Callard RE, Hoare RL, Molina-París C. 2016 How many TCR clonotypes does a body maintain? *J. Theor. Biol*. **389**, 214–224. (doi:10.1016/j.jtbi.2015.10.016)
111. Sinclair C, Bains I, Yates AJ, Seddon B. 2013 Asymmetric thymocyte death underlies the CD4:CD8 T-cell ratio in the adaptive immune system. *Proc. Natl Acad. Sci. USA* **110**, E2905–E2914. (doi:10.1073/pnas.1304859110)
112. Comans-Bitter WM, de Groot R, van den Beemd R, Neijens HJ, Hop WC, Groeneveld K, Hooijkaas H, van Dongen JJM. 1997 Immunophenotyping of blood lymphocytes in childhood: reference values for lymphocyte subpopulations. *J. Pediatr*. **130**, 388–393. (doi:10.1016/S0022-3476(97)70200-2)
113. Saini M, Sinclair C, Marshall D, Tolaini M, Sakaguchi S, Seddon B. 2010 Regulation of *Zap70* expression during thymocyte development enables temporal separation of CD4 and CD8 repertoire selection at different signaling thresholds. *Sci. Signal*. **3**, ra23. (doi:10.1126/scisignal.2000702)
114. Stritesky GL, Xing Y, Erickson JR, Kalekar LA, Wang X, Mueller DL, Jameson SC, Hogquist KA. 2013 Murine thymic selection quantified using a unique method to capture deleted T cells. *Proc. Natl Acad. Sci. USA* **110**, 4679–4684. (doi:10.1073/pnas.1217532110)
115. Wolthers KC *et al.* 1996 T cell telomere length in HIV-1 infection: no evidence for increased $CD4^+$ T cell turnover. *Science* **274**, 1543–1547. (doi:10.1126/science.274.5292.1543)
116. Sachsenberg N *et al.* 1998 Turnover of $CD4^+$ and $CD8^+$ T lymphocytes in HIV-1 infection as measured by Ki-67 antigen. *J. Exp. Med*. **187**, 1295–1303. (doi:10.1084/jem.187.8.1295)
117. Hazenberg MD *et al.* 2000 T-cell division in human immunodeficiency virus (HIV)-1 infection is mainly due to immune activation: a longitudinal analysis in patients before and during highly active antiretroviral therapy (HAART). *Blood* **95**, 249–255.
118. Hellerstein M *et al.* 1999 Directly measured kinetics of circulating T lymphocytes in normal and HIV-1-infected humans. *Nat. Med*. **5**, 83–89. (doi:10.1038/4772)
119. Mohri H *et al.* 2001 Increased turnover of T lymphocytes in HIV-1 infection and its reduction by antiretroviral therapy. *J. Exp. Med*. **194**, 1277–1287. (doi:10.1084/jem.194.9.1277)
120. Ribeiro RM, Mohri H, Ho DD, Perelson AS. 2002 In vivo dynamics of T cell activation, proliferation, and death in HIV-1 infection: why are $CD4^+$ but not $CD8^+$ T cells depleted? *Proc. Natl Acad. Sci. USA* **99**, 15 572–15 577. (doi:10.1073/pnas.242358099)
121. Ribeiro RM, Mohri H, Ho DD, Perelson AS. 2002 Modeling deuterated glucose labeling of T-lymphocytes. *Bull. Math. Biol*. **64**, 385–405. (doi:10.1006/bulm.2001.0282)
122. Wolthers KC, Noest AJ, Otto SA, Miedema F, De Boer RJ. 1999 Normal telomere lengths in naive and memory $CD4^+$ T cells in HIV type 1 infection: a mathematical interpretation. *AIDS Res. Hum. Retroviruses* **15**, 1053–1062. (doi:10.1089/088922299310340)
123. Asquith B, Debacq C, Macallan DC, Willems L, Bangham CR. 2002 Lymphocyte kinetics: the interpretation of labelling data. *Trends Immunol*. **23**, 596–601. (doi:10.1016/S1471-4906(02)02337-2)
124. De Boer RJ, Mohri H, Ho DD, Perelson AS. 2003 Turnover rates of B cells, T cells, and NK cells in simian immunodeficiency virus-infected and uninfected rhesus macaques. *J. Immunol*. **170**, 2479–2487. (doi:10.4049/jimmunol.170.5.2479)
125. Borghans JA, de Boer RJ. 2007 Quantification of T-cell dynamics: from telomeres to DNA labeling. *Immunol. Rev*. **216**, 35–47. (doi:10.1111/j.1600-065X.2007.00497.x)
126. Ganusov VV, Borghans JA, De Boer RJ. 2010 Explicit kinetic heterogeneity: mathematical models for interpretation of deuterium labeling of heterogeneous cell populations. *PLoS Comput. Biol*. **6**, e1000666. (doi:10.1371/journal.pcbi.1000666)
127. De Boer RJ, Perelson AS, Ribeiro RM. 2012 Modelling deuterium labelling of lymphocytes with temporal and/or kinetic heterogeneity. *J. R. Soc. Interface* **9**, 2191–2200. (doi:10.1098/rsif.2012.0149)
128. Ganusov VV, De Boer RJ. 2012 A mechanistic model for bromodeoxyuridine dilution naturally explains labelling data of self-renewing T cell populations. *J. R. Soc. Interface*. **9**, 20120617. (doi:10.1098/rsif.2012.0617)
129. Hellerstein MK *et al.* 2003 Subpopulations of long-lived and short-lived T cells in advanced HIV-1 infection. *J. Clin. Invest*. **112**, 956–966. (doi:10.1172/JCI200317533)
130. Vrisekoop N *et al.* 2008 Sparse production but preferential incorporation of recently produced naive T cells in the human peripheral pool. *Proc. Natl. Acad. Sci. USA* **105**, 6115–6120. (doi:10.1073/pnas.0709713105)
131. Westera L *et al.* 2013 Closing the gap between T-cell life span estimates from stable isotope-labeling studies in mice and humans. *Blood* **122**, 2205–2212. (doi:10.1182/blood-2013-03-488411)
132. Vrisekoop N *et al.* 2015 Quantification of naive and memory T-cell turnover during HIV-1 infection. *AIDS*. **29**, 2071–2080. (doi:10.1097/QAD.0000000000000822)
133. De Boer RJ, Perelson AS. 2013 Quantifying T lymphocyte turnover. *J. Theor. Biol*. **327**, 45–87. (doi:10.1016/j.jtbi.2012.12.025)
134. Meyer-Hermann M, Mohr E, Pelletier N, Zhang Y, Victora GD, Toellner KM. 2012 A theory of germinal center B cell selection, division, and exit. *Cell Rep*. **2**, 162–174. (doi:10.1016/j.celrep.2012.05.010)
135. Dustin ML, Meyer-Hermann M. 2012 Antigen feast or famine. *Science* **335**, 408–409. (doi:10.1126/science.1218165)
136. Smith AM, Adler FR, Ribeiro RM, Gutenkunst RN, McAuley JL, McCullers JA, Perelson AS. 2013 Kinetics of coinfection with influenza A virus and Streptococcus pneumoniae. *PLoS Pathog*. **9**, e1003238. (doi:10.1371/journal.ppat.1003238)
137. Gillard JJ, Laws TR, Lythe G, Molina-París C. 2014 Modeling early events in *Francisella tularensis*

pathogenesis. *Front. Cell Infect. Microbiol*. **4**, 169. (doi:10.3389/fcimb.2014.00169)

138. O'Neill LA, Golenbock D, Bowie AG. 2013 The history of Toll-like receptors—redefining innate immunity. *Nat. Rev. Immunol*. **13**, 453–460. (doi:10.1038/nri3446)
139. Netea MG, Latz E, Mills KH, O'Neill LA. 2015 Innate immune memory: a paradigm shift in understanding host defense. *Nat. Immunol*. **16**, 675–679. (doi:10.1038/ni.3178)
140. Heath WR, Carbone FR. 2013 The skin-resident and migratory immune system in steady state and memory: innate lymphocytes, dendritic cells and T cells. *Nat. Immunol*. **14**, 978–985. (doi:10.1038/ni.2680)
141. Chylek LA, Harris LA, Faeder JR, Hlavacek WS. 2015 Modeling for (physical) biologists: an introduction to the rule-based approach. *Phys. Biol*. **12**, 045007. (doi:10.1088/1478-3975/12/4/045007)
142. Stites EC, Aziz M, Creamer MS, Von Hoff DD, Posner RG, Hlavacek WS. 2015 Use of mechanistic models to integrate and analyze multiple proteomic datasets. *Biophys. J*. **108**, 1819–1829. (doi:10.1016/j.bpj.2015.02.030)
143. Munsky B, Neuert G, van Oudenaarden A. 2012 Using gene expression noise to understand gene regulation. *Science* **336**, 183–187. (doi:10.1126/science.1216379)
144. Korem Y, Szekely P, Hart Y, Sheftel H, Hausser J, Mayo A, Rothenberg ME, Kalisky T, Alon U. 2015 Geometry of the gene expression space of individual cells. *PLoS Comput. Biol*. **11**, e1004224. (doi:10.1371/journal.pcbi.1004224)
145. Intosalmi J, Ahlfors H, Rautio S, Mannerstom H, Chen ZJ, Lahesmaa R, Stockinger B, Lähdesmäki H. 2015 Analyzing Th17 cell differentiation dynamics using a novel integrative modeling framework for time-course RNA sequencing data. *BMC Syst. Biol*. **9**, 81. (doi:10.1186/s12918-015-0223-6)
146. Chylek LA, Wilson BS, Hlavacek WS. 2014 Modeling biomolecular site dynamics in immunoreceptor signaling systems. *Adv. Exp. Med. Biol*. **844**, 245–262. (doi:10.1007/978-1-4939-2095-2_12)
147. Lloyd CM, Halstead MD, Nielsen PF. 2004 CellML: its future, present and past. *Prog. Biophys. Mol. Biol*. **85**, 433–450. (doi:10.1016/j.pbiomolbio.2004.01.004)
148. The systems biology markup language. See http://sbml.org.
149. JSim. See http://physiome.org/jsim/.
150. Spitzer MH *et al.* 2015 An interactive reference framework for modeling a dynamic immune system. *Science* **349**, 1259425. (doi:10.1126/science.1259425)
151. Perie L, Hodgkin PD, Naik SH, Schumacher TN, de Boer RJ, Duffy KR. 2014 Determining lineage pathways from cellular barcoding experiments. *Cell Rep*. **6**, 617–624. (doi:10.1016/j.celrep.2014.01.016)
152. Gerlach C *et al.* 2013 Heterogeneous differentiation patterns of individual $CD8^+$ T cells. *Science* **340**, 635–639. (doi:10.1126/science.1235487)
153. Quantitative T cell Immunology. Marie Curie Initial Training Network. See http://www1.maths.leeds.ac.uk/applied/QUANTI/.
154. European Network Linking Informatics and Genomics of Helper T Cells Comprising TEN Beneficiaries. See http://www.enlight-ten.eu/.
155. Indo-European Network: Mathematics for Health and Disease. See http://www1.maths.leeds.ac.uk/applied/INDOMATH/.
156. International Network in Theoretical Immunology. See http://www1.maths.leeds.ac.uk/applied/INTI/.