Mathematical and statistical analysis of the *Trypanosoma brucei* slender to stumpy transition

N. J. SAVILL[1]* *and* J. R. SEED[2]

[1] *Department of Zoology, Cambridge University, Downing Street, Cambridge CB2 3EJ, UK*
[2] *Department of Epidemiology, School of Public Health, University of North Carolina, Chapel Hill, North Carolina 27599–7435, USA*

(*Received 28 April 2003; revised 4 August 2003; accepted 5 August 2003*)

SUMMARY

We propose a new model for the Stumpy Induction Factor-induced slender to stumpy transformation of *Trypanosoma brucei gambiense* cells in immunosuppressed mice. The model is a set of delay differential equations that describe the time-course of the infection. We fit the model, using a maximum-likelihood method, to previously published data on parasitaemia in four mice. The model is shown to be a good fit and parameter estimates and confidence intervals are derived. Our estimated parameter values are consistent with estimates from previous experimental studies. The model predicts the following. Slender cells can be classified as uncommitted, committed and dividing, and committed and non-dividing. A committed slender cell undergoes about 5 divisions before exiting the cell-cycle. Committed slender cells must produce SIF, and stumpy cells must not produce SIF. There are two mechanisms for differentiation, a background differentiation rate, and a SIF-concentration-dependent differentiation rate, which is proportional to SIF concentration. SIF has a half-life of about 1·4 h in mice. We also show, with suitable changes in the parameter values, that the model reflects behaviours seen in other host species and trypanosome strains.

Key words: SIF, differentiation, delay differential equation.

INTRODUCTION

African trypanosomes cause widespread disease in humans and livestock. They are transmitted between mammalian hosts by the bite of the tsetse fly. When first transmitted to a mammal, trypanosomes have a long, slender morphology and these forms are not infectious to the tsetse. Slender cells divide rapidly in the bloodstream and other organs before transforming into non-dividing stumpy forms. Stumpy cells are infectious to the tsetse.

Within a host, successive waves of parasitaemia are observed. As the host's immune system fights off one wave, trypanosomes change their antigenic coat and escape to start a new wave (Barry & Turner, 1991). However, the host's immune system is not the only mechanism that causes parasitaemic waves. Trypanosomes are able to regulate their own density by slender to stumpy cell differentiation when their density passes a critical threshold. It has been speculated that this is an evolved trait in order to keep the host alive for prolonged periods – a necessity when tsetse bites are infrequent (Vassella *et al*. 1997).

Black *et al*. (1985) and Seed & Sechelski (1989*b*) first proposed that the slender to stumpy transformation is triggered by a cell density-sensing mechanism, and further work has supported this idea (Seed & Black, 1997; Reuner *et al*. 1997; Vassella *et al*. 1997; Seed & Black, 1999; Tyler *et al*. 2001). Hesse *et al*. (1995) first suggested that a trypanosome-derived factor or metabolite is the mediator of the density-sensing mechanism. Then, in a set of elegant experiments, Vassella *et al*. (1997) demonstrated that the factor is produced by the trypanosomes and that it acts via the cAMP pathway. They termed this factor SIF, short for stumpy induction factor. They also demonstrated that cell-cycle arrest precedes morphological change.

Several mathematical models have been proposed to describe slender to stumpy differentiation (Turner, Aslam & Dye, 1995; Seed & Black, 1997, 1999; Tyler *et al*. 2001). However, none has performed a rigorous statistical analysis on their fit to data. That is, they have not been tested for goodness-of-fit. Nor have their parameters been tested for significance and, in some cases, no confidence limits determined.

In this paper we propose a new model, based on elements of the previous models and extended to include SIF-induced differentiation. The model is fitted to data from previous experiments with immunosuppressed mice (Seed & Sechelski, 1988, 1989*a*; Seed & Black, 1997). By removing the effects of the immune system we can get a better understanding of the density-sensing mechanism.

* Corresponding author: Department of Zoology, Cambridge University, Downing Street, Cambridge CB2 3EJ, UK. Tel: +01223 330933. Fax: +01223 336676. E-mail: njs@zoo.cam.ac.uk

Parasitology (2004), **128**, 53–67.
DOI: 10.1017/S0031182003004256 Printed in the United Kingdom

Our model is fitted, using the maximum-likelihood method, to reconstructed data from the experiments. This gives quantitative estimates of the model parameters. Such estimates give our model strong predictive capacity and the potential to be rigorously tested.

MATERIALS AND METHODS

Previous experiments

Five mice of strain C3HeB/FeJ were irradiated to suppress their immune systems (more details are available in Seed & Black (1997)). Each mouse was infected with 1000 or 5000 *T. brucei gambiense* cells of clone NCTat-1.36. At various times, the tail of each mouse was snipped and a drop of blood obtained. The blood was diluted in 1% glucose phosphate-buffered saline plus 1% sodium citrate (GPBS + C) as an anticoagulant. Whole blood was diluted 1 part blood to 49 parts GPBS + C. If the number of trypanosomes in this suspension was too high to read, further dilutions were made. Two aliquots from each mouse were analysed in a haemacytometer. A second drop of tail blood was smeared on a glass slide, stained with Giemsa, and a minimum of 100 trypanosomes were examined for their morphological type.

One mouse died prematurely so we will not use its data in this paper.

Reconstructing the original data

Unfortunately, the original cell counts from the haemacytometer and staining experiments have been lost. To do a statistical analysis based on the maximum-likelihood method we need to reconstruct these lost cell counts. Fortunately, this is possible because we know the experimental procedure.

The data consist of total cell concentrations and three cell-type concentrations – slender, intermediate and stumpy – at various times. The concentrations are measured in units of number of cells per millilitre of blood. Due to problems of distinguishing between intermediate and stumpy cells we combine these two classes into a single stumpy class.

To reconstruct the total cell counts let us define d to be the dilution made, v_s to be the volume of the dilution counted in the haemacytometer and r to be the number of aliquots. The total volume of blood analysed is, therefore, $v = rv_s/d$. If n_t is the total number of cells counted then the expected total concentration of cells in the blood is $\hat{C}_t = n_t/v$. We want to find n_t having been given v and $\hat{C}_t$. Therefore

$$n_t = v\hat{C}_t, \tag{1}$$

where n_t is rounded to the nearest integer.

To reconstruct the counts for the two cell-types let us define m_t to be the total number of cells counted.

Table 1. Model variables and parameters

	Definition	Dimensions
Variable		
l_1	Concentration of dividing slender cells not committed to differentiate	Cells/ml
l_2	Concentration of dividing slender cells committed to differentiate	Cells/ml
l_3	Concentration of non-dividing slender cells committed to differentiate	Cells/ml
l	Concentration of all slender cells $(l_1 + l_2 + l_3)$	Cells/ml
s	Concentration of stumpy cells	Cells/ml
f	Concentration of SIF	No unit
Parameter		
t_1	Time of first differentiation event	h
t_2	Period from commitment to cell-cycle exit	h
t_3	Period from cell-cycle exit to stumpy morphology	h
a_1	Slender cell birth rate	h^{-1}
a_2	Slender cell background differentiation rate	h^{-1}
a_3	Slender cell, SIF-induced differentiation rate	h^{-1}
a_4	Stumpy cell death rate	h^{-1}
a_5	SIF degradation rate	h^{-1}
c_1	Uncommitted slender cell concentration at time t_1	Cells/ml
c_2	Michaelis–Menten constant	No unit

Let n_l and n_s be the counts of slender and stumpy cells respectively. The expected frequencies of the cell-types are $\hat{f}_l = n_l/m_t$ and $\hat{f}_s = n_s/m_t$. The expected concentrations of the two types are, therefore, $\hat{C}_l = \hat{f}_l\hat{C}_t$ and $\hat{C}_s = \hat{f}_s\hat{C}_t$. We want to find n_l and n_s having been given m_t, $\hat{C}_l$, $\hat{C}_s$ and $\hat{C}_t$. Therefore

$$n_i = m_t\frac{\hat{C}_i}{\hat{C}_t}, \tag{2}$$

where i is l or s, and n_i is rounded to the nearest integer.

As well as reconstructing the cell counts we can calculate the standard errors in the expected total cell concentrations and the cell-type frequencies and concentrations. Although these are not needed for the statistical analysis, they are instructive when plotting the data.

There are two main sources of error in the experiments. The first is due to the difficulty of counting highly motile cells in a haemacytometer. Cells may be counted more than once or not at all if they migrate into adjacent squares. This source of error was minimized by counting a slide 2 or 3 times and taking an average count, and by the same experienced person examining all the slides. In our analysis, we assume that this error is negligible. The second source of error is a sampling error, which we can calculate as follows.

https://doi.org/10.1017/S0031182003004256 Published online by Cambridge University Press

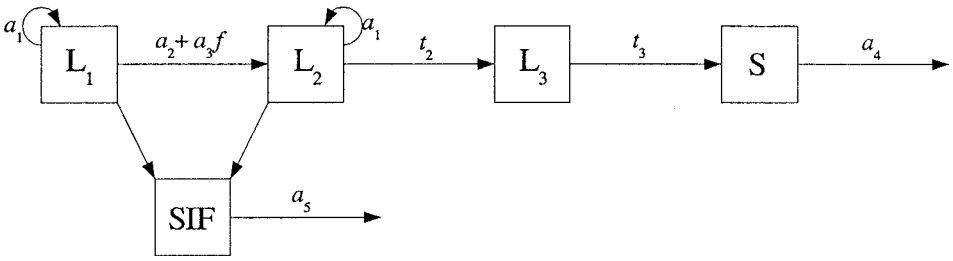

Fig. 1. Schematic representation of the model. Slender cells, denoted L, are split into 3 subclasses: those not committed to differentiate, denoted L_1, those committed to differentiate but still dividing, denoted L_2, and those committed to differentiate and not dividing, denoted L_3. Stumpy cells are denoted S. The parameters are defined in Table 1 and the model details are given in the text.

The standard error of the estimated total cell concentration is

$$e_{C_t} = \frac{\hat{C}_t}{\sqrt{n_t}}. \tag{3}$$

A detailed derivation of this equation, and the ones that follow, are given in Appendix A. The standard errors of the estimated frequencies are

$$e_{f_i} = \sqrt{\frac{\hat{f}_i(1-\hat{f}_i)}{m_t}}, \tag{4}$$

where i is l or s. Thus the standard errors in the estimated concentrations of the two cell-types are

$$e_{C_i} = \sqrt{(\hat{f}_i e_{C_t})^2 + (\hat{C}_t e_{f_i})^2}. \tag{5}$$

The model

The model is a set of delay differential equations (DDEs) that describe how the cell and SIF concentrations change over time. We define the model's variables and parameters in Table 1, and a schematic representation of the model is shown in Fig. 1.

The model equations are

$$\frac{d}{dt}l_1(t) = a_1 l_1(t) - w(t)l_1(t), \tag{6}$$

$$\frac{d}{dt}l_2(t) = a_1 l_2(t) + w(t)l_1(t) - e^{a_1 t_2} w(t-t_2)l_1(t-t_2), \tag{7}$$

$$\frac{d}{dt}l_3(t) = e^{a_1 t_2} w(t-t_2)l_1(t-t_2) - e^{a_1 t_2} w(t-t_2-t_3)l_1(t-t_2-t_3), \tag{8}$$

$$\frac{d}{dt}s(t) = e^{a_1 t_2} w(t-t_2-t_3)l_1(t-t_2-t_3) - a_4 s(t), \tag{9}$$

$$\frac{d}{dt}f(t) = l_1(t) + l_2(t) - a_5 f(t), \tag{10}$$

where

$$w(t) = a_2 + a_3 f(t), \tag{11}$$

is the differentiation rate of uncommitted slender cells. The initial conditions are

$$l_1(t_1) = c_1, \tag{12}$$

$$l_2(t_1) = l_3(t_1) = s(t_1) = f(t_1) = 0. \tag{13}$$

Because we are assuming that no differentiation occurs before time t_1 (see below), we do not need to specify conditions before this time.

The model is based on the following assumptions. (1) The cells and SIF are well-mixed in the blood. (2) The influx and outflux of cells between the blood and any other organs infected by trypanosomes is in equilibrium, that is, can be ignored. (3) The injected slender cells do not immediately begin to differentiate or grow exponentially. Only after t_1 h does the slender cell population begin to grow exponentially, produce SIF and differentiate. The concentration of uncommitted slender cells is set to c_1 cells/ml at this time. (4) Slender cell birth rate is proportional to the instantaneous, dividing slender cell concentration l_1+l_2. This gives exponential growth of the slender cell population as observed experimentally. The birth rate is a_1, and the population doubling time, in the absence of differentiation, is given by $\ln(2)/a_1$. (5) On receiving the differentiation signal, the transitions from one class to the next take fixed periods of time. The uncommitted (L_1) to committed slender cell (L_2) transition is instantaneous. The dividing committed to non-dividing committed slender cell (L_3) transition takes t_2 h. The non-dividing slender cell to stumpy cell (S) transition takes t_3 h. So, once a slender cell is committed to differentiate, it can take several hours for it to become recognizable as a stumpy cell. (6) Because there is a time lag between receiving a differentiation signal and exiting the cell-cycle, a committed slender cell may have many descendants. We assume that all descendants are also committed to differentiate. (7) Stumpy cell death rate is proportional to the instantaneous stumpy cell concentration. The death rate is a_4 and the half-life in h is given by $\ln(2)/a_4$. (8) SIF is produced at a rate proportional to the instantaneous concentration of dividing slender cells. We do not know the production rate of SIF, nor do we know at what

https://doi.org/10.1017/S0031182003004256 Published online by Cambridge University Press

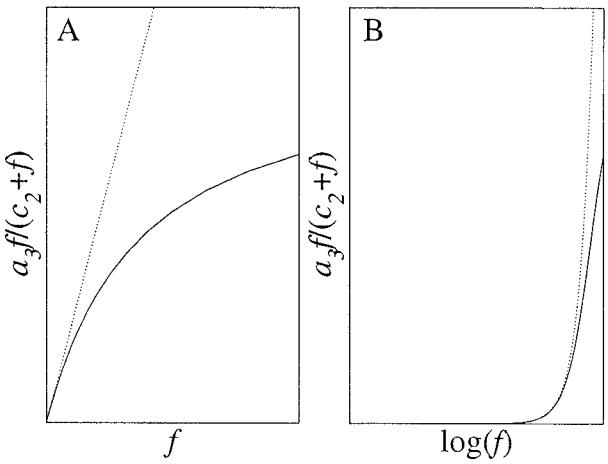

Fig. 2. Solid lines: (A) Michaelis–Menten kinetics of slender cell differentiation rate. (B) As SIF concentration is increasing exponentially, Michaelis–Menten kinetics gives switch like behaviour in the differentiation rate. The dotted lines represent differentiation rate proportional to SIF concentration i.e., $a_3 f$. For an exponentially increasing population, such a rate also gives switch-like behaviour.

concentration it induces differentiation. Therefore, we can normalize the production rate to 1, giving us one less parameter to fit (see Appendix B). (9) SIF degrades exponentially at a rate a_5. Its half-life in hours is given by $\ln(2)/a_5$. (10) There is a background level of differentiation of slender cells, irrespective of SIF concentration. This is necessary because we observe stumpy cells at all times, in particular, during the exponential growth phase. This background rate is constant and occurs at a rate a_2. (11) As cell concentration reaches about 10^8 cells/ml there is a marked increase in the differentiation rate. We must assume something about the functional form for this differentiation rate. The simplest form is linear: that is, proportional to SIF concentration. Another, more biologically motivated mechanism, is described by Michaelis–Menten kinetics. This has the form

$$\frac{a_3 f}{c_2+f}, \tag{14}$$

where a_3 is the maximum differentiation rate, and c_2 the Michaelis–Menten constant. A sketch of this equation is shown in Fig. 2A (solid line). As f increases, differentiation rate saturates at the value a_3. Biologically, this can be interpreted as a saturation of membrane receptors by ligand (e.g., SIF) as the ligand concentration increases.

Because the slender cell concentration and hence SIF concentration, increases exponentially, differentiation rate switches from background to above background very rapidly. This is demonstrated by the sketches in Fig. 2B by using a logarithmic x-axis.

Numerical schemes

A fourth-order Runge–Kutta scheme with fixed time-step h was used to solve the equations. To use this scheme we must convert our model from a finite set of delay differential equations to an infinite set of ordinary differential equations. The reformulated model is given in Appendix C. All numerical solutions were done with $h=0{\cdot}25$. This was checked to make sure the solutions were stable and accurate.

Curve fitting

The model was fitted to the data using the maximum-likelihood method. For a particular set of parameters and data we calculate the likelihood of these parameters given the data. We then adjust the parameters in order to maximize the likelihood, thus giving us our maximum-likelihood parameter estimates. The likelihood is found as follows.

We know that, if the total cell concentration is C_t and if we analyse a volume v of blood in a haemacytometer then, the probability of counting n_t cells is given by a Poisson distribution

$$\text{Prob}(n_t|C_t, v)=\frac{(vC_t)^{n_t}}{n_t!e^{vC_t}}. \tag{15}$$

We also know that, if the frequencies of the two cell-types are f_l and f_s then, the probability of counting n_l and n_s cells, given that we count m_t cells in total, is given by a binomial distribution

$$\text{Prob}(n_l, n_s|f_l, f_s, m_t)=f_l^{n_l}f_s^{n_s}\frac{m_t!}{n_l!n_s!}. \tag{16}$$

The likelihood of a set of parameters given the data is proportional to the product of all the probabilities of the data given the simulated results of the model. That is,

$$L\propto\prod_{j=1}^{7}\frac{(v_jC_t(t_j))^{n_{t_j}}}{n_{t_j}!e^{v_jC_t(t_j)}}f_l(t_j)^{n_{l_j}}f_s(t_j)^{n_{s_j}}\frac{m_{t_j}!}{n_{l_j}!n_{s_j}!}. \tag{17}$$

where j denotes a particular data point. C_t, f_l and f_s are found by simulating the model for a particular parameter set. v_j, m_{t_j}, n_{t_j}, n_{l_j}, n_{s_j} are the data at time t_j.

For convenience the natural logarithm of the likelihood is calculated. The maximum-(log)-likelihood, $l_{\max}=\ln(L_{\max})$, is found using the Marquis–Levenberg algorithm (Press *et al.* 1992) with a tolerance of 10^{-6}.

Statistical analysis

Goodness-of-fit. Once the best-fit parameters for a particular data set have been found we want to test the goodness-of-fit of these parameters. We do this by comparing the observed maximum-likelihood ($l_{\max}$) to its expected distribution. The expected distribution is found by simulating artificial data sets from the best-fit model. We can do this because we know the experimental procedure. We simulate 100 data sets and calculate their maximum likelihoods.

https://doi.org/10.1017/S0031182003004256 Published online by Cambridge University Press

Table 2. Given cell concentrations and experimental parameters

Time (h)	$\hat{C}_t$ (cells/ml)	$\hat{C}_l$ (cells/ml)	$\hat{C}_s$ (cells/ml)	Haemacytometer	v_s (ml)	d	r	v (ml)	m_t (cells)
Mouse 1									
70·5	$2{\cdot}50\times10^{5}$	$2{\cdot}14\times10^{5}$	$3{\cdot}38\times10^{4}$	All	9×10^{-4}	50	2	$3{\cdot}6\times10^{-5}$	100
75·5	$2{\cdot}15\times10^{6}$	$1{\cdot}99\times10^{6}$	$1{\cdot}44\times10^{5}$	Central	1×10^{-4}	50	2	$4{\cdot}0\times10^{-6}$	100
96·0	$3{\cdot}93\times10^{7}$	$3{\cdot}43\times10^{7}$	$4{\cdot}83\times10^{6}$	5 small	2×10^{-5}	50	2	$8{\cdot}0\times10^{-7}$	100
124·5	$7{\cdot}26\times10^{8}$	$3{\cdot}59\times10^{8}$	$3{\cdot}64\times10^{8}$	5 small	2×10^{-5}	2500	2	$1{\cdot}6\times10^{-8}$	100
137·0	$9{\cdot}80\times10^{8}$	$2{\cdot}76\times10^{7}$	$9{\cdot}54\times10^{8}$	5 small	2×10^{-5}	2500	2	$1{\cdot}6\times10^{-8}$	100
148·0	$1{\cdot}14\times10^{9}$	$3{\cdot}19\times10^{7}$	$1{\cdot}11\times10^{9}$	5 small	2×10^{-5}	2500	2	$1{\cdot}6\times10^{-8}$	100
163·0	$1{\cdot}25\times10^{9}$	$1{\cdot}49\times10^{8}$	$1{\cdot}10\times10^{9}$	5 small	2×10^{-5}	2500	2	$1{\cdot}6\times10^{-8}$	100
Mouse 2									
70·5	$2{\cdot}88\times10^{5}$	$2{\cdot}49\times10^{5}$	$3{\cdot}86\times10^{4}$	All	9×10^{-4}	50	2	$3{\cdot}6\times10^{-5}$	100
75·5	$9{\cdot}00\times10^{5}$	$8{\cdot}07\times10^{5}$	$9{\cdot}00\times10^{4}$	All	9×10^{-4}	50	2	$3{\cdot}6\times10^{-5}$	100
96·0	$4{\cdot}30\times10^{7}$	$3{\cdot}47\times10^{7}$	$7{\cdot}96\times10^{6}$	5 small	2×10^{-5}	50	2	$8{\cdot}0\times10^{-7}$	100
124·5	$9{\cdot}20\times10^{8}$	$5{\cdot}19\times10^{8}$	$3{\cdot}96\times10^{8}$	5 small	2×10^{-5}	2500	2	$1{\cdot}6\times10^{-8}$	100
137·0	$1{\cdot}72\times10^{9}$	$2{\cdot}41\times10^{7}$	$1{\cdot}69\times10^{9}$	5 small	2×10^{-5}	2500	2	$1{\cdot}6\times10^{-8}$	100
148·0	$1{\cdot}60\times10^{9}$	$1{\cdot}92\times10^{7}$	$1{\cdot}58\times10^{9}$	5 small	2×10^{-5}	2500	2	$1{\cdot}6\times10^{-8}$	100
163·0	$8{\cdot}24\times10^{8}$	$3{\cdot}79\times10^{7}$	$7{\cdot}86\times10^{8}$	5 small	2×10^{-5}	2500	2	$1{\cdot}6\times10^{-8}$	100
Mouse 3									
70·5	$6{\cdot}25\times10^{5}$	$5{\cdot}19\times10^{5}$	$1{\cdot}04\times10^{5}$	All	9×10^{-4}	50	2	$3{\cdot}6\times10^{-5}$	100
75·5	$2{\cdot}15\times10^{6}$	$2{\cdot}06\times10^{6}$	$8{\cdot}17\times10^{4}$	Central	1×10^{-4}	50	2	$4{\cdot}0\times10^{-6}$	100
96·0	$4{\cdot}80\times10^{7}$	$3{\cdot}16\times10^{7}$	$1{\cdot}63\times10^{7}$	5 small	2×10^{-5}	50	2	$8{\cdot}0\times10^{-7}$	100
124·5	$1{\cdot}18\times10^{9}$	$3{\cdot}54\times10^{8}$	$8{\cdot}18\times10^{8}$	5 small	2×10^{-5}	2500	2	$1{\cdot}6\times10^{-8}$	100
137·0	$1{\cdot}32\times10^{9}$	$9{\cdot}24\times10^{6}$	$1{\cdot}31\times10^{9}$	5 small	2×10^{-5}	2500	2	$1{\cdot}6\times10^{-8}$	100
148·0	$1{\cdot}16\times10^{9}$	$2{\cdot}20\times10^{7}$	$1{\cdot}14\times10^{9}$	5 small	2×10^{-5}	2500	2	$1{\cdot}6\times10^{-8}$	100
163·0	$8{\cdot}00\times10^{8}$	$5{\cdot}68\times10^{7}$	$7{\cdot}41\times10^{8}$	5 small	2×10^{-5}	2500	2	$1{\cdot}6\times10^{-8}$	100
Mouse 4									
70·5	$5{\cdot}75\times10^{5}$	$4{\cdot}93\times10^{5}$	$8{\cdot}05\times10^{4}$	All	9×10^{-4}	50	2	$3{\cdot}6\times10^{-5}$	100
75·5	$5{\cdot}75\times10^{5}$	$5{\cdot}29\times10^{5}$	$4{\cdot}60\times10^{4}$	All	9×10^{-4}	50	2	$3{\cdot}6\times10^{-5}$	100
96·0	$2{\cdot}36\times10^{7}$	$1{\cdot}53\times10^{7}$	$8{\cdot}28\times10^{6}$	5 small	2×10^{-5}	50	2	$8{\cdot}0\times10^{-7}$	100
124·5	$7{\cdot}72\times10^{8}$	$4{\cdot}26\times10^{8}$	$3{\cdot}45\times10^{8}$	5 small	2×10^{-5}	2500	2	$1{\cdot}6\times10^{-8}$	100
137·0	$1{\cdot}18\times10^{9}$	$9{\cdot}68\times10^{7}$	$1{\cdot}10\times10^{9}$	5 small	2×10^{-5}	2500	2	$1{\cdot}6\times10^{-8}$	100
148·0	$1{\cdot}54\times10^{9}$	$5{\cdot}24\times10^{7}$	$1{\cdot}48\times10^{9}$	5 small	2×10^{-5}	2500	2	$1{\cdot}6\times10^{-8}$	100
163·0	$9{\cdot}32\times10^{8}$	$7{\cdot}34\times10^{7}$	$8{\cdot}56\times10^{8}$	5 small	2×10^{-5}	2500	2	$1{\cdot}6\times10^{-8}$	100

These are then binned to create a distribution. The observed $l_{\max}$ is compared to its expected distribution. If the observed $l_{\max}$ falls above 5% of the values in the expected distribution then the parameters are considered a good fit. Otherwise they are considered a poor fit.

Parameter significance. Parameter significance is tested by setting each one to zero and recalculating the maximum-likelihood. However, to test the significance of the initial slender cell concentration, c_1 is set to 1000, otherwise there would be no cells in the blood. Significance is determined by using the likelihood-ratio test. If the difference between the original and recalculated maximum-likelihoods is greater than 1, then the parameter is significant at the 5% level.

Parameter error estimates. To estimate the errors in the parameter values we assume that the best fit parameters to a particular data set are the true parameters. Using these parameters, we simulate 100 artificial data sets and calculate their maximum-likelihoods. This gives 100 sets of estimated parameters from which we can construct a probability distribution for each parameter.

RESULTS

The reconstructed data

In Tables 2 and 3 we tabulate the original and reconstructed data for each mouse. The total number of cells counted to determine the frequency of each type (m_t) is unknown. Therefore, we assumed that 100 cells were counted for each case. The data and error bars for the four mice are shown in Fig. 3.

Best-fit curves

The best-fit curves for all mice are shown in Fig. 4. The maximum-likelihoods are tabulated in Table 4.

Goodness-of-fit

The goodness-of-fit of each mouse is given in Table 4.

The fits are good for mice 1 and 2, but very poor for mice 3 and 4. This might be caused by outliers in

https://doi.org/10.1017/S0031182003004256 Published online by Cambridge University Press

Table 3. Reconstructed raw data

Time (h)	$\hat{f}_l$	$\hat{f}_s$	n_t (cells)	n_l (cells)	n_s (cells)	e_{C_t} (cells/ml)	e_{C_l} (cells/ml)	e_{C_s} (cells/ml)
Mouse 1								
70·5	0·86	0·14	9	86	14	$8{\cdot}33\times10^{4}$	$7{\cdot}25\times10^{4}$	$1{\cdot}42\times10^{4}$
75·5	0·93	0·07	9	93	7	$7{\cdot}33\times10^{5}$	$6{\cdot}86\times10^{5}$	$7{\cdot}32\times10^{4}$
96·0	0·88	0·12	31	88	12	$7{\cdot}01\times10^{6}$	$6{\cdot}28\times10^{6}$	$1{\cdot}56\times10^{6}$
124·5	0·50	0·50	12	50	50	$2{\cdot}13\times10^{8}$	$1{\cdot}12\times10^{8}$	$1{\cdot}13\times10^{8}$
137·0	0·03	0·97	16	3	97	$2{\cdot}47\times10^{8}$	$1{\cdot}76\times10^{7}$	$2{\cdot}41\times10^{8}$
148·0	0·03	0·97	18	3	97	$2{\cdot}67\times10^{8}$	$2{\cdot}03\times10^{7}$	$2{\cdot}60\times10^{8}$
163·0	0·12	0·88	20	12	88	$2{\cdot}80\times10^{8}$	$5{\cdot}26\times10^{7}$	$2{\cdot}49\times10^{8}$
Mouse 2								
70·5	0·87	0·13	10	87	13	$8{\cdot}94\times10^{4}$	$7{\cdot}81\times10^{4}$	$1{\cdot}55\times10^{4}$
75·5	0·90	0·10	32	90	10	$1{\cdot}58\times10^{5}$	$1{\cdot}45\times10^{5}$	$3{\cdot}14\times10^{4}$
96·0	0·81	0·19	34	81	19	$7{\cdot}33\times10^{6}$	$6{\cdot}19\times10^{6}$	$2{\cdot}16\times10^{6}$
124·5	0·57	0·43	15	57	43	$2{\cdot}40\times10^{8}$	$1{\cdot}43\times10^{8}$	$1{\cdot}13\times10^{8}$
137·0	0·01	0·99	28	1	99	$3{\cdot}28\times10^{8}$	$2{\cdot}08\times10^{7}$	$3{\cdot}24\times10^{8}$
148·0	0·01	0·99	26	1	99	$3{\cdot}16\times10^{8}$	$1{\cdot}78\times10^{7}$	$3{\cdot}13\times10^{8}$
163·0	0·05	0·95	13	5	95	$2{\cdot}27\times10^{8}$	$2{\cdot}02\times10^{7}$	$2{\cdot}17\times10^{8}$
Mouse 3								
70·5	0·83	0·17	23	83	17	$1{\cdot}32\times10^{5}$	$1{\cdot}12\times10^{5}$	$3{\cdot}20\times10^{4}$
75·5	0·96	0·04	9	96	4	$7{\cdot}33\times10^{5}$	$7{\cdot}06\times10^{5}$	$4{\cdot}98\times10^{4}$
96·0	0·66	0·34	38	66	34	$7{\cdot}75\times10^{6}$	$5{\cdot}59\times10^{6}$	$3{\cdot}48\times10^{6}$
124·5	0·30	0·70	19	30	70	$2{\cdot}72\times10^{8}$	$9{\cdot}83\times10^{7}$	$1{\cdot}97\times10^{8}$
137·0	0·01	0·99	21	1	99	$2{\cdot}87\times10^{8}$	$1{\cdot}12\times10^{7}$	$2{\cdot}85\times10^{8}$
148·0	0·02	0·98	19	2	98	$2{\cdot}69\times10^{8}$	$1{\cdot}66\times10^{7}$	$2{\cdot}65\times10^{8}$
163·0	0·07	0·93	13	7	93	$2{\cdot}24\times10^{8}$	$2{\cdot}60\times10^{7}$	$2{\cdot}09\times10^{8}$
Mouse 4								
70·5	0·86	0·14	21	86	14	$1{\cdot}26\times10^{5}$	$1{\cdot}10\times10^{5}$	$2{\cdot}67\times10^{4}$
75·5	0·92	0·08	21	92	8	$1{\cdot}26\times10^{5}$	$1{\cdot}17\times10^{5}$	$1{\cdot}86\times10^{4}$
96·0	0·65	0·35	19	65	35	$5{\cdot}43\times10^{6}$	$3{\cdot}70\times10^{6}$	$2{\cdot}22\times10^{6}$
124·5	0·55	0·45	12	55	45	$2{\cdot}20\times10^{8}$	$1{\cdot}27\times10^{8}$	$1{\cdot}05\times10^{8}$
137·0	0·08	0·92	19	8	92	$2{\cdot}72\times10^{8}$	$3{\cdot}90\times10^{7}$	$2{\cdot}52\times10^{8}$
148·0	0·03	0·97	25	3	97	$3{\cdot}10\times10^{8}$	$3{\cdot}00\times10^{7}$	$3{\cdot}01\times10^{8}$
163·0	0·08	0·92	15	8	92	$2{\cdot}41\times10^{8}$	$3{\cdot}15\times10^{7}$	$2{\cdot}24\times10^{8}$

the data. For mouse 3, potential outliers are the stumpy cell concentrations at 70·5 and 75·5 h. For mouse 4, potential outliers are the slender and stumpy cell concentrations at 70·5 and 75·5 h.

For mouse 3, removing the stumpy cell concentration at 75·5 h gives a much better fit (see Table 5 and the dotted lines in Fig. 4C). Removing the stumpy cell concentration at 70·5 h also gives a good fit (result not shown). For mouse 4, removing all data points at 75·5 h gives a much better fit (see Table 5 and the dotted lines in Fig. 4D). Removing all data points at 70·5 h also gives a good fit (result not shown). We denote these modified fits as corresponding to mouse 3′ and mouse 4′.

Parameter significance of the best-fit model

The maximum-likelihoods of the model with constrained parameters for all mice are shown in Table 6. Parameter significance for a particular mouse is denoted by a tick.

Parameters t_1 and c_1 set up the initial conditions. Parameter t_1 is marginally significant, which probably implies that differentiation does not begin immediately at the start of the experiment, but it can be compensated for by changes in c_1 during the fitting procedure. Parameter c_1 is not significant: setting it to 1000 cells/ml can be compensated for by changes to t_1. Parameter a_1, the slender cell birth rate, is clearly significant. Parameter a_2, the background slender cell differentiation rate is significant. This is because differentiation is taking place even before SIF concentration has reached levels high enough to limit the parasitaemia after 124·5 h. Parameter a_3, the above-background slender cell differentiation rate, is highly significant. Without this parameter there would be no self-limitation of the parasitaemia, that is, only exponential growth. Parameter a_4, the stumpy cell death rate, is only marginally significant in mouse 3′. This is because the drop in stumpy cell concentration is only captured by the last data point at 163 h. Parameter t_2 is significant as without it there are no oscillations in slender cell concentration (see below). Parameter t_3 is marginally significant as removing it only has a small effect on the dynamics. Parameter a_5 is significant, as without SIF degradation the slender cell concentration would not recover after its drop.

https://doi.org/10.1017/S0031182003004256 Published online by Cambridge University Press

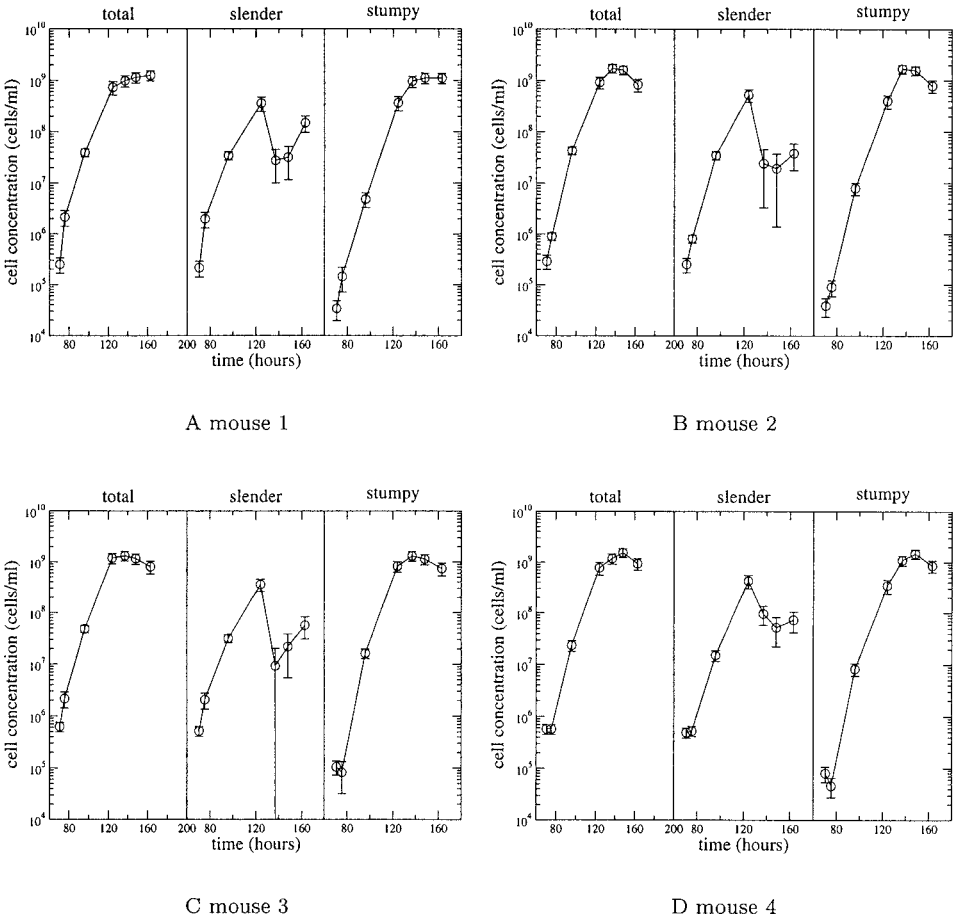

Fig. 3. Experimental results (circles) with error bars (±1 S.E.M.). One error bar extends to zero. The x-axis gives the time since the start of the experiment, and the y-axis the number of cells/ml on a log scale.

Parameter estimates

The parameter estimates for each mouse are given in Table 7. In Fig. 5 we show boxplots of the parameter distributions. The median is shown as a horizontal line contained within a box that bounds 50% of the values. Small circles show outliers in the distributions. Even though the stumpy cell death rate is not a significant parameter, it is included because there is strong evidence that stumpy cells do die (Black, Hewett & Sendashonga, 1982; Turner *et al.* 1995).

Many of the parameters show discrepancies between mice 1 and 2 and mice 3′ and 4′. This is most likely due to the removal of the outliers at 75·5 h for mice 3′ and 4′. Parameter a_4 is much smaller for mouse 1 than for the other mice. This is because the stumpy cell concentration in this mouse does not fall near the end of the experiment as it does in the others.

There is quite a large variation in parameters a_3 and a_5. By de-dimensionalizing the model we can show that the differential equation describing SIF rate of change can be written as

$$\frac{df}{dt} = l_1 + l_2 - \frac{a_5}{a_3} f. \tag{18}$$

Thus, changing a_3 and a_5 by an equal factor does not change the solution. For the best-fit model, a_5/a_3 varies from 3 to 5×10^8 across the four mice. This is a much smaller variation than when the parameters are treated separately. This suggests that the variations in the separate parameters are an artifact of the fitting procedure and not true variations between the mice.

Having found our best-fit model we can now make some further predictions. We can calculate differentiation rate, extrapolate the simulations forward in time, and change parameter values and observe their effects on the dynamics.

Differentiation rate

We can determine the instantaneous differentiation rate of slender cells from class L_1 to L_2. This is given by $a_2 + a_3 f(t)$. In Fig. 6 we plot these curves for the four mice. Even though there is large variation in the SIF-concentration-dependent differentiation rate (a_3), there is very little variation in the overall rates shown in Fig. 6. This is due to the large variation in the amount of SIF produced – which is a function of the slender cell concentration. This again suggests that the variation in parameters a_3 and a_5 is a consequence of the fitting procedure and not a true indication of variation between the mice.

https://doi.org/10.1017/S0031182003004256 Published online by Cambridge University Press

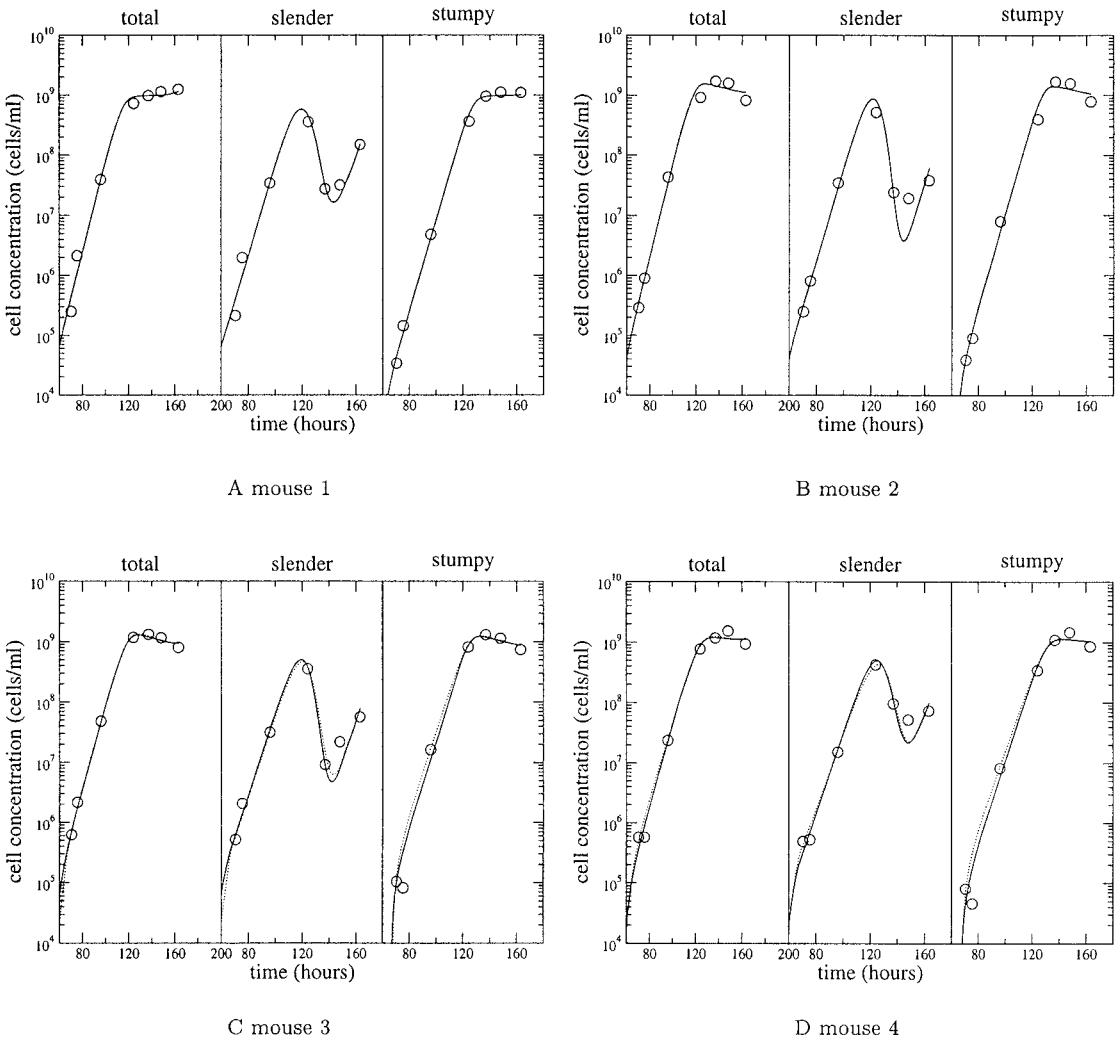

Fig. 4. Best-fit curves. Circles are observed values, lines are best fits. The x-axis shows time since beginning of experiment, the y-axis shows cell concentration in cells/ml on a log scale. Dotted lines for mouse 3 shows best-fit where cell type concentrations at 75·5 h have been omitted from the fit. Dotted lines for mouse 4 shows best-fit where total and cell type concentrations at 75·5 h have been omitted from the fit.

Table 4. The maximum-likelihoods and goodness-of-fits for each mouse

Mouse	l_{max}	gof
1	−36	15
2	−36	15
3	−45	0
4	−45	0

Table 5. The maximum-likelihoods and goodness-of-fits for mouse 3′ and 4′

Mouse	l_{max}	gof
3′	−29	72
4′	−30	36

Peak differentiation rate is 3–5 times higher than background differentiation rate. Which are both several orders of magnitude higher than stumpy cell death rate.

Extrapolating the simulations

The mice experiments were halted at 163 h. The solid line in Fig. 7 shows what happens when we extrapolate the model past this time using the best-fit parameters for mouse 3. The oscillations continue indefinitely, and the total cell concentration plateaus at around 2×10^9 cells/ml. Numerical analysis (not shown) suggests that the dynamics undergo a Hopf bifurcation as t_2 is increased, that is, an abrupt change from stable to decaying oscillations. The dotted line in Fig. 7 demonstrates this when t_2 is increased by 40%.

Which cells produce SIF?

To determine which cells produce SIF we calculated the maximum-likelihood for each mouse with different combinations of cells producing SIF. We assume

https://doi.org/10.1017/S0031182003004256 Published online by Cambridge University Press

Table 6. Parameter significance

Constrained parameter(s)	l_{max}				Significant			
	1	2	3′	4′	1	2	3′	4′
None	−36	−36	−29	−30				
t_1	−40	−37	−33	−34	✓		✓	✓
c_1	−36	−37	−30	−30				
a_1	−1380	−685	−1311	−794	✓	✓	✓	✓
a_2	−153	−189	−147	−143	✓	✓	✓	✓
a_3	−240	−292	−165	−150	✓	✓	✓	✓
a_4	−36	−37	−31	−30			✓	
t_2	−44	−48	−40	−42	✓	✓	✓	✓
t_3	−46	−39	−30	−34	✓	✓		✓
a_5	−80	−57	−96	−51	✓	✓	✓	✓
M–M	−36	−36	−29	−30				

Table 7. Estimated parameter values for best-fit model

Parameter	Value			
	1	2	3′	4′
t_1	46	48	53	52
a_1	0·33	0·32	0·42	0·38
a_2	0·15	0·15	0·26	0·24
a_3	$5{\cdot}6\times10^{-10}$	$8{\cdot}6\times10^{-10}$	$1{\cdot}5\times10^{-9}$	$2{\cdot}8\times10^{-9}$
a_4	$1{\cdot}2\times10^{-7}$	$1{\cdot}2\times10^{-2}$	$1{\cdot}6\times10^{-2}$	$1{\cdot}2\times10^{-2}$
t_2	8·1	11	12	12
t_3	7·9	6·4	2·9	3·6
a_5	0·17	0·40	0·62	1·4

that each cell-type produces the same amount of SIF. The results are given in Table 8. For all mice, when stumpy cells produce SIF, there is a significant decrease in the maximum-likelihood compared to the best-fit model. In Fig. 8 the dashed line shows that when stumpy cells produce SIF, too much is produced and the slender cell concentration drops too low. When only uncommitted slender cells produce SIF, there is a significant decrease in the maximum-likelihood compared to the best-fit model (Table 8). In Fig. 8 the dotted line shows that when only uncommitted cells produce SIF, too little is produced and the slender cell concentration does not drop far enough. Whether or not non-dividing slender cells produce SIF is inconclusive (Table 8).

Michaelis–Menten kinetics

We have assumed that the SIF-concentration-dependent differentiation rate is proportional to SIF concentration i.e. $a_3 f$. Although this is the simplest assumption we could make, it is not necessarily correct. As mentioned above, a more biologically motivated example is Michaelis–Menten kinetics, modelled by equation 14. Note that equation 14 approximates a linear differentiation rate if c_2 is very much larger than peak SIF concentration.

We tested the significance of Michaelis–Menten kinetics (Table 6, last row). It did not give a significantly better fit to the data for any mouse. We conclude, therefore, that a linear SIF-concentration-dependent differentiation rate gives the best fit to the data.

Effects of changing parameter values

By changing individual parameter values we can understand how these parameters influence the dynamics.

In Fig. 9A the solid lines are the best-fit curves of mouse 2. The dotted lines demonstrate what happens when slender cell birth rate is increased by 20%. As expected, the slender cell population increases more rapidly. More interestingly, this causes a higher parasitaemia, which leads to a huge production of SIF and consequently a massive decline in slender cell concentration. Stumpy cell concentration is, therefore, greatly increased. The dashed line demonstrates the case when birth rate is reduced by a factor of 2. The oscillatory dynamics are lost, both cell types increase less rapidly and stumpy cells dominate much earlier.

Changing the background differentiation rate (a_2) has a similar but opposite effect to changing

https://doi.org/10.1017/S0031182003004256 Published online by Cambridge University Press

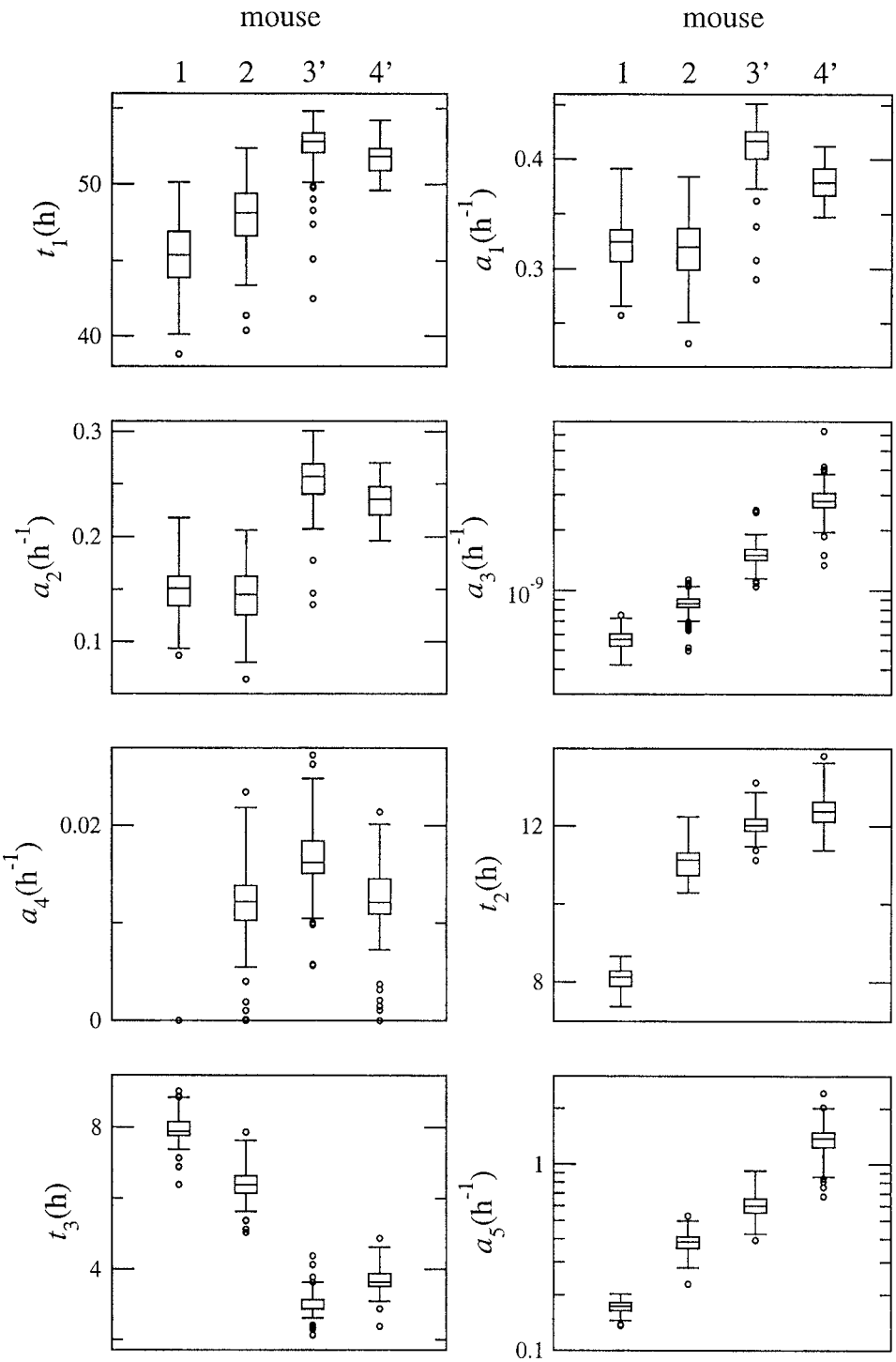

Fig. 5. Boxplots of the estimated parameter distributions. The horizontal line is the median, the box bounds 50% of the values. Circles show outliers.

parameter a_1. This is not unexpected because committed slender cell growth rate is proportional to both (see equation 6). The only major difference is that when $a_2 = 0$, stumpy cells appear a little later in the simulation.

The dotted line in Fig. 9C shows the case when parameter a_3 has been increased 500-fold over the best-fit model. This is equivalent to slender cells producing 500 times more SIF or being 500 times more sensitive to SIF. Peak parasitaemia is reduced to about 10^6 cells/ml. This is closer to peak parasitaemia seen in domestic cattle. The dashed line represents a 5-fold decrease in parameter a_3 compared to the best-fit model. Peak parasitaemia is increased and the peak is delayed for both slender and stumpy cells.

The dotted line in Fig. 9D shows what happens when the stumpy half-life is 1 h. Slender cell concentration is not affected, but because stumpy cells die so rapidly, their concentration oscillates and

https://doi.org/10.1017/S0031182003004256 Published online by Cambridge University Press

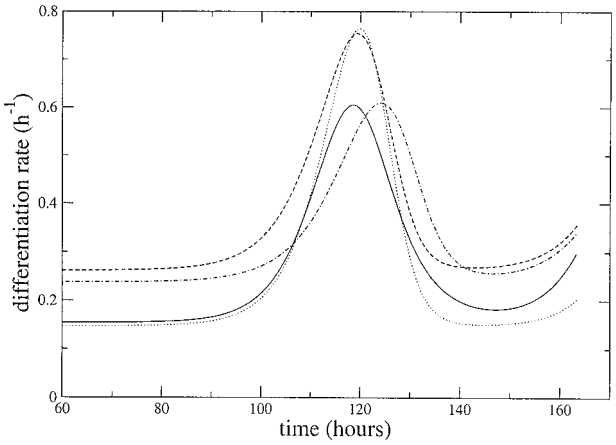

Fig. 6. Instantaneous differentiation rate from L_1 to L_2 class ($a_2+a_3f(t)$). Solid line: mouse 1; dotted line: mouse 2; dashed line: mouse 3′; dot-dashed line: mouse 4′.

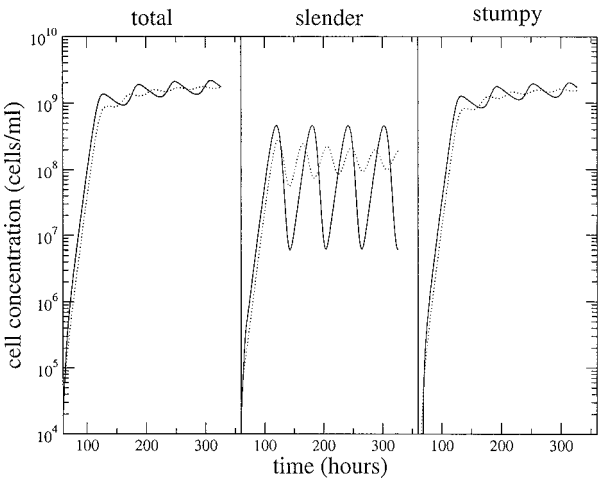

Fig. 7. Extrapolation of the best-fit model for mouse 3 (solid line) shows stable oscillations. Increasing the L_2 to L_3 transition time by 40% causes the oscillations to decay (dotted line).

never passes slender cell concentration. When stumpy cells do not die (dashed line) then their concentration saturates. This is seen in the data of mouse 1.

In Fig. 9E we have increased the L_2 to L_3 transition period (t_2) by 100% (dotted lines) and set it to zero (dashed lines). A longer transition period implies more slender cells, hence the curves shifting left. This also means more SIF is produced and the slender population does not recover after its decline. For instantaneous transition the slender cells do not show oscillatory behaviour.

In Fig. 9F and G we have increased the L_3 to stumpy transition period (t_3) by 10-fold (dotted lines) and set it to zero (dashed lines). In Fig. 9F only L_1 and L_2 cells produce SIF. In Fig. 9G L_1, L_2 and L_3 cells produce SIF. In both cases, an increased transition period (dotted lines) prevents oscillations and delays the occurrence of stumpy cells. These curves could be interpreted as arising from monomorphic strains if the host died before peak

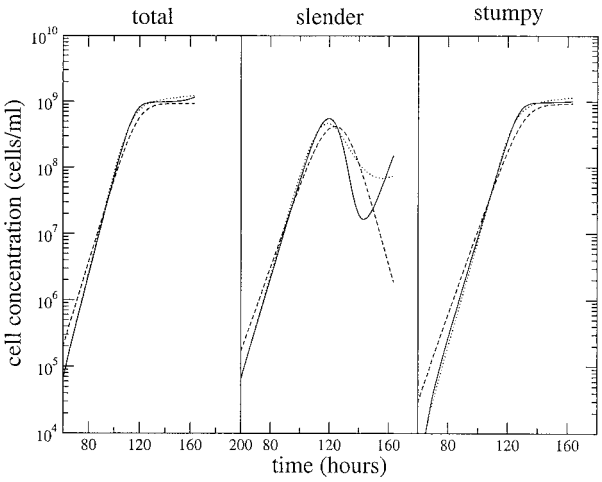

Fig. 8. If stumpy cells produce SIF then too much is produced and the slender cell concentration drops too low (dashed line). If only uncommitted cells produce SIF then too little is produced and the slender cell concentration does not drop far enough (dotted line). The best-fit model for mouse 1 was used.

Table 8. Maximum-likelihoods for different combinations of cells producing SIF

Mouse	L_1	L_1, L_2 (best)	L_1, L_2, L_3	L_1, L_2, L_3, S
1	−48	−36	−36	−100
2	−48	−36	−37	−101
3′	−167	−29	−30	−77
4′	−157	−30	−30	−92

parasitaemia. Instantaneous transition (dashed lines) causes a drop in peak parasitaemia and earlier occurrence of stumpy cells. Moreover, if L_3 cells produce SIF then the amplitude and period of the oscillations are reduced.

In Fig. 9H we increased SIF degradation rate (a_5) by 100% (dotted lines) and set it to zero (dashed lines). This parameter affects the peak parasitaemia and its timing, and the magnitude and period of slender cell oscillations. When set to zero all slender cells differentiate as one would expect.

DISCUSSION

From the best-fit model we can make the following conclusions. Slender cells divide at a rate of approximately 0·33/h. This corresponds to a doubling time of about 2·1 h. This is much shorter than previously reported: 4·18 h by Seed & Black (1997) and about 5 h for ANTat-1.1 and MITat-1.2 cells by Vassella *et al.* (1997). Part of this discrepancy may be because these estimates were based on the growth rate of slender cells. However, growth rate equals birth rate minus differentiation rate. For example, the birth rate of cells in mouse 1 is 0·33/h and the differentiation rate, when cell concentration is low, is 0·15/h. Thus the growth rate is 0·18/h giving a

https://doi.org/10.1017/S0031182003004256 Published online by Cambridge University Press

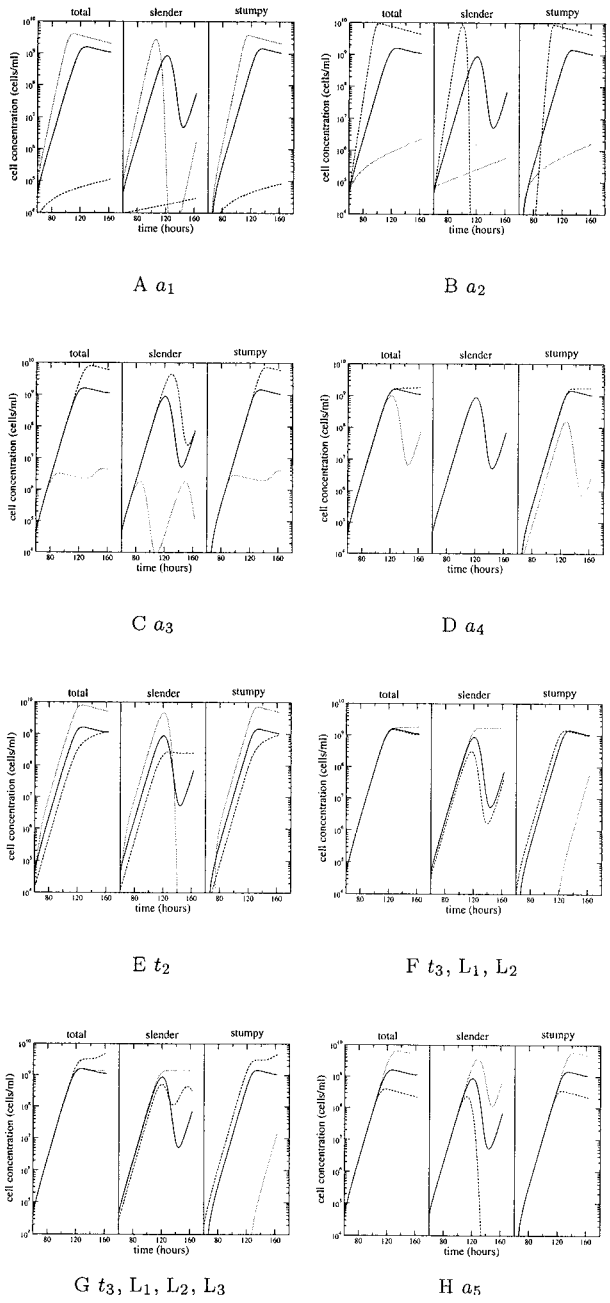

Fig. 9. Effect of changing the parameter values for the best fit model of mouse 2. (A) Dotted: $a_1 \times 1{\cdot}2$, dashed: $a_1/2$. (B) Dotted: $a_2 \times 2$, dashed: $a_2 = 0$. (C) Dotted: $a_3 \times 500$, dashed: $a_3/5$. (D) Dotted: $a_4 = 0{\cdot}7$, dashed: $a_4 = 0$. (E) Dotted: $t_2 \times 2$, dashed: $t_2 = 0$. (F) Dotted: $t_3 \times 10$, dashed: $t_3 = 0$, L_1 and L_2 produce SIF. (G) Dotted: $t_3 \times 10$, dashed: $t_3 = 0$, L_1, L_2 and L_3 produce SIF. (H) Dotted: $a_5 \times 2$, dashed: $a_5 = 0$.

doubling time of 3·9 h. This is much closer to the previously published 4–5 h. The other mice give similar results.

There appear to be two mechanisms that can induce slender cell differentiation. The first is what we have termed a background rate. This rate is constant (parameter a_2 in the model) and is, therefore, independent of SIF concentration. However, this does not mean that this mechanism is independent of the existence of SIF – just its concentration. The rate of background differentiation is approximately 0·15/h. The second is a SIF concentration-dependent mechanism. We found that a linear dependence with SIF concentration is sufficient to explain the observed data. The more complicated Michaelis–Menten kinetics did not significantly improve the fits. This second mechanism only becomes effective at high slender cell concentrations. It acts somewhat like a switch, causing a 3 to 5-fold increase in the differentiation rate when the exponentially expanding slender cell population passes 10^8 cells/ml. (Hence the need for the background rate a_2, otherwise we would not observe differentiated, stumpy cells before this time.) The nature of these two mechanisms is unknown. We can speculate that background differentiation is caused by an autocrine signal since slender cells produce SIF which causes them to differentiate, and that SIF concentration-dependent differentiation is caused by a paracrine signal. More experimental work is needed to test this hypothesis.

SIF is produced only by slender cells. We predict that stumpy cells should not produce SIF. We also predict that committed slender cells should produce SIF. Whether or not non-dividing slender cells produce SIF is inconclusive.

It has been shown that commitment to differentiation precedes the final cell division (Tyler *et al.* 1997). Our best-fit model is consistent with this view, as parameter t_2 is significant. However, our model predicts that the time-period from commitment to cell-cycle exit is around 10 h. This implies about 5 cell divisions between commitment and cell-cycle exit. This is an unexpected result and requires further theoretical and experimental analysis to test its validity.

Once committed cells exit the cell-cycle we predict that they keep a slender morphology for about another 6 h (parameter t_3). This is slightly shorter than that found by Vassella *et al.* (1997). They estimated from data a period of 8–10 h from exit of the cell-cycle to mitochondrion metabolic activity.

We predict that stumpy cells die at a rate of approximately 0·012/h (parameter a_4). This is a half-life of about 58 h; higher than the estimate of 24–36 h reported by Black *et al.* (1982), and between the estimates of 48 and 72 h reported by Turner *et al.* (1995). However, the non-significance of this parameter, due to insufficient data, means that its estimate should be used with caution. If data had been collected after 163 h it may have increased the significance of this parameter.

We predict that SIF decays at a rate of approximately 0·5/h. This is a half-life of about 1·4 h. This is consistent with SIF being a pheromone-like factor or a small catabolite (Vassella *et al.* 1997).

This, and previous models (Seed & Black, 1997, 1999; Tyler *et al.* 2001), predict oscillations in the slender and stumpy cell concentrations past the end

https://doi.org/10.1017/S0031182003004256 Published online by Cambridge University Press

of the experiments. The Tyler *et al*. (2001) model exhibits decaying oscillations, but phase-plane analysis shows that it cannot exhibit stable oscillations. Stable oscillations appear to be a consequence of a finite delay between receiving the differentiation signal and exiting the cell-cycle. Further experiments may help to validate this prediction. Decaying oscillations in *in vitro* experiments were observed by Hesse *et al*. (1995). However, in those experiments culture medium was replaced every 24 h. Therefore, they cannot be compared to our results without further analysis.

In this paper we have only shown results from a single model. This is not to imply that this is the only model we have studied. There are many models that could be used to fit the data. Many of these can be ruled out on biological grounds and many of these will not give good fits. Hence, the process of finding a good model is based partly on trial and error. For example, the previous models by Turner *et al*. (1995) and Tyler *et al*. (2001) assumed that the transition from slender to stumpy cells is proportional to the instantaneous slender cell concentration. This is a perfectly valid assumption and one we tested against our data. We found that this type of model gave poor fits to our data: the maximum-likelihoods were at least 8 lower than our best-fit model. Thus we tried the more realistic behaviour of a delay between commitment to differentiate and morphological change, which resulted in good fits to the data. Our assumption can be relaxed even further to allow for a distribution of transition times. This, however, becomes much harder and slower to solve numerically.

The model can be tested in various ways. Data with more frequent observations would be an especially good test. Longer time series are also a possibility, but these run the risk of being affected by later stage processes in the host not taken into account in the model. Our parameter estimates can also be tested. However, these will vary not only between different strains and host species but also between replicates of the same experiment. Some of our predictions may also be used as tests for the model. For example, observing multiple cell divisions between commitment and cell-cycle arrest would be a strong validation of the model.

Although in the experiments, slender, intermediate and stumpy cells were counted, we choose to include intermediate cells in the stumpy class. Even for experienced observers, distinguishing the intermediate morphology is very difficult. In this analysis, intermediate cells were classed as stumpy because only a small percentage are dividing, as opposed to the 30% or so of slender cells that are dividing at any given time. However, we have built another model that incorporates intermediate cells as a separate class. We assumed that cells remain in this class for a fixed period (necessitating an additional parameter). We got very good fits to the data *if* we hypothesized two slender cell subpopulations. One subpopulation that differentiated at the background rate into intermediate cells but never fully developed into stumpy cells. And another subpopulation that differentiated at the SIF-concentration dependent rate and developed into stumpy cells. Although an intriguing finding, there is no evidence for 2 subpopulations of cells. Thus we conclude that this result is an artifact of including an intermediate class into the model.

Evidence for the existence of SIF is strong (Vassella *et al*. 1997), and thus we have assumed it to be the differentiation signal. However, our model does not verify its existence. This is because if we model the differentiation signal as a response to slender cell density instead of SIF concentration we get just as good fits to the data as our best-fit model. This is because of two factors. Firstly, because we assume that SIF production is proportional to slender cell concentration, and secondly, because SIF degradation is fast. These two factors mean that the strength of the differentiation signal is similar whether or not the signal is SIF derived or slender cell derived.

Changes in single parameter values can cause multiple, large and counter-intuitive changes in the dynamical behaviour. This is not unexpected for a non-linear model. Of particular interest is the observation that we can simulate behaviours of other model systems with simple parameter value changes. For example, increasing SIF production or sensitivity to SIF reduces peak parasitaemia to levels seen in cattle. And increasing slender to stumpy transition period causes behaviour similar to monomorphic trypanosome strains. However, just because our model can replicate these behaviours does not imply that the model is correct for these cases. More work is needed to explore these predictions.

In conclusion, we have developed a simple model that is consistent with the available experimental data. It predicts many facets of slender to stumpy transformation that are testable. Having found such a model and determined some key parameter values, the model can be used to predict the behaviour of other systems (e.g. cattle and monomorphic strains). Moreover, the model can be extended in the future to include the interaction of the host immune system with the trypanosomes, and to model *in vitro* experiments, for example, those done by Hesse *et al*. (1995).

Nick Savill is supported by SHEFC and The Wellcome Trust. The authors thank John Sechelski for his technical assistance, and Gavin Gibson for advice on the statistical analysis.

REFERENCES

BARRY, J. D. & TURNER, C. M. R. (1991). The dynamics of antigenic variation and growth of African Trypanosomes. *Parasitology Today* **7**, 207–211.

BLACK, S. J., HEWETT, R. S. & SENDASHONGA, C. N. (1982). *Trypanosoma brucei* variable surface antigen is released by degenerating parasites but not by actively dividing parasites. *Parasite Immunology* **4**, 233–244.

BLACK, S. J., SENDASHONGA, C. N., O'BRIAN, C., BOROWY, N. K., NAESSENS, M., WEBSTER, P. & MURRAY, M. (1985). Regulation of parasitemia in mice infected with *Trypanosoma brucei*. *Current Topics in Microbiology and Immunology* **117**, 93–118.

HESSE, F., SELZER, P. M., MÜHLSTÄDT, K. & DUSZENKO, M. (1995). A novel cultivation technique for long-term maintenance of bloodstream form trypanosomes *in vitro*. *Molecular and Biochemical Parasitology* **70**, 157–166.

PRESS, W. H., TEUKOLSKY, S. A., VETTERLING, W. T. & FLANNERY, B. P. (1992). *Numerical Recipies in C : The Art of Scientific Computing*, 2nd Edn. Cambridge University Press, Cambridge.

REUNER, B., VASSELLA, E., YUTZY, B. & BOSHART, M. (1997). Cell density triggers to stumpy differentiation of *Trypanosoma brucei* bloodstream forms in culture. *Molecular and Biochemical Parasitology* **90**, 269–280.

SEED, J. R. & BLACK, S. J. (1997). A proposed density-dependent model of long slender to short stumpy transformation in the African trypanosomes. *Journal of Parasitology* **83**, 656–662.

SEED, J. R. & BLACK, S. J. (1999). A revised arithmetic model of long slender to short stumpy transformation in the African trypanosomes. *Journal of Parasitology* **85**, 850–854.

SEED, J. R. & SECHELSKI, J. B. (1988). Growth of pleomorphic *Trypanosoma brucei rhodesiense* in irradiated inbred mice. *Journal of Parasitology* **74**, 781–789.

SEED, J. R. & SECHELSKI, J. B. (1989*a*). African trypanosomes: Inheritance of factors in resistance to the African trypanosomes. *Experimental Parasitology* **69**, 1–8.

SEED, J. R. & SECHELSKI, J. B. (1989*b*). Mechanism of long slender (LS) to short stumpy (SS) transformation in the African trypanosomes. *Journal of Protozoology* **36**, 572–577.

SQUIRES, G. L. (1985). *Practical Physics*, 3rd Edn. Cambridge University Press, Cambridge.

TURNER, C. M. R., ASLAM, N. & DYE, C. (1995). Replication, differentiation, growth and the virulence of *Trypanosoma brucei* infections. *Parasitology* **111**, 289–300.

TYLER, K. M., HIGGS, P. G., MATTHEWS, K. R. & GULL, K. (2001). Limitation of *Trypanosoma brucei* parasitemia results from density-dependent parasite differentiation and parasite killing by the host immune response. *Proceedings of the Royal Society London, B* **268**, 2235–2243.

TYLER, K. M., MATTHEWS, K. R. & GULL, K. (1997). The bloodstream differentiation-division of *Trypanosoma brucei* studied using mitochondrial markers. *Proceedings of the Royal Society London, B* **264**, 1481–1490.

VASSELLA, E., REUNER, B., YUTZY, B. & BOSHART, M. (1997). Differentiation of African trypanosomes is controlled by a density sensing mechanism which signals cell cycle arrest via the cAMP pathway. *Journal of Cell Science* **110**, 2661–2671.

APPENDIX A: DERIVATION OF EQUATIONS 1, 3, 4 AND 5

The likelihood of the total cell concentration being C_t given that n_t cells were counted in a volume of blood v is proportional to the probability of counting n_t cells in a volume of blood v containing a concentration of cells C_t. This probability has a Poisson distribution, therefore

$$L(C_t|n_t, v) \propto \frac{(vC_t)^{n_t}}{n_t! e^{vC_t}}. \tag{19}$$

The most likely value of C_t, denoted $\hat{C}_t$, is the value of C_t that maximizes this equation. So, taking the derivative of equation 19, setting it to zero and solving for C_t we get

$$\hat{C}_t = \frac{n_t v}{v}. \tag{20}$$

Rearranging gives equation 1. The standard error in C_t, that is, the uncertainty around the expected value $\hat{C}_t$ is given by

$$e_{C_t} = -\left(\frac{d^2 L}{dC_t^2}\bigg|_{\hat{C}_t}\right)^{-1}. \tag{21}$$

By solving the second derivative of equation 19 and substituting in equation 20 we arrive at equation 3.

The expected values of the frequencies of the two cell-types, $\hat{f}_l$ and $\hat{f}_s$, are found in a similar fashion, but now with the likelihood proportional to equation 16. To find $\hat{f}_l$ and e_{f_l}, $f_s = 1 - f_l$ and $n_s = m_t - n_l$ have to be substituted into equation 16. Similarly for $\hat{f}_s$ and e_{f_s}.

The expected slender cell concentration, $\hat{C}_l$, is given by $\hat{C}_l = \hat{f}_l \hat{C}_t$. The standard error in $\hat{C}_l$ is given by the relation (Squires, 1985)

$$\left(\frac{e_{C_l}}{\hat{C}_l}\right)^2 = \left(\frac{e_{f_l}}{\hat{f}_l}\right)^2 + \left(\frac{e_{C_t}}{\hat{C}_t}\right)^2. \tag{22}$$

This is easily rearranged to give equation 5.

APPENDIX B: NORMALIZING SIF PRODUCTION

We assume that SIF is produced at a rate a_{11} and degraded at a rate a_5, that is

$$\frac{df}{dt} = a_{11}(l_1 + l_2) - a_5 f. \tag{23}$$

SIF induces L_1 differentiation

$$\frac{dl}{dt} = a_1 l_1 - (a_1 + a_3 f) l_1. \tag{24}$$

Because we do not have any indication of SIF production rate or its quantitative effect on L_1 cells, we

https://doi.org/10.1017/S0031182003004256 Published online by Cambridge University Press

can manipulate the equations to subsume parameter a_{11} into parameter a_3 like so.

Let SIF concentration be composed of a constant factor F that represents the units of measurement, and a number f^* that has no dimensions i.e., $f = Ff^*$. Substituting this into equations 23 and 24, setting $F = a_{11}$ and dividing through gives

$$\frac{df^*}{dt} = l_1 + l_2 - a_5 f^*, \tag{25}$$

$$\frac{dl_1}{dt} = a_1 l_1 - (a_2 + a_3 a_{11} f^*) l_1. \tag{26}$$

For convenience we write f^* as f, and subsume parameter a_{11} into a_3 with no loss of generality.

APPENDIX C: CONVERTING THE DDE MODEL TO AN ODE MODEL

To convert the DDE model into an ODE model we have to have two classes of cells: an uncommitted class and a differentiated class. The former is just L_1 of the DDE model. The latter includes classes L_2, L_3 and S. The concentration of the uncommitted class is just l_1, as before. However, cells in the differentiated class must be traced back from when they differentiated. We do this by defining the variable $y(\tau, t)\delta\tau$, which is the concentration of differentiated cells, at time t, that differentiated between the times τ and $\tau + \delta\tau$. Depending on how long in the past a cell differentiated it will be a dividing slender cell, an arrested slender cell, or a stumpy cell. The model is defined as follows

$$\frac{d}{dt} l_1(t) = a_1 l_1(t) - w(t) l_1(t), \tag{27}$$

$$\frac{d}{dt} y(\tau, t)\delta\tau = \begin{cases} w(t) l_1(t) & \text{if } \tau = t, \\ a_1 y(\tau, t)\delta\tau & \text{if } t - t_2 \leqslant \tau < t, \\ 0 & \text{if } t - t_2 - t_3 \leqslant \tau < t - t_2, \\ -a_4 y(\tau, t)\delta\tau & \text{if } a_0 \leqslant \tau < t - t_2 - t_3, \end{cases} \tag{28}$$

$$\frac{d}{dt} f(t) = l_1(t) + l_2(t) - a_5 f(t), \tag{29}$$

where

$$w(t) = a_2 + a_3 f(t), \tag{30}$$

$$l_1(t) = \int_{t-t_2}^{t} y(\tau, t)\, d\tau, \tag{31}$$

$$l_2(t) = \int_{t-t_2-t_3}^{t-t_2} y(\tau, t)\, d\tau, \tag{32}$$

$$s(t) = \int_{a_0}^{t-t_2-t_3} y(\tau, t)\, d\tau. \tag{33}$$

https://doi.org/10.1017/S0031182003004256 Published online by Cambridge University Press