

## 1 Sensitivity of hybrid model to length of time step

2 To ensure the Strang splitting error is sufficiently small, we ran the model for multiple  
3 sets of parameters (representing “growth”, “equilibrium” and “death”) and multiple time  
4 step lengths (from 10 days to one day, and also 0.5, 0.25, and 0.1 days). In each  
5 instance, we recorded a number of observables, and chose  $h$  such that decreasing  
6 the time step further did not substantially affect these observables. We chose the  
7 observables to be the number of clones at a given duration of the hybrid (500 days);  
8 the number of infected cells at this duration; and the minimum number of clones over  
9 the model run. S4 Fig plots each observable as a function of the inverse of  $h$ . For all  
10 parameter sets, each observable is flat when plotted against time step. We conclude  
11 that a time step of  $h = 1$  day results in a sufficiently small splitting error.

12

## 13 Comparison of hybrid and purely stochastic model

14 It is not possible to model the entire system HTLV-1 T cell clones stochastically, due  
15 to the prohibitively high frequencies of large clones. However, we consider a greatly  
16 reduced system for which a purely stochastic model and a hybrid stochastic and  
17 deterministic model can be compared.

18

19 The reduced system omits all clones of frequency greater than 460 cells, which  
20 substantially reduces the number of infected cells, but reduces the number of clones  
21 less drastically. When considered as a hybrid, clones above  $F = 100$  are considered  
22 deterministically. We fit the purely stochastic simulation and the hybrid as before for  
23 example patients for  $t_{Dur} = 100$  days.

24

25 We find that that the purely stochastic and the hybrid models give very similar values  
26 of both the values of  $r_l$  and expected number of clones over time (S5 Fig), for each  
27 example patient. It must be stressed that in these simulations, the fitted value of  $r_l$  is  
28 much higher than for the actual system of HTLV-1 clones, because the number of  
29 infected cells in the reduced system is approximately 5 orders of magnitude lower, and  
30 new clones are modelled as a mass-action process.

31

32 We would expect the errors associated with using a hybrid instead of a purely  
33 stochastic model to be lower for the actual system of HTLV-1 clones because there  
34 clones and cells are much larger and less susceptible by random variation. We  
35 conclude that the error associated with using a hybrid model is acceptable.