

## Activation-Induced B Cell Fates Are Selected by Intracellular Stochastic Competition

Ken R. Duffy,<sup>1</sup> Cameron J. Wellard,<sup>2,3</sup> John F. Markham,<sup>4</sup> Jie H. S. Zhou,<sup>2,3</sup> Ross Holmberg,<sup>2</sup> Edwin D. Hawkins,<sup>5</sup> Jhagvaral Hasbold,<sup>2,3</sup> Mark R. Dowling,<sup>2,3\*</sup> Philip D. Hodgkin<sup>2,3\*†</sup>

<sup>1</sup>Hamilton Institute, National University of Ireland, Maynooth, Ireland. <sup>2</sup>Walter and Eliza Hall Institute of Medical Research, Parkville, Melbourne, Victoria 3052, Australia. <sup>3</sup>Department of Medical Biology, University of Melbourne, Melbourne, Victoria 3052, Australia. <sup>4</sup>Victoria Research Laboratory, National Information and Communications Technology (ICT) Australia, University of Melbourne, Victoria 3010, Australia. <sup>5</sup>Immune Signalling Laboratory, Peter MacCallum Cancer Centre, Melbourne, Victoria 3002, Australia.

\*These authors contributed equally to this work.

†To whom correspondence should be addressed. E-mail: hodgkin@wehi.edu.au

**In response to stimulation, B lymphocytes pursue a large number of distinct fates important for immune regulation. Whether each cell's fate is determined by external direction, internal stochastic processes, or directed asymmetric division is unknown. Measurement of times to isotype switch, to develop into a plasmablast, and to divide or to die for thousands of cells indicated that each fate is pursued autonomously and stochastically. As a consequence of competition between these processes, censorship of alternative outcomes predicts intricate correlations that are observed in the data. Stochastic competition can explain how the allocation of a proportion of B cells to each cell fate is achieved. The B cell may exemplify how other complex cell differentiation systems are controlled.**

Production of antibody by B cell-derived plasma cells is critical for an effective immune response (1), but B cell activation mechanisms leading to the formation of plasma cells are numerous and poorly understood. One known mechanism is for the B cell to use its surface receptor to capture and internalize antigen, which leads to presentation of T cell epitopes on its cell surface (2). Upon detection, a T cell delivers cell contact- and cytokine-mediated signals (3) that lead to B cell proliferation or changes in antibody type (isotype switching) (4), as well as differentiation into dividing plasmablasts (PBs) and sessile, long-lived plasma cells, both of which secrete antibody (5, 6). The heterogeneous B cell fates resulting from isotype switching and development into PBs can be replicated in vitro by stimulating naïve B cells through CD40 in addition to the cytokines interleukin 4 (IL-4) and IL-5 to simulate T cell interaction (7, 8), a method we used. As Blimp1 is a transcription factor that is selectively required for differentiation to PB, we used a Blimp1-GFP reporter mouse (9) to identify PBs by green fluorescent protein (GFP) expression and fluorescently labeled antibody

against immunoglobulin G1 (IgG1) to identify cells switched to IgG1. After an initial 3-day culture, single-cell video microscopy was used to observe cells sorted from generations 1, 3, 5, and 7 that do not express Blimp1 or IgG1 and to optically track their times to isotype switch to IgG1, to differentiate to PBs, and ultimately to division or death (Fig. 1A and fig. S1) (10). Only one division round was followed because of the strong homotypic adhesion of B lymphoblasts, which leads to a loss of identity.

As found previously (11–15), division and death times were highly variable, and this was also true for times to isotype switching and commitment to become a PB (Fig. 1, B and C). Consistent with earlier population studies (11, 16), the dependence of frequency of isotype switching, as well as other parameters, on generation is apparent (Fig. 1, D and E). Despite the diverse range of experience of individual cells, the population-level response is insensitive to this variability (12, 13, 17–20), which presents the conundrum of how to reconcile the single-cell and population-level responses. Furthermore, these data have complex correlation structures, both within single cells (intracellular) and between siblings (intercellular), which pose an additional challenge to any paradigm of understanding.

As a representative of intracellular correlation, for all cells that differentiate and go on to divide, Fig. 2A presents a scatter plot of times to these events, as well as estimates of Pearson's correlation coefficient ( $\rho$ ). The latter reveals positive correlation coefficients for cells of each generation ( $\rho$ : 0.54, 0.59, 0.56, and 0.80). Analysis of other combinations shows that time to differentiate to PBs is positively correlated with time to death (fig. S2A) ( $\rho$ : 0.70, 0.54, 0.81, and 0.54), and with time to isotype switch (Fig. 2C) ( $\rho$ : –, 0.34, 0.29, and 0.51), where a dash indicates an insufficient number of observations to form an estimate. There is little evidence for correlation in time to isotype

switch and time to division (Fig. 2B) ( $\rho$ :  $-$ , 0.17,  $-$ 0.05, and  $-$ 0.08) or death (fig. S2B) ( $\rho$ :  $-$ , 0.15, 0.08, and 0.11).

For intercellular dependencies, we can investigate the existence of concordance in sibling fates. Visual inspection of Fig. 1B suggests strong positive relatedness between siblings. To quantify the strength of concordance in opposing fates of siblings, we use Yule's  $Q$  (21). It takes a value in  $[-1, 1]$ , with 1 corresponding to perfect positive correlation, and 0 corresponding to no correlation in sibling fate (10). For our system, the opposing fates of siblings are death versus division, differentiation to PBs versus no differentiation to PBs, and isotype switching versus no isotype switching. Figure 2D plots Yule's  $Q$  for the division versus death outcome of siblings. It is high for all generations ( $Q$ : 0.97, 0.93, 0.90, and 0.96), which confirms strong sibling concordance in division or death fates: If a cell dies or divides, the likelihood that its sibling experiences the same fate is substantially higher than the likelihood of a cell chosen uniformly at random from the population at large has the same fate. All other fates display similar evidence of strong concordance: differentiate to PB versus not (Fig. 2E) ( $Q$ : 0.98, 0.98, 0.92, and 0.93) and isotype switch versus not (Fig. 2E) ( $Q$ :  $-$ , 0.98, 0.99, and 0.94).

This strong concordance justifies investigating correlations within the times to fates of siblings. We found strong correlation in time to fate between siblings (Fig. 2, F and G):  $\rho$  of 0.90, 0.93, 0.84, and 0.82 for division and  $\rho$  of 0.85, 0.79, 0.72, and 0.77 for death. For differentiation to PB and isotype switching, the former exhibits strong correlation (fig. S2C) ( $\rho$ : 0.75, 0.88, 0.92, and 0.94), whereas the latter is the exception and does not (fig. S2D) ( $\rho$ :  $-$ , 0.18, 0.29, and 0.28). Unexpectedly, correlation structure is also found in the nonconcordant fates of siblings (Fig. 2, H to J, and fig. S2, E to G).

These numerous sibling correlation structures would appear challenging for any simple mechanistic theory to explain them, as they suggest involved sharing of molecular machinery regulating features of isotype switching, development into a PB, division, and death. We challenged this deduction by questioning whether all of these features could be explained by a simple hypothesis of autonomous functional units in cells that are subject to competition. Such a hypothesis was proposed for division and death of lymphocytes in the cyton model (11, 13, 22–25), whose fundamental tenet is that each cell's death or division fate is a consequence of competition between independent times for division and death. It posits the existence of two independent random variables in each cell determining times from division to fates: a time to division,  $T_{\text{div}}$ , and a time to death,  $T_{\text{death}}$ . We postulate the existence of two additional independent random variables per cell. One,  $T_{\text{diff}}$ , determines a time to differentiate to PB, and the other,  $T_{\text{switch}}$ , determines time to isotype

switching. Consistent with observations of  $T_{\text{div}}$  (13, 14, 20), with the exception of death, the corresponding process may not be in operation in every cell. This is encapsulated mathematically by assuming that the associated random variables can have a positive probability of being infinite.

These times are subject to competition and censorship. If  $T_{\text{death}} < T_{\text{div}}$ , then the cell's fate is death at  $T_{\text{death}}$ , and the value  $T_{\text{div}}$  is not observed and vice versa. Thus, the larger value is censored. If  $T_{\text{diff}} < \min(T_{\text{div}}, T_{\text{death}})$ , then the cell is observed to differentiate to PB at time  $T_{\text{diff}}$  from birth, and otherwise the time is censored. If  $T_{\text{switch}} < \min(T_{\text{diff}}, T_{\text{div}}, T_{\text{death}})$ , then isotype switching is observed at  $T_{\text{switch}}$ , and otherwise its value is censored. Even though the random variables  $\{T_{\text{switch}}, T_{\text{diff}}, T_{\text{div}}, T_{\text{death}}\}$  are independent, it is of fundamental importance that the distributions of the observed variables  $\{T_{\text{switch}}^{\text{obs}}, T_{\text{diff}}^{\text{obs}}, T_{\text{div}}^{\text{obs}}, T_{\text{death}}^{\text{obs}}\}$  are distinct and correlated, as they have been altered by competition and censorship (10).

Because the times to concordant fate of siblings are related (Fig. 2, F and G, and fig. S2, C and D), although the model assumes that uncensored times within a cell are independent, we must allow for correlations between times to concordant fate in siblings. With a superscript 1 indicating times for one sibling and 2 for the other, we posit the existence of pairs  $(T_{\text{switch}}^1, T_{\text{switch}}^2)$ ,  $(T_{\text{diff}}^1, T_{\text{diff}}^2)$ ,  $(T_{\text{div}}^1, T_{\text{div}}^2)$  and  $(T_{\text{death}}^1, T_{\text{death}}^2)$ , where the two random variables within each pair may be correlated, but they are independent across pairs. For each pair, we adopt the class of bivariate log-normal distributions with symmetric marginal distributions whose correlation structure is defined through a correlated Gaussian exponent (10). The model has 15 parameters, four means, four variances, four cross-correlations, and three probabilities of finiteness:

$$\theta = \left( \begin{array}{c} \mu_{\text{switch}}^1, \mu_{\text{diff}}^1, \mu_{\text{div}}^1, \mu_{\text{death}}^1, \sigma_{\text{switch}}^2, \sigma_{\text{diff}}^2, \sigma_{\text{div}}^2, \sigma_{\text{death}}^2, \rho_{\text{switch}}^1, \rho_{\text{diff}}^1 \\ \rho_{\text{div}}^1, \rho_{\text{death}}^1, \rho_{\text{switch}}^1, \rho_{\text{diff}}^1, \rho_{\text{div}}^1 \end{array} \right)$$

Consider a parameterization,  $\theta$ , defining the model. For a set of data,  $D$ , consisting of the times of fates of sibling cells, we calculate the likelihood that the model would generate the data  $L(D|\theta)$  (10) and then determine  $\theta_{\text{MAP}} = \arg \sup L(D|\theta)$ , where "arg sup" reads "arguments of the supremum," the maximum a posteriori parameterization defining our best-fit model. For each generation, we fit to a data set  $D$ , consisting of all siblings that share the same fate. This gives us sufficient data (382, 458, 330, and 254 observations in generations 1, 3, 5, and 7, respectively) to estimate all components of  $\theta_{\text{MAP}}$ , but it means that the model's cross-fate predictions of the remaining data (145, 193, 234, and 131 cells, respectively) are extrapolations.

For generation 5, we plotted the uncensored marginal distributions for the independent random variables within a single sibling, as well as its censored match to observed data (Fig. 3A). The impact of competition is apparent in the

substantially reduced mean and variance in the latter graphs and also in the reshaping of the distributions (see other generations in fig. S3). One indicator of whether the model's structure can capture key features of the data is comparison of average time to fate with the best-fit model. This is shown in Fig. 3, B to E, along with the measured and modeled proportion of cells that undergo division and isotype switching (death and differentiation to PB) (fig. S4). A second indicator is its ability to capture the correlation observed in the times to concordant sibling fate (Fig. 3, H and I, and fig. S5). These good fits suggest that this simple model with four pairs of independent random variables has sufficient flexibility to mimic the data.

Although intercellular concordant fate correlations are to be expected because of the cross-correlation parameters in our model, perhaps counterintuitively, it also predicts correlations within cells and between siblings. The uncensored variables are uncorrelated, so these correlations arise in the model exclusively, as a consequence of its postulated competition and censorship (Fig. 3, F and G). With the exception of generation 7 (Fig. 3F), the predictions all lie within the 95% confidence intervals (CIs) of the measured values. Other fate combinations are shown in fig. S5, all of which also fall within 95% CIs of the measured values, with the exception of generation 5 death and differentiation to PB (10).

For asymmetric intercellular correlations (where one sibling divides and the other dies), division times across siblings are positively correlated, as are death times. Thus, the model predicts that one only observes asymmetric fates when time to division and time to death are close to each other, and hence, these are predicted to be positively correlated when observed (Fig. 3J). The model predicts that the correlations between time to division of one sibling and to isotype switch of the other are less correlated, as these two processes are typically separated in time and so do not strongly influence each other (Fig. 3K and fig. S5). Remarkably, with the exception of generation 5 death versus differentiation to PB, the model's predicted correlations lie within the 95% CIs for Pearson's  $\rho$  as measured in the data. Thus, not only is this simple mechanistic model of competing independent random variables capable of fitting the heterogeneous times, its predicted nontrivial cross-correlation structures that arise as a consequence of competition and censorship quantitatively match observed phenomena.

This analysis offers a mechanism to explain how lymphocytes reliably allocate a consistent proportion of cells to various populations even in the absence of specific signals, such as antigen, which may induce asymmetric cell division (26). This inherent randomness provides robustness to loss of cells of a given class, which seems desirable in a system experiencing decentralized control. High-level order then

appears not by explicit individual instructions to each cell, but by chemical signaling that shapes probability distributions and influences resultant laws of large numbers. As a consequence, this study draws attention to the stochastic epigenetic processes that influence the expression levels of the many molecular regulators of division, death, and differentiation within each cell (27–29), as it is likely that these quantitative differences will ultimately explain this shaping of distributions. Competing intracellular processes may also underlie the robust allocation of multiple fates during proliferation and differentiation in other cell systems from the embryonic to hematopoietic.

## References and Notes

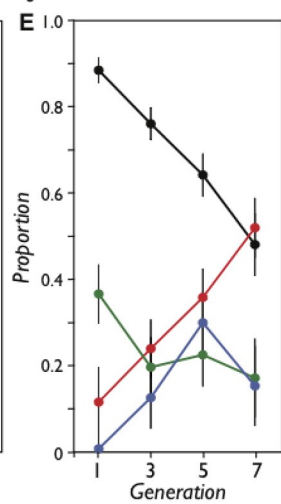
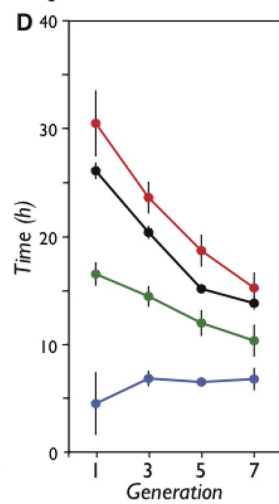
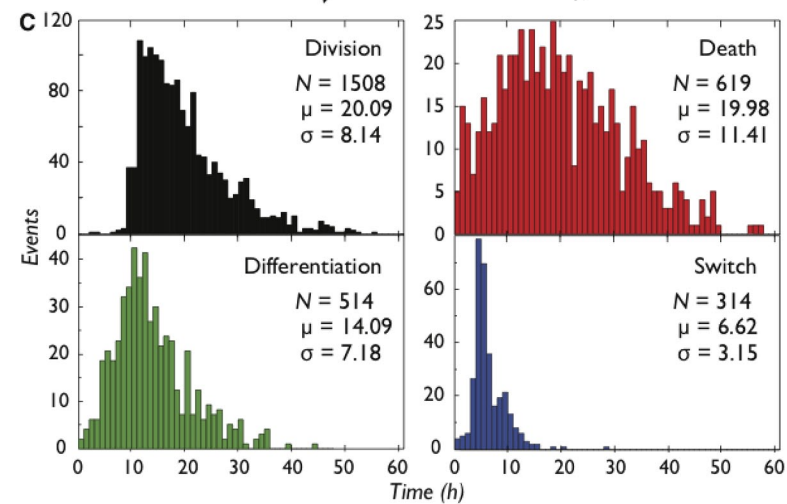
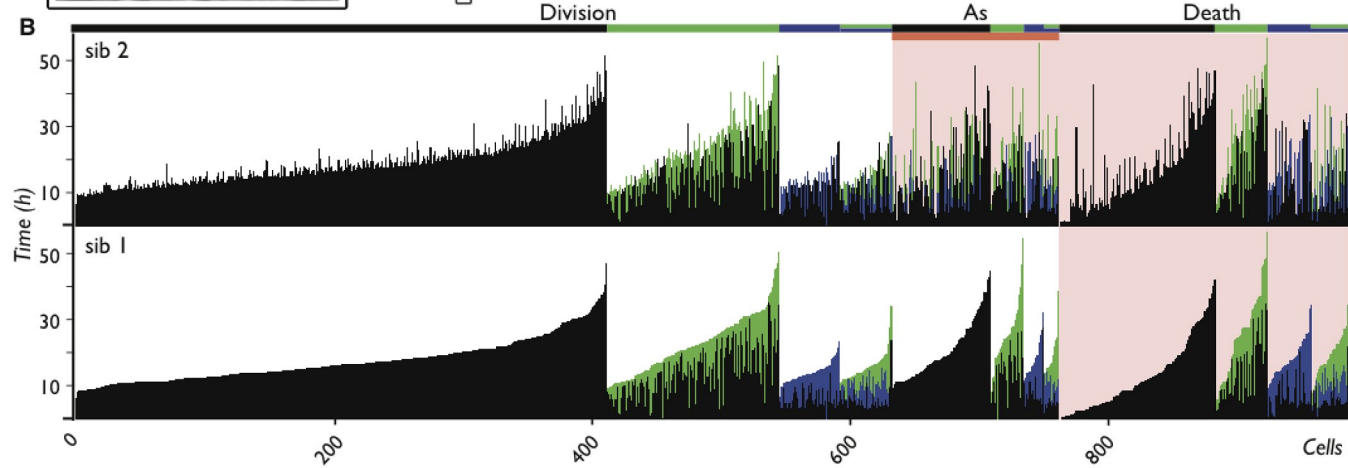
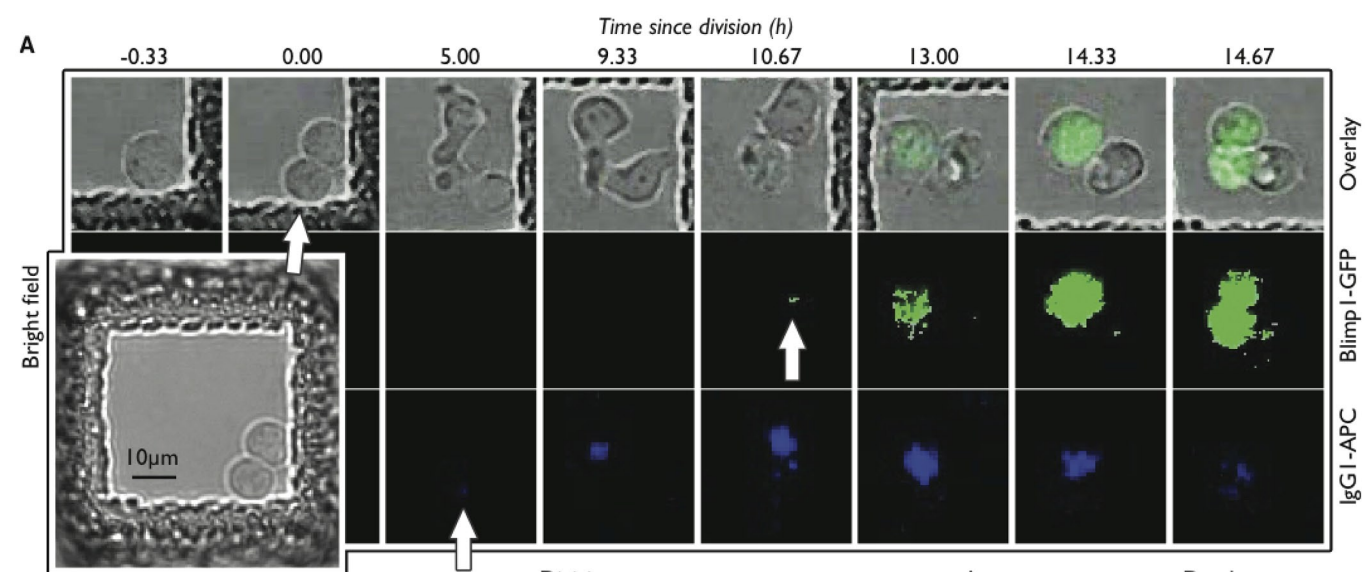
1. M. K. Slifka, R. Ahmed, Long-lived plasma cells: A mechanism for maintaining persistent antibody production. *Curr. Opin. Immunol.* **10**, 252 (1998).
2. A. Lanzavecchia, Antigen-specific interaction between T and B cells. *Nature* **314**, 537 (1985).
3. D. C. Parker, T cell-dependent B cell activation. *Annu. Rev. Immunol.* **11**, 331 (1993).
4. R. L. Coffman, D. A. Leberman, P. Rothman, Mechanism and regulation of immunoglobulin isotype switching. *Adv. Immunol.* **54**, 229 (1993).
5. I. C. MacLennan *et al.*, Extrafollicular antibody responses. *Immunol. Rev.* **194**, 8 (2003).
6. D. Paus *et al.*, Antigen recognition strength regulates the choice between extrafollicular plasma cell and germinal center B cell differentiation. *J. Exp. Med.* **203**, 1081 (2006).
7. S. L. Nutt, N. Taubenheim, J. Hasbold, L. M. Corcoran, P. D. Hodgkin, The genetic network controlling plasma cell differentiation. *Semin. Immunol.* **23**, 341 (2011).
8. J. Hasbold, L. M. Corcoran, D. M. Tarlinton, S. G. Tangye, P. D. Hodgkin, Evidence from the generation of immunoglobulin G-secreting cells that stochastic mechanisms regulate lymphocyte differentiation. *Nat. Immunol.* **5**, 55 (2004).
9. A. Kallies *et al.*, Plasma cell ontogeny defined by quantitative changes in Blimp-1 expression. *J. Exp. Med.* **200**, 967 (2004).
10. Materials and methods are available as supporting material on *Science Online*.
11. A. V. Gett, P. D. Hodgkin, A cellular calculus for signal integration by T cells. *Nat. Immunol.* **1**, 239 (2000).
12. E. K. Deenick, A. V. Gett, P. D. Hodgkin, Stochastic model of T cell proliferation: A calculus revealing IL-2 regulation of precursor frequencies, cell cycle time, and survival. *J. Immunol.* **170**, 4963 (2003).
13. E. D. Hawkins, M. L. Turner, M. R. Dowling, C. van Gend, P. D. Hodgkin, A model of immune regulation as a consequence of randomized lymphocyte division and

- death times. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 5032 (2007).
14. E. D. Hawkins, J. F. Markham, L. P. McGuinness, P. D. Hodgkin, A single-cell pedigree analysis of alternative stochastic lymphocyte fates. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 13457 (2009).
  15. S. L. Spencer, S. Gaudet, J. G. Albeck, J. M. Burke, P. K. Sorger, Non-genetic origins of cell-to-cell variability in TRAIL-induced apoptosis. *Nature* **459**, 428 (2009).
  16. P. D. Hodgkin, J. H. Lee, A. B. Lyons, B cell differentiation and isotype switching is related to division cycle number. *J. Exp. Med.* **184**, 277 (1996).
  17. V. P. Badovinac, B. B. Porter, J. T. Harty, Programmed contraction of CD8(+) T cells after infection. *Nat. Immunol.* **3**, 619 (2002).
  18. S. G. Tangye, D. T. Avery, E. K. Deenick, P. D. Hodgkin, Intrinsic differences in the proliferation of naive and memory human B cells as a mechanism for enhanced secondary immune responses. *J. Immunol.* **170**, 686 (2003).
  19. R. J. De Boer, D. Homann, A. S. Perelson, Different dynamics of CD4+ and CD8+ T cell responses during and after acute lymphocytic choriomeningitis virus infection. *J. Immunol.* **171**, 3928 (2003).
  20. M. L. Turner, E. D. Hawkins, P. D. Hodgkin, Quantitative regulation of B cell division destiny by signal strength. *J. Immunol.* **181**, 374 (2008).
  21. G. U. Yule, On the methods of measuring association between two attributes. *J. R. Stat. Soc.* **75**, 579 (1912).
  22. V. G. Subramanian, K. R. Duffy, M. L. Turner, P. D. Hodgkin, Determining the expected variability of immune responses using the cyton model. *J. Math. Biol.* **56**, 861 (2008).
  23. K. R. Duffy, V. G. Subramanian, On the impact of correlation between collaterally consanguineous cells on lymphocyte population dynamics. *J. Math. Biol.* **59**, 255 (2009).
  24. C. Wellard, J. Markham, E. D. Hawkins, P. D. Hodgkin, The effect of correlations on the population dynamics of lymphocytes. *J. Theor. Biol.* **264**, 443 (2010).
  25. J. F. Markham, C. J. Wellard, E. D. Hawkins, K. R. Duffy, P. D. Hodgkin, A minimum of two distinct heritable factors are required to explain correlation structures in proliferating lymphocytes. *J. R. Soc. Interface* **7**, 1049 (2010).
  26. J. T. Chang *et al.*, Asymmetric T lymphocyte division in the initiation of adaptive immune responses. *Science* **315**, 1687 (2007).
  27. R. N. Germain, The art of the probable: System control in the adaptive immune system. *Science* **293**, 240 (2001).
  28. H. Cedar, Y. Bergman, Epigenetics of haematopoietic cell development. *Nat. Rev. Immunol.* **11**, 478 (2011).
  29. D. A. Hume, Probability in transcriptional regulation and its implications for leukocyte differentiation and inducible gene expression. *Blood* **96**, 2323 (2000).
  30. D. Day *et al.*, A method for prolonged imaging of motile lymphocytes. *Immunol. Cell Biol.* **87**, 154 (2009).
- Acknowledgments:** The data presented in this paper are tabulated in the main paper and in the Supporting Online Material. This work was supported by the Australian National Health and Medical Research Council (NHMRC) (program grant 575500; fellowships, E.D.H., M.R.D. and P.D.H.), Victorian State Government Operational Infrastructure support, Australian Government NHMRC Independent Research Institute Infrastructure Support Scheme (IRIS), Science Foundation Ireland (grant RFP-ENEF530, K.R.D.). J.F.M. was supported by National ICT Australia (NICTA), which is funded by the Australian Government's Backing Australia's Ability initiative, in part through the Australian Research Council. The authors thank D. Day (Swinburne University of Technology) and Microsurfaces Pty. Ltd. for providing the microgrids (30), S. Nutt and A. Kallies for Blimp1 reporter mice, and R. DeBoer and D. Tarlinton for helpful feedback on the manuscript. "A cell population system and process" (U.S. Patent Application 12/312,933) describes computational algorithms based on competing fates that can be used to calculate and predict immune outcomes. Authors who are listed inventors are P.D.H., M.R.D., C.J.W, J.F.M., and E.D.H.
- Supporting Online Material**  
[www.sciencemag.org/cgi/content/full/1213230/DC1](http://www.sciencemag.org/cgi/content/full/1213230/DC1)  
 Materials and Methods  
 Figs. S1 to S5  
 References
- 26 August 2011; accepted 29 November 2011  
 Published online 5 January 2012; 10.1126/science.1213230
- Fig. 1.** Individual B cell fate. (A and B) Cells were observed to divide, and times to subsequent events (division, death, switch to IgG1, and expression of Blimp1) were recorded for both siblings (sib) (2127 cells pooled from four cultures tracked to final death or division fate). Illustrated cell becomes IgG1+ at 5 hours, GFP+ at 10.67 hours, and divides at 14.67 hours. (B) Collection of all sibling pairs illustrating fates. Sib 1 is chosen as the first in a pair to divide or die. Sib 1's are sorted into fate categories and then rank-ordered in increasing time within their category, with each cell shown as a column. Sib 2's are not ranked but plotted in the same order as paired sibling to illustrate the close relation between fate and times to fate of siblings. Black line shows time spent unswitched or undifferentiated. Green line in the column shows time as Blimp1+ PB, blue as IgG1+ switched cells.

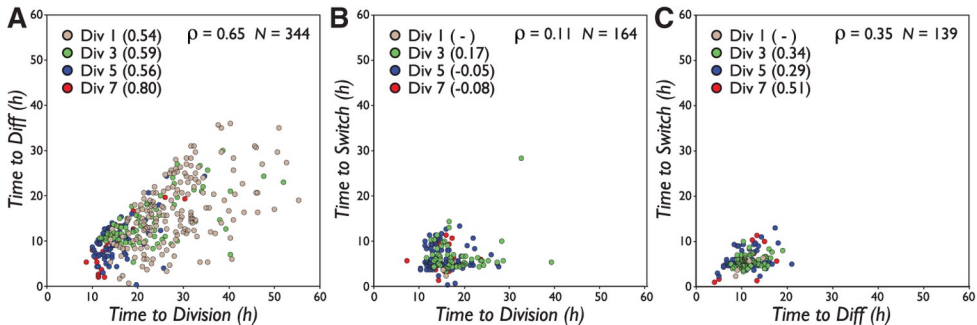
The fate categories: divide only; become PB and divide; isotype switch and divide; switch, then become PB, then divide. Sequence repeated for cells that die. 'As' shows asymmetric siblings (one died and one divided). (C) Times to fate. Number of cells ( $N$ ), mean ( $\mu$ ), and standard deviation ( $\sigma$ ) are shown. (D) Average time to divide (black), die (red), differentiate to PB (green), and switch (blue) per generation. (E) Proportion of cells undergoing each fate.

**Fig. 2.** Fate concordance and correlations. (A to C) Intracellular correlations (correlations in times to distinct fates within individuals) for each generation. Pearson's  $\rho$  is shown per generation and for cohort inclusive of all generations. (D) Measure of concordance of siblings in division versus death fate (red) shown as Yule's  $Q$  with 95% CIs. (E) Measure of concordance of siblings in isotype switch versus no isotype switch (blue) and development to PB versus no development to PB fate (green). (F and G) Times to intercellular concordant fates (siblings that share the same fate) and their correlation. (H to J) Times to intercellular nonconcordant fates and their correlation.

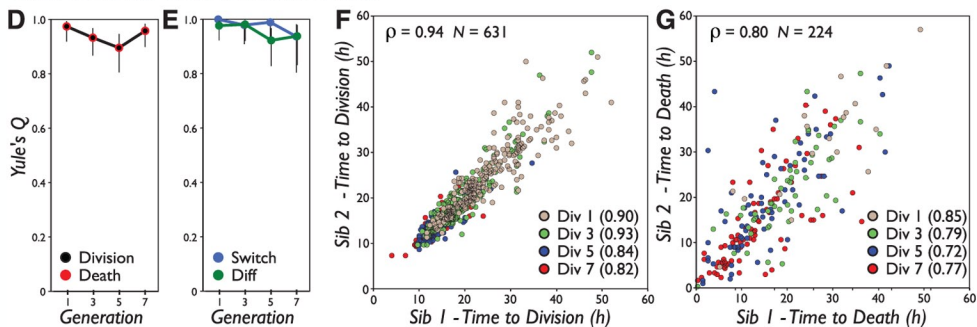
**Fig. 3.** Model of four independent fate times. (A) Fit of model to generation 5 data. (Top) Predicted uncensored distribution for each fate (lognormal mean,  $\mu$ ; variance,  $\sigma$ ; likelihood of being finite,  $p$ ). (Bottom) Predicted censored distribution overlaid on data. (B to K) Calculated model values (red) to data (blue). (B) Average time to divide. (C) Proportion of cells dividing per generation. (D) Time to switch. (E) Proportion switch per generation. (F and G) Intracellular nonconcordant fates. (H and I) Intercellular concordant correlations. (J and K) Intercellular fates.



## Intracellular fates



## Intercellular concordant fates



## Intercellular non-concordant fates

