Nongenetic origins of cell-to-cell variability in B lymphocyte proliferation

Article  (Accepted Version)


This version is available from Sussex Research Online: http://sro.sussex.ac.uk/id/eprint/84322/

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher’s version. Please see the URL above for details on accessing the published version.

Copyright and reuse:
Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

http://sro.sussex.ac.uk
Non-genetic origins of cell-to-cell variability in B-lymphocyte proliferation

Short Title: Rendering B-cell proliferation predictable

Simon Mitchell\textsuperscript{a,1}, Koushik Roy\textsuperscript{a,1}, Thomas A. Zangle\textsuperscript{b}, and Alexander Hoffmann\textsuperscript{a}\textsuperscript{*}

\textsuperscript{a} Institute for Quantitative and Computational Biosciences and Department of Microbiology, Immunology, and Molecular Genetics, University of California, Los Angeles, CA 90095

\textsuperscript{b} Department Chemical Engineering and Huntsman Cancer Institute, University of Utah, Salt Lake City, UT 84112

*Correspondence to: ahoffmann@ucla.edu

\textsuperscript{1}These authors contributed equally to this work.

Keywords: B Cells, Variability, Heterogeneity, Systems Biology, Proliferation, Lymphocytes.
Rapid antibody production in response to invading pathogens requires the dramatic expansion of pathogen-derived antigen-specific B-lymphocyte populations. Whether B-cell population dynamics are based on stochastic competition between competing cell fates, as in the development of competence by the bacterium *Bacillus subtilis*, or on deterministic cell fate decisions that execute a predictable program, as during the development of the worm *Caenorhabditis elegans*, remains unclear. Here, we developed long term live-cell microscopy of B-cell population expansion and multi-scale mechanistic computational modeling in order to characterize the role of molecular noise in determining phenotype heterogeneity. We show that the cell lineage trees underlying B-cell population dynamics are mediated by a largely predictable decision-making process where the heterogeneity of cell fates at any given timepoint largely derives from non-genetic heterogeneity in the founder cells. This means that contrary to previous models only a minority of genetically identical founder cells contribute the majority to the population response. We computationally predict and experimentally confirm non-genetic molecular determinants that are predictive of founder cells’ proliferative capacity. While founder cell heterogeneity may arise from different exposure histories, we show that it may also be due to the gradual accumulation of small amounts of intrinsic noise during the lineage differentiation process of hematopoietic stem cells to mature B cells. Our finding of the largely deterministic nature of B-lymphocyte responses may provide opportunities for diagnostic and therapeutic development.

**Significance Statement**

This study addresses why splenic B lymphocytes have differential cell fates and whether these may be predictable. Biology provides examples of both stochastic decision making and highly deterministic developmental programs. Prior studies of B-lymphocytes suggested these cells make stochastic decisions, but the key experiment to settle this question – long term tracing of individual cell lineages - has not been done. Overcoming the technical challenges, we found that B-cell fate decisions are largely non-stochastic. Using a mathematical model, we found that they are predictable, as long as the state of the molecular network in founder cells is known. That in turn allowed us to identify the molecular determinants of proliferative fate decisions, which potentially constitute novel drug targets and biomarkers for B-cell mediated diseases.
Introduction

How similar cells give rise to distinct fates is a fundamental question with different answers in different biological contexts. Cell fates are established deterministically for some cell types or organisms, such as Caenorhabditis elegans (1). In other cases, cell fates are established seemingly stochastically, such as for the development of competence by the bacterium Bacillus subtilis (2) or the generation of alternative color vision photoreceptors in Drosophila melanogaster (3), and are thereby independent of cellular history (4). Here, we examined whether B-lymphocyte proliferation decisions are result of stochastic or deterministic fate decisions, and whether molecular network determinants may be identified.

B lymphocytes are an essential component of the adaptive immune response and source of antibody producing cells. In response to invading pathogens B lymphocytes rapidly proliferate, differentiate into antibody producing cells and produce antigen-specific antibodies which are essential for an effective immune response. B-cells genetically diversify by rearranging the immunoglobulin locus to produce a diverse antibody repertoire and therefore diverse B-cell receptor (BCR)-antigen affinities, which control mitogenic signals.

While genetic heterogeneity arising from BCR diversification has the potential to be a source of heterogeneity of B-cell fate, BCR-antigen affinity is a poor predictor of B-cell proliferative expansion (5), and snapshot flow-cytometry measurements reveal a high degree of cell-to-cell generational heterogeneity even in response to BCR-independent stimuli (6). This led to the notion that B-cell fate decision-making is highly stochastic. Indeed, direct measurement of division times at single cell resolution revealed a highly variable first division (7, 8), consistent with a stochastic division-making process. Based on these observations, Hodgkin and colleagues developed a phenotypic model of lymphocyte proliferation using probability distributions of division and death times (9). The Cyton model has shown remarkable ability to fit dye dilution measurements by flow cytometry and derive corresponding cell biological parameters (such as division and death times) (9-14).

Whereas a key assumption of the Cyton model is the independent stochastic decision-making of each cell at each generation, direct observation of sibling cell behavior revealed correlations in cell fate decisions and division times (8, 10, 12, 15). This has prompted revisions of the model to consider heritability. Thus lymphocyte population dynamics models have been proposed that structure cell decisions by age (9, 16, 17) or division number (18) (or technical aspects,(19, 20)). However, the degree to which fate decisions are non-stochastic remains unclear (21). Recently
developed approaches combining multiple division-tracking dyes revealed that clonal populations were all of a similar generation at given time-points during the proliferative expansion phase (22). To mathematically account for these results one recent study proposed a distributed division destiny time that is inherited through cell division, controlled in part by the proto-oncoprotein Myc and a separate ‘time-to-die’ mechanism (23).

Prior studies therefore provide the basis for considering the molecular mechanisms underlying B-cell decision-making and thereby quantify the degree of inheritance versus intrinsic noise. Generally, progeny cells are thought to inherit proteomic networks that mediate decisions (24, 25). Indeed, direct observation of protein abundances, indicated that the mixing time of inherited proteins exceeds two generations (more than 40 hours) (26). However, in studies of TRAIL-induced death, the concordance of cell fates among siblings decayed rapidly (with a half-life of 1.5 hrs) (24). Blocking protein synthesis slowed this loss of concordance, indicating a substantial role for intrinsic gene expression noise (27). To what degree gene expression noise or other sources of intrinsic molecular variability affect phenotypic heterogeneity of B-cell decision-making remains to be determined.

In the present study, we addressed the molecular underpinnings of the heterogeneity of cell fate decisions during B-cell expansion, and examined the roles of heritability and intrinsic noise. In order to obtain accurate, longitudinal, single-cell lineage information we established an experimental workflow for long-term live cell microscopy of primary B-cells and a computational workflow for image analysis and data processing. Resulting data was then used to parameterize a multi-scale mechanistic model that accounts for B-cell proliferation as a function of the interplay of cell cycle and apoptosis molecular networks, each receiving inputs from a multi-dimeric NFκB signaling model (25). Using an iterative systems biology approach, we could thereby quantify the degree of heritability and stochasticity in cell fate decisions, identify the sources of phenotypic cell-to-cell heterogeneity, and then computationally predict and experimentally confirm cell-intrinsic determinants of proliferative capacity.

Results

B lymphocyte proliferative decisions are predictable

To study the heterogeneity of underlying B-cell fate choices, we circumvented the complexity of differential BCR-antigen affinities by using the BCR-independent stimulus, toll-like receptor 9 ligand, CpG in dye dilution assays with Cell Trace Red (CTR) monitored at several time points
by flow cytometry (Fig. 1A). The diversity of cellular responses within the population could be determined by fitting the probability distributions of the established non-heritable phenotypic-scale Cyton model to the data (9, 14), revealing that cell division and cell death times of proliferating cells vary between approximately 4 to 12 hours and 8 to 48 hours, respectively (Fig. 1B). These distributed timings underlie the successive peaks of individual generations while the maximum total cell count was reached at about 72 hours (Fig. 1C). We extended the Cyton model to a single-cell probabilistic formulation (referred to as the phenotypic model), in which each simulation sampled from these distributions of division and death times (Fig. 1D), and showed that it also recapitulated the total and generation-specific population dynamics (Fig. 1E). An alternative, multi-scale model (referred to as the mechanistic model) of single-cell fate decision-making as a function of the molecular networks that control cell cycle, cell-survival/death, and the NFκB signaling network (Fig. 1F, Fig. S1 and S2) (25), also adequately recapitulated the total and generation-specific population dynamics (Fig. 1G).

Although the two models (phenotypic and mechanistic) account for dye dilution flow cytometry data, they are based on fundamentally different assumptions. The phenotypic model assumes that each cell at each generation makes a stochastic decision about its division and death times. The mechanistic model assumes that cellular decisions are predictably determined by a molecular network, whose constituents are naturally inherited from their predecessor; in this case variable cell fates are then largely due to pre-existing differences in the molecular networks of founder cells. Given these different assumptions, we sought to produce distinguishable predictions by plotting respective lineage trees. We found the lineage trees were fundamentally distinct: the phenotypic model yields many similar looking lineage trees, with a range of terminal generations within each single lineage (Fig. 1H); the mechanistic model yields highly diverse lineage trees but with all progeny within a lineage sharing the same terminal generation (Fig. 1I). The fundamental difference in the lineage trees generated by each modeling approach is a reflection of the degree of inheritance within each lineage. The phenotypic model (without inheritance) assumes division and death mechanisms reset with each division and generates lineage trees consistent with a mechanistic model in which signaling molecule concentrations are resampled from a wide distribution with each division (Fig. S3 A). Conversely, the recently reformulated phenotypic model with heritable division destiny (23) generates lineage trees consistent with the mechanistic model with perfect inheritance (Figure S3 B).
To quantify what degree of heritability best reflects B-cell fate choices during population expansion, we developed a long-term time-lapse microscopy workflow coupled to accurate single-cell tracking (Fig. 2A, Movie S1). The resulting lineage trees showed a high degree of regularity as each founder cell gave rise to progeny of generally equal terminal generations (Fig. 2B and C). However, founder cells differed dramatically in their proliferative capacity, or the terminal generation of the lineage tree. Such intra-lineage regularity and inter-lineage heterogeneity was also observed within specific B cell subtypes (Fig. 2D, E, F and G and Fig. S4). In sum, mathematical models of B-cell proliferative decisions must contain heritable mechanisms to accurately predict lineage data.

**Heterogeneity between lineages exceeds intrinsic noise**

To quantify the relative contribution of inter-lineage (pre-existing) and intra-lineage (intrinsic) heterogeneity, we used cell lineage data to parameterize these two sources of heterogeneity in the mechanistic model (Fig. 3A) such that simulated cell lineages (Fig. 3B) and generation-specific population dynamics (Fig. 3C) provide optimal fits to the experimental data. We considered two metrics in particular, inter-lineage differences in the terminal generation (Fig. 3D), and intra-lineage differences in terminal generation (Fig. 3E). Whereas the phenotypic model that lacks inheritance (probabilistic model) showed distributions that are quite distinct from the experimental result, the mechanistic model enabled tuning of pre-existing heterogeneity and intrinsic noise to fit the experimental data. We first determined how widely biochemical parameters within molecular networks would need to be distributed between founder cells to optimally account for the observed distributions in terminal generations (Fig. S5), and found it to correspond to an average coefficient of variation (CV) of 32% (Fig. 3F). Then we determined the width of the distribution of perturbations added to progeny molecular networks at each division that optimally accounted for differences in terminal generation within lineages, and found it to correspond to a CV of 7.9% (Fig. 3G and Fig. S6 and S7). The resulting, parameterized heritable mechanistic model produces lineage trees and generational distributions (Fig. 3B and C) that closely match those observed by live cell tracking (Fig. 2B-G).

The model predicts an average 32% coefficient of variation in founder cell metabolites however, due to molecular network architecture, not all signaling molecules are predicted to have the same magnitude of heterogeneity. We generated individual predictions for three key molecules (Bcl-2, Myc and CycD) that showed distinct CVs (Fig. S8). Using intra-cellular antibody-staining followed by flow-cytometry, an experimental technique that introduces an unknown amount of
technical noise, we obtained higher CVs, but found, remarkably, that the relative CV associated with Myc and CyclD normalized to the lowest observed (associated with Bcl2) was accurately predicted (Fig. S8C).

As modeling the experimentally observed lineage trees suggested that the symmetry of division must be tightly controlled (CV 7.9%), while signaling network states are widely distributed between founder cells (CV 32%), we asked whether this heterogeneity is also manifest in measurements of cell mass. Cell mass may be measured precisely by adapting quantitative phase microscopy, which measures the phase shift of light as it passes through and interacts with the material inside of cells (28), to live cell tracking experiments. We quantified mass distributions in founder cells (at 5 h post-stimulation) and in daughter cells following division (Fig. 4A and B) and found that founder cell masses were widely distributed (CV ~35%) but that cell division was generally highly symmetric with daughter cell masses having a CV of 4.8% (Fig. 4C). While it is unclear how cell mass asymmetry relates to B cell fate divergence during the proliferative expansion phase, our model simulations employ a shuffling (at each division) of protein concentrations (Fig. S6) or kinetic rate constants (Fig. S7) (that are mediated by enzymes not explicitly represented in the model). The concordance between molecular network distributions and mass distributions may suggest that they are the result of the same sources of heterogeneity and tightly controlled within lineages but distributed among founder cells.

**Proliferative fates are distinct due to non-genetic heterogeneity of founder cells**

We wondered whether differences in the molecular networks in founder cells that determine cell fate may be a consequence of genetic BCR diversity. (Though the TLR9 stimulus CpG is of course a BCR-independent stimulus, distinct BCRs may produce distinct signaling histories and hence molecule networks.) To address this question, we utilized BCR-transgenic mice in which >90% of B-cells carry the same BCR specific for the hen egg lysosome (HEL) antigen (29). Both B-cell subtypes showed differences in terminal generation between founder cell lineages that is largely indistinguishable from non-transgenic controls, which is evident from dye dilution assay (Fig. S9) and single-cell lineage tracking (Fig. 5A and B). This indicates that heterogeneity in the molecular networks of founder cells has non-genetic origins.

Many potential sources of non-genetic differences in founder cells are possible including, but not limited to, distinct histories of migration from the bone marrow, residence times in microenvironments, and prior antigen exposures. One possible source of the non-genetic differences in the state of molecular networks in mature splenic B cells is simply the
accumulation of intrinsic noise through successive generations in each founder cell’s history (Fig. 5C). Based on our estimates of intrinsic noise and measurements of division asymmetry, we found that more than 30 generations would be required to distribute the molecular network sufficiently to explain the distribution of terminal generations observed by lineage tracking (Fig. 5D). This number is compatible with estimates of the number of divisions separating hematopoietic stem cells and splenic B-cells (30, 31, 32 and personal communication Höfer T. 2017). Thus, while distinct founder cell histories are more than likely, they may not be a required driver of founder cell heterogeneity.

The population response is dominated by founder cells with low propensity for apoptosis

A consequence of the non-heritable phenotypic model, in which differences in terminal generation are predominantly determined by intrinsic noise during proliferation, is a prediction that many founder cells contribute similarly to population responses (Fig. 6A). In contrast, the mechanistic model, in which cell fate differences are predominantly driven by heritable differences in the molecular networks of founder cells, predicts that a remarkably small proportion of founder cells contribute to the expanded population. The lineages of these founder cells survive for a disproportionate number of generations, resulting in a large number of descendant cells that dominate the population response (Fig. 6B). Given the largely deterministic nature of cell fate decisions characterized by the mechanistic model, it ought to be possible to predict which founder cells will contribute highly to the population and undergo many divisions. Using multi-variate statistical regression we sought molecular network components that are associated with founder cells that are dominant contributors to the overall population response. Though no single reaction rate was a reliable predictor of proliferative capacity, in principle, knowledge of the founder cell molecular network could predict its proliferative capacity ($R^2 \sim 0.6$) (Fig. S10 A). While each individual network component is a poor predictor, combinations such as low procaspase 8, high Bar, and high XIAP are predictive of high proliferative capacity (Fig. 6C and Fig. S11). Indeed, classifying the state of founder cell molecular networks, in terms of propensity for apoptosis and cell cycle activity, (Fig. 6D, E and F) demonstrated that low apoptotic propensity was a strong indicator of founder cell terminal generation with non-dividing cells. Cells which did not divide could be separated from cells that divided 3 or more times by difference in the state of the founder cell’s apoptosis regulatory network (Fig. 6F).

Key molecular regulators in founder cells are identifiable
To characterize how non-genetic heterogeneity within the signaling networks of founder cells affects their proliferative capacity, we compared the distributions of terminal generations resulting from simulations containing only intrinsic noise with those that included distributed founder cell molecular network states within the apoptosis and cell cycle networks (Fig. 7A). We found that distributing the components of the apoptosis control network resulted in a bimodal distribution of proliferative capacities, including not only cells with a higher propensity to die curtailing even a first division, but also founder cells of lineages with much higher terminal generations. Additionally, distributing the components of the cell cycle network could result in a further increase, but also decrease in terminal generations (Fig. 7A and consistent with Fig. 6A).

A sensitivity analysis found that many reaction rates within the signaling network controlling cell fate could perturb the terminal generation of stimulated lymphocytes; however, terminal generation was predicted to be most sensitive to the state of the apoptosis control network in founder cells (Fig. 7B and Fig. S11). The sensitivity predictions are consistent with B-cell proliferation control by cRel and IkBε seen in previous experiments (33, 34). Perturbations targeting the cell-cycle network (such as Myc overexpression) are predicted to increase proliferation (Fig. S12 A) and the computational model recapitulated increases in total cell number and mean division number reported recently (Fig. S12B) (23). The prediction that terminal generation was most sensitive to small perturbations within the apoptosis control network such as those controlling the level and activity of caspases (Fig. S9), rather than cell growth and cell cycle control networks was surprising as inhibition of the mitochondrial amplification loop was not predicted by the model to substantially alter proliferation, a result also confirmed in a recent experimental study of Bcl-2 over-expression (23, 34). However, the sensitivity analysis predicted that inhibition of the central apoptotic caspase cascade would significantly enhance proliferation (Fig. S11 and S12D).

To generate a testable quantitative prediction of the degree of increased proliferation resulting from caspase inhibition, we simulated the effect of a 25 and 50% reduction in caspase activity, predicting a dose-dependent increase in the fraction of proliferating cells at late time points (Fig. 7C). To test this prediction experimentally, naïve B cells were treated with pan-caspase inhibitor Z-VAD (35, 36) prior to CpG stimulation; indeed, we found increased B cell proliferation at late time points in a dose-dependent manner (Fig. 7D), based on an increased fraction of cells positive for the proliferation marker Ki67 (Fig. S13). Examining why survival signals would enhance proliferative capacity in more detail, we found that while growing B-cells are protected from cell death in early divisions, as previously documented (34), during later divisions cell
death can occur even in growing cells, resulting in censoring of subsequent generations (Fig. S14). Thus cell death regulation emerges as a critical determinant of proliferative capacity of founder cells during the B cell immune response.

Discussion

In this study we showed that the diversity of B cell proliferative fate decision-making is largely the result of non-genetic differences between founder cell molecular network states. Indeed, utilizing a mechanistic mathematical model of the molecular interaction networks that determine cell proliferation decisions we were able to identify predictive markers of a high proliferative capacity and potential drug targets that control lymphocyte proliferation with high sensitivity.

The non-heritable phenotypic models of lymphocyte proliferation have demonstrated great utility in quantifying cell-biological parameters (such as division and death time) from snap-shot flow cytometry data. These models were based on the assumption that cells at each generation engage in stochastic decision-making process (7, 8, 10, 13, 34, 37). However, recent studies suggested revealed concordance between recently divided siblings (15) and a recent revision of the phenotypic model proposed to assign division destiny times to founder cells that is then inherited to descendants through the lineage . The present study extends these findings by demonstrating that heritability of phenotypic cell fate from one generation to the next can be understood as a function of the heritability of the underlying molecular networks. Leveraging the availability of sufficient high quality lineage data (complete lineage trees with no missing branches) we could quantitatively distinguish between the heritability vs seeming stochasticity of cell fate decisions through the lineage. Using the molecular mechanistic model of B-cell fate decisions we could then the quantify the contributions of heterogeneity in the molecular networks of the starting founder cell population (“extrinsic noise”) vs randomization of abundances of molecular network components at each division (“Intrinsic noise”) and how each noise source accounts for the heterogeneity of B-cell fates in the population. We found that the former with a CV of 32% substantially exceeded the latter with a CV of 7.9%.

Our findings are surprising in the context of previous studies quantifying the heritability of fate decisions by HeLa cells treated with the apoptosis-inducing TNF receptor superfamily ligand TRAIL (38). These studies also found cell fate concordance of recently divided siblings but reported a decay of this concordance with a half-life of 1.5 hours (a longer half-life when gene
expression was blocked) (38). Including such a substantial degree of intrinsic noise into our B-cell model would produce lineage trees indistinguishable from the non-heritable models (Fig. 1H, Fig. S3A), but not account for the experimentally observed lineages. Our study therefore revealed an unexpected degree of heritability suggesting that intrinsic noise is tightly controlled during lymphocyte proliferation. This may indicate that gene expression noise itself is limited or that topological characteristics of the network minimize the effect of that noise, as for example observed for the MyD88 pathway in macrophages (39) or the chemotaxis network of Escherichia coli (40). This is supported by the high degree of symmetry in the division mass (CV ~5%). Given that even this small asymmetry of division nevertheless may contribute to randomization of cell fate decisions, we surmise that the gene expression noise in proliferating B lymphocytes is remarkably small.

What may be the reasons for the discrepancies between HeLa cell apoptosis and B-cell proliferation? Aside from invoking cell type differences, we are wondering whether gene expression noise may be a function of whether cells are in a basal steady state or in a highly activated state in which protein production is maximized. The distinction between basal vs activated states may also pertain to the covariance of gene expression, a mechanism proposed to limit phenotypic heterogeneity in T lymphocytes (41); it is possible that gene expression variance relevant to HeLa cell apoptosis control is largely independent, whereas there is a greater degree of co-variance in the expression of genes regulating lymphocyte proliferation. We note that key regulators of lymphocyte proliferation are targets of the same stimulus-induced transcription factor, NFκB.

As B-cell expansion is largely determined by the highly heterogeneous proliferative capacity of mature B-cells, we sought to identify the underlying molecular determinants. We focused our studies on the molecular networks that are engaged regardless of the mitogenic stimulus, that is the IKK-NFκB signaling network, the cell cycle control network, and the apoptosis control network. Among these, expression variance that determined apoptotic susceptibility was found to be the most predictive of proliferative capacity (Fig. 6 and 7). However, when including receptors and their proximal signaling pathways, we may indeed find other prominent molecular sources of phenotypic heterogeneity. Indeed, in studies of T cell responses, interleukin-2 receptor α (IL-2Rα) was identified as a significant predictor of cell fate, with its expression varying across 5 orders of magnitude. In contrast, TLR9 (the cognate receptor for CpG) shows an expression variability of one order of magnitude (41-44), insufficient to be a key determinant in the stimulation conditions employed in our studies. In the absence of one key determinant of
phenotypic heterogeneity, a regression analysis of high-throughput simulations of the mechanistic model suggests that many individual molecules contribute to the cell’s overall propensity for proliferation and survival (Fig. 6C-F). Measuring expression variances directly is challenging as the functionally relevant variability in expression (e.g. 32% CV) is comparable to the technical variability of in single-cell assays such as flow-cytometry or fluorescence microscopy using antibodies or fluorescent reporters (42).

The mechanisms by which the distinct molecular network states of founder-cells are acquired in vivo likely involve a combination of intrinsically stochastic processes and distinct environmental exposure histories for each cell. In the absence of any other sources of variability the small degree of intrinsic noise quantified here may be sufficient to establish heterogeneous founder cell states only after more than 30 generations (Fig. 5D). Low abundance components such as mitochondria have recently been identified recently as a source of variability in TRAIL-induced apoptosis (45). Heterogeneous chromatin states may also contribute to the gene expression heterogeneity among founder cells. For example in T-lymphocytes, chromatin states have been found to contribute to IL-4 expression variability and maintain its state over ~3 days (46). This relatively long maintenance of chromatin state could represent a cellular memory of distinct exposure histories that contributes to founder-cell heterogeneity but is unlikely to contribute to intra-lineage heterogeneity.

What may be the physiological consequences of non-genetic differences in splenic founder B-cell molecular networks? Our work suggests the hypothesis that founder cell molecular heterogeneity contributes significantly to noise in lymphocyte clonal selection based on antigen-antibody affinity; as such, it might increase the diversity of the antibody repertoire. Antibody diversity is critical for an effective immune response (47) and a key indicator of immune health during aging (48). Our results therefore suggest that non-genetic heterogeneity of splenic B-cell plays an important physiological role in immune system function.

Following the proliferative expansion studied here, single lymphocytes can produce distinct cellular progeny, such as memory or plasma B-cells (10, 49, 50). These cell fate decisions are also not primarily stochastic but are thought to reflect the uneven impact of the germinal center microenvironment that fosters asymmetric segregation of signaling molecules, transcription factors and growth factors into daughter cell (49). Thus, despite the great degree of phenotypic heterogeneity, rigorous characterization of molecular interaction networks along with cellular interactions renders both B-lymphocyte proliferation and differentiation decisions surprisingly
predictable. Indeed, the origins of the phenotypic heterogeneity, that are thus identifiable, are of critical relevance to achieving the ambitious goals of personalized precision medicine.

**Methods**

**Experimental Procedures**

**B cell isolation and culture**

Spleen were harvested from 10-12 week old female C57BL/6 mice and Hemizygote BCR transgenic mice (The Jackson Laboratory, stock no: 002595). Homogenized spleen were incubated with anti-CD43 magnetic bead for 15 min at 4-8°C, washed with MACS buffer (Phosphate buffer saline, (pH 7.4), 0.5% Bovine serum albumin and 2 mM Ethylenediaminetetraacetic acid (pH 8)) and pass through LS column (Miltenyi Biotech). The purity of B cell was >95% based on B220 staining.

Marginal zone (MZ) and Follicular (FO) B cells were isolated using magnetic sorting followed by fluorescence activated cell sorting (FACS). Homogenized spleen were incubated with MZ and FO B cell biotin-antibody cocktail for 5 min at 4-8°C followed by incubation with Streptavidin microbeads for 10 min at 4-8°C, then washed with MACS buffer and passed through LS column. The eluted B cell population was enriched with MZ and FO B cells. MZ and FO B cells purified by FACS sorting. Briefly, the enriched B cell population was stained with B220-eF450 (eBioscience, #48-0452-82), CD23-PE (Biolegend, #101607), CD21-APC (Biolegend, #123411), CD93-FITC (eBioscience, # 11-5892-81), CD3e-FITC (eBioscience, #11-0031-82), CD11b-FITC (eBioscience, # 11-0112-82) and Ly-6C-Alexa 488 (eBioscience, # 53-5932-82). The stained cells were gated on B220⁺CD93⁺CD3e⁺CD11b⁺Ly-6C⁻ population to further exclude immature B cell and if any residual non-B cell population is present in the enriched B cell population. MZ and FO B cells were sorted based the expression CD21 and CD23 on B220⁺ cells (51, 52). The purity of MZ and FO B cells were ~ 98% (Extended data Fig. 3).

B cells were grown in fresh media with 1% penicillin streptomycin solution (Mediatech Inc.), 5 mM L-glutamine (Mediatech Inc.), 20 mM HEPES buffer (Mediatech Inc.), 1mM MEM non-essential amino acid (Sigma) and 1 mM Sodium pyruvate (Sigma), 10% FBS, and 55 µM 2-ME (Fisher Scientific).

**CellTrace far red (CTR) labelling and cell biological parameter estimation**
Purified B cells were labeled with CTR (1 μM) and stimulated with 250 ng/mL CpG ODN 1668 (Invivogen, tlr-1668-1). The cells were grown in 48 well tissue-culture plates (2X10^5 cells/mL) at 37°C, 5% CO₂ containing humidified chamber for a period of 5 days. Stimulated B cells were collected at different time points and stained with 7 aminoacitomycine D (7AAD) (Biolegend, #420404) to exclude dead cell. B cell proliferation and survival were measured in C6 Accuri Flow Cytometry (BD Bioscience) and 100 μL volume counted in each time point. After manual gating on viable (7AAD-negative) B cells FlowMax software was used to construct CTR histograms (14). All assays were performed in duplicate. The fluorescence of the undivided peak was manually identified for each time point and maximum-likelihood fcyton model parameter distributions were identified from 500 best fit solutions.

**Quantitative phase microscopy mass measurement**

Quantitative phase microscopy measures the phase shift of light as it passes through and interacts with the material inside of cells. Purified naïve B cells were stimulated with CpG (250 ng/mL) and grown in PLL treated microgrid array (Microsurface, MGA-250) in an environment controlled microscope as described previously. Images were acquired on an Axio Observer Z1 inverted microscope (Carl Zeiss Microscopy GmbH, Germany) using a SID4-Bio quadriwave lateral shearing interferometry (53) camera (Phasics) and an LD Plan Neofluor objective at 40X, 0.6 NA. Images were acquired with illumination from a 660 nm collimated LED (Thorlabs) every 3 minutes for 5 days.

**Z-VAD treatment and proliferation of B cells**

Proliferation of B cells were measured by Ki67 expression. Ki67 expressed in nucleus of cells which are in active cell cycle thus Ki67 used proliferation marker (54-56). Purified naïve B cells were pretreated with Z-VAD(OMe)-FMK (10 and 20 μM) (Enzo Life Science International, # BML-P416-0001) for 90 minutes followed by stimulation with CpG (250 ng/mL) for 4 days. Cell were harvested at 4 days, washed twice with PBS, stained with fixable viability dye eF520 (eBioscience, # 65-0867-14) for 30 min in 4°C, washed with PBS, permeabilization and staining with Ki67-PerCp-eFluoro 710 (eBioscience, # 46-5698-82) were performed using Fix and Perm cell permeabilization kit (ThermoFisher, # GAS003) as recommended. Briefly, cell were fixed with reagent A for 15 min at RT, washed with PBS, permeabilized and stained with reagent B for 20 min at RT followed by washing with PBS. Cell proliferation was measured by AccuriC6 and data were analyzed with FlowJo. Cell proliferation was determined by increase of cell
fluorescence index over the naïve B cells. The data are representative of three biological replicates.

**Time-lapse microscopy**

Purified naïve B cells, MZ and FO B cells were stimulated with CpG (250 ng/mL) and grown in a microgrid array (Microsurface, MGA-250) in an environment controlled microscope. Images were acquired on an Axio Observer Z1 inverted microscope (Carl Zeiss Microscopy GmbH, Germany) with a Pln Apo 20×, 0.8 NA objective to a Coolsnap HQ2 CCDcamera (Photometrics, Canada) using ZEN imaging software (Carl Zeiss Microscopy GmbH). Environmental conditions were maintained at 37°C, 8.5% CO₂ with a heated enclosure, and CO₂ controller (Pecon, Germany). Brightfiled images were taken every 3 minutes for 5 days. Microgrid arrays were coated with poly-L-lysine (PLL) before incubating cells to reduce the movement of the cells during image acquisition to facilitate lineage tracking.

**Cell tracking**

Cell tracking was performed on phase-contrast images using semi-automated software (34). A representative example of cell tracking and lineage tree generation can be seen in Supplementary Movie 1 (created in MATLAB). Briefly, founder cells were annotated in the first image and approximate linear paths were manually drawn for each identified founder cell. While PLL coating and low cell density enables many frames to be approximated to a linear path at early time points the temporal resolution of cell path annotations necessarily increased in later time points. Founder cells were tracked until a division or death event was manually identified and annotated. Death events were recognized by morphological changes in cell morphology consistent with apoptotic blebbing. Division events were manually annotated such that each daughter cell stores the cell ID of its mother cell enabling complete lineage reconstruction. The cell tracking software was further developed to report terminal generation, Δ terminal generation, generation-specific cell count, total cells within lineage and generation-specific standard deviation of division times for each founder cell. To aid comparison of computational models and live-cell microscopy the experimental cell tracking software was modified to take model-generated lineage information as input and perform identical statistical analysis and lineage tree generation for both *in vitro* and *in silico* experiments.

**Measurement of heterogeneity in intra-cellular protein expression**
Naïve mature B cells isolated from spleen as described above. Cell were washed twice with PBS, stained with fixable viability dye eF660 (eBioscience, # 65-0864-14) for 30 min in 4°C, washed with PBS, permeabilization and staining with primary antibody followed by secondary antibody were performed using Fix and Perm cell permeabilization kit (ThermoFisher, # GAS003) as recommended. Cell permeabilization and staining protocol followed as described above with little modification in staining time. Cells were stained 15 min with primary antibody and 15 min with secondary antibody. Cell proliferation was measured by CytoFLEX flow cytometry (Beckman Coulter) and data were analyzed with FlowJo. Protein abundances were graphed by fluorescence histogram and the CV was determined. CycD1 (SantaCruz, #sc-718), Myc (Abcam, #ab32072) and Bcl-2 (Cell Signaling, #2870S) were chosen as these are key regulators of cell cycle and cell survival.

**Mathematical Modeling**

**Single-cell Cyton model**

A single-cell computational model of the Cyton assumptions was created using MATLAB. Parameters describing the distributions of division and death times (Tdiv and Tdie respectively) in first generation and subsequent generation cells, along with proliferative fate parameters (pF) from FlowMax were used as input parameters into the model to recapitulate CTR population distributions. 10,000 founding cells were simulated with identical parameter distributions. To simulate the experimental time course each initial cell sampled from both division and death distributions created by the input parameters and the lowest of the two sampled times determined the fate of the cell. To recapitulate the proportion of cells progressing to the next generation (pF) as identified using FlowMax every cell sampled from a uniform distribution and determined if the value was greater than the appropriate progression factor for the cells generation; only if the sampled value was less than the progression factor was a division time sampled, otherwise the cell would die with the sampled Tdie. Cells which achieved a division fate triggered the introduction of 2 new cells that then independently sampled from the distributions corresponding to division 1+ events.

**Mechanistic mathematical model**

Mechanistic computational modeling was performed as described previously (34). Briefly; established ordinary differential equation (ODE) models of the NFκB signaling network (33), apoptotic control network (57) and cell cycle control network (58) were combined into single cell
simulations (Fig. S1 and S2). 125 starting cells were used for all simulations except where indicated, for example to generate data for identification of fate predictors in which 20,000 starting cells were simulated to improve statistical power. All parameters were maintained as published previously with the exception of coefficients of variance as indicated in Fig. 3 and Fig S4, S5 and S6. Cleaved PARP (cParp) and cadherin-1 (Cdh1) concentration thresholds triggered cell death and division respectively (for parameters see Dataset S1). Division events triggered the addition of two new daughter cells which inherited reaction rate parameters and metabolite concentrations from the mother cell. The timing of the cell cycle in individual cellular agents is determined by cyclin expression and resultant dynamics in the established mammalian cell-cycle module that in turn controls the timepoint at which cdh1 crosses the previously defined threshold (25, 58). Similarly cell death timings are determined by the dynamics within the established apoptosis module that determines death when cParp exceeds the published threshold (25, 57).

**Founder cell variance**

To distribute the state of starting cells all initial conditions were sampled from lognormal distributions and synthesis and degradation rate constants were sampled from a normal distribution with mean value from the established model and coefficient of variance indicated. An initial simulation-phase was run to enable individual cells to obtain a steady-state after which IKK activity was transiently increased recapitulating TLR9 activation. The coefficient of variance for the distribution from which founder cell parameters were sampled was fit using particle swarm optimization (max velocity: 2%, max CV: 50, min CV: 0, population size: 20) (Fig. S4D) in which each particle is a simulation with 125 founder cells distributed with a chosen CV and scored by the Euclidian distance of each generation and total cell count time courses to the experimentally measured data shown in Fig. 2A.

**Intrinsic noise**

Simulations of asymmetric distribution were performed by sampling the concentration of metabolites inherited by daughter cells. A division ratio was sampled with mean of 1 and coefficient of variance as indicated. The initial condition of all metabolites in one daughter cell were scaled uniformly by the sampled division ratio and inversely scaled in the other daughter cell. The coefficient of variance for the distribution from which daughter cell parameters were sampled was fit using particle swarm optimization (max velocity: 1%, max CV: 30, min CV: 0, population size: 20) (Fig. S5C) in which each particle is a simulation with 125 founder cells.
distributed with a chosen CV and scored by the KolChi-Square statistic comparing the distribution of intra-lineage Δ terminal generation between the simulated results and the single-cell tracked data (Fig. 3E). Simultaneous particle swarm optimization of both intrinsic and extrinsic noise was also performed to account for potential confounding effects between intrinsic and extrinsic noise, however a similar optimal CV for both values was found (Fig S5D). For simulations in which noise was added with division events parameters inherited by daughter cells were sampled from with mean value inherited from the mother cell and coefficient of variance as indicated (Extended data Fig. S6). Estimates of intrinsic noise accumulation were generated by repeated sampling distribution of all metabolite concentrations prior to simulation.

**Sensitivity analysis**

Sensitivity analysis of terminal generation was performed on a simulation without division-introduced variance. Scaling factors were sampled at 10 points logarithmically chosen from 10 fold lower and higher than the standard parameter value. A founder cell was simulated for each scaled parameter value leading to a simulation of 1680 distinct founder cells. The terminal generation reached by each founder cell was stored in order to quantify the effect of perturbations in each parameter on terminal generation.

**Statistical analysis**

To identify the best predictors of terminal generation a simulation with distributed founder cell parameters was performed with 20,000 founder cells. Parameter values, initial (steady-state) metabolite concentrations (i.e. Figure S4A x₁,₁, x₂,₁, x₃,₁) and terminal generation were stored for each founder cell. Least absolute shrinkage and selection operator (LASSO) regression was performed using either founder cell parameter values initial (steady-state) metabolite concentrations (i.e. Figure S4A x₁,₁, x₂,₁, x₃,₁) as predictors and maximum depth for each founder cell as observations using MATLAB’s statistical toolbox (1·10⁻⁶ relative convergence threshold for coordinate descent algorithm and 50-fold cross validation). 100 regularization coefficients λ were fitted and cross-validation mean square error was plotted for each λ penalty values.

**Caspase Inhibition Simulation**

125 founder cells were simulated with, distributed molecular networks. All caspase inhibition reactions were scaled as indicated, and Cdh1 positivity was used as a marker for proliferating cells.
**Founder cell module analysis**

125 single-cell simulations were run with founder cells distributed as described previously but with distributions of kinetic rates restricted to reactions only in apoptosis, cell-cycle or NFκB regulatory modules, or combinations thereof.

To visualize the space of founder-cell networks and the relation to terminal generation 5000 single-cell simulation were run with distribution of both cell cycle and apoptosis networks in founder-cells. Scores for "apoptotic propensity" were generated by identifying 5 chemical species within the apoptosis network that were significantly different between founder cells that divided one or more times and those that died (phosphorylated caspase 8, Bar, phosphorylated caspase 3, phosphorylated caspase 6 and XIAP). The concentration of these metabolites in founder cells was then normalized to unit variance with negative regulators of terminal generation inverted to match positive regulators and the average of the 5 values was the founder cell’s apoptotic propensity score. An identical approach was taken for cell cycle activity scores of chemical species showing significant differences within the cell cycle network of founder cells that divided to late terminal generation (3+) compared to those that did not (Cdh1, cyclin A, IEP, Exopolyphosphatase and Growth).

**Author Contributions**

S.M. performed the computational modeling, statistical analysis and figure preparation. K.R. performed the experiments, and cell tracking with S.M. T.Z. performed quantitative phase microscopy with K.R. A.H. coordinated the study. All authors designed the study, interpreted the findings, and wrote the paper.

**Acknowledgements**

The computational modeling code is available in the supplementary material. We thank lab members Yi Liu and Diane Lefaudeux, as well as Dinesh Rao (UCLA), Roy Wollman (UCLA), and Ken Duffy (Hamilton Institute) for critical discussions and/or reading of the manuscript.
References


Fig. 1.

Phenotypic and mechanistic cell fate models of lymphocyte expansion predict similar cell population dynamics but distinct single cell lineage trees. **A)** Histogram of cell generations revealed by a dye dilution timecourse (38 h, 48 h, 70 h and 90 h) of CTR-labeled, CpG-stimulated B cells. Plots indicate undivided founder cells (generation “0”, red line), and total cell count (black histogram), and best fit of the probabilistic Cyton model using computational tool FlowMax(14) (green). **B)** Probability density plots of log normal distributions of division and death time for founder cells (Tdiv0/Tdie0) and subsequent generations (Tdiv1+/Tdie1+) following FlowMax fitting to the data in panel A. **C)** Line graphs of total cell count (black) and individual generations (indicated colors) over time following FlowMax fitting from panel A, compared to measured cell counts (orange crosses). **D)** Schematic of the single-cell phenotypic cell-fate model in which cell fate is determined by independently sampling division and death times at each generation from appropriate distributions, the shortest of which determines cell...
fate. **E)** Stacked area plot of total cell counts of the probabilistic single-cell model with the contribution of each generation marked with the same color as used in panel C. **F)** Schematic of the mechanistic single-cell model in which molecular signaling networks determine cell-fate and are inherited during cell division(34). **G)** Stacked area plot of total cell counts of the mechanistic single-cell model with the contribution of each generation (as in panel E). **H)** Lineage tree generated using the probabilistic single-cell model. **I)** Lineage tree generated using the mechanistic single-cell model.
Fig. 2.

Lymphocyte lineage trees produced by long-term live cell microscopy. A) Schematic of the live-cell time-lapse microscopy pipeline. See Movie S1 for annotated example of single-cell lineage tree construction. B) Representative lineage tree for CpG stimulated splenic mature B cells. C) Stacked area plot of total cell count of the tracked cells with the contribution of each generation marked with similar color used in Fig. 1C. D) Representative lineage tree for CpG stimulated Follicular B cells. E) Stacked area plot of total cell count of the tracked cells with the contribution of each generation marked with similar color used in Fig. 1C. F) Representative lineage tree for
CpG stimulated splenic marginal zone B cells. G) Stacked area plot of total cell count of the tracked cells with the contribution of each generation marked with similar color used in Fig. 1C.
Quantification of extrinsic and intrinsic cell-to-cell variability by interpreting lymphocyte lineage trees produced by long-term live cell microscopy. **A)** Schematic of the optimization approach used to quantify the level of cell-to-cell variability between founder cells and within lineages. All parameters are taken from (25) and the coefficient of variance (CV) between founder cells and between cells within lineages are optimized using particle swarm optimization to minimize the distance between the live-cell tracking and the simulation. **B and C)** Representative lineage tree and cell counts generated by the mechanistic model with 32% coefficient of founder-cell variance and 7.9% coefficient of asymmetric division-associated variance of the daughter cells. **D)** Violin plot of the distribution, mean (○) and median (—) of the maximal terminal generation obtained from 125 founder cells in experimentally tracked cells (Exp), mechanistic model simulations with indicated extrinsic coefficient of variance (Mech), and the single-cell
probabilistic model (Prob). **E**) Violin plot, mean (○) and median (—), of the distribution of Δ (difference between the maximum and minimum) terminal generation within each lineage in 125 experimentally tracked lineages (Exp), mechanistic simulations with 31% extrinsic coefficient of variance and indicated intrinsic coefficient of variance (Mech), and the single-cell probabilistic model (Prob). **F**) Violin plots of the distributions of each molecular species concentration within 125 simulated founder cells required to recapitulate the experimental lineage tree. Median coefficient of variance for all metabolites: 32%. **G**) Violin plots of the distributions of each molecular species concentrations in the 100 simulations with the lowest distance between the Δ terminal generation distribution of experimentally tracked cells and the model (optimum CV 7.9%).
Fig. 4.

Validation of extrinsic and intrinsic cell-to-cell variability predictions by live cell interferometry. **A)** Representative quantitative phase measurement of a single founder cell (M) and the corresponding daughter cells (D1 and D2) at 0 min, 1998 min and 2004 min after the start of imaging. The color code represents area density [mass (pg) / area (μm²)] with blue color as lower area density and red as higher area density. **B)** Biomass over time for the annotated cells. **C)** Violin plot of the distribution of biomass ratios for 188 founder cells at 5 h post-stimulation and daughter cells resulting from 93 division events tracked by live cell quantitative phase microscopy.
Non-genetic heterogeneity is the source of cell-to-cell heterogeneity. 

**A)** Lineage trees of HEL transgenic follicular B cells (HEL-Tg FO B cells).

**B)** The terminal generation reached by each dividing founder cell of wild type (WT) FO B cell and HEL-Tg FO B cell.

**C)** Schematic depicting the origins of the splenic founder B lymphocyte, and the potential accumulation of intrinsic noise at each division. Hematopoietic stem cells (HSC) differentiate into multi-potent progenitor cells (MPP), MPPs differentiate into common lymphoid progenitors (CLP), CLPs differentiate into progenitor B cells (Pro), Pro cells differentiate into precursor B cells (Pre), Pre cells differentiate into transitional 1 B cells (T1), T1 cells differentiate into transitional 2 B cells (T2), T2 cells differentiate into follicular B cell (F). HSC to Pre cell differentiation occurs in bone marrow and T1 to F differentiation occurs in spleen.

**D)** Violin plot of the distribution of terminal generations reached by 125 simulated founder cells (gray) with the indicated number of partitioning steps in which the molecular network is distributed (CV 7.2%) prior to stimuli and experimentally tracked lineage information (red).

**Fig. 5.**
Identifying non-genetic determinants of founder B cell proliferative capacity. A) and B) Line-plots of the number of descendant cells from 250 founder cells over time (0-120 h). (A) Descendant cell numbers generated using the single-cell probabilistic model and (B) descendant cell numbers generated using the mechanistic model, colored to indicate the contribution of each founder cell’s descendant relative to the maximum. C) Three-dimensional scatterplot of the terminal generation (color code) reached by 20000 founder cells with distributed kinetic parameters against the values for the first 3 non-redundant predictors identified by LASSO (XIAP degradation rate, pro-caspase 8 degradation rate and Bar expression rate). D) Scatter plot of the terminal generation reached by single-cell simulations. The position of each cell is determined by state of 5 key signaling molecules in each of the apoptosis and cell cycle networks of each founder cell. The color code for terminal generation is shown on the right. E) Heatmap of the data in D) binned into 8 equally spaced bins for each module’s activity. F) Density contour map for the data in D). Contours are drawn at 0.85, 0.95 and 0.99 indicating the location of the highest density regions for each generation in the space created by the apoptosis and cell-cycle signaling network states.
Fig. 7.
Identifying and targeting the molecular determinants of lymphocyte proliferation. A) Violin plots, mean (○) and median (—), of the distributions of terminal generations reached in a simulation of
250 founder cells with identical starting molecular network states and subsequent asymmetric division (intrinsic noise), with the addition of distributed starting state of the apoptosis network, cell cycle network, NFκB network and combinations of these. B) Sensitivity analysis heatmap of change in terminal generation reached by the mean founder cell when parameters within each indicated network are scaled by the indicated parameter modifiers across 2 orders of magnitude. C) Bar graph of the proliferation increase predicted by the mechanistic simulation at day 4 following caspase inhibition. 125 founder cells were simulated with caspase inhibition reactions scaled as indicated, and Cdh1 positivity was used as a marker for proliferating cells. ** indicates p≤0.01 by Student’s t-test. D) Bar graph of experimentally measured proliferation increase following inhibition of caspase activity using Z-VAD (10 and 20 µM) at day 4. Ki67 expression was used as a marker for proliferating cells. Data shown is from simulation and biological triplicates at the 4 day timepoint following CpG stimulation. ** indicates p≤0.01 by Student’s t-test.
Supporting Information

Fig. S1.

Schematic of the mechanistic computational model in which intra-cellular signaling-networks determine cell fate by thresholding cPARP and Cdh1 for death and division decisions respectively (34). Models recapitulating NFκB signaling (33), the cell cycle (58) and apoptosis (59) are connected through NFκB target genes (34). Schematic shows only a few representative molecular regulators, e.g. αβε: NFκB inhibitors α/β/ε, C3/6/8: caspase-3/6/8, cPARP: cleaved poly (ADP-ribose) polymerase, CyC: cytochrome C, DISC: death-induced signaling complex, GM: general machinery, CycA/B/D/E: Cyclin A/B/D/E, NEMO: NFκB essential modulator, Rb: Retinoblastoma, XIAP: X-linked inhibitor of apoptosis protein.
Fig. S2.

Complete diagram of the mechanistic computational model in Systems Biology Markup Language (SBML). Each chemical species is represented by a rectangle the concentration of which is calculated by solving an ordinary differential equation. The terms of each differential equation are determined by the reactions each metabolite is involved in represented by arrows. If a reaction rate is increased by a metabolite a line ending with a circle is drawn to the reaction
from the promoting metabolite. Inhibition is denoted by a bar-headed line from the inhibiting metabolite to the reaction. The concentration of all chemical species at each time point is calculated by the set of differential equations. The complete model consists of 144 ordinary differential equations representing the rates of change of the 144 chemical species and 329 parameters controlling reaction kinetic rates. The differential equation system is solved using MATLAB’s ODE15s solver in two phases, the first to obtain a steady state and the second in which NEMO/IKK activity follows an input curve representing CpG activation as defined previously (33).
Fig. S3.

Schematics (top) and single-cell lineage trees (bottom) generated by computational models in which cell fate mechanisms are reset with each generation due to the magnitude of intrinsic noise (A) and models in which the cell fate mechanism is inherited by daughter cells (B). A) The Cyton model (left) (9) and the mechanistic model (right) (60) produce similar lineage trees if the mechanistic model resamples initial conditions (denoted $x_n$ as in Fig. S6), or kinetic parameters (as in Fig. S7) are resampled from a wide distribution with each generation. B) The revised phenotypic model (left) (23) and the mechanistic model (right) produce similar lineage trees if the resampling is reduced to zero such that the mechanistic model permits daughter cells to perfectly inherits kinetic parameters and initial conditions from the mother cell.
Fig. S4.

Purity for B cell subtypes. FACS plots for mature splenic B-cells (A) with subtypes (B and C). The purity is identified by anti-CD21 and anti-CD23 staining. Follicular B-cells (B) are identified as CD23^{hi} CD21^{hi} while marginal zone B-cells (C) are CD23^{low} CD21^{hi}. Bottom:
Fig S5.
Mechanistic computational model quantification of the distribution of founder cell molecular network states. A) Schematic of mechanistic single-cell simulation with extrinsic noise in which each founder cell samples parameters from a distribution around the standard value before being simulated until steady-state is achieved and subsequently stimulated to proliferate. B) Stacked area plot of total cell counts and contribution from each generation for increasing coefficient of variance (CV) for founder cell parameter distributions. C) Violin plots, mean (o)
and median (—), of the distribution of terminal generation and $\Delta$ (difference between the maximum and minimum) terminal generation for each founder cell in the mechanistic simulation with 125 founder cells kinetic parameters distributed by the indicated coefficient of variance. **D)** Scatter plot of particle swarm optimization convergence of the Euclidian distance to the experimental generational cell counts over time (Fig. 3C) by optimizing founder cell distribution coefficient of variance. Each point represents a simulation of 125 cells and 1072 total simulations were performed. **E)** Stacked area plot of total cell count of the population with the contribution of each generation marked with a distinct color for a simulation with optimal fit to experimental data. **F)** Representative lineage tree for the optimal model.
Fig. S6.

Quantification of intrinsic noise introduced through asymmetric division of molecular regulators to daughter cells. **A)** Schematic of mechanistic computational model in which, during a division event, the initial metabolite conditions of all metabolites are partitioned unevenly between the two daughter cells. **B)** Bar graph, mean (○) and median (—), of distribution of difference in intralineage maximum and minimum terminal generation for the mechanistic model with indicated coefficient of asymmetric division-associated variance, and the probabilistic model (prob). Boxes represent 25th and 75th percentiles respectively. Whiskers extend to the furthest observation with 1.5 times the interquartile range from the box. Outliers which lay outside the whisker length are plotted individually. 125 founder cells were simulated per condition. **C)** Scatter plot of particle swarm optimization convergence of Chi-Square statistic between computational model to
experimentally measured distribution of Δ terminal generation (Fig. 3E) by optimizing the coefficient of variance of daughter cell metabolite partitioning in the computational model. Each point represents a simulation of 125 cells and 1072 total simulations were performed. D) Simultaneous optimization of founder-cell variability and asymmetric division by minimizing the metrics described in Fig. S4 and panel C). Each point represents a simulation of 125 cells and 1072 total simulations were performed.
Fig. S7.

Intrinsic noise introduced through division-associated kinetic rate variance. **A)** Schematic of mechanistic computational model in which each division introduces variance of kinetic rate parameters. **B)** Box plot, mean (○) and median (―), of the distribution of difference in intra-lineage maximum and minimum terminal generation for the mechanistic model with indicated coefficient of division-associated kinetic rate variance, and the probabilistic model (Prob). Boxes represent 25th and 75th percentiles respectively. Whiskers extend to the furthest observation with 1.5 times the interquartile range from the box. Outliers which lay outside the whisker length are plotted individually. 125 founder cells were simulated per condition.
Fig S8.

A) Violin plot of the distribution of CycD (green), Myc (blue) and Bcl2 (red) in simulated founder cells. Three key metabolites extracted from (Figure 3F).

B) FACS staining for Bcl2, Myc and CycD1 in naïve mature B cells.

C) Bar graph of the coefficient of variance (CV) of Bcl2, Myc and CycD in the model (left) and experimentally determined by FACS (right) relative to Bcl2. Model errors bars are from three replicates while experimental results are from two replicates.
Fig. S9.

Heterogeneity of B-cell proliferative response is maintained in BCR transgenic mice. Histograms of CellTrace Far Red (CTR) labelled marginal zone (top) and follicular (bottom) B-cells from wild type (WT) and HEL-Tg mice were stimulated with CpG. Proliferation of B cells were measured at 14 h, 36 h, 48 h, 67 h, 88 h and 110 h. Peaks corresponding to distinct generations are indicated.
Fig. S10.

Identifying non-genetic determinant of B cells proliferative capacity. **A)** Line plot of unexplained variance (1- $R^2$) for each metabolite included in regression performed by LASSO with kinetic parameters as predictors. **B)** Scatter plot of the terminal generation reached by 20000 progenitor cells with distributed kinetic parameters against the value for the first 3 non-redundant predictors identified by LASSO (XIAP degradation rate, pro-caspase 8 degradation rate and Bar expression rate).
Fig. S11.

Sensitivity of terminal generation to individual parameter perturbation. Heatmap of terminal generation reached by the mean founder cell when the indicated reaction parameters are perturbed as indicated. Caspase 3, 6 and 8 degradation show identical sensitivity.
Fig. S12.

Computational predictions of Bcl-2 and Myc over expression recapitulate experimental observations. A) Line plot of mean division number in the mechanistic model with WT parameter (gray), 10x Bcl-2 expression (Bcl-2-Tg, black) and 10x Myc expression (Myc, red) (Recapitulates experimental data in Figure 2 in Heinzel et al. (23)). B) Stacked area plot of total cell counts and contribution from each generation. (Recapitulates experimental data in Figure 2 Heinzel et al. (23)) C,D,E) Line plots of the number of descendant cells resulting from each founder cells over time in Myc overexpression (10x) (C), Z-VAD caspase inhibition (50%) (D) and Bcl-2 overexpression (10x) (E) simulations, cells dividing at least once (solid lines), cells dividing to generation 5 (dotted lines) in the perturbed condition indicated (red) and in the wild-type simulation for reference (black). All simulations are of 125 founder cells.
Fig. S13.

Pan-caspase inhibitor (Z-VAD) increases the number of proliferating cells. Flow cytometry plot of proliferating follicular B cells identified by Ki67 expression at 4 days after stimulation. Dead cells were excluded from the analysis using fixable viability dye eF520. A) Naïve follicular B cells, used as negative control for Ki67 staining, B) follicular B cells stimulated with CpG, C) follicular B cells treated with 10 μM Z-VAD prior to CpG stimulation and D) follicular B cells treated with 20 μM Z-VAD prior to CpG stimulation.
Fig. S14.

Bar graph of the fate of the growing cells within a mechanistic simulation of 1000 founder cells with standard parameters (A, recapitulates cell-tracking data in Figure 2C Shokhirev et al.(34)) and 10% inhibition of caspase expression (B).
Movie S1. Representative time-lapse microscopy movie showing tracking of a single founder B cell through its entire lineage.

**Left:** The spatial location of a single follicular founder B cell stimulated with CpG tracked through its entire lineages using semi-automated cell tracking software. Division events are indicated by the annotation of two daughter cells of distinct generation (color) from the mother cell. Death events are indicated by a white highlight in the frames in which apoptosis is annotated. **Right:** Complete lineage tree constructed over time from left (0h) to right (82h).