Quantifying the dynamics of hematopoiesis by *in vivo* IdU pulse-chase, mass cytometry and mathematical modeling

Amir Erez[1*], Ratnadeep Mukherjee[1*], Alexandre Day[2], Pankaj Mehta[2], Grégoire Altan-Bonnet[1†]

1 Immunodynamics Group, Cancer and Inflammation Program, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA **2** Department of Physics, Boston University, Boston MA, USA

* These authors contributed equally to this work.

†gregoire.altan-bonnet@nih.gov

Acknowledgments

This work was supported by Human Frontier Science Program grant LT000123/2014 (Amir Erez), the Gordon Moore foundation and by the Intramural Research Program of the NCI, NIH. We thank April Huang for her help with tagging the mice.

Abstract

We present a new method to directly quantify the dynamics of differentiation of multiple cellular subsets in unperturbed mice. We combine a pulse-chase protocol of IdU injections with subsequent analysis by mass cytometry (CyTOF), and mathematical modeling of the IdU dynamics. Measurements by CyTOF allow for a wide range of cells to be analyzed at once, due to the availability of a large staining panel without the complication of fluorescence spill-over. These are also compatible with direct detection of integrated iodine signal, with minimal impact on immunophenotyping based on surface markers. Mathematical modeling beyond a binary classification of surface marker abundance allows for a continuum of cellular states as the cells transition from one state to another. Thus, we present a complete and robust method for directly quantifying differentiation at the systemic level, allowing for system-wide comparisons between different mouse strains and/or experimental conditions.

keywords: IdU, Differentiation, CyTOF, Hematopoiesis, Bone-marrow, B-Cells, Neutrophils

Introduction

The origin, development and turnover of leukocytes is a longstanding yet fundamental question in immunology. Fifty years ago, Ralph van Furth and Zanvil Cohn (1) used a pulse-chase technique to monitor the development of monocytes and macrophages. By using ^3H incorporation and meticulous microscopy, they shaped our understanding of innate immunity by pointing out that tissue macrophages most likely originated from blood monocytes, that themselves originated from promonocytes in the bone marrow. Much has changed since the 1960's yet the study of hematopoiesis and leukocyte development continues to provide interesting questions and intriguing answers. Importantly, the field of differentiation and hematopoiesis suffers from a noticeable gap between in-vivo and in-vitro studies. Whereas many in-vitro and cell-transplantation assays revealed what hematopoietic stem cells (HSC) and other leukocytes are *capable* of differentiating to, these studies do not tell us what actually happens in a healthy adult. Indeed, since cell-transplantation experiments usually involve depleting the host populations (by eg., irradiation), the resulting "empty" mouse needs to rapidly reconstruct its entire immune system, operating in highly inflammatory conditions. It is thus unclear how to apply knowledge gained from such perturbed conditions on healthy organisms at homeostasis.

Several approaches currently exist to study hematopoiesis, each with its merits and limitations. Ablation of relevant populations following a knockout of a gene, eg., a transcription factor such as FoxP3 (2), typically serve as conclusive evidence to define master regulators of leukocyte differentiation. Such genetic approach to probe leukocyte development, can be deceptive, as was recently demonstrated in the context of thymic negative selection of self-reactive T-cells (3). Alternatively, much can be

gained by studying unperturbed healthy adult mice. Computational approaches (4, 5, 6, 7) use an algorithm for lineage inference without requiring time-series data, by hypothesizing that there is a continuous relationship between maturing cell subsets. As hematopoiesis is both continuous and asynchronous, the full spectrum of cell types exists in a single bone-marrow sample from a healthy individual. Though powerful, such approaches are currently limited to particular differentiation scenarios and lack a direct observation of the dynamics. Indeed, most of these approaches map out differentiation along a pseudo-time dimension based on the proximities of cell states.

In contrast, others (8) have combined inducible in-situ labeling (the so-called barcoding) and mathematical modeling to reveal the nature of the HSC stem cell contribution to hematopoietic maintenance. Specific labeling of HSC cells in mice is achieved by expression of a modified Cre recombinase from the Tie2(Tek) locus. Subsequent articles (9) have dissected the benefits of such in-situ genetic tagging in comparison to traditional techniques. An important advantage for in-situ labeling is its non-perturbative nature and its flexibility in assigning labeling targets. Crucially, to interpret the resulting time-series it has been essential to use mathematical modeling. However, such genetic manipulation is not easy or quick, and can severely limit the applicability of the approach on different strains or species.

A giant leap in the field of single-cell RNA sequencing demonstrated that dynamic maps of differentiation could be inferred using a clever model of RNA maturation and splicing. Kharchenko and colleagues (10) introduced a powerful approach to reconstitute the whole dynamics of RNA transcription and maturation in entire tissues or organisms. In turn, they used their RNA dynamic map to infer the dynamic relationship between the cells themselves. This method (called *velocyto*) indeed delivers

"real-time" measurements of cell differentiation, when all previous methods in scRNAseq were confined to inferring pseudo-times. Yet, the discrepancies between RNA and protein studies(11) are warranting the development of dynamic methods in the field of mass cytometry (CyTOF) to leverage the protein-resolution of cytometry with the labeling depth of CyTOF.

Proliferation specific tags such as BrdU (*5-Bromo-2-Deoxyuridine*), EdU (*5-Ethynyl-2-Deoxyuridine*) and IdU (*5-Iodo-2'-deoxyuridine*), jointly referred to as XdU in this publication, offer an alternative for protein-based cytometry (FACS or mass cytometry). XdU gets inserted in the genomic material of cells, when they are undergoing DNA replication. Accordingly, proliferating cells (and only them) get tagged with XdU, and, upon division after tagging, cells pass on half their XdU to their daughters. The tagged cells can also differentiate by changing their phenotype (by *e.g.*, expressing different surface markers), in which case they carry the XdU content, undiluted, to their new phenotypic compartment. XdU can be administered in the drinking water (12), or injected intraperitoneally or intravenously, followed by flow-cytometry and fitting of XdU time-series experiments to a mathematical model. The resulting time-series give insight about proliferating populations and their progeny, *in vivo*, without perturbing the animal and without requiring the elaborate genetics used in the in-situ fate-mapping techniques outlined in the previous section. Moreover, because administration of the tag stops at some point, the time-series can observe both the labeling and the de-labeling phases, thereby providing more information than the equivalent time-series from in-situ tags.

Upon intraperitoneal injection in mice, XdU is made bio-available for less than an hour after which it gets metabolized. Cells that are undergoing proliferation get tagged practically instantaneously; upon

proliferation, they dilute their XdU content; upon differentiation, they bequeath the XdU tag to their descendant. In the context of monocyte differentiation, such a flow cytometry (FCM) based method has been previously used (13). However, FCM is disadvantaged by lack of breadth of markers to simultaneously study different cell groups. Moreover, antibody staining for BrdU requires harsh permeabilization of the cells (for the anti-BrdU antibody to reach its nuclear target) and this can be detrimental to the detection of various epitopes and surface markers. EdU requires a simpler protocol, yet relies on "click" chemistry which is incompatible with certain common dyes. Instead, here, we propose to use a combination of IdU injection followed by analysis using mass cytometry (CyTOF), which circumvents these problems. IdU is immediately detected as elemental Iodine in the CyTOF without the requirement of any additional experimental step. The CyTOF allows for a wide staining panel, capturing several cell groups in detail and providing a systemic view of differentiation. Whereas IdU injection in mice has been used in microscopy imaging, for neogenesis (14), cell cycle analysis (15), and proliferation (16), its use in single-cell mass cytometry remains limited. IdU has been used *in vitro* on CyTOF measurements (17) to study cell cycle and to immunologically profile proliferating versus non-proliferating cells in the mouse bone marrow (18), however, it remains an untapped resource to study differentiation. Even so, studying hematopoiesis by an IdU pulse-chase experiment on the CyTOF carries with it significant qualitative advantages with respect to existing methods. We develop a combination of a pulse-chase IdU experiment and model-based data analysis to study hematopoiesis *in vivo* in wild-type adult mice. Our method allows to simultaneously track multiple populations, thereby giving a systemic view of differentiation, making full use of the power of mass cytometry. Our method can be applied on a variety of systems to

generate systemic, quantitative comparisons, directly observing and quantifying differentiation.

Here we present our method, a combination of an *in vivo* mouse experiment and computational analysis and model fitting which yields direct observation of the differentiation flux during hematopoiesis. We first demonstrate the power of mass cytometry by identifying immune cell subsets by high dimensional clustering followed by plotting changes in their IdU profile over time. Then we illustrate our mathematical modelling approach by using conventional Boolean gating on bivariate scatterplots to track three different transitions in mouse bone marrow: neutrophil maturation, Pre-B to Immature-B cell development, and Immature-B to Mature-B cell development.

Methods

Experimental sketch

A schematic of the experimental plan is depicted in Fig. 1A. To track the differentiation of monocytes we injected IdU into the peritoneum of 48 C57BL-6J age-matched mice at 12 hour intervals. Mice are then sacrificed in cohorts of 16 on consecutive days, and their bone marrows were harvested, prepared as single-cell suspensions, stained with antibodies, and analyzed by CyTOF mass cytometry.

Mice

6-8 weeks old female C57BL/6J mice were used for the study. All animals received food and water *ad libitum*. All animal experiments were in accordance with the ethical approval of the institutional animal ethics review board.

IdU Injection

IdU (*5-Iodo-2'-deoxyuridine*) was purchased from Sigma-Aldrich (ref. I7125). All mice were injected in their peritoneum with a solution of 0.4mg of IdU diluted in 200μl of sterile PBS, adjusted to a pH of 8.5-9. The injection times were set to sacrifice 16 mice per day over consecutive days, with each day covering as evenly as possible the time-range between 1 and 228 hours thus effectively representing a separate experiment. Each day included an un-injected control mouse in the cohort. Each time-point is represented by 2-3 different mice; the injection/collection regimen was structured to prevent batch effects in the pulse-chase dynamics.

Cell preparation for CyTOF analysis

We euthanized the mice at various time points post IdU-injection, harvested and prepared their femoral bone marrow as single-cell suspension. Cells were then stained with metal-chelated antibodies according to manufacturer's protocols (Fluidigm). Fc block antibody was included in the staining cocktail to minimize nonspecific antibody binding on Fc-expressing cells. The panel of antibodies can be found in the supplementary information.

Data pre-processing

We used the data normalization package from the CyTOF manufacturer (Fluidigm) to correct for signal drift during acquisition. We then use manual gating to remove the calibration beads and dead cells (Cisplatin^{+}) and to select singlet cells (peak at 2n for ^{191}Ir-DNA intercalator). Further details of the gating strategy can be found in the supplementary information.

High-dimensional analysis of mass cytometry data

We applied HAL-x (Hierarchical Agglomerative Learning), a custom-designed Python pipeline (https://pypi.org/project/hal-x/), to globally define cluster of differentiations (CD) within leukocyte populations and to analyze their dynamics of IdU intake and dilution. HAL-x uses cross-validated random-forest classification to identify in an unsupervised yet tunable manner CD based on the levels of surface markers. Upon automatically clustering the CyTOF data into leukocyte sub-populations, we then measured their frequency in each of the harvested bone marrow and quantified their IdU intake. This resulted in a global map of the dynamics of differentiation within leukocytes of the bone marrow.

Defining a phenotypic continuum of differentiation

Classical FCM analysis involves displaying a density map of observed fluorescence variables two at a time, in a two-dimensional (bi-axial) plot, and then manually selecting sub-populations (gating) based on local abundance peaks. Such an approach is valid when dealing with large differentiation steps (19). More often than not, cells in fact differentiate and change their surface marker expression profile, in a more gradual and continuous manner. Previously, a *slice* processing helped gain insight into the heterogeneous T-cell response to stimulus (20, 21). We employed this *slicing* method to study differentiation along phenotypic continua of leukocyte subpopulations within the bone marrow of mice.

We selected our focus population in a two-dimensional grid based on surface expression of markers of interest (Fig. 1B) whereby each grid voxel would then be treated as a phenotypically distinct population. Each voxel's full time-series is then averaged over experimental replicates. Next, we define a one-dimensional direction in this phenotypic continuum by

looking at the direction where the density distribution changes the most - namely, the first principle component (PC1) of the principal component analysis of the 2D density maps. We are now in a position to interrogate the differentiation dynamics along this path. For that purpose, we first develop our mathematical model and then fit it to the data.

Mathematical model

Our IdU experimental setup presents a unique modeling opportunity to quantify the dynamics of differentiation and proliferation. Our modeling accounts for the simple transfer of the IdU from the proliferating populations to their progenies. Such transfer-function models are ubiquitous and their time-series representations have been well-studied in many scenarios (22). From the modeling perspective, our innovation lies in the application of this simple formalism to the question of cell differentiation, and the reliance on high-quality dynamics afforded by the CyTOF/IdU pulse chasing. In Fig. 1C, we show a schematic of the model which captures the flow of IdU from a "Source" to a "Destination" population. k is the rate going from the source population to the destination population; d is the "disappearance rate" of the destination, either through cell death or differentiation to different cell types or tissues, these cells leave our "Destination" population. The schematic in Fig. 1C implies an ordinary-differential-equation for the rate of change of IdU signal intensity,

$$\frac{dDest(t)}{dt} = k \cdot Source(t) - d \cdot Dest(t) \tag{1}$$

which when inverted produces the integral equation

$$Dest(t) \quad = \quad k \int_{t_0}^{t} e^{-d(t-t')} Source(t')\, dt' + e^{-d(t-t_0)} Dest(t_0) \qquad (2)$$

This integral equation takes the (experimental) source population's IdU dual count, and fits values for k and d which best explain the destination's IdU signal. To apply our model on the experimental data, we relied on the time-discrete version of the integral equation, where the integral is approximated as a sum of discrete time-points, a well-established technique (22). A similar mathematical structure was previously used to model genetic fate-mapping data (8, 23) but with crucial differences: (i) whereas the genetic fate-mapping captures the rise in label, experimentally it does not capture the subsequent decay that exists in IdU pulse-chase time-series. The direct observation of the decay thanks to the long-term monitoring of the IdU signal after a short pulse, better constrains the model. In Ref (13) it is precisely the arrest of the experiment before the decay can be measured in the Ly6C^{low}population that rendered those data less attractive to modeling; (ii) In Refs. (8, 23), homeostasis (steady-state) is assumed and is utilized to resolve the two rates k, d whose sum is the only experimentally-accessible observable (steady-state distribution of cells determines the ratio k/d). Conversely, our IdU pulse-chasing resolves both incorporation and decay, k and d are independent rates that are fitted independently, and the resulting values can be checked for compatibility with the homeostasis hypothesis. Hence, IdU pulse-chasing is advantageous as it affords stronger experimentally-constrained fit of our model.

[Figure 1 about here.]

Data and code availability

All data used in this manuscript, including an MIFlowCyt checklist (24) is available in https://flowrepository.org/experiments/2015. All code used to process the data and generate the figures, including the fitting of the data to our model, is freely available in https://github.com/AmirErez/IdUPulseChase.

Results

Clustering of high-dimensional data allows for parallel visualization of differentiation dynamics of multiple immune cell types

In order to demonstrate the capability of mass cytometry to simultaneously track the differentiation dynamics of multiple immune cell types, we first projected the high-dimensional data into two-dimensional space using a custom made clustering pipeline. Fig. 2A shows a t-SNE map of 2-D projection of the obtained clusters based on the surface expression of individual markers. The clustering heatmap in figure 2B shows differential expression of cell surface receptors on the clusters that is used to identify classes of immune cells. We first tested whether IdU injection would induce inflammation and/or perturb hematopoietic homeostasis in mice. We applied HAL-x to measure the frequency of leukocyte subpopulation as a function of time post–IdU injection. We found that representative trajectories of cluster frequencies plotted over time demonstrate a high degree of stability, thus showing that the IdU injection did not induce any fluctuations in cluster frequencies over time (figure 2C). Finally, we plot the full time course of IdU positive cells for each of the identified clusters (figure 2D), demonstrating the ability of our

method to observe differentiation dynamics of immune cells in the mouse bone marrow on a global scale.

[Figure 2 about here.]

Phenotypic Stability for leukocyte subpopulations

To further confirm a lack of any notable change in marker expression profiles across the studied populations of interest as a result of IdU injection, we also examined the geometric-mean dual count of MHC-II levels for each mouse, as a function of time post ^{127}IdU injection (Fig. 3). We found that, for 0.4mg of IdU as validated in our protocol, MHC-II expression does not vary with time in all leukocyte subpopulation. In particular, hand-gated neutrophils maintain lower MHC-II expression in comparison with the B-cells, as expected (Fig. 3). This strongly suggests that IdU injection in mice does not elicit any substantial inflammatory response. Additionally, for each population of interest, the markers used for studying differentiation dynamics were also found to be stable throughout the entire time series (Fig. 3).

Taken together, these results support the temporal stability in leukocyte phenotype *i.e.* minimal inflammation throughout the IdU pulse-chase experiment. Thus, we demonstrate that the use of IdU pulse-chase allows to monitor the dynamics of leukocyte differentiation while maintaining system homeostasis.

[Figure 3 about here.]

Fitting the time-series on select groups of immune cells with known paths of differentiation

We then moved on to our primary objective, which was to illustrate the strength of our modeling methodology. To this end, we focused on

individual tracks of differentiation within well-characterized cell lineages (e.g. Neutrophils, Immature B cells, and mature B cells), as defined by manual gating (see Supp Figures). We further analyzed each cell lineage by reducing its differentiation path within, to a one-dimensional path as defined by the first *principal component* (PC1) of the principle-component-analysis for the density distribution in a 2D marker space (see Figs. 4B, 5B and 6B with the black asterisks). We then decomposed the dynamics of differentiation into individual steps between nearest-neighbor voxels in the marker space, such that we could define, for each individual subpopulation a *progenitor* population, with the *progeny* as the voxel next to it along the phenotypic continuum. We applied our modeling approach for the IdU dynamics within each subpopulation, using the pairwise fitting as applied to each subpopulation along PC1. In Figs. 4C, 5C and 6C, we compare the data (points) to the best fit according to the model (lines). The color code reflects the location along PC1 which each curve captures. Overall, our model captures well the dynamics of differentiation in the data, directly modeling and quantifying the dynamics of maturation within each cell lineage.

(i) Neutrophils

We begin by observing the maturation dynamics of neutrophils. As demonstrated in Fig. 4A, the IdU peak time increases as we progress diagonally across from Ly-6G^{lo}/CD11b^{lo} to Ly-6G^{hi}/CD11b^{hi} populations, suggesting an increase in both Ly-6G and CD11b surface expression as neutrophils mature, as reported previously (25). Moreover, proliferating neutrophils were found to express lower Ly-6G than non–proliferating ones, as indicated by the noticeably decreased Ly-6G expression amongst IdU time points (Fig. 4A). This observation is also in accordance with a recently published study (18), where the authors reported decreased Ly-6G

surface expression on proliferating neutrophils. Modeling the time-series by following the trajectory along the first principal component axis of the density distribution reveals two major peaks of differentiating populations. As can be seen in Fig. 4B, Ly-6G^{lo}/CD11b^{lo} neutrophils peak at around 40 to 50 hours followed by a gradual decay by 100 hours, by which time the Ly-6G^{hi}/CD11b^{hi} populations peak followed by a sharper decline.

[Figure 4 about here.]

(ii) B cells

The dynamics of B cell development in mice is a well-studied question (26, 27, 28). We took advantage of this to further validate the efficiency of our approach and to recapitulate known biology. CD43 is a marker that is expressed on all B cell precursors - Pre-Pro-B, Pro-B, and Pre-B cells (29, 30, 31). As Pre-B cells differentiate into immature B cells they gradually gain surface IgM (31). Fig. 5 shows the transition dynamics of Pre-B cells to immature B cells, which displays two distinct differentiating cell types - an early CD43^{pos}/IgM^{neg} Pre-B cell population followed later by IgM^{pos} Immature B cells Fig. 5C. Interestingly, the two differentiation peaks are separated by only a few hours, suggesting a rapid transition from precursors to immature B cells.

We then sought to model the transition of immature B cells to mature B cells Fig. 6. Immature B cells are high on CD93 and sIgM while low on CD21 (26). As they mature, they progressively downregulate expression of CD93 while upregulating expression of CD21 (26). Here again, such an elementary step in differentiation can be delineated using a one-dimensional path in the (CD21, CD93) plane along which we can monitor the dynamics of IdU pulse-chasing. As depicted in Fig. 6C, the immature B cells appear to peak in IdU uptake around 50 hours post-IdU

pulsing, followed by a decline. As observed with the pre-B to immature B cell transition, the mature B cell differentiation peak almost overlapped with the immature B cells, indicating a very rapid transition.

[Figure 5 about here.]

[Figure 6 about here.]

From our fits, we observe that it is by no means obvious that Eq. 1 would indeed fit the time-series data so well as is apparent. From that, we can conclude that the differentiation process convincingly follows along PC1 and that it does not involve significant nonlinear effects, e.g., secretion of iodine by the cells. This serves as a powerful proof of concept for the IdU pulse-chase framework we propose here. Moreover, we are now in the position to compare the different k, d fitted rates. In Fig. 7A we plot k, d for the three compartments (circles) together with the $k = d$ line (dashed). It is apparent that $k \approx d$ with $d \geq k$. The latter is not surprising: by definition our model allows k from only the progenitor ("Source") to go to the progeny ("Destination"), whereas d is not limited in such a way: cells can die or otherwise exit outside PC1, resulting in $d \geq k$. Indeed, it is precisely this flexibility that differs from a previous modeling approach (8). Of further note is that the rates appear comparable across the three transitions, suggesting that it is not a particular compartment's transition that is fast or slow, but rather, that within a compartment, different parts of a transition are faster or slower.

The modeling strategy presented here can also be used to glean information about the cellular fluxes. The rates k, d shown in Fig. 7 are agnostic to number of cells in progenitor and progeny. To account for the cells, we multiply k by the fraction of cells that belong to each progenitor and d by the progeny's fraction. These are normalized so that the total cell

count in focus along PC1 has a total fraction of 1. We plot the fluxes in Fig. 7B, with the lines tracking the fluxes along PC1. It is immediately apparent that at the edges of the compartment, the cellular flux is low, consistent with the lower cell density at those regions. All three transitions show a pattern of maximal flux around the center of the transition, with the B cells replacing a larger overall percentage per hour than the neutrophils.

Together, the above results clearly demonstrate that our method, which combines the power of mass cytometry with mathematical modeling can faithfully recapitulate previously reported facets of hematopoiesis in the mouse bone marrow and add to our understanding of the dynamics of hematopoiesis.

[Figure 7 about here.]

Discussion

Discrete gating by Boolean logic is standard practice for identifying cell populations of interest among heterogeneous cell samples by FCM. However, expression of a certain marker on a given cell type is highly variable even within a given population positive for that marker. To better capture this cell-to-cell heterogeneity we parsed cell populations into several subpopulation (binned voxels) based on expression of markers on a bivariate plot, thereby allowing us to model relative changes in marker expressions as the cells transition from one marker state to another. More recent advances have introduced dimensional reduction to delineate simple differentiation paths (Diffusion Maps (32), Isomap (33), Wishbone (7)). Both approaches (manual gating and binning, vs high-dimensional reduction) organize cells along a low-dimensional path of differentiation based on phenotypic similarities, thus defining a pseudo-time of

differentiation.

We introduced here a methodology combining IdU pulse-chasing, CyTOF monitoring and mathematical modeling to experimentally monitor the dynamics of differentiation. Using this approach, we quantitatively monitored the dynamics of three major cell type transitions during hematopoiesis in the bone marrow of adult mice. This enabled us to measure the real dynamics of differentiation rather than a pseudo-time of purely-phenotypic approaches. We confirmed the simultaneous increase of Ly-6G and CD11b expression in the neutrophil compartment as the cells mature. Our observations are in agreement with a previously published report (34). In addition, we observed two temporally well separated peaks of differentiating populations in the neutrophil compartment with the peak differentiation of fully mature neutrophils happening at around day 4.

In contrast, our analysis suggested that the maturation of B cells from their precursors happens on a much shorter time scale with very little time lag between development of early and late stages. Our methodology will be of particular interest when studying how such differences in maturation dynamics between immune cell types shape an immune response, e.g. during an inflammatory response.

Mass cytometry is fast gaining popularity for immune cell phenotype analysis because of the number of simultaneously measurable markers it allows without running into the technical difficulty introduced by fluorescence spillover between channels in classical fluorescence-based cytometry. In the present study, we took advantage of our monitoring of *in vivo* IdU uptake, and our direct detection of iodine by mass cytometry, to study hematopoietic dynamics in mouse bone marrow at a more global scale. This enjoys the advantage of easy preparation and minimal perturbation of the sample by reagents required to access the nucleus in its FCM counterparts, EdU and BrdU. Although studying progression of

immune cells through stages of cell cycle by IdU incorporation followed by analysis by mass cytometry has been introduced before for *in vitro* setting (17), its use to study differentiation *in vivo* has been limited. In a recent study, Evrard *et al.* used *in vivo* administration of IdU to phenotypically and functionally distinguish proliferating and non-proliferating populations of immune cells (18). Here, we built on these two papers to perform IdU pulse-chase in live animals over multiple time points and to address the dynamics of differentiation.

To analyze multiple time-points, we introduced a simple mathematical modeling scheme. Our model benefits from the fact that tracking a tracer molecule (here IdU) is not a biological phenomenon (with all implied complexity), but rather, is a simple physical phenomenon readily captured by a set of linear pairwise transfer relations. And so, it is the combination of, which provides the strength and flexibility for monitoring hematopoietic differentiation: the IdU allows CyTOF acquisition with its technical advantages; The slicing objectively treats surface protein expression heterogeneity without imposing on it a binary classification; The time-series and its model-based analysis provide the direct quantification of the differentiation dynamics. The simplicity of each of these three aspects is what, when taken together, renders them a general and effective technique. To conclude, we introduced a quantitative framework for analyzing the dynamics of cell maturation that can be leveraged to study differentiation of multiple immune cell types in parallel. It will be interesting to see whether such a broad and detailed mapping of hematopoiesis at homeostasis or under inflammation would reveal unexpected systemic differences in mouse strains and/or well-used transgenic models.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

[1] van Furth R and Cohn ZA. The origin and kinetics of mononuclear phagocytes. J Exp Med 1968;128:415.

[2] Fontenot JD, Gavin MA, and Rudensky AY. Foxp3 programs the development and function of cd4+cd25+ regulatory t cells. Nat Immunol 2003;4:330–336.

[3] Yu W, Jiang N, Ebert PR, Kidd B, Müller S, Lund P, Juang J, Adachi K, Tse T, Birnbaum M, Newell E, Wilson D, Grotenbreg G, Valitutti S, Quake S, and Davis M. Clonal deletion prunes but does not eliminate self-specific cd8+ t lymphocytes. Immunity 2016; 42:929–941.

[4] Inokuma MS, Maino VC, and Bagwell CB. Probability state modeling of memory cd8+ t-cell differentiation. Journal of immunological methods 2013;397:8–17.

[5] Bendall S, Davis K, Amir Ea, Tadmor M, Simonds E, Chen T, Shenfeld D, Nolan G, and Pe?er D. Single-cell trajectory detection uncovers progression and regulatory coordination in human b cell development. Cell 2014;157:714–725.

[6] Bagwell CB, Hunsberger BC, Herbert DJ, Munson ME, Hill BL, Bray CM, and Preffer FI. Probability state modeling theory. Cytometry Part A 2015;87:646–660.

[7] Setty M, Tadmor MD, Reich-Zeliger S, Angel O, Salame TM, Kathail P, Choi K, Bendall S, Friedman N, and Pe'er D. Wishbone identifies

bifurcating developmental trajectories from single-cell data. Nature Biotechnology 2016;34:637.

[8] Busch K, Klapproth K, Barile M, Flossdorf M, Holland-Letz T, Schlenner SM, Reth M, Hofer T, and Rodewald HR. Fundamental properties of unperturbed haematopoiesis from stem cells in vivo. Nature 2015;518:542–546.

[9] Höfer T, Barile M, and Flossdorf M. Stem-cell dynamics and lineage topology from in vivo fate mapping in the hematopoietic system. Current Opinion in Biotechnology 2016;39:150–156.

[10] La Manno G, Soldatov R, Zeisel A, Braun E, Hochgerner H, Petukhov V, Lidschreiber K, Kastriti ME, Lönnerberg P, Furlan A, Fan J, Borm LE, Liu Z, van Bruggen D, Guo J, He X, Barker R, Sundström E, Castelo-Branco G, Cramer P, Adameyko I, Linnarsson S, and Kharchenko PV. Rna velocity of single cells. Nature 2018;560:494–498.

[11] Oyler-Yaniv J, Oyler-Yaniv A, Shakiba M, Min NK, Chen YH, Cheng SY, Krichevsky O, Altan-Bonnet N, and Altan-Bonnet G. Catch and release of cytokines mediated by tumor phosphatidylserine converts transient exposure into long-lived inflammation. Molecular cell 2017; 66:635–647.e7.

[12] Boer D, Rob J, and Perelson AS. Quantifying t lymphocyte turnover. Journal of Theoretical Biology 2013;327:45–87.

[13] Yona S, Kim KW, Wolf Y, Mildner A, Varol D, Breker M, Strauss-Ayali D, Viukov S, Guilliams M, Misharin A, Hume D, Perlman H, Malissen B, Zelzer E, and Jung S. Fate mapping reveals origins and dynamics of monocytes and tissue macrophages under homeostasis. Immunity 2013;38:79 – 91.

[14] Van de Casteele M, Leuckx G, Baeyens L, Cai Y, Yuchi Y, Coppens V, De Groef S, Eriksson M, Svensson C, Ahlgren U, Ahnfelt-Ronne J, Madsen OD, Waisman A, Dor Y, Jensen JN, and Heimberg H. Neurogenin 3+ cells contribute to beta-cell neogenesis and proliferation in injured adult mouse pancreas. Cell Death &Amp; Disease 2013;4:e523.

[15] Hija A, Salpeter S, Klochendler A, Grimsby J, Brandeis M, Glaser B, and Dor Y. G0-g1 transition and the restriction point in pancreatic beta-cells in vivo. Diabetes 2014;63:578–584.

[16] Wang F, Zhan R, Chen L, Dai X, Wang W, Guo R, Li X, Li Z, Wang L, Huang S, Shen J, Li S, and Cao C. Rhoa promotes epidermal stem cell proliferation via pkn1-cyclin d1 signaling. PLOS ONE 2017; 12:1–13.

[17] Behbehani GK, Bendall SC, Clutter MR, Fantl WJ, and Nolan GP. Single-cell mass cytometry adapted to measurements of the cell cycle. Cytometry Part A 2012;81A:552–566.

[18] Evrard M, Kwok IW, Chong SZ, Teng KW, Becht E, Chen J, Sieow JL, Penny HL, Ching GC, Devi S, et al. Developmental analysis of bone marrow neutrophils reveals populations specialized in expansion, trafficking, and effector functions. Immunity 2018;48:364–379.

[19] Erez A, Vogel R, Mugler A, Belmonte A, and Altan-Bonnet G. Modeling of cytometry data in logarithmic space: When is a bimodal distribution not bimodal? Cytometry Part A ;93:611–619.

[20] Feinerman O, Veiga J, Dorfman JR, Germain RN, and Altan-Bonnet G. Variability and robustness in t cell activation from regulated heterogeneity in protein levels. Science 2008;321:1081–1084.

[21] Cotari JW, Voisinne G, Dar OE, Karabacak V, and Altan-Bonnet G. Cell-to-cell variability analysis dissects the plasticity of signaling of common gamma chain cytokines in t cells. Sci Signal 2013;6:ra17–.

[22] Box GEP, Jenkins GM, and Reinsel GC. *Time Series Analysis: Forecasting and Control*. Hoboken, N.J: Wiley, 4 edition edition, 2008.

[23] Höfer T, Busch K, Klapproth K, and Rodewald HR. Fate mapping and quantitation of hematopoiesis in vivo. Annual Review of Immunology 2016;34:449–478.

[24] Lee JA, Spidlen J, Boyce K, Cai J, Crosbie N, Dalphin M, Furlong J, Gasparetto M, Goldberg M, Goralczyk EM, Hyun B, Jansen K, Kollmann T, Kong M, Leif R, McWeeney S, Moloshok TD, Moore W, Nolan G, Nolan J, Nikolich-Zugich J, Parrish D, Purcell B, Qian Y, Selvaraj B, Smith C, Tchuvatkina O, Wertheimer A, Wilkinson P, Wilson C, Wood J, Zigon R, Scheuermann RH, and Brinkman RR. Miflowcyt: The minimum information about a flow cytometry experiment. Cytometry Part A 2008;73A:926–930.

[25] Kim MH, Yang D, Kim M, Kim SY, Kim D, and Kang SJ. A late-lineage murine neutrophil precursor population exhibits dynamic changes during demand-adapted granulopoiesis. Scientific reports 2017;7:39804.

[26] Allman D and Pillai S. Peripheral b cell subsets. Current opinion in immunology 2008;20:149–157.

[27] Pieper K, Grimbacher B, and Eibel H. B-cell biology and development. Journal of Allergy and Clinical Immunology 2013; 131:959–971.

[28] Melchers F. Checkpoints that control b cell development. The Journal of clinical investigation 2015;125:2203–2210.

[29] Hardy RR, Carmack CE, Shinton SA, Kemp JD, and Hayakawa K. Resolution and characterization of pro-b and pre-pro-b cell stages in normal mouse bone marrow. Journal of Experimental Medicine 1991; 173:1213–1225.

[30] Wells SM, Kantor AB, and Stall AM. Cd43 (s7) expression identifies peripheral b cell subsets. The Journal of Immunology 1994; 153:5503–5515.

[31] Tung JW, Mrazek MD, Yang Y, Herzenberg LA, and Herzenberg LA. Phenotypically distinct b cell development pathways map to the three b cell lineages in the mouse. Proceedings of the National Academy of Sciences 2006;103:6293–6298.

[32] Haghverdi L, Buettner F, and Theis FJ. Diffusion maps for high-dimensional single-cell analysis of differentiation data. Bioinformatics 2015;31:2989–2998.

[33] Thomas GD, Hamers AA, Nakao C, Marcovecchio P, Taylor AM, McSkimming C, Nguyen AT, McNamara CA, and Hedrick CC. Human blood monocyte subsets. Arteriosclerosis, Thrombosis, and Vascular Biology 2017;37:1548–1558.

[34] Rožman S, Yousefi S, Oberson K, Kaufmann T, Benarafa C, and Simon HU. The generation of neutrophils in the bone marrow is controlled by autophagy. Cell death and differentiation 2015;22:445.

List of Figures

1 **Summary experimental and modeling method: (A)** Experimental setup: mice are injected *i.p.* with IdU, sacrificed at indicated time intervals for harvest of bone marrow, which is then stained with metal-conjugated antibodies followed by acquisition on a Helios mass cytometer. **(B)** For modeling differentiation along a phenotypic continuum, we zoom in on a transition of interest based on bivariate scatterplots and define a grid along its phenotype directions to create a map of voxels. **(C)** Schematic of pair-wise differentiation model and the transfer function equation which describes it. . 26

2 **Global analysis of IdU pulse-chase by CyTOF. (A)** t-SNE map of $CD45^+$ Live Single cells in the bone marrow. **(B)** Automatic binning of $CD45^+$ Live Single cells using HAL-x defines 50 subpopulations of cells whose immunophenotype can be characterized by the mean signal of surface epitopes. These subpopulations can be mapped to classical hematopoietic ontology (dendritic cells, granulocytes, etc.). **(C)** Stability of subpopulation frequencies over time after injection of IdU. **(D)** Time dynamics of IdU uptake (normalized to maximum) for each leukocyte subpopulation. The color coding represents the range of IdU signal (red: high; blue: low). . . . 27

3 **Phenotypic stability.** Top row: neutrophils; middle: pre-B to immature B; bottom: immature-B to mature-B. None of the selected populations display any noticeable structure in their gate frequency (left column) nor any noticeable structure in their marker's fluorescence intensity (two center columns) nor signs of inflammation in MHC-II expression (right column). This confirms that IdU injection does not perturb the bone marrow composition. 28

4 **Neutrophil maturation (A)** IdU Time-series for each voxel, colored according to its peak IdU time. A voxel's position corresponds to a range of Ly6G (vertical) and CD11b (horizontal) abundances which define the population. Each voxel shows an entire time-series (228 hours) of IdU mean-fluorescence intensity of the voxel's population. **(B)** Density heatmap showing relative abundance for each voxel. The black asterisks mark the high-density differentiation path along the first principle component (PC1) of the density distribution. **(C) Time-series fit along PC1:** we follow the time-series for each of the voxels marked with an asterisk, showing both data (dots) and transfer-function (Eq. 1) fit (lines). Colors correspond to progression along PC1. 29

5 **Pre-B To Immature B Cell transition. (A)** IdU Time-series for each voxel, colored according to its peak IdU time. A voxel's position corresponds to a range of IgM (vertical) and CD43 (horizontal) abundances which define the population. Each voxel shows an entire time-series (228 hours) of IdU mean-fluorescence intensity of the voxel's population. **(B)** Density heatmap showing relative abundance for each voxel with asterisks showing PC1. **(C) Time-series fit:** data (dots) and transfer-function (Eq. 1) fit (lines). Colors correspond to progression along PC1. 30

6 **Immature to Mature B Cell development. (A)** IdU Time-series for each voxel, colored according to its peak IdU time. A voxel's position corresponds to a range of CD93 (vertical) and CD21 (horizontal) abundances which define the population. Each voxel shows an entire time-series (228 hours) of IdU mean-fluorescence intensity of the voxel's population. **(B)** Density heatmap showing relative abundance for each voxel with asterisks showing PC1. **(C) Time-series fit:** data (dots) and transfer-function (Eq. 1) fit (lines). Colors correspond to progression along PC1. 31

7 **Comparison of differentiation rates. (A)** Bare rates k, d for the neutrophil maturation (red), pre-B to immature B (green), and immature B to mature (blue). Dashed black line corresponds to $k = d$. **(B)** Differentiation fluxes along PC1. % Flux in (out) is calculated as k (d) multiplied by the cell count fraction along the progenitor (progeny) of the respective voxels. 32

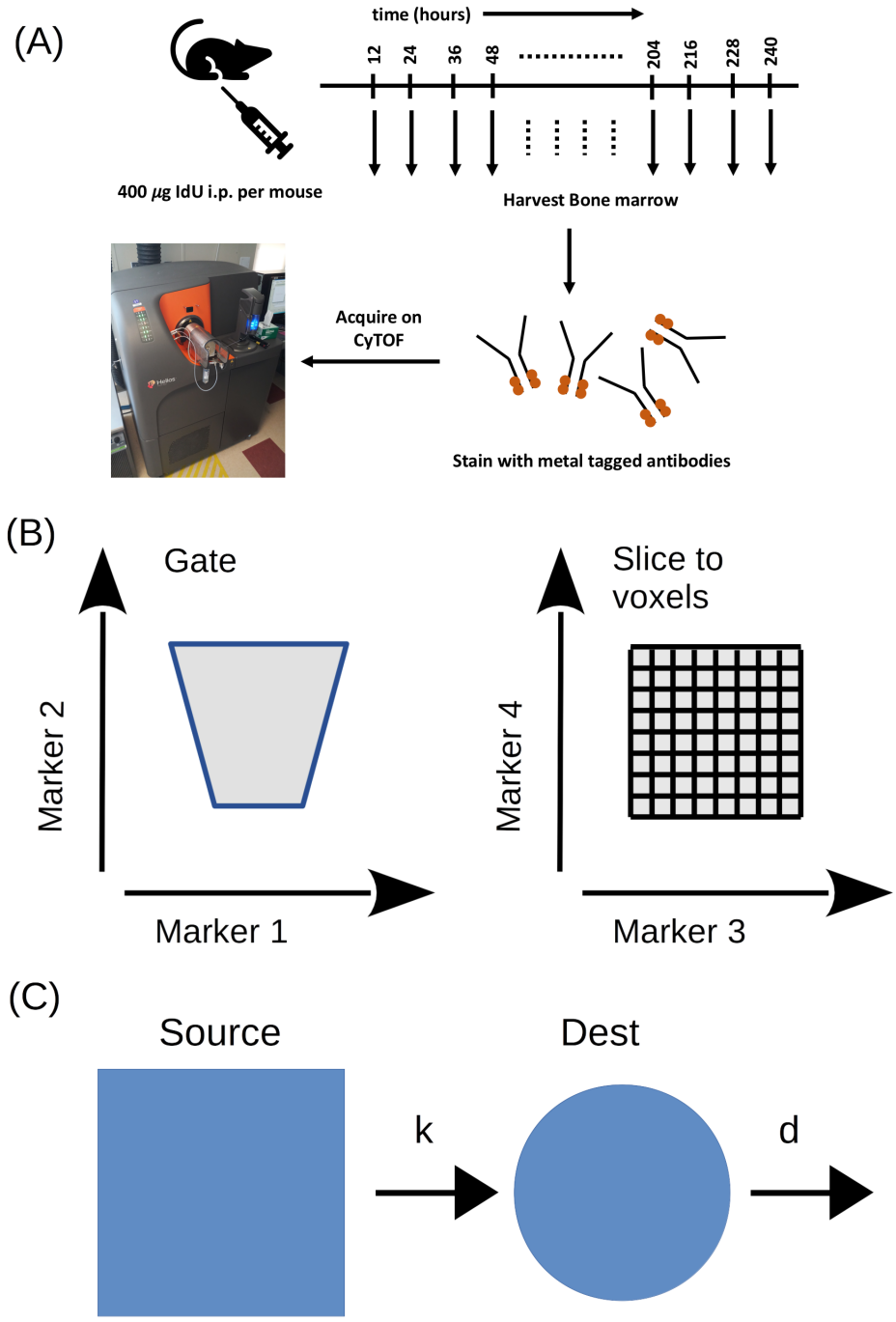

Figure 1. Summary experimental and modeling method: (A) Experimental setup: mice are injected *i.p.* with IdU, sacrificed at indicated time intervals for harvest of bone marrow, which is then stained with metal-conjugated antibodies followed by acquisition on a Helios mass cytometer. **(B)** For modeling differentiation along a phenotypic continuum, we zoom in on a transition of interest based on bivariate scatterplots and define a grid along its phenotype directions to create a map of voxels. **(C)** Schematic of pair-wise differentiation model and the transfer function equation which describes it.

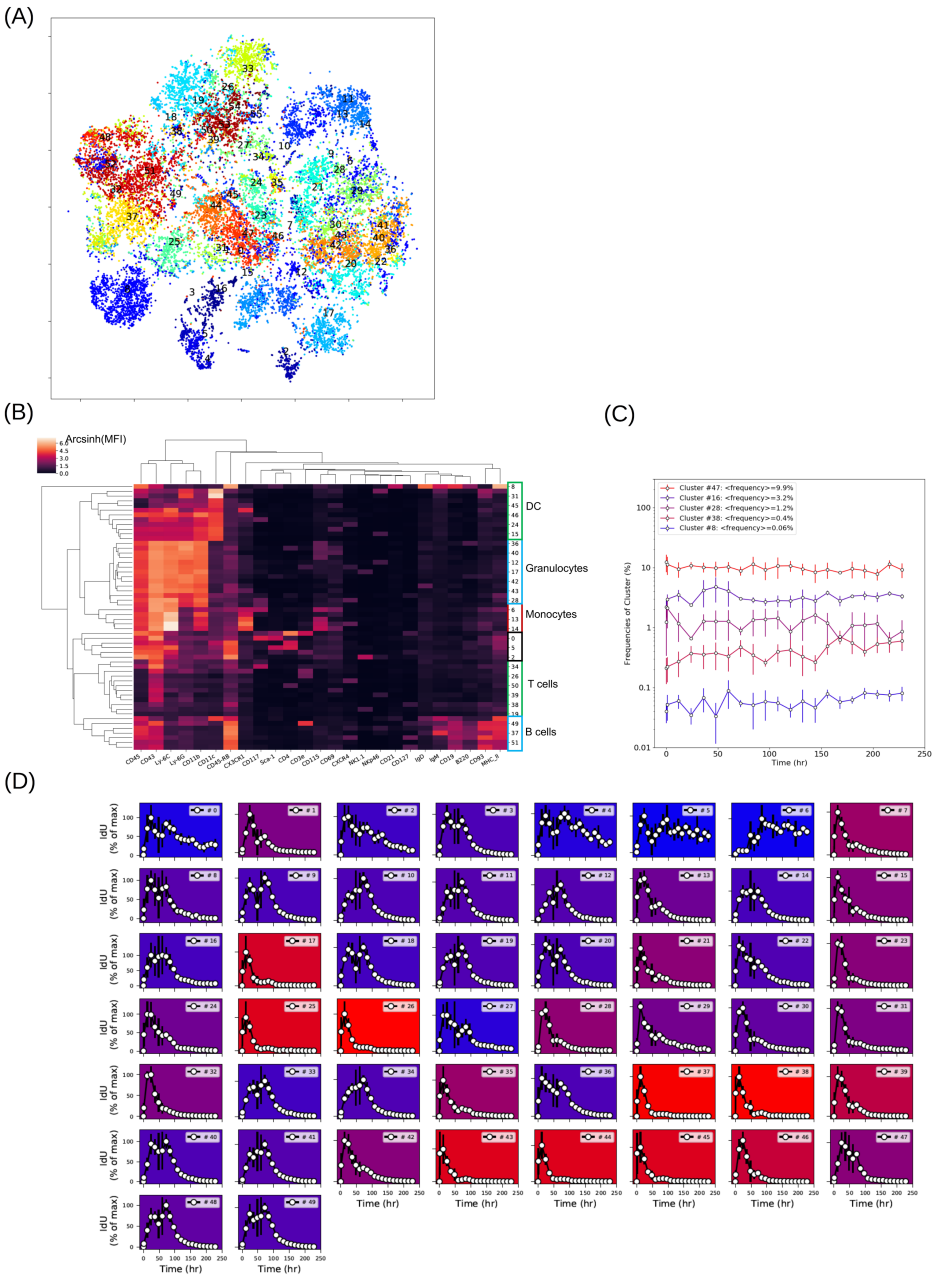

Figure 2. Global analysis of IdU pulse-chase by CyTOF. (A) t-SNE map of CD45$^+$ Live Single cells in the bone marrow. **(B)** Automatic binning of CD45$^+$ Live Single cells using HAL-x defines 50 subpopulations of cells whose immunophenotype can be characterized by the mean signal of surface epitopes. These subpopulations can be mapped to classical hematopoietic ontology (dendritic cells, granulocytes, etc.). **(C)** Stability of subpopulation frequencies over time after injection of IdU. **(D)** Time dynamics of IdU uptake (normalized to maximum) for each leukocyte subpopulation. The color coding represents the range of IdU signal (red: high; blue: low).

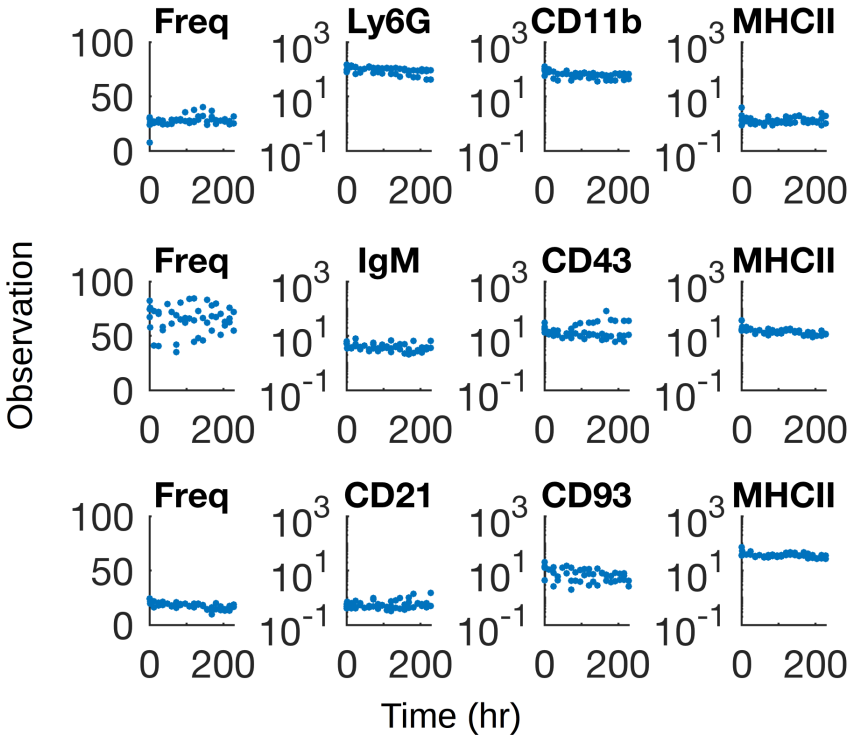

Figure 3. Phenotypic stability. Top row: neutrophils; middle: pre-B to immature B; bottom: immature-B to mature-B. None of the selected populations display any noticeable structure in their gate frequency (left column) nor any noticeable structure in their marker's fluorescence intensity (two center columns) nor signs of inflammation in MHC-II expression (right column). This confirms that IdU injection does not perturb the bone marrow composition.

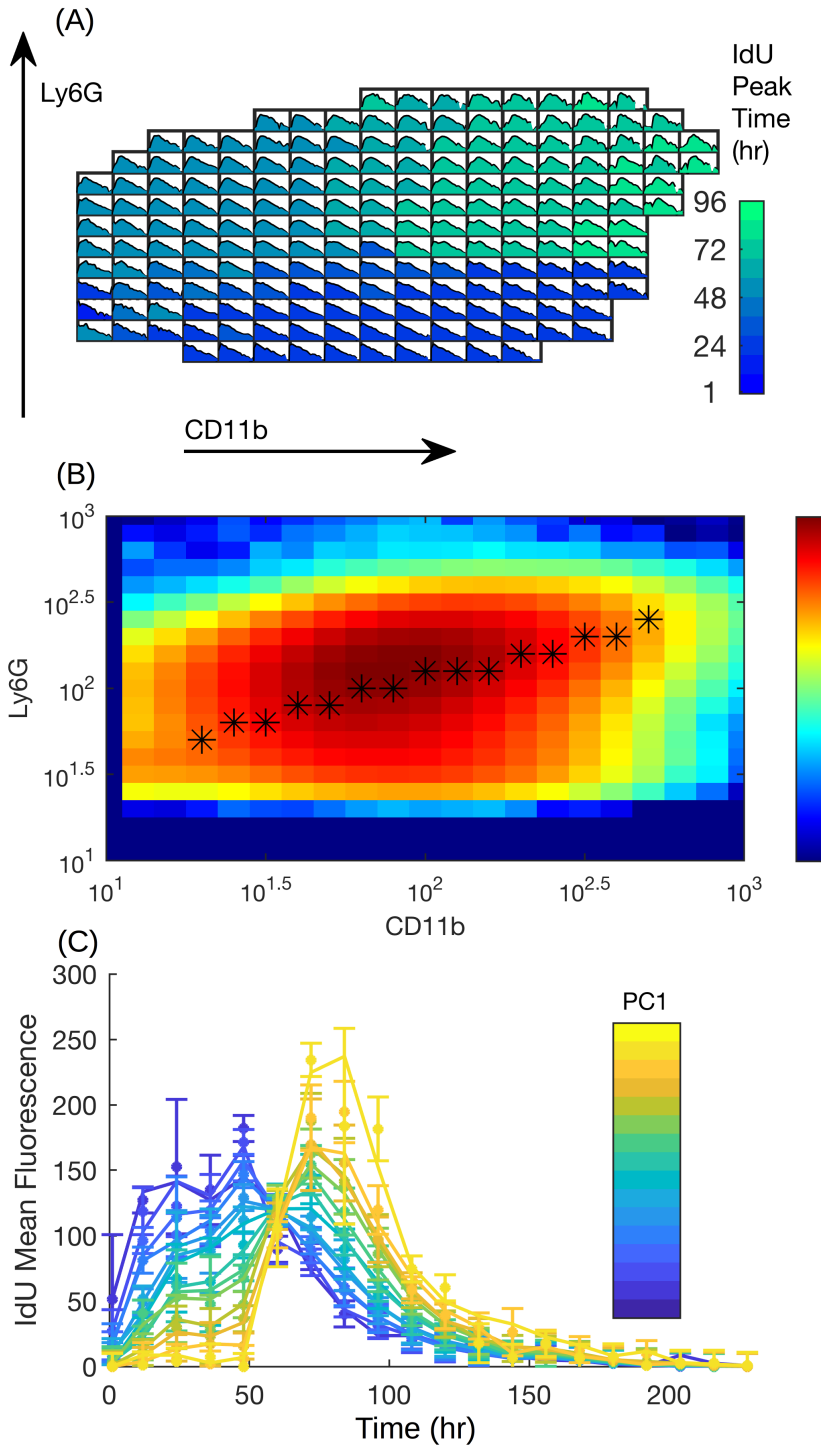

Figure 4. Neutrophil maturation (A) IdU Time-series for each voxel, colored according to its peak IdU time. A voxel's position corresponds to a range of Ly6G (vertical) and CD11b (horizontal) abundances which define the population. Each voxel shows an entire time-series (228 hours) of IdU mean-fluorescence intensity of the voxel's population. **(B)** Density heatmap showing relative abundance for each voxel. The black asterisks mark the high-density differentiation path along the first principle component (PC1) of the density distribution. **(C) Time-series fit along PC1:** we follow the time-series for each of the voxels marked with an asterisk, showing both data (dots) and transfer-function (Eq. 1) fit (lines). Colors correspond to progression along PC1.

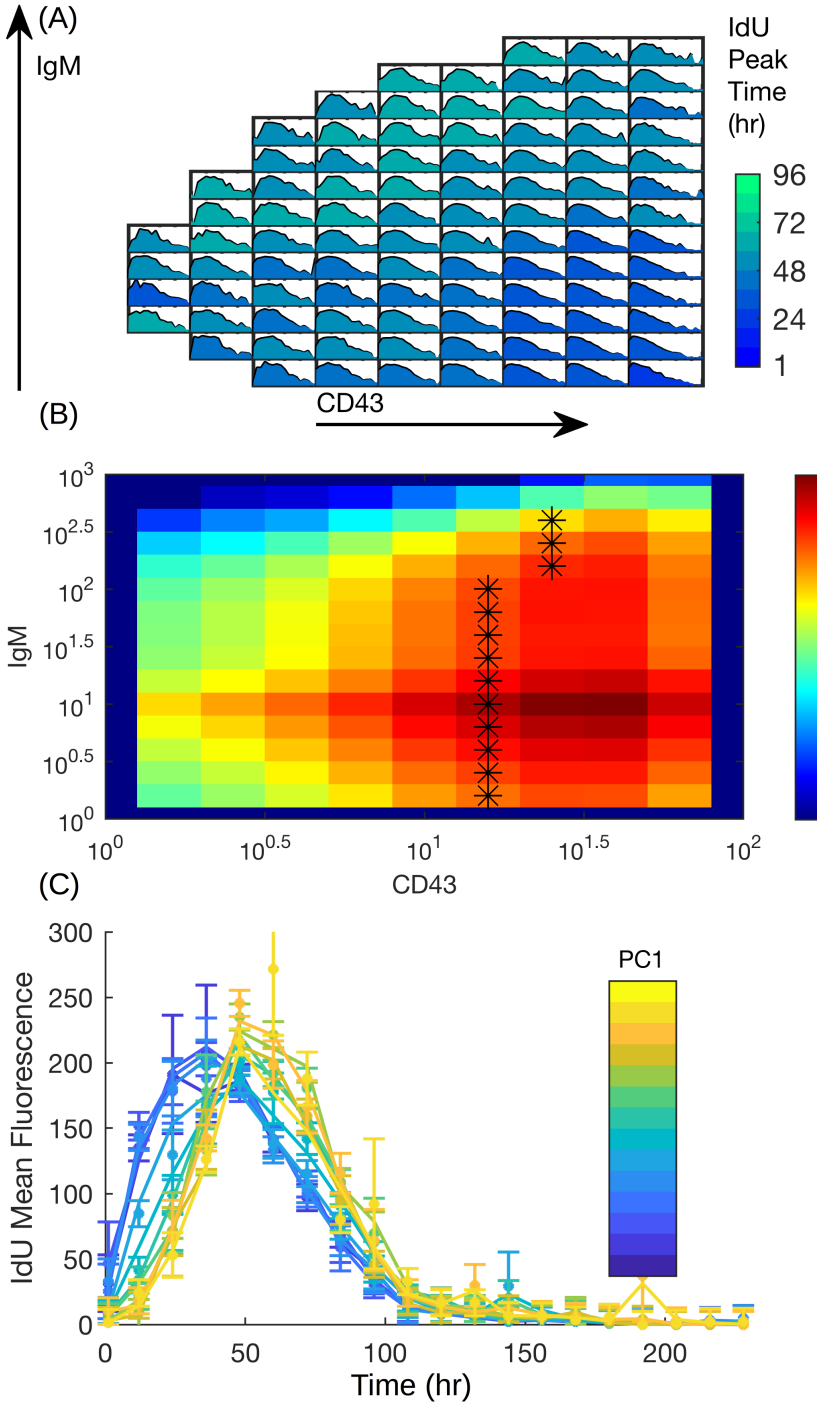

Figure 5. Pre-B To Immature B Cell transition. (A) IdU Time-series for each voxel, colored according to its peak IdU time. A voxel's position corresponds to a range of IgM (vertical) and CD43 (horizontal) abundances which define the population. Each voxel shows an entire time-series (228 hours) of IdU mean-fluorescence intensity of the voxel's population. **(B)** Density heatmap showing relative abundance for each voxel with asterisks showing PC1. **(C) Time-series fit:** data (dots) and transfer-function (Eq. 1) fit (lines). Colors correspond to progression along PC1.

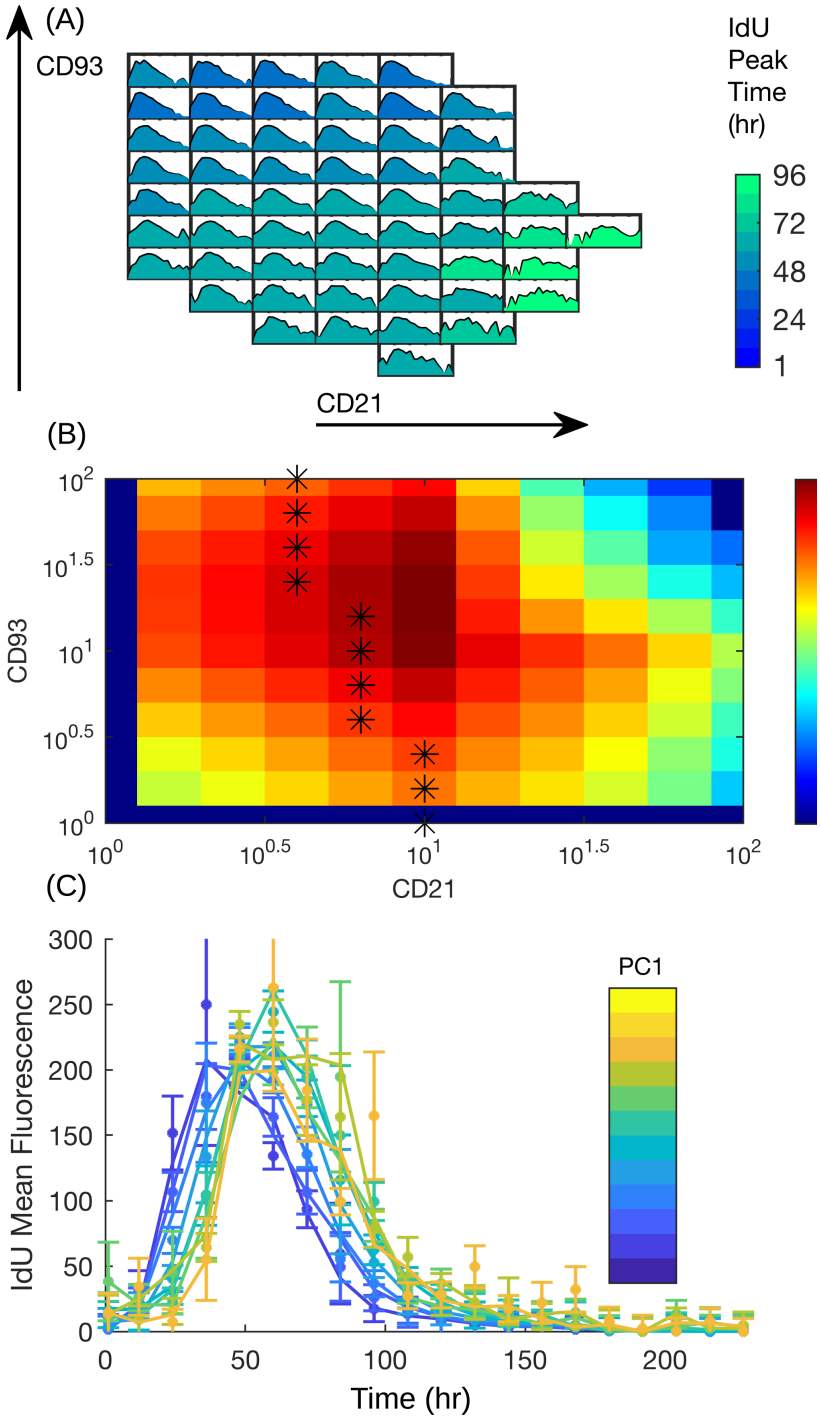

Figure 6. Immature to Mature B Cell development. (A) IdU Time-series for each voxel, colored according to its peak IdU time. A voxel's position corresponds to a range of CD93 (vertical) and CD21 (horizontal) abundances which define the population. Each voxel shows an entire time-series (228 hours) of IdU mean-fluorescence intensity of the voxel's population. **(B)** Density heatmap showing relative abundance for each voxel with asterisks showing PC1. **(C) Time-series fit:** data (dots) and transfer-function (Eq. 1) fit (lines). Colors correspond to progression along PC1.

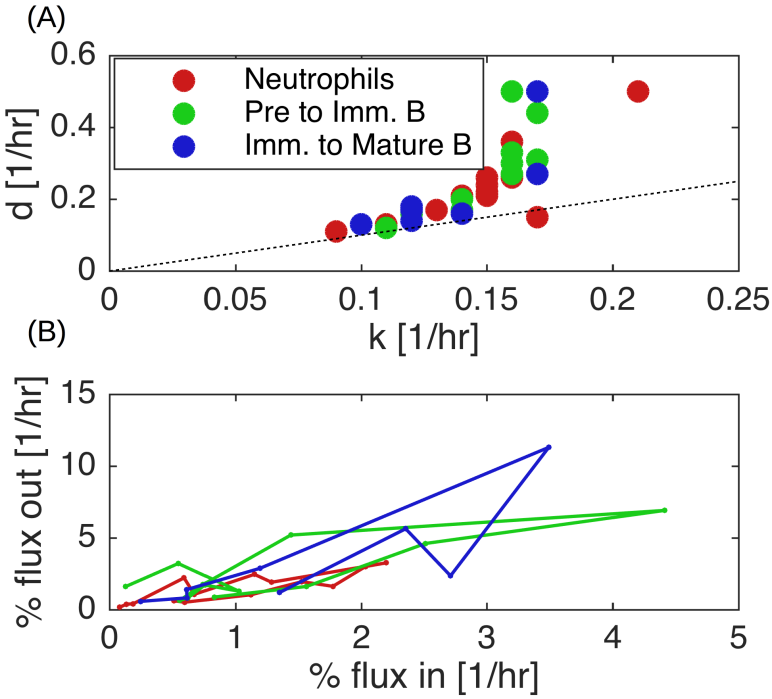

Figure 7. Comparison of differentiation rates. (A) Bare rates k, d for the neutrophil maturation (red), pre-B to immature B (green), and immature B to mature (blue). Dashed black line corresponds to $k = d$. **(B)** Differentiation fluxes along PC1. % Flux in (out) is calculated as k (d) multiplied by the cell count fraction along the progenitor (progeny) of the respective voxels.