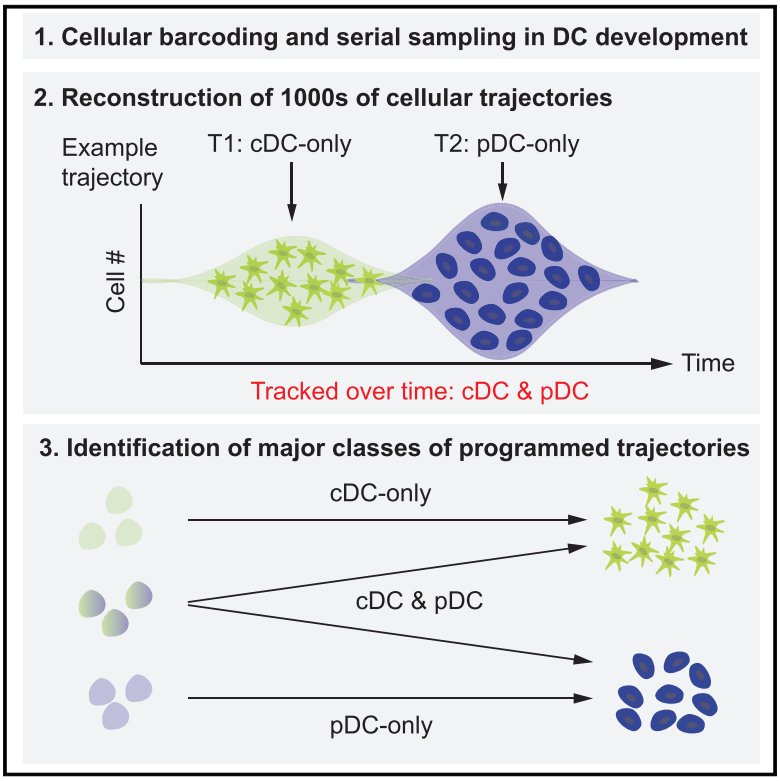
Cell Reports

DiSNE Movie Visualization and Assessment of Clonal Kinetics Reveal Multiple Trajectories of Dendritic Cell Development

Graphical Abstract



Highlights

- A cellular trajectory is defined by a clone's fate and cell production over time
- DiSNE movies allow dynamic visualization of clonal cellular trajectories
- Single HSPCs follow different cellular trajectories in DC development
- Cellular trajectories of individual clones are intrinsically programmed

Authors

Dawn S. Lin, Andrey Kan, Jerry Gao, Edmund J. Crampin, Philip D. Hodgkin, Shalin H. Naik

Correspondence

naik.s@wehi.edu.au

In Brief

Lin et al. develop a framework to longitudinally track and visualize clonal cellular trajectories during dendritic cell development. The authors demonstrate that properties including fate bias and the timing and size of clonal bursts are heterogeneous within populations yet largely imprinted in single progenitors at early developmental stages.

Lin et al., 2018, Cell Reports *22*, 2557–2566
March 6, 2018
https://doi.org/10.1016/j.celrep.2018.02.046

CellPress

Cell Reports
Report

DiSNE Movie Visualization and Assessment of Clonal Kinetics Reveal Multiple Trajectories of Dendritic Cell Development

Dawn S. Lin,[1,2,3,7] Andrey Kan,[2,3,7] Jerry Gao,[1] Edmund J. Crampin,[4,5,6] Philip D. Hodgkin,[2,3] and Shalin H. Naik[1,2,3,8,*]
[1]Molecular Medicine Division, Walter and Eliza Hall Institute, Parkville, VIC 3052, Australia
[2]Immunology Division, Walter and Eliza Hall Institute, Parkville, VIC 3052, Australia
[3]Faculty of Medicine, Dentistry & Health Sciences, University of Melbourne, Parkville, VIC 3010, Australia
[4]Systems Biology Laboratory, University of Melbourne, Parkville, VIC 3010, Australia
[5]Centre for Systems Genomics, University of Melbourne, Parkville, VIC 3010, Australia
[6]ARC Centre of Excellence in Convergent Bio-Nano Science and Technology, Melbourne School of Engineering, University of Melbourne, Parkville, VIC 3010, Australia
[7]These authors contributed equally
[8]Lead Contact
*Correspondence: naik.s@wehi.edu.au
https://doi.org/10.1016/j.celrep.2018.02.046

SUMMARY

A thorough understanding of cellular development is incumbent on assessing the complexities of fate and kinetics of individual clones within a population. Here, we develop a system for robust periodical assessment of lineage outputs of thousands of transient clones and establishment of bona fide cellular trajectories. We appraise the development of dendritic cells (DCs) in *fms*-like tyrosine kinase 3 ligand culture from barcode-labeled hematopoietic stem and progenitor cells (HSPCs) by serially measuring barcode signatures and visualize these multidimensional data using developmental interpolated t-distributed stochastic neighborhood embedding (DiSNE) time-lapse movies. We identify multiple cellular trajectories of DC development that are characterized by distinct fate bias and expansion kinetics and determine that these are intrinsically programmed. We demonstrate that conventional DC and plasmacytoid DC trajectories are largely separated already at the HSPC stage. This framework allows systematic evaluation of clonal dynamics and can be applied to other steady-state or perturbed developmental systems.

INTRODUCTION

Dendritic cells (DCs) represent a distinct branch of hematopoiesis and are responsible for pathogen sensing and activation of the adaptive immune response (Merad et al., 2013). There are three major subtypes, including plasmacytoid DCs (pDCs), type 1 conventional DCs (cDC1s), and type 2 cDCs (cDC2s) (Guilliams et al., 2014). DC development is relatively well established at the population level and can be recapitulated in *fms*-like tyrosine kinase 3 ligand (FL) cultures (Naik et al., 2005). According to the current hierarchical model of hematopoiesis (Guo et al., 2013; Månsson et al., 2007), all DC subtypes can be generated from a restricted common DC progenitor (CDP) population downstream of hematopoietic stem and progenitor cells (HSPCs) (Naik et al., 2007, 2013; Onai et al., 2007) via discrete subtype-committed precursor stages (Grajales-Reyes et al., 2015; Naik et al., 2006; Onai et al., 2013; Schlitzer et al., 2015). However, clonal evidence has suggested earlier lineage imprinting (Ema et al., 2014; Lee et al., 2017; Naik et al., 2013; Notta et al., 2016; Sanjuan-Pla et al., 2013; Yamamoto et al., 2013) or even DC subtype imprinting (Helft et al., 2017; Lee et al., 2017; Naik et al., 2007; Onai et al., 2007) within individual HSPCs. Importantly, most lineage tracing studies measured clonal fate at only a single time point. Therefore, questions remain as to whether the fate bias observed at one snapshot in time is consistent over time.

Some studies have assessed clonal contribution longitudinally (e.g., by serially sampling progeny derived from HSPCs in the blood) and have been instrumental in highlighting clonal properties, including repopulation kinetics and lineage bias (Dykstra et al., 2007; Kim et al., 2014; Naik et al., 2013; Sun et al., 2014; Verovskaya et al., 2013; Wu et al., 2014; Yamamoto et al., 2013). However, these approaches are not feasible in tracking DC development, as DCs are rare, and most are residential in lymphoid tissues such that serial sampling *in vivo* is not possible (Shortman and Naik, 2007). Long-term imaging can allow accurate reconstruction of pedigrees from transient progenitors that produce rare progeny such as DCs *in vitro*. However, because of technical demands it generally only allows assessment of 10–100 s clones for a short period of days to weeks, which might not be sufficient to allow full differentiation (Dursun et al., 2016; Skylaki et al., 2016). Recent "pedigree" tools that measure evolving barcodes in progeny can infer developmental history (Frieda et al., 2017; Kalhor et al., 2017; McKenna et al., 2016) but are limited in their assessment of clonal kinetics.

Another method that aims to recapitulate the dynamic aspects of development and differentiation is "pseudo-time" analyses,

which infer developmental trajectories by assuming that single cells within a population represent different "snapshots" along archetypal paths and align cells on the basis of their proteomic or transcriptomic profiles (Wagner et al., 2016). These models can be of great benefit in understanding the order of gene and protein expression in developmental pseudo-time. A confounding factor, however, is the inability to assess individual clones, as data are derived from a snapshot assessment or with no lineage connection when assessed between time points. Therefore, such archetypal trajectories may mask heterogeneity at the clonal level, including features such as kinetics, lineage bias, and division destiny (Marchingo et al., 2014).

Cellular barcoding allows tracking of clonal fate by differential tagging of individual progenitors with unique and heritable DNA barcodes (Bystrykh et al., 2012; Naik et al., 2014). Quantification and barcode comparison between progeny cell types allows inference of lineage relationships (i.e., barcodes shared between cell types implies common ancestors, whereas differing barcodes implies separate ancestors). Here we combine cellular barcoding and DC development in FL cultures to facilitate longitudinal assessment of clonal kinetics in a robust, controlled, and high-throughput manner by serially sampling progeny from the same wells at multiple time points. Our results highlight that there are several distinct classes of cellular trajectories in DC development: each consists of clones with a similar pattern of DC subtypes produced over time but with varying properties including the timing, duration, and magnitude of clonal waves. Importantly, using clone-splitting experiments, we demonstrate that many of these cellular trajectories are "programmed" within individual HSPCs. Furthermore, we demonstrate that pDC and cDC development has already largely diverged at the HSPC stage, not downstream in the CDPs, as is currently assumed. Our results offer a powerful analytical and visualization framework that reveals the diversity of clonal kinetics and cellular trajectories.

RESULTS

Longitudinal Tracking of Clonal DC Development Reveals Time-Varying Patterns

To track clonal DC development longitudinally, we barcode-labeled mouse Sca1^{+} cKit^{hi} cells that contained early HSPCs and cultured them with FL to allow DC generation (Figures 1A and S1A). The cultures were serially split in two at various times such that half of the cells were sorted for the DC subtypes using flow cytometry for subsequent barcode analysis, and half were kept in culture with a compensating amount of fresh media (Figure 1A). To accurately define the DC subtypes, we used CD11c, major histocompatibility complex class II (MHCII), Siglec-H, C-C chemokine receptor type 9 (CCR9), Sirpα, and CD24 (Figure 1B). In addition, we sorted cells that were outside these DC gates collectively as "non-DCs" to allow estimation of the recovery of barcodes in the culture at any given time points and track clones that still contained DC progenitors. CCR9 inclusion was critical to define bona fide pDCs as Siglec-H^{+}CCR9^{-} cells generated cDCs upon re-culture (Figure S2) (Schlitzer et al., 2011). Importantly, individual samples were separated into technical replicates after sorting and cell lysis to allow assessment of technical variation of barcode recovery (Figure 1A). Furthermore, control experiments (Figure S1B) were performed and demonstrated that serial sampling of barcoded progeny at the indicated time intervals was a robust approach to measure DC clonal kinetics (Figures S1C and S1D).

Our method assessed DC developmental dynamics and revealed time-varying patterns. The percentages of barcodes detected at each time point over total seeded barcoded cells varied and were consistently lower than the percentages of detected barcodes across time (Table S1). Stacked histograms of the number of cells produced by each detected clone over time (Figures 1C and 1D) showed a temporal shift of DC contribution by a spectrum of large and small clones, and this pattern was apparent for all DC subtypes. This indicated that DC generation was not sustained by a set of "stable" clones within the tracking period, and the contribution by different clones was not equal. We also generated a heatmap showing the barcode contribution to the number of DCs (biomass) from all cell types at all time points to capture the entirety of the data (Figure 1E). Again, the shift of clonal contribution to cell types over time was apparent, as was their bias.

Next, we reasoned that the asynchronous waves of clonal contribution could affect classification of clone output. For example, if a multipotent clone generated pDCs at an early time point and cDCs later, it would be classified as having a pDC-only or cDC-only fate depending on which time point was assessed. To test this, we first categorized clones into four classes (no DCs, pDC only, cDC only, and pDC/cDC) and determined that only ~30%–40% of clones generated DCs when considered at any given single time point (Figures 2A and S1E). However, when we compared the "across time" fate, taking into account a clone's capacity to produce DCs at multiple time points, that proportion of DC-generating clones increased to nearly 90%. In addition, ~20% of clones were re-classified from unipotent (pDC or cDC only) when measured at single time points to multipotent (pDC/cDC) when all time points were considered (Figure 2A). The asynchronous contribution to different DC subtypes over time was indeed apparent in the majority of clones using violin plots (Figure 2C). Therefore, fate should be considered in the context of time for a full appreciation of a clone's potential. We further quantified the contribution to the number of DC subtypes by different classes of clones on the basis of the definition "across time" and observed lower contribution by multipotent (~40%) than unipotent (~60%) clones to both pDCs and cDCs (Figure 2B). These results highlight the importance of tracking development longitudinally to accurately and thoroughly interpret cellular output. Furthermore, our results indicate that cDCs and pDCs are largely generated by progenitors that have already branched.

DiSNE Movies Allow Visualization of Clonal Dynamics

To facilitate interpretation of the kinetics of clonal contribution, we developed a dynamic visualization tool termed developmental interpolated t-distributed stochastic neighborhood embedding (t-SNE) (DiSNE) time-lapse movies (Movies S1, S2, and S3). First, t-SNE (Van der Maaten and Hinton, 2008) was applied to reduce the dimensionality of the dataset to a two-dimensional (2D) map in which the properties of clones in terms of subtypes and number of cells produced at different times

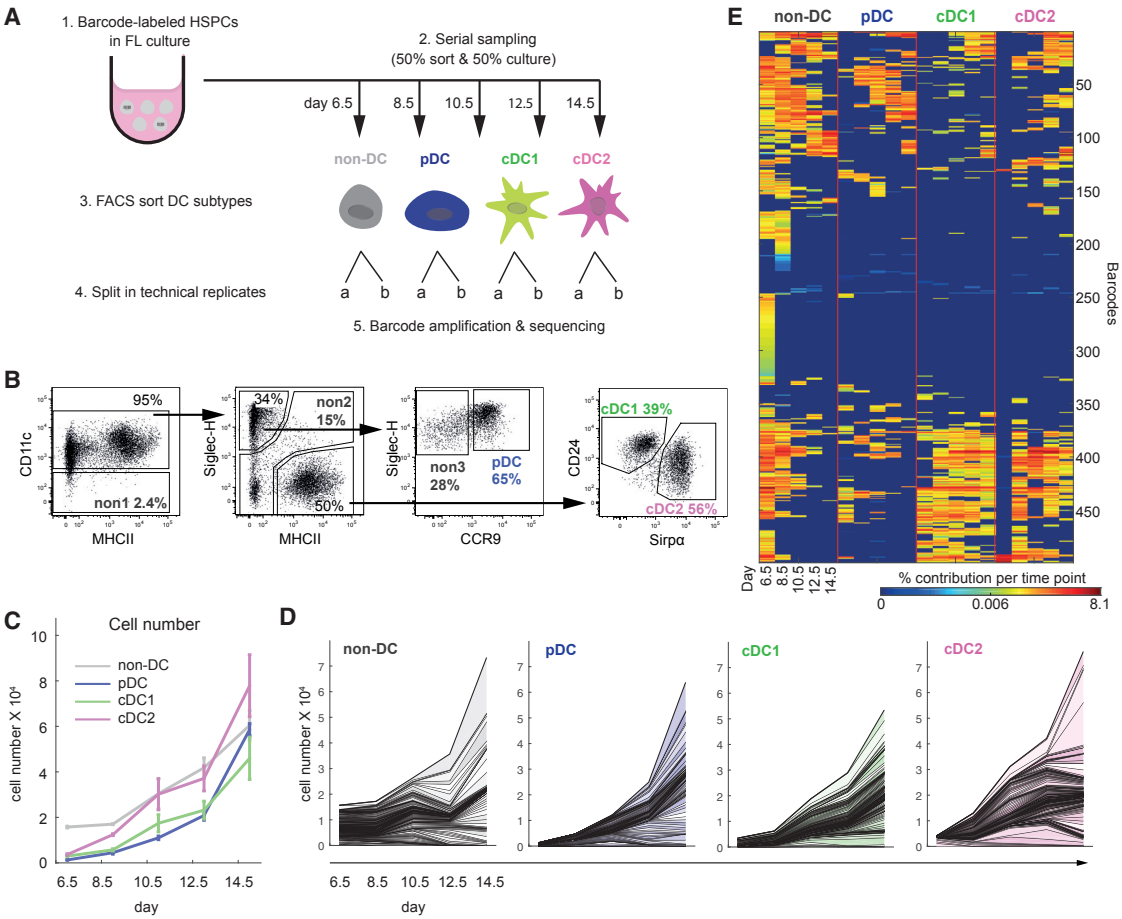


Figure 1. Longitudinal Tracking Reveals Asynchronous Waves of DC Generation

(A) Experimental setup. HSPCs (cKit$^+$Sca1$^+$) from mouse bone marrow (BM) were transduced with lentivirus containing DNA barcodes and cultured in FL-supplemented DC conditioned medium. At each time point, cells were equally split for either DC subtype isolation or further development in culture. Non-DCs (non1 + non2 + non3), pDCs, cDC1s, and cDC2s were sorted as in (B) at each time point. Samples were then lysed and split into technical replicates, and barcodes were amplified and sequenced.

(B) Gating strategy to isolate pDCs, cDC1s, cDC2s, and non-DCs using CD11c, MHCII, Siglec-H, CCR9, Sirpα, and CD24. Numbers represent percentages of cells from parent gate.

(C) Number of DC subtype generation over time at the population level.

(D) Stacked histogram showing clonal contribution (i.e., per barcode) to each DC subtype number over time. Clones are shown in the same order for each subtype. It is apparent that clones differ in size and also in timing of expansion.

(E) Heatmap representation of clonal output to DC subtypes from individual time points.

Data in (C) are average ± SEM of three independent cultures from one experiment, representative of three independent experiments. (D) and (E) show all clones from one representative culture.

dictated the position of each barcoded HSPC. To visualize clonal fate and DC biomass, we created "t-SNE pie maps" by generating a pie chart representing the proportional output to different DC subtypes and altering point size, respectively (Figure 3). Finally, changes in pie size and composition were interpolated between flanking data points during DC development for dynamic visualization.

We performed DiSNE visualization on data pooled from three independent wells, incorporating all time points available (Figure 3; Movie S1). Similar to the heatmap representation, heterogeneity was observed, but patterns were more easily distinguishable considering that the bias was incorporated into one pie, rather than four elements, and that clone size was better represented through dot size rather than color. These DiSNE movies (Movies S1, S2, and S3) portrayed the dynamic process of DC development encompassing the complexities of qualitative, quantitative, and now temporal characteristics of each clone underlying development. Therefore, DiSNE movies are an effective and powerful tool for visualization of clonal dynamics, and this technique has been packaged into a standalone software package PieMaker (https://data.mendeley.com/datasets/9mkz5n9jtf/1).

Multiple Trajectories of DC Development

To further characterize the clonal dynamics of DC development, we compared several clustering methods and observed similar

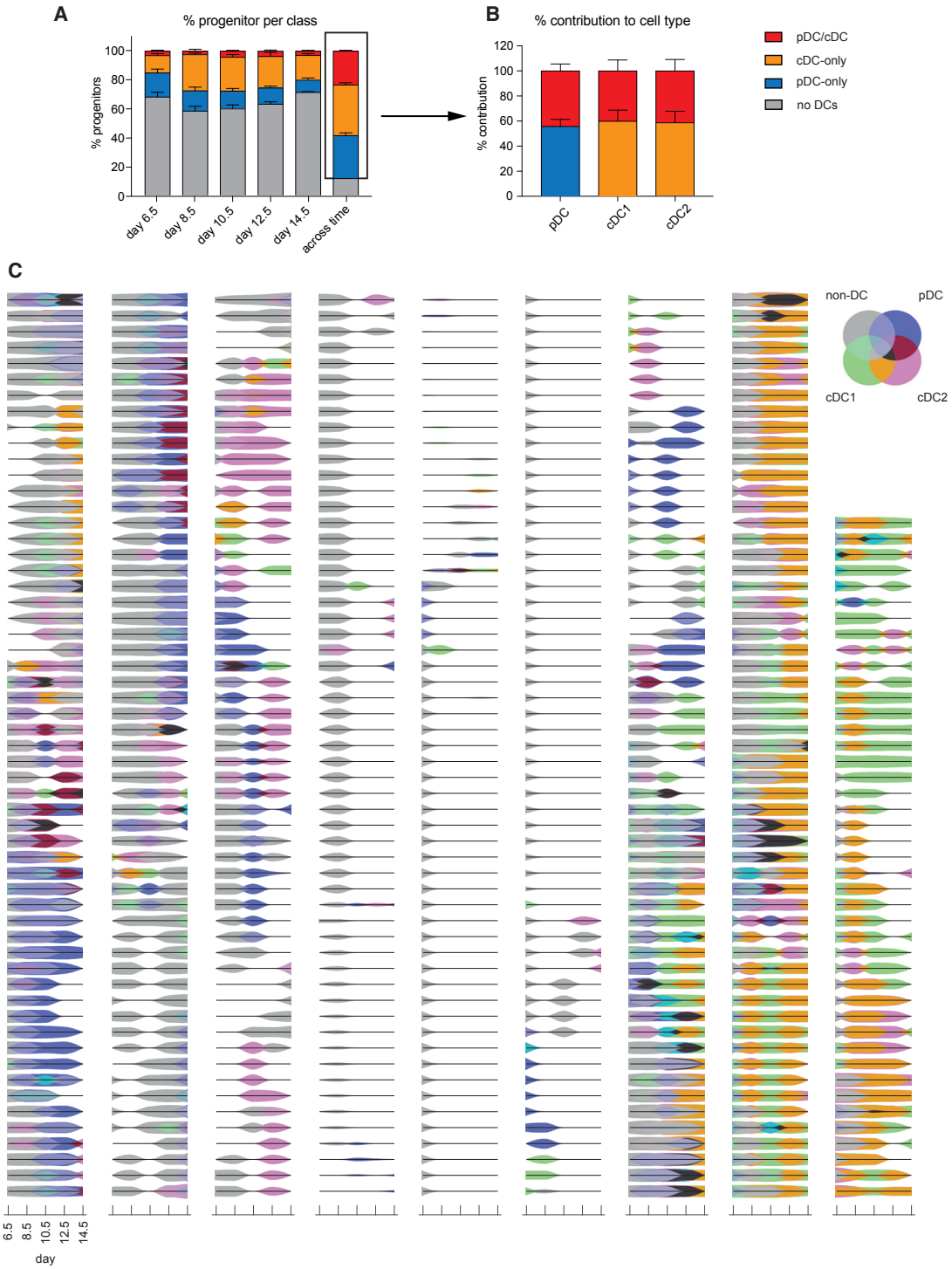


Figure 2. Longitudinal Tracking Allows Accurate Interpretation of Clonal Fate

(A) Categorization of clones into four classes, including no DCs, pDC only, cDC only, and pDC/cDC, on the basis of subtype output at a single time point or across time.

(B) Percentage contribution to cell types from three classes of clones on the basis of across-time definition.

(C) Violin plots showing clonal output of individual barcodes over time. The width of the violin is proportional to the contribution of the clone to the corresponding cell type at that time point.

Data in (A) and (B) are average + SEM of three independent cultures from one experiment, representative of three independent experiments. (C) shows the same clones as in Figures 1D and 1E, from one representative culture.

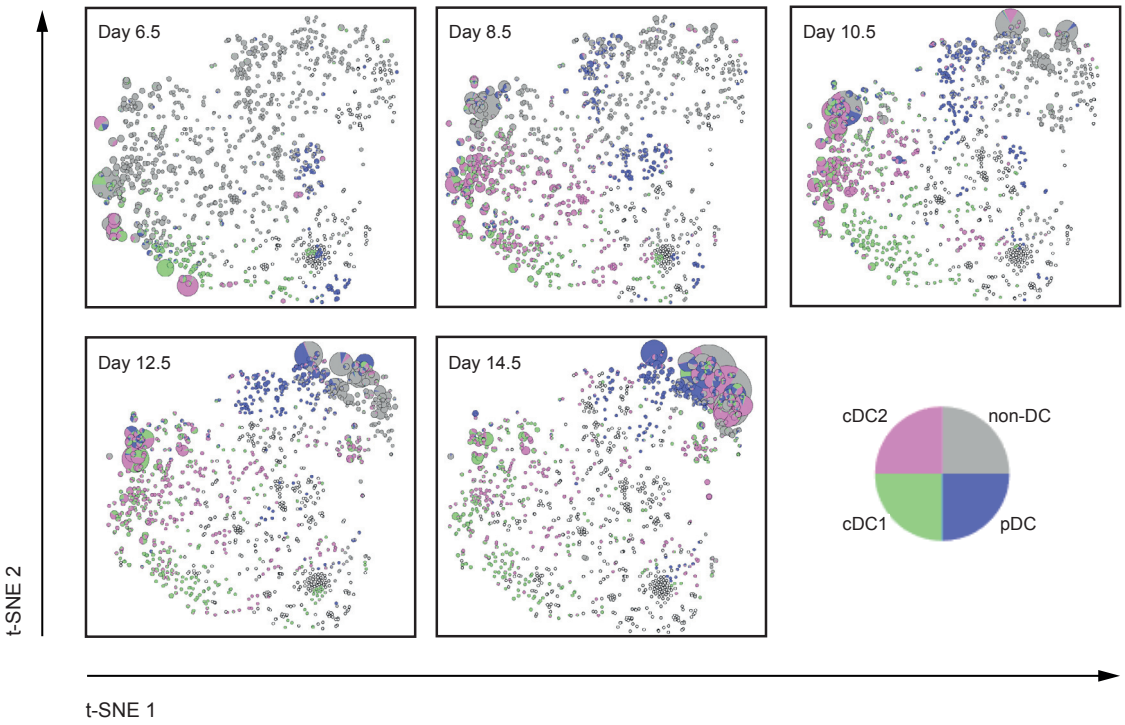


Figure 3. Visualizing Diversity of DC Cellular Trajectories Using DiSNE

Static t-SNE pie map at each time point (see Movie S1 for a dynamic visualization). Each circle represents a barcode-labeled progenitor and the size scaled to the number of cells produced by that clone per time point. Each sector in the pie chart represents the proportion of each cell type produced. Data are pooled from three independent cultures (368 data points each [out of 368, 410, and 384] for equivalence) from one experiment, representative of three independent experiments. Movies S2 and S3 show results from the other two experiments.

results (Figure S3). We then applied density-based spatial clustering of applications with noise (DBSCAN) (Ester et al., 1996), which best cohered with our DiSNE visualization and identified 16 clusters (Figure 4A). Visualization using spindle plots for each cluster showed that clusters were mainly separated by distinct fate bias or timing of contribution with similar fate output (Figure 4B). Interestingly, t-SNE mostly positioned clusters with a similar fate but asynchronous waves of contribution across a band in the plot (see manually annotated circles in Figure 4A) to form four major groups of trajectories, including cDC biased, pDC biased, multipotent, and a group of very small clones with mixed output (Figures 4B and 4F). There was large variation in the number of clones and DC biomass produced by each cluster (Figure 4C). The most prominent trajectory was cDC biased, which comprised ~43% of clones that contributed ~60% of cDC generation (Figures 4D and 4E). Similarly, ~33% of clones followed a pDC-biased trajectory, which generated more than half of pDCs (Figures 4D and 4E). Only 12% of clones were identified in the multipotent clusters, which contributed to 36% of pDCs, 31% of cDC1s, and 39% of cDC2s (Figures 4D and 4E). In addition, cluster 2 was in a region containing very small clones that were mostly unipotent. These represented 12% of total clones and fewer than 1% of the total number of DCs generated (Figures 4D and 4E). Importantly, independent wells within the same experiment were reproducible by comparing the occurrence of barcodes in each cluster (Figure 4G) and between experiments using Jensen-Shannon (JS) divergence (Amir et al., 2013) to holistically assess similarity between datasets (Figure 4H). Thus, we have identified multiple major trajectories of DC development and demonstrated the majority of clones within the HSPC fraction, but not all, follow cDC- or pDC-biased trajectories that contribute to the majority of their biomass.

Cellular Trajectories Are Intrinsically Programmed

Next, we asked whether the cellular trajectories of siblings derived from a single clone are highly correlated. To this end, we applied clone splitting by first pre-expanding barcoded progenitors for 4.5 days and then equally split the wells into two parallel FL cultures (Figure 5A). We then performed serial sampling and barcode analysis on both arms of the experiment as described. We compared the fate and clone size of shared barcodes (58% in experiment 1 and 73% in experiment 2) in parallel cultures across all time points (Figures 5B–5D and S4). Fate conservation was defined using JS divergence or cosine similarity, in which both measured similarity in clonal kinetics (types of progeny produced and the order) and produced similar results (Figure S4B). Size conservation was measured as the base two logarithm of the ratio of biomass between the shared barcodes, which essentially measured the discrepancy in division number between splits. Interestingly, we found that many sisters were concordant in their cellular trajectories, implying that descendant cells carried a "memory" of what DCs to make, when to

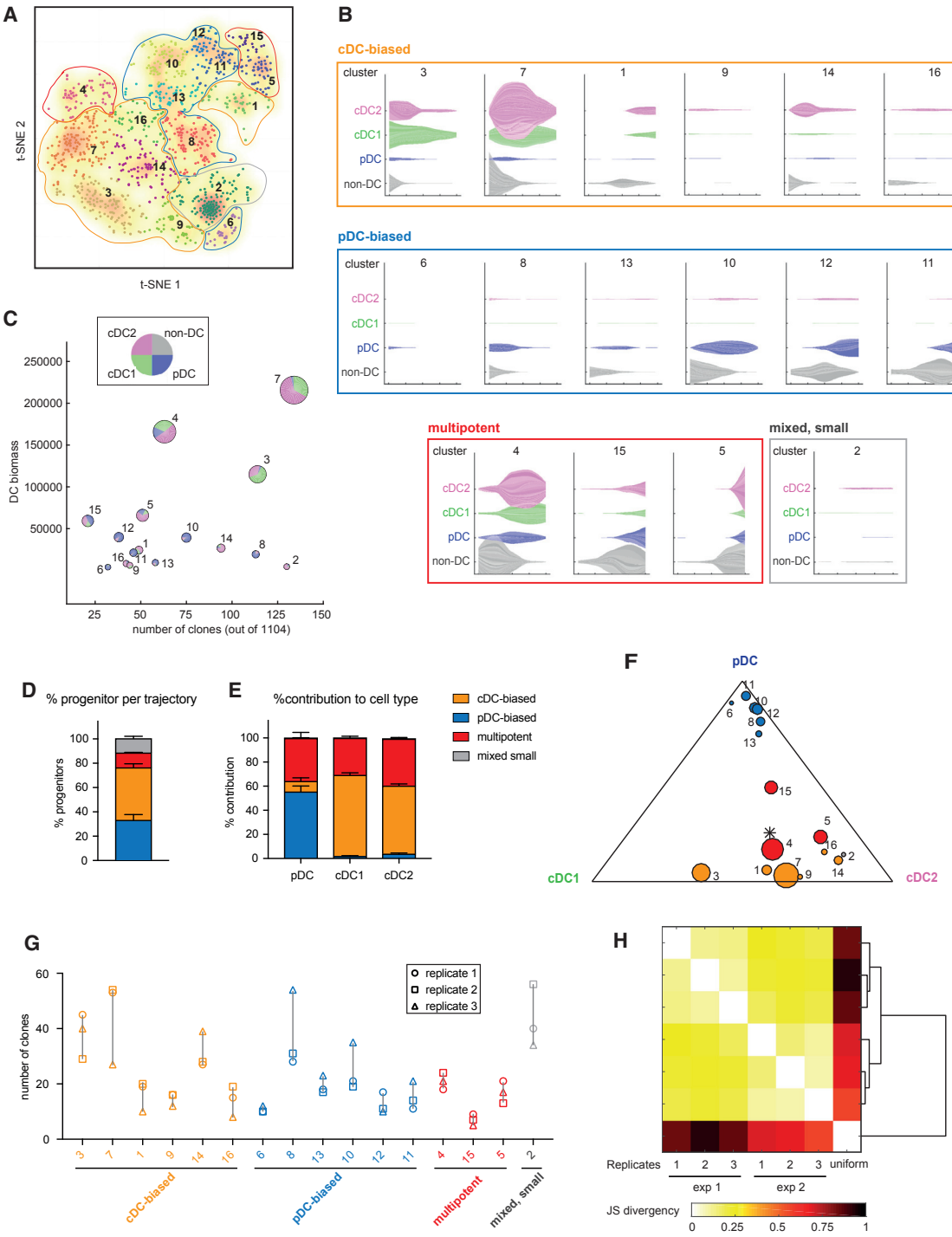


Figure 4. Identification of Major Classes of DC Cellular Trajectories

(A) DBSCAN-based algorithm identifies 16 clusters on the t-SNE map (as in Figure 3). Most clusters correlate well with the overlaid barcode density heatmap. The clusters are manually annotated into four major classes of trajectories on the basis of distinct fate output.

(B) Spindle plots showing contribution to each subtype over time by clones from individual clusters. The width of the spindle is proportional to the contribution of the cluster to the corresponding cell type at that time point, and each partition of the spindle (varying color shades) represents individual clones within the cluster.

(C) Each cluster is quantified in terms of both the number of clones (out of a total of 1,104, pool of three independent cultures) it includes (x axis) and DC biomass (the number of DCs it contributes) (y axis, pie radius). Pie charts show cluster compositions in terms of DC subtypes.

(D) Percentage progenitors from each trajectory class as defined in (A) and (B).

(E) Percentage contribution to cell types by each trajectory class. In (D) and (E), average + SEM of three independent cultures is shown.

(legend continued on next page)

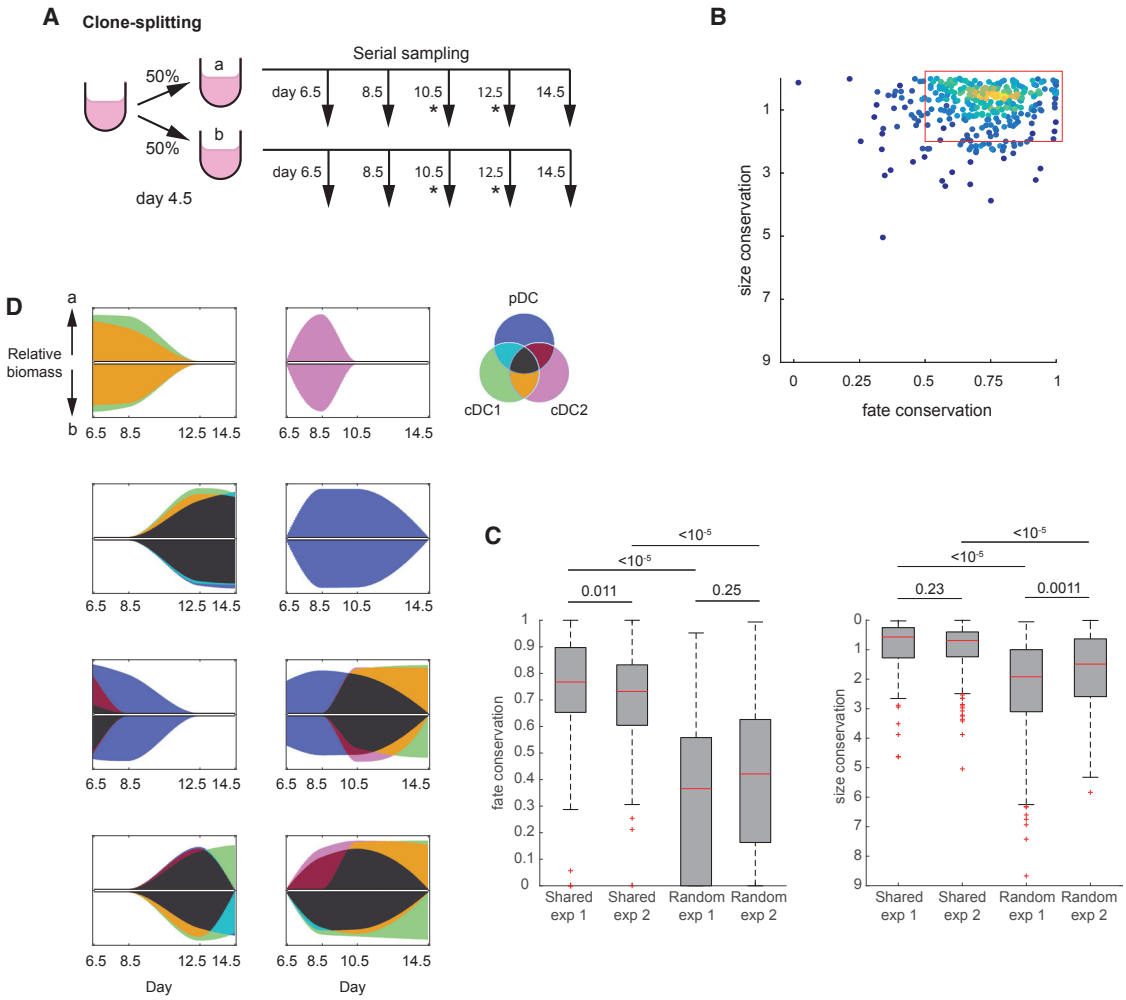


Figure 5. Clonal Cellular Trajectories Are Largely Programmed

(A) Schematic of clone-splitting experiment. Barcoded progenitors were pre-expanded in FL culture for 4.5 days and split into two parallel cultures (a and b). Serial sampling was then performed from both arms as described in Figure 1A. Asterisk, data from day 10.5 are lacking in experiment 1, and data from day 12.5 are lacking in experiment 2 because of technical issues.

(B) Conservation of shared barcodes across all time points. Each point represents a barcode with reads detected in both halves of the split culture. For each barcode, size conservation is defined as the base 2 logarithm of ratio of total read counts, and fate conservation is defined as JS divergence. Clones inside the gate represent 80% of total shared barcodes, which contributes to 80% of total biomass. Data are a pool of two sets of parallel cultures from experiment 1, representative of two independent experiments.

(C) Summary of fate and clone conservation value comparing split barcodes with randomly paired unrelated barcodes. Boxplots span interquartile range: The central line indicates the median; the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively; the whiskers extend to the most extreme data points not considered outlier; an outlier is a point further than 1.5 interquartile ranges from the box in either direction. Pooled data from both independent experiments are shown. Statistical significance is measured using Mann-Whitney U test.

(D) Paired violin plots comparing cellular trajectories from two arms of split culture (a versus b). Eight examples of clones with high conservation values are shown. Full list of clones from experiment 1 is shown in Figure S4.

make them, and how many cells to produce (Figures 5 and S4A). These results are consistent with fate being a heterogeneous, yet intrinsic and heritable property of individual founder cells when measured in similar environments.

DISCUSSION

The framework developed here provides a statistically robust, quantitative, visually intuitive approach for high-throughput

(F) Ternary plot showing subtype bias of each cluster. Circle size is proportional to DC biomass of the cluster. Asterisk denotes the population average.

(G) Barcode representation from the three independent cultures (replicates) in each cluster.

(H) JS divergence measuring the similarity between independent cultures within the same experiment (very low value, highly similar pattern), between two independent experiments (low value, reproducible pattern), and between uniformly distributed pattern on the defined t-SNE region (high value, dissimilar pattern).

tracking of clonal kinetics. It allows systematic examination of lineage trajectories of any developmental system, whereby cells can be cultured *ex vivo* and subsampled at desired time intervals. Our results indicate that assessment of bona fide clonal cellular trajectories is crucial to accurately determine clonal fate, as opposed to measuring fate at a fixed time point. In addition, by incorporating clone splitting, we demonstrate that clonal fate and waves of contribution to DCs is heterogeneous yet largely programmed early in the developing clone. This provides the rationale to combine our method with other approaches such as single-cell RNA sequencing in parallel to not only measure cellular trajectories but the underlying molecular trajectories that guide these properties or to test the effect of biological variation or perturbation such as drug treatment and genetic manipulation on one arm of the clone-splitting experiment.

Importantly, we demonstrate that the majority of HSPCs already have a cDC- or pDC-biased fate by measuring clonal output across multiple time points in FL cultures of DC development. Our results do not support the current model, which implies a common origin of cDCs and pDCs from CDPs (Guilliams et al., 2014). This could be partly explained in that many prior studies do not incorporate CCR9 to define pDCs, leading to possible misallocation of cDC precursors as pDCs. Similarly, two recent studies (See et al., 2017; Villani et al., 2017) using single-cell profiling of the human DC compartment independently identified contaminating DC precursors within the phenotypically defined pDCs. Future studies should determine whether those populations and murine Siglec-H^{+}CCR9^{-} cells represent the same precursor population. Our observation of early cDC and pDC bifurcation is also partly supported by the identification of cDC-, cDC1-, and cDC2-committed progenitors in various fractions of HSPCs and downstream progenitors (Grajales-Reyes et al., 2015; Schlitzer et al., 2015; Schraml et al., 2013). Importantly, our results indicate that a pDC-committed progenitor population likely exists within HSPC fraction, indicative of early branching, similar to a recent study (Lee et al., 2017).

Our results highlight a remarkable degree of heterogeneity within early HSPC population. Longer term efforts should appraise not only progenitors but also their progeny at a single-cell level to determine how origin dictates functional heterogeneity. This information, combined with the molecular drivers that underlie true cellular trajectories, and within an *in vivo* context, are necessary for a full understanding of development.

EXPERIMENTAL PROCEDURES

Mice

All mice were bred and maintained under specific pathogen-free conditions at the Walter and Eliza Hall Institute (WEHI), according to institutional guidelines. Either C57BL/6 (CD45.2) or C57BL/6 Pep^{3b} (CD45.1) male mice aged 8–16 weeks were used.

Cellular Barcoding

Barcode transduction, amplification, and data processing were performed largely as described previously (Naik et al., 2013). See the Supplemental Information for detailed procedures.

FL Culture and Serial Sampling

Labeled cells (5×10^3) were cultured with 200 μL DC conditioned medium supplemented with hFL (BioXcell, 800 ng/mL) per well in a 96-well round-bottom plate. After 6.5 days of culture, cells were gently mixed a few times with a pipette, half were removed for subtype isolation by flow cytometry, and another half were kept in culture with medium topped up to 200 μL. The same procedure was repeated every 2 days. At the last time point, all cells from each well were harvested and sorted.

Clone-splitting experiments were performed to assess conservation of fate between shared barcodes over time. Briefly, wells were split into two at day 4.5 and both cultured until day 6.5. After that, serial sampling was performed on both splits every 2 days, and analysis was performed from each split well as the other samples.

Barcode Categorization

Each barcode was categorized into four classes: "no DCs" (produced no mature DCs), "pDC only" (produced pDCs but no cDC1s or cDC2s), "cDC only" (produced cDC1 and/or cDC2 cells but no pDCs), and "pDC/cDC" (produced pDCs and either one or both cDC types). Barcode categorization was done on the basis of minimal read count and minimal proportion. In Figure 2A, minimal read count was set to 750, and minimal proportion was set to 5%. For example, if barcode A had 1,000 reads in pDC, 600 reads in cDC1, and 90,000 reads in cDC2, the cDC1 reads was first set to zero, as it did not pass the minimal read count threshold. As this barcode produced 1% pDCs (< 5%) and 99% cDC2s, it was classified as a cDC-only clone. Categorization was performed on the basis of data at each time point independently or on the basis of data across time. For categorization across time, a barcode was considered to produce a certain subtype X if it produced subtype X at any of the time points. Categorization was repeated using varying value combination of the two thresholds to verify that small changes in the values of these parameters qualitatively resulted in similar outcomes (Figure S1E).

Visualization Using t-SNE, Static Pie Maps, and DiSNE Movies

First, t-SNE was performed with default parameters to reduce the dimensions of the dataset to 2D (Van der Maaten and Hinton, 2008). Hyperbolic arcsine-transformed biomass counts from all time points were pooled from three independent cultures as input for the t-SNE algorithm. Barcodes that did not produce any DCs were excluded. The output of t-SNE was used for downstream visualization and clustering. To visualize clonal output and size on the t-SNE map, each barcode was represented as a pie chart (t-SNE pie maps). The segments of the chart depict the proportion of each subtype at a particular time point. The radius of the pie chart reflects the total biomass of the given barcode at the given time point. For the purpose of visualizing individual cellular trajectories (developmental changes over time) and clusters (see below), a cubic spline-based interpolation for time values between experimental time points was applied. Depending on the settings for DiSNE movie generation, linearly interpolated frames can be added between frames that correspond to experimental time points (see manual for PieMaker software at: https://data.mendeley.com/datasets/9mkz5n9jtf/1).

Clustering

To identify major patterns, several clustering methods were applied, including DBSCAN (Ester et al., 1996), Gaussian-mixture clustering, and the affinity propagation algorithm (Frey and Dueck, 2007). These methods were applied on both raw data and using scatterplots derived from t-SNE as input. These methods were capable of producing similar results. For example, raw data-based clustering resulted in clusters that were spatially consistent when projected onto t-SNE plots (Figure S3A) and produced trajectories with similar patterns (data not shown).

A barcode density plot using kernel density estimation via diffusion (Botev et al., 2010) was generated to assess the feasibility of each particular clustering method by first running each of the three algorithms on grids of parameter values and visually inspecting how well the resulting clustering aligned with the barcode density plot. DBSCAN-based clustering was found to align with the density plot best. Therefore, DBSCAN was used to identify cluster centroids, and each unassigned point was assigned to the cluster with the nearest centroids. The resulting clusters were manually categorized a "cDC biased," "pDC biased," "multipotent," or "mixed, small" on the basis of visual inspection of the corresponding DiSNE movie, visualization of subtype output per cluster using spindle plots (Figure 3B), and visualization of fate bias per

cluster using ternary plots (Figure 3F). The spindle plots are stacks of biomasses of barcodes included in the corresponding cluster. Individual barcodes can be distinguished by varying color shades. The ternary plot was generated using proportions of pDC, cDC1, and cDC2 biomasses (non-DCs excluded) to define coordinates for each cluster in the equilateral triangles.

Conservation between Shared Barcodes in Split Cultures

Given a split culture, barcodes without any DC biomass in each of the split parts (non-DCs only) and barcodes that were present only in one of the splits were excluded, and the rest were identified as shared barcodes. Fate conservation was computed to measure similarity between kinetics of DC subtype production (e.g., whether split parts of the same clone produced same types of DCs and in the same order). First, biomass values were hyperbolic arcsine-transformed, and JS divergence and cosine similarity were computed. Both methods produced very similar results (Figure S4B), and hence JS divergence was used to estimate fate conservation in Figure 5. Size conservation was computed to measure similarity in clonal expansion between the split parts of the same clone. First, total biomass per barcode was calculated (sum of biomass from all subtypes from all time points) for each split part. Next, the ratio of the smaller total biomass to the larger was calculated, and the base 2 logarithm of this ratio was computed as a measure of size conservation. For example, a difference of 1 could be interpreted that one part of the clone made on average one more division round. Note that biomass of non-DCs was excluded during computation of both fate and size conservation. Random controls were generated by randomly paired unrelated barcodes in the same culture to assess whether the observed conservation was due to chance.

Statistical Analysis

The Mann-Whitney U test was performed to measure the significance of the observed difference between groups. All data are presented in boxplots that span the interquartile range.

DATA AND SOFTWARE AVAILABILITY

The data reported in this study have been deposited to Mendeley Data at 10.17632/9mkz5n9jtf/1.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, four figures, one table, and three movies and can be found with this article online at https://doi.org/10.1016/j.celrep.2018.02.046.

ACKNOWLEDGMENTS

We thank the Walter and Eliza Hall fluorescence-activated cell sorting (FACS) laboratory, Dr. Stephen Wilcox, and Dr. Tom Weber for technical support. We thank Dr. Samir Taoudi, Prof. Gabrielle Belz, and Prof. Stephen Nutt for insightful discussions. This work was supported by grants from the National Health & Medical Research Council, Australia, GNT1062820, GNT1100033, and GNT1145184, and the Australia Research Council's special initiative Stem Cells Australia.

AUTHOR CONTRIBUTIONS

Conceptualization, D.S.L., A.K., J.G., and S.H.N.; Methodology, D.S.L., A.K., J.G., and S.H.N.; Software, A.K. and J.G.; Formal Analysis, A.K.; Investigation, D.S.L.; Visualization, D.S.L. and A.K.; Writing – Original Draft, D.S.L., A.K., and S.H.N.; Writing – Review & Editing, D.S.L., A.K., J.G., E.J.C., P.D.H., and S.H.N.; Supervision, S.H.N., E.C., and P.D.H.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: August 20, 2017
Revised: November 21, 2017
Accepted: February 9, 2018
Published: March 6, 2018

REFERENCES

Amir, el-A.D., Davis, K.L., Tadmor, M.D., Simonds, E.F., Levine, J.H., Bendall, S.C., Shenfeld, D.K., Krishnaswamy, S., Nolan, G.P., and Pe'er, D. (2013). viSNE enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia. Nat. Biotechnol. *31*, 545–552.

Botev, Z.I., Grotowski, J.F., and Kroese, D.P. (2010). Kernel density estimation via diffusion. Ann. Stat. *38*, 2916–2957.

Bystrykh, L.V., Verovskaya, E., Zwart, E., Broekhuis, M., and de Haan, G. (2012). Counting stem cells: methodological constraints. Nat. Methods *9*, 567–574.

Dursun, E., Endele, M., Musumeci, A., Failmezger, H., Wang, S.-H., Tresch, A., Schroeder, T., and Krug, A.B. (2016). Continuous single cell imaging reveals sequential steps of plasmacytoid dendritic cell development from common dendritic cell progenitors. Sci. Rep. *6*, 37462.

Dykstra, B., Kent, D., Bowie, M., McCaffrey, L., Hamilton, M., Lyons, K., Lee, S.-J., Brinkman, R., and Eaves, C. (2007). Long-term propagation of distinct hematopoietic differentiation programs in vivo. Cell Stem Cell *1*, 218–229.

Ema, H., Morita, Y., and Suda, T. (2014). Heterogeneity and hierarchy of hematopoietic stem cells. Exp. Hematol. *42*, 74–82.e2.

Ester, M., Kriegel, H.P., Sander, J., and Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, pp. 226–231.

Frey, B.J., and Dueck, D. (2007). Clustering by passing messages between data points. Science *315*, 972–976.

Frieda, K.L., Linton, J.M., Hormoz, S., Choi, J., Chow, K.K., Singer, Z.S., Budde, M.W., Elowitz, M.B., and Cai, L. (2017). Synthetic recording and in situ readout of lineage information in single cells. Nature *541*, 107–111.

Grajales-Reyes, G.E., Iwata, A., Albring, J., Wu, X., Tussiwand, R., Kc, W., Kretzer, N.M., Briseño, C.G., Durai, V., Bagadia, P., et al. (2015). Batf3 maintains autoactivation of Irf8 for commitment of a CD8α(+) conventional DC clonogenic progenitor. Nat. Immunol. *16*, 708–717.

Guilliams, M., Ginhoux, F., Jakubzick, C., Naik, S.H., Onai, N., Schraml, B.U., Segura, E., Tussiwand, R., and Yona, S. (2014). Dendritic cells, monocytes and macrophages: a unified nomenclature based on ontogeny. Nat. Rev. Immunol. *14*, 571–578.

Guo, G., Luc, S., Marco, E., Lin, T.-W., Peng, C., Kerenyi, M.A., Beyaz, S., Kim, W., Xu, J., Das, P.P., et al. (2013). Mapping cellular hierarchy by single-cell analysis of the cell surface repertoire. Cell Stem Cell *13*, 492–505.

Helft, J., Anjos-Afonso, F., van der Veen, A.G., Chakravarty, P., Bonnet, D., and Reis E Sousa, C. (2017). Dendritic cell lineage potential in human early hematopoietic progenitors. Cell Rep. *20*, 529–537.

Kalhor, R., Mali, P., and Church, G.M. (2017). Rapidly evolving homing CRISPR barcodes. Nat. Methods *14*, 195–200.

Kim, S., Kim, N., Presson, A.P., Metzger, M.E., Bonifacino, A.C., Sehl, M., Chow, S.A., Crooks, G.M., Dunbar, C.E., An, D.S., et al. (2014). Dynamics of HSPC repopulation in nonhuman primates revealed by a decade-long clonal-tracking study. Cell Stem Cell *14*, 473–485.

Lee, J., Zhou, Y.J., Ma, W., Zhang, W., Aljoufi, A., Luh, T., Lucero, K., Liang, D., Thomsen, M., Bhagat, G., et al. (2017). Lineage specification of human dendritic cells is marked by IRF8 expression in hematopoietic stem cells and multipotent progenitors. Nat. Immunol. *15*, 3221.

Månsson, R., Hultquist, A., Luc, S., Yang, L., Anderson, K., Kharazi, S., Al-Hashmi, S., Liuba, K., Thorén, L., Adolfsson, J., et al. (2007). Molecular evidence for hierarchical transcriptional lineage priming in fetal and adult stem cells and multipotent progenitors. Immunity *26*, 407–419.

Marchingo, J.M., Kan, A., Sutherland, R.M., Duffy, K.R., Wellard, C.J., Belz, G.T., Lew, A.M., Dowling, M.R., Heinzel, S., and Hodgkin, P.D. (2014). T cell signaling. Antigen affinity, costimulation, and cytokine inputs sum linearly to amplify T cell expansion. Science *346*, 1123–1127.

McKenna, A., Findlay, G.M., Gagnon, J.A., Horwitz, M.S., Schier, A.F., and Shendure, J. (2016). Whole-organism lineage tracing by combinatorial and cumulative genome editing. Science *353*, aaf7907.

Merad, M., Sathe, P., Helft, J., Miller, J., and Mortha, A. (2013). The dendritic cell lineage: ontogeny and function of dendritic cells and their subsets in the steady state and the inflamed setting. Annu. Rev. Immunol. *31*, 563–604.

Naik, S.H., Proietto, A.I., Wilson, N.S., Dakic, A., Schnorrer, P., Fuchsberger, M., Lahoud, M.H., O'Keeffe, M., Shao, Q.X., Chen, W.F., et al. (2005). Cutting edge: generation of splenic CD8+ and CD8- dendritic cell equivalents in Fms-like tyrosine kinase 3 ligand bone marrow cultures. J. Immunol. *174*, 6592–6597.

Naik, S.H., Metcalf, D., van Nieuwenhuijze, A., Wicks, I., Wu, L., O'Keeffe, M., and Shortman, K. (2006). Intrasplenic steady-state dendritic cell precursors that are distinct from monocytes. Nat. Immunol. *7*, 663–671.

Naik, S.H., Sathe, P., Park, H.Y., Metcalf, D., Proietto, A.I., Dakic, A., Carotta, S., O'Keeffe, M., Bahlo, M., Papenfuss, A., et al. (2007). Development of plasmacytoid and conventional dendritic cell subtypes from single precursor cells derived in vitro and in vivo. Nat. Immunol. *8*, 1217–1226.

Naik, S.H., Perié, L., Swart, E., Gerlach, C., van Rooij, N., de Boer, R.J., and Schumacher, T.N. (2013). Diverse and heritable lineage imprinting of early haematopoietic progenitors. Nature *496*, 229–232.

Naik, S.H., Schumacher, T.N., and Perié, L. (2014). Cellular barcoding: a technical appraisal. Exp. Hematol. *42*, 598–608.

Notta, F., Zandi, S., Takayama, N., Dobson, S., Gan, O.I., Wilson, G., Kaufmann, K.B., McLeod, J., Laurenti, E., Dunant, C.F., et al. (2016). Distinct routes of lineage development reshape the human blood hierarchy across ontogeny. Science *351*, aab2116.

Onai, N., Obata-Onai, A., Schmid, M.A., Ohteki, T., Jarrossay, D., and Manz, M.G. (2007). Identification of clonogenic common Flt3+M-CSFR+ plasmacytoid and conventional dendritic cell progenitors in mouse bone marrow. Nat. Immunol. *8*, 1207–1216.

Onai, N., Kurabayashi, K., Hosoi-Amaike, M., Toyama-Sorimachi, N., Matsushima, K., Inaba, K., and Ohteki, T. (2013). A clonogenic progenitor with prominent plasmacytoid dendritic cell developmental potential. Immunity *38*, 943–957.

Sanjuan-Pla, A., Macaulay, I.C., Jensen, C.T., Woll, P.S., Luis, T.C., Mead, A., Moore, S., Carella, C., Matsuoka, S., Bouriez Jones, T., et al. (2013). Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. Nature *502*, 232–236.

Schlitzer, A., Loschko, J., Mair, K., Vogelmann, R., Henkel, L., Einwächter, H., Schiemann, M., Niess, J.-H., Reindl, W., and Krug, A. (2011). Identification of CCR9- murine plasmacytoid DC precursors with plasticity to differentiate into conventional DCs. Blood *117*, 6562–6570.

Schlitzer, A., Sivakamasundari, V., Chen, J., Sumatoh, H.R.B., Schreuder, J., Lum, J., Malleret, B., Zhang, S., Larbi, A., Zolezzi, F., et al. (2015). Identification of cDC1- and cDC2-committed DC progenitors reveals early lineage priming at the common DC progenitor stage in the bone marrow. Nat. Immunol. *16*, 718–728.

Schraml, B.U., van Blijswijk, J., Zelenay, S., Whitney, P.G., Filby, A., Acton, S.E., Rogers, N.C., Moncaut, N., Carvajal, J.J., and Reis e Sousa, C. (2013). Genetic tracing via DNGR-1 expression history defines dendritic cells as a hematopoietic lineage. Cell *154*, 843–858.

See, P., Dutertre, C.-A., Chen, J., Günther, P., McGovern, N., Irac, S.E., Gunawan, M., Beyer, M., Händler, K., Duan, K., et al. (2017). Mapping the human DC lineage through the integration of high-dimensional techniques. Science *356*, eaag3009.

Shortman, K., and Naik, S.H. (2007). Steady-state and inflammatory dendritic-cell development. Nat. Rev. Immunol. *7*, 19–30.

Skylaki, S., Hilsenbeck, O., and Schroeder, T. (2016). Challenges in long-term imaging and quantification of single-cell dynamics. Nat. Biotechnol. *34*, 1137–1144.

Sun, J., Ramos, A., Chapman, B., Johnnidis, J.B., Le, L., Ho, Y.-J., Klein, A., Hofmann, O., and Camargo, F.D. (2014). Clonal dynamics of native haematopoiesis. Nature *514*, 322–327.

Van der Maaten, L., and Hinton, G. (2008). Visualizing data using t-SNE. J. Mach. Learn. Res. *9*, 2579–2605.

Verovskaya, E., Broekhuis, M.J.C., Zwart, E., Ritsema, M., van Os, R., de Haan, G., and Bystrykh, L.V. (2013). Heterogeneity of young and aged murine hematopoietic stem cells revealed by quantitative clonal analysis using cellular barcoding. Blood *122*, 523–532.

Villani, A.-C., Satija, R., Reynolds, G., Sarkizova, S., Shekhar, K., Fletcher, J., Griesbeck, M., Butler, A., Zheng, S., Lazo, S., et al. (2017). Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. Science *356*, eaah4573.

Wagner, A., Regev, A., and Yosef, N. (2016). Revealing the vectors of cellular identity with single-cell genomics. Nat. Biotechnol *34*, 1145–1160.

Wu, C., Li, B., Lu, R., Koelle, S.J., Yang, Y., Jares, A., Krouse, A.E., Metzger, M., Liang, F., Loré, K., et al. (2014). Clonal tracking of rhesus macaque hematopoiesis highlights a distinct lineage origin for natural killer cells. Cell Stem Cell *14*, 486–499.

Yamamoto, R., Morita, Y., Ooehara, J., Hamanaka, S., Onodera, M., Rudolph, K.L., Ema, H., and Nakauchi, H. (2013). Clonal analysis unveils self-renewing lineage-restricted progenitors generated directly from hematopoietic stem cells. Cell *154*, 1112–1126.