Combination image flow cytometry reveals novel methods for isolating megakaryocyte progenitors cells

Lucas M. Bush, Connor P. Healy, James E. Marvin and Tara L. Deans

Abstract

Megakaryocytes (MKs) are rare cells that develop in the bone marrow and give rise to platelets. A major challenge in studying megakaryocyte development, and the diseases that arise from these cells, is the identification and classification of MK progenitor cells that are produced during hematopoiesis. Here, we utilize combination image flow cytometry and principal component analysis (PCA) to definitively identify MK progenitor cell populations originating from a heterogeneous sample of bone marrow cells using minimal antibody panels.

Introduction

Hematopoietic stem cells (HSCs) are capable of producing all cells of the blood lineage throughout the lifetime of an individual by balancing self-renewal with differentiation[1]. In the mouse, HSCs are broadly defined by the $Lin^{-}Sca1^{+}cKit^{+}$ (LSK) immunophenotype[2,3], where HSCs capable of long-term bone marrow (BM) reconstitution (LT-HSCs) are further defined by expression of Thy1.1 and CD150, but lack expression of CD34 and FLT3. HSCs capable of short-term hematopoietic reconstitution (ST-HSCs) are defined by expression of CD34 and FLT3, but have lost expression of CD150. In the classic model of hematopoietic differentiation, ST-HSCs progress towards a multipotent progenitor (MPP) that has no self-renewal capacity, actively differentiates into all cells of the blood lineage, and is further defined by loss of Thy1.1 expression[2-4]. Recent studies suggest an alternative to the classical model of hematopoiesis; HSCs may undergo myeloid bypass to directly yield lineage-restricted myeloid progenitors (LRMPs) capable of clonal expansion and terminal differentiation into mature cell types[5,6]. Interestingly, three distinct subsets of LRMPs capable of prolonged MK repopulation have been identified: the common myeloid repopulating progenitor (CMRP); the megakaryocyte-erythroid repopulation progenitor (MERP); and the megakaryocyte repopulating progenitor (MKRP)[5]. To study these unique progenitor subpopulations, we capitalized on thrombopoietin (TPO)/c-Mpl signaling within the hematopoietic niche, or compartment, of the mouse bone marrow. TPO/c-Mpl signaling is essential for HSC survival, proliferation and differentiation into mature MKs[1,5,7-11]. Commitment down the MK lineage is characterized by an increase in cell size, endoreplication to increase DNA content, upregulation of MK-specific surface markers, and cytoplasmic remodeling to prepare for platelet production[9,10]. Although fundamental to the health of the animal, HSCs and MKs are extremely rare cell

types that constitute approximately 0.5% and 0.01% of the mouse BM, respectively, presenting complications for large-scale studies[12,13].

A challenge with identifying and isolating LRMP is that their expression of canonical MK markers are often similar to other myeloid cells, making it difficult to distinguish these subpopulations from other hematopoietic cells without using large antibody panels. As cells undergo differentiation, most cells develop distinct morphological differences when they are at various stages of cell fate. These individual morphological changes can be observed using image flow cytometry, where an image of each individual cell is captured as it flows through the cytometer. Altogether, each captured image contains data of multiple variables that can be used to characterize each individual cell by cell size, shape, and membrane texture, as well as expression levels of fluorescently-labeled surface markers. In all, these data enable the identification of cells at distinct stages of lineage development.

Here, we quantified the effect of TPO on HSCs, MK progenitors, and mature MKs stained with two different antibody panels using the ImageStream®X Mark II Imaging Flow Cytometer. Using only LSK antibodies, we identified an HSC-derived LRMP population (R4 cells) that proliferates and increases in size in response to TPO/c-Mpl signaling. To validate that R4 is a distinct subpopulation of cells in an unbiased way and identify the morphological characteristics which may identify the R4 subpopulation, principal components analysis (PCA) was used to mathematically capture variance in morphological characteristics of each cell in brightfield images. PCA is a statistical method to group highly correlated data together on the basis of phenotypic characteristics by converting a set of possibly correlated variables in a multivariate dataset into a set of values called principal components, which are displayed as axes. The input variables contribute to each principal component and the first principal component explains the largest possible variance in the dataset. For cells stained with LSK antibodies, unsupervised PCA determined that diameter/area and circularity, two morphological characteristics that arise as cells commit to the MK lineage, were the variables that best distinguished cells in the dataset. Lastly, image flow cytometry was also used to classify the phenotypically linked MK progenitor subsets and mature MKs using two antibodies or less. Collectively, we demonstrate that R4 cells represent a novel MK progenitor subset in the LSK compartment that preferentially gives rise to mature MKs in culture.

Methods

Bone marrow dissection and cell culture

Mice were maintained at the University of Utah under IACUC approved guidelines. The use of animals and/or animal-derived materials in this study were carried out in accordance with relevant guidelines and regulations.

Mice were euthanatized by CO_2 asphyxiation, followed by cervical dislocation. Femurs were dissected from the mouse, excess tissue was removed, and the epiphyses were snipped off. Bone marrow was flushed with IMDM (ThermoFisher Scientific, #12440053) + 10% FBS (Gibco, #16000044) and 1X Penicillin-Streptomycin (ThermoFisher Scientific, #15140122) using a 26.5-guage needle and collected in a 35mm dish. Red blood cells were lysed using ACK Lysis Buffer (ThermoFisher Scientific, # A1049201) for 5 min at room temperature, then the solution was centrifuged at 300xg for 5 min. The cells were washed 2x with HBSS (Gibco, #14025092) then counted using a hemocytometer. If cells were cultured before analysis, $1x10^7$ cells were incubated in IMDM + 10% FBS supplemented with 50 ng/ml of recombinant murine TPO (PeproTech, #315-14) in a T75 flask at 37 °C in 5% CO_2 for 72 h.

Preparation for image flow cytometry

After isolation from the bone marrow or culture with TPO for 72 h, cells were washed 2x with PBS. For the LSK stain, cells were resuspended in PBS + 3% BSA (ThermoFisher Scientific, #BP1600-100) with the addition of PE-Cy™7 Rat anti-Mouse CD117, PE Rat Anti-Mouse Ly-6A/E, and APC Mouse Lineage Antibody Cocktail, and incubated on ice protected from light for 30 min. Cells were washed 2x with PBS and resuspended at a final concentration of $2x10^7$ cells/ml in PBS. DAPI was added at least 5 min prior to running the sample on the image flow cytometer.

For the MK stain, cells were resuspended in prewarmed IMDM + 10% FBS containing Hoechst 33342 solution, then incubated at 37 °C for 45 min. Cells were washed with PBS then resuspended in PBS + 3% BSA with the addition of PE-Cyanine7 CD45 Monoclonal Antibody, PE Rat Anti-Mouse CD41, and APC CD42d Monoclonal Antibody and incubated on ice, protected from light for 30 min. Cells were washed 2x with PBS then resuspended at a final concentration of $2x10^7$ cells/ml in PBS. 7-AAD was added at least 5 min prior to running the sample on the image flow cytometer. For a complete description of staining reagents, see Table 1.

Image flow cytometry data acquisition and analysis

All experiments used the ImageStream®X Mark II Imaging Flow Cytometer and all image flow cytometry analysis was performed in the IDEAS® Software. Data from image flow cytometry experiments was acquired by gating for events that were in focus, did not contain calibration focus beads, and excluded viability dyes (supplementary Figure 1, 6). For a full description of data acquisition

and analysis methods, consult the INSPIRE™ ImageStream®X Mark II Imaging Flow Cytometer user's manual and the IDEAS® software user manual available online.

Image Analysis and Morphological Feature Extraction

In this analysis, brightfield images of lineage-negative cells from each experimental group (BM cells not treated TPO (no TPO) and BM cells treated with 50ng/ml TPO were first isolated from the full dataset. Individual cells in each image were distinguished from the background and defined automatically using the default brightfield settings in the IDEAS® software to produce masked images. Next, the IDEAS® software's built-in feature analysis utility was applied to the masked images and used to measure various morphological features of each cell. In total, twenty-nine morphological features were computed for each cell. Thirteen size features were computed: area; diameter, height; length; major/minor axis; major/minor axis weighted by intensity; perimeter; spot area min; thickness min/max; and width. Next, ten shape features were computed: aspect ratio; intensity weighted aspect ratio; circularity; compactness; elongatedness; lobe count; shape ratio; two-lobe symmetry (symmetry 2); three lobe symmetry (symmetry 3); and four lobe symmetry (symmetry 4). Lastly, six texture features were computed: the intensity of bright details with radii less than 3 (BDI.R3) and radii less than 7 (BDI.R7); contrast; modulation; spot count; and standard deviation. Full descriptions of these features and the means by which they are computed can be found in the IDEAS® software user manual. Finally, each observed cell was assigned a categorical variable, R1, R2, R3, or R4, based on the previously specified gates for Sca1 and cKit expression. For the surface marker expression analysis, each cell was assigned its corresponding Sca1 and cKit expression level instead of a categorical variable.

Principal Components Analysis

The twenty-nine morphological features from 31,265 observations in the no TPO dataset were centered, and scaled/normalized such that the values of each feature ranged from 0 to 1 (i.e. min-max normalization). Next, the principal components of the normalized data set were computed using R's **prcomp** function[14]. The number of principal components needed to describe the dataset was determined using the broken stick model and David Zelený's **evplot** function[15,16]. Observed principal components that describe more variance than the amount of variance predicted under the broken-stick model are considered necessary to effectively interpret the data. In this experiment, it was determined that the first three principal components were needed to adequately describe the data. Finally, bi-plots (plots with principal components on each axes) and correlation plots were generated with the **ggplot2** library[17] and the categorical variables (R1 through R4) were used to color each observation. For the surface marker expression bi-plots, the color of each observation was determined based on the relative expression of the surface markers cKit (green) and Sca1 (blue) (Fig. 4). The dynamic range of this color map was defined to capture 99% of the observations.

Using min-max normalization, the biplots for the 50ng/ml TPO dataset were generated by normalizing the 10,926 50ng/ml TPO observations and mapping them to the principal components of the no TPO dataset. This was achieved by scaling the 50ng/ml TPO feature values in relation to the no TPO PCA's center, and applying the rotation of the no TPO PCA matrix to the new observations. This plots the 50ng/ml TPO dataset using the same coordinate system as the no TPO dataset. Biplots for the 50ng/ml TPO dataset were generated in the same manner as the biplots for the no TPO dataset.

Fluorescence activated cell sorting (FACS), reculture, and antibody staining of R4 cells

Cells were prepared for FACS using LSK antibodies as described above. Cells were sorted into IMDM + 10% FBS with the addition of 50ng/ml TPO using the BD FACSAria™ III, then plated into a cell culture dish and incubated at 37 °C for an additional 48 h. After 48 h, cells were stained with PE-Cyanine7 CD45 Monoclonal Antibody and PE Rat Anti-Mouse CD41 and run on the BD FACSCanto™ II. All BD FACSAria™ III and BD FACSCanto™ II data analysis was performed using FlowJo®.

Results

Identification of a novel MK progenitor subpopulation

To quantify the effect of TPO on the hematopoietic compartment, cells were stained with the LSK panel of antibodies (Table 1) and analyzed using the image flow cytometer before and after culture with TPO (for gating strategy, see supplementary Figure 1). In response to culture with TPO, the subpopulations of Lin^- cells shifted, changing the landscape of Sca1 vs. cKit intensity plots (Figure 1A-B, supplementary Figure 2-3). Near complete loss of the R1 subpopulation indicates that R1 cells no longer retain the capacity for self-renewal but are primed for differentiation into cells expressing lineage markers. The R2 subpopulation is enriched after culture with TPO, suggesting an LRMP phenotype that has symmetrically divided to self-renew. The R3 subpopulation are the classic LSK cells that proliferate and retain high levels of Sca1/cKit expression in response to TPO/c-Mpl signaling. The R4 subpopulation also self-renews in response to TPO, suggesting a second, separate LRMP phenotype within the Lin^- BM compartment[3,5,6]. Quantitative inspection of image files associated with discrete flow cytometric events led to the observation of drastic morphological changes in R4 cells after culture with TPO (Figure 1C-D). To elucidate these changes, we applied PCA to the morphological features dataset acquired by the image flow cytometer. In this dataset, the observations are the lineage negative cells in the LSK stain and the variables are the 29-morphological features corresponding to the shape, size, and texture of those cells acquired by the image flow cytometer (Supplementary Figure 4A). The interpretation of this PCA is that the closer a given cell's principal component score is to another cell's principal component score, the more similar those two cells are morphologically.

Performing this analysis, principal components 1 through 3 (PC1-PC3, respectively) explain 74.0% of the observed variance in the dataset cumulatively—or 38.9%, 24.5%, and 10.6% of the observed variance in the dataset, respectively. Comparison of the percent variance explained by each principal component to the broken-stick model [16] confirmed that the first 3 principal components were sufficient to interpret the full dataset (Figure 2A). Size features contributed the most to PC1, shape features contributed the most to PC2, and texture features contributed the most to PC3 (Figure 2B). The top three morphological features that contributed to each principal component were diameter, area, and major axis for PC1; circularity, aspect ratio, and bilateral symmetry for PC2; and standard deviation, modulation, and contrast for PC3 (Figure 2B).

Converting the no TPO observations into the principal component space suggests that the R2 and R4 subpopulations contain cells that are morphologically distinct (Figure 2C). R4 cells are extended along the PC1 axes and are, therefore, are largely distinguished by their unique size. To a lesser extent, these cells also extend along the PC2 axes, which correlates mostly to shape. In contrast, R2 cells strongly extend along the PC2 axes indicating that these cells also have very distinct morphologies compared to cells in the other subpopulations. The R3 subpopulation extended roughly equally between PC1 and PC2 suggesting it may contain morphological intermediates between the R2 and R4 populations. PC3 did not distinguish any subpopulations but rather suggests the variation in texture between all the subpopulations is roughly the same (supplementary Figure 4B-C). The spread in the R3 subpopulation also suggests it has the most morphological variation overall, followed by the R4 population, the R2 population, and finally the R1 population which has the most uniform morphology.

Transforming the 50ng/ml TPO observations into the principal component space suggests that treatment with TPO exaggerates morphological differences between the R4 subpopulation and the other subpopulations (Figure 2D). In particular, the morphological differences in the R4 subpopulation become more consistent and distinctly extend along the PC1 axis. The R1, R2, and R3 subpopulations are also more internally consistent across all principal components, indicating these cells are more uniform in morphology after culturing with TPO. Taken together, these results indicate that the size and shape of individual R4 cells can be used to distinguish the R4 subpopulation from the rest of the cells in the LSK stain, and that R4 cell morphology becomes more uniform when cultured with TPO.

To illustrate the results of the PCA, plots of brightfield circularity vs. area of all lineage negative image files were generated for the no TPO and 50ng/ml TPO groups, respectively (Figure 2E-F, and Supplementary Figure 5). These plots establish the R4 subpopulation as the largest, most circular cell type of the lineage negative cells, a property that was exaggerated after culture with TPO (Figure 2E-F,

supplementary Figure 6). Because MK maturation is marked by an increase in cell size[10], and cKit^{high} cells exhibit an intrinsic MK bias[18], R4 cells were considered a novel, HSC-derived MK progenitor subset that is susceptible to further expansion via stem cell factor/cKit receptor signaling and differentiation into mature MKs.

Classification of MK progenitor subpopulations using image flow cytometry

To study the effect of TPO on MK development and maturation, mouse BM cells were stained with MK antibodies (Table 1) before and after exposure to 50ng/ml TPO. Flow cytometry plots of CD45 vs. CD41 intensity were generated for both time points and seven subpopulations (R1-R7, respectively) with distinct expression levels were identified (Figure 3A-B, for gating strategy see, supplementary Figure 7). In response to TPO/c-Mpl signaling, the cell fractions shifted by decreasing the percentage of cells in subpopulations R1 and R2, while the percentage of cells in R3-R7 increased, respectively (supplementary Figure 8).

The R1 cells are CD45^{-}CD41^{-}CD42d^{-}, have distinct nuclei, do not proliferate in response to TPO, and are highly prevalent in unstimulated BM, representative of a subpopulation of cells that does not self-renew but is primed for differentiation. R2 cells express only CD45 and do not proliferate in response to TPO, indicating an MPP that dies or differentiates but does not self-renew. R3 cells proliferate massively in response to TPO and upregulate CD45, CD41, and CD42d expression. This demonstrates an initial CMRP phenotype that can self-renew or directly gives rise to MKs when cultured with TPO. R4 cells emerge as a true subpopulation after TPO stimulation and are marked by low surface marker expression and DNA content, indicating a subpopulation that can self-renew in response to TPO but have low MK production potential. R5 cells are CD45^{-}CD41^{low}CD42d^{-} and lack nuclei, which is typical of maturing reticulocytes. R6 cells are CD45$^{-/low}$CD41$^{mid\rightarrow high}$CD42d^{+} cells with low DNA content, exemplary of an LT-HSC-derived MKRP that can self-renew but has not yet undergone endoreplication to increase ploidy [2,3,5]. R7 cells are enriched after TPO stimulation, express the highest levels of all surface markers and are highly ploidy, indicative of mature MKs produced through asymmetric cell division of CMRPs, MERPs, or MKRPs, respectively (supplementary Figure 9).

Inspection of image files associated with individual events in each subpopulation confirms the phenotypic identities determined by flow cytometry. MKs are large cells that adhere to the vascular endothelial surface and extend cytoplasmic protrusions into sinusoidal blood vessels to produce platelets[10]. Analysis of R7 image files confirms R7 cells as a true mature MKs (Fig. 3C-D, additional image files found in supplementary figure 10). R7 cells were the subpopulation with the lowest circularity score and largest area both before and after TPO exposure (supplementary Figure 11), characteristics expected of mature MKs. Additionally, the associated fluorescent images revealed bona

fide co-expression of surface markers and multilobed, polyploid nuclei within the confines of the cell membrane. The R7 subpopulation is phenotypically linked to the R3 and R6 subpopulations by surface marker expression and image analysis data, respectively, indicating that the R3 and R6 subpopulations are indeed LRMPs that directly mature into MKs via asymmetric cell division.

R4 cells in the LSK stain preferentially differentiate into mature MKs

To confirm that LSK surface marker expression can be used to predict MK cell development, the previous PCA projections of the no TPO Lin^- experimental group were re-colored according to expression of cKit and Sca1 (Figure 4A). R4 cells are distinguished from R2 cells, and both subpopulations are distinguishable from R1 cells. Furthermore, R4 cells with greater cKit expression extend further along PC1, while R4 cells with lower Sca1 expression do not extend as uniformly along PC2 (Figure 4B). Because area strongly correlates with PC1 (Supplementary Figure 13A), R4 cells with high cKit expression and corresponding low PC1 scores, are more likely to be larger in size compared to the other subpopulations. Similarly, because circularity is stongly negatively correlated with PC2 (Supplementary Figure 13B), R4 cells with low Sca1 expression and corresponding low PC2 scores are more likely to be circular in shape. Performing the same analysis for the 50 ng/mL TPO Lin^- experimental group results in a stronger separation between the four subpopulations (Figure 4C). The trends observed in the no TPO Lin^- analysis are amplified with the addition of TPO. Taken together, these results indicate that R4 cells with similar cKit/Sca1 expression also have similar morphologies, and that culture with TPO exaggerates both the expression (or lack thereof) of these markers and the morphologies to which they correlate.

Because the PCA results indicate that the R4 subpopulation is independently grouped as $cKit^+/Sca1^-$ cells, and previous studies have identified that $cKit^{high}$ cells exhibit an intrinsic bias towards the MK lineage[18], this subpopulation was further explored for MK potential. To determine if R4 cells preferentially differentiate into mature MKs, fresh BM cells or BM cells precultured with 50ng/ml of TPO for 72 h were stained with LSK antibodies (Figure 4D-E). Similar to previous results, R1 cells did not proliferate in culture, while R2-R4 cells were able to self-renew and expand in response to TPO/c-Mpl signaling. R4 cells were sorted and recultured with TPO for an additional 48 h, then stained for CD45 and CD41 expression. Almost 36% of fresh BM cells were double-positive for CD45 and CD41 expression, while close to 80% of BM cells precultured with TPO were double-positive for CD45 and CD41 expression (Figure 4F-G), indicating preferential differentiation of R4 cells into mature MKs, particularly after pretreatment with TPO.

Discussion

In the past several years, the hierarchical model of hematopoiesis has been reconfigured, transitioning away from the linear, sequential development of terminally differentiated cells from HSCs towards one that includes branches, bypasses and feedback loops[2,3,5]. One of the most highly debated aspects of the continuously evolving hematopoietic organization is the origin and development of MK progenitors capable of giving rise to mature MKs. HSCs and MKs have many features in common; most appreciably are their mutual dependence on TPO/c-Mpl stimulation [1,5,7-11,19] and surface marker expression [20-24]. As a result, drawing the line between HSCs and MK progenitors to systematically study these cells has been exceedingly difficult, especially since recent evidence supports direct MK progenitor development from a subset of lineage-biased HSCs[18]. Current methods lack the ability to image cells while simultaneously assessing surface marker expression in response to culture conditions, ultimately limiting our analytical capabilities. The experiments performed here use the ImageStream®X Mark II Imaging Flow Cytometer to collect morphology data that is intrinsically coupled with surface marker expression to identify elusive cell types that respond to TPO/c-Mpl signaling at single-cell and whole-population resolution. Furthermore, PCA confirmed that these minimal panels produce subpopulations with distinct morphologies characteristic of MK progenitors. Taken together, these experiments and the corresponding PCA justify the use of minimal antibody panels and simple cell culture techniques to isolate fleeting MK progenitor populations.

Previous studies have identified MK progenitors in the adult mouse but, to our knowledge, this is the first report to use only LSK antibodies to quantify early enrichment of MK progenitors after stimulation with TPO. By culturing whole mouse BM with 50ng/ml of TPO for just 3 days, we observed enrichment of the cells of interest, which can be easily distinguished by morphological characteristics and surface marker expression. In summary, these R4 cells proliferate, increase in size and circularity, develop a distinctly unique texture, and upregulate cKit expression in response to TPO/c-Mpl signaling. The absence of lineage markers and expression of cKit indicates that these cells are still immature and capable of further proliferation. Furthermore, sorted R4 cells that have been previously exposed to TPO readily differentiate into mature MKs at ~80% efficiency when recultured. These culture methods serve as a novel procedure for isolating MK progenitors that may provide insight to the early stages of MK commitment and development. Applying these culture methods to the disease setting can provide us with a better understanding of the molecular mechanisms underlying atypical MK development to improve targeted therapies for diseases that arise from skewed cell fate decisions, in addition to informing alternative approaches to produce MKs in vitro when using genetic approaches to drive cell fate decisions[25,26].

Acknowledgements

We gratefully acknowledge the funding from the University of Utah startup funds, the National Science Foundation CAREER Program (CBET-1554017), the Office of Naval Research Young Investigator Program (N00014-16-1-3012), the National Institute of Health Trailblazer Award (1R21EB025413-01), and the NIH Ruth L. Kirschstein National Research Service Award NIH from the National Institute of Diabetes and Digestive and Kidney Diseases (5T32DK007115-42). This work was also supported by the University of Utah Flow Cytometry Facility, the University of Utah Fluorescence Microscopy Core Facility (1S10RR024761-01), the National Cancer Institute (5P30CA042014-24), and the National Center for Research of the National Institutes of Health (1S10RR026802-01). We thank Dr. Charles Voyton, Dr. Ken A. Christensen and Dr. Josh Andersen of the BYU Department of Chemistry & Cancer Discovery for use of the ImageStream®X Mark II Imaging Flow Cytometer. For technical expertise with imaging flow cytometry analyses, we thank Dr. Amal El-Mabhouh, Ph.D. – MilliporeSigma, USA.

Author contributions

LMB conducted all the experiments and analyzed the data. CPH performed the principal components analysis. TLD and JEM helped design experiments and analyze data. LMB, CPH, and TLD wrote and edited the manuscript.

Competing interests

The authors declare no competing financial interest.

Tables and Figures

Stain		Reagent	Specificity/Target Information	Brand	Clone	Cat/Ref#	Dilution
LSK Stain	Antibodies	PE-Cy™7 Rat anti-Mouse CD117	CD117, also known as cKit. cKit is a transmembrane tyrosine-kinase receptor expressed on hematopoietic progenitor cells and the cKit ligand, also known as stem cell factor (SCF), is a co-stimulatory factor for hematopoietic stem and progenitor cells.	BD Pharmingen	2B8	561681	1:133
		PE Rat Anti-Mouse Ly-6A/E	Ly-6E.1 and Ly-6A.2 of the Ly-6 family of GPI-linked surface proteins, collectively known as Sca1. Sca1 is expressed by hematopoietic stem cells, myeloid cells, and peripheral T and B cells.	BD Pharmingen	D7	562059	1:100
		APC Mouse Lineage Antibody Cocktail	CD3e, expressed on mouse T cells; CD11b, expressed on granulocytes, macrophages, dendritic cells, and natural killer cells; CD45R/B220, mainly expressed by the B cell lineage from early Pro-B to mature B cells; Ly-76, expressed on erythroid cells; Ly-6G and Ly-6C, expressed on neutrophils, eosinophils, and monocytes. Collectively known as Lineage Antibody Cocktail.	BD Pharmingen	N/A	51-9003632	1:50
	DNA Dye(s)	DAPI	DAPI binds to AT rich regions of DNA along the minor grove. The dye can be excited by a UV laser with an emission maximum of 461nm. DAPI is predominantly impermeant to live cells at low concentration allowing it to be used as a viability dye for unfixed cells.	Thermo Scientific	N/A	62247	1:1000
MK Stain	Antibodies	CD45 Monoclonal Antibody (30-F11), PE-Cyanine7, eBioscience™	CD45, also known as the leukocyte common antigen (LCA). CD45 is a receptor-type protein tyrosine phosphatase expressed on all cells of the blood lineage, excluding mature erythrocytes and platelets. CD45 is a regulator of Src-family kinases, making it crucial in lymphocyte development and antigen signaling.	eBioscience	30-F11	25-0451-81	1:133
		PE Rat Anti-Mouse CD41	CD41, also known as glycoprotein IIb (GP IIb), integrin α IIb chain, or integrin alpha 2b. CD41 associates with integrin β3 chain (GPIIIa or CD61) form a complex that binds to fibrinogen, fibronectin, vitronectin, von Willebrand Factor, and thrombospondin. CD41/CD61 complexes are expressed on platelets, megakaryocytes, and early hematopoietic progenitors.	BD Pharmingen	MWReg30	561850	1:100
		CD42d Monoclonal Antibody (1C2), APC, eBioscience™	CD42d, also known as platelet glycoprotein V (GPV). CD42d interacts with CD42a, CD42b, and CD42c to form the receptor for von Willebrand Factor and thrombin. CD42d expression is restricted to resting platelets and megakaryocytes.	eBioscience	1C2	17-0421-80	1:50
	DNA Dye(s)	Hoechst 33342 Solution	Hoechst 33342 preferentially binds to AT regions of DNA. The dye is a cell-permeable DNA stain that excited with a UV laser and emits blue fluorescence at 460 to 490nm. Hoechst 33342 is used to stain nuclei of eukaryotic cells.	BD Pharmingen	N/A	561908	1:1000
		eBioscience™ 7-AAD Viability Staining Solution	7-AAD binds to GC-rich regions of DNA. The dye is a fluorescent intercalator that undergoes a spectral shift upon association with DNA. 7-AAD/DNA complexes can be excited by the 488nm laser and has an emission maxima of 647nm. 7-AAD is generally excluded from live cells, allowing it to be used as a viability dye for unfixed cells.	eBioscience	N/A	00-6993-50	1:100

Table 1. Reagents used in flow cytometry experiments. A comprehensive list of reagents used in each of the flow cytometry experiments, including reagent name, specificity/target information, brand, clone (if applicable), catalog and/or reference number, and dilution.

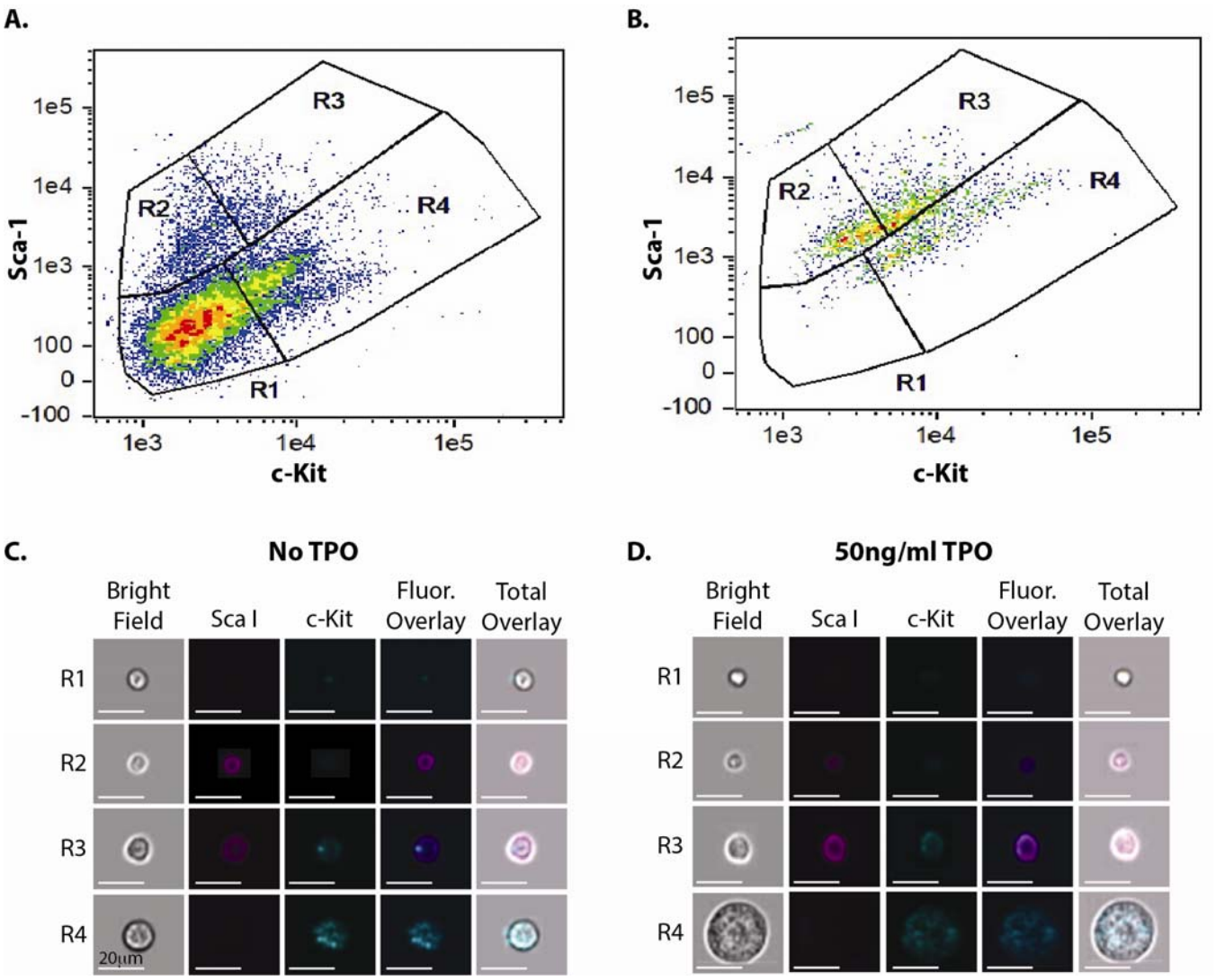


Figure 1. Image flow cytometry of BM cells stained with LSK antibodies identifies MK progenitors in the hematopoietic compartment. (A-B) To identify the major subpopulation in the hematopoietic compartment, flow cytometry plots of Sca1 intensity vs. cKit intensity for Lin^- cells before and after culture with 50ng/ml of TPO were generated. Flow cytometric events were divided into four distinct subpopulations: the R1 subpopulation ($Lin^-Sca1^-cKit^-$) is lost after culture with TPO; the R2 subpopulation ($Lin^-Sca1^+cKit^-$) self-renews and maintains high Sca1 expression in response to culture with TPO; the R3 subpopulation ($Lin^-Scat1^+cKit^+$, LSK) self-renews and maintains high levels of Sca1 and cKit expression in response to culture with TPO; the R4 subpopulation ($Lin^-Sca1^-cKit^+$) self-renews and maintains high levels of cKit expression in response to culture with TPO. (C-D) Representative image panels of each subpopulation before and after culture with 50ng/ml of TPO and stained with LSK antibodies. Scale bars, 20µm.

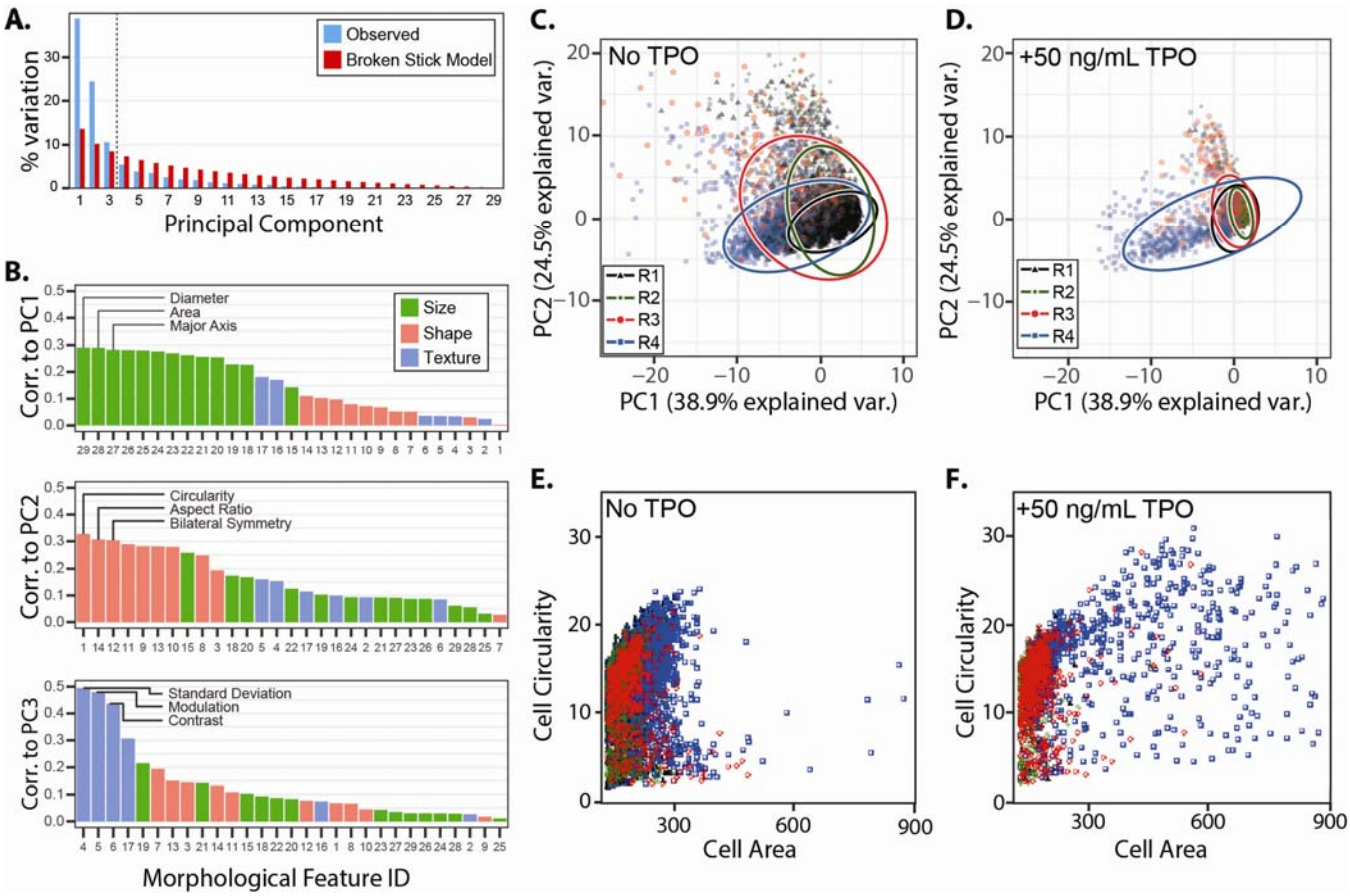


Figure 2. Morphology predicts MK progenitor phenotype. (A) Scree plot showing % eigenvalue for each principal component in the LSK morphology dataset and theoretical % variation under the broken-stick model. Dashed line indicates the boundary between interpretable and uninterpretable principal components. (B) Absolute contribution/correlation of each morphological feature to each principal component. Color indicates the type of feature; the top 3 most correlated features are listed (see supplementary Figure 11a for the features corresponding to each ID). (C-D) 1st and 2nd principal component projections of the 31,265 Lin$^-$ cells with No TPO (C) and the 10,926 Lin$^-$ cells with 50 ng/mL TPO. Colors indicate subpopulation. Ellipses contain 99% of the cells in each subpopulation (see supplementary figure 11 for remaining principal component projections). (E-F) Plots of brightfield circularity vs. area for all Lin$^-$ cells before and after culture with TPO. Virtually all cells in the R4 population are large, highly circular cells after culture with TPO. Black triangles = R1; green diamonds = R2; red circles = R3; blue squares = R4.

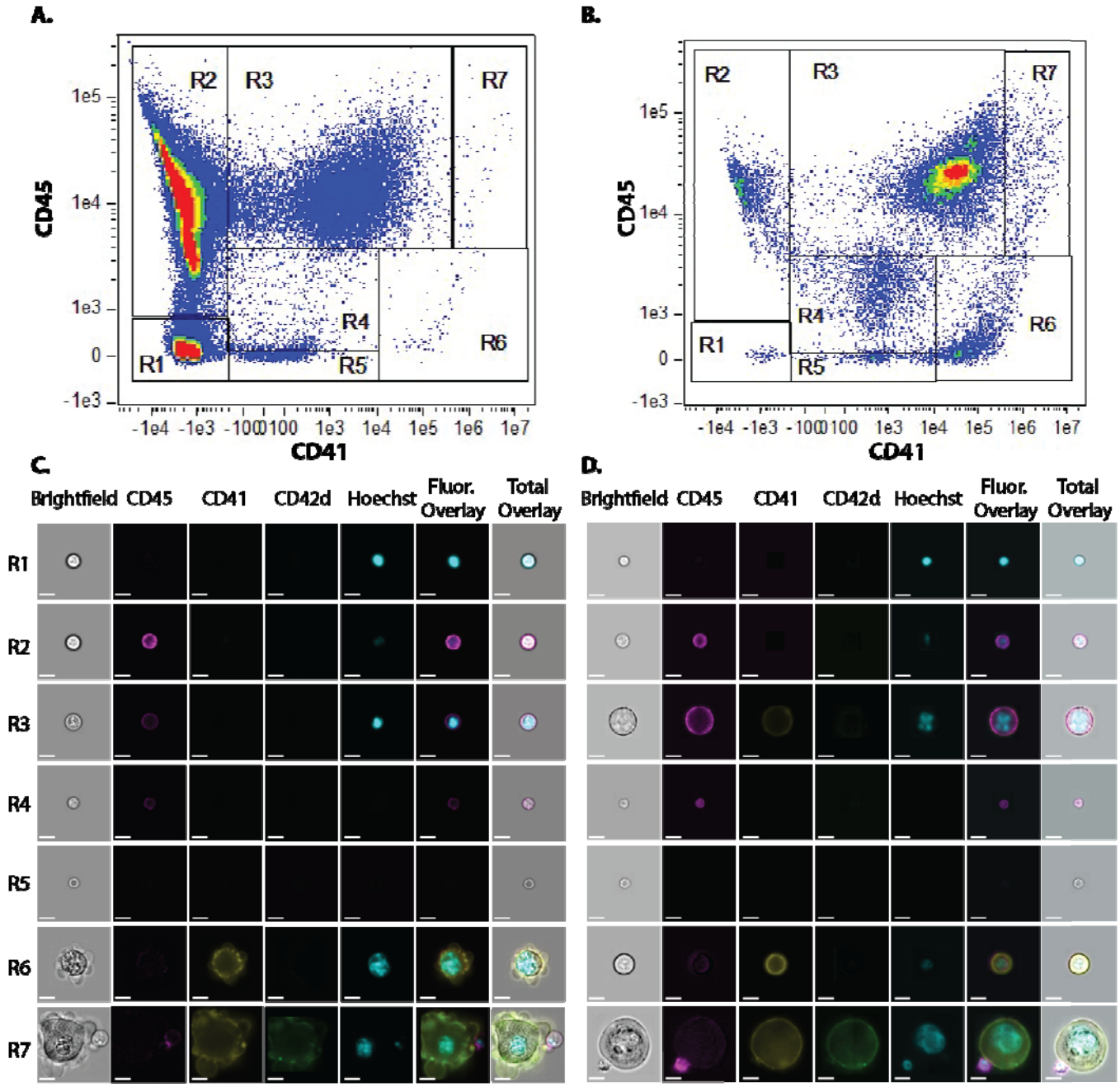
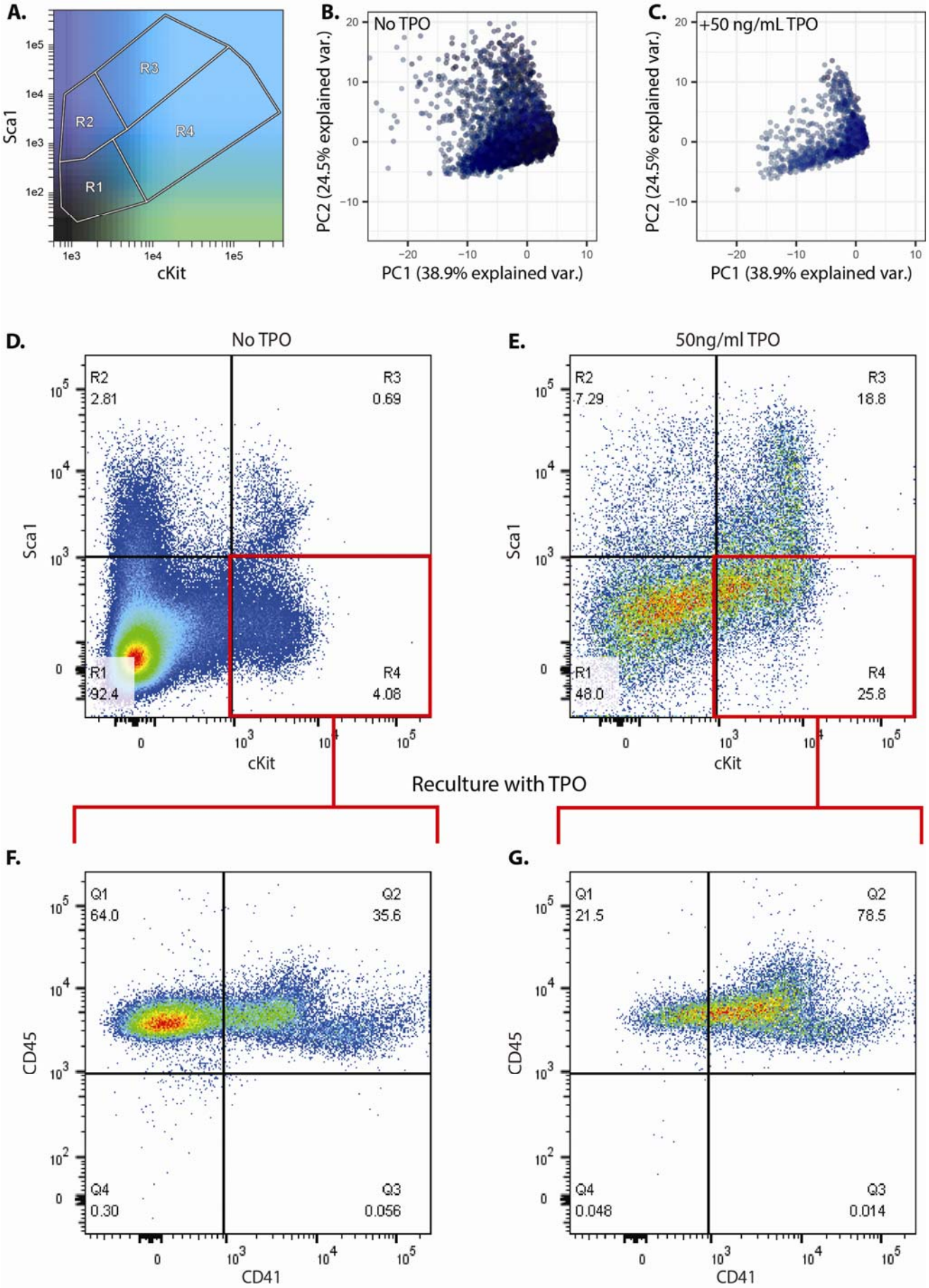


Figure 3.

Image flow cytometry of BM cells stained with MK antibodies phenotypically connects MK progenitors with mature MKs. (A-B) To demonstrate the effect of TPO on BM cells stained with MK antibodies, flow cytometry plots of CD45 intensity vs. CD41 intensity before and after culture with 50ng/ml of TPO were generated. Flow cytometric events were divided into seven distinct subpopulations: the R1 subpopulation ($CD45^{-}CD41^{-}$) is highly prevalent in unstimulated BM, but is lost after culture with TPO; the R2 subpopulation ($CD45^{mid/high}CD41^{-}$) expresses CD45, but does not self-renew when cultured with TPO; the R3 subpopulation ($CD45^{mid\rightarrow high}CD41^{low\rightarrow mid}$) upregulates CD45/CD41 expression and proliferates in response to culture with TPO; the R4 subpopulation

($CD45^{low}CD41^{low}$) is characterized by low surface marker expression and emerges only after culture with TPO; the R5 subpopulation ($CD45^{-}CD41^{low}$) expresses low levels of CD41 and proliferates modestly when cultured with TPO; the R6 subpopulation ($CD45^{-/low}CD41^{mid\rightarrow high}$) self-renews in response to culture with TPO; the R7 subpopulation ($CD45^{high}CD41^{high}$) expresses the highest levels of CD45/CD41 and is enriched after culture with TPO. (C-D) Representative image panels of each subpopulation before (C) and after culture with 50ng/ml of TPO (D) and stained with MK antibodies. R7 cells are the largest subpopulation, followed by the R3 subpopulation. Scale bars, 10μm.


A.
Sca1
cKit
R1
R2
R3
R4
B.
No TPO
PC2 (24.5% explained var.)
PC1 (38.9% explained var.)
C.
+50 ng/mL TPO
PC2 (24.5% explained var.)
PC1 (38.9% explained var.)
D.
No TPO
R2 2.81
R3 0.69
R1 92.4
R4 4.08
E.
50ng/ml TPO
R2 7.29
R3 18.8
R1 48.0
R4 25.8
Reculture with TPO
F.
Q1 64.0
Q2 35.6
Q4 0.30
Q3 0.056
G.
Q1 21.5
Q2 78.5
Q4 0.048
Q3 0.014
CD45
CD41

1

Figure 4.

R4 cells preferentially give rise to mature MKs, especially after preculture with TPO. (A) Colormap for subpopulation mask overlay. The dynamic ranges of the colormap were chosen to capture 99% of the Lin^- cells. (B-C) 1st and 2nd principal component projections of No TPO cells (B) and 50 ng/mL TPO cells (C) stained with LSK antibodies. Color indicates expression of the surface markers cKit and Sca1. R4 cells group by cKit surface marker expression, particularly after culture with TPO. (D-E) Flow cytometry plots of Sca1 vs. cKit expression of Lin^- cells before and after culture with TPO. R4 ($cKit^+$) cells were sorted (red boxed cells), then recultured with 50ng/ml TPO for 2 additional days. (F-G) Flow cytometry plots of CD45 vs. CD41 expression. Close to 36% of cell express high levels of CD45 and CD41 with No TPO preculture, while nearly 80% of cells express CD45 and CD41 when precultured with TPO.

1

References:

1. Seita J, Ema H, Ooehara J, et al. Lnk negatively regulates self-renewal of hematopoietic stem cells by modifying thrombopoietin-mediated signal transduction. *Proc Natl Acad Sci U S A*. 2007;104(7):2349-2354.
2. Chang Y, Bluteau D, Debili N, Vainchenker W. From hematopoietic stem cells to platelets. *J Thromb Haemost*. 2007;5 Suppl 1:318-327.
3. Seita J, Weissman IL. Hematopoietic stem cell: self-renewal versus differentiation. *Wiley Interdiscip Rev Syst Biol Med*. 2010;2(6):640-653.
4. Reya T, Morrison SJ, Clarke MF, Weissman IL. Stem cells, cancer, and cancer stem cells. *Nature*. 2001;414(6859):105-111.
5. Yamamoto R, Morita Y, Ooehara J, et al. Clonal analysis unveils self-renewing lineage-restricted progenitors generated directly from hematopoietic stem cells. *Cell*. 2013;154(5):1112-1126.
6. Paul F, Arkin Y, Giladi A, et al. Transcriptional Heterogeneity and Lineage Commitment in Myeloid Progenitors. *Cell*. 2015;163(7):1663-1677.
7. Buza-Vidas N, Antonchuk J, Qian H, et al. Cytokines regulate postnatal hematopoietic stem cell expansion: opposing roles of thrombopoietin and LNK. *Genes Dev*. 2006;20(15):2018-2023.
8. Kirito K, Fox N, Kaushansky K. Thrombopoietin stimulates Hoxb4 expression: an explanation for the favorable effects of TPO on hematopoietic stem cells. *Blood*. 2003;102(9):3172-3178.
9. Kaushansky K, Lok S, Holly RD, et al. Promotion of megakaryocyte progenitor expansion and differentiation by the c-Mpl ligand thrombopoietin. *Nature*. 1994;369(6481):568-571.
10. Machlus KR, Italiano JE, Jr. The incredible journey: From megakaryocyte development to platelet formation. *J Cell Biol*. 2013;201(6):785-796.
11. Yoshihara H, Arai F, Hosokawa K, et al. Thrombopoietin/MPL signaling regulates hematopoietic stem cell quiescence and interaction with the osteoblastic niche. *Cell Stem Cell*. 2007;1(6):685-697.
12. Nakeff A, Maat B. Separation of megakaryocytes from mouse bone marrow by velocity sedimentation. *Blood*. 1974;43(4):591-595.
13. Wilson A, Trumpp A. Bone-marrow haematopoietic-stem-cell niches. *Nat Rev Immunol*. 2006;6(2):93-106.
14. Team RC. R: A Language and Environment for Statistical Computing Vienna, Austria: R Foundation for Statistical Computing; 2018.
15. Zelený D. Analysis of community ecology data in R; 2017.
16. Jackson DA. Stopping Rules in Principal Components Analysis: A Comparison of Heuristical and Statistical Approaches. *Ecology*. 1993;74:2204-2214.
17. Wickham H. ggplot2: Elegant Graphics for Data Analysis. New York: Springer-Verlag New York; 2016.
18. Shin JY, Hu W, Naramura M, Park CY. High c-Kit expression identifies hematopoietic stem cells with impaired self-renewal and megakaryocytic bias. *J Exp Med*. 2014;211(2):217-231.
19. Huang H, Cantor AB. Common features of megakaryocytes and hematopoietic stem cells: what's the connection? *J Cell Biochem*. 2009;107(5):857-864.
20. Woolthuis CM, Park CY. Hematopoietic stem/progenitor cell commitment to the megakaryocyte lineage. *Blood*. 2016;127(10):1242-1248.
21. Pronk CJ, Rossi DJ, Mansson R, et al. Elucidation of the phenotypic, functional, and molecular topography of a myeloerythroid progenitor cell hierarchy. *Cell Stem Cell*. 2007;1(4):428-442.

22. Wang JF, Liu ZY, Groopman JE. The alpha-chemokine receptor CXCR4 is expressed on the megakaryocytic lineage from progenitor to platelets and modulates migration and adhesion. *Blood*. 1998;92(3):756-764.
23. Sugiyama T, Kohara H, Noda M, Nagasawa T. Maintenance of the hematopoietic stem cell pool by CXCL12-CXCR4 chemokine signaling in bone marrow stromal cell niches. *Immunity*. 2006;25(6):977-988.
24. Kiel MJ, Yilmaz OH, Iwashita T, Yilmaz OH, Terhorst C, Morrison SJ. SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells. *Cell*. 2005;121(7):1109-1121.
25. Deans TL, Cantor CR, Collins JJ. A tunable genetic switch based on RNAi and repressor proteins for regulating gene expression in mammalian cells. *Cell*. 2007;130(2):363-372.
26. Fitzgerald M, Gibbs C, Shimpi AA, Deans TL. Adoption of the Q Transcriptional System for Regulating Gene Expression in Stem Cells. *ACS Synth Biol*. 2017;6(11):2014-2020.