



OPEN

# High-throughput enrichment and isolation of megakaryocyte progenitor cells from the mouse bone marrow

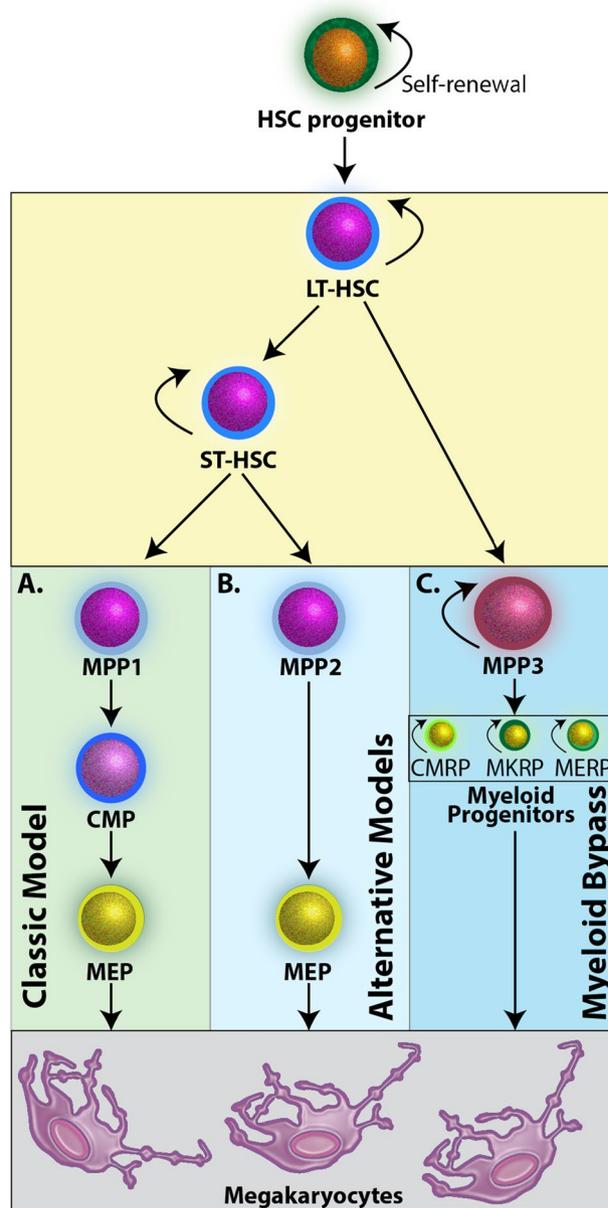
Lucas M. Bush<sup>1</sup>, Connor P. Healy<sup>1</sup>, James E. Marvin<sup>2</sup> & Tara L. Deans<sup>1</sup>✉

Megakaryocytes are a rare population of cells that develop in the bone marrow and function to produce platelets that circulate throughout the body and form clots to stop or prevent bleeding. A major challenge in studying megakaryocyte development, and the diseases that arise from their dysfunction, is the identification, classification, and enrichment of megakaryocyte progenitor cells that are produced during hematopoiesis. Here, we present a high throughput strategy for identifying and isolating megakaryocytes and their progenitor cells from a heterogeneous population of bone marrow samples. Specifically, we couple thrombopoietin (TPO) induction, image flow cytometry, and principal component analysis (PCA) to identify and enrich for megakaryocyte progenitor cells that are capable of self-renewal and directly differentiating into mature megakaryocytes. This enrichment strategy distinguishes megakaryocyte progenitors from other lineage-committed cells in a high throughput manner. Furthermore, by using image flow cytometry with PCA, we have identified a combination of markers and characteristics that can be used to isolate megakaryocyte progenitor cells using standard flow cytometry methods. Altogether, these techniques enable the high throughput enrichment and isolation of cells in the megakaryocyte lineage and have the potential to enable rapid disease identification and diagnoses ahead of severe disease progression.

Megakaryocytes (MKs) are important cells within the hematopoietic lineage because they produce platelets, the anucleate cells that facilitate healing. Platelets are released into the bloodstream from MKs where they play a central role in homeostasis, immune responses, and hemostasis<sup>1,2</sup>. Platelets also play an important role in thrombosis, which underlies heart attacks and stroke, making them the target of many pharmaceuticals and therapies. MKs differentiate from hematopoietic stem cells (HSCs)<sup>3</sup>, but the mechanisms of MK and subsequent platelet development are not fully understood, leading to difficulties when identifying the circumstances that cause disease<sup>4-7</sup>. MK maturation is characterized by an increase in cell size and DNA content to develop multilobed polyploid nuclei, an upregulation of MK-specific surface markers, and cytoplasmic remodeling to prepare for platelet production<sup>5,8</sup>. A significant challenge with studying MK development, and the diseases that arise from disrupted MK development, is the identification and classification of the progenitor cells that are produced from HSCs and subsequently give rise to MKs<sup>9</sup>. Although MKs are continually produced through adulthood, HSCs and MKs account for only about 0.01% and 0.05–0.1% of cells, respectively, in the mouse bone marrow. The frequency of MK progenitors is expected to be similarly infrequent which presents significant challenges for large-scale studies of this intermediate cell population<sup>10,11</sup>.

Recent studies suggest that MKs are derived from HSCs through multiple parallel lineage pathways<sup>12-15</sup>. While these lineage pathways are not fully elucidated, their diversity highlights the need to identify and isolate MK progenitor cells in order to understand MK development and maturation. However, because the bone marrow contains a heterogeneous population of cells, it is difficult to identify MK progenitor cells. Single-cell transcriptomics have provided biological snapshots that depict hematopoietic development as a continuum of differentiation, where a subpopulation of HSCs develops along a trajectory with an intrinsic bias to become platelets<sup>16,17</sup>. These quantitative approaches provide a high-resolution transcriptomic landscape of single timepoints throughout hematopoietic development that have helped to inform more standardized approaches for distinguishing individual cells in this lineage by using a combination of antibodies to recognize cell surface specific markers to

<sup>1</sup>Department of Biomedical Engineering, University of Utah, Salt Lake City, UT 84112, USA. <sup>2</sup>Flow Cytometry Core Facility, University of Utah Health Sciences Center, Salt Lake City, UT 84112, USA. ✉email: tara.deans@utah.edu



**Figure 1.** Models of MK differentiation from HSCs. **(A)** In the classic model, cell-fate decisions progress in a step-wise fashion where ST-HSCs differentiate toward an MPP1 cell population that has no self-renewal capacity but can further commit to CMP cells, MEP cells, and finally MKs. **(B)** The alternative model yields MEP cells directly from a non-self-renewing MPP2 cell type. **(C)** The myeloid bypass model suggests that LT-HSCs are primed to directly give rise to an MPP3 cell type capable of producing myeloid progenitor populations that all have the capacity to self-renew and maintain specific cell populations. *HSC* hematopoietic stem cell; *LT-HSC* long-term hematopoietic stem cell; *ST-HSC* short-term hematopoietic stem cell; *MPP* multipotent progenitor; *CMP* common myeloid progenitor; *MEP* megakaryocyte/erythroid progenitor; *MK* megakaryocyte; *CMRP* common myeloid repopulating progenitor; *MKRP* megakaryocyte repopulating progenitor; *MERP* megakaryocyte-erythroid repopulating progenitor.

identify cells based upon their immunofluorescent profiles when analyzed by flow cytometry. Using this method, cellular subpopulations can be sorted and re-cultured for further investigation.

In the mouse, HSCs are broadly defined by the Lin<sup>-</sup> Sca-1<sup>+</sup> c-Kit<sup>+</sup> (LSK) immunophenotype<sup>18,19</sup>, and have two distinct populations: (1) HSCs that are capable of long-term bone marrow reconstitution (LT-HSCs) and (2) HSCs capable of short-term hematopoietic reconstitution (ST-HSCs). Two different differentiation pathways, classical and alternative, have been proposed where ST-HSCs give rise to MK cells in which precursor cells progress through multiple differentiation states in a step-wise fashion before becoming a MK cell<sup>19</sup> (Fig. 1A,B). Notably, in neither of these pathways are the MK progenitor cells (MPP1, MPP2, respectively) thought to have any proliferative potential. In contrast, the recently proposed myeloid bypass model suggests that a population

of self-renewing multipotent progenitor cells (MPP3), which are formed directly from LT-HSCs, give rise to an MK repopulating progenitor (MKRP) population of cells<sup>15,20</sup> (Fig. 1C). The myeloid bypass model suggests that the formation and development of MKs may be dynamic and sensitive to the ever-changing biological environments present during inflammation, infection, and wound healing<sup>15,21</sup>. Identifying and isolating this MKRP cell population holds the potential to rapidly advance our understanding of MK development and maturation.

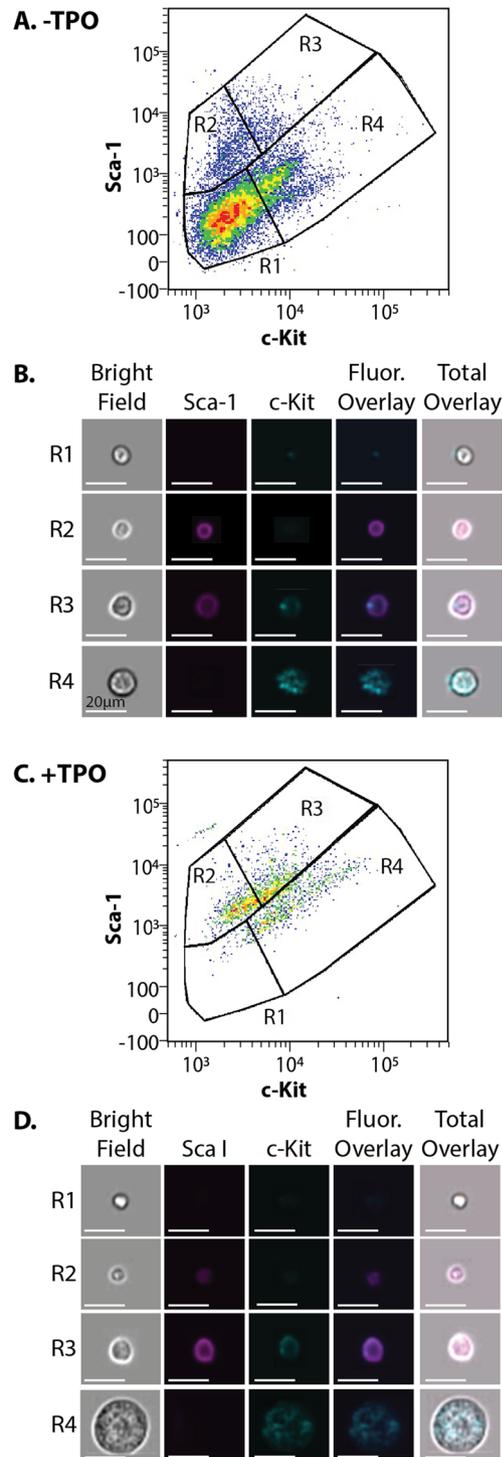
Because most cells develop distinct morphological characteristics at various stages of differentiation, we hypothesize that cellular morphology may be an important parameter to identify the MKRP cell population. Consistent with this hypothesis, commitment down the MK lineage is characterized by an increase in cell size, endoreplication that increases DNA content, and cytoplasmic remodeling to prepare for platelet production<sup>5,8</sup>. These individual morphological changes, along with fluorescently-labeled cell markers, can be observed using image flow cytometry, where an image of each individual cell is captured as it flows through the cytometer<sup>22</sup>. Each captured image contains data of multiple variables that can be used to characterize individual cells by cell size, shape, and membrane texture, as well as expression levels of fluorescently-labeled cell surface markers. These characteristics can be quantified using computational approaches like principal component analysis (PCA). PCA is a statistical method that groups highly correlated data, such as phenotypic characteristics, by converting a set of possibly correlated variables in a multivariate dataset into sets of values called principal components. Principal components allow multidimensional observations to be grouped based upon overall similarity in order to identify trends in datasets that would otherwise be intractable due to their high dimensionality. In short, this method captures variance in a multivariate dataset to identify the most correlated features that can be used to distinguish cells from different stages of lineage development. Here, we couple conventional antibody labeling, image flow cytometry, and PCA analysis to identify a repopulating population of MK progenitor cells. By coupling these analyses with a cell culture method that enriches for MKRP cells we can reliably identify these cells from heterogeneous bone marrow samples using both image flow and conventional flow cytometry. We envision that this will facilitate studies leading to a better understanding the development of MKRP cells, in addition to approaches for increasing MK and platelet production for disease treatment.

## Results

**Image flow cytometry analysis of bone marrow cells following TPO-mediated expansion.** MKRP cells often express canonical MK markers that are similar to other myeloid cells, making it difficult to distinguish early progenitor subpopulations from other hematopoietic cells without using large antibody panels. Since most cells develop distinct morphological differences when they are at various stages of cell fate, cell morphology is an additional parameter that can be used to identify MKRP cells. Image flow cytometry is an ideal tool for identifying cell populations and distinct stages of lineage development based upon these differences because it enables the simultaneous analysis of morphologic characteristics and cell surface marker expression.

To identify the cells that participate in the commitment to MK cells, mouse bone marrow cells were harvested and characterized by the expression of c-Kit, Sca-1, and lineage (Lin) surface markers. This immunofluorescence labeling strategy, known as LSK staining (Lin<sup>-</sup> Sca-1<sup>+</sup> c-Kit<sup>+</sup>), is historically used to identify HSCs that can be distinguished by analysis using flow cytometry. The TPO/c-Mpl signaling within the hematopoietic niche of the mouse bone marrow is essential for HSC survival, proliferation and differentiation into mature MKs<sup>5,8,12,23–26</sup>. Therefore, the TPO cytokine is commonly used for stimulating and differentiating HSCs to become MKs<sup>27</sup>. In all experimental groups, the – TPO condition are cells harvested directly from the bone marrow, then immediately prepared for analysis. In the + TPO condition, bone marrow cells are harvested and grown in TPO for 72 h prior to analysis. Cells in the – TPO condition were initially stained with LSK antibodies and run on the image flow cytometer (Fig. 2A). After culturing with TPO, the + TPO cells were also stained with LSK antibodies and run on the image flow cytometer. The + TPO cells experience a shift in Lin, Sca-1, and c-Kit expression (Fig. 2C), indicating these cells are primed for differentiation. In both the – TPO and + TPO conditions, four subpopulations of cells were defined (R1–R4) based on their Sca-1 and c-Kit expression. The boundaries for these subpopulations were defined by confirming cell surface labeling in the captured images. For example, all cells in the R2 gate were confirmed to be positive for only Sca-1, R3 for both Sca-1 and c-Kit, R4 for only c-Kit, and all cells in the R1 gate were negative for both Sca-1 and c-Kit labeling (Supplementary Figs. S1–S3). Comparing image files associated with the cells in each condition shows drastic morphological changes of the cells in the R4 subpopulation of cells (Fig. 2B,D). Cells in the R4 subpopulation are substantially larger and have increased in Sca-1 and c-Kit expression after exposure to TPO. These cells are present in both growth conditions and enrichment upon TPO treatment suggests that TPO may prime the cells in the R4 subpopulation for lineage commitment into MKs.

**Principal component analysis identifies MKRP cells.** Because HSCs have been shown to differentiate into MK cells in the presence of TPO, we aimed to use PCA to identify transitions in cell morphologies to locate the MKRP cells. To validate that distinct subpopulations of cells can be characterized morphologically, PCA was used to mathematically capture the variance in image datasets obtained from the image flow cytometer for both the – TPO and + TPO conditions stained with the LSK immunofluorescence panel. With PCA, the input variables contribute to each principal component such that the first principal component explains the largest possible variance in the dataset<sup>28</sup>. 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. This approach determined the most correlated morphological characteristics of each cell using the bright field images for – TPO cells (Fig. 2B) and + TPO (Fig. 2D), respectively. Twenty-nine morphological features corresponding to the shape, size, and texture of the cells in the dataset were considered in the principal component analysis (Supplementary Fig. S4A–E). To determine which principal component to keep, each principal component was compared to the broken stick model, which is a probabilistic model that determines the principal components needed to effectively inter-



**Figure 2.** Characterization of progenitor cell subpopulations of mouse bone marrow cells. **(A)** Cells were harvested from the mouse bone marrow ( $-$  TPO), stained with LSK antibodies, and run on an image flow cytometer. Cells were gated for  $\text{Lin}^-$  cells and the total population was divided into four distinct subpopulations (R1–R4). **(B)** As cells ran through the image flow cytometer, an image was captured of each cell. Representative images of the cells ( $-$  TPO) in each of each of the four subpopulations are represented. **(C)** Bone marrow cells grown in the presence of TPO ( $+$  TPO), stained with LSK antibodies and run on the image flow cytometer. Cells were divided into four distinct subpopulations (R1–R4). **(D)** Representative images of the cells ( $+$  TPO) panels of each of the four subpopulations are represented. Scale bars for all images, 20  $\mu\text{m}$ .

pret the data. When using the broken stick model, any principal component that exceeds the model indicates that principal component should be used for further analysis. This comparison showed that the – TPO dataset, PC1 (size) explained 60% of the morphological variation between cells while PC2 (shape) explained 14% of the morphological variation between cells, both exceeding the broken stick model (Fig. 3A)<sup>29</sup>. In the + TPO dataset, only PC1 (size) exceeded the broken stick model but PC1 also explained over 75% of the morphological variation in the dataset (Fig. 3B). Therefore, TPO treatment resulted in a population of cells that were more similar in shape, but more variable in size. In both datasets, PC1 correlated most highly with the size parameters: diameter and area (Fig. 3C,D). This suggests that the morphology of cells in both datasets varied most widely based on these two categories.

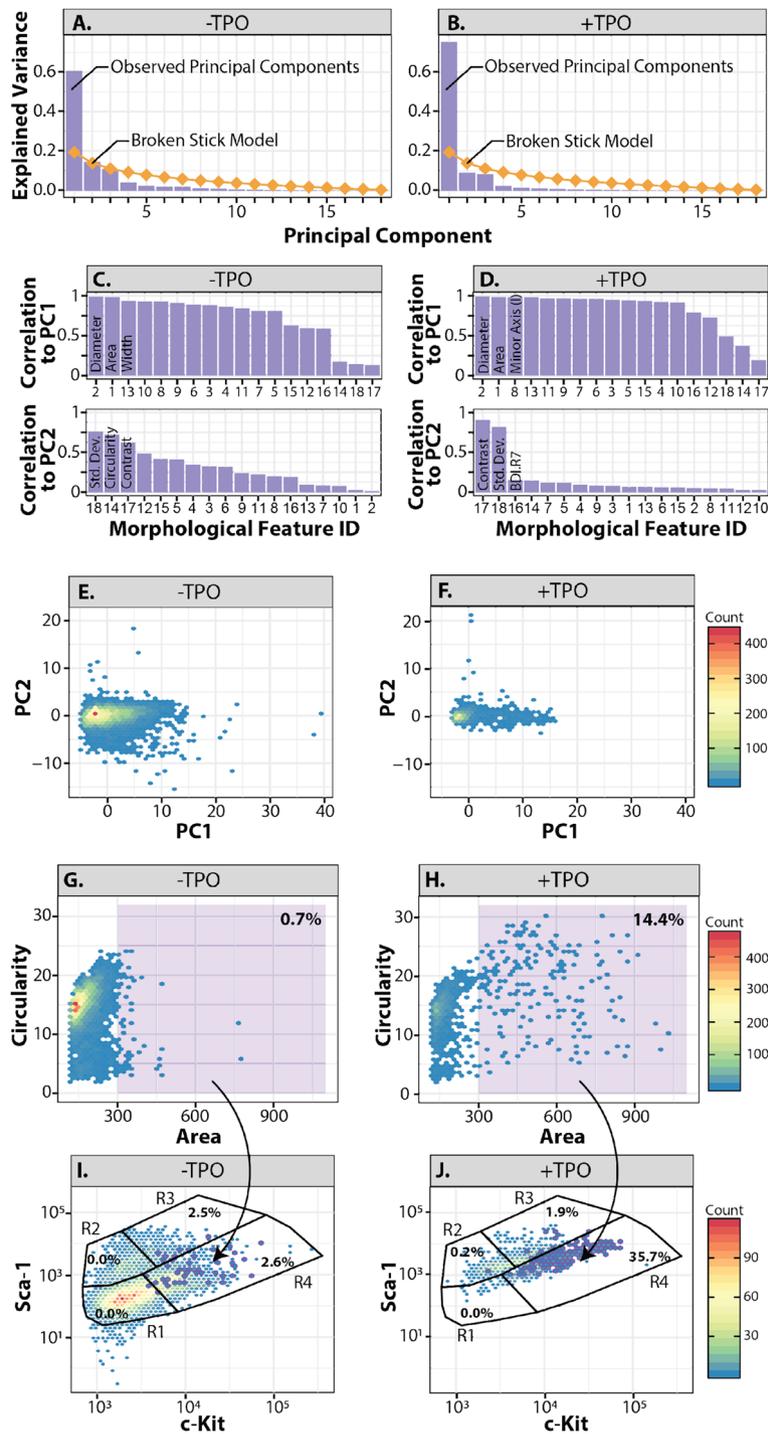
Projecting the principal component scores of each cell from the – TPO and + TPO conditions onto a plot shows that TPO treatment reduces morphological variation due to shape (PC2) while enhancing morphological variation due to size (PC1) (Fig. 3E,F). Furthermore, the remaining variation in the + TPO dataset is primarily due to differences in size (Fig. 3F). Overall, PCA revealed that cell size, particularly in + TPO populations, may be sufficient to distinguish cells in the R4 subpopulation. To test this, plots of cell area versus circularity were generated and gating was applied to cells with an area greater than 300 (Fig. 3G,H). By plotting the cells identified in this gate back onto the flow plots, we demonstrate that TPO treatment increases the size of a specific population of cells and that this population belongs primarily to the R4 subpopulation of cells identified in the LSK stain image flow cytometry dataset (Fig. 3I, J). Because MK maturation is marked by an increase in cell size<sup>5</sup>, and high levels of c-Kit expression indicates an intrinsic MK bias<sup>30</sup>, we hypothesized that the R4 subpopulation of cells is a LT-HSC-derived MKRP subset of cells capable of self-renewal and differentiating into mature MKs.

**Flow cytometry analysis of MKRP cells.** To investigate the PCA prediction that MKRP cells exist in the R4 subpopulation, cells from the – TPO and + TPO conditions were stained with antibodies to identify MKRP cells<sup>12,31,32</sup> using flow cytometry (Fig. 4A,B). Initially, a gate was drawn based on the Lin, Sca-1, and c-Kit expression to identify the R4 subpopulation of cells identified by the PCA analyses as being the potential MKRP population of cells in both growth conditions (Fig. 4A,B). To determine if this population of cells is undergoing MK differentiation and are of HSC origin, a gate was drawn to include CD9 positive and Thy1.1 negative cells. To confirm that these cells were committing to the MK lineage and not the B-cell lineage, a gate was drawn to include cells that were IL7R $\alpha$  negative and CD41 positive (Fig. 4A,B). Comparing the percentage of MKRP cells in each condition, we showed that a subset of the R4 subpopulation contains MKRP cells that can be enriched when cultured with TPO (Fig. 4C,D).

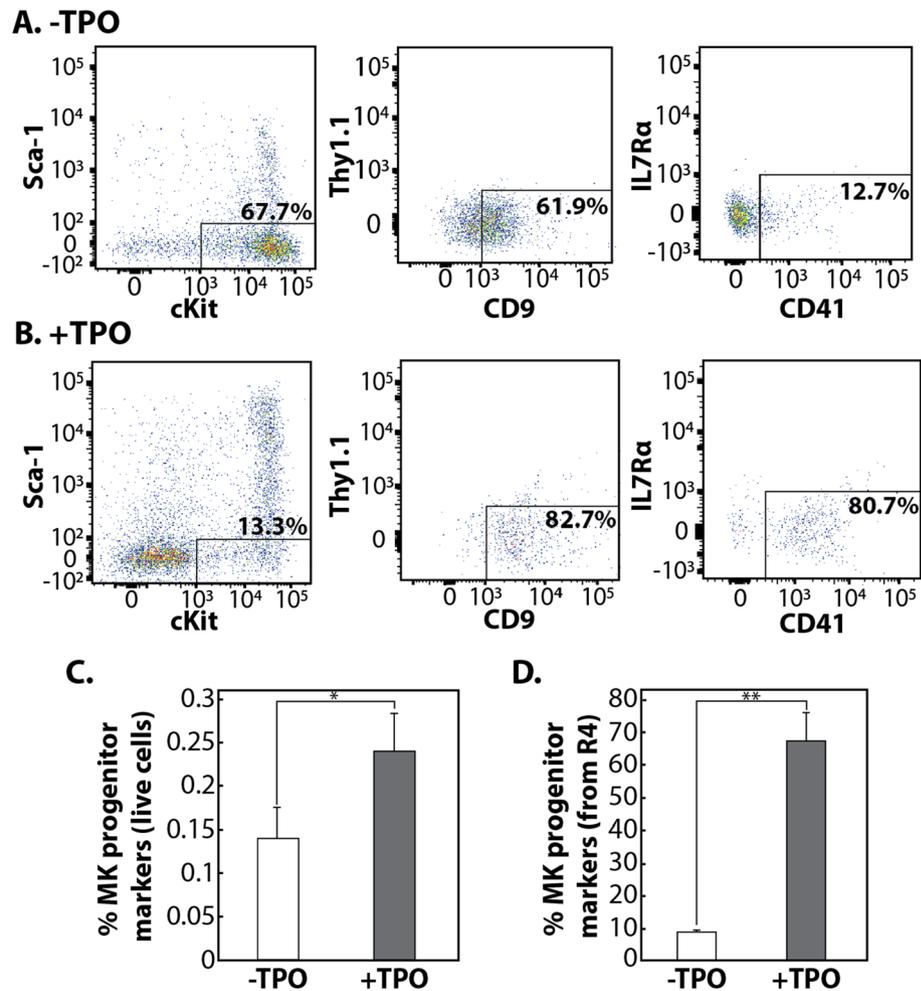
**Image flow cytometry analysis of MK lineage subpopulations.** To further classify the potential of the MKRP cell population, we investigated the development and maturation of these cells. To observe the transition from HSCs to MKs, mouse bone marrow cells from the – TPO and + TPO conditions were stained with MK antibodies (Supplementary Table S1) and image flow cytometry was performed to assess the expression of the MK surface markers (Fig. 5A,B)<sup>33</sup>. In both experimental cases, flow cytometry analysis revealed seven subpopulations of cells (R1–R7) with distinct CD45 and CD41 expression levels (Fig. 5A,B, Supplementary Fig. S5). In the + TPO condition, the cell subpopulations shifted (Supplementary Fig. S6) and increased in MK surface marker expression levels, indicating cells differentiating from a multipotent state HSC state to a more committed MK state.

Qualitative inspection of image files associated with individual flow cytometric events led to the observation that cells in the R7 subpopulation of cells were exceptionally large with the co-expression of MK surface markers and multilobed, polyploid nuclei within the confines of the cell membrane (Fig. 5C,D). To quantify this, the mean fluorescence intensity (MFI) of the three MK surface markers and the total area of the Hoechst 33342 signal were calculated for each subpopulation, revealing that the subpopulation of cells in the R7 gate expressed the highest levels of all three surface markers and had the largest nuclear area of the seven subpopulations (Supplementary Fig. S7). When plotting brightfield circularity versus bright field area, the R7 population was easily distinguishable from the other subpopulations in both culture conditions (Supplementary Fig. S8), further confirming the R7 subpopulation as mature MKs.

**Enrichment and isolation of a repopulating MKRP progenitor cell population.** The ability to identify and isolate the MKRP cell population in an easy and high throughput fashion would significantly improve studies on MK and platelet development. Since our PCA analysis strongly suggested that the R4 subpopulation of LSK cells potentially contain a repopulating MK progenitor cell population, we directly tested this using standard fluorescent activated cell sorting (FACS). The defined R4 subpopulation of cells from the – TPO condition were sorted using FACS, then grown with TPO for 48 h, and labeled with MK antibodies prior to analysis on the flow cytometer (Fig. 6A, – TPO). The defined R4 subpopulation of cells from the + TPO group were also sorted using FACS, and grown with TPO for an additional 48 h. The + TPO cells were labeled with MK antibodies and the expression of these cell surface markers were assessed on the flow cytometer (Fig. 6A, + TPO). In the – TPO group, it was found that 30% of cells were c-Kit<sup>+</sup> (Fig. 6B, D), consistent with our previous results. When the R4 cells from the – TPO group were sorted using FACS, then cultured with TPO for an additional 48 h, the cells continued to proliferate, however, only 30% of the cells differentiated into MKs (Fig. 6B,D). Also consistent with previous results, 30% of cells in the + TPO group remained c-Kit<sup>+</sup> after the first 72 h of culture with TPO (Fig. 6C,D). However, 80% of the R4 cells from the + TPO group differentiated into MKs (Fig. 6C,D) while continuing to proliferate (Fig. 6E). These data demonstrate that cells pre-treated with TPO had a much higher propensity to differentiate into mature MKs compared to the cells not pre-treated with TPO, and indi-



**Figure 3.** Morphology predicts MKRP phenotype using PCA. (A,B) Scree plots showing percent eigenvalue for each principal component in the LSK morphology dataset in the (A) – TPO and the (B) + TPO conditions. The theoretical percent variation under the broken-stick model is represented as a diamond dashed line, which indicates the boundary between interpretable and uninterpretable principal components. (C,D) The absolute contribution/correlation of each morphological feature to each morphological principal component in the (C) – TPO and (D) + TPO conditions. (E,F) Projections of LSK cells grown in the (E) – TPO and (F) + TPO conditions on the space defined by the first and second principal components. (G,H) Cell area vs. circularity of LSK cells in the (G) – TPO and the (H) + TPO conditions. A gate is included (shaded purple) to mark particularly large and circular cells with the percent of cells present in each gate. (I,J) PCA data mapped back onto the experimental flow data in the (I) – TPO and (J) + TPO conditions.



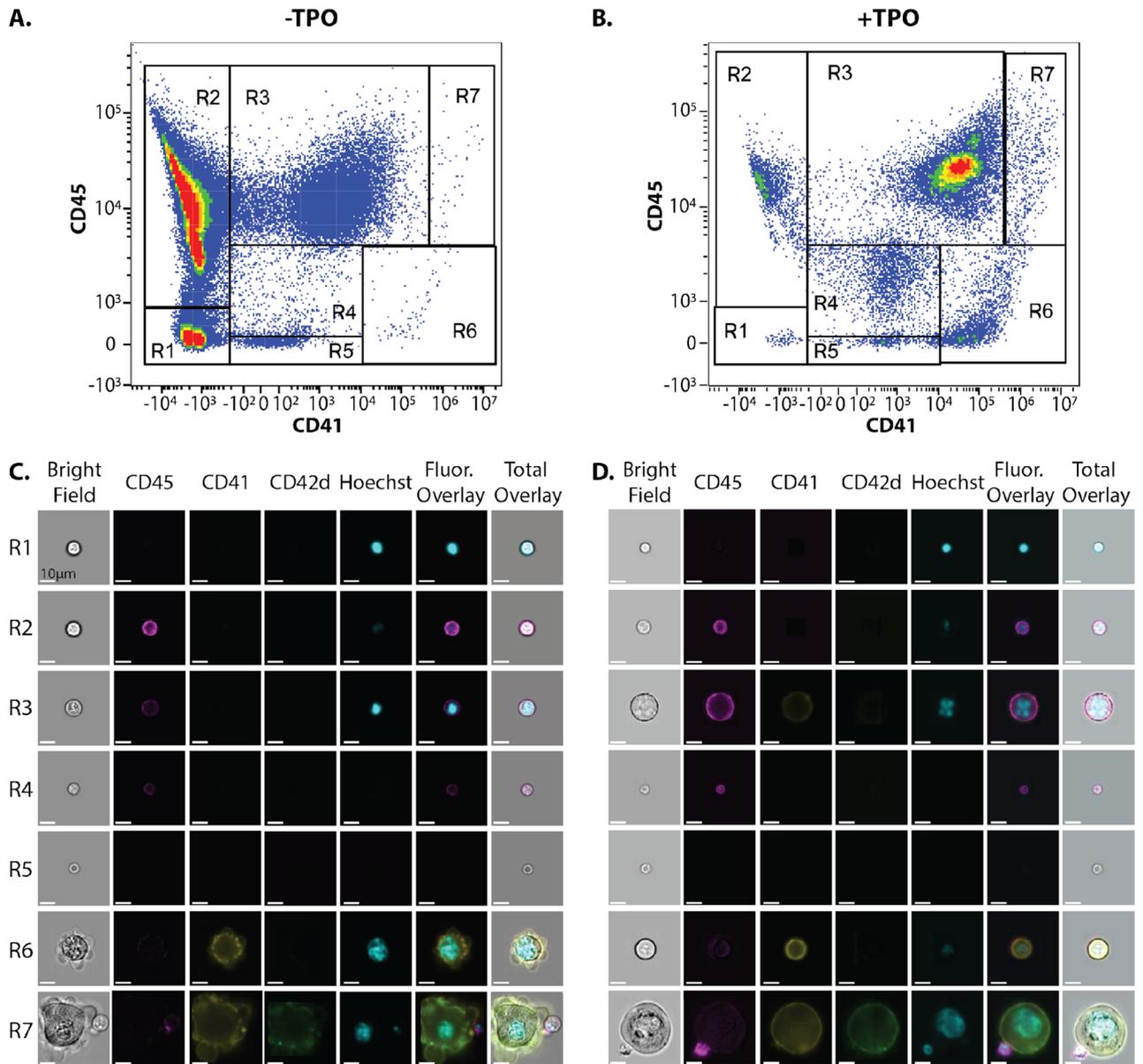
**Figure 4.** A subset of bone marrow cells are MKRP cells. (A) Representative flow cytometry panel of – TPO cells after gating on live, lineage negative cells. Boxes represent gates to identify the potential MKRP cells. (B) Representative flow cytometry panel of + TPO cells after gating on live, lineage negative cells. Boxes represent gates to identify the potential MKRP cells. (C) The percentage of live cells that stain for MKRP cell markers in the –TPO and + TPO groups, respectively. (D) The percentage of R4 cells (Lin<sup>–</sup>Sca1<sup>–</sup>cKit<sup>+</sup>) that stain for MKRP progenitor cell markers in the – TPO and + TPO groups, respectively. For (C) and (D) the percent of specified cells represents the average from three independent mice. The error bars represent the standard deviation of the mean. \* $p < 0.05$ ; \*\* $p < 0.005$ .

ates TPO pre-treatment as a method for the identification and enrichment of an MKRP cell population capable of self-renewal.

## Discussion

The classic model of hematopoiesis has been challenged over the last several years, transitioning away from the traditional, sequential differentiation pathway toward one that includes branches, bypassing and feedback loops<sup>12,13,18,25,34</sup>. One of the most highly debated aspects of the continuously evolving hematopoietic cell lineage organization is the origin and development of unipotent MKRP cells capable of giving rise to mature MKs. LT-HSCs and MKs have many features in common; most appreciably are their mutual dependence on TPO/c-Mpl stimulation and surface marker expression<sup>5,8,18,23–26,35</sup>. As a result, drawing the line between LT-HSCs, MKRP cells, and mature MKs to systematically study these cells has been exceedingly difficult. Previous methods have lacked the ability to image cells while simultaneously assessing the surface marker expression in response to culture conditions, ultimately limiting analytical capabilities. This study coupled cell surface antibody labeling and image flow cytometry to collect morphology data that is intrinsically coupled with surface marker expression to identify elusive MKRP cell types that can be further isolated to study the mechanisms of MK maturation and platelet development.

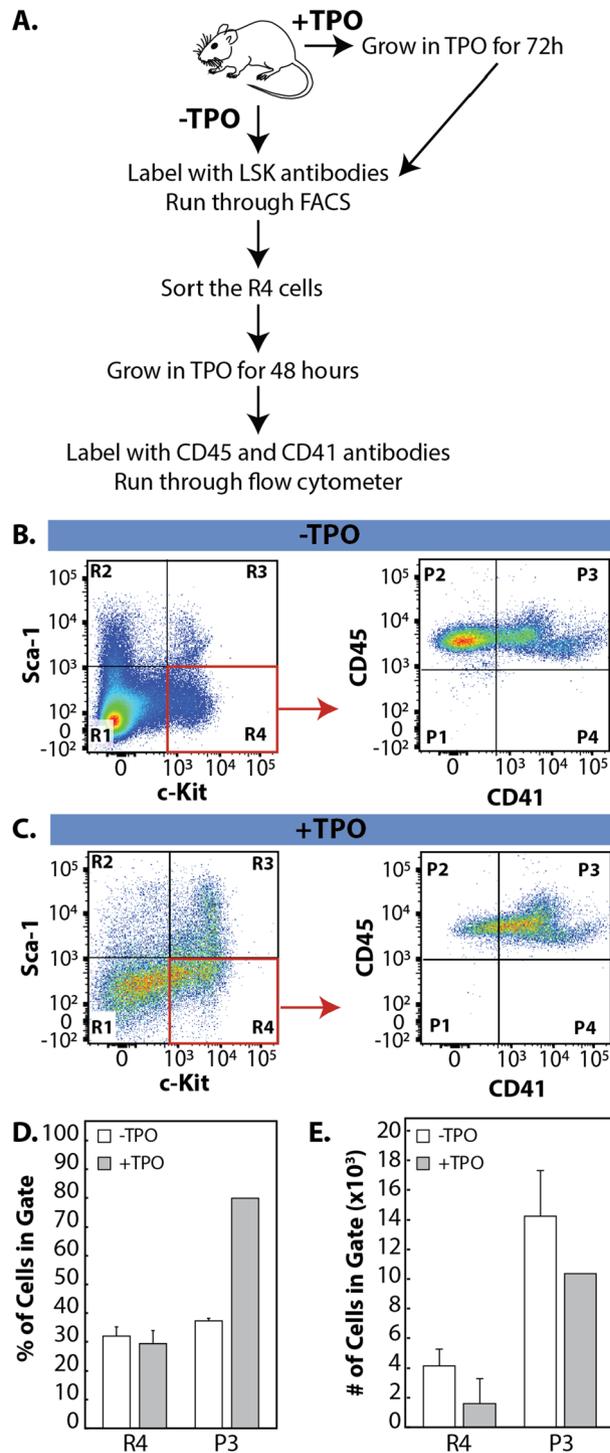
By combining mathematical approaches with experimental design, we identified, enriched, and isolated a population of MKRP cells. Subsequent characterization of these cells revealed that the c-Kit<sup>+</sup>Sca-1<sup>–</sup>Lin<sup>–</sup> fraction of cultured bone marrow cells was highly enriched for unipotent LT-HSC-derived MKRP cells as defined by the CD9<sup>+</sup>CD41<sup>+</sup>c-Kit<sup>+</sup>Sca-1<sup>–</sup>IL7Ra<sup>–</sup>Thy1.1<sup>–</sup>Lin<sup>–</sup> immunotype. Furthermore, sorted MKRP cells that had



**Figure 5.** Classification of MK lineage subpopulations. (A,B) Flow cytometry plots displaying CD45 and CD41 intensity of bone marrow cells stained with MK antibodies in the (A) – TPO and (B) +TPO conditions. Seven distinct subpopulations were identified (R1–R7). (C) Representative images of cells in the – TPO and (D) + TPO conditions. Scale bars, 10  $\mu$ m for all images.

been previously exposed to TPO retained the ability to self-renew and readily differentiate into mature MKs when re-cultured in TPO, similar to platelet-biased HSCs or stem-like MK progenitor cells identified in other studies<sup>20,21</sup>. Because these cells continue to proliferate when recultured with TPO after FACS, they can be considered a LT-HSC that has gone through the myeloid bypass pathway of differentiation to become MKRP cells.

Altogether, this study validates the use of minimal antibody panels and simple cell culture techniques for the enrichment and isolation of MKRP cells. Applying these methods to the disease setting may improve insights to the early stages of MK commitment and development. Additionally, these methods can be used to better understand the molecular mechanisms underlying atypical MK and platelet development to improve targeted therapies for diseases that arise from skewed cell fate decisions. Finally, this work sets the stage to better understand the role of dynamic gene expression patterns during the development of MKs and platelets to inform alternative approaches for directing stem cell fate that may include reprogramming progenitor cells with gene circuits to drastically increase the production of MKs and platelets *in vitro*<sup>36–39</sup>.



**Figure 6.** Identification of a repopulating MKRP cell population. **(A)** Overview of experimental design. Bone marrow cells were harvested and split into two populations: 1. - TPO cells were labeled with LSK antibodies, run on the FACS and the R4 subpopulation was sorted, then grown in TPO for 48 h before labeling with the MK cell markers, CD45 and CD41. 2. + TPO cells were grown in TPO for 72 h, labeled with LSK antibodies, run on the FACS, and the R4 subpopulation was sorted then regrown in TPO for an additional 48 h before staining with the MK cell markers CD45 and CD41. **(B)** Flow cytometry plot of - TPO cells stained with LSK antibodies. The defined subpopulations were identified (R1–R4). R4 cells were sorted (red box) and regrown in TPO for an additional 48 h, then stained for MK markers and run on the flow cytometer. New subpopulations were identified (P1–P4) by the expression of MK surface markers. **(C)** Flow cytometry plot of + TPO cells stained with LSK antibodies to identify the defined subpopulations (R1–R4). R4 cells were sorted (red box) and regrown in TPO for an additional 48 h then stained for MK markers and run on the flow cytometer. New subpopulations were identified (P1–P4) by the expression of MK surface markers. **(D)** Percentage of cells in the R4 (MKRP cells) and P3 (MKs) gates. **(E)** Number of cells in the R4 (MKRP cells) and P4 (MKs) gate. For **(D)** and **(E)** error bars represent the standard deviation of the mean from three independent bone marrow harvests from mice.

## Materials and methods

The methods and experimental protocols in this manuscript were carried out in accordance with ethics approval from the University of Utah under Institutional Animal Care and Use Committees (IACUC) approved guidelines. The use of animals and/or animal-derived materials in this study were carried out in accordance with relevant ARRIVE and University of Utah guidelines and regulations.

**Bone marrow dissection and cell culture.** Mice were euthanatized by CO<sub>2</sub> asphyxiation, followed by cervical dislocation based upon IACUC approved guidelines. Femurs were dissected from the mouse, excess tissue was removed, and the epiphyses were snapped 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-gauge needle and collected in a 35 mm 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 300 × g for 5 min. The cells were washed 2 × with HBSS (Gibco, #14025092) then counted using a hemocytometer. If cells were cultured before analysis, 1 × 10<sup>7</sup> cells were incubated in IMDM + 10% FBS supplemented with 50 ng/mL of recombinant murine thrombopoietin (TPO) (PeproTech, #315–14) in a T75 flask at 37 °C in 5% CO<sub>2</sub> for 72 h.

**Preparation for image flow cytometry.** After direct isolation from the bone marrow (– TPO) or culture with TPO for 72 h (+ TPO), cells were washed 2 × with PBS. For the LSK stain, cells were resuspended in PBS + 3% BSA (ThermoFisher Scientific, #BP1600-100) with the addition of PE-Cy7 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 2 × with PBS and resuspended at a final concentration of 2 × 10<sup>7</sup> 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 2 × with PBS then resuspended at a final concentration of 2 × 10<sup>7</sup> 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 Supplementary Table S1.

**Image flow cytometry data acquisition and analysis.** All experiments used the ImageStreamX 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. 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. Gating strategies used in all experiments are described in the Supplemental Information (Supplementary Figs. S1,S5).

**Image analysis and morphological feature extraction.** With the LSK stain dataset, brightfield images of lineage-negative cells from each experimental group (– TPO and + TPO, respectively) 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 Sca-1 and c-Kit expression. For the surface marker expression analysis, each cell was assigned its corresponding Sca-1 and c-Kit expression level instead of a categorical variable.

**Principal component analysis.** The variance of all twenty-nine morphological features was computed and those with variances less than 0.1 were omitted from further analysis (Supplementary Fig. S4). These features were aspect ratio intensity, aspect ratio, compactness, elongatedness, lobe count, shape ratio, symmetry 2–4, modulation, and spot count. The remaining eighteen morphological features from 31,265 observations in the – 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 dataset were computed using R's **prcomp** function<sup>40</sup>. The number of principal components needed to describe each dataset was determined using the broken stick model and David Zelený's **evplot** function<sup>41,42</sup>. 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 two 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<sup>35</sup>. These steps were repeated for

the + TPO dataset to generate the biplots (Fig. 3 E, F). For the surface marker expression bi-plots, the color of each observation was determined based on the relative expression of the surface markers c-Kit (green) and Sca-1 (blue) (Fig. 6, Supplementary Fig. S4). The dynamic range of this color map was defined to capture 99% of the observations. These same steps were repeated to generate the principal components for the 10,926 observations from the + TPO dataset.

### Fluorescence activated cell sorting (FACS), re-culture, and antibody staining of the MKRP sub-population.

For the – TPO group, cells were directly isolated from the mouse bone marrow and prepared for FACS using LSK antibodies as described above. For the + TPO group, cells were isolated from the mouse bone marrow and cultured with 50 ng/mL of TPO for 72 h, then prepared for FACS using LSK antibodies as described above. Cells were then sorted into IMDM + 10% FBS with the addition of 50 ng/mL TPO using the BD FACSAria III, 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.

### Data availability

Data in used in the PCA in this manuscript were obtained from the ImageStream X Mark II Imaging flow cytometer, as detailed in the Methods section. We will freely share data and code upon request.

Received: 5 November 2020; Accepted: 17 March 2021

Published online: 15 April 2021

### References

- Laffont, B. *et al.* Activated platelets can deliver mRNA regulatory Ago2\* microRNA complexes to endothelial cells via microparticles. *Blood* **122**, 253–261. <https://doi.org/10.1182/blood-2013-03-492801> (2013).
- Leiter, O. & Walker, T. L. Platelets: The missing link between the blood and brain?. *Prog. Neurobiol.* **183**, 101695. <https://doi.org/10.1016/j.pneurobio.2019.101695> (2019).
- Nishikii, H., Kurita, N. & Chiba, S. The road map for megakaryopoietic lineage from hematopoietic stem/progenitor cells. *Stem Cells Transl. Med.* **6**, 1661–1665. <https://doi.org/10.1002/sctm.16-0490> (2017).
- Kuter, D. J., Beeler, D. L. & Rosenberg, R. D. The purification of megapoietin: A physiological regulator of megakaryocyte growth and platelet production. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 11104–11108 (1994).
- Machlus, K. R. & Italiano, J. E. Jr. The incredible journey: From megakaryocyte development to platelet formation. *J. Cell Biol.* **201**, 785–796. <https://doi.org/10.1083/jcb.201304054> (2013).
- Patel, S. R., Hartwig, J. H. & Italiano, J. E. Jr. The biogenesis of platelets from megakaryocyte proplatelets. *J. Clin. Investig.* **115**, 3348–3354. <https://doi.org/10.1172/JCI26891> (2005).
- Thon, J. N. & Italiano, J. E. Platelet formation. *Semin. Hematol.* **47**, 220–226. <https://doi.org/10.1053/j.seminhematol.2010.03.005> (2010).
- Kaushansky, K. *et al.* Promotion of megakaryocyte progenitor expansion and differentiation by the c-Mpl ligand thrombopoietin. *Nature* **369**, 568–571. <https://doi.org/10.1038/369568a0> (1994).
- Xavier-Ferruccio, J. & Krause, D. S. Concise review: Bipotent megakaryocytic-erythroid progenitors: Concepts and controversies. *Stem Cells* **36**, 1138–1145. <https://doi.org/10.1002/stem.2834> (2018).
- Nakeff, A. & Maat, B. Separation of megakaryocytes from mouse bone marrow by velocity sedimentation. *Blood* **43**, 591–595 (1974).
- Wilson, A. & Trumpp, A. Bone-marrow haematopoietic-stem-cell niches. *Nat. Rev. Immunol.* **6**, 93–106. <https://doi.org/10.1038/nri1779> (2006).
- Yamamoto, R. *et al.* Clonal analysis unveils self-renewing lineage-restricted progenitors generated directly from hematopoietic stem cells. *Cell* **154**, 1112–1126. <https://doi.org/10.1016/j.cell.2013.08.007> (2013).
- Paul, F. *et al.* Transcriptional heterogeneity and lineage commitment in myeloid progenitors. *Cell* **163**, 1663–1677. <https://doi.org/10.1016/j.cell.2015.11.013> (2015).
- Woolthuis, C. M. & Park, C. Y. Hematopoietic stem/progenitor cell commitment to the megakaryocyte lineage. *Blood* **127**, 1242–1248. <https://doi.org/10.1182/blood-2015-07-607945> (2016).
- Noetzli, L. J., French, S. L. & Machlus, K. R. New insights into the differentiation of megakaryocytes from hematopoietic progenitors. *Arterioscler. Thromb. Vasc. Biol.* **39**, 1288–1300. <https://doi.org/10.1161/ATVBAHA.119.312129> (2019).
- Grover, A. *et al.* Single-cell RNA sequencing reveals molecular and functional platelet bias of aged haematopoietic stem cells. *Nat. Commun.* **7**, 11075. <https://doi.org/10.1038/ncomms11075> (2016).
- Laurenti, E. & Gottgens, B. From haematopoietic stem cells to complex differentiation landscapes. *Nature* **553**, 418–426. <https://doi.org/10.1038/nature25022> (2018).
- Chang, Y., Bluteau, D., Debili, N. & Vainchenker, W. From hematopoietic stem cells to platelets. *J. Thromb. Haemostasis* **5**(Suppl 1), 318–327. <https://doi.org/10.1111/j.1538-7836.2007.02472.x> (2007).
- Seita, J. & Weissman, I. L. Hematopoietic stem cell: self-renewal versus differentiation. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **2**, 640–653. <https://doi.org/10.1002/wsbm.86> (2010).
- Sanjuan-Pla, A. *et al.* Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. *Nature* **502**, 232–236. <https://doi.org/10.1038/nature12495> (2013).
- Haas, S. *et al.* Inflammation-induced emergency megakaryopoiesis driven by hematopoietic stem cell-like megakaryocyte progenitors. *Cell Stem Cell* **17**, 422–434. <https://doi.org/10.1016/j.stem.2015.07.007> (2015).
- Buckman, C. *et al.* High throughput, parallel imaging and biomarker quantification of human spermatozoa by ImageStream flow cytometry. *Syst. Biol. Reprod. Med.* **55**, 244–251. <https://doi.org/10.3109/19396360903056224> (2009).
- Buza-Vidas, N. *et al.* Cytokines regulate postnatal hematopoietic stem cell expansion: Opposing roles of thrombopoietin and LNK. *Genes Dev.* **20**, 2018–2023. <https://doi.org/10.1101/gad.385606> (2006).
- Kirito, K., Fox, N. & Kaushansky, K. Thrombopoietin stimulates Hoxb4 expression: an explanation for the favorable effects of TPO on hematopoietic stem cells. *Blood* **102**, 3172–3178. <https://doi.org/10.1182/blood-2003-03-0944> (2003).
- Seita, 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.* **104**, 2349–2354. <https://doi.org/10.1073/pnas.0606238104> (2007).
- Yoshihara, H. *et al.* Thrombopoietin/MPL signaling regulates hematopoietic stem cell quiescence and interaction with the osteoblastic niche. *Cell Stem Cell* **1**, 685–697. <https://doi.org/10.1016/j.stem.2007.10.020> (2007).

27. Ng, A. P. *et al.* Characterization of thrombopoietin (TPO)-responsive progenitor cells in adult mouse bone marrow with in vivo megakaryocyte and erythroid potential. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 2364–2369. <https://doi.org/10.1073/pnas.1121385109> (2012).
28. Attema, J. L. *et al.* Epigenetic characterization of hematopoietic stem cell differentiation using miniChIP and bisulfite sequencing analysis. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 12371–12376. <https://doi.org/10.1073/pnas.0704468104> (2007).
29. Wang, J. F., Liu, Z. Y. & Groopman, J. E. The alpha-chemokine receptor CXCR4 is expressed on the megakaryocytic lineage from progenitor to platelets and modulates migration and adhesion. *Blood* **92**, 756–764 (1998).
30. Shin, J. Y., Hu, W., Naramura, M. & Park, C. Y. High c-Kit expression identifies hematopoietic stem cells with impaired self-renewal and megakaryocytic bias. *J. Exp. Med.* **211**, 217–231. <https://doi.org/10.1084/jem.20131128> (2014).
31. Weiss-Gayet, M., Starck, J., Chaabouni, A., Chazaud, B. & Morle, F. Notch stimulates both self-renewal and lineage plasticity in a subset of murine CD9High committed megakaryocytic progenitors. *PLoS ONE* **11**, e0153860. <https://doi.org/10.1371/journal.pone.0153860> (2016).
32. Nakorn, T. N., Miyamoto, T. & Weissman, I. L. Characterization of mouse clonogenic megakaryocyte progenitors. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 205–210. <https://doi.org/10.1073/pnas.262655099> (2003).
33. Hashimoto, K. *et al.* Distinct hemogenic potential of endothelial cells and CD41+ cells in mouse embryos. *Dev. Growth Differ.* **49**, 287–300. <https://doi.org/10.1111/j.1440-169X.2007.00925.x> (2007).
34. Reya, T., Morrison, S. J., Clarke, M. F. & Weissman, I. L. Stem cells, cancer, and cancer stem cells. *Nature* **414**, 105–111. <https://doi.org/10.1038/35102167> (2001).
35. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* (Springer-Verlag, 2016).
36. Deans, T. L., Cantor, C. R. & Collins, J. J. A tunable genetic switch based on RNAi and repressor proteins for regulating gene expression in mammalian cells. *Cell* **130**, 363–372. <https://doi.org/10.1016/j.cell.2007.05.045> (2007).
37. Fitzgerald, M., Gibbs, C., Shimpi, A. A. & Deans, T. L. Adoption of the Q transcriptional system for regulating gene expression in stem cells. *ACS Synth. Biol.* **6**, 2014–2020. <https://doi.org/10.1021/acssynbio.7b00149> (2017).
38. Healy, C. P. & Deans, T. L. Genetic circuits to engineer tissues with alternative functions. *J Biol Eng* **13**, 39. <https://doi.org/10.1186/s13036-019-0170-7> (2019).
39. MacDonald, I. C. & Deans, T. L. Tools and applications in synthetic biology. *Adv. Drug Deliv. Rev.* **105**, 20–34. <https://doi.org/10.1016/j.addr.2016.08.008> (2016).
40. Team, R. C. R. *A Language and Environment for Statistical Computing* <http://www.R-project.org/> (2018).
41. Zelený, D. *Analysis of community ecology data in R*, <https://www.davidzeleny.net/anadat-r/doku.php/en:numecolr:evplot> (2017).
42. Jackson, D. A. Stopping rules in principal components analysis: A comparison of heuristical and statistical approaches. *Ecology* **74**, 2204–2214 (1993).

## Acknowledgements

This work was funded by the National Science Foundation CAREER Program (CBET-1554017), the Office of Naval Research Young Investigator Program (N00014-16-1-3012), the National Institutes of Health Trailblazer Award (1R21EB025413-01), and the National Institutes of Health Director's New Innovator Award (1DP2CA250006-01) awarded to T.L.D. We would also like to acknowledge funding from the Hematology Training Grant (5T32DK007115) awarded to L.M.B. Finally, T.L.D. would like to acknowledge support by the University of Utah Flow Cytometry Facility in addition to the National Cancer Institute through Award Number 5P30CA042014-24. Research reported in this publication was supported by the National Center for Research Resources of the National Institutes of Health under Award Number 1S10RR026802-01. We are also grateful for the helpful discussions about PCA with Dr. Jeff Weiss (U. Utah).

## Author contributions

L.M.B. conducted all of the experiments. C.H.P. performed the principal components analysis. L.M.B., J.E.M., and T.L.D. designed the experiments and analyzed the data. L.M.B., C.H.P., and T.L.D. wrote the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-87681-2>.

**Correspondence** and requests for materials should be addressed to T.L.D.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021