Clonal tracking of erythropoiesis in rhesus macaques

Xing Fan,1 Chuanfeng Wu,1 Lauren L. Truitt,1 Diego A. Espinoza,1,2 Stephanie Sellers,1 Aylin Bonifacino,1 Yifan Zhou,1,3 Stefan F. Cordes,1 Allen Krouse,1 Mark Metzger,1 Robert E. Donahue,1 Rong Lu4 and Cynthia E. Dunbar4

1Translational Stem Cell Biology Branch, National Heart, Lung, and Blood Institute, National Institute of Health, Bethesda, MA, USA; 2Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA; 3Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, UK and 4Eli and Edythe Broad Center for Regenerative Medicine and Stem Cell Research, University of Southern California, Los Angeles, CA, USA

©2020 Ferrata Storti Foundation. This is an open-access paper. doi:10.3324/haematol.2019.231811

Received: July 8, 2019.
Accepted: October 3, 2019.
Pre-published: October 3, 2019.
Correspondence: CYNTHIA DUNBAR - dunbarc@nhlbi.nih.gov
Supplementary Materials

Supplementary methods

Barcoded library diversity

Diversity of each barcoded lentiviral library preparation was validated using Monte Carlo simulations of retrieved barcodes to determine the number of CD34+ target cells able to be transduced with each library preparation resulting in a greater than 95% probability that more than 95% of barcodes represent transduction of single engrafting cells (2, 5, 6).

Barcode retrieval

DNA was extracted from all cell collections with the DNeasy Blood & Tissue kit (Qiagen) and quantified by Qubit (Invitrogen). RNA from nucleated cells was extracted using RNAzol®RT (Molecular Research Center)(8). RNA from enucleated mature RBC and reticulocytes was extracted using PAXgene Blood RNA Tubes and PAXgene Blood RNA Kits (PreAnalytiX). RNA was reverse transcribed using the SuperScript™ IV First-Strand Synthesis System (Thermo Fisher). 200 ng (ZH33, ZG66, ZJ31, ZK22, JD76 and ZL40) or 500 ng (ZH19) DNA, or cDNA reverse transcribed from 12ng RNA underwent PCR with primers (Table S2) bracketing the barcode and multiplex sequencing as described (2, 6). Equal amounts of gel-purified barcode PCR product from individual samples were pooled together for multiplex sequencing (Illumina HiSeq 3000).

Sequencing data analysis

Data analysis, Pearson correlations, Euclidean distances, P values, plot generation and statistical analyses were performed using R (Foundation for Statistical Computing) and Prism (GraphPad Software). Python code and associated R functions used for analyses are available on Github (www.github.com/dunbarlabNIH/).
References

1. Lu R, Neff NF, Quake SR, Weissman IL. Tracking single hematopoietic stem cells in vivo using high-throughput sequencing in conjunction with viral genetic barcoding. Nat Biotechnol. 2011 Oct 2;29(10):928-33.
2. Wu C, Li B, Lu R, Koelle SJ, Yang Y, Jares A, et al. Clonal tracking of rhesus macaque hematopoiesis highlights a distinct lineage origin for natural killer cells. Cell Stem Cell. 2014 Apr 3;14(4):486-99.
3. Uchida N, Washington KN, Hayakawa J, Hsieh MM, Bonifacino AC, Krouse AE, et al. Development of a human immunodeficiency virus type 1-based lentiviral vector that allows efficient transduction of both human and rhesus blood cells. J Virol. 2009 Oct;83(19):9854-62.
4. Uchida N, Hargrove PW, Lap CJ, Evans ME, Phang O, Bonifacino AC, et al. High-efficiency transduction of rhesus hematopoietic repopulating cells by a modified HIV1-based lentiviral vector. Mol Ther. 2012 Oct;20(10):1882-92.
5. Koelle SJ, Espinoza DA, Wu C, Xu J, Lu R, Li B, et al. Quantitative stability of hematopoietic stem and progenitor cell clonal output in rhesus macaques receiving transplants. Blood. 2017 Mar 16;129(11):1448-57.
6. Wu C, Espinoza DA, Koelle SJ, Potter EL, Lu R, Li B, et al. Geographic clonal tracking in macaques provides insights into HSPC migration and differentiation. J Exp Med. 2018 Jan 2;215(1):217-32.
7. Donahue RE, Kuramoto K, Dunbar CE. Large animal models for stem and progenitor cell analysis. Curr Protoc Immunol. 2005 Nov;Chapter 22:Unit 22A 1.
8. Chomczynski P WW, Kennedy A, Rymaszewski M, Mackey K. RNAzol® RT: a new single-step method for isolation of RNA. Nat Methods. 2010.
Table S1: Antibodies used for flow cytometric sorting of blood and bone marrow cell populations.

| Antigen | Conjugation | Vendor        | Catalog number | Clone       |
|---------|-------------|---------------|----------------|-------------|
| CD3     | APC-Cy7     | BD Pharmingen | 557757         | SP34-2      |
| CD3     | BV786       | BD Pharmingen | 563918         | SP34-2      |
| CD14    | Pacific blue| Invitrogen    | MHCD1428       | TUK4        |
| CD20    | BV650       | BD Pharmingen | 563780         | 2H7         |
| CD20    | APC Cy7     | BD Pharmingen | 335794         | L27         |
| CD34    | PE          | BD Pharmingen | 550761         | 563         |
| CD34    | Purified (for CD34 selection) | N/A | N/A | 12.8         |
| CD45    | BV510       | BD Horizon    | 563830         | D058-1283   |
| CD71    | PE          | GeneTex       | GTX43030       | DF1513      |
Table S2: Sequences of primers used for barcodes retrieval.

| Primer Names               | 5'----------------3' sequences                                      | Index          |
|----------------------------|-----------------------------------------------------------------|----------------|
| Universal reverse primer   | CAAGCAGAAGACGGCATACGAGATCGTACACGAGCTCTTCCGATCT                  |                |
| Barcode New Forward i501   | AATGATACGGGCAACCGAGATCTACACCTTACAGTTACACGCTCTTCCGATCT          | TCGCTTTA       |
| Barcode New Forward i502   | AATGATACGGGCAACCGAGATCTACACCTTACAGTTACACGCTCTTCCGATCT          | CTAGTACG       |
| Barcode New Forward i503   | AATGATACGGGCAACCGAGATCTACACCTTACAGTTACACGCTCTTCCGATCT          | TCTGCTCT       |
| Barcode New Forward i504   | AATGATACGGGCAACCGAGATCTACACCTTACAGTTACACGCTCTTCCGATCT          | GCTAGAGG       |
| Barcode New Forward i505   | AATGATACGGGCAACCGAGATCTACACCTTACAGTTACACGCTCTTCCGATCT          | AGGAGTCC       |
| Barcode New Forward i506   | AATGATACGGGCAACCGAGATCTACACCTTACAGTTACACGCTCTTCCGATCT          | CATGCCTA       |
| Barcode New Forward i507   | AATGATACGGGCAACCGAGATCTACACCTTACAGTTACACGCTCTTCCGATCT          | GCTAGAGG       |
| Barcode New Forward i508   | AATGATACGGGCAACCGAGATCTACACCTTACAGTTACACGCTCTTCCGATCT          | AGGAGTCC       |
| Barcode New Forward i509   | AATGATACGGGCAACCGAGATCTACACCTTACAGTTACACGCTCTTCCGATCT          | CATGCCTA       |
| Barcode New Forward i510   | AATGATACGGGCAACCGAGATCTACACCTTACAGTTACACGCTCTTCCGATCT          | GCTAGAGG       |
| Barcode New Forward i511   | AATGATACGGGCAACCGAGATCTACACCTTACAGTTACACGCTCTTCCGATCT          | AGGAGTCC       |
| Barcode New Forward i512   | AATGATACGGGCAACCGAGATCTACACCTTACAGTTACACGCTCTTCCGATCT          | CATGCCTA       |
| Barcode New Forward i513   | AATGATACGGGCAACCGAGATCTACACCTTACAGTTACACGCTCTTCCGATCT          | GCTAGAGG       |
| Barcode New Forward i514   | AATGATACGGGCAACCGAGATCTACACCTTACAGTTACACGCTCTTCCGATCT          | AGGAGTCC       |
| Barcode New Forward i515   | AATGATACGGGCAACCGAGATCTACACCTTACAGTTACACGCTCTTCCGATCT          | CATGCCTA       |
| Barcode New Forward i516   | AATGATACGGGCAACCGAGATCTACACCTTACAGTTACACGCTCTTCCGATCT          | GCTAGAGG       |
| Barcode New Forward i517   | AATGATACGGGCAACCGAGATCTACACCTTACAGTTACACGCTCTTCCGATCT          | GCTAGAGG       |
| Barcode New Forward i518   | AATGATACGGGCAACCGAGATCTACACCTTACAGTTACACGCTCTTCCGATCT          | GCTAGAGG       |
| Barcode New Forward i519   | AATGATACGGGCAACCGAGATCTACACCTTACAGTTACACGCTCTTCCGATCT          | GCTAGAGG       |
| Barcode New Forward i520   | AATGATACGGGCAACCGAGATCTACACCTTACAGTTACACGCTCTTCCGATCT          | GCTAGAGG       |
| Barcode New Forward i521   | AATGATACGGGCAACCGAGATCTACACCTTACAGTTACACGCTCTTCCGATCT          | GCTAGAGG       |
| Barcode New Forward i522   | AATGATACGGGCAACCGAGATCTACACCTTACAGTTACACGCTCTTCCGATCT          | GCTAGAGG       |
| Barcode New Forward i523   | AATGATACGGGCAACCGAGATCTACACCTTACAGTTACACGCTCTTCCGATCT          | GCTAGAGG       |
| Barcode New Forward i524   | AATGATACGGGCAACCGAGATCTACACCTTACAGTTACACGCTCTTCCGATCT          | GCTAGAGG       |
| Barcode New Forward i525   | AATGATACGGGCAACCGAGATCTACACCTTACAGTTACACGCTCTTCCGATCT          | GCTAGAGG       |
| Barcode New Forward i526   | AATGATACGGGCAACCGAGATCTACACCTTACAGTTACACGCTCTTCCGATCT          | GCTAGAGG       |
| Barcode New Forward i527   | AATGATACGGGCAACCGAGATCTACACCTTACAGTTACACGCTCTTCCGATCT          | GCTAGAGG       |
| Barcode New Forward i528   | AATGATACGGGCAACCGAGATCTACACCTTACAGTTACACGCTCTTCCGATCT          | GCTAGAGG       |
| Barcode New Forward i529   | AATGATACGGGCAACCGAGATCTACACCTTACAGTTACACGCTCTTCCGATCT          | GCTAGAGG       |
Supplementary Figure Legends

Figure S1: Purification strategies for HSPCs and specific hematopoietic lineages

(A) Representative purity check by flow cytometric analysis of BM CD34+ HSPCs from RM ZK22 15.5 months post-transplantation selected by using MACS beads.

(B) Schematic of flow cytometric gating strategy used for sorting lineage-specific cells from the BM samples (upper panel), and the FACS plots (lower panel) of ZL40 BM sample at 15m.

(C) Schematic of flow cytometric gating strategy used for sorting lineage-specific cells from the PB sample (upper panel), and the FACS plots (lower panel) of PB sample from ZL40 at 15m.

Figure S2: Lineage bias of contributing clones

Bar plots displaying the bias and relative size of each barcode. Each barcode is represented as a grey box with a black outline so that very small clones appear only as a black line. The categories in which clones fall are determined by comparing the percent contribution of the given lineage on the left to the maximum percent contribution of all other lineages (NRBC, Gr, Mono, CD34+, T, B without the given lineage). The positive sign (+) indicates bias towards the given lineage (i.e. present more in the NRBC lineage) and the negative sign (-) indicates bias away from the given lineage (i.e. present more in other lineages).

Figure S3: GFP expression in individual blood cell lineages

(A) The FACS plots of CD45 and GFP expression of the PB RBC and granulocytes (Gr) in both MSCV- and EF1a-driven barcode lentivirus transduced HSPC RM.
(B) Bar plot summarized the GFP% in RBCs and Gr in individual rhesus macaques. Different colors indicate different cell types.

**Figure S4: Nuclear red blood cell (NRBC) percentage in bone marrow samples**

The FACS plots show NRBCs’ sorting strategy of CD45-CD71+ in different RM’s bone marrow samples.
Figure S1

A

MACS purified CD34+ cells

ZK22 11.5m

BM

CD71

CD45

CD3

CD20

SSA

Mono

CD14

B

BM

ZL40 15m

C

PB

ZL40 15m

CD3

T

B

SSA

Mono

PB

CD20

CD14

Unstained

stained
Figure S2

ZK22 11.5m LBM
ZJ31 3.5m RBM
RQ3600 48m BM
ZH19 45.5m RBM

JD76 3.5m LBM
JD76 3.5m RBM
ZH33 46m LBM
ZL40 10.5m BM

ZL40 12m BM
ZL40 15.5m BM
ZK22 15.5m BM

CD71

CD45
Figure S3
