Full title:

Single cell analysis of bone marrow derived CD34+ cells from children with sickle cell disease and thalassemia

Short title:

Analysis of CD34+ cells in SCD and thalassemia

Peng Hua[1,2], Noemi Roy[3,6*], Josu de la Fuente[4*], Guanlin Wang[1], Supat Thongjuea[1], Kevin Clark[1], Anindita Roy[5], Bethan Psaila[1,3,6], Neil Ashley[1], Yvonne Harrington[4], Claus Nerlov[1], Suzanne M. Watt[2], Irene Roberts[1,5,6] and James O. J. Davies[1,3,6]

1. MRC Molecular Haematology Unit, MRC Weatherall Institute of Molecular Medicine, Radcliffe Department of Medicine, John Radcliffe Hospital, Oxford, UK.
2. Nuffield Division of Clinical Laboratory Sciences, Radcliffe Department of Medicine, John Radcliffe Hospital, Oxford, UK.
3. Department of Haematology, Oxford University NHS Foundation Trust, Oxford, UK
4. Department of Paediatrics, St. Marys Hospital, Imperial College Healthcare NHS Trust, London, UK.
5. Department of Paediatrics, University of Oxford, Childrens Hospital, John Radcliffe Hospital, Oxford, UK.
6. Oxford Biomedical Research Centre, University of Oxford

* These authors contributed equally.

Authors' contact emails:

Peng Hua: peng.hua@ndcls.ox.ac.uk

Noemi Roy: noemi@romahi.com

Josu de la Fuente: josu.delafuente@imperial.nhs.uk

Guanlin Wang: guanlin.wang@imm.ox.ac.uk

Supat Thongjuea: supat.thongjuea@ndcls.ox.ac.uk

Kevin Clark: kevin.clark@imm.ox.ac.uk

Anindita Roy: anindita.roy@paediatrics.ox.ac.uk

Bethan Psaila: bethan.psaila@ndcls.ox.ac.uk

Neil Ashley: neil.ashley@ndcls.ox.ac.uk

Yvonne Harrington: yvonne.harrington@nhs.net

Claus Nerlov: claus.nerlov@imm.ox.ac.uk

Suzanne M. Watt: suzanne.watt@ndcls.ox.ac.uk

Irene G. Roberts: irene.roberts@paediatrics.ox.ac.uk

James Davies: james.davies@ndcls.ox.ac.uk

Corresponding author: James O. J. Davies (james.davies@ndcls.ox.ac.uk)
MRC Molecular Haematology Unit

MRC Weatherall Institute of Molecular Medicine

Radcliffe Department of Medicine, University of Oxford

John Radcliffe Hospital

Headington, Oxford. OX3 9DS

Tel: +44 1865 222419

Word count:1377

TO THE EDITOR:

Sickle cell disease (SCD) and thalassemia are the commonest forms of inherited anemia, affecting around 300,000 births per annum world-wide. They result in progressive end organ damage mainly due to recurrent vasoocclusive crises (in SCD) and iron overload. Allogeneic hemopoietic stem cell transplantation (HSCT) has been the only curative option but generally is only suitable for a minority of young patients with severe disease due to the transplant-related mortality[1-3]. Recently, it has become possible to use gene therapy and genome editing to modify HSCs *ex-vivo* for autologous transplantation. Gene therapy approaches are based on using lentiviral vectors to insert exogenous copies of beta globin (HBB)[4] or to upregulate gamma globin (HBG) production[5]. Several genome editing strategies have been developed including induction of fetal hemoglobin[6]; reduction of alpha globin expression[7] and correction of the HBB gene *in situ* using homology directed repair[8].

A number of lines of evidence suggest that hemopoietic stem/progenitor cell (HSPC) function might be perturbed in SCD and thalassemia. SCD is an inflammatory condition[9], which results in chronic leukocytosis, particularly neutrophilia and monocytosis and this chronic inflammatory state may perturb HSPC function[10,11]. SCD is also associated with persistently elevated plasma erythropoietin (Epo)[12] and transferrin receptor[13] levels. Murine models show that high levels of Epo reprogram the transcriptome of HSPCs, supporting the development of committed erythroid progenitors[14]. This results in an erythroid lineage bias during differentiation and a decreased myeloid output. Interestingly, murine models also show that the increased reactive oxygen species in SCD perturb haemopoietic stem cell function[15].

To create a platform for these advanced cellular therapies, we set out to perform an in-depth characterization of the cellular and molecular composition of the CD34+ hemopoietic stem and progenitor cell (HSPC) in children with SCD and thalassemia.

Results and Discussion

CD34+ cells were MACS-purified (Miltenyi Biotech) from pre-transplant bone marrow harvests from five children with homozygous SCD (SCD1-5); three children with transfusion-dependent beta-thalassemia major (BT1-3) and four healthy controls (see supplemental Table 1 for clinical details). Adult bone marrow was used as the control because the majority of data supporting the use of CD34 for hemopoietic stem cell dosing in transplantation is derived from adult patients[16,17]. Further FACS sorted CD34+ cells were analyzed using a combination of single cell RNA sequencing (scRNA-seq) on the Chromium platform (10X Genomics), multiparameter flow cytometry and colony-forming assays (supplemental Methods).

Analysis of scRNAseq data using Seurat[18] and AUcell[19] (see supplemental Table 2 for QC data) showed a marked expansion of CD34+ B-lymphoid progenitors (Figure 1A&B)

in all 8 children with hemoglobinopathies compared to controls (supplemental Figures 1-6). Cells identified as B-lymphoid progenitors were enriched for expression of lymphoid genes[20,21] including CD79A, CD79B, VPREB1, VPREB3 and EBF1 (supplemental Table 3, AUC 0.25 was used for all samples). As predicted, patients with SCD appear to show an increased proportion of cells with an erythroid gene expression program although this did not reach statistical significance (Figure 1B; supplemental Table 3) (Control: 7.86%, Thalassemia: 6.97% and SCD: 9.67%). No expansion of erythroid gene expressing cells was seen in children with thalassemia consistent with adherence to a rigorous regular transfusion protocol designed to suppress endogenous erythropoiesis in these patients (Figure 1A&B, supplemental Figure 1-5).

Trajectory analysis was performed using an in-house R package[20], which allowed the expanded lymphoid progenitors to be subclassified into early B progenitors[22], expressing high levels of IL7R and IGHM, and late CD24 expressing B progenitors (supplemental Figure 5 and supplemental Tables 1-3). These data were corroborated using multiparameter flow-cytometry, which showed significant increases in the frequency of CD34+CD10+CD19+ cells in these patients (Figure 1C&D; supplemental Figures 7-8; supplemental Table 4). In some of the children (e.g. BT2, SCD2) over 65% of all CD34+ cells were CD19+CD10+CD34+ B-progenitors (Figure 1C&D). There was significant variation between patients, which did not clearly correlate with age, genotype or clinical factors (Supplemental Figure 6B, Supplemental Table 1). Previous studies show that children have higher numbers of CD10+CD19+ B cells in the bone marrow compared to adults but this rarely exceeds 20% of all cells[23], however, there are no systematic studies regarding the proportion of CD34+ B-progenitor cells in children. Our hypothesis is that a proinflammatory state in the children with hemoglobinopathies leads to these changes but it is possible that it relates to transfusion related alloimmunization or it may be an age related phenomenon.

scRNAseq analysis also suggested a relative reduction in the frequency of multipotent HSPC within the CD34+ compartment in all the children with thalassemia and SCD (Supplemental Figure 1&7-9). We carried out flow cytometric analysis, which confirmed a significant reduction in the proportion of CD38low\- cells as a proportion of total CD34+ cells in hemoglobinopathy patients compared to controls (Figure 2A&B; supplemental Figure 8). The proportion of HSC (Lin-CD10-CD34+CD38low\-CD45RA-CD90+) and MPP (Lin-CD10-CD34+CD38low\-CD45RA-CD90-) in the CD34+ compartment was also reduced in the hemoglobinopathy patients (Figure 2B) raising the possibility that such patients had a reduced frequency of editable long-term multipotent HSPC.

Despite this, when CD34+CD10+ lymphoid progenitors were excluded from the analysis by gating on the Lin-CD10-CD34+ compartment, the proportion of CD38low\-cells in the population was higher in SCD patients[24] (median 19.4±6%) compared to controls (9.7±0.55%) and thalassemia (7.7±1.9%) (supplemental Figure 8B). Flow cytometric analysis of the Lin-CD10-CD34+CD38+ myeloid progenitor compartment showed no significant differences in GMP, although CMP and MEP were slightly decreased in the patients (supplemental Figure 8C&E). Colony-forming assays from sorted progenitor populations showed no significant differences between the patients and controls (supplemental Figure 9B), suggesting no functional differences between the progenitor populations.

The majority of gene editing and gene therapy pre-clinical and clinical trials use total CD34+ cells to determine cell dose for transplantation. The massively expanded B-progenitor compartment in the bone marrow of children with hemoglobinopathies will significantly underestimate the proportion of editable long-term multipotent HSPC in

autologous bone marrow samples. Consequently, we went on to investigate alternative strategies for quantifying HSPC in these patients. First, we confirmed that the majority of CD34+ B-lymphoid progenitor cells were excluded by CD10 or CD19 included lineage cocktail (Figure 1B; supplemental Figure 8A). This was confirmed by paired Chromium 10x analysis of lin-CD34+ cells with unselected CD34+ cells from control sample (supplemental Figure 9A).

Our scRNA-seq data showed that CD133 (PROM1), a well-defined alternative to CD34 for quantification of HSPCs[21], is expressed by very few B-progenitors (supplemental Figure 10A&B). Addition of CD133 improved the reliability of flow cytometric measurement of HSC frequency. Gating on CD133+ cells removed CD34+ lymphoid progenitors and corrected the proportion of HSCs as quantified by CD133+Lin-CD10-CD34+CD38low\-CD45RA-CD90+ such that it was similar in both patients and controls (Figure 2C).

Our data show that the majority children with SCD and beta thalassemia major have a marked increase in the proportion of B-progenitor cells, which do not have long-term repopulating capacity as measured by xenograft assays[25]. As a result, an apparent rather than absolute reduction of HSC is observed in the CD34+ compartment. CD34 is therefore a potentially unreliable marker as a surrogate for HSC frequency in these patients. Data from gene therapy studies suggests that patients with hemoglobinopathies are often slow to engraft following transplantation with a gene therapy product. In the first patient treated with lentiviral vector mediated beta globin therapy, neutrophil engraftment occurred on day +38[4] and subsequent patients have also been slow to achieve neutrophil engraftment (median 22 days (range 17-29)[16]. We hypothesize that this may be in part due to the large numbers of CD34+ B-lymphoid cells causing the number of HSCs to be overestimated for a given dose of CD34+ cells/kg.

In conclusion, we have shown that, compared to healthy adults, a high percentage of CD34+ cells in bone marrow harvests taken from children with sickle cell disease and thalassemia are B lymphoid progenitor cells. This reduces the proportion of other stem and progenitor populations within the CD34 compartment, making it unreliable to use CD34 alone as a biomarker for quantifying HSCs in these patients. This could potentially be circumvented at least partially either by the use of lineage depletion or the addition of CD133 to exclude B progenitor cells. Further work is required to delineate the mechanisms behind this phenomenon, particularly since there are limited studies defining how the CD34+ stem and progenitor compartment changes with age. This finding has important implications for transplantation, gene therapy and genome editing for these diseases.

Acknowledgements

This work was funded by a Medical Research Council Discovery Award led by Prof. Doug Higgs (Developing an initiative in stem cell editing for human genetic diseases, MC_PC_15069). JD and PH are also funded by an MRC Clinician Scientist Award (MRC Clinician Scientist Fellowship ref. MR/R008108) to JD.

Authorship Contributions

P.H. designed, performed, analyzed experiments, performed bioinformatics analyses and wrote the manuscript. G. W., B. P. and A. R. contributed to experimental analyses. K.C. performed FACS sorting. N. A contributed to the single cell experiment. Y.H. supplied patients' details and assisted with obtaining consent. J.D., C. N., J. F., N. R., S. M. W., S. T., and I. R. conceived and supervised the project, designed and analyzed experiments and contributed to writing the manuscript.

Disclosure of Conflicts of Interest

Figure Legends

Figure 1. Marked expansion of lymphoid-primed progenitors in the bone marrow CD34+ cells of children with sickle cell diseases and thalassemia.

A) Dimensionality reduction using t-Distributed Stochastic Neighbour Embedding (t-SNE) on pooled data from four controls (n=29,086), three Thalassemia (n=11,034) and five SCD patients (n=11,776) cells identified distinct cell populations while preserving inter-cluster relationships. Cell type was annotated based on the expression of reported marker genes using AUcell (see supplemental Table 3). **B)** Components of cell types identified in single-cell sequencing data: Lymphoid Progenitors, Megakaryocyte/Erythroid Progenitors, Granulocyte/Monocyte Progenitors and Hematopoietic Stem/Progenitor Cells. **C)** Quantification of B progenitor cells by flow cytometry in normal donors and patients (SCD in red and Thalassemia in blue). **D)** Multiparameter flow cytometry plots showing the expansion of CD10+CD19+ cells in CD34+ cells derived from the bone marrow of patients with SCD and thalassemia.

Figure 2. CD34-based HSPC estimation overestimate the HSC quantity in patients. A) Flow plots show gating of Lin-CD10-CD34+cD38low\-CD45RA-CD90+HSCs from a normal donor and a child with sickle cell disease. **B)** The proportion of phenotypic Lin-CD10-CD34+CD38low\-CD45RA-CD90+ HSCs in total CD34+ cells was reduced in patients (Supplemental Figure 3B&D). This may be partially explained by the fact that she was the oldest and only post pubertal subject. **C**) The proportion of phenotypic Lin-CD10-CD34+CD133+CD38low\-CD45RA-CD90+HSCs in total CD133+ cells were not significantly different between patients and controls.

References:

1. Bolaños-Meade J, Brodsky RA. Blood and marrow transplantation for sickle cell disease: Is less more? *Blood Reviews*. 2014;28(6):243-248.
2. Talano JA, Cairo MS. Hematopoietic stem cell transplantation for sickle cell disease: state of the science. *Eur J Haematol*. 2015;94(5):391-399.
3. Walters MC, De Castro LM, Sullivan KM, et al. Indications and Results of HLA-Identical Sibling Hematopoietic Cell Transplantation for Sickle Cell Disease. *Biology of Blood and Marrow Transplantation*. 2016;22(2):207-211.
4. Ribeil J-A, Hacein-Bey-Abina S, Payen E, et al. Gene Therapy in a Patient with Sickle Cell Disease. *New England Journal of Medicine*. 2017;376(9):848-855.
5. Bauer DE, Kamran SC, Lessard S, et al. An erythroid enhancer of BCL11A subject to genetic variation determines fetal hemoglobin level. *Science (New York, NY)*. 2013;342(6155):253-257.
6. Lin MI, Paik E, Mishra B, et al. CRISPR/Cas9 Genome Editing to Treat Sickle Cell Disease and B-Thalassemia: Re-Creating Genetic Variants to Upregulate Fetal Hemoglobin Appear Well-Tolerated, Effective and Durable. *Blood*. 2017;130(Suppl 1):284-284.
7. Mettananda S, Fisher CA, Hay D, et al. Editing an α-globin enhancer in primary human hematopoietic stem cells as a treatment for β-thalassemia. *Nature Communications*. 2017;8(1):424.
8. Dever DP, Bak RO, Reinisch A, et al. CRISPR/Cas9 β-globin gene targeting in human haematopoietic stem cells. *Nature*. 2016;539(7629):384-389.
9. Hoppe CC. Inflammatory mediators of endothelial injury in sickle cell disease. *Hematol Oncol Clin North Am*. 2014;28(2):265-286.
10. Nogueira-Pedro A, Dias CC, Regina H, et al. Nitric Oxide-Induced Murine Hematopoietic Stem Cell Fate Involves Multiple Signaling Proteins, Gene Expression, and Redox Modulation. *STEM CELLS*. 2014;32(11):2949-2960.
11. Pietras EM, Mirantes-Barbeito C, Fong S, et al. Chronic interleukin-1 exposure drives haematopoietic stem cells towards precocious myeloid differentiation at the expense of self-renewal. *Nat Cell Biol*. 2016;18(6):607-618.
12. Croizat H, Nagel RL. Circulating cytokines response and the level of erythropoiesis in sickle cell anemia. *American Journal of Hematology*. 1999;60(2):105-115.
13. Lulla RR, Thompson AA, Liem RI. Elevated soluble transferrin receptor levels reflect increased erythropoietic drive rather than iron deficiency in pediatric sickle cell disease. *Pediatr Blood Cancer*. 2010;55(1):141-144.
14. Grover A, Mancini E, Moore S, et al. Erythropoietin guides multipotent hematopoietic progenitor cells toward an erythroid fate. *The Journal of Experimental Medicine*. 2014;211(2):181-188.
15. Javazon EH, Radhi M, Gangadharan B, Perry J, Archer DR. Hematopoietic stem cell function in a murine model of sickle cell disease. *Anemia*. 2012;2012.
16. Kanter J, Walters MC, Hsieh MM, et al. Interim Results from a Phase 1/2 Clinical Study of Lentiglobin Gene Therapy for Severe Sickle Cell Disease. *Blood*. 2016;128(22):1176.
17. Gorin NC. Bone Marrow Harvesting for HSCT. In: Carreras E, Dufour C, Mohty M, Kröger N, eds. The EBMT Handbook: Hematopoietic Stem Cell Transplantation and Cellular Therapies. Cham: Springer International Publishing; 2019:109-115.
18. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nature Biotechnology*. 2018;36:411.

19. Aibar S, González-Blas CB, Moerman T, et al. SCENIC: single-cell regulatory network inference and clustering. *Nature Methods*. 2017;14:1083.
20. Psaila B, Wang G, Meira AR, et al. Single-cell analyses reveal aberrant pathways for megakaryocyte-biased hematopoiesis in myelofibrosis and identify mutant clone-specific targets. *bioRxiv*. 2019:642819.
21. Hua P, Kronsteiner B, van der Garde M, et al. Single-cell assessment of transcriptome alterations induced by Scriptaid in early differentiated human haematopoietic progenitors during ex vivo expansion. *Scientific Reports*. 2019;9(1):5300.
22. Bendall SC, Davis KL, Amir E-AD, et al. Single-cell trajectory detection uncovers progression and regulatory coordination in human B cell development. *Cell*. 2014;157(3):714-725.
23. Andreoni C, Rigal D, Bonnard M, Bernaud J. Phenotypic analysis of a large number of normal human bone marrow sample by flow cytometry. *Blut*. 1990;61(5):271-277.
24. Luck L, Zeng L, Hiti AL, Weinberg KI, Malik P. Human CD34(+) and CD34(+)CD38(-) hematopoietic progenitors in sickle cell disease differ phenotypically and functionally from normal and suggest distinct subpopulations that generate F cells. *Exp Hematol*. 2004;32(5):483-493.
25. Byrne S, Elliott N, Rice S, et al. Discovery of a CD10 negative B-progenitor in human fetal life identifies unique ontogeny-related developmental programs. *Blood*. 2019:blood.2019001289.