The miR-144/451\textsuperscript{GFP} allele, a novel tool for resolving the erythroid potential of hematopoietic precursors

Kasper D Rasmussen and Dónal O’Carroll

European Molecular Biology Laboratory (EMBL), Mouse Biology Unit, Via Ramarini 32, Monterotondo Scalo, 00015, Italy.

Corresponding author:
Dónal O’Carroll
European Molecular Biology Laboratory (EMBL), Mouse Biology Unit, Via Ramarini 32, Monterotondo Scalo, 00015, Italy.
Tel +39 06 90091 222
Fax +39 06 90091 406
Email: ocarroll@embl.it
Abstract

A long outstanding problem is the resolution of the full potential of hematopoietic precursors. The commonly used allotypic marker Ly5 permits the tracing of lymphoid and granulocyte-macrophage (GM) output. Here we present a novel eGFP allele that allows the quantitative analysis of red blood cell (RBC) origin at the single-cell level. The miR-144/451 locus is required for erythroid development and homeostasis. Taking advantage of the fact that miR-451 is specifically and highly expressed in the erythroid lineage, we inserted an eGFP expression cassette into the miR-144/451 locus. In miR-144/451^+/eGFP animals, accumulation of eGFP is exclusively observed during terminal erythroid differentiation. Expression of miR-144/451^eGFP ignites immediately prior to the CFU-E stage and results in strong and complete labeling of all mature RBCs in circulation. Using competitive reconstitution experiments in the Ly5 transplant model, we show that eGFP linearly correlates with Ly5 expression. Thus, the miR-144/451^eGFP allele represents a novel tool for the resolution of erythroid potential.

Introduction

The study of blood cell formation has for many years spearheaded the investigation of adult stem cells. In the bone marrow an enigmatic, self-renewing, and rare hematopoietic stem cell (HSC) population gives rise to all blood cell lineages\(^1\). Traditionally, the study of hematopoietic stem and progenitor cells has profited from the relative ease of transplanting small amounts of stem cells into lethally irradiated recipients. Usage of mouse transplantation models enables the in vivo long-term qualitative and quantitative analysis of hematopoietic lineage determination\(^2\). The CD45/Ly5 congenic strains constitute the most widely adopted transplantation mouse model for tracking lineage output in peripheral blood\(^3\). However, the Ly5 antigen is expressed exclusively on nucleated cells, precluding the evaluation of donor-derived contribution to the erythrocytic and thrombocytic lineages. Several other transplant models are available that take advantage of naturally occurring isoforms (eg. \(Hbb^{sd}\), \(Gpi1^{ab}\))\(^4,5\), as well as transgenes under control of ubiquitous promoters (eg. \(Actb^{GFP}\), \(H2B^{GFP}\), \(H2K^{b-GFP}\))\(^6-8\). However, these methods have proved to be unsuitable or relatively inefficient in assessing erythroid output due to an absence of expression in mature RBCs or laborious methods of detection.
The vertebrate-specific miR-144/451 cluster, encoding miR-144 and miR-451, regulates erythroid development, homeostasis and susceptibility to oxidative stress\(^9\text{-}^{11}\). The maturation and accumulation of miR-451 is mediated by a unique biogenesis pathway that is strictly dependent upon Ago2’s endonuclease activity\(^12\text{-}^{14}\). Here we report a novel erythroid reporter allele based on the miR-144/451 non-coding RNA locus.

**Methods**

*Flow cytometry*

Cells were stained and analyzed and/or sorted using FACSCanto\(\text{I}\) or FACSARia\(\text{I}\) cytometers (BD) as described\(^15\text{-}^{23}\). Where appropriate, RBCs were lysed with an NH\(_4\)Cl-based solution (ACK lysis buffer) and non-specific binding was reduced by pre-incubation with Fc receptor block (Fc\(\gamma\text{II/III}\)). Dead cells and duplets were excluded by 7-Aminoactinomycin D (Sigma) staining and FSC/SSC profile, respectively. Data was analyzed off-line using FlowJo® Flow Cytometry analysis software (TreeStar, Inc.). Table S1 and S2 detail surface marker profiles and antibody information.

*Bone marrow transplantations*

Young (8-12 weeks) female C57Bl/6 \(\text{Ly5.1/Ly5.2}\) recipient mice were lethally irradiated in a single dose (875 Rad). Donor bone marrow cells were isolated from age- and sex-matched (8-10 weeks) congenic \(\text{Ly5.1 or miR-144/451}\text{^eGFP/Ly5.2}\) mice and 2x10\(^6\) nucleated cells were transplanted by tail vein injection. Bone marrow reconstituted mice were maintained on medicated water (Baytril; Bayer) for 4 weeks after reconstitution and peripheral blood was analyzed 8 weeks after transfer.

**Result and Discussion**

The miR-144/451 cluster represents a promising candidate locus to engineer a reporter for specifically marking terminal differentiated erythrocytes, and erythroid lineage-committed hematopoietic precursors. Firstly, expression of miR-144/451 are dependent on the erythroid transcription factor GATA-1\(^24\). Secondly, they are restricted to the erythroid lineage\(^9\). And finally, miR-451 accumulates to high levels in developing erythroblasts indicating extensive transcriptional activity of the locus\(^9\text{-}^{14}\). To investigate the expression pattern in detail, we
generated mice carrying an enhanced green fluorescent protein (eGFP) reporter cassette in the stem loop of precursor-miR-144 (Fig. 1A, Fig. S1). The insertion of the cassette disrupted expression of both miR-144 and miR-451 from the locus (Fig. S1). As expected, homozygous \textbf{miR-144/451\textasciitilde eGFP/eGFP} mice displayed the same phenotype as the previously described \textbf{miR-144/451\textasciitilde eGFP/eGFP} animals (Fig. S2). However, despite disruption of expression from one allele of the miR-144/451 locus and concomitant decrease in miRNA dosage, heterozygous \textbf{miR-144/451+/eGFP} animals did not show any physiological and erythropoietic abnormalities (Fig. S2). More importantly, we found that expression of the reporter faithfully distinguished hematopoietic cells expressing miR-144 and miR-451. Strong reporter expression was observed exclusively in Ter119\textsuperscript{+} erythroblasts among terminal differentiated blood cell lineages and conversely, miR-144 and miR-451 expression fractionated with eGFP expression in sorted \textbf{miR-144/451+/eGFP} bone marrow cells (Fig. S3).

We next wanted to perform an in-depth analysis of reporter expression in \textbf{miR-144/451+/eGFP} animals. Reporter expression was absent in the HSC population and all lymphoid- and GM-committed precursors in the bone marrow, as well as in developing T-cell progenitors in the thymus (Fig. S4). Instead, we could detect minor reporter expression in the common myeloid precursor (CMP) and pre-megakaryocyte-erythrocyte precursor (PreMegE) populations (Fig. 1B and 1C). Notably, as cells committed to the erythroid lineage by proceeding through PreMegE, PreCFU-E and finally CFU-E stages, a dramatic increase in eGFP expression from \textasciitilde 9\% to \textasciitilde 93\% of the population was seen (Fig. 1C). On the contrary, the low level of eGFP expression at the PreMegE stage was completely extinguished as cells differentiated further into megakaryocytic progenitors (MkP) (Fig. 1D). The strong reporter expression observed at the CFU-E stage was fully maintained as cells went through terminal erythroid differentiation in the bone marrow (Fig. 1E). Finally, we bled the mice and examined the peripheral blood for traces of eGFP. In agreement with the absence of reporter expression in the megakaryocytic precursors, also fully differentiated platelets were GFP-negative (Fig. 1F). Astoundingly, when examining mature erythrocytes in the peripheral blood we found that all cells (>99\%) were GFP-positive (Fig. 1G and 1H).

To examine whether the properties of the \textbf{miR-144/451+/eGFP} hematopoietic system could be faithfully recapitulated upon bone marrow transplantation we took advantage of the widely used CD45/Ly5\textsuperscript{5} transplantation model. Lethally irradiated recipients were injected with either control (Ly5.1) or experimental (\textbf{miR-144/451+/eGFP}; Ly5.2) bone marrow followed by peripheral blood analysis 8 weeks post-transplantation. Similar to WT, \textbf{miR-144/451+/eGFP} cells could fully reconstitute the GM lineage and the B- and T-cell lineage (Fig. 2A and 2B).
Importantly, the reconstitution with \( \text{miR-144/451}^{+/eGFP} \) cells also resulted in >99% GFP-positive erythrocytes in the peripheral blood of the recipients (Fig. 2C). Next, we wanted to examine the quantitative nature of the \( \text{miR-144/451}^{+/eGFP} \) reporter in a competitive transplantation assay. Bone marrow transplantations were performed with diverse mixtures of control (Ly5.1) or experimental (\( \text{miR-144/451}^{+/eGFP}; \text{Ly5.2} \)) cells (Fig. 2D). Remarkably, analysis of peripheral blood in four independent experiments revealed a near-perfect linear correlation \((r^2 > 0.9)\) between the relative contribution of leukocytes, as measured by Ly5.2 chimerism, and GFP-positive RBCs (Fig. 2E). The two parameters are directly proportional as the slopes approximate a value of 1 in each experiment (Fig. 2E). Hence, the relative fraction of GFP-erythrocytes can be used as a novel direct measure of the HSC output in the erythroid lineage.

In summary, analysis of the \( \text{miR-144/451}^{eGFP} \) allele revealed unprecedented reporter expression in the entire pool of mature erythrocytes both in terms of specificity and reporter intensity. High transcriptional activity of the locus during terminal erythroid differentiation results in a substantial accumulation of eGFP protein, which in turn concedes detection of RBCs throughout their extended lifetime in circulation. Equivalent to Ly5 mediated tracking of leukocyte output, the \( \text{miR-144/451}^{eGFP} \) reporter enables single-cell resolution and quantification of \textit{in vivo} erythroid lineage output from a subset of hematopoietic stem and progenitor cells. Moreover, the absence of reporter expression in early progenitors makes the allele compatible with the usage of already existing GFP-based reporters to isolate HSC subsets for subsequent analysis of erythroid potential. Thus, the \( \text{miR-144/451}^{eGFP} \) allele represents a novel quantitative \textit{in vivo} erythroid lineage assay that will be a valuable tool in the study of blood formation and hematopoietic stem and progenitor cell heterogeneity.

**Acknowledgements**

We acknowledge the services of Daniel Bilbao-Cortes and Jeanette Rientjes at EMBL Flow Cytometry Core facility and Monash University’s Gene Recombineering facility, respectively.

**Authorship Contributions**

KDR contributed to the design, execution and analysis of the experiments and wrote the final version of the manuscript. DOC conceived and supervised this study.
Disclosure of Conflicts of Interest

The authors declare no financial conflicts of interests.

References


Figure Legends

Fig. 1. miR-144/451<sup>eGFP</sup> expression in the hematopoietic system

(A) Schematic overview of the miR-144/451 locus before (upper) and after (lower) insertion of the eGFP expression cassette. The upstream GATA-1 cis-regulatory motif (CRM) and the transcriptional start site (TSS) are indicated. The integration site (chromosome 11: 77886538) of the eGFP cassette within the pre-miR-144 (blue) is shown. The black triangle symbolizes the remaining <i>frt</i> site after <i>Flp</i>-mediated excision of the <i>neo</i> cassette and the white rectangle indicates the SV40 poly(A) sequence. The synthetic intronic sequence introduced to ensure efficient nuclear processing and export of the transcript, and the newly introduced <i>AflII</i> restriction site used for Southern analysis is shown. (B) Small population of GFPlow cells observed in the CMP population. (i) Live Lin<sup>+</sup>, IL-7R<sup>−</sup>, Sca-1<sup>−</sup>, c-kit<sup>+</sup> gated cells stained for CD34 and CD16/32. The blue gate indicates the CMP subset. (ii) Cells gated as CMP as a function of FSC-A and GFP intensity. (C) MiR-144/451<sup>eGFP</sup> expression is strongly induced upon erythroid lineage commitment. (i) Live Lin<sup>+</sup>, IL-7R<sup>−</sup>, Sca-1<sup>−</sup>, c-kit<sup>+</sup>, CD41<sup>−</sup> gated cells stained for CD150 and CD105. The blue gates indicate granulo-monocyte precursor (GMP) and pre-granulo-monocyte precursor (Pre-GM), pre-megakaryocyte-erythroid precursor (Pre-MegE), Pre-colony forming-unit erythroid precursor (Pre-CFU-E), colony forming unit erythroid (CFU-E), respectively. (ii) The respective populations identified in (i) plotted as a function of FSC-A and GFP intensity. (D) Megakaryocytic precursors (MkP) do not express the reporter. (i) Live Lin<sup>+</sup>, IL-7R<sup>−</sup>, Sca-1<sup>−</sup>, c-kit<sup>+</sup> cells stained for CD150 and CD41. The blue gate indicates the MkP subset. (ii) Cells gated as MkP as a function of FSC-A and GFP. (E) All terminal erythropoietic subpopulations are strongly GFP-positive. (i) Live unfractionated bone marrow cells stained for CD71 and Ter119. The blue gates indicate proerythroblast (Pro), Basophilic erythroblast (Baso), Poly-chromatophilic erythroblast (Poly), and Orthochromatophilic erythroblast (Ortho) populations, respectively. (ii) The respective populations identified in (i) plotted as a function of FSC-A and GFP intensity. (F) Mature platelets in the peripheral blood are GFP-negative. (i) EDTA-anticoagulated blood cells plotted as a function of FSC-A and SSC-A. Note the logarithmic axis. (ii) FSC-A<sub>low</sub>, SSC-A<sub>low</sub> gated cells stained with CD41 and CD61. The blue gate identifies small mature platelets. (iii) Cells gated as platelets as a function of FSC-A and GFP. (G) Mature erythrocytes in peripheral blood are GFP-positive. (i) EDTA-anticoagulated peripheral blood cells stained for Ter119. The blue gate indicates mature RBCs. (ii) Cells gated as RBCs as a function of FSC-
A and GFP. (H) Blood smears of fixed peripheral blood show strong endogenous GFP signal from all RBCs. Fluorescence micrographs and corresponding bright field images are shown for wild-type and miR-144/451+/eGFP.

All FACS plot shown are representative of stained miR-144/451+/eGFP bone marrow cells, except (F) and (G) which are miR-144/451+/eGFP peripheral blood cell stains. The gates for GFP-positive cells in FSC-A vs GFP plots are defined by taking the respective cell population background fluorescence into account using wild-type littermate control stained cells. The percentages shown in the plots indicate the average GFP-positive fraction of the population in question, based on three independent experiments.

**Fig. 2. miR-144/451/eGFP is a quantitative marker of erythroid lineage output.**

MiR-144/451+/eGFP cells can give rise to (A) G/M cells, (B) B- and T-cells, and (C) erythroid cells, and marks all erythrocytes in the peripheral blood of bone marrow-transplanted lethally irradiated Ly5.1/5.2 recipients. (i) Representative FACS plots of peripheral blood cells stained for (A) Gr-1 and Mac-1, (B) B220 and CD5, and (C) Ter119. (A & B) Representative FACS plots show the gated populations stained for Ly5.1 and Ly5.2 isolated from animals reconstituted with (ii) Ly5.1 or (iii) miR-144/451+/eGFP:Ly5.2 cells respectively. Note that a minor population (~15%) of radioresistant recipient-derived T-cells can be detected in both wild-type and miR-144/451+/eGFP injected animals. (C) Lower histograms show gated RBCs plotted for GFP expression (x-axis) from animals reconstituted with (ii) Ly5.1 or (iii) miR-144/451+/eGFP:Ly5.2 cells respectively. (A, B & C) Numbers indicate mean fraction of the various gated cell populations based on five independent measurements. (D) Experimental setup to evaluate miR-144/451/eGFP expression as a quantitative erythroid lineage marker. Lethally irradiated Ly5.1/5.2 recipients were injected with 2 million nucleated whole bone marrow cells representing diverse mixtures of Ly5.1 or miR-144/451+/eGFP:Ly5.2 cells. Eight week post-transplantation, the fraction of GFP-positive RBCs and Ly5.2 single-positive leukocytes were measured in the peripheral blood. (E) Near-perfect linear correlation between Ly5 chimerism and GFP-positive RBCs. Plot indicates fractions of GFP-positive erythrocytes (x-axis) and Ly5.2 single-positive leukocytes (y-axis) for individual transplanted mice as in (D). Empty and filled symbols represent individual animals in four independent transplantation experiments. The solid line represents simple linear regression based on data points from all experiments. The slope, standard error and r squared (r²) value computed for each individual experiment as well as for all datapoints are listed.
Rasmussen KD and O’Carroll D.

Fig. 2

A

B

C

D

Mix of donor bone marrow

Ly5.1

miR-144/451-\textsuperscript{\textsubscript{wtGFP}}

Ly5.2

2\textsuperscript{6} cells

875 rad

8 weeks

Analysis

E

Total Ly5.2 chimerism (%)

GFP-positive erythrocytes (%)
The miR-144/451eGFP allele, a novel tool for resolving the erythroid potential of hematopoietic precursors

Kasper D Rasmussen and Dónal O'Carroll