Functional Profiling of Single CRISPR/Cas9-Edited Human Long-Term Hematopoietic Stem Cells

Elvin Wagenblast[1], Maria Azkanaz[1], Sabrina A. Smith[1], Lorien Shakib[1], Jessica L. McLeod[1], Gabriela Krivdova[1], Leonard D. Shultz[2], Olga I. Gan[1], John E. Dick[1,3,*] & Eric R. Lechman[1,*]

Affiliations:

[1]Princess Margaret Cancer Centre, University Health Network, Toronto, ON M5G 1L7, Canada

[2]The Jackson Laboratory, Bar Harbor, ME 04609, United States

[3]Department of Molecular Genetics, University of Toronto, Toronto, ON M5S 1A8, Canada

*Shared senior and corresponding authors: John E. Dick: (john.dick@uhnresearch.ca) and Eric R. Lechman (elechman@uhnresearch.ca)

Abstract:

In the human hematopoietic system, rare self-renewing multi-potent long-term hematopoietic stem cells (LT-HSCs) are responsible for the lifelong production of mature blood cells and are the rational target for clinical regenerative therapies. However, the heterogeneity in the hematopoietic stem cell compartment and variable outcomes of CRISPR/Cas9 editing make functional interrogation of rare LT-HSCs challenging. Here, we report high efficiency LT-HSC editing at single cell resolution using electroporation of modified synthetic gRNAs and Cas9 protein. Targeted short isoform expression of the GATA1 transcription factor elicited distinct differentiation and proliferation effects in single LT-HSC when analyzed with functional *in vitro* differentiation and long-term repopulation xenotransplantation assays. Our method represents a blueprint for systematic genetic analysis of complex tissue hierarchies at single cell level.

Introduction:

The hematopoietic stem and progenitor cell (HSPC) compartment is a functional continuum comprised of multiple stem and progenitor cell populations including abundant committed progenitors such as common myeloid progenitors (CMPs) and myelo–erythroid progenitors (MEPs), as well as rarer multipotent stem cells including short-term hematopoietic stem cells (ST-HSCs) and long-term hematopoietic stem cells (LT-HSCs, Supplementary Fig. 1)[1,2]. LT-HSCs are the only population that have the ability to permanently repopulate the entire hematopoietic system following transplantation[2]; they represent the key target for blood-based

regenerative therapies. Thus, LT-HSCs are essential for therapeutic genome editing to correct acquired and genetic hematopoietic disorders[3,4]. Furthermore, the pathogenesis of hematological malignancies like acute myeloid leukemia (AML) is associated with the presence of initiating mutations acquired in LT-HSCs, which lead to their competitive expansion[5,6]. Pre-leukemic LT-HSCs are a source of clonal evolution within blood malignancies and can act as a reservoir of relapse after chemotherapy treatment[7]. Thus, modeling and understanding the genetic complexity and cellular heterogeneity seen in human hematological malignancies requires novel methodologies that allow genome editing in LT-HSCs and their functional read-out[3,4].

Recently several studies have demonstrated efficient gene editing of bulk CD34$^+$ populations that are enriched for human HSPCs[8-16]. Highly efficient non-homologous end joining (NHEJ) mediated gene disruption of up to 80-90% efficiency has been reported in CD34$^+$ HSPCs[10,11,14,16]. In addition, homology-directed repair (HDR) mediated knock-ins, with or without selectable fluorescent reporter genes, have been established with an efficiency of up to 20% in CD34$^+$ HSPCs[8,9,11,13,15]. Stable integration of a fluorescent reporter using rAAV6 combined with flow cytometry-based sorting enabled enrichment of CRISPR/Cas9 edited HSPCs[8,9,16]. Because LT-HSCs represent only 0.1 – 1% of CD34$^+$ populations, these studies did not address LT-HSC targeting in the most direct manner. Previous studies have reported long term engraftment of up to 16 weeks following xenotransplantation of CRISPR/Cas9 edited human CD34$^+$ HSPCs[8,15,16], suggesting that rare LT-HSCs within the CD34$^+$

population can be gene edited. However, these studies utilized considerable numbers of $CD34^+$ HSPCs and did not have the resolution to functionally interrogate the differentiation and proliferation properties of individual LT-HSCs. In order to simplify LT-HSC targeting within heterogeneous $CD34^+$ HSPCs, we explored the possibility of CRISPR/Cas9 editing in highly purified LT-HSCs. This approach would enable direct functional characterization of LT-HSCs, rather than bulk populations. Here, we show successful editing of highly purified LT-HSCs via CRISPR/Cas9-mediated NHEJ or HDR and their subsequent functional investigation using single cell *in vitro* differentiation and near-clonal xenotransplantation assays.

Results:

CRISPR/Cas9-mediated GATA1 isoform expression in LT-HSCs

As a proof of principle, we modelled GATA1 isoform expression in LT-HSCs, ST-HSCs and MEPs from neonatal cord blood using CRISPR/Cas9. The GATA1 gene normally produces two protein isoforms as a result of alternative mRNA splicing – the GATA1 full length (GATA1-Long) and a truncated form (GATA1-Short). GATA1 isoform biology is particularly important for children with Down Syndrome. These children have a 150-fold higher risk of developing acute megakaryoblastic leukemia (AMKL), which is characterized by an abnormal proliferation of immature megakaryocytes[17-19]. Mutations in exon 2, which lead to the exclusive expression of GATA1-Short, are thought to be an essential driver of this disease. Previous work has shown effects of GATA1-Short on megakaryocytic proliferation, but these changes were only seen in fetal HSPCs and not in neonatal or adult HSPCs, implying a developmental stage-

specific effect[20]. To verify this hypothesis, we decided to test GATA1-Long versus GATA1-Short isoform expression in purified LT-HSCs, ST-HSCs and MEPs from neonatal cord blood.

Overall, our experimental scheme employed flow cytometric isolation of cord blood LT-HSCs for xenotransplantation and isolation of LT-HSCs, ST-HSCs and MEPs for single cell *in vitro* differentiation assays (Supplementary Fig. 2). Sorted cells were electroporated with modified synthetic gRNAs and Cas9 protein (Fig. 1a). Because of the transient delivery of the ribonucleoprotein complex and lack of stable integration of a selectable marker, all CRISPR/Cas9-mediated edits were subsequently genetically verified. To express the GATA1-Short isoform, we used two gRNAs targeting the 5' and 3' flanking regions of exon 2, resulting in the NHEJ-mediated dropout of the exon (Fig. 1b). By contrast, mutation of the GATA1-Short alternative start site on exon 3 from ATG to CTC via CRISPR/Cas9-mediated HDR led to the exclusive expression of the GATA1-Long isoform (Fig. 1c). Importantly, expression of both the GATA1-Short and -Long isoforms remained under the regulatory control of the endogenous GATA1 promoter. Because GATA1 is X-linked, all our studies utilized male cord blood samples. As a control, we used two gRNAs targeting exon 1 of the olfactory receptor OR2W5 that were designed with the CRoatan algorithm[21], resulting in a 150 bp dropout of the exon. After electroporation, individual LT-HSCs, ST-HSCs and MEPs were deposited into single cell *in vitro* assays under erythro-myeloid differentiation conditions (Fig. 1a)[22]. After 16-17 days, each single cell-derived colony was assessed by flow cytometry for

lineage output and the genotype was determined by PCR (Supplementary Fig. 3a-c). Moreover, LT-HSCs that were CRISPR/Cas9 edited with GATA1-Short or control gRNAs were transplanted into mice at a near-clonal level in order to detect lineage and proliferation biases (Fig. 1a).

Single cell *in vitro* differentiation of CRISPR/Cas9 edited LT-HSCs

CRISPR/Cas9 editing efficiency in LT-HSCs, ST-HSCs and MEPs was high; the percentage of single cell-derived colonies with homozygous deletion of OR2W5, GATA1-Short and GATA1-Long was 50-60%, 40% and 20%, respectively (Fig. 2a). Any control or GATA1-Short edited single cell colony with only one gRNA cut and no exon dropout was disregarded. Although statistically insignificant, CRISPR/Cas9-mediated HDR efficiencies were lower in LT-HSCs than in ST-HSCs and MEPs (Fig. 2a). While electroporation by itself did not drastically affect the efficiency of single cells to form colonies, LT-HSCs had slightly lower single cell colony formation efficiencies compared to ST-HSCs and MEPs (Fig. 2b). Finally, no off-target cleavage was detected at loci that were similar in sequence to the gRNA target sequence (Supplementary Fig. 4a). Western assay of bulk CRISPR/Cas9-edited MEPs cultured under erythro-myeloid conditions showed enrichment of either the GATA1-Long or GATA1-Short isoform (Fig. 2c). Culture of single LT-HSC, ST-HSC and MEP under erythro-myeloid differentiation conditions revealed a drastic shift towards megakaryocytic output upon exclusive expression of GATA1-Short, with a 2- and 4-fold increase in megakaryocytic colonies compared to LT-HSCs expressing control and GATA1-Long, respectively (Fig. 2d, e). Interestingly, only GATA1-Short edited

LT-HSCs produced bi-potent myelo-megakaryocytic colonies. ST-HSCs showed even higher fold increases towards megakaryocytic output compared to LT-HSCs. While CRISPR/Cas9 control edited MEPs did not possess any megakaryocytic differentiation capabilities, GATA1-Short edited MEPs were able to produce megakaryocytes, albeit with lower efficiency compared to LT-HSCs and ST-HSCs. In addition to the engineering of CRISPR/Cas9-mediated isoform re-arrangements, single gRNA mediated knock-outs can also be utilized in our method; for example, they could be used against STAG2 with single cell CRISPR/Cas9 efficiencies as high as 80-90% in LT-HSCs (Supplementary Fig. 4b, c). In summary, purified LT-HSCs and more committed stem and progenitor cells can be edited with high efficiency by CRISPR/Cas9-mediated NHEJ and HDR, and the effects of gene editing on differentiation can be read out reliably using single cell *in vitro* assays.

Long-term near-clonal xenotransplantation of CRISPR/Cas9 edited LT-HSCs

LT-HSC can only be tested functionally using the gold standard xenograft assay[23]. To investigate the functional consequences of exclusive GATA1-Short expression in LT-HSCs *in vivo*, we performed near-clonal xenotransplantation assays in NSG mice. Limiting dilution analysis[24] of CRISPR/Cas9 control-edited LT-HSCs injected into NSG mice for 24 weeks revealed a repopulating stem cell frequency of ~1/100 edited cells (Supplementary Fig. 5a). To achieve near-clonal xenotransplantation, we transplanted control- or GATA1-Short-edited LT-HSCs into NSG mice at an equivalent dose of 100-150 cells/mouse and after 24 weeks analyzed bone marrow harvested from the injected right femur (RF) and the left femur plus both tibias

(bone marrow - BM). Only mice with human $CD45^+$ engraftment levels of >5% in the RF and >90% CRISPR/Cas9 editing efficiency as determined by PCR and Sanger sequencing were included in our analysis: 40% of control mice and 35% of mice transplanted with GATA1-Short edited LT-HSCs fulfilled these criteria (Fig. 3a, Supplementary Fig. 5b-e). To precisely determine the genotype of the clonal progeny of injected LT-HSCs, secondary methylcellulose colony formation assays from cells of the RF were carried out. Analysis of CRISPR/Cas9 edits in individual colonies by Sanger sequencing revealed clonal engraftment in 3 out of 5 GATA1-Short edited LT-HSCs injected mice, highlighting that the xenotransplantations were indeed performed at a near clonal level (Supplementary Fig. 6). On average, human $CD45^+$ engraftment in the RF was 40%, both in control and GATA1-Short edited LT-HSCs injected mice (Fig. 3b). Interestingly, GATA1-Short edited LT-HSCs generated grafts with 2-fold higher percentage of human $CD41^+CD45^-$ megakaryocytes in the RF compared to controls (Fig. 3c). In addition, we observed an increase in the percentage of human $CD19^+CD45^+$ B-lymphoid cells in the RF at the expense of myeloid cells in these grafts (Fig. 3d, Supplementary Fig. 7a-d). GATA1-Short edited LT-HSCs generated grafts with higher absolute cell numbers, mainly due to increased numbers of B-lymphoid lineage cells (Fig. 3e, Supplementary Fig. 7e-i). The observed increase in megakaryocytic output in mice transplanted with GATA1-Short edited LT-HSCs (Fig. 3f) confirms our single cell *in vitro* findings, and demonstrates the feasibility of conducting near-clonal xenotransplantation assays using purified CRISPR/Cas9-edited LT-HSCs.

Normal human HSCs show a predominant B-lymphoid bias upon long-term engraftment in NSG mice[25]. We therefore repeated our xenotransplantation assays using c-kit-deficient NSGW41 recipients, which support enhanced erythropoiesis and megakaryocyte formation[26]. Limiting dilution analysis of CRISPR/Cas9-edited LT-HSCs in NSGW41 mice for 12 weeks revealed a repopulating cell frequency of ~1/175, compared to ~1/100 in the NSG background for 24 weeks (Supplementary Fig. 8a). Consequently, an equivalent dose of 200-250 control- or GATA1-Short-edited LT-HSCs were transplanted into NSGW41 mice and human engraftment in the RF and BM was analyzed after 12 weeks. 35% of mice transplanted with control edited LT-HSCs and 30% of mice transplanted with GATA1-Short edited LT-HSCs showed robust engraftment and successful CRISPR/Cas9 editing (Fig. 4a, Supplementary Fig. 8b, c). Human $CD45^+$ engraftment was observed at similar high levels as in NSG mice (Fig. 4b). Strikingly, a 5-fold increase in megakaryocytic output was observed in the RF and BM of mice transplanted with GATA1-Short edited LT-HSCs (Fig. 4c). No changes were seen in the percentage of B-lymphoid cells, but a decrease in human $GlyA^+CD45^-$ erythroid cells was detected in mice transplanted with GATA1-Short edited LT-HSCs (Fig. 4d, Supplementary Fig. 9a, b). Analysis of total cell numbers revealed a similar pattern, including a 3-fold increase in the number of megakaryocytes and, at the same time, a 6-fold reduction in erythroid cell numbers (Fig. 4e, Supplementary Fig. 9c-g). There were no differences in B-lymphoid proliferation in grafts generated by control- and GATA1-Short edited LT-HSCs in NSGW41 mice, in contrast to our observations in NSG recipients, which was possibly due to the reduced lymphoid bias observed in NSGW41 mice. In summary,

xenotransplantation into NSGW41 mice further augmented the megakaryocytic output of GATA1-Short edited LT-HSCs *in vivo*, with a concomitant decrease in erythroid cells (Fig. 4f).

Discussion:

In conclusion, we demonstrate that isolated LT-HSCs as well as more committed stem and progenitor cells can be edited with high efficiency using CRISPR/Cas9. Conventional CRISPR/Cas9 approaches on bulk $CD34^+$ populations yield a mixed population of edited and un-edited cells, making functional characterization difficult. Our approach, which requires retrospective verification of the CRISPR/Cas9 edits, permits the functional interrogation of LT-HSCs at a single cell level. Alternative approaches utilize HDR-mediated stable integration of selectable markers that allow prospective enrichment of CRISPR/Cas9 edited cells upon expression of the fluorescent marker[8,9]. The advantages of our method are that no exogenous sequences are introduced into the genomic DNA and all regulatory processes such as splicing and spatial control of promoters and enhancers are kept intact. Additionally, we show for the first time that GATA1-Short isoform expression in LT-HSCs under the endogenous promoter leads to an increase in megakaryocytic output; this was not previously observed in cord blood derived bulk $CD34^+$ HSPCs[20]. Aside from showing how a prototypical transcription factor functions in regulating lineage fate in LT-HSCs, this result is of great importance in modeling the first step in human Down Syndrome AMKL. Our method opens up the possibility of studying gene function relationships not only in LT-HSCs, but also in other stem and

progenitor cells to uncover cell type specific phenotypes. We believe that the continuous improvement of CRISPR/Cas9 editing efficiency, for example through chemically modified gRNAs[27-29] or different Cas9 variants[30,31], will further improve this approach. In the future, we envision that this method could potentially be adapted for cellular therapies.

Methods

Cord blood lineage depletion

Human cord blood samples were obtained from Trillium and William Osler hospitals with informed consent in accordance to guidelines approved by University Health Network (UHN) Research Ethics Board. Cord blood samples were processed 24 – 48 hours after birth. Male samples were exclusively utilized in this study because GATA1 and STAG2 are located on the X-chromosome. This aided the CRISPR/Cas9 efficiency because of the need to edit only one X-chromosome. Control OR2W5 is located on chromosome 1. Samples were diluted 1:1 with PBS and mononuclear cells were enriched using lymphocyte separation medium (Wisent, 305-010-CL). Subsequently, red blood cells were lysed using an ammonium chloride solution (StemCell Technologies, 07850). Then, lineage positive cells were depleted by negative selection with the StemSep Human Hematopoietic Progenitor Cell Enrichment Kit (StemCell Technologies, 14056) and Anti-Human CD41 TAC (Stem Cell, 14050) according to the manufacturer's protocol. Lineage depleted cells were stored in 50% PBS, 40% FBS (ThermoFisher, 12483-020) and 10% DMSO (FisherScientific, D128-500) at -150C.

Cord blood sorting

Lineage depleted cells were thawed via slow dropwise addition of X-VIVO 10 (Lonza, 04743Q) with 50% FBS (Sigma, 15A085) and DNaseI (200 ug/ml, Roche, 10104159001). Cells were spun at 1450RPM for 10 min at 4˚C and then

resuspended in PBS + 2.5% FBS. For all *in vitro* and *in vivo* experiments, the full stem and progenitor hierarchy sort as described in Notta et al[22] was utilized in order to sort LT-HSCs, ST-HSCs and MEPs. Lineage depleted cells were re-suspended in 100µl per 1x10^6 cells and stained in two subsequent rounds for 20 min at room temperature each. First, the following antibodies were used (volume per 1x10^6 cells, all from BD Biosciences, unless stated otherwise): CD45RA FITC (5µl, 555488, HI100), CD49f PE-Cy5 (3.5µl, 551129, GoH3), CD10 BV421 (4µl, 562902, HI10a), CD19 V450 (4µl, 560353, HIB19) and FLT3 CD135 biotin (12µl, clone 4G8, custom conjugation). After washing the cells, a second set of antibodies was used (volume per 1x10^6 cells, all from BD Biosciences, unless stated otherwise): CD45 V500 (4µl, 560777, HI30), CD34 APC-Cy7 (3µl, clone 581, custom conjugation), CD38 PE-Cy7 (2.5µl, 335825, HB7), CD90 APC (4µl, 559869, 5E10), CD7 A700 (10µl, 561603, M-T701) and Streptavidin Conjugate Qdot 605 (3µl, ThermoFisher, Q10101MP). Cell sorting was performed on the FACSAria III (BD Biosciences). LT-HSCs were sorted as $CD45^+CD34^+CD38^-CD45RA^-CD90^+CD49f^+$, ST-HSCs as $CD45^+CD34^+CD38^-CD45RA^-CD90^-CD49f^-$ and MEPs as $CD45^+CD34^+CD38^+CD10/19^-CD7^-CD45RA^-FLT3^-$ (Supplementary Fig. 1,2).

Pre-electroporation culture of sorted cells

Sorted LT-HSCs, ST-HSCs or MEPs were cultured for 36-48 hours in serum-free X-VIVO 10 (Lonza) media with 1% Bovine Serum Albumin Fraction V (Roche, 10735086001), 1X L-Glutamine (Thermo Fisher, 25030081), 1X Penicillin-Streptomycin (Thermo Fisher, 15140122) and the following cytokines (all from

Miltenyi Biotec): FLT3L (100ng/mL), G-CSF (10ng/mL), SCF (100ng/mL), TPO (15ng/mL) and IL-6 (10ng/mL). Cells were cultured in 96 well U-bottom plates (Corning, 351177).

gRNA and HDR template design

gRNAs for GATA1 Short and Long were designed on Benchling (http://www.benchling.com). For GATA1 Short, gRNAs sequences were considered that were flanking the 5' and 3' end of exon 2. Individual gRNAs targeting the 5' or 3' end were individually tested for cleavage efficiency and the best gRNA targeting each end was selected. Combined use of both gRNAs enabled complete excision of exon 2 (Fig. 1b). For GATA1 Long, gRNA sequences closest to the second ATG start codon were individually tested for cleavage efficiency and the best gRNA was selected. The GATA1 Long HDR template was designed with 60 bp homology ends at either side. For the template, the ATG (Methionine) start codon was mutated to CTC (Leucine) and the PAM sequence was mutated from GGG (Glycine) to GGC (Glycine) in order to avoid repeated cutting by the gRNA (Fig. 1c). The control gRNAs, which target exon 1 of the olfactory receptor OR2W5, were predicted by the CRoatan algrotihm[21]. The STAG2 gRNA was predicted with the same algorithm.

gRNA and HDR template sequences:

Control gRNA-1: GACAACCAGGAGGACGCACT

Control gRNA-2: CTCCCGGTGTGGACGTCGCA

GATA1 Short gRNA-1: TGGAACGGGGAGATGCAGGA

GATA1 Short gRNA-2: CCACTCAATGGAGTTACCTG

GATA1 Long gRNA: CATTGCTCAACTGTATGGAG

GATA1 Long HDR template:

TCTTTCCTCCATCCCTACCTGCCCCCAACAGTCTTTCAGGTGTACCCATTGCTCAACTGTC

TCGAGGGCATCCCAGGGGGCTCACCATATGCCGGCTGGGCCTACGGCAAGACGGGGCTCT

ACCCTGCC

STAG2 gRNA: AATGGTCATCACCAACAGAA

CRISPR/Cas9 RNP electroporation

gRNAs were synthesized from IDT as Alt-R CRISPR/Cas9 crRNA, which require annealing with Alt-R tracrRNA (IDT) to form a functional gRNA duplex. The HDR template was synthesized from IDT as a single-strand Ultramer. crRNAs and tracrRNAs were resuspended to 200μM with TE Buffer (IDT). Both RNA oligonucleotides were mixed 1:1 to a final concentration of 100uM and annealed at 95˚C for 5 min in a thermocycler, then cooled down to room temperature on the bench top. If using two gRNAs at the same time, both crRNAs were annealed to the tracrRNA in a single tube. For each reaction, 1.2μl crRNA:tracrRNA, 1.7μl Cas9 protein (IDT) and 2.1μl PBS were combined in a low binding Eppendorf tube (Axygen, MCT-175-C-S) and incubated for 15 min at room temperature. Subsequently, 1μl of 100μM electroporation enhancer (IDT) was added. Pre-electroporation cultured cells were washed in warm PBS and spun down at 1450RPM for 10 min at room temperature. Between 1x10^4 – 5x10^4 cells per condition were resuspended in 20μl of Buffer P3 (Lonza) per reaction and quickly

added to the Eppendorf tube containing the Cas9 gRNA RNP complex. The mixture was briefly mixed by pipetting and then added to the electroporation chamber (Lonza, V4XP3032). Cells were electroporated with the program DZ-100 and, immediately afterwards, 180µl of pre-warmed X-VIVO 10 media (as described above) was added. Cells were recovered overnight in the incubator before their use in *in vivo* or single cell *in vitro* assays.

Setup of a single cell *in vitro* assay

Single cell *in vitro* assays were performed according to Notta et al.[22] with slight modifications. Briefly, three days prior to the single cell assay, Nunc 96-well flat bottom plates (Thermo Fisher, 167008) were coated with 50µl per well of 0.2% Gelatin (G1393, Sigma-Aldrich) using a multi-channel pipette. The first and last column of each 96-well plate were not utilized and instead filled with 150µl PBS per well to support proper humidity within the plate. After one hour, the 0.2% Gelatin is removed and murine MS-5 stroma cells[32] were seeded at a density of 1500 cells per well in 100µl H5100 media (StemCell Technologies, 05150). One-day prior, the H5100 media was removed and replaced with 100µl of erythro-myeloid differentiation media. For this, StemPro-34 SFM media (Thermo Fisher, 10639011) is used with the supplied supplement, 1X L-Glutamine (Thermo Fisher, 25030081), 1X Penicillin-Streptomycin (Thermo Fisher, 15140122), 10µl Human LDL (StemCell Technologies, 02698) for every 50ml media and the following cytokines (all from Miltenyi Biotec unless stated otherwise): FLT3L (20ng/mL), GM-CSF (20ng/mL), SCF (100ng/mL), TPO (100ng/mL), EPO (3ng/mL, Eprex), IL-2 (10ng/mL), IL-3

(10ng/mL), IL-6 (50ng/mL), IL-7 (20ng/mL) and IL-11 (50ng/mL). On the day of the experiment, electroporated cells were stained with 1:2000 Sytox Blue (ThermoFisher, S34857) in PBS + 2.5% FBS and viable single cells were sorted and deposited into the MS-5 seeded 96 well plate using the FACSAria II (BD Biosciences). Single cells were cultured for 16-17 days with an addition of 100μl erythro-myeloid differentiation media per well at day 8.

Flow cytometry of single cell *in vitro* assay

Wells with hematopoietic cell content were marked one day prior and the total number of wells with colonies was used to assess CRISPR/Cas9 and single cell colony efficiencies. On the day of analysis, 140μl of media was removed from each well with a multi-channel pipette and the content was mixed well. Upon additional washing of the wells with PBS, the content in each well was transferred to a 96 well filter plate (8027, Pall) in order to remove stromal cells. For this, the filter plate was put on top of a 96 well U-bottom plate (Corning, 351177) and centrifuged at 1360RPM for 7 min at room temperature. After the cells were pelleted at the bottom of the U-bottom plate, the media was tossed by quickly inverting the U-bottom plate (~20ul of liquid per well remained). Then, 30μl of PBS + 2.5% FBS was added and mixed into each well with a multi-channel pipette and 25μl (half) was transferred to a 96 well PCR plate (Eppendorf, 951020362) for subsequent genomic DNA extraction (the PCR plate was sealed and stored at -80C). After that, 25μl of antibody mix in PBS + 2.5% FBS was added to each well, containing 25μl of hematopoietic cells, and stained for 45 min at 4˚C. The following antibodies were used (all

antibodies from BD Biosciences, unless stated otherwise): CD45 APC (1:100, 560777, HI30), CD34 APC-Cy7 (1:250, clone 581, custom conjugation), CD33 BV421 (1:50, Biolegend, 303416, WM53), CD71 FITC (1:100, 347513, L01.1), CD41 PE-Cy5 (1:200, Beckman Coulter, 6607116, P2) and GlyA PE (1:200, Beckman Coulter, IM2211U, KC16). Finally, 150μl of PBS + 2.5% FBS were added to each well with a multi-channel pipette and the cells were analyzed on the FACSCelesta with a high throughput sampler (HTS, BD Biosciences). All flow cytometry quantification was performed in a blinded manner. Generally, greater than 10 cells were required to call a positive lineage. Erythroid cells were defined as positive upon CD71 expression, with or without expression of GlyA.

Genotyping of single cell *in vitro* assay

The PCR plates containing 25μl of cells per well were thawed and a modified protocol of the Agencourt GenFind V2 (Beckman Coulter, A41499) was utilized to isolate genomic DNA. Wells with cell content were transferred to a new 96 well PCR plate (Eppendorf, 951020362) in order to utilize multi-channel pipetting for each future step. 25μl of lysis buffer and 1.2μl of Proteinase K (Zymo Research, D3001220) were pipetted into each well. After 30 min at room temperature, 50μl of magnetic particles were added. After 5 min, the PCR plate was placed on a magnetic stand (ThermoFisher, AM10027) for 10 min. The supernatant was removed with a multi-channel pipette and the plate was taken off the magnet. 200μl of wash buffer 1 were mixed into each well and the PCR plate was put back onto the magnetic stand. After 10 min, the supernatant was removed and each well was washed with 125μl of

wash buffer 2. Finally, after the last wash buffer was removed, the magnetic particles were resuspended with 60μl of TE buffer (IDT). The PCR plate was put back on the magnetic stand and after 10 min, 57μl of eluted genomic DNA was removed and put into a new PCR plate.

The CRISPR/Cas9 engineered genomic locus was amplified via PCR. For each PCR reaction, 23μl of eluted genomic DNA was mixed with 1μl of forward and reverse primer (10μM) and 25μl of AmpliTaq Gold 360 Master Mix (ThermoFisher, 4398881). The PCR program was: 95˚C for 10 min, followed by 95˚C for 30s, 56˚C for 30s and 72˚C for 1 min (40 cycles) and then 72˚C for 7 min. To identify colonies with the GATA1-Short genotype, 15μl of PCR product was run on a 1.5% agarose gel (ThermoFisher, 16500500). Control gRNA colonies were screened for homozygous deletion of OR2W5 (deletion within exon 1, 700bp to 500bp, (Supplementary Fig. 3a). Similarly, GATA1-Short colonies were screened for a shift in the size of the PCR product (deletion of exon 2, 1000 bp to 550 bp, Supplementary Fig. 3b). In order to identify colonies with the GATA1-Long genotype, PCR products were column purified using the ZR-96 DNA Clean-up Kit (Zymo Research, D4018) according to the manufacturer's protocol. The purified PCR product was Sanger sequenced using the reverse PCR primer and the chromatograms were inspected to identify colonies that contained the alternative start site mutation (Supplementary Fig. 3c). Finally, to identify STAG2 knock-out colonies, PCR products were column purified and sent for Sanger sequencing using the reverse PCR primer. Only colonies that showed a frame shift mutation were considered positive (Supplementary Fig. 4c).

PCR primers:

Control (OR2W5) forward primer: 5'-TCGGCCTGGACTGGAGAAAA-3'

Control (OR2W5) reverse primer: 5'-GAGACCACTGTGAGGTGAGA-3'

GATA1-Short forward primer: 5'-CAGGAGAGAATGAGAAAAGAGTGGA-3'

GATA1-Short reverse primer: 5'-ATTTCCAAGTGGGTTTTTGAGGAT-3'

GATA1-Long forward primer: 5'-GCCACACTGAGAGGCAATACT-3'

GATA1-Long reverse primer: 5'-AAAAGTCAGGGCCCCCATAAG-3'

STAG2-forward primer: 5'-CCACAAAGAGGCTGTCACAGTT-3'

STAG2-reverse primer: 5'-CATGCAGCAGAAAATGAATCAAAAC-3'

Western assay

MEPs were sorted and CRISPR/Cas9 RNP electroporated as described above. After electroporation, MEPs were cultured in erythro-myeloid differentiation media (as described above for single cells *in vitro* assays) for 3 days. 1x10^5 cells were then lysed in RIPA buffer (ThermoFisher, 899000) containing protease and phosphatase inhibitors (ThermoFisher, 78446). Subsequently, samples were centrifuged at 12,000g for 5 min at 4˚C and supernatants were utilized for Western assay. Western assay was performed on the size-based Wes capillary platform (ProteinSimple) using the 12-230 kDa capillary cartridge according to manufacturer's protocol. In order to determine the optimal antibody dilution, anti-GATA-1 antibody (Cell Signaling D24E4) was titrated prior to use on a lysate from $CD34^+$ cord blood and 1:5 dilution was used in the samples.

Animal studies

All mouse experiments were approved by the University Health Network (UHN) Animal Care Committee. All mouse transplants were performed with 8- to 12-week-old female NOD.*Cg-Prkdc*^{scid}*Il2rg*^{tm1Wjl}*/SzJ* (NSG) mice (JAX) that were sublethally irradiated with 225cGy, 24 hours before transplantation, or with 8- to 12-week-old female NOD.*Cg-Prkdc*^{scid}*Il2rg*^{tm1Wjl}*Kit*^{em1Mvw}*/SzJ* (NSGW41) mice that were not irradiated. Sample size was chosen to give sufficient power for calling significance with standard statistical tests. Intrafemoral injections were performed as previously described in Mazurier et al[33]. For this, mice were anesthetized with isoflurane and the right knee was secured in a bent position to drill a hole into the right femur with a 27gauge needle. Then, 100-250 CRISPR/Cas9 edited LT-HSCs were injected in 30µl PBS using a 28gauge ½ cc syringe (Becton Dickinson, 329461). LT-HSC cell numbers are based on the number of flow cytometry sorted cells at day 0. After 12 or 24 weeks, mice were sacrificed to obtain the right femur (RF) and bone marrow (left femur and both tibias, BM). Bones were flushed in 1mL PBS + 2.5% FBS and cells were centrifuged at 1450RPM for 10 min. Cells were resuspended in 500µl of PBS + 2.5% FBS. Subsequently, cells from BM and RF were counted in ammonium chloride (StemCell Technologies, 07850) using the Vicell XR (Beckman Coulter). 25µl of cells were used for flow cytometry analysis and another 25µl were frozen down for genomic DNA isolation in order to verify CRISPR/Cas9 edits (same protocol as above for single cell *in vitro* assays).

Limiting dilution *in vivo* assays

For limiting dilution transplantation assays, CRISPR/Cas9 control gRNA electroporated LT-HSCs were injected at defined doses (equivalent to 25, 50, 100 and 200 LT-HSCs) into 8- to 12-week-old female NSG or NSGW41 mice. LT-HSC cell numbers were based on the number of flow cytometry sorted cells at day 0. LT-HSC frequency was estimated using the online tool ELDA (http://bioinf.wehi.edu.au/software/elda/index.html)[24].

Flow cytometry of *in vivo* assays

Cells from the RF and BM of transplanted mice were stained in 96 well U-bottom plates (Corning, 351177) for 45 min at 4˚C. The following antibodies were used (all from BD Biosciences, unless stated otherwise): CD45 APC-Cy7 (1:100, 348795560566), CD45 A700 (1:100, 560566), CD33 APC (1:100, 340474) and CD19 V450 (1:100, 560353), CD41 PE-Cy5 (1:200, Beckman Coulter, 6607116), GlyA PE (1:100, Beckman Coulter, IM2211U) and CD3 FITC (1:100, 349201). For NSGW41 mice at 12 weeks, CD3 FITC staining was omitted. Cells were analyzed on the FACSCelesta (BD Biosciences).

CRISPR/Cas9 efficiency in electroporated cells and engrafted mice

After each CRISPR/Cas9 RNP electroporation, a small subset of cells was cultured in X-VIVO 10 media (as described above) for 5-7 days in order to validate CRISPR/Cas9 efficiency. Genomic DNA was isolated from bulk cells and the CRISPR/Cas9 engineered genomic locus was amplified via PCR as described above.

Sanger sequencing was carried out using the reverse PCR primer and the chromatograms were analyzed using the online tool TIDE (https://tide.deskgen.com/)[34] in order to verify CRISPR/Cas9 editing in control, GATA1-Short and GATA1-Long edited bulk cells. Because of the large deletion size of control (200bp) and GATA1-Short (400bp) edited cells, the CRISPR/Cas9 efficiency was evaluated based on the percentage of aberrant sequences after the gRNA cut site (Supplementary Fig. 5d).

For each transplanted mouse, CRISPR/Cas9 efficiency of control and GATA1-Short edited cells was evaluated in the RF using the same approach. Only mice that showed a CRISPR/Cas9 knock-out efficiency of >90% as determined by the percentage of aberrant sequences after the gRNA cut site and a $CD45^+$ engraftment level in the RF of >5% were utilized in the analysis. Because only one X-chromosome needed to be edited for GATA1-Short, single clonal engraftment was visible based on individual chromatograms (Supplementary Fig. 5e). GATA1-Short edited LT-HSCs transplanted mice with more than one clone were included into our near-clonal xenotransplantation analysis, as long as the CRISPR/Cas9 knock-out efficiency of >90% and engraftment criteria of >5% were satisfied.

Methylcellulose colony formation assay

1x10^5 cells from the RF of xenotransplanted mice were transferred to 1ml of MethoCult H4034 Optimum methylcellulose medium (StemCell Technologies, 04034) and plated onto a 35mm dish for human specific colony formation. After 10-11 days, individual colonies were collected, washed in PBS and genomic DNA was

isolated as described above. Individual CRISPR/Cas9 edits were determined using PCR amplification and Sanger sequencing with the reverse PCR primer as described above.

CRISPR/Cas9 Off-target analysis

Genomic loci that were similar to the gRNA target sequence were identified with Cas0-OFFinder (http://www.rgenome.net/cas-offinder)[35], using a mismatch number of 2-3 and a DNA/RNA bulge size of 0. Two genomic loci with 2 mismatches were chosen for GATA1-Short gRNA-1, four genomic loci with 3 mismatches for GATA1-Short gRNA-2, four genomic loci with 3 mismatches for GATA1-Long gRNA-1, three genomic loci with 2 mismatches for Control gRNA-1 and one genomic loci with 3 mismatches for Control gRNA-2. PCR primers were designed to amplify 500bp around these genomic loci. 5-10 single cell colonies that were positively identified with the correct CRISPR/Cas9 edit were selected from each cell type and genotype for PCR amplification. PCR products were column purified and subsequent Sanger sequencing with both the forward and reverse PCR primer was carried out. TIDE analysis was used to assess any CRISPR/Cas9 cleavage efficiency.

Statistical analysis

Error bars represent standard deviations. Statistical significance was assessed using two-tailed unpaired student's t-test.

Life Sciences Reporting Summary

Additional information on experimental design is available in the Nature Research Reporting Summary linked to this article.

Data availability

All datasets generated in this study are available within the paper or from the corresponding author upon reasonable request.

References

1. Laurenti, E. *et al.* CDK6 levels regulate quiescence exit in human hematopoietic stem cells. *Cell Stem Cell* **16,** 302–313 (2015).
2. Notta, F. *et al.* Isolation of single human hematopoietic stem cells capable of long-term multilineage engraftment. *Science* **333,** 218–221 (2011).
3. Porteus, M. H. A New Class of Medicines through DNA Editing. *N. Engl. J. Med.* **380,** 947–959 (2019).
4. Dever, D. P. & Porteus, M. H. The changing landscape of gene editing in hematopoietic stem cells: a step towards Cas9 clinical translation. *Curr. Opin. Hematol.* **24,** 481–488 (2017).
5. Shlush, L. I. *et al.* Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. *Nature* **506,** 328–333 (2014).
6. Corces-Zimmerman, M. R., Hong, W.-J., Weissman, I. L., Medeiros, B. C. & Majeti, R. Preleukemic mutations in human acute myeloid leukemia affect epigenetic regulators and persist in remission. *Proc. Natl. Acad. Sci. U.S.A.* **111,** 2548–2553 (2014).
7. Shlush, L. I. *et al.* Tracing the origins of relapse in acute myeloid leukaemia to stem cells. *Nature* **547,** 104–108 (2017).
8. Bak, R. O., Dever, D. P. & Porteus, M. H. CRISPR/Cas9 genome editing in human hematopoietic stem cells. *Nat Protoc* **13,** 358–376 (2018).
9. Bak, R. O. *et al.* Multiplexed genetic engineering of human hematopoietic stem and progenitor cells using CRISPR/Cas9 and AAV6. *Elife* **6,** 18 (2017).

10. Brunetti, L., Gundry, M. C., Kitano, A., Nakada, D. & Goodell, M. A. Highly Efficient Gene Disruption of Murine and Human Hematopoietic Progenitor Cells by CRISPR/Cas9. *J Vis Exp* (2018). doi:10.3791/57278

11. Gundry, M. C. *et al.* Highly Efficient Genome Editing of Murine and Human Hematopoietic Progenitor Cells by CRISPR/Cas9. *Cell Rep* **17,** 1453–1461 (2016).

12. Tothova, Z. *et al.* Multiplex CRISPR/Cas9-Based Genome Editing in Human Hematopoietic Stem Cells Models Clonal Hematopoiesis and Myeloid Neoplasia. *Cell Stem Cell* **21,** 547–555.e8 (2017).

13. Lattanzi, A. *et al.* Optimization of CRISPR/Cas9 Delivery to Human Hematopoietic Stem and Progenitor Cells for Therapeutic Genomic Rearrangements. *Mol. Ther.* **27,** 137–150 (2019).

14. Mandal, P. K. *et al.* Efficient ablation of genes in human hematopoietic stem and effector cells using CRISPR/Cas9. *Cell Stem Cell* **15,** 643–652 (2014).

15. Dever, D. P. *et al.* CRISPR/Cas9 β-globin gene targeting in human haematopoietic stem cells. *Nature* **539,** 384–389 (2016).

16. Wu, Y. *et al.* Highly efficient therapeutic gene editing of human hematopoietic stem cells. *Nat. Med.* **387,** 2554 (2019).

17. Hitzler, J. K., Cheung, J., Li, Y., Scherer, S. W. & Zipursky, A. GATA1 mutations in transient leukemia and acute megakaryoblastic leukemia of Down syndrome. *Blood* **101,** 4301–4304 (2003).

18. Roberts, I. *et al.* GATA1-mutant clones are frequent and often unsuspected in babies with Down syndrome: identification of a population at risk of

leukemia. *Blood* **122,** 3908–3917 (2013).

19. Gruber, T. A. & Downing, J. R. The biology of pediatric acute megakaryoblastic leukemia. *Blood* **126,** 943–949 (2015).

20. Gialesaki, S. *et al.* GATA1s exerts developmental stage-specific effects in human hematopoiesis. *Haematologica* **103,** e336–e340 (2018).

21. Erard, N., Knott, S. R. V. & Hannon, G. J. A CRISPR Resource for Individual, Combinatorial, or Multiplexed Gene Knockout. *Mol. Cell* **67,** 348–354.e4 (2017).

22. Notta, F. *et al.* Distinct routes of lineage development reshape the human blood hierarchy across ontogeny. *Science* **351,** aab2116–aab2116 (2016).

23. Dick, J. E. Stem cell concepts renew cancer research. *Blood* **112,** 4793–4807 (2008).

24. Hu, Y. & Smyth, G. K. ELDA: extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. *J. Immunol. Methods* **347,** 70–78 (2009).

25. Wunderlich, M. *et al.* Improved multilineage human hematopoietic reconstitution and function in NSGS mice. *PLoS ONE* **13,** e0209034 (2018).

26. Rahmig, S. *et al.* Improved Human Erythropoiesis and Platelet Formation in Humanized NSGW41 Mice. *Stem Cell Reports* **7,** 591–601 (2016).

27. Ryan, D. E. *et al.* Improving CRISPR-Cas specificity with chemical modifications in single-guide RNAs. *Nucleic Acids Res.* **46,** 792–803 (2018).

28. Hendel, A. *et al.* Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells. *Nat. Biotechnol.* **33,** 985–989 (2015).

29. Mir, A. *et al.* Heavily and fully modified RNAs guide efficient SpyCas9-mediated genome editing. *Nature Communications* **9,** 2641 (2018).

30. Gopalappa, R., Suresh, B., Ramakrishna, S. & Kim, H. H. Paired D10A Cas9 nickases are sometimes more efficient than individual nucleases for gene disruption. *Nucleic Acids Res.* **46,** e71–e71 (2018).

31. Vakulskas, C. A. *et al.* A high-fidelity Cas9 mutant delivered as a ribonucleoprotein complex enables efficient gene editing in human hematopoietic stem and progenitor cells. *Nat. Med.* **24,** 1216–1224 (2018).

32. Itoh, K. *et al.* Reproducible establishment of hemopoietic supportive stromal cell lines from murine bone marrow. *Exp. Hematol.* **17,** 145–153 (1989).

33. Mazurier, F., Doedens, M., Gan, O. I. & Dick, J. E. Rapid myeloerythroid repopulation after intrafemoral transplantation of NOD-SCID mice reveals a new class of human stem cells. *Nat. Med.* **9,** 959–963 (2003).

34. Brinkman, E. K., Chen, T., Amendola, M. & van Steensel, B. Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res.* **42,** e168–e168 (2014).

35. Bae, S., Park, J. & Kim, J.-S. Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics* **30,** 1473–1475 (2014).

Acknowledgements

We thank M. DSouza for support with mouse work; T. Velauthapillai, S. Ng, A. AuYeung, N. Simard and S. Zhao for assistance with flow cytometry; the labs of S. Chan and F. Notta for shared equipment. We thank M. Anders, J. Kennedy and J. Wang for comments on the manuscript. E.W. is supported by a long-term fellowship from the Human Frontier Science Program and a Banting postdoctoral fellowship. Work in J.E.D.'s laboratory is supported by grants from the Canadian Institutes for Health Research, Ontario Institute for Cancer Research through funding provided by the Government of Ontario, Terry Fox Foundation, Canadian Cancer Society and the University of Toronto's Medicine by Design initiative, which receives funding from the Canada First Research Excellence.

Author contributions

E.W., J.E.D and E.L. conceived the project, supervised research, and wrote the paper. E.W., O.I.G and E.L. analyzed experiments. E.W., M.A., S.S., L.S., and E.L. performed *in vitro* and *in vivo* experiments. J.L.M. assisted with intrafemoral injections. G.K. assisted with western assay. O.I.G. assisted with single cell *in vitro* assays and methylcellulose colony formation assays. L.D.S. provided NSGW41 mice.

Competing interests

The authors declare no competing financial interests.

Figure Legends

Figure 1: CRISPR/Cas9-mediated isoform expression of GATA1 at single cell level. a Experimental work-flow for single cell *in vitro* differentiation assay and near-clonal xenotransplantation into mice. **b** Two gRNAs targeting the 5' and 3' end of exon 2 of GATA1 led to the NHEJ-mediated dropout of this exon, resulting in the exclusive expression of GATA1-Short. **c** HDR-mediated mutation of the alternative start site from ATG to CTC using a single gRNA and a single-strand DNA template, resulting in the exclusive expression of GATA1-Long.

Figure 2: Functional interrogation of single CRISPR/Cas9 edited hematopoietic stem cells *in vitro*. a Percentage of CRISPR/Cas9 efficiency as determined by single cell-derived colonies that were positive for homozygous deletion of exon 1 in control OR2W5, positive for exclusive assignment to the GATA1-Long isoform through HDR-mediated mutation of the alternative start site or positive for exclusive assignment to the GATA1-Short isoform through deletion of exon 2 (n = 3 experiments with independent cord blood pools, error bars represent standard deviations, which is also the case for **b,d**). **b** Percentage of single cells that grew into a colony in the single cell *in vitro* differentiation assays (n = 3 experiments with independent cord blood pools, unpaired t-test $P < 0.05$ for % colony efficiency in LT-HSCs GATA1-Long versus ST-HSCs GATA1-Long). **c** Capillary-based western assay of GATA1 in bulk MEPs that were CRISPR/Cas9 edited with control, GATA1-Long and GATA1-Short gRNAs and cultured under erythro-myeloid conditions for three days (n = 2 experiments, only one shown). **d** Lineage output from the *in vitro* single cell

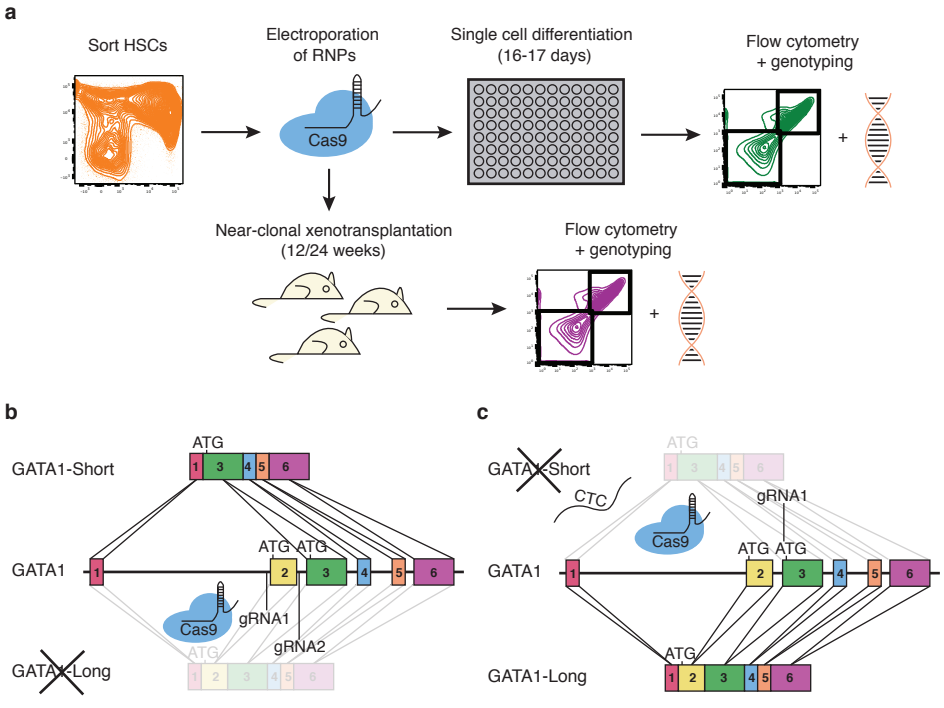
differentiation assay of individual CRISPR/Cas9 edited LT-HSCs, ST-HSCs and MEPs with control, GATA1-Long and GATA1-Short gRNAs. Numbers of single cell colonies with positive genotype are indicated at each condition (unpaired t-test $P < 0.005$ for E, M, Meg in GATA1-Short versus control and E, M, Meg in GATA1-Short versus GATA1-Long among all three cell types). **e** Overall percentage of erythroid, myeloid and megakaryocyte containing colonies from single cell *in vitro* differentiation assays ($n = 3$ experiments with independent cord blood pools, unpaired t-test $P < 0.0001$ % Meg in GATA1-Short versus Control or % Meg in GATA1-Short versus GATA1-Long among all three cell types).

Figure 3: Functional interrogation of single CRISPR/Cas9 edited hematopoietic stem cells in NSG mice. a Percentage of CRISPR/Cas9 edited LT-HSCs injected NSG mice with engraftment (>5% based on human $CD45^+$ expression in RF) and high CRISPR/Cas9 knock-out efficiency (>90% based on PCR and Sanger sequencing, $n = 3$ animal cohorts with independent cord blood pools, error bars represent standard deviations, which is also the case for **b-e**). **b** Engraftment levels of control and GATA1-short edited LT-HSCs injected NSG mice based on human $CD45^+$ expression in RF and BM. **c** Percentage of $CD45^-CD41^+$ megakaryocytes of control and GATA1-Short edited LT-HSCs injected NSG mice in RF and BM (unpaired t-test $P < 0.005$ for RF GATA1-Short versus RF control and unpaired t-test $P = 0.107$ for BM GATA1-Short versus BM control). **d** Percentage of $CD45^+CD19^+$ lymphoid cells in control and GATA1-Short edited LT-HSCs injected NSG mice in RF and BM (unpaired t-test $P < 0.005$ for RF GATA1-Short versus RF control). **e** Absolute cell numbers of different

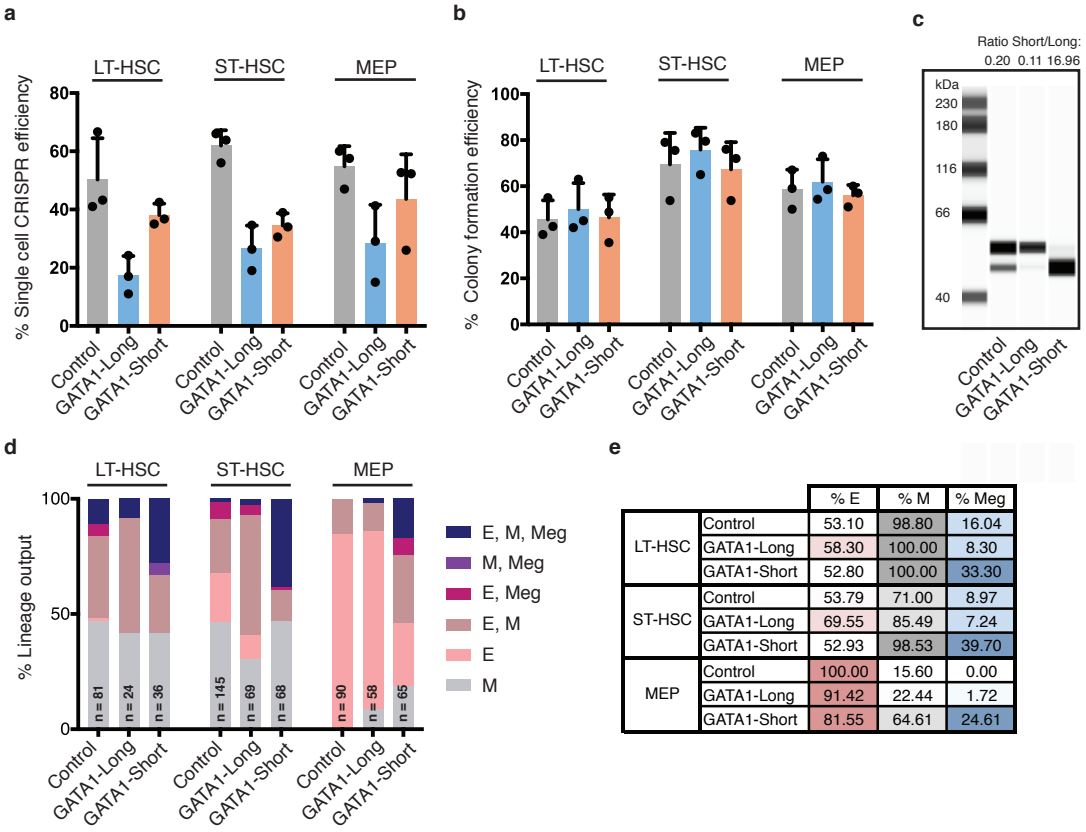
lineages in control and GATA1-Short edited LT-HSCs injected NSG mice in RF and BM (unpaired t-test $P < 0.01$ for B-lymphoid RF GATA1-Short versus RF control and unpaired t-test $P < 0.005$ for megakaryocytes RF GATA1-Short versus RF control). **f** Representative flow cytometry plots of control and GATA1-Short edited LT-HSCs injected NSG mice in RF showing $CD45^{-}CD41^{+}$ megakaryocytes and $CD45^{-}GlyA^{+}$ erythroid cells.

Figure 4: Functional interrogation of single CRISPR/Cas9 edited hematopoietic stem cells in W41 mice. a Percentage of CRISPR/Cas9 edited LT-HSCs injected NSGW41 mice with engraftment (>5% based on human $CD45^{+}$ expression in RF) and high CRISPR/Cas9 knock-out efficiency (>90% based on PCR and Sanger sequencing, $n = 3$ animal cohorts with independent cord blood pools, error bars represent standard deviations, which is also the case for **b-e**). **b** Engraftment levels of control and GATA1-short edited LT-HSCs injected NSGW41 mice based on human $CD45^{+}$ expression in RF and BM. **c** Percentage of $CD45^{-}CD41^{+}$ megakaryocytes of control and GATA1-Short edited LT-HSCs injected NSGW41 mice in RF and BM (unpaired t-test $P < 0.001$ for GATA1-Short versus control for both RF and BM). **d** Percentage of $CD45^{-}GlyA^{+}$ erythroid cells in control and GATA1-Short edited LT-HSCs injected NSGW41 mice in RF and BM (unpaired t-test $P < 0.0005$ for RF GATA1-Short versus RF control and unpaired t-test $P < 0.005$ for BM GATA1-Short versus BM control). **e** Total cell numbers of different lineages in control and GATA1-Short edited LT-HSCs injected NSGW41 mice in RF and BM (unpaired t-test $P < 0.005$ for megakaryocytes GATA1-Short versus control for both RF and BM and

unpaired t-test $P < 0.005$ for erythroid GATA1-Short versus control for both RF and BM). **f** Representative flow cytometry plots of control and GATA1-Short edited LT-HSCs injected NSGW41 mice in RF showing $CD45^{-}CD41^{+}$ megakaryocytes and $CD45^{-}GlyA^{+}$ erythroid cells.



bioRxiv preprint doi: https://doi.org/10.1101/609040; this version posted April 15, 2019.



e

		% E	% M	% Meg
LT-HSC	Control	53.10	98.80	16.04
	GATA1-Long	58.30	100.00	8.30
	GATA1-Short	52.80	100.00	33.30
ST-HSC	Control	53.79	71.00	8.97
	GATA1-Long	69.55	85.49	7.24
	GATA1-Short	52.93	98.53	39.70
MEP	Control	100.00	15.60	0.00
	GATA1-Long	91.42	22.44	1.72
	GATA1-Short	81.55	64.61	24.61