PRE-CLINICAL STUDY FOR THERAPEUTIC GENOME EDITING IN BETA-HEMOGLOBINOPATHIES

VIGNESH R

PhD THESIS
2023

SREE CHITRA TIRUNAL INSTITUTE FOR MEDICAL SCIENCES AND

TECHNOLOGY, TRIVANDRUM

An Institution of National Importance established by an Act of the Indian Parliament
(Act No.52 of 1980)

Dept. of Science and Technology, Govt. of India
www.sctimst.ac.in

PRE-CLINICAL STUDY FOR THERAPEUTIC GENOME EDITING IN BETA-HEMOGLOBINOPATHIES

A THESIS SUBMITTED BY

VIGNESH R

TO

SREE CHITRA TIRUNAL INSTITUTE FOR MEDICAL SCIENCES AND TECHNOLOGY, TRIVANDRUM.

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE AWARD OF

DOCTOR OF PHILOSOPHY

2023

DECLARATION BY THE STUDENT

CERTIFICATE

I, **VIGNESH R** hereby certify that I had personally carried out thework depicted in the thesis titled, **"PRE-CLINICAL STUDY FOR THERAPEUTIC GENOME EDITING IN BETA-HEMOGLOBINOPATHIES"**.

No part of this thesis has been submitted for the award of any other degree or diploma prior to this date.

Signature

Name of the Candidate

VIGNESH R

21/12/2023

Date

Institute for Stem Cell Science and Regenerative Medicine

An autonomous Institutee of Department of Biotechnology, Government of India.

inStem CSCR CENTER FOR STEM CELL RESEARCH (a unit of inStem, Bengaluru) Christian Medical College Campus Bagayam, Vellore - 632002.

Center for Stem Cell Research

(A unit of inStem, Bengaluru in Collaboration with DBT and CMC, Vellore)

CERTIFICATE

Name of the Guide: Dr. Mohankumar K M
Division/Department: Centre for Stem Cell Research (a unit of inStem, Bengaluru)

This is to certify that **VIGNESH R**, Centre for Stem Cell Research of Christian Medical College has fulfilled the requirements prescribed for the Ph.D. degree of the Sree Chitra Tirunal Institute for Medical Sciences and Technology, Trivandrum.

The thesis entitled, "**Pre-clinical study for therapeutic genome editing in beta-haemoglobinopathies**", was carried out under my direct supervision. No part of the thesis was submitted for the award of any degree or diploma prior to this date.

*Clearance was obtained from the Institutional Ethics Committee / Institutional Animal Ethics / Institutional Committee for Stem Cell Research / Other appropriate committees (if any, specify) for carrying out the study.

Signature *of the Guide*

Dr. Mohankumar K. Murugesan, Ph.D.,
Assistant Investigator
Centre for Stem Cell Research (CSCR),
(A Unit of inStem, Bengaluru)
Christian Medical College Campus,
Bagayam, Vellore - 632 002, Tamil Nadu, India.

Date:21/Dec/2023

Christian Medical College Campus,
Bagayam, Vellore - 632002. INDIA.

Phone : +91 - 416 - 228 5101 / 5107
Fax : +91 - 416 - 228 5103
E-mail : cscr@cmcvellore.ac.in

The thesis entitled

PRE-CLINICAL STUDY FOR THERAPEUTIC GENOME EDITING IN BETA-HEMOGLOBINOPATHIES

Submitted by

VIGNESH R

for the degree of

Doctor of Philosophy

of

SREE CHITRA TIRUNAL INSTITUTE FOR MEDICAL SCIENCES AND TECHNOLOGY, TRIVANDRUM

is evaluated and approved by

(Name & Signature of the Guide)

Dr. Mohankumar K. Murugesan, Ph.D.,
Assistant Investigator
Centre for Stem Cell Research (CSCR),
(A Unit of inStem, Bengaluru)
Christian Medical College Campus,
Bagayam, Vellore - 632 002, Tamil Nadu, India.

14.3.2024

(Name & Signature of thesis examiner)

Dr. E. KOKILADEVI, Ph.D.,
Professor and Head
Department of Plant Biotechnology
Centre for Plant Molecular Biology & Biotechnology
Tamil Nadu Agricultural University
Coimbatore -641 003

ACKNOWLEDGEMENTS

‘Calm sea never made good sailors.’
-Frank D Rosevelt

I would like express my deepest gratitude and acknowledge the following individuals and institutions who have played a significant role in the completion of my research and thesis.

I would like to extend my sincerest gratitude to **Dr. Mohankumar** for guiding me and supporting me in performing all my experiments. I am also thankful for all the work discussions we had which enabled me to learn a lot. His support and encouragement to pursue ideas which have formed a part of this thesis is deeply appreciated. I also thank him for supporting me with all the reagents and other amenities required to complete this study successfully. Since I was part of the initial set-up in the lab, I had the opportunity to learn, manage and be a part of a team going forward.

I would like to thank **Dr. Alok Srivastava** for all the insightful discussions during my work-in progress meeting and Journal Club presentations which helped me understand the larger goal of research. I thank him for his help in sample collection during the study. I also thank him for giving me this opportunity to work in this Center for Stem Cell Department.

My sincere gratitude to my Doctoral Advisory Committee members, **Dr. Shaji, Dr. Saravanabhavan, Dr. Poonkuzali** for keeping me on track throughout the study and for providing valuable suggestion/advice through the course.

I thank DBT JRF fellowship for supporting me financially throughout the study. I would also like to thank all the funding agencies for supporting my research all these years. I would also like to thank my parents for supporting me financially for my

conference visit in Germany. I would like to thank SCTIMST for providing this wonderful opportunity to pursue my degree through this University.

I also thank other faculty members like **Dr. Srujan, Dr. Gurbind Singh, Dr. Sanjay Kumar, Dr. Sandhiya, Dr. Sunil Martin** and **Dr. Aloukick Singh** for their valuable discussion/advice/suggestions in all the department presentations throughout the study.

I thank my international collaborators, **Dr. Merlin Crossley, Dr. Jacob E Corn, Dr. Joel Mackay, Dr. Mitchel Weiss** and their students **Mahdi Hadad, Dr. Stacia Wyman**, and **Dr. Thiyagarajan** for their invaluable help in the successful completion of these projects.

I thank my friend **Mr. Chandrasekar** and **Dr. Rajasekar** for helping me in performing all the computational analysis. I am thankful to him for helping me every time I approached him even at untimely hours.

I would also like to extend my sincerest thanks to **Mr. Aswin Anand, Ms. Agila**, from Dr. Poonkuzali's lab and **Mr. Neelakandan** and **Mr. Gopi** of the Department of Hematology for their help in HPLC analysis. I also extend my gratitude to **Mr. Naveen, Mr. Silambu** and others involved in obtaining samples from the hospital campus.

Special thanks to the **reviewers** of my manuscript for their insightful suggestions and remarks which have shaped my study and my research perspective effectively.

'Science is a team game' and I am happy and honored to have played that game with my lab mates, **Dr. Nithin Sam, Nivedhitha, Anila, Kirti Prasad, Lokesh, Joshua, Sandhiya, Dr. Anji, Unnati, Bhanu, Nilofer, Shivani** and **Naser**. I have learnt a lot from all of you and thank you for your help and understanding all these years. Special gratitude to **Ms. Nivedhitha** for helping me in successfully completing a

chapter in this work. I thank **Ms. Sandhiya** and **Mr. Naser** for making all the purchase requests and ordering reagents and following through with it.

I thank my project students, **Ms. Shanmathi, Dr. Gowri, Mr. Karthik Lokatiya, Ms. Akshaya, Ms. Nikitha Rao, Ms. Keerthiga, Mr. Dhiyanesh, Mr. Harish**, and **Ms. Anusiya** for helping me in my study and for teaching me an important skill of teaching others. I thank all of you for your role in my professional life in CSCR. It was a privilege both teaching and learning from all you.

'We see further because we stand on the shoulder of giants' by Aristotle is a phrase I truly acknowledge. I am extremely grateful to all my seniors **Mr. Bhanuprasad, Mr. Harish, Mrs. Abisha Crystal, Dr. Abhirup Bagchi, Dr. Karthikeyan, Dr. Sowmya, Dr. Aneesha, Dr. Anji Anura** and **Dr. Sumitha** chechi for teaching me many techniques, clarifying doubts, and troubleshooting experiments throughout the journey. The soft skills I learnt from you will stay with me forever.

Any journey would be incomplete without companions, I would like to extend my special gratitude to my batchmates and friends, **Aishwarya, Nithin Sam, Saranya, Vigneshwaran** and **Rashmi**. Thank you for all the scientific discussion, trekking, outings, and every activity. I thank you sincerely for making life at CSCR worth it.

I am grateful to have made lot of good friends throughout this journey, **Ms. Nivedhitha, Dr. Karthik V K, Ms. Anila, Ms. Kirti, Mr. Lokesh, Mr. Joshua, Mr. Ashis, Mr. Karthik C, Dr. Sarubala, Mrs. Swati, Mrs. Tamilselvi, Ms. Porkizhi, Ms. Antra, Mr. Manoj, Ms. Annlin, Ms. Prathiba, Ms. Nishka, Ms. Smitha Ijee, Mr. Muthu, Mrs. Nimmy, Mr. Ajay, Ms. Arulmozhi, Ms. Sonam, Mr. Rishab, Mrs. Unnati, Mrs. Seventhy** and so many others. I thank everyone for the many joyous memories which will be cherished always.

Special thanks to my friends from inStem, **Mr. Karthik** and **Ms. Radhika** from **Dr. Dhandapani** lab. It was a pleasure working and learning from you. I would also like to thank **Ms. Priya Thakur** from **Dr. Sivaprakasam** Lab in IGIB.

I would also like to thank the CSCR Core facility, **Dr. Sandhiya, Mr. Rajesh, Mrs. Saranya, Mr. Abdul, Mr. Joel** and **Ms. Imani Job** for their help in all the flowcytometry and imaging analysis. **Mr. Augustine** and **Ms. Aleeya** for their valuable tips to in troubleshooting cell culture related problems. I extend my gratitude to CSCR Animal facility **Dr. Vigneshwar, Dr. Joe Varghese, Mrs. Pavithra, Mr. Ashok** and **Mrs. Esther** for their help in animal studies.

I extend my deepest gratitude to all the Administration staff of Center for Stem Cell Research, **Mr. Tamil, Mrs. Geetha, Mrs. Anu, Mrs. Selvi, Mrs. Shirley, Ms. Priya, Mr. Rajesh**, and **Mr. Muthukrishnan** for providing pleasant atmosphere and facility in the department. I would like to thank all the members of the housekeeping for providing a clean and efficient environment.

I am extremely grateful and blessed to have my family, **Mrs. Shanthi (amma)**, **Mr. Rajendiran (appa), Mrs. Sujitha (akka), Mr. Senthilkumar (mama)**, have trusted and supported me in all my decisions throughout my life. My niece, **Ms. Arshitha (achu)**, was born during the first year of my studies. Seeing your beautiful smile and your ‘mame’ has been the most welcoming and rejuvenating sound at the end of every tiring day. It has been an absolute pleasure seeing you grow from day 1.

My wife, **Mrs. Aishwarya Prasannan**, has been an integral part of my research journey, from being my best friend to becoming the best part of my life. I thank you for being beside me through all the ups and downs of this journey. I also thank you for

understanding and tolerating me. I am eagerly awaiting to start our life together.

I would always cherish the soccer games played with my friends, **Dr. Abhirup, Mr. Debanjan, Mr. Anurag, Mr. Gaurav, Mr. Ashis, Dr. Shaji** and **Sagar, Dhanush and his brother**. Thank you for helping me stay healthy and enjoy the sport I love. I also thank **Dr. Franklin, Saranya, Nithin, Vigneshwaran, Ashis** and others for the cheerful badminton time.

My study would have been complete without the generous donation of cells from all the **Donors** and **Patients**. Special thanks and apologies to all the **NSG** and **NBSGW** mice involved in the study, your life means a lot to me.

I grateful to **GOD** almighty for this opportunity and for providing **me** with good health, perseverance, patience, determination, and the data (obviously!) obtained to complete this degree. Lessons learnt here will be remembered and will forever help in improving my career in research.

Thank you all for everything!

Vignesh Rajendiran

Table of Contents

List of figures

List of Tables

List of Abbreviations

S.No	Abbreviations	Expansion
1	ABE	Adenosine Base Editor
2	ANOVA	Analysis of variance
3	BCL11A	B-cell lymphoma/leukemia 11A
4	cas9	CRISPR associated protein 9
5	CBE	Cytosine Base Editor
6	CD	Cluster of Differentiation
7	cDNA	complementary DNA
8	COS 7	CV-1 in Origin with SV40 genes
9	COSMID	CRISPR Off-target Sites with Mismatches, Insertions, and Deletions
10	CRISPR	Clustered Regularly Interspaced Protospacer Repeats
11	dcas9	dead cas9
12	DMEM	Dulbecco's Modified Eagle Medium
13	DNA	Dioxyribo Nucleic Acid
14	DSB	Double Strand Break
15	EditR	Editing analysis Algorithm using R program
16	EMSA	Electrophoretic Mobility Shift Assay
17	EPO	Erythropoietin
18	FBS	Fetal Bovine Serum
19	FLT3	fms like tyrosine kinase 3
20	GATA1	GATA-binding factor 1
21	HbA	Adult Hemoglobin
22	HBB	Beta globin subunit
23	HbF	Fetal Hemoglobin
24	HBG	Gamma globin subunit
25	HPFH	Hereditary Persistence of Fetal Hemoglobin
26	HSC	Hematopoietic Stem Cell

27	HSPC	Hematopoietic Stem and Progenitor Cell
28	HUDEP2	Human Umblical Cord-derieved Erythroid Progenitor 2
29	ICE	Inference of CRISPR Edits
30	IL3	Interleukin 3
31	IL6	Interleukin 6
32	IMDM	Iscove's Modified Dulbecco's Medium
33	KLF1	Kruppel-like factor 1
34	LRF	Lymphoma/leukemia-related factor
35	mRNA	messenger RNA
36	NBSGW	NOD.Cg- Kit W-41J Tyr + Prkdc scid Il2rg tm1Wjl /ThomJ
37	NGS	Next Generation Sequencing
38	OT	Off-Target
39	PBMNCs	Peripheral Blood Mono Nuclear Cells
40	PCR	Polymerase Chain Reaction
41	qRT-PCR	quantitative Real-Time Polymerase Chain Reaction
42	RNA	Ribo Nucleic Acid
43	RP-HPLC	Reverse-phase High Pressure Liquid Chromatography
44	SCD	Sickle Cell Disease
45	SCF	Stem Cell Factor
46	sgRNA	single Guide RNA
47	TAL1	T-cell acute leukemia protein 1
48	TPO	Thrombopoietin
49	ZBTB7A	Zinc Finger And BTB Domain Containing 7A
50	ZnF	Zinc Finger

SYNOPSIS

β-Hemoglobinopathies are caused due to mutations in the β-globin gene resulting in moderate to severe anemia, hepatosplenomegaly, and other secondary complications affecting people in different parts of the world and are common in India. Individuals affected will have decreased quality of life and shortened life expectancy without treatment. Sickle Cell Disorder (SCD) and ß-Thalassemia are the major ß-hemoglobinopathies affecting the ß-chain of the hemoglobin. SCD is caused by an E6V mutation in the exon 1 of beta-chain resulting in the polymerization of the adult hemoglobin (HbA) causing it to exhibit a sickle-shape. The mal-functional HbA which is sickle-shaped is termed sickle hemoglobin (HbS). ß-Thalassemia is caused by a spectrum of mutations leading to quantitative deficiency in beta-globin production. There is a severe imbalance between the alpha and beta chains of the hemoglobin in this diseased condition. Both these conditions severely compromise the quality of life of the patients.

Current treatment strategies involve the use of Hydroxyurea, HDAC inhibitors, and DNMT inhibitors for ameliorating the disease condition but the major caveat is the continuous mode of treatment. Myeloablative allogeneic stem cell transplantation from a matched unrelated donor (MUD) offers a single-shot treatment opportunity. However, the major limitation of this approach is the lack of a human leukocyte antibody (HLA) matched donor, immunologic side effects due to Graft Versus Host Disease (GVHD) or graft rejection. Gene therapy offers the best way to ameliorate the SCD and ß-Thalassemia pathogenesis by inducing higher levels of HbF as an alternative to the defective HbA without any GVHD and donor limitations. It provides

the patient with a one-shot treatment for these SCD and ß-Thalassemia conditions resulting in significant improvement in their quality of life.

There are 3 major types of hemoglobin (Hb) produced based on the developmental stage of the individual. The switch from one type to another is termed globin switching which is tightly regulated by multiple factors. Two globin switching have been identified in humans, one in the embryonic stage and the other in the fetal stage with a change from embryonic hemoglobin to HbF($\alpha 2\gamma 2$) which is again switched to HbA($\alpha 2\beta 2$) after birth. Different globin chain genes related to Hb are present in chromosome 16(ζ, $\alpha 1$, $\alpha 2$) and chromosome 11 (ε, γ, δ, β). In normal adults, HbF levels are below 1% of the total hemoglobin and are hetero-cellular in distribution. Induction of HbF has been the cornerstone of the treatment strategies to deal with hemoglobinopathies, which have malfunctioned HbA. In the case of SCD, increased HbF decreases the polymerization of deoxy sickle hemoglobin by forming a hybrid tetramer with sickle β-hemoglobin ($\alpha 2\beta S\gamma$). In the ß-thalassemia condition, the production of the gamma-globin chain sequesters the excess alpha chain and restores the balance between the globin chains.

Strategies to utilize HbF induction for therapeutic purpose for beta-hemoglobinopathies have been inspired from individuals and certain patients, known as Hereditary Persistence of Fetal Hemoglobin (HPFH) individuals, exhibiting higher levels of fetal hemoglobin in the adult stage without any clinical variations. Genome-wide analysis Studies (GWAS) have identified that these HPFH individuals harbor mutations in certain genes contributing to elevated fetal hemoglobin. Based on this data, many transcriptional factors like KLF1, BCL11A, LRF, GATA-1, and GATA-2, which are involved in the hemoglobin switching from HbF to HbA and repression of

γ globin have been identified. Hence, gene therapy/editing strategies involving targeting these factors or their regulatory regions have been employed successfully to ameliorate the disease condition by de-repressing fetal hemoglobin. Current targets associated with gene editing strategies are the BCL11A erythroid-specific enhancer disruption using ZFNs, CRISPR-cas9, and base editors or disrupting the binding site of BCL11A at the -115 cluster in the HBG promoter using both CRISPR-cas9 and base editors. shRNA-mediated downregulation of BCL11A is also clinically employed highlighting the importance of BCL11A in fetal hemoglobin repression.

B-cell Lymphoma 11A protein (BCL11A) is a major HbF gene regulator through direct binding mainly to the HBG promoter. Even though repression of HbF involves many such proteins, suppressing BCL11A alone has resulted in high HbF induction. Outside the globin gene regulation, BCL11A has been implicated in many other important physiological functions like Hematopoietic Stem Cell (HSC) self-renewal and lymphopoiesis. So complete knock-out of BCL11A will disrupt the normal functioning HSC by affecting its multi-lineage differentiation potential and engraftment capability. Studies have identified an erythroid-specific enhancer region within the intronic region of the BCL11A gene. Mutations within this region have been implicated in the loss of BCL11A expression in the erythroid lineage alone in HPFH patients. This region has been dissected further to identify key motifs capable of inducing high levels of HbF.

In this study, we propose to identify and evaluate alternate strategies for impacting BCL11A to achieve therapeutically relevant fetal hemoglobin levels. Towards this end, I have studied two different strategies for impacting critical regions within the BCL11A gene resulting in the induction of fetal hemoglobin in the adult

developmental stage. The first strategy is to target the C-terminal Zinc finger domains (ZnF) of the BCL11A-XL isoform (present exclusively in the adult stage) to disrupt the DNA recognition of BCL11A. This approach inhibits the repressional function of BCL11A thus elevating fetal hemoglobin. The presence of mutations within this region in intellectual disability patients inspired us to alter this region to evaluate its potential as a clinical target in fetal hemoglobin induction. The second strategy is to study the erythroid-specific enhancer of BCL11A to identify other target sites and unravel novel mechanisms for the enhancer function.

In both strategies, I have utilized both CRISPR/Cas9 and base-editing mediated approaches to create controlled insults in the target DNA region. The region was first evaluated for fetal hemoglobin induction using the Human Umbilical-cord Derived Erythroid Progenitor (HUDEP2) immortalized cell line through lentiviral transduction. Later the confirmed targets were evaluated in hematopoietic stem and progenitor cells (HSPCs) using either protein or mRNA-mediated gene editing. The in vivo homing, engraftment, and sustenance of the edited cells were evaluated through an animal study involving the NBSGW mice model. The molecular mechanism was also evaluated through biochemical assays to establish the mechanism of action.

In my first study of targeting the BCL11A-XL specific ZnF domains, we identified that base editing mediated sequence alteration has lesser consequences within the erythroid maturation than the cas9 mediated approach. We also modelled the binding of the ZnF-mutated BCL11A to the known -115 cluster at the HBG promoter and observed reduced binding. Interestingly we observed that altering the ZnF mutation had minimal impact on the transcriptome profile within the erythroid lineage compared to the total knockout of BCL11A and other clinical trial targets.

During engraftment analysis, we observed a significant loss of editing within the B-cell compartment compared to the bone marrow. This data suggests previously unreported function of the c-terminal ZnF domains of BCL11A-XL in B-cell production and HSC engraftment and maintenance.

In my other approach of identifying the alternate targets within the BCL11A erythroid-specific enhancer, I reported that the motif present +33bp upstream of the GATA motif (1617; clinical trial site) is not occupied by KLF1 protein yet it accommodates an unknown activator binding as reported by the ATAC-seq. cas9-mediated disruption of this site elevates significant levels of fetal hemoglobin as validated in multiple healthy donors. Interestingly we observed that on base-editing at this site using hypermutant adenosine base editor, we created a de-novo binding site for ZBTB7A factor. This ZBTB7A is a known transcriptional repressor. Through this approach, we interlink two independent gamma-globin repressors to mediate fetal hemoglobin de-repression. This strategy of targeting the BCL11A enhancer identified a novel mechanism through which the clinically utilized BCL11A erythroid-specific enhancer functions and new approaches for its fetal hemoglobin regulation function.

Both my studies will help in understanding the different aspects of the BCL11A gene from the perspective of its role in fetal hemoglobin regulation. Despite the therapeutic induction of HbF, the safety concerns caused by the impact on other lineages are a major limitation in clinical application. Overall, my study provides foundational knowledge on the functions of BCL11A and its erythroid-specific enhancer. Further optimizations within this region might result in the identification of efficient clinical trial targets bypassing the safety concerns highlighted in my work.

INTRODUCTION

Section-1

Introduction

β-Hemoglobinopathies are caused due to mutations in β globin gene resulting in moderate to severe anemia, hepatosplenomegaly, and other secondary complications affecting people in different parts of the world and are common in India. Individuals affected will have decreased quality of life and shortened life expectancy without treatment. Major examples of hemoglobinopathies are Sickle Cell Disorder (SCD) caused by an E6V mutation in chromosome 11 and ß-Thalassemia caused by multiple mutations leading to deficiency in the beta-globin. There is polymerization of the adult hemoglobin (HbA) causing it to exhibit a sickle- shape. The mal-functional HbA which is sickle-shaped is termed as HbS.

Hemoglobin is a tetrameric protein found in the erythrocytes helping in oxygen transport. Individuals with hemoglobin levels below the normal range are, by definition, anemic. There are 3 major types of hemoglobin (Hb) based on the stage of growth of the individual. The switch from one type to another is termed globin switching. Two globin switching have been identified in humans, one in the embryonic stage and the other in the fetal stage. During the embryonic stage, embryonic hemoglobin switches to fetal hemoglobin (HbF). After birth fetal hemoglobin (HbF: $\alpha 2\gamma 2$) switches to Adult hemoglobin (HbA: $\alpha 2\beta 2$). Different globin chain genes related to Hb are present in chromosome 16(ζ, $\alpha 1$, $\alpha 2$) and chromosome 11 (ε, γ, δ, β). In normal adults, HbF levels are below 1% of the total hemoglobin and are hetero-cellular in distribution. Induction of HbF has been the corner stone of treatment strategy to deal with hemoglobinopathies, which have malfunctioned HbA. Increased HbF decreases the polymerization of deoxy sickle hemoglobin. HbF forms a hybrid tetramer with

sickle β hemoglobin ($\alpha_2\beta_S\gamma$) which will not undergo sickling when compared to the hybrid tetramer with sickle β_S and normal β_A hemoglobin.

Myeloablative allogeneic stem cell transplantation from a matched related donor remains the current mode of treatment in use for β-hemoglobinopathies. The progression of SCD toward end-organ damage is reversed by allogeneic stem cell transplantation. However, the major limiting factors in this approach include the lack of a HLA matched donor, and immunologic side effects due to GVHD or graft rejection.

Gene therapy offers the best way to ameliorate the SCD and ß-Thalassemia pathogenesis by improving the HbF levels without graft versus host disease and donor limitations. Increasing the HbF levels greatly improves the quality of life in patients suffering from SCD and B-Thalassemia. Strategies to increase the HbF expression levels have been identified through clinical features of patients exhibiting Hereditary Persistence of Fetal Hemoglobin (HPFH) and other compound mutations. Based on these data many transcriptional factors like KLF1, BCL11A, LRF, GATA-1, and GATA-2, which are involved in the switching of expression from HbF to HbA and repression of γ globin have been identified. Targeting these factors for the elevation of HbF levels is being utilized clinically through compounds like Hydroxyurea, HDAC inhibitors, and DNMT inhibitors even though their exact mechanism of action is unknown.

Advances in gene-editing like CRISPR/cas9 have been explored to disrupt these factors for therapeutic induction of HbF levels. The creation of INDELS (INsertions and DELetions) within the target region has knocked out the functionality of the gene. Modifications in the cas9 like nickase which is capable of cleaving single

strands of DNA and dead cas9 (dCas9) where the catalytic activity of the cas9 protein has been removed, thus making it a homing molecule for compounds. Base Editors are nickase cas9 variants conjugated with cytosine deaminase or adenine binding enzyme which converts Cytosine to Thymine and Adenine to Guanine respectively. Modulating transcriptional repressor(s) expression using CRISPR/Cas9 is a major treatment strategy for beta-hemoglobinopathies.

B-cell Lymphoma 11A protein (BCL11A) is a well-known fetal hemoglobin regulator. Its role in γ to β switching has been reported and widely studied. Even though repression of HbF involves many such proteins, suppressing BCL11A alone has resulted in high HbF induction. BCL11A has been implicated in many important physiological functions like Hematopoietic Stem Cell (HSC) self-renewal, lymphopoiesis, and differentiation into other lineages. So complete knock-out of BCL11A leads to disruption of the normal functioning HSC by affecting its multi-lineage differentiation potential and engraftment ability. GWAS studies have identified an erythroid-specific enhancer region within the BCL11A intronic region and mutations within this region have been implicated in loss of BCL11A expression within the erythroid lineage in HPFH patients. This region has been dissected further using CRISPR/cas9 to identify key motifs capable of inducing high levels of HbF which can be utilized as a safe site for therapeutic fetal hemoglobin induction. The GATA motif present within the DHS58 of the BCL11A enhancer has been implicated in regulating the enhancer. Thus, disruption of this site using various gene editing strategies has been utilized to achieve therapeutic benefits in beta-globin disorders. The precise regulatory mechanism of the BCL11A enhancer is still elusive outside the GATA motif present within the DHS58 region.

Recent studies have identified the direct repression of fetal hemoglobin by BCL11A through direct binding in the *HBG*1/2 promoter. This study has reported that disrupting this binding site (-115 cluster: TGACCA) within the *HBG*1/2 promoter induces therapeutically relevant levels of fetal hemoglobin. The highly homologous region of the *HBG*1 and *HBG*2 promoter sequence causes significant levels of 4.9kb intermittent region deletion while using CRISPR/cas9. This was avoided or reduced drastically while using the base editor system. The usage of an adenosine base editor is reported to generate a de novo GATA1 binding site within the BCL11A binding site. This leads to displacement of a repressor with an activator resulting in elevated levels of *HBG* activation. BCL11A-XL is a major transcriptional factor that binds to many regulatory sites other than the *HBG*1/2 promoter site. The presence of 3 unique ZnF domains within the BCL11A-XL gene, provides a unique opportunity to disrupt all interactions of BCL11A with DNA sequence. Mutations within these ZnF domains and the resulting BCL11A variants are yet to be understood.

Towards this end, my study is designed to understand the function of critical regions within the BCL11A gene furthering our existing understanding of the BCL11A in globin regulation through genome editing platforms. We aimed to achieve this through 2 studies, first by disrupting the BCL11A-XL specific ZnF domains using various gene editors and second by disrupting the other transcription factor binding sites within the BCL11A erythroid-specific intronic enhancer to understand its regulatory mechanism.

We performed genome editing in the Human Umbilical Derived Erythroid Progenitors (HUDEP-2) cell line and later validated the results in HSPCs using both CRISPR/cas9 and base editors. The long-term engraftment and differentiation

potential of these genome-edited HSPCs were evaluated in the NBSG mice model. We performed various biochemical assays to understand the variant produced or the mechanisms through which the target site functions.

Overall, this study sheds light on the role of various domains within the BCL11A gene namely the Zinc Finger binding domains and the intronic enhancer.

Aim of the study

To understand the role of various critical regulatory regions within the BCL11A gene using genome editing platforms.

Objectives

Study 1:

To study the structure-function relation of the BCL11A-XL-specific ZnF domains

1. Evaluating the targets in HUDEP2 and HSPCs using CRISPR/cas9
2. Evaluating the targets in HUDEP2 and HSPCs using Base editing
3. Engraftment and multi-lineage differentiation potential in NBSGW mice
4. Molecular mechanism identification

Study 2:

To study the regulatory role of the putative KLF1 binding site(s) present adjacent to the GATA-motif within the DHS58 of the BCL11A intronic enhancer

1. Targeted disruption of KLF sites in HUDEP2 cells
2. Validation in HSPCs
3. Elucidating the regulatory mechanism
4. Evaluation of base editing at the target site

LITERATURE REVIEW

Section-2

Literature Review

Beta-Hemoglobinopathies

The most common genetic blood disorders are sickle cell disease and beta-thalassemia. Sickle cell disease is a monogenic disorder caused by a qualitative defect in the adult hemoglobin caused by a single mutation(E6V) in the beta chain resulting in the production of a defective variant of the adult hemoglobin, HbS ($\alpha_2\beta_2^s$). This single amino acid change causes the hemoglobin to polymerize with each other, a condition known as sickling, reducing the life span and the oxygen-carrying capacity of the erythrocytes. Beta-thalassemia is a quantitative disorder caused by a spectrum of mutations in the beta gene resulting in the reduced production of functional beta-chain. The incidence rate of these genetic blood disorders is high among the common genetic disorders. The imbalance between the alpha and the beta globin chain of the adult hemoglobin is the major causative reason for this severe condition among patients. Based on the amount of beta-chain produced beta thalassemia condition is clinically classified as either $\beta^0\beta^0$, $\beta^0\beta^{+,}$ and β^+ β^+B+. The hemoglobin content in the erythrocytes is severely reduced and the elevated free iron and the resultant reactive oxygen species produced also adds to the severity of the diseased phenotype. The onset of beta-hemoglobinopathy symptoms is observed more in adults than in the fetal stages due to the presence of compensating fetal hemoglobin. Certain individuals harboring mutations in certain regions of the genome have been identified to exhibit persistent expression of fetal hemoglobin in the adult stage. These individuals are termed as HPFH individuals or Hereditary Persistence of Fetal Hemoglobin individuals. In certain rare situations, Beta-hemoglobinopathy patients harboring these beneficial

HPFH mutations have exhibited milder clinical features or are asymptomatic of their condition. This led to the idea that the reversal of the fetal hemoglobin switch would ameliorate the beta-hemoglobinopathy condition.

Globin Switching

Human beings exhibit 3 different types of hemoglobin during various developmental stages. Embryonic hemoglobin (Hb Gower 1, Hb Gower 2, and Hb Portland, sequential appearance) switches to fetal hemoglobin (Hb F). After birth, there is a hemoglobin switch from fetal hemoglobin (HbF) to Adult hemoglobin (HbA1 and HbA2). The Gamma globin expressed during the fetal stages till birth is gradually replaced by the Beta-globin resulting in the switching from fetal hemoglobin HbF ($\alpha 2\gamma 2$) with adult hemoglobin HbA ($\alpha 2\beta 2$). These different hemoglobins have different affinities to the oxygen molecule and are present in different stages of development ranging from embryonic to fetal to the adult stage. Genes related to Hb are present in chromosome 16(ζ, $\alpha 1$, $\alpha 2$) and chromosome 11 (ε, γ, δ, β). In normal adults, HbF levels are below 1% of the total HbF and are hetero-cellular in distribution (**Fig 1**).

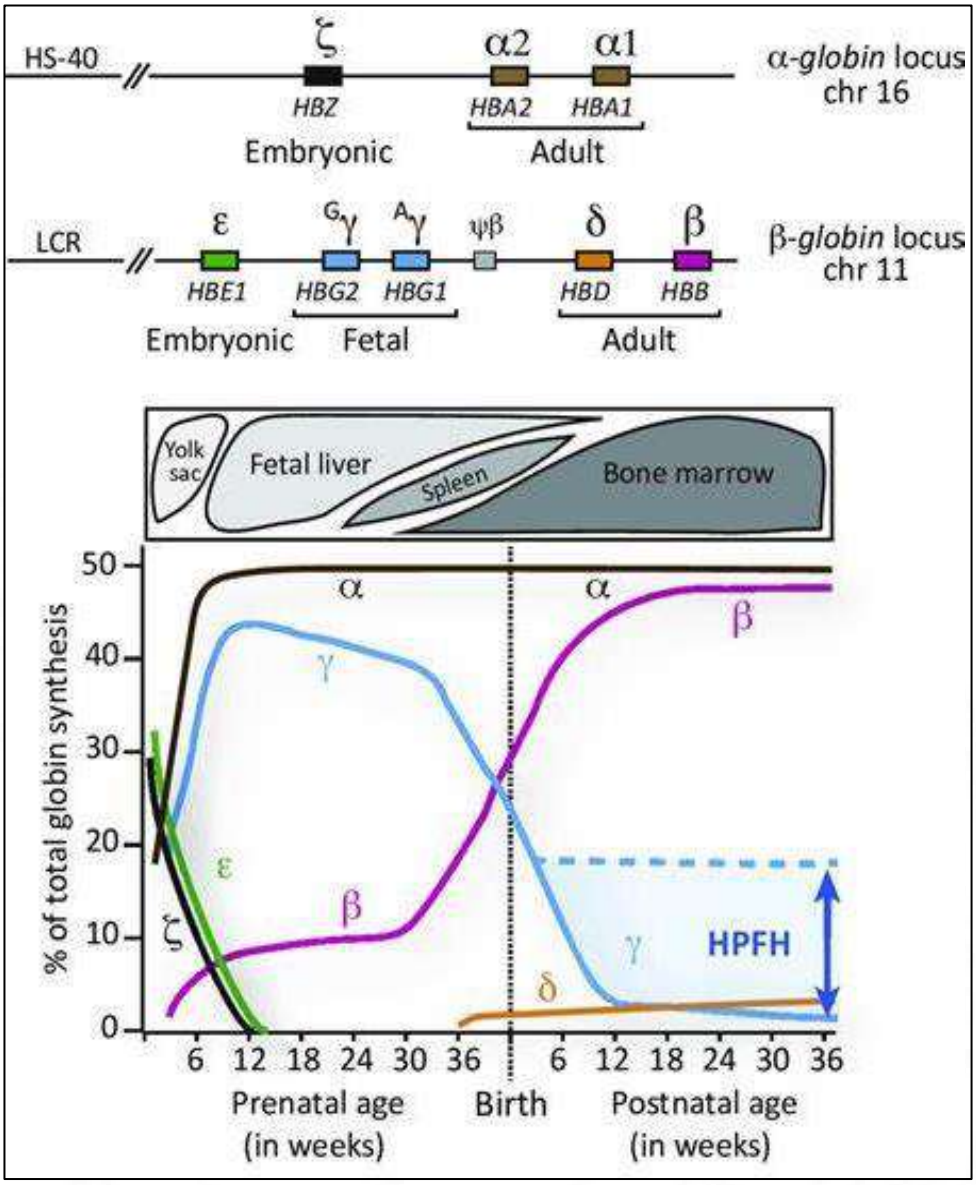


Fig 1: Globin switching during developmental stages also depicts the HPFH condition. Source: (Wienert et al., 2018)

Current treatment strategies

The major therapeutic strategy for the treatment of beta-hemoglobinopathy is the re-induction of fetal hemoglobin as a functional alternative for the defective beta-globin chain (**Fig 2**). Many compounds like Hydroxyurea (HU), HDAC inhibitors, and DNMT inhibitors have FDA approval for patient treatment but the effectiveness of treatment is varied due to the genetic background of the patient. It is reported that

individuals with (Xmn1 polymorphism) have better responses to the HU treatment than non- carriers of this mutation. Allogenic Stem cell transplantation offers a more comprehensive treatment strategy to overcome this disorder, but the major limitation is the availability of allogenic matched donors and graft acceptance.

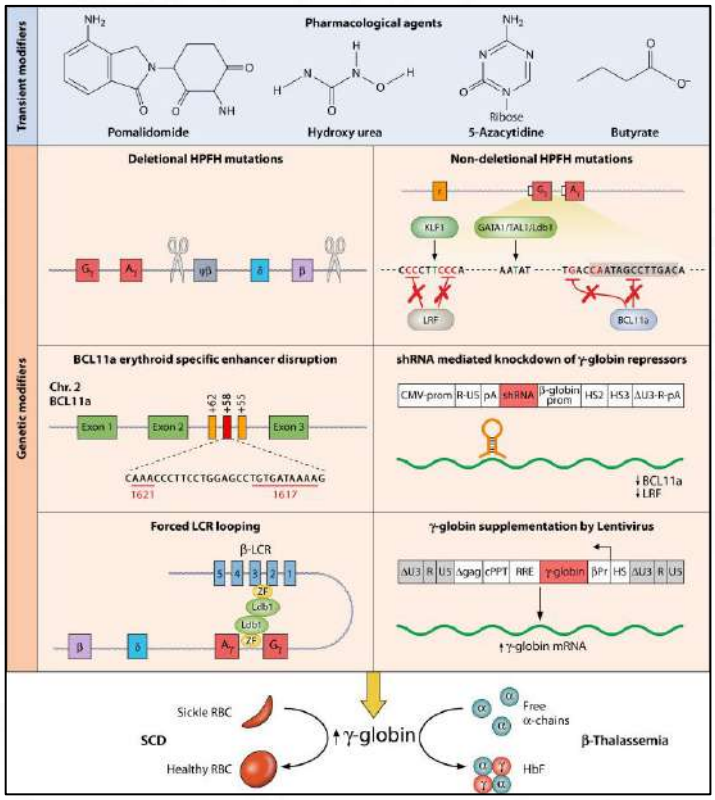


Fig 2: Current treatment strategies for Beta-globinopathies.
Source:(Venkatesan et al., 2021)

Gene therapy/editing with autogenic transplantation offers a promising approach to solving the major limitations of allogenic stem cell transplantation. Gene correction is an efficient strategy to correct genetic mutations like SCD but in the case of thalassemia which is caused by a spectrum of mutations a more generalised approach is required. The addition of a complete gene (Beta or Gamma globin gene) can be an alternative strategy to treat these patients but the efficiency of inserting and the mode of insertion (viral transduction) is a major concern of this approach though

there are multiple clinical trials in progress. The therapeutic induction of fetal hemoglobin is the third and most sought-after approach in the gene therapy strategies for beta-hemoglobinopathies. Fetal-to-adult hemoglobin switch is a tightly regulated process (**Fig 3**), involving many transcription factors some of which are still being identified. Hence targeting these transcription factors has proven to be an efficient way of elevating fetal hemoglobin. This is the most promising approach currently in clinical trials but the major caveat in this approach is identifying the targetable transcription factor with a limited range of activity outside the erythroid-lineage and the genotoxicity of the approach.

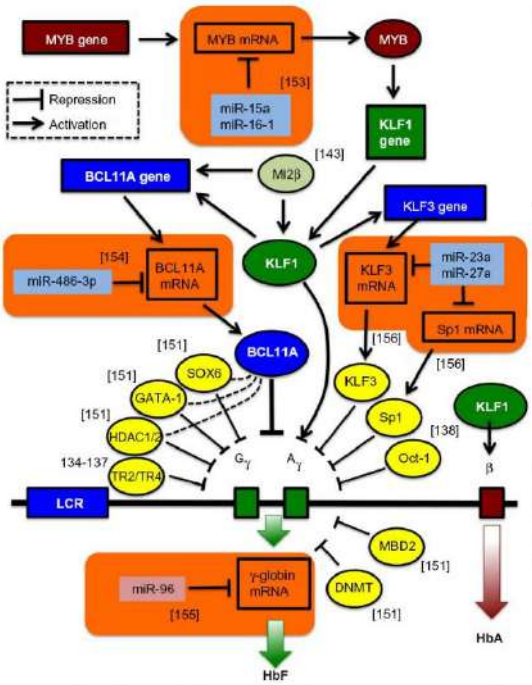


Fig 3: Tight regulation of Fetal-Adult switching by a multitude of TFs.
Source:(Finotti et al., 2015)

The two major strategies to induce fetal hemoglobin levels in adult erythroid cells are focused on either modulating the level of gamma-globin repressors or designed to alter the binding sites of gamma-globin regulators by mimicking the naturally occurring HPFH mutations.(Demirci et al., 2021a) The resulting elevation in

fetal globin expression exhibits anti-sickling potential in SCD and decreases the globin chain imbalance in β-thalassemia.(Davis et al., 2019) Thus, an ex vivo genetic modification of hematopoietic stem and progenitor cells (HSPCs) to reactivate the fetal hemoglobin expression would be an effective therapeutic strategy for the treatment of beta-hemoglobinopathies.(Ferrari et al., 2021)

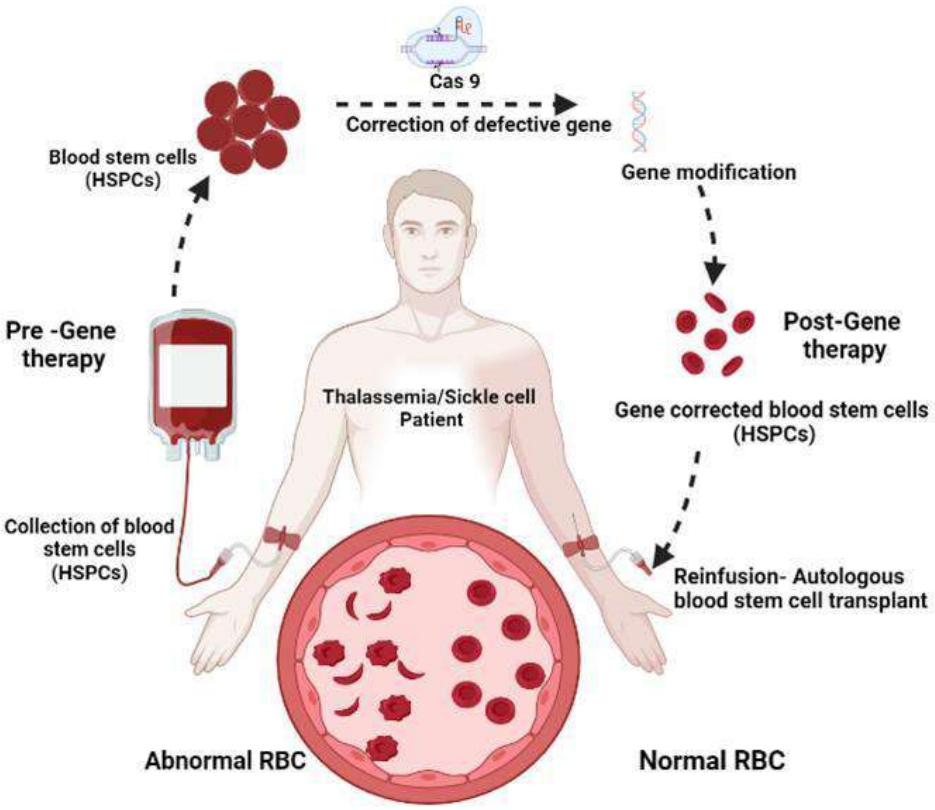


Fig 4: Ex-vivo gene therapy strategy

Role of BCL11A in Transcriptional Regulation of fetal hemoglobin

BCL11A (also known as CTIP1) is one of the major repressors involved in the gamma-globin expression. It is a highly conserved regulatory C2H2 type zinc finger protein expressed from Chromosome 2 of the human genome with substantial expression in multiple tissues such as brain, liver, B-cells, bone marrow, erythroid, and hematopoietic stem cells (Yin et al., 2019). Four major isoforms of BCL11A are generated including BCL11A XL, BCL11A L, BCL11A S, and BCL11A XS by alternative splicing (**Fig 5**).(Satterwhite et al., 2001) All the isoforms share a common C2HC domain responsible for the protein-protein interactions at the N terminus but have a variable number of C2H2 zinc fingers domains at the C terminus required for binding to the target DNA motif. The longer isoforms are present in the nucleus and the shorter isoforms are localized in the cytoplasm which modulates the transcription by the formation of co-repressor complexes(Liu et al., 2006a). Interestingly, the differential expression of BCL11A isoforms is observed during development which could explain their role in gamma-globin switching. Specifically, the shorter isoforms S and XS are shown to be expressed in the yolk sac and fetal liver, whereas the longer isoforms XL and L are expressed in the adult bone marrow (Sankaran et al., 2008).

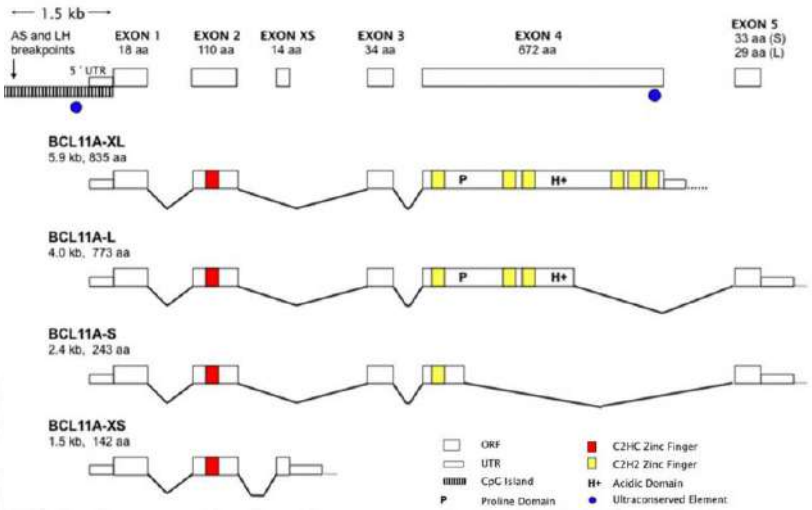


Fig 5: Various isoforms of BCL11A. Source: (Liu et al., 2006)

QTL study identifying genetic mutations associated with increased F cell production reported the existence of an intronic region later recognized and studied as an erythroid-specific enhancer of the BCL11A gene (Menzel et al., 2007). This identification of the role of BCL11A in fetal hemoglobin repression opened the flood gates for further research in unraveling its role in developmental hemoglobin switch. ChIP-seq analysis of the beta-globin cluster reveals that the BCL11A binds to many regions like the beta-globin, LCR, embryonic globin, and the intergenic regions between gamma and delta globin (Xu et al., 2010). It has also been identified that loss of BCL11A changes the chromatin looping formation at the gamma-globin locus. Systemic loss of BCL11A resulted in the phenotypic changes of HSC causing it to be more aged-like and have biased differentiation(Luc et al., 2016). Lin28b expression represses the expression of BCL11A-XL in the fetal stage by binding to the XL motif (Basak et al., 2020). Further, the functional importance of the role of BCL11A isoforms in gamma-globin regulation is evident by the shRNA-mediated specific

knockdown of BCL11A XL for the treatment of beta hemoglobinopathies. (Brendel et al., 2020)

Genome-wide association studies of sickle cell patients with HPFH conditions revealed that BCL11A as the major quantitative trait loci responsible for the higher levels of HbF.(Basak et al., 2015; Funnell et al., 2015; Uda et al., 2008) Three clusters of genetic variants in BCL11A are associated with a higher level of HbF includes the intronic SNPs in BCL11A enhancer (Bauer et al., 2013), exonic SNPs at ZNF 4-5 domain in BCL11A XL isoform (Shen, Li, et al., 2021) and the regulatory SNPs in *HBG* promoter that disrupts the binding site of BCL11A XL. The intronic SNPs located at the +55, +58, and +62 DHS of BCL11A erythroid enhancer, encompass the regulatory regions required for binding of various erythroid transcriptional factors like GATA1 and TAL1.(Canver et al., 2015) Recent studies revealed that targeting the +58 DHS core motif of the BCL11A enhancer region downregulates the BCL11A expression in erythroid lineage which causes simultaneous reinduction of HbF clinically without altering the expression of BCL11A in non-erythroid cells.(Demirci et al., 2020; Psatha et al., 2018; Wu et al., 2019) These regions were further subjected to cas9-mediated screening to identify the critical bases required within these. The DHS58 region was identified to be the most crucial in humans having the maximum impact in BCL11A expression. Recently the DHS+55 region has also been implicated to have a prominent effect in the enhancer activity through the presence of KLF1 sites and the HIC2 interaction(Huang et al., 2022). The GATA motif present in the DHS58 has been targeted extensively, for its effective role in BCL11A enhancer function, to elevate fetal hemoglobin expression within the erythroid lineage. This site has been tried and tested through various modalities of gene editing like CRISPR-cas9 and base

editing. Currently, Vertex therapeutics and CRISPR Therapeutics autologous product, Exa-cel or Casgevy, HSPCs with gene editing at the BCL11A enhancer site has been approved by both the FDA and the EMA.

The regulatory SNPs located at the -115 cluster upstream of the gamma-globin promoter including –114C>A, –114C>T, –114C>A, –117G>A, and Δ13bp deletion are shown to alter the core binding consensus motif of BCL11A and substantially elevate the expression of HbF.(Demirci et al., 2021b; Doerfler et al., 2021; Lux et al., 2019; Martyn et al., 2019; NS Ravi et al., 2022; Wang et al., 2020)

Current clinical trials (NCT03745287, NCT04211480, NCT04208529, NCT0443904, NCT04390971, NCT03653247, NCT03432364, NCT03655678) (**Table 1**) in patients with beta hemoglobinopathies for the reactivation of HbF by genome editing emphasize the significance of the SNPs in BCL11A enhancer and BCL11A binding site regions.(Zittersteijn et al., 2021)

Phase	Product name	Clinical trial	Target cell	Delivery	Nuclease	Target gene & effect	Disorder	Status	Sponsor	Results
I/II	ST-400	NCT03432364	CD34+	Electroporated mRNA	ZFN	Disruption of erythroid enhancer of *BCL11A* gene	TM	Active/Not recruiting	Sangamo	**Five patients (3/5 prelim. data): 1/3**–23% on-target indels 2.7 g/dl HbF (0.9 g/dl at baseline); **2/3**–73% on-target indels <1 g/dl HbF; **3/3**–54% on-target indels 2.8 g/dl HbF
I/II	BIVV003	NCT03653247	CD34+	Electroporated mRNA	ZFN	Disruption of erythroid enhancer of *BCL11A* gene	SCD	Recruiting	Bioverativ (Sanofi)	No results published yet
I/II	CTX001 (CLIMB-111)	NCT03655678	CD34+	Electroporated RNP	CRISPR-Cas9	Disruption of erythroid enhancer of *BCL11A* gene	TM	Recruiting	Vertex pharmaceuticals inc.	13 patients treated. TM: 7/13 TI with 3-18 months of follow-up. Total Hb from 9.7 - 14.1 g/dL and fetal Hb levels from 40.9 - 97.7%
I/II	CTX001 (CLIMB-121)	NCT03745287	CD34+	Electroporated RNP	CRISPR-Cas9	Disruption of erythroid enhancer of *BCL11A* gene	SCD	Recruiting	Vertex pharmaceuticals inc.	6 patients treated. VOC-free: 3/3; 3-15 months after CTX001 infusion. Total Hb from 11.5 - 13.2 g/dL and fetal Hb levels from 31.3 - 48%.
I/II	-	NCT04211480	CD34+	Electroporated RNP	CRISPR-Cas9	*BCL11A* binding site disruption *HBG* promoter	TM	Recruiting	Shanghai Bioray Laboratory Inc.	**2 patients:** 1 MPT HbF levels increased to 76 and 97 g/L HbF (total Hb 129 and 115 g/L, resp.) TI at 75 DPT
Long-term follow-up	CTX001	NCT04208529	CD34+	Electroporated RNP	CRISPR-Cas9	Disruption of erythroid enhancer of *BCL11A* gene	TM & SCD	-	Vertex pharmaceuticals inc.	-

TM, beta thalassemia major; TI, Transfusion independent; DPT, Days post-treatment; MPT, Months post-treatment; VOC, Vaso-occlusive crises.

Table 1: Clinical Trials based on Gene Editing. Source:(Zittersteijn et al., 2020)

With the advent of advanced gene editing tools and the extensive possibility of applications, the need to understand the mechanism of regulation of the target gene is

the need of the hour. This will enable us to validate and evaluate the outcomes of gene editing modulating the target gene in patients. Thus, improving the safety of the available treatment strategies.

Recent alternatives for downregulation of BCL11A like the use of PROTAC system (Mehta et al., 2022) or the nanobody (Shen et al., 2022a) (Yin, Izadi, Tenglin, Viennet, Zhai, Zheng, Arthanari, Laura M. K. Dassama, et al., 2023) mediated downregulation system provide various alternatives for the existing gene editing system. The models hypothesizing the mechanism of action of BCL11A have been described. From this we can strongly agree that despite our knowledge of the BCL11A gene, and its enhancer, the precise knowledge of its regulation is still elusive, probing more studies on the regulatory regions within the gene.

MATERIAL AND METHODS

Section-3

Material and Methods

3.1 sgRNA Plasmids and lentivirus production

The target gRNAs (Table:1) were cloned into the pLKO5.sgRNA.EFS.GFP plasmid (a gift from Benjamin Ebert, Addgene #57822). The gRNA-containing plasmids were then co-transfected with the VSV-G plasmid, pMD2.G (a gift from Didier Trono, Addgene #12259), and the envelope plasmid, psPAX2 (a gift from Didier Trono, Addgene #12260), at a 2:1:1 ratio in HEK293T cells using the Mirus transfection reagent (1:3 ratio of DNA: reagent). Viral supernatants were collected at 48- and 72-hours post-transfection, concentrated using an in-house preparation of PEG-6000 concentrator, and stored at -80°C until use.

Table 2: List of guide RNAs utilized

Guide RNA	sgRNA seq 5'-> 3'
AAVS1	**GGGGCCACTAGGGACAGGAT**
CCR5	**AAAATATAATCTTTAAGATA**
BCL11A In Cd (ABE8e)	**AGACATGGTGGGCTGCGGGG**
BCL11A exon2 (cas9)	**TGAACCAGACCACGGCCCGT**
Enhancer(cas9)	**CTAACAGTTGCTTTTATCAC**
Enhancer(ABE8e)	**TTTATCACAGGCTCCAGGAA**
ZnF4/41	**GTAGCAATCTCACTGTCCAC**
ZnF42	**TGTCCACAGGAGAAGCCACA**
ZnF5/51	**CCCAGAGTAGCAAGCTCACC**
ZnF52	**ACCAGGCACATGAAAACGCA**
ZnF6/61	**TGGAGAAACACATGAAAAAA**
ZnF62	**AGGGTACTGTACACGCTAAA**
BS(cas9)	**CTTGTCAAGGCTATTGGTCA**
BS(ABE8e)	**CTTGACCAATAGCCTTGACA**
+33bp GATA	**GGTTTGGCCTCTGATTAGGG**
CB11	**TTGGCCTCTGATTAGGGTGG**
CB13	**GATTAGGGTGGGGGCGTGGG**
CB15	**TTAGGGTGGGGGCGTGGGTG**

CB18	**CACGCCCCCACCCTAATCAG**
Hoxa13	**CGCCTTCGCGGACAAGTACA**
HOXb2	**CCCCGGCCCCCGAGTTCCCT**
HOXd3	**TCCTGAGTGCACAATGCAGA**
GSX1	**GGACTCGCTAGTGCTGCGCG**

3.2 Cell culture: Cell lines and Primary Cells

HEK293T was cultured using 10% fetal bovine serum supplemented DMEM media with 1% Penstrep antibiotics and incubated at 37°C in 5% CO2. Subculture was performed by trypsinized in 0.05% Trypsin (Gibco) at 37°C for 1 min for passaging.

HUDEP2 (immortalized Human Umbilical cord-derived erythroid progenitor cells) cells were cultured and differentiated as described previously. HUDEP2 cells, stably expressing Cas9 or ABE8e were generated and validated.(N Ravi et al., 2022) HUDEP2 lines were transduced with the gRNA lentivirus, and the transduction efficiency was determined by flow cytometry analysis of GFP expression after 48 hours.

$CD34^+$ HSPCs were isolated from the leftover G-CSF mobilized blood of healthy donors after infusion to patients with informed consent as per clinical guidelines authorized by the Intuitional Review Board, Christian Medical College, Vellore. The PBMNCs were separated using density gradient centrifugation (Lymphoprep Density gradient medium; Stemcell Technologies) and washed with 1XRBC lysis buffer to remove any residual RBCs present. $CD34^+$ HSPCs were purified using EasySep Human CD34 Positive Selection Kit II (Stemcell Technologies). The isolated $CD34^+$ cells were expanded and differentiated into the erythroid lineage as reported earlier

and the growth kinetics were measured by cell counting at each media change.(Devaraju et al., 2022)

COS-7 cells were cultured in DMEM (Gibco) supplemented with 10% FBS (Gibco) and 1% Penicillin-Streptomycin-Glutamine (PSG, Gibco) and incubated at 37°C in 5% CO2. Cells were lifted for passaging by incubation in 0.05% TrypLE™ Express Enzyme (Gibco) at 37°C for 5 min.

Table 3: HUDEP2 cell culture Media components

Cell Type	Media Component	Working Concentration
HUDEP2 Expansion media	Stemspan SFEM II	Basal
	Penstrep	1%
	L-Glutamine	10%
	Doxycycline	1ug/ml
	Dexamethasone	1uM
	Erythropoietin	3U/ml
	Stemcell Factor	50ng/ml
HUDEP2 Erythroid Differentiation Media	IMDM-GlutaMax	Basal
	Penstrep	1%
	AB serum	3%
	Fetal Bovine Serum	2%
	Holotransferin	200ug/ml (Increased to 500ug/ml on day 6)
	Erythropoietin	3u/ml
	Stemcell Factor	5ng/ml
	Heparin	
	Doxycycline	1ug/ml
	IL3	10ng/ml

Table 4: HSPC Culture Media Components

Cell Type	Media Component	Working Concentration
HSC Expansion media	Stemspan SFEM II	Basal
	Penstrep	1%
	IL6	10%
	Stemcell Factor	240ng/ml
	TPO	16ng/ml
	FLT3	240ng/ml
HSPC Erythroid Differentiation Media	IMDM-GlutaMax	Basal
	Penstrep	1%
	AB Serum	5%
	Holotransferin	330ug/ml (Increased to 500ug/ml)
	Erythropoietin	3U/ml
	Stemcell Factor	100ng/ml
	Heparin	
	Hydrocortisone	
	IL3	50ng/ml

3.3 ABE8e mRNA production

To produce ABE8e mRNA, we used the ABE8e plasmid (a gift by David Liu, Addgene#138495)(Richter et al., 2020) as the template and performed in vitro transcription as previously described. The resulting mRNA was stored in smaller aliquots at -80°C until use.

3.4 Nucleofection:

For Cas9 experiments, 0.2 million HSPCs were electroporated with a Cas9: gRNA ratio of 50:100pmoles using a 4D-Lonza nucleofector (pulse code: DZ100). For base editing experiments, we nucleofected 1 million HPSCs with 5microgram of ABE8e

mRNA with 100pmoles of gRNA using a GTx electroporator (MaxCyte Inc., pulse code: HSC-3).

3.5 Gene modification analysis:

Genomic DNA from the edited and unedited cells was isolated and the target regions were PCR amplified for Sanger sequencing using the primers listed in the supplementary table 2. The efficiency of gene modification generated by Cas9 was analyzed using the Synthego ICE program.(Conant et al., 2022) To estimate the base conversion percentage, the Sanger sequencing files were analyzed using the EditR analysis software.(Kluesner et al., 2018)

Table 5: Primers used for genotyping.

	Sample	Fwd primer	Rev primer
Sanger Sequencing	AAVS1	GAGATGGCTCCAGGAAATGG	ACCTCTCACTCCTTTCATTTGG
	Initiation CD	CATCTTCCCTGCGCCATCTTTG	CTCTCTCCCCCTCGCTTTTG
	Exon2	ATGGGGTTGAGATGTGCTTC	ACTGCTTGGCTACAGCACCT
	BCL11A enhancer	TCAAACCACAGGGATCACAA	AGAGAGCCTTCCGAAAGAGG
	ZnF4/5	CCCTTCTCTAAGCGCATCAA	TCCAGGGGTACTGTACACGCTAA
	ZnF6	GCCCAGAGTAGCAAGCTCAC	ATCCTACAGGGAGTGGGGCT
	BS	ACAAAAGAAGTCCTGGTATC	CTTCCCAGGGTTTCTCCTCC
NGS Primers	BCL11A enhancer	TACACGACGCTCTTCCGATCT CAGTGCAAAGTCCATACAGGT	AGACGTGTGCTCTTCCGATCTGG CAAACGGCCACCGAT
	ZnF4/5	TACACGACGCTCTTCCGATCT TGGAGGGAGCACGCCCCATAT	AGACGTGTGCTCTTCCGATCTCTC CAGGGTACTGTACACGC

COS cell transfections and nuclear extractions:

COS-7 cells were transiently transfected with pcDNA3-based plasmids to overexpress BCL11A zinc fingers (ZnF) 4-6 and mutants. Cells were transfected in 10 cm plates with 5 μg of plasmid using FuGENE® 6 Transfection Reagent (Promega). Mammalian expression plasmids used are listed in Table 4. Cells were washed in PBS after 48 h incubation at 37°C. Cells were harvested and resuspended in hypotonic lysis buffer (10 mM HEPES-KOH, pH 7.9, 1.5 mM MgCl2, 10 mM KCl, 5 mM dithiothreitol (DTT), 1 mM phenylmethylsulphonide (PMSF), 0.01 mg/mL aprotinin, 0.01 mg/mL leupeptin) and incubated on ice for 10 min. Cells were spun down and the pellets were resuspended in extraction buffer (20 mM HEPES-KOH pH 7.9, 25% glycerol, 420 mM NaCl, 1.5 mM MgCl2, 0.2 mM EDTA, 5 mM DTT, 1 mM PMSF, 0.01 mg/mL aprotinin, 0.01 mg/mL leupeptin) on ice for 20 min. The suspension was centrifuged at 13,000 rpm for 3 min at 4°C and the supernatant containing nuclear extracts recovered.

Electrophoretic mobility shift assay

Oligonucleotides used in the synthesis of radio-labeled probes are listed in Table 2. The sense oligonucleotide was labeled with ^{32}P from γ-^{32}P ATP (Perkin Elmer) using T4 PNK (New England Biolabs) before the antisense oligonucleotide was annealed by slow cooling from 100°C to room temperature. The probe was purified using a Quick Spin Column (Roche). The nuclear extracts were harvested from COS-7 cells. Empty extract from Cos-7 cells was used as a control to check endogenous protein binding. Antibody against V5 (Invitrogen) was used to identify and super-shift of BCL11A ZnF (4-6) and mutations. Samples were complexed in gel shift buffer (10 mM HEPES (pH

7.8), 50 mM KCl, 5 mM $MgCl_2$, 1 mM EDTA, 0.05 mg/mL poly(dI-dC), 1 mM DTT, 0.1 mg/mL bovine serum albumin, 5% glycerol) with and without V5 antibody and loaded on 6% native polyacrylamide gel in TBE buffer (45mM Tris, 45mM boric acid and 1mM EDTA). Electrophoresis was performed at 4°C, for 105 min, at 250V followed by drying under vacuum. The gel was exposed to a FUJIFILM BAS CASSETTE2 2025 phosphor screen overnight and visualized using a GE Typhoon FLA 9500 fluorescent image analyzer.

Table 6: Probe used for EMSA

Gel Shift Probe	-128 to -100 bp fetal globin promoter (canonical binding motif underlined)	**Forward Probe:** AGCCTTGCCTTGACCAATAGCCTTGACAA	**Reverse Probe:** TTGTCAAGGCTATTGGTCAAGGCAAGGCT

Table 7: Plasmid used for EMSA

Mammalian Expression Plasmids	**Species**	**Source**
pcDNA3-Empty	N/A	Invitrogen
pcDNA3-FLAG BCL11A ZF (4-6) 740-835aa-V5	Human	(Martyn et al., 2018)
pcDNA3-FLAG BCL11A ZF (4-6) 740-835aa 755S>G-V5	Human	Mahdi Haddad
pcDNA3-FLAG BCL11A ZF (4-6) 740-835aa 756N>G-V5	Human	Mahdi Haddad
pcDNA3-FLAG BCL11A ZF (4-6) 740-835aa 755S>G, 756N>G -V5	Human	Mahdi Haddad

Off-target analysis:

We used the COSMID web tool to predict Cas-dependent DNA off-target sites with a stringency of no DNA and RNA bulges. Using the primers listed in Supplementary Table 2, we amplified the predicted off-target sites and carried out 150bp-paired-end Illumina sequencing with the specified adapter sequences. The resulting data was analyzed using the default settings of CRISPRESSOv2.(Clement et al., 2019) The transcriptome-wide A-to-I (or T-to-C conversion) conversion in the ABE8e base-edited erythroid cells was calculated by REDItools v2. All the nucleotides other than A were removed from the analysis with the previously reported coverage and quality criteria.(Koblan et al., 2018) The frequency of A converted to I (or T to C conversion) was calculated by dividing the total number of edited A by the overall counts of A after filtering (A-to-I)/A*100. The experiment was carried out as two biological replicates.

Table 8: Off-target Primers used

Sample	Forward primer	Reverse primer
ZnF4/5 OT1	TACACGACGCTCTTCCGATCTTCATACCATGGACAAAGACATTCCT	AGACGTGTGCTCTTCCGATCTACTAGGCTTCTACCATCCCTTGC
ZnF4/5 OT2	TACACGACGCTCTTCCGATCTGCAGCTGGAGTACCTGGTCT	AGACGTGTGCTCTTCCGATCTGAAGCAAAGTCCCATTCCCTCC
ZnF4/5 OT3	TACACGACGCTCTTCCGATCTTGAAGAAAGAATGGCTCAGAGGGC	AGACGTGTGCTCTTCCGATCTAGGGTTATGTTCTTTCTGGACCATCTCCTT
ZnF4/5 OT4	TACACGACGCTCTTCCGATCTAGGGTTGCAGATCCCACCAG	AGACGTGTGCTCTTCCGATCTTGTAGACTCTGACGCACCCC
ZnF4/5 OT5	TACACGACGCTCTTCCGATCTCCCAGTGAAATGTGTAAGTGAATTC	AGACGTGTGCTCTTCCGATCTGGACTTTCAGCCCTTTGCCTTG
ZnF4/5 OT6	TACACGACGCTCTTCCGATCTCAAGAACTCAGTTTTCAAGGTTACT	AGACGTGTGCTCTTCCGATCTGGATTGTGCAAGGTACATGCATTTAATATTAGATTGG
ZnF4/5 OT7	TACACGACGCTCTTCCGATCTGCCTATCACTACCCAACAAGAAGTA	AGACGTGTGCTCTTCCGATCTCTGTAAGTTTCCTGAGGCCTCC
ZnF4/5 OT8	TACACGACGCTCTTCCGATCTTGGGCTTCGTCTTTAGGATATGGG	AGACGTGTGCTCTTCCGATCTATCCTTCAGGACATATACCCCTGG
+33bp GATA OT1	TACACGACGCTCTTCCGATCTCTCTGTGCCTCAGAGGAATGAC	AGACGTGTGCTCTTCCGATCTACCTGTCTCCTTGGCACATTATCC
+33bp GATA OT2	TACACGACGCTCTTCCGATCTTGTGAGCAAATGCACATGGCTCTG	AGACGTGTGCTCTTCCGATCTTCAGCACTAAGTCCCAGAGATGAG
+33bp GATA OT3	TACACGACGCTCTTCCGATCTCCAGTCTCTGCCCAGGAAAG	AGACGTGTGCTCTTCCGATCTGGACTCAGATCAATAATTACATGCTTCCCAG
+33bp GATA OT4	TACACGACGCTCTTCCGATCTCTTATGTCTCCGTCCCTCCTC	AGACGTGTGCTCTTCCGATCTTGTCATGACACGAGGGCTG
+33bp GATA OT5	TACACGACGCTCTTCCGATCTGAGCGGACATCTCCTTGAAGC	AGACGTGTGCTCTTCCGATCTCAGAGACTGATATTTCCAGGGTCC
+33bp GATA OT6	TACACGACGCTCTTCCGATCTGGGGCGTCATTGGTGTGTTTTTC	AGACGTGTGCTCTTCCGATCTGTCAGGTAGAGAGATGCCACAG
+33bp GATA OT7	TACACGACGCTCTTCCGATCTCTTTGGTTTTCTAAAGCGGTGATGA	AGACGTGTGCTCTTCCGATCTAGCTCACATCTTCAAAGACAGGGG
+33bp GATA OT8	TACACGACGCTCTTCCGATCTGTTAACAGGTTTCAGAACTGGGATT	AGACGTGTGCTCTTCCGATCTCCACACCAAATCTAAACTGCCAGG
+33bp GATA OT9	TACACGACGCTCTTCCGATCTGGGGCTGGGTGAGTGAGATA	AGACGTGTGCTCTTCCGATCTCTGCTCACACTCTCTGACCC
+33bp GATA OT10	TACACGACGCTCTTCCGATCTCATTCTCTGCGAATTCTGCAAC	AGACGTGTGCTCTTCCGATCTTGATGGATTTAGTTGCGGCGCAGTG
+33bp GATA OT11	TACACGACGCTCTTCCGATCTAGTTGTGGCAGGTGCAGCAG	AGACGTGTGCTCTTCCGATCTCAGTCTTCAGGCCTCAGGCA
+33bp GATA OT12	TACACGACGCTCTTCCGATCTGTTTCTCCACATTGAGCCAACAGC	AGACGTGTGCTCTTCCGATCTAGGGGAAGCTCCTGGTGACA

Flow cytometry analysis:

The erythroid cells were stained with antibodies (CD235a and CD71) for 15 minutes, washed twice with PBS, and analyzed. The intracellular HbF staining and mouse bone marrow cells were processed as previously described. All antibodies used are listed in supplementary table 3. BD celesta and CytoFLEX-LX (Beckman Coulter) flow cytometers were used for sample analysis.

Table 9: Flow-cytometry antibodies used

Antibodies	Company	Cat No
CD71	BD	562995 and 555536
CD235a	BD	551336 and 240947
HbF APC	THERMO-Invitrogen	MHFH05
NucRed	Thermo Fischer Scientific	R37106
CD45 human	BD	555482
CD45 Mouse	BD	559864
CD34	BD	348057
CD13	BD	555394
CD33	BD	340474
CD16	BD	347617
CD19	BD	340364

Western blot:

Whole-cell protein lysates were prepared by resuspending the cell pellet with 1xRIPA Buffer along with Halt Protease Inhibitor Cocktail (ThermoFisher Scientific). Approximately 35µg of protein was used for western blot analysis.

Table 10: Western Antibodies used

Antibodies	Company	Cat No
BCL11A (D4E3P)	Cell Signalling Technology	75432
β-Actin	MP Biological	AS003
HRP Goat antimouse	Abclonal	A6154
Anti-Rabbit IgG	Sigma	69001

Quantitative RT-PCR:

RNA form the erythroid differentiated cells were isolated using the Machery Nagel RNA isolation kit. cDNA from 500ng of the isolated RNA is reverse transcribed using a TAKARA Primescript kit. The obtained cDNA is diluted 4 times and used in qPCR using Taqman Primescript II kit.

Table 11: RT-qPCR primers used

GENE	Forward Primer	Reverse Primer
BCL11A	AACCCCAGCACTTAAGCAAA	GGAGGTCATGATCCCCTTCT
HBB	ACCTTTGCCACACTGAGTGAG	TTTGCCAAAGTGATGGGCCA
*HBG*1	Qiagen primers. Cat No: QT00041384	
*HBG*2	Qiagen primers. Cat No: QT00040068	
BGLT3	AAGATAATCTTGGTTTTGCCTCAAA	TCTACTTGATATAGTTGAGAGGCAGTTACC
HBBP1	TTATGCTCACGGATGACCTCAAAG	AACAATCAATATCACGTTGCCTAAGAG
GAPDH	CTGCACCACCAACTGCTTAG	GTCTTCTGGGTGGCAGTGAT

HPLC:

At the terminal stage of differentiation, the differentiated cells were sonicated and centrifuged with maximum rpm at 4°C for 15-20mins, to obtain the clear lysate.(Bagchi et al., 2021) Analysis of total hemoglobin (TOSOH HPLC variant analyzer) was performed using the lysates. The percentage of total globin was calculated as [HbF/HbF+HbA] *100.

***In-silico* modeling:**

The BCL11A protein-DNA structure obtained from RCSB PDB-6U9Q was used as a reference.(Yang et al., 2019) We used Pymol software to visualize the amino acids involved in interaction with the DNA bases.

RNA-Seq:

Total RNA isolated during phase-2 of erythroid differentiation (Day-8) of edited HSPCs was processed for RNA-sequencing using Illumina (Nova-seq6000). The raw reads were filtered using Trimmomatic for quality scores and adapters. Filtered reads were aligned to the Human genome (hg38) using splice-aware aligners like HISAT2 to quantify reads mapped to each transcript. The alignment percentage of reads ranged between 95.28 - 98.52 percentage for all the samples. The total number of uniquely mapped reads was counted using feature counts. The uniquely mapped reads were then subjected to differential gene expression using DeSeq2.

ATAC-seq:

Omni-ATAC protocol was followed to perform ATAC-seq. Buffer containing 0.1% NP-40, 0.1% Tween-20, and 0.01% digitonin is used for each sample (5×10^4–6×10^4 cells) to lyse the cells. The lysate was then washed and incubated with Tagment DNA TDE1 enzyme (Illumina) for 30 min at 37 °C. DNA was purified with a Qiagen MinElute PCR Purification kit. Library fragments were amplified using NEBNext High-Fidelity 2× PCR master mix (NEB, M0541) and custom primers with unique dual indexes and PCR amplified. PCR products were purified using AMPure XP beads (Beckman Coulter, A63881) at a 1:1 ratio according to the manufacturer's instructions. Purified PCR products sized from 150 bp to 1 kb were purified from 2% E-Gel EX agarose gels (Thermo Fisher, G401002). Libraries were pooled and sequenced in paired-end mode using an Illumina HiSeq kit. Raw reads were trimmed to remove the Tn5 adaptor sequence using a skewer (v.0.2.2) and then mapped to hg19 using BWA-MEM (v.0.7.16a). Duplicated multi-mapped reads were removed with SAMtools (v.0.17). ATAC-seq peaks were called using MACS2 (v.2.1.1) with the following parameters: macs2 callpeak –nomodel –shift –100 –extsize 200. BigWiggle files were generated using DeepTools (v.3.2.0) (Feng et al., 2022)

Cut and Run:

The CUT&RUN assay was performed to detect transcription factor occupancy and histone modification profiles. Approximately 5×10^5 cells were analyzed for each reaction. Library construction for DNA sequencing was performed using an NEBNext UltraII DNA Library Prep kit from NEB (E7645S). Indexed samples were analysed by paired-end NGS, using an Illumina NextSeq Mid-Output (150 cycles) kit. FASTQ files

were mapped to hg19 using BWA-MEM (v.0.7.16a). Reads that could not be uniquely mapped to the human genome were removed by SAMtools (v.0.17). Peaks were called using MACS2 (v.2.1.1). BigWiggle files were generated using DeepTools (v.3.2.0). Footprint analysis was performed using CUT&RUNTools. Data were normalized using S3norm(v.2)**.** (Feng et al., 2022)

Statistical Analysis:

All graphs were generated using GraphPad Prism v8.1. One-way ANOVA was used in all the analyses and the comparison is with the edited control unless otherwise mentioned.

RESULTS

Section-4

Results

4.1 Base editing of key residues in the BCL11A-XL-specific zinc finger domains de-represses fetal globin expression.

BCL11A-XL directly binds and represses the *γ-globin (HBG1/2)* gene promoters, using three Zinc-Finger domains (ZnF4, ZnF5, ZnF6), and is a potential target for β-hemoglobinopathy treatments. Disrupting BCL11A-XL results in de-repression of *fetal globin* and high HbF, but also affects Hematopoietic Stem and Progenitor Cells (HSPC) engraftment and erythroid maturation. Intriguingly, neurodevelopmental patients with ZnF domain mutations have elevated HbF with normal hematological parameters. Inspired by this natural phenomenon, we employed both CRISPR/Cas9 and base editing at specific ZnF domains and assessed the impacts on HbF production and hematopoietic differentiation. Generating indels in the various ZnF domains by CRISPR/Cas9 prevented the binding of BCL11A-XL to its site in the *HBG1/2* promoters and elevated HbF levels but impacted normal hematopoiesis. Far fewer side effects were observed with base editing, for instance, erythroid maturation *in vitro* was near normal. However, we observed a modest reduction in HSPC engraftment and a complete loss of B-cell development *in vivo*, presumably because current base editing is not capable of precisely recapitulating the mutations found in patients with BCL11A-XL-associated neurodevelopment disorders. Overall, our results reveal that disrupting different ZnF domains has different impacts. Disrupting ZnF4 elevated HbF levels significantly, while leaving many other erythroid target genes unaffected, and interestingly, disrupting ZnF6 also elevated HbF levels, which was unexpected as this region does not directly interact with the *HBG1/2* promoters. This first

structure/function analysis of ZnF4-6 provides important insights into the domains of BCL11A-XL required to repress *γ-globin* expression and provides a framework for exploring the introduction of natural mutations that may enable de-repression of single genes while leaving other functions unaffected.

4.1.1 CRISPR/Cas9 mediated disruption of BCL11A-XL specific ZnF domains in adult human erythroid cells de-represses fetal *globin* expression.

Previous studies have shown that BCL11A-XL interacts with the γ-*globin* (*HBG1/2*) promoters in adult erythroid cells using its C-terminal ZnF domain, containing ZnFs 4, 5, and 6. However, the roles of each ZnF domain specific to BCL11A-XL in fetal hemoglobin regulation and erythroid maturation have not been established. Therefore, we performed CRISPR/Cas9-mediated disruption of each ZnF (gRNA: ZnF4,5 and 6) and compared the impacts with the other BCL11A associated targets, like disruption of the intronic erythroid-specific *BCL11A* Enhancer (Enh), and its binding site in the *HBG1/2* promoters (BS). In addition, gRNAs designed to ablate BCL11A (targeting BCL11A exon2) and AAVS1 were selected as positive and negative controls respectively **(Fig. 6a)**. HUDEP2 cells stably expressing Cas9 were transduced with the specific gRNAs and around 70-90% indels were achieved at all the intended target sites **(Fig. 6b and c)**. Though editing at ZnF4, 5, and 6 (759, 785, and 810 amino acids respectively) did not alter the protein expression, it led to the formation of altered BCL11A-XL protein with varying lengths due to frameshift mutations **(Fig. 6d-f).** In contrast, downregulation of BCL11A protein was observed when targeting BCL11A exon 2 and the enhancer. These results suggest that targeting the BCL11A exon2 (N-terminal) leads to the degradation of the protein while targeting

the XL-specific ZnF-domains (C-terminal) produced truncated variants of BCL11A(Shen, Verboon, et al., 2021). Nevertheless, as expected the truncated BCL11A variants failed to repress *HBG* at the adult stage and produced a noticeably higher number of HbF+ cells than the enhancer-modified cells and comparable to binding site disruption at the *HBG* promoters after erythroid differentiation **(Fig. 6g-i).** Despite previous reports demonstrating that ZnF6 has a weak direct interaction with the *HBG* promoter compared to ZnF4 and ZnF5, we unexpectedly observed a significant elevation in fetal hemoglobin upon ZnF6 disruption(Liu et al., 2018). These results highlight the importance of each ZnF-domain specific to BCL11A-XL in the repression of the fetal *globin* genes.

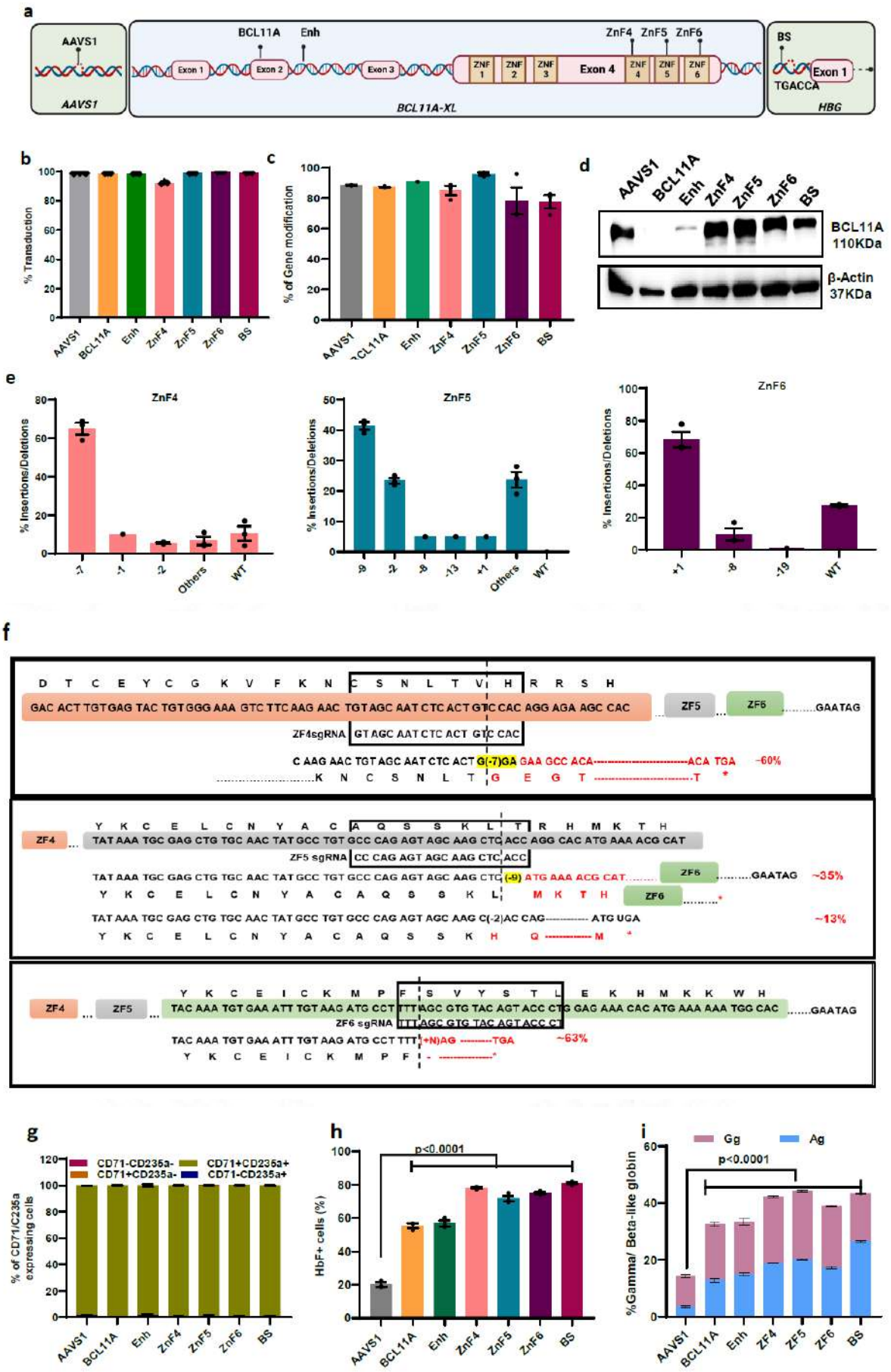

a
AAVS1
AAVS1
BCL11A
Enh
Exon 1
Exon 2
Exon 3
ZNF 1
ZNF 2
ZNF 3
Exon 4
ZnF4
ZnF5
ZnF6
ZNF 4
ZNF 5
ZNF 6
BCL11A-XL
BS
TGACCA
Exon 1
HBG
b
% Transduction
AAVS1
BCL11A
Enh
ZnF4
ZnF5
ZnF6
BS
c
% of Gene modification
d
AAVS1
BCL11A
Enh
ZnF4
ZnF5
ZnF6
BS
BCL11A
110KDa
β-Actin
37KDa
e
ZnF4
% Insertions/Deletions
-7
-1
-2
Others
WT
ZnF5
% Insertions/Deletions
-9
-2
-8
-13
+1
Others
WT
ZnF6
% Insertions/Deletions
+1
-8
-19
WT
f
D T C E Y C G K V F K N C S N L T V H R R S H
GAC ACT TGT GAG TAC TGT GGG AAA GTC TTC AAG AAC TGT AGC AAT CTC ACT GTC CAC AGG AGA AGC CAC
ZF5
ZF6
GAATAG
ZF4sgRNA
GT AGC AAT CTC ACT GTC CAC
C AAG AAC TGT AGC AAT CTC ACT G(-7)GA GAA GCC ACA
ACA TGA
~60%
K N C S N L T G E G T
T
*
Y K C E L C N Y A C A Q S S K L T R H M K T H
ZF4
TAT AAA TGC GAG CTG TGC AAC TAT GCC TGT GCC CAG AGT AGC AAG CTC ACC AGG CAC ATG AAA ACG CAT
ZF5 sgRNA
CC CAG AGT AGC AAG CTC ACC
TAT AAA TGC GAG CTG TGC AAC TAT GCC TGT GCC CAG AGT AGC AAG CTC (-9) ATG AAA ACG CAT
ZF6
GAATAG
~35%
Y K C E L C N Y A C A Q S S K L M K T H
ZF6
*
TAT AAA TGC GAG CTG TGC AAC TAT GCC TGT GCC CAG AGT AGC AAG C(-2)AC CAG ATG TGA
~13%
Y K C E L C N Y A C A Q S S K H Q M
*
Y K C E I C K M P F S V Y S T L E K H M K K W H
ZF4
ZF5
TAC AAA TGT GAA ATT TGT AAG ATG CCT TTT AGC GTG TAC AGT ACC CTG GAG AAA CAC ATG AAA AAA TGG CAC
ZF6 sgRNA
TTT AGC GTG TAC AGT ACC CT
GAATAG
TAC AAA TGT GAA ATT TGT AAG ATG CCT TTT +N)AG TGA
~63%
Y K C E I C K M P F
*
g
CD71-CD235a-
CD71+CD235a+
CD71+CD235a-
CD71-CD235a+
% of CD71/C235a expressing cells
AAVS1
BCL11A
Enh
ZnF4
ZnF5
ZnF6
BS
h
p<0.0001
HbF+ cells (%)
AAVS1
BCL11A
Enh
ZnF4
ZnF5
ZnF6
BS
i
Gg
Ag
p<0.0001
%Gamma/ Beta-like globin
AAVS1
BCL11A
Enh
ZF4
ZF5
ZF6
BS

Fig 6: CRISPR/Cas9 mediated disruption of BCL11A-XL specific ZnF domains in adult human erythroid cells de-represses fetal *globin* expression. a) Schematic representation of the target location of gRNA in different BCL11A-XL specific ZnF domains (ZnF 4, 5, and 6), Exon2, Enhancer, -115 cluster of *HBG* promoter, and the AAVS1 site; b)The transduction efficiency of each gRNA measured in flow cytometry in HUDEP2 cells expressing cas9. Mean ± SEM of n=3 independent replicates. c)The frequency of insertions and deletions (InDELs) was measured by ICE analysis at the respective target site in HUDEP-2 cells during expansion; Mean ± SEM of n=3 independent replicates. d) The protein expression analysis by Western blot on the edited cells during differentiation (Day 4) depicts the generation of BCL11A variants protein on targeting the XL-specific ZnF domains in contrast to the reduction of BCL11A protein while targeting Exon-2 and the enhancer; e) The frequency of multiple insertion or deletion patterns generated at each target site; f) An illustration depicting the insertion or deletion with the highest frequency that could result in the generation of BCL11A ZnF variants at ZnF 4, 5, and 6; g) Erythroid differentiation profile analyzed by the expression of CD235a and CD71 marker by flowcytometry in BCL11A ZnF edited HUDEP2 cells at the day 8 of erythroid differentiation; h) Intracellular fetal hemoglobin analysis using flow cytometry demonstrates the potential of ZnF mutants on HbF reactivation at the terminal stage of erythroid differentiation (Day 8); Mean ± SEM of n=3 independent replicates. i) Analysis of globin chain production through reverse-phase HPLC at the terminal stage of erythroid differentiation (Day 8); Mean ± SEM of n=3 independent replicates.

4.1.2 BCL11A ZnF editing in CD34$^+$ HSPCs through CRISPR/Cas9 elevates HbF but also affects erythroid maturation

Besides the major role in globin regulation, *BCL11A* is essential for human hematopoietic stem cell maintenance and erythropoiesis. Therefore, we assessed the effect of BCL11A-ZnF disruption on cell survival and erythroid maturation in CD34$^+$ HSPCs obtained from a healthy donor by nucleofecting the CRISPR/Cas9-gRNA complex. Based on our HUDEP2 results where enhancer disruption showed moderate HbF elevation, we determined the impact of the generating BCL11A-ZnF variants in erythroid cells derived from HSPCs, along with the binding site disruption at the *HBG1/2* promoters, and BCL11A knockout at exon2. HSPCs exhibiting efficient gene modifications **(Fig. 7a)** at the respective target sites were subjected to erythroid

differentiation to assess the impact of the modifications on functions beyond HbF elevation during *in vitro* erythropoiesis.

In line with previous research demonstrating the importance of *BCL11A* in cell maintenance and proliferation, the growth kinetics of committed erythroid progenitors from the edited samples were analyzed. The total BCL11A-knockout and ZnF-editing in $CD34^+$ HSPCs affected cell proliferation, whereas editing other targets did not affect survival **(Fig. 7b)**. The frequency of gene perturbations was assessed at the erythroid terminal differentiation stage. This revealed that except for ZnF4 and total knockout edited cells, gene modifications in all other targets persisted during differentiation **(Fig. 7a)**. This indicates that cells carrying alterations at the terminal sequences coding for ZnF5 and ZnF6 retain edits and persist across multiple cell divisions. A previous study(Shen, Verboon, et al., 2021) reported that disruption of BCL11A exon-4 (ZnF) did not impede early erythroid differentiation, whereas our study targeting the ZnF4, 5 and 6 domains showed a decline in the percentage of terminal-stage erythroid cells ($CD235a^+CD71^-$), **(Fig. 7c)** with a significant reduction in the enucleation **(Fig. 7d)**. However, a remarkable level of fetal *globin gene* reactivation was seen in adult erythroid cells on targeting the BCL11A ZnF domains **(Fig. 7e and f)**, equivalent to levels obtained with the binding site disruption, despite the differences observed in cell proliferation and differentiation. Overall, these results demonstrate the *in vitro* consequence of disrupting either exon2 (N-terminal) or major disruption in the C-terminal ZnF domains of BCL11A protein impact cell survival and erythroid maturation as well as globin regulation.

BCL11A isoforms have been implicated in HSC engraftment and B-lymphocyte maturation.(Liu et al., 2003; Yu et al., 2012) Therefore we investigated the engraftment potential of BCL11A-XL ZnF-modified cells in the NBSGW mouse model. Since ZnF5 had exhibited significant erythroid maturation defects during *in vitro* experiments, we infused only the ZnF4 and ZnF6 modified HSPCs. As controls, we also infused BCL11A binding site disrupted HSPCs and total BCL11A knockout (exon 2) edited HSPCs. All the transplanted animals showed chimerism of human cells ($hCD45^+$) in the bone marrow after 16 weeks of infusion, even the cohorts with total BCL11A knockout cells **(Fig. 7g)**. Therefore, we genotyped the engrafted cells to study the long-term survival of gene-modified HSPCs, which revealed that the indel frequency in the *HBG1/2* promoter BCL11A binding site edited cells remained unchanged, whereas HSPCs with BCL11A complete knockout did not survive in mouse bone marrow (BM), consistent with the previous findings. We had speculated originally that BCL11A ZnF domain modification would have lesser impacts on HSPCs, but we observed a drastic decrease in ZnF4 and ZnF6 gene modifications in the BM cells **(Fig. 7h)**. Similar to total BCL11A knockout, we observed that the CRISPR/Cas9-mediated disruption of BCL11A-XL specific ZnF domains does not support HSC engraftment. Furthermore, due to the inadequate engraftment potential of the BCL11A ZnF mutants, we were unable to investigate its multilineage differentiation ability. These findings demonstrate that even a minimal loss of the BCL11A-XL C-terminal coding region functions like a total protein loss and can negatively affect the engraftment of HSCs.

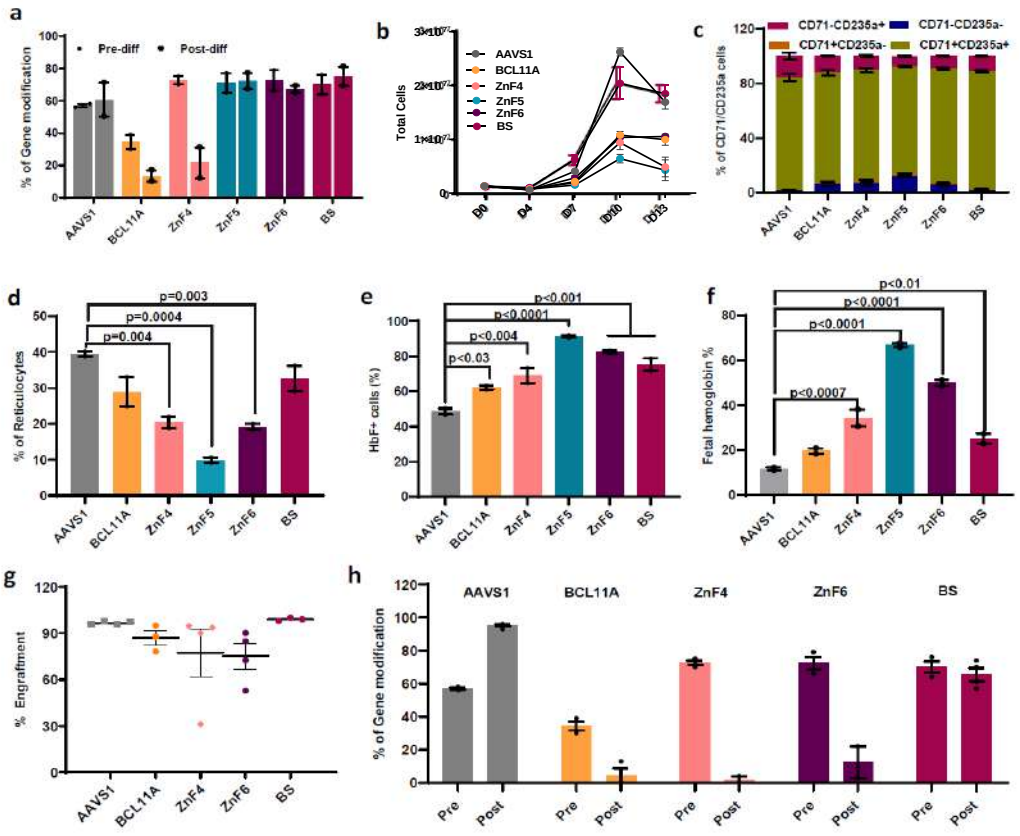


Fig 7: BCL11A ZnF editing in CD34^{+} HSPCs through CRISPR/Cas9 elevates HbF but also affects erythroid maturation. a) The percentage of gene modifications during HSPCs expansion after 48 hours of electroporation and after erythroid differentiation day 21 respectively; Mean ± SEM of n=2 replicate on 1 donor. b) The cell proliferation profile of all the gene-modified HSPCs after electroporation during the three-phase erythroid differentiation. c) The change in erythroid differentiation ability of the BCL11A modified HSPCs was analyzed by flow cytometry using two erythroid-specific surface markers (CD71-Transferin; CD235a–Glycophorin A) at the terminal stage of differentiation (Day 21); Mean ± SEM of n=2 replicate of 1 donor. d) The measure of mature enucleated cells formed by the edited HSPCs during the terminal stage of erythroid differentiation (Day 21) using NucRed stain and CD235a marker; Mean ± SEM of n=2 replicate on 1 donor. e and f) The number of HbF+ cells expressed, and the total fetal hemoglobin tetramer in edited cells formed on BCL11A ZnF domain modification by FACS and HPLC respectively measured during terminal erythroid differentiation (Day 21); Mean ± SEM of n=2 replicate on 1 donor. g) The long-term engraftment of BCL11A total knockout, binding site, and ZnF domain modified HSPCs after 16 weeks of transplantation in NBSGW mice; Mean ± SEM of n≥3 mice per group. h) The persistence of gene modification in the edited cells depicted by editing during both pre-infusion and after 16 weeks of transplantation.

4.1.3 Specific base substitutions in the BCL11A-XL C-terminal ZnF domains reactivate HbF expression in adult erythroid cells

CRISPR/Cas9-mediated targeting of BCL11A-XL specific C-terminal ZnF domains resulted in multiple truncated variants hampering the concise understanding of the role of BCL11A-XL ZnF domains in HbF repression, engraftment, and erythroid maturation. To address this, we introduced more subtle mutations at specific amino acid residues in the ZnFs using an adenosine base editor in HUDEP2 cells anticipating fetal hemoglobin elevation without compromising other functional roles of *BCL11A.*

The key nucleotides at ZnF4 (742-765 amino acids) and ZnF5 (772-795 amino acids) were targeted to alter crucial motifs required for DNA interaction and where possible to resemble as closely as possible known missense variants associated with higher HbF levels in patients with neurodevelopmental disorders but essentially normal hematology. With current base editing technology, it is not possible to recreate the exact mutations, but we targeted mutations as closely as possible to the sites found in patients. No naturally occurring mutations have been reported in ZnF6 (803-827 amino acids) but we designed a gRNA that would result in a base substitution at a potential DNA-contact region in the ZnF6 domain **(Fig. 8a)**. The fetal globin reactivation potential of base substitutions in each ZnF domain was compared with the same *BCL11A* control targets used in previous experiments, along with the gRNA at the initiation codon for the total knockout of *BCL11A***.** Subsequently, all these gRNAs were individually transduced into HUDEP2 cells with more than 90% efficiency.

After 8 days of expansion, the efficiency of base conversion at the target sites was evaluated by Sanger sequencing **(Fig. 8b)**. In the case of the ZnF4 domain, the ZnF4-1 gRNA specifically converted the nucleotides that encode the amino acids S755 and N756 with 40% and 90% base conversion efficiency, respectively. S755 makes van der Waals interactions with the -119 (T: A) base and the DNA backbone at the *HBG1/2* promoters, whereas N756 forms a specific double hydrogen bond with -118 (A: T) (**Fig 8c: ZnF4-1**). The ZnF4-2 gRNA generated novel mutants H760R and R761G with 80% efficiency. Although these amino acids do not contact DNA, H760 is one of the zinc-ligating residues in ZnF4, and its mutation likely perturbs the structure of the domain (**Fig 8c: ZnF4-2**). Alteration of the ZnF5 domain was achieved using ZnF5-1 gRNA to efficiently (70%) convert bases that encode Q781 and S782. Q781 forms a hydrogen bond with -116 (T: A) (**Fig 8c: ZnF5-1**). Similarly, the ZnF5-2 gRNA generated specific nucleotide perturbations that mutate T786, R787, and H788; R787 forms hydrogen bonds with -114 (G: C), whereas H788 is one of the zinc-coordinating residues in ZnF5 (**Fig 8c: ZnF5-2**). The ZnF6 gRNA exhibited a less efficient (40%) conversion of bases and encodes changes to E816 and K817; the latter residue is predicted to make non-specific electrostatic interactions with the DNA backbone (**Fig 8c: ZnF6**).

The control sites, the -115 cluster in the *HBG1/2* promoter and the *BCL11A* enhancer were substantially modified with overall efficiencies greater than 70%. Unintended indels were less than 5%, despite the high editing efficiency at all the target sites. The consequence of the generated BCL11A-XL mutants on protein expression was investigated to analyze the impact of pathogenic SNVs on the stability of the BCL11A protein. Interestingly, the XL-specific ZnF mutants exhibited no effect

on protein synthesis and were expressed similarly to the AAVS1 negative control editing, whereas the alteration of initiation codon affected the BCL11A protein expression mimicking BCL11A knockout as expected **(Fig. 8d)**.

We assessed the functional consequence of the *BCL11A* mutations on fetal *globin* expression during terminal erythroid differentiation **(Fig. 8e)**. Interestingly, the HbF reactivation observed in all ZnF mutants was comparable to complete BCL11A knockout (BCL11A). The increase in the number of HbF+ cells ranged from 60-90% across the ZnF mutants which may be due to a difference in editing efficiency or the impact of the different mutations **(Fig. 8f)**. The ZnF4 mutations increased the number of HbF+ cells more than the ZnF5 or ZnF6 mutations. Although the editing frequency at the ZnF6 target site was lower, significant HbF elevation was observed, so we hypothesize that ZnF6 also influences BCL11A binding to the *HBG1/2* promoters. Consistent with the increase in HbF+ cells, the ZnF4 mutants produced higher levels of total HbF relative to mutations in other ZnF domains (ZnF5 and ZnF6).

Importantly, the majority of ZnF-specific mutations outperformed the disruption of the BCL11A binding site in the *HBG1/2* promoters in HbF reactivation **(Fig. 8g)**. It may be that BCL11A binds at additional sites in the *globin* locus, so mutating the DNA-binding domain has a more profound effect than mutating a single binding motif. These findings confirm that induction of HbF does not require haploinsufficiency or the total loss of BCL11A but can be mediated by targeted mutations.

Together, both the *in vitro* and *in silico* results imply that altering the C-terminal XL-specific ZnF domains involved in binding to the target DNA (such as the

motifs in the *HBG1/2* promoters) exhibited a similar HbF upregulation to that observed with total BCL11A loss. Further, unique base conversations that alter key amino acid residues are sufficient to significantly impact ZnF domain function.

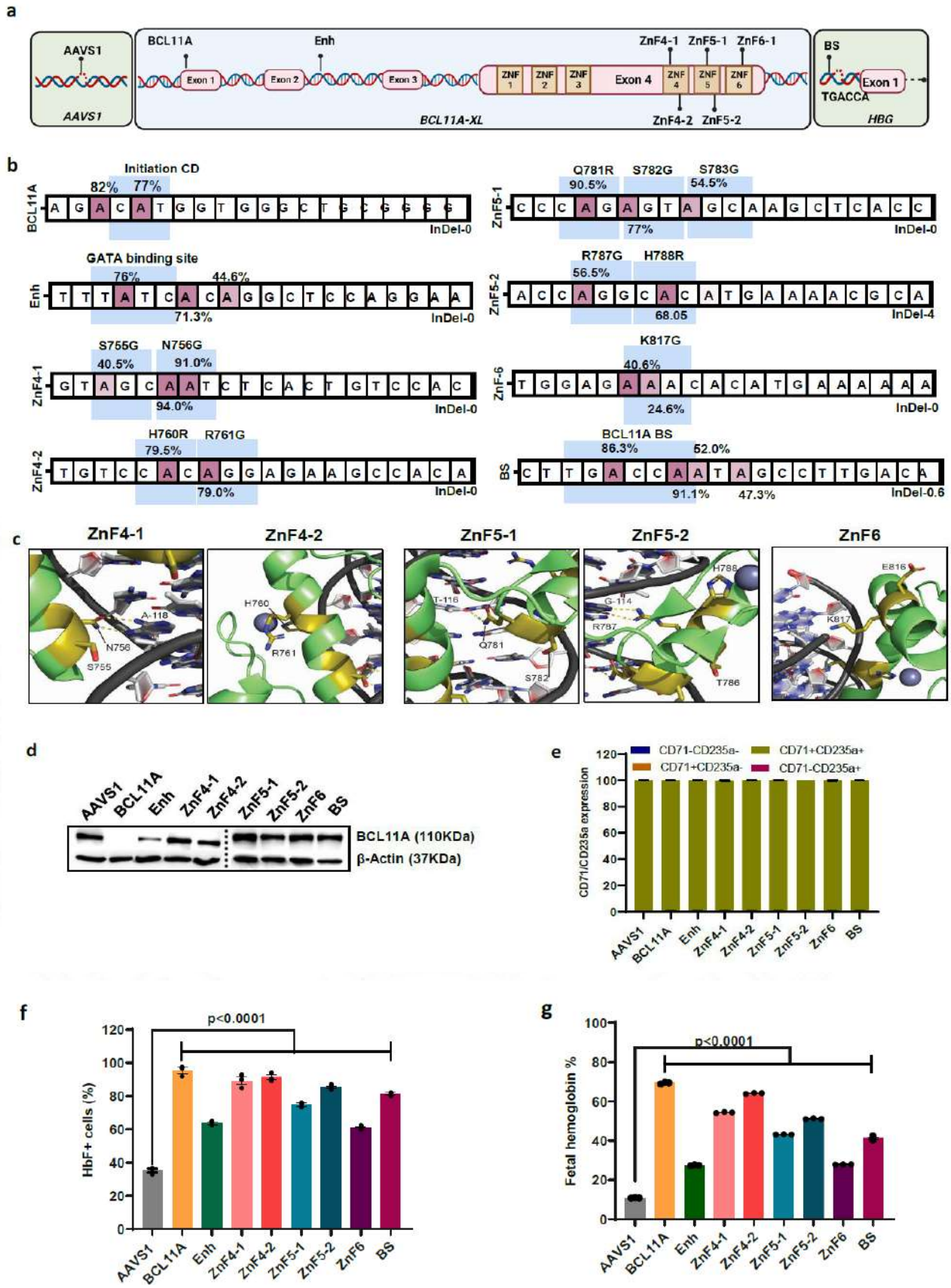

a
AAVS1
AAVS1
BCL11A
Enh
Exon 1
Exon 2
Exon 3
ZNF 1
ZNF 2
ZNF 3
Exon 4
ZNF 4
ZNF 5
ZNF 6
ZnF4-1
ZnF5-1
ZnF6-1
ZnF4-2
ZnF5-2
BCL11A-XL
BS
TGACCA
Exon 1
HBG
b
BCL11A
Initiation CD
82%
77%
InDel-0
Enh
GATA binding site
76%
44.6%
71.3%
InDel-0
ZnF4-1
S755G
N756G
40.5%
91.0%
94.0%
InDel-0
ZnF4-2
H760R
R761G
79.5%
79.0%
InDel-0
ZnF5-1
Q781R
S782G
S783G
90.5%
54.5%
77%
InDel-0
ZnF5-2
R787G
H788R
56.5%
68.05
InDel-4
ZnF-6
K817G
40.6%
24.6%
InDel-0
BS
BCL11A BS
86.3%
52.0%
91.1%
47.3%
InDel-0.6
c
ZnF4-1
ZnF4-2
ZnF5-1
ZnF5-2
ZnF6
A-118
N756
S755
H760
R761
T-116
Q781
S782
G-114
R787
H788
T786
E816
K817
d
AAVS1
BCL11A
Enh
ZnF4-1
ZnF4-2
ZnF5-1
ZnF5-2
ZnF6
BS
BCL11A (110KDa)
β-Actin (37KDa)
e
CD71-CD235a-
CD71+CD235a+
CD71+CD235a-
CD71-CD235a+
CD71/CD235a expression
120
100
80
60
40
20
0
AAVS1
BCL11A
Enh
ZnF4-1
ZnF4-2
ZnF5-1
ZnF5-2
ZnF6
BS
f
p<0.0001
HbF+ cells (%)
120
100
80
60
40
20
0
AAVS1
BCL11A
Enh
ZnF4-1
ZnF4-2
ZnF5-1
ZnF5-2
ZnF6
BS
g
p<0.0001
Fetal hemoglobin %
100
80
60
40
20
0
AAVS1
BCL11A
Enh
ZnF4-1
ZnF4-2
ZnF5-1
ZnF5-2
ZnF6
BS

Fig 8: Specific base substitutions in the BCL11A-XL C-terminal ZnF domains reactivate HbF expression in adult erythroid cells. a) A diagrammatic representation of each of the target sites to induce nucleobase modifications by using an adenosine base editor. BCL11A- gRNA targeting initiation codon of BCL11A, Enh-gRNA targeting the BCL11A erythroid-enhancer region, two gRNAs each targeting ZnF 4(ZF4-1, -2) and 5(ZF 5-1, -2) respectively, and one gRNA targeting ZF6(ZF6) to produce missense mutations, BS (Binding Site) gRNA to alter the BCL11A-XL binding site in the -115 cluster of *HBG* promoter; AAVS1-gRNA targeting the safe harbor site at AAVS1 locus for negative control. b) Base conversion efficiency at all the target sites on day 8 expansion analyzed using EditR *in silico* tool after sanger sequencing; Mean ± SEM of n=3 independent replicates. c) Closeup views of the BCL11A-XL specific ZnF (ZnF 4, 5, and 6) structures, illustrating the amino acids targeted in each of the indicated mutants. Yellow dashed lines indicate hydrogen bonds and bases are named according to the *HBG1/2* promoter numbering. d) The protein expression analysis by Western blot confirms no significant change in BCL11A expression on ZnF editing compared to the reduction of BCL11A protein observed on disrupting the Initiation codon and the BCL11A-enhancer; 35ug of protein used for loading. e) Analysis of the erythroid differentiation profile (CD71 and CD235a) of the base edited HUDEP2 cells on the end stage of differentiation (Day 8) showing no significant variation among the edited samples; Mean ± SEM of n=3 independent replicates. f) Increase in percentage intracellular HbF levels measured end stage of differentiation (Day 8) by FACS (HbF+ cells) on generated BCL11A ZnF mutants; Mean ± SEM of n=3 independent replicates. g) The elevation in the total fetal hemoglobin tetramer formation in adult erythroid cells in the BCL11A modified cells analyzed at the end stage of differentiation (Day 8). Mean ± SEM of n=3 independent replicates.

4.1.4 Generation of BCL11A-ZnF mutants in HSPCs leads to HbF upregulation during *ex vivo* erythropoiesis with reduced hematological impacts

A previous study reported that eliminating hydrogen bonds formed between N756 and the -118 (A: T) base by creating N756A mutation reduces the affinity of BCL11A for DNA by eight-fold. To confirm that changing N756 to a Glycine as achieved via base editing had the same effect we validated the loss of binding of the ZnF4-1 mutant to the –115 site in the *HBG1/2* promoters by gel shift assays using proteins containing S755G, N756G, and S755G and N756G double mutants (**Fig. 9a and b**). We observed that the N756G mutant, both individually and in combination with S755G, abrogated the binding to the target site, whereas the S755G mutant alone

did not significantly affect binding. Further to phenotypically evaluate the contribution of each mutant on fetal globin de-repression, we performed single-cell sorting in HUDEP2 cells and observed that N756G contributed majorly to the HbF induction, consistent with the gel shift assay results (**Fig. 9c**).

Given the results in HUDEP2 cells, we next tested the effect of base editing at BCL11A-XL ZnFs in healthy donor $CD34^+$ HSPCs, using ABE8e mRNA, and again compared with the control *BCL11A* associated targets. We obtained base conversions at the respective target sites with overall efficiencies of 60-70% in HSPCs **(Fig. 9d)**. As expected, the ZnF4-1 gRNA generated two missense mutations at N756 and S755 at ZnF4 whereas the gRNAs targeting the other *BCL11A* associated targets altered the *BCL11A* start codon to achieve total knockout or the disruption at the intronic enhancer, or at the -115 site in the *HBG1/2* promoters. Notably, all the base modifications persisted throughout erythroid differentiation **(Fig. 9d)**.

To investigate the effect of the *BCL11A* ZnF4 mutant HSPCs on cell survival and differentiation, we monitored the base-modified HSPCs throughout differentiation to the erythroid lineage. Cell proliferation in the *BCL11A* knockout mutant was considerably lower than that in the ZnF4 mutant confirming that the BCL11A-XL isoform has a role in erythroid cell proliferation and survival **(Fig. 9e)**. Importantly, the specific mutations in ZnF4 domain did not impair cell viability.

We anticipated that disrupting multiple binding of BCL11A in the globin locus could lead to stronger upregulation than the ablation of the single BCL11A binding site at *HBG1/2* promoters. Therefore, we assessed the potency of BCL11A ZnF4 missense mutant on de-repression of fetal *globin* expression in erythroid progenitors derived

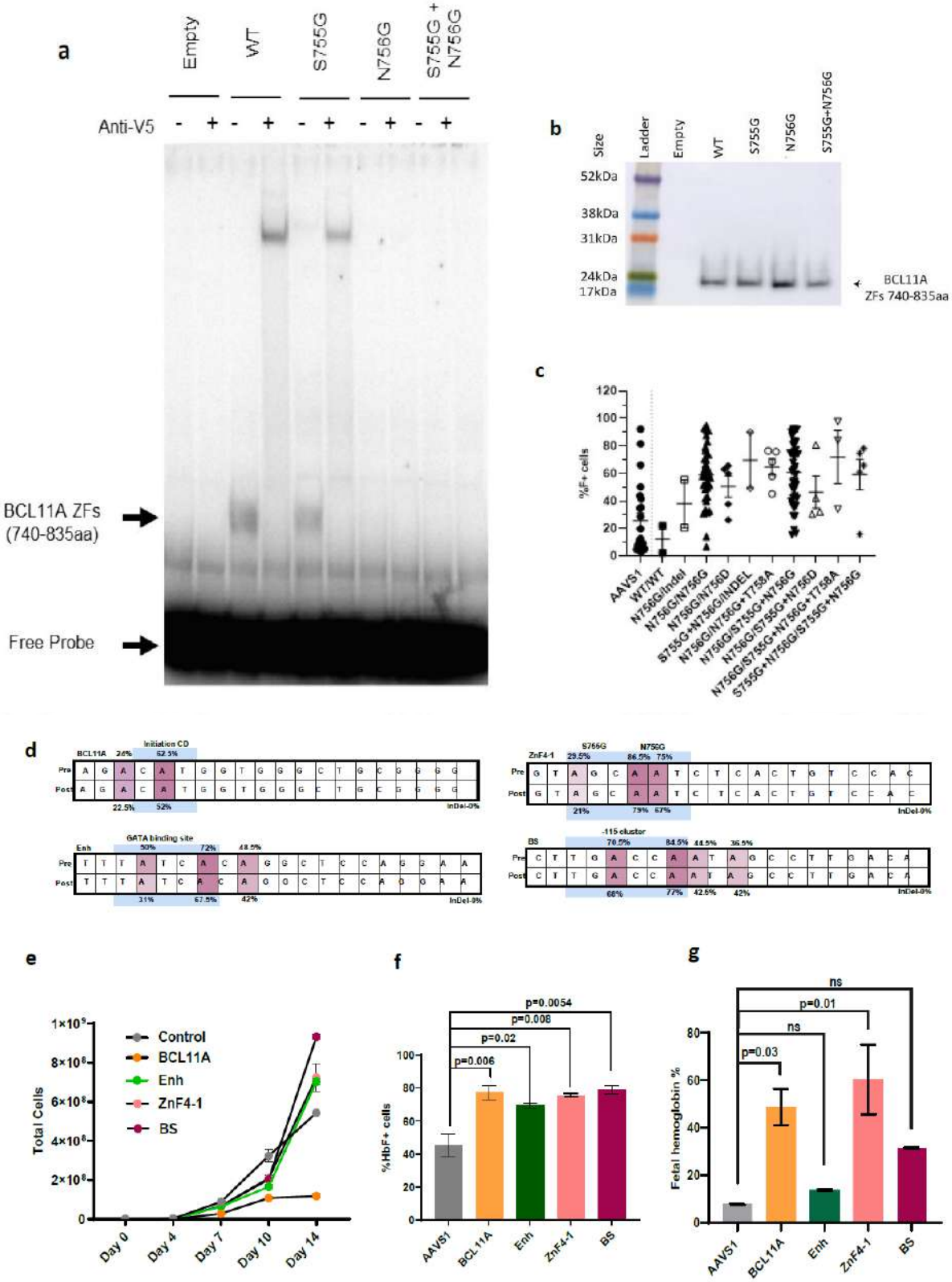
from base-edited HSPCs. Similar to the total knockout, the *BCL11A* ZnF4 mutation resulted in high HbF+ cells **(Fig. 9f)** and high total HbF production **(Fig. 9g)**.

We then investigated the effect of the *BCL11A* ZnF4 missense mutant on terminal erythropoiesis as our CRISPR/Cas9 cutting study showed a significant erythroid maturation delay in BCL11A ZnF truncated protein-expressing erythroid cells during the terminal stages. Interestingly, BCL11A ZnF4 base edited cells underwent terminal erythroid maturation ($CD235a^+/NucRed^-$) more efficiently than total knockout cells and truncated proteins generated in the CRISPR/Cas9 cutting experiments on day 21 of HSPC differentiation **(Fig. 9h)**. Specifically, we also observed higher HbF-expressing enucleated cells at the erythroid terminal maturation phase in the BCL11A ZnF mutant **(Fig. 9i)**. Notably, the BCL11A ZnF4 mutant exhibited a modest decrease in the frequency of cells expressing erythroid differentiation markers ($CD235a^+/CD71^-$) with normal enucleation **(Fig. 9j)**. However, this effect was modest when compared to CRISPR/Cas9 mediated truncation of the BCL11A ZnF4 domain. Altogether, these data demonstrate that the fetal *globin* reactivation induced by altering the individual amino acids specific to the BCL11A-XL isoform involved in binding to the *HBG1/2* promoters is preferable to indel formation via CRISPR/Cas9 and if refined could ultimately serve as a strategy for fetal globin induction.

We then investigated the repopulation ability of the BCL11A ZnF4 base-edited human $CD34^+$ HSPCs by transplanting into immunodeficient NBSGW mice, noting that BCL11A truncated proteins from the CRISPR/Cas9 experiment exhibited defective engraftment. We observed efficient chimerism of human cells (hCD45+) in

the control edited cells whereas the ZnF4 mutant exhibited reduced engraftment **(Fig. 9k)**, but importantly the ZnF4-1 mutant cells were able to survive and repopulate over time. As the neurodevelopmental patients harboring certain missense mutations in BCL11A-XL specific ZF domain exhibit essentially normal hematopoiesis, we analyzed the impact of *BCL11A* ZnF4 mutant on multilineage repopulation potential. Consistent with a moderate reduction in engraftment, ZnF4-1 mutants also showed altered levels of lineage distribution, notably a significant decline in B-cells, with a bias towards myeloid cells **(Fig. 9k)**. Although re-creating the exact patient SNVs using base editor is not feasible, we aimed to mimic the heterozygous nature of these BCL11A-ID patients by reducing the editing efficiency at the target site. Genotypic analysis of the engrafted human cells showed an evident drop in the editing frequency of the ZnF4-1 mutant within both the bone marrow (2.5-fold) and total loss in the B-cell engrafted compartment **(Fig. 9l).** Interestingly, we observed that N756G amino acid production to be prevalent over N756D or N756S amino acids. Further N756G amino acid also persisted at the 15 weeks of *in vivo* engraftment in NBSGW mice **(Fig. 9m).** Importantly, the CD235a+ erythroblasts present in the mouse bone marrow of the ZnF4-1 mutant engrafted mice showed an increase in HbF+ cells compared to edited control **(Fig. 9n)**.

Further, the persistence of single mutants in the bone marrow population suggests that creating mono-allelic single mutations resembling the patient SNVs at the ZnF domains elevates fetal hemoglobin. However, the data also suggests that C-terminal ZnF domains have important functions in HSPC engraftment and B-Cell production outside the globin regulation within the erythroid lineage.


a
Empty
WT
S755G
N756G
S755G + N756G
Anti-V5
BCL11A ZFs (740-835aa)
Free Probe
b
Size
Ladder
Empty
WT
S755G
N756G
S755G+N756G
52kDa
38kDa
31kDa
24kDa
17kDa
BCL11A ZFs 740-835aa
c
%F+ cells
AAVS1
WT/WT
N756G/Indel
N756G/N756G
N756G/N756D
N756G/INDEL
S755G+N756G+T758A
N756G/N756G+T758A
N756G/S755G+N756G
N756G/S755G+N756D
N756G/S755G+N756G+T758A
S755G+N756G/S755G+N756G
d
BCL11A
Initiation CD
Enh
GATA binding site
ZnF4-1
S766G
N756G
BS
-115 elustor
Pre
Post
InDel-0%
e
Total Cells
Control
BCL11A
Enh
ZnF4-1
BS
Day 0
Day 4
Day 7
Day 10
Day 14
f
%HbF+ cells
p=0.0054
p=0.008
p=0.02
p=0.006
AAVS1
BCL11A
Enh
ZnF4-1
BS
g
Fetal hemoglobin %
ns
p=0.01
ns
p=0.03
AAVS1
BCL11A
Enh
ZnF4-1
BS

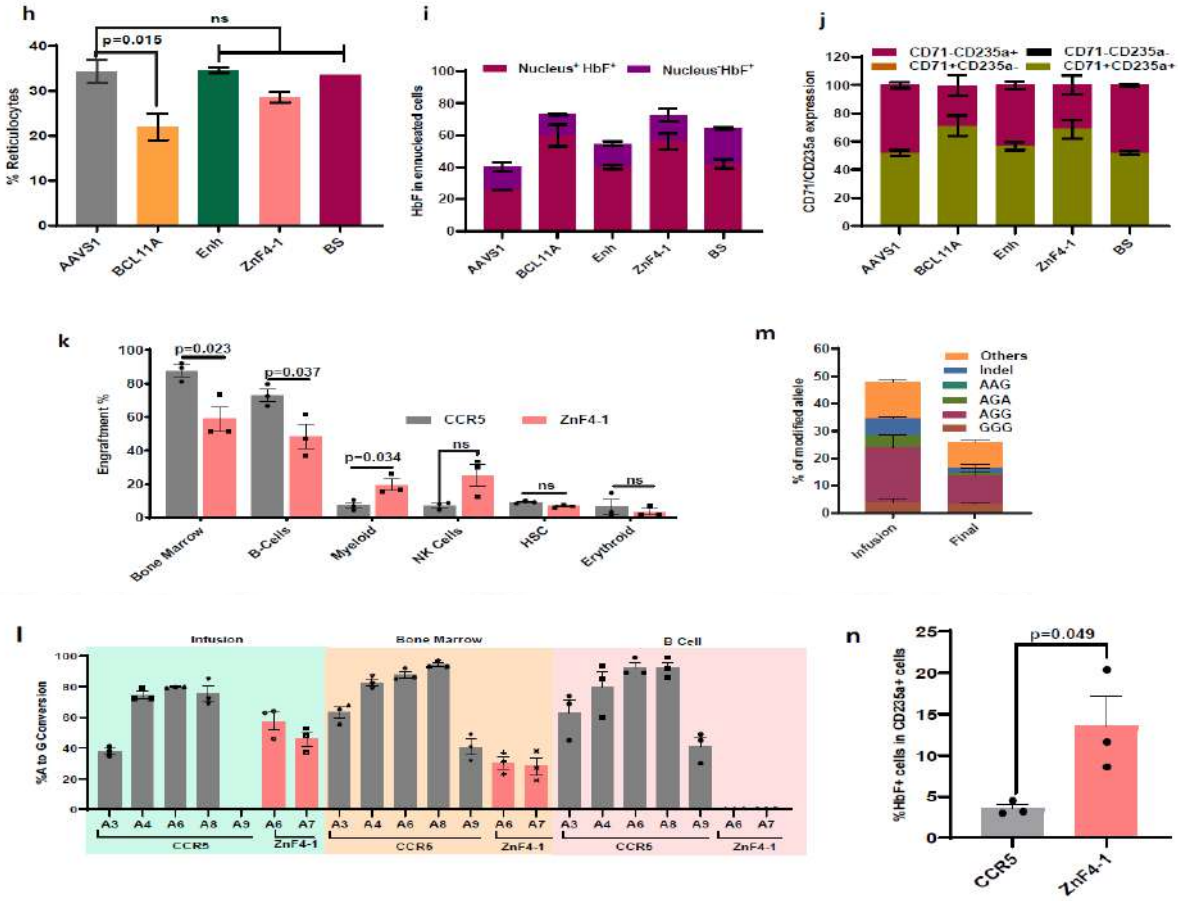


Fig 9: Generation of BCL11A-ZnF mutants in HSPCs leads to HbF upregulation during *ex vivo* erythropoiesis with reduced hematological impacts. a) EMSA depicting the loss of interaction between the ZnF4-1 (N756 mutants) mutants to a TGACC motif. Nuclear extract of COS-7 cells expressing an epitope-tagged fused zinc finger (ZF) 4-6 of WT BCL11A and mutants with radio-labeled fetal globin promoter BCL11A binding motif TGACC. The black arrow shows the BCL11A ZF4-6 binding to the probe and super-shifted bands are observed when antibody to the V5 epitope is added. b) Western blot showing equivalent amounts of recombinant V5 tagged BCL11A ZF4-6 proteins were used in gel shift assays. The Western blot was carried out with the same amount of COS-7 nuclear extract used in the gel shift experiment in (A) using an antibody against V5 and standard methods. c) Percentage of HbF+ cells exhibited in the erythroid cells exhibiting various genotypes due to ZnF4-1 editing in HUDEP2 cells obtained through single cell analysis followed by 8 days of erythroid differentiation. Single-cell clones of AAVS1 edited HUDEP2 cells were also analyzed separately for HbF+ cells after differentiation. d) Target base perturbation efficiency at each target site was analyzed after 2 days of electroporation in EditR software. Mean ± SEM of n=2 duplicates of 1 donor. e) The cell proliferation curve of the AAVS1(control),BCL11A total loss(BCL11A), Enhancer(Enh), Binding Site(BS) and the ZnF4(ZnF4-1) base edited HSPCs during erythroid differentiation. f) The reactivation of total HbF measured by HPLC at the terminal stage of erythroid differentiation (Day 21). Mean ± SEM of n=2 duplicates of 1 donor. g) The percentage

of erythroid cells expressing HbF measured by FACS at the terminal day of differentiation (day 21). Mean ± SEM of n=2 duplicates of 1 donor. All samples are compared with the AAVS1 control with one-way ANOVA analysis. h) The percentage of edited cells capable of maturing into enucleated reticulocytes in vitro measured by CD235a expression and NucRed stain. Both erythroid differentiation and maturation analysis were measured at the terminal stage of erythroid differentiation (Day 21). i) The measure of HbF expression in two different populations ($Nucleus^{+}HbF^{+}$ and $Nucleus^{-}HbF^{+}$) in differentiated erythroid cells. j) The erythroid differentiation potential of base-edited HSPCs was examined using CD235a and CD71 surface markers at the terminal stage of erythroid differentiation (Day 21). k) Engraftment percentage of base edited HSPCs and multilineage repopulation in NBSGW mice after 15 weeks of transplantation. l) The percentage of base editing observed before infusion, after 15 weeks in the mouse bone marrow and in the B-cell subset. B-cells were flow-sorted using human CD45 and CD19 surface markers. m) NGS analysis depicting the multiple editing patterns within engrafted ZnF4-1 edited cells derived from the mouse bone marrow after 15 weeks of engraftment. A AA- First 'A' denotes the adenine base in the 755^{th} position and 'AA' denotes the dual Adenine bases present in the N756 position which gets converted to G with different frequencies. Others in the graph indicate the summation of all the non-specific mutations. n) Percentage of HbF+ cells in the CD235a+ cell population within the NBSGW bone marrow after 15 weeks. n=3; Mean ± SEM.

4.1.5 Global transcriptomic analysis of the BCL11A ZnF4 base edited erythroid cells derived from $CD34^{+}$ HSPCs

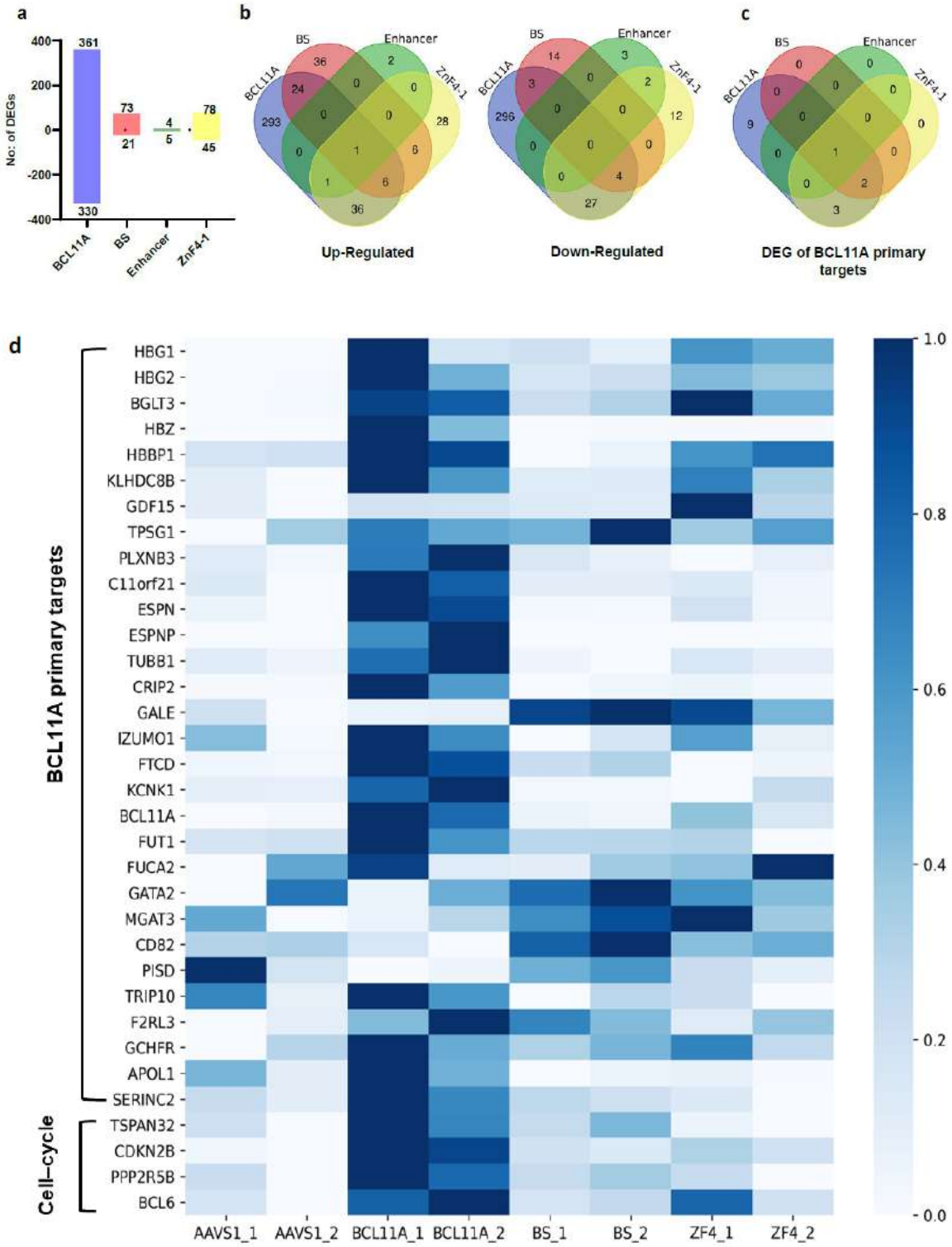
As BCL11A regulates multiple pathways involved in the development and differentiation of various cell types, any alteration of *BCL11A* requires extensive characterization.(Luc et al., 2016; Tsang et al., 2015) Therefore, we compared genome editing at the XL-specific ZnF domain along with the other BCL11A-associated targets. A total of 691 genes were differentially expressed after the complete loss of *BCL11A*, whereas disruption of ZnF4 (123 genes) and binding site (94 genes) showed significantly fewer differentially expressed genes. The expression of only 9 genes was observed to be significantly altered in the case of erythroid-enhancer disruption, consistent with the modest fetal hemoglobin induction observed *in vitro* after erythroid

differentiation. (**Fig. 10a and b)**. The transcriptional repressor activity of BCL11A is evident based on the fact that more genes were upregulated on modifying BCL11A compared to the binding site disruption in the *HBG1/2* mutation (BS). This demonstrates that the BCL11A ZnF4 mutant generation has fewer impacts beyond the upregulation of fetal globin expression than observed with the total ablation of *BCL11A* expression.

Next, we examined the known primary targets[41] of BCL11A exhibiting significant differential expression among our study group **(Fig. 10c)**. In all the target sites *HBG2* was the commonly upregulated gene as expected. *BGLT3* and *HBG1* are the two other genes commonly upregulated among the ZnF4-1, BS, and BCL11A edited cells. Interestingly, the ZnF4 editing shared upregulation of the *HBZ*, *HBBP1*, and *KLHDC8B* genes with BCL11A total knockout. Consistent with the whole transcriptome profile, the disruption of the BCL11A binding site generates no other significant alterations in the primary targets **(Fig. 10d)**. Moreover, the significant increase in the expression of the cell cycle-associated genes, such as *CDKN2B, PPP2R5B*, and *BCL6* occurs only in BCL11A complete knockout cells, and not in the ZnF4 mutant cells. This may explain why the ZnF4 mutation has modest impacts on erythroid function other than globin regulation. We further validated the expression of these globin locus genes by qRT-PCR **(Fig. 10e)**. These results again confirm that specifically disrupting the ZnF4 domain has fewer impacts on other erythroid genes compared to the total ablation of BCL11A.

Next, to evaluate whether the observed induction of fetal hemoglobin is specific to the on-target editing activity, we performed targeted deep sequencing at the potential off-target sites (with up to 3 mismatches) predicted by the *in silico* algorithm,

COSMID (**Fig. 10f)**. The top 9 sites were deep sequenced in the base edited HSPCs and we observed no significant DNA variation between the control and the edited samples indicating that the gRNA used is highly specific to the *BCL11A* ZnF4 locus. The CIRCOS plot depicting the predicted off-target (OT) locations shows that none of the OT sites are in the coding region, providing further evidence that the gRNA is highly specific **(Fig. 10g)**. To further rule out the impact of unintended RNA deamination caused by ABE8e, we analyzed the whole transcriptome RNA-seq information of the ZnF4-1 edited HSPCs during erythroid differentiation and compared it with the other BCL11A associated targets. We observed no significant difference in the A-to-I conversion among the gRNAs used in this study **(Fig. 10h)**. These results further confirm the specificity of the gRNA used in targeting the ZnF4-1 using ABE8e.


a
No: of DEGs
400
200
0
-200
-400
361
330
73
21
4
5
78
45
BCL11A
BS
Enhancer
ZnF4-1
b
BS
Enhancer
BCL11A
ZnF4-1
Up-Regulated
Down-Regulated
c
DEG of BCL11A primary targets
d
BCL11A primary targets
Cell-cycle
HBG1
HBG2
BGLT3
HBZ
HBBP1
KLHDC8B
GDF15
TPSG1
PLXNB3
C11orf21
ESPN
ESPNP
TUBB1
CRIP2
GALE
IZUMO1
FTCD
KCNK1
BCL11A
FUT1
FUCA2
GATA2
MGAT3
CD82
PISD
TRIP10
F2RL3
GCHFR
APOL1
SERINC2
TSPAN32
CDKN2B
PPP2R5B
BCL6
AAVS1_1
AAVS1_2
BCL11A_1
BCL11A_2
BS_1
BS_2
ZF4_1
ZF4_2
1.0
0.8
0.6
0.4
0.2
0.0

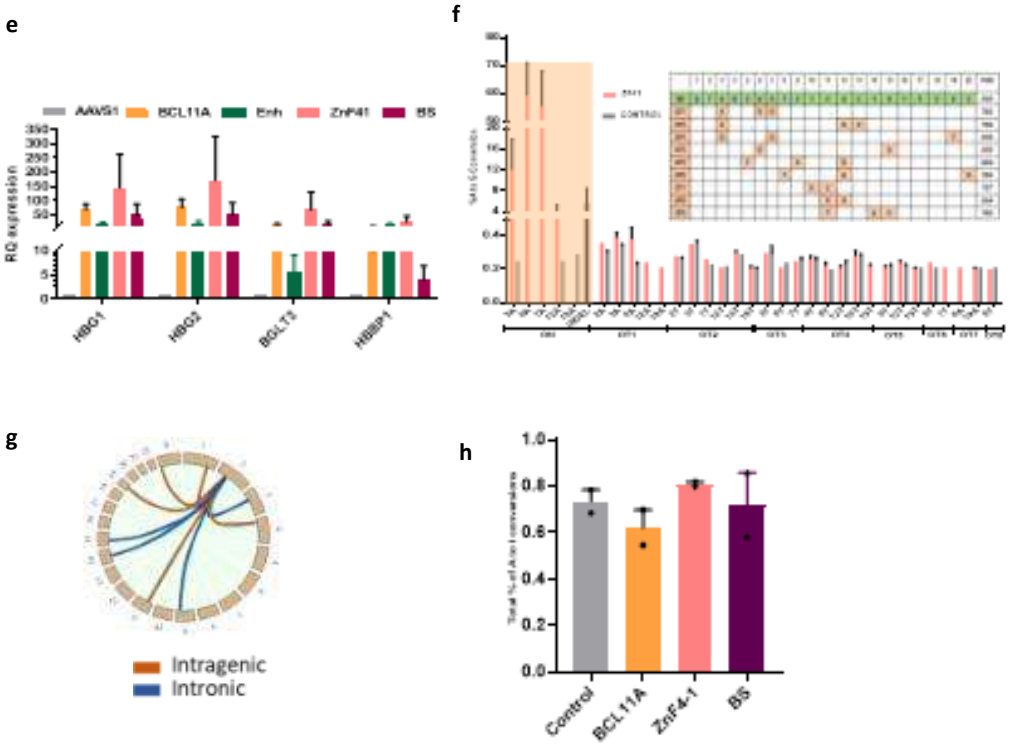


Fig 10: Global transcriptomic analysis of the BCL11A ZnF4 base edited erythroid cells derived from CD34^{+} HSPCs. a) Total number of differentially expressed genes in total knockout, enhancer edited, ZnF4-1 modification, and BS disruption compared to the edited control. b) Correlation between the different transcriptomes among upregulated and downregulated genes in total knockout, enhancer edited, ZnF4-1 modification, and BS disruption compared to the edited control. Significant gene expression is maintained at FDR-5% and -1>Fold Change >+1. c) Correlation showing the significantly expressed (FDR-5% and -1>Fold Change >+1) primary targets among the different edited groups. d) Heatmap showing the BCL11A primary target gene counts among all the edited HSPCs. e) *HBG*1, *HBG*2 and BGLT3 and HBBP1 upregulation validated by qPCR. Each sample is performed in biological duplicates and each is replicated as technical triplicates. f) Table displaying all possible predicted off-target sites based on the mismatches in the ZnF4-1 sgRNA; The percentage of on-target editing compared to all the predicted off-targets measured by targeted amplicon NGS followed by CRISPResso2 analysis. g) Circos plot depicting the chromosomal position of each off-target site in the human genome and the nature of the site. Red denotes Intragenic and Blue denotes Intronic regions. h) Total percentage of A-to-I (T-to-C) conversion at the RNA level in all the edited samples. Performed as biological duplicates.

4.2 Identification of novel co-activator mediated regulation of DHS58 intronic BCL11A enhancer

4.2.1 Editing the core DHS58 region within the erythroid-specific BCL11A enhancer in HUDEP2 cells.

BCL11A is a tightly regulated transcription factor with high expression in neuronal and hematopoietic cells, especially the erythroid and the B-cell lineages. Within the erythroid lineage, the elevated expression is attributed to the 3 DNA hypersensitivity sites (DHS 55, DHS 58, and DHS 62) of the enhancer present in the intron 2 of BCL11A gene. The core region within the +DHS58 of BCL11A enhancer contains many transcriptional activator binding sites like the Tal1, GATA1, E-Box, and eKLF1 which enhances the expression of BCL11A specifically within the human adult erythroid lineage. Of these, the GATA1 motif, present between the Tal1 and E-box element, is the most characterized target taken to clinical trials for beta-hemoglobinopathies. There are 2 eKLF1 sites present approximately 35bp upstream of the GATA motif. Previous studies have reported that the KLF1 binding stabilizes the interaction of GATA1 with its binding site (1617 sgRNA site). These eKLF1 sites have not been extensively studied for their functional role in BCL11A enhancer activation.

We designed sgRNAs targeting these GATA1 and the putative eKLF1 sites to evaluate the effect of binding site alteration on BCL11A expression and the resultant fetal hemoglobin elevation **(Fig. 11a)**. 1617 is the previously studied GATA motif disruption target that is currently pursued in the clinics. +33bp GATA and the CB11 sgRNA disrupt the proximal KLF1 site whereas the CB13 and the CB15 disrupt the

distal KLF1 site from the GATA motif. HUDEP-2 cells stably expressing cas9 were transduced with lentivirus containing these sgRNAs and evaluated for its impact on BCL11A expression and fetal hemoglobin elevation.

We achieved more than 90% editing at all the target sites with the indel profile disturbing the desired KLF1 sites **(Fig. 11b and c)**. The edited samples were subjected further to erythroid differentiation. As expected, we observed significant loss of BCL11A protein on targeting the GATA motif (sgRNA-1617) **(Fig. 11d)**. Surprisingly, there was no significant loss of BCL11A protein on targeting the KLF1 sites **(Fig. 11e)**. We also evaluated the BCL11A mRNA expression profile and observed significant loss of BCL11A expression in all the target sites **(Fig. 11f)** which contradicts the protein expression profile. This might suggest that novel regulation pattern through the KLF1 sites, though further investigations are required to prove this. The induction of fetal hemoglobin was used as an evaluation parameter for functional BCL11A production. We observed a significant elevation of fetal hemoglobin levels **(Fig. 11g and h)** in all the edited samples. On targeting the +33bp GATA region, we observed prominent and comparable induction to that of GATA motif disruption. The other proximal KLF1 site disruption sgRNA CB11 showed lower levels of elevation compared to the +33bp GATA guide. The distal KLF1 site disruption resulted in a significant elevation of fetal hemoglobin yet lower than the +33bp GATA and the GATA motif target. These data suggest that targeting the KLF1 site(s) elevates fetal hemoglobin levels through the production of non-functional BCL11A protein denoting an alternate mechanism of action of these sites or guides compared to the canonical GATA site regulation.

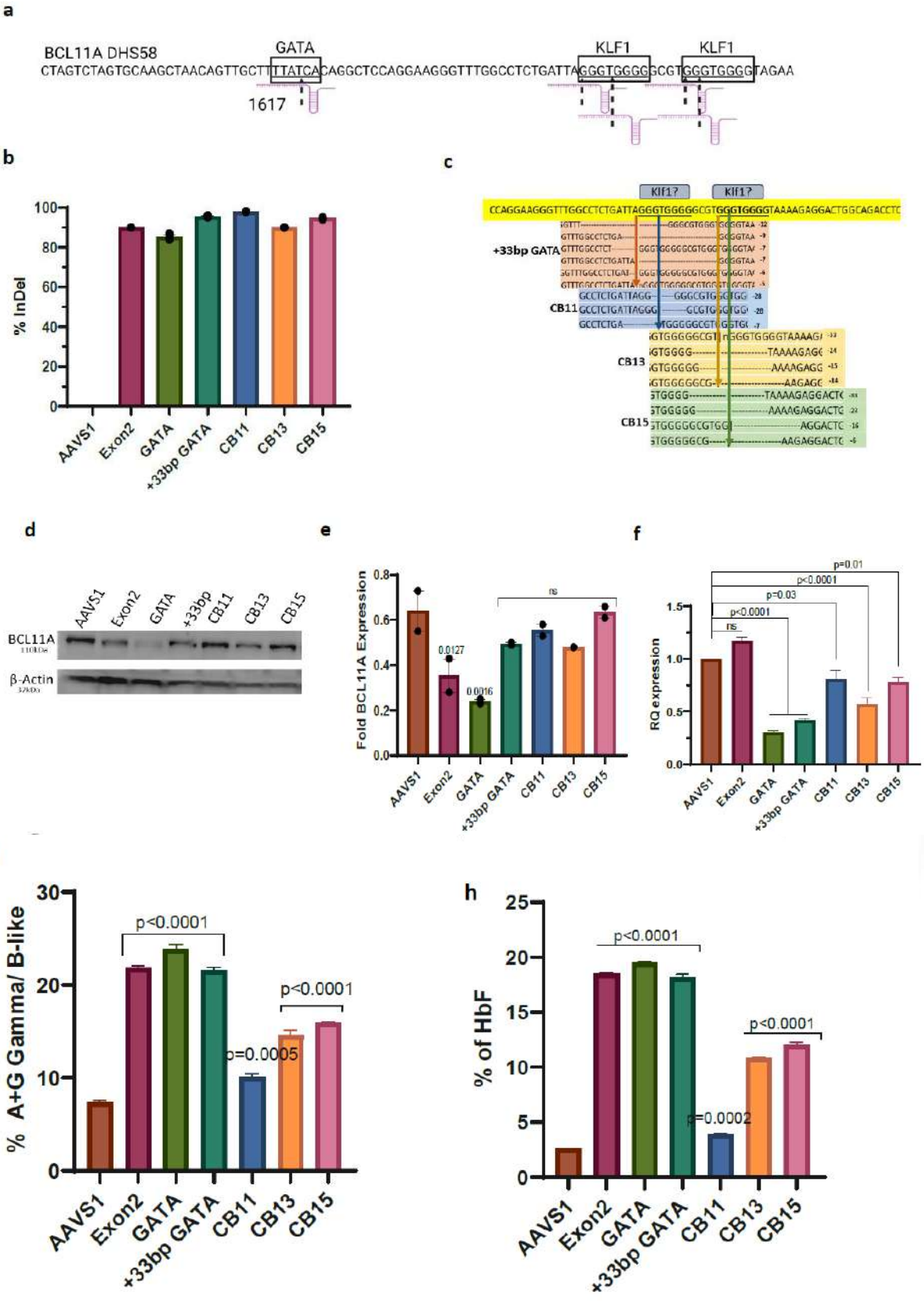

a
BCL11A DHS58
GATA
KLF1
KLF1
CTAGTCTAGTGCAAGCTAACAGTTGCTTTTATCACAGGCTCCAGGAAGGGTTTGGCCTCTGATTAGGGTGGGGGCGTGGGTGGGGTAGAA
1617
b
% InDel
AAVS1
Exon2
GATA
+33bp GATA
CB11
CB13
CB15
c
Klf1?
Klf1?
+33bp GATA
CB11
CB13
CB15
d
AAVS1
Exon2
GATA
+33bp
CB11
CB13
CB15
BCL11A
β-Actin
e
Fold BCL11A Expression
ns
0.0127
0.0016
f
RQ expression
p=0.01
p<0.0001
p=0.03
p<0.0001
ns
g
% A+G Gamma/ B-like
p<0.0001
p<0.0001
p=0.0005
h
% of HbF
p<0.0001
p<0.0001
p=0.0002

Fig 11: CRISPR/cas9-based disruption of KLF1 sites within the GATA motif proximity in HUDEP2 erythroid cell line. a) Schematic representation of guide RNAs target position. 1617 disrupts the known GATA motif and the +33bp GATA and the CB11 guide disrupts the proximal KLF1 site. CB13 and CB15 disrupts the distal KLF1 site. All these sites are present within the DHS58 of BCL11A enhancer. b) Gene editing percentage acquired after 8 days of lentiviral transduction observed through Sanger sequencing. n=3/sample. c) The editing profile at each target site caused by the respective guide RNAs. d) Representative Immunoblot depicting the presence of BCL11A protein and B-Actin among the edited samples. e) Quantification of protein expression obtained from western blot using ImageJ software. 2 duplicate blots were used for SEM calculation. f) BCL11A mRNA expression profiles among the edited samples were validated by qPCR. Each sample was performed in biological triplicates and each replicate as technical triplicates. g&h) The elevation in the individual gamma chains and total fetal hemoglobin tetramer formation in adult erythroid cells in the enhancer-modified cells analyzed at the end stage of differentiation (Day 8). Mean ± SEM of n=3 independent replicates.

4.2.2 Validating the +33bp GATA site disruption in healthy donor CD34+ HSPCs.

Next, the disruption of the +33bp GATA site was performed in multiple healthy donor HSPCs to evaluate its impact on fetal hemoglobin elevation. We also included the 1621 sgRNA previously reported to effectively reduce BCL11A expression along with deletion of the entire 33bp region spanning from the GATA motif to the +33 bp GATA site. Using these guide RNAs along with CRISPR/cas9 protein (as RNP complex) we evaluated the editing efficiency achieved in nucleofecting multiple healthy donor CD34+ cells. We successfully obtained consistent editing of more than 80% in all the donors **(Fig. 12a)**. Surprisingly, on analyzing the indel profile of the +33bp GATA sgRNA, we noted that the proximal KLF1 site was not as efficiently lost as previously observed while editing the HUDEP2 cell **(Fig. 12b)**. It was observed that the observed reduction in BCL11A expression on

targeting the 1617 and 1621 sgRNA can be associated with the disruption of the known GATA1 and the E-box domains along with the intermediate regions as previously reported in multiple studies. To evaluate the effect of this region between the GATA1 motif and the +33bp eKLF1 site, we employed multiplex editing using both these sgRNAs simultaneously to achieve deletion of this region. To compare our enhancer disruption targets with the total knockout, we also targeted the BCL11A exon2 to achieve the complete loss of BCL11A expression. The impact of editing on *in vitro* multi-lineage differentiation potential was analyzed by colony-forming assay. We observed no difference between the edited samples compared to the edited control apart from BCL11A total knockout cells **(Fig. 12c).**

Next, the edited CD34+ cells were subjected to erythroid differentiation to functionally characterize the target site disruption. No significant variation in the growth kinetics of the edited cells during differentiation was observed with the expected exception of the total BCL11A knockout showing a slight reduction in cell count at the terminal stages of differentiation **(Fig. 12d)**. The imminent reduction in BCL11A mRNA due to the DNA disruption achieved was demonstrated by the RT-qPCR highlighting the importance of these motifs for the BCL11A expression **(Fig. 12e)**. On targeting the exon2 region, the drop in BCL11A mRNA is not observed suggesting that the cDNA produced in the total BCL11A knockout cells were defective. Further, we observed no significant difference in the differentiation profile, CD71, and CD235a levels, highlighting that disrupting the targets didn't affect the erythroid differentiation. Enucleation potential was better when the +33bp GATA site was disturbed compared to the GATA motif disruption and the total BCL11A loss as compared to the control edited cells **(Fig. 12f)**. The F^{+ve} cells were analyzed at the end

stage of differentiation **(Fig. 12g)** and showed significant elevation on editing the target sites corroborating the loss of BCL11A due to editing. HPLC chains showed that there is no bias in the production of *HBG*1 and *HBG*2 **(Fig. 12h)** and through variant analysis elevation in the fetal hemoglobin tetramer was also evident **(Fig. 12i)**.

From these results, we note that targeting the +33bp GATA motif in the BCL11A enhancer core region is capable of significantly elevating fetal hemoglobin levels with efficient terminal differentiation and enucleation of the edited erythroblasts. This showcases the presence of other regulatory regions within the BCL11A enhancer suggesting that complex activation pathway for the BCL11A enhancer.

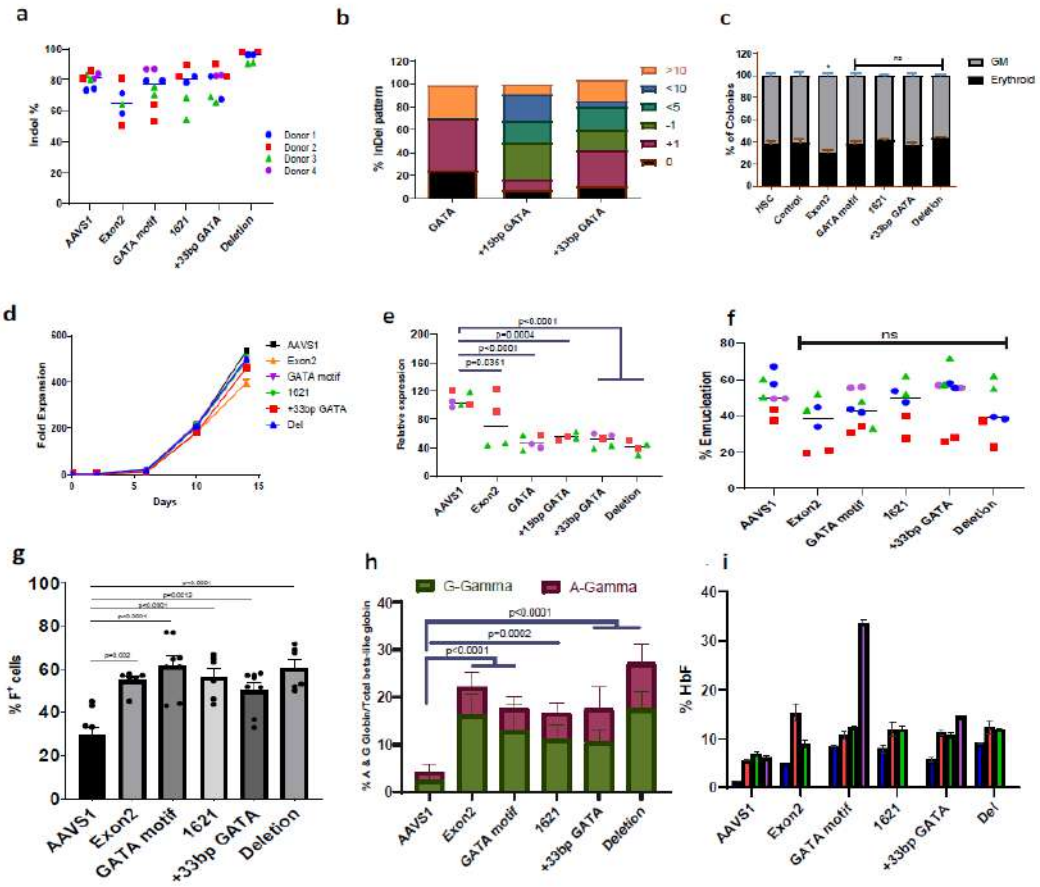


Fig 12: Validating the +33bp GATA site disruption in healthy donor CD34+ HSPCs. a) Percentage of gene editing achieved among 4 different donors. b) InDel profile of 1617, 1621 and +33bp GATA. c) Invitro multi-lineage differentiation capability was analyzed through methylcellulose assay. 500 cells/ml were seeded 48-72 hours after electroporation and scored on days 14-17 based on their morphological differences. d) Representative growth curve showing fold increase from the initial day of seeding. All the edited cells show similar growth profiles during erythroid differentiation. e) BCL11A mRNA expression profile during erythroid differentiation was validated by qPCR among the edited samples on day 8 of phase-1 differentiation. Each sample was performed in biological triplicates and each replicate as technical triplicates. f) The percentage of edited cells capable of maturing into enucleated reticulocytes in vitro measured by CD235a expression and NucRed stain. Both erythroid differentiation and maturation analysis were measured at the terminal stage of erythroid differentiation (Day 21). g) The percentage of erythroid cells expressing HbF measured by FACS at the terminal day of differentiation (day 21). All samples are compared with the AAVS1 control with one-way ANOVA analysis. h) Analysis of globin chain production through reverse-phase HPLC at the terminal stage of erythroid differentiation (Day 21). i) The reactivation of total HbF measured by HPLC at the terminal stage of erythroid differentiation (Day 21). n=2replicates/donor.

4.2.3 In vivo characterization of BCL11A enhancer edited cells in NBSGW mice.

To evaluate the homing and engraftment efficiency of the +33bp GATA-edited HSPCs, we performed xenotransplantation experiments in NBSGW mice with the edited HSPCs. The NBSGW mice were pre-conditioned with 12.5mg/kg body weight of Busulfan 48 hours before the infusion. 16 weeks after the infusion, the NBSGW mice showed similar engraftment of the human cells between the control edited and target edited samples **(Fig. 13a)**. We also observed female mice showing better engraftment than the male counterparts across all the study groups. The multilineage differentiation potential of the target-edited HSPCs was similar to the control-edited HSPCs in both the study cohorts **(Fig. 13a)**. These results clearly showed that there is no bias in the control edited and the +33bp GATA-edited HSPCs in forming multi-lineage cells within the in vivo system. The persistence of gene editing throughout the engraftment process was analyzed by genotyping the bone marrow after 16 weeks. We noted an increase in the gene editing percentage after 16 weeks of engraftment *in vivo* **(Fig. 13b)**. This can be partially explained by the survival advantage the edited cells might hold inside the *in vivo* system compared to the *in vitro* culture conditions due to the presence of a multitude of factors within the mice serum/plasma. Serial transplantation was performed, and we observed no reduction in engraftment compared to the control sample upon analyzing the bone marrow of the secondary recipient after 8-10 weeks elucidating that the current procedure was not toxic to the long-term HSCs within the system **(Fig. 12c)**. Upon bone marrow collection from primary recipients, we performed ex-vivo erythropoiesis **(Fig. 13d and e)** and

observed significant elevation in the fetal hemoglobin levels between the control edited and the +33bp GATA edited cells **(Fig. 13g)**.

Together with the in vitro data, we can strongly conclude that disrupting the +33bp GATA site in the BCL11A enhancer using cas9 significantly elevates fetal hemoglobin levels. The innate functions of the HSPCs like engraftment and multi-lineage differentiation potential are also unaffected implying the specificity of the BCL11A enhancer.

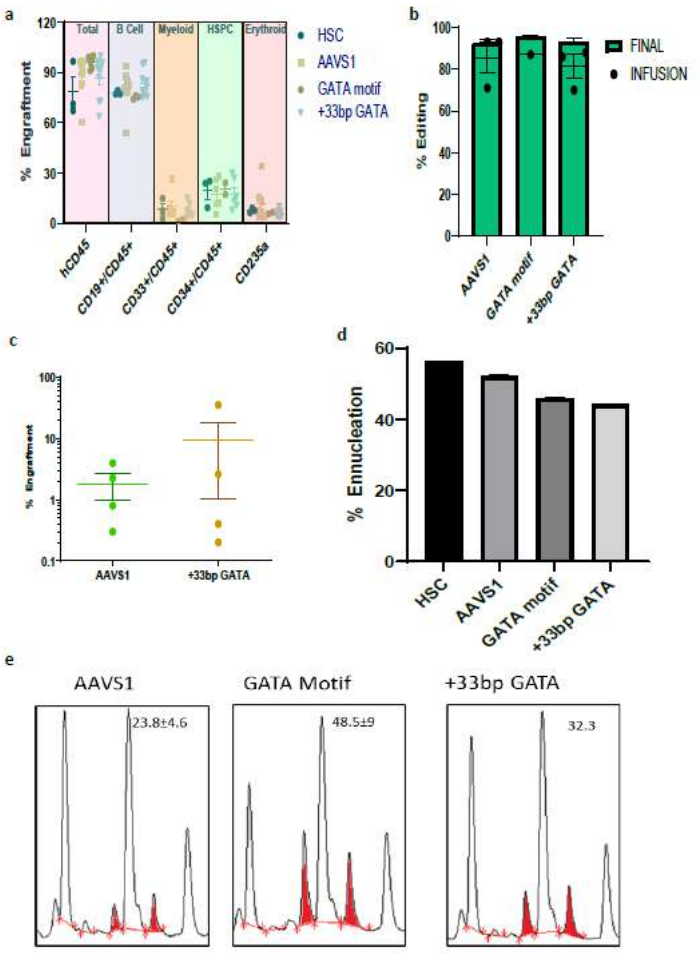


Fig 13: In vivo characterization of BCL11A enhancer edited cells in NBSGW mice. a) Engraftment percentage of base edited HSPCs and multilineage repopulation in NBSGW mice after 16 weeks of transplantation. b) The percentage of gene editing observed before infusion, after 16 weeks in the mouse bone marrow c) Percentage of secondary engraftment observed 10 weeks after primary bone marrow infusion in secondary recipients. d) The percentage of enucleated reticulocytes observed on ex vivo erythroid differentiation of bone marrow cells measured by CD235a expression and NucRed stain. e) Induction of gamma-globin chains on the terminal differentiation (day 26) of ex vivo erythroid differentiation among gene-edited samples.

4.2.4 Understanding the mechanism of action of the +33bp GATA site in BCL11A regulation.

Surprisingly, editing at the +33bp site has downregulated BCL11A mRNA in CD34+ HSPCs resulting in elevated fetal hemoglobin levels without disrupting the proximal KLF1 site. To understand the molecular mechanism of the +33bp GATA site that regulates the BCL11A enhancer, we performed a whole transcriptomic analysis of GATA and the +33bp GATA edited samples along with the control edited cells. We noted that +33bp GATA disruption elevated 100 unique genes among the 1617 and the 1621 site disruption. Surprisingly, we also observed variation in the total number of genes dysregulated between the GATA edited and the +33bp GATA edited samples **(Fig. 14a and b)**. This data suggests that disruption of +33bp GATA has an alternate mode of regulating the BCL11A DHS58 enhancer.

To understand the mechanism of regulation on the +33bp GATA site effects on the BCL11A enhancer, we performed ATAC-seq on the edited HSPCs to evaluate the conformational changes of the region post-editing. We observed that, on disrupting the GATA motif the entire enhancer region becomes closed shutting off BCL11A expression. But on disrupting the +33bp region, we observed that the enhancer region persists in the open conformation **(Fig. 14c)**. Next, we performed CUT&RUN to analyze the presence of Tal1 and GATA1 binding to the adjacent GATA1 and Tal1 motifs. We observed that, despite the +33bp GATA site being disrupted, there is no loss of GATA1 **(Fig. 14d)** or Tal1 **(Fig. 14e)** binding as observed with the GATA motif disruption. This suggests the presence of a novel regulatory pathway through

which the +33bp region regulates the BCL11A erythroid-specific enhancer in a GATA-independent way.

Further, to eliminate the role of off-target activity of the +33bp GATA sgRNA in elevating fetal hemoglobin, we performed targeted deep-sequencing at the top predicted off-target sites. Recent understanding of the human genome and the off-target probability suggest that the off-target proclivity of each guide RNA used varies with individuals. Hence, we evaluated the effect of the unintended editing by the sgRNA (+33bp GATA) predicted by the *in silico* COSMID algorithm in 2 different CD34+ HSPCs donors. Most of these predicted sites were either intergenic (4 sites) or intronic (12 sites) and none were in the coding region of any known gene **(Fig. 14f)**. These predicted sites were deep sequenced in 2 donor post-erythroid differentiation, and we observed no significant editing at these sites compared to the control edited cells **(Fig. 14g)**.

Since, disrupting the nucleobase sequence at the target site, using cas9, disrupts the expression of the gene, we hypothesized the presence of an activator binding to the target region. Hence, to test this hypothesis, we masked this site using the dead cas9 variant and the +33bp GATA sgRNA in HUDEP2 cells. We achieved significant transduction of sgRNA within the HUDEP-dcas9 stable **(Fig. 14h)**. The associated dCas9 molecule prevents the binding of any transcriptional activators (GATA1 in case of 1617 and an unknown factor at +33bp GATA site). We observed that on blocking the GATA1 motif and the +33bp GATA site, significant elevation in F+ve cells **(Fig. 14i)** was observed. As expected, on blocking the -115 cluster in the *HBG*1/2 promoter, a reduction in F+ve cells was observed due to the non-binding of NFY because of the

hindrance caused by dCas9. These data suggest the presence of an activator binding at the +33bp GATA region. We utilized the JASPER algorithm and identified 5 factors (Hoxa13, Hoxb2, Hoxa2, Hoxd3, and GSX1) having a binding site that closely resembles the +33bp GATA site **(Fig. 14j)**. Next, to evaluate the occupancy of these factors at the target site, we performed knockout of these factors and checked for the elevation of F^+ cells. We observed that there was no increase in the percentage of F^+ cells whereas the total loss of BCL11A and enhancer disruption showed higher F+ cells **(Fig. 14k)**. This suggests that these predicted transcriptional factors may not be responsible for the activation of this +33bp GATA site within the BCL11A enhancer.

Together, these data suggest the presence of a novel regulatory mechanism at the +33bp GATA region through which functional activity of BCL11A DHS58 erythroid-specific enhancer is regulated in a GATA-independent manner.

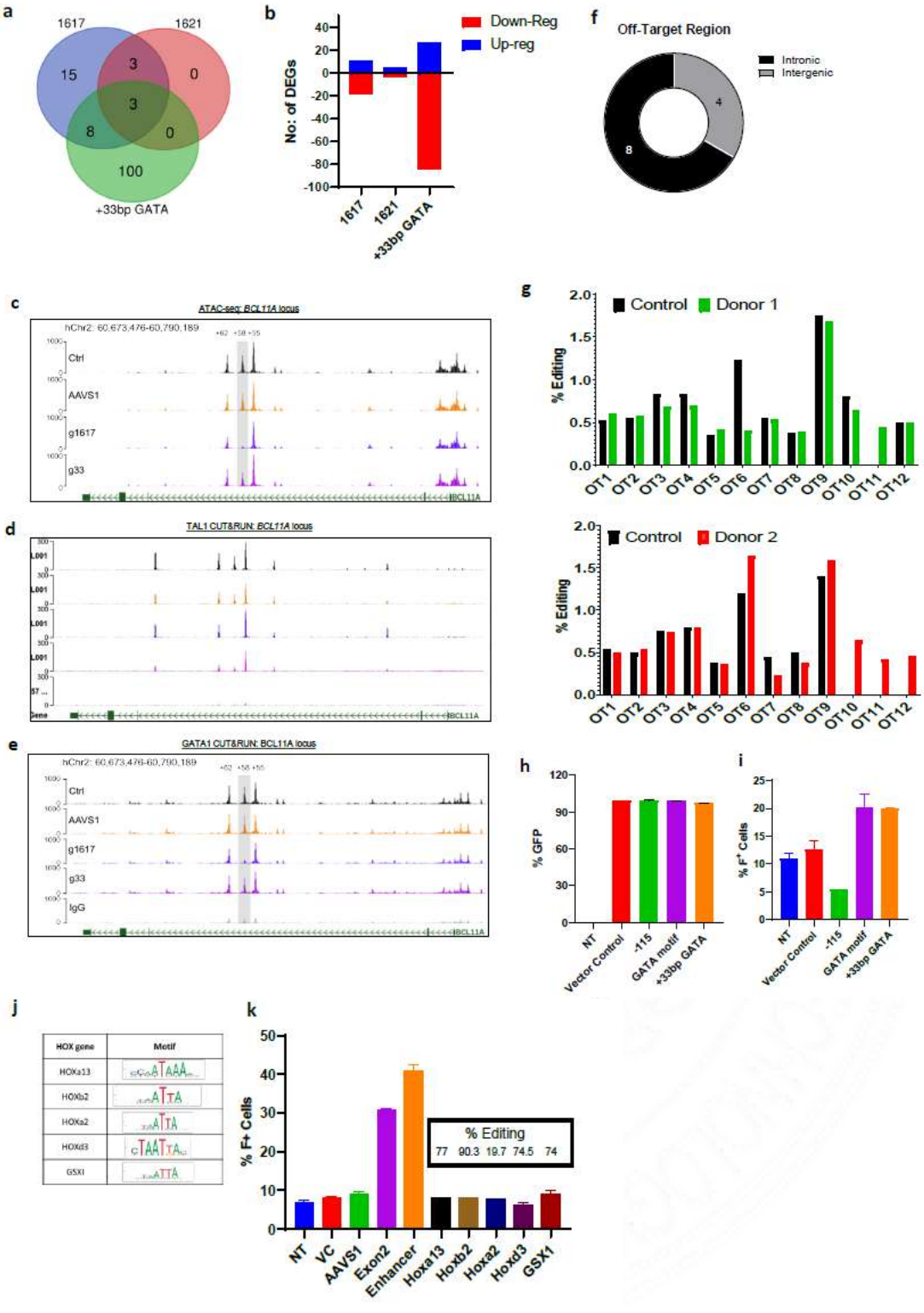

a
1617
1621
15
3
0
3
8
0
100
+33bp GATA
b
Down-Reg
Up-reg
No: of DEGs
40
20
0
-20
-40
-60
-80
-100
1617
1621
+33bp GATA
f
Off-Target Region
Intronic
Intergenic
4
8
c
ATAC-seq: BCL11A locus
hChr2: 60,673,476-60,790,189
Ctrl
AAVS1
g1617
g33
BCL11A
d
TAL1 CUT&RUN: BCL11A locus
BCL11A
e
GATA1 CUT&RUN: BCL11A locus
hChr2: 60,673,476-60,790,189
Ctrl
AAVS1
g1617
g33
IgG
BCL11A
g
Control
Donor 1
% Editing
2.0
1.5
1.0
0.5
0.0
OT1
OT2
OT3
OT4
OT5
OT6
OT7
OT8
OT9
OT10
OT11
OT12
Control
Donor 2
% Editing
h
% GFP
120
90
60
30
0
NT
Vector Control
-115
GATA motif
+33bp GATA
i
% F+ Cells
25
20
15
10
5
0
NT
Vector Control
-115
GATA motif
+33bp GATA
j
HOX gene
Motif
HOXa13
HOXb2
HOXa2
HOXd3
GSX1
k
% F+ Cells
50
40
30
20
10
0
% Editing
77
90.3
19.7
74.5
74
NT
VC
AAVS1
Exon2
Enhancer
Hoxa13
Hoxb2
Hoxa2
Hoxd3
GSX1

Fig 14: Understanding the mechanism of action of the +33bp GATA site in BCL11A regulation. a) Global transcriptomic profile of the gene-edited cells during erythroid differentiation showing variation between the different samples. b) Total number of dysregulated genes among the edited samples analyzed through RNA-seq. c) Evaluation of conformation of the gene post-editing analyzed through ATAC-seq. Evaluation of variation of TAL1 (d) and GATA1 (e) binding at the target locus caused by gene editing. f) Predicted positions of off-target sites for the +33bp GATA identified through the COSMID algorithm. g) Evaluation of unintended editing at the predicted off-target sites compared with the control unedited cells. The evaluation was performed in 2 donors post-erythroid differentiation. h) The transduction efficiency of each gRNA measured in flow cytometry in HUDEP2 cells expressing dead cas9. Mean ± SEM of n=3 independent replicates. i) Intracellular fetal hemoglobin analysis using flow cytometry demonstrates the potential of blocking the target sites using guide RNAs on HbF reactivation at the terminal stage of erythroid differentiation (Day 8); -115 denoted the sgRNA used to block the -115 cluster in *HBG*1/2 promoter. Mean ± SEM of n=3 independent replicates. j) Possible transcription factor binding site matches that resemble the +33bp GATA motif. Data was obtained through the JASPER algorithm. k) Validation of intracellular HbF induction through KO of these transcription factors.

4.2.5 Precise base conversion at the +33bp site avoiding double-strand break to evaluate the impact of KLF1 binding to the target site.

Having proved that cas9-mediated editing at the desired target sites within the enhancer elevated fetal hemoglobin, albeit, through disruption of an activator site, we performed base-editing using both cytosine and adenosine base editors to evaluate precise changes. We observed that the cytosine base editor was more efficient compared to the adenosine base editor which falls in line with the previously reported publications. Thus, we employed the hyper-variant mutant of Adenosine base editor, ABE8e, to evaluate the effect of targeted disruption within these motifs. We included multiple sgRNAs to disturb the proximal KLF1 site through both the base editors **(Fig. 15a)**. We achieved more than 80% editing at the intended sites through all the guide RNAs. Surprisingly, we also observed base conversion outside the CB15 guide RNA region while using the ABE8e variant. We observed that altering the proximal KLF1 motif did not elevate fetal hemoglobin expression **(Fig. 15b)**. The conversion of bases

within the adjacent 'ATTA' (in the complementary strand) elevated the total F^+ cells significantly compared to the control. The base sequence after the conversion of ATTA to ATTG by the CB13 sgRNA resulted in the generation of a de novo binding site of ZBTB7A, a known repressor. Hence, to evaluate the binding of this repressor, we also performed a gel-shift assay and observed a clear interaction of the ZBTB7A with that of the converted base sequence **(Fig. 15c)**. This data suggests that base editing at the +33bp site, results in the de-novo creation of ZBTB7A binding thereby impacting the BCL11A erythroid-specific enhancer function resulting in elevated F+ cells.

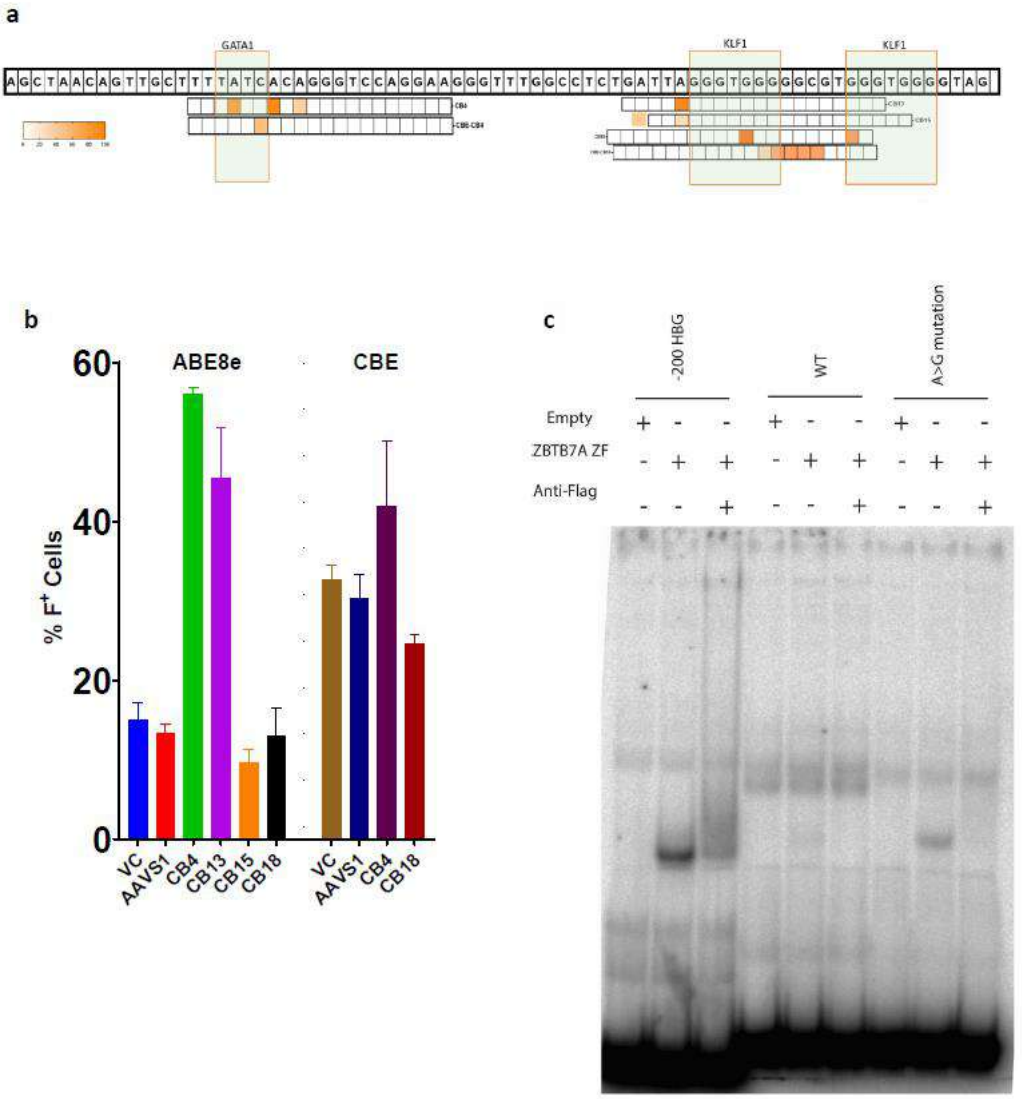


Fig 15: Precise base conversion at the +33bp site avoiding double-strand break to evaluate the impact of KLF1 binding to the target site. a) Position and the editing efficiency of the guide RNAs targeting the proximal KLF1 site in the DHS58 region of the BCL11A enhancer. The mean editing of the target bases is plotted. b) Intracellular fetal hemoglobin analysis using flow cytometry demonstrates the potential of base conversion of target sites using CBE and ABE8e base editor on HbF reactivation at the terminal stage of erythroid differentiation (Day 8). c) Gel-shift assay showing the potential binding of ZBTB7A transcription factor after base conversion at the +33bp site using ABE8e.

DISCUSSION

Section-5

Discussion

5.1 Base editing of key residues in the BCL11A-XL-specific zinc finger domains de-represses fetal globin expression.

In this study, we investigated a novel strategy to specifically alter the DNA-binding domains of BCL11A-XL to functionally mimic naturally occurring mutations affecting the BCL11A-XL isoform found in patients with normal or near normal hematopoiesis. One of the major advantages of this strategy is that it generates very high levels of HbF induction, presumably because it impairs the binding of BCL11A-XL not just to its site at -115 in the fetal *globin* promoters, but also to other sites that may lie elsewhere in the *globin* locus and be required for complete fetal *globin* silencing.(Ghedira et al., 2013) Specifically, modification of ZnF4 led to very high levels of HbF in the adult stage, similar to those seen with the total loss of BCL11A protein.

BCL11A-XL has been demonstrated to interact with the 'TGACCA' site in the *HBG1/2* promoters (-115 element) via primarily the ZnF4 and ZnF5 domains.(Liu et al., 2018; Yang et al., 2019) Several microRNAs and transcriptional regulators, including LIN28B, have been shown to affect the production of XL-isoform by binding to BCL11A exon4.(Smith et al., 2020) Many studies have focused on targeting the BCL11A recognition motif at the *HBG* promoter rather than the endogenous BCL11A-XL ZnF domains to eliminate the BCL11A: *HBG* interaction.

To refine our understanding of ZnF domains in globin regulation, we first targeted the endogenous BCL11A XL-specific ZnF (ZnF4, 5 and 6) domains in adult erythroid

cells using CRISPR/Cas9 mediated editing and indel formation. Gene perturbations at each ZnF4, ZnF5, and ZnF6 domains efficiently generated ZnF truncated mutants that produced elevated levels of HbF expression, comparable with changes seen by targeting other BCL11A-associated regions like the enhancer or the -115 element in the *HBG1/2* promoters. Although ZnF6 has not been observed to directly interact with the DNA binding site, we observed HbF upregulation on ZnF6 disruption, suggesting that it makes an indirect contribution and is required for proper fetal *globin* silencing. The truncated BCL11A-XL mutants in HSPCs were defective in cell survival, and erythroid maturation, and also failed to survive in an *in vivo* mouse bone marrow environment.

To further explore the role of specific subdomains within each ZnF and to attempt to reduce side effects, we used a high-fidelity base editor to alter the key bases, present in XL-specific ZnF domains, that are known to interact with the target DNA strand. Together, with the available database on SNVs, we particularly aimed to target the reported BCL11A SNVs associated with neurodevelopmental disorders, where patients show de-repressed fetal *globin,* but otherwise have normal hematology. Current base editors have limited capability to generate desired heterozygous mutations without any bystander conversion. Nevertheless, we generated different ZnF-specific missense mutants in BCL11A-XL, using an ABE8e, and confirmed the reduced binding of BCL11A-XL to its DNA recognition site in a gel shift assay, demonstrating that specific mutations, with the expected compromised functions, could be made.

Among the ZnF domains, the reactivation of HbF was highest with ZnF4 and ZnF5 mutants, which was equivalent to total *BCL11A* suppression, followed by binding site disruption at the *HBG1/2* promoters, and the enhancer-mediated *BCL11A* downregulation. This emphasizes that base modifications at the ZnF domains are sufficient to produce significant levels of HbF. Importantly, however, we also observed a reduction in engraftment and B cell production. This data is in line with another study that reported the role of ZnF domains in the immunomodulatory activity of B-cells against HIRRV in Flounder fish.(Tang et al., 2022) The exact effect of the ZnF4-1 mutation on the functional activity of B-cells and HSPC engraftment needs to be studied to extend this approach clinically. However, the high efficiency of ABE8e resulted in more than one missense mutation in each of the ZnF domains creating double mutants, unlike the natural observed variants. Nevertheless, the ZnF4 mutant resulted in effective HbF re-induction with near-normal erythroid maturation in vitro and thus generated fewer side effects than CRISPR/Cas9 mediated disruption of the domain or other major perturbations of the BCL11A gene.

Overall, altering the BCL11A ZnF4 domain might disrupt all the potential interactions with the globin cluster leading to effective induction of HbF compared to the previously reported(Han et al., 2022) multiplex editing at both the BCL11A enhancer and binding site. Further, we observed minimal transcriptomic changes in altering the ZnF4 domain compared to total BCL11A loss. The transcriptomic profiling demonstrated the upregulation of globin locus genes other than *HBG*1/2, such as HBBP1 and BGLT3, consistent with reports that the latter genes are also regulated by BCL11A.

A recent study highlights the homology between the BCL11A and BCL11B proteins,(Shen et al., 2022a; Yin, Izadi, Tenglin, Viennet, Zhai, Zheng, Arthanari, Laura M.K. Dassama, et al., 2023) but due to the codon degeneracy, the sequences encoding the two distinct proteins vary. Interestingly, the non-occurrence of the predicted off-target site at BCL11B reconfirmed that the target is specific to BCL11A-XL and does not affect the function of BCL11B. Collectively, at both DNA and RNA levels, the occurrence of negligible levels of off-targets suggests that the observed fetal globin de-repression is solely due to the base alterations at the ZnF4 domain.

To conclude, our approach has shown that one can achieve HbF reactivation through minimal genetic alterations at key bases specifically in the BCL11A-XL with fewer transcriptomic changes. We believe that our approach of altering the DNA recognition domain could further be extended to other transcription factors that use multiple ZnFs to bind different sets of target genes. Our structure/function analysis of ZnF4-6 using base editor approaches provides important insights on the domains of BCL11A-XL required to repress -globin expression and its role in HSPCs. Further, refining our editing strategy would act as a framework for exploring other natural mutations that may selectively reverse the repression of individual genes while preserving their other functions.

5.2 Identification of novel co-activator mediated regulation of DHS58 intronic BCL11A enhancer

BCL11A has long been implicated in fetal *globin* regulation and most of the gene therapy clinical trials for β-hemoglobinopathies modulating BCL11A expression are a testament to this statement. Complete systemic knockout of BCL11A has shown adverse effects through improper lineage differentiation and proliferation bias. The disruption of BCL11A erythroid-specific enhancer and erythroid lineage-specific shRNA mediated downregulation of BCL11A expression results in the de-repression of fetal *globin*. Disrupting the BCL11A binding sites at the -115 site in the *HBG1/2* promoters by both CRISPR/Cas9 mediated disruption and base editing is also currently being pursued clinically. In addition, to these studies involving *BCL11A* at either genomic or transcript level, small molecules-like nanobodies have been developed to selectively reduce BCL11A-XL at the protein level.(Mehta et al., 2022; Shen et al., 2022b; Yin, Izadi, Tenglin, Viennet, Zhai, Zheng, Arthanari, Laura M.K. Dassama, et al., 2023) Even PROTAC-mediated degradation of BCL11A-XL results in efficient downregulation of BCL11A-XL isoform resulting in elevated fetal hemoglobin. The major caveat in these studies is the spatial and temporal window for optimal delivery of cargo is yet to be optimized. Agencies like the FDA and EMA have approved for commercial utilization of EXA-CEL product which is HSPCs exhibiting BCL11A enhancer disruption (GATA motif by 1617) by Vertex pharmaceuticals and CRISPR therapeutics. Most of the studies, even the newer ones reporting novel regulators of fetal hemoglobin inducers, use the BCL11A enhancer disruption as a positive indicator for their study testament to the depth of the data that we have on GATA motif (1617)

disruption as a therapeutic target for beta-hemoglobinopathies. Despite this, the exact mechanism of regulation of a gene through an intronic enhancer present within the gene itself is still elusive. This renders the need to understand erythroid-specific enhancers of therapeutically relevant genes like BCL11A. In our study, we aimed to evaluate the impact of the usually overlooked transcriptional regulator, KLF1, binding site present adjacent to the GATA motif (1617 site) within the DHS58 of erythroid-specific enhancer of BCL11A.

First, to identify the impact of disrupting these sites, we designed 2 sgRNAs and performed cas9-mediated editing in HUDEP2 cells. We observed a significant elevation in fetal hemoglobin in 3 of these guides albeit without the loss of BCL11A protein expression. One of the 3 guides, +33bp GATA, showed an elevation comparable to that of GATA (1617) disruption. +33bp GATA guide RNA exhibited significant disruption of the proximal KLF1 site. Despite the CB11 guide also disrupting the proximal KLF1 site, it showed lower levels of fetal hemoglobin induction. This probably suggests that the 5' extension of the InDel profile played an important role in the regulation. As expected, the CB13 and CB15 sgRNA showed significant yet modest elevation in fetal hemoglobin. Implying the role of the KLF1 site in co-activating the BCL11A enhancer. A significant loss of BCL11A mRNA was observed even though the protein loss was not appreciable. Despite lower BCL11A protein repression, fetal hemoglobin was significantly elevated suggesting an alternate pathway of enhancer regulation.

The +33bp GATA site was evaluated in multiple healthy donor CD34+ HSPCs to verify its fetal hemoglobin induction potential along with other erythroid parameters

in vitro. From these results, we conclude that the fetal hemoglobin induction potential of +33bp GATA was second only to the actual GATA motif disruption despite not disturbing the proximal KLF1 site. The whole transcriptome profile highlighted a lot of variation between the GATA motif disruption and the +33bp GATA disruption. GATA1-Tal1 binding was not altered through the disruption of the +33bp site GATA motif. The open confirmation of the enhancer and the persistent GATA1-TAL1 binding even after +33bp disruption suggests that the +33bp GATA site has a novel regulatory mechanism. Blocking the +33bp GATA site using dcas9-+33bp GATA gRNA resulted in a significant elevation of F+ cells. This suggests that the +33bp GATA site is an activator binding site that probably activating in a GATA-independent manner.

Base-editing at this site highlighted that the proximal KLF1 site is not occupied by the KLF1 transcriptional activator yet +33bp GATA guide RNA created a de novo LRF/ZBTB7A repressor binding site. This binding of ZBTB7A after +33bp GATA motif editing was confirmed by biochemical assay like gel-shift. Proving that it is possible to interlink two previously independent transcriptional regulators and produce an effective and therapeutically relevant phenotype.

Identifying the putative transcription factor that binds at the +33bp GATA site and further imploring the role of KLF1 transcription factor in regulating the BCL11A enhancer function will collectively improve the understanding of our BCL11A enhancer and the various ways of its regulatory mechanisms. Collectively this data suggests that there are still many unknown regulatory site(s) that are yet to be identified within the well-studied clinically approved BCL11A enhancer region.

SUMMARY & CONCLUSION

Section-6

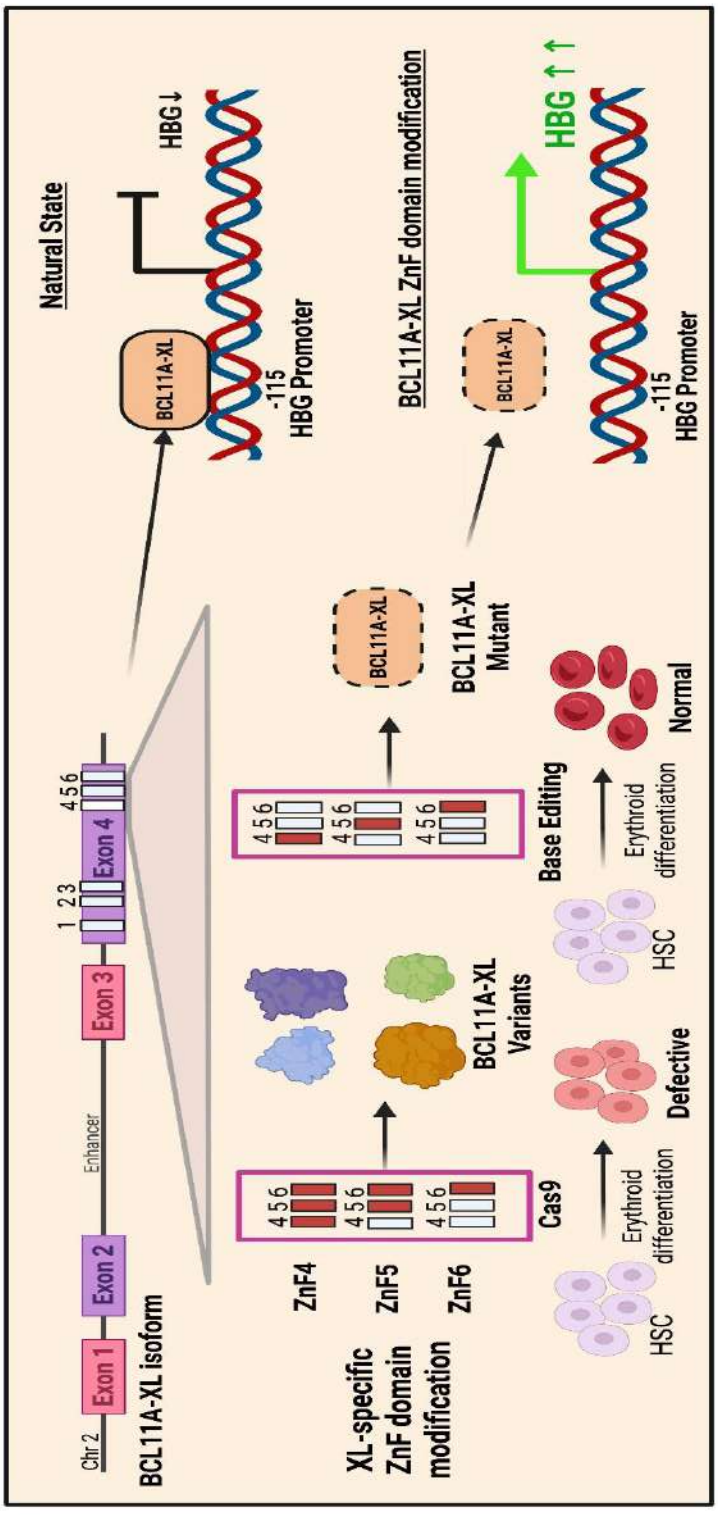
Summary

In the first study, on disrupting the BCL11A-XL-specific ZnF domains, we observed the following key findings.

1. BCL11A-XL-specific targeted disruption using CRISPR/cas9 affects terminal erythropoiesis. Indicating the important role of BCL11A-XL in enucleation.
2. Role of BCL11A-XL in engraftment highlighted.
3. The efficiency of different base editors varies within the BCL11A region. Cytosine and Adenosine base editors are not as efficient as CRISPR/cas9 in regulatory regions like ZnF domains. The hyper-variant of the adenosine base editor, ABE8e efficiently converts bases in this region suggesting that the processivity of the gene editors plays an important role in editing efficiency.
4. Base conversions at the target site elevate fetal hemoglobin at significantly higher levels with almost normal terminal erythropoiesis/enucleation.
5. The ZnF domains present at the C-terminal domains also play a crucial role in engraftment and multi-lineage differentiation in vivo.

This study highlights the previously unreported precise role of ZnF domains present at the C-terminal region specific to the BCL11A-XL isoform in terminal erythropoiesis, engraftment, and multi-lineage differentiation.

Fig 16: BCL11A ZnF disruption study overview



In the second study, elucidating the role of co-activator binding sites within the BCL11A intronic enhancer, we observed the following key findings.

1. Observed that disruption of KLF1 sites present upstream of GATA motif elevated fetal hemoglobin. The role of each site varies, the proximal KLF site had a more prominent effect towards its 5' end whereas disruption of the distal KLF site showed significant elevation.
2. Disrupting these sites did not impact erythropoiesis or engraftment despite elevating fetal hemoglobin through different mechanisms.
3. The GATA-independent role of the +33bp GATA site in regulating the BCL11A expression is validated through ATAC-seq and Cut&Run experiments.
4. Creating a binding site of repressors using base editors is another viable option for downregulating fetal hemoglobin elevation.

This study shed light on the role of other regulatory sites within the BCL11A intronic enhancer which co-regulates the expression of BCL11A along with the reported GATA1-TAL1 axis.

Conclusion

This study on both the BCL11A-XL specific ZnF domains and the BCL11A intronic enhancer has elevated the understanding and role of critical regions within the BCL11A gene.

Despite closely mimicking BCL11A-IDD mutations near the ZnF domains, the innate functions of the gene were affected which is not observed in the patients harboring these mutations. Hence, a replica of these mutations needs to be achieved in highly conservative regulatory regions to study its therapeutic benefits. Yet the use of base editors in elucidating the precise functions of regulatory genes can be implied from this study. This approach of altering the binding domain can be utilized in research areas where a single transcription factor has multiple binding sites and altering each site is not practically feasible.

Understanding the role of co-activators and the other important regulatory elements within the enhancer leads to a precise understanding of its functioning. This study highlights the importance of further probing this enhancer region to understand the finer details of its regulatory mechanisms. This aspect of the study will help in improving our understanding of currently approved gene editing targets improving the safety and better understanding of the SCD and Beta-Thalassemia patient outcomes in the future.

BIBLIOGRAPHY

Section-7

Bibliography

Bagchi A, Nath A, Thamodaran V, et al. (2021) Direct generation of immortalized erythroid progenitor cell lines from peripheral blood mononuclear cells. *Cells* 10(3): 1–18.

Basak A, Hancarova M, Ulirsch JC, et al. (2015) BCL11A deletions result in fetal hemoglobin persistence and neurodevelopmental alterations. *Journal of Clinical Investigation* 125(6). American Society for Clinical Investigation: 2363–2368.

Basak A, Munschauer M, Lareau CA, et al. (2020) Control of human hemoglobin switching by LIN28B-mediated regulation of BCL11A translation. *Nature Genetics* 52(2). Springer US: 138–145.

Bauer DE, Kamran SC, Lessard S, et al. (2013) An erythroid enhancer of BCL11A subject to genetic variation determines fetal hemoglobin level. *Science* 342(6155). American Association for the Advancement of Science: 253–257.

Brendel C, Negre O, Rothe M, et al. (2020) Preclinical Evaluation of a Novel Lentiviral Vector Driving Lineage-Specific BCL11A Knockdown for Sickle Cell Gene Therapy. *Molecular Therapy - Methods and Clinical Development* 17. Cell Press: 589–600.

Canver MC, Smith EC, Sher F, et al. (2015) BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis. *Nature* 527(7577). Nature Publishing Group: 192–197.

Clement K, Rees H, Canver MC, et al. (2019) CRISPResso2 provides accurate and rapid genome editing sequence analysis. *Nature Biotechnology* 37(3). Springer US: 224–226.

Conant D, Hsiau T, Rossi N, et al. (2022) Inference of CRISPR Edits from Sanger Trace Data. *CRISPR Journal* 5(1): 123–130.

Davis R, Gurumurthy A, Hossain MA, et al. (2019) Engineering Globin Gene Expression. *Molecular Therapy - Methods and Clinical Development*. Cell Press.

Demirci S, Zeng J, Wu Y, et al. (2020) BCL11A enhancer–edited hematopoietic stem cells persist in rhesus monkeys without toxicity. *Journal of Clinical Investigation* 130(12). American Society for Clinical Investigation: 6677–6687.

Demirci S, Leonard A, Essawi K, et al. (2021) CRISPR-Cas9 to induce fetal hemoglobin for the treatment of sickle cell disease. *Molecular Therapy - Methods and Clinical Development*. Cell Press.

Devaraju N, Rajendiran V, Ravi NS, et al. (2022) Genome Engineering of Hematopoietic Stem Cells Using CRISPR/Cas9 System. *Methods in Molecular Biology* 2429(Lv): 307–331.

Doerfler PA, Feng R, Li Y, et al. (2021) Activation of γ-globin gene expression by GATA1 and NF-Y in hereditary persistence of fetal hemoglobin. *Nature Genetics* 53(8). Nature Research: 1177–1186.

Feng R, Mayuranathan T, Huang P, et al. (2022) Activation of γ-globin expression by hypoxia-inducible factor 1α. *Nature* 610(7933). Springer US: 783–790.

Ferrari S, Vavassori V, Canarutto D, et al. (2021) Gene Editing of Hematopoietic Stem Cells: Hopes and Hurdles Toward Clinical Translation. *Frontiers in Genome Editing* 3. Frontiers Media SA.

Finotti A, Breda L, Lederer CW, et al. (2015) Recent trends in the gene therapy of β-thalassemia. *Journal of Blood Medicine* 6: 69–85.

Funnell APW, Prontera P, Ottaviani V, et al. (2015) P15-p16.1 microdeletions encompassing and proximal to BCL11A are associated with elevated HbF in addition to neurologic impairment. *Blood* 126(1). American Society of Hematology: 89–93.

Ghedira ES, Lecerf L, Faubert E, et al. (2013) Estimation of the difference in HbF expression due to loss of the 5' δ-globin BCL11A binding region. *Haematologica* 98(2): 305–308.

Han Y, Tan X, Jin T, et al. (2022) CRISPR/Cas9-based multiplex genome editing of BCL11A and HBG efficiently induces fetal hemoglobin expression. *European Journal of Pharmacology* 918. Elsevier: 174788.

Huang P, Peslak SA, Ren R, et al. (2022) HIC2 controls developmental hemoglobin switching by repressing BCL11A transcription. *Nature Genetics* 54(9). Springer US: 1417–1426.

Kluesner MG, Nedveck DA, Lahr WS, et al. (2018) EditR: A Method to Quantify Base Editing from Sanger Sequencing. *The CRISPR Journal* 1(3): 239–250.

Koblan LW, Doman JL, Wilson C, et al. (2018) Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. *Nature Publishing Group* (May). Nature Publishing Group. Epub ahead of print 2018. DOI: 10.1038/nbt.4172.

Liu H, Ippolito GC, Wall JK, et al. (2006) Functional studies of BCLIIA: Characterization of the conserved BCLIIA-XL splice variant and its interaction with BCL6 in nuclear paraspeckles of germinal center B cells. *Molecular Cancer* 5.

Liu N, Hargreaves V V., Zhu Q, et al. (2018) Direct Promoter Repression by BCL11A Controls the Fetal to Adult Hemoglobin Switch. *Cell* 173(2). Elsevier Inc.: 430-442.e17.

Liu P, Keller JR, Ortiz M, et al. (2003) Bcl11a is essential for normal lymphoid development. *Nature Immunology* 4(6): 525–532.

Luc S, Huang J, McEldoon JL, et al. (2016) Bcl11a Deficiency Leads to Hematopoietic Stem Cell Defects with an Aging-like Phenotype. *Cell Reports* 16(12). The Authors: 3181–3194.

Lux CT, Pattabhi S, Berger M, et al. (2019) TALEN-Mediated Gene Editing of HBG in Human Hematopoietic Stem Cells Leads to Therapeutic Fetal Hemoglobin Induction. *Molecular Therapy - Methods and Clinical Development* 12. Cell Press: 175–183.

Martyn GE, Wienert B, Yang L, et al. (2018) Natural regulatory mutations elevate the fetal globin gene via disruption of BCL11A or ZBTB7A binding. *Nature Genetics* 50(4). Springer US: 498–503.

Martyn GE, Wienert B, Kurita R, et al. (2019) *A natural regulatory mutation in the proximal promoter elevates fetal globin expression by creating a de novo GATA1 site*. Available at: http://ashpublications.org/blood/article-pdf/133/8/852/1557834/blood863951.pdf.

Mehta S, Buyanbat A, Kai Y, et al. (2022) Temporal resolution of gene derepression and proteome changes upon PROTAC-mediated degradation of BCL11A protein in erythroid cells. *Cell Chemical Biology* 29(8). The Authors: 1273-1287.e8.

Menzel S, Garner C, Gut I, et al. (2007) A QTL influencing F cell production maps to a gene encoding a zinc-finger protein on chromosome 2p15. *Nature Genetics* 39(10): 1197–1199.

Psatha N, Reik A, Phelps S, et al. (2018) Disruption of the BCL11A Erythroid Enhancer Reactivates Fetal Hemoglobin in Erythroid Cells of Patients with β-Thalassemia Major. *Molecular Therapy - Methods and Clinical Development* 10. Cell Press: 313–326.

Ravi N, Wienert B, Wyman S, et al. (2022) Identification of novel HPFH-like mutations by CRISPR base editing that elevate the expression of fetal hemoglobin. *eLife*. Epub ahead of print 2022. DOI: 10.7554/eLife.65421.

Richter MF, Zhao KT, Eton E, et al. (2020) Phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity. *Nature Biotechnology* 38(7). Springer US: 883–891.

Sankaran VG, Menne TF, Xu J, et al. (2008) Human fetal hemoglobin expression is regulated by the developmental stage-specific repressor BCL11A. *Science* 322(5909): 1839–1842.

Satterwhite E, Sonoki T, Willis TG, et al. (2001) *The BCL11 gene family: involvement of BCL11A in lymphoid malignancies*. Available at: http://cgap.nci.nih.gov/.

Shen F, Zheng G, Setegne M, et al. (2022) A Cell-Permeant Nanobody-Based Degrader That Induces Fetal Hemoglobin. *ACS Central Science* 8(12): 1695–1703.

Shen Y, Verboon JM, Zhang Y, et al. (2021) A unified model of human hemoglobin switching through single-cell genome editing. *Nature Communications* 12(1). Springer US: 1–12.

Shen Y, Li R, Teichert K, et al. (2021) Pathogenic BCL11A variants provide insights into the mechanisms of human fetal hemoglobin silencing. *PLoS Genetics* 17(10). Public Library of Science.

Smith PH, Bessette AJ, Weinberger AH, et al. (2020) Control of human hemoglobin switching by LIN28B-mediated regulation of BCL11A translation. *Nature Genetics* 52(2): 138–145.

Tang X, Sun P, Wang H, et al. (2022) Zinc Finger Protein BCL11A Contributes to the Abortive Infection of Hirame novirhabdovirus (HIRRV) in B Lymphocytes of Flounder (Paralichthys olivaceus) . *Journal of Virology* 96(24). American Society for Microbiology: 1–19.

Tsang JCH, Yu Y, Burke S, et al. (2015) Single-cell transcriptomic reconstruction reveals cell cycle and multi-lineage differentiation defects in Bcl11a-deficient hematopoietic stem cells. *Genome Biology* 16(1). Genome Biology: 1–16.

Uda M, Galanello R, Sanna S, et al. (2008) *Genome-wide association study shows BCL11A associated with persistent fetal hemoglobin and amelioration of the phenotype of-thalassemia*. Available at: www.pnas.org/cgi/content/full/.

Venkatesan V, Srinivasan S, Babu P, et al. (2021) Manipulation of Developmental Gamma-Globin Gene Expression: an Approach for Healing Hemoglobinopathies. *Molecular and Cellular Biology* 41(1): 1–18.

Wang L, Li L, Ma Y, et al. (2020) Reactivation of γ-globin expression through Cas9 or base editor to treat β-hemoglobinopathies. *Cell Research*. Springer Nature.

Wienert B, Martyn GE, Funnell APW, et al. (2018) Wake-up Sleepy Gene: Reactivating Fetal Globin for β-Hemoglobinopathies. *Trends in Genetics* 34(12). Elsevier Ltd: 927–940.

Wu Y, Zeng J, Roscoe BP, et al. (2019) Highly efficient therapeutic gene editing of human hematopoietic stem cells. *Nature Medicine* 25(5). Nature Publishing Group: 776–783.

Xu J, Sankaran VG, Ni M, et al. (2010) Transcriptional silencing of γ-globin by BCL11A involves long-range interactions and cooperation with SOX6. *Genes and Development* 24(8): 783–789.

Yang Y, Xu Z, He C, et al. (2019) Structural insights into the recognition of γ-globin gene promoter by BCL11A. *Cell Research* 29(11): 960–963.

Yin J, Xie X, Ye Y, et al. (2019) BCL11A: A potential diagnostic biomarker and therapeutic target in human diseases. *Bioscience Reports*. Portland Press Ltd.

Yin M, Izadi M, Tenglin K, Viennet T, Zhai L, Zheng G, Arthanari H, Dassama Laura M. K., et al. (2023) Evolution of nanobodies specific for BCL11A. *Proceedings of the National Academy of Sciences* 120(3): 2017.

Yu Y, Wang J, Khaled W, et al. (2012) Bcl11a is essential for lymphoid development and negatively regulates p53. *Journal of Experimental Medicine* 209(13): 2467–2483.

Zittersteijn HA, Harteveld CL, Klaver-Flores S, et al. (2020) A Small Key for a Heavy Door: Genetic Therapies for the Treatment of Hemoglobinopathies. *Frontiers in Genome Editing* 2(February).

LIST OF PUBLICATIONS

Section-8

List of Publications

1. Rajendiran V, Devaraju N, Haddad M, et al. Base editing of key residues in the BCL11A-XL-specific zinc finger domains derepresses fetal globin expression.*Mol Ther*.2024;32(3):663-677. doi:10.1016/j.ymthe.2024.01.023.
2. Devaraju N, Rajendiran V, Ravi NS, Mohankumar KM. Genome Engineering of Hematopoietic Stem Cells Using CRISPR/Cas9 System. Methods Mol Biol. 2022;2429:307-331. doi:10.1007/978-1-0716-1979-7_20

ANNEXURES

Section-8

Publication from Thesis

Base editing of key residues in the BCL11A-XL-specific zinc finger domains de-represses fetal globin expression **Molecular Therapy** (2024) 10.1016/j.ymthe.2024.01.023 Epub ahead of print.

Genome Engineering of Hematopoietic Stem Cells Using CRISPR/Cas9 System. ***Methods in Molecular Biology*** *(2022)* Book-chapter: 10.1007/978-1-0716-1979-7_20

Curriculum Vitae

Vignesh Rajendiran

20, Vivekanandhar street,Thirunagar,
Vellore -632 006, Tamil Nadu, India.
DOB: 9th Dec, 1992.
Orcid ID: 0000-0002-5037-7020

E-mail: rvix92@gmail.com
Mobile: + 91 75984 84051

Academic Profile

Level	Institution/Board/University	Year	Percentage
M.Tech Industrial Biotechnology	SASTRA University, Thanjavur, Tamil Nadu, India	2015-2017	8.05
B.Tech Biotechnology	Bharathidasan University, Tiruchirapalli, Tamil Nadu, India	2011-2015	8.03

Exams Cleared:

1. **GRE-** 309/340 (2015)(159Q, 150V,3/6W)
2. **IELTS-** 7.5/9 (2015)
3. **Graduate Aptitude Test in Engineering**
    a. **2016 -** 94 percentile with All India Rank 660
    b. **2017 -** 95.7 percentile with All India Rank 336
4. **Department of Biotechnology-Junior Research Fellowship**-2018.

Languages Known

- English (Read, Write & Speak)
- Tamil (Read, Write & Speak)

Research Experience

1. **Tenure** : August 2017-**Present [PhD]**

   Host Lab: Dr. Mohankumar K Murugesan
   Scientist E,
   Center for Stem Cell Research (a unit of inStem, Bengaluru)
   Christian Medical College, Vellore, Tamil Nadu, India.

   Topic : Pre-clinical study for Therapeutic Genome Editing in β-hemoglobinopathies

2. **Tenure** : June 2016-April 2017 **[Master Dissertation]**

   Host Lab: Dr. K Umamaheshwari

   Associate Dean,

   Center for Nanotechnology and Advanced Biomaterials

   School of Chemical and Biotechnology

   SASTRA University, Tamil Nadu, India.

   Topic : Effect of Rutin in Differentiated IMR32 Cell Line

3. **Tenure** : December, 2014-May, 2015 **[Bachelor Dissertation]**

   Host Lab: Dr.Dipshikha Chakravortty, Ph.D

   Associate Professor

   Department of Microbiology and Cell Biology

   Center for Infectious disease

   Indian Institute of Science, Bangalore, India.

   Topic : Characterization of rfaG mutant of *Salmonella typhimurium*

4. **Tenure** : May-July 2014 **[Indian Academy of Science- Summer Research Fellowship)**

   Host Lab: Dr. Ravinder Goswami,

   Dept. of Endocrinology and Metabolism,

   All India Institute of Medical Sciences, New Delhi, India.

   Topic : Standardisation of PCR conditions for Fetuin-A

Research Skills

1. **Molecular Biology Techniques:** Cloning, Genomic DNA, RNA and Protein isolation, Plasmid DNA extraction, PCR, Sanger and NGS sequencing, RT-PCR and HPLC and Western Blotting.
2. **Cell Culture:** HEK293T, K562, HUDEP2, Cos7, HeLa cell line culturing, Basic Microscopy, Lentivirus production, Healthy and Patient Blood sample processing, CD34 isolation from PBMNCs, primary Hematopoietic Stem Cells culturing and erythroid differentiation.

3. **Gene Editing:** Gene knockout, knockdown and base alteration using crispr/cas9, base editor and HDR system, arrayed screening of multiple guide RNAs, Lonza (2b & 4D) nucleofector, Neon Transfection system and Maxcyte Electroporator.
4. **Animal Handling:** NSG and NBSGW mice infusion and complete engraftment studies.

Awards and Fellowships

- Cleared DBT-JRF national level exam in 2018 securing funding for the entire PhD tenure.
- Awarded **Summer Research Fellowship** under the guidance of Prof. Ravinder Goswami, Associate Professor, **All India Institute of Medical Sciences**, New Delhi by the Indian Academy of Sciences, Bangalore.

Achievements:

- Provisionally filed patent 'A METHOD FOR ELEVATION OF GAMMA GLOBIN' Patent No: 202241055876 at Indian Patent Office (IPO).
- PhD work selected as Late-breaking abstract and presented as poster in International Symposium in Stem Cell Research (ISSCR) –June 2023.
- Poster presentation at CRISPR-2023 conference in Wurzburg, Germany in June 2023
- Oral presentation at Genome Writers Guild- in July 2023.
- Presented poster in Annual research day at Christian Medical College (Vellore) and InStem (Bengaluru) India.
- Multiple poster presentations at Cell and Gene Therapy Symposium at Center for Stem Cell Research.

Publications:

1. Base editing of key residues in the BCL11A-XL-specific zinc finger domains de-represses fetal globin expression **Molecular Therapy** (2024) 10.1016/j.ymthe.2024.01.023 Epub ahead of print.
2. Correlating the differences in the receptor binding domain of SARS-CoV-2 spike variants on their interactions with human ACE2 receptor. **Scientific reports** (2023)10.1038/s41598-023-35070-2
3. Efficient and error-free correction of sickle mutation in human erythroid cells using prime editor-2. ***Frontiers in Genome Editing*** *(2022)*10.3389/fgeed.2022.1085111
4. Erythroid lineage-specific lentiviral RNAi vectors suitable for molecular functional studies and therapeutic applications. ***Scientific Reports*** *(2022)* 10.1038/s41598-022-13783-0
5. Identification of novel HPFH-like mutations by CRISPR base editing that elevate the expression of fetal hemoglobin. ***eLife*** *(2022)* 10.7554/elife.65421
6. Preferential Expansion of Human CD34+CD133+CD90+ Hematopoietic Stem Cells Enhances Gene-Modified Cell Frequency for Gene Therapy. ***Human Gene Therapy*** *(2022).* 10.1089/hum.2021.089
7. Genome Engineering of Hematopoietic Stem Cells Using CRISPR/Cas9 System. ***Methods in Molecular Biology*** *(2022)* Book-chapter: 10.1007/978-1-0716-1979-7_20
8. Direct Generation of Immortalized Erythroid Progenitor Cell Lines from Peripheral Blood Mononuclear Cells. **Cells** (2021) 10.3390/cells10030523
9. Amelioration of oxidative stress in differentiated neuronal cells by rutin regulated by a concentration switch. ***Biomedicine & Pharmacotherapy*** *(2018)* 10.1016/j.biopha.2018.09.021

Extracurricular Activity

- Member of Bharat Scouts program during school.
- Active member of the University Soccer team during my Undergraduation.
- Won multiple awards for athletics in interschool level and inter-department level.

Declaration:

I hereby declare that the above statements are true and correct to the best of my knowledge.

Yours Sincerely,

Vignesh Rajendiran

APPENDIX A

ETHICS COMMITTEE APPROVAL

Approval Documents: Study 1

Base editing of key residues in the BCL11A-XL-specific zinc finger domains de-represses fetal globin expression

OFFICE OF RESEARCH
INSTITUTIONAL REVIEW BOARD (IRB)
CHRISTIAN MEDICAL COLLEGE, VELLORE, INDIA

CDSCO – Ethics Committee Registration No: ECR/326/INST/TN/2013/RR-2019, DHR Provisional Registration No: EC/NEW/INST/2020/818

Dr. George Thomas, D. Ortho., Ph.D.,
Chairperson, Ethics Committee

Dr. Prasanna Samuel, M.Sc., Ph.D.,
Secretary, Research Committee

Prof. Keith Gomez, MA (S.W), M.Phil.,
Deputy Chairperson, Ethics Committee

Dr. Biju George, MD., DM.,
Chairperson, Research Committee

Dr. Suceena Alexander, MD., DM., FASN.,
Secretary, Ethics Committee, IRB
Additional Vice-Principal (Research)

September 09, 2022

Dr. Mohankumar K. Murugesan,
Scientist,
Centre for Stem Cell Research,
Christian Medical College,
Vellore – 632 002.

Sub: External Research Grant: New Proposal: ICMR
Targeted disruption on binding domain in the BCL11A for therapeutic reactivation of fetal haemoglohin in beta haemoglobinopathies.
Dr. Mohankumar K. Murugesan, Scientist E, Centre for Stem Cell Research, Dr. Saravanabhavan, Dr. Srujan Kumar Marepally, Dr. RV Shaji , Dr. Alok Srivastava, CSCR.

Ref: IRB Min. No. 14628 [OTHER] dated 27.04.2022

Dear Dr. Mohankumar K. Murugesan,

The Institutional Review Board (Silver, Research and Ethics Committee) of the Christian Medical College, Vellore, reviewed and discussed your project titled "Targeted disruption on binding domain in the BCL11A for therapeutic reactivation of fetal haemoglohin in beta haemoglobinopathies" to define recto-sigmoid junction" to define recto-sigmoid junction" on April 27, 2022.

The Committee reviewed the following documents:

1) IRB Application Format
2) Patient information sheet and consent form (Tamil, English, Malayalam, Hindi, Odiya, Telugu, Bengali)
3) Cvs. of Dr Rini Bandyopadhyay, Dr. Abi Manesh, Dr. Rajiv Karthik, Dr. Dan Barnabas Inja, Dr. Sandeep Albert, Dr. Jeremy Bilss, Dr. Anil Thomas Ommen, Dr. V J Chandy, Dr. V V Ramu, Dr. Viju Daniel Varghese, Dr. Thomas Mathai, Dr. Thilak Jepegnanam, Dr. Samuel C Raj, Dr. Anil Mathew, Dr. Venkatesh K, Dr. Justin Arockiaraj, Dr. Divya Elizabeth Muliyil, Dr. Prasanna Samuel Premkumar.
4) No. of Documents 1 – 3.

Ethics Committee Silver, Office of Research, I Floor, Carman Block, Christian Medical College, Vellore, Tamil Nadu 632 002
Tel: 0416 – 2284294, 2284508 Fax: 0416 – 2262788 E-mail: research@cmcvellore.ac.in

ETHICS COMMITTEE MEMBERS			
Name	Qualification	Designation	Affiliation
Dr. George Thomas	D. Ortho, PhD	Orthopaedic Surgeon, St. Isabella Hospital, Chennai, Chairperson, Ethics Committee, IRB, Chennai	External, Clinician
Dr. Suceena Alexander	MD, DM, FASN.	Professor, Nephrology, Additional Vice-Principal (Research), Member Secretary (Ethics Committee), IRB, CMC, Vellore.	Internal, Clinician
Rev. Joseph Devaraj	BSc, BD	Chaplaincy Department, CMC, Vellore	Internal, Social Scientist
Prof. Keith Gomez	MA (S.W), M. Phil (Psychiatry Social Work)	Deputy Chairperson, Ethics Committee, IRB Student counselor, Loyola College, Chennai.	External, Lay Person & Social Scientist
Mr. C. Sampath	BSc, BL	Advocate, Vellore	External, Legal Expert
Mr. Samuel Abraham	MA, PGDBA, PGDPM, M. Phil, BL.	Sr. Legal Officer, Vellore	External Legal Expert
Mrs. Ilavarasi Jesudoss	M Sc (Nursing)	Deputy Nursing Superintendent, College of Nursing, CMC, Vellore	Internal, Nurse
Dr. Sujith J. Chandy	MD., PhD., FRCP (E)	Professor, Clinical Pharmacology, CMC, Vellore	Internal, Pharmacologist
Dr. Jayaprakash Muliyil	MD, MPH, Dr PH (Epid), DMHC	Retired Professor, CMC, Vellore	External, Scientist & Epidemiologist
RESEARCH COMMITTEE MEMBERS			
Dr. Biju George	MD, DM	Chairperson (Research Committee), Professor, Haematology, CMC, Vellore.	Internal, Clinician

Ethics Committee Silver, Office of Research, I Floor, Carman Block, Christian Medical College, Vellore, Tamil Nadu 632 002
Tel: 0416 – 2284294, 2284508 Fax: 0416 – 2262788 E-mail: research@cmcvellore.ac.in

Dr. Santhanam Sridhar	DCH, DNB	Professor, Neonatology, CMC, Vellore	Internal, Clinician
Dr. Prasanna Samuel	M. Sc, PhD	Secretary (Research Committee) Associate Professor, Biostatistics, CMC, Vellore	Internal, Statistician
Dr. Sathya Subramani	MD, PhD	Professor, Physiology, CMC, Vellore	Internal, Clinician
Dr. D. J. Christopher	DTCD DNB, FRCP(Glasg), FCCP(USA)	Professor, Pulmonary Medicine, CMC, Vellore	Internal, Clinician
Dr. Nihal Thomas	MBBS, MD MNAMS DNB (Endo) FRACP (Endo) FRCP (Edin) FRCP (Glas) FRCP (London) FACP PhD (Copenhagen)	Professor & Head Department of Endocrinology, Diabetes and Metabolism	Internal, Clinician
Dr. Rohin Mittal	MS, DNB	Professor, Department of General Surgery, CMC Vellore	Internal Clinician
Dr. Joe Varughese	MBBS, MD	Professor, Biochemistry, CMC	Internal, Clinician
Dr. Christhunesa S. Christudass	M.Sc., P.hd.,	Professor, Neurochemistry, CMC	Internal, Basic Medical Scientist
Dr. Sridhar Gibikote	DMRD, DNB	Professor, Radiology, CMC, Vellore	Internal, Clinician
Dr. Winsely Rose	MBBS, MD (Paed)	Professor, Paediatrics, CMC Vellore	Internal, Clinician

We approve the project to be conducted as presented for the duration 03 Years.

Kindly provide the total number of patients enrolled in your study and the total number of Withdrawals for the study entitled: "Targeted disruption on binding domain in the BCL11A for therapeutic reactivation of fetal haemoglohin in beta haemoglobinopathies" on a monthly basis. Please send copies of this to the Research Office (research@cmcvellore.ac.in).

Ethics Committee Silver, Office of Research, I Floor, Carman Block, Christian Medical College, Vellore, Tamil Nadu 632 002
Tel: 0416 – 2284294, 2284508 Fax: 0416 – 2262788 E-mail: research@cmcvellore.ac.in

The Institutional Ethics Committee expects to be informed about the progress of the project, Any adverse events occurring in the course of the project, any amendments in the protocol and the patient information / informed consent. On completion of the study, you are expected to submit a copy of the final report. Respective forms can be downloaded from the following link: http://172.16.11.136/Research/IRB_Polices.html in the CMC Intranet and in the CMC website link address: http://www.cmch-vellore.edu/static/research/Index.html.

Yours sincerely,

Dr. Suceena Alexander
Secretary (Ethics Committee)
Institutional Review Board

Dr. Suceena Alexander, MD.,DM.,FASN.
Secretary - (Ethics Committee)
Institutional Review Board
Christian Medical College,
Vellore - 632 002, Tamil Nadu, India.

IRB Min. No. 14628 [OTHER] dated 27.04.2022

Ethics Committee Silver, Office of Research, I Floor, Carman Block, Christian Medical College, Vellore, Tamil Nadu 632 002
Tel: 0416 – 2284294, 2284508 Fax: 0416 – 2262788 E-mail: research@cmcvellore.ac.in

INSTITUTIONAL ANIMAL ETHICS COMMITTEE
CHRISTIAN MEDICAL COLLEGE, VELLORE

Dr. Solomon Sathishkumar
Chairperson

Dr. Joe Varghese
Member Secretary

Date:15.10.2022

Dr. Mohankumar Murugesan,
Scientist,
Centre for Stem Cell Research,
CMC, Vellore.

Sub: Project Proposal for IAEC approval.: "Targeted disruption of binding domain in the BCL11A for therapeutic reactivation of fetal hemoglobin in beta hemoglobinopathies."

Dear Dr. Mohankumar,

Your application to the Institutional Animal Ethics Committee (IAEC) titled "***Targeted disruption of binding domain in the BCL11A for therapeutic reactivation of fetal hemoglobin in beta hemoglobinopathies.***" has been reviewed by the IAEC at the meeting held on 15/10/2022.

The following members of the IAEC were present at the meeting:

1. Dr. Solomon Sathishkumar. (Chairperson, IAEC)
2. Dr. P. Kumarasamy, CPCSEA Main Nominee (External member)
3. Dr. R. Ananda Raja, CPCSEA Link Nominee (External member)
4. Dr. P. Patric Joshua, Scientist from outside the institute (External member)
5. Dr. R. Prakash, Socially aware member (External member)
6. Dr. Christhunesa Soundararajan C, Scientist from different discipline (Internal member)
7. Dr. Eunice Sindhuvi, Scientist from different biological discipline (Internal member)
8. Dr. K. Imayarasi, Veterinarian (Internal member)
9. Dr. Joe Varghese, Member Secretary IAEC, CMC, Vellore (Internal member)

After discussion, the project proposal was approved. One hundred and fifty-two (152) female NBSGW mice have been sanctioned for 24 months. Seventy-two (72) mice have been allotted for the first year of the study.

The IAEC approval number for the study is **11/2022**

Please see the overleaf for the responsibilities of the principal investigator.

With best wishes,

Yours sincerely,

Dr. Solomon Sathishkumar
Chairperson

CHAIRPERSON
Institutional Animal Ethics Committee,
Christian Medical College,
Bagayam, Vellore - 632 002, Tamil Nadu.

Dr. Joe Varghese
Member Secretary

SECRETARY
Institutional Animal Ethics Committee,
Christian Medical College,
Bagayam, Vellore - 632 002, Tamil Nadu.

Dr. Solomon Sathishkumar
Chairperson

Dr. Joe Varghese
Member Secretary

Responsibilities of the Principal Investigator:

1. Progress Report / Project completion report should be submitted for consideration of IAEC at prescribed intervals for review and it should not exceed the timeline of the research as mentioned in the Form B / IAEC approval certificate
2. Final report should be submitted to the IAEC of the establishment at the end of study.
3. All Serious Adverse Events (SAEs) and the interventions undertaken should be intimated to IAEC.
4. Protocol deviation, if any, should be informed with adequate justifications to IAEC and in case of large animals, it should be intimated to CPCSEA immediately for consideration of CPCSEA.
5. The procedural deviations in research protocol shall be treated as a fresh research protocol by CPCSEA.
6. Any new information related to the study should be communicated to IAEC immediately.
7. Premature termination of study should be notified to the IAEC with reasons along with summary of the data obtained so far.
8. Change of investigators/sites should be informed and approval of IAEC should be undertaken first.

Dr. Solomon Sathishkumar
Chairperson

Dr. Joe Varghese
Member Secretary

OFFICE OF THE VICE PRINCIPAL (RESEARCH)
INSTITUTIONAL BIO-SAFETY COMMITTEE (IBSC)
CHRISTIAN MEDICAL COLLEGE, VELLORE - 632 002, INDIA

31st March, 2023

To,

Dr. Mohankumar K. Murugesan,
Scientist E,
Centre for Stem Cell Research
CMC

Dear Dr. Mohankumar K. Murugesan,

The Institutional Bio Safety Committee of the Christian Medical College, Vellore, reviewed and discussed your project entitled *"Erythroid-lineage specific de-functionalization of BCL11A for therapeutic induction of fetal hemoglobin in Thalassemia"*.

IRB Min. No: 14628 dated: 27.04.2022

The following IBSC members were present at the meeting held on January 28, 2023 at 9.00 am in the BRTC Room, College Campus Bagayam, Vellore-632002.

S.No	Name	Designation	IBSC Affiliation
1	Dr. Solomon SathishKumar	Principal	Chairperson
2	Dr. G. Dhinakar Raj	Professor and Head, Department of Animal Biotechnology Madras Veterinary College	DBT Nominee
3	Dr. Suceena Alexander	Professor & Addl. Vice Principal (Research), CMC.	Member Secretary
4	Dr. K A Balasubramanian	Emeritus Professor, Sri Narayani Hospital and Research Centre	External Expert
5	Dr. Joy S Michael	Professor, Microbiology, CMC.	Biosafety Officer
6	Dr. Alok Srivastava	Professor & Head, CSCR, CMC.	Internal Expert
7	Dr. Molly Jacob	Professor, Department of Biochemistry, CMC.	Internal Expert

IBSC – Members

Dr. Solomon Sathishkumar
Chairman

Dr G. Dhinakar Raj
DBT Nominee

Dr. Suceena Alexander
Member Secretary

Dr. K. A. Balasubramaniam
Outside expert

Dr. Joy S. Michael
Biosafety Officer

Dr. Alok Srivastava
Internal Expert

Dr. Molly Jacob
Internal Expert

Dr. Rajesh Kannangai
Internal Expert

Dr. R V Shaji
Internal Expert

Dr. Sitara Swarna Rao Ajjampur
Internal Expert

IBSC Decision: Dr. Mohankumar K. Murugesan was present during the presentation and responded to the queries raised by the Members. After discussion, it was resolved to provide IBSC approval.

Yours sincerely,

SECRETARY
Institutional Bio-Safety Committee
Christian Medical College,
Vellore - 632 002, Tamil Nadu, India.

Dr. Suceena Alexander
Secretary, Institutional Bio Safety Committee
Christian Medical College, Vellore

Approval Documents: Study 2

Identification of novel co-activator mediated regulation of DHS58 intronic BCL11A enhancer

OFFICE OF RESEARCH
INSTITUTIONAL REVIEW BOARD (IRB)
CHRISTIAN MEDICAL COLLEGE, VELLORE, INDIA

Ethics Committee Registration No: ECR/326/INST/TN/2013 issued under Rule 122D of the Drugs & Cosmetics Rules 1945, Govt. of India

Dr. George Thomas, D Ortho Ph.D.
Chairperson, Ethics Committee

Dr. B. Antonisamy, M.Sc., Ph.D., FSMS, FRSS.
Secretary, Research Committee

Prof. Keith Gomez, B.Sc., MA (S.W), M.Phil.
Deputy Chairperson, Ethics Committee

Dr. Alfred Job Daniel, D Ortho MS Ortho DNB Ortho.
Chairperson, Research Committee & Principal

Dr. Nihal Thomas,
MD, MNAMS, DNB (Endo), FRACP (Endo), FRCP (Edin), FRCP (Glasg)
Deputy Chairperson,
Secretary, Ethics Committee, IRB
Additional Vice-Principal (Research)

January 04, 2015

Dr. R. V. Shaji,
Professor / Adjunct Scientist
Department of Haematology,
Christian Medical College,
Vellore – 632 002.

Sub: **External Research Grant NEW PROPOSAL:**
Department of Biotechnology (DBT)
Preclinical model for gene therapy for Thalassemia and Sickle Cell Disease (SCD
Dr. R. V. Shaji, Employment Number: 30127, Haematology and Adjunct Scientist,
CSCR, Dr. Alok Srivastava, CSCR / Hematology, Dr. Mohankumar Murugesan CSCR ,
Dr. Saravanabhavan Thangavel CSCR

Ref: IRB Min No: 9644 [OTHER] dated 23.09.2015

Dear Dr. R. V. Shaji,
The Institutional Review Board (Silver, Research and Ethics Committee) of the Christian Medical College, Vellore, reviewed and discussed your project titled "Preclinical model for gene therapy for Thalassemia and Sickle Cell Disease (SCD" on September 23rd 2015.

The Committee reviewed the following documents:

1. IRB Application format
2. DBT Proforma
3. Information Sheet and Informed Consent Form (English)
4. Cvs of Drs. . R. V. Shaji, Alok Srivastava, . Mohankumar Murugesan
5. No. of documents 1 – 4

The following Institutional Review Board (Silver, Research & Ethics Committee) members were present at the meeting held on September 23rd 2015 in the CREST/SACN Conference Room,Christian Medical College, Bagayam, Vellore 632002.

Name	Qualification	Designation	Affiliation
Dr. B. Antonisamy	MSc, PhD, FSMS, FRSS	Professor, Biostatistics, Secretary (Research Committee), IRB, CMC, Vellore	Internal, Statistician

Ethics Committee Silver, Office of Research, I Floor, Carman Block, Christian Medical College, Vellore, Tamil Nadu 632 002
Tel: 0416 – 2284294, 2284202 Fax: 0416 – 2262788 E-mail: research@cmcvellore.ac.in

Dr. George Thomas, D Ortho Ph.D.
Chairperson, Ethics Committee

Dr. B. Antonisamy, M.Sc., Ph.D., FSMS, FRSS.
Secretary, Research Committee

Prof. Keith Gomez, B.Sc., MA (S.W), M.Phil.
Deputy Chairperson, Ethics Committee

Dr. Alfred Job Daniel, D Ortho MS Ortho DNB Ortho.
Chairperson, Research Committee & Principal

Dr. Nihal Thomas,
MD, MNAMS, DNB (Endo), FRACP (Endo), FRCP (Edin), FRCP (Glasg)
Deputy Chairperson,
Secretary, Ethics Committee, IRB
Additional Vice-Principal (Research)

Dr. Prasanna Samuel	MSc, PhD	Lecturer, Biostatistics, CMC, Vellore	Internal, Statistician
Dr. B. Poonkuzhali	MSC, PhD	Professor, Haematology, CMC, Vellore	Internal, Basic Medical Scientist
Dr. Vinod Joseph Abraham	MBBS, MD, MPH	Professor, Community Medicine, CMC, Vellore	Internal, Clinician
Dr. Biju George	MBBS, MD, DM	Professor, Haematology, CMC, Vellore	Internal, Clinician
Dr. Deepak Abraham	MBBS, MS	Professor, Endocrine Surgery, CMC, Vellore	Internal, Clinician
Dr. Sathya Subramani	MD, PhD	Professor, Physiology, CMC, Vellore	Internal, Clinician
Dr. George Thomas	MBBS, D Ortho, PhD	Orthopaedic Surgeon, St. Isabella Hospital, Chennai, Chairperson, Ethics Committee, IRB.	External, Clinician
Prof. Keith Gomez	BSc, MA (S.W), M. Phil (Psychiatry Social Work)	Student counselor, Loyola College, Chennai, Deputy Chairperson, Ethics Committee, IRB	External, Lay Person & Social Scientist
Dr. P. Zachariah	MBBS, PhD	Retired Professor, Vellore	External, Clinician
Mrs. Pattabiraman	BSc, DSSA	Social Worker, Vellore	External, Lay person
Dr. Shirley David	MSc, PhD	Professor, Head of Fundamentals Nursing Department, College of Nursing, CMC, Vellore	Internal, Nurse
Mr. Samuel Abraham	MA, PGDBA, PGDPM, M. Phil, BL.	Sr. Legal Officer, CMC, Vellore	Internal, Legal Expert
Dr. Vinitha Ravindran	PhD (Nursing)	Professor & Addl. Deputy Dean, College of Nursing, CMC, Velore	Internal, Nurse
Mr. C. Sampath	BSc, BL	Advocate, Vellore	External, Legal Expert
Rev. Dr. T. Arul Dhas	MSc, BD, DPC, PhD(Edin)	Chaplaincy Department, CMC, Vellore	Internal, Social Scientist

Ethics Committee Silver, Office of Research, I Floor, Carman Block, Christian Medical College, Vellore, Tamil Nadu 632 002
Tel: 0416 – 2284294, 2284202 Fax: 0416 – 2262788 E-mail: research@cmcvellore.ac.in

Dr. Binu Susan Mathew	MBBS, MD	Associate Professor, Clinical Pharmacology, CMC, Vellore	Internal, Pharmacologist
Mrs. Ruma Nayak	M Sc (Nursing)	Professor, Head of Paediatric Nursing & Deputy Nursing Superintendent, College of Nursing, CMC, Vellore	Internal, Nurse
Dr. Nihal Thomas	MD, MNAMS, DNB(Endo), FRACP (Endo) FRCP(Edin) FRCP (Glasg)	Professor & Head, Endocrinology, Additional Vice Principal (Research), Deputy Chairperson (Research Committee), Member Secretary (Ethics Committee), IRB, CMC, Vellore	Internal, Clinician

We approve the project to be conducted as presented.

Administrative Committee's approval is to be obtained for opening the account-head, employing any personnel or purchasing any equipment. The investigator also needs to present to Administrative Committee, the terms and condition of the Funding agency for approval.

Kindly provide the total number of patients enrolled in your study and the total number of withdrawals for the study entitled: "Preclinical model for gene therapy for Thalassemia and Sickle Cell Disease (SCD " on a monthly basis. Please send copies of this to the Research Office (research@cmcvellore.ac.in)

Yours sincerely

Dr. Nihal Thomas
Secretary (Ethics Committee)
Institutional Review Board
Christian Medical College, Vellore

Dr. NIHAL THOMAS
MD.,MNAMS.,DNB(Endo),FRACP(Endo),FRCP(Edin),FRCP(Glasg)
SECRETARY - (ETHICS COMMITTEE)
Institutional Review Board,
Christian Medical College, Vellore - 632 002.

IRB Min No: 9644 [OTHER] dated 23.09.2015

Ethics Committee Silver, Office of Research, I Floor, Carman Block, Christian Medical College, Vellore, Tamil Nadu 632 002
Tel: 0416 – 2284294, 2284202 Fax: 0416 – 2262788 E-mail: research@cmcvellore.ac.in

No. BT/PR17316/MED/31/326/2015
(e-file No. RAD-28/35/2020-MED-DBT)
Government of India
Ministry of Science & Technology
Department of Biotechnology

Block -2&3, 6-8 Floor,
CGO Complex, Lodi Road,
New Delhi 110003

Dated: 04.02.2021

EXTENSION ORDER

In continuation of this Department's Sanction Order of even number dated 30.12.2015, sanction of the President is hereby accorded to the extension of the project **"Accelerating the application of Stem cell technology in Human Disease"** including the two components; "Accelerator program for discovery in brain disorders using stem cells (ADBS)", and "Novel approaches to haematological disorders (NAHD)", as a collaborative project with four partnering institutes being implemented by Project Coordinators Prof. Mahendra Rao & Prof. Apurva Sarin, inStem, Bangalore; Prof. Raghu Padinjat, NCBS, Bangalore; Prof. Sanjeev Jain, NIMHANS, Bangalore; and Dr. Alok Srivastava, CSCR, CMC, Vellore, for a period of **1 Year 6 months** from **30.12.2020** to **30.06.2022** within the overall sanctioned budget.

2. The other terms and conditions of the project will remain unchanged.

3. This issues with the approval of Competent Authority.

(Karthikeya K)
Scientist 'C'

To
The Pay & Accounts Officer,
Department of Biotechnology,
New Delhi

Copy to:
1. The Director of Audit (CW & MII) IP Estate, AGCR Building, New Delhi – 110002
2. Prof. Sanjeev Jain, Director, Department of Psychiatry, National Institute of Mental Health & Neuro Science, NIMHANS, Bangalore – 560065, Karnataka
3. The Director, National Institute of Mental Health & Neuro Science, NIMHANS, Bangalore - 560029, Karnataka
4. Prof. Mahendra Rao and Dr. Apurva Sarin (Project Co-ordinator), inStem Bangalore
5. The Dean, Institute For Stem Cell Biology And Regenerative Medicine, National Centre for Biological Sciences, GKVK, Bellary Road, Bangalore - 560065, Karnataka
6. Cash Section DBT (2 copies)
7. IFD, DBT
8. File Copy

(Karthikeya K)
Scientist 'C'

INSTITUTIONAL ANIMAL ETHICS COMMITTEE (IAEC)
CHRISTIAN MEDICAL COLLEGE, VELLORE, INDIA

Dr. Anna B. Pulimood
Chairperson

Dr. Joe Varghese
Member Secretary

Date: 11.03.2021

Dr. Mohankumar KM,
Assistant Investigator,
Centre for Stem Cell Research,
CMC, Vellore.

Sub: Project Progress report - "***Pre-clinical model for gene therapy for thalassemia and sickle cell disease***"

Ref: IAEC no. **15/2017**

Dear Dr. Mohankumar,

The progress report submitted by you to the Institutional Animal Ethics Committee (IAEC) for the project titled "***Pre-clinical model for gene therapy for thalassemia and sickle cell disease***" has been reviewed by the IAEC at the meeting held on 27/02/2021.

The following members of the IAEC were present at the meeting:

1. Dr. Anna B. Pulimood, Chairperson, IAEC (Internal member)
2. Dr. A. Yasotha, CPCSEA Main Nominee (External member)
3. Dr. D. Sivaraman, Scientist from outside the institute (External member)
4. Dr. B.R. Senthilkumar, Socially aware member (External member)
5. Dr. Christhunesa Christudass, Scientist (Internal member)
6. Dr. Suceena Alexander, Scientist (Internal member)
7. Dr. K. Imayarasi, Veterinarian (Internal member)
8. Dr. Joe Varghese, Member Secretary (Internal member)

After discussion, ninety-six (96) NBSGW mice have been sanctioned for a period of 12 months.

Please see overleaf for responsibilities of the principal investigator.

With best wishes,

Yours sincerely,

Dr. Anna B. Pulimood
Chairperson

PRINCIPAL
Christian Medical College
Vellore - 632 002, Tamil Nadu, India.

Dr. Joe Varghese
Member Secretary

SECRETARY
Institutional Animal Ethics Committee,
Christian Medical College,
Bagayam, Vellore - 632 002, Tamil Nadu.

Dr. Anna B. Pulimood
Chairperson

Dr. Joe Varghese
Member Secretary

Responsibilities of the Principal Investigator:

1. Progress Report / Project completion report should be submitted for consideration of IAEC at prescribed intervals for review and it should not exceed the timeline of the research as mentioned in the Form B / IAEC approval certificate
2. Final report should be submitted to the IAEC of the establishment at the end of study.
3. All Serious Adverse Events (SAEs) and the interventions undertaken should be intimated to IAEC.
4. Protocol deviation, if any, should be informed with adequate justifications to IAEC and in case of large animals, it should be intimated to CPCSEA immediately for consideration of CPCSEA.
5. The procedural deviations in research protocol shall be treated as a fresh research protocol by CPCSEA.
6. Any new information related to the study should be communicated to IAEC immediately.
7. Premature termination of study should be notified to the IAEC with reasons along with summary of the data obtained so far.
8. Change of investigators / sites should be informed and approval of IAEC should be undertaken first.

Dr. Anna B. Pulimood
Chairperson

Dr. Joe Varghese
Member Secretary

PRINCIPAL
Christian Medical College
Vellore-632 002, Tamil Nadu, India

SECRETARY
Institutional Animal Ethics Committee,
Christian Medical College,
Bagayam, Vellore - 632 002, Tamil Nadu.

APPENDIX B

Publications

Molecular Therapy
Original Article

Base editing of key residues in the BCL11A-XL-specific zinc finger domains derepresses fetal globin expression

Vignesh Rajendiran,[1,2,12] Nivedhitha Devaraju,[1,3,12] Mahdi Haddad,[4] Nithin Sam Ravi,[1,2] Lokesh Panigrahi,[1,3] Joshua Paul,[1,3] Chandrasekar Gopalakrishnan,[5] Stacia Wyman,[6] Keerthiga Ariudainambi,[7] Gokulnath Mahalingam,[1] Yogapriya Periyasami,[1] Kirti Prasad,[1,3] Anila George,[1,2] Dhiyaneshwaran Sukumaran,[5] Sandhiya Gopinathan,[1] Aswin Anand Pai,[8] Yukio Nakamura,[9] Poonkuzhali Balasubramanian,[8] Rajasekaran Ramalingam,[5] Saravanabhavan Thangavel,[1] Shaji R. Velayudhan,[1,8] Jacob E. Corn,[6,10] Joel P. Mackay,[11] Srujan Marepally,[1] Alok Srivastava,[1,8] Merlin Crossley,[4] and Kumarasamypet M. Mohankumar[1]

[1]Centre for Stem Cell Research (a Unit of inStem, Bengaluru), Christian Medical College Campus, Bagayam, Vellore, Tamil Nadu 632002, India; [2]Sree Chitra Tirunal Institute for Medical Sciences and Technology, Thiruvananthapuram, Kerala 695 011, India; [3]Manipal Academy of Higher Education, Manipal, Karnataka 576104, India; [4]School of Biotechnology and Biomolecular Sciences, University of New South Wales, Sydney, NSW, Australia; [5]Department of Integrative Biology, School of Bioscience and Technology, Vellore Institute of Technology (VIT, Deemed to be University), Vellore, Tamil Nadu 632014, India; [6]Innovative Genomics Institute, University of California, Berkeley, Berkeley, CA 94704, USA; [7]Department of Biochemistry, Auxilium College, Vellore, Tamil Nadu 632006, India; [8]Department of Haematology, Christian Medical College & Hospital, Vellore, Tamil Nadu 632 004, India; [9]Cell Engineering Division, RIKEN BioResource Center, 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan; [10]Institute of Molecular Health Sciences, Department of Biology, Zurich, Switzerland; [11]School of Life and Environmental Sciences, University of Sydney, Sydney, NSW 2006, Australia

BCL11A-XL directly binds and represses the fetal globin *(HBG1/2)* gene promoters, using 3 zinc-finger domains (ZnF4, ZnF5, and ZnF6), and is a potential target for β-hemoglobinopathy treatments. Disrupting BCL11A-XL results in derepression of fetal globin and high HbF, but also affects hematopoietic stem and progenitor cell (HSPC) engraftment and erythroid maturation. Intriguingly, neurodevelopmental patients with ZnF domain mutations have elevated HbF with normal hematological parameters. Inspired by this natural phenomenon, we used both CRISPR-Cas9 and base editing at specific ZnF domains and assessed the impacts on HbF production and hematopoietic differentiation. Generating indels in the various ZnF domains by CRISPR-Cas9 prevented the binding of BCL11A-XL to its site in the *HBG1/2* promoters and elevated the HbF levels but affected normal hematopoiesis. Far fewer side effects were observed with base editing- for instance, erythroid maturation *in vitro* was near normal. However, we observed a modest reduction in HSPC engraftment and a complete loss of B cell development *in vivo,* presumably because current base editing is not capable of precisely recapitulating the mutations found in patients with BCL11A-XL-associated neurodevelopment disorders. Overall, our results reveal that disrupting different ZnF domains has different effects. Disrupting ZnF4 elevated HbF levels significantly while leaving many other erythroid target genes unaffected, and interestingly, disrupting ZnF6 also elevated HbF levels, which was unexpected because this region does not directly interact with the *HBG1/2* promoters. This first structure/function analysis of ZnF4–6 provides important insights into the domains of BCL11A-XL that are required to repress fetal globin expression and provide framework for exploring the introduction of natural mutations that may enable the derepression of single gene while leaving other functions unaffected.

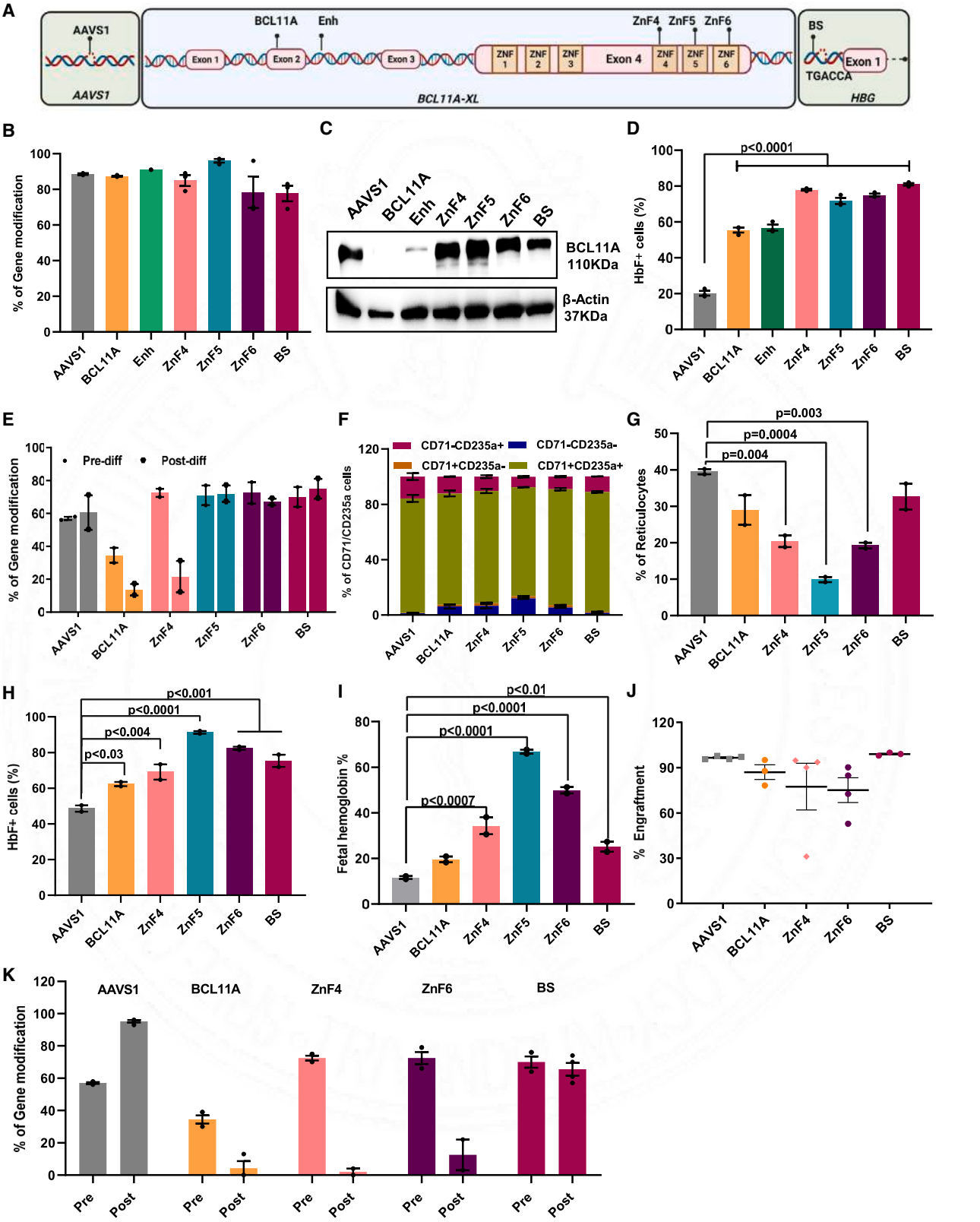
INTRODUCTION

BCL11A is a transcription factor that directly binds the fetal globin (*HBG1/2*) promoters and represses gene expression. It contains several classical C_2H_2-type zinc finger (ZnF) domains and is expressed in multiple tissues, such as the brain, liver, and bone marrow (BM), and in erythroid cells, hematopoietic stem cells (HSCs), and B cells.[1,2] Alternative splicing generates 4 major isoforms: BCL11A-XL, BCL11A-L, BCL11A-S, and BCL11A-XS, which retain a variable number of the ZnF domains. Only the longest isoform, BCL11A-XL, is implicated in fetal globin silencing. The shorter isoforms S and XS are expressed in the yolk sac and fetal liver, whereas the longer isoforms XL and L are more highly expressed in the adult BM.[3,4] The differential expression of BCL11A isoforms in the definitive erythroid compartment fits with a role in fetal globin silencing during

Received 28 April 2023; accepted 18 January 2024;
https://doi.org/10.1016/j.ymthe.2024.01.023.

[12]These authors contributed equally

Correspondence: Kumarasamypet M. Mohankumar, PhD, Centre for Stem Cell Research (a Unit of inStem, Bengaluru), Christian Medical College Campus, Bagayam, Vellore, Tamil Nadu 632002, India.
E-mail: mohankumarkm@cmcvellore.ac.in



(legend on next page)

development.[3] The functional importance of the BCL11A-XL isoform in fetal globin repression was firmly established by short hairpin RNA (shRNA)-mediated BCL11A-XL knockdown for the treatment of β-hemoglobinopathies, and BCL11A-XL is now regarded to be a promising target for β-hemoglobinopathy treatments.[5–8]

Genome-wide association studies initially implicated the *BCL11A* locus as a major quantitative trait associated with variation in fetal hemoglobin (HbF) levels.[9–11] Three sets of genetic variants are associated with high HbF, including single nucleotide variants (SNVs) in the intronic *BCL11A* enhancer,[12–15] regulatory SNVs in the *HBG1/2* promoters that disrupt the binding site (BS) of BCL11A-XL, and the exonic SNVs in the ZnF4 and ZnF5 domains contained specifically in the BCL11A-XL isoform.[16] Recent studies revealed that targeting the SNVs located in the +58 DNase hypersensitive site of the *BCL11A* intronic enhancer downregulates *BCL11A* expression in the erythroid lineage, which causes simultaneous derepression of HbF clinically, without altering the expression of *BCL11A* in non-erythroid cells.[14,15,17–21] The regulatory SNVs located in the *HBG1/2* promoters at the −115 site alter the core binding consensus motif of BCL11A-XL and substantially elevate the expression of HbF.[22,23] Ongoing clinical trials for the reactivation of HbF by genome editing emphasize the significance of the SNVs in the *BCL11A* intronic enhancer and BCL11A-XL BS in the *HBG1/2* promoters.[24] Structural studies have indicated that the BCL11A ZnF domains ZnF4 and ZnF5 interact directly with the *HBG1/2* promoter; however, the potential role of targeting these BCL11A-XL-specific ZnF domains on HbF regulation and other functions of BCL11A remains unexplored.[25]

In this study, we test the effect of targeting different ZnF domains specific to the BCL11A-XL isoform: ZnF4, ZnF5, and ZnF6. We initially disrupted each ZnF using CRISPR-Cas9-based cutting. Next, taking cues from the available *BCL11A* SNVs[10,16,26] in exon4 and the structural studies depicting the specific amino acid residues involved in the interaction of BCL11A-XL with DNA,[25,27] single nucleotide substitutions were generated using base editors. We compared our disruptions of the XL-specific ZnF domain with the other BCL11A-associated HbF-inducing targets such as the erythroid-specific enhancer and −115 site in *HBG1/2* promoters, and ablation of the *BCL11A* gene, both *in vitro* and *in vivo*. Furthermore, we analyzed the transcriptomic variations caused by the specific ZnF alterations and compared the effects with the other BCL11A-associated approaches. Our results provide important insights about the ZnF domains of BCL11A-XL required for fetal globin repression and other functions such as hematopoietic stem and progenitor cell (HSPC) engraftment and B cell production.

RESULTS

CRISPR-Cas9-mediated disruption of BCL11A-XL specific ZnF domains in adult erythroid cells derepresses fetal globin expression

Previous studies have shown that BCL11A-XL interacts with the *HBG1/2* promoters in adult erythroid cells using its C-terminal ZnF domain, which contains ZnFs 4, 5, and 6 (Figure S1A)[28]. However, the roles of each ZnF domain specific to BCL11A-XL in HbF regulation and erythroid maturation have not been established. Therefore, we performed CRISPR-Cas9-mediated disruption of each ZnF (guide RNA (gRNA): ZnFs 4, 5, and 6), and compared the effects with the other BCL11A-associated targets, such as disruption of the intronic erythroid-specific *BCL11A* enhancer , and its −115 cluster binding site (BS) in the *HBG1/2* promoters. In addition, gRNAs designed to ablate BCL11A (targeting *BCL11A* exon2) and adeno-associated virus integration site1 (*AAVS1*) were chosen as positive and negative controls, respectively (Figure 1A). Immortalized human umbilical cord derived erythroid progenitor 2 (HUDEP2) cells stably expressing Cas9 were transduced with the specific gRNAs (Figure S1B). A total of 70%–90% indels were achieved at all of the intended target sites (Figure 1B). Editing at ZnFs 4, 5, and 6 (amino acids 759, 785, and 810, respectively) did not alter BCL11A levels but resulted in the truncation of BCL11A-XL proteins to various lengths due to frameshift mutations (Figures 1C, S1C, and S1D). In contrast, significant downregulation of BCL11A protein was observed when targeting *BCL11A* exon2 (*BCL11A* knockout) or the enhancer. These results indicate that targeting *BCL11A* exon2 (N-terminal) may also reduce expression by nonsense-mediated RNA decay, whereas targeting the XL-specific ZnF domains (C-terminal) produced truncated variants of BCL11A.[29] As expected, the truncated BCL11A variants failed to repress *HBG1/2* at the adult stage. After erythroid differentiation (Figure S1E), the ZnF alterations produced a higher frequency of HbF^{+}

Figure 1. Targeted disruption of BCL11A-XL specific ZnF domains using Cas9 to upregulate fetal globin expression in adult erythroid cells

(A) Schematic representation of the target location of gRNA in different BCL11A-XL-specific ZnF domains (ZnF4, ZnF5, and ZnF6), Exon2, Enhancer (Enh), −115 cluster of *HBG1/2* promoters (BS), and the *AAVS1* site. (B) The frequency of indels was measured by ICE analysis at the respective target site in HUDEP2 cells during expansion. Mean ± SEM of n = 3 independent experiments. (C) The protein expression analysis by Western blot on the edited cells during differentiation (day 4) depicts the generation of BCL11A variants protein on targeting the XL-specific ZnF domains in contrast to the reduction of BCL11A protein while targeting Exon2 and the enhancer. (D) Intracellular HbF analysis using flow cytometry demonstrates the potential of ZnF mutants on HbF reactivation at the terminal stage of erythroid differentiation (day 8). Mean ± SEM of n = 3 independent experiments. (E) The percentage of gene modifications during HSPCs expansion after 48 h of electroporation and after erythroid differentiation (day 21), respectively. Mean ± SEM of n = 2 experiments on one donor. (F) The change in erythroid differentiation ability of the BCL11A-modified HSPCs was analyzed by flow cytometry using 2 erythroid-specific surface markers (CD71-transferrin; CD235a-glycophorin A) at the terminal stage of differentiation (day 21). Mean ± SEM of n = 2 experiments of one donor. (G) The measure of mature enucleated cells formed by the edited HSPCs during the terminal stage of erythroid differentiation (day 21) using NucRed stain and CD235a marker. Mean ± SEM of n = 2 experiments on one donor. (H) The number of HbF^{+} cells expressed, and (I) the total HbF tetramer in edited cells formed on BCL11A ZnF domain modification by flow cytometry and HPLC, respectively, measured during terminal erythroid differentiation (day 21). Mean ± SEM of n = 2 experiments on one donor. (J) The long-term engraftment of *BCL11A* total knockout, BS and ZnF domain-modified HSPCs after 16 weeks of transplantation in NBSGW mice. Mean ± SEM of $n \geq 3$ mice per group. (K) The persistence of gene modification in the edited cells is depicted by editing during both pre-infusion and after 16 weeks of transplantation.

cells (Figure 1D) and higher fetal globin derepression than seen in control cells (Figure S1F). The HbF^+ cells produced were noticeably more numerous than observed after modification of the enhancer (Enh vs. ZnF, $p < 0.0001$) and at a level comparable to BS disruption at the *HBG1/2* promoters. Despite previous reports demonstrating that ZnF6 has a weak direct interaction with the *HBG1/2* promoters compared to ZnF4 and ZnF5[28], we observed a significant elevation in HbF upon ZnF6 disruption. These results highlight the importance of each ZnF-domain specific to BCL11A-XL in the repression of the fetal globin genes.

BCL11A ZnF editing in $CD34^+$ HSPCs through CRISPR-Cas9 elevates HbF but also affects erythroid maturation

In addition to its major role in globin regulation, BCL11A is essential for human HSC maintenance and erythropoiesis. Therefore, we assessed the effect of BCL11A-ZnF disruption on cell survival and erythroid maturation in $CD34^+$ HSPCs obtained from a healthy donor by nucleofecting the CRISPR-Cas9-gRNA complex. Based on our HUDEP2 results in which enhancer disruption showed moderate HbF elevation, we determined the impact of the generating BCL11A-ZnF variants in erythroid cells derived from HSPCs, along with the BS disruption at the *HBG1/2* promoters, and *BCL11A* knockout at exon2. HSPCs exhibiting efficient gene modifications (Figure 1E) at the respective target sites were subjected to erythroid differentiation to assess the impact of the modifications on functions beyond HbF elevation during *in vitro* erythropoiesis.

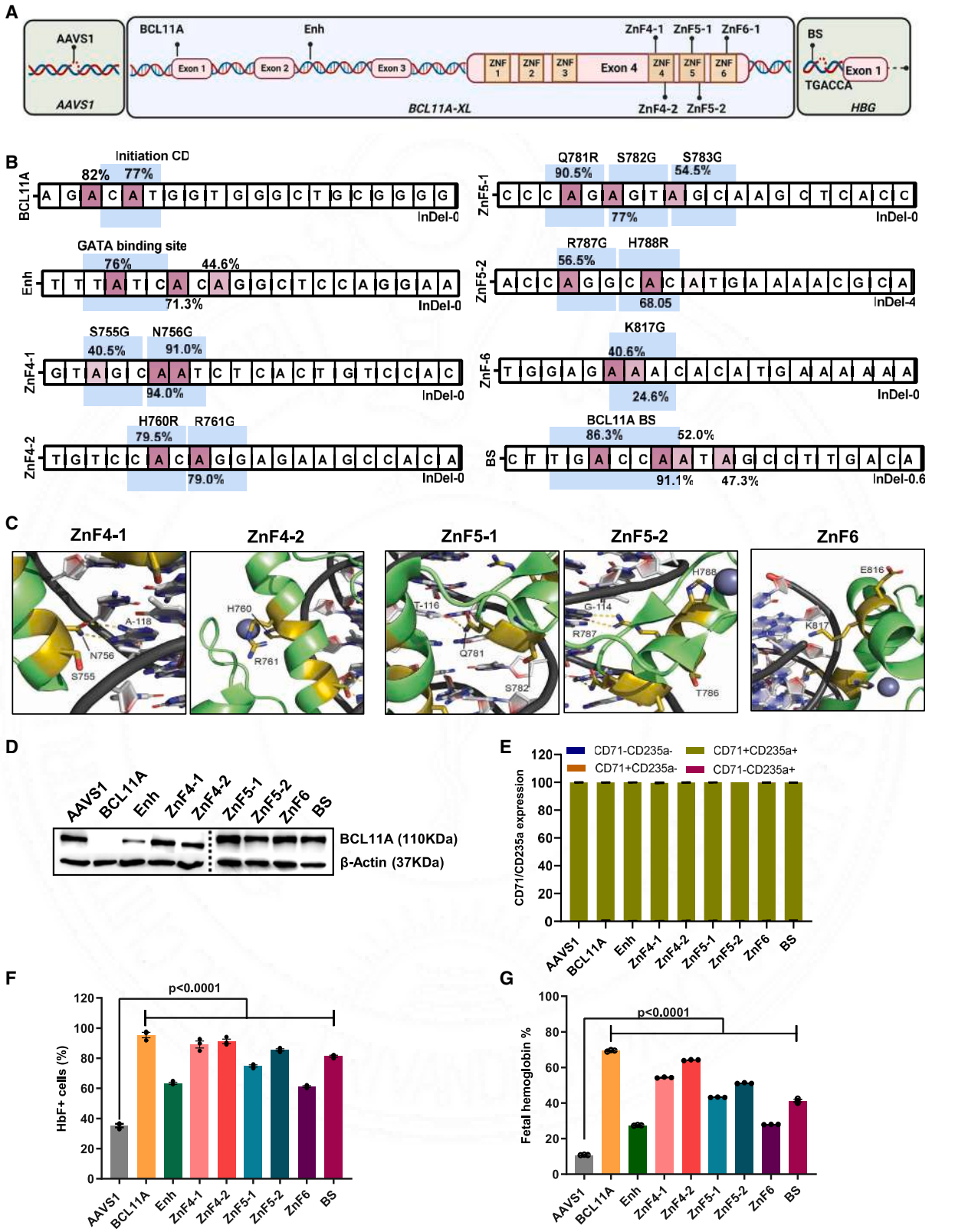
In line with previous research demonstrating the importance of BCL11A in cell maintenance and proliferation[30], the growth kinetics of committed erythroid progenitors from the edited samples were analyzed. The total *BCL11A* knockout and ZnF editing in $CD34^+$ HSPCs affected cell proliferation, whereas editing other targets did not affect survival (Figure S1G). The frequency of gene perturbations was assessed at the erythroid terminal differentiation stage. This revealed that except for ZnF4 and total knockout edited cells, gene modifications in all of the other targets persisted during differentiation (Figure 1E). This indicates that cells carrying alterations at the terminal sequences coding for ZnF5 and ZnF6 retain edits and persist across multiple cell divisions. A previous study reported that disruption of *BCL11A* exon4 (ZnF) did not impede early erythroid differentiation[29], whereas in our study, targeting the ZnF4, ZnF5, and ZnF6 domains showed a decline in the percentage of terminal-stage erythroid cells ($CD235a^+CD71^-$) (Figure 1F), with a significant reduction in the enucleation (Figure 1G). However, a remarkable level of fetal globin reactivation was seen in adult erythroid cells on targeting the BCL11A ZnF domains (Figures 1H and 1I), which was equivalent to levels obtained with the BS disruption, despite the differences observed in cell proliferation and differentiation. Overall, these results demonstrate the *in vitro* consequence of disrupting either at the exon2 (N-terminal) or the ZnF domains (C-terminal) of BCL11A on cell survival and erythroid maturation along with globin regulation.

BCL11A isoforms have been implicated in HSC engraftment and B lymphocyte maturation.[31,32] Therefore, we investigated the engraftment potential of BCL11A-XL ZnF-modified cells in the NBSGW mouse model. Since ZnF5 had exhibited significant erythroid maturation defects during *in vitro* experiments, we infused only the ZnF4 and ZnF6-modified HSPCs. As controls, we also infused BCL11A BS-disrupted HSPCs and total *BCL11A* knockout (exon2-edited) HSPCs. All of the transplanted animals showed chimerism of human cells ($hCD45^+$) in the BM after 16 weeks of infusion, even the cohorts with total *BCL11A* knockout cells (Figure 1J). Therefore, we genotyped the engrafted cells to study the long-term survival of gene-modified HSPCs, which revealed that the indel frequency in the *HBG1/2* promoter (BCL11A BS) edited cells remained unchanged, whereas HSPCs with *BCL11A* complete knockout did not survive in mouse BM, consistent with the previous findings.[30] We had speculated originally that BCL11A ZnF domain modification would have lesser impacts on HSPCs, but we observed a drastic decrease in ZnF4 and ZnF6 gene modifications in the BM cells (Figure 1K). Similar to total *BCL11A* knockout, we observed that the CRISPR-Cas9-mediated disruption of BCL11A-XL-specific ZnF domains does not support HSC engraftment. Furthermore, due to the inadequate engraftment potential of the BCL11A ZnF mutants, we were unable to investigate its multilineage differentiation ability. These findings demonstrate that even a minimal loss of the BCL11A-XL C-terminal coding region functions similar to total protein loss and can negatively affect the engraftment of HSCs.

Specific base substitutions in the BCL11A-XL C-terminal ZnF domains reactivate HbF expression in adult erythroid cells

CRISPR-Cas9-mediated targeting of BCL11A-XL-specific C-terminal ZnF domains resulted in multiple truncated variants hampering the concise understanding of the role of BCL11A-XL ZnF domains in HbF repression, engraftment, and erythroid maturation. To address this, we introduced more subtle mutations at specific amino acid residues in the ZnFs using an adenosine base editor in HUDEP2 cells anticipating HbF elevation without compromising other functional roles of BCL11A.

The key nucleotides at ZnF4 (amino acids 742–765) and ZnF5 (amino acids 772–795) were targeted to alter crucial motifs required for DNA interaction and where possible to resemble as closely as possible known missense variants associated with higher HbF levels in patients with neurodevelopmental disorders but essentially normal hematology (Figure S2A)[10,16,25–27]. With the current base editing technology, it is not possible to recreate the exact mutations, but we targeted mutations as closely as possible to the sites found in patients. No naturally occurring mutations have been reported in ZnF6 (amino acids 803–827), but we designed a gRNA that would result in a base substitution at a potential DNA contact region in the ZnF6 domain (Figure 2A). The fetal globin reactivation potential of base substitutions in each ZnF domain was compared with the same *BCL11A* control targets used in previous experiments, along with the gRNA at the initiation codon for the total knockout of *BCL11A*. Subsequently, all of these



(legend on next page)

gRNAs were individually transduced into HUDEP2 cells with more than 90% efficiency (Figure S2B).

After 8 days of expansion, the efficiency of base conversion at the target sites was evaluated by Sanger sequencing (Figure 2B). In the case of the ZnF4 domain, the ZnF4-1 gRNA specifically converted the nucleotides that encode the amino acids S755 and N756 approximately with 40% and 90% base conversion efficiency, respectively. S755 makes van der Waals interactions with the −119 (T:A) base and the DNA backbone at the *HBG1/2* promoters, whereas N756 forms a specific double hydrogen bond with −118 (A:T) (Figure 2C, ZnF4-1). The ZnF4-2 gRNA generated novel mutants H760R and R761G with ∼80% efficiency. Although these amino acids do not contact DNA, H760 is one of the Zn-ligating residues in ZnF4, and its mutation likely perturbs the structure of the domain (Figure 2C, ZnF4-2). Alteration of the ZnF5 domain was achieved using ZnF5-1 gRNA to efficiently convert bases that encode Q781 (∼90%), S782 (∼75%) and S783 (∼55%). Q781 forms a hydrogen bond with −116 (T:A) (Figure 2C, ZnF5-1). Similarly, the ZnF5-2 gRNA generated specific nucleotide perturbations that mutates R787 (∼55%), and H788 (∼70%); R787 forms hydrogen bonds with −114 (G:C), whereas H788 is one of the Zn-coordinating residues in ZnF5 (Figure 2C, ZnF5-2). The ZnF6 gRNA exhibited a less efficient (∼40%) conversion of bases and encoded changes to K817; this residue is predicted to make nonspecific electrostatic interactions with the DNA backbone (Figure 2C, ZnF6).

The control sites including the −115 cluster in the *HBG1/2* promoter, *BCL11A* initiation codon and the *BCL11A* enhancer were substantially modified, with overall efficiencies greater than 70%. Unintended indels were less than 5%, despite the high editing efficiency at all of the target sites. The consequence of the generated BCL11A-XL mutants on protein expression was investigated to analyze the impact of pathogenic SNVs on the stability of the BCL11A protein. Interestingly, the XL-specific ZnF mutants exhibited no effect on protein synthesis and were expressed similarly to the *AAVS1* negative control editing, whereas the alteration of initiation codon affected the BCL11A protein expression mimicking *BCL11A* knockout as expected (Figure 2D).

We assessed the functional consequence of the *BCL11A* mutations on fetal globin expression during terminal erythroid differentiation (Figure 2E). Interestingly, the HbF reactivation observed in all of the ZnF mutants was comparable to complete *BCL11A* knockout. The increase in the number of HbF^+ cells ranged from 60% to 90% across the ZnF mutants, which may be due to a difference in editing efficiency or the impact of the different mutations (Figure 2F). The ZnF4 mutations increased the number of HbF^+ cells more than the ZnF5 or ZnF6 mutations. Although the editing frequency at the ZnF6 target site was lower, significant HbF elevation was observed, so we hypothesize that ZnF6 also influences BCL11A binding to the *HBG1/2* promoters. Consistent with the increase in HbF^+ cells, the ZnF4 mutants produced higher levels of total HbF relative to mutations in other ZnF domains (ZnF5 and ZnF6).

Importantly, the majority of ZnF-specific mutations outperformed the disruption of the BCL11A BS in the *HBG1/2* promoters in HbF reactivation (Figure 2G). It may be that BCL11A binds at additional sites in the globin locus; therefore, mutating the DNA-binding domain has a more profound effect than mutating a single binding motif. These findings confirm that the induction of HbF does not require haploinsufficiency or the total loss of BCL11A but can be mediated by targeted mutations.

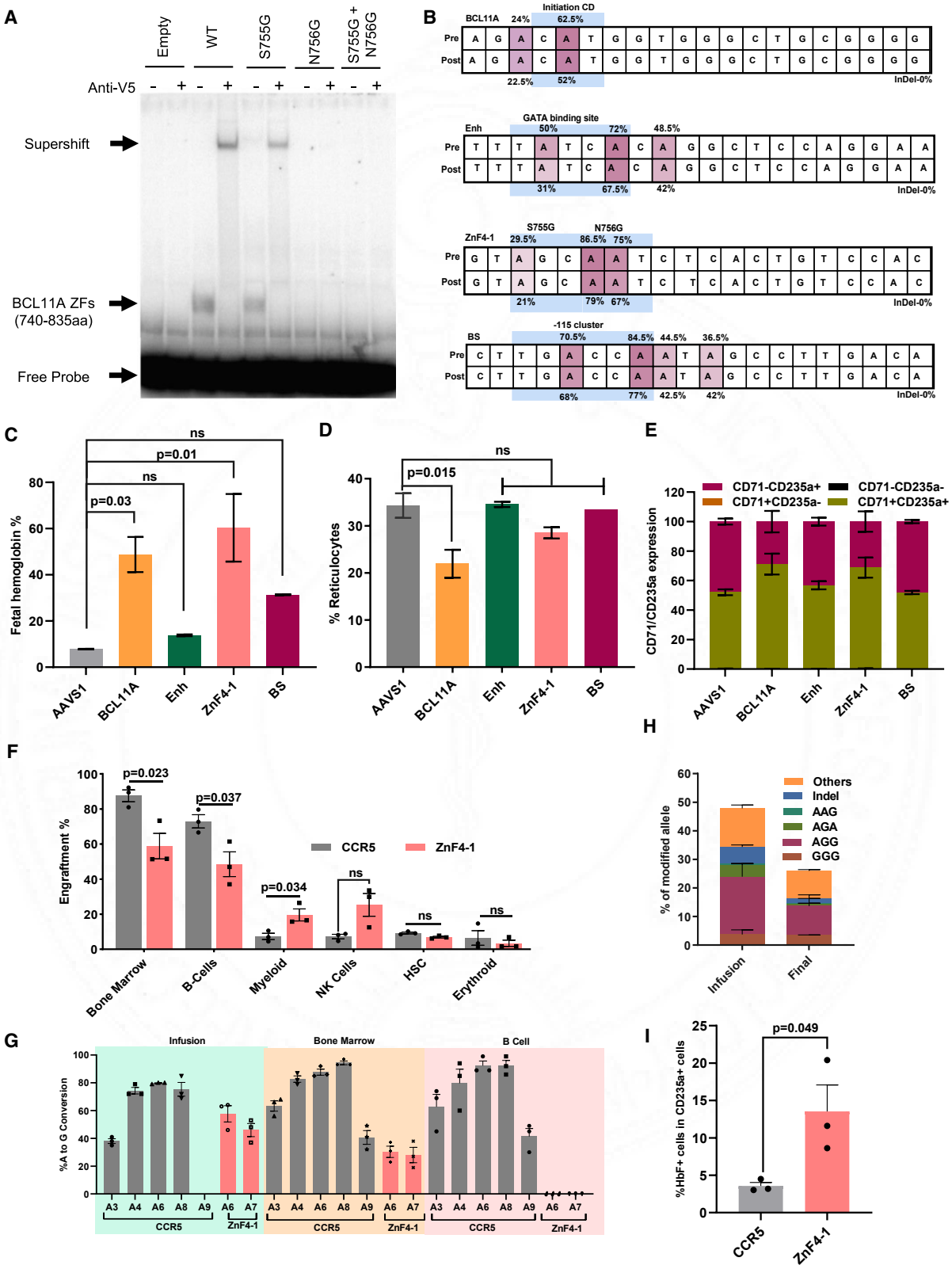
Together, both the *in vitro* and *in silico* results imply that altering the C-terminal XL-specific ZnF domains involved in binding to the target DNA (e.g., the motifs in the *HBG1/2* promoters) exhibited an HbF upregulation similar to that observed with total BCL11A loss. Furthermore, unique base conversations that alter key amino acid residues are sufficient to significantly affect ZnF domain function.

Generation of BCL11A-ZnF mutants in HSPCs leads to HbF upregulation during *ex vivo* erythropoiesis with reduced hematological effects

A previous study reported that eliminating hydrogen bonds formed between N756 and the −118 (A:T) base by creating N756A mutation reduces the affinity of BCL11A for DNA by 8-fold.[25] To confirm that changing N756 to glycine as achieved via base editing had the same effect, we validated the loss of binding of the ZnF4-1 mutant to the −115 site in the *HBG1/2* promoters by gel shift assays using proteins containing S755G, N756G, and S755G and N756G double mutants (Figures 3A and S3A). We observed that the N756G mutant, both individually and in combination with S755G, abrogated the binding

Figure 2. The key base substitutions in BCL11A-XL-specific ZnF domains reactivate HbF expression in adult erythroid cells

(A) A diagrammatic representation of each of the target sites to induce nucleobase modifications by using ABE. BCL11A-gRNA targeting initiation codon of BCL11A, Enh-gRNA targeting the *BCL11A* erythroid-enhancer region, 2 gRNAs each targeting ZnF4 (ZnF4-1 and -2) and ZnF5 (ZnF5-1 and -2), respectively, and 1 gRNA targeting ZnF6 to produce missense mutations. BS gRNA to alter the BCL11A-XL BS in the −115 cluster of *HBG* promoter. AAVS1-gRNA targeting the safe harbor site at the *AAVS1* locus for negative control. (B) Base conversion efficiency at all of the target sites on day 8 of expansion was analyzed using the EditR *in silico* tool after Sanger sequencing. Mean ± SEM of n = 3 independent experiments. (C) Close-up views of the BCL11A-XL-specific ZnF (ZnF4, ZnF5, and ZnF6) structures, illustrating the amino acids targeted in each of the indicated mutants. Yellow dashed lines indicate hydrogen bonds, and bases are named according to the *HBG1/2* promoter numbering. (D) The protein expression analysis by western blot confirms no significant change in BCL11A expression on ZnF editing compared to the reduction of BCL11A protein observed on disrupting the initiation codon and the BCL11A-enhancer; 35 μg of protein used for loading. (E) Analysis of the erythroid differentiation profile (CD71 and CD235a) of the base-edited HUDEP2 cells on the end stage of differentiation (day 8) showed no significant variation among the edited samples. Mean ± SEM of n = 3 independent experiments. (F) Increase in percentage of intracellular HbF levels measured at the end stage of differentiation (day 8) by flow cytometry (HbF^+ cells) on generated BCL11A ZnF mutants. Mean ± SEM of n = 3 independent experiments. (G) The elevation in the total HbF tetramer formation in adult erythroid cells in the BCL11A-modified cells was analyzed at the end stage of differentiation (day 8). Mean ± SEM of n = 3 independent experiments.



(legend on next page)

to the target site, whereas the S755G mutant alone did not significantly affect binding. Furthermore, to phenotypically evaluate the contribution of each mutant on fetal globin derepression, we performed single-cell sorting in HUDEP2 cells and observed that N756G contributed significantly to the HbF induction, consistent with the gel shift assay results (Figure S3B).

Given the results in HUDEP2 cells, we next tested the effect of base editing at BCL11A-XL ZnFs in healthy donor CD34$^+$ HSPCs, using adenine base editor 8e (ABE8e) mRNA and again compared them with the *BCL11A* associated targets (Figure S3C). We obtained base conversions at the respective target sites with overall efficiencies of 60%–70% in HSPCs (Figure 3B). As expected, the ZnF4-1 gRNA generated 2 missense mutations at N756 and S755 at ZnF4, whereas the gRNAs targeting the other *BCL11A*-associated targets altered either the *BCL11A* start codon (to achieve total knockout) or the intronic enhancer or the −115 site in the *HBG1/2* promoters. Notably, all of the base modifications persisted throughout erythroid differentiation (Figure 3B).

To investigate the effect of the BCL11A ZnF4 mutant HSPCs on cell survival and differentiation, we monitored the base-modified HSPCs throughout differentiation to the erythroid lineage. Cell proliferation in the *BCL11A* knockout mutant was considerably lower than that in the ZnF4 mutant, confirming that the BCL11A-XL isoform has a role in erythroid cell proliferation and survival (Figure S3D). More important, the specific mutations in ZnF4 domain did not impair cell viability.

We anticipated that disrupting multiple binding of BCL11A in the globin locus could lead to stronger upregulation than the ablation of the single BCL11A BS at *HBG1/2* promoters. Therefore, we assessed the potency of the BCL11A ZnF4 missense mutant on the derepression of fetal globin expression in erythroid progenitors derived from base-edited HSPCs. Similar to the total knockout, the *BCL11A* ZnF4 mutation resulted in high HbF$^+$ cells (Figure S3E) and high total HbF production (Figure 3C).

We then investigated the effect of the BCL11A ZnF4 missense mutant on terminal erythropoiesis because our CRISPR-Cas9 cutting study showed a significant erythroid maturation delay in BCL11A ZnF truncated protein-expressing erythroid cells during the terminal stages. Interestingly, BCL11A ZnF4 base-edited cells underwent terminal erythroid maturation (CD235a$^+$/NucRed$^-$) more efficiently than total knockout cells and truncated proteins generated in the CRISPR-Cas9 cutting experiments on day 21 of HSPC differentiation (Figure 3D). Specifically, we also observed higher HbF-expressing enucleated cells at the erythroid terminal maturation phase in the BCL11A ZnF mutant (Figure S3F). Notably, the BCL11A ZnF4 mutant exhibited a modest decrease in the frequency of cells expressing erythroid differentiation markers (CD235a$^+$/CD71$^-$) with normal enucleation (Figure 3E). However, this effect was modest when compared to CRISPR-Cas9-mediated truncation of the BCL11A ZnF4 domain. Altogether, these data demonstrate that the fetal globin reactivation induced by altering the individual amino acids specific to the BCL11A-XL isoform involved in binding to the *HBG1/2* promoters is preferable to indel formation via CRISPR-Cas9, and if refined, could ultimately serve as a strategy for fetal globin induction.

We then investigated the repopulation ability of the BCL11A ZnF4 base-edited human CD34$^+$ HSPCs by transplanting them into immunodeficient NBSGW mice, noting that BCL11A truncated proteins from the CRISPR-Cas9 experiment exhibited defective engraftment. We observed efficient chimerism of human cells (hCD45$^+$) in the control edited cells, whereas the ZnF4 mutant exhibited reduced engraftment (Figure 3F), but more important, the ZnF4-1 mutant cells were able to survive and repopulate over time. Because the neurodevelopmental patients harboring certain missense mutations in BCL11A-XL-specific ZnF domain exhibit essentially normal hematopoiesis, we analyzed the impact of the BCL11A ZnF4 mutant on multilineage repopulation potential. Consistent with a moderate reduction in engraftment, ZnF4-1 mutants also showed altered levels of lineage distribution, notably a significant decline in B cells, with a bias toward myeloid cells (Figure 3F). Although re-creating the exact patient SNVs using base editor is not feasible, we aimed to mimic the heterozygous nature of these BCL11A-ID patients by reducing the editing efficiency at the target site. Genotypic analysis of the engrafted human cells showed an evident drop in the editing frequency of the ZnF4-1 mutant within both the BM (2.5-fold) and total loss in the B cell engrafted compartment (Figure 3G). We observed that N756G amino

Figure 3. Generation of BCL11A ZnF mutants in HSPCs exhibit elevated levels of HbF on erythroid differentiation

(A) Electrophoretic mobility shift assay depicting the loss of interaction between the ZnF4-1 (N756 mutants) mutants to a TGACC motif. Nuclear extract of COS-7 cells expressing epitope tagged fused ZnF4–6 of wild-type BCL11A and mutants with radiolabeled fetal globin promoter BCL11A binding motif TGACC. The black arrow shows the BCL11A ZF4-6 binding to the probe, and super-shifted bands are observed when antibody to the V5 epitope is added. (B) Target base perturbation efficiency at each target site was analyzed after 2 days of electroporation in EditR software. Mean ± SEM of n = 2 experiments of one donor. (C) The reactivation of total HbF measured by HPLC at the terminal stage of erythroid differentiation (day 21). Mean ± SEM of n = 2 experiments of one donor. (D) The percentage of edited cells capable of maturating into enucleated reticulocytes *in vitro* measured by CD235a expression and NucRed stain. Both erythroid differentiation and maturation analysis were measured at the terminal stage of erythroid differentiation (day 21). (E) The erythroid differentiation potential of base-edited HSPCs was examined using CD235a and CD71 surface markers at the terminal stage of erythroid differentiation (day 21). (F) Engraftment percentage of base-edited HSPCs and multilineage repopulation in NBSGW mice after 15 weeks of transplantation. (G) The percentage of base editing observed before infusion, after 15 weeks in the mouse BM and in the B cell subset. B cells were flow sorted using human CD45 and CD19 surface markers. (H) NGS analysis depicting the multiple editing pattern within engrafted ZnF4-1 edited cells derived from the mouse BM after 15 weeks of engraftment. The first A denotes the adenine base in the S755 position and subsequent AA denotes the dual adenine bases present in the N756 position, which is converted to G with different frequencies. Others in the graph indicate the summation of all of the nonspecific mutations. (I) Percentage of HbF$^+$ cells in the CD235a$^+$ cell population within the NBSGW BM after 15 weeks. n = 3; mean ± SEM.

acid production was prevalent over N756D or N756S amino acids. Furthermore, N756G amino acid also persisted at the 15 weeks of *in vivo* engraftment in NBSGW mice (Figure 3H). The CD235a$^+$ erythroblasts present in the mouse BM of the ZnF4-1 mutant engrafted mice showed increase in HbF$^+$ cells compared to edited controls (Figure 3I).

Furthermore, the persistence of single mutants in the BM population suggests that creating monoallelic single mutation resembling the patient SNVs at the ZnF domains elevates HbF, but the data also suggest that C-terminal ZnF domains have an important function in HSPC engraftment and B cell production outside the globin regulation within the erythroid lineage.

Global transcriptomic analysis of the BCL11A ZnF4 base-edited erythroid cells derived from CD34$^+$ HSPCs

Since BCL11A regulates multiple pathways involved in the development and differentiation of various cell types, any alteration of *BCL11A* requires extensive characterization.[30,33] Therefore, we compared genome editing at the XL-specific ZnF domain along with the other BCL11A-associated targets. A total of 691 genes were differentially expressed after the complete loss of *BCL11A*, whereas disruption of ZnF4 (123 genes) and BS (94 genes) showed significantly fewer differentially expressed genes. The expression of only 9 genes was observed to be significantly altered in the case of erythroid-enhancer disruption, consistent with the modest HbF induction observed *in vitro* after erythroid differentiation (Figures 4A and 4B). The transcriptional repressor activity of BCL11A is evident based on the fact that more genes were upregulated on modifying BCL11A compared to the BS disruption in the *HBG1/2* mutation. This clearly demonstrates that BCL11A ZnF4 mutant generation has fewer effects beyond the upregulation of fetal globin expression than are observed with the total ablation of *BCL11A* expression.

Next, we examined the known primary targets[34] of BCL11A exhibiting significant differential expression among our study group (Figure 4C). In all of the target sites, *HBG2* was the commonly upregulated gene, as expected. *BGLT3* and *HBG1* are the 2 other genes commonly upregulated among the ZnF4-1, BS, and BCL11A edited cells. Interestingly, the ZnF4 editing shared upregulation of the *HBZ, HBBP1*, and *KLHDC8B* genes with *BCL11A* total knockout. Consistent with the whole transcriptome profile, the disruption of BCL11A BS generates no other significant alterations in the primary targets (Figure 4D). Moreover, the significant increase in the expression of the cell-cycle-associated genes, such as *CDKN2B*, *PPP2R5B*, and *BCL6* occurs only in *BCL11A* complete knockout cells and not in the ZnF4 mutant cells. This may explain why the ZnF4 mutation has modest effects on erythroid function other than globin regulation.

We further validated the expression of these globin locus genes by real time qRT-PCR (Figure S4A). These results confirm that specifically disrupting the ZnF4 domain has fewer effects on other erythroid genes compared to the total ablation of BCL11A.

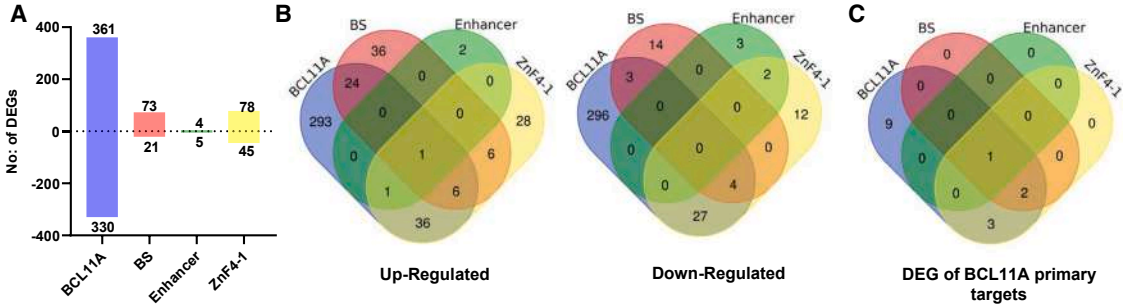
Next, to evaluate whether the observed induction of HbF is specific to the on-target editing activity, we performed targeted deep sequencing at the potential off-target (OT) sites (with up to 3 mismatches) predicted by the *in silico* algorithm COSMID (Figure S4B). The top 9 sites were deep sequenced in the base edited HSPCs and we observed no significant DNA variation between the control and the edited samples indicating that the gRNA used is highly specific to the *BCL11A* ZnF4 locus. The CIRCOS plot depicting the predicted OT locations shows that none of the OT sites are in the coding region, providing further evidence that the gRNA is highly specific (Figure S4C). To further rule out the impact of unintended RNA deamination caused by ABE8e, we analyzed the whole transcriptome RNA sequencing (RNA-seq) information of the ZnF4-1 edited HSPCs during erythroid differentiation and compared it with the other BCL11A-associated targets. We observed that no significant difference in the adenosine-to-inosine (A-to-I) conversion among the gRNAs used in this study (Figure S4D). These results confirm the specificity of targeting the ZnF4 using ABE8e.

DISCUSSION

From the time that BCL11A was implicated in fetal globin regulation, most of the gene therapy clinical trials for β-hemoglobinopathies have focused on the modulation of *BCL11A* through various approaches. The erythroid-specific enhancer has been targeted, and shRNA-mediated downregulation of *BCL11A* expression specifically within the erythroid lineage also results in the derepression of fetal globin. Disrupting the BCL11A BS at the −115 site in the *HBG1/2* promoters by both CRISPR-Cas9-mediated disruption and base editing is being pursued. In addition to these studies that target *BCL11A* at the genomic or transcript level, small molecules such as nanobodies have been developed to selectively reduce BCL11A-XL at the protein level.[35–37] However, the spatiotemporal window for optimal nanobody delivery is still being explored. Although these studies emphasized the clinical importance of BCL11A in the treatment of β-hemoglobinopathies, they also provide insights into the role of BCL11A-XL protein in blood cell gene expression.

Here, we investigated a novel strategy to specifically alter the DNA-binding domains of BCL11A-XL to functionally mimic naturally occurring mutations affecting the BCL11A-XL isoform found in patients with normal or near-normal hematopoiesis. One of the major advantages of this strategy is that it generates very high levels of HbF induction, presumably because it impairs the binding of BCL11A-XL, not only to its site at −115 in the fetal globin promoters but also to other sites that may lie elsewhere in the globin locus and be required for complete fetal globin silencing.[38] Specifically, modification of ZnF4 led to very high levels of HbF in the adult stage, similar to those seen with the total loss of BCL11A protein.

BCL11A-XL has been demonstrated to interact with the TGACCA site in the *HBG1/2* promoters (−115 cluster) via primarily the ZnF4 and ZnF5 domains.[25,28] Several microRNAs and transcriptional regulators, including LIN28B, have been shown to affect the production of the XL isoform by binding to *BCL11A* exon4.[34] Many studies



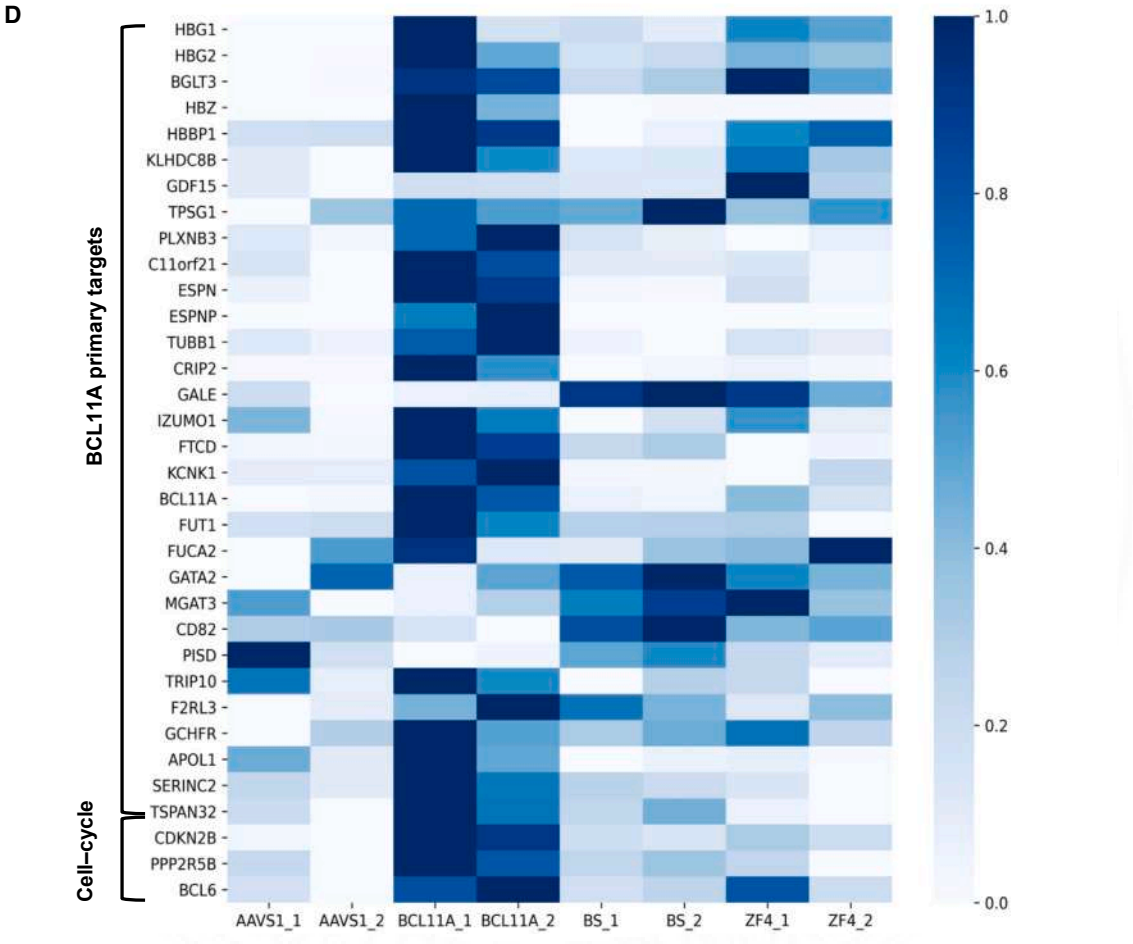


Figure 4. Transcriptomic analysis of ZnF4-1 mutant in erythroid cells derived from HSPCs

(A) Total number of differentially expressed genes in total knockout (*BCL11A*), enhancer-edited (Enhancer), ZnF4-1 modification, and −115 site at *HBG1/2* (BS) disruption compared to the edited control. (B) Correlation between the different transcriptome among upregulated and downregulated genes in total knockout, enhancer edited, ZnF4-1 modification, and BS disruption compared to the edited control. Significant gene expression is maintained at false discovery rate (FDR) 5% and −1>fold change >+1. (C) Correlation showing the significantly expressed (FDR 5% and −1>fold change >+1) primary targets among the different edited groups. (D) Heatmap showing the BCL11A primary target gene counts among all of the edited HSPCs.

have focused on targeting the BCL11A recognition motif at the *HBG* promoter rather than the endogenous BCL11A-XL ZnF domains to eliminate the BCL11A:*HBG* interaction.

To refine our understanding of ZnF domains in globin regulation, we targeted the endogenous BCL11A XL-specific ZnF (ZnF4, ZnF5, and ZnF6) domains in adult erythroid cells using CRISPR-Cas9-mediated editing and indel formation. Gene perturbations at the ZnF4, ZnF5, and ZnF6 domains efficiently generated ZnF truncated mutants that produced elevated levels of HbF expression, comparable with changes seen by targeting other BCL11A-associated regions such as the enhancer or the −115 cluster in the *HBG1/2* promoters. Although ZnF6 has not been observed to interact directly with the DNA BS, we observed HbF upregulation on ZnF6 disruption, suggesting that it makes an indirect contribution and is required for proper fetal globin silencing. The truncated BCL11A-XL mutants in HSPCs were defective in cell survival and erythroid maturation, and they also failed to survive in an *in vivo* mouse BM environment.

To further explore the role of specific subdomains within each ZnF and to attempt to reduce side effects, we used a high-fidelity base editor to alter the key bases, present in XL-specific ZnF domains, that are known to interact with the target DNA strand. Together with the available database on SNVs, we particularly aimed to target the reported *BCL11A* SNVs associated with neurodevelopmental disorders, in which patients show derepressed fetal globin but otherwise have normal hematology. The current base editors have limited capability to generate desired heterozygous mutations without any bystander conversion. Nevertheless, we generated different ZnF-specific missense mutants in BCL11A-XL using an ABE8e and confirmed the reduced binding of BCL11A-XL to its DNA recognition site in a gel shift assay, demonstrating that specific mutations, with the expected compromised functions, could be made.

Among the ZnF domains, the reactivation of HbF was highest with ZnF4 and ZnF5 mutants, which was equivalent to total *BCL11A* suppression, followed by BS disruption at the *HBG1/2* promoters, and the enhancer-mediated *BCL11A* downregulation. This emphasizes that base modifications at the ZnF domains are sufficient to produce significant levels of HbF. However, we also observed a reduction in engraftment and B cell production. These data are in line with another study that reported the role of ZnF domains in the immunomodulatory activity of B cells against Hirame novirhabdovirus in flounder.[39] The exact effect of the ZnF4-1 mutation in the functional activity of B cells and HSPC engraftment need to be studied to extend this approach clinically. However, the high efficiency of ABE8e resulted in more than 1 missense mutation in each of the ZnF domains, creating double mutants, unlike the natural observed variants. Nevertheless, the ZnF4 mutant resulted in effective HbF reinduction with near-normal erythroid maturation *in vitro* and thus generated fewer side effects than CRISPR-Cas9-mediated disruption of the domain or other major perturbations of the *BCL11A* gene.

Overall, altering the BCL11A ZnF4 domain may disrupt all of the potential interactions with the globin cluster, leading to effective induction of HbF compared to the previously reported multiplex editing at both the *BCL11A* enhancer and BS.[40] Furthermore, we observed minimal transcriptomic changes on altering the ZnF4 domain compared to total *BCL11A* loss. The transcriptomic profiling clearly demonstrated the upregulation of globin locus genes other than *HBG1/2*, such as *HBBP1* and *BGLT3*, consistent with reports that the latter genes are also regulated by BCL11A.

A recent study highlighted the homology between the BCL11A and BCL11B proteins,[35,37] but due to the codon degeneracy, the sequences encoding the 2 distinct proteins varied. It is interesting that the nonoccurrence of the predicted OT site at BCL11B reconfirmed that the target is specific to BCL11A-XL and does not affect the function of BCL11B. Collectively, at both DNA and RNA levels, the occurrence of negligible levels of OTs suggests that the observed fetal globin derepression is solely due to the base alterations at the ZnF4 domain.

To conclude, our approach has shown that one can achieve HbF reactivation through minimal genetic alterations at key bases specifically in the BCL11A-XL with fewer transcriptomic changes. We believe that our approach of altering the DNA recognition domain could be extended further to other transcription factors that use multiple ZnFs to bind different sets of target genes. Our structure/function analysis of ZnF4–6 using base editor approaches provides important insights into the domains of BCL11A-XL required to repress fetal globin expression and its role in HSPCs. Furthermore, refining our editing strategy would act as a framework for exploring other natural mutations that may selectively reverse the repression of individual genes while preserving its other functions.

MATERIALS AND METHODS

Plasmids and lentivirus production

The target guide RNAs (Table S1) were cloned into the pLKO5.sgRNA.EFS.GFP plasmid (a gift from Benjamin Ebert, Addgene no. 57822). The gRNA-containing plasmids were then cotransfected with the VSV-G (vesicular stomatitis virus G protein) plasmid, pMD2.G (a gift from Didier Trono, Addgene no. 12259) and the envelope plasmid psPAX2 (a gift from Didier Trono, Addgene no. 12260) at a 2:1:1 ratio in HEK293T cells using the Mirus transfection reagent (1:3 ratio of DNA:reagent). Viral supernatants were collected at 48 and 72 h posttransfection, concentrated using an in-house preparation of PEG-6000 concentrator, and stored at −80°C until use.

Cell culture

HUDEP2 cells were cultured and differentiated as described previously. HUDEP2 cells, stably expressing Cas9 or ABE8e, were generated and validated.[22] HUDEP2 lines were transduced with the gRNA lentivirus, and the transduction efficiency was determined by flow cytometry analysis of GFP expression after 48 h.

$CD34^+$ HSPCs were isolated from the leftover G-CSF (granulocyte-colony-stimulating factor) mobilized blood of healthy donors after

infusion to patients with informed consent as per clinical guidelines authorized by the institutional review board of Christian Medical College. The peripheral blood mononuclear cells were separated using density gradient centrifugation (Lymphoprep Density Gradient Medium, STEMCELL Technologies) and washed with 1× red blood cell (RBC) lysis buffer to remove any residual RBCs present. $CD34^+$ HSPCs were purified using EasySep Human CD34 Positive Selection Kit II (STEMCELL Technologies). The isolated $CD34^+$ cells were expanded and differentiated into the erythroid lineage as reported earlier, and the growth kinetics were measured by cell counting at each media change.[41]

COS-7 cells were cultured in DMEM (Gibco) supplemented with 10% fetal bovine serum (Gibco) and 1% penicillin-streptomycin-glutamine (Gibco) and incubated at 37°C in 5% CO_2. Cells were lifted for passaging by incubation in 0.05% TrypLE Express Enzyme (Gibco) at 37°C for 5 min.

ABE8e mRNA production

To produce ABE8e mRNA, we used the ABE8e plasmid (a gift by David Liu, Addgene no. 138495) as the template and performed *in vitro* transcription as previously described.[42] The resulting mRNA was stored in smaller aliquots at −80°C until use.

Nucleofection

For Cas9 experiments, 0.2 million HSPCs were electroporated with a Cas9:gRNA ratio of 50:100 pmol using a Lonza 4D-Nucleofector (pulse code: DZ100). For the base-editing experiments, we nucleofected 1 million HPSCs with 5 μg ABE8e mRNA with 100 pmol gRNA using a GTx electroporator (MaxCyte, pulse code: HSC-3).

Gene modification analysis

Genomic DNA from the edited and unedited cells was isolated and the target regions were PCR amplified for Sanger sequencing using the primers listed in Table S2. The efficiency of gene modification generated by Cas9 was analyzed using the Synthego Inference of CRISPR Edits (ICE) program.[43] To estimate the base conversion percentage, the Sanger sequencing files were analyzed using the EditR analysis software.[44]

COS-7 cell transfections and nuclear extractions

COS-7 cells were transiently transfected with pcDNA3-based plasmids to overexpress BCL11A ZnF4–6 and mutants. Cells were transfected in 10-cm plates with 5 μg plasmid using FuGENE 6 Transfection Reagent (Promega). The mammalian expression plasmids used are listed in Table S3. Cells were washed in PBS after 48 h incubation at 37°C. Cells were harvested and resuspended in hypotonic lysis buffer (10 mM HEPES-KOH, pH 7.9, 1.5 mM $MgCl_2$, 10 mM KCl, 5 mM DTT, 1 mM PMSF, 0.01 mg/mL aprotinin, and 0.01 mg/mL leupeptin) and incubated on ice for 10 min. Cells were spun down and the pellets were resuspended in extraction buffer (20 mM HEPES-KOH, pH 7.9, 25% glycerol, 420 mM NaCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 5 mM DTT, 1 mM PMSF, 0.01 mg/mL aprotinin, and 0.01 mg/mL leupeptin) on ice for 20 min. The suspension was centrifuged at 13,000 rpm for 3 min at 4°C and the supernatant containing nuclear extracts recovered.

Electrophoretic mobility shift assay

Oligonucleotides used in the synthesis of radiolabeled probe are listed in Table S2. The sense oligonucleotide was labeled with ^{32}P from γ-^{32}P ATP (PerkinElmer) using T4 PNK (New England Biolabs) before the antisense oligonucleotide was annealed by slow cooling from 100°C to room temperature. The probe was purified using a Quick Spin Column (Roche). The nuclear extracts were harvested from COS-7 cells. Empty extract from COS-7 cells was used as a control to check endogenous protein binding. Antibody against V5 (Invitrogen) was used to identify and super-shift of BCL11A ZnF4–6 and mutations. Samples were complexed in gel shift buffer (10 mM HEPES, pH 7.8, 50 mM KCl, 5 mM $MgCl_2$, 1 mM EDTA, 0.05 mg/mL poly(dI-dC), 1 mM DTT, 0.1 mg/mL BSA, and 5% glycerol) with and without V5 antibody and loaded on 6% native polyacrylamide gel in Tris-borate-EDTA buffer (45 mM Tris, 45 mM boric acid, and 1 mM EDTA). Electrophoresis was performed at 4°C for 105 min at 250 V followed by drying under vacuum. The gel was exposed to a FUJIFILM BAS CASSETTE2 2025 phosphor screen overnight and visualized using a GE Typhoon FLA 9500 fluorescent image analyser.

OT analysis

We used the COSMID web tool to predict Cas-dependent DNA OT sites with a stringency of no DNA and RNA bulges. Using the primers listed in Table S2, we amplified the predicted OT sites and carried out 150-bp paired-end Illumina sequencing with the specified adapter sequences. The resulting data were analyzed using the default settings of CRISPRESSOv2.[45] The transcriptome-wide A-to-I (or threonine-to-cysteine [T-to-C]) conversion in the ABE8e base-edited erythroid cells was calculated by REDItools version 2. All of the nucleotides other than A were removed from the analysis with the previously reported coverage and quality criteria.[46] The frequency of A converted to I (or T to C) was calculated by dividing the total number of edited A by the overall counts of A after filtering (A-to-I)/A × 100. The experiment was carried out as 2 biological experiments.

Flow cytometry analysis

The erythroid cells were stained with antibodies for 15 min, washed twice with PBS, and analyzed. The intracellular HbF staining and NBSGW mouse BM cells were processed as previously described[41] in accordance with the Institutional Animal Ethics Committee. All of the antibodies used are listed in Table S4. BD Celesta and CytoFLEX-LX (Beckman Coulter) flow cytometers were used for sample analysis.

Western blot

Whole-cell protein lysates were prepared by resuspending the cell pellet with 1× radioimmunoprecipitation assay buffer along with Halt Protease Inhibitor Cocktail (Thermo Fisher Scientific). Approximately 35 μg protein was used for western blot analysis using the antibodies listed in Table S4.

High-performance liquid chromatography (HPLC)

At the terminal stage of erythroid differentiation, the differentiated cells were sonicated and centrifuged with maximum revolutions per minute at 4°C for 15–20 min to obtain the clear lysate.[47] Analysis of total hemoglobin (TOSOH HPLC variant analyser) was performed using the lysates. The percentage of total globin was calculated as [HbF/(HbF+HbA)] × 100.

In silico modeling

The BCL11A protein-DNA structure obtained from RCSB PDB-6U9Q was used as a reference.[25] We used Pymol software to visualize the amino acids involved in interactions with the DNA bases.

RNA-seq

Total RNA isolated during the second phase of erythroid differentiation (day 8) of edited HSPCs was processed for RNA-seq using Illumina (Novaseq 6000). The raw reads were filtered using Trimmomatic for quality scores and adapters. Filtered reads were aligned to the human genome (hg38) using splice aware aligners such as HISAT2 to quantify reads mapped to each transcript. The alignment percentage of reads ranged between 95.28% and 98.52% for all of the samples. The total number of uniquely mapped reads was counted using feature counts. The uniquely mapped reads were then subjected to differential gene expression using DeSeq2. Real time qRT-PCR was performed to validate the results using primers from Table S2.

Statistical analysis

All of the graphs were generated using GraphPad Prism version 8.1. One-way ANOVA was used in all of the analyses, and the comparison is with the control unless otherwise noted.

DATA AND CODE AVAILABILITY

Next-generation sequencing (NGS) data were deposited to the NCBI Sequence Read Archive using accession no. PRJNA952324. Transcriptome analysis raw files can be accessed through GSE: 229989.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.ymthe.2024.01.023.

ACKNOWLEDGMENTS

All of the human samples used in this study were obtained with informed consent as per clinical guidelines authorized by the institutional review board of Christian Medical College. The mice used for the study were with approval from the Institutional Animal Ethics Committee of Christian Medical College. This work was supported by NAHD grant BT/PR17316/MED/31/326/2015 (Department of Biotechnology [DBT], New Delhi, India), EMR grant EMR/2017/004363 (Science and Engineering Research Board, New Delhi, India), DBT grant BT/PR38392/GET/119/301/2020 and BMS grant 2019-0916/SCR/ADHOC-BMS (Indian Council of Medical Research (ICMR), India), and India Alliance Team Science grant IA/TSG/22/1/600410. M.C. is supported by an Australian National Health and Medical Research Council grant 2020861. M.H. is supported by the Australian Government Research Training Program. We sincerely acknowledge the Centre for Stem Cell Research (CSCR) (a unit of inStem, Christian Medical College [CMC] Campus, Vellore, India) for providing the startup funds. V.R. is supported by Senior Research Fellowship DBT India. N.S.R. and A.G. are supported by the Senior Research Fellowship from Council of Scientific & Industrial Research India. N.D. and K.P. are supported by the Senior Research Fellowship from ICMR. We thank Mr. Ashis Kumar S, CSCR, for his help in R platform-based analysis; Mr. Gopinath and Mr. Neelagandan at the Department of Hematology, CMC, for help with the HPLC variants. Also, we acknowledge the CSCR core facility and CSCR animal facility for supporting us with all of the required instrumentations.

AUTHOR CONTRIBUTIONS

V.R., N.D., and K.M.M. conceived and designed the experiments and wrote the manuscript. V.R. and N.D. performed the experiments and analyzed the data. N.S.R. performed the base editing nucleofection. L.P., J.P., K.A., D.S., A.A.P., and S.G. helped in data collection. G.M., Y.P., and S.M. helped in the mRNA production. J.P.M., C.G., and R.R. helped in the *in silico* modeling and transcriptome data analysis. S.W. and J.E.C. helped in the RNA editing analysis. K.P., A.G., S.T., S.R.V., S.M., P.B., and A.S. helped in revision of the manuscript. A.S. helped in the donor samples acquisition. Y.N. provided the HUDEP2 cells. M.H. and M.C. helped in the DNA-binding assays and manuscript revision. K.M.M. supervised the project and acquired the funding.

DECLARATION OF INTERESTS

All of the authors declare no competing interests.

REFERENCES

1. Pulford, K., Banham, A.H., Lyne, L., Jones, M., Ippolito, G.C., Liu, H., Tucker, P.W., Roncador, G., Lucas, E., Ashe, S., et al. (2006). The BCL11A XL transcription factor: Its distribution in normal and malignant tissues and use as a marker for plasmacytoid dendritic cells. Leukemia *20*, 1439–1441. https://doi.org/10.1038/sj.leu.2404260.

2. Liu, H., Ippolito, G.C., Wall, J.K., Niu, T., Probst, L., Lee, B.S., Pulford, K., Banham, A.H., Stockwin, L., Shaffer, A.L., et al. (2006). Functional studies of BCLIIA: Characterization of the conserved BCLIIA-XL splice variant and its interaction with BCL6 in nuclear paraspeckles of germinal center B cells. Mol. Cancer *5*, 1–16. https://doi.org/10.1186/1476-4598-5-18.

3. Sankaran, V.G., Menne, T.F., Xu, J., Akie, T.E., Lettre, G., Van Handel, B., Mikkola, H.K.A., Hirschhorn, J.N., Cantor, A.B., and Orkin, S.H. (2008). Human fetal hemoglobin expression is regulated by the developmental stage-specific repressor BCL11A. Science *322*, 1839–1842. https://doi.org/10.1126/science.1165409.

4. Xu, J., Sankaran, V.G., Ni, M., Menne, T.F., Puram, R.V., Kim, W., and Orkin, S.H. (2010). Transcriptional silencing of γ-globin by BCL11A involves long-range interactions and cooperation with SOX6. Genes Dev. *24*, 783–798. https://doi.org/10.1101/gad.1897310.

5. Esrick, E.B., Lehmann, L.E., Biffi, A., Achebe, M., Brendel, C., Ciuculescu, M.F., Daley, H., MacKinnon, B., Morris, E., Federico, A., et al. (2021). Post-Transcriptional Genetic Silencing of BCL11A to Treat Sickle Cell Disease. N. Engl. J. Med. *384*, 205–215. https://doi.org/10.1056/nejmoa2029392.

6. Guda, S., Brendel, C., Renella, R., Du, P., Bauer, D.E., Canver, M.C., Grenier, J.K., Grimson, A.W., Kamran, S.C., Thornton, J., et al. (2015). MiRNA-embedded shRNAs for Lineage-specific BCL11A Knockdown and Hemoglobin F Induction. Mol. Ther. *23*, 1465–1474. https://doi.org/10.1038/mt.2015.113.

7. Brendel, C., Guda, S., Renella, R., Bauer, D.E., Canver, M.C., Kim, Y.J., Heeney, M.M., Klatt, D., Fogel, J., Milsom, M.D., et al. (2016). Lineage-specific BCL11A knockdown

circumvents toxicities and reverses sickle phenotype. J. Clin. Invest. *126*, 3868–3878. https://doi.org/10.1172/JCI87885.

8. Bagchi, A., Devaraju, N., Chambayil, K., Rajendiran, V., Venkatesan, V., Sayed, N., Pai, A.A., Nath, A., David, E., Nakamura, Y., et al. (2022). Erythroid lineage-specific lentiviral RNAi vectors suitable for molecular functional studies and therapeutic applications. Sci. Rep. *12*, 14033. https://doi.org/10.1038/s41598-022-13783-0.

9. Uda, M., Galanello, R., Sanna, S., Lettre, G., Sankaran, V.G., Chen, W., Usala, G., Busonero, F., Maschio, A., Albai, G., et al. (2008). Genome-wide Association Study Shows BCL11A Associated with Persistent Fetal Hemoglobin and Amelioration of the Phenotype Of-Thalassemia.

10. Basak, A., Hancarova, M., Ulirsch, J.C., Balci, T.B., Trkova, M., Pelisek, M., Vlckova, M., Muzikova, K., Cermak, J., Trka, J., et al. (2015). BCL11A deletions result in fetal hemoglobin persistence and neurodevelopmental alterations. J. Clin. Invest. *125*, 2363–2368. https://doi.org/10.1172/JCI81163.

11. Funnell, A.P.W., Prontera, P., Ottaviani, V., Piccione, M., Giambona, A., Maggio, A., Ciaffoni, F., Stehling-Sun, S., Marra, M., Masiello, F., et al. (2015). P15-p16.1 microdeletions encompassing and proximal to BCL11A are associated with elevated HbF in addition to neurologic impairment. Blood *126*, 89–93. https://doi.org/10.1182/blood-2015-04-638528.

12. Bauer, D.E., Kamran, S.C., Lessard, S., Xu, J., Fujiwara, Y., Lin, C., Shao, Z., Canver, M.C., Smith, E.C., Pinello, L., et al. (2013). An erythroid enhancer of BCL11A subject to genetic variation determines fetal hemoglobin level. Science *342*, 253–257. https://doi.org/10.1126/science.1242088.

13. Basak, A., Munschauer, M., Lareau, C.A., Montbleau, K.E., Ulirsch, J.C., Hartigan, C.R., Schenone, M., Lian, J., Wang, Y., Huang, Y., et al. (2020). Control of human hemoglobin switching by LIN28B-mediated regulation of BCL11A translation. Nat. Genet. *52*, 138–145. https://doi.org/10.1038/s41588-019-0568-7.

14. Vierstra, J., Reik, A., Chang, K.-H., Stehling-Sun, S., Zhou, Y., Hinkley, S.J., Paschon, D.E., Zhang, L., Psatha, N., Bendana, Y.R., et al. (2015). Functional footprinting of regulatory DNA HHS Public Access. Nat. Methods *12*, 927–930. https://doi.org/10.1038/nmeth.3554.

15. Canver, M.C., Smith, E.C., Sher, F., Pinello, L., Sanjana, N.E., Shalem, O., Chen, D.D., Schupp, P.G., Vinjamur, D.S., Garcia, S.P., et al. (2015). BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis. Nature *527*, 192–197. https://doi.org/10.4315/0362-028X.JFP-13-395.Knowledge.

16. Shen, Y., Li, R., Teichert, K., Montbleau, K.E., Verboon, J.M., Voit, R.A., and Sankaran, V.G. (2021). Pathogenic BCL11A variants provide insights into the mechanisms of human fetal hemoglobin silencing. Plos Genet. *17*, e1009835. https://doi.org/10.1371/journal.pgen.1009835.

17. Wu, Y., Zeng, J., Roscoe, B.P., Liu, P., Yao, Q., Lazzarotto, C.R., Clement, K., Cole, M.A., Luk, K., Baricordi, C., et al. (2019). Highly efficient therapeutic gene editing of human hematopoietic stem cells. Nat. Med. *25*, 776–783. https://doi.org/10.1038/s41591-019-0401-y.

18. Demirci, S., Zeng, J., Wu, Y., Uchida, N., Shen, A.H., Pellin, D., Gamer, J., Yapundich, M., Drysdale, C., Bonanno, J., et al. (2020). BCL11A enhancer–edited hematopoietic stem cells persist in rhesus monkeys without toxicity. J. Clin. Invest. *130*, 6677–6687. https://doi.org/10.1172/JCI140189.

19. Psatha, N., Reik, A., Phelps, S., Zhou, Y., Dalas, D., Yannaki, E., Levasseur, D.N., Urnov, F.D., Holmes, M.C., and Papayannopoulou, T. (2018). Disruption of the BCL11A Erythroid Enhancer Reactivates Fetal Hemoglobin in Erythroid Cells of Patients with β-Thalassemia Major. Mol. Ther. Methods Clin. Dev. *10*, 313–326. https://doi.org/10.1016/j.omtm.2018.08.003.

20. Chang, K.H., Smith, S.E., Sullivan, T., Chen, K., Zhou, Q., West, J.A., Liu, M., Liu, Y., Vieira, B.F., Sun, C., et al. (2017). Long-Term Engraftment and Fetal Globin Induction upon BCL11A Gene Editing in Bone-Marrow-Derived CD34+ Hematopoietic Stem and Progenitor Cells. Mol. Ther. Methods Clin. Dev. *4*, 137–148. https://doi.org/10.1016/j.omtm.2016.12.009.

21. Smith, E.C., Luc, S., Croney, D.M., Woodworth, M.B., Greig, L.C., Fujiwara, Y., Nguyen, M., Sher, F., Macklis, J.D., Bauer, D.E., and Orkin, S.H. (2016). Strict in vivo specificity of the Bcl11a erythroid enhancer. Blood *128*, 2338–2342. https://doi.org/10.1182/blood-2016-08-736249.

22. Ravi, N.S., Wienert, B., Wyman, S.K., Bell, H.W., George, A., Mahalingam, G., Vu, J.T., Prasad, K., Bandlamudi, B.P., Devaraju, N., et al. (2022). Identification of novel HPFH-like mutations by CRISPR base editing that elevate the expression of fetal hemoglobin. Elife *11*, e65421. https://doi.org/10.7554/eLife.65421.

23. Martyn, G.E., Wienert, B., Kurita, R., Nakamura, Y., Quinlan, K.G.R., and Crossley, M. (2019). A natural regulatory mutation in the proximal promoter elevates fetal globin expression by creating a de novo GATA1 site. Blood *133*, 852–856. https://doi.org/10.1182/blood-2018-07-863951.

24. Zittersteijn, H.A., Harteveld, C.L., Klaver-Flores, S., Lankester, A.C., Hoeben, R.C., Staal, F.J.T., and Gonçalves, M.A.F.V. (2020). A Small Key for a Heavy Door: Genetic Therapies for the Treatment of Hemoglobinopathies. Front. Genome Ed. *2*, 617780. https://doi.org/10.3389/fgeed.2020.617780.

25. Yang, Y., Xu, Z., He, C., Zhang, B., Shi, Y., and Li, F. (2019). Structural insights into the recognition of γ-globin gene promoter by BCL11A. Cell Res. *29*, 960–963. https://doi.org/10.1038/s41422-019-0221-0.

26. Cai, T., Chen, X., Li, J., Xiang, B., Yang, L., Liu, Y., Chen, Q., He, Z., Sun, K., and Liu, P.P. (2017). Identification of novel mutations in the HbF repressor gene BCL11A in patients with autism and intelligence disabilities. Am. J. Hematol. *92*, E653–E656. https://doi.org/10.1002/ajh.24902.

27. Abdulazeez, S., Sultana, S., Almandil, N.B., Almohazey, D., Bency, B.J., and Borgio, J.F. (2019). The rs61742690 (S783N) single nucleotide polymorphism is a suitable target for disrupting BCL11A-mediated foetal-to-adult globin switching. PLoS One *14*, e02124922. https://doi.org/10.1371/journal.pone.0212492.

28. Liu, N., Hargreaves, V.V., Zhu, Q., Kurland, J.V., Hong, J., Kim, W., Sher, F., Macias-Trevino, C., Rogers, J.M., Kurita, R., et al. (2018). Direct Promoter Repression by BCL11A Controls the Fetal to Adult Hemoglobin Switch. Cell *173*, 430–442.e17. https://doi.org/10.1016/j.cell.2018.03.016.

29. Shen, Y., Verboon, J.M., Zhang, Y., Liu, N., Kim, Y.J., Marglous, S., Nandakumar, S.K., Voit, R.A., Fiorini, C., Ejaz, A., et al. (2021). A unified model of human hemoglobin switching through single-cell genome editing. Nat. Commun. *12*, 4991–5012. https://doi.org/10.1038/s41467-021-25298-9.

30. Luc, S., Huang, J., McEldoon, J.L., Somuncular, E., Li, D., Rhodes, C., Mamoor, S., Hou, S., Xu, J., and Orkin, S.H. (2016). Bcl11a Deficiency Leads to Hematopoietic Stem Cell Defects with an Aging-like Phenotype. Cell Rep. *16*, 3181–3194. https://doi.org/10.1016/j.celrep.2016.08.064.

31. Liu, P., Keller, J.R., Ortiz, M., Tessarollo, L., Rachel, R.A., Nakamura, T., Jenkins, N.A., and Copeland, N.G. (2003). Bcl11a is essential for normal lymphoid development. Nat. Immunol. *4*, 525–532. https://doi.org/10.1038/ni925.

32. Yu, Y., Wang, J., Khaled, W., Burke, S., Li, P., Chen, X., Yang, W., Jenkins, N.A., Copeland, N.G., Zhang, S., and Liu, P. (2012). Bcl11a is essential for lymphoid development and negatively regulates p53. J. Exp. Med. *209*, 2467–2483. https://doi.org/10.1084/jem.20121846.

33. Tsang, J.C.H., Yu, Y., Burke, S., Buettner, F., Wang, C., Kolodziejczyk, A.A., Teichmann, S.A., Lu, L., and Liu, P. (2015). Single-cell transcriptomic reconstruction reveals cell cycle and multi-lineage differentiation defects in Bcl11a-deficient hematopoietic stem cells. Genome Biol. *16*, 178–216. https://doi.org/10.1186/s13059-015-0739-5.

34. Basak, A., Munschauer, M., Lareau, C.A., Montbleau, K.E., Ulirsch, J.C., Hartigan, C.R., Schenone, M., Lian, J., Wang, Y., Huang, Y., et al. (2020). Control of human hemoglobin switching by LIN28B-mediated regulation of BCL11A translation. Nat. Genet. *52*, 138–145. https://doi.org/10.1038/s41588-019-0568-7.

35. Shen, F., Zheng, G., Setegne, M., Tenglin, K., Izadi, M., Xie, H., Zhai, L., Orkin, S.H., and Dassama, L.M.K. (2022). A Cell-Permeant Nanobody-Based Degrader That Induces Fetal Hemoglobin. ACS Cent. Sci. *8*, 1695–1703. https://doi.org/10.1021/acscentsci.2c00998.

36. Mehta, S., Buyanbat, A., Kai, Y., Karayel, O., Goldman, S.R., Seruggia, D., Zhang, K., Fujiwara, Y., Donovan, K.A., Zhu, Q., et al. (2022). Temporal resolution of gene derepression and proteome changes upon PROTAC-mediated degradation of BCL11A protein in erythroid cells. Cell Chem. Biol. *29*, 1273–1287.e8. https://doi.org/10.1016/j.chembiol.2022.06.007.

37. Yin, M., Izadi, M., Tenglin, K., Viennet, T., Zhai, L., Zheng, G., Arthanari, H., Dassama, L.M.K., and Orkin, S.H. (2023). Evolution of nanobodies specific for BCL11A. Proc. Natl. Acad. Sci. USA *120*, e2218959120. https://doi.org/10.1073/pnas.2218959120.

38. Ghedira, E.S., Lecerf, L., Faubert, E., Costes, B., Moradkhani, K., Bachir, D., Galactéros, F., and Pissard, S. (2013). Estimation of the difference in HbF expression due to loss of the 5' δ-globin BCL11A binding region. Haematologica *98*, 305–308. https://doi.org/10.3324/haematol.2012.061994.

39. Tang, X., Sun, P., Wang, H., Cao, J., Xing, J., Sheng, X., Chi, H., and Zhan, W. (2022). Zinc Finger Protein BCL11A Contributes to the Abortive Infection of Hirame novirhabdovirus (HIRRV) in B Lymphocytes of Flounder (Paralichthys olivaceus). J. Virol. *96*, e0147022. https://doi.org/10.1128/jvi.01470-22.

40. Han, Y., Tan, X., Jin, T., Zhao, S., Hu, L., Zhang, W., Kurita, R., Nakamura, Y., Liu, J., Li, D., et al. (2022). CRISPR/Cas9-based multiplex genome editing of BCL11A and HBG efficiently induces fetal hemoglobin expression. Eur. J. Pharmacol. *918*, 174788. https://doi.org/10.1016/J.EJPHAR.2022.174788.

41. Devaraju, N., Rajendiran, V., Ravi, N.S., and Mohankumar, K.M. (2022). Genome Engineering of Hematopoietic Stem Cells Using CRISPR/Cas9 System. Methods Mol. Biol. *2429*, 307–331. https://doi.org/10.1007/978-1-0716-1979-7_20.

42. Richter, M.F., Zhao, K.T., Eton, E., Lapinaite, A., Newby, G.A., Thuronyi, B.W., Wilson, C., Koblan, L.W., Zeng, J., Bauer, D.E., et al. (2020). Phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity. Nat. Biotechnol. *38*, 883–891. https://doi.org/10.1038/s41587-020-0453-z.

43. Conant, D., Hsiau, T., Rossi, N., Oki, J., Maures, T., Waite, K., Yang, J., Joshi, S., Kelso, R., Holden, K., et al. (2022). Inference of CRISPR Edits from Sanger Trace Data. CRISPR J. *5*, 123–130. https://doi.org/10.1089/crispr.2021.0113.

44. Kluesner, M.G., Nedveck, D.A., Lahr, W.S., Garbe, J.R., Abrahante, J.E., Webber, B.R., and Moriarity, B.S. (2018). EditR: A Method to Quantify Base Editing from Sanger Sequencing. CRISPR J. *1*, 239–250. https://doi.org/10.1089/crispr.2018.0014.

45. Clement, K., Rees, H., Canver, M.C., Gehrke, J.M., Farouni, R., Hsu, J.Y., Cole, M.A., Liu, D.R., Joung, J.K., Bauer, D.E., and Pinello, L. (2019). CRISPResso2 provides accurate and rapid genome editing sequence analysis. Nat. Biotechnol. *37*, 224–226. https://doi.org/10.1038/s41587-019-0032-3.

46. Koblan, L.W., Doman, J.L., Wilson, C., Levy, J.M., Maianti, J.P., Tay, T., Newby, G.A., Raguram, A., and Liu, D.R. (2018). Improving Cytidine and Adenine Base Editors by Expression Optimization and Ancestral Reconstruction (Nature Publishing Group). https://doi.org/10.1038/nbt.4172.

47. Bagchi, A., Nath, A., Thamodaran, V., Ijee, S., Palani, D., Rajendiran, V., Venkatesan, V., Datari, P., Pai, A.A., Janet, N.B., et al. (2021). Direct generation of immortalized erythroid progenitor cell lines from peripheral blood mononuclear cells. Cells *10*, 523. https://doi.org/10.3390/cells10030523.

YMTHE, Volume *32*

Supplemental Information

Base editing of key residues in the BCL11A-XL-specific zinc finger domains derepresses fetal globin expression

Vignesh Rajendiran, Nivedhitha Devaraju, Mahdi Haddad, Nithin Sam Ravi, Lokesh Panigrahi, Joshua Paul, Chandrasekar Gopalakrishnan, Stacia Wyman, Keerthiga Ariudainambi, Gokulnath Mahalingam, Yogapriya Periyasami, Kirti Prasad, Anila George, Dhiyaneshwaran Sukumaran, Sandhiya Gopinathan, Aswin Anand Pai, Yukio Nakamura, Poonkuzhali Balasubramanian, Rajasekaran Ramalingam, Saravanabhavan Thangavel, Shaji R. Velayudhan, Jacob E. Corn, Joel P. Mackay, Srujan Marepally, Alok Srivastava, Merlin Crossley, and Kumarasamypet M. Mohankumar

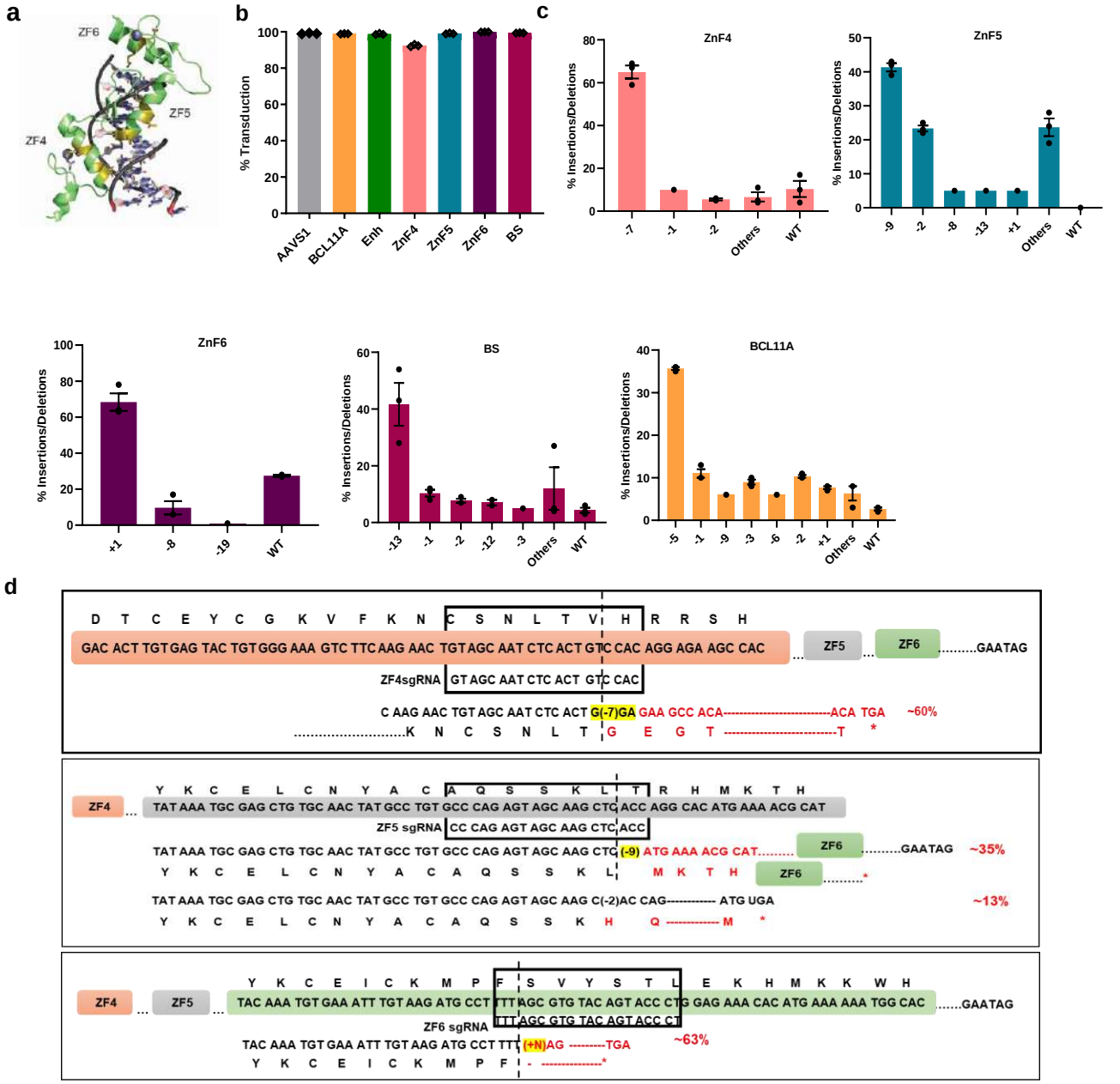

a
ZF6
ZF5
ZF4
b
% Transduction
AAVS1
BCL11A
Enh
ZnF4
ZnF5
ZnF6
BS
c
ZnF4
% Insertions/Deletions
-7
-1
-2
Others
WT
ZnF5
-9
-2
-8
-13
+1
Others
WT
ZnF6
+1
-8
-19
WT
BS
-13
-1
-2
-12
-3
Others
WT
BCL11A
-5
-1
-9
-3
-6
-2
+1
Others
WT
d
D T C E Y C G K V F K N C S N L T V H R R S H
GAC ACT TGT GAG TAC TGT GGG AAA GTC TTC AAG AAC TGT AGC AAT CTC ACT GTC CAC AGG AGA AGC CAC
ZF5
ZF6
.....GAATAG
ZF4sgRNA GT AGC AAT CTC ACT GTC CAC
C AAG AAC TGT AGC AAT CTC ACT G(-7)GA GAA GCC ACA----------ACA TGA
~60%
.....K N C S N L T G E G T ----------T *
ZF4
Y K C E L C N Y A C A Q S S K L T R H M K T H
TAT AAA TGC GAG CTG TGC AAC TAT GCC TGT GCC CAG AGT AGC AAG CTC ACC AGG CAC ATG AAA ACG CAT
ZF5 sgRNA CC CAG AGT AGC AAG CTC ACC
TAT AAA TGC GAG CTG TGC AAC TAT GCC TGT GCC CAG AGT AGC AAG CTC (-9) ATG AAA ACG CAT.....
ZF6
.....GAATAG
~35%
Y K C E L C N Y A C A Q S S K L M K T H
ZF6
.....*
TAT AAA TGC GAG CTG TGC AAC TAT GCC TGT GCC CAG AGT AGC AAG C(-2)AC CAG---------- ATG UGA
~13%
Y K C E L C N Y A C A Q S S K H Q ---------- M *
Y K C E I C K M P F S V Y S T L E K H M K K W H
ZF4
ZF5
TAC AAA TGT GAA ATT TGT AAG ATG CCT TTT AGC GTG TAC AGT ACC CTG GAG AAA CAC ATG AAA AAA TGG CAC
.....GAATAG
ZF6 sgRNA TTT AGC GTG TAC AGT ACC CT
TAC AAA TGT GAA ATT TGT AAG ATG CCT TTT (+N) AG ---------TGA
~63%
Y K C E I C K M P F - ----------*
e
% of CD71/C235a expressing cells
CD71-CD235a-
CD71+CD235a-
CD71+CD235a+
CD71-CD235a+
AAVS1
BCL11A
Enh
ZnF4
ZnF5
ZnF6
BS
f
Gg
Ag
p<0.0001
%Gamma/ Beta-like globin
AAVS1
BCL11A
Enh
ZF4
ZF5
ZF6
BS
g
Total Cells
AAVS1
BCL11A
ZnF4
ZnF5
ZnF6
BS
D0
D4
D7
D10
D13

Figure S1:Targeted disruption of BCL11A-XL specific Zinc finger domains using Cas9 to upregulate fetal globin expression in adult erythroid cells

a. Pictorial representation of the X-ray crystal structure of the BCL11A ZnF(4-6) bound to its canonical binding site at HBG1/2 promoter (PDB:6U9Q).

b. The transduction efficiency of each gRNA measured in flow cytometry in HUDEP2 cells expressing cas9. Mean ± SEM of n=3 independent experiments.

c. The frequency of multiple insertion or deletion patterns generated at each target site;

d. An illustration depicting the insertion or deletion with the highest frequency that could result in the generation of BCL11A ZnF variants at ZnF 4, 5, and 6;

e. Erythroid differentiation profile analyzed by the expression of CD235a and CD71 marker by flowcytometry in BCL11A ZnF edited HUDEP2 cells at the day 8 of erythroid differentiation;

f. Analysis of globin chain production through reverse-phase HPLC at the terminal stage of erythroid differentiation (Day 8); Mean ± SEM of n=3 independent experiments.

g. The cell proliferation profile of all the gene-modified HSPCs after electroporation during the three-phase erythroid differentiation.

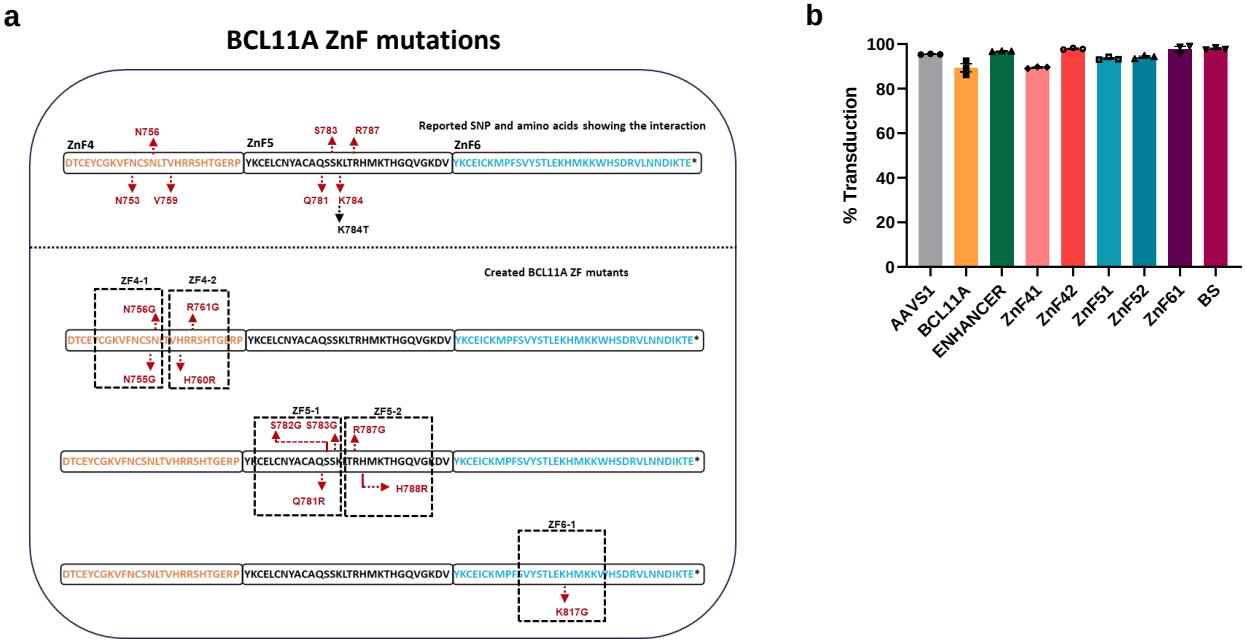
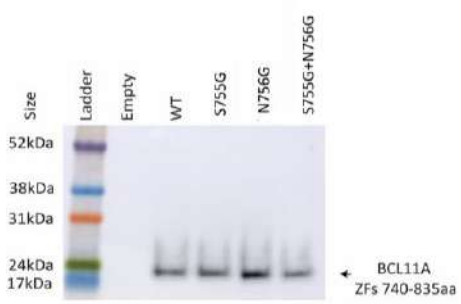


Figure S2: The key base substitutions in BCL11A-XL specific ZnF finger domains reactivate HbF expression in adult erythroid cells.

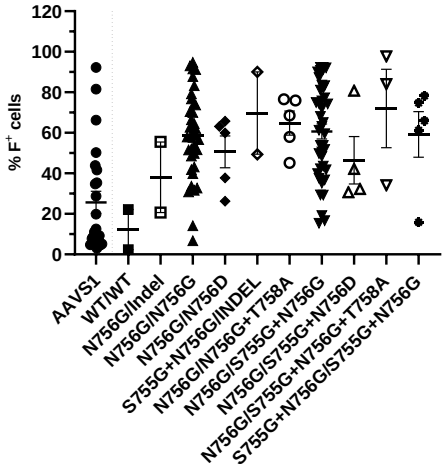
a. Pictorial representation of reported SNPs(black) and amino acids involved in interaction(red) present within the ZnF domains. The amino acid alteration within the BCL11A ZnF created using the respective sgRNAs in the study.

b. The transduction efficiency of each gRNA measured in flowcytometry as GFP percentage in HUDEP2 cells stably expressing ABE8e editor. Mean ± SEM of n=3 independent experiments.

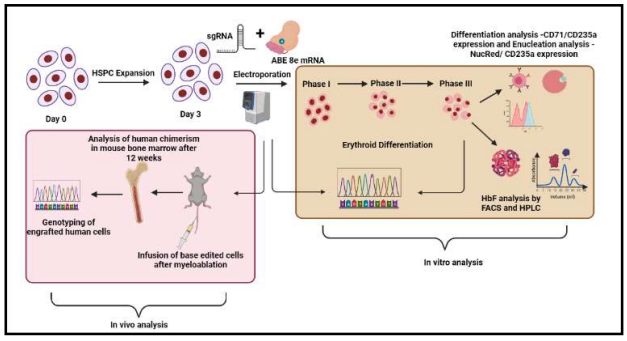
a


Size
Ladder
Empty
WT
S755G
N756G
S755G+N756G
52kDa
38kDa
31kDa
24kDa
17kDa
BCL11A
ZFs 740-835aa

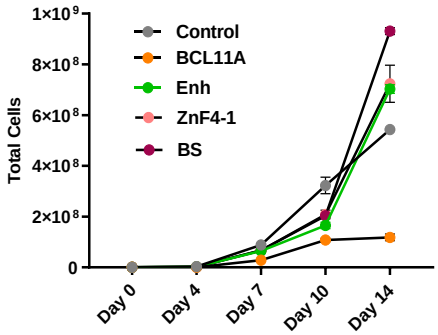
b


% F+ cells
120
100
80
60
40
20
0
AAVS1
WT/WT
N756G/Indel
N756G/N756G
N756G/N756D
N756G/N756G/INDEL
S755G+N756G+T758A
N756G/N756G+T758A
N756G/S755G+N756G
N756G/S755G+N756D
N756G/S755G+N756G+T758A
N756G/S755G+N756G
S755G+N756G/S755G+N756G

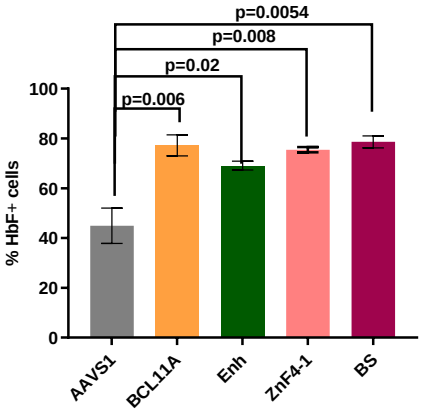
c


sgRNA
ABE 8e mRNA
HSPC Expansion
Electroporation
Day 0
Day 3
Differentiation analysis -CD71/CD235a expression and Enucleation analysis - NucRed/ CD235a expression
Phase I
Phase II
Phase III
Erythroid Differentiation
HbF analysis by FACS and HPLC
Analysis of human chimerism in mouse bone marrow after 12 weeks
Genotyping of engrafted human cells
Infusion of base edited cells after myeloablation
In vivo analysis
In vitro analysis

d


Total Cells
1×10⁹
8×10⁸
6×10⁸
4×10⁸
2×10⁸
0
Control
BCL11A
Enh
ZnF4-1
BS
Day 0
Day 4
Day 7
Day 10
Day 14

e


% HbF+ cells
p=0.0054
p=0.008
p=0.02
p=0.006
100
80
60
40
20
0
AAVS1
BCL11A
Enh
ZnF4-1
BS

f

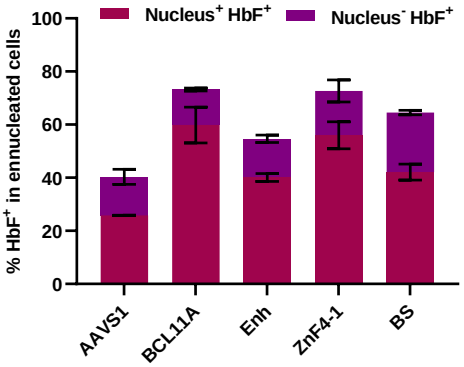

% HbF+ in ennucleated cells
Nucleus+ HbF+
Nucleus- HbF+
100
80
60
40
20
0
AAVS1
BCL11A
Enh
ZnF4-1
BS

Figure S3: Generation of BCL11A ZnF mutant in HSPCs exhibit elevated levels of HbF on erythroid differentiation.

a. Western blot showing equivalent amounts of recombinant V5 tagged BCL11A ZF4-6 proteins were used in gel shift assays. The Western blot was carried out with the same amount of COS-7 nuclear extract used in the gel shift experiment in (A) using an antibody against V5 and standard methods.[24]

b. Percentage of HbF^+ cells exhibited in the erythroid cells exhibiting various genotypes due to ZnF4-1 editing in HUDEP2 cells obtained through single cell analysis followed by 8 days of erythroid differentiation. Single cell clones of AAVS1 edited HUDEP2 cells were also analyzed separately for HbF^+ cell after differentiation.

c. A pictorial illustration of the experimental workflow for the creation of BCL11A ZnF4-1 mutant in healthy donor-derived HSPCs and its downstream analysis

d. The cell proliferation curve of the AAVS1(control),BCL11A total loss(BCL11A), Enhancer(Enh), Binding Site(BS) and the ZnF4(ZnF4-1) base edited HSPCs during erythroid differentiation.

e. The percentage of erythroid cells expressing HbF measured by FACS at the terminal day of differentiation (day 21). Mean ± SEM of n=2 independent experiments in one donor. All samples are compared with the AAVS1 control with one-way ANOVA analysis.

f. The measure of HbF expression in two different populations ($Nucleus^+HbF^+$ and $Nucleus^-HbF^+$) in differentiated erythroid cells.

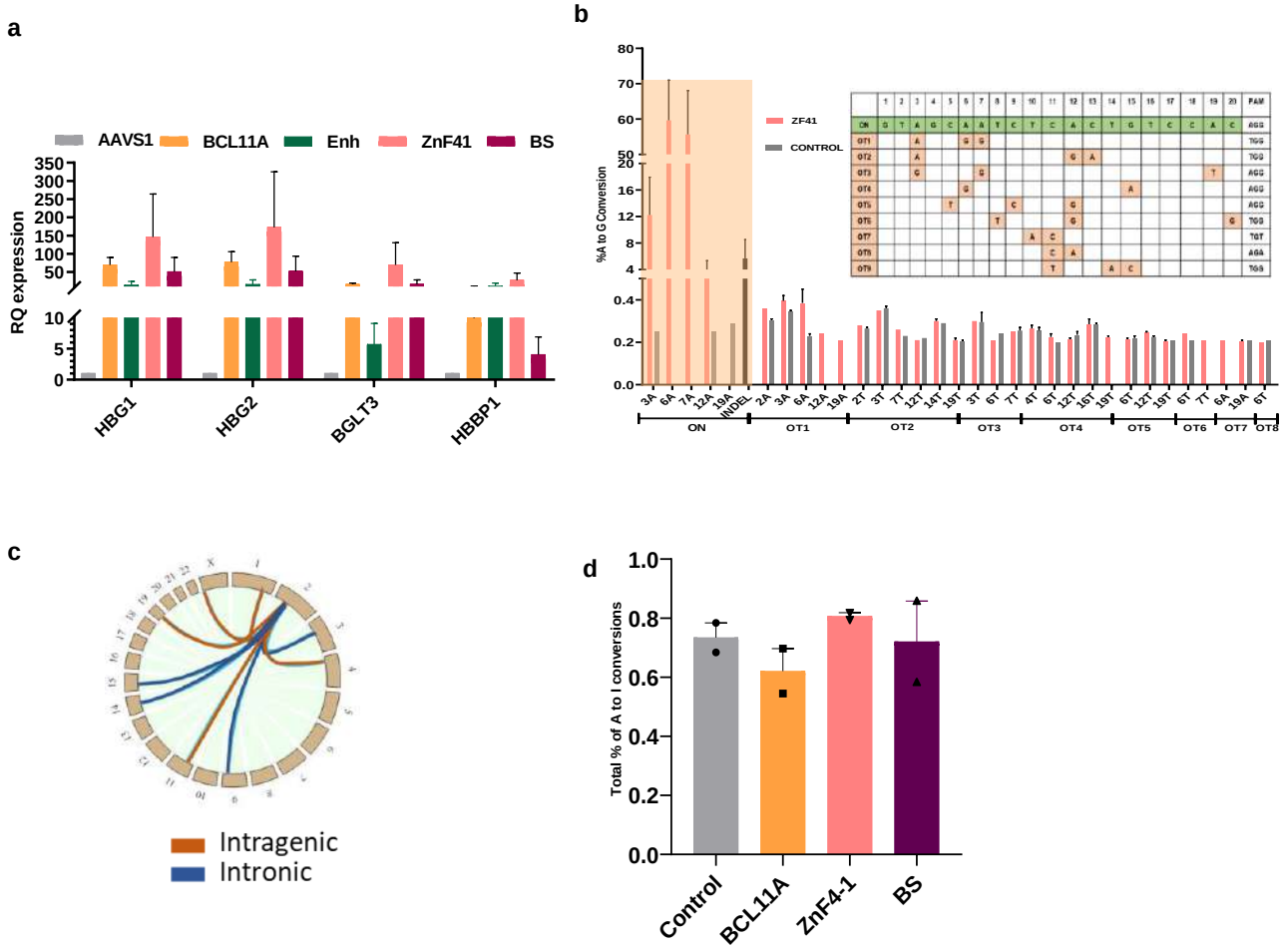


Figure S4: Global transcriptomic analysis of ZnF4-1 base edited erythroid cells derived from CD34⁺ HSPCs

a. *HBG1*, *HBG2* and *BGLT3* and *HBBP1* upregulation validated by real time quantitative reverse transcriptase PCR. Each sample performed in biological duplicates and each replicate as technical triplicates

b. Table displaying all possible predicted off-target sites based on the mismatches in the ZnF4-1 sgRNA; The percentage of on-target editing compared to all the predicted off-targets measured by targeted amplicon NGS (Next Generation Sequencing) followed by CRISPResso2 analysis.

c. Circos plot depicting the chromosomal position of each off-target site in the human genome and the nature of the site. Red denotes Intragenic and Blue denotes Intronic regions.

d. Total percentage of A-to-I (T-to-C) conversion at the RNA level in all the edited samples. Performed as biological duplicates.

Table S1: The sgRNAs used in this study

Guide RNA	sgRNA seq 5'-> 3'
AAVS1	GGGGCCACTAGGGACAGGAT
CCR5	AAAATATAATCTTTAAGATA
BCL11A In Cd (ABE8e)	AGACATGGTGGGCTGCGGGG
BCL11A exon2 (cas9)	TGAACCAGACCACGGCCCGT
Enhancer(cas9)	CTAACAGTTGCTTTTATCAC
Enhancer(ABE8e)	TTTATCACAGGCTCCAGGAA
ZnF4/41	GTAGCAATCTCACTGTCCAC
ZnF42	TGTCCACAGGAGAAGCCACA
ZnF5/51	CCCAGAGTAGCAAGCTCACC
ZnF52	ACCAGGCACATGAAAACGCA
ZnF6/61	TGGAGAAACACATGAAAAAA
ZnF62	AGGGTACTGTACACGCTAAA
BS(cas9)	CTTGTCAAGGCTATTGGTCA
BS(ABE8e)	CTTGACCAATAGCCTTGACA

Table S2: Primers and Probes used in this study

	Sample	Fwd primer	Rev primer
Sanger Primers	AAVS1	GAGATGGCTCCAGGAAATGG	ACCTCTCACTCCTTTCATTTGG
	Initiation CD	CATCTTCCCTGCGCCATCTTTG	CTCTCTCCCCCTCGCTTTTG
	Exon2	ATGGGGTTGAGATGTGCTTC	ACTGCTTGGCTACAGCACCT
	BCL11A enhancer	TCAAACCACAGGGATCACAA	AGAGAGCCTTCCGAAAGAGG
	ZnF4/5	CCCTTCTCTAAGCGCATCAA	TCCAGGGTACTGTACACGCTAA
	ZnF6	GCCCAGAGTAGCAAGCTCAC	ATCCTACAGGGAGTGGGGCT
	BS	ACAAAAGAAGTCCTGGTATC	CTTCCCAGGGTTTCTCCTCC
NGS Primers	ZnF4/5	TACACGACGCTCTTCCGATCTTGGAGGGAGCACGCCCCATAT	AGACGTGTGCTCTTCCGATCTCTCCAGGGTACTGTACACGC
	OT1	TACACGACGCTCTTCCGATCTTCATACCATGGACAAAGACATTCCTAAGGG	AGACGTGTGCTCTTCCGATCTACTAGGCTTCTACCATCCCTTGC
	OT2	TACACGACGCTCTTCCGATCTGCAGCTGGAGTACCTGGTCT	AGACGTGTGCTCTTCCGATCTGAAGCAAAGTCCCATTCCCTCC
	OT3	TACACGACGCTCTTCCGATCTTGAAGAAAGAATGGCTCAGAGGGC	AGACGTGTGCTCTTCCGATCTAGGGTTATGTTCTTTCTGGACCATCTCCTT
	OT4	TACACGACGCTCTTCCGATCTAGGGTTGCAGATCCCACCAG	AGACGTGTGCTCTTCCGATCTTGTAGACTCTGACGCACCCC
	OT5	TACACGACGCTCTTCCGATCTCCCAGTGAAATGTGTAAGTGAATTCTGTTCC	AGACGTGTGCTCTTCCGATCTGGACTTTCAGCCCTTTGCCTTG
	OT6	TACACGACGCTCTTCCGATCTCAAGAACTCAGTTTTCAAGGTTACTCTGGG	AGACGTGTGCTCTTCCGATCTGGATTGTGCAAGGTACATGCATTTAATATTAGATTGG
	OT7	TACACGACGCTCTTCCGATCTGCCTATCACTACCCAACAAGAAGTATATCAC	AGACGTGTGCTCTTCCGATCTCTGTAAGTTTCCTGAGGCCTCC
	OT8	TACACGACGCTCTTCCGATCTTGGGCTTCGTCTTTAGGATATGGG	AGACGTGTGCTCTTCCGATCTATCCTTCAGGACATATACCCCTGG
qRT-PCR Primers	BCL11a	AACCCCAGCACTTAAGCAAA	GGAGGTCATGATCCCCTTCT
	HBB	ACCTTTGCCACACTGAGTGAG	TTTGCCAAAGTGATGGGCCA
	HBG1	Qiagen primers. Cat No: QT00041384	
	HBG2	Qiagen primers. Cat No: QT00040068	
	BGLT3	AAGATAATCTTGGTTTTGCCTCAAA	TCTACTTGATATAGTTGAGAGGCAGTTACC
	HBBP1	TTATGCTCACGGATGACCTCAAAG	AACAATCAATATCACGTTGCCTAAGAG
	GAPDH	CTGCACCACCAACTGCTTAG	GTCTTCTGGGTGGCAGTGAT
Gel Shift Probe	-128 to -100 bp fetal globin promoter (canonical binding motif underlined)	AGCCTTGCCTTGACCAATAGCCTTGACAA	TTGTCAAGGCTATTGGTCAAGGCAAGGCT

Table S3: Mammalian Expression plasmids used in this study

Mammalian Expression Plasmids	Species	Source
pcDNA3-Empty	N/A	Invitrogen
pcDNA3-FLAG BCL11A ZF (4-6) 740-835aa-V5	Human	Ref.[24]
pcDNA3-FLAG BCL11A ZF (4-6) 740-835aa 755S>G-V5	Human	Mahdi Haddad
pcDNA3-FLAG BCL11A ZF (4-6) 740-835aa 756N>G-V5	Human	Mahdi Haddad
pcDNA3-FLAG BCL11A ZF (4-6) 740-835aa 755S>G, 756N>G -V5	Human	Mahdi Haddad

Table S4: Antibodies used in this study

Antibodies	Company	Cat No
CD71	BD	562995 and 555536
CD235a	BD	551336 and 240947
HbF APC	THERMO-Invitrogen	MHFH05
NucRed	Thermo Fischer Scientific	R37106
BCL11a (D4E3P)	Cell Signalling Technology	75432
β-Actin	MP Biological	AS003
HRP Goat antimouse	Abclonal	A6154
Anti-Rabbit IgG	Sigma	69001
CD45 human	BD	555482
CD45 Mouse	BD	559864
CD34	BD	348057
CD13	BD	555394
CD33	BD	340474
CD16	BD	347617
CD19	BD	340364

Chapter 20

Genome Engineering of Hematopoietic Stem Cells Using CRISPR/Cas9 System

Nivedhitha Devaraju, Vignesh Rajendiran, Nithin Sam Ravi, and Kumarasamypet M. Mohankumar

Abstract

Ex vivo genetic manipulation of autologous hematopoietic stem and progenitor cells (HSPCs) is a viable strategy for the treatment of hematologic and primary immune disorders. Targeted genome editing of HSPCs using the CRISPR-Cas9 system provides an effective platform to edit the desired genomic locus for therapeutic purposes with minimal off-target effects. In this chapter, we describe the detailed methodology for the CRISPR-Cas9 mediated gene knockout, deletion, addition, and correction in human HSPCs by viral and nonviral approaches. We also present a comprehensive protocol for the analysis of genome modified HSPCs toward the erythroid and megakaryocyte lineage in vitro and the long-term multilineage reconstitution capacity in the recently developed NBSGW mouse model that supports human erythropoiesis.

Key words HSPCs, Nucleofection, Transduction, Erythroid differentiation, Hemoglobinopathies, Megakaryopoiesis, NHEJ, HDR, Base editing, Clonal analysis, NBSGW mouse

1 Introduction

Hematopoietic stem and progenitor cell (HSPC) transplantation is the current primary treatment strategy for many genetic disorders of blood and immunologic origin [1]. The significant limitations of finding a matched-unrelated donor (MUD) for allogeneic transplantation have shifted the focus towards autogenic ex vivo gene therapy as an option for the treatment of monogenic disorders. HSPCs are readily isolated from mobilized peripheral blood from the donor after G-CSF (granulocyte-colony stimulating factor) stimulation and are able to tolerate gene modification in ex vivo conditions before reinfusion into the patients [2]. Most of the current gene therapy studies in HSPCs are performed using lentiviral vectors (LV), as these vectors are most efficient in transfecting the transgene into recipient cells. Gene therapy based on viral vectors can be effectively used to improve a spectrum of disorders

Nagarajan Kannan and Philip Beer (eds.), *Stem Cell Assays: Methods and Protocols*, Methods in Molecular Biology, vol. 2429, https://doi.org/10.1007/978-1-0716-1979-7_20,

[3]; however, insertional mutagenesis with viral vectors leading to oncogenesis is a major hurdle for translating this approach into the clinic.

Genome editing technologies based on programmable nucleases are emerging as a potential alternative to viral-mediated gene modification. Programmable nucleases such as meganucleases, zinc finger nucleases (ZFN), transcription activator-like effector nucleases (TALENS), and clustered regularly interspaced short palindromic repeat (CRISPR)-associated nuclease Cas9 have enabled the possibility of site-specific genome editing by correction of the disease-causing mutation, the addition of transgenes, or the disruption of regulatory regions. Double-strand DNA breaks caused by gene modification are typically processed by the endogenous DNA repair pathways like nonhomologous end-joining (NHEJ) and homology directed repair (HDR) [4]. NHEJ is the predominant repair pathway occurring throughout the cell cycle phase, whereas the HDR pathway is used in G2 and S phase preferentially.

The target specific endonuclease activity of the cas9 sgRNA (RNP) complex results in double stranded DNA breaks which activate NHEJ, resulting in gene disruption and deletion. The modular nature of the CRISPR system and the ease of designing RNA sequences over the protein sequences underline the popularity of CRISPR over other gene-editing tools. The major limiting factor of protospacer adjacent motif (PAM) availability is addressed by the newer variants of cas protein which recognizes DNA sequences with any specific PAM. The efficiency of the cas9 protein in creating double stranded DNA breaks is determined by the efficiency of the protein to localize into the nucleus. The function is enhanced by coupling a nuclear localizing signal (NLS) gene along with the cas9 gene. Fusion of 3×NLS to the cas9 protein is the commonly used to achieve higher efficiency. In this chapter, we describe the optimized protocol to obtain higher editing efficiency in the primary cells with the 1×NLS-Cas9 protein.

Gene correction and addition are achieved through the HDR pathway, which requires a donor DNA cassette containing homology arms to act as a template for the repair mechanism during the double-strand break. Donor template size, mode of delivery, and the dependence on the cell cycle phase are the major limiting factor for the efficiency of HDR. Introduction or correction of point mutation is mostly preferred by single-stranded oligo donor (ssODN) delivery (with asymmetric homology arms [5]) even though recent studies report higher efficiency of recombinant adeno-associated virus serotype 6 (rAAV6) [6]. Large gene delivery (up to 4 kb), with or without its promoter, in the intronic or promoter regions of the target gene is facilitated by the rAAV6 donor delivery system [7].

The fusion of cas9 nickase with specific deaminase enzyme is employed in base editing to create or correct certain transition mutations like C*G to T*A or A*T to G*C. Cytosine base editor (CBE) converts cytosine to thymine base, and the adenosine base editor (ABE) converts adenosine to guanine in the target region [8]. The major advantages of base editing over HDR are the independence of cell cycle phase and higher desired editing with lower off-target activity (due to single-strand nick instead of DSB). Base editing is limited by the transversions, bystander effect, editing window, and PAM availability which has been overcome by the recently developed base editor variants.

The delivery of the editing tools is achieved by viral (lentivirus and AAV), nonviral (electroporation), or by a hybrid system involving both (viral and electroporation) into HSPCs. Each delivery method has advantages (and disadvantages) which need to be considered cautiously based on the desired editing requirement to achieve high gene editing efficiency. Various delivery strategies and their applications are listed in Table 1, along with the optimal off-target analysis suitable for each method and its application.

This chapter provides different strategies for efficient genome engineering of human HSPCs for therapeutic purposes. The possible effects of targeted genome engineering in HSPCs are evaluated by multiple functional assays during in vitro differentiation (myeloid lineage-erythroid/megakaryocyte) in pooled cell population and at the single-cell level to understand the biology of the gene modification in the individual cells.

Further, we also elucidate the methodology to determine the long-term engraftment and multilineage repopulation efficiency of genome engineered HSPCs in NBSGW mouse model which exhibits higher human chimerism and supports better erythropoiesis of engrafted gene-modified cells than NSG mice [13].

2 Materials

2.1 Equipment

1. Laminar airflow.
2. Magnetic separator.
3. Cell strainer (40 μ).
4. FACS tubes.
5. FACS analyzer.
6. Lonza 4D nucleofector.
7. 16-well nucleofection strips.
8. Maxcyte electroporator.
9. OC25 × 3 MaxCyte cuvette and buffer.
10. Swinging bucket centrifuge.

Table 1
Different genome engineering strategies in HSPCs and their applications

Scheme	Approach	Delivery Strategy	Pathway	Application	Efficiency	Off-target analysis
Gene disruption	Cas9-sgRNA	Electroporation	NHEJ	Knockout of gene of interest with high efficiency and specificity. Usually preferred for exons, promoter regions, and enhancers	Very high efficiency	Guide seq [9]
Gene deletion	Cas9-sgRNA(s)	Electroporation	NHEJ	Deletion of gene or gene sequences	Efficiency depends on the deletion fragment size	Guide Seq Chromosome karyotyping
Gene correction/ addition	Cas9-sgRNA+ ssODN	Electroporation	HDR	• Point mutation insertion, correction up to 200bases • Usually preferred at coding/ regulatory regions	Low efficiency	Guide seq
	Cas9-sgRNA+ rAAV6	Electroporation and viral transduction	HDR	Point mutations insertions and also to insert any gene of interest up to 4 kb	Low efficiency	Guide seq
Base editing	ABE-nCas9+sgRNA	Lentiviral		Conversion of A*T to G*C in regulatory and CDS regions Usually preferred at the coding regions to introduce a desired mutation	High efficiency	Orthogonal R-loop assay [10], RNA-seq, CIRCLE-seq [11] LAM-PCR [12]
		mRNA			High efficiency	Orthogonal R-loop assay RNA-seq CIRCLE-seq

11. Cytospin centrifuge.
12. CO_2 incubator.
13. Inverted microscope.
14. Thermocycler.
15. Falcon tubes.
16. Serological pipettes.
17. Culture dishes.
18. Glass slides.

19. Dissection scissors.
20. Scalpel and forceps.
21. Beckman Coulter "Class S" Ultracentrifuge.

2.2 Reagents

1× PBS.

1× RBC lysis buffer.

Easy sep CD34 positive selection kit (Stem cell technologies, 17856).

Trypan Blue stain.

Giemsa stain.

Propidium iodide.

NucRed (Thermo Scientific, R37106).

1 × NLS Cas9 protein (TAKARA, 632640).

Chemically modified sgRNAs (Synthego), ssODN (Integrated DNA Technologies).

P3 Nucleofection solution (Lonza, V4XP-3024).

Methocult Optimum medium (Stem cell technologies, H4034).

Cyclosporine H (Sigma, SML1575).

Megakaryocyte expansion supplement (Stem cell technologies #02696).

Sodium Meta Bisulfide.

Transfection reagents—polyethylenimine (PEI) (Polysciences), TransIT-LT1(Mirus Bio).

0.45 μm filter (PVDF).

HEPES buffer.

Polybrene.

Opti-MEM (Gibco).

Benzonase, Tris–HCl.

NaCl.

HiScribe mRNA synthesis Kit, CleanCap AG (Trilink).

2.3 Media

1. HSC expansion medium [14]—Stemspan STEM II medium (Stem cell technologies,09655) supplemented with 240 ng/ml SCF (300-07), 40 ng/ml IL6 (200-06), 240 ng/ml FLT3 (300-19), and 80 ng/ml TPO (300-18) (All the cytokines are obtained from Pepro Tech).
2. Erythroid differentiation medium [15]—Step I medium—1× IMDM–GlutaMAX (Thermo Scientific, 31980030), 100 U/ml penicillin–streptomycin (Thermo Scientific 10378-016), 5% Human AB serum (MP Biomedicals, 2930649), 330 mg/ml human holotransferrin (BBI Solutions, T101-5), 20 mg/ml

human insulin (Sigma, SLCC5730), 2 U/ml heparin (Sigma, H3149), 1 mM hydrocortisone (MP Biomedicals, 101996), 3 U/ml recombinant human erythropoietin (Zyrop 4000), 100 ng/ml stem cell factor, and 5 ng/ml interleukin-3 (Peprotech, 200-03). Step II medium—Step I medium without IL-3 and hydrocortisone. Step III medium—Step II medium without SCF (*see* **Note 1**).

3. Megakaryocyte expansion medium: 100 μl of 100× megakaryocyte supplement in 10 ml of SFEM II medium.
4. Freezing Medium—7% IMDM (Hyclone SH30228.01), 2% Filtered FBS (Thermo Scientific, A4766801) and 1% DMSO (7:2:1) or Cryostor (Stemcell Technologies, 07941).
5. AAV lysis buffer: 50 mM Tris–HCl, 150 mM NaCl adjusted to pH 8.0.

2.4 Antibodies

1. hCD34 PE (550761), hCD90 BV421(562556), hCD133 APC (372806) for HSC characterization; hCD45 FITC (55548), mCD45 APC (559864) for engraftment analysis, multi lineage reconstitution analysis requires CD19 PE (340364) and hCD3 APC (340440) [lymphoid markers], hCD56 APC (341025) and hCD16 PE (347617) [NK cells markers], hCD33 APC (340474) and hCD3APC (340440) [myeloid markers], hCD71 BV421 (562995) and hCD235a PE (21812354) [erythroid markers], hCD41a APC (559777) [platelet marker]. Annexin V,7-AAD. All these antibodies are obtained from BD biosciences. Human fetal HB APC (Thermo Scientific, MHFH05) (*see* **Note 2**).

2.5 Animal Requirement

Female Nonirradiated NOD, B6. SCID *Il2rγ*$^{-/-}$*Kit*$^{W41/W41}$ (NBSGW) at 6–8 weeks of age weighing around 25 g

3 Methods

All the cell culture experiments are to be followed under aseptic conditions inside the laminar airflow chamber.

3.1 PBMNC Isolation

3.1.1 Normal Donors

Peripheral blood mononuclear cells (PBMNCs) comprises of lymphocytes, monocytes, and dendritic cells. These cells are fractionated from mobilized peripheral blood by density gradient centrifugation method.

1. Transfer the G-CSF mobilized peripheral blood from blood bag into a 50 ml falcon tube. Make a note of donor details for future reference.

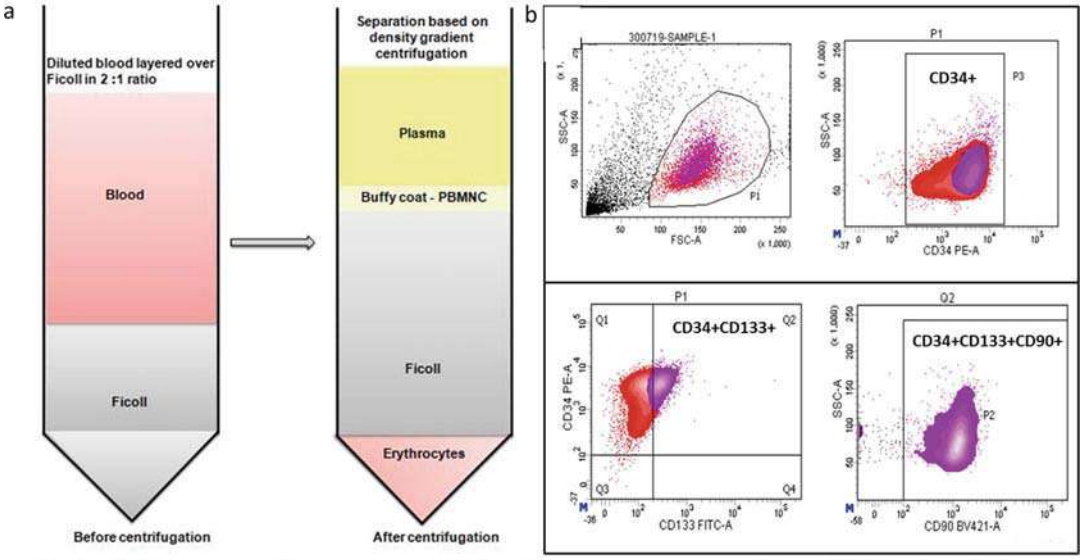


Fig. 1 (**a**) Isolation of PBMNC from mobilized peripheral blood by density gradient centrifugation using Ficoll-Paque. (**b**) Determination of long term and primitive HSC population by flow cytometry

2. Dilute the blood with an equal volume of 1× PBS (*see* **Note 3**). Gently layer the diluted blood above Ficoll-Paque in 2:1 ratio in a 50 ml falcon tube.
3. Spin the tubes at 400 × *g* for 30–40 min at room temperature with minimal acceleration and deacceleration (Fig. 1a).
4. Carefully discard the upper layer containing serum and aspirate the white buffy coat comprising PBMNCs into a new falcon tube (*see* **Note 4**).
5. In case of RBC contamination, incubate the cells with 1× RBC lysis buffer for 10 min on ice and spin at 200 × *g* for 10 min (*see* **Note 5**).
6. Wash the PBMNC cell pellet with 1× PBS and count the cells using trypan blue staining (*see* **Note 6**).
7. Cryopreserve the isolated PBMNCs in ice-cold freezing medium or proceed for HSPC isolation immediately.

3.1.2 Beta-Hemoglobinopathy Donors

1. Mobilized peripheral blood of β-hemoglobinopathy patient is processed directly by centrifuging at 600 × *g* for 10 min to separate out the buffy coat followed by RBC lysis for 5mins (twice) to remove the RBCs and obtain PBMNCs population.
2. This is because Ficoll-Plaque layering would result in thin PBMNC layer contaminated with defective erythrocytes.

3.2 HSPC Isolation

Hematopoietic stem and progenitor cells (HSPCs) constitute about 1–3% of PBMNCs derived from mobilized peripheral blood. These HSPCs are heterogeneous in population, and the state of each cell is determined by their differential expression of surface antigens. Standard and rapid enrichment of HSPCs from PBMNC is based on immunomagnetic positive selection of CD34 antigen expressing cells.

1. Resuspend PBMNC pellet in 1× PBS thoroughly to generate a homogeneous suspension.
2. Add 100 μl of CD34 positive selection antibody cocktail to 100 million PBMNCs and incubate for 30 min at room temperature with periodic tapping. Subsequently, add 60 μl of magnetic nanoparticles and incubate for another 20 min. These magnetic nanoparticles will bind to the CD34 antibody-bound target cells.
3. Transfer the labelled cells to an FACS tube and place inside the magnetic stand followed by 5 min incubation at room temperature. Invert the FACS tube along with the magnetic stand into a discard jar to remove the unbound cells after the incubation period.
4. Rinse the sides of the tube with 1× PBS and repeat **step 3** for a total of 4–5 washes with 1× PBS until the discard solution becomes transparent (*see* **Note 7**).
5. Remove the FACS tube from the magnetic stand and collect the CD34 positive cells bound to magnetic nanoparticles attached to the tube walls by flushing them with 1 × PBS.
6. Transfer the cells into a 15 ml centrifuge tube and pellet down the CD34 positive cells at 200 × *g* for 5 min, resuspend them in HSPC expansion medium and seed them in a culture dish at appropriate cell density, preferably 0.5 million/ml medium.

3.3 Long Term HSC Characterization

The isolated CD34 positive cells are further characterized for the presence of primitive and long-term repopulating HSCs based on the expression of other surface antigens such as CD90 (Thy-1) and CD133 (Prominin-1) as their role is elucidated in maintaining HSPC at quiescence state and enhancing engraftment capacity respectively [16–18].

1. Add 2 μl of CD34, CD90 and CD133 antibody to approximately 50,000 isolated cells after CD34 enrichment and incubate for 15–20 min at room temperature.
2. Wash the cells with 1×PBS to remove unbound antibodies. Resuspend the cells with 200 μl of 1×PBS and analyze the cells by flow cytometry (Fig. 1b)

3.4 Genome Engineering of HSPCs

Based on the desired type of target modification, the suitable strategy for genome editing is used to achieve maximum efficiency and for desired modification. Selecting the optimal procedure is the most vital step in gene editing.

3.4.1 Gene Disruption/Deletion

Gene disruption is performed by combining the endonuclease activity of cas9 with the endogenous NHEJ repair pathway. RNP complex electroporation is the most preferred method of cas9-sgRNA delivery, resulting in higher on-target and low off-target activity. Dual sgRNAs flanking the target region can be used to delete the desired loci.

1. Reconstitution of sgRNA—Spin the vial containing lyophilized sgRNA. Reconstitute the sgRNA with the appropriate volume of 1× TE buffer to achieve 100pmole per microliter concentration. Mix thoroughly and make aliquots of the sgRNA in smaller volumes to avoid multiple freeze-thawing (*see* **Note 8**). Nucleofection solution (NS) preparation—Add 16.4 μl P3 solution with 3.6 μl of supplement to prepare 20 μl of nucleofection solution per reaction (*see* **Note 9**).
2. Ribonucleoprotein (RNP) complex assembly—Mix 50 pmoles of Cas9 with 100 pmoles of reconstituted sgRNA and incubate at room temperature for 15 min (*see* **Note 10**).
3. Pellet 0.2 million of HSPCs and wash the cell pellet with 1× PBS to remove medium components completely. Gently resuspend the cell pellet in nucleofection solution and mix thoroughly with cas9: sgRNA complex. Carefully dispense the cells onto the sides of a well in a 16-well strip without generating air bubbles (*see* **Note 11**).
4. Cover the strip with a lid and insert into the X-unit of Lonza 4D nucleofector. Set up the Nucleofector program by choosing the appropriate wells and the buffer used. The cells are subjected to high-voltage electrical pulse using the defined pulse code DZ100 [19].
5. After the successful electroporation (*see* **Note 12**), add 50 μl of culture medium without antibiotic to each well and leave the cells undisturbed (this step will stabilize the edited cells from the shock created) for 10–15 min. Transfer the cells into a 12-well plate supplemented with complete expansion medium and incubate in a CO2 incubator at 37 °C.

3.4.2 Gene Correction/Addition

The donor DNA provides the template for the DNA repair mechanism to follow after the double-strand break by the cas9-sgRNA complex. The mode of delivery of the donor DNA is by either ssODN or rAAV6.

ssODN-Based Donor DNA Delivery

1. ssODN Design: DNA sequences containing the desired alteration is flanked by 111-27 bp homology arms (asymmetric) [5] from the antisense strand (against the sgRNA sequence) is designed along with a silent mutation in the sgRNA recognition domain or PAM motif to prevent recutting of the edited DNA sequence (*see* **Note 13**).
2. Reconstitution of ssODN—Dissolve the lyophilized ssODN contents from Integrated DNA Technologies (IDT) with the appropriate volume of 1× TE buffer to achieve 100 pmoles per microliter concentration. Mix thoroughly and make aliquots of the ssODN in smaller volumes to avoid multiple freeze-thawing (*see* **Note 8**).
3. Reconstitution of sgRNA—Spin the vial containing lyophilized sgRNA. Reconstitute the sgRNA with the appropriate volume of 1× TE buffer to achieve 100 pmoles per microliter concentration. Mix thoroughly and make aliquots of the sgRNA in smaller volumes to avoid multiple freeze-thawing.
4. Nucleofection solution (NS) preparation—Add 16.4 μl P3 solution with 3.6 μl of supplement to prepare 20 μl of nucleofection solution per reaction (*see* **Note 9**).
5. Ribonucleoprotein (RNP) complex assembly—Mix 50 pmoles of Cas9 with 100 pmoles of reconstituted sgRNA and incubate at room temperature for 15 min. 100 pmoles of ssODN (1 μl) is added along with the RNP complex right before mixing with the cells for nucleofection.
6. Pellet 0.2 million of HSPCs and wash the cell pellet with 1× PBS to remove medium components completely. Gently resuspend the cell pellet in nucleofection solution and mix thoroughly with cas9–sgRNA complex. Carefully dispense the cells onto the sides of a well in a 16-well strip without generating air bubbles (*see* **Note 11**).
7. Cover the strip with the lid and insert it into the X-unit of Lonza 4D nucleofector. Set up the Nucleofector program by choosing the appropriate wells and the buffer used. The cells are subjected to high-voltage electrical pulse using the defined pulse code CM149 [6].
8. After the successful electroporation (*see* **Note 12**), add 50 μl of culture medium without antibiotic to each well and leave the cells undisturbed (this step will stabilize the edited cells from the shock created) for 10–15 min. Transfer the cells into a 12-well plate supplemented with complete expansion medium and incubate in a CO_2 incubator at 37 °C.

rAAV-Based Donor DNA Delivery

1. Clone the donor DNA template in rAAV6 vector by Gibson assembly. More detailed donor DNA cloning in rAAV is available in [7].

2. Plate approximately 25–30 million HEK293T cells in a 15 cm dish on day 1. Sixteen hours later when the cells are at 80% confluency transfect them using 15 μg of serotype helper, rAAV6 vector, and adeno helper plasmids with 135 μl of PEI (transfection reagent) [6].
3. After 72 h, harvest the cells using a cell scraper and centrifuge for 5 min at 3000 × *g*.
4. Discard the supernatant and resuspend the cell pellet in 2 ml of AAV lysis buffer.
5. Perform three cycles of freeze–thaw using dry ice and 37 °C water bath. Add benzonase (10 U/ml of lysate) and incubate at 37 °C for 30 min.
6. Centrifuge at 10,000 × *g* for 10 min and perform an iodixanol density gradient ultracentrifugation with the supernatant to purify the virions with recombinant rAAV6 genomes.
7. Wash the AAV virus pellet with 1×PBS (supplemented with 0.001% Tween-20) thrice and concentrate to 200 μl and store at −80 °C.
8. After nucleofection of the cells with RNP complex, seed the edited cells in medium containing the AAV virus (at 3% of culture volume) [6] and incubate in a 5% CO_2 incubator.
9. After 24 h, change the medium to HSC expansion medium supplemented with the cytokines.

3.5 Base Editing Mediated Gene Modification

The base editor is involved in the conversion of nucleobase (C*G to T*A or A*T to G*C) within the editing window of the target region. sgRNAs for base editing is designed containing the target base within the editing window for ABE8e (4–8 bases from the opposite end of PAM) and CBE (2–9 bases from the opposite end of PAM). The base conversion protocol for the recently evolved ABE 8 through lentiviral and mRNA delivery is given below *(see* **Note 14**).

3.5.1 Lentiviral Approach

Plasmid Details: Cas9 coding sequence in the Addgene Plasmid (#57818) [20] is replaced by the ABE8e segment from the Addgene Plasmid (#138495) [21] by restriction digestion and subsequent cloning to achieve a lentiviral vector containing ABE8e and sgRNA scaffold along with the GFP reporter gene.

1. Plate approximately 4.5million 293 T cells in a 100 mm dish on day 1 (*see* **Note 15**).
2. Sixteen hours later, when the cells are at 80% confluency transfect them using second-generation lentiviral packaging plasmids and the plasmid of interest with 50 μl of TransIT LT1 (transfection reagent) mixed in 500 μl Opti-MEM medium.

3. Gently add the solution dropwise onto the plate and incubate at 37 °C.
4. After 24 h, replace the plate with new medium. Forty-eight hours after transfection, collect the supernatant and store it at 4 °C. In the same way, collect 60 h and 72 h supernatant.
5. Combine the supernatants in a 50 ml falcon tube and filter using a 0.45 μm filter (PVDF). Carefully add the filtered supernatant to the ultracentrifugation tube and close it using a cap.
6. Measure the weight of the ultracentrifuge tubes to ensure balancing and centrifuge at 40,000 × *g* using 70Ti rotor at 4 °C for 2 h.
7. After 2 h, remove the supernatant and carefully resuspend the virus pellet in IMDM–GlutaMAX medium (*see* **Note 16**).
8. Incubate the virus on ice for 1 h for homogeneous suspension and store at −80 °C.
9. Seed around 0.2million Day 1 HSPCs after characterization in one well of 12-well plate. Add 0.8 μl polybrene (6 mg/ml) and 10 μl of HEPES per ml of medium to aid transduction.
10. To cells add 40 μl of 300 × concentrated virus and plate centrifuge at 800 × *g* for 30 min at room temperature.
11. To enhance transduction efficiency, preincubate the isolated HSPCs in medium containing CsH (Concentration—2 μM (micromolar) of medium) for 16 h and perform transduction (*see* **Note 17**).
12. Change medium after 24 h of transduction and seed the cells in a 6-well plate at a seeding density of 0.2 million per ml.
13. Observe selection marker (GFP) in the transduced cells by FACS after 48 h and then set up differentiation using the required protocol.

3.5.2 Electroporation [8]

1. Amplify the template for mRNA synthesis from plasmid comprising base editor using forward and reverse primer with Q5 Hot Start 2× Master Mix (*see* **Note 18**).
2. Purify the PCR product using Zymo Research 25 μg DCC column.
3. With the obtained product as a template, synthesize mRNA using HiScribe High-Yield Kit with uridine modification and capping with CleanCap AG (Trilink) following the protocol given in the manual.
4. Pellet down 1.25 million Day 2 HSPCs cells and wash with 1 ml of MaxCyte buffer (HyClone) with 0.1% HAS.
5. Resuspend the cells in 1 ml ice-cold MaxCyte buffer and mix thoroughly. Split this cell suspension into multiple 20 μl aliquots.

6. Add 0.15 μM ABE8 mRNA and 4.05 μM sgRNA in a tube and make the final volume to 5 μl using MaxCyte buffer.
7. Load a set of 3 reactions containing 20 μl of cells and 5 μl of RNA composition are mixed uniformly by pipetting up and down without creating air bubbles into the individual chamber of an OC25 × 3 MaxCyte cuvette.
8. Select HSC-4 code to electroporate the RNA mixture into the cells.
9. Soon after electroporation transfer the cells into a 24well plate containing basal medium.
10. After 20 min of recovery, add complete medium with cytokines and incubate 37 °C (*see* **Note 19**).

The above protocol can be adapted for base conversion using cytosine base editors.

3.6 Ex Vivo Clonal Analysis

The pool of genome-modified HSPCs displays molecular and functional heterogeneity which creates a challenge in understanding the biology of individual cells. The development of methods to isolate and culture individual gene-modified cells by flow cytometry based single-cell index sorting is preferred for understanding heterogeneity. The homozygous and heterozygous edited clonal biology can be studied by analyzing the clonal editing percentage [22].

1. Add 150 μl 1×PBS to the side wells of 96-well Nunc flat bottom plate to maintain sufficient humidity within the plate.
2. Add 120 μl of MethoCult (semisolid) medium or differentiation medium to the other wells and incubate at 37 °C for 1 h before sorting.
3. Test sort using a test plate (96-well Nunc flat bottom) and verify proper placement of the cells (toward the center of a well). Place the cells to be sorted on ice before taking for sorting.
4. Gate only the viable cells by adding 2 μl of 7 AAD (viability dye) and index sort the cells using 100 μm diameter nozzle. Incubate the plate 37 °C with 5% CO2 (*see* **Note 20**).
5. Leave the plate undisturbed for 7 days and then visualize under a microscope for the appearance colonies. Score the colonies based on their morphology after 14 days of culture and genotype individual colonies.
6. Monitor the sorted gene-modified cells cultured in differentiation medium on alternative days. Based on the size and number of cells expanded, score the clones as small, medium or large. Top up the well with another 50 μl of medium if required (*see* **Note 21**).

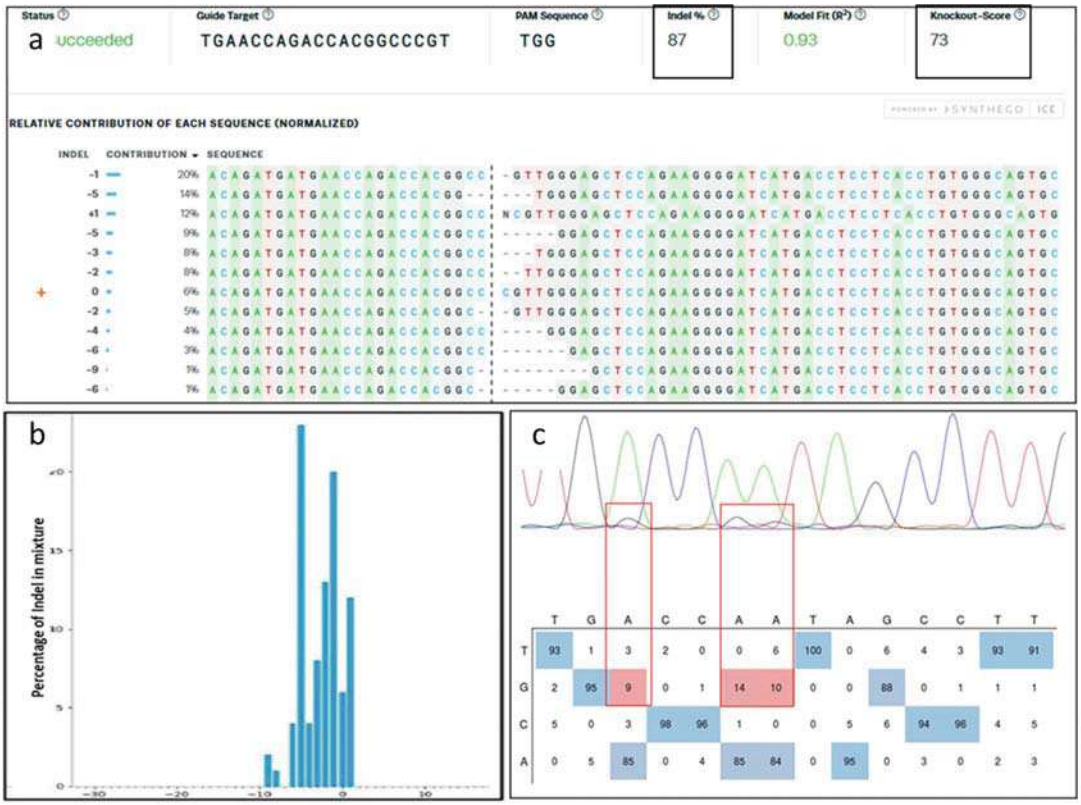


Fig. 2 (**a**) Cas9 RNP–mediated disruption of target region in HSPCs by Synthego ICE software. (**b**) Editing frequency and size of the Insertions/Deletions generated at the target region. (**c**) Adenosine base editor mediated A to G substitution in the target region in human HSPCs

7. Transfer the cells to a 48-well plate followed by 12-well as the cell number increases and perform genotyping and functional assays.

3.7 Analysis of Genome Engineered HSPCs

1. Isolate DNA from approximately 50,000 nucleofected HSPCs (from pooled/clonal populations obtained).
2. Perform Sanger sequencing of the target region using suitable primer and quantify the level of on-target gene editing efficiency using Synthego ICE software (for NHEJ and HDR) (Fig. 2 a, b) and EditR (for Base editing) (Fig. 2c) by aligning with the unedited reference sequence.
3. To analyze the in vitro colony-forming potential of gene-edited HSPCs by MethoCult colony-forming assay, premix the thawed MethoCult (CFU assay medium) with optimized cell number (500 cells/ml of medium) vigorously to obtain homogeneous cell suspension.
4. Keep the cell suspension undisturbed for approximately 30 min at room temperature for the air bubbles to settle.

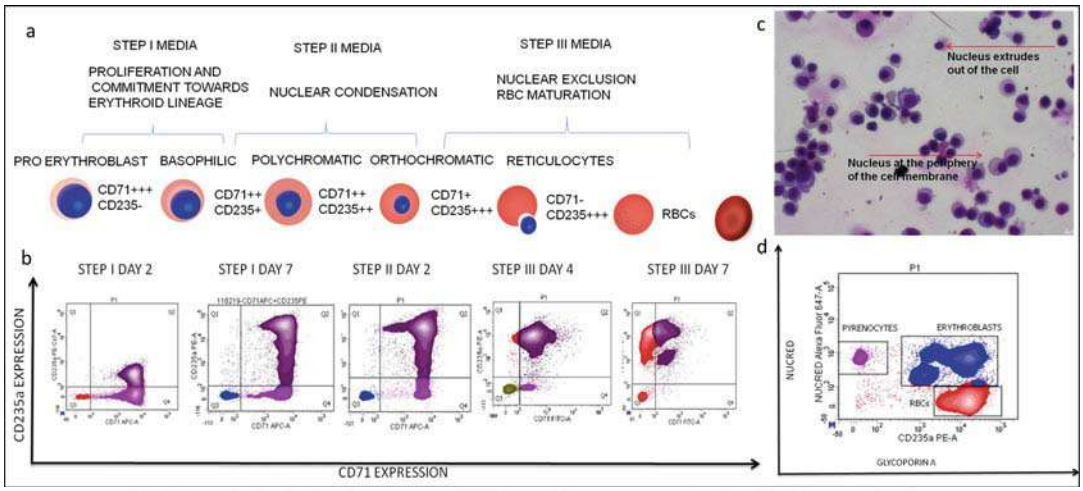


Fig. 3 (**a**) Image representing the sequential stages of HSC toward erythroid lineage differentiation. (**b**) Differential expression of erythroid markers during in vitro erythropoiesis. c) Morphology analysis of erythroid maturation using Giemsa stain. (**d**) Evaluation of percentage of nucleated/enucleated cells with erythroid differentiation marker CD235a

5. Layer the cell suspension evenly throughout the MethoCult plate (or 35 mm TC dish) without generating air bubbles. Incubate the plates at 37° C for 14–16 days.
6. Examine the plates under a light field microscope with 10× magnification and score the colonies formed based on their morphology, colour, and number (*see* **Note 22**) (Fig. 5a).
7. Genotype the individual colonies to verify the sustainability of the editing during in vitro multilineage differentiation.

3.8 Ex Vivo Differentiation of Genome Engineered HSPCs

The differentiation of edited HSPCs into different lineages is achieved by altering the culture conditions. The sequential maturation of the edited HSPCs in the myeloid lineage (Erythroid, Fig. 3a, and Megakaryocyte, Fig. 4a) is characterized by analyzing the differential expression of lineage-specific surface markers.

3.8.1 In Vitro Erythropoiesis

1. Gene modified HSPCs are seeded in varying culture conditions with different concentrations of cytokines that support erythrocyte maturation (*see* **Note 23**).
2. The gene-modified HSPCs proliferate and expand logarithmically in Step-1 erythroid differentiation medium for 7 days (48 h after nucleofection) with a regular medium change every 3–4 days.
3. On day 8, replace the culture medium into the Step 2 erythroid differentiation medium and culture for another 4–5 days. Subsequently, replate the cells in a Step-3 medium for 7–10 days or until enucleation is observed.

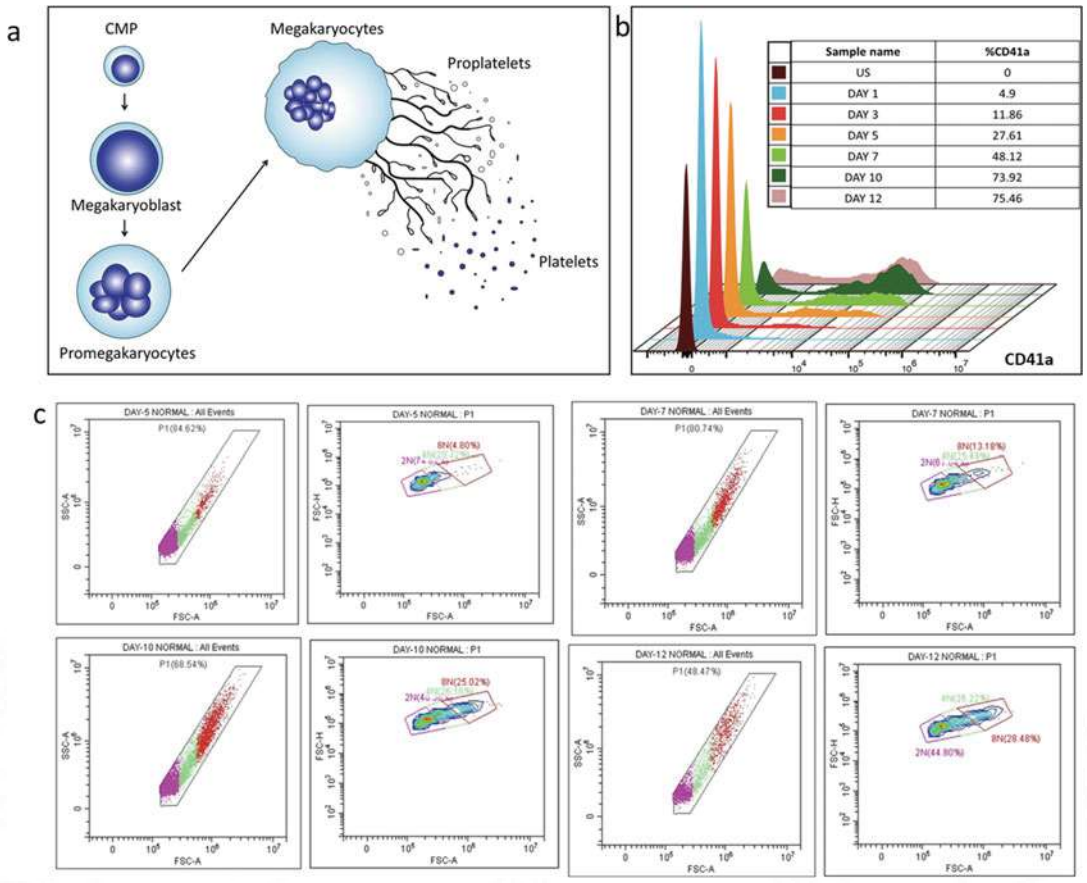


Fig. 4 (**a**) Schematic representation of various stages of megakaryocytic differentiation from human HSPCs. (**b**) Differential expression of megakaryocyte markers during in vitro megakaryocyte differentiation. (**c**) Assessment of DNA ploidy during in vitro megakaryocyte differentiation

4. Monitor the differential expression level of surface markers like transferrin (CD71) and Glycophorin A (CD235a) using labelled antibodies through FACS analysis at different stages of Erythroid maturation (Fig. 3b).
5. Assess the level of enucleation, by adding 10 μl of NucRed Live-cell stain along with 3 μl of CD235a antibody (Erythroid surface marker) (*see* **Note 24**).
6. Incubate the cells in the dark for 15 min and wash them with 1×PBS to remove unbound antibodies.
7. Analyze the percentage of enucleated cells in flow cytometry [14]. Cells positive for the Glycophorin A surface marker and negative for the nuclear stain are denoted as enucleated cells (CD235a + veNucRed-ve) (Fig. 3d).
8. Evaluate the growth profile of the gene-edited cells during erythroid expansion by counting the number of viable cells

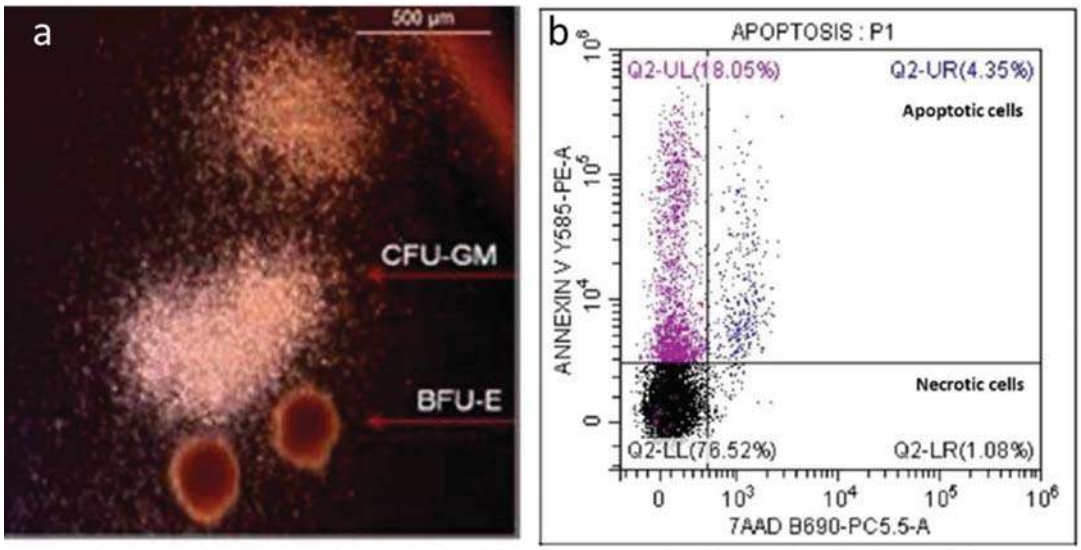


Fig. 5 (**a**) Representative images of colonies observed in Methocult media after 14 days. (**b**) Determination of cell viability using AnnexinV/7AAD

periodically using Trypan blue staining at every medium change.

9. Calculate fold expansion based on the ratio of total number cells at a particular time point to the initial number of cells seeded at the beginning of the experiment.
10. To determine the percentage of cell death during differentiation via apoptosis, add 2 µl of Annexin V and incubate for 15 min.
11. Acquire the sample immediately 5 min after the addition of 2 µl 7AAD in flow cytometry. Cells positive for Annexin V and 7 AAD undergo apoptosis, whereas cells positive for 7AAD and negative for Annexin V undergo necrosis (*see* **Note 25**) (Fig. 5b).
12. Isolate the RNA and protein from the gene-modified cells by using either commercially available kits or standard laboratory-based methods before the cell undergoes terminal differentiation.
13. Perform qRT-PCR and western blot analysis to validate the effect of gene modification on transcriptome and translatome level.
14. Morphological changes of cells during erythroid differentiation are observed by cytospin analysis.
15. Cytocentrifuge approximately 50,000 erythroid cells onto the slides at 200 × *g* for 5 min and fix them using ice-cold

methanol. Perform Giemsa staining to visualize the morphological changes (Fig. 3c).

16. Terminate the erythroid differentiation culture by pelleting down the erythroid cells at 200 × *g* for 5 min. Resuspend the cell pellet in Milli-Q water of appropriate volume and sonicate to extract the cell lysate. Remove the cell debris by spinning at 4 °C and store the supernatant at −80 °C until HPLC is run.

3.8.2 In Vitro Megakaryopoiesis

1. Maintain 0.5million HSPCs after gene modification in SFEM II medium supplemented with megakaryocyte supplement (100 μl of 100× supplement in 10 ml of medium). Carry out half medium change every alternative day up to Day 14.
2. On different days of differentiation stain around 50,000 cells with CD41a for 15 min and wash with 1 × PBS containing 0.5% BSA and 2 mM EDTA (Fig. 4b).
3. To determine the ploidy level of differentiating cells, wash the cells after CD41a staining with 1 × PBS containing 0.5% BSA and 2 mM EDTA (Fig. 4c).
4. Resuspend the pellet in 200 μl of buffer containing 1.25 mM sodium citrate, 2.5 mM sodium chloride, 3.5 mM dextrose along with 20 μg/ml propidium iodide and 0.05% Triton X-100 followed by 15–20 min incubation at 4 °C in the dark.
5. To this add RNase at 0.03 mM concentration and incubate in the dark for 20–30 min. Analyse the intensity of propidium iodide in CD41a expressing cells by flow cytometry.

3.9 Engraftment of Genome Engineered Cells in NBSGW Mice

Current mouse models used for transplantation studies show lower chimerism of human HSCs even after myeloablation. Further, the analysis of multilineage differentiation potential of engrafted genome modified cells is challenging as they do not support in vivo human erythropoiesis, and the mortality rate is also higher due to damage caused during whole-body irradiation. The NBSGW (NOD, B6. SCID Il2rg-/-Kit-W41/W41) mouse model is superior to NSG ((NOD SCID gamma mouse), exhibiting higher chimerism without irradiation along with enhanced human erythropoiesis. For long-term engraftment analysis, harvest the spleen, blood, and bone marrow of the transplanted mice after sacrificing 16 weeks of transplantation.

3.9.1 Xeno-transplantation

1. Select 6–8-week-old female NBSGW mice for transplantation studies.
2. Transplant around 0.5-1 million of gene-modified HSPCs through the tail vein or retro-orbital route of administration

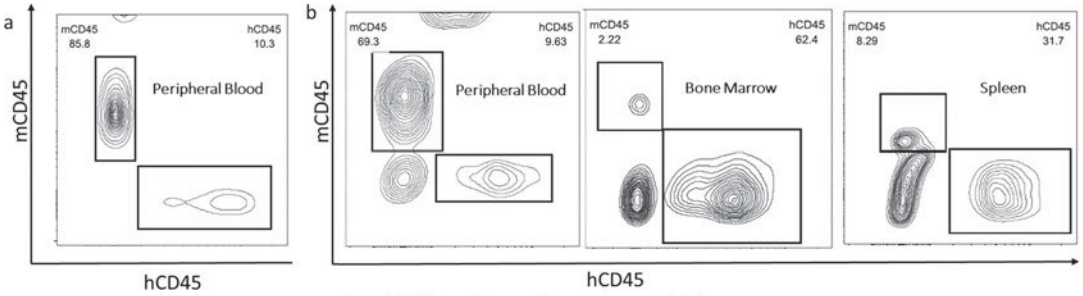


Fig. 6 (**a**) Short term engraftment potential of genome edited HSPCs in NBSGW mice after 8 weeks of transplantation. (**b**) Long term repopulation capacity of genome edited HSPCs in the Peripheral blood, bone marrow and spleen after 16 weeks of transplantation in NBSGW mice

using an insulin syringe with a 31 gauge X ¼ (0.25 mm × 6 mm) inch needle.

Short-Term Engraftment Analysis (Fig. 6a)

1. Collect peripheral blood via retro-orbital route 8 weeks after transplantation using a heparin-coated capillary tube and transferring blood into a tube containing heparin.
2. Wash the blood sample with 1×PBS.
3. To the red pellet obtained, add 1×RBC lysis buffer and place the tube on ice for 10 min followed by centrifugation at 110 × *g* for 5 min.
4. Repeat RBC lysis until the pellet becomes transparent.
5. Resuspend the pellet in 150 μl of 1×PBS; add 3 μl mCD45 APC antibody and 5 μl of hCD45 FITC antibody (*see* **Note 26**).
6. Incubate the tubes in the dark at room temperature with periodic tapping for 20 min.
7. Remove unbound antibodies by washing the cells with 1× PBS.
8. Re suspended the final cell pellet in 1× PBS and analyze using FACS.

Long-Term Bone Marrow Engraftment Analysis (Fig. 6b)

1. Collect both the tibia and femur bones of transplanted mice and remove the muscle along with other tissues using a sterile scalpel or blade.
2. Collect the bone marrow cells by syringe flushing or by centrifugation process (*see* **Note 27**) and seed the cells in a 65 mm dish with IMDM–GlutaMAX medium.
3. Pipette gently and make sure cells are mixed uniformly then incubate at 37 °C.

Table 2
Antibodies combination to analyse for multilineage engraftment potential

S.No	Lineage	Antibody combination
1	Lymphoid	hCD45 + hCD3 + hCD19
2	NK cells	hCD45 + hCD56 + hCD16
3	HSPCs	hCD45 + hCD133 + hCD34
4	Myeloid cells	hCD45 + hCD13 + hCD33
5	Megakaryocyte–erythroid progenitors	hCD45 + hCD41a(platelets) + hCD71
6	Erythroid–erythroid progenitors	hCD45 + hCD235a + hCD71

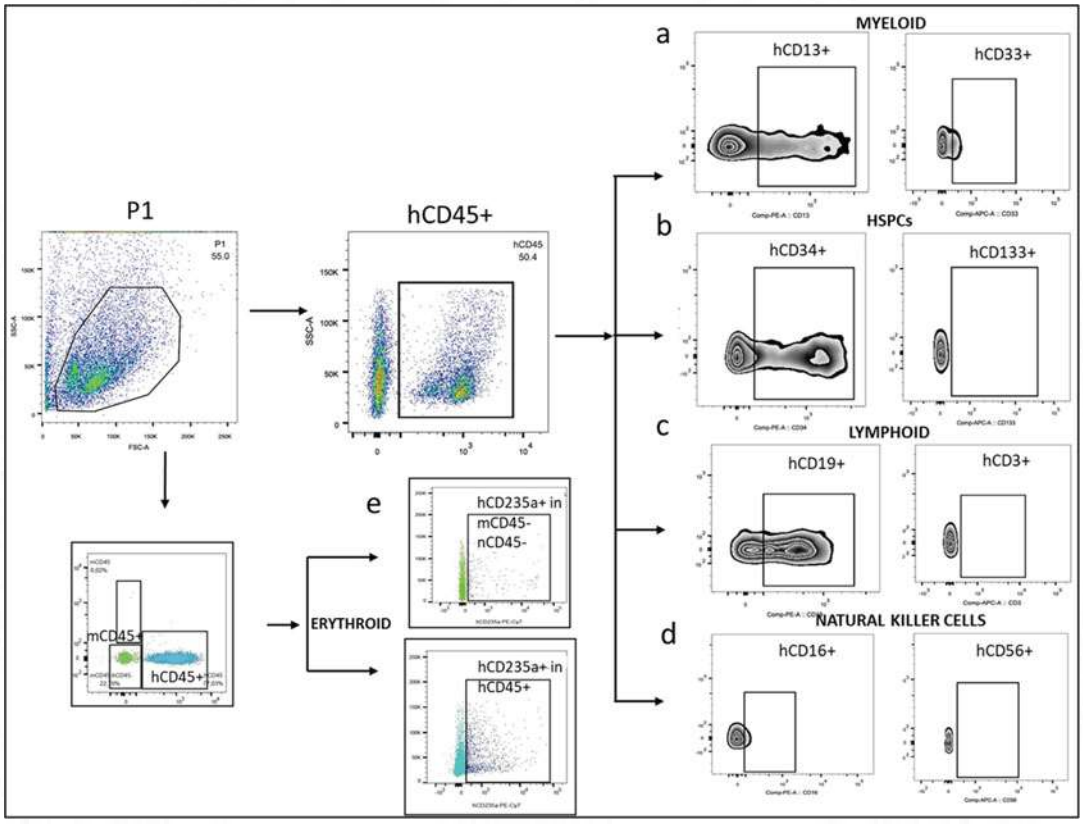


Fig. 7 Multilineage reconstitution potential of genome-edited HSPCs in bone marrow compartment of NBSGW mice after 16 weeks of transplantation depicting the frequency of (**a**) myeloid cells, (**b**) HSPC, (**c**) lymphoid, (**d**) natural killer cells, (**e**) erythroid

4. To 50,000 cells, add 5 μl of hCD45 FITC and 3 μl of mCD45 APC antibody. Wash the cells after 15 min of incubation and analyze by flow cytometry (*see* **Note 28**).
5. Incubate 50,000 bone marrow cells with appropriate antibodies as depicted in Table 2 for multilineage analysis and analysis in flow cytometry (Fig. 7) (*see* **Note 29**).
6. Stain 50,000 bone marrow cells with 2 μl of CD235a antibody and analysis through flow cytometry to estimate the percentage of human erythroid cells present in mouse bone marrow (NBSGW is reported to support human hematopoiesis due to hypomorphic mutation in the c-kit gene [23, 24]).

Long-Term Peripheral Blood Engraftment Analysis

1. Collect blood by cardiac puncture and process similarly as in short term engraftment analysis.

Long-Term Spleen Engraftment Analysis

1. Collect the spleen and place it in a well of a 12-well plate with 500 μl 1 XPBS.
2. Crush the spleen tissue with the end of the syringe plunger and wash with RBC lysis buffer.
3. Add human CD45 and mouse CD45 antibodies and analyze by flow cytometry.

4 Notes

1. Store all the supplements, including cytokines, at an appropriate temperature in multiple small aliquots to minimize repeated freeze–thaw cycles.
2. Fluorescence conjugated antibodies are highly sensitive to temperature and light; therefore, maintain them at 4 °C in the dark.
3. Increase the purity of PBMNCs by diluting the blood sample with a higher volume of 1×PBS.
4. Aspirate carefully the buffy coat containing the PBMNCs to minimize the contamination with other cells. The presence of Ficoll in the buffy coat hinders the cells from pelleting down efficiently during subsequent processes. Eliminate the Ficoll contamination by multiple washes with 1×PBS.
5. Avoid longer incubation of cells in 1× RBC lysis buffer to reduce cellular stress.
6. To prevent clump formation, perform DNAse-1 treatment for 10mins during the initial resuspension of PBMNCs in 1×PBS. Remove persisting clumps after DNAse-1 treatment by filtering the cell suspension through the 40 μm cell strainer to generate a homogeneous cell suspension.

7. Decreasing the number of washes with 1×PBS generates a higher cell counts. On the other hand, the purity of isolated cells dramatically improves by increasing the number of washes with 1× PBS.
8. For long term storage, store sgRNA aliquots in −80 °C immediately after reconstitution. Handle the sgRNA aliquots in an RNAse-free environment to prevent degradation.
9. Equilibrate nucleofection solution at room temperature for 15 min before the experiment. Adding cold nucleofection solution will affect the viability of the cells and decreases the editing efficiency.
10. Deletion depends on the individual sgRNA efficiency in causing double-strand break and also the total size of the deletion. RNP complex for each sgRNA is prepared separately and mixed before electroporation.
11. Wash the cell pellet with 1× PBS to remove any medium component like salts that would interfere with electroporation. Incubation of cells in nucleofection solution for an extended period is toxic to cells. Make sure that the final solution makes proper contact with the sides of the well, and the volume does not exceed 20–25 μl inside each well.
12. After the electroporation, a summary of the nucleofection process results will be displayed on the Lonza 4D nucleofector system. A colour code depicts the overall result: Green "+" symbol denotes successful nucleofection, whereas yellow "+" means nucleofection passed with reduced efficiency, and red "+" means unsuccessful nucleofection.
13. The total size of ssODN ranges from 50 to 400 bp inclusive of the homology arms but rAAV6 is approximately 4.2 kb with ~400 bp homology arm. Titrate the ssODN concentration to achieve good viability after nucleofection. Large gene insertions can be performed efficiently with the rAAV6 donor delivery system. Successive HR with two delivery vectors can be performed for the addition of genes larger than 4 kb. The efficiency of integration will be higher when the modification is proximal to the DNA cleavage site.
14. IDLV (integrase-deficient lentivirus) based delivery of ABE 8 is better suited in clinical applications over lentiviral delivery.
15. For efficient virus production, low passage 239 T cells are preferred, and transfection is carried out at optimal confluency (80–90%).
16. Make a circular mark in the ultracentrifugation tubes at one bottom corner and place the tube accordingly so that the virus pellet would be confined with a circle and is easy to obtain.

17. Adding Cyclosporin H to HSPCs before transduction at a higher concentration will affect the viability of cells.
18. Linearizing the plasmid using suitable restriction enzymes (giving blunt end or 5′ staggered cut) before amplification would be a better strategy to obtain template mRNA synthesis.
19. The rate of base conversion varies between different versions of base editors. ABE7.1 takes 5 days, whereas ABE 8e shows optimal conversion after 24 hrs of nucleofection.
20. Exclude preapoptotic cells while performing single-cell sorting by gating only the live population using a viability stain.
21. Monitor every alternative day for the appearance of colonies and check for the level of PBS in the side wells of the plate.
22. Methocult medium is highly susceptible to bacterial contamination it is enriched with nutrients and cytokines. Therefore, handling the medium in sterile conditions with utmost care is essential. Add water to outer space surrounding the wells in the Methocult plates or place additional wells containing water to prevent evaporation. Stemgrid (#27000) from Stemcell Technologies provides a grid background for easy counting of the colonies.
23. The cell pellet becomes red in colour during the initial erythroid expansion in Step 2 medium, and progressively the colour changes into dark brick red colour during the terminal differentiation stage. Cells are maintained at high culture density in subsequent stages of erythroid differentiation. Each donor exhibits differences in the duration of erythroid expansion and erythroid differentiation profile.
24. Hoechst staining can be used as an alternative to NucRed to stain the nucleus.
25. Cells positive for Annexin V and 7 AAD undergo apoptosis, whereas cells positive for 7AAD and negative for Annexin V undergo necrosis.
26. The total number of events in the flow-cytometry analysis increases if there are any residual RBCs present in the sample, which leads to the reduced live cell population. To overcome the issue, increase the number of PBS washes and acquire a higher number of total events.
27. Syringe flushing of bone marrow is performed by dissecting the ends of the bone and flushing the cells out using an insulin syringe with 1× PBS. Pellet down the bone marrow cells from the final soup. In the centrifugation method, the bones are cut down proximal to the knee-end and placed downward in a 0.5 ml centrifuge tube perforated at the bottom. Stack the perforated tube inside a 1.5 ml centrifuge tube and spin down for 30 s at maximum RPM. Culture the red pellet

containing bone marrow cells in the appropriate medium for further processing. [25].

28. The ratio of human CD45 to the total CD45 expressing cells (including mouse and human) determines the percentage of human engraftment [26].
29. Proceed with multilineage engraftment analysis only if there is a significant amount of human CD45 expression.

References

1. Morgan RA, Gray D, Lomova A, Kohn DB (2018) Hematopoietic stem cell gene therapy—Progress and lessons learned. Cell Stem Cell 21(5):574–590
2. Melve GK, Ersvaer E, Eide GE, Kristoffersen EK, Bruserud O (2018) Peripheral blood stem cell mobilization in healthy donors by granulocyte colony-stimulating factor causes preferential mobilization of lymphocyte subsets. Front Immunol 9:845
3. Cavazzana M, Bushman FD, Miccio A, André-Schmutz I, Six E (2019) Gene therapy targeting haematopoietic stem cells for inherited diseases: progress and challenges. Nat Rev Drug Discov 18(6):447–462
4. Gaj T, Staahl BT, Rodrigues GMC, Limsirichai P, Ekman FK, Doudna JA, Schaffer DV (2017) Targeted gene knock-in by homology-directed genome editing using Cas9 ribonucleoprotein and AAV donor delivery. Nucleic Acids Res 45(11):e98
5. Richardson CD, Ray GJ, DeWitt MA, Curie GL, Corn JE (2016) Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. Nat Biotechnol 34(3):339–344
6. Pattabhi S, Lotti SN, Berger MP, Singh S, Lux CT, Jacoby K, Lee C, Negre O, Scharenberg AM, Rawlings DJ (2019) In vivo outcome of homology-directed repair at the HBB gene in HSC using alternative donor template delivery methods. Mol Ther Nucleic Acids 17:277–288
7. Bak RO, Dever DP, Porteus MH (2018) CRISPR/Cas9 genome editing in human hematopoietic stem cells. Nat Protoc 13(2): 358–376
8. Gaudelli NM, Lam K, Rees HA, Solá-Esteves NM, Barrera LA, Born DA, Edwards A, Gehrke JM, Lee S-J, Liquori AJ, Murray R, Packer MS, Rinaldi C, Slaymaker IM, Yen J, Young LE, Ciaramella G (2020) Directed evolution of adenine base editors with increased activity and therapeutic application. Nat Biotechnol 38(7):892–900
9. Tsai SQ, Zheng Z, Nguyen NT, Liebers M, Topkar VV, Thapar V, Wyvekens N, Khayter C, John Iafrate A, Le LP, Aryee MJ, Joung JK (2014) GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. Nat Biotechnol 33(2):187–197
10. Doman JL, Raguram A, Newby GA, Liu DR (2020) Evaluation and minimization of Cas9-independent off-target DNA editing by cytosine base editors. Nat Biotechnol 38(5): 620–628
11. Tsai SQ, Nguyen NT, Malagon-Lopez J, Topkar VV, Aryee MJ, Keith Joung J (2017) CIRCLE-seq: a highly sensitive *in vitro* screen for genome-wide CRISPR-Cas9 nuclease off-targets. Nat Methods 14(6):607–614
12. Gabriel R, Kutschera I, Bartholomae CC, von Kalle C, Schmidt M (2014) Linear amplification mediated PCR – localization of genetic elements and characterization of unknown flanking DNA. J Vis Exp (88):e51543
13. McIntosh BE, Brown ME, Duffin BM, Maufort JP, Vereide DT, Slukvin II, Thomson JA (2015) Nonirradiated NOD,B6.SCID Il2rγ-/-KitW41/W41 (NBSGW) mice support multilineage engraftment of human hematopoietic cells. Stem Cell Reports 4(2):171–180
14. Tajer P, Pike-Overzet K, Arias S, Havenga M, Staal FJT (2019) Ex vivo expansion of hematopoietic stem cells for therapeutic purposes: lessons from development and the niche. Cell 8:169
15. Psatha N, Reik A, Phelps S, Zhou Y, Dalas D, Yannaki E, Levasseur DN, Urnov FD, Holmes MC, Papayannopoulou T (2018) Disruption of the BCL11A erythroid enhancer reactivates fetal hemoglobin in erythroid cells of patients with b-thalassemia major. Mol Ther Methods Clin Dev 10:313–326

16. Radtke S, Pande D, Cui M, Perez AM, Chan Y-Y, Enstrom M, Schmuck S, Berger A, Eunson T, Adair JE, Kiem H-P (2020) Purification of human CD34+ CD90+ HSCs reduces target cell population and improves lentiviral transduction for gene therapy. Mol Ther Methods Clin Dev 18:679–691
17. Drake AC, Khoury M, Leskov I, Iliopoulou BP, Fragoso M, Lodish H, Chen J (2011) Human $CD34^{+}CD133^{+}$hematopoietic stem cells cultured with growth factors including Angptl5 efficiently engraft adult NOD-SCID $Il2r\gamma^{-/-}$ (NSG) mice. PLoS One 6:e18382
18. Radtke S, Görgens A, Kordelas L, Schmidt M, Kimmig KR, Köninger A, Horn PA, Giebel B (2015) CD133 allows elaborated discrimination and quantification of hematopoietic progenitor subsets in human hematopoietic stem cell transplants. Br J Haematol 169:868–878
19. Gomez-Ospina N, Scharenberg SG, Mostrel N, Bak RO, Mantri S, Quadros RM, Gurumurthy CB, Lee C, Bao G, Suarez CJ, Khan S, Sawamoto K, Tomatsu S, Raj N, Attardi LD, Aurelian L, Porteus MH (2019) Human genome-edited hematopoietic stem cells phenotypically correct Mucopolysaccharidosis type I. Nat Commun 10(1):4045
20. Heckl D, Kowalczyk MS, Yudovich D, Belizaire R, Puram RV, McConkey ME, Thielke A, Aster JC, Regev A, Ebert BL (2014) Generation of mouse models of myeloid malignancy with combinatorial genetic lesions using CRISPR-Cas9 genome editing. Nat Biotechnol. 10(1):4045
21. Richter MF, Zhao KT, Eton E, Lapinaite A, Newby GA, Thuronyi BW, Wilson C, Koblan LW, Zeng J, Bauer DE, Doudna JA, Liu DR (2020) Phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity. Nat Biotechnol 38(7): 883–891
22. Wagenblast E, Azkanaz M, Smith SA, Shakib L, McLeod JL, Krivdova G, Araújo J, Shultz LD, Gan OI, Dick JE, Lechman ER (2019) Functional profiling of single CRISPR/Cas9-edited human long-term hematopoietic stem cells. Nat Commun 10(1):4730
23. Hess NJ, Lindner PN, Vazquez J, Grindel S, Hudson AW, Stanic AK, Ikeda A, Hematti P, Gumperz JE (2020) Different human immune lineage compositions are generated in non-conditioned NBSGW mice depending on HSPC source. Front Immunol 11:573406
24. Fiorini C, Abdulhay NJ, McFarland SK, Munschauer M, Ulirsch JC, Chiarle R, Sankaran VG (2017) Developmentally-faithful and effective human erythropoiesis in immunodeficient and Kit mutant mice. Am J Hematol 92(9):E513–E519
25. Amend SR, Valkenburg KC, Pienta KJ (2016) Murine hind limb long bone dissection and bone marrow isolation. J Vis Exp (110):53936
26. Futrega K, Lott WB, Doran MR (2016) Direct bone marrow HSC transplantation enhances local engraftment at the expense of systemic engraftment in NSG mice. Sci Rep 6:23886

APPENDIX C

Plagiarism Check Report

Institute for Stem Cell Science and Regenerative Medicine

An autonomous Institutee of
Department of Biotechnology,
Government of India.

inStem
CSCR
CENTER FOR STEM CELL RESEARCH
(a unit of inStem, Bengaluru)
Christian Medical College Campus
Bagayam, Vellore - 632002

Center for Stem Cell Research

(A unit of inStem, Bengaluru in Collaboration with DBT and CMC, Vellore)

CERTIFICATE

This is to certify that the thesis submitted by **Mr. VIGNESH R**, affiliated with the Center for Stem Cell Research at Christian Medical College, has undergone a thorough plagiarism check utilizing iThenticate software. We are pleased to declare that the thesis is deemed acceptable for submission. While the overall similarity index was 18%, it is noteworthy that the individual primary source index fell within the permissible range of 1-5%.

Title: Pre-clinical study for therapeutic genome editing in beta-hemoglobinopathies

Registration No: **2019/PhD/09**

Name of the Guide: **Dr. Mohankumar K M**

Date: 05/Dec/2023

Christian Medical College Campus,
Bagayam, Vellore - 632002. INDIA.

Phone : +91 - 416 - 228 5101 / 5107
Fax : +91 - 416 - 228 5103
E-mail : cscr@cmcvellore.ac.in

SREE CHITRA TIRUNAL INSTITUTE FOR MEDICAL SCIENCES AND

TECHNOLOGY, THIRUVANANTHAPURAM - 695 011, KERALA, INDIA.

Email :sct@sctimst.ac.in
Phone : 91-471-2443152 Fax : 91-471-2446433

An Institution of National Importance established by an Act of the Indian Parliament
(Act No.52 of 1980)