UCLA
UCLA Electronic Theses and Dissertations

Title
Gene Editing of Bruton's Tyrosine Kinase for Treatment of X-Linked Agammaglobulinemia

Permalink
https://escholarship.org/uc/item/9w25s8dg

Author
Gray, David

Publication Date
2020

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA

Los Angeles

Gene Editing of Bruton's Tyrosine Kinase for Treatment of X-Linked Agammaglobulinemia

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Molecular Biology

by

David Gray

2020

© Copyright by

David Gray

2020

ABSTRACT OF THE DISSERTATION

Gene Editing of Bruton's Tyrosine Kinase for

Treatment of X-Linked Agammaglobulinemia

by

David Gray
Doctor of Philosophy in Molecular Biology
University of California, Los Angeles, 2020
Professor Donald Barry Kohn, Chair

X-Linked Agammaglobulinemia (XLA) is a primary immunodeficiency characterized by a lack of mature B lymphocytes and antibody production. Patients with XLA have loss of function mutations in the Bruton's Tyrosine Kinase (*BTK*) gene. The standard of care for XLA is immunoglobulin supplementation, which has a profound effect on patient wellbeing and life expectancies. However, immunoglobulin supplementation requires frequent, expensive injections throughout a patient's life and patients remain susceptible to certain recurring illnesses. The only permanent cure for XLA is an allogeneic hematopoietic stem cell (HSC) transplant, though it is rarely performed due to the associated risks. Gene therapy-based methods to replace or repair the *BTK* gene in autologous HSCs provide an alternative with the benefits of a permanent cure for XLA while circumventing much of the risk.

While previous efforts to deliver a functional copy of the *BTK* gene using viral vector mediated gene transfer have shown promise, these vectors carry a risk of insertional oncogenesis

which may not be tolerated for treatment of XLA. Instead, this dissertation lays a foundation for targeted integration of a functional copy of the *BTK* gene into HSCs using Cas9 endonuclease mediated gene editing to drastically reduce those risks.

Initial work identified and optimized a target site for integration into both cell lines and primary cells. However, integration of the *BTK* sequence alone generated insufficient transgene expression. We identified three modifications that improved integration and expression of the construct: mutation of the protospacer adjacent motif, re-addition of the *BTK* terminal intron, and addition of the woodchuck hepatitis virus posttranscriptional regulatory elemental. Together, these modifications generated expression of transgenic BTK nearing wildtype levels in cell lines with seamless DNA integration. Finally, an unbiased comparison of three different methods of targeted integration (homology directed repair, homology independent targeted integration, and precise integration into target chromosome) was performed to identify the optimal donor design for treatment of XLA. This work has produced a novel treatment for XLA that is ready to progress into *in vivo* models and towards the clinic. The new, effective donor modifications may be invaluable for similar treatments in development for other genetic disorders.

The dissertation of David Gray is approved.

Gay Crooks

Kenneth Dorshkind

Owen Witte

Dinesh Rao

Donald Barry Kohn, Committee Chair

University of California, Los Angeles

2020

DEDICATION

This dissertation is dedicated to my family. Steve Gray, Terri Gray, Christie Gray, Pat Thomas, Dave Thomas, Gerry Gray, Harry Gray, and Meztli España and a remaining list too long to name. Their constant support, compassion, and love has made this possible.

TABLE OF CONTENTS

List of Figures and Tables vii

Acknowledgements ix

VITA xiii

CHAPTER 1 1

Introduction: Hematopoietic Stem Cell Gene Therapy – Progress and Lessons Learned

CHAPTER 2 51

Optimizing Integration and Expression of Transgenic Bruton's Tyrosine Kinase for CRISPR/Cas9 Mediated Gene Editing of X-Linked Agammaglobulinemia

CHAPTER 3 93

A Comparison of DNA Repair Pathways to Achieve Site-Specific Correction in Bruton's Tyrosine Kinase

CHAPTER 4 130

Discussion and Closing Remarks: Major Steps Towards a Cure for XLA

LIST OF FIGURES AND TABLES

CHAPTER 1

Figure 1-1. Overview of Targets for Gene Therapy 5

Figure 1-2. Autologous hematopoietic stem cell transplantation combined with gene addition or editing 6

Table 1-1. Genetic Diseases of Blood Cells and the Transplantation Modalities that Have Been Applied Clinically as Therapies or Are in Pre-Clinical Development 15

Figure 1-3. Summary of Gene Editing Pathways 19

CHAPTER 2

Figure 2-1. Targeted integration of *BTK* cDNA into intron 1 of the gene leads to sub-physiological levels of mRNA and protein expression 80

Figure 2-2. Addition of a terminal intron of the woodchuck hepatitis virus post-transcriptional regulatory element (WPRE) to donor templates increases transgenic BTK expression 82

Figure 2-3. Combining successful donor modifications leads to an additive effect on transgene expression 84

Figure 2-4. Integration of WPRE containing donors led to elevated BTK expression in non-BTK expressing Jurkat T lymphocytes 86

Figure 2-5. Chemical modification of sgRNA and Cas9 delivered as mRNA produced efficient allelic disruption and integration rates in human CD34+ mobilized peripheral blood hematopoietic stem and progenitor cells 87

Figure 2-6. Targeted integration of a corrective cDNA donor into the BTK exon 2 target 88

site led to efficient integration and expression of the transgene

Figure 2-S1. Production and verification of a BTK-deficient K562 cell sub-line 90

Figure 2-S2. sgRNA variants demonstrated improved cutting efficiency when delivered as ribonucleoprotein to human CD34+ peripheral blood stem cells 92

CHAPTER 3

Figure 3-1. Editing schema and methods of targeted DNA integration 118

Figure 3-2. Identifying integration events of plasmid donors in BTK deficient K562 cells 120

Figure 3-3. Synchronizing K562 cells in G_1 increases corrective donor integration via HITI while decreasing integration from HDR and PITCh 122

Figure 3-4. Targeted integration rates in bulk and population-sorted primary human Mobilized peripheral blood stem cells 123

Figure 3-S1. Schematics of donor template and targeted integration pathways 124

Figure 3-S2. K562 cells treated with only donor plasmid or only endonuclease produced no integration events 126

Figure 3-S3. Hydroxyurea cell cycle synchronization efficacy in K562 cells 127

Figure 3-S4. Optimizing reagents for delivery into human CD34+ mobilized peripheral blood stem and progenitor cells 128

ACKNOWLEDGEMENTS

There are many people I need to thank for their personal and professional support through my graduate studies. First and foremost are my mentors, Dr. Donald Kohn and Dr. Caroline Kuo. Dr. Kohn first accepted me into the lab as an undergraduate in 2011 with minimal prior research experience. All said and done, I have spent nearly a decade under Dr. Kohn's tutelage and he really taught most of what I know about scientific research. I have been incredibly lucky to have such a brilliant, kind, patient, and thoughtful mentor. I have always admired Dr. Kohn. As a scientist, he is unbelievably accomplished and internationally renowned, and yet he is very down to earth and always has time for those who need it whether it is a graduate student, a patient, or anyone else. He is always eager to help teach and support every member of his lab. I could not have asked for a better environment to do my doctoral studies in.

The other research mentor I want to thank is Dr. Caroline Kuo. Dr. Kuo has been a constant source of guidance and support. Nearly every experiment I have begun and every project I have undertaken was bounced off her first, and her insight has made me a much better scientist. Outside of the scientific help Dr. Kuo has always provided me with, working with her and being a part of her team has been wonderful and an absolute joy. Our casual conversations shared over chocolate about anything from recent scientific papers to the best hamburgers in Los Angeles, has made the lab feel like home and will be one of the things I miss the most.

I need to thank the entire Kohn Lab, as well. The lab has such a friendly and collaborative atmosphere, in large part due to the leaders within it. Dr. Beatrice Campo Fernandez is an absolute wizard in her mastery of essentially everything in the lab, and her contributions have been invaluable to so many aspects of my work. Dr. Zulema Romero has been a constant help throughout my time in the lab and her input and questions have helped me deepen my

understanding of the field and my own project. Dr. Roger Hollis has been another deep well of knowledge, and he often comes up with perfect solutions when things seem to be otherwise stuck. Joseph Long, a technician in the lab, has been a wonderful collaborator and a great friend throughout my graduate career.

I also want to thank the students and technicians who have assisted me on the project and that I have had the opportunity to mentor. Ali Abele, Jasmine Santos, Alex Keir, Aayushi Kapadia, Simon Maddock, and Isaac Villegas have all been huge helps to me in different ways and mentoring them has been one of my favorite parts of my time in the lab. All of them showed their ingenuity, determination, and eagerness each day and I am grateful for all the work they have performed over the years. I would particularly like to thank the two that were with me the longest: Jasmine and Isaac. Jasmine worked in the lab tirelessly for most of her undergraduate career, starting with zero research experience and leaving as a seasoned contributor. Her growth was tremendous and all thanks to her hard work and dedication. Isaac, a technician in the lab, joined with a lot of experience and quickly became my partner and right hand on the project. His independence, work ethic, and creativity were a huge asset to the science at hand.

The Broad Stem Cell Research Center's flow cytometry core run by Felicia Codrea, Jessica Scholes, and Jeffrey Calimim, taught me most of what I know about flow cytometry. They worked tirelessly to help me figure out an unusual assay and have an unbelievable amount of expertise between the three of them. Thank you to all three members of the core.

Thank you to my funding sources: the Cellular and Molecular Biology Training Program and the Virology and Gene Therapy Training Grant (Ruth L. Kirschstein National Research Service Awards GM007185 and AI060567). As well as funding sources that have supported the project: The Broad Foundation Innovative Pilot Stem Cell Research Project Award, I3T-CTSI

Discovery Award/David Geffen School of Medicine Seed Grant Program, and Orchard Therapeutics.

A huge thank you to my doctoral committee for their excellent suggestions and questions that have helped me grow and have improved the work I have undertaken. They are truly experts at the top of their respective fields, and their comments and input have helped shape the direction of the project. Additionally, the advice of Dr. Owen Witte led me to pursue Institutional Research and Academic Career Development Award (IRACDA) postdoctoral programs, one of which is where I will be doing my postdoctoral fellowship.

Another important thank you is to Dr. Zoran Galic. I had the opportunity to serve as a teaching assistant to Dr. Galic for two quarters. Dr. Galic's passion for teaching was infectious and he really helped me develop my skills. On top of that, Dr. Galic has been so supportive of me and on many occasions has been willing to share his wisdom and career advice. Getting to learn from one of the very best, kindest instructors on campus helped cement my decision to pursue collegiate teaching as a career path. I owe Dr. Galic a tremendous amount of thanks and I hope to someday be as great of an educator as he is.

A big thank you to my friends, both at UCLA and outside. Through these last six years, I have forged friendships with so many amazing people that I know will last a lifetime. They have been peers, board game opponents, travel companions, pseudo group therapists, and so much more. I could not possibly name everyone here that I want to, but I want to give a special thank you to Guillaume Urtecho for always being willing to help with things like paper editing and constructive criticism. Having a tight knit circle of friends to share the triumphs and tribulations of graduate school made the experience infinitely more enjoyable.

A special thank you to my wonderful significant other, Meztli España. She has put up with my long nights in the lab and been incredibly supportive during this difficult process. She always makes me laugh and she is always game to go on adventures big and small. She is the most thoughtful and compassionate person I have ever met. Meztli has been my teammate, my partner, and my best friend. I am so grateful to have her in my life and I look forward to taking the next our big steps together as we move together to New York.

Finally, I want to thank my family. I am so lucky to have the phenomenal family I have. It is cliché to say, but I could not have become the man I am today without their love and help. I have learned everything from them, and my parents and grandparents really instilled my passion for education. All four of my grandparents have always looked out for me and helped me become the best I could be with their advice and their examples. My parents have always been my rock. From late nights spent reading to me as a kid and homework help to seeing me off to college and graduate school, they have always been there for me. I am thankful to have them both every day.

VITA

EDUCATION:
Ph.D. Candidate, Molecular Biology, University of California, Los Angeles, Sept 2014 – Present

Bachelor of Science (B.S.), Microbiology, Immunology and Molecular Genetics, University of California, Los Angeles, Sept 2009 – June 2013

RESEARCH EXPERIENCE:
University of California, Los Angeles, Dr. Donald Kohn Lab, March 2015 – Present
Graduate Student Researcher
Targeted Correction of X-Linked Agammaglobulinemia

- Designed targeted endonucleases for use in gene therapy of primary immunodeficiencies
- Achieved site-specific stable integration and expression of cDNA donors into cell lines
- Demonstrated integration of a potentially therapeutic construct in up to 30% of primary human hematopoietic stem and progenitor cells
- Performed work enabling a sponsored research agreement with Orchard Therapeutics to continue to progress this treatment towards clinical trials

Concortis Biosystems and Moradec LLC, Jan 2014 – Aug 2014
Intern
Antibody-Drug Conjugations

- Made chemically and electrically competent cells
- Prepared buffers for HPLC
- Ran buffer exchanges to remove unconjugated drug from solution

University of California, Los Angeles, Dr. Donald Kohn Lab, Sept 2011 – June 2013
Undergraduate Researcher
Sickle Cell Gene Correction using Zinc Finger Nucleases

- Focus of improving targeted integration of corrective genes into the HPRT locus with 6-Thioguanine selection
- Maintained and harvested cell cultures under sterile conditions,
- Isolated genomic DNA, ran PCRs, transfections, Mini/Maxi Preps, Flow Cytometry

TEACHING EXPERIENCE
University of California, Los Angeles, January–March 2016 and September–December 2016
Microbiology, Immunology and Molecular Genetics 185A (Immunology)
Teaching Assistant

- Averaged at least 8.50 out of a possible 9 across every category evaluated by students
- Overall rating from student evaluations was 8.62 out of a possible 9

PUBLICATIONS

Morgan RA, Gray D, Lomova A, and Kohn DB. Hematopoietic Stem Cell Gene Therapy: Progress and Lessons Learned. 2017; 21:574-590. doi: 10.1016/j.stem.2017.10.010. PMID:29100011.

Hoban MD, Cost GJ, Mendel MC, Romero Z, Kaufman ML, Joglekar AV, Ho M, Lumaquin D, Gray D, Lill GR, Cooer AR, Urbinati F, Senadheera S, Zhu A, Liu PQ, Paschon DE, Zhang L, Rebar Ej, Wilber A, Wang X, Gregory PD, Holmes MC, Reik A, Hollis RP, and Kohn DB. Correction of the sickle cell disease mutation in human hematopoietic stem/progenitor cells. Blood. 2015; 125(17):2597-2604. doi: 10.1182/blood-2014-12-615948. PMID: 25733580.

York AG, Williams KJ, Argus JP, Zhu Q, Brar G, Vergnes L, Gray EE, Zhen A, Wu NC, Yamada DH, Cunningham CR, Tarling EJ, Wilks MQ, Casero D, Gray DH, Yu A, Brooks DG, Sun R, Kitchen SG, Wu TT, Reue K, Stetson DB, and Bensinger SJ. 2015. Limiting cholesterol biosynthetic flux engages type I IFN signaling in a STING-dependent manner. Cell. 163: 1716–1729. doi: 10.1016/j.cell.2015.11.045. PMID: 26686653.

AWARDS AND PATENTS

Sidney Rittenberg Award recognizing outstanding achievement in research presented by the UCLA department of Microbiology, Immunology, and Molecular Genetics October 23rd, 2019.

John-Eiserling-Lengyel Teaching Award presented by the UCLA Molecular Biology Institute December 18th, 2018

UCLA Microbiology, Immunology, and Molecular Genetics Department Teaching Assistant Award March 23rd, 2018

UCLA Molecular Biology Institute’s Annual Teaching Assistant Award March 17th, 2018

Virology and Gene Therapy T32 Training Grant September 2016 – August 2017
(Ruth L. Kirschstein National Research Service Award AI060567.)

Cellular and Molecular Biology T32 Training Grant July 2015 – June 2016 (Ruth L. Kirschstein National Research Service Award GM007185.)

Kohn DB, Kuo CY, Gray DH, Garcia R, Hollis RP, Lomova A. Gene Editing of Monogenic Disorders in Human Hematopoietic Stem Cells – Correction of x-Linked Agammaglobulinemia (XLA). WO/2019/232253. Filed May 2019.

CHAPTER 1

INTRODUCTION

Hematopoietic Stem Cell Gene Therapy – Progress and Lessons Learned

Richard A. Morgan[1], David Gray[2], Anastasia Lomova[1] and Donald B. Kohn[1,3,4,5]

Affiliations:

1 Department of Microbiology, Immunology & Molecular Genetics, University of California, Los Angeles, CA 90095, USA

2 Molecular Biology Institute, University of California, Los Angeles, CA 90095, USA

3 Department of Pediatrics, University of California, Los Angeles, CA 90095, USA

4 Department of Molecular and Medical Pharmacology, University of California, Los Angeles, CA 90095, USA

5 Eli & Edythe Broad Center of Regenerative Medicine & Stem Cell Research, University of California, Los Angeles, CA 90095, USA

Reprinted with permission from: Morgan, R.A., Gray, D., Lomova, A., and Kohn D.B. (2017). Hematopoietic Stem Cell Gene Therapy – Progress and Lessons Learned. *Cell Stem Cell*, 21(5), 574-590. Copyright 2017 Elsevier Inc.

1.1 ABSTRACT:

The use of allogeneic hematopoietic stem cells (HSCs) to treat genetic blood cell diseases has become a clinical standard but is limited by availability of suitable matched donors and potential immunologic complications. Gene therapy using autologous HSCs should avoid these limitations and thus may be safer. Progressive improvements in techniques for genetic correction of HSCs, by either vector gene addition or gene editing, are facilitating successful treatments for an increasing number of diseases. We highlight the progress, successes, and remaining challenges toward development of HSC gene therapies and discuss lessons they provide for development of future clinical stem cell therapies.

1.2 INTRODUCTION

Most inherited blood cell diseases, such as primary immune deficiencies, hemoglobinopathies, storage and metabolic disorders, congenital cytopenias and stem cell defects, can be treated by transplantation of allogeneic hematopoietic stem cells (HSCs) (Table 1).[1,2] The transplanted genetically normal HSCs can serve as an ongoing source of blood cells of all lineages, eliminating these disorders from a single treatment with benefits lasting life-long.

While there are generally high rates of success when an HLA-identical sibling donor is available, the outcomes of hematopoietic stem cell transplantation (HSCT) are usually not as successful with less well-matched allogeneic donors (either haplo-identical family members or unrelated donors).[1,2] Reduced HLA matching between recipient and donor increases the risks of graft rejection and graft versus host disease (GVHD). Rejection of an HSC graft generally leaves the patient in a perilous position, with an urgent need to restore hematopoiesis to prevent complications from prolonged pancytopenia (anemia, infection, bleeding). The primary donor may

not be available (e.g. cord blood units are not linked to their source) and a suitably matched second donor may not be identified. GVHD is a major cause of transplant morbidity and even mortality, and can impose a chronic rheumatologic-like inflammatory/fibrotic disease, with need for persistent immune suppression and the attendant risks of infection and toxicities.[3] Immediately before and after the allogeneic transplant, high levels of immune suppression are necessary to reduce immunological risks but these treatments also add to morbidity. There has been ongoing progress with methods to reduce GVHD in allogeneic HSCT, including improved graft engineering by removal of selective T cell populations (TCR α/β depletion, naïve T cell depletion), and by use of post-transplant cyclophosphamide.[4,5] Nonetheless, immune complications and lack of suitable matched donors present significant clinical barriers to successful application of allogeneic HSCT for a wider range of disorders.

Autologous HSCT in which the patient's HSCs are gene-modified should offer complete avoidance of the major immunological complications of allogeneic HSCT, which may contribute to better outcomes for patients with genetic blood cell disorders. For specific disorders, expression of the gene introduced into HSCs is needed in cells of one or more hematopoietic lineages (e.g. red blood cells, neutrophils, lymphocytes) (Figure 1). The lack of immunogenicity with autologous cells allows the use of reduced intensity of the pre-transplant conditioning to make space in the marrow niche to facilitate HSC engraftment, compared to what is required for effective allogeneic HSCT (Figure 2).

Current approaches to autologous transplant/gene therapy using lentiviral vectors (LVs) have produced clinical benefits similar to those from allogeneic transplant for several disorders.[6–9] In multiple clinical trials (Table 1), this approach has consistently achieved quite stable frequencies of gene-corrected blood cells of all lineages, indicating engraftment, long-term

persistence and ongoing generative capacity of gene-modified HSCs, with no significant diminution observed over time in human subjects.[10,11]

Recent developments in gene editing have led to investigations toward its application for *ex vivo* gene correction in HSCs, which may have advantages compared to integrating viral vector-mediated gene addition.[12,13] This review will present the primary approach that is currently being used for gene modification of HSCs for clinical applications and gene addition using integrating viral vectors, as well as discuss the current status of gene editing in human HSCs for autologous transplantation. Lessons learned from advancing HSC therapies to the clinic may help inform the development of other stem cell therapies.

1.3 HSCs FOR GENE THERAPY

HSCs are long-lived and multipotent, so gene correction in HSCs should lead to persistent gene correction among the different lineages.[14] The hematopoietic system is an ideal target for gene therapy because of the ease with which HSCs can be accessed for *ex vivo* gene manipulation, effective gene-modification, and re-administration as an intravenous infusion. HSCs are traditionally harvested from bone marrow derived from the iliac crests under general anesthesia. Multiple aspirations are performed with the goal of collecting 10–20 ml of bone marrow per kilogram of recipient body weight. Alternatively, HSCs can be obtained as cytokine (e.g. G-CSF)-mobilized peripheral blood stem cells (PBSC) collected by leukopheresis. Hematopoietic growth factors, including GM-CSF and G-CSF, or CXCR4 inhibitors have been shown to increase the numbers of circulating hematopoietic stem and progenitor cells (HSPC) by 30–1000 fold.[15] PBSCs are now the predominant clinical HSC source used for allogeneic and autologous transplants to routinely and successfully treat multiple blood cell disorders using current techniques.

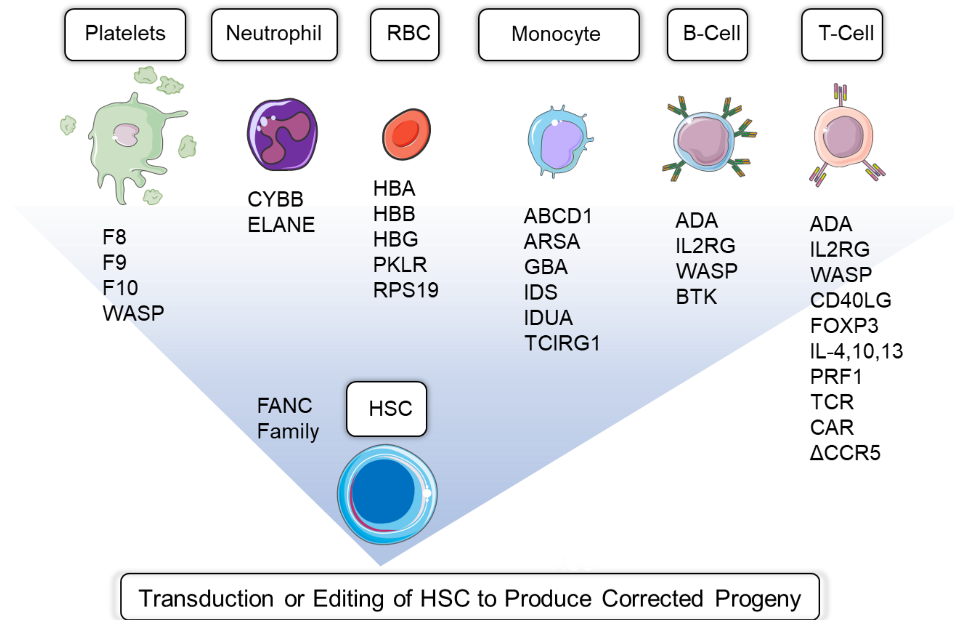


Figure 1. Overview of Targets for Gene Therapy

Hematopoietic stem cells (HSCs) isolated from bone marrow can be modified *ex vivo* and transferred back to the recipient to produce functional, terminally differentiated cells. Specific cellular targets and the relevant diseases and genes for gene therapy include the following. HSCs: Fanconi anemia (FANC A-F). Platelets: Hemophilia A (Factor VIII (F8)); Hemophilia B (Factor IX (F9)); Factor X deficiency (Factor X (F10)); Wiskott-Aldrich Syndrome (Wiskott Aldrich Syndrome Protein (WASP)). Neutrophils: X-linked Chronic Granulomatous Disease (Cytochrome B-245 Beta Chain (CYBB)); Kostmann's Syndrome (Elastase Neutrophil Expressed (ELANE)). Erythrocytes: Alpha-Thalassemia (Hemoglobin Subunit Alpha (HBA)); Beta-Thalassemia and Sickle Cell Disease (Hemoglobin Subunit Beta (HBB)); Pyruvate Kinase Deficiency (Pyruvate Kinase, Liver and RBC (PKLR)); Diamond-Blackfan Anemia (Ribosomal Protein S19 (RPS19)). Monocytes: X-linked Adrenoleukodystrophy (ATP Binding Cassette Subfamily D Member 1 (ABCD1)); Metachromatic Leukodystrophy (Arylsulfatase A (ARSA)); Gaucher disease (Glucosylceramidase Beta (GBA)); Hunter Syndrome (Iduronate 2-Sulfatase (IDS)); Mucopolysaccharidosis type I (Iduronidase, Alpha-L (IDUA)); Osteopetrosis (T Cell Immune Regulator 1 (TCIRG1)). B Cells: Adenosine deaminase (ADA)-deficient Severe Combined Immunodeficiency (Adenosine Deaminase (ADA)); X-linked severe combined immunodeficiency (Interleukin 2 Receptor Subunit Gamma (IL2RG)); Wiskott-Aldrich Syndrome (Wiskott Aldrich Syndrome Protein (WASP)); X-linked agammaglobulinemia (Bruton's Tyrosine Kinase (BTK)). T Cells: Adenosine Deaminase (ADA)-deficient Severe Combined Immunodeficiency (ADA); X-linked severe combined immunodeficiency (IL2RG); Wiskott-Aldrich Syndrome Protein (WASP); X-linked Hyper IgM syndrome (CD40 Ligand (CD40LG)); IPEX Syndrome (Forkhead Box P3 (FOXP3)); Early Onset Inflammatory Disease (Interleukin 4, 10, 13 (IL-4, 10, 13)); Hemophagocytic Lymphohistiocytosis (Perforin 1 (PRF1)); Cancer (Artificial T cell receptors (TCR), Cancer; Chimeric Antigen Receptor (CAR)); Human immunodeficiency virus (C-C Motif Chemokine Receptor 5 (CCR5)).

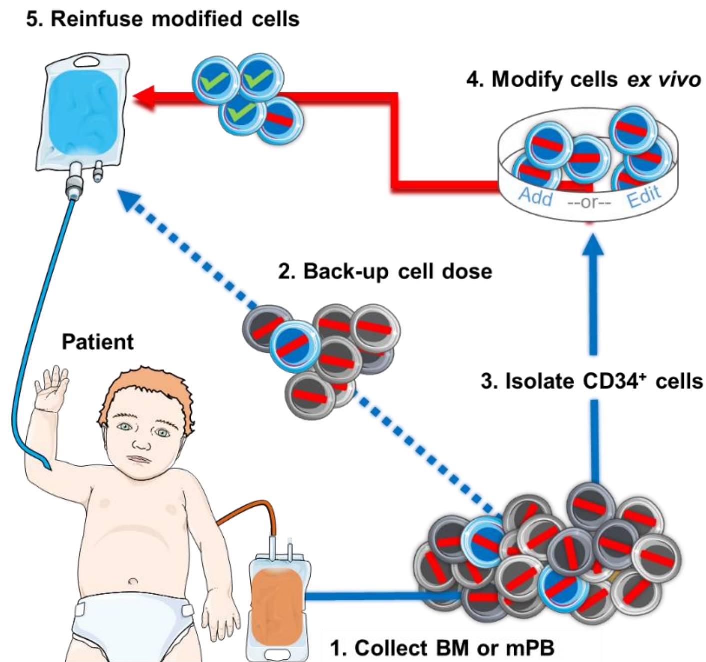


Figure 2. Autologous hematopoietic stem cell transplantation combined with gene addition or editing
(1) Bone marrow (BM) or mobilized peripheral blood (mPB) cells are collected from the patient (red line represents a disease-causing mutation). Typically, 15–20ml of BM/Kg is an acceptable harvest target. While collecting HSCs by mobilization and apheresis is less invasive than BM aspiration, infants have small blood volumes making leukapheresis challenging. Failure to harvest adequate cell numbers can prevent therapy. (2) Modification of HSCs may reduce stem cell capacity. A back-up cell dose of non-modified cells is apportioned to restore native hematopoiesis in the event of graft failure. (3) CD34+ cells are isolated in a GMP-compliant, closed system. Purification of HSCs may reduce total cell number as CD34+ HSCs represent less than one percent of total cells. Alternatively, a CD34+/CD38− enrichment strategy may be employed to further purify HSCs and lower the amount vector required for modification. CD34+ cells may be pre-stimulated *ex vivo* for 1–3 days prior to modification, depending on the protocol. (4) Gene modification of HSCs must be permanent so as to be passed down to all progeny. Cells are modified by either a viral vector to add a gene (typically requires high concentration vector), or targeted nucleases with/without a donor template to disrupt, correct, or insert a gene. After *ex vivo* modification, the cell product undergoes release testing to assess purity, identity, safety, potency (transduction/editing efficiency), and other characteristics. If the modification strategy requires selection of corrected cells, low cell yield may prevent transplantation. (5) Prior to receiving the cell product, the patient undergoes conditioning to “make space” for engraftment of modified HSCs (green check represents successful modification of a disease-causing gene). Modified cells may be reinfused fresh or cryopreserved for delivery at a later time. While high-levels of cytoreductive agents may be toxic, inadequate conditioning may result in poor engraftment.

However, the use of HSCs for gene therapy presents several challenges. HSCs are rare and delicate and are found among large numbers of more committed progenitors and mature blood cells that do not have long-term repopulating activity. While the immunophenotypic definition of unitary human HSCs has been well-developed, (e.g. CD34+, CD38−, CD45RA−, CD90+, CD49f+, purification to high levels at clinical scale may entail significant losses of cells and impair their stem cell capacity.[16] In current clinical practice for gene therapy, the HSCs from the clinical source (bone marrow or mobilized peripheral blood stem cells) are enriched, rather than purified, usually by isolating the CD34+ fraction using immunomagnetic separation. The CD34+ population (~1% of cells in adult bone marrow) contains most long-term engrafting multipotent HSCs, but also far more numerous short-term progenitor cells. CD34 selection enables ~30–50-fold enrichment of HSCs, removing the majority of highly numerous mature blood cells and enriching the HSC targets to culture for *ex vivo* gene modification. The dosages of CD34-selected cells typically used for transplantation range from 2 to 20 million/kg, necessitating efficient processing of relatively large numbers of cells.

Because they will divide many times, any gene modification of HSCs needs to be permanent and heritable to be passed on to all successive generations of progeny cells. Currently this necessitates making changes in the genome, either by covalent gene addition with an integrating vector or direct genome editing. The critical technical challenge for successful HSC gene therapy is performing sufficient gene engineering of the autologous HSCs to provide a therapeutic level of permanent genetic correction without impairing their stem cell capacity or causing adverse effects.

Thresholds for sufficiency can be based on observations from cases where patients, allo-transplanted for these disorders, develop mixed chimerism with only a sub-fraction of the

hematopoiesis coming from donor cells. Clinical improvement has been reported with donor chimerism as low as 10–30% for sickle cell disease, thalassemia, SCID, and other PIDs, making this level a reasonable target for engrafted, gene-corrected HSCs.[17,18]

1.4 VECTORS

1.4.1 VECTOR CHOICE AND DESIGN

An attractive property of retroviruses is their ability to convert their RNA genome into proviral DNA through reverse transcription and integration into the DNA of the host cell's genome in a quasi-random fashion. This integrating property of retroviruses allows the transmission of therapeutic information to all progeny of a transduced HSC. The initial retroviral systems used were derived from Murine Leukemia Viruses (MLV) a class of simple gammaretroviruses (gRV) that were well-known from studies of their oncogenic properties. Transduction of human HSCs with gRV vectors has remained challenging due to the quiescent nature of HSCs that typically cycle infrequently during steady state hematopoiesis.[19–21] MLV requires the breakdown of the nuclear envelope and cellular progression through mitosis to stably integrate into host cell genome as the virus lacks active nuclear localization elements.[22,23] Therefore, HSCs need to be cultured for several days with multiple cytokines to induce cycling for retroviral transduction, and this may lead to loss of stem cell capacity.

Lentiviral vectors (LVs) have subsequently become the vector platform of choice because they do not require the cells to undergo mitosis for the breakdown of the nuclear membrane to efficiently integrate their proviral DNA into host cells. Rather, LVs transit through the nuclear pores by recruitment of host cell proteins. The most widely used design of LV system used for transduction of HSCs was first developed by Naldini et al in 1996 and was subsequently shown to

efficiently transduce HSCs by Miyoshi et al and Case et al, among others.[24–26] While clinically-effective methods have been developed for gene introduction to human HSCs using retroviral and lentiviral vectors, HSCs are relatively resistant to transduction, requiring the use of high multiplicities of infection of vector (e.g. 10–100, based on titers measured on permissive cells) to effectively modify the majority of HSCs.

1.4.2 SAFETY ISSUES

The first clinical trial utilizing gene therapy to modify autologous HSCs with curative intent began in 1992. The goal of this first trial conducted by Bordignon et al was to correct severe combined immune deficiency (SCID) syndrome caused by deficiency in adenosine deaminase (ADA).[27] SCID patients experience severe, recurrent and persistent infections resulting from immunodeficiency and, prior to the availability of HSCT options, the disease was lethal. This first clinical trial employing autologous gene-corrected HSCs to correct ADA-SCID utilized a vector derived from MLV to introduce an ADA cDNA into HSCs isolated from afflicted patients.[28,29] These investigators and other groups in the U.K. and the U.S. have gone on to treat more than 45 ADA SCID patients with gRV vectors, with good immune recovery in most and no complications from the vectors.[30–32]

Additional trials using MLV-based gRV vectors, such as those conducted by groups in France and the U.K. for SCIDX1 (X-linked SCID) to correct interleukin 2 common gamma chain (IL2Rg) deficiency, demonstrated both the utility and the limitations of gRV vectors.[33] Although curative in the majority of patients, five (of 20 total) patients developed T cell acute lymphoblastic leukemia (T-ALL), two to six years post treatment, as a result of the action of vectors that had integrated near proto-oncogenes.[34,35] Similar leukoproliferative complications were seen in other clinical trials using gRV vector for X-linked Chronic Granulomatous Disease (X-CGD), and

Wiskott-Aldrich Syndrome (WAS).[36,37] The occurrence of T-ALL or myeloid malignancies in subsets of patients from each of these clinical trials was a result of the LTR driven gRV vector landing upstream of proto-oncogenes and ectopically activating their expression.[34,35] It is now known that gRVs tend to land near transcriptional start sites of genes, CpG islands, and DNAse 1 hypersensitive sites (which tend to be transcriptionally active).[38–42] The LTRs of these gRV act as strong enhancers that recruit a number of transcription factors capable of overriding innate cellular transcriptional control of neighboring genes, promoting leukemogenesis.[43]

To address MLV's propensity to induce enhancer-mediated insertional mutagenesis, Gilboa et al developed the first self-inactivating (SIN) vector by introducing a deletion within the 3′ U3 that abolishes enhancer activity. During reverse transcription, the deleted 3′ U3 is copied to both ends of provirus DNA and deprives the provirus of LTR mediated enhancer and promoter activities.[44] Instead, internal promoters can be introduced to drive transgene expression with higher regulated/tissue specific expression. This alteration provides SIN vectors with increased safety by reducing cellular gene activation when in proximity to neighboring promoters.[45] SIN gRV vectors have been used safely for subsequent studies of gene therapy for SCIDX1, indicating that this modification did achieve its goal of greatly reducing risks of insertional transformation.[46]

These studies informed the development of LV as vehicles for gene delivery. Self-inactivating deletions were introduced into the viral LTRs of LV and all sequences encoding proteins supporting HIV virulence were deleted from the provirus and all packaging constructs to create second (-Vpr, -Vif, -Vpu, -Nef) and third (also -Tat) generation LV vector systems.[47,48] The VSV-G glycoprotein is most commonly used to pseudotype lentiviral vectors, although other envelope proteins have shown some favorable properties.[49] Lentiviral vectors can transduce non-dividing cells via several mechanisms they have for nuclear import of their viral cores.[50] They also

have somewhat larger carrying capacity than gRV vectors (6–9 kb) and are generally more robust for transducing human cells. They have mostly become the vector of choice for stable gene addition to human HSCs.

1.4.3 METHODS TO OPTIMIZE VECTOR DELIVERY

A major limitation of gRV vectors is their inability to transduce non-dividing cells efficiently. Addition to culture dishes of a specific adhesion domain of fibronectin in a recombinant protein, CH-206, which recruits virus particles to HSCs (resulting in a higher MOI at the interface between virus and cell) was found to significantly increase transduction efficiency.[51] Ex-vivo culture conditions were also found to influence HSC proliferation and transduction efficiency.[52,53] Addition of recombinant human hematopoietic growth factors (typically ckit ligand, Flt-T ligand, thrombopoietin, and Interleukin-3) during transduction resulted in activation of CD34+ HSCs and therefore higher transduction rates. The combination of fibronectin and optimal ex-vivo culture conditions greatly improved transduction with proven success in the clinic, as discussed below.[54]

Another method that has been used to enhance transduction is to alter the cell target specificity of a viral vector by exchanging the innate envelope protein for one derived from an alternative virus allows researchers to alter the tropism of resultant vector particle (called pseudotyping). Typically, pseudotype is chosen based on expression level and exclusivity of the envelope protein's cognate receptor (higher receptor levels equal greater gene transfer levels). HSCs can be transduced with RD114, GALV, BaEV and VSV-G pseudotype viruses, among others. Other, more specific pseudotyping strategies are being developed that include the use of diverse viral envelopes and even fusions proteins with antibodies or cytokines to target specific cell types.[55,56]

1.4.4 ADVANCES IN VECTOR DESIGN

Surprisingly, there have been no significant improvements to the basic design of lentiviral vectors since the so-called third generation vectors were introduced almost two decades ago.[48,57] They were designed based on a decade of experience with gammaretroviral vector design and production and have met all the safety expectations, with no report of emergence of replication-competent lentivirus during packaging in research or clinical manufacturing.

Lentiviral vectors with relatively small and simple gene cassettes (e.g. human phosphoglycerate kinase gene or elongation alpha-1 gene minimal promoters and a cDNA) are readily produced to titers sufficiently high for effective gene modification of human HSCs at clinical-scale. However, low vector titer remains a significant problem with some LVs, especially those tasked to carry larger transgene cassettes, such as the human beta-globin gene. Several studies have shown that increasing viral RNA genome length negatively affects both titer and transduction efficiency.[58,59] Reduction of viral RNA length through removal of non-essential sequences is a viable strategy for improving LV titer for large transgene cassettes. Additionally, codon optimization may be used to improve titer by depleting secondary structures detrimental to mRNA stability (with the added benefit of improving transgene expression levels).[60] Other strategies for increasing LV titer include the addition of the Woodchuck Hepatitis Virus Post-transcriptional Regulatory Element (wPRE), which has been shown to increase vector titer through increasing vector genomic RNA stability, export and translation or addition of elements known to improve polyadenylation of vector mRNA during packaging, such as the Bovine Growth Hormone Polyadenylation Sequence.[61–63] Recently, Vink et al described a reconfiguration of the basic arrangement of the key cis-regulatory elements of HIV included in the vector backbone (LTRs, primer binding site, rev-responsive element) to simplify the process of reverse transcription, which

may be a limit to transduction with large vectors.[64] The advantages of this design in clinical applications remain to be tested.

To date, lentiviral vectors have been produced by transient transfection of multiple plasmids (vector, packaging proteins, envelope), which is cumbersome and challenging to scale-up to levels that would be needed for commercial production. There have been extensive and long-standing efforts to develop stable lentiviral vector packaging lines, similar to those routinely produced for gammaretroviral vectors, which could provide a master cell bank for much simpler production of vector lots. It has proven to be challenging to repress expression of the VSV-G glycoprotein (which remains the most effective pseudotype for lentiviral vectors in most cases) and some of the HIV-1 genes which may be cytotoxic to the packaging cells for cell passaging, but then rapidly and robustly inducing expression of these genes for vector production. One stable packaging cell line has been used to produce lentiviral vector for clinical trials.[65,66] It was made by successively transfecting plasmids encoding each gene cassette needed to make the virus proteins (HIV gag/pol, VSV-G) under tight, inducible expression control, and then concatamers of the vector plasmid to obtain multiple copies. This led to a cell line that was capable of scale-up to production volumes needed for production of clinical lots without loss of the packaging capacity.

1.4.5 GENE EXPRESSION FROM LENTIVIRAL VECTORS

For some gene therapy applications, unregulated constitutive, ubiquitous expression of the transgene is acceptable. For example, adenosine deaminase (ADA) is expressed in all cell types and a broad range of ADA enzyme activity in all blood cell lineages is safe and sufficient to allow immune reconstitution. Thus, the vectors deployed for gene therapy of ADA SCID have used constitutive promoters, such as the MLV LTR or the Elongation Factor 1-Alpha gene promoter.[67,68] Other genes may require precise lineage, temporal or physiological-responsive expression patterns

to be safe and effective. For example, molecules involved in signal transduction (receptors, intracellular signaling molecules, transcription factors) may be expressed in only specific cell types, or under specific physiologic states or in response to specific stimuli (e.g. BTK, CD40 ligand, JAK3, Stat proteins).[69] In some cases, it has been possible to build vectors using transcriptional control elements from endogenous cellular genes to apply regulated expression of transgenes (e.g. Beta-globin transcriptional control elements directing erythroid-specific expression of beta-globin.[70] However, it may not be possible to incorporate into a vector the necessary regulatory sequences to recapitulate endogenous gene expression patterns for vectors integrated at an array of chromosomal sites in different cells. Insulator elements have been incorporated into some vectors to attempt to mitigate potential silencing of vector expression by heterochromatinization or trans-activation of adjacent cellular genes.[71,72] The benefits of insulators remain theoretical and, in fact, silencing and trans-activation has been a problem with current lentiviral vectors. Endogenous gene expression is often controlled by enhancer and promoter interactions that occur over long distances (up to 100kb) away.[73] Regulated gene expression can be achieved to variable extents by adding a gene's known enhancer(s) upstream of a minimal promoter within the LV. These enhancers recruit and bind specific sets of transcription factors to cause cis activation of the promoter.[74] The best studied example is the use of multiple elements from the β-globin gene locus to achieve erythroid-specific expression for the treatment of hemoglobinopathies.[70,75] However, it is sometimes difficult to fit all of the necessary cis-acting genetic elements required for precise gene expression within the size limit of vector genomes. In some cases, the function of enhancers can be antagonized by repressive chromatin structure at some LV integration sites.[76,77] Chromatin domain insulators can be added into the LV's LTRs to

overcome these positional effects and may also reduce risks from internal enhancers of vectors from affecting neighboring genes.[78]

Table 1. Genetic Diseases of Blood Cells and the Transplantation Modalities that Have Been Applied Clinically as Therapies or Are in Pre-clinical Development

		Transplantation Modalities Applied			
Category of Disease	Specific Conditions	Allogeneic HSCT	γ-Retroviral Gene Therapy	Lentiviral Gene Therapy	Genome Editing
Primary immune deficiencies	ADA-deficient severe combined immune deficiency	+	+	+	n.d
	X-linked severe combined immune deficiency	+	+	+	pre-clinical
	other genetic forms of SCID (Artemis, Rag1/2).	+	n.d	pre-clinical	n.d
	Wiskott-Aldrich Syndrome	+	+	+	pre-clinical
	chronic granulomatous disease	+	+	+	pre-clinical
	leukocyte adhesion deficiency	+	+	pre-clinical	n.d
	hemophagocytic lymphohistiocytosis	+	n.d	pre-clinical	n.d
	X-linked hyper IgM syndrome	+	n.d	n.d	pre-clinical
	X-linked lymphoproliferative disease	+	n.d	n.d	n.d
	X-linked Agammaglobulinemia	Few	n.d	n.d	pre-clinical
	common variable immunodeficiency	heterogeneous genetic etiologies, often unknown; may require editing	heterogeneous genetic etiologies, often unknown; may require editing	heterogeneous genetic etiologies, often unknown; may require editing	heterogeneous genetic etiologies, often unknown; may require editing
Hemoglobinopathies	sickle cell disease	+	n.d	+	pre-clinical
	beta-thalassemia	+	n.d	+	pre-clinical
Storage and metabolic disorders	Gaucher Disease and other lipidoses	+	+	n.d	n.d
	mucopolysaccharidoses (I-VII)	+	n.d	+	n.d
	X-linked Adrenoleukodystrophy	+	n.d	+ recent phase II/III	n.d
	metachromatic leukodystrophy	+	n.d	+	n.d
	osteopetrosis	+	n.d	pre-clinical	n.d
Congenital cytopenias and stem cell defects	Fanconi anemia	+	+	+	pre-clinical
	Schwachman-Diamond Syndrome	+	n.d	pre-clinical	pre-clinical
	Kostmann's Syndrome	+	n.d	n.d	n.d

1.4.6 LENTIVIRAL VECTORS IN CLINICAL TRIALS

Clinical trials using LVs began in the mid 2000's and LV have now been used safely and effectively in multiple studies for almost a dozen disorders using hematopoietic stem cells (Table 1) (and in many of the T cell-based immunotherapies with Chimeric Antigen Receptors and T Cell Receptor genes). [75,79,80] In most trials, gene delivery to HSCs was at sufficiently high levels to produce clinically beneficial levels of gene-modified HSCs and relevant mature hematopoietic

cells for the treatment of disease. The absence of GVHD and the reduced amounts of conditioning chemotherapy needed for engraftment of the autologous HSCs has allowed the predicted improved safety profiles. Analyses of LV integration site in the blood cells of subjects in the different clinical trials have shown a remarkably consistent pattern, with no predilection for insertion near proto-oncogenes and no clinically significant clonal expansions.[81,82] LV gene therapy is being developed for several other disease indications (Table 1), including additional primary immune deficiencies, storage and metabolic diseases, and stem cell defects such as Fanconi's Anemia. Limitations to wider applications for more clinical indications are now less biological and more logistic, as each disorder requires its specific vector and the entire pre-clinical drug development pathway. Funding by research grants to perform vector development for additional indications may become more difficult to obtain, as the scientific novelty is diminishing; while funding from pharmaceutical companies and venture capitalists relies on the expectations of financial returns, which are unknown for these relatively rare orphan disorders. Issues related to commercial marketing and reimbursement for these cell and gene therapies are complex.[83,84] Additionally, the capacity to produce the large volumes of clinical-grade LVs for these studies (and also for the larger studies of CAR T cells), is limited and may slow progress. Nonetheless, gene therapy using HSCs is continuing to advance and provide effective and safe therapies for a growing list of disorders.

1.4.7 X-LINKED ADRENOLEUKODYSTROPHY (X-ALD) HSC GENE THERAPY

A recent report described efficacious HSC gene therapy for X-linked Adrenoleukodystrophy (X-ALD), a progressive neurodegenerative disorder with onset in boys during the first 1–2 decades.[85] The major premise for the approach is that the gene-corrected engrafted HSC will produce cells that become CNS microglia and provide essential enzyme activity that can rescue very long chain fatty acid catabolism and prevent demyelination. Graft

versus host disease, which seems to accelerate progression of X-ALD, is avoided with the use of autologous cells.[86] The findings of beneficial stabilization of neurologic status now extend and expand upon those of the initial report of two X-ALD patients treated using a lentiviral vector into CD34+ PBSC**1.[79] The larger cohort studied by Eichler et al. (n=17) showed a high rate of response, with only one patient having progressive neurologic deterioration. HSC gene therapy effectively arrested disease progression in 88%, either without symptoms or with only initial progression to early neurological impairments followed by stabilization of neuro-imaging studies and clinical function.

This was the first commercially-sponsored clinical trial of HSC gene therapy (bluebird bio), with essentially all prior trials done at academic medical centers as research investigations. It was performed as a Phase II/II trial intended to obtain data to support applications for regulatory approval to market this stem cell gene therapy. Centralized manufacturing of the stem cell product was performed under full GMP conditions. Stem cell mobilization and leukopheresis was performed at several academic clinical sites enrolling subjects and shipped to a contract manufacturing organization for stem cell enrichment and transduction, yielding a cryopreserved product, which was shipped back to the clinical site following completion of product release testing. They produced consistently high quality cell products with good cell dosages (6–19x10^6 CD34+ cells/kg) and gene transduction levels (0.5–2.5 vector copies per cell). Patients were treated within 2–3 months from enrollment, which is an excellent time-frame to produce and certify a gene-modified stem cell product and perform all the clinical evaluations prior to a HSC transplant. The high rate of successful outcomes in terms of halting neurologic progression meet or exceed those of the clinical alternative of unrelated or haplo-identical donor transplants, strongly supporting the use of autologous gene therapy for this disorder (and by extension to many other

storage diseases). Bluebird bio is pursuing regulatory drug approval for this combined cell and gene therapy product, which would represent one of the first approved gene therapies (bluebird bio, Inc., 2017).

1.5 GENE EDITING

Viral mediated gene transfer has well established benefits and has demonstrated clinical efficiency, as described above. However, risks and drawbacks of these methods remain, such as insertional oncogenesis and modified transgene expression pattern.[87–89] Targeted gene editing allows for site-specific genome modification, and thereby eliminates the risks posed by randomly inserting genes. Another advantage of targeted gene editing over viral vectors is the ability to retain endogenous control of gene expression.[90,91] Targeted gene editing can be achieved by employing site-specific endonucleases to induce a double-stranded break (DSB) in the DNA near the mutation site. The recruitment of DNA repair proteins to the site of DNA damage stimulates DNA DSB repair via one of two main pathways: non-homologous end joining (NHEJ), or homology-directed repair (HDR) (Figure 3). NHEJ can be used to achieve gene disruption with site-specific induction of a DSB by a targeted nuclease and NHEJ-mediated introduction of small insertions and deletions (Indels). NHEJ is a quick but error-prone pathway; HDR is precise, but is dependent on cell cycle phase.[92,93] Both of these endogenous DNA repair mechanisms can be harnessed for therapeutic benefit. In mammalian cells, NHEJ is more prevalent than HDR.[94] For a more detailed review on these DNA damage repair pathways, please see the included references.[95–99]

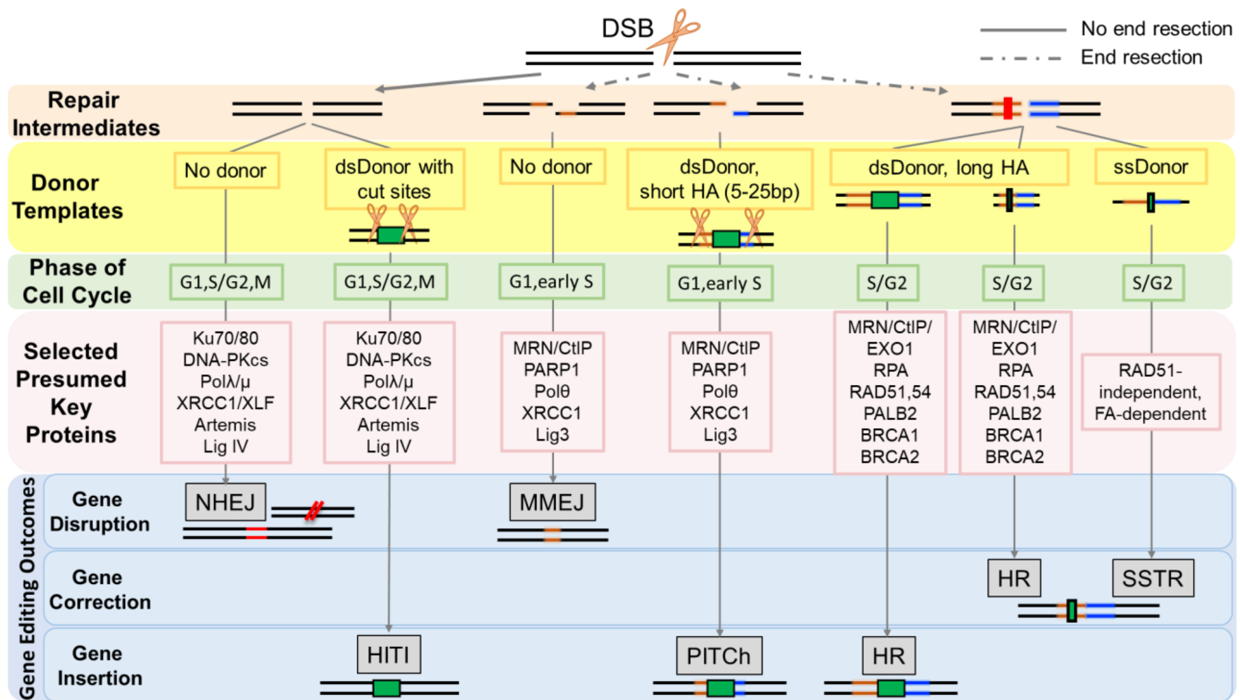


Figure 3: Summary of Gene Editing Pathways

Double stranded break (DSB) is induced by a targeted nuclease (represented by scissors). DSB ends may or may not be resected (dashed or solid line, respectively),. The ultimate gene editing outcome (light blue boxes on the bottom) depends on several factors: the type of donor template provided (yellow box), the phase of cell cycle (light green box) and the presumed DNA repair proteins available (pink box). Gray boxes indicate the names of the repair mechanisms. It should be noted that the figure illustrates the common pathways described to date, however modification of DNA repair pathways and their utilization for gene editing purposes is an area of active research. (From left to right). A DSB with no end resection and no donor available is likely to result in insertions and deletions (indels) and lead to gene disruption via the non-homologous end joining (NHEJ) pathway. NHEJ may occur in any phase of cell cycle. Exogenously providing a double-stranded donor (dsDonor), which contains nuclease cut sites (scissors) around the gene of interest (green rectangle), may result in homology-independent targeted integration (HITI). The presence of microhomology on opposite strands of DNA around the cut site may result in gene disruption via the microhomology-mediated end joining (MMEJ) pathway. A recently reported method of gene integration, termed precise integration into target chromosome (PITCh), utilizes MMEJ machinery to integrate a gene of interest, which is provided by dsDonor with short homology arms (HA) to the DNA (HA are highlighted in orange and blue). The three pathways on the right are generally only active in S/G2 phases of cell cycle and may be used to correct a single nucleotide mutation in the DNA (represented by a red line). Exogenously providing a dsDonor with long homology arms may lead to either gene integration or gene correction via homologous recombination (HR) mechanism, depending on the length of the donor template. A new type of repair mechanism for gene correction was recently described, termed single stranded template repair (SSTR). Although resulting in the same outcome as HR-mediated gene correction, SSTR is presumed to utilize the Fanconi Anemia (FA) pathway and be RAD-51 independent.

1.5.1 FOUR MAIN CLASSES OF ENDONUCLEASES FOR GENE EDITING

1.5.1.1 HOMING ENDONUCLEASES

The first, homing endonucleases recognize DNA sequences up to 40 bp long. These proteins are naturally found in six structural families.[100] The LAGLIDADG endonuclease family can be engineered to modify the sequence of DNA they recognize. The process is difficult and time consuming. An alternative being explored to facilitate retargeting HEs has been dubbed "megaTAL".[101] These megaTALs feature a DNA binding domain composed of transcriptional activator like (TAL) effector DNA recognition motifs with the active endonuclease domain of a meganuclease.

1.5.1.2 ZINC FINGER NUCLEASES (ZFNs)

Zinc finger nucleases (ZFNs), the second generation of engineered targeted endonucleases, provide a more easily modified system than homing endonucleases.[102] The enzyme functions as a dimer, with each ZFN containing three to five Zinc Finger protein motifs which recognize 3 base pair sequences of DNA and half of the FokI endonuclease complex. When a pair of ZFNs with appropriate target sequence binds closely enough for their FokI domains to dimerize, they make a DSB. ZFNs have promising function but are limited by the complexity of engineering new pairs.

1.5.1.3 TRANSCRIPTION ACTIVATOR LIKE EFFECTOR NUCLEASES (TALENs)

Transcription activator like effector nucleases (TALENs) function similarly to ZFNs but use a different mechanism to recognize specific regions of DNA.[103] Instead of the zinc fingers, TALENs have 15–30 repeats of a 35 amino acid transcription activator like effector (TALE). Each TALE is composed of mostly invariable regions with only two amino acid differences known as repeat variable di-residues (RVDs). A TALE recognizes one base pair determined by which RVD a TALE contains. Adding a number of these TALEs together, fusing them to a FokI domain, and

administering them in pairs, allows similar DSB formation to ZFNs but with an easier and more modular assembly.

1.5.1.4 CRISPR/CAS9

The most recently described targeted endonuclease, Cas9, is a monomeric protein guided by a specific type of RNA, known as a CRISPR guide.[104] The guide RNA (gRNA) contains an 18–21-nucleotide long target sequence attached to a 3′ RNA scaffold loop for Cas9 protein binding. The target region must be complementary to a region in the DNA immediately upstream of a 2–5 base pair proto-spacer adjacent motif (PAM) which depends on the species of bacteria from which the Cas protein is derived. The most obvious advantage of CRISPR/Cas9 over the other nucleases is the ease and flexibility of developing guides to target new sites.

1.5.2 GENE EDITING STRATEGIES

Depending on the disease being targeted, the type of targeted editing required may fall into one of the three categories: gene disruption, gene correction, or gene insertion (Figure 3).

1.5.2.1 GENE DISRUPTION

In certain cases, knocking out a regulatory element, viral receptor or a pathogenic gene may be sufficient to ameliorate the disease-causing phenotype. High levels of gene disruption may be achieved in hematopoietic stem and progenitor cells because this type of editing does not require a donor template and can be done via the NHEJ pathway. For example, disruption of the BCL11A erythroid enhancer (a repressor of fetal globin expression) can increase levels of fetal hemoglobin for the treatment of sickle cell disease and beta-thalassemia.[105–108] Alternatively, knockout of the CCR5 gene in cells from HIV-infected individuals can prevent ongoing infection by the virus.[109–116] Trials targeting BCL11A are approaching the clinic and several early phase clinical trials have

been completed using ZFNs to modify the CCR5 gene in HIV-infected patient peripheral blood T-cells or HSCs.[108,117,118]

1.5.2.2 GENE CORRECTION

Diseases that result from a single nucleotide substitution or other small genetic lesions may be corrected by providing a homologous donor with the corrective sequence to serve as a template for DNA repair via the HDR pathway. Applications of this approach include correction of the sickle cell mutation in the beta-globin gene and restoration of beta-globin expression in beta-thalassemia. It is challenging to engage HDR-mediated repair in the quiescent, long term HSC population. In pre-clinical studies, gene correction levels of ~10%, ~25%, and ~35% using TALENs, ZFNs and CRISPR/Cas9, respectively, at the HBB locus were reported *in vitro*.[119–121] However once the gene-edited cells were transplanted into immunocompromised NOD-scid-IL2Rgnull (NSG) mice, the gene correction levels decreased to below 10%. These data suggest that correction was less efficient in the true stem cell population than in more differentiated progenitor cells. Further studies must be performed to improve HDR in HSCs. Recently, a new mechanism of gene correction using a single-stranded donor template was reported, which is thought to act in a RAD-51 independent, and fanconi anemia (FA)-depended manner.[122] Modulation of FA pathway may be beneficial for improving the efficiency of gene correction.

1.5.2.3 GENE INSERTION

In many human genetic diseases, there are a variety of different pathogenic mutations spread across the relevant gene in different patients. In general, the lengths of the gene repair tracts mediated by HDR are relatively short (<40 bp), so that it may be necessary to develop a panel of nuclease/donor template combinations capable of performing efficient editing to cover an entire gene region.[123] If this was accomplished using multiple CRISPR guides targeting different sites

along a target gene or the genome, each guide might require validation for levels of activity and specificity, which could be impractical for Good Manufacturing Practice (GMP) production of the cells.

Instead, activity of the whole gene can be restored by the targeted insertion of a corrective cDNA of the relevant gene into the start of the endogenous gene locus. Examples of genes being targeted in HSCs by this method include: IL2Rg for X-linked Severe Combined Immune Deficiency (X-SCID), BTK for X-Linked Agammaglobulinemia (XLA), CD40L for X-Linked Hyper IgM Syndrome (XHIM), and CYBB for X-Linked Chronic Granulomatous Disease (XCGD).[120,124–127] Gene cassettes may be inserted into specific gene loci, such as CCR5 for anti-HIV strategies to knock-out the viral co-receptor gene and simultaneously insert another anti-HIV gene; or into "safe harbors" such as the AAVS1 site which safely supports sustained transgene expression.[104,125,127–130]

The donor template or cassette is flanked by homology arms surrounding the nuclease cut site and generally consists of the gene's full length cDNA complete with a stop codon and a 3′ untranslated region containing the polyadenylation signal. This cDNA donor template is most often delivered via an adeno-associated virus (AAV) or other non-integrating vector.[131] A major challenge of this approach is to achieve efficient delivery and integration of these larger donor template. In pre-clinical studies, gene insertion rates of up to 43% at the IL2RG, AAVS1 or CYBB loci were achieved *in vitro*; however, as with gene correction, the levels of gene insertion in HSCs decreased *in vivo*.[9,120,131,132]

Recently, two new methods of gene integration were described: homology-independent targeted integration (HITI) and precise integration into target chromosome (PITCh), which use

NHEJ and MMEJ machinery, respectively.[133–135] The advantage of these strategies over HDR-mediated gene integration include being able to target cells outside the S/G2 phases of cell cycle.

1.5.3 QUIESCENCE

HSCs are quiescent and mostly reside in the G0/G1 phase of cell cycle, providing a unique challenge for gene editing, since cell cycle phase is a major factor determining which DNA repair pathway is utilized to repair DSB. NHEJ occurs throughout the cell cycle while HDR is mostly restricted to S/G2 phases, when a sister chromatid is available to serve as a homologous template for repair.[21,136] This pattern of DNA repair has caused the rates of targeted gene correction and insertion in primary human HSCs to remain relatively low and rates of gene disruption by indels too high. Methods to increase HDR and decrease NHEJ are being developed. Current gene editing protocols utilize culture with a combination of recombinant hematopoietic growth factors (e.g. ckit ligand, FLT3 ligand, thrombopoietin and others) to induce cell cycling 24–72 hours prior to the delivery of nuclease and donor template. However, even with pre-stimulation the majority of HSPCs are not in S/G2 phases. Cell synchronization agents have been used successfully in 293T cells and the H9 embryonic cell line to temporarily arrest the cells in S/G2 phases of cell cycle during DNA repair to increase HDR.[137] However, the downstream effects of cell synchronization on the self-renewal and differentiation potential of HSCs is not yet known.

An alternative method being explored to improve the precision of gene insertion and gene correction is by simply reducing NHEJ. One such method reported to reduce NHEJ is the inhibition of Ligase IV, which is involved in the final step in the NHEJ pathway.[138–140] However, other groups have not been able to achieve a significant decrease in NHEJ using this inhibitor.[141–143] A potential concern with the inhibition of Ligase IV is that decreasing NHEJ levels in the cells may not result in an increase in HDR if the cells have already committed to the end-joining pathway. The effects

of this type of late repair pathway blockade are still unknown, but it may lead to lower correction efficiency or even induce apoptosis.

Rather than preventing NHEJ via its terminal step, another possibility is to control the DNA repair pathway choice more upstream at the decision-making stage. For instance, formation of the BRCA1-PALB2-BRCA2 protein complex is crucial for HDR to occur; however, it is inhibited during G1.[144] Modulating the interaction between BRCA1-PALB2-BRCA2 in U2OS cells allowed Orthwein and colleagues to initiate HDR in the G1 phase of cell cycle.[144] Whether this approach can be translated to primary human HSPCs remains to be tested. Another possibility is to decrease nuclease cutting in the G1 phase of the cell cycle by adding to Cas9 a fragment of the Geminin protein that causes Cas9 degradation during G1 when only the NHEJ repair pathway is available.[141]

1.5.4 Methods of Delivery

One recurring challenge of targeted editing is how best to deliver the endonuclease and homologous donor template (if necessary for the particular treatment) to HSCs. Primary human HSCs are notoriously resistant to transfection methods of gene delivery.[145] Electroporation methods to deliver nucleic acids have improved over time and effectively transfer nucleic acid to the majority of HSCs in a treated sample. However, there is often a mild-moderate degree of toxicity from electroporation and this is significantly worsened by delivery of plasmid DNA, in some cases resulting in up to 60% cell death 24 hours post electroporation.[110] The delivery of *in vitro* transcribed mRNA encoding the nuclease and either *in vitro transcribed* or chemically-synthesized short guide RNA is better tolerated. Co-delivery of recombinant Cas9 protein complexed to short guide RNA as ribonucleoprotein (RNP) complexes has also been shown to be effective.[93]

Delivery of homologous donor sequences has been achieved with multiple modalities. Chemically-synthesized oligonucleotides (e.g. 50–200 bp in length) are effective donors for small sequence changes, although they may cause moderate toxicity when introduced by electroporation.[110] Integrase-defective lentiviral vectors (IDLV), and the more effective adeno-associated virus (AAV) vectors can efficiently deliver donor sequences of variable lengths (up to several kb) to HSCs with lower cytotoxicity than oligonucleotides or plasmids.[121,131,146]

1.6 GENE EDITING OF HSCs FOR CLINICAL APPLICATIONS

For clinical applications, multiple reagents are thus needed to perform gene editing (nuclease and donor) and each will need to be produced under standardized GMP conditions. Research scale editing is typically done with 0.2–1x106 CD34+ cells per experimental arm; clinical scale will involve at least 5–10x that many CD34+ cells per kg, and thus 50–1,000 times more cells. Although standards for acceptable levels of off-target cutting by a nuclease for clinical editing have not been defined, it is incumbent to investigate their occurrence with the most sensitive and relevant assays that can be practically done as part of pre-clinical toxicology assessments.

1.7 CHALLENGES TO CLINICAL APPLICATION OF HSC GENE THERAPY

1.7.1 HSC HARVEST AND EXPANSION

While there has been much progress in applications of HSC gene therapy, many challenges remain. The numbers of HSCs that can be obtained from a patient are limited by the yields that can be isolated by bone marrow harvest or mobilization, although the combination of G-CSF and a CXCR4 inhibitor (plerixafor) generally leads to abundant cell collections.[15] Some specific diseases may limit the numbers of HSCs that can be isolated, such as Fanconi anemia, which results

in progressive HSC failure, or osteopetrosis, where the marrow space is progressively reduced by the accumulated bone.[147,148]

The cell processing manipulations, including stem cell enrichment or gene modification--particularly when using electroporation--may lead to significant cell losses. Efforts to expand the numbers of true transplantable HSCs have been made, with several small molecules (such as SR-1, UM171, PGE2) holding some promise, although no massive HSCs expansion has been achieved.[149–153] The goal of producing transplantable HSCs from pluripotent stem cells is advancing, with direct reprogramming to HSCs from endothelial cells also showing promise.[154,155]

1.7.2 GENE TRANSFER

Gene transfer to HSCs has also advanced to a large degree, with current protocols of hematopoietic growth factor stimulation and transduction with lentiviral vectors reaching therapeutic efficacy for many disorders. Nonetheless, here too, improvements are needed. Human HSCs are relatively resistant to lentiviral vectors, evidenced by the seemingly high multiplicities of infection (M.O.I., the vector/cell ratio) needed to effectively transduce HSCs, compared to the relatively easier transduction of the cell lines typically used to gauge vector titers. The carrying capacity of lentiviral vectors has limitations, with vector titers falling off sharply as the size of the gene cassette increases. In our hands, a lentiviral vector at the small end of the size range (e.g. with a simple cDNA or transgene like GFP and a small promoter at ~4 kb proviral length) has a titer 10–30-fold higher than a vector at the large end of the size range (e.g. with a beta-globin gene cassette with exons, introns, upstream locus control region segments at ~9kb). The lower production titer necessitates a proportionately higher volume of vector preparation to produce a patient dose, increasing the costs. Additionally, the bigger vectors do not transduce HSCs as well as smaller vectors, even when adjusted to matching MOI. Improved transduction of HSCs with

lentiviral vectors using small molecules (proteasome inhibitors, cyclosporine A, rapamycin) has been reported in pre-clinical studies, but the effects have not been clinically validated.And, of course, the semi-random integration of the vectors throughout the genome continue to pose genotoxicity risks, although these are greatly diminished with current generation vectors that lack the strong long terminal repeat enhancers that were the major cause of insertional oncogenesis with the first generation of gRV's.[156–158]

1.7.3 GENE EDITING

Gene editing may avoid many of the problems specific to viral vectors, but it too remains less than ideal in several ways. The various site-specific endonucleases (HE, ZFN, TALEN, or CRISPR) are fairly efficient and targeted gene disruption in HSCs is now in clinical trials for HIV (targeting CCR5, HIV co-receptor) and sickle cell disease (targeting BCL11a, repressor of fetal globin); clinical results have not yet been reported.[108,118] However, the more elegant goals of targeted gene correction and gene insertion are more complicated to achieve, relying on the HDR pathway to perform the desired edits and thus requiring co-delivery of a homologous donor with the nuclease. Before clinical translation of gene editing, GMP methods of gene editing combining multiple GMP-grade reagents (e.g. CRISPR RNP or mRNA and short-guide RNA; AAV vector homologous donor) will need to be established. Pre-clinical work editing human HSCs using a clinically-scale approach has been reported.[125]

1.7.4 *EX VIVO* PROCESSING

The *ex vivo* processing to enrich HSCs for clinical gene therapy has mostly been limited to CD34+ cell selection. This achieves a moderate (30–50-fold) decrease in total numbers of cells that need to be exposed to vector or gene edited (akin to a lineage-negative {lin−} murine population), but yet retains most of the HSCs. However, the CD34+ cell population is still quite

heterogeneous with only a small fraction of cells being the target long-lived HSCs. Thus, a large proportion of the vector or gene editing reagents are wasted modifying the more abundant but short-term progenitor cells. Efforts to further enrich for HSCs using additional markers such as CD38(−), CD90(+), CD133(+), etc, (akin to a murine "LSK" {lin−/Sca1+/ckit+ fraction}) have been reported, but require FACS sorting which may entail long processing times and subject the cells to damaging shear forces.[159] Newer sorting methodologies based on microfluidics or using serial immuno-affinity bead processes may be beneficial if they provide further enrichment without undue losses of cells.[160,161]

The cell culture methods used in current clinical trials are relatively standardized, using static culture in gas permeable bags or flasks in serum-free medium supplemented with multiple hematopoietic growth factors. The use of continuous feed bioreactors and/or lower partial pressures of oxygen may provide more optimal conditions for HSC modification and preservation. Small molecules such as PGE2 may also support HSC survival *ex vivo* improving the level of engraftment of gene-modified HSCs.[152] Additionally, the current *ex vivo* processing of HSCs is often done in multiple open systems, but new closed systems that continuously contain the cells are being developed that may allow processing to be done in environments less demanding than the current GMP "clean rooms".

1.7.5 UNIVERSAL DONOR PRODUCTS

Universal donor cells could largely supplant use of autologous cell products if able to achieve the ideal properties of immunogenicity absence. They can be banked as an off-the-shelf, immediately-ready source of compatible normal cells, including regenerative stem cells. Universal donor cells would have a major advantage in that they could be produced in multi-patient dose lots vs. patient-specific single lots using autologous cell products. It is possible to engineer the cells to

have favorable properties, e.g. produce a therapeutic protein such as clotting factors and other serum proteins, lysosomal enzymes, anti-tumor T cell receptor or Chimeric Antigen Receptor or an immunomodulative cytokine or chemokine. Allogeneic HSCT sources may continue to have advantages for HSCT for hematologic malignancies due to their potential graft-versus-leukemia effects, although it should become possible to augment specific immune effector cell products for the positive anti-leukemia effect, but without risks for GVHD.

1.7.6 PRE-TRANSPLANT CYTOREDUCTIVE CONDITIONING

Finally, the pre-transplant cytoreductive conditioning used to "make space" for engraftment of the isolated and reinfused HSCs is finally advancing beyond the use of cytotoxic chemotherapy drugs or radiation. These agents are effective at ablating the marrow stem cells (myeloablation), which is necessary for engraftment of gene-modified HSCs and for suppressing the recipient's immune system (immunoablation), which is necessary to avoid immunologic rejection of the graft. However, they may have severe acute toxicities in multiple organ systems (heart, lungs, liver, kidney, GI) and may produce infertility or sterility, due to toxicity to germ cells. Monoclonal antibodies to HSC surface proteins (e.g. ckit, CD47, CD45) have been shown in murine models to allow improved engraftment without apparent toxicity.[162–165] These efforts are now being translated to the clinic and may eliminate the need to use toxic preparative regimens to facilitate engraftment.

1.8 LESSONS LEARNED

1.8.1 IT TAKES A LONG TIME

Lessons learned from the almost 30-year history of developing clinical HSC gene therapy products can inform emerging stem cell-based cellular therapies for myriad other non-

hematopoietic diseases such as Duchene's Muscular Dystrophy, Huntington's Disease, Parkinson's Disease, Diabetes Mellitus and others. One clear lesson has been that development of novel therapies takes a long time. Methods for effective gene transfer to HSCs were initially developed in the 1980s and clinical trials started in the 1990's. The initial trials yielded no evidence of efficacy; the first clinical successes were not seen until the 2000's and only in the last decade are therapeutic benefits being conferred consistently for multiple disorders. Several promising HSC gene therapy cell products are advancing through early phase clinical trials (for indications including X-adrenoleukodystrophy, Metachromatic Leukodystrophy, Beta-thalassemia, Sickle Cell Disease, ADA-deficient and X-linked forms of SCID, Wiskott-Aldrich Syndrome, Chronic Granulomatous Disease) and towards licensure for commercial manufacture and sales, with the first (Strimvelis for ADA-deficient SCID) approved by the European Medicines Agency (European Medicines Agency, 2016). While it may be expected that therapies using other stem cell types will be derived in a shorter time-frame, drug development remains a slow process.

1.8.2 SUPPORT ACADEMIC MEDICAL CENTERS TO DEVELOP NOVEL CELL THERAPIES

HSC gene therapies were incubated at academic medical centers in multiple countries often at innovative HSCT programs, not via the traditional pharmaceutical company model of drug development. Universities and other research centers need to have sufficient infrastructure for early phase clinical trial performance and GMP cell processing to achieve similar academic pioneering for other stem cell therapies. Indeed, the centers that have had strong gene therapy programs with the necessary cell processing and regulatory infrastructure have been the leaders in this field (e.g. TIGET, Milan Italy; Hôpital NeckerEnfants Malade, Paris, France; University College London, London, UK; The National Institutes of Health, Bethesda MD; University of

California, Los Angeles, Los Angeles CA; Boston Children's Hospital, Boston MA; St. Jude Children's Research Hospital, Memphis TN). Ongoing support of this type of research will be essential to continue the innovation of new therapies.

1.8.3 DO IT RIGHT

In the early days for the field of gene therapy, the NIH RAC-provided public oversight to review clinical protocols to be performed in academic medical centers was an important forum to discuss the novel potential biohazard issues, as well as scientific and ethical concerns. This may have tempered or even slowed some advancements, but it helped provide another level of expertise, in addition to that provided by local IRB and FDA (and EMA in Europe), to ensure that trials were based on sound scientific principles, had adequate supporting pre-clinical data on potential efficacy and safety, and were well-designed and monitored. The role of the RAC in overseeing individual trials has decreased, but it still serves its federal advisory role in assessing novel biosafety issues. Other forms of stem cell therapy should also proceed with a base of strong pre-clinical data, careful consideration of the clinical setting and approach, as well as well-controlled cell manufacturing and regulatory oversight, to provide maximum safety for subjects and quality of data derived.

1.8.4 PROTECT SUBJECT SAFETY BY STRICT COMPLIANCE

Any clinical trial with a novel major intervention, such as cell therapy, and especially with subjects with organ dysfunction caused by their disease, can have unexpected and potentially fatal events, either related to the cell product or not. The death of a volunteer subject in a gene therapy trial in 1999 shocked the field and the effects extended throughout much of academic clinical research.[166,167] The response was to strengthen the quality of clinical trial performance to provide maximum protection to subjects and preserve the integrity of the data they contribute.[167] The field of HSCT began clinical investigations in the 1960's-70's, prior to the establishment of IRB and

the other oversight bodies, as best available clinical practice for severe, generally fatal disorders. Since that time, the standards for clinical investigations have been greatly expanded for a much more complex regulatory environment. A typical clinical trial of gene therapy may undergo review by a dozen or more entities, including IRB, IBC, ISPRC, DSMB, NIH RAC, FDA, and one or more funding agencies. HSCT came to full maturity in the U.S. by the development, initiated by members of the ASBMT academic society, of the Foundation for the Accreditation of Cellular Therapy (FACT), which brought uniformity and high standards for cell processing, clinical operations, data management, regulatory management and other clinical trial activities. Commercial cell processing methods and standards have also been developed supporting several cell products that advanced to relatively late stage of investigation (e.g. neural progenitor cells derived from fetal tissue or pluripotent stem cells) and producing a licensed dendritic cell vaccine (the marketed prostate cancer dendritic cell Sipuleucel-T from Dendreon Corp.). This industry has a strong base to produce high quality cell products, but each new cell product type developed brings unique challenges, including details of the cell processing protocol, the release testing, storage, transport, and therapy administration.

1.8.5 TRANSLATIONAL RESEARCH IS DRUG DEVELOPMENT

It is important to keep in mind that the goal of translational research is to develop a drug that is effective and safe to achieve licensure, be it a cell and/or gene therapy product such as lentiviral-transduced HSCs, iPSC derived myoblasts, dopaminergic neurons, shRNA, etc. This necessitates basic scientists learning fundamental principles of drug development, such as Good Laboratory Practices, Good Manufacturing Practice and Good Clinical practice. An important tool commonly used in drug development is the Target Product Profile (TPP), which sets goals for attributes like: clinical indication, patient population, administration route and schedule, clinical

efficacy targets, potential risks, drug quality and testing methods. Early drafting of a TPP can guide subsequent studies to keep focused on the drug development goals.

1.8.6 TRIALS SHOULD BE DESIGNED TO BE INFORMATIVE ABOUT THE CELL

While the primary end-points for early phase trials mostly relate to safety, it should be possible to incorporate secondary end-points for efficacy and exploratory end-points for biomarkers that can be assessed for potential suitability as eventual primary end-points for drug approval. In the clinical trials we have done testing new vectors, there has been a primary end-point assessing safety, which is typical for a Phase I study, but also secondary end-points assessing efficacy. For ADA SCID, this has involved safety assessments by documentation of clinical adverse events, as well as ensuring absence of replication-competent viral vector emergence and absence of vector-driven clonal expansion. Efficacy assessments involved measuring expression of ADA enzyme activity in mature blood cells, quantifying engraftment of gene-modified stem cells by measuring vector copy number in cells by quantitative PCR, and performing standard clinical tests of immune function, as well as recording clinical health.

In some instances, especially with orphan diseases, non-traditional pathways of clinical trials may be accepted by regulators, with even small trials used as pivotal for registration, assuming they were done with appropriate design and rigor. The EMA approval of Strimvelis for ADA SCID was based on a single center's Phase I/II clinical data involving 12 patients which served as a pivotal clinical trial.

1.8.7 GET THE MOST FROM PRE-CLINICAL STUDIES

The other major serious complication in the gene therapy field was the development of leukemia in subjects in several primary immune deficiency trials from insertional oncogenesis by

retroviral vectors introduced into HSCs.[168–170] The relatively high frequency of the development of leukemia in some trials (25–75%) was not predicted by pre-clinical studies. However, pre-clinical models may not detect clinical risks that can occur in patients with much larger absolute cell dosages and longer post-treatment time periods. It may be difficult to test cell therapies by the parameters traditionally applied to drug therapies (pharmacokinetics, biodistribution, toxicity), but these aspects can be often assessed using PCR methods to quantify transgenes or cellular markers, such as human genomes in human cells against the background of murine host genomes.

Nonetheless, pre-clinical studies that are performed should be optimized to provide as much relevant information as possible. Principles of Good Laboratory Practices should be applied whenever possible, even at early stages of discovery and prior to formal IND-enabling studies. These include such key elements of GLP as following a detailed pre-defined plan for the studies, statistical plan, data capture forms, with formal data reporting. Again, even during early phases of product development, it important to include toxicology analysis within efficacy studies to obtain initial information that can be used in the design of definitive studies for IND application. Preliminary proof-of-principles studies can also be used to begin to investigate cell dosages, potential toxicities, as well as disease-modifying activity.

1.8.8 ADVANCING CLINICAL CELL THERAPIES IS CHALLENGING

Clinical cell therapy requires point-to-point control of the manufacturing process and starting materials (e.g. from skin biopsy to delivery of iPSC-derived somatic cell product, whether it is HSCs or other cell product). The GMP process requires highly trained staff and SOPs, materials specification, batch records, personnel training, in addition to the highly-controlled environment and regulated processes. To characterize the cell product for human administration, it is necessary to define release criteria – identity, purity, potency, and safety. A Certificate of

Analysis is completed for each batch of cell product and the testing for each critical attribute is required with full documentation. The analytic testing for aspects of cell quality (e.g. cell counts and viability, immunohistochemistry or flow cytometry, PCR, RNA-SEQ, etc.) should be performed using well-characterized assays, which should be made more robust with advancing stages of investigation.

1.9 CONCLUSION

In conclusion, gene therapy using HSCs has progressed over three decades from ineffectiveness to being able to essentially cure several different disorders. The pathway was not linear, but required multiple iterative bench-to-bedside cycles. It is likely that therapies using other stem cells will also have progress and set-backs. But, because the underlying hypotheses for cellular therapies are so convincing, it is highly likely that multiple novel stem cell-based therapies will be developed. The lessons from the field of HSC gene therapy may provide some guidance for investigators pursuing the translational process.

1.10 ACKNOWLEDGEMENTS

DBK is a member of the UCLA Broad Stem Cell Center, the Jonsson Comprehensive Cancer Center and the Children's Discovery and Innovation Institute. DBK is a consultant to Orchard Therapeutics Ltd. and a member of its Scientific Advisory Board. Orchard has licensed a lentiviral vector for gene therapy of ADA SCID from the University of California Regents on which DBK is an inventor. Training grants provided support to R.M. (NIH/NHLBI F31 HL134313; PhRMA Foundation Paul Calabresi Predoctoral Fellowship; UCLA MSTP T32 GM00804), DG (UCLA Virology and Gene Therapy Training Grant Ruth L. Kirschstein National Research Service Award T32 AI060567) and AL (NIH/NHLBI F31 HL134208).

1.11 REFERENCES

1. Boelens, JJ, Aldenhoven, M, Purtill, D, Ruggeri, A, DeFor, T, Wynn, R, *et al.* (2013). Outcomes of transplantation using various hematopoietic cell sources in children with Hurler syndrome after myeloablative conditioning. *Blood*doi:10.1182/blood-2012-09-455238.

2. Walters, MC (2015). Update of hematopoietic cell transplantation for sickle cell disease. *Curr. Opin. Hematol.*doi:10.1097/MOH.0000000000000136.

3. Cooke, KR, Luznik, L, Sarantopoulos, S, Hakim, FT, Jagasia, M, Fowler, DH, *et al.* (2017). The Biology of Chronic Graft-versus-Host Disease: A Task Force Report from the National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host Disease. *Biol. Blood Marrow Transplant.*doi:10.1016/j.bbmt.2016.09.023.

4. Fuchs, EJ (2015). HLA-haploidentical blood or marrow transplantation with high-dose, post-transplantation cyclophosphamide. *Bone Marrow Transplant.*doi:10.1038/bmt.2015.92.

5. Muccio, L, Bertaina, A, Falco, M, Pende, D, Meazza, R, Lopez-Botet, M, *et al.* (2016). Analysis of memory-like natural killer cells in human cytomegalovirus-infected children undergoing αβ+T and B cell-depleted hematopoietic stem cell transplantation for hematological malignancies. *Haematologica*doi:10.3324/haematol.2015.134155.

6. Aiuti, A, Biasco, L, Scaramuzza, S, Ferrua, F, Cicalese, MP, Baricordi, C, *et al.* (2013). Lentiviral hematopoietic stem cell gene therapy in patients with wiskott-aldrich syndrome. *Science (80-.).*doi:10.1126/science.1233151.

7. Biffi, A, Montini, E, Lorioli, L, Cesani, M, Fumagalli, F, Plati, T, *et al.* (2013). Lentiviral hematopoietic stem cell gene therapy benefits metachromatic leukodystrophy. *Science (80-.).*doi:10.1126/science.1233158.

8. Cartier, N and Aubourg, P (2010). Hematopoietic stem cell transplantation and hematopoietic stem cell gene therapy in X-linked adrenoleukodystrophy. *Brain Pathol.*doi:10.1111/j.1750-3639.2010.00394.x.

9. De Ravin, SS, Reik, A, Liu, PQ, Li, L, Wu, X, Su, L, *et al.* (2016). Targeted gene addition in human CD34 + hematopoietic cells for correction of X-linked chronic granulomatous disease. *Nat. Biotechnol.* **34**: 424–429.

10. Cartier, N, Hacein-Bey-Abina, S, Bartholomae, CC, Bougnres, P, Schmidt, M, Von Kalle, C, *et al.* (2012). Lentiviral hematopoietic cell gene therapy for X-linked adrenoleukodystrophy. *Methods Enzymol.*doi:10.1016/B978-0-12-386509-0.00010-7.

11. Enssle, J, Trobridge, GD, Keyser, KA, Ironside, C, Beard, BC and Kiem, HP (2010). Stable marking and transgene expression without progression to monoclonality in canine long-term hematopoietic repopulating cells transduced with lentiviral vectors. *Hum. Gene Ther.*doi:10.1089/hum.2009.076.

12. Carroll, D (2016). Genome editing: Progress and challenges for medical applications.

*Genome Med.*doi:10.1186/s13073-016-0378-9.

13. Wright, A V., Nuñez, JK and Doudna, JA (2016). Biology and Applications of CRISPR Systems: Harnessing Nature's Toolbox for Genome Engineering. *Cell*doi:10.1016/j.cell.2015.12.035.

14. Kondo, M, Wagers, AJ, Manz, MG, Prohaska, SS, Scherer, DC, Beilhack, GF, *et al.* (2003). B IOLOGY OF H EMATOPOIETIC S TEM C ELLS AND P ROGENITORS : Implications for Clinical Application . *Annu. Rev. Immunol.*doi:10.1146/annurev.immunol.21.120601.141007.

15. Brave, M, Farrell, A, Ching Lin, S, Ocheltree, T, Pope Miksinski, S, Lee, SL, *et al.* (2010). FDA review summary: Mozobil in combination with granulocyte colony-stimulating factor to mobilize hematopoietic stem cells to the peripheral blood for collection and subsequent autologous transplantation. *Oncology*doi:10.1159/000315736.

16. Notta, F, Doulatov, S, Laurenti, E, Poeppl, A, Jurisica, I and Dick, JE (2011). Isolation of single human hematopoietic stem cells capable of long-term multilineage engraftment. *Science (80-.).*doi:10.1126/science.1201219.

17. Chaudhury, S, Ayas, M, Rosen, C, Ma, M, Viqaruddin, M, Parikh, S, *et al.* (2017). A Multicenter Retrospective Analysis Stressing the Importance of Long-Term Follow-Up after Hematopoietic Cell Transplantation for β-Thalassemia. *Biol. Blood Marrow Transplant.*doi:10.1016/j.bbmt.2017.06.004.

18. Hsieh, MM, Wu, CJ and Tisdale, JF (2011). In mixed hematopoietic chimerism, the donor red cells win. *Haematologica*doi:10.3324/haematol.2010.035576.

19. Cheshier, SH, Morrison, SJ, Liao, X and Weissman, IL (1999). In vivo proliferation and cell cycle kinetics of long-term self-renewing hematopoietic stem cells. *Proc. Natl. Acad. Sci. U. S. A.*doi:10.1073/pnas.96.6.3120.

20. Passegué, E, Wagers, AJ, Giuriato, S, Anderson, WC and Weissman, IL (2005). Global analysis of proliferation and cell cycle gene expression in the regulation of hematopoietic stem and progenitor cell fates. *J. Exp. Med.*doi:10.1084/jem.20050967.

21. Pietras, EM, Warr, MR and Passegué, E (2011). Cell cycle regulation in hematopoietic stem cells. *J. Cell Biol.* **195**: 709–720.

22. Lewis, PF and Emerman, M (1994). Passage through mitosis is required for oncoretroviruses but not for the human immunodeficiency virus. *J. Virol.*doi:10.1128/jvi.68.1.510-516.1994.

23. Miller, DG, Adam, MA and Miller, AD (1990). Gene transfer by retrovirus vectors occurs only in cells that are actively replicating at the time of infection. *Mol. Cell. Biol.*doi:10.1128/mcb.10.8.4239.

24. Case, SS, Price, MA, Jordan, CT, Yu, XJ, Wang, L, Bauer, G, *et al.* (1999). Stable transduction of quiescent CD34+CD38- human hematopoietic cells by HIV-1-based lentiviral vectors. *Proc. Natl. Acad. Sci. U. S. A.*doi:10.1073/pnas.96.6.2988.

25. Miyoshi, H, Smith, KA, Mosier, DE, Verma, IM and Torbett, BE (1999). Transduction of

human CD34+ cells that mediate long-term engraftment of NOD/SCID mice by HIV vectors. *Science (80-.)*.doi:10.1126/science.283.5402.682.

26. Naldini, L, Blömer, U, Gallay, P, Ory, D, Mulligan, R, Gage, FH, *et al.* (1996). In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science (80-.)*.doi:10.1126/science.272.5259.263.

27. Bordignon, C, Notarangelo, LD, Nobili, N, Ferrari, G, Casorati, G, Panina, P, *et al.* (1995). Gene therapy in peripheral blood lymphocytes and bone marrow for ADA-immunodeficient patients. *Science (80-.)*.doi:10.1126/science.270.5235.470.

28. Aiuti, A and Roncarolo, MG (2009). Ten years of gene therapy for primary immune deficiencies. *Hematology Am. Soc. Hematol. Educ. Program*doi:10.1182/asheducation-2009.1.682.

29. (2014). 276. Immunological and Metabolic Correction After Lentiviral Vector Mediated Haematopoietic Stem Cell Gene Therapy for ADA Deficiency. *Mol. Ther.*doi:10.1016/s1525-0016(16)35289-3.

30. Candotti, F, Shaw, KL, Muul, L, Carbonaro, D, Sokolic, R, Choi, C, *et al.* (2012). Gene therapy for adenosine deaminase-deficient severe combined immune deficiency: Clinical comparison of retroviral vectors and treatment plans. *Blood*doi:10.1182/blood-2012-02-400937.

31. Gaspar, HB, Cooray, S, Gilmour, KC, Parsley, KL, Zhang, F, Adams, S, *et al.* (2011). Immunodeficiency: Hematopoietic stem cell gene therapy for adenosine deaminase-deficient severe combined immunodeficiency leads to long-term immunological recovery and metabolic correction. *Sci. Transl. Med.*doi:10.1126/scitranslmed.3002716.

32. Shaw, KL, Garabedian, E, Mishra, S, Barman, P, Davila, A, Carbonaro, D, *et al.* (2017). Clinical efficacy of gene-modified stem cells in adenosine deaminase-deficient immunodeficiency. *J. Clin. Invest.*doi:10.1172/JCI90367.

33. Hacein-Bey-Abina, S, Le Deist, F, Carlier, F, Bouneaud, C, Hue, C, De Villartay, JP, *et al.* (2002). Sustained correction of X-linked severe combined immunodeficiency by ex vivo gene therapy. *N. Engl. J. Med.*doi:10.1056/NEJMoa012616.

34. Hacein-Bey-Abina, S, Garrigue, A, Wang, GP, Soulier, J, Lim, A, Morillon, E, *et al.* (2008). Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. *J. Clin. Invest.*doi:10.1172/JCI35700.

35. Howe, SJ, Mansour, MR, Schwarzwaelder, K, Bartholomae, C, Hubank, M, Kempski, H, *et al.* (2008). Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. *J. Clin. Invest.*doi:10.1172/JCI35798.

36. Ott, MG, Schmidt, M, Schwarzwaelder, K, Stein, S, Siler, U, Koehl, U, *et al.* (2006). Correction of X-linked chronic granulomatous disease by gene therapy, augmented by insertional activation of MDS1-EVI1, PRDM16 or SETBP1. *Nat. Med.*doi:10.1038/nm1393.

37. Boztug, K, Schmidt, M, Schwarzer, A, Banerjee, PP, Díez, IA, Dewey, RA, *et al.* (2010).

Stem-cell gene therapy for the Wiskott-Aldrich syndrome. *N. Engl. J. Med.*doi:10.1056/NEJMoa1003548.

38. DeRijck, J, deKogel, C, Demeulemeester, J, Vets, S, ElAshkar, S, Malani, N, *et al.* (2013). The BET Family of Proteins Targets Moloney Murine Leukemia Virus Integration near Transcription Start Sites. *Cell Rep.*doi:10.1016/j.celrep.2013.09.040.

39. Derse, D, Crise, B, Li, Y, Princler, G, Lum, N, Stewart, C, *et al.* (2007). Human T-Cell Leukemia Virus Type 1 Integration Target Sites in the Human Genome: Comparison with Those of Other Retroviruses. *J. Virol.*doi:10.1128/jvi.02752-06.

40. Mitchell, RS, Beitzel, BF, Schroder, ARW, Shinn, P, Chen, H, Berry, CC, *et al.* (2004). Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. *PLoS Biol.*doi:10.1371/journal.pbio.0020234.

41. Wu, X, Li, Y, Crise, B and Burgess, SM (2003). Transcription start regions in the human genome are favored targets for MLV integration. *Science (80-.).*doi:10.1126/science.1083413.

42. Liu, M, Li, CL, Stamatoyannopoulos, G, Dorschner, MO, Humbert, R, Stamatoyannopoulos, JA, *et al.* (2012). Gammaretroviral vector integration occurs overwhelmingly within and near DNase hypersensitive sites. *Hum. Gene Ther.*doi:10.1089/hum.2010.177.

43. Modlich, U, Navarro, S, Zychlinski, D, Maetzig, T, Knoess, S, Brugman, MH, *et al.* (2009). Insertional transformation of hematopoietic cells by self-inactivating lentiviral and gammaretroviral vectors. *Mol. Ther.*doi:10.1038/mt.2009.179.

44. Yu, SF, von Ruden, T, Kantoff, PW, Garber, C, Seiberg, M, Rüther, U, *et al.* (1986). Self-inactivating retroviral vectors designed for transfer of whole genes into mammalian cells. *Proc. Natl. Acad. Sci. U. S. A.*doi:10.1073/pnas.83.10.3194.

45. Nienhuis, AW, Dunbar, CE and Sorrentino, BP (2006). Genotoxicity of Retroviral Integration In Hematopoietic Cells. *Mol. Ther.*doi:10.1016/j.ymthe.2006.03.001.

46. Hacein-Bey-Abina, S, Pai, SY, Gaspar, HB, Armant, M, Berry, CC, Blanche, S, *et al.* (2014). A modified γ-retrovirus vector for X-linked severe combined immunodeficiency. *N. Engl. J. Med.*doi:10.1056/NEJMoa1404588.

47. Zufferey, R, Nagy, D, Mandel, RJ, Naldini, L and Trono, D (1997). Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo. *Nat. Biotechnol.*doi:10.1038/nbt0997-871.

48. Dull, T, Zufferey, R, Kelly, M, Mandel, RJ, Nguyen, M, Trono, D, *et al.* (1998). A Third-Generation Lentivirus Vector with a Conditional Packaging System. *J. Virol.*doi:10.1128/jvi.72.11.8463-8471.1998.

49. Girard-Gagnepain, A, Amirache, F, Costa, C, Levý, C, Frecha, C, Fusil, F, *et al.* (2014). Baboon envelope pseudotyped LVs outperform VSV-G-LVs for gene transfer into early-cytokine-stimulated and resting HSCs. *Blood*doi:10.1182/blood-2014-02-558163.

50. Matreyek, KA and Engelman, A (2013). Viral and cellular requirements for the nuclear

entry of retroviral preintegration nucleoprotein complexes. *Viruses*doi:10.3390/v5102483.

51. Hanenberg, H, Xiao, XL, Dilloo, D, Hashino, K, Kato, I and Williams, DA (1996). Colocalization of retrovirus and target cells on specific fibronectin fragments increases genetic transduction of mammalian cells. *Nat. Med.*doi:10.1038/nm0896-876.

52. Barrette, S, Douglas, JL, Seidel, NE and Bodine, DM (2000). Lentivirus-based vectors transduce mouse hematopoietic stem cells with similar efficiency to Moloney murine leukemia virus-based vectors. *Blood*doi:10.1182/blood.v96.10.3385.

53. Sutton, RE, Reitsma, MJ, Uchida, N and Brown, PO (1999). Transduction of Human Progenitor Hematopoietic Stem Cells by Human Immunodeficiency Virus Type 1-Based Vectors Is Cell Cycle Dependent. *J. Virol.*doi:10.1128/jvi.73.5.3649-3660.1999.

54. Millington, M, Arndt, A, Boyd, M, Applegate, T and Shen, S (2009). Towards a clinically relevant lentiviral transduction protocol for primary human CD34+ hematopoietic stem/progenitor cells. *PLoS One*doi:10.1371/journal.pone.0006461.

55. Gennari, F, Lopes, L, Verhoeyen, E, Marasco, W and Collins, MK (2009). Single-chain antibodies that target lentiviral vectors to MHC class II on antigen-presenting cells. *Hum. Gene Ther.*doi:10.1089/hum.2008.189.

56. Verhoeyen, E, Wiznerowicz, M, Olivier, D, Izac, B, Trono, D, Dubart-Kupperschmitt, A, *et al.* (2005). Novel lentiviral vectors displaying ‘early-acting cytokines’ selectively promote survival and transduction of NOD/SCID repopulating human hematopoietic stem cells. *Blood*doi:10.1182/blood-2004-12-4736.

57. Zufferey, R, Dull, T, Mandel, RJ, Bukovsky, A, Quiroz, D, Naldini, L, *et al.* (1998). Self-Inactivating Lentivirus Vector for Safe and Efficient In Vivo Gene Delivery. *J. Virol.*doi:10.1128/jvi.72.12.9873-9880.1998.

58. Canté-Barrett, K, Mendes, RD, Smits, WK, Van Helsdingen-Van Wijk, YM, Pieters, R and Meijerink, JPP (2016). Lentiviral gene transfer into human and murine hematopoietic stem cells: Size matters. *BMC Res. Notes*doi:10.1186/s13104-016-2118-z.

59. Kumar, M, Keller, B, Makalou, N and Sutton, RE (2001). Systematic determination of the packaging limit of lentiviral vectors. *Hum. Gene Ther.*doi:10.1089/104303401753153947.

60. Moreno-Carranza, B, Gentsch, M, Stein, S, Schambach, A, Santilli, G, Rudolf, E, *et al.* (2009). Transgene optimization significantly improves SIN vector titers, gp91phox expression and reconstitution of superoxide production in X-CGD cells. *Gene Ther.*doi:10.1038/gt.2008.143.

61. Hope, T (2002). Improving the post-transcriptional aspects of lentiviral vectors. *Curr. Top. Microbiol. Immunol.*doi:10.1007/978-3-642-56114-6_9.

62. Schambach, A, Wodrich, H, Hildinger, M, Bohne, J, Kräusslich, HG and Baum, C (2000). Context dependence of different modules for posttranscriptional enhancement of gene expression from retroviral vectors. *Mol. Ther.*doi:10.1006/mthe.2000.0191.

63. Woychik, RP, Lyons, RH, Post, L and Rottman, FM (1984). Requirement for the 3’ flanking region of the bovine growth hormone gene for accurate polyadenylylation. *Proc.*

*Natl. Acad. Sci. U. S. A.*doi:10.1073/pnas.81.13.3944.

64. Vink, CA, Counsell, JR, Perocheau, DP, Karda, R, Buckley, SMK, Brugman, MH, *et al.* (2017). Eliminating HIV-1 Packaging Sequences from Lentiviral Vector Proviruses Enhances Safety and Expedites Gene Transfer for Gene Therapy. *Mol. Ther.*doi:10.1016/j.ymthe.2017.04.028.

65. De Ravin, SS, Wu, X, Moir, S, Anaya-O'Brien, S, Kwatemaa, N, Littel, P, *et al.* (2016). Lentiviral hematopoietic stem cell gene therapy for X-linked severe combined immunodeficiency. *Sci. Transl. Med.*doi:10.1126/scitranslmed.aad8856.

66. Throm, RE, Ouma, AA, Zhou, S, Chandrasekaran, A, Lockey, T, Greene, M, *et al.* (2009). Efficient construction of producer cell lines for a SIN lentiviral vector for SCID-X1 gene therapy by concatemeric array transfection. *Blood*doi:10.1182/blood-2008-11-191049.

67. Aiuti, A, Cattaneo, F, Galimberti, S, Benninghoff, U, Cassani, B, Callegaro, L, *et al.* (2009). Gene therapy for immunodeficiency due to adenosine deaminase deficiency. *N. Engl. J. Med.*doi:10.1056/NEJMoa0805817.

68. Carbonaro, DA, Zhang, L, Jin, X, Montiel-Equihua, C, Geiger, S, Carmo, M, *et al.* (2014). Preclinical demonstration of lentiviral vector-mediated correction of immunological and metabolic abnormalities in models of adenosine deaminase deficiency. *Mol. Ther.*doi:10.1038/mt.2013.265.

69. Brown, MP, Topham, DJ, Sangster, MY, Zhao, J, Flynn, KJ, Surman, SL, *et al.* (1998). Thymic lymphoproliferative disease after successful correction of CD40 ligand deficiency by gene transfer in mice. *Nat. Med.*doi:10.1038/3233.

70. Chad, M, Rivella, S, Callegari, J, Heller, G, Gaensler, KML, Luzzatto, L, *et al.* (2000). Therapeutic haemoglobin synthesis in β-thalassaemic mice expressing lentivirus-encoded human β-globin. *Nature*doi:10.1038/35017565.

71. Browning, DL and Trobridge, GD (2016). Insulators to improve the safety of retroviral vectors for HIV gene therapy. *Biomedicines*doi:10.3390/biomedicines4010004.

72. Emery, DW (2011). The use of chromatin insulators to improve the expression and safety of integrating gene transfer vectors. *Hum. Gene Ther.*doi:10.1089/hum.2010.233.

73. West, AG and Fraser, P (2005). Remote control of gene transcription. *Hum. Mol. Genet.*doi:10.1093/hmg/ddi104.

74. Spitz, F and Furlong, EEM (2012). Transcription factors: From enhancer binding to developmental control. *Nat. Rev. Genet.*doi:10.1038/nrg3207.

75. Cavazzana-Calvo, M, Payen, E, Negre, O, Wang, G, Hehir, K, Fusil, F, *et al.* (2010). Transfusion independence and HMGA2 activation after gene therapy of human β-thalassaemia. *Nature*doi:10.1038/nature09328.

76. Hofmann, A, Kessler, B, Ewerling, S, Kabermann, A, Brem, G, Wolf, E, *et al.* (2006). Epigenetic regulation of lentiviral transgene vectors in a large animal model. *Mol. Ther.*doi:10.1016/j.ymthe.2005.07.685.

77. Yao, S, Sukonnik, T, Kean, T, Bharadwaj, RR, Pasceri, P and Ellis, J (2004). Retrovirus silencing, variegation, extinction, and memory are controlled by a dynamic interplay of multiple epigenetic modifications. *Mol. Ther.*doi:10.1016/j.ymthe.2004.04.007.

78. Ramezani, A, Hawley, TS and Hawley, RG (2008). Combinatorial Incorporation of Enhancer-Blocking Components of the Chicken β-Globin 5′HS4 and Human T-Cell Receptor α/δ BEAD-1 Insulators in Self-Inactivating Retroviral Vectors Reduces Their Genotoxic Potential. *Stem Cells*doi:10.1634/stemcells.2008-0258.

79. Cartier, N, Hacein-Bey-Abina, S, Bartholomae, CC, Veres, G, Schmidt, M, Kutschera, I, *et al.* (2009). Hematopoietic stem cell gene therapy with a lentiviral vector in X-linked adrenoleukodystrophy. *Science (80-.).*doi:10.1126/science.1171242.

80. Morgan, RA and Boyerinas, B (2016). Genetic modification of T cells. *Biomedicines*doi:10.3390/biomedicines4020009.

81. Biasco, L, Pellin, D, Scala, S, Dionisio, F, Basso-Ricci, L, Leonardelli, L, *et al.* (2016). In Vivo Tracking of Human Hematopoiesis Reveals Patterns of Clonal Dynamics during Early and Steady-State Reconstitution Phases. *Cell Stem Cell*doi:10.1016/j.stem.2016.04.016.

82. Biffi, A, Bartholomae, CC, Cesana, D, Cartier, N, Aubourg, P, Ranzani, M, *et al.* (2011). Lentiviral vector common integration sites in preclinical models and a clinical trial reflect a benign integration bias and not oncogenic selection. *Blood*doi:10.1182/blood-2010-09-306761.

83. Brennan, TA and Wilson, JM (2014). The special case of gene therapy pricing. *Nat. Biotechnol.*doi:10.1038/nbt.3003.

84. Orkin, SH and Reilly, P (2016). Paying for future success in gene therapy. *Science (80-.).*doi:10.1126/science.aaf4770.

85. Eichler, F, Duncan, C, Musolino, PL, Orchard, PJ, De Oliveira, S, Thrasher, AJ, *et al.* (2017). Hematopoietic stem-cell gene therapy for cerebral adrenoleukodystrophy. *N. Engl. J. Med.*doi:10.1056/NEJMoa1700554.

86. Peters, C, Charnas, LR, Tan, Y, Ziegler, RS, Shapiro, EG, DeFor, T, *et al.* (2004). Cerebral X-linked adrenoleukodystrophy: The international hematopoietic cell transplantation experience from 1982 to 1999. *Blood*doi:10.1182/blood-2003-10-3402.

87. Cesana, D, Ranzani, M, Volpin, M, Bartholomae, C, Duros, C, Artus, A, *et al.* (2014). Uncovering and Dissecting the Genotoxicity of Self-inactivating Lentiviral Vectors In Vivo. *Mol. Ther.* **22**: 774–785.

88. Ng, YY, Baert, MRM, Pike-Overzet, K, Rodijk, M, Brugman, MH, Schambach, a, *et al.* (2010). Correction of B-cell development in Btk-deficient mice using lentiviral vectors with codon-optimized human BTK. *Leuk. Off. J. Leuk. Soc. Am. Leuk. Res. Fund, U.K* **24**: 1617–1630.

89. Zhou, S, Fatima, S, Ma, Z, Wang, Y-D, Lu, T, Janke, LJ, *et al.* (2016). Evaluating the safety of retroviral vectors based on insertional oncogene activation and blocked differentiation in cultured thymocytes. *Mol. Ther.* **24**: 1090–1099.

90. Barzel, A, Paulk, NK, Shi, Y, Huang, Y, Chu, K, Zhang, F, *et al.* (2015). Promoterless gene targeting without nucleases ameliorates haemophilia B in mice. *Nature*doi:10.1038/nature13864.

91. Li, H, Haurigot, V, Doyon, Y, Li, T, Wong, SY, Bhagwat, AS, *et al.* (2011). In vivo genome editing restores haemostasis in a mouse model of haemophilia. *Nature*doi:10.1038/nature10177.

92. Metzger, L and Iliakis, G (1991). Kinetics of DNA double-Strand break repair throughout the cell cycle as assayed by pulsed field gel electrophoresis in CHO cells. *Int. J. Radiat. Biol.* **59**: 1325–1339.

93. Kim, S, Kim, D, Cho, SW, Kim, J and Kim, J-S (2014). Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res.* **24**: 1012–1019.

94. Chiruvella, KK, Liang, Z and Wilson, TE (2013). Repair of double-strand breaks by end joining. *Cold Spring Harb. Perspect. Biol.*doi:10.1101/cshperspect.a012757.

95. Chapman, JR, Taylor, MRG and Boulton, SJ (2012). Playing the End Game: DNA Double-Strand Break Repair Pathway Choice. *Mol. Cell*doi:10.1016/j.molcel.2012.07.029.

96. Heyer, W-D, Ehmsen, KT and Liu, J (2010). Regulation of Homologous Recombination in Eukaryotes. *Annu. Rev. Genet.* **44**: 113–139.

97. Hustedt, N and Durocher, D (2017). The control of DNA repair by the cell cycle. *Nat. Cell Biol.*doi:10.1038/ncb3452.

98. Lieber, MR, Ma, Y, Pannicke, U and Schwarz, K (2003). Mechanism and regulation of human non-homologous DNA end-joining. *Nat. Rev. Mol. Cell Biol.*doi:10.1038/nrm1202.

99. Shrivastav, M, De Haro, LP and Nickoloff, JA (2008). Regulation of DNA double-strand break repair pathway choice. *Cell Res.*doi:10.1038/cr.2007.111.

100. Jasin, M (1996). Genetic manipulation of genomes with rare-cutting endonucleases. *Trends Genet.* **12**: 224–228.

101. Boissel, S, Jarjour, J, Astrakhan, A, Adey, A, Gouble, A, Duchateau, P, *et al.* (2014). MegaTALs: A rare-cleaving nuclease architecture for therapeutic genome engineering. *Nucleic Acids Res.* **42**: 2591–2601.

102. Kim, YG, Cha, J and Chandrasegaran, S (1996). Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc. Natl. Acad. Sci.* **93**: 1156–1160.

103. Cermak, T, Doyle, EL, Christian, M, Wang, L, Zhang, Y, Schmidt, C, *et al.* (2011). Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting (Nucleic Acids Research (2011) 39 (e82) DOI: 10.1093/nar/gkr218). *Nucleic Acids Res.* **39**: 7879.

104. Mali, P, Aach, J, Stranges, PB, Esvelt, KM, Moosburner, M, Kosuri, S, *et al.* (2013). CAS9 transcriptional activators for target specificity screening and paired nickases for

cooperative genome engineering. *Nat. Biotechnol.* **31**: 833–8.

105. Bauer, DE, Kamran, SC, Lessard, S, Xu, J, Fujiwara, Y, Lin, C, *et al.* (2013). An erythroid enhancer of BCL11A subject to genetic variation determines fetal hemoglobin level. *Science (80-.).*doi:10.1126/science.1242088.

106. Bjurström, CF, Mojadidi, M, Phillips, J, Kuo, C, Lai, S, Lill, GR, *et al.* (2017). Reactivating Fetal Hemoglobin Expression in Human Adult Erythroblasts Through BCL11A Knockdown Using Targeted Endonucleases. *Mol. Ther. - Nucleic Acids* **5**.

107. Canver, MC, Smith, EC, Sher, F, Pinello, L, Sanjana, NE, Shalem, O, *et al.* (2015). BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis. *Nature*doi:10.1038/nature15521.

108. Chang, KH, Smith, SE, Sullivan, T, Chen, K, Zhou, Q, West, JA, *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.*doi:10.1016/j.omtm.2016.12.009.

109. Cradick, TJ, Fine, EJ, Antico, CJ and Bao, G (2013). CRISPR/Cas9 systems targeting β-globin and CCR5 genes have substantial off-target activity. *Nucleic Acids Res.* **41**: 9584–9592.

110. Hendel, A, Bak, RO, Clark, JT, Kennedy, AB, Ryan, DE, Roy, S, *et al.* (2015). Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells. *Nat. Biotechnol.*doi:10.1038/nbt.3290.

111. Holt, N, Wang, J, Kim, K, Friedman, G, Wang, X, Taupin, V, *et al.* (2010). Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 in vivo. *Nat. Biotechnol.*doi:10.1038/nbt.1663.

112. Li, L, Krymskaya, L, Wang, J, Henley, J, Rao, A, Cao, LF, *et al.* (2013). Genomic editing of the HIV-1 coreceptor CCR5 in adult hematopoietic stem and progenitor cells using zinc finger nucleases. *Mol. Ther.*doi:10.1038/mt.2013.65.

113. Mandal, PK, Ferreira, LMR, Collins, R, Meissner, TB, Boutwell, CL, Friesen, M, *et al.* (2014). Efficient ablation of genes in human hematopoietic stem and effector cells using CRISPR/Cas9. *Cell Stem Cell*doi:10.1016/j.stem.2014.10.004.

114. Perez, EE, Wang, J, Miller, JC, Jouvenot, Y, Kim, KA, Liu, O, *et al.* (2008). Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. *Nat. Biotechnol.*doi:10.1038/nbt1410.

115. Saydaminova, K, Ye, X, Wang, H, Richter, M, Ho, M, Chen, HZ, *et al.* (2015). Efficient genome editing in hematopoietic stem cells with helper-dependent Ad5/35 vectors expressing site-specific endonucleases under microRNA regulation. *Mol. Ther. - Methods Clin. Dev.*doi:10.1038/mtm.2014.57.

116. Miller, JC, Tan, S, Qiao, G, Barlow, KA, Wang, J, Xia, DF, *et al.* (2011). A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol.*doi:10.1038/nbt.1755.

117. Tebas, P, Stein, D, Tang, WW, Frank, I, Wang, SQ, Lee, G, *et al.* (2014). Gene editing of

CCR5 in autologous CD4 T cells of persons infected with HIV. *N. Engl. J. Med.*doi:10.1056/NEJMoa1300662.

118. DiGiusto, DL, Cannon, PM, Holmes, MC, Li, L, Rao, A, Wang, J, *et al.* (2016). Preclinical development and qualification of ZFN-mediated CCR5 disruption in human hematopoietic stem/progenitor cells. *Mol. Ther. - Methods Clin. Dev.*doi:10.1038/mtm.2016.67.

119. DeWitt, MA, Magis, W, Bray, NL, Wang, T, Berman, JR, Urbinati, F, *et al.* (2016). Selection-free genome editing of the sickle mutation in human adult hematopoietic stem/progenitor cells. *Sci. Transl. Med.*doi:10.1126/scitranslmed.aaf9336.

120. Genovese, P, Schiroli, G, Escobar, G, Di Tomaso, T, Firrito, C, Calabria, A, *et al.* (2014). Targeted genome editing in human repopulating haematopoietic stem cells. *Nature* **510**: 235–240.

121. Hoban, MD, Cost, GJ, Mendel, MC, Romero, Z, Kaufman, ML, Joglekar, a. V., *et al.* (2015). Correction of the sickle-cell disease mutation in human hematopoietic stem/progenitor cells. *Blood*doi:10.1182/blood-2014-12-615948.

122. Richardson, C, Kazane, K, Feng, S, Bray, N, Schäfer, A, Floor, S, *et al.* (2017). CRISPR-Cas9 genome editing in human cells works via the Fanconi Anemia pathway. *Cris. genome Ed. Hum. cells Occur. via Fanconi anemia Pathw.*doi:10.1101/136028.

123. Paquet, D, Kwart, D, Chen, A, Sproul, A, Jacob, S, Teo, S, *et al.* (2016). Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature* **533**: 125–129.

124. Clough, C, Wang, Y, Khan, IF, Singh, S, Hung, K and Rawlings, DJ (2016). 132. Targeting the BTK Locus in Primary Human Hematopoietic Cells with TALENs and AAV Donor Template. *Mol. Ther.*doi:10.1016/s1525-0016(16)32941-0.

125. De Ravin, SS, Li, L, Wu, X, Choi, U, Allen, C, Koontz, S, *et al.* (2017). CRISPR-Cas9 gene repair of hematopoietic stem cells from patients with X-linked chronic granulomatous disease. *Sci. Transl. Med.*doi:10.1126/scitranslmed.aah3480.

126. Hubbard, N, Hagin, D, Sommer, K, Song, Y, Khan, I, Clough, C, *et al.* (2016). Targeted gene editing restores regulated CD40L function in X-linked hyper-IgM syndrome. *Blood* **127**: 2513–2522.

127. Lombardo, A, Genovese, P, Beausejour, CM, Colleoni, S, Lee, YL, Kim, KA, *et al.* (2007). Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. *Nat. Biotechnol.*doi:10.1038/nbt1353.

128. DeKelver, RC, Choi, VM, Moehle, EA, Paschon, DE, Hockemeyer, D, Meijsing, SH, *et al.* (2010). Functional genomics, proteomics, and regulatory DNA analysis in isogenic settings using zinc finger nuclease-driven transgenesis into a safe harbor locus in the human genome. *Genome Res.*doi:10.1101/gr.106773.110.

129. Hockemeyer, D, Soldner, F, Beard, C, Gao, Q, Mitalipova, M, Dekelver, RC, *et al.* (2009). Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. *Nat. Biotechnol.*doi:10.1038/nbt.1562.

130. Wang, J, Exline, CM, Declercq, JJ, Llewellyn, GN, Hayward, SB, Li, PWL, *et al.* (2015). Homology-driven genome editing in hematopoietic stem and progenitor cells using ZFN mRNA and AAV6 donors. *Nat. Biotechnol.* **33**: 1256–1263.

131. Wang, J, Exline, CM, DeClercq, JJ, Llewellyn, GN, Hayward, SB, Li, PW-L, *et al.* (2015). Homology-driven genome editing in hematopoietic stem and progenitor cells using ZFN mRNA and AAV6 donors. *Nat. Biotechnol.* **33**: 1256–1263.

132. Schiroli, G, Genovese, P, Castiello, MC, Capo, V, Gregory, PD, Holmes, MC, *et al.* (2016). 481. Targeted Genome Editing in Mouse Hematopoietic Stem/Progenitor Cells (HSPC) To Model Gene Correction of SCID-X1. *Mol. Ther.* **23**: S191.

133. Nakade, S, Tsubota, T, Sakane, Y, Kume, S, Sakamoto, N, Obara, M, *et al.* (2014). Microhomology-mediated end-joining-dependent integration of donor DNA in cells and animals using TALENs and CRISPR/Cas9. *Nat. Commun.* **5**.

134. Sakuma, T, Nakade, S, Sakane, Y, Suzuki, KIT and Yamamoto, T (2016). MMEJ-Assisted gene knock-in using TALENs and CRISPR-Cas9 with the PITCh systems. *Nat. Protoc.*doi:10.1038/nprot.2015.140.

135. Suzuki, K, Tsunekawa, Y, Hernandez-Benitez, R, Wu, J, Zhu, J, Kim, EJ, *et al.* (2016). In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration. *Nature* **540**: 144–149.

136. Branzei, D and Foiani, M (2008). Regulation of DNA repair throughout the cell cycle. *Nat. Rev. Mol. Cell Biol.* **9**: 297–308.

137. Lin, S, Staahl, BT, Alla, RK and Doudna, JA (2014). Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *Elife*doi:10.7554/eLife.04766.

138. Maruyama, T, Dougan, SK, Truttmann, MC, Bilate, AM, Ingram, JR and Ploegh, HL (2015). Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat. Biotechnol.*doi:10.1038/nbt.3190.

139. Srivastava, M, Nambiar, M, Sharma, S, Karki, SS, Goldsmith, G, Hegde, M, *et al.* (2012). An inhibitor of nonhomologous end-joining abrogates double-strand break repair and impedes cancer progression. *Cell*doi:10.1016/j.cell.2012.11.054.

140. Chu, VT, Weber, T, Wefers, B, Wurst, W, Sander, S, Rajewsky, K, *et al.* (2015). Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nat. Biotechnol.* **33**.

141. Gutschner, T, Haemmerle, M, Genovese, G, Draetta, GF and Chin, L (2016). Post-translational Regulation of Cas9 during G1 Enhances Homology-Directed Repair. *Cell Rep.* **14**: 1555–1566.

142. Pinder, J, Salsman, J and Dellaire, G (2015). Nuclear domain ‘knock-in’ screen for the evaluation and identification of small molecule enhancers of CRISPR-based genome editing. *Nucleic Acids Res.*doi:10.1093/nar/gkv993.

143. Yang, D, Scavuzzo, MA, Chmielowiec, J, Sharp, R, Bajic, A and Borowiak, M (2016).

Enrichment of G2/M cell cycle phase in human pluripotent stem cells enhances HDR-mediated gene repair with customizable endonucleases. *Sci. Rep.*doi:10.1038/srep21264.

144. Orthwein, A, Noordermeer, SM, Wilson, MD, Landry, S, Enchev, RI, Sherker, A, *et al.* (2015). A mechanism for the suppression of homologous recombination in G1 cells. *Nature* **528**: 422–426.

145. Van Tendeloo, VFI, Ponsaerts, P, Lardon, F, Nijs, G, Lenjou, M, Van Broeckhoven, C, *et al.* (2001). Highly efficient gene delivery by mRNA electroporation in human hematopoietic cells: Superiority to lipofection and passive pulsing of mRNA and to electroporation of plasmid cDNA for tumor antigen loading of dendritic cells. *Blood* **98**: 49–56.

146. Dever, DP, Bak, RO, Reinisch, A, Camarena, J, Washington, G, Nicolas, CE, *et al.* (2016). CRISPR/Cas9 β-globin gene targeting in human haematopoietic stem cells. *Nature* **539**: 384–389.

147. Daneshbod-Skibba, G, Martin, J and Shahidi, NT (1980). Myeloid and Erythroid Colony Growth in Non-anaemic Patients with Fanconi's Anaemia. *Br. J. Haematol.* **44**: 33–38.

148. Giri, N, Batista, DL, Alter, BP and Stratakis, CA (2007). Endocrine abnormalities in patients with Fanconi anemia. *J. Clin. Endocrinol. Metab.* **92**: 2624–2631.

149. Boitano, AE, Wang, J, Romeo, R, Bouchez, LC, Parker, AE, Sutton, SE, *et al.* (2010). Aryl hydrocarbon receptor antagonists promote the expansion of human hematopoietic stem cells. *Science (80-.).*doi:10.1126/science.1191536.

150. Fares, I, Chagraoui, J, Gareau, Y, Gingras, S, Ruel, R, Mayotte, N, *et al.* (2014). Pyrimidoindole derivatives are agonists of human hematopoietic stem cell self-renewal. *Science (80-.).*doi:10.1126/science.1256337.

151. Goessling, W, Allen, RS, Guan, X, Jin, P, Uchida, N, Dovey, M, *et al.* (2011). Prostaglandin E2 enhances human cord blood stem cell xenotransplants and shows long-term safety in preclinical nonhuman primate transplant models. *Cell Stem Cell*doi:10.1016/j.stem.2011.02.003.

152. Hoggatt, J, Singh, P, Sampath, J and Pelus, LM (2009). Prostaglandin E2 enhances hematopoietic stem cell homing, survival, and proliferation. *Blood*doi:10.1182/blood-2009-01-201335.

153. North, TE, Goessling, W, Walkley, CR, Lengerke, C, Kopani, KR, Lord, AM, *et al.* (2007). Prostaglandin E2 regulates vertebrate haematopoietic stem cell homeostasis. *Nature*doi:10.1038/nature05883.

154. Lis, R, Karrasch, CC, Poulos, MG, Kunar, B, Redmond, D, Duran, JGB, *et al.* (2017). Conversion of adult endothelium to immunocompetent haematopoietic stem cells. *Nature*doi:10.1038/nature22326.

155. Sugimura, R, Jha, DK, Han, A, Soria-Valles, C, Da Rocha, EL, Lu, YF, *et al.* (2017). Haematopoietic stem and progenitor cells from human pluripotent stem cells. *Nature*doi:10.1038/nature22370.

156. Petrillo, C, Cesana, D, Piras, F, Bartolaccini, S, Naldini, L, Montini, E, *et al.* (2015). Cyclosporin A and rapamycin relieve distinct lentiviral restriction blocks in hematopoietic stem and progenitor cells. *Mol. Ther.*doi:10.1038/mt.2014.193.

157. Santoni De Sio, FR, Cascio, P, Zingale, A, Gasparini, M and Naldini, L (2006). Proteasome activity restricts lentiviral gene transfer into hematopoietic stem cells and is down-regulated by cytokines that enhance transduction. *Blood*doi:10.1182/blood-2005-10-4047.

158. Wang, CX, Sather, BD, Wang, X, Adair, J, Khan, I, Singh, S, *et al.* (2014). Rapamycin relieves lentiviral vector transduction resistance in human and mouse hematopoietic stem cells. *Blood*doi:10.1182/blood-2013-12-546218.

159. Baum, CM, Weissman, IL, Tsukamoto, AS, Buckle, AM and Peault, B (1992). Isolation of a candidate human hematopoietic stem-cell population. *Proc. Natl. Acad. Sci. U. S. A.*doi:10.1073/pnas.89.7.2804.

160. Masiuk, KE, Brown, D, Laborada, J, Hollis, RP, Urbinati, F and Kohn, DB (2017). Improving Gene Therapy Efficiency through the Enrichment of Human Hematopoietic Stem Cells. *Mol. Ther.*doi:10.1016/j.ymthe.2017.05.023.

161. Wu, HW, Hsu, RC, Lin, CC, Hwang, SM and Lee, G Bin (2010). An integrated microfluidic system for isolation, counting, and sorting of hematopoietic stem cells. *Biomicrofluidics*doi:10.1063/1.3454767.

162. Chhabra, A, Ring, AM, Weiskopf, K, Schnorr, PJ, Gordon, S, Le, AC, *et al.* (2016). Hematopoietic stem cell transplantation in immunocompetent hosts without radiation or chemotherapy. *Sci. Transl. Med.*doi:10.1126/scitranslmed.aae0501.

163. Czechowicz, A, Kraft, D, Weissman, IL and Bhattacharya, D (2007). Efficient transplantation via antibody-based clearance of hematopoietic stem cell niches. *Science (80-.).*doi:10.1126/science.1149726.

164. Palchaudhuri, R, Saez, B, Hoggatt, J, Schajnovitz, A, Sykes, DB, Tate, TA, *et al.* (2016). Non-genotoxic conditioning for hematopoietic stem cell transplantation using a hematopoietic-cell-specific internalizing immunotoxin. *Nat. Biotechnol.*doi:10.1038/nbt.3584.

165. Xue, X, Pech, NK, Shelley, WC, Srour, EF, Yoder, MC and Dinauer, MC (2010). Antibody targeting KIT as pretransplantation conditioning in immunocompetent mice. *Blood*doi:10.1182/blood-2010-07-295949.

166. Gelsinger, P and Shamoo, AE (2008). Eight Years after Jesse's Death, Are Human Research Subjects Any Safer? *Hastings Cent. Rep.*doi:10.1353/hcr.2008.0022.

167. Wilson, JM (2018). Lessons learned from the gene therapy trial for ornithine transcarbamylase deficiency. *Get. to Good Res. Integr. Biomed. Sci.*doi:10.1016/j.ymgme.2008.12.016.

168. Braun, CJ, Boztug, K, Paruzynski, A, Witzel, M, Schwarzer, A, Rothe, M, *et al.* (2014). Gene therapy for Wiskott-Aldrich syndrome-long-term efficacy and genotoxicity. *Sci. Transl. Med.*doi:10.1126/scitranslmed.3007280.

169. Hacein-Bey-Abina, S, Von Kalle, C, Schmidt, M, Le Deist, F, Wulffraat, N, McIntyre, E, *et al.* (2003). A serious adverse event after successful gene therapy for X-linked severe combined immunodeficiency [1]. *N. Engl. J. Med.*doi:10.1056/NEJM200301163480314.

170. Stein, S, Ott, MG, Schultze-Strasser, S, Jauch, A, Burwinkel, B, Kinner, A, *et al.* (2010). Genomic instability and myelodysplasia with monosomy 7 consequent to EVI1 activation after gene therapy for chronic granulomatous disease. *Nat. Med.*doi:10.1038/nm.2088.

1.12 WEB RESOURCES

bluebird bio Inc. 2017 Published online September 06, 2017
http://www.businesswire.com/news/home/20170906005645/en/
European Medicines Agency. 2016 Published online April 01, 2016
http://www.ema.europa.eu/docs/en_GB/document_library/Press_release/2016/04/WC500204145.pdf

CHAPTER 2

RESEARCH PROJECT

Optimizing Integration and Expression of Transgenic Bruton's Tyrosine Kinase for CRISPR/Cas9 Mediated Gene Editing of X-Linked Agammaglobulinemia

David H. Gray[1], Isaac Villegas[2], Joseph Long[2], Jasmine Santos[2], Alexandra Keir[2], Alison Abele[2], Caroline Y. Kuo[3,4], and Donald B. Kohn[2,4-6].

Affiliations:

1 Molecular Biology Institute, University of California, Los Angeles, CA 90095, USA

2 Department of Microbiology, Immunology & Molecular Genetics, University of California, Los Angeles, CA 90095, USA

3 Division of Allergy & Immunology, David Geffen School of Medicine at the University of California, Los Angeles, CA 90095, USA

4 Department of Pediatrics, David Geffen School of Medicine at the University of California, Los Angeles, CA 90095, USA

5 Department of Molecular and Medical Pharmacology, University of California, Los Angeles, CA 90095, USA

6 Eli & Edythe Broad Center of Regenerative Medicine & Stem Cell Research, University of California, Los Angeles, CA 90095, USA

Abstract:

X-Linked Agammaglobulinemia (XLA) is a monogenic primary immune deficiency characterized by very low levels of immunoglobulins and greatly increased risks for recurrent and severe infections. Patients with XLA have a loss of function mutation in the Bruton's Tyrosine Kinase (*BTK*) gene and fail to produce mature B lymphocytes. Gene editing in the hematopoietic stem cells (HSCs) of XLA patients to correct or replace the defective gene should restore B cell development and the humoral immune response. We used the CRISPR/Cas9 (clustered regularly interspaced short palindromic repeats) platform to precisely target integration of a corrective, codon-optimized *BTK* complementary DNA (cDNA) cassette into its endogenous locus. This process is driven by homologous recombination and should place the transgenic *BTK* under transcriptional control of its endogenous regulatory elements. Each integrated copy of this cDNA in BTK-deficient K562 cells produced only 11% as much BTK protein as the wildtype gene. The donor cDNA was modified to include the terminal intron of the *BTK* gene. Successful integration of the intron-containing *BTK* donor led to a nearly two-fold increase in BTK expression per cell over the base donor. However, this donor variant was too large to package into an adeno-associated viral vector for delivery into primary cells. Donors containing truncated variants of the terminal intron also produced elevated expression, although to a lesser degree than the full intron. Addition of the Woodchuck Hepatitis Virus Posttranscriptional Regulatory Element (WPRE) led to a large boost in *BTK* transgene expression. Combining these modifications led to a *BTK* donor template that generated nearly physiological levels of BTK expression in cell lines. These reagents were then optimized to maximize integration rates into human hematopoietic stem and progenitor cells, which have reached potentially therapeutic levels *in vitro*. The novel donor modifications support

effective gene therapy for XLA and will likely assist in the development of other gene editing-based therapies for genetic disorders.

Introduction:

Bruton's Tyrosine Kinase (BTK) is a cytoplasmic kinase that is a lynchpin of multiple signaling pathways including B cell receptor (BCR) signaling.[1–4] Defective BCR signaling halts development of B lymphocytes and results in the absence of mature B lymphocytes and antibody production, a characteristic of X-linked Agammaglobulinemia (XLA).[2,5] Without the protection from functional antibodies, patients are susceptible to infections and have reduced life expectancies.[6] The current standard of care for XLA is subcutaneous or intravenous antibody supplementation from healthy donors. This treatment provides a substantial improvement to the patient's quality of life and dramatically increases patient life expectancies. However, the treatment requires ongoing immunoglobulin injections for life that are expensive and imperfect, leaving susceptibility to recurrent pulmonary infections, invasive viral infections, and inflammatory bowel disease.[6]

Allogeneic hematopoietic stem cell (HSC) transplant is currently the only permanent cure for XLA, although it is very rarely performed due to the transplant-associated risks. Graft vs host disease and graft rejection are generally considered unacceptably high risks when compared to the moderate severity of clinically managed XLA. Gene therapy using autologous HSCs would maintain the curative effects of allogeneic transplants while reducing the risks. Multiple groups have made progress towards viral vector mediated gene transfer to deliver a functional *BTK* gene to human HSCs.[7,8] While lentiviral vectors have improved dramatically in recent years, there remains some inherent risk of insertional oncogenesis (IO) with any semi-randomly integrating vector.[9,10] In some cases, that risk can be tolerated because of the extreme severity of the disease.

However, due to the relatively effective current treatment for XLA, any appreciable risk of oncogenesis may be unacceptable. Lentiviral based XLA therapies have also run into hurdles restoring endogenous expression patterns. Using the natural *BTK* promoter and enhancer sequences to drive transgene expression produced much lower than wildtype protein levels.[8] Stronger promoters and enhancers increased this expression, but made it exceedingly difficult to get appropriate expression in all the relevant cell types and may elevate IO risks.[7]

It remains somewhat unclear what range of BTK expression is required to restore B cell development and produce protective levels of antibodies. Previous work has demonstrated that BTK expression near physiological levels leads to the most efficient signaling.[11] Overexpression of BTK is correlated with some types of B lymphoid leukemias (e.g. chronic lymphocytic leukemia) and BTK inhibitors like ibrutinib are revolutionizing treatment for many of these patients.[12,13] Though BTK overexpression does not seem to be sufficient for transformation alone, the correlation is worrisome for XLA gene therapies. These data together suggest a relatively narrow window of BTK expression will be clinically beneficial; too little BTK expression may not restore B lymphopoiesis while too much may affect signaling efficiency or even carry risks of oncogenesis. Our approach for correcting XLA instead utilizes the CRISPR/Cas9 (Clustered Regularly Interspaced Short Palindromic Repeats) platform to improve the fidelity of treatment by first creating a targeted double stranded DNA break (DSB) at the *BTK* locus.[14] Following Cas9-mediated DNA cleavage at the target site, the cell can use one of multiple mechanisms to repair the DSB. The most notable of these pathways are non-homologous end joining (NHEJ), which results in deletions or insertions of random nucleotides at the repair site, or homologous recombination using a template DNA molecule to guide repair, which is the basis of this method of gene therapy. Homology-directed repair (HDR) of BTK mutants can occur if high numbers of

a corrective *BTK* donor DNA are present in the nucleus during DSB repair. These donor molecules contain the *BTK* cDNA sequence flanked by "homology arms" that parallel the cut site and serve as templates for homologous recombination.[15,16]

Addition of a corrective copy of the *BTK* gene into the start of the gene could be an effective treatment for every patient with exonic mutations anywhere downstream. We utilized the CRISPR/Cas9 to integrate a potentially therapeutic, human *BTK* cDNA sequence into the 5' end of the endogenous *BTK* locus. We initially observed sub-optimal BTK protein production from the wild-type cDNA and identified several modifications to the *BTK* transgene cassette to dramatically improve expression levels. Integration and expression from donor integration at multiple target sites were assessed and optimized to produce a novel therapy that may provide a safe, effective gene therapy for XLA.

Results:

A BTK cDNA integrated efficiently into the endogenous locus but produced sub-physiological levels of BTK expression

Initially, a Cas9 target site at the 3' end of *BTK* intron 1 was chosen because it could be curative to patients with any loss of function mutations in the BTK protein coding region (Fig 1a). Multiple guide RNAs were tested in cell lines to identify the optimal site for allelic disruption (data not shown). A donor plasmid was designed to achieve targeted integration of the full *BTK* cDNA sequence to the DSB created by the Cas9/single guide RNA (sgRNA) at this site. This *BTK* sequence was codon optimized to reduce homology to the intact genomic locus without changing the amino acids encoded.[17] Three C-terminal hemagglutinin tags were added upstream of the stop codon to differentiate between wildtype (endogenous) and transgenic BTK. The full *BTK* 3'

untranslated region (UTR) was also included following the stop codon to support transcriptional termination and polyadenylation. Five hundred bp long homology arms were added to either side of the donor construct that matched the sequences flanking the target site in the genomic DNA. The sgRNA/Cas9 was delivered as an expression plasmid in conjunction with a plasmid containing the described *BTK* donor. BTK-deficient K562 cells were created using Cas9 to generate a homozygous 4 bp deletion in exon 15 of the *BTK* gene and absence of BTK protein was verified via immunoblot (Suppl. Fig 1).

Delivery of the Cas9/sgRNA endonuclease alone to the BTK-deficient K562 cells generated allelic disruption in 24.3% ± 5.1% of cells (Fig 1b). Adding only the Cas9/sgRNA endonuclease plasmid or only the *BTK* cDNA donor plasmid led to no detectable targeted integration, while adding both nuclease and donor together achieved integration in 13.0% ± 1.1% of cells (Fig 1c). Immunoblots probing for the HA tags on the *BTK* donor demonstrated that the transgenic protein was present in the BTK-disrupted K562 cells (Fig 1d). However, probing for BTK protein showed that the actual expression per cell was only 11.1% ± 1.4% of the wildtype protein in parental K562 cells (Fig 1d). Transgenic *BTK* mRNA normalized for the DNA integration rate was only at 20.2% ± 1.5% of wildtype levels (Fig 1e). To identify off-target cleavage events, genome-wide unbiased identification of double-stranded breaks enabled by high throughput sequencing (GUIDE-seq) was performed in K562 cells and detected only on-target cleavage events with this sgRNA to *BTK* intron 1 (Fig 1f).[18]

To test for function of the transgenic *BTK* construct, a Ramos B cell subline was engineered using Cas9-mediated gene knock-out to have a frameshift mutation in *BTK* exon 15 that led to an early stop codon, using Cas9-mediated gene knock-out (data not shown). The BTK-deficient Ramos line was then used to create a clonal BTK rescue line with the Cas9/sgRNA and *BTK* cDNA

donor described above integrated into the intron 1 site. Wildtype, BTK-deficient, and BTK rescued (gene edited) Ramos clonal cell lines were stimulated with a BCR-crosslinking antibody fragment and immunoblotted for phosphorylated BTK (Fig 1g). The BTK rescued cell line responded to the BCR-crosslinking with an increase in BTK protein phosphorylation, suggesting that the transgene folds and localizes correctly.

Addition of the BTK terminal intron or the WPRE to *BTK* donor templates boosted transgene expression

Previous work identified the importance of introns for optimal transgene expression.[19,20] Multiple mechanisms function together for this phenomenon; intron splicing improves nuclear export of the transcript and polyadenylation of the nascent RNA.[21] Additionally, introns can include important enhancer sequences. The terminal intron of a gene plays a particularly important role in efficient transcription termination and 3' end processing of transcripts; removal of the 3' splice site immediately preceding the terminal exon led to a substantial drop in polyadenylation of the resulting transcript.[22]

The terminal intron of *BTK*, intron 18, is 3.2 kb long, which would be too large to package with the rest of the *BTK* cDNA donor in an adeno-associated virus serotype 6 (AAV6) vector.[23] These AAV6 vectors are ideal for efficient donor delivery into primary cells, while plasmids yield very high toxicity with low efficacy in HSPCs.[24] However, the donor can still be tested in cell lines via delivery as a plasmid. To the best of our knowledge, BTK intron 18 has no known or putative enhancer binding sites and no patients have been identified with pathogenic mutations solely in intron 18 (except mutations that disrupted splice sites). Three donor variants were created: one

with the full intron 18 (I18), another with a 679 bp truncated variant of intron 18 (I18t), and finally one with the much smaller intron 6 of *BTK* intron (I6), each inserted between exons 18 and 19 of the cDNA sequence. Importantly, all three variants retained the 5' and 3' splicing signals (Fig 2a). These donors, as well as the intron-less base donor ("cDNA"), were assessed as electroporated plasmids in BTK-deficient K562 cells.

The full intron 18 donor generated the lowest rates of integration at 7.2% ± 1.7%, while the other three donors had similar integration frequencies between 10.8% and 15.3% (Fig 2b). The presence of any of the three intron variants led to increased BTK expression per modified cell (Fig 2c). The donors containing the I6, I18, and I18t intron variants produced 2.2, 3.4, and 2.8 times as much of the transgenic protein as the base cDNA donor. The I18t donor markedly increased BTK expression while remaining of a size compatible with packaging into AAV6 vectors.

Another element worth assessing for the potential to increase expression levels is the Woodchuck Hepatitis Virus Posttranscriptional Regulatory Element (WPRE). WPRE is often added to retro- and lentiviral vectors to improve titers by increasing the abundance of the viral RNA transcript in the packaging cells.[25] It is also used in some types of vectors to increase transgene expression levels. This element has been used for decades, although there remains some degree of uncertainty about its mechanisms of action. The most commonly described mechanism of action is its formation of a tertiary structure, sterically blocking the 3' end of the mRNA strand from degradation. Other proposed mechanisms include improved transcriptional terminal or more efficient nuclear export of transcripts. We designed a new *BTK* cDNA donor variant for integration into *BTK* intron 1 with the WPRE added immediately after the 3' UTR, just before the polyadenylation signal (Fig 2d). This new WPRE containing donor was tested head to head with the base cDNA donor in BTK-deficient K562 cells. While integration rates were similar between

the two donors, the WPRE-containing *BTK* donor produced nearly tripled BTK protein expression (Fig 2e,f).

Simultaneous addition of a PAM mutation, a truncated intron, and the WPRE to *BTK* donor templates produced nearly wildtype levels of BTK expression

While the codon optimization of exonic donors usually disrupts the sgRNA binding site, intronic regions like the BTK intron 1 site have no codons to be optimized, so the full Cas9 binding site and PAM sequence remain intact. The preservation of this Cas9/sgRNA binding site allowed the integrated donor to be re-cleaved and repaired via NHEJ, generating indels at the integration junction of every sample tested (data not shown). Removing the PAM sequence within the *BTK* donor (ΔPAM) resulted in 100% base perfect 5' integration junctions without affecting total integration rates in K562 cells.

Donor plasmids were designed containing combinations of the three modifications described: ΔPAM, truncated intron 18 addition, and WPRE addition. When both the truncated intron and the WPRE were both added to the donor plasmid, the resultant size was too large to be efficiently packaged into an AAV vector, necessitating the creation of a further truncated intron 18 variant, dubbed micro intron 18 (I18u). The new variants created were: ΔPAM-WPRE, ΔPAM-I18t-WPRE, and ΔPAM-I18u-WPRE (Fig 3a). Each of these donor templates were electroporated as plasmid into BTK deficient K562 cells in conjunction with the BTK intron 1 targeting sgRNA/Cas9 expression plasmid. The integration rates of each donor were comparable except the base ΔPAM donor, which integrated at slightly higher efficiency (Fig 3b).

RNA from these cells was collected and analyzed via reverse transcription ddPCR to determine relative quantities of the BTK transcripts (normalized to the housekeeping gene IPO8).

This expression was also scaled via the previously determined integration rates to effectively compare expression per integrated (or wildtype) copy of the BTK gene. The base ΔPAM donor produced only 9.1% ± 0.2% of wildtype expression (Fig 3c). The I18t donor generated 19.2% ± 1.9% of wildtype levels, while WPRE alone led to 22.4% ± 0.8% and 23.9% ± 4.1% with the PAM mutation. Combining the three elements led to the highest expression—46.8% ± 1.1% of wildtype BTK. The packageable combination of the elements (ΔPAM-I18u-WPRE) generated second highest RNA levels at 38.1% ± 0.5%.

Protein lysates from each treated population were prepared and analyzed via immunoblot for BTK, with Actin as a loading control. The results reflected the trend seen in the RNA expression data (Fig 3d). The base donor produced 16.8% the amount of BTK protein per copy of the gene when compared to wildtype cells. Integration of the I18t containing donor led to an increase in expression to 24.9% of wildtype. Donors with the WPRE element yielded 49.0% of wildtype expression with the PAM intact and 41.6% for the ΔPAM-WPRE variant. Again, the highest expression came from the ΔPAM-I18t-WPRE donor at 80.4% of wildtype, while the packageable-sized variant (ΔPAM-I18u-WPRE) performed nearly as well, with 74.6% of wildtype expression.

Addition of the WPRE, but not the truncated intron 18, produced elevated BTK expression in a T lymphocyte line

While the integration of a donor with both the truncated intron 18 and WPRE led to nearly wildtype levels of BTK mRNA and protein, the inclusion of these elements raised the question of whether the lineage-specificity of the donor was maintained. While BTK is found in most hematopoietic lineages, there are notable exceptions such as T lymphocytes.[26] Four BTK donors were compared in the Jurkat T cell line by electroporation of the Cas9/sgRNA and donor plasmids:

ΔPAM, I18t, ΔPAM-WPRE, and ΔPAM-I18u-WPRE. Both WPRE-containing donors yielded about three times higher rates of integration than the donors lacking the WPRE (Fig 4a). Immunoblot analysis of these treated cells normalized for percent integration demonstrated BTK protein expression from Jurkat T cells treated with the WPRE-containing donors but not from the base or truncated intron 18 donors (Fig 4b).

Optimization of donor integration into BTK intron 1 in human CD34+ hematopoietic stem and progenitor cells (HSPCs)

Primary human granulocyte colony-stimulating factor (G-CSF) mobilized peripheral blood CD34+ HSPCs were thawed and pre-stimulated for two days before electroporation with the sgRNA/Cas9. For experiments assessing integration rates, cells were transduced with an AAV6 vector containing the ΔPAM donor for 24 hours immediately following electroporation (Fig 5a). Cas9 was delivered as either a ribonucleoprotein (RNP) pre-complexed with the sgRNA or as Cas9-encoding mRNA. The sgRNA was either *in vitro* transcribed (IVT) or chemically synthesized with 2'-O-methyl and 3' phosphorothioate inter-nucleotide linkages on the three terminal residues on both the 5' and 3' ends of the molecule. These chemical modifications have been reported to increase the stability of the sgRNA molecules and lead to more efficient editing.[27]

Initial experiments compared the different endonuclease delivery methods for the ability to induce allelic disruption in the absence of a donor template. In this experiment, Cas9 mRNA vastly outperformed Cas9 RNP at the *BTK* intron 1 site (Fig 5b), which is an observation we have not found true for any other target sites. Chemically-modified sgRNA co-electroporated with Cas9 mRNA was by far the most effective combination of reagents, producing nearly 80% allelic

disruption (Fig 5b). The modified sgRNA used with recombinant Cas9 in RNP was less effective for allelic disruption at the intron 1 site, and the IVT sgRNA and Cas9 mRNA gave the lowest activity.

When a donor template was introduced to the cells via AAV6 transduction post electroporation, successful integration events could be identified. The combination of chemically-modified sgRNA with Cas9 mRNA also led to the highest frequencies (~20%) of targeted integration (Fig 5c). The levels of targeted donor integration with IVT sgRNA and Cas9 mRNA and with chemically-modified sgRNA and Cas9 protein paralleled the activity of these nuclease combinations for gene disruption in the absence of the donor. Titration of the AAV6 donor vectors led to a clear dose response for targeted gene integration, with increasing amounts of donor generating increasing donor integrations (Fig 5d). However, higher AAV6 vector multiplicities of infection (MOI) also reduced viability and expansion of the treated cells, suggesting cytotoxicity at the higher vector doses (data not shown).

While delivering the Cas9 as mRNA with the sgRNA to intron 1 led to efficient editing of PBSCs, the inefficiency of Cas9 delivered as RNP at this site prompted a search for additional target sites. Seven additional sgRNAs were assessed for their abilities to induce allelic disruption in *BTK* intron 1 or exon 2 in primary human CD34+ HSPC, using RNP made with IVT sgRNA and Cas9 protein (Supp Fig 2). This search identified a prime candidate sgRNA targeting *BTK* exon 2 that targets 7 bp after the *BTK* translational start codon, with high *in silico* predicted nuclease specificity.

Donor integration into *BTK* exon 2 maintains similar BTK expression levels with improved integration into human PBSCs with RNP

Plasmid donors were developed with homology arms for the *BTK* exon 2 site, following the same principles as for the intron 1 donors (Fig 6a). Three variants were evaluated for expression levels generated from a *BTK* donor integrated at the *BTK* exon 2 target site: a base donor with the codon-optimized *BTK* cDNA, a *BTK* donor with I18t, and a donor containing both the WPRE and the I18u fragment. The codon optimization of the cDNA in the exon 2 donor disrupts the sgRNA binding site, so further PAM modification was not necessary. Each of these donor plasmids were delivered to BTK-deficient K562 cells via electroporation along with the exon 2 targeting sgRNA/Cas9 expression plasmid.

The base cDNA and I18t donors each generated over 10% *BTK* integration, while the I18u-WPRE donor integrated in 6% of cells (Fig 6b). Probing protein lysates from these cells for BTK expression yielded a pattern similar to the pattern seen with the intron 1 reagents (Fig 6c). The base donor generated much lower levels of BTK expression than wild-type K562 cells. Addition of the truncated intron 18 variant yielded a slight increase in BTK expression, while the micro intron 18 fragment and WPRE together provided a large boost to expression, almost to wildtype BTK levels. The RNA expression levels by the *BTK* transgene very closely mirrored the quantified BTK protein levels from the immunoblot (Fig 6d).

The exon 2 sgRNA was analyzed for its off-target nuclease activity in the K562 cell line using GUIDE-seq.[18] Two off-target loci yielded substantial cleavage events: intron 5 of the thrombospondin type 1 domain containing 4 gene and intron 21 of the low density lipoprotein receptor related protein 5 with 26.0% and 4.2% of the total cleavage events, respectively (Fig 6e). To assess whether the off-target activity could be reduced, the exon 2 sgRNA was tested with three

engineered Cas9 variants (eSP, VP12 and Alt-R) reported to have higher fidelity than wildtype Cas9.[28–30] All three variants nearly eliminated off-target activity, with 99.97%, 99.97%, or 99.41% on-target activity instead of only 70.88% with the wildtype protein (Fig 6e).

These Cas9 variants were then tested with the exon 2 sgRNA for the ability to achieve targeted integration in human CD34+ PBSCs. The exon 2 sgRNA and each of the Cas9 variants were electroporated as RNP while the exon 2 WPRE-I18u donor template was transduced as an AAV6 vector. All three Cas9 variants led to integration levels comparable to the wildtype Cas9 protein (Fig 6f). Together, the specificity and activity data support the function of all three high fidelity Cas9 variants for use with the sgRNA to exon 2 and identify this site with the I18u-WPRE donor as one achieving clinically-relevant levels of integration and BTK expression.

Discussion:

XLA is a promising candidate for gene therapy via targeted gene insertion for multiple reasons. Firstly, the current standard of care has room for improvement, both in efficacy and duration. The only cure currently available, allogeneic hematopoietic stem cell transplantation is rarely performed for XLA patients due to the associated risks of graft vs host disease, graft rejection and transplant-related morbidities that generally outweigh the benefits of the treatment. Autologous gene therapy offers similar potential benefits as allogeneic transplant with potentially far lower risks. Lentiviral vectors can be used to deliver a functional copy of BTK, although the drawbacks of non-physiological regulation of the transgene and the added risk of insertional oncogenesis may be untenable for treatment of a relatively mild disease. We propose that gene editing is the safest route to provide an enduring therapy for XLA.

Another aspect of XLA that makes it a good candidate for gene therapy is the strong selective advantage that B cells have with functional BTK protein. In a murine model of the disease, low quantities of healthy HSPCs engrafted alongside BTK-deficient HSPCs were able to produce a disproportionally large percentage of mature B cells. [31] As few as 0.5% BTK-intact hematopoietic stem cells (out of $5x10^6$ total cells) produced splenic B cells at frequencies higher than 10% of levels found in a healthy mouse as well as antibody responses to vaccine challenge within the normal range.[31] If this trend applies as expected in XLA patients, who have a more severe B cell deficiency than this murine model, the edited cells may have an even stronger selective advantage such that even a small fraction of edited HSC would be sufficient for clinical benefit. This trait will help overcome one of the main challenges for current HDR-based gene editing therapies: achieving sufficient rates of gene modification in long-term HSCs.

The work presented here demonstrates the potential for *BTK* gene editing to achieve a lasting therapy for XLA. Initial work focused on achieving targeted integration into intron 1 of the BTK gene. The sgRNA used for this site led to reasonable allelic disruption frequencies when delivered as a plasmid in cell lines as well as no detectable off-target cleavage by GUIDE-seq. Experiments in K562 cells yielded consistently high levels of integration at this site. However, levels of BTK RNA and protein in cells were only 10-20% of wildtype levels. The initial hypothesis of the project was that site-specific integration of the corrective *BTK* donor template into its endogenous locus would lead to nearly physiological expression of the gene, which these data do not fully support. The reduction in expression prompted a search for mechanisms to improve transgenic BTK expression without further disrupting the regulation of the gene. Although it remains unclear how much BTK expression per cell is necessary to provide a clinical

benefit to XLA patients, previous work suggested that optimal BTK signaling efficiency occurs near wildtype expression levels.[11]

The two main donor modifications that we identified to achieve that goal were the additions of a terminal intron and the WPRE. The terminal intron of the *BTK* gene improved expression of the transgene while preserving its lineage specificity. While the terminal intron has been reported to have distinct functions, it may be interesting to explore whether the presence of additional BTK introns further boosts protein expression. The WPRE also boosted expression levels, though it may have altered the lineage-specificity of expression, with the unexpected observation of expression from the *BTK* transgenes that included the WPRE in Jurkat T cells. Re-addition of elements from the endogenous *BTK* locus, such as the *BTK* cDNA, 3' UTR, or introns, may have less intrinsic risk than adding exogenous elements such as the WPRE. The specificity of expression from these BTK transgenes in gene-edited HSCs after transplantation and multi-lineage differentiation needs to be examined to determine if it is both beneficial for restoring B lymphocyte function and safe without skewing leukocyte differentiation or activity. Animal models of the therapy are likely the best method to assess the safety of the treatment.

Integration of the *BTK* donors into human PBSCs also presented some unexpected difficulties. When the nuclease was delivered as IVT sgRNA with Cas9 protein, the intron 1 donor integrated at much lower than expected levels. Transitioning to a chemically synthesized (but unmodified) guide made no difference for gene editing frequencies. However, adding 2'-O-methyl and 3' phosphorothioate inter-nucleotide linkages on the terminal residues to stabilize the sgRNA tripled integration levels of the donor template.[32]

The intron 1 site also had the unusual property that using Cas9 mRNA instead of recombinant Cas9 protein led to a large increase in *BTK* donor integration frequencies. We have

not previously observed this strong preference for Cas9 mRNA over Cas9 protein for RNP with any other target site. With these two changes to the delivery scheme for the intron 1 site (Cas9 mRNA and synthetic sgRNA with modified bases), integration levels of up to 30% were achieved. This integration frequency should be well above the threshold for clinical efficacy of the gene editing for XLA. However, clinical grade mRNA may be more difficult than recombinant Cas9 protein to procure in the quantity and quality necessary for use in clinical-scale editing. This concern spawned a search for alternative target sites near the 5' end of the *BTK* locus and yielded a highly efficacious sgRNA targeting exon 2 of *BTK*.

The exon 2 sgRNA led to similar levels of editing in cell lines as the intron 1 sgRNA. However, unlike the intron 1 guide, the exon 2 guide had substantial off-target activity. Three recently developed high-fidelity Cas9 variants were tested for their abilities to reduce the off-target activity of the exon 2 sgRNA while maintaining the on-target activity. All three variants succeeded at nearly eliminating off-target activity while retaining most, if not all, of the on-target integration efficacy. Any of these variants could be considered for translation of *BTK* gene editing towards the clinic.

In all, the work here constitutes major progress towards a clinically relevant gene editing scheme for an enduring treatment for XLA. Editing the *BTK* locus came with unanticipated hurdles that required the development of novel modifications to the *BTK* donor template as well as fine tuning of the reagent delivery scheme. The addition of the *BTK* terminal intron to the donor template led to a significant increase in BTK expression. We expect terminal intron addition to cDNA cassettes for improved expression to also be applicable and beneficial for other gene editing based therapies that are in development. The WPRE provided an even larger expression boost, but it may have led to non-physiological regulation of the *BTK* gene. To fully determine whether this

expression we observed in T lymphoid lineages with *BTK* donors containing the WPRE is problematic, further studies in murine models of XLA will be performed. This work is a step towards improving the lives of XLA patients and provides insights that may facilitate the development of therapies for other genetic disorders.

Materials and Methods:

Donor Template Assembly: The human *BTK* cDNA synthesis was codon optimized via the GeneOptimizer web tool [ThermoFisher Scientific, Waltham, MA] and commercially synthesized by IDT [Integrated DNA Technologies, Coralville, IA]. All of the donor templates contains *BTK* exons 2 through 19 (2010 bp), the 3' untranslated region (428 bp), three C terminal HA epitope tags attached by a linker (99 bp), and two homology arms that match the sequences flanking the respective target site in the genomic DNA (500 bp each). Donors for the intron 1 target site also included a small portion of the intron 1 sequence between the cut site and the start of exon 2, to ensure efficient splicing (33 bp).

Guide RNA and Cas9: For K562 and Jurkat experiments, single gRNA was delivered via the pX330-U6-Chimeric_BB-CBhhSpCas9 expression plasmid, which was a gift from Feng Zhang [Addgene plasmid #42230; http://n2t.net/addgene:42230; RRID:Addgene_42230].[14] Paired oligonucleotides representing the gRNA sequences were synthesized [Integrated DNA Technologies, Coralville, IA], annealed, and cloned into the gRNA/Cas9 expression plasmid. For experiments with gRNA delivered as RNA (rather than on an expression plasmid), it was generated via IVT or chemically synthesized [Synthego Corporation, Menlo Park, CA] as specified in each experiment. IVT guide RNA was produced from a PCR-generated template using the HiScribe T7 Quick High Yield RNA Synthesis Kit [Cat: E2050S, New England Biolabs, Ipswich, MA].[33] Chemically synthesized gRNA was purchased as either unmodified or chemically modified with

2'-O-methyl and 3' phosphorothioate inter-nucleotide linkages on the three terminal residues [Synthego Corporation, Menlo Park, CA].[27] All Cas9 used in this work was *Streptococcus pyogenes* Cas9 (SpCas9) with the exception of the high fidelity engineered variants. Cas9 mRNA was also produced via IVT. A plasmid containing wildtype Cas9 was cloned with the protein under the control of the T7 promoter. The plasmid was linearized via restriction digest and the IVT was performed using the mMESSAGE mMACHINE T7 Ultra Transcription kit [Cat: AM1345, ThermoFisher Scientific, Waltham, MA]. The resulting mRNA was purified using the RNEasy Minelute kit [Cat: 74204. Qiagen, Germantown, MD]. Cas9 protein for wildtype SpCas9, VP12 Cas9, and eSpCas9 were obtained from the University of California, Berkeley QB3 MacroLab.[28,29] Alt-R Cas9 protein was purchased from IDT [Cat: 1081058, Integrated DNA Technologies, Coralville, IA].[30]

<u>Electroporation of Cell lines:</u> K562, Jurkat, and Ramos cells were cultured in "R10": RPMI 1640 [Cat: 15-040-CV, Corning, Corning, NY] supplemented with 1x penicillin/streptomycin/glutamine [Cat: 10378016, ThermoFisher Scientific, Waltham, MA] and 10% fetal bovine serum [Omega Scientific, Tarzana, CA]. Each cell type required different electroporation conditions. K562 and Jurkat cell electroporations were carried out at 2e5 cells/condition. Cells were spun at 90g for 10 minutes, the supernatant was aspirated, and the cell pellet was resuspended in Lonza buffer (SF buffer for K562, SE buffer for Jurkat) at 20 uL/condition [Lonza, Basel, Switzerland]. The resuspended cells were transferred to pre-aliquoted tubes containing the payload to be delivered into the cells. For plasmid delivery, 500 ng of pX330 sgRNA/Cas9 expression plasmid was used with 3 ug of *BTK* donor template plasmid in the TOPO 2.1 plasmid backbone. With endonucleases delivered as ribonucleoprotein, 200 pmol of Cas9 were added to 9 ug of sgRNA and incubated at room temperature for 15 minutes before adding to cell

mixtures. The resulting reaction mix was electroporated using the Lonza 4D Nucleofector system (FF120 protocol for K562, CL120 for Jrukat) and allowed to rest for 10 minutes before plating in 480 uL of R10 medium (described above) per sample. Ramos cell line electroporations were carried out using the Invitrogen Neon electroporation system [Cat: MPK5000, ThermoFisher Scientific, Waltham, MA]. For the Ramos cells, 7.6 x 10^4 cells/condition were resuspended in Resuspension Buffer R [ThermoFisher Scientific, Waltham, MA]. 2.25 µg of gRNA and 50 pmol of Cas9 protein were pre-complexed for 10 minutes and added to the resuspension mixture. Samples were electroporated at 1350 V, 30 ms, 1 pulse before being plated in 250 mL of R10 containing the relevant AAV donor. One day post electroporation, the medium was changed and the transducing vector removed.

Electroporation/Transduction of Human CD34+ Cells: Granulocyte colony-stimulating factor (G-CSF) mobilized peripheral blood CD34+ cells [HemaCare, Van Nuys, CA] were pre-stimulated for 2 days in X-Vivo 15 medium [Cat: BEBP04-744Q, Lonza, Basel, Switzerland] supplemented with 1x penicillin/streptomycin/glutamine [Cat: 10378016, ThermoFisher Scientific, Waltham, MA], 50 ng/uL recombinant human (rh) stem cell factor [Cat: 300-07, PeproTech, Rocky Hill, NJ], 50 ng/uL rh Flt-3 ligand [Cat: 300-19, PeproTech, Rocky Hill, NJ], and 50 ng/uL of rh thrombopoietin [Cat: 300-18, PeproTech, Rocky Hill, NJ]. 2 x 10^5 cells were resuspended in 100 uL of BTXpress electroporation solution [Cat: 45-0802, BTX, Holliston, MA] and electroporated using the BTX ECM 830 Square Wave Electroporator set to 255 V, 5 ms, and 1 pulse [Cat: 45-0661, BTX, Holliston, MA]. 5 µg of gRNA and 5 µg Cas9 mRNA or 200 pmol of Cas9 protein were used, as specified by the delivery method in each figure. Following electroporation, cells were left to rest for 10 minutes in the supplemented X-Vivo 15 culture condition described above before a 24 hour transduction with AAV6 donor vector custom ordered from Virovek [Virovek,

Hayward, CA] or Vigene [Vigene Biosciences, Rockville, MD]. After transduction, cells were counted to determine the viability and fold expansion before being re-plated in outgrowth media: IMDM [Cat: 12440053, ThermoFisher Scientific, Waltham, MA], 10% fetal bovine serum, 1x penicillin/streptomycin/glutamine [Cat: 10378016, ThermoFisher Scientific, Waltham, MA], 5 ng/mL of rhIL-3, 10 ng/mL of rhIL-6, and 25 ng/mL of rhSCF [Cat: 200-03, 200-06, and 300-07, PeproTech, Rocky Hill, NJ].

Measuring Allelic Disruption: A 638 bp amplicon that encompasses BTK exon 2 as well as portions of introns 1 and 2 was amplified from genomic DNA taken from edited cells (primer sequences: "GCCATTTAACATCTAGAGCATTCC", "GTGGCTTCTTAGGACCTTTGAC"). The resulting amplicon was Sanger sequenced and analyzed via the Synthego ICE analysis web tool [Synthego Corporation, Menlo Park, CA].

GUIDE-seq: GUIDE-seq was performed in K562 cells evaluating the specificity of the sgRNA as described by Kuo *et al*.[34]

Integration Analysis: Genomic DNA was extracted from edited cells for integration site analysis using the Invitrogen Purelink Genomic DNA kit [Cat:K182002, Thermofisher Scientific, Waltham, MA] and quantified using the Nanodrop system [Cat:ND-2000, Thermofisher Scientific, Waltham, MA]. DNA samples were then analyzed by droplet digital PCR to measure integration rates. Two sets of primers were duplexed, each with their own fluorescent probe (FAM/HEX). One primer was complementary to a *BTK* gene sequence upstream from the left homology arm of the donor. The second primer bound to the codon-optimized *BTK* donor sequence, to allow for specific measurement of the integrated *BTK* transgene distict from the endogenous *BTK* gene. The FAM conjugated nucleotide probe with a quencher also bound to the minus strand DNA near the second primer. A reference primer/probe set was also delivered to recognize the UC462 region of

the X chromosome. 1 ul of EcoRV-HF endonuclease [Cat: R3195S, New England Biolabs, Ipswich, MA] was added to the reaction mix (an enzyme that does not disrupt the experimental or reference amplicons) to reduce background. Each sample was digested at 37°C for 1 hour before droplet generation with the BioRad QX200 Droplet Generator [Cat: 186-4002, BioRad, Hercules, CA]. The prepared samples were then assayed via the QX200 BioRad Droplet Reader on the "Absolute" measurement setting [Cat: 186-4003, BioRad, Hercules, CA].

ddPCR Primer/Probe	**Sequence**
BTK Fwd	5'-AGCAGTTAGTGTGTGTCCAGAAC-3'
BTK Rev	5'-CCTTATTAGTTCCCTTGGTTACAGA-3'
BTK Probe	5'-TCGAAGTCGTACTCGTAGTAGCTCAGCTTG-3'

BTK phosphorylation in Ramos Cells: 2 x 10^6 Ramos cells per condition were stimulated with 10 ug/mL of goat $F(ab')_2$ specific to human IgM [Cat:H15100, Thermofisher Scientific, Waltham, MA]. The cells were incubated at 37°C for 30 minutes before protein lysates were prepared using RIPA Lysis and Extraction Buffer [Cat:89900, Thermofisher Scientific, Waltham, MA] supplemented with 1x HALT Protease and Phosphatase Inhibitor Cocktail [Cat:78444, Thermofisher Scientific, Waltham, MA].

Immunoblot Analysis: For immunoblots probing for non-phosphorylated proteins, cells were lysed in Denaturing Cell Extraction Buffer [Cat:FNN0091, ThermoFisher Scientific, Waltham, MA] with added HALT protease inhibitor [Cat: 87786, ThermoFisher Scientific, Waltham, MA] at 1x concentration following manufacturer protocols. Cells were first spun at 500xg for 5 minutes at 4°C, the supernatant was aspirated, and the cell pellet was washed with chilled DPBS [Cat: 14190250, ThermoFisher Scientific, Waltham, MA]. The spin, aspirate, and wash steps were repeated. Then, chilled DCEB + HALT was added to the cell pellet at a concentration of 100 uL

per 2e6 cells and incubated at 4ºC for 30 minutes, vortexing every 10 minutes. The samples were then spun at 17,000xg for 15 minutes at 4ºC before transferring the supernatant to a fresh tube. Processing lysates for analysis of phosphorylation sites used the same procedure except lysis was performed using RIPA buffer with HALT Protease and Phosphatase Inhibitor was added [Cat: 78440, ThermoFisher Scientific, Waltham, MA]. Lysate concentrations were determined using the Pierce BCA protein assay [Cat: 23227, ThermoFisher Scientific, Waltham, MA] following the manufacturer's protocol. Samples were treated for SDS-PAGE with NuPAGE LDS Sample Buffer [Cat: NP0007, Thermofisher Scientific, Waltham, MA] and NuPAGE Sample Reducing Agent [Cat: NP0009, Thermofisher Scientific, Waltham, MA], each to a 1x concentration. An equivalent amount of "BTK intact" (either gene edited or wildtype BTK) was loaded for each sample, calculated using integration rates determined by droplet digital PCR. Lysate previously verified to be negative (untreated BTK-deficient K562 lysate) was added to samples control for total protein added and allow for a loading control—lest samples with high editing rates would have far lower actin signals due to a higher concentration of BTK intact cells, which would make it impossible to identify whether differences in signal intensity were due to real phenomenon or loading errors. Control samples such as those from endonuclease only or donor only treated cells were normalized only for total lysate concentration, as they did not have any intact BTK for experiments using BTK deficient K562 cells. These samples were heated to 95ºC for 10 minutes before loading into a 4-12% Bis-Tris Gel [Cat: NP0322BOX, Thermofisher Scientific, Waltham, MA]. Samples were transferred to Polyvinylidene fluoride (PVDF) transfer membrane [Cat:88520, ThermoFisher Scientific, Waltham, MA]. The membrane was blocked with 5% milk in 1x Pierce TBS Tween 20 Buffer [Cat:28360, ThermoFisher Scientific, Waltham, MA] for 1 hour. Membranes were split into discrete portions to probe for different proteins and incubated in a 1:1000 dilution of primary

antibody for BTK [Cat:56044S, Cell Signaling Technology, Danvers, MA], phosphorylated BTK [Cat:5082S, Cell Signaling Technology, Danvers, MA], HA [Cat:NBP2-43714, Novus, Irvine, CA], or Actin [Cat:3700S, Cell Signaling Technology, Danvers, MA] shaking overnight at 4°C. Membranes were washed three times with TBS Tween and incubated in 5% milk with secondary antibody for 90 minutes [Cat: 554002, BD Biosciences, San Diego, CA][Cat:A-21236, Thermofisher Scientific, Waltham, MA]. Secondary antibodies were washed off three times. Blots stained with horse radish peroxidase conjugated antibodies were treated with Pierce ECL Plus Western Blotting Substrate [Cat:32132, Thermofisher Scientific, Waltham, MA]. All blots were imaged via phosphoimager and quantified via densitometry.

Quantifying RNA Abundance via reverse-transcription digital droplet PCR

Total RNA was extracted from treated cells, 1e6 cells per sample, with RNeasy Mini Plus Kit [Cat: 74136, Qiagen, Germantown, MD] following the manufacturer's protocol. Purified RNA was transferred to 0.2 mL PCR tubes with 0.1 volume 10X Turbo DNase buffer [Cat: AM1907, ThermoFisher Scientific, Waltham, MA] and incubated at 37°C for 30 minutes. We used 0.1 volume DNase Inactivation Reagent and incubated at room temperature for 5 minutes then centrifugated samples at 10,000 x g for 1.5 minutes taking the supernatant. cDNA was synthesized from purified DNase-treated RNA with SuperScript III First-Strand Synthesis System for RT-PCR (Invitrogen: Cat# 18080-051) following the manufacturer's protocol. To analyze mRNA expression of the BTK transgene donor, primers were designed to be specific to the codon optimized construct and the WT sequence. IPO8 was chosen as the reference gene. The cDNA were diluted 25-fold and gene expression was determined by reverse transcription digital droplet PCR (RT-ddPCR).

RT-ddPCR Primer	Sequence

Co-Opt BTK-Fwd	5’-AAGTACACCGTGTCCGTGTT-3’
Co-Opt BTK-Rev	5’-ATTGTGCTGAACAGGTGCTT-3’
WT BTK-Fwd	5’-CAATGGCTGCCTCCTGAA-3’
WT BTK-Rev	5’-TGCCAGGTCTCGGTGAA-3’
IPO8-Fwd	5’-TTTGAATACTTTACAGACATGATGC-3’
IPO8-Rev	5’-AACGAAGAGTGGAATGCACTG-3’

References

1. Futatani, T, Miyawaki, T, Tsukada, S, Hashimoto, S, Kunikata, T, Arai, S, *et al.* (1998). Deficient expression of Bruton’s tyrosine kinase in monocytes from X-linked agammaglobulinemia as evaluated by a flow cytometric analysis and its clinical application to carrier detection. *Blood* **91**: 595–602.

2. Vetrie, D, Vorechovsky, I, Sideras, P, Holland, J, Davies, A, Flinter, F, *et al.* (1993). The gene involved in X-linked agammaglobulinaemia is a member of the src family of protein-tyrosine kinases. *Nature* **361**: 226–233.

3. Siegal, FP, Pernis, B and Kunkel, HG (1971). Lymphocytes in human immunodeficiency states: a study of membrane-associated immunoglobulins. *Eur. J. Immunol.* **1**: 482–486.

4. de Weers, M, Brouns, GS, Hinshelwood, S, Kinnon, C, Schuurman, RK, Hendriks, RW, *et al.* (1994). B-cell antigen receptor stimulation activates the human Bruton’s tyrosine kinase, which is deficient in X-linked agammaglobulinemia. *J. Biol. Chem.* **269**: 23857–23860.

5. BRUTON, OC (1952). AGAMMAGLOBULINEMIA. *Pediatrics* **9**: 722 LP – 728.

6. Hermaszewski, RA and Webster, AD (1993). Primary hypogammaglobulinaemia: a survey of clinical manifestations and complications. *Q. J. Med.* **86**: 31–42.

7. Ng, YY, Baert, MRM, Pike-Overzet, K, Rodijk, M, Brugman, MH, Schambach, a, *et al.* (2010). Correction of B-cell development in Btk-deficient mice using lentiviral vectors with codon-optimized human BTK. *Leuk. Off. J. Leuk. Soc. Am. Leuk. Res. Fund, U.K* **24**: 1617–1630.

8. Sather, BD, Ryu, BY, Stirling, B V, Garibov, M, Kerns, HM, Humblet-Baron, S, *et al.* (2011). Development of B-lineage predominant lentiviral vectors for use in genetic therapies for B cell disorders. *Mol. Ther.* **19**: 515–525.

9. Themis, M, Waddington, SN, Schmidt, M, von Kalle, C, Wang, Y, Al-Allaf, F, *et al.* (2005). Oncogenesis following delivery of a nonprimate lentiviral gene therapy vector to fetal and neonatal mice. *Mol. Ther.* **12**: 763–771.

10. Cavazza, A, Moiani, A and Mavilio, F (2013). Mechanisms of retroviral integration and mutagenesis. *Hum. Gene Ther.*doi:10.1089/hum.2012.203.

11. Satterthwaite, AB, Cheroutre, H, Khan, WN, Sideras, P and Witte, ON (1997). Btk dosage determines sensitivity to B cell antigen receptor cross-linking. *Proc. Natl. Acad. Sci. U. S. A.* **94**: 13152–13157.

12. Kokabee, L, Wang, X, Sevinsky, CJ, Wang, WLW, Cheu, L, Chittur, S V, *et al.* (2015). Bruton's tyrosine kinase is a potential therapeutic target in prostate cancer. *Cancer Biol. Ther.* **16**: 1604–15.

13. Honigberg, LA, Smith, AM, Sirisawad, M, Verner, E, Loury, D, Chang, B, *et al.* (2010).

The Bruton tyrosine kinase inhibitor PCI-32765 blocks B-cell activation and is efficacious in models of autoimmune disease and B-cell malignancy. *Proc. Natl. Acad. Sci. U. S. A.*doi:10.1073/pnas.1004594107.

14. Cong, L, Ran, FA, Cox, D, Lin, S, Barretto, R, Habib, N, *et al.* (2013). Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* **339**: 819–823.

15. Hoban, MD, Cost, GJ, Mendel, MC, Romero, Z, Kaufman, ML, Joglekar, a. V., *et al.* (2015). Correction of the sickle-cell disease mutation in human hematopoietic stem/progenitor cells. *Blood*doi:10.1182/blood-2014-12-615948.

16. Genovese, P, Schiroli, G, Escobar, G, Di Tomaso, T, Firrito, C, Calabria, A, *et al.* (2014). Targeted genome editing in human repopulating haematopoietic stem cells. *Nature* **510**: 235–240.

17. Wei, W, Kim, JM, Medina, D, Lakatta, EG and Lin, L (2014). GeneOptimizer program-assisted cDNA reengineering enhances sRAGE autologous expression in Chinese hamster ovary cells. *Protein Expr. Purif.*doi:10.1016/j.pep.2013.12.006.

18. Tsai, SQ, Zheng, Z, Nguyen, NT, Liebers, M, Topkar, V V, Thapar, V, *et al.* (2014). GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat. Biotechnol.* **33**: 187–198.

19. Xu, DH, Wang, XY, Jia, YL, Wang, TY, Tian, ZW, Feng, X, *et al.* (2018). SV40 intron, a potent strong intron element that effectively increases transgene expression in transfected Chinese hamster ovary cells. *J. Cell. Mol. Med.*doi:10.1111/jcmm.13504.

20. Choi, T, Huang, M, Gorman, C and Jaenisch, R (1991). A generic intron increases gene

expression in transgenic mice. *Mol. Cell. Biol.*doi:10.1128/mcb.11.6.3070.

21. Valencia, P, Dias, AP and Reed, R (2008). Splicing promotes rapid and efficient mRNA export in mammalian cells. *Proc. Natl. Acad. Sci. U. S. A.*doi:10.1073/pnas.0800250105.

22. Dye, MJ and Proudfoot, NJ (1999). Terminal exon definition occurs cotranscriptionally and promotes termination of RNA polymerase II. *Mol. Cell*doi:10.1016/S1097-2765(00)80464-5.

23. Grieger, JC and Samulski, RJ (2005). Packaging Capacity of Adeno-Associated Virus Serotypes: Impact of Larger Genomes on Infectivity and Postentry Steps. *J. Virol.*doi:10.1128/jvi.79.15.9933-9944.2005.

24. De Ravin, SS, Reik, A, Liu, P-Q, Li, L, Wu, X, Su, L, *et al.* (2016). Targeted gene addition in human CD34(+) hematopoietic cells for correction of X-linked chronic granulomatous disease. *Nat. Biotechnol.* **34**: 424–429.

25. Zufferey, R, Donello, JE, Trono, D and Hope, TJ (1999). Woodchuck Hepatitis Virus Posttranscriptional Regulatory Element Enhances Expression of Transgenes Delivered by Retroviral Vectors. *J. Virol.*doi:10.1128/jvi.73.4.2886-2892.1999.

26. Tomlinson, MG, Kane, LP, Su, J, Kadlecek, TA, Mollenauer, MN and Weiss, A (2004). Expression and Function of Tec, Itk, and Btk in Lymphocytes: Evidence for a Unique Role for Tec. *Mol. Cell. Biol.*doi:10.1128/mcb.24.6.2455-2466.2004.

27. Hendel, A, Bak, RO, Clark, JT, Kennedy, AB, Ryan, DE, Roy, S, *et al.* (2015). Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells. *Nat Biotech* **advance on**: 985–989.

28. Slaymaker, IM, Gao, L, Zetsche, B, Scott, DA, Yan, WX and Zhang, F (2016). Rationally engineered Cas9 nucleases with improved specificity. *Science (80-.).* **351**: 84–88.

29. Kleinstiver, BP, Pattanayak, V, Prew, MS, Tsai, SQ, Nguyen, NT, Zheng, Z, *et al.* (2016). High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* **529**: 490–495.

30. Vakulskas, CA, Dever, DP, Rettig, GR, Turk, R, Jacobi, AM, Collingwood, MA, *et al.* (2018). A high-fidelity Cas9 mutant delivered as a ribonucleoprotein complex enables efficient gene editing in human hematopoietic stem and progenitor cells. *Nat. Med.*doi:10.1038/s41591-018-0137-0.

31. Rohrer, J and Conley, ME (1999). Correction of X-linked immunodeficient mice by competitive reconstitution with limiting numbers of normal bone marrow cells. *Blood*doi:10.1182/blood.v94.10.3358.422k04_3358_3365.

32. Hendel, A, Bak, RO, Clark, JT, Kennedy, AB, Ryan, DE, Roy, S, *et al.* (2015). Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells. *Nat. Biotechnol.*doi:10.1038/nbt.3290.

33. Dewitt, M and Wong, J (2015). In vitro transcription of guide RNAs. *protocols.io*.

34. Kuo, CY, Long, JD, Campo-Fernandez, B, de Oliveira, S, Cooper, AR, Romero, Z, *et al.* (2018). Site-Specific Gene Editing of Human Hematopoietic Stem Cells for X-Linked Hyper-IgM Syndrome. *Cell Rep.* **23**: 2606–2616.

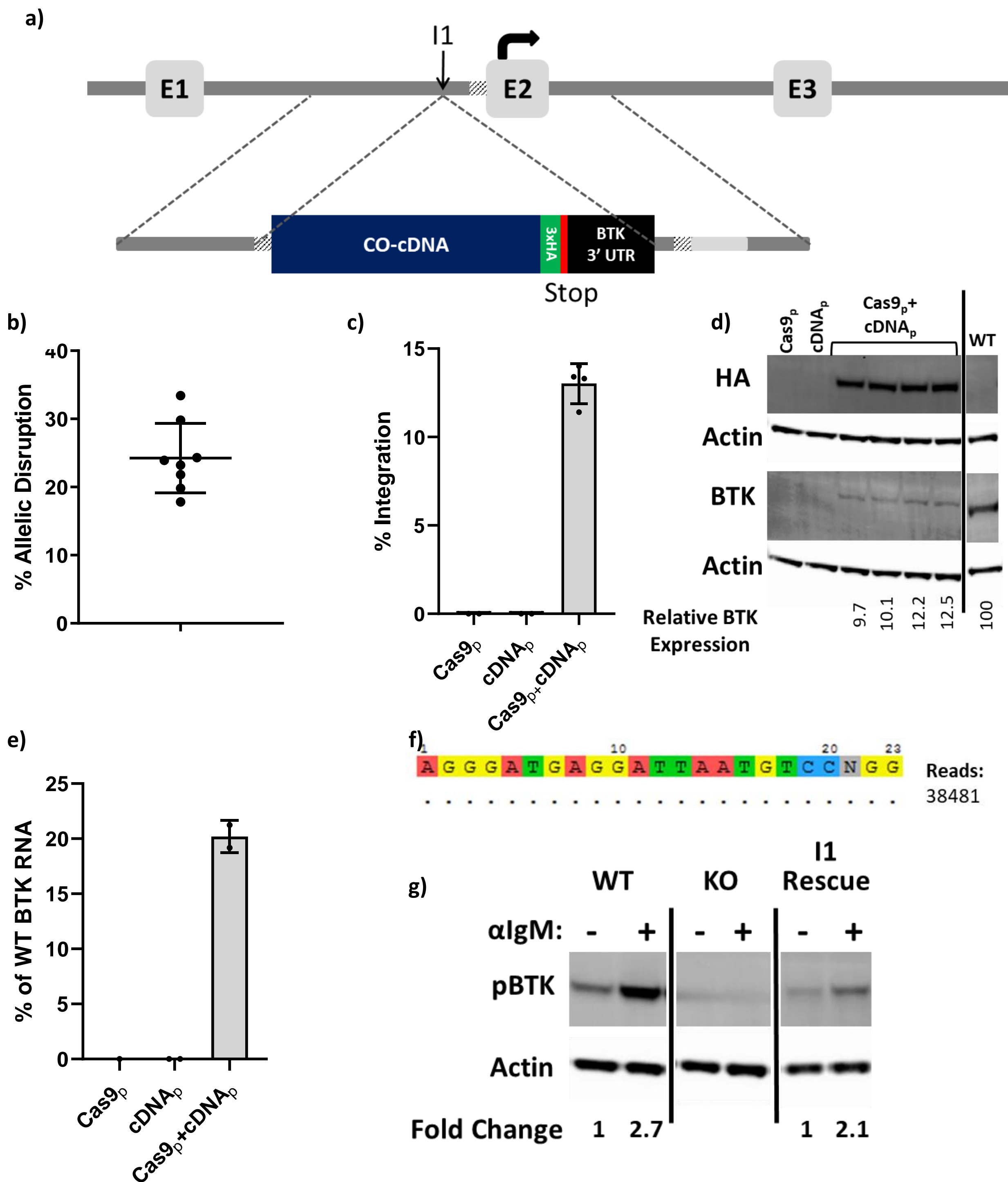
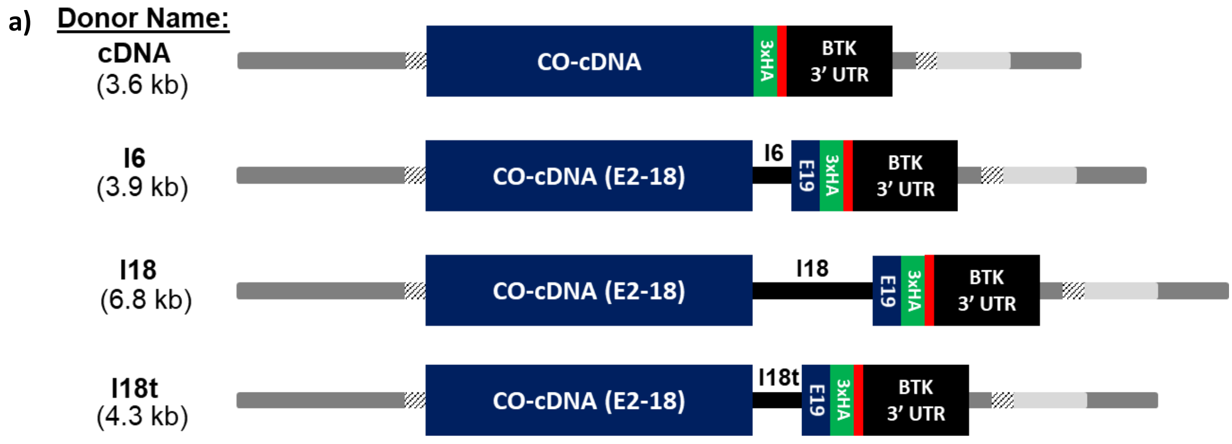
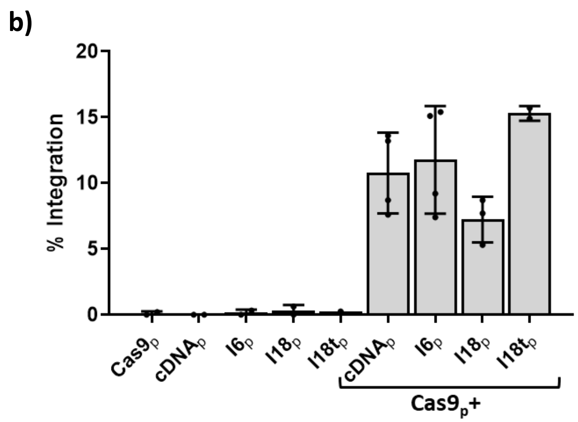


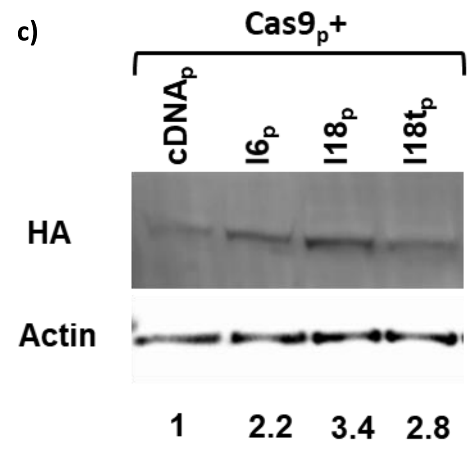
Figure 1: Targeted integration of *BTK* cDNA into intron 1 of the gene leads to sub-physiological levels of mRNA and protein expression

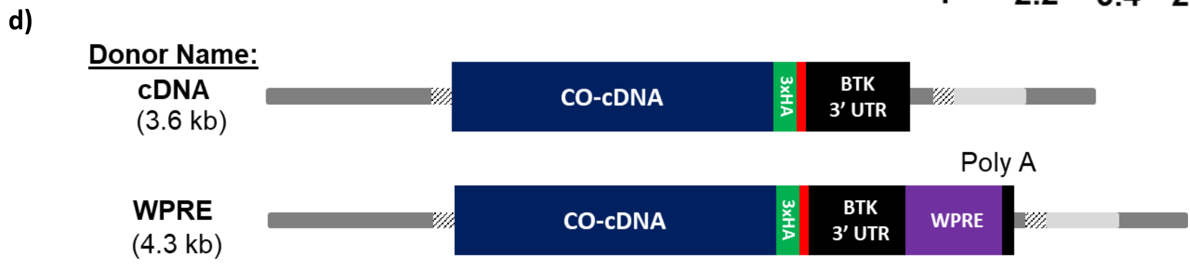
a) Diagram of the BTK intron 1 editing scheme and donor. The donor template consists of the codon-optimized (CO) cDNA sequence, 3 C-terminal hemagglutinin tags (3xHA), a stop codon, and the BTK 3' untranslated region (UTR), all flanked by 500 bp homology arms that correspond

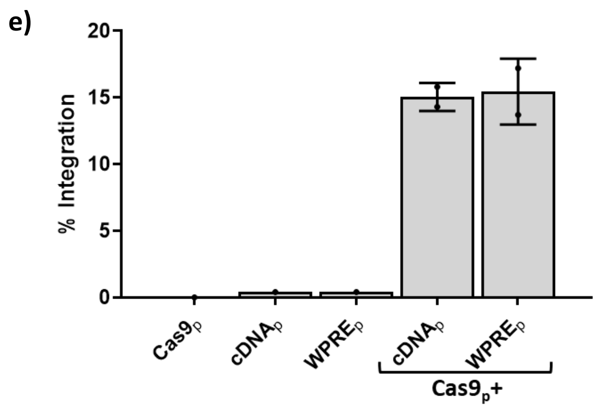
to the DNA sequence on either side of the cut site. **b)** Single guide RNA (sgRNA) targeting BTK intron 1 and Cas9 were delivered via plasmid electroporation to BTK-deficient K562 erythroleukemia cells. Genomic DNA (gDNA) from the treated cells was harvested and analyzed for allelic disruption at the intron 1 site. **c)** BTK-deficient K562s were elecroporated with either the sgRNA/Cas9 expression plasmid ($Cas9_p$), the donor template plasmid ($cDNA_p$), or both together. Targeted integration of the donor template was measured using droplet digital PCR (ddPCR) to quantify the frequency of 5' integration junction in the resultant gDNA. **d)** Immunoblot analysis of BTK deficient K562 cells treated with $Cas9_p$, $cDNA_p$, or both and probed for actin and the 3xHA tag or BTK protein. Samples were normalized based on their respective integration rates for a comparison of expression per integrated copy. Wildtype (WT) lysate was diluted with BTK-deficient cell lysate to have an equivalent fraction of intact BTK as the experimental samples. **e)** RNA expression from treated, BTK-deficient K562 cells was analyzed via reverse transcription ddPCR and normalize to the IPO8 housekeeping gene and compared to wildtype BTK RNA levels. **f)** Genome-wide unbiased identification of double stranded breaks enabled by high throughput sequencing (GUIDE-seq) analysis of the BTK intron 1 sgRNA/Cas9 plasmid in K562 cells. The top line represents the inputed sgRNA target sequence and protospacer adjacent motif in BTK intron 1. The dotted line beneath it represents the only detected cutting event—each dot corresponds to a perfect match to the target sequence, while base pair differences would be illustrated with nucleotide symbols differring from the target sequence. Only on target cleavage was detected. **g)** Two Ramos B lymphocyte cell sublines were engineered using gene editing: one with a clonal, homozygous deletion in BTK exon 15 (KO), and another with the same deletion as well as a successfully integrated BTK cDNA donor template in the intron 1 locus (I1 Rescue). Each of these lines as well as the parental (WT) line were stimulated with a human IgM crosslinking antibody fragment to induce B cell receptor signaling. Phosphorylation of the BTK protein was measured via immunoblot and quantified via densitometry.


a)
Donor Name:
cDNA
(3.6 kb)
CO-cDNA
3xHA
BTK
3' UTR
I6
(3.9 kb)
CO-cDNA (E2-18)
I6
E19
I18
(6.8 kb)
I18
I18t
(4.3 kb)
I18t


b)
% Integration
Cas9p
cDNAp
I6p
I18p
I18tp
Cas9p+


c)
Cas9p+
cDNAp
I6p
I18p
I18tp
HA
Actin
1
2.2
3.4
2.8


d)
Donor Name:
cDNA
(3.6 kb)
CO-cDNA
3xHA
BTK
3' UTR
WPRE
(4.3 kb)
Poly A
WPRE


e)
% Integration
Cas9p
cDNAp
WPREp
Cas9p+

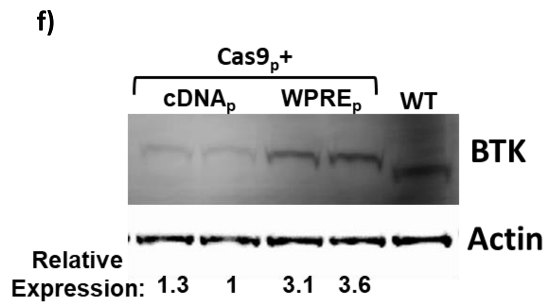
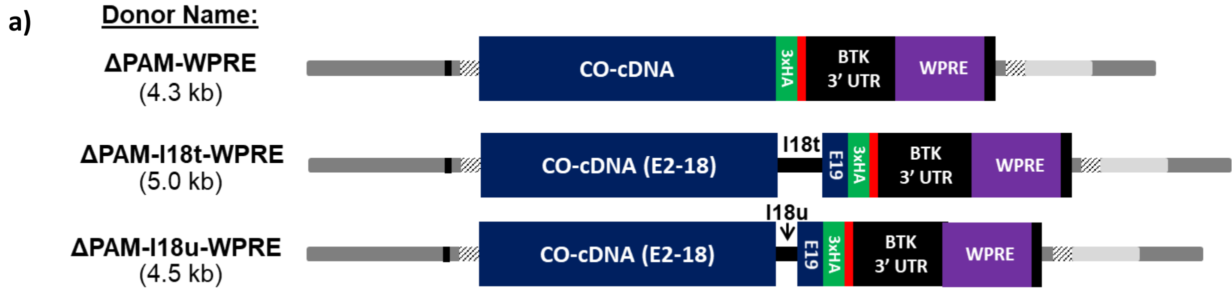
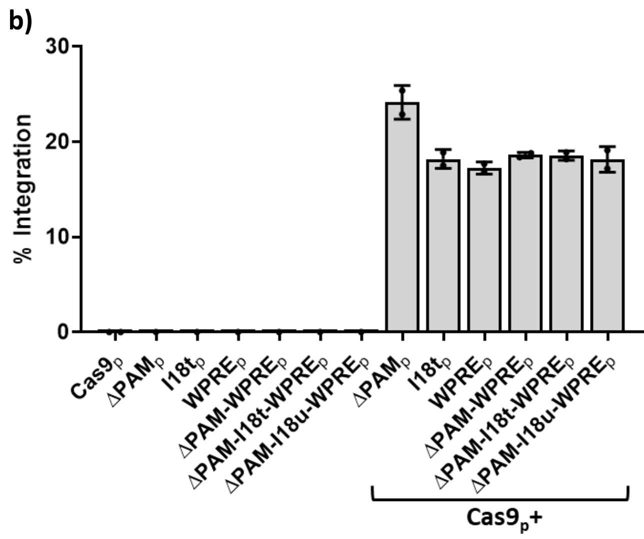
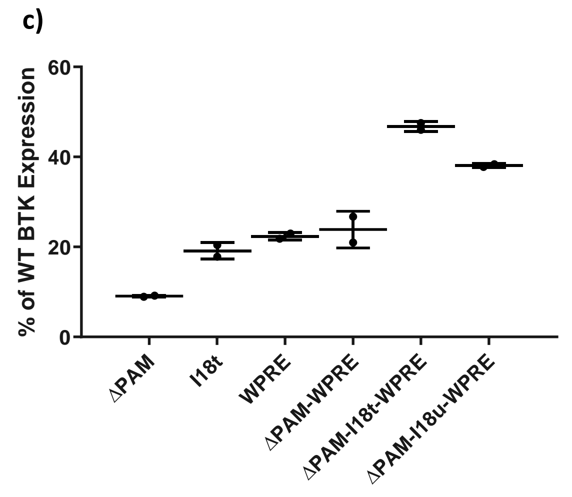

f)
Cas9p+
cDNAp
WPREp
WT
BTK
Actin
Relative
Expression: 1.3 1 3.1 3.6

Figure 2: Addition of a terminal intron or the woodchuck hepatitis virus post-transcriptional regulatory element (WPRE) to donor templates increases transgenic BTK expression

a) Schematics of the cDNA, Intron 6 (I6), Intron 18 (I18), and Intron 18 truncated (I18t) homologous donor templates designed to integrate into BTK intron 1. All the intron containing donors share the same base as the cDNA donor with only the added intronic sequence between exons 18 and 19 of the cDNA sequence. **b)** BTK-deficient K562 cells were electroporated with the sgRNA/Cas9 plasmid ($Cas9_p$), one of the four donor templates, or $Cas9_p$ and donor template. Targeted integration rates of each donor template into the BTK intron 1 locus were measured by ddPCR. **c)** Immunoblot analysis of the treated, BTK deficient K562 cells probed for actin and the 3xHA tag on each donor construct. Samples were normalized based on their respective integration rates for a comparison of expression per integrated copy. **d)** Schematic of the WPRE containing donor (WPRE) compared to the base cDNA donor. The WPRE was added at the 3' end of the BTK 3' untranslated region, just before the poly adenylation (Poly A) signal. **e)** Targeted integration of the $cDNA_p$ donor compared to the $WPRE_p$ donor in BTK-deficient K562 cells as measured by ddPCR. **f)** Immunoblot analysis of BTK deficient K562 cells treated with $Cas9_p$, one of the donor plasmids, or both and probed for actin and BTK protein. Samples were normalized based on their respective integration rates for a comparison of expression per integrated copy. Wildtype (WT) lysate was diluted with BTK-deficient cell lysate to have an equivalent fraction of intact BTK as the experimental samples.


a)
Donor Name:
ΔPAM-WPRE
(4.3 kb)
ΔPAM-I18t-WPRE
(5.0 kb)
ΔPAM-I18u-WPRE
(4.5 kb)
CO-cDNA
CO-cDNA (E2-18)
I18t
I18u
E19
3xHA
BTK
3' UTR
WPRE


b)
% Integration
Cas9p+


c)
% of WT BTK Expression
ΔPAM
I18t
WPRE
ΔPAM-WPRE
ΔPAM-I18t-WPRE
ΔPAM-I18u-WPRE

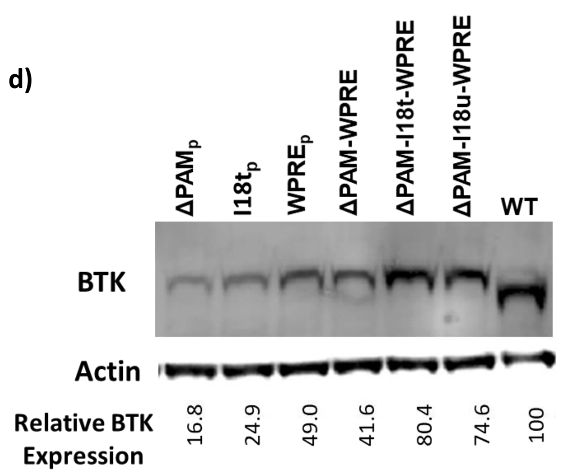

d)
WT
BTK
Actin
Relative BTK Expression
16.8
24.9
49.0
41.6
80.4
74.6
100

Figure 3: Combining successful donor modifications leads to an additive effect on transgene expression

a) Schematics of BTK intron 1 donors containing three combinations of modifications for improved integration junction fidelity or expression. The ΔPAM modification is represented by a black bar in the 5' homology arm. The new combinations were ΔPAM-WPRE, ΔPAM-I18t-WPRE, and ΔPAM-I18u-WPRE where I18u is a further truncated variant of I18t. **b)** BTK-deficient K562 cells were electroporated with BTK intron 1 targeting sgRNA/Cas9 plasmid ($Cas9_p$), one of six homologous donor plasmids, or $Cas9_p$ and donor template together. Targeted integration rates of each donor template into the BTK intron 1 locus were measured by ddPCR. **c)** RNA expression from treated, BTK-deficient K562 cells was analyzed via reverse transcription ddPCR and normalize to the IPO8 housekeeping gene and compared to wildtype BTK RNA levels. **d)** Immunoblot analysis of BTK deficient K562 cells treated with $Cas9_p$, one of the donor plasmids, or both. Cell lysates were probed for actin and BTK protein. Samples were normalized based on their respective integration rates for a comparison of expression per integrated copy. Wildtype (WT) lysate was diluted with BTK-deficient cell lysate to have an equivalent fraction of intact BTK as the experimental samples.

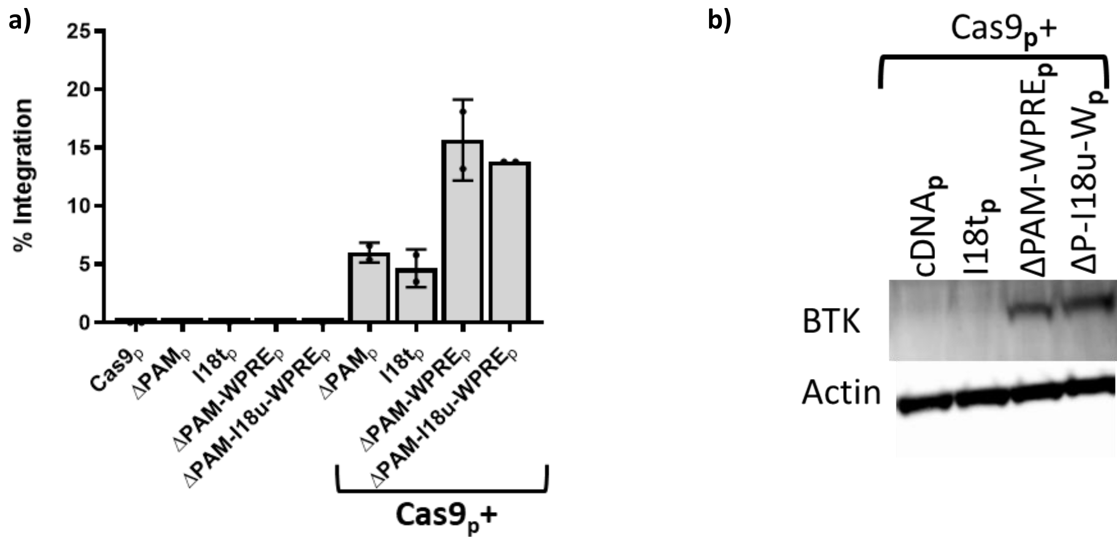


Figure 4: Integration of WPRE containing donors led to elevated BTK expression in non-BTK expressing Jurkat T lymphocytes

a) Jurkat cells were electroporated with BTK intron 1 targeting sgRNA/Cas9 plasmid ($Cas9_p$), one of four previously described homologous donor plasmids, or $Cas9_p$ and donor plasmid together. Targeted integration rates of each donor template into the BTK intron 1 locus were measured by ddPCR. **b)** Immunoblot analysis of Jurkat T lymphocytes treated with $Cas9_p$ and each of the donor plasmids. Cell lysates were probed for actin and BTK protein. Samples were normalized based on their respective integration rates for a comparison of expression per integrated copy.

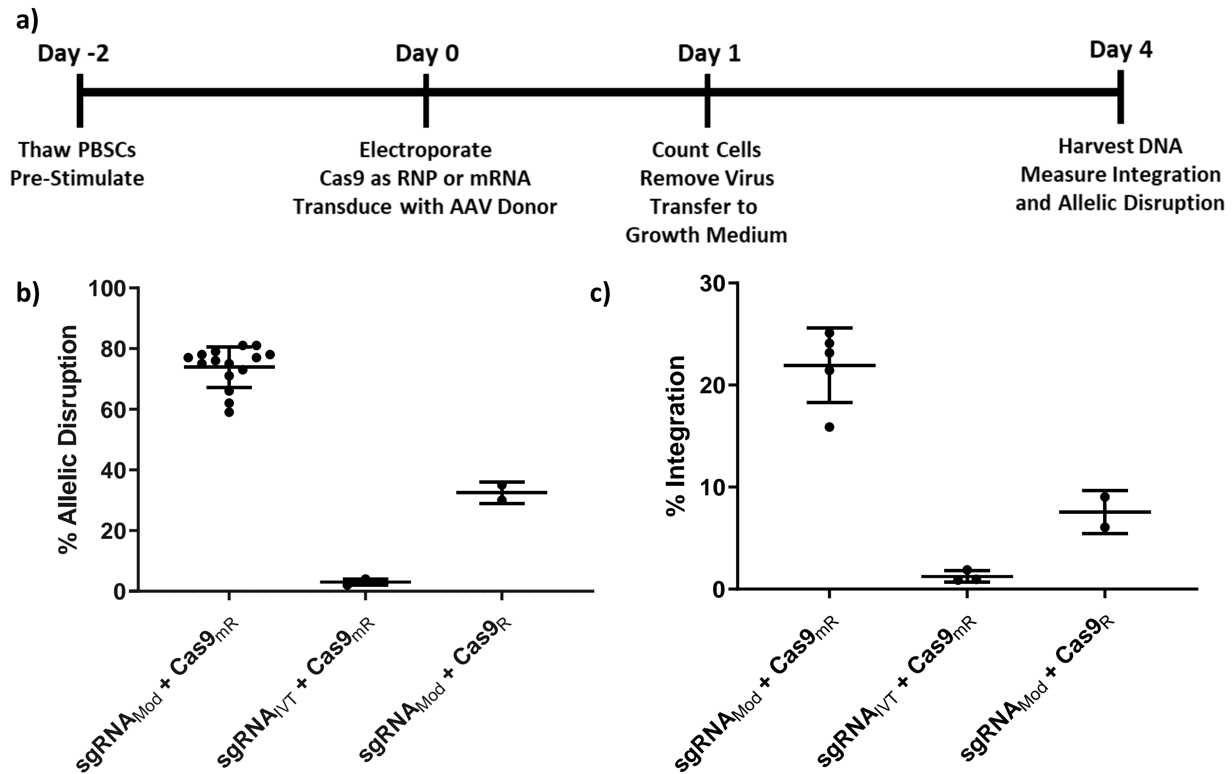
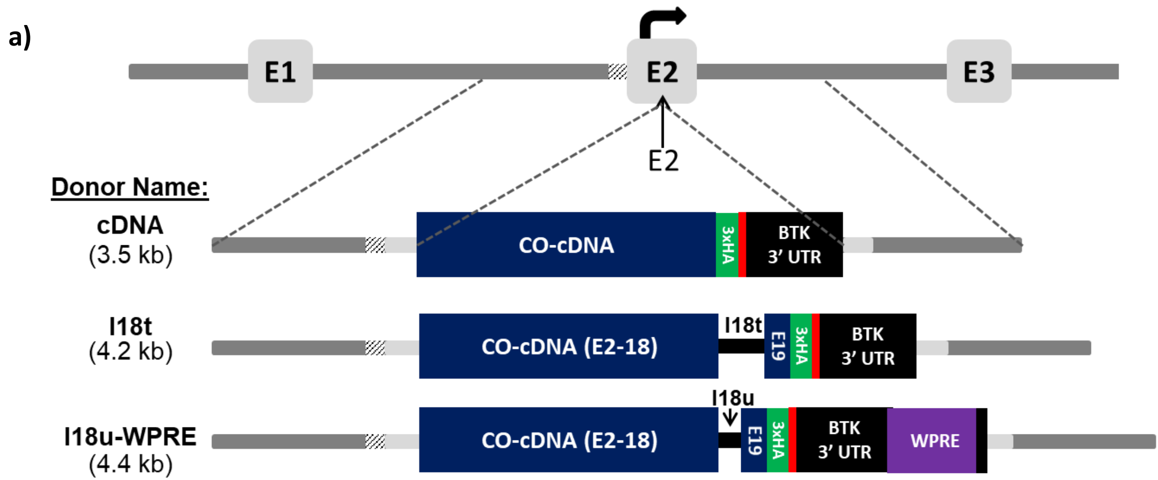
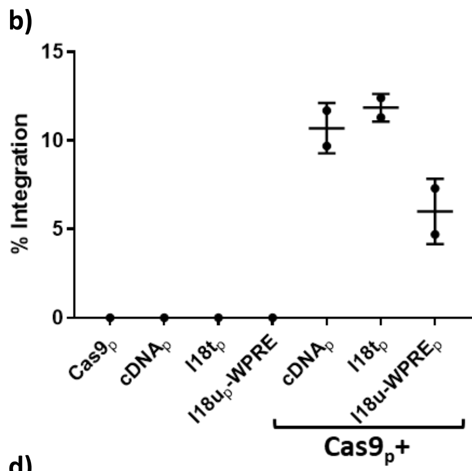


Figure 5: Chemical modification of sgRNA and Cas9 delivered as mRNA produced efficient allelic disruption and integration rates in human CD34+ mobilized peripheral blood hematopoietic stem and progenitor cells

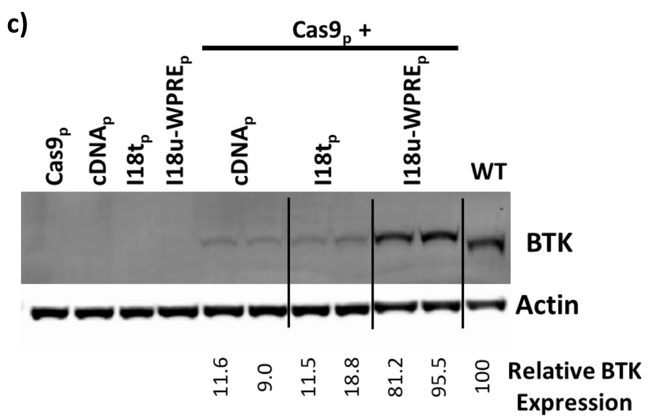
a) Timeline of electroporation of peripheral blood stem cells (PBSCs) from thaw to genomic DNA harvest. **b)** A comparison of allelic disruption rates in human PBSCs when the BTK intron 1 targeting single guide RNA (sgRNA)/Cas9 was delivered with different combinations of starting reagents. sgRNA was delivered as chemically modified ($sgRNA_{Mod}$) with 2'-O-methyl and 3' phosphorothioate inter-nucleotide linkages, or *in vitro* transcribed ($sgRNA_{IVT}$) with no modifications. Cas9 was delivered as either mRNA ($Cas9_{mR}$) or precomplexed with the sgRNA as ribonucleoprotein ($Cas9_{R}$). **c)** The same endonuclease reagent permutations were compared for the amount of targeted integration found at the intron 1 site with the addition of a 24 hour transduction with adeno-associated viral vector serotype 6 (AAV6) homologous donor template.



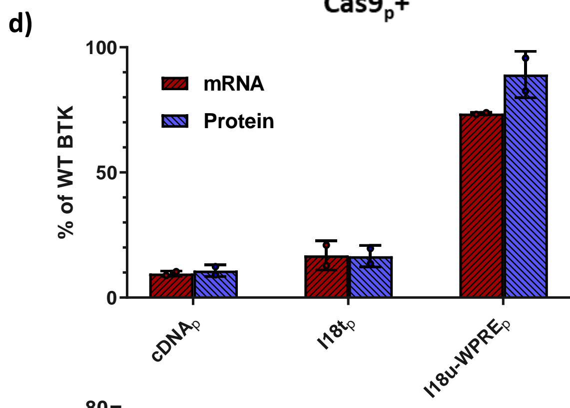
a)
E1
E2
E3
E2
Donor Name:
cDNA
(3.5 kb)
CO-cDNA
3xHA
BTK 3' UTR
I18t
(4.2 kb)
CO-cDNA (E2-18)
I18t
E19
3xHA
BTK 3' UTR
I18u-WPRE
(4.4 kb)
CO-cDNA (E2-18)
I18u
E19
3xHA
BTK 3' UTR
WPRE



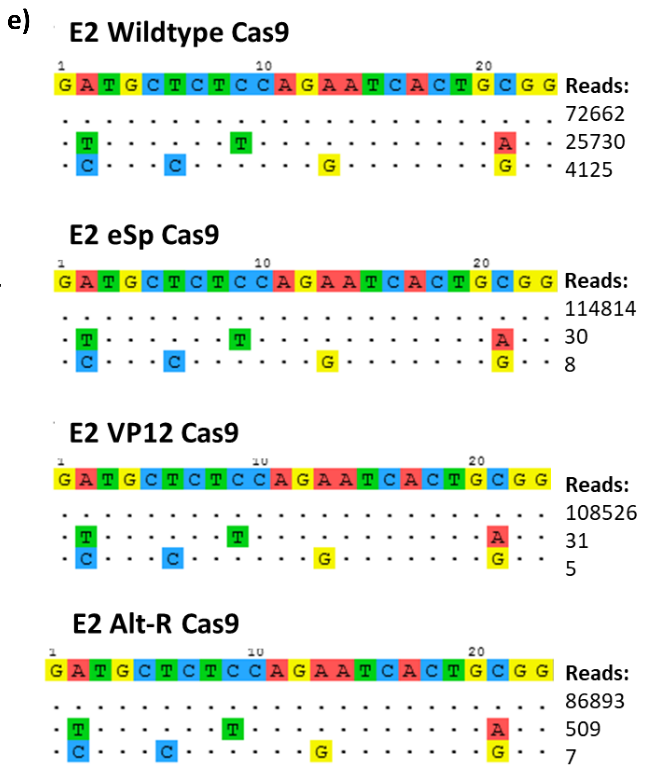
b)
% Integration
15
10
5
0
Cas9p
cDNAp
I18tp
I18up-WPRE
cDNAp
I18tp
I18u-WPREp
Cas9p+



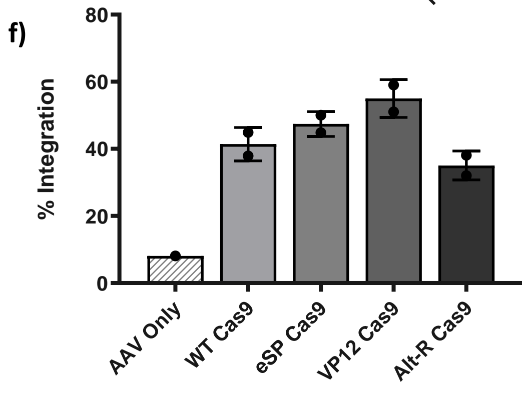
c)
Cas9p +
Cas9p
cDNAp
I18tp
I18u-WPREp
cDNAp
I18tp
I18u-WPREp
WT
BTK
Actin
11.6
9.0
11.5
18.8
81.2
95.5
100
Relative BTK Expression



d)
% of WT BTK
100
50
0
mRNA
Protein
cDNAp
I18tp
I18u-WPREp



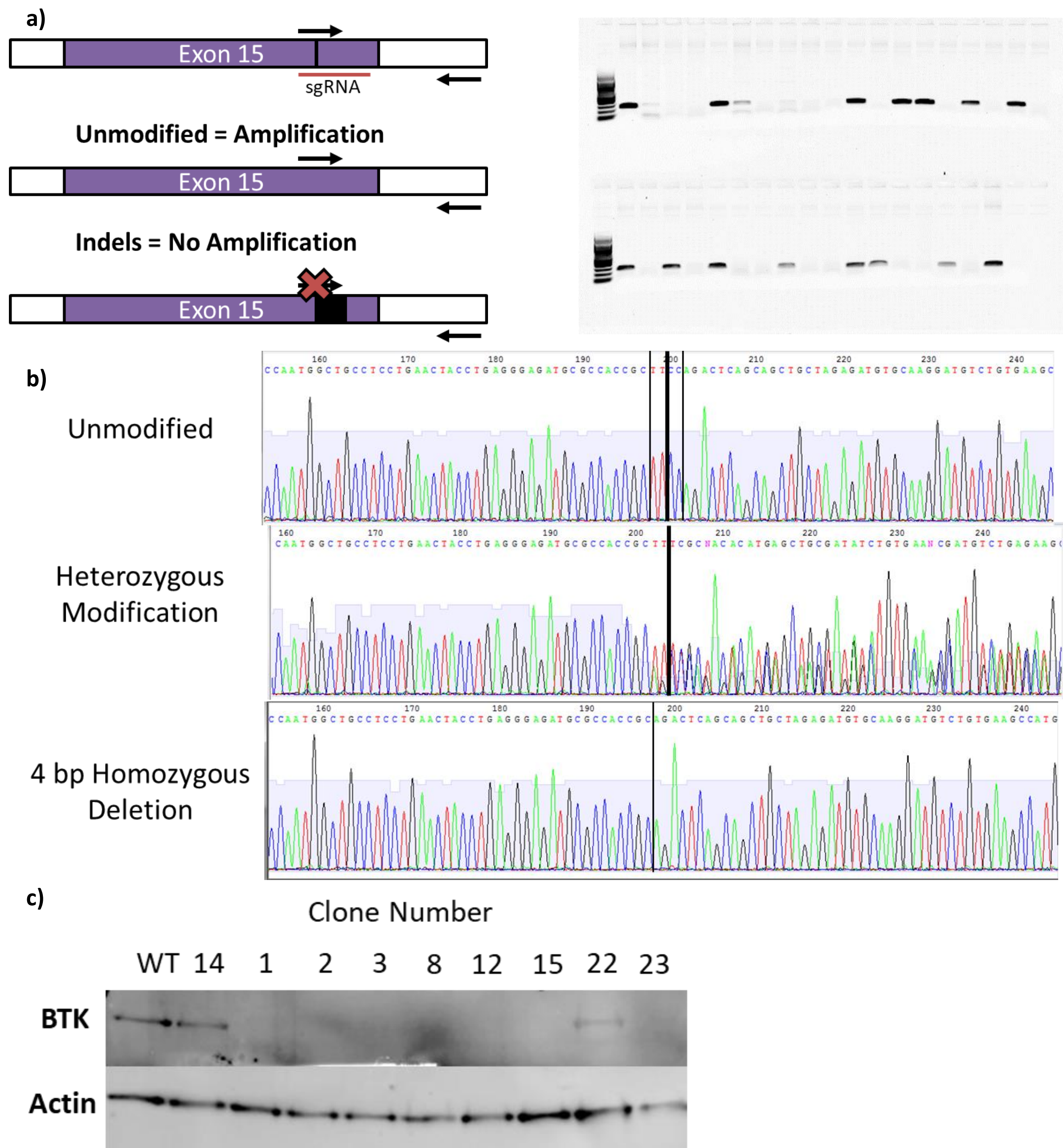
e)
E2 Wildtype Cas9
GATGCTCTCCAGAATCACTGCGG
Reads:
72662
25730
4125
E2 eSp Cas9
GATGCTCTCCAGAATCACTGCGG
Reads:
114814
30
8
E2 VP12 Cas9
GATGCTCTCCAGAATCACTGCGG
Reads:
108526
31
5
E2 Alt-R Cas9
GATGCTCTCCAGAATCACTGCGG
Reads:
86893
509
7



f)
% Integration
80
60
40
20
0
AAV Only
WT Cas9
eSP Cas9
VP12 Cas9
Alt-R Cas9

Figure 6: Targeted integration of a corrective cDNA donor into the BTK exon 2 target site led to efficient integration and expression of the transgene

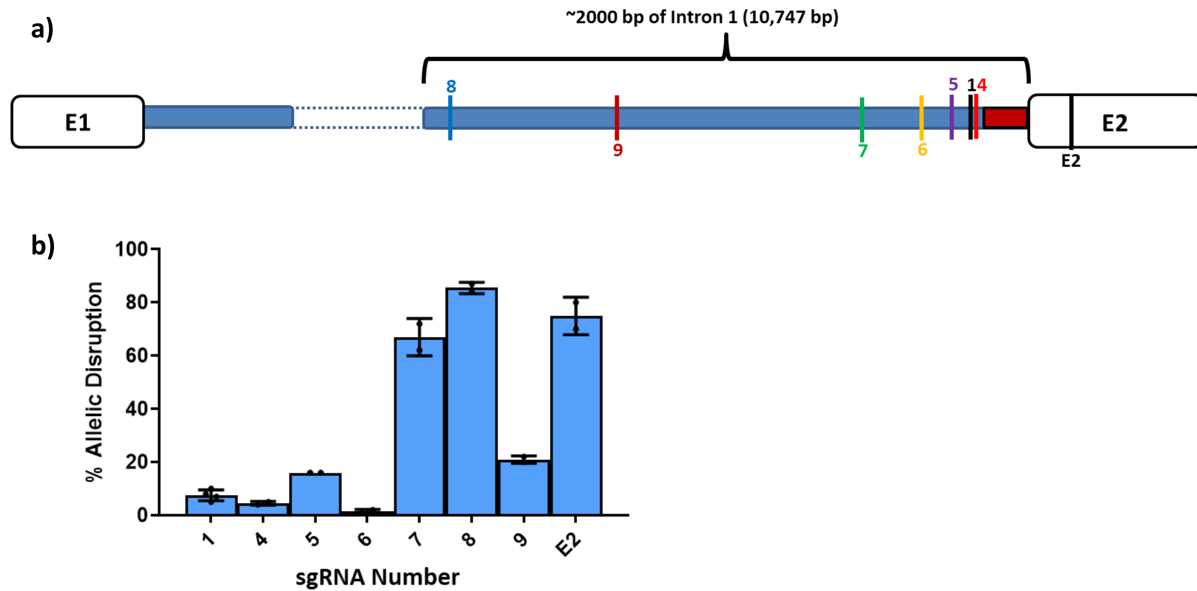
a) Schematic of three donor homologous templates designed to integrate into BTK exon 2. The three donor variants featured optimizations originally assessed for efficacy at the intron 1 target site: a base cDNA donor, a donor with I18t, and a donor with I18u-WPRE. **b)** BTK-deficient K562 cells were electroporated with the sgRNA/Cas9 plasmid ("p" subscript denoting plasmid), one of the three BTK exon 2 specific donor templates, or $Cas9_p$ together with a donor template. Targeted integration rates of each donor template into the BTK exon 2 locus were measured by ddPCR. **c)** Immunoblot analysis of the treated, BTK deficient K562 cells probed for actin and the BTK protein. Samples were normalized based on their respective integration rates for a comparison of expression per integrated copy. Wildtype (WT) lysate was diluted with BTK-deficient cell lysate to have an equivalent fraction of intact BTK as the experimental samples. **d)** Transgenic BTK mRNA (red) and protein (blue) levels in treated cells compared to the levels found in wildtype cells. mRNA concentrations were measured by reverse transcription ddPCR while protein levels were obtained via densitometry of a BTK immunoblot. **e)** Genome-wide unbiased identification of double stranded breaks enabled by high throughput sequencing (GUIDE-seq) analysis of the BTK exon 2 sgRNA/Cas9 delivered as ribonucleoprotein in K562 cells. The top line represents the inputed sgRNA target sequence and protospacer adjacent motif in BTK exon 2. Each row beneath represents a detected cutting event—each dot corresponds to a perfect match to the target sequence, while base pair differences are illustrated with nucleotide symbols differring from the target sequence. **f)** Human CD34+ G-CSF mobilized peripheral blood stem cells were electroporated with BTK exon 2 tageting sgRNA/Cas9 RNP (either Wildtype, eSpCas9, VP12 Cas9, or Alt-R Cas9) and transduced with an AAV6 vector containing the I18u-WPRE exon 2 donor template. Targeted integration rates of each donor template into the BTK exon 2 locus were measured by ddPCR.



Supplemental Figure 1: Production and verification of a BTK-deficient K562 cell sub-line

K562 cells were electroporated with a sgRNA/Cas9 expression plasmid that targeted *BTK* exon 15. Following electroporation, fluorescence activated cell sorting was used to sort single cells into individual wells to establish clonal populations. After the cell populations were re-established, gDNA from each clonal population was harvested. **a)** A qualitative PCR screen was run to identify clones that did not form indels around the cut site. One primer was designed to bind directly over the exon 15 target site. A lack of indels on either copy of the BTK gene would allow for efficient amplification. If both copies of the gene had indels at the junction, no PCR

amplicon would form. The resulting gel identified a subset of unedited cells that were excluded from further screens. **b)** New PCR reactions using primers that encompassed the target site were performed and sent for Sanger sequencing. Three illustrative examples of sequencing results are shown. **c)** Clones that were identified as having homozygous deletions were analyzed via immunoblot to confirm the absence of the BTK protein. Clone 14 was an unmodified control sample. Clone 3 was selected as the optimal BTK-deficient clone due to it's small indel size (4 bp) that resulted in a premature stop codon and it produced no BTK protein.



Supplemental Figure 2: sgRNA variants demonstrated improved cutting efficiency when delivered as ribonucleoprotein to human CD34+ peripheral blood stem cells

a) Schematic of the 5' end of the *BTK* gene. Exon 1 (E1) and exon 2 (E2) represented by white boxes. The red section indicates the *BTK* intron 1 3' splice signal. Each vertical line represents an sgRNA assembled and tested. The black line labeled "1" represents the intron 1 target site that has been previously described. **b)** Each sgRNA was *in vitro* transcribed and pre-complexed with wildtype Cas9 protein before electroporation into human CD34+ peripheral blood stem cells. Indel formation at each target site was measured via the ICE webtool.

CHAPTER 3

RESEARCH PROJECT

A Comparison of DNA Repair Pathways to Achieve Site-Specific Correction in Bruton's Tyrosine Kinase

David H. Gray[1], Jasmine Santos[2], Alexandra Grace Keir[2], Isaac Villegas[3], Simon Maddock[2], Caroline Y. Kuo[3]

1 Molecular Biology Institute, University of California, Los Angeles, CA 90095, USA

2 Department of Microbiology, Immunology & Molecular Genetics, University of California, Los Angeles, CA 90095, USA

3 Department of Pediatrics, David Geffen School of Medicine at the University of California, Los Angeles, CA 90095, USA

4 Division of Allergy & Immunology, David Geffen School of Medicine at the University of California, Los Angeles, CA 90095, USA

Abstract:

Gene editing utilizing homology directed repair has advanced significantly for many monogenic hematopoietic diseases in recent years but have also been hindered by decreases between *in vitro* and *in vivo* gene integration rates. Homology directed repair occurs primarily in S/G_2 phases of cell cycle whereas long term engrafting hematopoietic stem cells are typically quiescent. Alternative methods for targeted integration have been proposed including homology independent targeted integration and precise integration into target chromosome which utilize non-homologous end joining and microhomology mediated end joining, respectively. Non-homologous end joining occurs throughout the cell cycle, while microhomology mediated end joining occurs predominantly in S phase. We compared these pathways for integration of a corrective DNA cassette at the Bruton's Tyrosine Kinase gene for treatment of X-Linked Agammaglobulinemia. Homology directed repair generated the most integration in K562 cells, however, synchronizing the cells into G_1 resulted in homology independent targeted integration producing the most integration. Only homology directed repair produced seamless integration junctions, making it optimal for targets where insertions and deletions are impermissible. Bulk CD34+ cells were best edited by homology directed repair and precise integration into target chromosome, while sorted hematopoietic stem cells contained similar integration rates from all three donors.

Introduction:

The identification of targeted endonucleases, particularly the RNA-guided clustered regularly interspaced short palindromic repeats (CRISPR)-associated proteins, has brought genome editing to the forefront of the gene therapy field.[1,2] These endonucleases can create double stranded breaks (DSBs) in genomic DNA with high efficiency and specificity, initiating one of several cellular repair pathways. The most thoroughly described DSB repair pathways are non-homologous end joining (NHEJ) and homology-directed repair (HDR). NHEJ is an error prone process that repairs the DSB while introducing random nucleotide insertions and deletions ('indels') at the repair junction.[3] In contrast, HDR uses a homologous template, typically a sister chromatid during cell replication, for seamless repair of the break.[4] Each of these two methods has been harnessed for different purposes. NHEJ is most commonly used to disrupt elements in the genome through the introduction of indels at the DSB, while HDR is used for targeted correction of single base pairs (bp) or insertion of large sequences of DNA for site-specific gene replacement.[5–9]

Ex vivo gene editing for diseases of the hematopoietic system, such as primary immunodeficiencies and hemoglobinopathies, is progressing towards the clinic with many promising therapeutic avenues underway. However, murine studies of HDR-based gene editing in hematopoietic stem and progenitor cells (HSPCs) by many groups have come upon a common hurdle: rates of editing via the HDR pathway in human HSPC often drop substantially over time in murine xenograft models, while NHEJ-mediated editing typically persists at stable rates.[10–13] This discrepancy is likely due to the cell cycle dependency of each repair pathway. HDR is generally restricted to the S/G_2 phases of the cell cycle while NHEJ is active throughout all phases of the cell cycle.[14] However, true hematopoietic stem cells (HSCs) exist mostly in G_0/G_1 and may

be inherently less capable of HDR than actively dividing progenitor cells.[15] Targeted integration into HSCs could be enhanced by either improving the effectiveness of HDR in quiescent cells or identifying methods of integration that utilize alternative mechanisms of DNA repair. Homology-independent targeted integration (HITI) and precise integration into target chromosome (PITCh) are two such alternative pathways for site-specific integration into DSBs which rely on NHEJ and microhomology mediated end joining (MMEJ), respectively.[16,17] HITI is expected to function throughout all phases of the cell cycle while PITCh occurs primarily in late G_1/S phases due to its dependence on MMEJ.[3,18]

In this report, each of these three methods of DNA integration--HDR, HITI, and PITCh--was assessed for their potential to achieve a permanent gene therapy based cure for X-Linked Agammaglobulinemia (XLA) through gene editing of autologous HSPC. XLA is a primary immunodeficiency resulting from a block in the pre-B cell phase of early B-lymphocyte development, which leads to a lack of mature B-lymphocytes and antibody production in most patients.[19–22] The current standard of care for XLA is immunoglobulin replacement with prompt treatment of infections using antibiotics. This treatment is life-saving and typically leads to a substantially improved quality of life, although it remains imperfect. Even with immunoglobulin supplementation, patients can experience recurrent bacterial infections, are susceptible to certain life-threatening viral infections such as enterovirus, and have reduced life expectancies. Additionally, immunoglobulin replacement is expensive and requires weekly to monthly infusions for a patient's lifetime.

Patients with XLA have loss of function mutations in the Bruton's Tyrosine Kinase (*BTK*) gene, which encodes a cytoplasmic signaling protein found in most hematopoietic lineages.[20] While previous work using integrating viral vectors to deliver new copies of the *BTK* gene to

patient HSPCs has produced promising results,[23,24] this approach has potential drawbacks to consider as well. *BTK* is a tightly regulated gene with evidence showing that both too little and too much expression can reduce protein efficacy.[25] In addition, the semi-random integration from viral vector mediated gene transfer may affect gene dosage and the regulation of the transgenic BTK depending on the insertion locus. Finally, the risk of insertional oncogenesis remains despite significant improvements with the current generation viral vectors. As an alternative to integrating viral vectors, targeted endonucleases such as Cas9 can be used to integrate a corrective, promoter-less copy of *BTK* at its endogenous locus in the genome presents an attractive way to sidestep the potential problems of viral vectors. Site-specific integration should retain most of the endogenous control of the *BTK* gene, leading to regulated expression near wildtype levels. Furthermore, using a DSB to guide integration with a highly specific endonuclease should nearly eliminate all off-target integration events and drastically reduce the risk of insertional oncogenesis.

While some genes can be repaired by replacing only specific bases or even small regions of the gene, neither of these approaches are feasible for BTK as pathogenic mutations span the entire locus without any noteworthy hotspots. Instead, a universal strategy that addresses nearly all known XLA-causing mutations is to integrate a nearly full length BTK cDNA donor at the start of the endogenous BTK locus. When integrated correctly, this donor supersedes the patient's mutated copy of the gene and leads to expression of only the corrected transgene. Here we compare three methods of targeted insertion--HDR, HITI, and PITCh--to achieve integration of a corrective cassette into intron 1 of BTK in cell lines and primary human PBSCs.

Results:

HDR, HITI, and PITCh plasmid donors achieve targeted integration and expression in K562 Cells

Corrective BTK cDNA cassettes were designed to integrate at intron 1 because intronic regions are permissive to imperfect junctions that are expected to result from HITI and PITCh mediated integrations (Fig. 1A, Supplemental Fig. 1). A single guide RNA (sgRNA) specific to intron 1 of *BTK* was delivered along with Cas9 either from an expression plasmid, as *in vitro* transcribed mRNA, or pre-complexed to Cas9 protein as a ribonucleoprotein (RNP) to introduce a targeted DSB. The resultant DSB can then be resolved via various integration pathways (Fig. 1B). HDR and PITCh require end-resection to expose homologous regions between the template and the target site and are limited to the S/G_2 or S phases of the cell cycle, respectively. In contrast, HITI does not require end-resection and can occur during any stage of the cell cycle.

Plasmid donor repair templates were assembled to assess the potential for each method of DNA integration to incorporate a DNA cassette in cell lines (Fig. 1C). HDR donors contained the corrective cassette flanked by 500 bp homology arms that parallel the sequences immediately adjacent to the Cas9 cut site. The cassette itself included the BTK cDNA sequence, three C terminal hemagglutinin (HA) tags to facilitate detection of the transgene in wildtype cells, a stop codon, and the BTK 3' untranslated region (UTR) (Supplemental Fig. 1A). Finally, to maintain splicing elements following donor integration, the 3' region of intron 1 was included immediately preceding the cDNA sequence. In the HDR donor ("HDR"), retention of this intronic sequence also preserved a functional Cas9 binding site, which could be re-cleaved and lead to additional integration events through pathways besides HDR as well as result in unwanted indels following seamless HDR-mediated donor integration. Therefore, a second variant of the HDR donor ("HDR ΔPAM") was generated with a 2 bp modification to the protospacer adjacent motif (PAM) site from NGG to NAA that abrogates Cas9 mediated cleavage of the donor after successful integration.

As HITI-mediated integration requires linearized donor cassettes, donor templates were designed to contain Cas9 binding sites flanking the cDNA sequence to allow plasmid linearization. CRISPR/Cas9 sites were included at either one or both ends of the donor template in an orientation opposite that found in the target genomic DNA. This reversal of Cas9 binding sites within the donor prevents restoration of sgRNA binding sites upon gene integration to eliminate unwanted re-cleavage events (Supplemental Fig. 1D). However, if donor cassettes are end-captured in reverse orientation, re-cleavage can continue until either the transgenic donor integrates in the desired orientation or indels form at the cut site. Although reversal of Cas9 binding site orientation favors correct integration events, proper integration of this cassette will nevertheless result in a genomic 'scar' of either 17 bp (the protospacer side of the cut site) or 6 bp (the PAM side of the cut site), making this strategy most suitable for gene editing approaches at intronic sequences.[16] BTK-specific HITI donors had a cut site on both ends of the corrective cassette ("2c"), a cut site on either the 5' or 3' end of the cassette ("1c5" or "1c3"), or no cut sites flanking the cassette ("0c").

PITCh donors contained the corrective cDNA sequence surrounded by microhomology regions of 20 bp that mirror the genomic DNA sequence surrounding the sgRNA cut site, all of which was flanked by two sgRNA cut sites. The small regions of homology guide the MMEJ repair process, although this donor may also integrate via NHEJ through end-trapping of the donor. Three PITCh donors were assembled: the first ("PITCh-1") had Cas9 binding sites positioned to minimize the length of residual DNA in the donor template after cleavage, the second ("PITCh-2") had both Cas9 binding sites positioned in the same orientation, and finally a third template ("PITCh-0c") lacked cut sites entirely.

Each donor plasmid was delivered to K562 human erythroleukemia cells via electroporation in conjunction with a Cas9 expression plasmid (pX330) encoding a sgRNA specific to *BTK* intron 1. Four days post-electroporation, integration rates of each donor were quantified using droplet digital PCR (ddPCR). The HDR donor achieved the highest integration rates of 18.2% ± 5.6%, while 11.4% ± 0.1% integration was observed using the HDR ΔPAM donor (Fig. 1D). The HITI-2c donor resulted in 4.2% ± 2.4% integration compared to 9.7% ± 3.0% using the HITI-1c5 donor. The HITI-1c3 donor had no detectable integration, although it would have produced a much larger amplicon outside the optimal range for the ddPCR assay. Likewise, the HITI donor without any cut sites had no detectable integration. Both PITCh-1 and PITCh-2 integrated at comparable rates, 6.2% ± 0.9% and 5.7% ± 1.9% respectively. The PITCh-0c donor achieved integration into 2.4% of cells. Immunoblot analysis of the gene-modified cells post electroporation demonstrated that the amount of transgenic protein closely followed the pattern of integration seen at the DNA level when probing for the C terminal HA tags (Fig. 1E). These data support the potential use for all three pathways to achieve targeted integration and expression of transgenic constructs in cell lines.

Targeted integration via HDR produces fewer indels in the integration junction than HITI or PITCh

To further characterize the integration patterns of the different donors, In/Out PCR was performed using four primer pairs designed to detect integration and orientation of the cDNA cassette or the TOPO 2.1 plasmid backbone. (Fig. 2A) Each primer pair shared one primer outside of the donor constructs oriented towards the target site, represented in black. Row 1 used the same primer pair used for ddPCR, and the resulting products demonstrated integration in the correct

orientation with each donor except HITI-1c3 and HITI-0c (Fig. 2B, Supplemental Fig. 2). Row 2 used a primer pair to amplify only events where the corrective cassette was integrated in reverse orientation. The four detectable products are from cells treated with HITI-2c, HITI-1c3, PITCh-1, and PITCh-2. These four donors all contain a cut site immediately 3' of the cassette. Rows 3 and 4 represent integration of the 5' or 3' ends of the TOPO 2.1 plasmid backbone at the target site. Every donor had detectable integration of the plasmid backbone. The donors that yielded the least backbone integration were those without an intact cut site in the donor: HDR ΔPAM, HITI-0C, and PITCh-0c. Of these three donors, only the HDR ΔPAM donor had substantial on-target integration.

Four of the PCR products from Row 1 (on-target cDNA integration events in proper orientation) were then TOPO TA cloned and analyzed by Sanger sequencing (n=30-44). (Fig. 2C,D) The HDR ΔPAM integration events all displayed seamless integration junctions whereas the HDR donor, which contains an intact PAM sequence, yielded no perfect clones. Instead, these sequences predominantly featured small indels of 1 to 5 bp. The HITI-2c donor had 10% of sequences with base perfect junctions. The most commonly found sequence was a single 8bp deletion in 30% of samples. The same 6 bp sequence, ATTAAT, is present on either side of the cut site and the deleted region corresponds to a deletion of one of those repeats and the intervening bases by a recombination event. The PITCh-2 sequences demonstrated integration via both NHEJ and MMEJ. NHEJ mediated integration of the PITCh-2 donor occurred in 9.1% of clones due to the flanking cut sites on the donor. In these cases, the whole construct and the flanking microhomology arms are end captured, thereby leaving residual bases from two copies of each microhomology arm being integrated. Another 90.9% of the clones had sequences that integrated

via MMEJ mediated integration with regions of microhomology flanking the cut site, similar to the HITI clones.

Synchronizing K562 cells into G_1 reduces rates of HDR and PITCh based integration while increasing HITI mediated integration

While K562 cells serve as a valuable tool for basic donor evaluation, they do not fully recapitulate primary human HSCs. HSCs are quiescent and primarily in G_0, whereas K562s are rapidly cycling and have a much higher prevalence of cells in S/G_2. To evaluate whether these donors may be effective in HSCs, K562 cells were reversibly synchronized to the G_1 phase of the cell cycle using hydroxyurea. Cells were first treated with hydroxyurea for 24 hours and analyzed by flow cytometry to identify changes in cell cycle. (Fig. 3A) Treatment with hydroxyurea resulted in a reduction of cells in the G_2/M phases of the cell cycle. (Supplemental Fig. 3) Both hydroxyurea treated and untreated cells were then electroporated with gRNA/Cas9 expression plasmids and one of the HDR, HDR ΔPAM, HITI-1c5, or PITCh-2 donors. Post-electroporation, cells were cultured in hydroxyurea for 3 days before being transferred to fresh medium lacking hydroxyurea for recovery and outgrowth. Integration rates in each population were measured via ddPCR.

Untreated K562 cells showed an integration pattern similar to previous experiments: the HDR ΔPAM, HDR, and HITI-1c5 donors yielded 11.4% ± 5.4%, 11.4% ± 4.4%, and 7.8% ± 1.9% integration, respectively, while the PITCh-2 donor produced only 3.0% ± 0.0% integration. K562 cells treated with hydroxyurea demonstrated a markedly different pattern of integration. (Fig. 3C) The frequency of integration of the HDR ΔPAM donor dropped to 1.1% ± 1.3% and the PITCh-2 donor produced 0.7% ± 0.4% integration. HDR dropped less precipitously to 4.9% ± 3.5% while the HITI-1c5 donor increased substantially to 14.7% ± 5.5% integration. These represent fold

changes of 0.43, 0.09, 1.75, and 0.235 for HDR, HDR ΔPAM, HITI-1c5, and PITCh-2 following hydroxyurea treatment.

HDR donor achieves the highest integration in bulk edited human PBSCs, however all three pathways yield similar integration in primitive human HSCs populations

After optimizing donors for each integration method in cell lines, we developed a strategy to deliver donor constructs to primary cells. The best performing donor for each of the three pathways (HDR ΔPAM, HITI 1c5, and PITCh-2) was packaged as AAV serotype 6. Mobilized PBSC from healthy donors were electroporated with gRNA/Cas9 as RNP following two days of pre-stimulation. Immediately following electroporation, cells were transduced with one of the three AAV6 donors for 24 hours. One week post electroporation, cells were harvested for genomic DNA and assessed for integration efficiency. Different donor multiplicities of infection (MOIs) were tested to maximize gene editing while minimizing toxicity of the donor template (Supplemental Fig. 4). While the HDR donor showed a very clear dose response to increasing MOI, the HITI and PITCh donors consistently yielded lower rates of integration. At MOIs above 1e5, the expansion of the cells dropped precipitously despite high viability rates (data not shown). In addition, three methods of gRNA/Cas9 delivery were assessed for integration of the AAV donors: chemically synthesized gRNA with chemical modifications (2'-O-methyl analogs and 3' phosphorothioate internucleotide linkage modifications on the first three 5' and 3' terminal RNA residues) with Cas9 mRNA, *in vitro* transcribed (IVT) gRNA with Cas9 mRNA, and the chemically modified gRNA precomplexed as Cas9 RNP (Fig. 4A). The HDR donor integrated at 21.9% ± 3.6% of cells treated with modified gRNA and Cas9 mRNA. However, the HITI and

PITCh donors both demonstrated their highest levels of editing with modified gRNA and Cas9 protein (1.7% ± 0.7% and 6.9% ± 3.5%). These conditions also produced 7.6% ± 2.1% editing with the HDR donor. Taking these results into consideration, modified gRNA with Cas9 as an RNP with a donor MOI of 1e5 was used for further comparisons.

To better understand editing in different HSC populations, bulk edited PBSCs were sorted into different HSC subpopulations. 24 hours following CRISPR/Cas9 RNP electroporation and AAV6 transduction, PBSCs were separated via fluorescence activated cell sorting into two different populations: HSCs (CD34+/CD38-/CD90+/CD45RA-) and progenitors (CD34+/CD38+). (Fig. 4B,C) After expansion in culture, gDNA was harvested and analyzed for donor integration via ddPCR. The HDR donor integrated in 5.2% ± 2.2% of bulk PBSCs, 7.4% ± 2.3% of sorted progenitors, and 4.8% ± 4.3% of sorted HSCs. The HITI donor yielded integration rates of 0.7% ± 0.2% in bulk, 1.7% ± 0.9% in progenitors, and 6.8% ± 0.2% in HSCs. Finally, the PITCh donor produced 1.2% ± 0.3% integration in bulk, 2.6% ± 1.9% in progenitors, and 5.7% ± 1.7% in HSCs. Despite large differences in editing efficiency in bulk PBSCs and sorted progenitors, all three methods led to comparable levels of integration in HSCs.

Discussion

The results here offer an unbiased comparison of three different methods of targeted integration at the *BTK* locus. In this study, nine donors (2 HDR, 4 HITI, and 3 PITCh) were evaluated for integration efficiency, aberrant integrations, and integration junction fidelity. In cell lines, HDR based methods led to the highest levels of integration, likely due to the continuously cycling nature of these cells. HDR, and particularly the HDR donor with a PAM mutation, led to the fewest undesired integration events detected by PCR. The same donor led to 100% of sequence

junctions being base perfect. HITI donors at the same target site had the second highest rates of integration. The addition of only one cut site on the 5' end of the donor achieved the highest integration rates. Notably, the addition of a second cut site on the 3' end of the donor led to a drop in integration of about 50%. This is likely because addition of a second cut site can split the donor into two similarly sized fragments (one being the corrective cassette and the other being the plasmid backbone) which compete to integrate into available nuclease-mediated DSB. In this scenario, each fragment would integrate approximately half of the time, thereby decreasing the rate of desirable integrations.

When examining the fidelity of the 5' integration junctions, some donor variants led to higher precision. Notably the HDR ΔPAM donor yielded 100% seamless junctions while the almost identical HDR donor had 0% sequence-perfect junctions. There are likely two different mechanisms at play to explain the discrepancy that results from only a 2 bp change in these donors. The first is that having an intact binding site for the gRNA/Cas9 in the donor plasmid made the donor capable of end capture integration via NHEJ, similar to the HITI donors, resulting in indels in the junction. The second mechanism is that even a perfect integration of the HDR donor restores the Cas9 binding site in the genomic DNA, allowing for the endonuclease to re-cleave an already integrated product. In this case, NHEJ would occur following the desired HDR event to produce indels at the junction. For HITI donors, even a "perfect" integration event with no additional indels will leave a scar by design. As expected, all of the HITI donors produced high levels of aberrant integration events via qualitative PCR screening and the HITI 2c donor had very few base perfect junctions identified via sequencing. Surprisingly, many of the sequenced junctions had a deletion of a region of microhomology, suggesting they may have integrated via the MMEJ pathway in some cases rather than the NHEJ pathway alone. PITCh donors consistently resulted in the lowest

integration rates with the majority of integration junctions having the same mid-sized deletion driven by microhomology.

HDR donors are ideal for use in rapidly cycling cell lines, however, in instances where homology arms cannot be practically added, such as large payloads approaching the maximum capacity of a delivery vector, using a HITI donor can be an acceptable replacement. It is important to note that these HITI donors leave DNA scarring in the form of duplicated portions of the protospacer and or PAM sequences, so they can only be used effectively in regions or cases where imperfect junctions can be tolerated, such as introns, UTRs, or with the addition of a transgenic promoter that eliminates donor reliance on endogenous sequences for regulation. In this study PITCh consistently performed the worst of the three tested methods and is not a recommended method of integration into cell lines.

Synchronizing the K562 cells into G_1 with hydroxyurea changed the dynamics of the donors and reflected the different mechanisms by which of the donors integrated. Pure HDR and PITCh donors both dropped to nearly 0% targeted integration while HITI doubled in integration. An interesting case in this experiment is that the HDR donor with an intact cut site only lost half of its integration compared to unsynchronized cells. This may be due to the ability of that donor to integrate via a HITI-like pathway rather than just HDR like the HDR ΔPAM donor. These experiments together support the previously published cell cycle dependencies of each pathway and are consistent with the initial hypothesis that HITI may be an ideal mechanism for targeted integration for non-replicating cells.

Transitioning into primary human PBSCs with the optimized donor for each integration method led to predicted integration rates in bulk cells. HDR led to the best overall integration, although that data is likely skewed by the rapidly cycling, larger proportion of progenitor cells in

bulk populations. Gene correction in more differentiated progenitor cells is generally not considered therapeutically relevant because they do not engraft long term and self-renew like true HSCs. Bulk edited human PBSCs were sorted into stem and progenitor populations to determine rates of gene integration in each cell type. While HDR vastly outperformed the two other pathways in bulk cells, all three donors led to nearly the same level of targeted integration in immunophenotypic HSCs. A particularly surprising finding was that HDR integration levels were maintained between bulk edited PBSCs and sorted HSCs. The fact that PITCh also led to similar rates of integration in HSCs went against the initial hypothesis due to MMEJ's dependence on the S phase of the cell cycle. One possible explanation is that the PITCh donor can also integrate via HITI due to the flanking cut sites beyond the microhomology arms.

In all, these findings suggest that HDR, HITI, and PITCh mediated targeted integration can be effective for editing in primary human HSCs. While levels of bulk editing vary substantially, integration rates are equivalent in the most quiescent HSC populations. Integration into an exon or sequence that requires a seamless junction is likely best approached with an HDR donor. However, integration into non-coding regions may tolerate all three approaches. HITI and PITCh donors do not require the full homology arms present in an HDR donor, which can conserve sequence length in larger donors that are nearing the size limits for viral packaging. At the BTK locus, HITI donors achieved similar or better results than PITCh donors in each experiment, making it the preferable alternative to HDR. While there are likely some locus dependent variables to integration proficiency that could impact the efficiency of certain pathways, the data presented here support the potential for all three pathways to lead to functional integration in HSCs and as potential avenues for clinically relevant gene therapies.

Materials and Methods:

Guide RNA design: A gRNA sequence targeting BTK intron 1 was initially identified using the design algorithm from CRISPRscan.[26] The gRNAs was assembled and cloned into a gRNA/Cas9 expression plasmid, referred to here as the pX330 plasmid. The pX330-U6-Chimeric_BB-CBh-hSpCas9 expression plasmid backbone was a gift from Feng Zhang [Addgene plasmid # 42230 ; http://n2t.net/addgene:42230 ; RRID:Addgene_42230].[27] The gRNA targeted the sequence AGGGATGAGGATTAATGTCC at the 3' end of *BTK* Intron 1.

Donor design: HDR donors featured *BTK* cDNAs from exon 2 to exon 19 that were codon-optimized using the GeneOptimizer online tool [Thermo Fisher Scientific, Waltham, MA]. BTK exon 2 features a small portion of untranslated sequence that was included in each donor, but not codon optimized, as there are no coding regions in this sequence. The donor template contains the 3' splice site from the end of *BTK* intron 1, the exon 2 portion of the 5' UTR, codon optimized exon 2-19 *BTK* cDNA, three C terminal HA tags, a stop codon, the BTK 3' UTR and a short sequence to improve polyadenylation.[28] HDR donors were flanked by 500 bp of homology matching the sequences immediately adjacent to the target site in the genomic DNA. PITCh donors were flanked by 20 bp of homologous sequence. HITI donors had no homologous flanking regions. HDR ΔPAM had a 2 bp mutation in the PAM sequence (NGG to NAA). For all cell line experiments, the donors were delivered as plasmids in the TOPO 2.1 backbone created with the Invitrogen TOPO TA cloning kit [Cat: K450002. Thermo Fisher Scientific, Waltham, MA]. Sequences were confirmed via Sanger sequencing.

K562 Electroporation and Hydroxyurea synchronization– 2e5 K562 cells per condition were nucleofected with 500 ng of gRNA/Cas9 expression plasmid and 3 ug of the respective donor

plasmid using the Lonza 4D Nucleofector X Unit [Cat: AAF-1002X. Lonza, Basel, Switzerland] in 20 uL of Lonza SF buffer [Cat: V4XC-2032. Lonza, Basel, Switzerland]. After electroporation, cells were plated in 500 uL of "R10": RPMI 1640 [Cat: 15-040-CV, Corning, Corning, NY] supplemented with 1x penicillin/streptomycin/glutamine [Cat: 10378016, ThermoFisher Scientific, Waltham, MA] and 10% fetal bovine serum [Omega Scientific, Tarzana, CA]. For cell cycle synchronization, K562 cells were pretreated with 100 ug/mL of hydroxyurea for 1 day before nucleofection [Cat: H8627, Millipore Sigma, Burlington, MA]. Samples of the treated and untreated cells were stained with Hoescht dye [Cat: 565877, BD Biosciences, San Diego, CA] following manufacturer protocol and analyzed via Flow Cytometry pre-electroporation to assess G_0/G_1 synchronization. Cells were electroporated using the above conditions and plated into R10 either with or without 100 ug/mL of Hydroxyurea. Three days after nucleofection, the treated cells were washed with Dulbecco's Phosphate-Buffered Saline [Cat: 14190250, ThermoFisher Scientific, Waltham, MA] and resuspended in R10 without hydroxyurea. Genomic DNA was harvested from the cells using the Invitrogen PureLink Genomic DNA Mini Kit [Cat: K182002, ThermoFisher Scientific, Waltham, MA]. Protein lysates from 2×10^6 cells per condition was harvested one to three weeks post-nucleofection using Denaturing Cell Extraction Buffer [Cat FNN0091, ThermoFisher Scientific, Waltham, MA] + 1x Halt Protease Inhibitor [Cat: 87786, ThermoFisher Scientific, Waltham, MA]. Lysate concentrations were quantified via Bicinchoninic Acid (BCA) assay [Cat: 23225, ThermoFisher Scientific, Waltham, MA].

<u>ddPCR</u>– Integration rates of the donors were quantified using ddPCR. For the all donors, the forward ddPCR primer was designed to bind just beyond the 5' end of the 500 bp homology arm and the reverse ddPCR primer was designed to bind in the 5' end of the cDNA sequence. A probe oligonucleotide labeled with 5' 6-FAM (Fluorescein) and a ZEN quencher that bound within the

PCR product was used to quantify the integrated product. For comparison, primers and a probe (conjugated with 5' Hexachlorofluorescein and a ZEN quencher) that bound to housekeeping gene on the X-chromosome were also run. All primers and probes were ordered from IDT [Integrated DNA Technologies, Coralville, IA]. Samples that produced fewer than 10,000 valid droplets were removed from analysis.

<u>Immunoblotting</u> –The Western Blot Assay was used to confirm the expression of exogenous BTK due to the integrated cDNA. Equal amounts of protein lysates per sample were loaded in each well of 4-12% Bis-Tris gel [Cat: NP0322BOX, Thermo Fisher Scientific, Waltham, MA]. The gel ran at 200 V for 1 hour. Transfer of the proteins in the gel to a PVDF membrane [Cat: 88520, Thermo Fisher Scientific, Waltham, MA] was performed at 30V for 75 minutes. The membrane was blocked in 5% milk in 1x TBST [Cat: 28360, Thermo Fisher Scientific, Waltham, MA] for 1 hour. The membrane was then cut in half to separate the HA bands from the actin bands for staining. The constructs were designed to have an HA tag for differentiation between endogenous and exogenous BTK using a primary antibody specific for HA [Cat: NBP2-43714, Novus Biologicals, Centennial, CO]. Beta actin was used as a loading control [MA5-15739 Thermo Fisher Scientific, Waltham, MA], and samples were diluted 1:1000 in 5 mL of 5% milk in TBST and allowed to bind overnight at 4°C with light shaking. The secondary antibodies were conjugated with horse radish peroxidase [Cat: 31460, ThermoFisher Scientific, Waltham, MA] for the HA secondary and Alexa Fluor 647 for the actin secondary [A-21236 ThermoFisher Scientific, Waltham, MA]. The secondaries were diluted 1:750 for HA and 1:1000 for actin, both in 5% milk mass/volume in TBST [28360 Thermo Fisher Scientific, Waltham, MA]. The HA half of the PVDF membrane was stained with Pierce ECL plus [Cat: 32132, Thermo Fisher Scientific, Waltham, MA]. Both halves were imaged using a Typhoon phosphoimager.

Qualitative Analysis of Integration Events

Genomic DNA harvested from electroporated K562 cells was PCR amplified using 4 different primer pairs. Each reaction had one primer in common – the forward primer used in ddPCR that binds just beyond the 500 bp of homology in the HDR donors. The four reverse primers used were designed to detect integration of the donor cassette in forward or reverse or the plasmid backbone in forward or reverse. The primer to detect the donor cassette integrating correctly into its target locus was the same reverse primer used in ddPCR. Each gDNA sample was amplified with all four primer pairs using the AccuPrime Taq HiFi system [ThermoFisher Scientific, Waltham, MA]. These products were run on agarose gels and imaged.

PCR products from the correct integration events amplified with the same primers used in ddPCR were cloned into a plasmid backbone using the TOPO TA 2.1 cloning kit [ThermoFisher Scientific, Waltham, MA]. Chemically competent *E. coli* cells [Cat: D5T09C-050-100, Moradec LLC, San Diego, CA] were transformed with 2 uL of the resultant material. 50 of the resultant bacterial colonies were grown in 3 mL of Luria broth [Cat: 10855001, ThermoFisher Scientific, Waltham, MA] overnight and plasmid was extracted the following morning [PureLink Quick Plasmid Miniprep Kit, ThermoFisher Scientific, Waltham, MA]. The plasmids were Sanger sequenced and analyzed to identify the sequences of each integration junction.

Human Peripheral Blood Stem Cell Electroporations:

Human CD34+ Granulocyte colony-stimulating factor (G-CSF) mobilized peripheral blood stem and progenitor cells were thawed two days pre-electroporation. Cells were pre-stimulated at a density of 5 x 10^5 cells/mL in X-Vivo 15 medium [Cat: BEBP04-744Q, Lonza, Basel, Switzerland] supplemented with 1x penicillin/streptomycin/glutamine [Cat: 10378016, ThermoFisher Scientific, Waltham, MA], 50 ng/uL recombinant human (rh) stem cell factor [Cat: 300-07,

PeproTech, Rocky Hill, NJ], 50 ng/uL rh Flt-3 ligand [Cat: 300-19, PeproTech, Rocky Hill, NJ], and 50 ng/uL of rh thrombopoietin [Cat: 300-18, PeproTech, Rocky Hill, NJ]. After two days of prestimulation, cells were electroporated using the BTX system [BTX, Holliston, MA]. 2e5 cells/condition were electroporated in 100 uL of BTX buffer (settings: 155 V, 30 ms, 1 pulse). Guide RNA and Cas9 were pre-aliquoted and either complexed ribonucleoprotein or as RNA. Concentrations for RNP complexing were 9 ug of chemically modified gRNA 2'-O-methyl analogs and 3' phosphorothioate internucleotide linkages at the first three 5' and 3' terminal RNA residues [Synthego, Silicon Valley, CA] with 200 pmol of Cas9 protein. For Cas9 mRNA electroporations, 5 ug of the chemically modified gRNA was used with 5 ug of WT SpCas9 mRNA. Mock treated cells were not electroporated but were otherwise handled the same as experimental conditions. Immediately post electroporation, cells were allowed to recover 10 minutes in the cuvette before being transferred into wells of 500 uL of fresh X-Vivo 15 medium supplemented as described above. Each sample was then transduced with the relevant multiplicity of infection of AAV6 donor vector, between 1e4 and 1e6 [Vigene Biosciences, Rockville, MD]. One day post electroporation, cells were counted, spun and the supernatant was aspirated. The cells were resuspended in 500 uL of outgrowth medium: IMDM [Cat: 12440053, ThermoFisher Scientific, Waltham, MA], 10% fetal bovine serum, 1x penicillin/streptomycin/glutamine [Cat: 10378016, ThermoFisher Scientific, Waltham, MA], 5 ng/mL of rhIL-3, 10 ng/mL of rhIL-6, and 25 ng/mL of rhSCF [Cat: 200-03, 200-06, and 300-07, PeproTech, Rocky Hill, NJ]. Four days post electroporation, cells were harvested for genomic DNA extraction.

Sorting Human Hematopoietic Stem and Progenitor Cells into Different Stem Populations

Primary CD34+ HSPCs from G-CSF mobilized peripheral blood were electroporated and transduced following the protocols listed above except with 6e6 cells per condition. Samples

treated with Cas9 mRNA were split into 6 cuvettes and electroporated at 1e6 cells/cuvette (each using previously described concentrations of Cas9 mRNA and modified gRNA) before being consolidated back together immediately before the 24-hour transduction. At the end of the transduction period (24 hours post electroporation), cells were harvested and counted. 5e4 cells/condition were plated in IMDM [Cat: 12440053, ThermoFisher Scientific, Waltham, MA], 10% fetal bovine serum, 1x penicillin/streptomycin/glutamine [Cat: 10378016, ThermoFisher Scientific, Waltham, MA], 5 ng/mL of rhIL-3, 10 ng/mL of rhIL-6, and 25 ng/mL of rhSCF [Cat: 200-03, 200-06, and 300-07, PeproTech, Rocky Hill, NJ] to grow out and measure bulk editing rates. The rest of the cells were pelleted, and the supernatants were aspirated. Cells were sorted using an Aria FACS sorter into two populations: HSC (CD34+, CD38-, CD45RA-, CD90+) and Progenitors (CD34+, CD38+). Each population was allowed to recover in IMDM [Cat: 12440053, ThermoFisher Scientific, Waltham, MA], 10% fetal bovine serum, 1x penicillin/streptomycin/glutamine [Cat: 10378016, ThermoFisher Scientific, Waltham, MA], 5 ng/mL of rhIL-3, 10 ng/mL of rhIL-6, and 25 ng/mL of rhSCF [Cat: 200-03, 200-06, and 300-07, PeproTech, Rocky Hill, NJ]. 3-10 days post sort, gDNA was extracted from each condition for analysis.

Acknowledgments:

The authors thank the UCLA Broad Stem Cell Center Flow Cytometry Core for their assistance with these studies. The authors thank Jessica Scholes, Felicia Codrea, and Jeffrey Calimlim for their expertise and assistance. This work was supported by research grants from the Ruth L. Kirschstein National Research Service Awards AI060567 and GM007185, the Broad Foundation

Innovative Pilot Stem Cell Research Project Award, and I3T-CTSI Discovery Award/David Geffen School of Medicine Seed Grant Program.

Author Contributions:

Conceptualization, D.H.G.; Methodology, D.H.G., J.S., A.G.K., and I.V.; Investigation, D.H.G., J.S., A.G.K, I.V., and S.M.; Resources, C.Y.K.; Writing – Original Draft, D.H.G; Writing – Review & Editing, D.H.G. and C.Y.K.; Visualization, D.H.G.; Supervision, C.Y.K.; Funding Acquisition, D.H.G and C.Y.K.

References

1. Jinek, M, Chylinski, K, Fonfara, I, Hauer, M, Doudna, JA and Charpentier, E (2012). A Programmable Dual-RNA – Guided DNA Endonuclease S figs. *Science (80-.).* **337**: 816–822.

2. Cong, L, Ran, FA, Cox, D, Lin, S, Barretto, R, Habib, N, *et al.* (2013). Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* **339**: 819–823.

3. Metzger, L and Iliakis, G (1991). Kinetics of DNA double-Strand break repair throughout the cell cycle as assayed by pulsed field gel electrophoresis in CHO cells. *Int. J. Radiat. Biol.* **59**: 1325–1339.

4. Kim, S, Kim, D, Cho, SW, Kim, J and Kim, J-S (2014). Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res.* **24**: 1012–1019.

5. Bauer, DE, Kamran, SC, Lessard, S, Xu, J, Fujiwara, Y, Lin, C, *et al.* (2013). An erythroid enhancer of BCL11A subject to genetic variation determines fetal hemoglobin level. *Science (80-.).* **342**: 253–257.

6. Bjurström, CF, Mojadidi, M, Phillips, J, Kuo, C, Lai, S, Lill, GR, *et al.* (2017). Reactivating Fetal Hemoglobin Expression in Human Adult Erythroblasts Through BCL11A Knockdown Using Targeted Endonucleases. *Mol. Ther. - Nucleic Acids* **5**.

7. Cradick, TJ, Fine, EJ, Antico, CJ and Bao, G (2013). CRISPR/Cas9 systems targeting β-globin and CCR5 genes have substantial off-target activity. *Nucleic Acids Res.* **41**: 9584–9592.

8. Hoban, MD, Cost, GJ, Mendel, MC, Romero, Z, Kaufman, ML, Joglekar, a. V., *et al.* (2015). Correction of the sickle-cell disease mutation in human hematopoietic stem/progenitor cells. *Blood*doi:10.1182/blood-2014-12-615948.

9. Kuo, CY, Long, JD, Campo-Fernandez, B, de Oliveira, S, Cooper, AR, Romero, Z, *et al.* (2018). Site-Specific Gene Editing of Human Hematopoietic Stem Cells for X-Linked Hyper-IgM Syndrome. *Cell Rep.* **23**: 2606–2616.

10. De Ravin, SS, Reik, A, Liu, P-Q, Li, L, Wu, X, Su, L, *et al.* (2016). Targeted gene addition in human CD34(+) hematopoietic cells for correction of X-linked chronic granulomatous disease. *Nat. Biotechnol.* **34**: 424–429.

11. Genovese, P, Schiroli, G, Escobar, G, Di Tomaso, T, Firrito, C, Calabria, A, *et al.* (2014). Targeted genome editing in human repopulating haematopoietic stem cells. *Nature* **510**: 235–240.

12. Schiroli, G, Genovese, P, Castiello, MC, Capo, V, Gregory, PD, Holmes, MC, *et al.*

(2016). 481. Targeted Genome Editing in Mouse Hematopoietic Stem/Progenitor Cells (HSPC) To Model Gene Correction of SCID-X1. *Mol. Ther.* **23**: S191.

13. Wang, J, Exline, CM, DeClercq, JJ, Llewellyn, GN, Hayward, SB, Li, PW-L, *et al.* (2015). Homology-driven genome editing in hematopoietic stem and progenitor cells using ZFN mRNA and AAV6 donors. *Nat. Biotechnol.* **33**: 1256–1263.

14. Branzei, D and Foiani, M (2008). Regulation of DNA repair throughout the cell cycle. *Nat. Rev. Mol. Cell Biol.* **9**: 297–308.

15. Pietras, EM, Warr, MR and Passegué, E (2011). Cell cycle regulation in hematopoietic stem cells. *J. Cell Biol.* **195**: 709–720.

16. Suzuki, K, Tsunekawa, Y, Hernandez-Benitez, R, Wu, J, Zhu, J, Kim, EJ, *et al.* (2016). In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration. *Nature* **540**: 144–149.

17. Nakade, S, Tsubota, T, Sakane, Y, Kume, S, Sakamoto, N, Obara, M, *et al.* (2014). Microhomology-mediated end-joining-dependent integration of donor DNA in cells and animals using TALENs and CRISPR/Cas9. *Nat. Commun.* **5**.

18. Truong, LN, Li, Y, Shi, LZ, Hwang, PY-H, He, J, Wang, H, *et al.* (2013). Microhomology-mediated End Joining and Homologous Recombination share the initial end resection step to repair DNA double-strand breaks in mammalian cells. *Proc. Natl. Acad. Sci.* **110**: 7720–7725.

19. Futatani, T, Miyawaki, T, Tsukada, S, Hashimoto, S, Kunikata, T, Arai, S, *et al.* (1998). Deficient expression of Bruton's tyrosine kinase in monocytes from X-linked agammaglobulinemia as evaluated by a flow cytometric analysis and its clinical application to carrier detection. *Blood* **91**: 595–602.

20. Vetrie, D, Vorechovsky, I, Sideras, P, Holland, J, Davies, A, Flinter, F, *et al.* (1993). The gene involved in X-linked agammaglobulinaemia is a member of the src family of protein-tyrosine kinases. *Nature* **361**: 226–233.

21. Siegal, FP, Pernis, B and Kunkel, HG (1971). Lymphocytes in human immunodeficiency states: a study of membrane-associated immunoglobulins. *Eur. J. Immunol.* **1**: 482–486.

22. BRUTON, OC (1952). AGAMMAGLOBULINEMIA. *Pediatrics* **9**: 722 LP – 728.

23. Ng, YY, Baert, MRM, Pike-Overzet, K, Rodijk, M, Brugman, MH, Schambach, a, *et al.* (2010). Correction of B-cell development in Btk-deficient mice using lentiviral vectors with codon-optimized human BTK. *Leuk. Off. J. Leuk. Soc. Am. Leuk. Res. Fund, U.K* **24**: 1617–1630.

24. Sather, BD, Ryu, BY, Stirling, B V, Garibov, M, Kerns, HM, Humblet-Baron, S, *et al.*

(2011). Development of B-lineage predominant lentiviral vectors for use in genetic therapies for B cell disorders. *Mol. Ther.* **19**: 515–525.

25. Satterthwaite, AB, Cheroutre, H, Khan, WN, Sideras, P and Witte, ON (1997). Btk dosage determines sensitivity to B cell antigen receptor cross-linking. *Proc. Natl. Acad. Sci. U. S. A.* **94**: 13152–13157.

26. Moreno-Mateos, MA, Vejnar, CE, Beaudoin, J, Fernandez, JP, Mis, EK, Khokha, MK, *et al.* (2015). CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. *Nat. Methods* **12**: 982–8.

27. Le Cong (2013). Multiplex Genome Engineering Using CRISPR-Cas Systems (Feng Zhang). *Science (80-.).* **339**: 53–68.

28. Levitt, N, Briggs, D, Gil, A and Proudfoot, NJ (1989). Definition of an efficient synthetic poly(A) site. *Genes Dev.* **3**: 1019–1025.

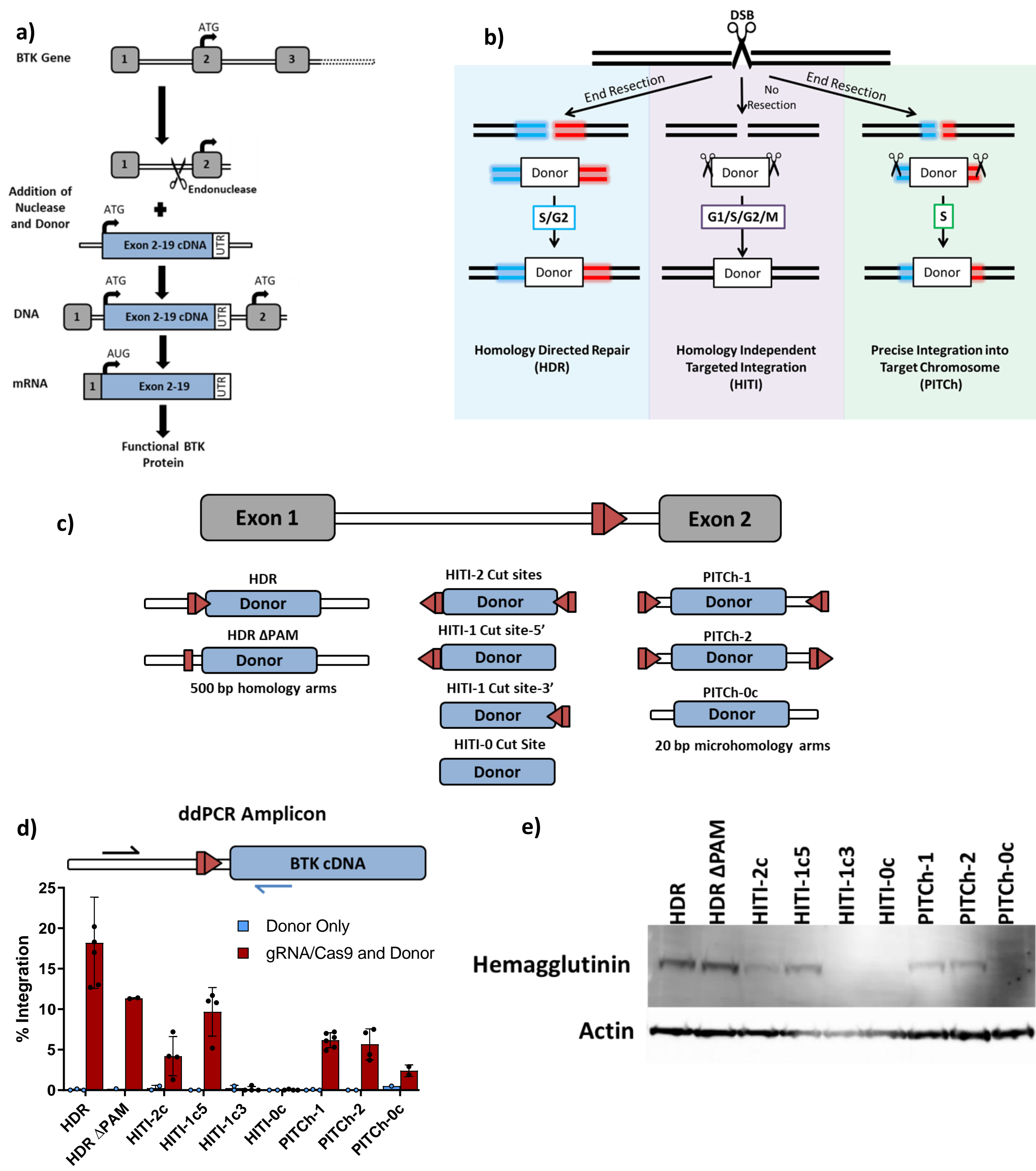


Figure 1: Editing schema and methods of targeted DNA integration

a) Diagram of proposed editing of the Bruton's Tyrosine Kinase (*BTK*) gene. Boxes represent exons while double lines are introns. Cas9/guide RNA (gRNA) targeted to *BTK* intron 1 creates a double stranded break (DSB) at the 3' end of the intron. Simultaneous addition of an exogenous donor allows for one of the cell's natural DSB repair pathways to be harnessed for targeted integration of the construct into the open cut site. The integrated product should be incorporated into RNA transcribed from the *BTK* locus and lead to production of functional protein. **b)** DSBs

can be repair through multiple pathways, at least three of which can be harnessed for targeted integration of exogenous DNA. When end resection occurs, homology directed repair (HDR) or Precise Integration into Target Chromosome (PITCh) can use homologous DNA sequences of varying lengths to guide integration into the DSB. These pathways occur predominantly in the S/G_2 and S phases of the cell cycle, respectively. HDR requires longer tracks of homology while PITCh has short microhomology regions flanked by regions that will bind to the introduced RNP and be cut. If no end resection occurs, homology independent targeted integration (HITI) can be harnessed for integration. HITI donors have no homology to the cut site, however they are flanked by 1 or 2 cut sites that match the RNP introduced. HITI can occur throughout all phases of the cell cycle. **c)** Simplified diagrams of the 9 donor variants evaluated. The red arrows together represent guide RNA (gRNA)/Cas9 binding sites with the rectangle representing the 17 bp of the protospacer on one side of the cut site while the triangle represents the remaining 3 bp of protospacer and the protospacer adjacent motif (PAM). Two HDR donor variants were created that differed only in the presence or absence of an intact PAM sequence. Four HITI donors were made with gRNA/Cas9 binding sites on both, either, or neither sides of the corrective donor template. Three PITCh donors were made: two with gRNA/Cas9 binding sites in differing orientations, and one that lacked these binding sites entirely. In addition to the donor sequence, each contains 3 C terminal hemagglutinin epitope tags to allow for identification of transgenic protein products. **d)** Diagram of the droplet digital PCR (ddPCR) amplicon used to measure integration rates of each donor template in K562 erythroleukemia cell lines. One primer bound the codon optimized sequence of the donor template while other bound *BTK* intron 1 outside of the homology arm sequence found in any of the donors. Integration results from each condition are displayed when either the donor plasmid was electroporated into K562 cells alone (blue bars) or in conjunction with the gRNA/Cas9 expression plasmid (red bars). **e)** Immunoblot analysis of cells treated with both gRNA/Cas9 and the respective donor template listed above. Lysates were probed for the hemagglutinin epitope tags and for an actin loading control.

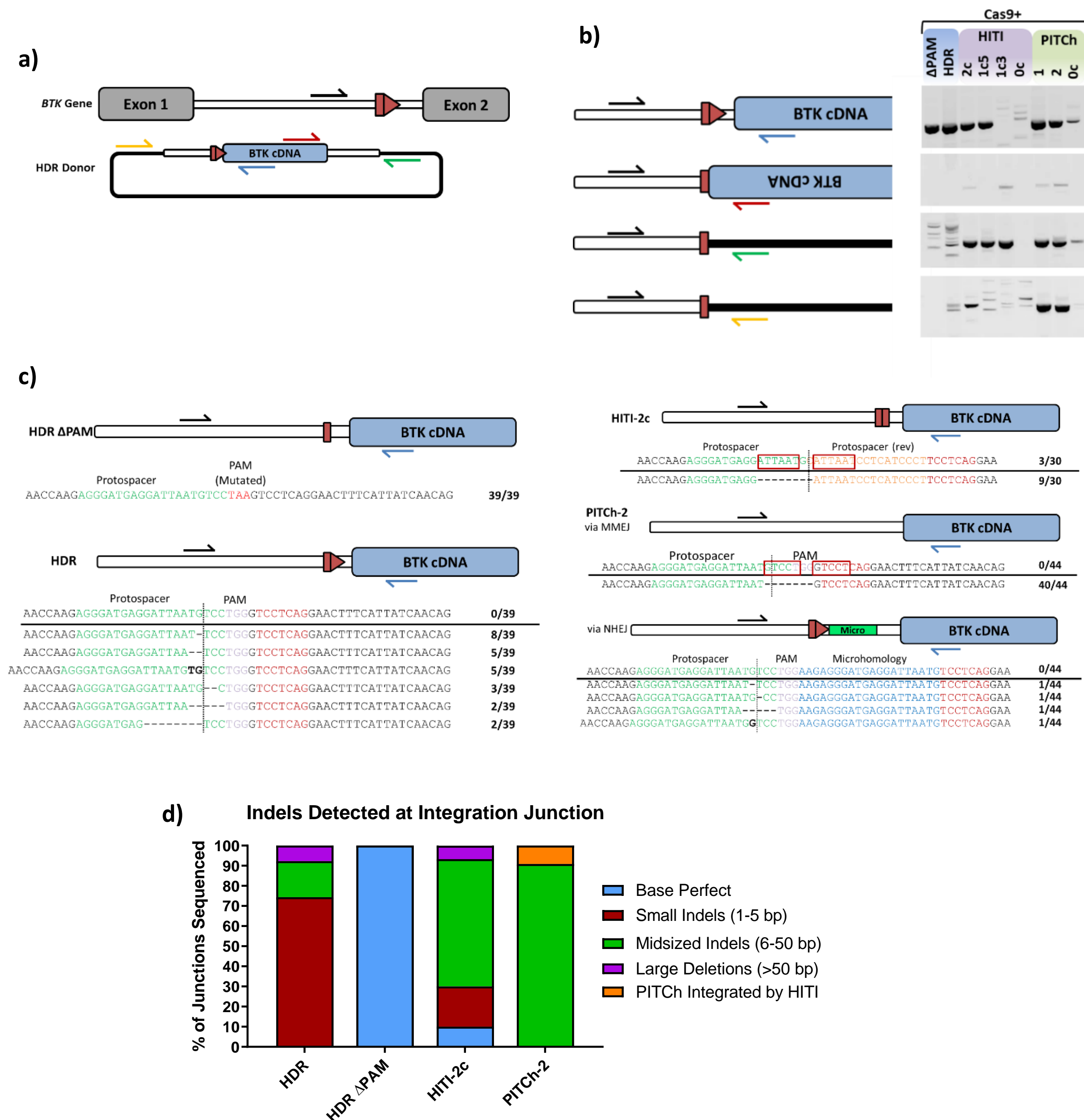


Figure 2: Identifying integration events of plasmid donors in BTK deficient K562 cells

a) Primers were designed to amplify and identify various integration events via In/Out PCR. One primer bound to the genomic DNA beyond the homology arms of any of the donors. Two primers bound within the donor cDNA, one on either side of the construct. The final two bound to either end of the plasmid backbone. **b)** Schematics for of each PCR amplicon and the resultant gel electrophoresis. BTK deficient K562 cells were electroporated with 500 ng of pX330 gRNA/Cas9 expression plasmid and 3 ug of the respective donor plasmid in the TOPO 2.1 backbone. The first two rows represent the cDNA integrated in the correct or reverse orientations. The last two rows show the plasmid backbone integrated in either potential orientation into the gDNA. **c)** PCR products from the HDR ΔPAM, HDR, HITI 2c, and PITCh-2 donors were TOPO TA cloned and

the resulting plasmids were Sanger sequenced to identify the integration junctions produced by each integration method. **d)** Summary of Sanger sequencing data.

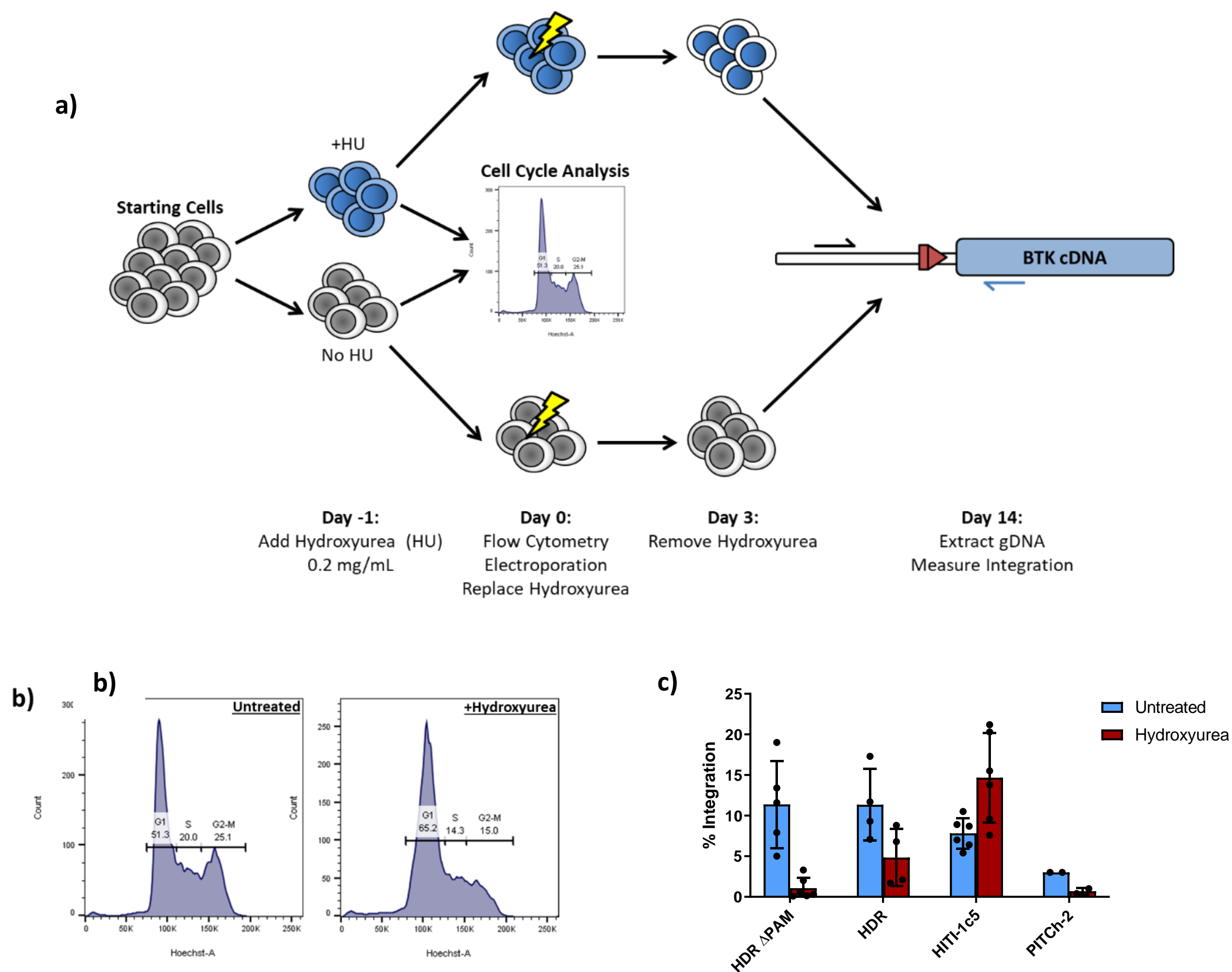


Figure 3: Synchronizing K562 cells in G_1 increases corrective donor integration via HITI while decreasing integration from HDR and PITCh

a) Workflow of hydroxyurea (HU) synchronization. Cells were cultured with HU for 1 day before electroporation. Immediately preceding electroporation, a split of cells were analyzed for cell cycle distributions via flow cytometry for Hoechst stain. Following electroporation, cells were cultured for 3 days in media containing HU. The HU was then removed and the cells grown in fresh media for 11 days before harvesting for targeted integration quantification via droplet digital PCR. **b)** Representative flow cytometry plots for untreated and HU treated K562 cell populations on day 0. **c)** Targeted integration rates of corrective donors into the BTK locus in K562 cells. The blue bar represents cells grown without HU, while the red bar represents cells treated with HU.

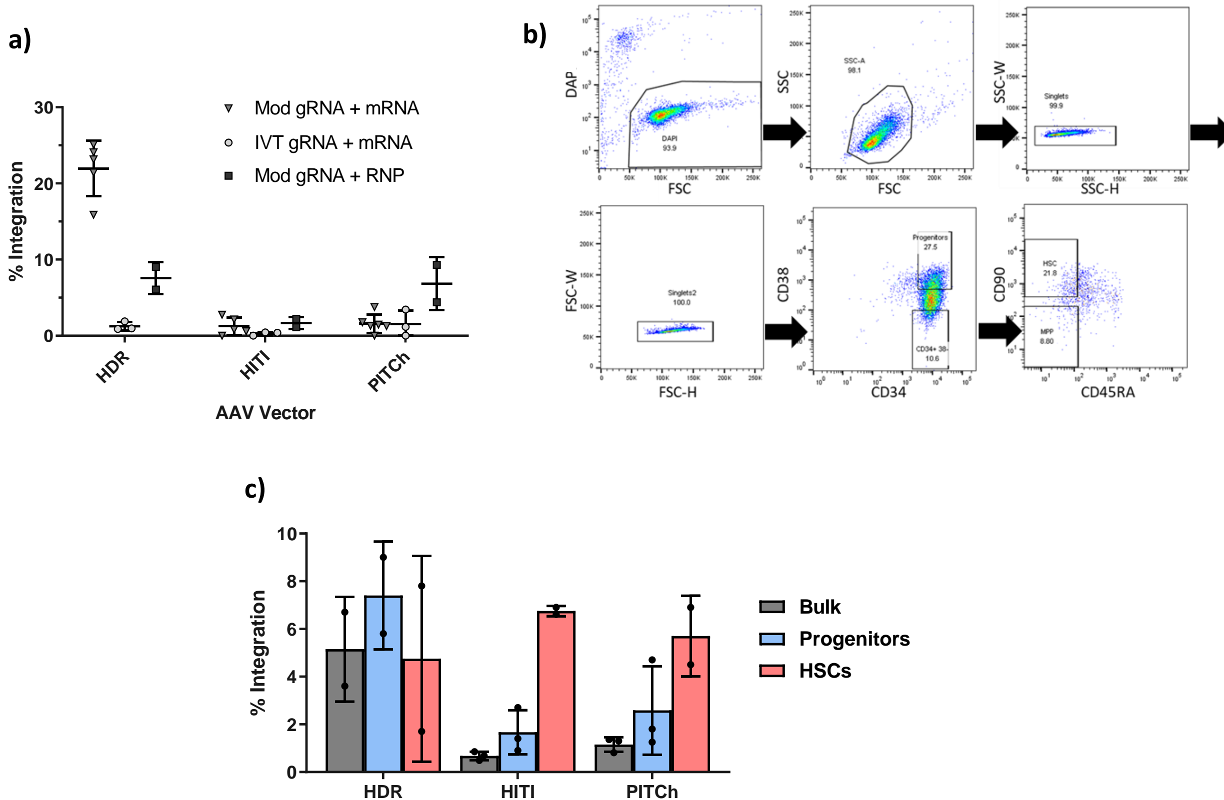
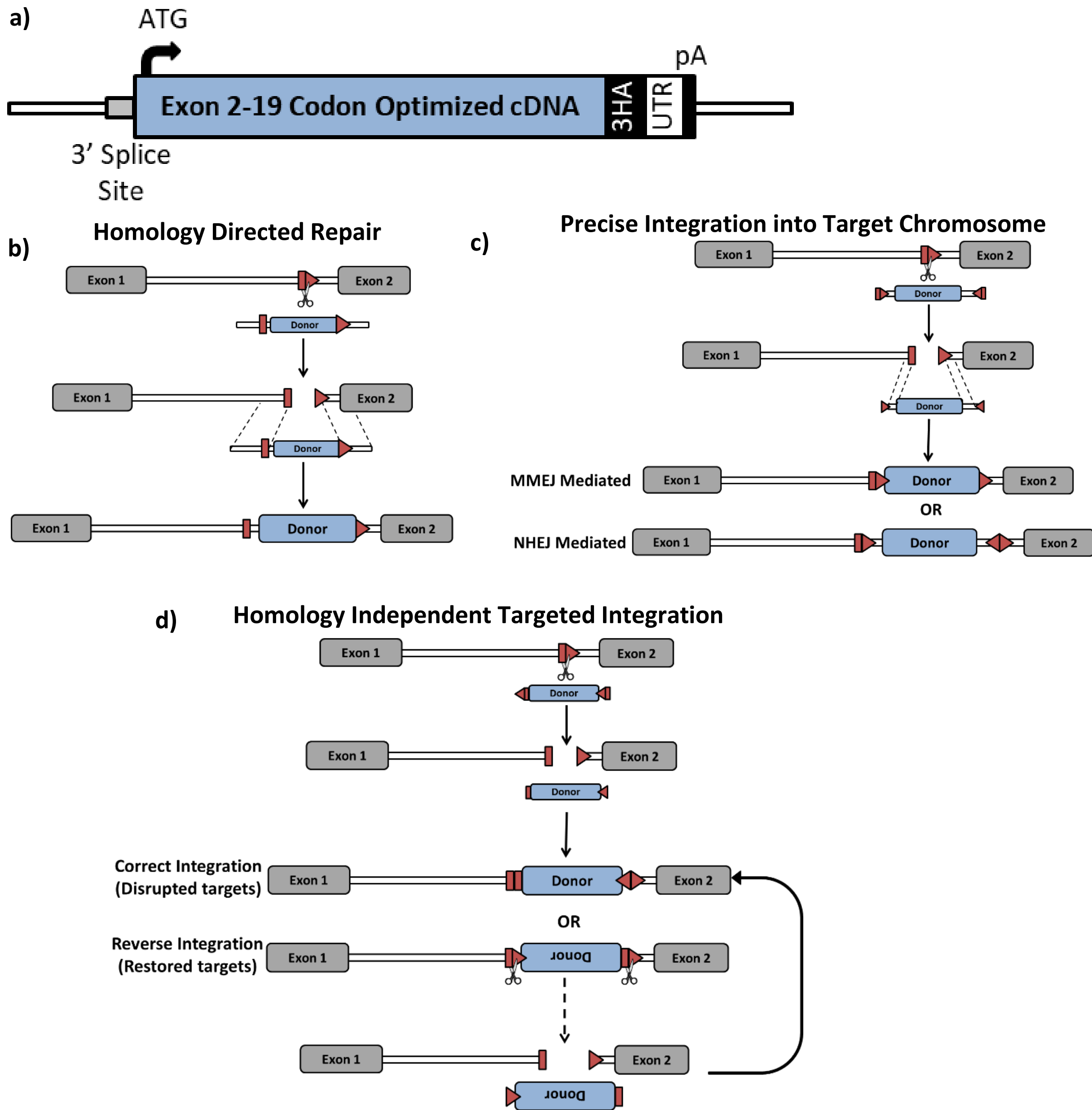


Figure 4: Targeted integration rates in bulk and population sorted primary human mobilized peripheral blood stem cells

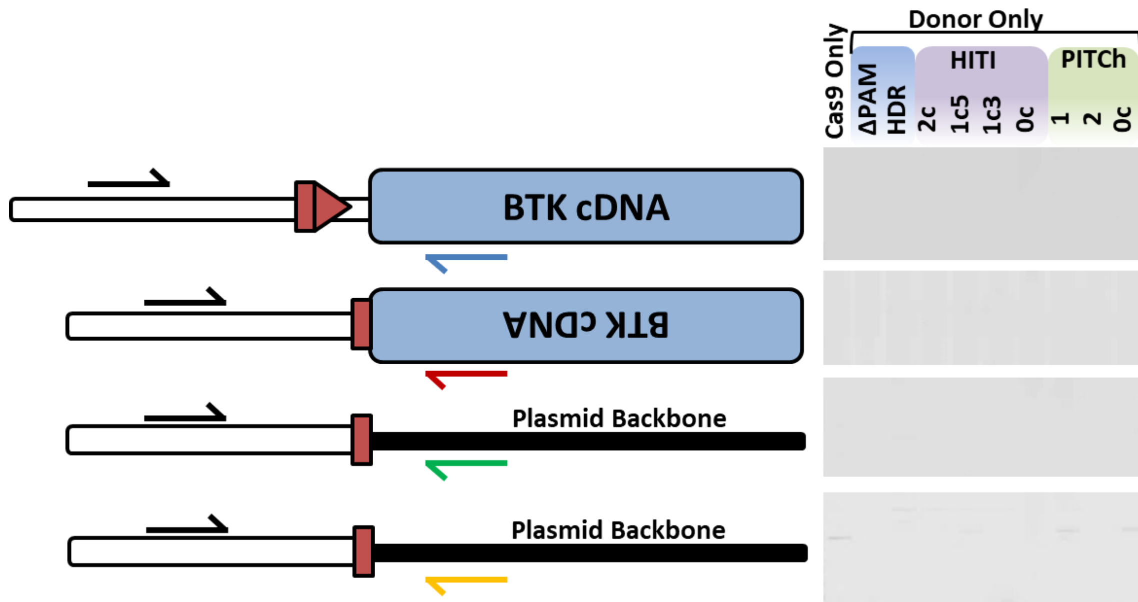
a) Targeted integration rates of each donor type in bulk CD34+ human mobilized peripheral blood stem cells (PBSCs). The three best performing adeno associated viral (AAV) donor vectors were tested for integration potential. Single guide RNA (gRNA) was produced via *in vitro* transcription (IVT) or chemically synthesized with 2'-O-methyl analogs and 3' phosphorothioate internucleotide linkages at each of the three 5' and 3' terminal RNA residues (labeled "Mod"). Cas9 was delivered either as a protein precomplexed with the gRNA (ribonucleoprotein, or "RNP") or as messenger RNA. Different combinations of these reagents were tested for targeted integration of each donor variant. **b)** Representative flow cytometry plot showing the sorting scheme to differentiate between hematopoietic stem cells (HSCs: CD34+, CD38-, CD45RA-, CD90+) and progenitors (CD34+, CD38+). **c)** Targeted integration rates of each donor type in bulk PBSCs (gray), Progenitors (blue), or HSCs (red)



Supplemental Figure 1: Schematics of donor template and targeted integration pathways

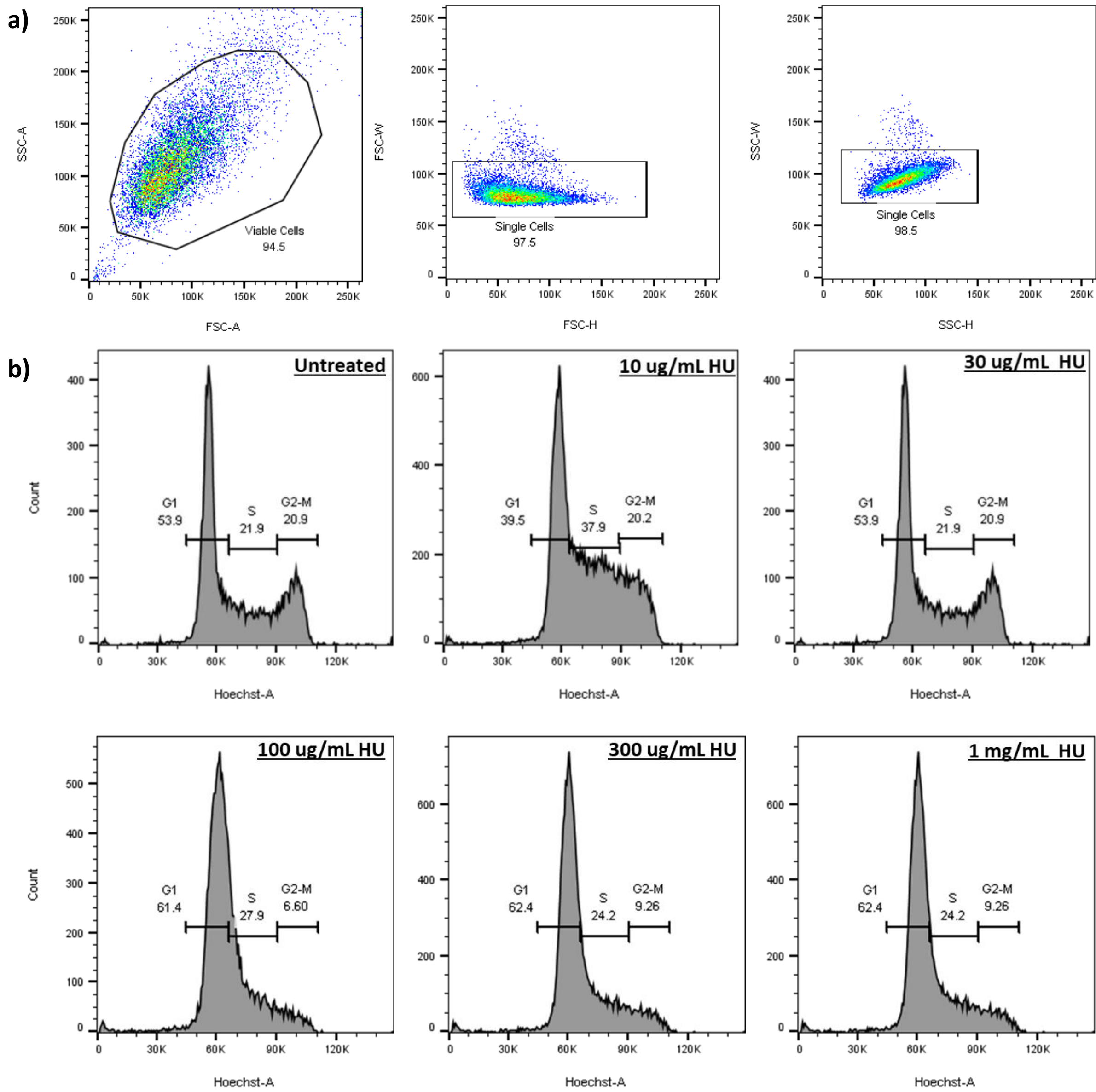
a) Detailed diagram of an HDR donor template. Each donor has the full *BTK* coding sequence (exons 2-19) that have been codon optimized. There are 3 C terminal hemagglutinin (3HA) tags immediately after the coding sequence before the stop codon. Following the stop codon is a copy of the *BTK* 3' untranslated region (3UTR) including the poly adenylation site (pA). On the 5' end of the donor sequence is the 3' splice site from *BTK* intron 1 to allow for efficient splicing of the transgenic sequence. **b)** Targeted integration via homology directed repair (HDR). The red rectangle represents the 17 bp of the protospacer that are PAM (protospacer adjacent motif) distal, while the red triangle represents the 3 PAM proximal bp and the PAM. When both pieces are together in the correct orientation (a red arrow), there is an intact target site matching the delivered single guide RNA (represented by scissors). A cDNA donor template flanked by 500 bp of sequence homologous to the region surrounding the double stranded break can be integrated via

the HDR pathway. **c)** Integration via precise integration into target chromosome (PITCh). A donor template was generated flanked by 20 bp of microhomology matching the regions adjacent to the target DSB. Beyond the microhomology regions were two Cas9 binding sites matching the target site in the genomic DNA. As the DSB is generated in the genomic DNA, the target sites in the donor template can also be cleaved. This cleaved donor template can integrate into the genomic DNA DSB via either microhomology mediated end joining (MMEJ) or non-homologous end joining (NHEJ). MMEJ integration is expected to leave a seamless repair while NHEJ integration would produce a duplicate of the microhomology regions and potentially a portion of the donor template's Cas9 binding regions. **d)** Integration via homology independent targeted integration. A donor template with no added homology arms was flanked by one or two Cas9 binding sites that are each the reverse complement of the sequence found in the genomic DNA. Upon delivery of the reagents, both the genomic DNA and the donor template will be cleaved and the resulting cleaved donor can be integrated via NHEJ. Without homology to guide the directionality of the integration, both the in frame and reverse oriented integration events are possible. However, if the donor is integrated in reverse, the Cas9 binding sites will be restored, barring insertions or deletions (indels) at the junction. The integrated donor may then be excised until it orients in the forward orientation or indels form at either junction, preventing re-cleavage.



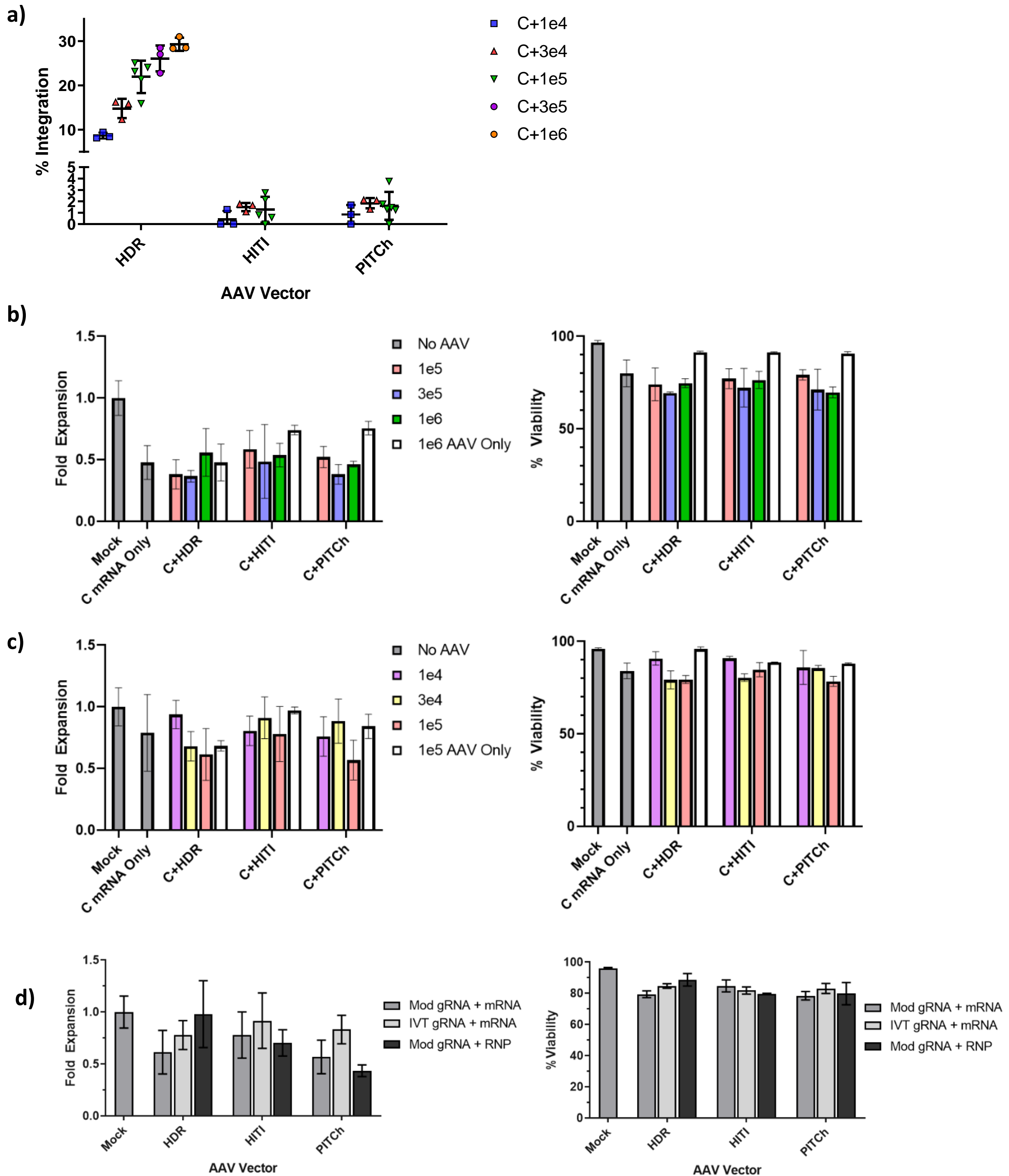
Supplemental Figure 2: K562 cells treated with only donor plasmid or only endonuclease produced no integration events

While Fig 2B demonstrates that K562 cell treatment with both a plasmid donor and the appropriate gRNA/Cas9 leads to substantial integration events of both the desired event and various undesired outcomes (reverse orientation or plasmid backbone integration), adding only gRNA/Cas9 expression plasmid or only the donor templates generates no detectable amplification events.



Supplemental Figure 3: Hydroxyurea cell cycle synchronization efficacy in K562 cells

a) Representative gating scheme for viability and single cells during cell cycle analysis. **b)** Hoechst dye analysis of cell cycle performed on K562 cells after incubation with increasing doses of hydroxyurea (HU).



Supplemental Figure 4: Optimizing reagents for delivery into human CD34+ mobilized peripheral blood stem and progenitor cells

a) Human CD34+ cells taken from Granulocyte colony-stimulating factor mobilized peripheral blood were electroporated with Cas9 mRNA and single guide RNA (sgRNA) targeting BTK intron 1 before transduction with donor template adeno-associated viral vectors (AAVs) at various multiplicities of infections (MOIs). Data from two experiments combined (higher MOIs vs lower MOIs). **b)** Viability and fold expansion measured 1-day post electroporation. Donor

templates AAVs were delivered at higher MOIs of 1e5, 3e5, or 1e6. **c)** Viability and fold expansion measured 1-day post electroporation. Donor templates AAVs were delivered at lower MOIs of 1e4, 3e4, or 1e5. **d)** Viability and fold expansion measured 1-day post electroporation from Fig 4A. Human PBSCs were treated with either *in vitro transcribed* (IVT) gRNA or chemically synthesized and modified (2'-O-methyl analogs and 3' phosphorothioate internucleotide linkages at each of the three 5' and 3' terminal RNA residues) gRNA. Cas9 was delivered either as mRNA or as RNP precomplexed with the respective gRNA. Cells were then transduced with donor AAV at an MOI of 1e5.

CHAPTER 4

DISCUSSION AND CLOSING REMARKS

Major Steps Towards a Cure for XLA

X-Linked Agammaglobulinemia (XLA) is in a somewhat difficult position for the development of new treatments. The current standard of care is protective enough to make any risks that come with a new treatment very hard to tolerate—a good problem to have. However, the treatment undeniably remains imperfect. The more permanent, more complete benefits that would come with a true cure would be a boon for patients in need of better options. Allogeneic hematopoietic stem cell (HSC) transplantation can and has been curative for patients with XLA, however the risks of graft vs host disease and graft rejection that come with the treatment generally outweigh the potential benefits under normal circumstances.[1,2]

Strategies that utilize autologous hematopoietic stem and progenitor cells (HSPCs) have the potential to provide the same benefit as allogeneic transplants but without the risks that come with an imperfect haplotype match. The big question for autologous therapies remains how best to replace or repair the non-functioning Bruton's Tyrosine Kinase (BTK) gene in patients. Earlier attempts to develop an XLA gene therapy utilized viral vectors to integrate *BTK* cDNA sequence into patient cells with promising results.[3,4] However, these viral vectors intrinsically carry a risk of insertional oncogenesis from the semi-random integration pattern that may disrupt tumor suppressing genes or activate proto-oncogenes by placing promoters or enhancers in their vicinity.[5,6] And while the current generation of lentiviral vectors have made substantial leaps in safety and have proven efficacious in clinical trials, some amount of risk remains. For diseases with particularly poor prognoses this risk may be acceptable, whereas it likely is not for XLA patients.

Another difficulty with viral vector mediated gene transfer is restoring the normal lineage specificity and expression levels of the gene.[7] In the case of BTK, this may be particularly important. The BTK protein is expressed in the majority of hematopoietic lineages, and past work

has demonstrated that BTK expression levels that are either too high or too low can lead to reduced signaling efficiency.[8] There is also a link between BTK overexpression and oncogenesis, though thus far it has not been demonstrated to be a causative relationship.

Targeted gene editing allows for integration of a potentially therapeutic *BTK* cDNA with very high precision, substantially reducing the risk of insertional oncogenesis. The development of the clustered regularly interspaced short palindromic repeats (CRISPR)/Cas9 gene editing platform has made it possible to evaluate more potential gene editing targets than ever before. One of the biggest hurdles of Cas9 mediated targeted insertion is the percentage of cells that can be successfully modified, especially for long term HSCs. And on that front, XLA is a particularly advantageous target. B lymphocyte progenitors with functional BTK protein have a strong selective advantage over cells with loss of function mutations.[9] Murine mixed chimerism experiments suggest that low single digit percentages of BTK containing HSCs may be sufficient for restoration of a robust antibody response to challenges.[9] That low bar for potential therapeutic efficacy can readily be achieved with today's gene editing technologies.

DONOR TEMPLATE OPTIMIZATION

Initially, the hypothesis was that addition of a corrective copy of the *BTK* cDNA sequence to the gene's endogenous locus would restore most of the normal regulatory control of the gene. However, this was shown to not be the case. While lineage specificity of BTK expression was preserved in all tested cell types, the quantities of both transgenic BTK mRNA and protein were only a fraction of those from the wildtype gene. While it remains unclear exactly what level of expression per edited cell is necessary for clinical benefit, BTK signals most effectively at concentrations near wildtype levels.[8] This prompted a search for ways to increase BTK expression

levels without further disrupting endogenous regulation. Some methods of increasing transgene expression, such as adding a strong promoter, would likely eliminate the safety benefits of targeted insertion. Ideally, expression would be increased only via re-addition of endogenous elements rather than including exogenous elements that could potentially dysregulate the gene.

Addition of the BTK Terminal Intron or the Woodchuck Hepatitis Virus Posttranscriptional Regulatory Element (WPRE) to the Donor Template Substantially Improved Transgene Expression

Previous work has shown that the terminal intron of genes can have an important role in transcription termination and poly adenylation.[10] Addition of introns to transgenes has also been linked to more efficient nuclear export of transcripts and higher protein expression.[11] While the two primary target sites discussed in this dissertation (*BTK* intron 1 and exon 2) both retain a functioning *BTK* intron 1 following targeted integration, it stood to reason that addition of more introns may further improve nuclear export. This hypothesis led to the creation of a donor variant that contains the *BTK* terminal intron between the cDNA sequence encoding *BTK* exons 18 and 19. However, the addition of the terminal intron 18 produced a donor variant that was too large to package into an adeno-associated virus (AAV) vector and would not be feasible to deliver to primary cells.[12] A search of existing literature on XLA and the *BTK* gene yielded no known enhancer sequences or pathogenic mutations within intron 18, suggesting that removing portions of the sequence may not impact BTK expression so long as the splicing signals were retained. A donor template with a truncated variant of the intron was created with only the ~300 bp from both the 5' and 3' ends of the intron. Both the full and the truncated terminal intron donors produced markedly improved expression in cell lines, more than doubling BTK expression levels. The full

intron led to 3.4 times wildtype expression, while the truncated intron produced 2.8 times wildtype protein.

While addition of the truncated intron 18 fragment led to improved expression at a packageable size, it still achieved less than 40% of wildtype BTK expression in K562 cells. Again, turning to the literature, we identified the WPRE as a commonly used element to improve mRNA and protein levels of transgenes.[13] A donor was constructed with the WPRE added immediately before the poly adenylation signal in the 3' untranslated region (UTR). This modification led to a 2.9-fold increase in expression over the base donor template in cell lines. Despite the success of the WPRE containing donors, it also raised some concerns. Adding elements that alter the normal BTK expression pattern necessitate extra caution to determine whether there are any adverse effects. Ultimately, murine models will be the optimal method to identify any such drawbacks.

Deletion of the PAM Sequence in the *BTK* Donor Template Prevented Formation of DNA Insertions and Deletions (Indels) in the Integration Junction

The cDNA sequence of each donor template was codon optimized for two main reasons. First, changing the codons to favor those with the highest transfer RNA (tRNA) abundance in human cells (without altering the amino acid sequence) is expected to increase rates of translation. Higher tRNA availability quickens translation of each protein, leading to more copies made per mRNA molecule before it is degraded.[14] Second, codon optimization of the sequence reduces the homology between the donor to be integrated and the gDNA sequence. This is important because without that distinction, the portion of the donor that is intended to be integrated into the cut site could serve as the homologous region directing the repair, leading to improper or incomplete integration events that would result in a non-functional protein.

An added benefit of codon optimization for exonic target sites is changing the codons can also disrupt the Cas9 binding site upon successful integration. The CRISPR/Cas9 platform is highly specific, so changing even a small number of base pairs can drastically reduce cleavage events. For intronic editing, there are no codons to optimize near the target site, so the guide RNA binding site may remain intact in the donor template. This cut site would then also be intact following a successful integration event. Restoring the target site post integration would allow the corrected gene to be re-cleaved, leading to the formation of indels at the integration junction. For *BTK* intron 1 gene editing, these indels were found in every integration junction sequenced. Should those indels interfere with splicing signals, they could lead to incorrect transcript processing and decrease overall BTK protein expression. Modifying the PAM sequence in the intronic portion of the donor template from NGG to NAA completely abrogated indel formation. The slight reduction in homology between the donor did not have any noticeable effect on integration rates—PAM modification led to an improved intronic donor with no identified drawbacks.

Combining Successful Modifications Produced a Donor Template that Integrated Efficiently and Expressed at Nearly Physiological Levels

The addition of the truncated terminal intron, WPRE, or the PAM mutation alone each led to improved donor expression or junction fidelity. The next step was to make donors featuring multiple of these improvements together. The simultaneous addition of all three modifications generated 80.4% as much BTK protein per cell as wildtype cells, the closest to physiological levels seen thus far. However, the presence of both the truncated intron and WPRE led to a donor template that was too large to be packaged into an AAV vector. This led to the creation of a further truncated intron fragment, dubbed “micro” intron 18. This new combination (ΔPAM, micro intron,

WPRE) produced 74.6% of wildtype BTK expression at a size packageable into a viral vector for efficient delivery into human primary cells.

Addition of the WPRE Element Led to Elevated BTK Expression in Jurkat T Cells

One possible pitfall to be cognizant of for XLA mediated gene therapy is aberrant expression in non-physiological cell lineages. It is unclear whether BTK expression in abnormal lineages would have detrimental effects, however, the correlation between BTK overexpression and oncogenesis is cause to be wary. Initial lineage specificity studies were performed in Jurkat cells, a T cell line that does not express BTK. In these cells, integration of the base donor with no modifications and the donor with the truncated intron both led to no detectable BTK expression. However, integration events from both donors variants that contained the WPRE sequence led to strong expression of BTK in Jurkat cells. This finding does not preclude the use of the WPRE sequence; but it does warrant further study to ensure the safety of the treatment.

HDR Narrowly Outperformed Two Alternative Methods to Achieve Targeted Integration of the BTK Transgene

Homology directed repair (HDR) has historically been the optimal method to achieve successful targeted integration events. However, recent publications have identified HDR rates are often lower in quiescent HSCs than the more rapidly dividing progenitors, a phenomenon thought to be caused by cell cycle dependent changes in DNA repair pathway utilization.[15–18] Interestingly, these findings are not universal—some groups do not see this discrepancy between editing in stem cells and progenitors at certain target sites. Recently, alternative methods to achieve endonuclease mediated targeted integration have been described that do not rely on HDR. In theory, these

pathways may be advantageous for HSC editing due to the ability to function in quiescent cells.[19,20] We compared HDR to two of these new methods, Homology Independent Targeted Integration (HITI) and Precise Integration into Target Chromosome (PITCh) for successful BTK integration into the intron 1 target site in both cell lines and primary HSPCs.[21,22] As expected, HDR led to the strongest integration rates in both K562 cells and unsorted HSPCs, both of which feature large numbers of dividing cells. For a deeper comparison, HSPCs were sorted into stem and progenitor populations following electroporation/transduction. In sorted HSCs, the most clinically relevant cell population, the integration rates of all three donors were comparable, suggesting any of the three could lead to a successful treatment. HDR retained one major advantage: with the PAM sequence disrupted, 100% of sequenced junctions were base perfect while the other methods produced indels in the junction. Because of that, HDR remained the preferred pathway for BTK targeted integration.

Combining Chemically Modified sgRNA and Cas9 mRNA with AAV6 Donor Vectors Produced High Rates of Targeted Integration into BTK Intron 1

While cell lines are invaluable tools for development of new gene editing reagents, one of their major drawbacks is that delivery of the reagents behaves differently than in primary HSPCs. The cell lines used in these studies are very durable and replicate quickly, which is a stark contrast to the fragile, largely quiescent, clinically relevant HSCs. The high toxicity of plasmids makes them an inefficient method of delivery into HSCs. Instead, donor templates are generally delivered as AAV6 vectors and the nuclease is electroporated as RNP. Surprisingly, using these conditions with reagents targeting the BTK intron 1 site, targeted integration was found in only 1.7% of cells, which is likely insufficient for clinical benefit if used as a cell product. Chemically modifying the

gRNA to protect it from degradation increased the integration rate to 5.7%.[23] Finally, delivering the Cas9 as mRNA rather than RNP provided a substantial boost to integration up to 24.7% of human PBSCs, which is expected to be more than enough to provide a clinical benefit for patients. Finding such a stark increase in integration with Cas9 mRNA is somewhat unique; to the best of our knowledge, similar results have not been reported for any other target sites and are not found with any other target site tested by our lab. While these optimized conditions are sufficient to advance into murine models and towards the clinic, concerns about the availability of clinical grade Cas9 mRNA prompted a search for alternative cut sites that function as effectively with RNP, which is the industry standard.

An Alternative Target Site for Integration in BTK Exon 2 Allowed for High Rates of Integration in Human HSPCs using Cas9 Protein with a Similar Expression as Donors Integrating in the Intron 1 Site

An additional set of target sites through BTK intron 1 and exon 2 were assessed for their ability to induce allelic disruption as RNP in human HSPCs, which identified a high performing gRNA targeting exon 2. The lessons learned from the intron 1 donor development were applied to the exon 2 target site. Three donor constructs were constructed for tests in cell lines: a base cDNA donor with no modifications, one featuring the truncated intron 18, and a third with both the truncated intron and the WPRE. These three donors yielded BTK protein expression at rates comparable to their respective intron 1 counterparts. Off-target analysis for this new target site was performed and identified substantial aberrant cleavage events. However, replacing the wildtype Cas9 protein with an engineered, high-fidelity variant nearly eliminated off target cutting. The I18t-WPRE donor for exon 2 was packaged as an AAV and tested for integration into human

HSPCs. The donor integrated in 14.4% of cells when delivered with high fidelity RNP; likely sufficient to provide clinical benefits. The reagents and conditions established here produce nearly physiological levels of BTK expression and successfully modify cells at high enough rates to be potentially therapeutic.

Next Steps: Continuing Down the Path Towards Benefitting Patients

The work outlined here provides the basis for a novel, potentially curative treatment for XLA. Despite unexpected hurdles in achieving sufficient editing rates and transgenic expression, these challenges have been surmounted and produced a treatment scheme that is ready for the next steps towards achieving clinical relevance. While bringing a new therapy from inception to animal models is a substantial step, there remain more hurdles to clear.

Murine models will be imperative; they will allow a more directly translatable model of the potential therapy. Severely immunodeficient murine models will allow evaluation of the human cell products via xenograft. It will be essential to test whether the edited cells retain engraftment potential, produce any abnormal proliferation, and can successfully advance past the early stages of B cell development. However, human B cells will not develop fully in this model, limiting the amount of functional studies that can be performed. Ideally, XLA patient mobilized peripheral blood stem cells will be utilized for murine xenograft studies, though healthy donor cells are an acceptable substitute if they cannot be obtained.

For functional analysis, Btk/Tek double knockout mice have a disease phenotype very similar to XLA. Successful integration of a donor human BTK donor template into the murine Btk locus would restore B cell development in the model due to the very high homology between the two proteins. Mice that receive the edited cells could be analyzed for engraftment potential, editing

frequencies in different lineages, and antibody responses posts treatment. The drawback of this model is that new murine reagents and delivery scheme will need to be optimized, modelling the human cell product as closely as possible. Together with the severely immunodeficient murine model, they will provide a very robust set of data to determine the potential efficacy of a human treatment.

The lessons learned here provide great promise as a novel curative treatment for XLA patients in need. Though there remains work to be done, these *in vitro* studies provide an optimistic outlook for the success of BTK gene editing. Additionally, the modifications to improve donor template design likely have more broad applicability towards gene therapy for other diseases. Addition of WPRE and a terminal intron or intron fragment to donor templates provided a robust increase in gene expression that could benefit a multitude of new treatments and help bring new gene therapies to patients worldwide.

REFERENCES

1. Abu-Arja, RF, Chernin, LR, Abusin, G, Auletta, J, Cabral, L, Egler, R, *et al.* (2015). Successful hematopoietic cell transplantation in a patient with X-linked agammaglobulinemia and acute myeloid leukemia. *Pediatr. Blood Cancer*doi:10.1002/pbc.25554.

2. Ikegame, K, Imai, K, Yamashita, M, Hoshino, A, Kanegane, H, Morio, T, *et al.* (2016). Allogeneic stem cell transplantation for X-linked agammaglobulinemia using reduced intensity conditioning as a model of the reconstitution of humoral immunity. *J. Hematol. Oncol.*doi:10.1186/s13045-016-0240-y.

3. Clough, C, Wang, Y, Khan, IF, Singh, S, Hung, K and Rawlings, DJ (2016). 132. Targeting the BTK Locus in Primary Human Hematopoietic Cells with TALENs and AAV Donor Template. *Mol. Ther.*doi:10.1016/s1525-0016(16)32941-0.

4. Ng, YY, Baert, MRM, Pike-Overzet, K, Rodijk, M, Brugman, MH, Schambach, A, *et al.* (2010). Correction of B-cell development in Btk-deficient mice using lentiviral vectors with codon-optimized human BTK. *Leukemia*doi:10.1038/leu.2010.140.

5. Nienhuis, AW, Dunbar, CE and Sorrentino, BP (2006). Genotoxicity of Retroviral Integration In Hematopoietic Cells. *Mol. Ther.*doi:10.1016/j.ymthe.2006.03.001.

6. Zhou, S, Fatima, S, Ma, Z, Wang, YD, Lu, T, Janke, LJ, *et al.* (2016). Evaluating the Safety of Retroviral Vectors Based on Insertional Oncogene Activation and Blocked Differentiation in Cultured Thymocytes. *Mol. Ther.*doi:10.1038/mt.2016.55.

7. Sather, BD, Ryu, BY, Stirling, B V, Garibov, M, Kerns, HM, Humblet-Baron, S, *et al.* (2011). Development of B-lineage predominant lentiviral vectors for use in genetic therapies for B cell disorders. *Mol. Ther.* **19**: 515–525.

8. Satterthwaite, AB, Cheroutre, H, Khan, WN, Sideras, P and Witte, ON (1997). Btk dosage determines sensitivity to B cell antigen receptor cross-linking. *Proc. Natl. Acad. Sci. U. S. A.* **94**: 13152–13157.

9. Rohrer, J and Conley, ME (1999). Correction of X-linked immunodeficient mice by competitive reconstitution with limiting numbers of normal bone marrow cells. *Blood*doi:10.1182/blood.v94.10.3358.422k04_3358_3365.

10. Dye, MJ and Proudfoot, NJ (1999). Terminal exon definition occurs cotranscriptionally and promotes termination of RNA polymerase II. *Mol. Cell*doi:10.1016/S1097-2765(00)80464-5.

11. Choi, T, Huang, M, Gorman, C and Jaenisch, R (1991). A generic intron increases gene expression in transgenic mice. *Mol. Cell. Biol.*doi:10.1128/mcb.11.6.3070.

12. Grieger, JC and Samulski, RJ (2005). Packaging Capacity of Adeno-Associated Virus Serotypes: Impact of Larger Genomes on Infectivity and Postentry Steps. *J. Virol.*doi:10.1128/jvi.79.15.9933-9944.2005.

13. Zufferey, R, Donello, JE, Trono, D and Hope, TJ (1999). Woodchuck Hepatitis Virus Posttranscriptional Regulatory Element Enhances Expression of Transgenes Delivered by Retroviral Vectors. *J. Virol.*doi:10.1128/jvi.73.4.2886-2892.1999.

14. Fath, S, Bauer, AP, Liss, M, Spriestersbach, A, Maertens, B, Hahn, P, *et al.* (2011). Multiparameter RNA and codon optimization: A standardized tool to assess and enhance autologous mammalian gene expression. *PLoS One*doi:10.1371/journal.pone.0017596.

15. De Ravin, SS, Reik, A, Liu, PQ, Li, L, Wu, X, Su, L, *et al.* (2016). Targeted gene addition in human CD34 + hematopoietic cells for correction of X-linked chronic granulomatous disease. *Nat. Biotechnol.*doi:10.1038/nbt.3513.

16. Genovese, P, Schiroli, G, Escobar, G, Di Tomaso, T, Firrito, C, Calabria, A, *et al.* (2014). Targeted genome editing in human repopulating haematopoietic stem cells. *Nature*doi:10.1038/nature13420.

17. Schiroli, G, Genovese, P, Castiello, MC, Capo, V, Gregory, PD, Holmes, MC, *et al.* (2016). 481. Targeted Genome Editing in Mouse Hematopoietic Stem/Progenitor Cells (HSPC) To Model Gene Correction of SCID-X1. *Mol. Ther.* **23**: S191.

18. Branzei, D and Foiani, M (2008). Regulation of DNA repair throughout the cell cycle. *Nat. Rev. Mol. Cell Biol.*doi:10.1038/nrm2351.

19. Suzuki, K, Tsunekawa, Y, Hernandez-Benitez, R, Wu, J, Zhu, J, Kim, EJ, *et al.* (2016). In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration. *Nature* **540**: 144–149.

20. Sakuma, T, Nakade, S, Sakane, Y, Suzuki, KIT and Yamamoto, T (2016). MMEJ-Assisted gene knock-in using TALENs and CRISPR-Cas9 with the PITCh systems. *Nat. Protoc.*doi:10.1038/nprot.2015.140.

21. Suzuki, K, Tsunekawa, Y, Hernandez-Benitez, R, Wu, J, Zhu, J, Kim, EJ, *et al.* (2016). In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration. *Nature*doi:10.1038/nature20565.

22. Nakade, S, Tsubota, T, Sakane, Y, Kume, S, Sakamoto, N, Obara, M, *et al.* (2014). Microhomology-mediated end-joining-dependent integration of donor DNA in cells and animals using TALENs and CRISPR/Cas9. *Nat. Commun.* **5**.

23. Hendel, A, Bak, RO, Clark, JT, Kennedy, AB, Ryan, DE, Roy, S, *et al.* (2015). Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells. *Nat. Biotechnol.*doi:10.1038/nbt.3290.