Master's Thesis of Science in Agriculture

New Platform for Gene Therapy using *In Vivo* Knock-In into Safe Harbor Locus

안전항구 위치에 생체 내 넉인을 이용한 새로운 유전자치료 플렛폼

August 2020

Se Jun Ahn

Department of International Agricultural Technology

Graduate School of International Agricultural Technology

Seoul National University

New Platform for Gene Therapy using *In Vivo* Knock-In into Safe Harbor Locus

A thesis
submitted in partial fulfillment of the requirements to the faculty
of Graduate School of International Agricultural Technology
for the Degree of Master of Science in Agriculture

By
Se Jun Ahn

Supervised by
Prof. Su Cheong Yeom

Major of International Agricultural Technology
Department of International Agricultural Technology
Graduate School of International Agricultural Technology
Seoul National University

August 2020

Approved as a qualified thesis
for the Degree of Master of Science in Agriculture
by the committee members

Chairman	**Tae Min Kim, Ph.D.**	
Member	**Joonghoon Park, Ph.D.**	
Member	**Su Cheong Yeom, Ph.D.**	

Abstract

As part of efforts to treat the genetic disorder, there have been many advances in the field of *in vivo* genome editing. However, the low editing efficiency due to technical problems acts as a limitation directly repairing the mutant gene. Thus, I tried to develop a novel gene-editing strategy for curing genetic disorders using knock-in (KI) into safe harbor locus, which exhibited high expression in the liver. Based on the microarray experiment using human hepatocyte, the Apoc3 gene was selected as the locus for KI. A portion of the human Apoc3 gene, which was harboring highly efficient sgRNA binding site of spCas9 and cjCas9, was used to generate genetically human-mimicking animals. B6.Apoc3-hApoc3 KI mice exhibited high gene expression in the liver (300 folds more elevated than the F9 gene) and no alternative splicing. Besides, hemophilia B was selected as a genetic disease to be treated using *in vivo* gene editing. The hemophilia B model mouse is generated through the deletion of the coagulation factor 9 gene using CRISPR/Cas9 and had a phenotype of clotting disorder similar to hemophilia B patients. After mouse production, three strategies for *in vivo* therapeutic gene KI into Apoc3 were applied using adeno-associated virus packing CRISPR/cas9 and donor template. First, the homologous recombination strategy was founded to be a failure due to donor auto-expression and extremely low KI efficiency. However, NHEJ mediated KI (HITI or AAV-mediated integration) presented human F9 concentration (about 50, 250ng/ml at six weeks after injection) without donor's auto-expression. In conclusion,

the model mice for *in vivo* genome editings such as mApoc3-hApoc3 KI and F9 KO were produced. In the attempts of *in vivo* genome editing, some indicators presenting the low KI efficiency were found through PCR and ELISA. To increase KI efficiency, a high dose experiment is on-going.

Keywords: Adeno-associated virus, CRISPR/cas9, Genetic disorder, *in vivo* genome editing, safe harbor gene

Student number : 2018-26740

Contents

List of Tables

List of Figures

List of Abbreviations

AA	amino acid
AAV	Adeno associated virus
Ad	Adenovirus
Amp	ampicillin
Apoc3	Apolipoprotein C3
aPTT	activated partial thromboplastin time
B6	C57BL/6
B6.WT	C57BL/6.wild type
bp	base pair
C+D	CRISPR+donor
Cas9	CRISPR associated protein9
CBC	complete blood count
cDNA	complementary Deoxy Nucleic Acid
Chr	chromosome
CMV Promoter	Cytomegalovirus promoter
CRISPR	Clustered regularly interspaced short palindromic repeats
crRNA	CRISPR RNA
D	donor
DMEM	Dulbecco modified eagle medium
DNA	Deoxy Nucleic Acid
DSBs	Double Strand Breaks
dsDNA	double strand DNA
eGFP	enhanced Green Fluorescent Protein
ELISA	Enzyme-Linked ImmunoSorbent Assay
EMH	Extramedullary hematopoiesis

En2Sa	Engrailed 2 gene Splice acceptor
EP	Electroporation
F9	Coagulation Factor 9
FRKM	fragments per kilobase of exon model per million reads mapped
GBA	Glucosidase-beta-acid
GC/ml	Genome copies per mL
GLA	Galactosidase-alpha
GOI	Gene of Interest
H&E stain	Hematoxylin and eosin stain
HA	Homology Arm
hApoc3	human Apolipoprotein C3
hCG	Human Chorionic Gonadotropin
Hct	Hematocrit
HDR	homology-directed repair
HITI	Homology-independent targeted integration
HR	Homologous Recombination
hseq	human Sequence
IDS	Iduronate-2-sulfatase
IDUA	Iduronidase-alpha-L
ITR	inverted terminal repeats
IU	International Unit
KI	Knock In
KO	Knock Out
MPLT	Manual Platelet Count
mRNA	messenger RNA
NHEJ	non-homologous end-joining
No.	Numder
PAM	Protospacer adjacent motif

PCR	Polymerase Chain Reaction
PMSG	Pregnant Mare's Serum Gonadotropin
PT	prothrombin time
RBC	red blood cell
RNA	Ribo Nucleic Acid
RNP	ribonucleoprotein
SEM	standard error of the mean
sgRNA	single guide RNA
spCas9	*Streptococcus pyogenes* Cas9
ssODN	single-stranded oligodeoxynucleotides
T7E1	T7 Endonuclease I
TALEN	transcription activator-like effector nuclease
TBG promoter	Human thyroxine binding globulin promoter
tracrRNA	trans acting crRNA
VLDL	Very-low-density lipoprotein
WT	Wild Type
ZFN	zinc-finger nuclease

1. Literature review

1.1 Genetic disorder

Genes contain information about proteins necessary for life support, and because they have individual genetic differences, they affect characteristics such as appearance and health. Genetic changes can induce expression levels and structural changes in the encoded protein. Genetic disease refers to the phenotype of health abnormalities and diseases caused by these changes. In genetic diseases, it is common to inherit the traits of parents, or there may be mutations in genes due to environmental and internal factors. Thanks to the development of sequencing technology, various genetic diseases have been discovered, and there are more than 7000 types [1]. Not all genetic diseases are fatal to health, and These disorders are generally known to occur in less than 1 in 2000 people, about 1 in 50 of whom require special care [2].

Genetic disorders include chromosomal disorders caused by structural and numerical abnormalities of chromosomes, multifactorial disorders caused by environmental factors along with complex genes, and single genes caused by mutations in one gene.

Monogeneic disorder refers to functional abnormalities and deficiencies in proteins caused by mutations in specific genes that indicate symptoms of the disease. This protein deficiency causes certain cascades to stop, causing problems such as accumulation of intermediates and delayed synthesis of substances. In particular, this single genetic disorder is known to cause various

types of diseases in the liver expressing numerous proteins (Table 1).

In humans, it consists of a total of 23 pairs of chromosomes, including 22 pairs of autosomes and 1 pair of sex chromosomes. In addition, genetic diseases are divided into recessive and dominant according to the characteristics of the gene, and the disease occurs according to the number of copies of the mutated gene (Figure 1).

Table.1 Type of protein deficiency in liver

Single- gene	Causing Gene	Chromosome	Symptoms	Incidence rate
Hemophilia A	FVIII	X	blood Coagulation disorder	5000:01:00
Hemophilia B	FIX	X	blood Coagulation disorder	20000 : 1
Fabry disease	GLA	X	Lysosomal storage disease	40000 : 1
Gaucher's disease	GBA	1	Lysosomal storage disease	100000-50000 : 1
Hurler syndrome	IDUA	X	Lysosomal storage disease	100000-50000 : 1
Hunter syndrome	IDS	X	Lysosomal storage disease	100000 : 1

FVIII : Coagulation Factor 8, FIX : Coagulation Factor 9, GLA : Alpha-galactosidase, GLA : Beta-alactosidase, IDUA : Alpha-L-iduronidase, IDS : Iduronidase 2-sulfatase

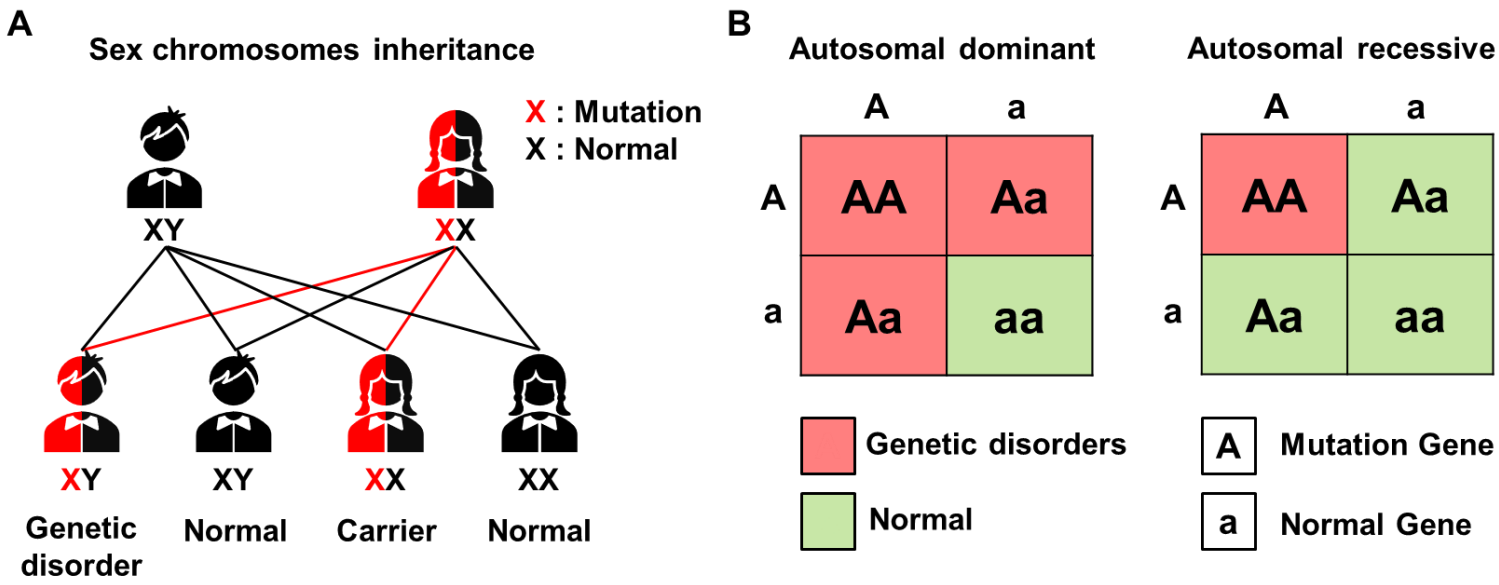


Figure 1. Chromosomal inheritance A) This is an example of X-linked recessive genetic disorder, a type of sex chromosomal inheritance. Carrier refers to a person who has a mutant gene but does not show the phenotype of the disease. B) This is a schematic showing the autosomal inheritance. The genetic disorder caused by one mutant gene is called dominant, and what happens when two are called recessive.

1.2 Treatment for single-gene disorder

The current treatment of genetic diseases used in the market is aimed at alleviating symptoms. However, because the current treatment is not a method for treating the mutant gene, there is a half-life depending on the blood concentration of the treatment. Since the half-life of blood products and recombinant proteins is about 2-3 days, periodic administration is necessary, which places an economic burden on patients. In the United States, the economic burden of hemophiliacs is known to be around 400,000 USD per year[3]. These high costs for therapy indicate the need to develop methods of treatment with a longer half-life.

1.2.1 Plasma derivatives

This is the earliest method of protein deficiency, followed by the separation of plasma from normal blood and infusion into the patient. However, due to the nature of blood products, it is true that they are vulnerable to human immunodeficiency virus and hepatitis C virus infections. Indeed, in the 1980s, large-scale infections occurred in patients injected with contaminated blood products. Because of this, despite much effort to inactivate the virus, the potential risk to new infectious agents has led to the use of this agent.

1.2.2 Protein therapy

The most common way to treat monogenetic disorders is by injecting a therapeutic protein. The difference between plasma-derived factors and

protein-based treatments is that proteins are derived from animal cells and microorganisms, not humans. The advantage of the recombinant protein developed in 1997 is that it can predict the infectious factors and stability. However, protein therapy has a short half-life of about 1 to 3 days [4, 5]. In order to increase the half-life, attempts are made to improve protein residue changes and formulations, but the current situation is to have a holding power of about 3 to 7 days without dramatic progress [6, 7]. This short half-life not only causes the patient to have an economic burden caused by continuous injection, but also has a problem that the effect is gradually reduced by making the protein resistant.

1.2.3 Gene transfer

Gene transfer refers to the introduction of a gene of interesting (GOI) into a cell, and a template consisting of a promoter and a therapeutic gene produces a therapeutic protein. The vector used for gene introduction determines the efficiency and duration. Various vector features are listed in Table 2.

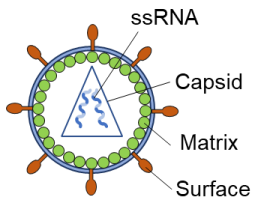
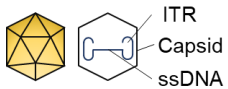
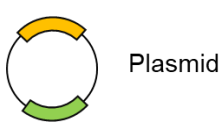
Adeno associated virus was discovered as a contaminant in the adeno virus preparation in 1965 and is one of the very small viruses with a capsid of about 22 nm [8]. This virus does not cause disease because it cannot replicate on its own without helper virus. In addition, the AAV has a genome of approximately 4.8 kb, a linear single-stranded DNA that includes reversed terminal repeats (ITR) and replication genes (Rep) and capsid genes (Cap) at both ends [9]. Interestingly, various serotypes of AAV have a tropism that binds

to a specific receptor, and it is possible to target a specific organ using these properties.

Recombinant AAV (rAAV) is a method designed to apply AAV to therapeutic agents using these features. Because cloning and capsid are generated based on ITR, the desired gene can be cloned and AAV packed by replacing Rep and Cap Gene with the gene of interest (GOI). Instead, a plasmid of Helper virus and Rep/Cap Gene are cotransfected into HEK293, making replication and packaging possible. After separation and purification, virus particles packed with the desired gene in AAV can be obtained.

Gene introduction using AAV begins with binding to a receptor. AAV, which has entered the cell through endocytosis, is released into the cytoplasm through a golgi-mediated capsid treatment. The virus particles are then introduced into the nucleus by the nuclear pore complex (NPC), releasing ssDNA. It is then converted to dsDNA through annealing, producing mRNA and protein using RNA polymerase from the host. This exogenous expression occurs without integration on the host genome and is lost in the process of dividing and replicating cells, gradually diluting and disappearing [9](Figure 2). This causes AAV gene therapy to have a half-life of about 6 months to 1 year[10]. Overall, treatments that use transient expression of AAV have succeeded in sufficiently increasing the half-life.

Table 2. Various kinds of vector

Vector	Structure	Genetic material	Immune response	Introduction efficiency	Feature
Lentivirus		RNA	Low	Very high	Host genome Integration Packing size : 9.4kb
Adeno-associated virus		ssDNA	Very low	High	Specific Integration(AAVS1) Packing size : 4.8kb
Naked plasmid		Plasmid(dsDNA)	-	Extremly low	Plasmid only
Lipofectamine		Genetic material	-	Moderate	Form of micelle

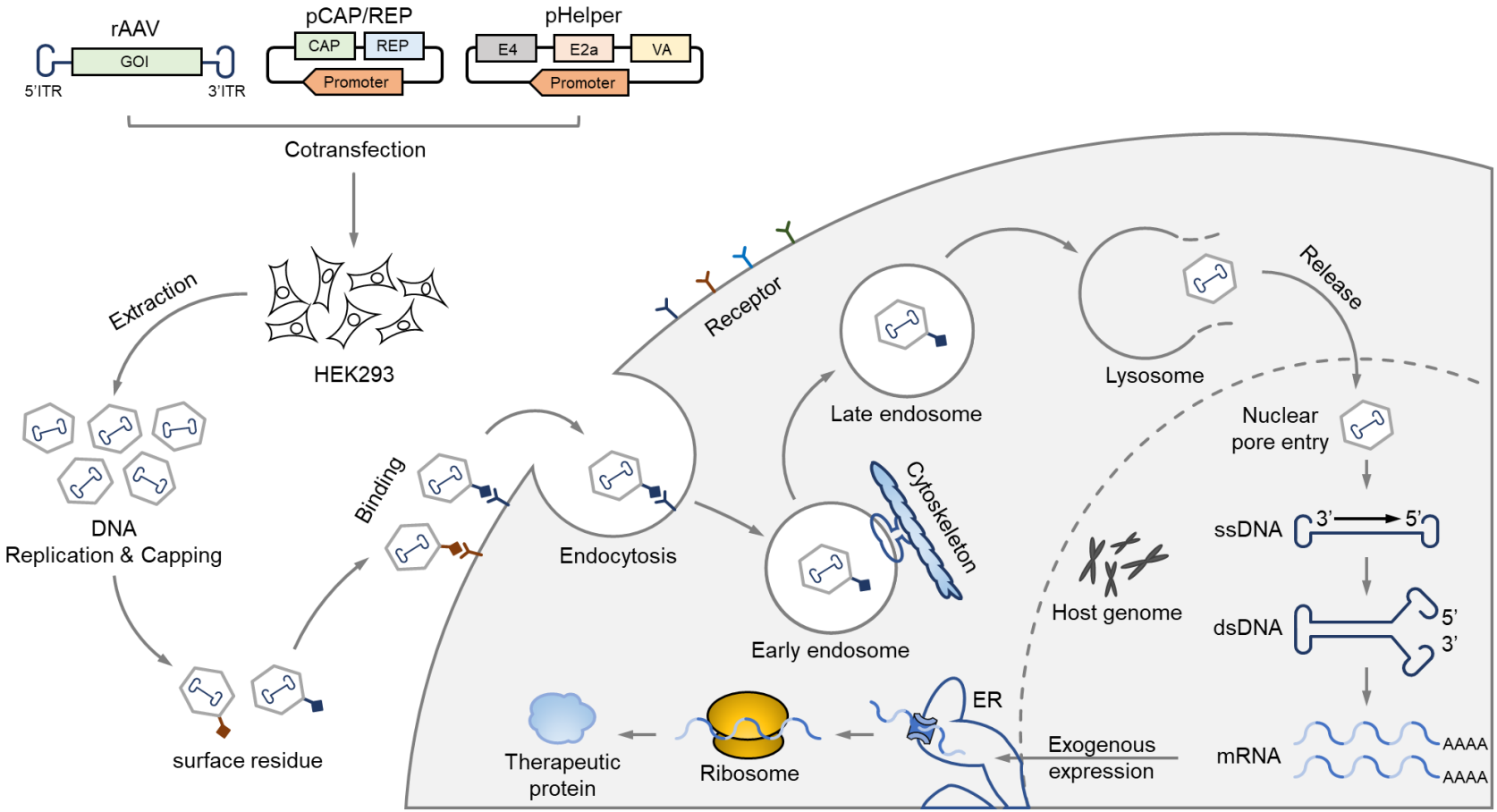

rAAV
GOI
5'ITR
3'ITR
pCAP/REP
CAP
REP
Promoter
pHelper
E4
E2a
VA
Promoter
Cotransfection
Extraction
HEK293
DNA
Replication & Capping
surface residue
Binding
Endocytosis
Receptor
Late endosome
Lysosome
Release
Nuclear
pore entry
Cytoskeleton
Early endosome
Host genome
ssDNA
3'
5'
dsDNA
5'
3'
mRNA
AAAA
AAAA
ER
Exogenous
expression
Ribosome
Therapeutic
protein

Figure 2. Synthesis and transduction process of adeno associated virus. ITR: includes reversed terminal repeats. GOI: gene of interesting. CAP: capsid gene. Rep: replication. HEK293: human embryonic kidney 293 cells. ER: endoplasmic reticulum.

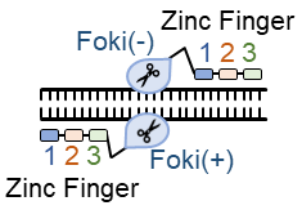
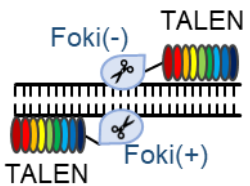
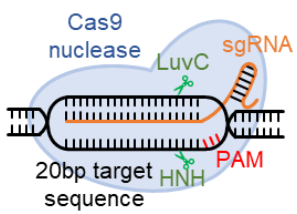
1.3. genome editing

Despite efforts to increase the therapeutic half-life, these treatments are not a permanent solution because they are not a way to repair genetic mutations in the host genome. To directly edit the mutant gene, we needed a technique that could be edited at the desired sequence. Programmable nucleases, including zinc-finger nuclease (ZFN), transcription activator-like effector nuclease (TALEN), and clustered regularly interspaced short palindromic repeats (CRISPR), made it possible to edit in the desired position. The characteristics of each programmable nucleases are listed in Table 3.

Gene editing begins with the creation of a double-strand break (DSB), which can make insertion and deletion of sequences generated during the DNA repairing process. The gene repair system is divided into 1) homologous recombination and 2) non-homologous end binding pathway. 1) Homologous recombination occurs predominantly in the S and G2 phases rich in sister chromosomes before entering the mitotic state (M state) and is known as an error-free recovery method. Destruction sites are recognized by the MRE11-RAD51-NBS1 complex and Ataxia telangiectasia mutated Kinase. Afterwards, CtlP nuclease processes the end of the site where DSB has occurred to form it in the form of ssDNA. This ssDNA is coated with replication protein A, and RAD51 undergoes invasion and DNA synthesis to homologous sequencing on sister chromosomes. Finally, recovery is completed through Ligation. 2) On the other hand, in the case of NHEJ, it is characterized by non-dependent recovery in the cell cycle and is an error-prone recovery method because it is

not affected by the replication template. It is recognized by the Ku70/80 and DNA-PKcs complex, and the ends are trimmed by Mre11 and AREMIS. Unlike HR, invasion does not occur, and it is bound by the DNA ligase IV complex composed of DNA ligase IV and cofactor XRCC4 [11](Figure 3)

Table 3. Programmable nucleases

Nuclease	Structure	Feature
ZFN		ZFN molecule capable of targeting three nucleotide sequences. A total of three ZFN molecules track specific sequences. Double strand break: 2 Foki nucleases. Truncated end: sticky end. difficult and expensive to synthesize and use.
TALEN		TALEN molecule capable of targeting one nucleotide sequence. Track specific sequences with a total of 9 or more TALEN molecules. Double strand break: 2 Foki nucleases. Truncated end: sticky end. Ease of access to synthesis and use.
CRISPR		Sequence tracking using a RNA. Approximately 20 nucleotide sequences can be traced. 2-6 base pair PAM Double strand break: LuvC and NHN domains of Cas9 protein. Truncated end: blunt end. Easy to synthesize and use, low cost

ZFN: zinc finger, TALEN: transcription activator-like effector nuclease, CRISPR: clustered regularly interspaced short palindromic. PAM: protospacer adjacent motif

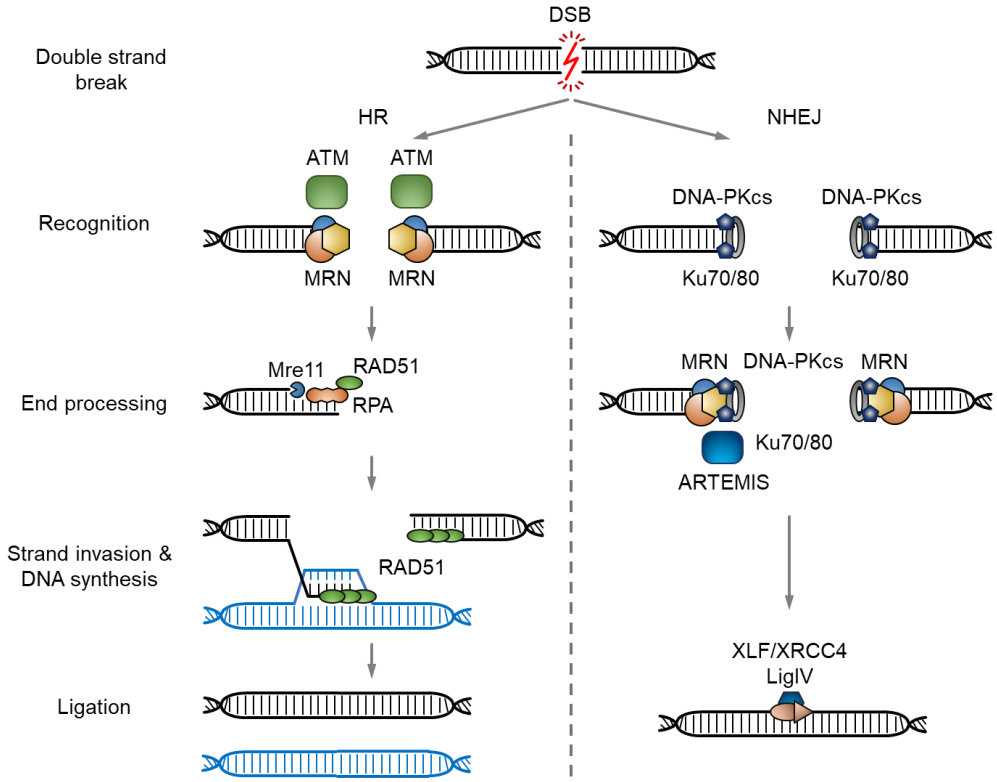


Figure 3. The process of repairing double strand breaks. To repair double strand breaks, there are pathways of homology recombination and non-homologous end joining. Subsequently, it is recovered through the recognition of the breakdown site, end processing, strand invasion & DNA synthesis, and ligation. The blue gene represents the sister chromatid strand. ATM: Ataxia telangiectasia mutated (Serine/Threonine Kinase). MRN: MRE11-RAD51-NBS1 complex **.**

1.3.1 Ex vivo genome editing

Application methods using gene-editing technology are divided into *ex vivo* and *in vivo*. Ex vivo gene editing refers to a method of genetically editing cells derived from a patient in vitro and re-transplanting them into a patient. Currently, the most commonly used ex vivo gene editing target is induced pluripotent stem cells (iPSc), which have differentiated epithelial cells derived from patients. The iPSc from a patient is corrected to normal cells through gene editing and has the advantage of being selectable after editing. Also, it does not cause an autoimmune reaction as it is a cell derived from a patient. On the other hand, iPSc, a type of stem cell, has the potential to develop into cancer.

1.3.2 *in vivo* genome editing

In vivo uses a method of injecting a gene editing tool directly into the body. This method includes the possibility of causing side effects such as off-target effects and vector immune responses. However, it has the advantage that the procedure of the procedure is simple because there is no surgical and implantation process. In addition, the editing of cells with excellent division ability, such as hepatocytes, has the possibility that cells corrected through cell division will continue to survive. It is known that the method of directly correcting a mutated gene using gene editing *in vivo* is affected by the efficiency of transduction, indel, knock in, and less than 1% of cells are corrected. However, there are currently technical limitations to raising this low efficiency.

2. Introduction

Single genetic diseases are caused by a defect in a gene encoding a specific protein and a protein's functional abnormality due to a gain of function or a loss of function. These single-gene disorders are generally known to occur in less than 1 in 2000 people, about 1 in 50 of whom require special care[2]. Not all of these genetic diseases are linked to fatal, but they have a significant impact on the life of the patient. Furthermore, some genetic diseases have a severe effect on lifespan[12-15].

Currently, recombinant proteins have been used for single-gene disorders with loss of function treatment. Treatments with recombinants are less likely to cause side effects, and adequate replacement treatment is possible[5]. However, there are still limitations, such as short duration and high cost[16]. Despite various studies for increasing the half-life were conducted[6], but there was a little dramatic effect. AAVs were considered as the right candidate in developing long-term stable therapeutic gene expression. Note, AAV was approved for gene therapy tools from the FDA and is highlighted as a new strategy[9]. This treatment has the advantage of long-term expression, low immune response, and high transduction efficiency due to the nature of AAV[17]. Despite these advantages, there is a problem that the treatment effect decreases over time[7, 17].

For overcoming those limitations, genome editing based therapy is suggested as an alternative. Genome editing utilizes programmable nucleases, including zinc-finger nuclease (ZFN), transcription activator-like effector nuclease (TALEN), and clustered regularly interspaced short palindromic repeats (CRISPR) to develop a double-strand break (DSB) at the target site. DSB is repaired via two different

pathways, such as non-homologous end-joining (NHEJ) or homology-directed repair (HDR). Also, DSB can be utilized for developing knock-in on the presence of a donor template[18, 19]. In gene therapy, there are two primary strategies of *in vivo* and ex vivo approach with the direct repair of mutant sequence or therapeutic gene insertion. Notably, Host genome editing offers the potential for patients with genetic disorders to get potentially permanent treatment. However, genome editing efficiency is extremely low, especially in the *in vivo* trial; thus, a new strategy overcomes this is needed[20-23].

Recently, many groups tried to find tools for efficient *in vivo* therapeutic gene insertion into specific target locus. Here I present the novel approach for curing genetic disorder with therapeutic gene knock-in. First, the target locus for gene insertion was selected using a microarray from human hepatocyte and developed an animal model using the CRISPR/Cas9 system. Moreover, the human Factor IX gene was chosen as a candidate therapeutic gene for *in vivo* insertion.

3. Materials and methods

3.1 Animal

C57BL/6 (B6) mouse was supplied by Koatech (Pyeongteak, Korea) and was used to construct mouse models through transformation. One cage contains 4-6 mice, and newborn mice were weaned after four weeks. All mice have maintained in individually ventilated cages and were provided with food and water *ad libitum*. The study was approved by the Institutional Animal Care and Use Committees of Seoul National University (s-170816-5 and SNU-170816-6). Animal experiments were conducted according to recommended guidelines.

3.2 Preparation of sgRNA and ssODN

single guide RNAs (sgRNA) were designed with a 20 bp binding sequence with PAM sequence (5'-NGG-4') of the Streptococcus pyogens Cas9 (SpCas9). They were synthesized using an *in vitro* RNA synthesis kit (Thermo Fisher Scientific, Waltham, MA, USA) after PCR amplification. Briefly, two oligonucleotides with overlapping were designed and synthesized as forward primers containing T7 promoter and sgRNA binding sequence and reverse primer, which containing remained sequences (Cosmo Genetech, Seoul, Korea). Next, PCR amplification was conducted using a premix (Takara Bio Inc., Shiga, Japan) and a PCR machine (BioRad, Contra

Costa County, CA, USA), and followed by *in vitro* transcription and purification. Synthesized sgRNA was stored at -80 ºC until use. ssODN consists of a total of 200 bp sequences including 40 bp homology arms on both sides and a 120 bp human Apoc3 sequence in the middle (HA-hApoc3 sequence-HA). Homology arms were designed to be homologous to mouse Apoc3 intron regions. The designed template was ordered from Cosmo Genetech in the form of single-stranded DNA. Detail sequence of sgRNAs and ssODN was listed in table 4-5.

Table 4. sgRNA sequence

sgRNA	No.	Target sequence	location	Intron	GC (%)	Off-target		
						0	1	2
Factor9	1	ACCCAGCCCCCCAAGTGGGGTGG	Chr 9	1	75	1	0	0
	2	AGCCCCCCAAGTGGGGTGGAAGG	Chr 9	1	70	1	0	0
Apoc3	1	AGCCGGGGATTCTGCCATGACGG	Chr X	1	60	1	0	0
	2	CTACTCAGTACCGAATGTGCAGG	Chr X	1	50	1	0	0
	3	TTCTGTGCAGGCTACCGTGAAGG	Chr X	8	55	1	0	0
	4	GGAGGCAAAGATTCGTGTGAAGG	Chr X	8	50	1	0	0

GC : Guanine-cytocine

Table 5. ssODN sequence

Feature	Sequence	Size
HA	5'-GTCCCGTTCTGCAGAGTATTTCTATACTCCACCTTCCACC-	40bp
Human sequence	CAGGTTCCCCCTCATTCTTCAGGCTTAGGGCTGGAGGAA- GCCTTAGACAGCCCAGTCCTACCCCAGACAGGGAAACTGA- GGCCTGGAGAGGGCCAGAAATCACCCAAAGACACACAGCA-	120bp
HA	CCACTTGGGGGGCTGGGTCTACTGTAGTTCTCTATCTAAT-3'	40bp

Red - PAM, Black - sgRNA binding site, Blue - Homology Arm

3.3 Animal production with CRISPR/cas9

In order to estrus synchronization and superovulation, B6 female mice were treated with five IU of serum gonadotropin (Prospec Bio, East Brunswick, USA) and 5 IU of hCG (Prospec) with 48 h intervals. Then, the female mice were mated with the sperm donor mice, and embryos were collected from oviduct on the following day. Embryos with normal morphology were cultured in KSOM medium (Merck Millipore, Billerica, MA, USA). Next, SpCas9 ribonucleoprotein (RNP) and ssODN were transfected into the embryos using electroporation [31]. Shortly, embryos were washed three times with Opti MEM 1 medium (Invitrogen, Carlsbad, CA, USA), and were transferred into the electroporation buffer on the electrode. The final concentration of buffer is 200 ng/μL of SpCas9 protein (Toolgen Inc, Seoul, Korea), 50 ng/μL of each sgRNA, and 200 ng/μL of ssODN. The electroporation pulse conditions were as follows: 7 cycles of 30 V with 3 ms ON and 97 ms OFF. After washing with M2 medium (MTI-GlobalStem, Rockville, MD, USA), the embryos were cultured under the KSOM medium (Merck Millipore, St. Louis, MO, USA).

Recipient donor mice were prepared by mating with a vasectomized male, and mice with plug were selected. After that, about twenty embryos were transferred to the oviduct of ICR recipient donor. At around ten days old of obtained pups, genomic DNAs (gDNA) were extracted from the toe clip and subjected to further PCR for genotyping. Information about the primer sequence was shown in table 6. The PCR amplicons through gel running and

gel extraction(New England Biolab, Ipswich, MA, USA) was used for cloning. Cloning used a T-blunt PCR cloning kit, and the protocol provided by the manufacturer was applied. PCR product (20ng at least) and buffer and T-blunt vector were reacted at a ratio of 4:1:1 for 1 hour to form a plasmid. To transfect the plasmid into the competent cell (IntronBio), heat shock was used (90 sec, 42°C). The competent cells with the plasmid were grown on Amp+ LB agar plates (37 °C, 16-18 hours) and Amp+ LB broth (37 °C, 12-14 hours, 250 rpm). Afterward, the plasmids from the competent cells were extracted using a mini-prep kit (Cosmo Genetech, Seoul, Korea). The plasmid was used for Sanger sequencing (Cosmo Genetech) with the universal primer M13

Table 6. Genotyping primer information

Target gene	Primer	Product size (bp)	Tm(℃)
Factor 9	F: 5'-CCTCCAAAACCCCCCTATC R: 5'-AAAAATGCACCACCAGCCC	KI : 809	58
Apoc3	F: 5'-GGAGAGGAAGGAAGGGAAGA R: 5'-TGCGAGAGAGCAGAGTTGG	KI : 821 WT : 701	58

KI: knock in, WT: wild type

3.4 RT-PCR and quantitative RT PCR

qRT-PCR was performed to compare the expression level of each gene and follow the protocol of the manufacturer (ThermoFisher Scientific). The method used in this experiment is a two-step RT-PCR, which includes reverse transcription PCR to make mRNA into cDNA and quantitative-Real Time PCR to amplify a desired portion of cDNA. the mRNA extracted from the tissue was synthesized into cDNA using an mRNA to cDNA kit (ThermoFisher Scientific). Primer information was obtained from the primer bank (https://pga.mgh.harvard.edu/primerbank/), and primers were ordered from the bioneer based on sequence information. the cDNA and primer were used for qRT-PCR with SYBR green (ThermoFisher Scientific), and qRT-PCR conditions are Tm: 60°C, El: 1&5min, and 50cycle. Information on the primer sequence is shown in Table 7.

3.5 Blood chemistry

CBC is a general blood test, a type of blood test used for disease diagnosis and follow-up. Sodium-citrate blood is used for the analysis, and each cell number is measured using an automated blood cell analyzer (mslabos, Chaussée Jules César, French). To obtain the numerical information of blood cells, diluent (Fluxionbio, Alameda, CA, USA) is diluted 15-50 times, and the test is performed using 50ul of diluted blood. Test results include leukocytes, erythrocytes, and thrombocytes.

Table 7. Primer sequence for RT-PCR

Target	length	Sequence Fwd	Tm	Sequence Rev	Tm
q-PCR Apoc3	374	AGCTGAAGAGGTAGAGGGA	58.83	AGGTGAGATCTAGGGAGGG	58.36
q-PCR F9	210	AGCAGCCCCTTCCCTAAGA	62	AGTGTGTCCGTCCTCCGAA	62.5
q-PCR IDUA	60	GCTGACCAGTACGACCTTAGT	60.9	TACGGCACCTATGTAGGCAAG	61.3
q-PCR IDS	150	AACCAGCAGCAAAACTACCAA	60.4	AGGCAGTGAAACCTGATATGATG	60.1
q-PCR GLA	209	GGCCCTAGAAGACATGATCCT	60.4	GGAGACGGATTGCTTGGAGG	60.4
q-PCR GBA	105	GGCTGTGCCGAAAGAGATTC	61.1	ATCCTGGGGTCCACTATCCTC	61.9
q-PCR GAPDH	123	AGGTCGGTGTGAACGGATTTG	62.6	TGTAGACCATGTAGTTGAGGTCA	60.2

3.6 PT and aPTT analysis

First, mice were anesthetized with an intraperitoneal injection of 2.5 % of Avertin solution. Using a syringe prefilled with 50µl of a 3.2% sodium citrate solution (Medicago, Durham, NC, USA), 450 µl of blood was collected from *inferior vena cava* to obtain a total of 500 µl of non-coagulated blood. In microplate-based activated partial thromboplastin time (aPTT) analysis, 30ul of plasma, and aPTT reagent (ThermoFisher Scientific) was mixed well in the 96 microplates and were incubated at 37°C for 5 min. Next, 30 µl of CaCl2 (26 µM) was added to the incubated serum-reagent mixture and measured absorbance at 405 nM in every 10 seconds for 8 minutes at Shaking (high, 10 sec). In prothrombin time (PT) analysis, 30 µl of plasma and PT reagent (ThermoFisher Scientific) were mixed and measured OD ratio immediately for 8 minutes with a 10-second interval and without shaking. PT reagent already contains an excess of CaCl2, and the coagulation reaction begins immediately. The time point with the highest value of △OD(Time(n+1), Time(n)) was selected as the result value of PT and aPTT test.

3.7 *In vivo* bleeding test

The tail bleeding test was performed to measure blood clotting time by trauma with previously reported protocol[24, 25]. Seven weeks old, littermate mice were divided according to their genotype (B6.WT and B6.F9 KO). After

anesthesia using 2.5% of an avertin, mice were placed above the warm plate and cut its tail with a sharp blade. The next cut tail end was dipped into warm saline in a 50 ml conical tube and measured bleeding time. Bleeding time was calculated only for 20 minutes. After *in vivo* bleeding analysis, the blood in the warm saline was collected and subjected to hemoglobin assay. The collected blood was centrifuged at 3500 rpm for 5 min, and the supernatant was removed. The pellet is then wholly dissolved using the 2 ml of RBC lysis buffer (ThermoFisher Scientific), and insoluble material was removed. The absorbance of each 200 ml solution was measured at 570 nm.

3.8 Histology of Spleen in B6. F9 KO mouse

Spleen tissues were fixed by 10% formalin, and H&E stain was conducted.

3.9 AAV-CRISPR and AAV-donor preparation

AAV packaging uses a triple transfection strategy, which is included the essential rep and cap genes (AAV Rep/Cap vector) and an adenovirus-derived helper plasmid (Ad Help vector) that supplies the genes required for replication. Also, the construction of a plasmid (pAAV-ITR-GOI-ITR vector) containing the gene of interest (GOI). GOI includes CRISPR templates, including sgRNA and cjCas9, and donor templates of various design designs.

The three plasmids are co-transfected with HEK293T, and the GOI is packed by AAV. Transfection was carried out on the 4th, followed by the concentration and purification of AAV particles. Based on standard curves of known concentrations of plasmid samples, measuring viral titer is quantified using qPCR with primers targeting ITR.

A mixture of AAV-CRISPR and AAV-donor was injected intravenously via the tail vein to a 6-8 weeks old mouse in a 1:1 ratio for delivery of AAV, and the control group was administered with AAV-donor alone. All the AAV solutions were diluted to 200ul using sterile saline. Information on the AAV dose is shown in Table 8.

Table 8. AAV dose by experiments

Experiment	No.	GC/mice	
		CRISPR	**Donor**
Normal dose (HR, C+D group)	#1~6	$5x10^{13}$	$5x10^{13}$
Normal dose (HR D group)	#1~5	-	$5x10^{13}$
Normal dose (Integration, C+D group)	#1~6	$5x10^{13}$	$5x10^{13}$
Normal dose (Integration, D group)	#1~3	-	$5x10^{13}$
Normal dose (HITI, C+D group)	#1~6	$5x10^{13}$	$5x10^{13}$
Normal dose (HITI, D group)	#1~3	-	$5x10^{13}$
No Treated (Control group)	#1~3	-	-
High dose (Integration, C+D group)	#1~4	$5x10^{13}$	$5x10^{13}$

GC: genome copy , C: CRISPR treated, D: donor treated, No:Number

3.10 Human Coagulation Factor 9 ELISA

Approximately 70 μl of blood was collected using the heparin micro-tube, and plasma was separated by centrifugation under 3000 rpm for 10 minutes at 4 °C. Blood concentration of human F9 is measured using the Factor 9 human simple step ELISA kit (Abcam, Cambridge, UK). The experiment was conducted as a manufacturer's instruction. After the 200-fold dilution process, the plasma was loaded into a well using 50 ul, mixed with 50 ul of antibody cocktail, and incubated at 400 rpm for 1 hour. After the washing process, the absorbance was measured at 450 nm through a reaction and stop the process. Based on the standard curve, the absorbance value could be calculated as the protein concentration. The result was recorded as a result multiplied by 200 times to correct the dilution rate. A value of 0 or less was judged to be unable to detect the protein and was displayed as 0.

3.11 Statistical analysis

Statistical analysis was conducted using unpaired Student t-test (Graphpad prism ver. 5.02, La Jolla, California, USA)

4. Result

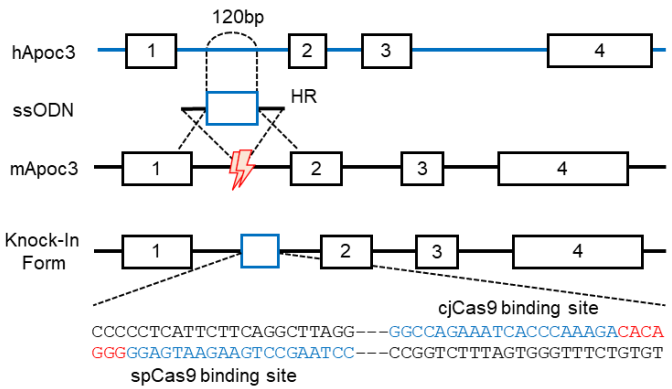
4.1 Producing mice for a new *in vivo* genome editing strategy

In strategies to repair or insert genes that cause disease directly, the reduced efficacy of *in vivo* gene editing is a significant problem of low efficiency [20, 21, 26], and this is caused by technical limitations such as small transfection and KI frequency. The strategy designed to overcome this technical limitation is to insert the therapeutic gene into a gene that has a higher expression level than the gene which causes the disease[27, 28]. Thus, it was necessary to find a gene that showed a high level of expression in the target organ, and apolipoprotein C3(Apoc3) was chosen after analyzing with microarray in human hepatocyte. Apoc3 gene is known to encode a major protein of VLDL and apolipoprotein, a small component of HDL[29]. The FRKM value of Apoc3 in human hepatocyte is 7998, which is higher than the expression level of other diseases such as coagulation factor 9(F9) : 42, iduronidase-alpha-L(IDUA) : 27, iduronate-2-sulfatase(IDS) : 35, galactosidase-alpha(GLA) : 18, glucosidase-beta-acid(GBA) : 29 (data from Toolgen). To utilize the expression level of Apoc3, the place for inserting the therapeutic gene was selected as Intron 1 of Apoc3. This is because intron 1 is a position where the indel caused by DSB will have little effect on expression and overlaps with various alternative splicing of Apoc3. Since indel affects the efficacy of *in vivo* gene editing, many sgRNAs were screened in the Apoc3 gene intron 1 region, and finally selected two sgRNAs

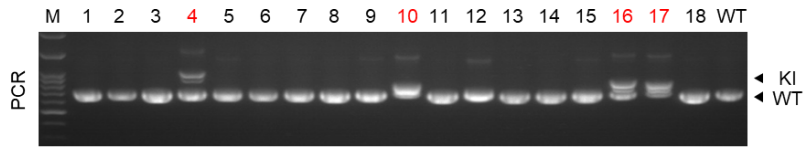
with high DNA cleavage potential for SpCas9 and *Campylobacter jejuni* Cas9 (CjCas9). The selected sgRNA sequences are only presented in human Apoc3, so it was necessary to generate a new mice model for further study. Since candidate sgRNAs were chosen from in intron one region, it was planned to inserted selected sgRNA sequences into intron 1 of the mouse Apoc3 (mApoc3) gene.

Based on previous reports of an efficient method for inserting sequences of 100 bp or more[30, 31], single-stranded template repair (SSTR) using single-stranded oligodeoxynucleotide (ssODN) was applied for animal production. ssODN consists of homologous arms of each 40bp on both sides and a human Apoc3 sequence of 120 bp in the middle. (ssODN: HA- human Apoc3 sequence -HA). RNP complex was created by incubating spCas9 and two sgRNAs targeting Apoc3 Intron 1, and electroporation was conducted to transfect RNP complex and ssODN into embryos. Then, the embryos were transferred to the oviduct of pseudopregnant females (Figure 4a). Obtained pups were subjected to PCR based genotyping using the toe clip, and knock-in efficiency is proven to 22% (4/18)(Figure 4b). Although no founder pups exhibited precise 120 bps knock-in, the #16 mice contained an intact sgRNA binding site of spCas9 and cjCas9. Besides, the others were just born with partial knock-in (including +70bp, +22bp, and +101bp, respectively) (Figure 4c). Based on the sequencing results, #16 mouse was selected as the founder for further germline transmission to make B6.Apoc3$^{hseq/hseq}$. There were no abnormalities in the founder mouse about reproduction or growth.

a



b



Knock-In efficiency : 22%(4/18)

c

No.	Band	Sequence
Expected Mutant	WT	TCCACC----------CCACT
	KI	TCCACCCAGGTTCCCCCCTCATTCTTCAGGCTTAGGGCTGGAGGAAGCCTTAGACAGCCCAGTCCTACCCCAGACAGGGAAACTGAGGCCTGGAGAGGGCCAGAAATCACCCAAAGACACACAGCACCACT
4	WT	TCCACC----------CCACT
	KI	[+29/-6bp]----------TCATTCTTCAGGCTTAGGGCTGGAGGAAGCCTTAGACAGCCCAGTCCTACCCCAGACAGGGAAACTGAGG[+157bp]----------CCACT
10	WT	TCCACC----------CCACT
	KI	[+65/-12bp]----------AATCACCCAAAGACACACAGCACCAC
16	WT	TCCACC----------CCACT
	KI	TC----CAGGTTCCCCCCTCATTCTTCAGGCTTAGGGCTGGAGGAAGCCTTAGACAGCCCAGTCCTACCCCAGACAGGGAAACTGAGGCCTGGAGAGGGCCAGAAATCACCCAAAGACACACAG-ACCACT
17	WT	TCCACC----------CCACT
	KI	[-17bp]----------TCAGGCTTAGGGCTGGAGGAAGCCTTAGACAGCCCAGTCCTACCCCAGACAGGGAAACTGAGGCCTGGAGAGGGCCAGAAATCACCCAAAGACACACAGCACCACT

Figure 4. Generation of KI mice with human sgRNA binding sequences **a**) Schematics for human Apoc3 sequence knock-in mouse generation. Two sgRNAs and 200 nt ssODN were transfected into one-cell stage embryos. Blue box: 120 nt human sequences, Lightening symbol: DSB, Black box: intron, Black bar in ssODN: 40 nt sized homology arm, Red alphabet: PAM sequences of each Cas9, and Blue alphabet: sgRNA binding sequences. **b**) Genotyping for human Apoc3 sequence knock-in(KI). Red alphabet: KI mice, Black triangle: target specific size. (WT; 701bp and KI;821bp). M: size marker, WT: wild type, KI: knock-in. **c**) Sequencing results of obtained pups. Blue alphabets: knock-in sequences, Red alphabets: summary of indel.

4.2 B6 Apoc3$^{hseq/hseq}$ as a suitable model to mimic *in vivo* gene editing for human

Mutations in the intron region have the possibility of inducing alternative splicing[32], and alternative splicing can cause disease[33, 34]. It was examined whether the insertion of the human sequence into the Apoc3 intron region causes alternative splicing. To find alternative splicing, the primer was designed to recognize sequences on exon 1 and exon 2, and RT-PCR was performed using cDNA obtained from B6. Apoc3$^{hseq/hseq}$ liver. The control group for comparison included cDNA and gDNA of B6. Each amplicon of B6.Apoc3$^{hseq/hseq}$ and B6 were presented at the same size and position, and it indicates that no evidence has been found that inserting human sequence affects splicing (Figure 5a). Besides, even the Apoc3 gene is highly expressed in human liver tissue, mouse Apoc3 (mApoc3) gene expression in the liver needs to be confirmed. qRT-PCR was performed using the cDNA of B6.Apoc3$^{hseq/hseq}$ mouse liver. Gene expression of mApoc3 is compared with Coagulation Factor 9(F9), Iduronate-2-sulfatase(IDS), Iduronidase-alpha-L(IDUA), Galactosidase-alpha(GLA), and Glucosidase-beta-acid(GBA), all of which are genes relating with the mono-genetic disorder. As a result, the expression level of Apoc3 was 40-300 times higher than causing-disease gene(Apoc3: 325.4, IDS: 1.8, F9: 1.0, GLA: 0.1, GBA: 0.2, IDUA: 6.7)(Figure 5b). Because the mApoc3 gene is highly expressed in the liver, B6.Apoc3$^{hseq/hseq}$ could be a suitable animal model for studying *in vivo* gene editing for humans.

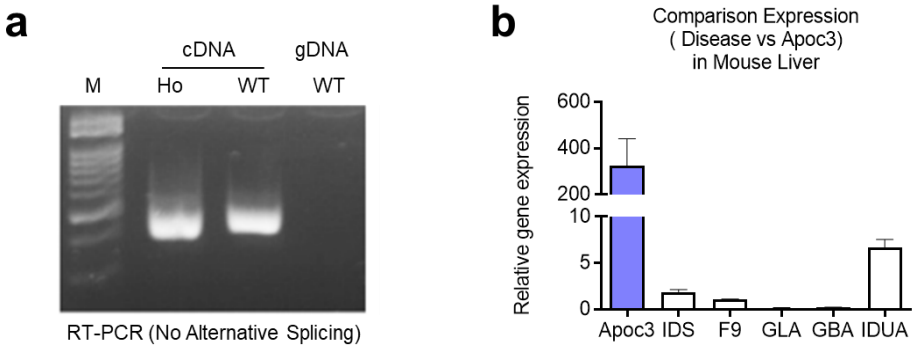
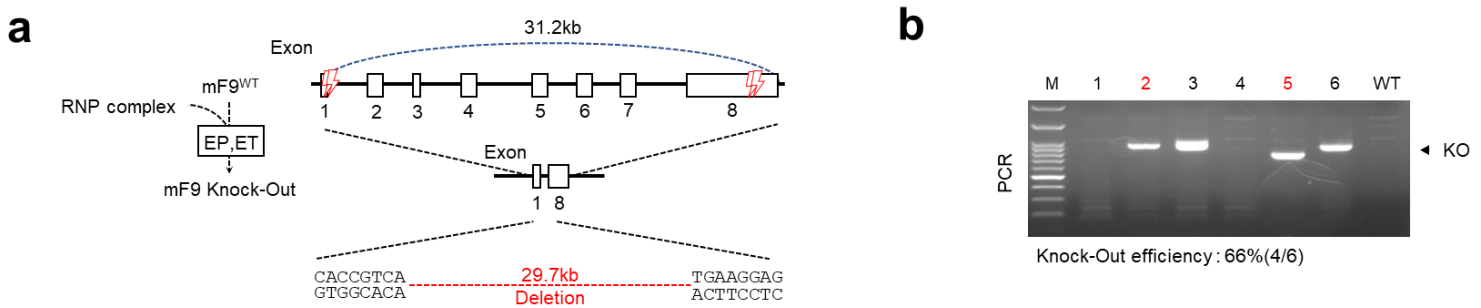


Figure 5. Analysis of gene expression of B6. Apoc3$^{hseq/hseq}$ mouse **a**) RT-PCR and PCR was performed using cDNA and gDNA from liver tissue. Primer was designed to recognize sequences in exon1 and exon2 to confirm the existence of alternative splicing. Ho: homozygote B6.Apoc3$^{hseq/hseq}$, WT: B6, target size: 374 bp **b**) Quantitative RT-PCR was conducted for Apoc3 and other disease-causing genes using the cDNA from the liver tissues of B6.Apoc3$^{hseq/hseq}$ (n=8). Gene expression was normalized to that fo GAPDH and presented as mean ± SEM. IDS: Iduronate-2-sulfatase, F9: Coagulation factor 9, GLA: Galactosidase-alpha, GBA: Glucosidase-beta-acid, IDUA: Iduronidase-alpha-L.

4.3 Generation of a mouse model with factor 9 deficiency using CRISPR/Cas9

After producing a mouse model that is capable of studying *in vivo* gene editing, it required another model with a genetic disease to evaluate whether *in vivo* gene editing has a therapeutic effect. In this study, hemophilia B is chosen as a candidate genetic disease, which has the advantage of high expression in hepatocyte and a treatment standard with a particular phenotype. Hemophilia B is known to have blood clotting disorders due to the X-chromosome factor 9 gene mutation with loss of function and occurs in about 1 in 20000-30000 infants. Several strategies, such as the formation of nonsense mutation, frameshift mediated premature termination codon, and whole gene exon deletion, could be applied in establishing a genetically engineered animal model with loss of function. However, nonsense or frameshift mutations could avoid nonsense-mediated decay[35, 36] and develop genetic compensation[37, 38]. Thus, strategy with the whole exon deletion was selected, and sgRNAs were designed to exon 1 and 8 of the F9 gene. Also, a pair of sgRNAs in the back of the start codon and front of the stop codon were designed to preserve the start codon and stop codon. The overall process of the F9 knock out (KO) mouse is similar to that of B6.Apoc3$^{hseq/+}$ except no applying ssODN(Figure 6a). After obtaining pups, PCR and sequencing were conducted and were proven to 4 out of 6 (66%)

mice were mutant with complete exon deletion(Figure 6b). Mutant mice had deletion with 29 kb or more, and among the 472 translated amino acids (AA) of the mF9 gene, 465 amino acids were deleted in #2 and #5 mice. In detail, additional amino acids were translated by deletion mediated frameshift at #2 mouse(+7 AA) and #5 mouse(+12 AA), but both had termination codon sequence at 3' end. However, #3 mice did not induce termination codon at 3' end. Thus #2 male mouse and #5 female mouse were chosen as founders for the production of hemizygote or homozygote mouse(B6.F9 KO) by the germline transmission (Figure 6c). Also, there were no remarkable abnormalities in breeding.



c

No.	Sequence	Amino acid
Expected	caccgtca----------Del----------tgaaggag	MKHLNTV
2	caccgtca----------ccggctctcatcaccatctt----------Del---gaaggaga	MKHLNTVNRLSSPS*
3	caccgtca-----gaatccccggctctcatcaccatcttcctttaggatatctactcagta---Del----------	MKHLNTVRIPGSHHHLPFRISR
5	caccgtcatg----------Del----------aaggagat	MKHLNTVMKEIVGDPMLLK*

Figure 6. Generation of Factor 9 knock out mouse **a**) Schematics of whole exon deletion for mF9 gene. Dual DNA breakage was induced by RNP (400 μg of SpCas9 and each 100 μg of four sgRNAs). Thunder symbols: double-stranded DNA breakage b) PCR genotyping for produced pups (target size: 865 bp), Red alphabets: founder mice with whole exon deletion. C) Sequencing results for produced pups. Red alphabets: additional translated by deletion meditated frameshift. *: termination codon

4.4 Similar hematological phenotype in B6.F9 KO with human hemophilia B patient

In blood clotting cascade, bleeding causes fibrin production via extrinsic and intrinsic pathway, and begin coagulation. Whereas, patients with Hemophilia B induces a slow fibrin production and a blood clotting disorder due to FIX deficiency in the intrinsic pathway. A bleeding disorder is assessed by blood chemistry and coagulation ability. Among the standard protocol for diagnosis of a bleeding disorder, PT and aPTT are used to find the cause of bleeding, such as defective in the intrinsic or extrinsic pathway. In PT experiment, mF9 KO and B6 mice exhibited coagulation of average 17.5 sec similarly. The coagulation time of the F9 KO mouse is close to the normal PT range of humans (11~13.5 sec) [39]; it reveals that F9 KO mice have a normal extrinsic pathway for blood clotting. Normal aPTT range is 30-45 sec in humans[40], and this is similar to that of wild type B6. On the other hand, B6 mice exhibited 47.5 sec of aPTT coagulation time, but F9 KO mice presented maximum value as 480 sec. Because coagulation time calculation in aPTT experiment only for 480 sec, this indicated severe bleeding disorder via the intrinsic pathway in F9 KO mice.

Delayed coagulation time in aPTT is a hallmark of hemophilia in a human patient [40] and mouse model[23, 41](Figure 7a). Next, a tail bleeding test was conducted to reconfirm coagulation disorder [25]. B6 mice stopped bleeding around at 5.2 min after trauma, but F9 KO mice did not terminate

bleeding until 20 min. In an additional analysis for blood loss using hemoglobin assay and weighting of collected blood, the F9 KO mouse exhibited significantly higher than B6. The clotting time of F9 KO in aPTT was delayed three times or more than that of B6. (Figure 7b). Delayed coagulation and increased bleeding volume in F9 KO mice tend to be the same as other papers in tail bleeding experiments using hemophilia mice. [41-43]

Next, a complete blood count (CBC) test was performed to check for hematological abnormalities. Hematocrit, which represents the volume of blood cells, decreased by about 10%, and the eosinophil percentage increased three-fold in the F9 KO mice than that of B6. But, they were all within the normal range in humans [40, 44]. Also, the ratio of RBC and MPLT did not develop a significant difference between F9 KO and B6 mice. In particular, a normal MPLT value means that the cause of the clotting disorder is not due to the absence of platelets (Figure 7c). Comprehensively, this clotting disorder is a phenotype caused by the lack of the intrinsic pathway, not the absence of platelets and the extrinsic pathways. This symptom is similar to patients with hemophilia B in humans[40]. Finally, the F9 KO mouse with complete exon deletion is an appropriate model mimicking to human hemophilia B.

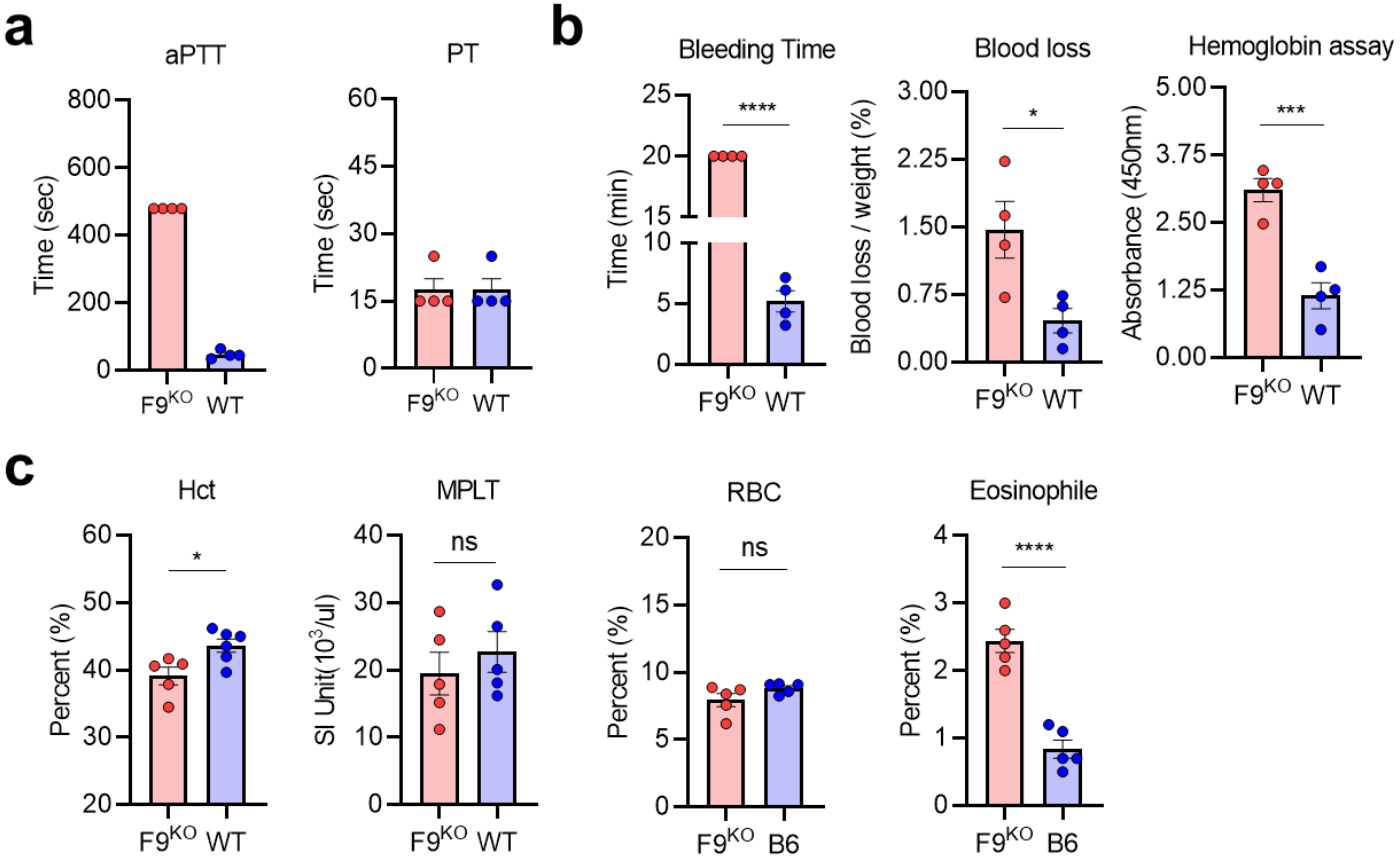

a
aPTT
Time (sec)
800
600
400
200
0
F9KO
WT
PT
Time (sec)
60
45
30
15
0
F9KO
WT
b
Bleeding Time
Time (min)
25
20
15
10
5
0
****
F9KO
WT
Blood loss
Blood loss / weight (%)
3.00
2.25
1.50
0.75
0.00
*
F9KO
WT
Hemoglobin assay
Absorbance (450nm)
5.00
3.75
2.50
1.25
0.00
***
F9KO
WT
c
Hct
Percent (%)
60
50
40
30
20
*
F9KO
WT
MPLT
SI Unit(10^3/ul)
40
30
20
10
0
ns
F9KO
WT
RBC
Percent (%)
20
15
10
5
0
ns
F9KO
B6
Eosinophile
Percent (%)
4
3
2
1
0
****
F9KO
B6

Figure 7. The hematological phenotype of the mouse with complete F9 exon deletion **a)** Results of PT / aPTT test. B6 (male, n=3) F9 KO mouse (male, n=3) **b)** Tail bleeding test, B6 (male, n=4) F9 KO mouse (male, n=4) **c)** Blood chemistry was conducted for complete blood count (CBC). Hematocrit (Hct) Manual Platelet Count (MPLT), Red blood cells (RBC), and eosinophils. B6 (male, n=6) F9 KO mouse (male, n=6)) each dot indicates a value from an individual mouse, and data were presented as means ± SEM. Statistical analysis was performed using Student's *t*-test. *: $p < 0.05$, **: $p < 0.01$ and ***: $p < 0.001$

4.5 Extramedullary hematopoiesis and decreased survival rate in F9 KO mice

Next, additional histological and survival rate was analyzed. Interestingly, the F9 KO mouse exhibited enlarged spleen and weighted about 24 % higher than B6(Figure 8a). Further histological analysis reveals that there was an increased number of RBC and myeloid precursor cells. This indicated symptoms of extramedullary hematopoiesis (EMH), with an [45] (Figure 8b). Pathological EMH can occur when physiological hematopoiesis is not working correctly and is also found in severe chronic anemia. Although F9 KO mice develop EMH in the spleen, EMH is not a common symptom in human hemophilia B patients. Thus, this still needs to investigate the correlation and mechanism between hemophilia and EMH.

Hemophilia B caused a shorten life span in humans; thus, the survival rate was analyzed in the F9 KO mice. For 150 days, the death of mice was recorded, and this reveals that the F9 KO mouse exhibited approximately 25% higher motility that controls B6(Figure 8c). In comparison with a human patient, the sever hemophilia B induced about 20 % of decreased life span, and the leading cause of death is intracranial and cardiovascular hemorrhage. [13]. As in hemophilia patients, a short life span in the F9 KO mouse would be caused by repeated bleeding, but there was no typical event of bleeding or trauma. However, it cannot be ruled out the possibility of mild and repeated bleeding, which is difficult to find.

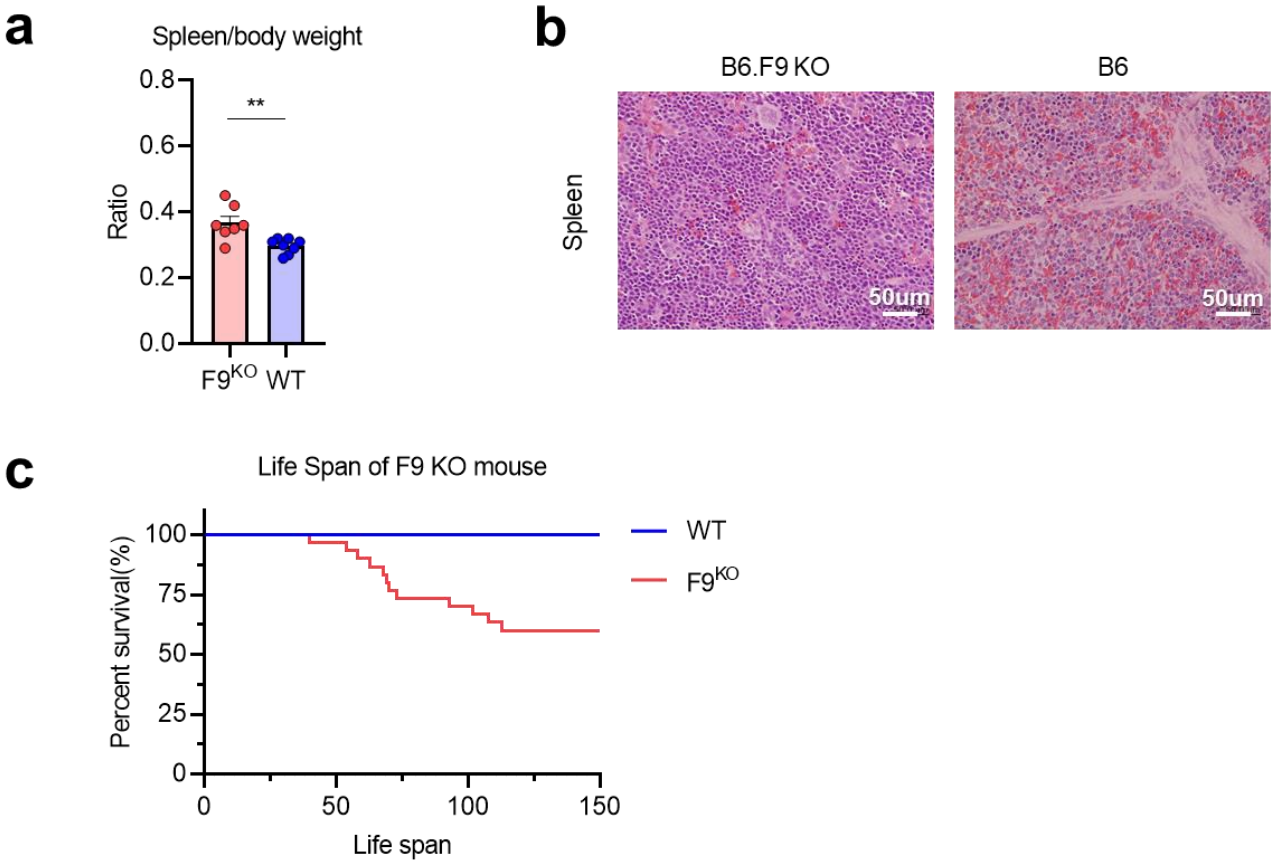

a
Spleen/body weight
0.8
0.6
0.4
0.2
0.0
Ratio
**
F9KO
WT
b
B6.F9 KO
B6
Spleen
50um
50um
c
Life Span of F9 KO mouse
Percent survival(%)
100
75
50
25
0
0
50
100
150
Life span
WT
F9KO

Figure 8. Splenomegaly and Extramedullary hematopoiesis F9 KO mice **a)** Comparison of weight (spleen/body weight) between B6 and F9 KO mice. B6 (male, n=8) F9 KO mouse (male, n=7). Each dot indicates an individual mouse. * in the graph indicates statistical significance (** : $p < 0.01$) **b**) H&E stain was conducted with spleen tissue. The scale bar represents 50 μm in the white box. **c)** The lifespan of B6 and F9 KO mice was calculated for 150 days. B6 (n=29) F9 KO mouse (male, n=30)

4.6 Effective transduction of AAV serotype 8 into the liver

Transduction of exogenous genetic material is one of the critical factors affecting the efficiency of gene editing. Methods for introducing exogenous genetic material into cells are various, and commonly include viral vector and nonviral vector. Currently, viral vectors are widely utilized for *in vivo* transfection due to its high potential of delivery and sustainability. Selecting a viral vector with advantages, such as low immune response and long-term expression, is essential for the successful delivery of genetic material[46]. Recently, gene therapy using adeno-associated virus (AAV) has been approved by the FDA with its low immunogenicity and long-term expression charasteristics[47, 48]. AAV has tropism that shows different transduction efficiency for each organ depending on the serotype[49]. Thus, I tried to select the AAV serotype that would be used for *in vivo* gene editing by comparing the transduction efficiency of AAV8 and AAV9, known to have tropism in the liver[50]. AAVs packaged with an enhanced green fluorescent protein (eGFP) (AAV8/9-CMV-eGFP-pA) were prepared by synthesis with commercial service. Next, $5x10^{13}$ GC/kg of AAV particles were injected via the intravenous route, and eGFP signal intensity was analyzed by immunofluorescence staining in the liver after four weeks. Immunofluorescence result presents a similar pattern with those of the paper comparing tropism of AAV serotype[49, 50], and it was found that both AAV8 and AAV9 transduced foreign genes into the liver. However, it was found that AAV8 exhibits higher transduction efficiency than AAV9[51]

(Figure 9). Therefore, AAV8 was selected as a target serotype for packing the CRISPR and donor template, which would induce therapeutic gene KI into the target locus.

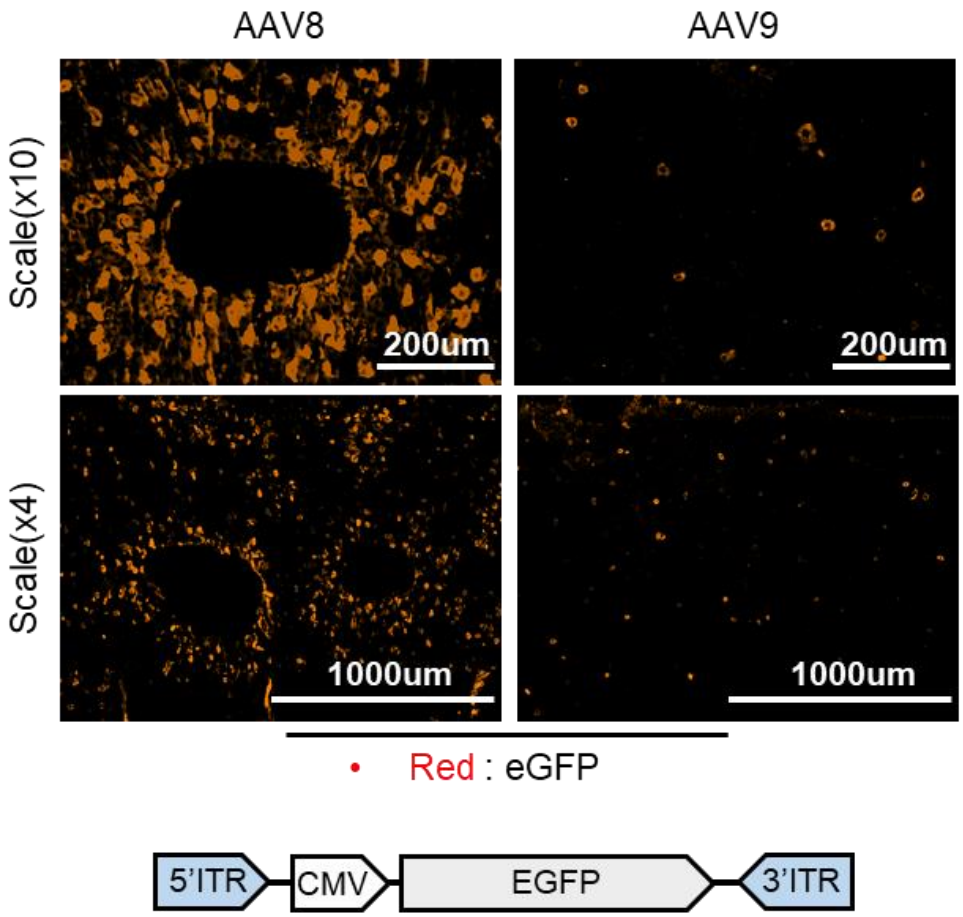


Figure 9. Comparison of transduction potential in hepatocyte between AAV8 and AAV9. After intravenous injection of AAV8-CMV-eGFP and AAV9-CMV-eGFP into mice at a dose of $5x10^{13}$GC/ kg, GFP expression in the liver was compared four weeks later. For each AAV treated group, 4 B6 mice were applied. The scale bars are 1000 μm (x4) and 200 μm (x10) in the white box. The red color represents the expression of eGFP, and the diagram shown below is the vector map of the AAV-eGFP.

4.7 Extremely low efficiency of homologous recombination-mediated *in vivo* gene KI

Because hF9 is selected as a therapeutic gene for KI, it was needed to decide the strategy for KI. As a method for KI, there is a method using HR and NHEJ. NHEJ-mediated KI produces a direction-independent insertion, but HR-mediated KI has the advantage that the therapeutic gene can be inserted in the desired direction. Thus, HR strategy was selected as a method for *in vivo* gene KI into the liver. For *in vivo* gene editing using the AAV8, it was necessary to design the cassette of CRISPR and donor template. The target location of DSB for *in vivo* genome editing is the sgRNA binding site for cjCas9 and spCas9 in the Human Apoc3 sequence. However, the sgRNA binding site for spCas9 could not be used due to the off-target effect in the human genome. Thus, the targeted place for DSB is selected as the sgRNA binding site for cjCas9, and it was necessary to make a vector encoding CjCas9 protein and sgRNA. A vector for AAV production was designed with an all-in-one type as a simultaneous expression of CjCas9 protein and sgRNA (ITR-U6-gRNA: hApoc3-TBG-cjCas9-ITR). In detail, the U6 promoter was used for RNA synthesis, and the TBG promoter was applied for enhancing liver-specific expression[52]. The DSB region caused by cjCas9 could be the insertion site of the therapeutic gene. Additionally, to induce this insertion, the donor Template needed to include a homology arms to trigger homology recombination(HR). HA has homology to the vicinity of the gRNA binding site where DSB will occur. The length of HA was selected as 1 kb,

considering the length(500 bp-1 kb) used in the paper[22, 53] that experimented with other HR-mediated KI. It includes a splicing acceptor(En2sa) that receives a promoter signal to express when KI occurs, and hF9 cDNA form and poly-A tail to treat human hemophilia. (ITR-HA-En2sa-hF9 cDNA-P.A-HA-ITR). (Figure 7a). The cassette of cjCas9 and donor were packed by AAV8.

B6.Apoc3$^{hseq/hseq}$ mouse was randomly divided with two groups, one group was transduced AAV-CRISPR and AAV-donor (C+D), and the other group was applied with AAV-donor (D). For similar conditions to other experiments using AAV[22, 53], 5x10^{13}GC/kg dose, and 8-10 weeks old mouse and male were used. Next, ELISA for human F9 was conducted at various time points from 2 weeks to 22 weeks (2, 6, 10, 14, 18, and 22 weeks) to evaluate the recovery of symptoms. At two weeks after AAVs transduction, the human F9 concentration in the "C + D" and "D" was similar at different time points of analysis. Also, there were several remarkable observations such as 1) the highest AAV mediated gene expression in 6 weeks after transduction, 2) relatively high hF9 plasma concentration in the group "D," and 3) high hF9 concentration up to 5,000 ng/ml which is average F9 concentration in human (Figure 7b). These results suggest an auto-expression of HR-donor, which could not distinguish the success of HR-mediated KI. Besides, the verification through PCR could not confirm whether the therapeutic gene was inserted. Briefly, there was no undeniable

evidence of HR mediated therapeutic KI, and KI efficiency seemed to be lower than common expectation.

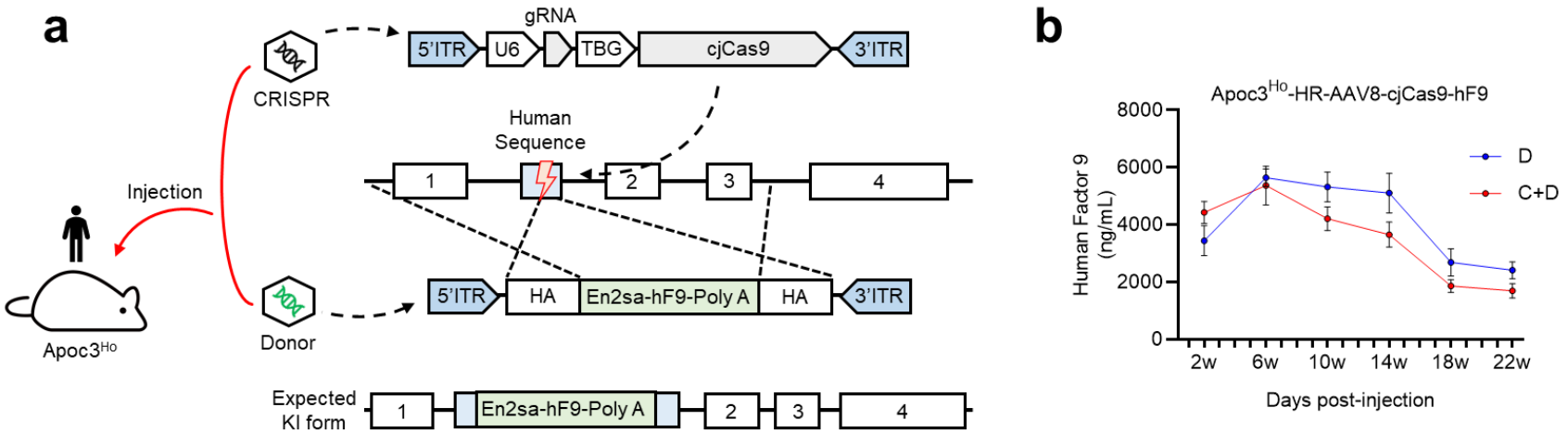

a
gRNA
5'ITR
U6
TBG
cjCas9
3'ITR
CRISPR
Human
Sequence
1
2
3
4
Injection
Apoc3Ho
Donor
5'ITR
HA
En2sa-hF9-Poly A
HA
3'ITR
Expected
KI form
1
En2sa-hF9-Poly A
2
3
4
b
Apoc3Ho-HR-AAV8-cjCas9-hF9
Human Factor 9
(ng/mL)
8000
6000
4000
2000
0
D
C+D
2w
6w
10w
14w
18w
22w
Days post-injection

Figure 10. *in vivo* HR-mediated gene editing in B6. Apoc3$^{hseq/hseq}$ **a)** This diagram shows the *in vivo* gene editing strategy using homologous recombination in B6. Apoc3$^{hseq/hseq}$ mouse. Adult B6. Apoc3$^{hseq/hseq}$ received a mixture of AAV8 (CRISPR and donor) via the intravenous route. **b)** Plasma samples were analyzed by ELISA to measure changes in the concentration of hF9 in the blood. The period lasted for 22 weeks, and sampling was performed at two weeks and every four weeks. The experiments was conducted in CRISPR+donor (B6. Apoc3$^{hseq/hseq}$, Male, n = 6, age = 7 weeks old) and donor only (B6. Apoc3$^{hseq/hseq}$, Male, n = 6, age = 7 weeks old) groups. Dose of AAV used in the experiment is CRISPR+donor (C: D = 5x10^{13}GC/Kg: 5x10^{13}GC/Kg), and donor only (D = 5x10^{13}GC/Kg).

4.8 High cross-reactivity between human and mouse F9

It was found that the AAV8 HR donor showed a peak of hF9 concentration at week 6 in the experiment of HR-mediated knock-In. It was assumed that the high hF9 concentration, which is produced by the AAV8-HR donor, was a suitable target for detecting cross-activity with mF9. The experiment included ELISA and PT / aPTT tests, and the experiment was conducted in three groups: B6.F9 KO -AAV8-HR donor, B6.F9 KO -no treated, and B6. WT-no treated. B6.F9 KO -AAV8-HR donor group was injected with the same dose($5x10^{13}$GC/kg) as in the previous experiment, and then Sampling was performed at six weeks. In the B6.F9 KO -AAV8-HR donor group, the hF9 concentration of 2935 ng/ml was confirmed, and B6. WT and B6.F9 KO group used as controls showed 0 ng/ml (Figure 8a). Through this, it was confirmed that hF9 and mF9 could be distinguished by ELISA. In aPTT, this expression level of hF9 made the coagulation disorder of the F9 KO mice group to have a coagulation time similar to the B6. WT group. Besides, PT results were found to show similar mean values in all groups, which was not statistically significant(Figure 8b). Consequently, this means that hF9 has cross-activity with mF9 and may rescue the clotting disorder of B6.F9 KO mouse. However, the dose correlation between hF9 and mF9 was not known. Also, since there was still no evidence of the *in vivo* therapeutic gene KI occurring, it was necessary to more increase the KI efficiency.

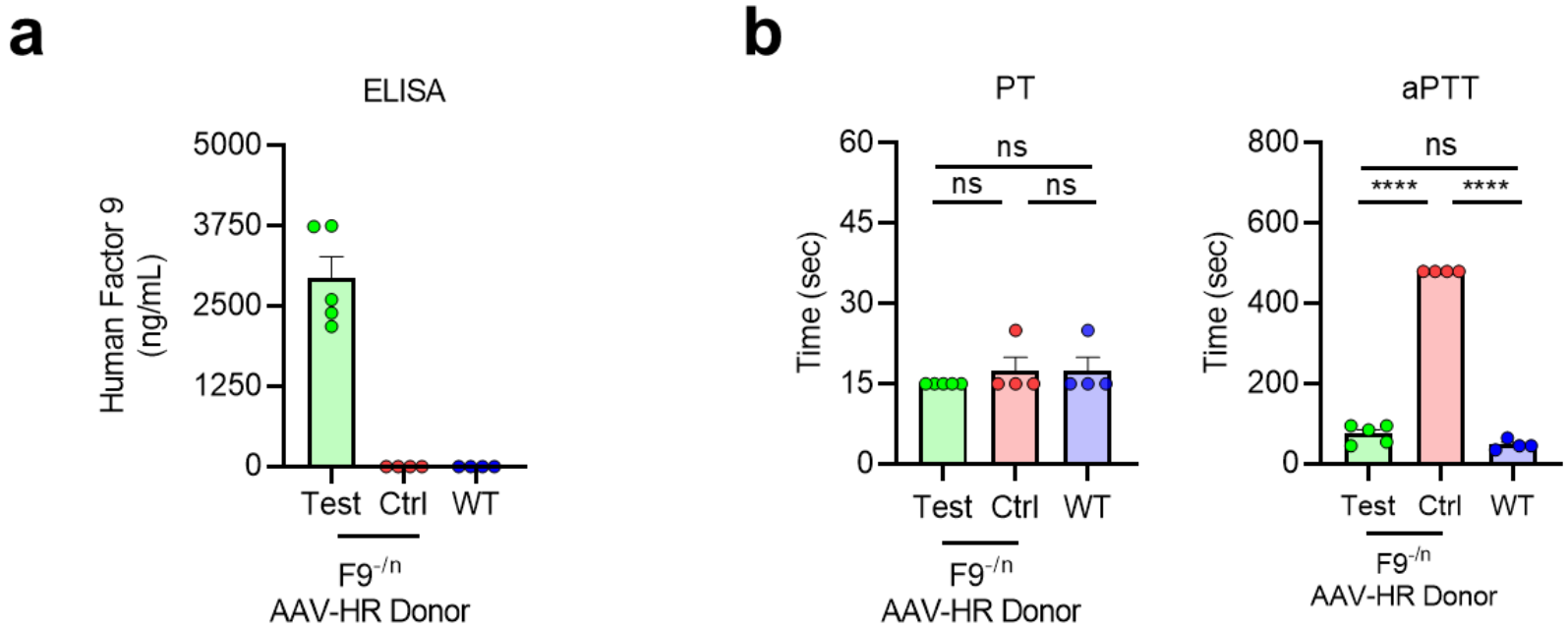

a
ELISA
Human Factor 9 (ng/mL)
5000
3750
2500
1250
0
Test
Ctrl
WT
F9-/n
AAV-HR Donor
b
PT
Time (sec)
60
45
30
15
0
ns
ns
ns
Test
Ctrl
WT
F9-/n
AAV-HR Donor
aPTT
Time (sec)
800
600
400
200
0
ns
****
****
Test
Ctrl
WT
F9-/n
AAV-HR Donor

Figure 11. Auto-expressing hF9 HR donor alleviates coagulation disorders in F9 KO mice **a)** $5x10^{13}$GC/kg of AAV8-hF9 HR donor were transduced into F9 KO mice of each group (male, age: 7~8 weeks old, n = 4~5). **a**) Next, plasma hF9 concentration was measured by ELISA at six weeks after AAV transduction. Each dot indicated an individual mouse. **b**) PT and aPTT experiment was conducted. Data were presented as mean ± SEM, and statistical analysis was conducted with unpaired Student *t*-test. ***: $p < 0.0001$, NS: not significant.

4.9 Evidence of therapeutic gene expression in NHEJ mediated KI strategy

Homologous recombination is known to be cell cycle-dependent[54], which mainly occurs in the S phase and G2 phase[11, 55]. However, hepatocytes remain in the G0 cell phase in the hepatocyte of adult mice or are replaced very slowly or for 180-400days[56]. These reports would support the low efficiency of HR-mediated KI in a previous trial to B6.Apoc3$^{hseq/hseq}$ mouse. On the other hand, NHEJ is known to be a cell cycle independent DNA repair pathway[11, 55], and this means that NHEJ mediated KI has a possibility of success of *in vivo* gene editing in adult mice. In this study, two different methods of NHEJ mediated KI were applied. Homology-independent targeted integration (HITI) is a KI method that does not use the HR pathway, characterized by the DSB of the cassette and the absence of HA[57, 58]. Also, due to the nature of NHEJ, it is inserted regardless of the direction. On the other hand, Integration is characterized by the insertion of a gene including ITR as a method using the characteristics of AAV[59], and has treatment genes facing each other to solve the directional problem of HITI. Both HITI and integration strategy could share AAV-CRISPR; but, there was a need for new preparation of donor AAVs. In HITI, instead of homology sequences, the human Apoc3 sequence containing the cjCas9 gRNA binding site was designed to be included on both sides (HITI donor : AAV8-ITR-Human Apoc3-En2sa-hF9 cDNA-Poly A-Human Apoc3-ITR). This would induce three DNA breakages as one in host target locus and two

at both ends of donor DNA, and followed by KI with micro-homology end-joining. However, there would be undesired integration with reverse direction (Figure 9a), whereas ITR mediated integration expects high frequency and bidirectional gene expression. Forward-hF9 and reverse-hF9 were linked so that they can be expressed well even when knocked-in in any direction regardless of the direction (Integration donor: AAV8-ITR-En2sa-hF9-Poly A-polyA-hF9-En2SA-ITR). (Figure 9b).

To evaluate the *in vivo* KI and therapeutic effect, B6.Apoc3$^{hseq/hseq}$ were randomly divided into five groups. Then, 5 X 10^{13}GC/kg of AAV8-CRISPR and AAV8-donor (HITI or integration) were transduced into mice, and plasma hF9 was measured at 2, 6, 10, and 14 weeks after transduction. Based on the previously confirmed evidence, a comparison between each group was analyzed with plasma from 6 weeks after AAVs transduction. Briefly, groups with AAV8-donor did not induce hF8 production, and this indicated the little possibility of auto-expression. Whereas, groups with AAV8-CRISPR and AAV8-donor develop hF9 production. An average of 250 ng/ml of hF9 was detected in the plasma of the integration group, and an average of 50 ng/ml was found in the HITI group (Figure 9c). The integration group presents approximately five times higher than the HITI group. These results suggest that the integration strategy would be an appropriate strategy for *in vivo* gene editing. However, the production of the therapeutic gene is still low as 12 times smaller than that of auto-expression(Figure 9c). In human hemophilia

therapy, five percent recovery of F9 production is not enough; thus, another strategy for improving *in vivo* KI efficiency needs to be considered.

In the AAV mediated *in vivo* transduction, the expression level increased dose-dependent manner[60]. Thus, a high titer of AAVs with 2 X 10^{14} GC/kg, with which was four times higher than the first trial, was applied to therapeutic gene KI into Apoc3 locus. At two weeks after AAV transduction, plasma hF9 concentration was reached to approximately 926 ng/ml (Figure 9d). This is 3.7 times higher than the AAV transduction with 5 X 10^{13}GC/kg, and this seemed to be a dose-dependent pattern(Figure 9d). Although there was high plasma hF9 concentration, additional confirmation of KI with molecular screening is essential.

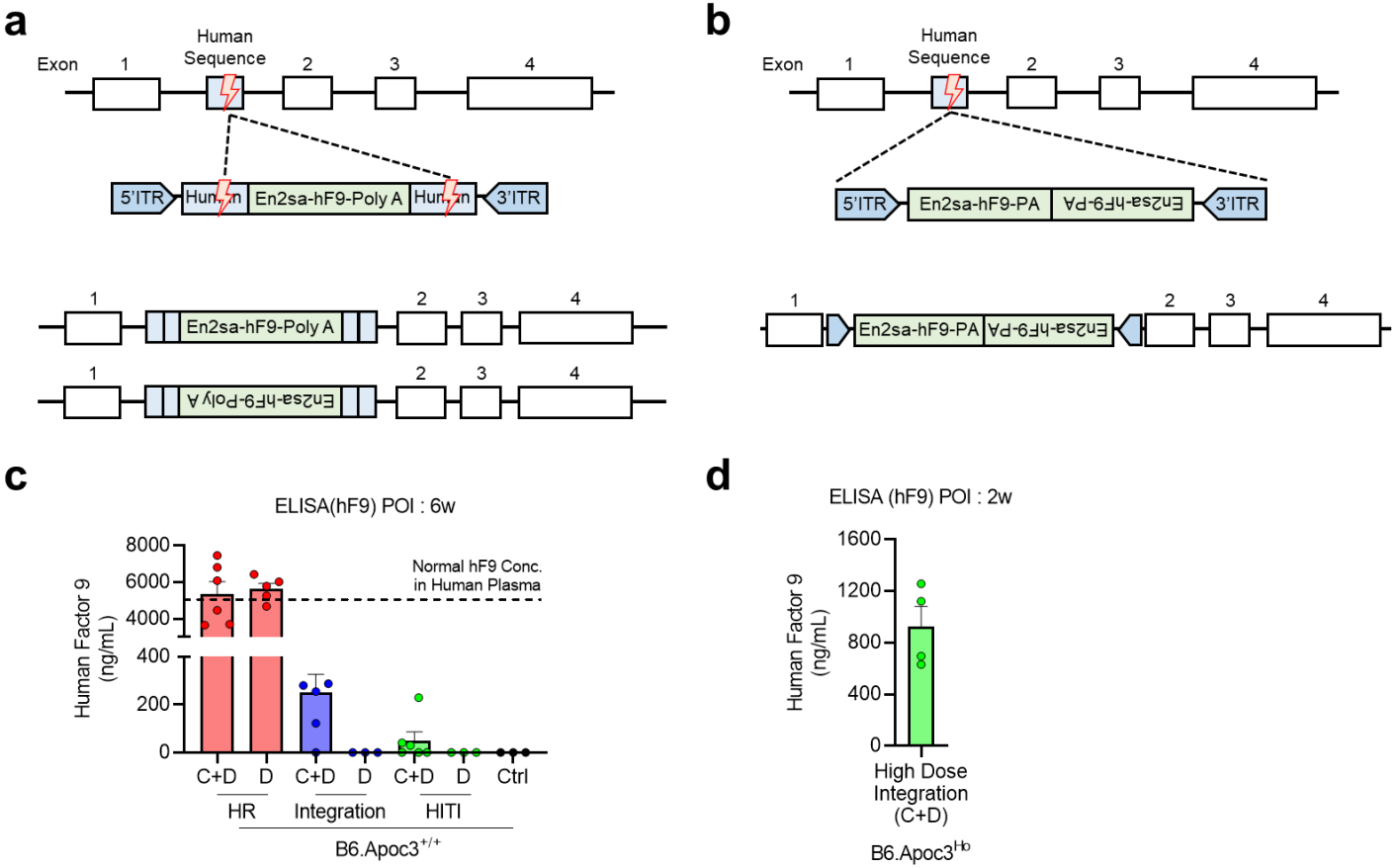

a
Exon
1
Human Sequence
2
3
4
5'ITR
En2sa-hF9-Poly A
3'ITR
En2sa-hF9-Poly A
b
Exon
1
Human Sequence
2
3
4
5'ITR
En2sa-hF9-PA
3'ITR
En2sa-hF9-PA
c
ELISA(hF9) POI : 6w
Human Factor 9 (ng/mL)
8000
6000
4000
400
200
0
Normal hF9 Conc. in Human Plasma
C+D
D
C+D
D
C+D
D
Ctrl
HR
Integration
HITI
B6.Apoc3+/+
d
ELISA (hF9) POI : 2w
Human Factor 9 (ng/mL)
1600
1200
800
400
0
High Dose Integration (C+D)
B6.Apoc3Hb

Figure 12. *in vivo* NHEJ-mediated gene editing in Apoc3$^{hseq/hseq}$ mouse **a)**Schematics showing NHEJ-mediated gene editing. a) HITI contains Human Apoc3 Sequence on both sides and has En2sa, hF9, and Poly A (AAV8-ITR-Human Apoc3-En2sa-hF9 cDNA-Poly A-ITR). **b)**The Integration Group has En2sa-hF9 cDNA-Poly A of the forward and reverses strand (AAV8-ITR-En2sa-hF9-Poly A-Poly A-hF9-En2SA-ITR). **c)**Results of *In vivo* NHEJ-mediated gene editing at week 6**.** To compare the expression level, ELISA results (post of injection: week 6) of NHEJ-mediated groups were collected. The mouse used in the experiment is B6. Apoc3$^{hseq/hseq}$ (Male or Female, N=3~6, Age=8~9w). **d)**The expression level of high dose integration at week 2. The injected dose is $2x10^{14}$GC / kg (CRISPR : Donor = 1 : 1). The mouse used in the experiment is B6. Apoc3$^{hseq/hseq}$ (Male, N=4, Age=8~9w). Each dot represents a mouse. C: CRISPR, D: donor.

4.10 sustain effects of *in vivo* AAV8 NHEJ-mediated gene editing in B6. Apoc3$^{hseq/hseq}$ mouse

The method of inserting the therapeutic gene into the host genome shows a tendency of decreasing or maintaining the decrease over time[23]. so, the tracking experiments using ELISA was performed to see if the expression level was maintained by NHEJ-mediated KI. hF9 expression was tracked through periodic plasma sampling (2, 6, 10, and 14 weeks). The expression of the integrated C + D and HITI C + D group showed a weak expression level of 250 ng/ml or less. The high dose integration C+D group showed 926 ng/ml at week 2 and an average of 619 ng/ml at week 6. From week 2 to 6, it appears to have decreased by about 34% (Figure 10).

Currently, it is difficult to judge whether KI happens or not, but positive indicators are found in the Integration C + D group. Besides, expression was confirmed in the donor only group, but the cause is unknown until now. The cause may be site-specific integration of AAV, such as AAVS1[61]. However, this hypothesis is not consistent with the report that rAAV has a lower rate of unexpected-integration than wtAAV, and integration issues are less frequent[59]. Overall, this decreasing trend may have been influenced by the immune response produced. Indeed, according to reports of the immune response to AAV capsid and hFIX[62, 63], it can be seen that in the early stage, the humoral immune response is very active. Since the results of *in vivo* genome editing were not the same for each period and showed repeated

recovery and reduction, continuous tracking is essential in NHEJ-mediated KI experiments.

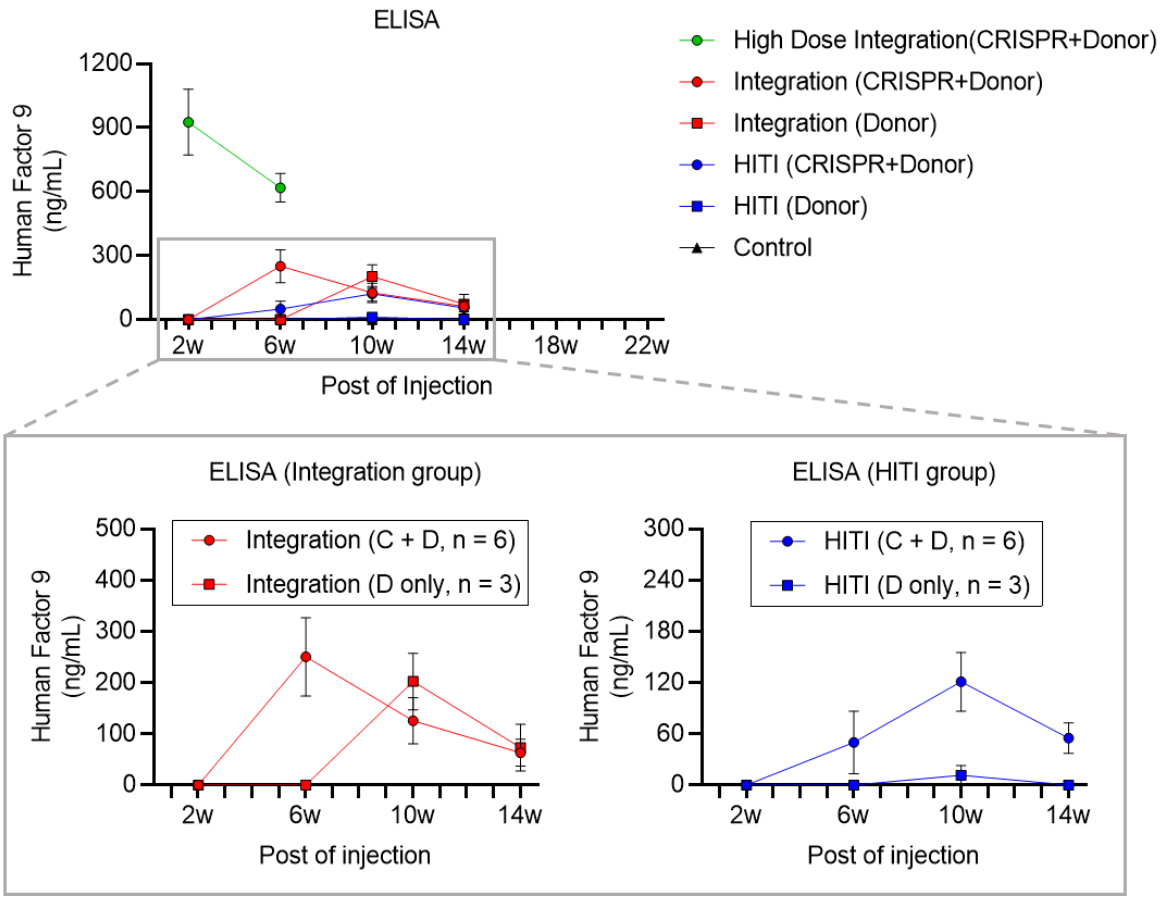

ELISA
High Dose Integration(CRISPR+Donor)
Integration (CRISPR+Donor)
Integration (Donor)
HITI (CRISPR+Donor)
HITI (Donor)
Control
Human Factor 9 (ng/mL)
1200
900
600
300
0
2w
6w
10w
14w
18w
22w
Post of Injection
ELISA (Integration group)
Integration (C + D, n = 6)
Integration (D only, n = 3)
500
400
300
200
100
0
Post of injection
ELISA (HITI group)
HITI (C + D, n = 6)
HITI (D only, n = 3)
300
240
180
120
60

Figure 13. ELISA results of NHEJ-mediated gene editing, including high dose integration (C+D) group. By tracking the hF9 concentration of Plasma through ELISA, NHEJ-mediated gene editing is compared. The mouse used for each group is B6. Apoc3$^{hseq/hseq}$ (N=3~6, Age=8~9w). Currently, a High dose Integration C+D experiment ($2x10^{14}$GC / kg) was conducted up to week 6, and the normal dose NHEJ experiment ($5x10^{13}$GC / kg) was conducted up to week 14. To confirm more specific changes, data for 2- 14w was divided into the Integration group and HITI group.

5. Discussion

This study focused on developing a new therapeutic platform for the genetic disorder. Here, I presented the production of an animal model for *in vivo* therapeutic gene insertion, and several strategies for inducing gene KI into Apoc3 locus. Although this study did not approach to strong evidence of recover the clinical symptom with *in vivo* gene KI, there were several products and interesting observations.

The previously produced animal model made possible an *in vivo* genome editing test. However, in the production of mice containing human genes, there was a problem that the size of hseq was 120 bp, which was too small to mimic humans. This size does not cover both HA of HR-donors. As such, KI attempts to insert human sequence into the mouse genome can be found in other papers, but there have been few successful cases of KI over 1kb[64]. This study using Apoc3hseq/hseq has a possibility of the *in vivo* genome editing result to be not completely mimic human study. To use for HR-mediated KI, a human sequence KI mouse with a size of HA may be required.

In addition, model mice KO-coagulation F9 has very similar to the phenotype of hemophilia b patients[40, 65]. However, the reduction of RBC value and the EMH symptoms were found in phenotypes of B6.F9 KO mouse, which is not a normal symptom of hemophilia patients[40]. However, it is well known that the decrease in RBC and EMH are mainly caused by anemia caused

by bleeding[45, 66], and mild bleeding within the hemophilia patients is frequently observed[67]. The exact cause of these symptoms has not been found but may be due to bleeding in a specific area, and bleeding in a specific area may be the leading cause of a short life expectancy of B6.F9 KO mouse.

The locus for KI in an attempt to insert a therapeutic gene through *in vivo* genome editing is the Apoc3 gene. Safe harbor gene has the advantage of having a high expression level without showing a pathological phenomenon by genetic destruction[27]. The strategy of KI into genes with high expression levels is used for albumin locus[23, 68], but it poses a risk of hypoalbuminemia due to mutation occurring in target locus[69, 70]. On the other hand, there have been positive reports that the decreasing expression level due to mutation in Apoc3 promotes the conversion of VLDL to LDL[71, 72]. Therefore, Apoc3 locus is judged to be safer in terms of side effects. Also, a change in the expression level of Apoc3 gene and a test for lipoprotein after genome editing will be necessary to reinforce these claims.

No evidence of the insertion of the hF9 gene through *in vivo* HR-mediated KI has been found. Interestingly, an unexpected auto-expression issue of the HR-donor cassette was found in the HR-mediated KI. This may be due to the left homology arm, which includes the promoter site and the components such as the TATA box and CAAT box, etc. This region is the same site mentioned in the paper comparing various species of Apoc3 promoter regions[73]. More interestingly, the donor only group maintains a high concentration for long time after injection. This concentration is the amount that can make a severe

hemophilia B patient into the range of healthy persons[40, 74]. The dose of AAV8-HR donor I used was twice as much as the amount suggested in other papers[75], but the hF9 level is 8 times higher . Also, considering the report that AAV maintains expression for more than eight years in experiments treating hemophilia B dog[76, 77], AAV8-HR donor has the potential to be used as gene therapy with a long-term therapeutic effect using a low dose.

In the NHEJ-mediated experiment, the Integration group presented a level of expression five times higher than the HITI group. That can be interpreted as an aspect of the structure. The integration donor will form a hairpin structure through homologous sequences because the two therapeutic genes are inverted and oppositely faced in single-strand form. This structure is similar to scAAV[74] and is considered to contribute to stability. Also, the knock-in of this structure contains the possibility of receiving a promoter signal irrespective of direction and the possibility of receiving a promoter signal of an anti-sense gene (Apoa1) adjacent to Apoc3. However, the HITI group has a half expression level because of dependence on directions[57, 58, 78]. In addition, Mutation cause by DSB in the HITI donor template will also affect lowering the expression level. These pieces of information support the reason why integration is higher in expression than HITI. So far, both groups have shown a decreasing level of expression during the period, so they show a completely different pattern from papers on gene editing *in vivo*. [22, 23, 53, 58]. It is difficult to predict the period of expression in the current situation because it is not known whether KI occurred or not.

The concept of this experiment is to compensate for low KI efficiency by inserting a therapeutic gene into a gene with high expression level. However, the limitation of this experiment was that we could not compare the method of KI to F9 locus and Apoc3 locus. This means that we cannot know how effective the strategy is. For this comparison, the production of mice with human sequence KI in F9 locus and *in vivo* genome editing should be included.

Through continuous sampling of groups that have attempted NHEJ-mediated KI, an experiment will be conducted to see if the expression of hF9 persists. Also, genome editing is performed on B6. Apoc3$^{hseq/hseq}$-F9 KO mouse and the therapeutic effect will also be verified. To be used to treat human genetic diseases in the future, stability evaluations will be required to include studies on off-target effects by CRISPR, untargeted integration caused by AAV, and immune responses.

Uncategorized References

1. Griggs, R.C., et al., *Clinical research for rare disease: opportunities, challenges, and solutions.* Mol Genet Metab, 2009. **96**(1): p. 20-6.
2. Kumar, P., et al., *Prevalence and patterns of presentation of genetic disorders in a pediatric emergency department.* Mayo Clin Proc, 2001. **76**(8): p. 777-83.
3. Rodriguez-Merchan, E.C., *The cost of hemophilia treatment: the importance of minimizing it without detriment to its quality.* Expert Review of Hematology, 2020. **13**(3): p. 269-274.
4. Egrie, J.C. and J.K. Browne, *Development and characterization of novel erythropoiesis stimulating protein (NESP).* Nephrol Dial Transplant, 2001. **16 Suppl 3**: p. 3-13.
5. Leader, B., Q.J. Baca, and D.E. Golan, *Protein therapeutics: a summary and pharmacological classification.* Nat Rev Drug Discov, 2008. **7**(1): p. 21-39.
6. Kontermann, R.E., *Strategies for extended serum half-life of protein therapeutics.* Curr Opin Biotechnol, 2011. **22**(6): p. 868-76.
7. Porter, S., *Human immune response to recombinant human proteins.* J Pharm Sci, 2001. **90**(1): p. 1-11.
8. Skubis-Zegadlo, J., A. Stachurska, and M. Malecki, *Vectrology of adeno-associated viruses (AAV).* Med Wieku Rozwoj, 2013. **17**(3): p. 202-6.
9. Smalley, E., *First AAV gene therapy poised for landmark approval.* Nat Biotechnol, 2017. **35**(11): p. 998-999.
10. Cunningham, S.C., et al., *Gene delivery to the juvenile mouse liver using AAV2/8 vectors.* Molecular Therapy, 2008. **16**(6): p. 1081-1088.
11. Mao, Z., et al., *DNA repair by nonhomologous end joining and homologous recombination during cell cycle in human cells.* Cell Cycle, 2008. **7**(18): p. 2902-6.
12. Block, R.C., et al., *Altered cholesterol and fatty acid metabolism in Huntington disease.* J Clin Lipidol, 2010. **4**(1): p. 17-23.
13. Darby, S.C., et al., *Mortality rates, life expectancy, and causes of death in people with hemophilia A or B in the United Kingdom who were not infected with HIV.* Blood, 2007. **110**(3): p. 815-825.
14. Muncie, H.L., Jr. and J. Campbell, *Alpha and beta thalassemia.* Am Fam

Physician, 2009. **80**(4): p. 339-44.
15. Platt, O.S., et al., *Mortality in sickle cell disease. Life expectancy and risk factors for early death.* N Engl J Med, 1994. **330**(23): p. 1639-44.
16. Saxena, K., *Barriers and perceived limitations to early treatment of hemophilia.* J Blood Med, 2013. **4**: p. 49-56.
17. Ginn, S.L., et al., *Gene therapy clinical trials worldwide to 2012 - an update.* J Gene Med, 2013. **15**(2): p. 65-77.
18. Cox, D.B., R.J. Platt, and F. Zhang, *Therapeutic genome editing: prospects and challenges.* Nat Med, 2015. **21**(2): p. 121-31.
19. Kim, H. and J.S. Kim, *A guide to genome engineering with programmable nucleases.* Nat Rev Genet, 2014. **15**(5): p. 321-34.
20. Kim, E., et al., *In vivo genome editing with a small Cas9 orthologue derived from Campylobacter jejuni.* Nat Commun, 2017. **8**: p. 14500.
21. Yin, H., et al., *Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype.* Nat Biotechnol, 2014. **32**(6): p. 551-3.
22. Wang, L., et al., *CRISPR/Cas9-mediated in vivo gene targeting corrects hemostasis in newborn and adult factor IX-knockout mice.* Blood, 2019. **133**(26): p. 2745-2752.
23. Chen, H., et al., *Hemophilia A ameliorated in mice by CRISPR-based in vivo genome editing of human Factor VIII.* Sci Rep, 2019. **9**(1): p. 16838.
24. Liu, Y., et al., *Standardizing a simpler, more sensitive and accurate tail bleeding assay in mice.* World J Exp Med, 2012. **2**(2): p. 30-6.
25. Saito, M.S., et al., *New approaches in tail-bleeding assay in mice: improving an important method for designing new anti-thrombotic agents.* Int J Exp Pathol, 2016. **97**(3): p. 285-92.
26. Dow, L.E., et al., *Inducible in vivo genome editing with CRISPR-Cas9.* Nat Biotechnol, 2015. **33**(4): p. 390-394.
27. Papapetrou, E.P. and A. Schambach, *Gene Insertion Into Genomic Safe Harbors for Human Gene Therapy.* Molecular Therapy, 2016. **24**(4): p. 678-684.
28. Sadelain, M., E.P. Papapetrou, and F.D. Bushman, *Safe harbours for the integration of new DNA in the human genome.* Nature Reviews Cancer, 2012. **12**(1): p. 51-58.

29. Jong, M.C., M.H. Hofker, and L.M. Havekes, *Role of ApoCs in lipoprotein metabolism - Functional differences between ApoC1, ApoC2, and ApoC3.* Arteriosclerosis Thrombosis and Vascular Biology, 1999. **19**(3): p. 472-484.
30. Chen, F.Q., et al., *High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases.* Nature Methods, 2011. **8**(9): p. 753-U96.
31. Renaud, J.B., et al., *Improved Genome Editing Efficiency and Flexibility Using Modified Oligonucleotides with TALEN and CRISPR-Cas9 Nucleases.* Cell Reports, 2016. **14**(9): p. 2263-2272.
32. Leroux, C., N. Mazure, and P. Martin, *Mutations Away from Splice Site Recognition Sequences Might Cis-Modulate Alternative Splicing of Goat Alpha-S1-Casein Transcripts - Structural Organization of the Relevant Gene.* Journal of Biological Chemistry, 1992. **267**(9): p. 6147-6157.
33. Tazi, J., N. Bakkour, and S. Stamm, *Alternative splicing and disease.* Biochim Biophys Acta, 2009. **1792**(1): p. 14-26.
34. Cieply, B. and R.P. Carstens, *Functional roles of alternative splicing factors in human disease.* Wiley Interdiscip Rev RNA, 2015. **6**(3): p. 311-26.
35. Chang, Y.F., J.S. Imam, and M.F. Wilkinson, *The nonsense-mediated decay RNA surveillance pathway.* Annu Rev Biochem, 2007. **76**: p. 51-74.
36. Holbrook, J.A., et al., *Nonsense-mediated decay approaches the clinic.* Nature Genetics, 2004. **36**(8): p. 801-808.
37. El-Brolosy, M.A., et al., *Genetic compensation triggered by mutant mRNA degradation.* Nature, 2019. **568**(7751): p. 193-197.
38. Rossi, A., et al., *Genetic compensation induced by deleterious mutations but not gene knockdowns.* Nature, 2015. **524**(7564): p. 230-3.
39. Kamal, A.H., A. Tefferi, and R.K. Pruthi, *How to interpret and pursue an abnormal prothrombin time, activated partial thromboplastin time, and bleeding time in adults.* Mayo Clinic Proceedings, 2007. **82**(7): p. 864-873.
40. Horava, S.D. and N.A. Peppas, *Recent advances in hemophilia B therapy.* Drug Deliv Transl Res, 2017. **7**(3): p. 359-371.
41. Ohmori, T., et al., *CRISPR/Cas9-mediated genome editing via postnatal administration of AAV vector cures haemophilia B mice.* Sci Rep, 2017. **7**(1): p. 4159.

42. Mohammed, B.M., et al., *Factor XI promotes hemostasis in factor IX-deficient mice.* J Thromb Haemost, 2018. **16**(10): p. 2044-2049.
43. Ostergaard, H., et al., *Prolonged half-life and preserved enzymatic properties of factor IX selectively PEGylated on native N-glycans in the activation peptide.* Blood, 2011. **118**(8): p. 2333-41.
44. Ross, D.W., et al., *Stability of Hematologic Parameters in Healthy-Subjects - Intraindividual Versus Interindividual Variation.* American Journal of Clinical Pathology, 1988. **90**(3): p. 262-267.
45. Kim, C.H., *Homeostatic and pathogenic extramedullary hematopoiesis.* J Blood Med, 2010. **1**: p. 13-9.
46. Waehler, R., S.J. Russell, and D.T. Curiel, *Engineering targeted viral vectors for gene therapy.* Nat Rev Genet, 2007. **8**(8): p. 573-87.
47. Chirmule, N., et al., *Immune responses to adenovirus and adeno-associated virus in humans.* Gene Ther, 1999. **6**(9): p. 1574-83.
48. Zaiss, A.K. and D.A. Muruve, *Immune responses to adeno-associated virus vectors.* Curr Gene Ther, 2005. **5**(3): p. 323-31.
49. Zincarelli, C., et al., *Analysis of AAV serotypes 1-9 mediated gene expression and tropism in mice after systemic injection.* Molecular Therapy, 2008. **16**(6): p. 1073-1080.
50. Michelfelder, S., et al., *Peptide Ligands Incorporated into the Threefold Spike Capsid Domain to Re-Direct Gene Transduction of AAV8 and AAV9 In Vivo.* Plos One, 2011. **6**(8).
51. Nakai, H., et al., *Unrestricted hepatocyte transduction with adeno-associated virus serotype 8 vectors in mice.* Journal of Virology, 2005. **79**(1): p. 214-224.
52. Yan, Z.H., H. Yan, and H.L. Ou, *Human thyroxine binding globulin (TBG) promoter directs efficient and sustaining transgene expression in liver-specific pattern.* Gene, 2012. **506**(2): p. 289-294.
53. T.Curiela, C.J.S.J.L.H.M.Y., *Long-term correction of hemophilia B using adenoviral delivery of CRISPR/Cas9.* Journal of Controlled Release, 2019. **Volume 298, 28 March 2019, Pages 128-141**.
54. Zhao, X., et al., *Cell cycle-dependent control of homologous recombination.* Acta Biochim Biophys Sin (Shanghai), 2017. **49**(8): p. 655-668.

55. Delacote, F. and B.S. Lopez, *Importance of the cell cycle phase for the choice of the appropriate DSB repair pathway, for genome stability maintenance: the trans-S double-strand break repair model.* Cell Cycle, 2008. **7**(1): p. 33-8.
56. Fausto, N. and J.S. Campbell, *The role of hepatocytes and oval cells in liver regeneration and repopulation.* Mech Dev, 2003. **120**(1): p. 117-30.
57. Shi, M., et al., *Targeted knock-in into the OVA locus of chicken cells using CRISPR/Cas9 system with homology-independent targeted integration.* J Biosci Bioeng, 2019.
58. Suzuki, K., et al., *In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration.* Nature, 2016. **540**(7631): p. 144-149.
59. Ward, P. and C.E. Walsh, *Targeted integration of a rAAV vector into the AAVS1 region.* Virology, 2012. **433**(2): p. 356-366.
60. Hu, H., et al., *Preclinical dose-finding study with a liver-tropic, recombinant AAV-2/8 vector in the mouse model of galactosialidosis.* Mol Ther, 2012. **20**(2): p. 267-74.
61. Schultz, B.R. and J.S. Chamberlain, *Recombinant adeno-associated virus transduction and integration.* Molecular Therapy, 2008. **16**(7): p. 1189-1199.
62. Mingozzi, F. and K.A. High, *Immune responses to AAV vectors: overcoming barriers to successful gene therapy.* Blood, 2013. **122**(1): p. 23-36.
63. Manno, C.S., et al., *Successful transduction of liver in hemophilia by AAV-Factor IX and limitations imposed by the host immune response.* Nat Med, 2006. **12**(3): p. 342-7.
64. Zhu, F., et al., *Humanising the mouse genome piece by piece.* Nat Commun, 2019. **10**(1): p. 1845.
65. Bowyer, A., S. Kitchen, and M. Makris, *The responsiveness of different APTT reagents to mild factor VIII, IX and XI deficiencies.* International Journal of Laboratory Hematology, 2011. **33**(2): p. 154-158.
66. Ward, J.M., J.E. Rehg, and H.C. Morse, *Differentiation of Rodent Immune and Hematopoietic System Reactive Lesions from Neoplasias.* Toxicologic Pathology, 2012. **40**(3): p. 425-434.
67. Perrin, G.Q., R.W. Herzog, and D.M. Markusic, *Update on clinical gene*

therapy for hemophilia. Blood, 2019. **133**(5): p. 407-414.

68. Sharma, R., et al., *In vivo genome editing of the albumin locus as a platform for protein replacement therapy.* Blood, 2015. **126**(15): p. 1777-1784.
69. Gatta, A., A. Verardo, and M. Bolognesi, *Hypoalbuminemia.* Intern Emerg Med, 2012. **7 Suppl 3**: p. S193-9.
70. Franch-Arcas, G., *The meaning of hypoalbuminaemia in clinical practice.* Clin Nutr, 2001. **20**(3): p. 265-9.
71. Garenc, C., et al., *Effect of the APOC3 Sst I SNP on fasting triglyceride levels in men heterozygous for the LPL P207L deficiency.* Eur J Hum Genet, 2005. **13**(10): p. 1159-65.
72. Reyes-Soffer, G., et al., *Effects of APOC3 Heterozygous Deficiency on Plasma Lipid and Lipoprotein Metabolism.* Arterioscler Thromb Vasc Biol, 2019. **39**(1): p. 63-72.
73. Januzzi, J.L., et al., *Characterization of the mouse apolipoprotein Apoa-1/Apoc-3 gene locus: genomic, mRNA, and protein sequences with comparisons to other species.* Genomics, 1992. **14**(4): p. 1081-8.
74. Wu, Z., et al., *Optimization of self-complementary AAV vectors for liver-directed expression results in sustained correction of hemophilia B at low vector dose.* Molecular Therapy, 2008. **16**(2): p. 280-289.
75. Kay, M.A., et al., *Evidence for gene transfer and expression of factor IX in haemophilia B patients treated with an AAV vector.* Nat Genet, 2000. **24**(3): p. 257-61.
76. Niemeyer, G.P., et al., *Long-term correction of inhibitor-prone hemophilia B dogs treated with liver-directed AAV2-mediated factor IX gene therapy.* Blood, 2009. **113**(4): p. 797-806.
77. Wang, L., et al., *Sustained expression of therapeutic level of factor IX in hemophilia B dogs by AAV-mediated gene therapy in liver.* Mol Ther, 2000. **1**(2): p. 154-8.
78. Nakade, S., et al., *Microhomology-mediated end-joining-dependent integration of donor DNA in cells and animals using TALENs and CRISPR/Cas9.* Nature Communications, 2014. **5**.

국문 초록

유전장애를 치료하는 노력의 일환으로 많은 증상을 환화하기 위한 방법들이 있으나, 주기적인 투여로 인해 환자의 불편 및 경제적 부담을 유발합니다. 이러한 문제점을 해결하기 위한, 생체 내 게놈 편집 분야의 많은 발전에도 불구하고, 여전히 편집 효율이 매우 낮아서 돌연변이 유전자를 직접적으로 수리하는 것에 어려움을 겪고 있습니다.

따라서, 앞선 한계점을 해결하기 위해서 간에서 높은 발현량을 나타내는 안전 항만 좌위로 치료유전자를 녹인하는 전략을 사용하여, 유전장애를 치료하기 위한 새로운 유전자 편집 전략을 개발하려고 노력하였습니다.

인간 간세포를 사용한 마이크로 어레이 실험의 결과를 토대로, 매우 높은 발현량을 가진 Apoc3 유전자가 치료 유전자를 삽입하기 위한 장소로 선택하였습니다. 그러나, 이러한 정보들은 사람에서 유래된 것이기 때문에 바로 동물실험에 적용하기에 적합하지 않았습니다. 사람의 Apoc3 유전자 내에서도, spCas9 및 cjCas9을 위한 높은 효율의 sgRNA 결합 부위를 함유하는 염기서열을 마우스 게놈내로 삽입하여, 유전적으로 인간-모방 마우스를 생산하였습니다. 생산된 B6.Apoc3-hApoc3 KI 마우스 간에서

Apoc3가 질병원인 유전자들보다 40-300배 높은 것을 알 수 있었고, 인간 염기서열의 넉인은 Apoc3의 대안적인 스플라이싱을 유발시키지 않았습니다.

이러한 B6.Apoc3-hApoc3 KI 마우스를 이용하여, 생체 내 편집 기술을 통해 치료할 유전질환은 혈우병 B입니다. 혈우병 B 마우스는 CRISPR/Cas9을 사용하여 응고인자 9 유전자의 결실을 통해 생산되었으며, 혈우병 환자와 비슷한 표현형을 나타내었습니다.

2가지의 모델 마우스의 생산 이후에, CRISPR 및 공여자 주형을 표적세포에 도입시키기 위해서 아데노 관련 바이러스를 치료 유전자의 넉인 전략에 사용하였습니다. 3가지의 전략중에 첫번째 시도에서, 동종 재조합 전략은 매우 낮은 넉인 효율과 공여자 주형에서 발견된 자동발현 문제로 인해 넉인의 증거를 확인할 수 없었습니다. 그러나, 2가지 비동종 말단 연력을 이용한 전략에서는 공여자 주형에서 자동-발현이 없는 발견되지 않았고, 약간의 긍정적인 결과를 얻을 수 있었습니다. 그러나 이 또한 낮은 발현양상을 나타낸다는 문제점을 가지고 있었습니다.

결론적으로 B6.Apoc3-hApoc3 KI 및 B6.F9 KO과 같은 생체 내 게놈 편집을 위한 모델 마우스의 생산에 성공하였고, 생체 내 게놈

편집을 시도하고 있습니다. 현재까지의 데이터에 따르면 낮은 효율로 인해 뚜렷한 넉인의 증거들을 얻지 못하였습니다. 이러한 넉인 효율을 높이기 위해 고용량 실험을 진행중에 있습니다.

감사의 글

저에게 논문은 2년간 밀린 일기를 쓰는 기분이 었습니다. 2년간 내가 무슨 일을 했는지 어떤 연구를 했는지 돌아보며, 좀더 노력했다면 좀더 많고 좋은 글을 쓸 수 있었을까 후회도 많이 했습니다. 대학원에 처음 입학했을 때에는 막막함보다는 설램이 그리고 논문을 쓸 때는 설렘보다는 막막함이 더 크게 느껴진다는 것이 제가 올바른 곳으로 가고 있는 지에 대해 고민하게 만들었습니다. 후회와 고민 속에서도 무사히 졸업논문을 다 쓸 수 있었던 것은 주변의 도움이 있었기 때문이라 확신합니다.

어찌보면 가장 아픈 손가락인 저에게 끝까지 포기하지 않아주신 염수청 교수님께 정말로 많은 것을 배웠습니다. 실험 뿐만 아니라 사람을 대하는 태도와 자세의 중요성, 존중과 배려까지 정말로 많은 것을 보여주셨습니다. 제일 사고뭉치지만 앞으로라도 교수님을 닮고자 노력하겠습니다. 항상 열정적으로 피드백 해주시는 박중훈 교수님께 살아가는 호탕함을 배우고, 항상 나긋하게 웃어주시는 김태민 교수님은 겸손을 배웠다고 생각합니다. 교수님들 정말로 감사합니다.

그리고 우리 사고뭉치 실험실 구성원들에게 항상 싸우지만 친구같이 지내줘서 감사합니다. 모두 앞으로도 잘부탁드립니다.