



Cytosine base editing systems with minimized off-target effect and molecular size

李, 昂

(Degree)

博士 (科学技術イノベーション)

(Date of Degree)

2024-03-25

(Date of Publication)

2025-03-01

(Resource Type)

doctoral thesis

(Report Number)

甲第8962号

(URL)

<https://hdl.handle.net/20.500.14094/0100490187>

※ 当コンテンツは神戸大学の学術成果です。無断複製・不正使用等を禁じます。著作権法で認められている範囲内で、適切にご利用ください。



Doctoral Thesis

博士論文

**Cytosine base editing systems with minimized off-
target effect and molecular size**

(小型かつ高精度な塩基編集技術の実現)

2024-1

Graduate School of Science, Technology and Innovation

Department of Science, Technology and Innovation

Kobe University

Li Ang

李昂

Abstract

Based on the further elucidation of bacterial nucleases, the past decades have witnessed rapid growth in the genome editing field. Genome editing has provided us with an extraordinary understanding of the relationship between genetic information and versatile disease by creating much more accurate cellular and animal models of pathological processes. These facile and efficient genome-editing platforms have begun to reveal exceptional potential in a variety of fields, ranging from the biotechnology industry to biomedical research.

In Chapter 1, we discussed the background and current research bottlenecks of genome editing field. As we know, the development of engineered nucleases, for instance, zinc-finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), and clustered regularly interspaced short palindromic repeat (CRISPR)–Cas-associated nucleases, has significantly accelerated the progress of genome editing from concept to clinical practice. However, the double-stranded DNA (dsDNA) breaks by these engineered nucleases result in high cellular toxicity limiting the applications of those tools. Cytidine base editor (CBE) as a novel CRISPR–Cas9 based genome editing platform that emerged in 2017 enables single or multiple targeted C•G-to-T•A conversions in a genomic DNA without introducing dsDNA breaks which promises much safer editing and extends applications in a wide variety of species like microbiome. Target-AID is one of the most researched CBEs, which employs sea lamprey cytidine deaminase PmCDA1 and SpCas9 (*staphylococcus pyogenes* Cas9). Owing to its low toxicity, this system has been used for precise point mutations in a variety of organisms such as bacteria, plants, and mammalian cells. Despite promising progress achieved, similar to the other type of base editing systems, Target-AID often exhibits the genome-wide off-target effect which seems to be involved in the non-specific DNA binding affinity of PmCDA1. Besides, the large molecular size of the current Target-AID also hinders its efficient loading to the drug delivery system.

In Chapter 2, we presented the experimental design and methods in this study. All evaluated target sites for both yeast and mammalian cells were shown, and primers for deep sequencing were also summarized in a table.

In Chapter 3, we showed the approaches and results of our experiments and discussed the findings in Chapter 4. In order to reduce its off-target effects for precise genome editing, we engineered PmCDA1 to diminish its non-specific binding affinity by intensive and careful truncation. Moreover, we investigated the overall interacting patterns between the engineered PmCDA1 and Cas9,

successfully generating two exquisite systems, Target-AID2S and Target-AID3S. These novel systems revealed a dramatically mitigated the gRNA-dependent/independent DNA off-targets, whilst exhibited equivalent or superior on-target editing efficiency. Also, the size of engineered PmCDA1 (121 amino acids) in Target-AID2S, 3S is only about half of conventional Target-AID (208 amino acids). In addition, in vivo delivery has been one of the major bottlenecks to achieving efficient and specific editing at the target tissue, and smaller molecular size is advantageous especially for in vivo delivery tools such as adeno-associated virus (AAV) vector. AAV is one of the promising delivery methods for gene therapy with greater safety and efficiency. However, the plasmid loading size is rigorously restrained, and the current Target-AID system is too big to be loaded. Therefore, we further expanded our protein engineering approaches into SaCas9 (*staphylococcus aureus* Cas9) and generated SaAID2S, 3S that revealed comparable on-target activity with much smaller DNA in size that able to successfully load onto AAV. Then, in order to verify the ability of engineered SaAID2S, 3S function in gene therapy, we generated a series of viral particles AAV-SaAID-2S (A2S-1 to A2S-8) and AAV-SaAID-3S (A3S-1 to A3S-7). Although no significant editing was observed for any of A3S series tested, we found significant editing efficiency for the A2S series, especially those with smaller compositions.

In Chapter 5, we proposed the innovation strategies for business based on the new technologies developed in this study. By analyzing the technology, intellectual property, business, and financial strategies as a series of strategic cycles and demonstrating the potential of Target-AID2S, 3S in the treatment of HBV, we believe our proposal will lead BioPalette to renegotiate a new cross-license agreement with Beam Therapeutics, granting them greater authority over the application of base editing technology to expand their business territory.

In Chapter 6, we provided brief conclusions about the findings of this study and the work performed. In conclusion, through protein engineering and in-depth screening in yeast and mammalian cells, we have successfully generated the novel base editing systems Target-AID2S, 3S, which are characterized by limited molecular size and high editing fidelity. In addition, by introducing SaCas9, we also successfully generated the first single-AAV-mediate cytosine base editing tool, and obtained the desirable on-target efficiency. Also, by reviewing and comparing the current gene editing technologies from different groups and Target-AID2S, 3S under the real gene therapy scenario, we believe our study contains a great potential for practical applications, and we hope our work could accelerate the application of gene editing technologies using in the therapy.

要約

細菌性ヌクレアーゼのより詳細な解明に基づき、過去数十年間でゲノム編集分野は急速に成長してきた。ゲノム編集は、細胞ゲノム上の特定の部位を改変する技術であり、疾患モデル細胞や動物を効率的に構築できるため、遺伝情報と多様な疾患との関係についてより深い理解を提供する。これらの簡便で効率的なゲノム編集プラットフォームは、バイオテクノロジー産業から生物医学研究まで、さまざまな分野での卓越した应用能力を示し始めている。

第1章では、ゲノム編集分野の背景と現在の研究のボトルネックについて述べた。すでに知られるように、ZFN (zinc-finger nuclease)、TALEN (transcription activator-like effector nuclease)、CRISPR (clustered regularly interspaced short palindromic repeat) -Cas-associated nucleases などの人工ヌクレアーゼの開発により、ゲノム編集技術の臨床への進展が著しく加速された。しかしながら、これらのツールにより生じる二本鎖 DNA の切断 (Double strand breaks: DSBs) は高い細胞毒性をもたらし、細胞の種類に依存するなど応用を制限している。2017 年に登場した CRISPR-Cas9 ベースの新規遺伝子編集プラットフォームとしての Cytidine base editor (CBE) は、DSB を導入することなく、ゲノム DNA 中の単一または複数の標的部位において C•G-to-T•A 変換を可能にし、より安全な編集を約束するとともに、より多くの種類の生物種への応用が拡大した。Target-AID は、ヤツメウナギのデアミナーゼ PmCDA1 と SpCas9 (*Streptococcus pyogenes* Cas9) を用いた、最も研究されている CBE の一つである。このシステムは細胞毒性が低く、バクテリア、植物、哺乳類細胞など、さまざまな生物において精密な点変異導入に用いられている。しかし、他の塩基編集技術と同様に、Target-AID は頻繁にゲノムワイドな off-target 効果を示し、これは PmCDA1 の非特異的 DNA 結合親和性に関連すると思われる。さらに、医療応用の面では現在の Target-AID は分子サイズが大きいため、ドラッグデリバリーシステムへの効率的な導入が難しい。

第2章では、本研究で行った実験デザインと方法を提示した。酵母と哺乳類細胞のすべての評価標的部位を示し、ディープシーケンスで使ったプライマーもテーブルにまとめた。

第3章では、我々の実験のアプローチと結果を示し、第4章ではその結果について考察した。まず、Off-target 効果を抑制し Target-AID による精密なゲノム編集を行う目的で、我々は PmCDA1 の非特異的な結合親和性を持っている部分に着目し、PmCDA1 のタンパク質エンジニアリングを行った。また、PmCDA1 と Cas9 の結合方法を検討し、最終的に Target-AID2S と Target-AID3S という2つの高性能・高精度システムを完成した。これらの新規シ

システムでは、on-target の編集効率を保ったまま、gRNA 依存的／非依存的な DNA off-target を劇的に減少させることを実現できた。さらに、Target-AID2S, 3S の PmCDA1 のサイズ (121 アミノ酸) は、従来の Target-AID (208 アミノ酸) の約半分まで小型化された。In vivo デリバリーは、標的組織で効率的かつ特定の編集を達成するための大きなボトルネックの一つとなっている。アデノ随伴ウイルス (AAV) は、安全で効率的な遺伝子治療のための有望なデリバリー方法の一つであるが、プラスミドの搭載サイズが厳しく制限されており、従来の Target-AID システムは搭載することができない。そこで、SpCas9 のオルソログである SaCas9 (*staphylococcus aureus* Cas9) を採用し、小型の SaAID2S, 3S を作製し、AAV への搭載に成功した。AAV に搭載した SaAID2S, 3S の機能を検証するために、一連のウイルス粒子 AAV-SaAID-2S (A2S-1～A2S-8) と AAV-SaAID-3S (A3S-1～A3S-7) を作製し、動物細胞で評価を行った。結果として、A3S シリーズでは編集が観察されなかったが、A2S シリーズでは有意な編集効率が観測できた。

第 5 章では、本研究で開発した新技術をベースに、事業化ためのイノベーション戦略を提案した。技術戦略、知財戦略、事業戦略、財務戦略の 4 つの戦略を一連のサイクルとして、Target-AID2S, 3S の遺伝子治療分野での優位性を分析・実証し、Beam therapeutics 社との新たなクロスライセンス契約の交渉につなげる。最終的には、バイオペレット社が塩基編集の事業化のために必要となる知財を既に集約しているマイクロバイオー姆分野に加えて、遺伝子治療分野においてアジア地域を対象とする知財基盤を確立することを目指す。

第 6 章では、この研究で得られた成果と実施した研究について簡単な総括を述べた。結論として、我々はタンパク質工学から始まり、酵母および哺乳類細胞でのインテンシブなスクリーニングにより、分子サイズが小さく、編集精度が高い新規塩基編集システム Target-AID2S, 3S の作製に成功した。さらに、SaCas9 を採用することで、世界初のシングル AAV 搭載可能なシトシン塩基編集ツールの作製にも成功し、望ましい on-target 効率を得ることができた。医療応用において、本研究で開発した Target-AID2S, 3S 技術は既存の主要なゲノム・塩基編集に対して高い技術的優位性を有しており、この技術は次世代の遺伝子治療プラットフォームの開発に貢献できると期待される。

Contents

1. Introduction

1-1. Research Background

1-1-1. Genome editing

1-1-2. Engineered nucleases

1-1-3. Zinc-finger nucleases (ZFNs)

1-1-4. Transcription activator-like effector nucleases (TALENs)

1-1-5. Clustered regularly interspaced short palindromic repeats (CRISPR-Cas9)

1-1-6. Base editing technology

1-1-7. Cytidine base editor (CBE)

1-1-8. Adenine base editor (ABE)

1-1-9. Delivery system

1-2. Challenges of base editing technology and our approach

1-2-1. Off-target effect

1-2-2. Large protein size

2. Methods

2-1. Protein modeling

2-2. Cloning and plasmid preparation

2-3. Yeast BY4741 culture and transformation

2-4. Yeast BY4741 transformation related media

2-5. Yeast BY4741 induction and spotting

2-6. Yeast BY4741 selection plate

2-7. Yeast BY4741 induction related media

2-8. Yeast assay targets

2-9. Mammalian HEK293T cell culture

2-10. Mammalian HEK293T cell transfections

2-11. Cell Sorting

2-12. Genome extraction

2-13. AAV production and infection

- 2-14. High-throughput sequencing of genomic DNA and data analysis
- 2-15. Mammalian HEK293T cell assay targets

3. Results

- 3-1. Protein engineering and yeast assay
 - 3-1-1. Elimination of DNA-binding region
 - 3-1-2. Target-AID2S, 3S
- 3-2. Mammalian HEK293T cell assay
 - 3-2-1. Variants and methodology for comparison
 - 3-2-2. On-target performance
 - 3-2-3. Editing window
 - 3-2-4. Off-target occurrence
 - 3-2-5. On/off-target profile
 - 3-2-6. Apply into SaCas9
 - 3-2-7. Load onto AAV vector

4. Discussion

5. Innovation strategy research (イノベーション・ストラテジー研究)

- 5-1. イノベーション・ストラテジー概要
- 5-2. ゲノム編集技術に関する知財の状況
 - 5-2-1. CRISPR 基本特許における特許紛争
 - 5-2-2. 塩基編集の基本特許
 - 5-2-3. バイオパレット社の設立と知財の状況
 - 5-2-4. 医療分野におけるゲノム編集スタートアップの概要と知財権利関係
 - 5-2-5. バイオパレット社と Beam 社のクロスライセンス
 - 5-2-6. 新たな特許出願：Target-AID2S, 3S 特許
- 5-3. バイオパレット社の事業戦略：塩基編集の事業環境分析とクロスライセンス再構築
 - 5-3-1. ヒューマンセラピー分野の事業環境分析
 - 5-3-2. ヒューマンセラピー分野のパイプライン分析
 - 5-3-3. 遺伝子細胞治療薬における需要と市場の継続的な成長（予測）

5-3-4. クロスライセンスの再構築

5-4. クロスライセンス再構築のための技術戦略と財務戦略

5-4-1. Target-AID2S, 3S×B 型肝炎ウイルス (HBV)

5-4-2. Target-AID2S, 3S 技術による HBV ウイルスの失活および感染力低下実験

5-4-3. 研究開発のロードマップ

5-4-4. インラボ研究開発コストの試算

5-4-5. 臨床試験パートナーの検討

5-4-6. 臨床試験コストの調査

6. Conclusion

- References

- Acknowledgments

1. Introduction

1-1. Research Background

< 1-1-1. Genome editing >

Genome editing also known as DNA/RNA manipulating technologies is a type of genetic engineering research method that helps researchers to better elucidate the function of target DNA. In the early development stage, to elicit the functional DNA, researchers commonly adopted external stimuli like UV to randomly insert genetic material into a host genome, but it is very ineffective. Due to the discovery of bacterial and engineered nucleases, genome editing technologies have been developed at a rapid pace over the past 10 years and allowed us to focus on the specific DNA rather than mutate them randomly. Starting from basic experimental strategy, agriculture, livestock industry, biotechnology industry, to pharmaceuticals and medicine, genome editing revealed a wide potential of using and we believe it will continue to attract world attention for the next decade.

< 1-1-2. Engineered nucleases >

The main concept and advantage of engineered nucleases is the ability to create DNA double-stranded break (DSB) and cleave DNA at a specific point. Thus far, in order to cleave DNA, the restriction enzymes that bacteria have in nature were commonly used, but generally recognition site is required and usually result in multiple sites cutting. To overcome this challenge, scientists developed engineered nucleases that can cleave arbitrary DNA, not only for lower animals such as nematodes but also for animal cells such as zebrafish, rats, and fertilized eggs. Currently, Zinc finger nucleases (ZFNs), transcription-activator like effector nucleases (TALEN), and the clustered regularly interspaced short palindromic repeats (CRISPR/Cas9) system are the most researched engineered nucleases.

< 1-1-3. Zinc-finger nucleases (ZFNs) >

ZFN is considered as the first-generation engineered nuclease reported in 1996. The concept behind ZFNs is based on the combination of a recognition zinc finger domain and a DNA nuclease FokI domain. Zinc ion helps to stabilize and organize the three-dimensional structure, while the C-terminal part of each finger is to target the specific DNA sequence. Besides, a nuclear localization signal is needed and is usually connected to the N-terminal. In the beginning, the recognized sequences consist

of around 3 base pairs, with recruiting multiple zinc finger domains it can elongate and obtain the competence to mark specific sequences around 18 to 36 base pairs. To date, ZFNs have been reported as the editing tool that can be used in cultured cells, microorganisms, animals, and plants. However, the requirement for designing two binding sites and difficulty to obtain desirable activity let ZFNs still hard to apply into industrial related fields¹. Over the past decade, ZFN-based genome editing has been reported to correct disease-causing mutations in sickle cell disease, antitrypsin deficiency, hemophilia B, and Parkinson's disease²⁻⁶.

< 1-1-4. Transcription activator-like effector nucleases (TALENs) >

TALENs share a similar idea to ZFN which also contains a DNA recognition TALEN domain and a cleaving domain. TALENs was developed in 2010, and it was produced by the plant pathogen *Xanthomonadaceae*. The DNA binding motif consists of highly conserved 34 amino acids as one unit and capable to target a single DNA nucleotide. FokI is introduced as the nuclease enzyme and connected to C-terminal, while the nuclear localization signal is designed to the N-terminal. In contrast to ZFNs, TALENs are easier to produce, retain the higher DNA binding specificity, and preserve relatively lower off-target effects¹. TALEN is widely used in genetic research and has shown potential for a variety of applications, including crop improvement, disease modeling, and potential therapeutic interventions⁷⁻¹⁰. On the other hand, due to the large size of TALEN, it cannot be easily packaged in viral vectors, which raises challenges for cell delivery.

< 1-1-5. Clustered regularly interspaced short palindromic repeats (CRISPR-Cas9) >

CRISPR-Cas9 is known as the powerful genome editing system derived from a bacterial adaptive immune defense system. This system maintains a robust and programmable activity to cleave the genome of eukaryotic cells via an RNA-guided (gRNA) module. Once it emerged in 2012, it gained high expectations from world-wide researches due to its facility and efficiency to use and served as the potential alternative method to ZFNs and TALENs. Among the Cas enzymes identified, currently most widely used as a genome editing technology is the one derived from *Streptococcus pyogenes* (SpCas9). To gain appreciable editing, a sequence called protospacer adjacent motif (PAM) is required adjacent to the target sequence. In addition, nCas9 and dCas9 which only conserved one of the two cleavage domains and inactivated both cleavage domains respectively have also been developed¹. Besides SpCas9, SaCas9 (*staphylococcus aureus* Cas9) has also been reported as a member of the

Cas9 family that can be used in gene editing systems. Due to its different recognize site and smaller in size, SaCas9 expands the application scope and reveals advantages for easier to deliver¹¹. The major advantages of CRISPR-Cas9 technology are its simplicity and versatility. Unlike previous genome editing technologies, CRISPR-Cas9 technology is relatively easy to design and implement, and has been used in a wide range of organisms, including bacteria, plants, animals, and human cells^{12–14}. It is also widely used in scientific research to study gene function, develop disease models, and elucidate genetic diseases^{15–19}. CRISPR-Cas9 also shows great potential for therapeutic applications, including the treatment of genetic diseases, immune cell engineering for cancer immunotherapy, and the development of new treatments for a variety of medical conditions. Doudna and Charpentier, who developed CRISPR-Cas9, will receive the Nobel Prize in Chemistry in 2020.

In addition to the commonly used CRISPR-Cas9 technology, other novel technologies have also been developed in recent years, including Cas12, also called CRISPR-Cas12 or CRISPR-Cpf1, a variant of the CRISPR technology derived from a different bacterial species than the well-known CRISPR-Cas9 technology. Cas12 is a single RNA-guided endonuclease that handles its own distinct gRNA, requiring only crRNA for targeting, and has the advantage of being a smaller overall package than Cas9²⁰.

< 1-1-6. Base editing technology >

Base editing is a fourth-generation genome editing technology developed based on CRISPR-Cas9. Unlike conventional genome editing methods such as CRISPR-Cas9, it does not require double-strand breaks in DNA. Instead, it allows for the direct modification of individual bases (nucleotides) within the genome. Base editing utilizes the enzymes dCas9 or nCas9, fused with a deaminase protein, to introduce single-base variations (SNVs). Since approximately half of the known gene mutations causing hereditary diseases are caused by SNVs, base editing holds significant potential in the treatment of many genetic disorders through the alteration of either temporary RNA or permanent DNA bases. There are various types of base editing techniques, broadly classified into DNA base editing and RNA base editing. The origins of base editing technology date back several decades, starting with RNA base editing, and more recently, both types of base editing have rapidly gained research popularity²¹. DNA base editing technologies can be further divided into cytosine base editing (CBE) and adenine base editing (ABE) technologies; both CBE and ABE are powerful genome editing tools for the permanent introduction of point mutations in DNA with high efficiency in living cells.

< 1-1-7. Cytidine base editor (CBE) >

Base editing tool CBEs employed cytidine deaminase as a functional group and combined with paired or impaired Cas9 protein was developed. Collaborating with gRNA to localize target, CBEs revealed the high conversion rate of C•G base pair (bp) to a T•A (bp)²² without introducing a DSB. In nature, DNA base conversions can act as mutagens under specific circumstances. For example, cytosine deamination can lead to the formation of uracil, which is structurally similar to thymine and might be mistakenly recognized, potentially causing a C to T mutation. However, such mutations are typically repaired through the base excision repair (BER) pathway, and only a small fraction becomes mutations. On the other hand, the acquired immune system of vertebrates is thought to actively utilize deamination as a mechanism. This serves as the basis for somatic hypermutation, enabling the variation of gene locus sequences to allow antibodies, as immunoglobulin molecules, to recognize various antigens. The enzyme central to this process is called Activation-Induced Cytidine Deaminase (AID). AID is specific to single-stranded DNA and targets regions where single-stranded DNA is exposed in R-loop structures when immunoglobulin domains are transcribed frequently²³. In this context, by using a deactivated Cas9 (dCas9) that retains the guide RNA binding capability without the intrinsic nucleolytic activity of CRISPR-Cas9 protein, and binding it with a deaminase, it becomes possible to achieve base conversions from C•G to T•A. Usually, CBEs fuse with one or more copies of uracil glycosylase inhibitor (UGI)²⁴ to disturb cell repairability for achieving high mutation efficiency. At the first stage, cytidine is deaminated into uracil and then replaced with thymine by conducting with cellular repair mechanism. Target-AID²⁵ and BE²² are the two most researched CBEs, Target-AID employs the deaminase PmCDA1 from lamprey eels, while Base editor incorporates a deaminase from rats. Depending on the nature of these deaminases, the DNA editing profiles are slightly different, the editing window is 16-20 bases upstream of the PAM sequence for Target-AID and 12–16 bases upstream of PAM for BE. CBEs have displayed remarkable single-nucleotide polymorphism (SNP) conversion with low levels of indels and deletions through editing numerous cell types. Collectively, CBEs exhibited better utilizability and safety compared with traditional CRISPR-Cas9 systems.

< 1-1-8. Adenine base editor (ABE) >

The ABE (adenine base editing) technology emerged in 2017. Inspired by CBE (cytosine base editing), ABE leverages the deamination reaction of adenosine to generate inosine. This inosine can be subsequently modified to guanine during the replication and transcription processes. One of the appealing aspects of ABE is its ability to correct mutations from C·G to T·A, which represents the most common pathogenic single-nucleotide variation (SNV) reported in the ClinVar database²⁶. While naturally occurring adenosine and adenosine deaminases exist, their substrates are limited to various forms of RNA, and they cannot target DNA. Consequently, the ABE activity of various naturally occurring adenosine deaminases, such as *Escherichia coli* TadA (or ecTadA), human ADAR2, mouse ADA, and human ADAT2, was tested, but none exhibited base editing from A to G above background levels. Therefore, researchers evolved the necessary enzymes from ecTadA, ultimately achieving successful editing. The invention of ABE further expanded the scope of the base editing application²⁷.

< 1-1-9. Delivery system >

When applying genome editing technology to in vivo practice, the way of delivering the editing proteins to the target cells/tissues is very important. Currently, viral delivery and non-viral delivery are the two main delivery systems. Among the viral delivery family, the most studied are lentiviruses, adenoviruses and adeno-associated viruses (AAV). Lentivirus vector has a cargo capacity of up to 8 kb and stable transduction activity, while their genome is randomly integrated into the host chromosome, which poses risks such as inactivation of tumor suppressor genes. Adenovirus vectors carry less risk of genomic insertion and allow for larger cargos, but their application has stalled due to their high immunogenicity. AAV vectors have shown appreciable transduction efficiency and long-term transgene expression in various tissues while maintaining a low level of immunogenicity, although its cargo capacity of the expression cassette is only to 4.7 kb. On the other hand, non-viral delivery systems such as nanoparticles can achieve DNA transport capabilities in excess of 10 kb, but the cytotoxicity and specificity of synthetic polymers is a major limiting factor for their in vivo application.

1-2. Challenges of base editing technology and our approach

< 1-2-1. Off-target effect >

Safety is always an important issue when it comes to the application of base-editing techniques. Although BE and Target-AID, introduced in the previous chapter, had greatly expanded their application in gene manipulation, it was also found that off-target also existed. Off-target refers to unintended editing that occurs at genomic sites other than the targeted target site during the genome editing process. Of course, there raises the possibility of various side effects if unintended edits are made, and this is an issue that must be resolved, especially for applications that require high editing precision like gene therapy.

Analogous to gRNA-dependent off-target which come from the miss-match binding of Cas9 (Cas9 is thought to be strictly regulated by the 20-nucleotide guide sequence of the gRNA and the presence of PAMs adjacent to the target sequence in the genome, potential off-target cleavage activity can also occur in DNA sequences with a 3-5 base pair mismatch in the PAM-distal portion of the gRNA guide sequence). In contrast to gRNA-dependent off-target, base editing also revealed genome-wide gRNA-independent off-target. (Fig. 1). The previous study was observed that cytosine base editing systems installed the genome-wide gRNA-independent off-target in both rice cell and mouse embryos^{28,29}. By reviewing Yang, Gao²⁹, Ye³⁰, and their respective coworkers' report, and the crystal structure of human AID protein³¹, we defined these off-target edits are possibly arisen from the intrinsic binding affinity of deaminase domain to non-specific dsDNA, consistent with the found demonstrated in Junjie Tan³² and their respective coworkers' report, obtained highly off-target reduction by truncating the C-terminal of PmCDA1 protein. Subsequently, we investigated the 3D model of AID protein (share a similar crystal structure to PmCDA1) and characterized the binding domain to non-specific dsDNA (Fig. 2). Interestingly, we found the truncation pattern reported in the previous study³² only removed a small amount of binding domain. Therefore, we hypothesize the further truncation of PmCDA1 could facilitate the reduction of off-target editing even more.

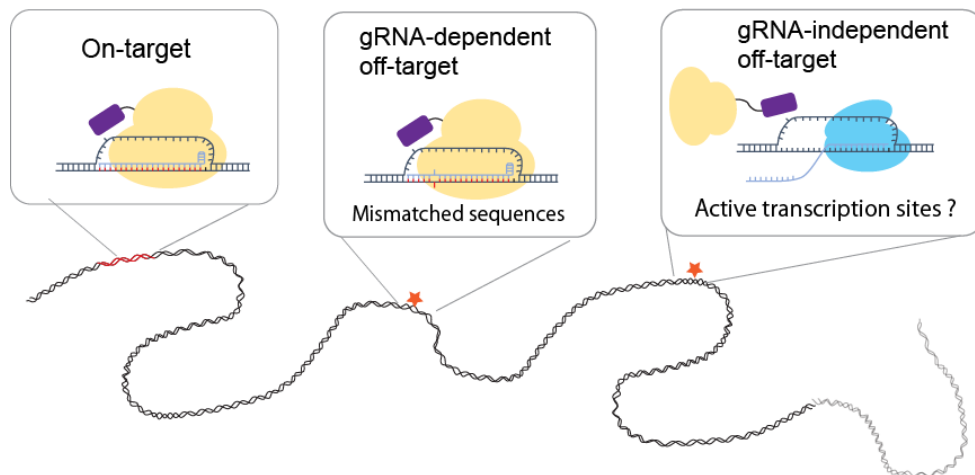


Fig. 1_Mechanism of gRNA-dependent off-target and gRNA-independent off-target

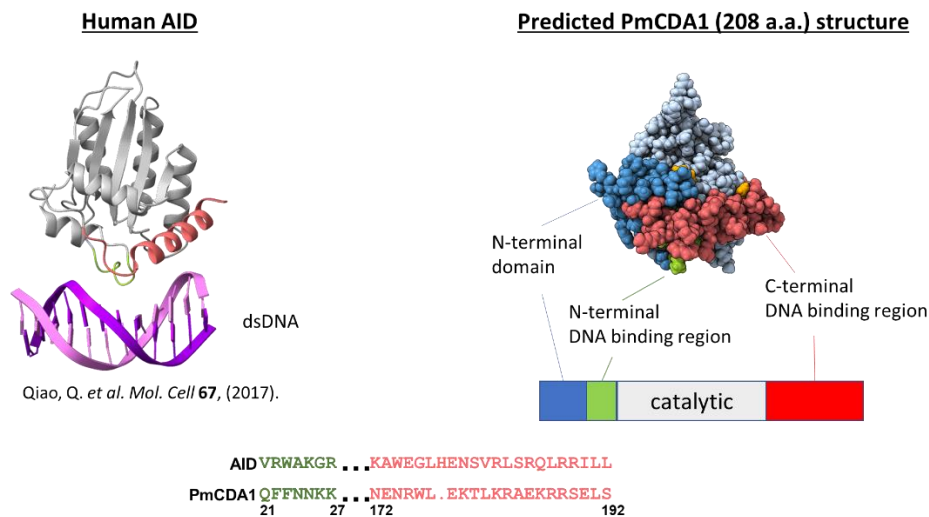


Fig. 2_Potential nonspecific dsDNA binding of PmCDA1 base on the AID-dsDNA interaction protein model (showing in green and red)

< 1-2-2. Large protein size >

By considering utilizing ability in clinics, the delivery method is the one problem obsessed with researches. As mentioned above, AAV presents low toxicity, lack of pathogenicity, and is capable of maintaining long-term transgene expression through chromosomal integration proved itself as a reliable method to use in gene therapy³³. However, the capacity of the plasmid loading size is strictly restricted (4.7 kb), and the current base editing systems are too big to be loaded (Target-AID, BE > 5 kb). For conventional genome editing, smaller CRISPR ortholog SaCas9 has led to the development

of single AAV vector^{11,34}, but adding a deaminase domain has been challenging³⁵. Instead of composing a single AAV vector, the base editing components could be split into two AAV vectors to circumvent the size limitation³⁶. However, in this case, the amount of AAV used for treatment must be doubled, thereby increasing the risk of immune response.

To tackle these two challenges and improve current base editing systems (we focus on Target-AID system), here we described a protein architecture rearrangement method that focusing on truncating and engineering PmCDA1 protein. In order to meet the desired plasmid loading size in AAV, we not only just completely removed the affinity domain to non-specific dsDNA but also truncated the non-functional domain as much as possible along with amino acid replacements. Encouragingly, the novel system Target-AID2S (Small and Specific) and Target-AID3S (Small, Specific and Superior) constructed with half more molecular size-down and facilitated spatial structure revealed a dramatically mitigated the off-target effect (down to 24~40 folds in yeast, 4~14 folds in HEK293 cell) compare to conventional Target-AID, whilst exhibited equivalent or superior on-target editing efficiency. We further combined these novel systems with SaCas9 to create SaAID2S and 3S, which maintain desirable on-targets at smaller molecular sizes. Finally, the world's first single AAV-mediated base editing system, AAV-SaAID-2S (A2S-8), was successfully generated, and revealed comparable editing efficiency with tradition method that recruits dual-AAV.

2. Methods

2-1. Protein modeling

Prediction of the protein structure of PmCDA1 was done by I-TASSER29 based on the homology with human AID³¹ (PDB: 5W1C). The analysis of protein model was done by ChimeraX³⁷.

2-2. Cloning and plasmid preparation:

JM109 chemically competent E. coli were used for cloning and preparation of the plasmids by using FastGene Plasmid Mini Kit or NucleoSpin® Plasmid Transfection-grade for transfection. For the yeast experiments, Target-AID variant constructs were made by modifying the original Target-AID vector pRS315e_pGal-nCas9(D10A)-PmCDA1 (Addgene #79617)²⁵. Plasmids for gRNA expression were made from p426-SNR52p-gRNA.CAN1.Y-SUP4t (Addgene #43803)³⁸ by replacing the target sequence. For mammalian transfection, plasmid constructs were generated based on YE1-BE4max (Addgene #138155)³⁹. SaCas9-AID versions were made from pX601-AAV-CMV::NLS-SaCas9-NLS-3xHA-bGHpA;U6::BsaI-sgRNA(Addgene #61591)⁴⁰. ScpI promoter⁴¹ and ploy-A tail⁴² were synthesized by Eurofin. For R-loop assay, nickase SaCas9(D10A) vector was developed from dead SaCas9 (Addgene #138162) by replacing the promoter with ScpI. The U6 promoter gRNA plasmid was constructed based on pU6-Sp-pegRNA-RNF2 (Addgene # 135957). The AAV related plasmid Split-532 (Addgene # 137182, 137183) and Split-739 (Addgene # 119943, 119944. Note: nSaCas9(KKH) arranged to nSaCas9) were introduced for comparison. DNA was PCR-amplified by PrimerSTARMax polymerase (TaKaRa) followed by gel extraction (FastGene Gel/PCR Extraction Kit). Gibson assembly follows the reported protocol³⁴ and Ligation high Ver.2 (TOYOBO) was used for ligation reactions. DNA was amplified using PrimeSTAR Max DNA Polymerase (Takara Bio) through PCR reaction.

< Reaction >

2×Prime STAR MAX	25 µl
dH ₂ O	22 µl
Template DNA	1 µl
10 µM forward primer	1 µl
<u>10 µM reverse primer</u>	<u>1 µl</u>

Total	50 μ l	
< PCR condition>		
98 °C	1 min	
98 °C	10 sec	 x 25 cycle
55 °C	5 sec	
72 °C	10 sec/kb	
72 °C	1 min	
10 °C	∞	

All plasmid in this study generated by using either Gibson-assembly or Restriction-enzyme ligation.

< Gibson-assembly >

DNA fragment 1	2 μ l
DNA fragment 2	2 μ l
DNA fragment 3	2 μ l
<u>4x Gibson buffer</u>	<u>2 μl</u>
Total	8 μ l

<Gibson-assembly condition>

50 °C	1 min
37 °C	30 sec
68 °C	30 sec
42 °C	30 sec
4 °C	∞

< Restriction-enzyme ligation >

DNA fragment 1	2 μ l
DNA fragment 2	2 μ l
<u>Ligation High Ver2</u>	<u>4 μl</u>
Total	8 μ l

< Restriction-enzyme condition >

16 °C	30 min
4 °C	∞

JM109 competent *E. coli* are selected to assemble and amplify plasmid by chemical transformation.

< Procedure >

1. Mix 2~5ul plasmid with 20~50ul JM109 competent cell.
2. On ice for 5 min, and then heat shock 42 °C for 40 sec.
3. After another on ice for 2 min, add 200~500ul SOC media and move to 37 °C incubator with shaking ~180 rpm for 1 hour to recover the cell.
4. Plating the cell onto LB based media plate contains antibiotic selection marker.

2-3. Yeast BY4741 culture and transformation:

< Day 1 >

1. Inoculate cells into 0.5ml liquid medium (YPDA or others) in >10ml tube. Shake overnight to grow to saturate at 30 °C.

< Day 2 >

1. Add 5ml YPDA and grow 3~4 hours at 30 °C. Collect cells by centrifugation at 3500 rpm for 15 min.
2. Suspend cells in 1ml TE buffer, transfer to round-bottomed 2ml tube. Collect cells by centrifugation at 10,000 rpm for 1 min, and discard fluid with the pipette.
3. Suspend cells in 500ul TE buffer, collect cells by centrifugation at 10,000 rpm for 1 min, and discard fluid with the pipette.
4. Suspend cells in 500 ul LiAc (100mM), collect cells by centrifugation at 5,000 rpm for 2 min, and discard fluid with the pipette.
5. Suspend cells in 20ul LiAc (100mM), add 5ul DTT (1M), 5ul freshly prepared ssDNA. (ssDNA needs to be denatured each time for use. Boiling in the water bath for 1 min, and in ice water for 2 min)
6. Aliquot mix-up (up to max 10ul) depending on the number of samples.
7. Add DNA solution (~100ng/ul or not change the conc) to 1/5 volume of mix-up and mix.
8. Add 2.7 volumes of 50%PEG, and 0.3 volumes of LiAc (1M).
9. Rotate gently for 1 hour at RT or 30 °C.
For quick protocol, this step can be skipped, which results in lower transformants.
10. Add 1/10 volume of DMSO, mix gently. (finger tapping, then spinning down)
11. Incubate at 42 °C for 15 min with gentle shaking.

12. Add an equal volume of water and mix by pipetting.
13. Spread cells onto appropriate selection plate in the clean bench.
14. Incubate at 30 °C for two days.

2-4. Yeast BY4741 transformation related media:

< ssDNA (10 mg/ml) >

ssDNA (ex. Salmon sperm)	0.1 g
ddH ₂ O (mili-Q)	10 ml

Heat then shakes overnight to dissolve.

- Aliquots (~100ul) of ssDNA in a 1.5ml tube can be stored at -30 °C and need to be freshly prepared each time as follows. Immediately before use, denature in boiling water in the microwave for 1 min then rapidly cool down in iced water.

< 1M LiAc >

Lithium Acetate	10.2 g
ddH ₂ O (mili-Q)	100 ml

Autoclave, store at RT.

< 50% PEG3500 >

PEG3500	50 g
ddH ₂ O (mili-Q)	30 ml

Heat and stir, bring total volume up to 100ml and stir to combine.

- Filter sterilizes (this might take about ~20 min). Aliquots can be stored at RT for 2 months or at -20 °C for a longer term.

2-5. Yeast BY4741 induction and spotting:

< Day 1 >

1. Inoculate cell in 100ul selection media (SD-Leu-Ura) in 96 well-plates. Seal around the plate to prevent it from drying out (normal tape should be ok). At 30 °C grow overnight with shaking ~100 rpm.

< Day 2 >

1. Dilute the 5ul culture into 155ul SRaf-Leu-Ura-Arg media, grow overnight.

< Day 3 >

1. Dilute the 5ul culture into proper 155ul induction solution (0.02%, 0.2%, 2%) Galactose-SRaf-Leu-Ura-Arg, grow overnight.

< Day 4 >

1. Warm-up and dry the appropriate control and selection plates (SD-Leu-Ura-Trp, SD-Leu-Ura-Arg+Can, SD-Leu-Ura-Lys+Thialysine) for 5~10min in clean bench.
2. Make serial dilution. Prepare 90ul dH2O (sample x 6) in line of 96 well. Transfer and mix well by pipetting 10ul culture from one lane to the next.
3. From the end of the dilution series, spot 10ul of dilutions in line and spot 10ul of culture for the last lane (sample x 7). Collect plate after drying.
4. Incubate at 30 °C for 2 days.

2-6. Yeast BY4741 selection plate:

< Control plate (SD-Leu-Trp-Ura) >

Premix SD-Leu-Trp-Ura	1 punch
Agar powder	10 g
dH2O	500 ml

*Note: dH2O use E-POD for better growth

Autoclave, store at 4 °C.

< On-target plate (SD-Leu-Ura-Arg) + CAN (60ug/ml) >

Yeast base w/o amino acid	3.35 g
Drop-out mix (-Leu-Ura-Arg)	1 g
Glucose	10 g
Agar powder	10 g
dH2O	500 ml

*Note: dH2O use E-POD for better growth

Adjust PH to 5.7 by KOH, autoclave, and add 300ul of Canavanine sulfate solution (100mg/ml).

Store at 4 °C.

< Off-target plate (SD-Leu-Ura-Lys) + Thialysine (100ug/ml) >

Yeast base w/o amino acid	3.35 g
Drop-out mix (-Leu-Ura-Lys)	1 g
Glucose	10 g
Agar powder	10 g
dH2O	500 ml

*Note: dH2O use E-POD for better growth

Adjust PH to 5.8 by KOH, autoclave, and add 500ul of Thialysine solution (100mg/ml).

Store at 4 °C.

2-7. Yeast BY4741 induction related media:

< S-Raffinose-Leu-Ura-Arg solution >

Yeast base w/o amino acid	3.35 g
Drop-out mix (-Leu-Ura-Arg)	1 g
Raffinose	10 g
dH2O	500 ml

*Note: dH2O use E-POD for better growth

Filter-sterilize, store at RT.

< 0.02% Gal-Raff-Leu-Ura-Arg solution >

5x -Arg-Ura-Leu	5 ml
20% Raffinose	2.5 ml
20% Galactose	25 µl
<u>E-POD-DW</u>	<u>17,475 µl</u>
Total	25 ml

Store at RT.

< 0.2% Gal-Raff-Leu-Ura-Arg solution >

5x -Arg-Ura-Leu	5 ml
20% Raffinose	2.5 ml
20% Galactose	250 µl
<u>E-POD-DW</u>	<u>17,250 µl</u>
Total	25 ml

Store at RT.

< 2% Gal-Raff-Leu-Ura-Arg solution >

5x -Arg-Ura-Leu	5 ml
20% Galactose	2,500 µl
<u>E-POD-DW</u>	<u>17,500 µl</u>
Total	25 ml

Store at RT.

< 20% Galactose solution >

Galactose	40 g
ddH ₂ O (mili-Q)	200 ml

Filter-sterilize, store at RT.

< 20% Raffinose solution >

Raffinose	40 g
ddH ₂ O (mili-Q)	200 ml

Filter-sterilize, store at RT.

< 5x -Arg-Ura-Leu >

Yeast base w/o amino acid	3.35 g
-Arg-Ura-Leu powder	1 g
dH ₂ O	100 ml

*Note: dH₂O use E-POD for better growth

Adjust pH to 5.8 (by 1M KOH), filter-sterilize, store at RT.

< 100mg/ml Canavanine sulfate solution >

Canavanine sulfate salt	0.5 g
ddH ₂ O (mili-Q)	5 ml

Store at -20 °C.

< 100mg/ml Thialysine solution >

Thialysine	0.5 g
ddH ₂ O (mili-Q)	5 ml

Store at -20 °C.

2-8. Yeast assay targets:

< On-target >

Site name	Protospacer	gRNA	PAM
CAN1	GCCCAGTTGGATTCCGTTAT	Site 1 (KN1086)	TGG
	CTCCAATAACGGAATCCAAC	Site 2 (KN1150)	TGG
	TTGGCCAAGTCATTCAATTT	Site 3 (KN1025)	TGG

< Off-target >

Site name	Protospacer
LYP1	AAGAATTTATGGGAGAAATTCTGGGC
	GCCCAGAATTTCTCCCATAAATTCTT

2-9. Mammalian HEK293T cell culture:

HEK293T cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin. Cells were incubated at 37 °C with 5% carbon dioxide and passaged every 3–4 days.

2-10. Mammalian HEK293T cell transfections:

< Cell pre-seeding >

- 24 well plate (Collagen I, Coated Plate): 1×10^5 cell/well --- 500 ul to each well with DMEM
- 48 well plate (Poly-D-lysine-coated plate): 5×10^4 cell/well --- 250 ul to each well with DMEM

*With overall 70~80% cell confluence

< Procedure >

Transfection proceeded 2 4h after pre-seeding

1. Distribute plasmids into a 1.5 ml tube within a total of 1,000ng
 - ① Normal assay.
 - 750 ng for effector, 250 ng for gRNA
 - ② R-loop assay
 - 300 ng for effector, 300 ng for nSaCas9
 - 200 ng for gRNA-nSpCas9, 200 ng for gRNA-nSaCas9
2. Add Opti-media with final volume 25 ul per well.
3. Add Fu-gene transfection agent with 2:1 to amount of plasmid.
4. Mix with finger tapping, and incubate 10 min at RT.
5. Transfer onto the cell, and softly shake the plate couple of times.
6. Change the media with DMEM after 24 h, and harvest after 72 h.

*For controls involving no base editor or no sgRNA, pUC19 DNA was used to maintain the total quantity of transfected DNA at 1,000 ng.

2-11. Cell Sorting:

< Sample preparation >

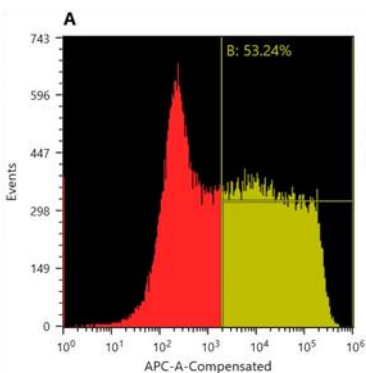
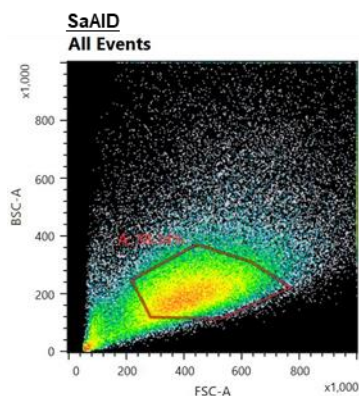
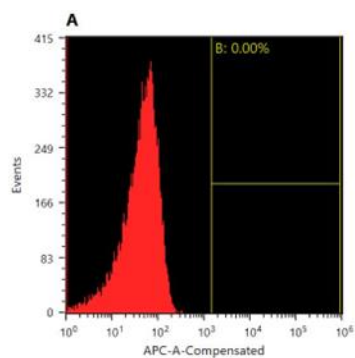
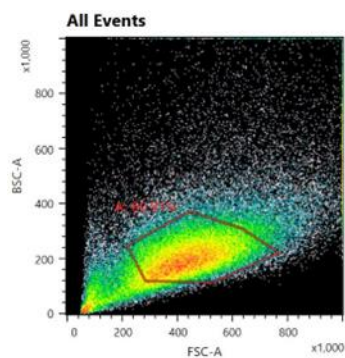
1. Wash the cells with 500ul PBS solution.
2. Add Trypsin 150ul, incubate 3 min, then add DENE media 150ul. Transfer the mixture into a 1.5ml tube, centrifuge 1,000rpm 2min.
3. Wash the cells with 800ul PBS solution, centrifuge 1,000 rpm 2min.
4. Suspend the cells with 800ul PBS solution and transfer them into a flex tube.
5. Cells were collected in a 1.5ml collection tube with 800ul PBS solution contained after sorted.

< Procedure >

1. Open the compressor. (wait until Pa in the middle of the green needle)
 2. Open the computer and cell sorter (Sony SH800 Cell Sorter).
 3. Wait until (Standby), opens cell sorting software.
 4. Opens cell sorter ship, and read QR code.
 5. Insert ship.
 6. Select 488: Green, 638: Red (for iRFP contained plasmid).
 7. Select Without 405 lasers.
 8. Sheath de-bubble
 - The inner tank with blue wire: knocks a couple of times with cap towards up, and release the air.
 9. Calibration
 - 8-10 drops of calibration beams into the empty tube (shaking bottle before use).
 - Calibration-Start-Target-OK.
 10. Successful-OK
 - Insert name, and create a new experiment.
 11. Adjust the measurement setting for the desired fluorochrome.
 12. Load the negative control to determine the sorting criterion.
 13. Load the sample (change the pressure to 7, select Sorting and Recording, click Auto-sorting if needed).
- ### **< Finished cleaning >**
1. Select: Cytometer-Soft and Hardware shut down---Start---Normal
 2. Prepare 50% bleach (fresh one every time, 5ml bleach + 5ml milli-Q) ---Start.

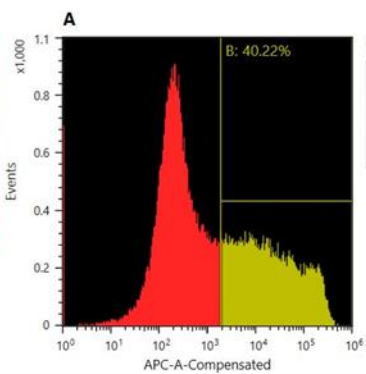
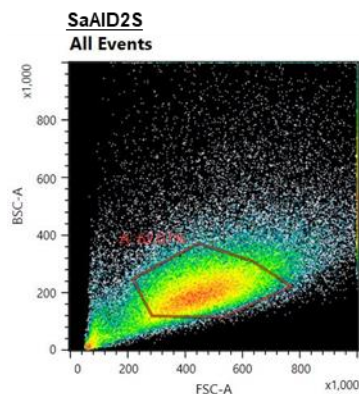
3. Prepare 12 ml milli-Q---15min normal---Start.
4. Cell Sorter shut down.
5. Close compressor.
6. Clean the machine.
 - Wrap with DI water.
 - Wrap with ethanol.
7. Exchange the ddH₂O (use a fresh one every time)
8. Fill the Sheath tank, and empty the waste tank.

*The gating strategy was shown in following graph



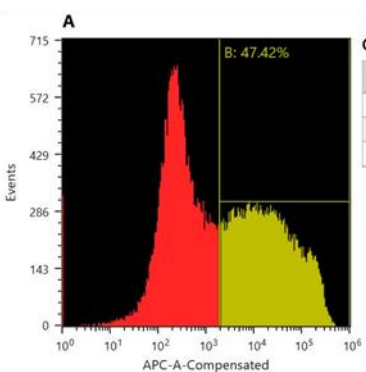
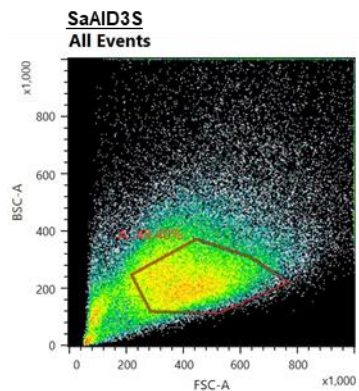
Gates and Statistics

Name	Events	%Parent	%Total
All Events	100,000	0.00%	100.00%
A	58,339	58.34%	58.34%
B	31,062	53.24%	31.06%



Gates and Statistics

Name	Events	%Parent	%Total
All Events	100,000	0.00%	100.00%
A	62,065	62.07%	62.07%
B	24,963	40.22%	24.96%



Gates and Statistics

Name	Events	%Parent	%Total
All Events	100,000	0.00%	100.00%
A	48,397	48.40%	48.40%
B	22,949	47.42%	22.95%

2-12. Genome extraction:

< Procedure >

1. Wash the cells with 500ul PBS solution.
2. Add 500ul PBS solution, and transfer cells into a 1.5ml tube.
3. Centrifuge 10,000 rpm for 2 min, and remove PBS solution.
4. Add 100ul Solution A (Kaneka Easy DNA Extraction Kit Version 2), and incubate 98 °C for 8 min.
5. Wait until cool down around RT, and then add 14ul Solution B.
6. Store genome in -20 °C

Genomic DNA was extracted by using Kaneka Easy DNA Extraction Kit (Version 2). Genomic RNA was extracted by using NucleoSpin RNA Plus, and cDNA was generated by using High-Capacity cDNA Reverse Transcription Kits (Applied Biosystems).

2-13. AAV production and infection

< Cell passage before seeding >

- 2-day gap: cell number: 7×10^5 , 7 well plates for 18 samples.
- 3-day gap: cell number: 5×10^5 , 7 well plates for 18 samples.

< AAV pre-seeding >

1. Trypsin and collect cell into single 50 ml tube and centrifuge 1000 rpm, 3 min.
2. Remove culture and add fresh 20 ml media.
3. Seed 30×10^5 cell/plate (Make 3×10^5 cell/ml mixture and aliquot 10 ml for each plate).
4. Make sure cell confluence around 60~70%.

< Transfection >

1. Effector: 6000 ng, pRQ-mi342: 6000 ng, pHelper: 6000 ng (SiRNA 600 ng, add if needed),
Mix together into 1.5 ml tube.
2. Add 750 ul Opti-media into tube.
3. Add 90 ul PEI-Max transfection agent (1:5 to the total vector (18 ug) used).
4. Vortex and centrifuge, then wait 5~10 min.
5. Dispense the mixture to cell, and softly shake the plate for a little bit.

< AAV harvest >

1. After 72h incubation, add 1/80 volume of 0.5M-pH 8 EDTA (125 ul for our petri-dish).
2. Shake and mix well, and wait 10 min at RT.
3. Collect in 50ml tube (combine the same sample), and centrifuge 1750 xg for 10 min at 4 degree and get rid of supernatant (as much as possible, however, supernatant also able to collect some AAV).
4. Vortex the tube to disperse the pellet well.
5. Add 500 ul AAV extraction solution A (Takara Kit) into tube.
6. Vortex about 15 seconds, and 2000-14000xg centrifuge for 10 min at 4 degrees.
7. Collect supernatant into 1.5 ml tube.
8. Add 50 ul AAV extraction solution B, and pipetting (proceed to purification or store at -80 degree, de-frozen in 37-degree water bath when using).

< AAV purification >

1. Add MgCl_2 to final conc ~80mM (1/12.5 volume of solution when using 1M MgCl_2).
2. Add cryonase cold-active nuclease to final conc 0.2 U/ul (200 U/ml) (20 U/ul enzyme in stock).
3. Keep in 37-degree for 30 min.
4. Add chloroform to 1:1 total volume ratio, and vortex vigorously for 10s.
5. Centrifuge 3000xg for 5 min at 4-degree.
6. Collect the upper water layer (pink) into Amicon Ultra column or a new 1.5ml tube.
7. Proceed Amicon Ultra centrifuge for concentrating (centrifuge ~60min to concentrate).

< AAV lysis >

1. Add 18 ul 10 mM Tris-HCl (pH 8.0) +0.05% tween 20 solution into 2ul AAV solution to dilute (x10).
2. Add 20 ul lysis buffer, and incubation 70 degree for 10min (x2).
3. Using 10 mM Tris-HCl (pH 8.0) +0.05% tween 20 to dilute AAV (x10000, x20000...), centrifuge before dilution to prevent any white precipitate taking.

< Droplet Digital PCR (ddPCR) for titer estimation >

Using Probe system

ddPCR Supermix for Probe (no dUTP)	10 ul
Forward Primer (10uM)	1.8 ul
Reverse Primer (10uM)	1.8 ul
Probe (10uM)	0.5 ul
Template DNA	1 ul
Water	4.9 ul
Total 20 ul (x20)	

*ITR Probe is blue cap, and sy946 sy947 are using for ITR PCR area

1. Open the switch of all machine like the seal machine.
2. Prepare the master mix of solution except template, and transfer 19ul into each well of middle line of DG8 Cartridge. Prevent bubble!
3. Add 1 ul sample into assigned well. Prevent bubble!
4. Add 70 ul generator oil into bottom line of cartridge.
5. Insert into holder and cover with orange gasket (single time of using cannot reuse).
6. Insert into droplet making machine, and prepare the PCR plate and assigned the sample area (the middle of plate using).
7. Cut the seal steel with best fitting size of the plate (The side contain red line is upper side and make it line up the top of plate).
8. Transfer the 40 ul of droplet into PCR plate softly and slowly, may taking 3~5s. The strong force may destroy the droplet. Prevent bubble!
9. After adding all droplet, seal the plate with steel sheet using sealing machine
10. Load the plate in middle of PCR machine.

< Droplet Digital PCR (ddPCR) setting >

1. Click saved filed.
2. Select AAV titration (40 ul 105-degree lid).
3. Start run (Takes around 2 hours).
4. After finished, transfer the plate into middle lines of QX200.
5. For shut down, click cancel and home bottom, then close the power of machine.

< QX200 software setting >

1. Click Quantasoft icon to start the program, click the machine condition the light (green, flashing green or yellow).
2. Click Prime (Run if not using in one week).
3. Select new template and double click to edit. Assign the using area: the middle region of plate.
4. Enter
 - a. Sample name: xx
 - b. Experiment: ABS
 - c. Supermix: ddPCR Supermix for Probe (no dUTP)
 - d. Target 2: For HEX/VIC, unknown for sample
5. Insert the plate within an appropriate position (middle) and cover the steel cover.
6. Save the template.
7. Click flush system.
8. Run (FAM/HEX) the machine.

*Note: the sum of positive and negative point of detected should >10000. The value of trustable data should around 5~500. (Conc: value x20 x20 xDilution x 1000 vg/ml).

2-14. High-throughput sequencing of genomic DNA and data analysis:

< 1st PCR >

Genomic DNA	1 µl
10 µM forward primer	1 µl
10 µM reverse primer	1 µl
KOD Fx Neo	0.5 µl
2x Buffer	10 µl
dNTP	2 µl
<u>dH₂O</u>	<u>4.5 µl</u>
Total	20 µl

< 1st PCR condition>

98 °C	2 min	
98 °C	10 sec	← × 32 cycle
58 °C	30 sec	
68 °C	30 sec/kb	
68 °C	5 min	
10 °C	∞	

< 1st PCR product purification>

PCR product	5 µl
<u>ExoSAP-IT Express</u>	<u>2 µl</u>
Total	7 µl
37 °C 4 min, 80 °C 1 min	

< 2nd PCR >

1 st PCR product	1 µl
10 µM forward primer	1 µl
10 µM reverse primer	1 µl
KOD Fx Neo	0.5 µl
2x Buffer	15 µl
dNTP	3 µl
<u>dH₂O</u>	<u>8.5 µl</u>
Total	30 µl

< 2nd PCR condition>

Same as the 1st PCR condition

< 2nd PCR product purification>

NucleoMag NGS Clean-up and Size Select kit.

➤ Elution with 30ul ddH₂O.

< 3rd PCR >

2 nd PCR product	5 µl
10 µM index primer 1	1 µl

10 μ M index primer 2	1 μ l
2x Q5 Polymerase	10 μ l
<u>dH₂O</u>	<u>3 μl</u>
Total	20 μ l

< 3rd PCR condition>

98 °C	10 sec	
98 °C	30 sec	← × 15 cycle
65 °C	30 sec	
72 °C	20 sec	
72 °C	5 min	
10 °C	∞	

< 3rd PCR product purification>

NucleoMag NGS Clean-up and Size Select kit.

➤ Elution with 30ul EB buffer.

Target region-containing fragment was first PCR-amplified using 1st primer pairs from the extracted genomic DNA. The second PCR was performed to obtain adapter added amplicon (~220 bp) by using the first PCR products as template and 2nd primer pairs containing adapter sequences. The amplicon was labeled using NEBNext Multiplex Oligos for Illumina. The DNA sequenced were analyzed by Miniseq system (Illumina, CA, USA) to obtain paired 2 × 150 bp read length by using MiniSeq Mid Output kit (300 cycle). Obtained Fastq data were processed and analyzed by Crispresso2⁴³.

< Primers used for deep-sequence>

Name	Sequence (5'-3')	Target	Experiment
HEK2 1st-1	TCCCACGTATTGCACTGCCATT	HEK2 gene	1st round PCR for deep-sequence
HEK2 1st-2	CCCCTCAGCATTCAGCCACTAATA	HEK2 gene	1st round PCR for deep-sequence
HEK2 2nd-1	TCTTTCCTACACGACGCTCTTCCGATCTCGG AATGAATGGATTCCTTG	HEK2 gene	2nd round PCR for deep-sequence
HEK2 2nd-2	GTGACTGGAGTTCAGACGTGTGCTCTTCCGA	HEK2 gene	2nd round PCR for

	TCTAGCCCCATCTGTCAAAC		deep-sequence
HEK3 1st-1	AGCTTGGCATGAGAAACCTTGGAG	HEK3 gene	1st round PCR for deep-sequence
HEK3 1st-2	TCTGGGTGCCCTGAGATCTTTT	HEK3 gene	1st round PCR for deep-sequence
HEK3 2nd-1	TCTTCCCTACACGACGCTCTTCCGATCTGGG AAACGCCCATGCAATTAGTCT	HEK3 gene	2nd round PCR for deep-sequence
HEK3 2nd-2	GTGACTGGAGTTCAGACGTGTGCTCTTCCGA TCTGCTGCACATACTAGCCCCTGTCT	HEK3 gene	2nd round PCR for deep-sequence
RNF2 1st-1	GTCTGTAAAGTCCATGGTTGGTGC	RNF2 gene	1st round PCR for deep-sequence
RNF2 1st-2	CTCCCAAAGTCCCCACCTCTTAAT	RNF2 gene	1st round PCR for deep-sequence
RNF2 2nd-1	TCTTCCCTACACGACGCTCTTCCGATCTCCA ACATACAGAAGTCAGGAATGC	RNF2 gene	2nd round PCR for deep-sequence
RNF2 2nd-2	GTGACTGGAGTTCAGACGTGTGCTCTTCCGA TCTTGTCTCAGGCTGTGCAGACAA	RNF2 gene	2nd round PCR for deep-sequence
VEGFA 1st-1	AAGGCTCCAATGCACCCAAGA	VEGFA gene	1st round PCR for deep-sequence
VEGFA 1st-2	CGCTCGGTGCTGGAATTTGATA	VEGFA gene	1st round PCR for deep-sequence
VEGFA 2nd-1	GTGACTGGAGTTCAGACGTGTGCTCTTCCGA TCTCTCTCCGAAGCGAGAACA	VEGFA gene	2nd round PCR for deep-sequence
VEGFA 2nd-2	TCTTCCCTACACGACGCTCTTCCGATCTACG GACAGACAGACAGACAC	VEGFA gene	2nd round PCR for deep-sequence
SaCas9 site_1 1st-1	GTGGAGTGCTCTGTGTTTGTCT	SaCas9 site_1 R-loop off-target	1st round PCR for deep-sequence
SaCas9 site_1 1st-2	TACCTGTGCACCTAAGGCATCA	SaCas9 site_1 R-loop off-target	1st round PCR for deep-sequence
SaCas9 site_1 2nd-1	TCTTCCCTACACGACGCTCTTCCGATCTTAC CAACATACATGCCCCTTTGCC	SaCas9 site_1 R-loop off-target	2nd round PCR for deep-sequence

SaCas9 site_1 2nd-2	GTGACTGGAGTTCAGACGTGTGCTCTTCCGA TCTTCTTGACTCATCCTGCAGTCTCCT	SaCas9 site_1 R- loop off-target	2nd round PCR for deep-sequence
SaCas9 site_2 1st-1	GATTTGCCAAGTTGCCTGTCCTTC	SaCas9 site_2 R- loop off-target	1st round PCR for deep-sequence
SaCas9 site_2 1st-2	TGGGTTTATGGAGGTCCTCTTGTG	SaCas9 site_2 R- loop off-target	1st round PCR for deep-sequence
SaCas9 site_2 2nd-1	TCTTCCCTACACGACGCTCTTCCGATCTTGC GGTGGGAGATCTGGTTTC	SaCas9 site_2 R- loop off-target	2nd round PCR for deep-sequence
SaCas9 site_2 2nd-2	GTGACTGGAGTTCAGACGTGTGCTCTTCCGA TCTAGGACATTTCCACCGCAAAATGGC	SaCas9 site_2 R- loop off-target	2nd round PCR for deep-sequence
SaCas9 site_3 1st-1	AGCTGTCCAGAGGTGTTAAGTGCC	SaCas9 site_3 R- loop off-target	1st round PCR for deep-sequence
SaCas9 site_3 1st-2	TTCACTGTGTTAGCCAGGATGGTC	SaCas9 site_3 R- loop off-target	1st round PCR for deep-sequence
SaCas9 site_3 2nd-1	TCTTCCCTACACGACGCTCTTCCGATCTTGG CATCCAGAGACATGGTTTC	SaCas9 site_3 R- loop off-target	2nd round PCR for deep-sequence
SaCas9 site_3 2nd-2	GTGACTGGAGTTCAGACGTGTGCTCTTCCGA TCTCTCTTTGCTCCAGATTTCCCTTCA	SaCas9 site_3 R- loop off-target	2nd round PCR for deep-sequence
SaCas9 site_4 1st-1	ACAACCTGTCCGCAAGGGAAAA	SaCas9 site_4 R- loop off-target	1st round PCR for deep-sequence
SaCas9 site_4 1st-2	AGTAGGGGTGAGGGAACAAACCA	SaCas9 site_4 R- loop off-target	1st round PCR for deep-sequence
SaCas9 site_4 2nd-1	TCTTCCCTACACGACGCTCTTCCGATCTAGG GGCCTCAGGGGAATAAATCAT	SaCas9 site_4 R- loop off-target	2nd round PCR for deep-sequence
SaCas9 site_4 2nd-2	GTGACTGGAGTTCAGACGTGTGCTCTTCCGA TCTTGGGATTGACAGAAATGTGAGCC	SaCas9 site_4 R- loop off-target	2nd round PCR for deep-sequence
SaCas9 site_5 1st-1	AGCTTGGCATGAGAAACCTTGGAG	SaCas9 site_5 R- loop off-target	1st round PCR for deep-sequence
SaCas9 site_5 1st-2	TCTGGGTGCCCTGAGATCTTTT	SaCas9 site_5 R- loop off-target	1st round PCR for deep-sequence
SaCas9	TCTTCCCTACACGACGCTCTTCCGATCTGGG	SaCas9 site_5 R-	2nd round PCR for

site_5 2nd-1	AAACGCCCCATGCAATTAGTCT	loop off-target	deep-sequence
SaCas9 site_5 2nd-2	GTGACTGGAGTTCAGACGTGTGCTCTTCCGA TCTGCTGCACATACTAGCCCCTGTCT	SaCas9 site_5 R- loop off-target	2nd round PCR for deep-sequence
SaCas9 site_6 1st-1	CATCTCGCACGTGGTTCCGGAA	SaCas9 site_6 R- loop off-target	1st round PCR for deep-sequence
SaCas9 site_6 1st-2	CGACCAAAGCGCCGATGGATGTG	SaCas9 site_6 R- loop off-target	1st round PCR for deep-sequence
SaCas9 site_6 2nd-1	TCTTTCCTACACGACGCTCTTCCGATCTGGC TTTAAAGTTGCCAGAG	SaCas9 site_6 R- loop off-target	2nd round PCR for deep-sequence
SaCas9 site_6 2nd-2	GTGACTGGAGTTCAGACGTGTGCTCTTCCGA TCTCCGCCAGAAGCTCGGAAAA	SaCas9 site_6 R- loop off-target	2nd round PCR for deep-sequence
SaCas9 site_7 1st-1	ACTGCCACTCGGTCTCTTCA	SaCas9 site_7 R- loop off-target	1st round PCR for deep-sequence
SaCas9 site_7 1st-2	AGGGGTAGGGATGGTCTCT	SaCas9 site_7 R- loop off-target	1st round PCR for deep-sequence
SaCas9 site_7 2nd-1	TCTTTCCTACACGACGCTCTTCCGATCTGGG AAAAGTTTGGGGTGGA	SaCas9 site_7 R- loop off-target	2nd round PCR for deep-sequence
SaCas9 site_7 2nd-2	GTGACTGGAGTTCAGACGTGTGCTCTTCCGA TCTCTCGCCACTCTGGGCTCCCCATT	SaCas9 site_7 R- loop off-target	2nd round PCR for deep-sequence
HEK2_off_s ite-1 1st-1	GCAGAAGTCGTACTATGCAAGC	HEK2 gene off- target 1	1st round PCR for deep-sequence
HEK2_off_s ite-1 1st-2	TGCATGGGCACATGTGTTTG	HEK2 gene off- target 1	1st round PCR for deep-sequence
HEK2_off_s ite-1 2nd-1	TCTTTCCTACACGACGCTCTTCCGATCTGAC AGATCTGTAAGTAGAACAATGGG	HEK2 gene off- target 1	2nd round PCR for deep-sequence
HEK2_off_s ite-1 2nd-2	GTGACTGGAGTTCAGACGTGTGCTCTTCCGA TCTCCATATCTACTATATACATATTAATGCTCC C	HEK2 gene off- target 1	2nd round PCR for deep-sequence
HEK2_off_s ite-2 1st-1	AGAGGCCTTCCAAGCCTGAAAATG	HEK2 gene off- target 2	1st round PCR for deep-sequence
HEK2_off_s	GCATTAATCCATATTGTATCACCTCC	HEK2 gene off-	1st round PCR for

ite-2 1st-2		target 2	deep-sequence
HEK2_off_s	TCTTTCCCTACACGACGCTCTTCCGATCTAGC	HEK2 gene off-	2nd round PCR for
ite-2 2nd-1	AGTGAGTTTGGGCACTTGT	target 2	deep-sequence
HEK2_off_s	GTGACTGGAGTTCAGACGTGTGCTCTTCCGA	HEK2 gene off-	2nd round PCR for
ite-2 2nd-2	TCTTATGATTCTAACTATCATGG	target 2	deep-sequence
VEGFA_off	AAAACAGTGCGGTGGAGAAAGG	VEGFA gene off-	1st round PCR for
_site-1 1st-1		target 1	deep-sequence
VEGFA_off	CAGAGGGAGAATCAAGCTGTACAC	VEGFA gene off-	1st round PCR for
_site-1 1st-2		target 1	deep-sequence
VEGFA_off	TCTTTCCCTACACGACGCTCTTCCGATCTAAG	VEGFA gene off-	2nd round PCR for
_site-1 2nd-1	GTTTCCCAATTCATCTTCAGAACAG	target 1	deep-sequence
VEGFA_off	GTGACTGGAGTTCAGACGTGTGCTCTTCCGA	VEGFA gene off-	2nd round PCR for
_site-1 2nd-2	TCTTCTCTGACTTGGACAGCTGCAGTA	target 1	deep-sequence
VEGFA_off	ACAGCACCATGGTTAGACTGAGGA	VEGFA gene off-	1st round PCR for
_site-2 1st-1		target 2	deep-sequence
VEGFA_off	CAACAGGATTGCTGAGGGATTGC	VEGFA gene off-	1st round PCR for
_site-2 1st-2		target 2	deep-sequence
VEGFA_off	TCTTTCCCTACACGACGCTCTTCCGATCTTTC	VEGFA gene off-	2nd round PCR for
_site-2 2nd-1	TCCCTACTCAACCCCGAGC	target 2	deep-sequence
VEGFA_off	GTGACTGGAGTTCAGACGTGTGCTCTTCCGA	VEGFA gene off-	2nd round PCR for
_site-2 2nd-2	TCTTTCGGCTGAGCAGTTAGAAGGAAG	target 2	deep-sequence
VEGFA_off	TGTGGCTGCCCCGCTTATTAAA	VEGFA gene off-	1st round PCR for
_site-3 1st-1		target 3	deep-sequence
VEGFA_off	CCGCCTGTGAGCTATTTCCACATT	VEGFA gene off-	1st round PCR for
_site-3 1st-2		target 3	deep-sequence
VEGFA_off	TCTTTCCCTACACGACGCTCTTCCGATCTTCC	VEGFA gene off-	2nd round PCR for
_site-3 2nd-1	TGGCCACCTCTTGCCTGTCA	target 3	deep-sequence
VEGFA_off	GTGACTGGAGTTCAGACGTGTGCTCTTCCGA	VEGFA gene off-	2nd round PCR for
_site-3 2nd-2	TCTGAGGATCGCAGCCTGGAGTTAAG	target 3	deep-sequence
VEGFA_off	ATCACACCCTTTGGGGATCTTTGG	VEGFA gene off-	1st round PCR for
_site-4 1st-1		target 4	deep-sequence

VEGFA_off _site-4 1st-2	AAAGGCTTCCTTGAGGCTGAAGTC	VEGFA gene off- target 4	1st round PCR for deep-sequence
VEGFA_off _site-4 2nd-1	TCTTTCCCTACACGACGCTCTTCCGATCTTGC TTTCCAGCTCTGCTTCATGT	VEGFA gene off- target 4	2nd round PCR for deep-sequence
VEGFA_off _site-4 2nd-2	GTGACTGGAGTTCAGACGTGTGCTCTTCCGA TCTGGAGAGTGCCATTCTCAGCCTAAA	VEGFA gene off- target 4	2nd round PCR for deep-sequence

2-15. Mammalian HEK293T cell assay targets:

< SpCas9 genomic loci for on-target evaluation >

Site name	Protospacer	PAM	Amplicon
On-target site 1 HEK2	GAACACAAAGCATAGACTGC	GGG	ATGAATGGATTTCCTTGGAACAATGATAACAA GACCTGGCTGAGCTAACTGTGACAGCATGTGGT AATTTTCCAGCCCGCTGGCCCTGTAAAGGAAA CTGGAACACAAAGCATAGACTGCGGGGCGGGC CAGCCTGAATAGCTGCAAACAAGTGCAGAATA TCTGATGATGTCATACGCACAGTTTGACAGATG GGGCTGGACAATTTTCCCCCTTTAA
On-target site 2 HEK3	GGCCCAGACTGAGCACGTGA	TGG	CGCCCATGCAATTAGTCTATTTCTGCTGCAAGT AAGCATGCATTTGTAGGCTTGATGCTTTTTTCT GCTTCTCCAGCCCTGGCCTGGGTCAATCCTTGG GGCCCAGACTGAGCACGTGATGGCAGAGGAAA GGAAGCCCTGCTTCCTCCAGAGGGCGTCGCAGG ACAGCTTTTCTAGACAGGGGCTAGTATGTGCA GCTCCTGCACCGGGATACTGGT
On-target site 3 RNF2	GTCATCTTAGTCATTACCTG	AGG	ACGTCTCATATGCCCTTGCCAGTCATCTTAGT CATTACCTGAGGTGTTTCGTTGTAACATCATATA ACTGAGTTCCTATGTTTTGCTTAATGGTTGAGT TCCGTTTGTCTGCACAGCCTGAGACATTGCTGG AAATAAAGAAGAGAGAAAAACAATTTTAGTA TTTGGAAGGGAAAGTGCTATGGTCTGAATGTATG TGCCCCACCAAATTCCTACGT
On-target site 4 VEGFA	GACCCCTCCACCCCGCCTC	CGG	CTCCGCTCGGCTCGGCTTCCCCCGCGCGGACCA CGGCTCCTCCGAAGCGAGAACAGCCCAGAAGT TGGACGAAAAGTTTCAGTGCAGCGCCGCGAGC CCCAGCCCTCCACCCCGCTCCGGGCGCGGG CTCCGGCCCTGCCCGCGGCTCGCCGCCGCGTC CACTGTCCGCGCCGCGCGGGAGGAGGTGGTA GCTGGGGCTGGGGCGGTGTCTGT
On-target site 5 HEK4	GGCACTGCGGCTGGAGGTGG	GGG	AGATGGCTGACAAAGGCCGGGCTGGGTGGAAG GAAGGGAGGAAGGGCGAGGCAGAGGGTCCAAA GCAGGATGACAGGCAGGGGCACCGCGGCGCCC CGGTGGCACTGCGGCTGGAGGTGGGGTTAAAG CGGAGACTCTGGTGCTGTGTGACTACAGTGGGG GCCCTGCCCTCTCTGAGCCCCGCCTCCAGGCC TGTTGTGTGTCTCCGTTCCGGTTG
On-target site 6 PPP1R12C_4	GACCCTCAGCCGTGCTGCTC	GGG	CCCTCCCGTGGGGTGAGGAGAAGGCTGGGAGG GATACATTGGTGGGGGATATTGGCTTGTAGGAT CAGCCAGGAACGAAGACCACTGGGGCCACTGG AAAGACCCTCAGCCGTGCTGCTCGGGGTGGGA CTCTGCGGCCCCAAATCACAAGGCGACTGCCT GTGACACCCTGTGACACCCGGCTCCACACAGG AGCCTCAGAGGGGCTGCAGTGCAGCAGG

On-target site 7 RPCI	GGCCTCCGTATCACTCTCTGA C	TGG	GTTGGGGGGGAGGTAGGGGCTATCAAACCTCAT GATTGGCCGCACACAAAGGTGTGAGTATGTGT ATATTTGAGGGTGGGTGGGAGTGGCACTTTCAC TAGGCCTCCGTATCACTCTCTGACTGGGGTATC TCCCAGCAAGCGAGACAGAGGCAGACACGCTT CCCAGACTGTCTTACTGGGTCTCTCTGTGTATT CTCTGCAGTGTCTGTGTGTATCGTG
On-target site 8 MALAT1	CTTCCAGTTGAATTCACCAG	TGG	ATAGCCTATTTACTTTAAATAAACCAACATT CCATTTTAAATGTGGGGATTGGGAACCACTAGT TCTTTCAGATGGTATTCTTCAGACTATAGAAGG AGCTTCCAGTTGAATTCACCAGTGGACAAAAT GAGGAAAACAGGTGAACAAGCTTTTTCTGTAT TTACATACAAAGTCAGATCAGTTATGGGACAA TAGTATTGAATAGATTTCAGCTTTAT

< SaCas9 genomic loci (R-loop assay and SaCas9 variants on-target) >

Site name	Protospacer	PAM	Amplicon
SaCas9 site 1	GTGGTAGACAGCATGTGTCCTA	AAGGGT	ATGCCCCCTTGCCTCAAGTCTGGTTATTTAGG GGGATGCTAGGTTGCTTTGGGTCTACCTTACTG AGAAAATGGCCCCAGGTCATTGTCATGTCCAG TTGTGGTAGACAGCATGTGTCCTAAAGGGTAT ATTACATGCATGTGCAAAAATACAGGGGTCC TTCTAACCCCTATCACAGAGAAGCAGGAGACTG CAGGATGAGTCAAGATGCTAAGTGATGA
SaCas9 site 2	ATTTACAGCCTGGCCTTTGGGG	TCGGGT	CCTCAGACCAGGGTGCGGTGGGAGATCTGGTTT CCGGAAGACGGAATGGGGAGAAGGGCAGGTTT CCCGAGGCGCCCAGACACCCAATCTCCCCGT GACATTTACAGCCTGGCCTTTGGGGTCGGGTCA ACGCTAGGCTGGCAGGGGAAGGGCGGGCCGT GAGGTGAGCCGGCGCTGCAGGAAGGGGCCACC ACCAGAGGGGCCATTTTGCCTGGAAT
SaCas9 site 3	GTGTCAGGTAATGTGCTAAACA	GAGAGT	GTGGCATCCAGAGACATGGTTTCTTATCTCCTT AAGTGTTCACTGCTTTTCTTTCATTTATTCCA CATATAATTACTATAATTGCTAAACATTTATTT AGTGTGAGGTAATGTGCTAAACAGAGAGTTAC TGCTCAGACATGTAATAATAATAAACACACA TCAAATAACCATAACCATTTTAAGCTGTAGTAT TATGAAGGGAAATCTGGAGCAAAGAGA
SaCas9 site 4	GGTGGAGGAGGGTGCATGGGGT	CAGAAT	TTTGTCTTATCCAGAAAAGGGAGTGATTGCTTC CAGGGGCTCAGGGGAATAAATCATAGAATCC TGGACAAGGTTTGAAGGACAGGTAGGATTTGG GTGGGTGGAGGAGGGTGCATGGGGTCAGAATT GTAACCGAAAACCTATTCCAGGTGGATAGAGA AAATTTCTAGTGTGTGTGTTTTAAACTATTTG GGGACTGGCACAGACCCTTTTGAATA

SaCas9 site 5	TCTGCTTCTCCAGCCCTGGC	CTGGGT	GATGAGAGAAGCCTGGAGACAGGGATCCCAGG GAAACGCCCATGCAATTAGTCTATTTCTGCTGC AAGTAAGCATGCATTTGTAGGCTTGATGCTTTT TTTCTGCTTCTCCAGCCCTGGCCTGGGTCAATC CTTGGGGCCCAGACTGAGCACGTGATGGCAGA GGAAAGGAAGCCCTGCTTCTCCAGAGGGCGT CGCAGGACAGCTTTTCTTAGACAGG
SaCas9 site 6 (FANCF)	GATGTTCCAATCAGTACGCA	GAGAGT	GGCTTTTAAGTTGCCAGAGTCAAGGAACACG GATAAAGACGCTGGGAGATTGACATGCATTTT GACCAATAGCATTGCAGAGAGGCGTATCATT CGCGGATGTTCCAATCAGTACGCAGAGAGTCG CCGTCTCCAAGGTGAAAGCGGAAGTAGGGCCT TCGCGCACCTCATGGAATCCCTTCTGCAGCAC CTGGATCGCTTTTCCGAGCTTCTGGCGG
SaCas9 site 7 (VEGFA)	GCCCGCCCTCCCCGCCCCG	CCGGGT	TGGGAAAAGTTTTGGGGTGATTGCTGCGGGGA CCCCCCTCCCTGCTGGGCCACCTGCGCCGCGC CAACCCCGCCCGTCCCCGCTCGCGTCCCGTCTG GTGCCCCCCTCCCCGCCCCGCGGGTGCGCG CGGCGCGGAGCCGATTACATCAGCCCCGGGCCT GGCCGCGCGCGTGTTCCTCGAGCCTCGGCTGCC CGAATGGGGAGCCCAGAGTGCCGAG

< SpCas9 genomic loci for SaCas9 R-loop off-target assay >

Site name	Protospacer	PAM	Amplicon
SpCas9 site 1	GAACACAAAGCATAGACTGC	GGG	ATGAATGGATTCCCTTGAAACAATGATAACAA GACCTGGCTGAGCTAACTGTGACAGCATGTGG TAATTTTCCAGCCCGCTGGCCCTGTAAAGGAA ACTGGAACACAAAGCATAGACTGCGGGGCGGG CCAGCCTGAATAGCTGCAAACAAGTGCAGAAT ATCTGATGATGTCATACGCACAGTTTGACAGA TGGGGCTGGACAATTTTCCCCCTTTAA
SpCas9 site 2	GGCCCAGACTGAGCACGTGA	TGG	CGCCCATGCAATTAGTCTATTTCTGCTGCAAGT AAGCATGCATTTGTAGGCTTGATGCTTTTTTTC TGCTTCTCCAGCCCTGGCCTGGGTCAATCCTTG GGGCCCAGACTGAGCACGTGATGGCAGAGGAA AGGAAGCCCTGCTTCCTCCAGAGGGCGTCGCA GGACAGCTTTTCTTAGACAGGGGCTAGTATGTG CAGCTCCTGCACCGGGATACTGGT
SpCas9 site 3	GTCATCTTAGTCATTACCTG	AGG	ACGTCTCATATGCCCCTTGGCAGTCATCTTAGT CATTACCTGAGGTGTTCTGTTGTAACATATATA ACTGAGTTCCCATGTTTGTCTTAATGGTTGAGT TCCGTTTGTCTGCACAGCCTGAGACATTGCTGG AAATAAAGAAGAGAGAAAAACAATTTTAGTA TTTGAAGGGAAGTGCTATGGTCTGAATGTATG TGTCACCAAAAATTCCTACGT

SpCas9 site 4	GACCCCCTCCACCCCGCCTC	CGG	CTCCGCTCGGCTCGGCTTCCCCGCGCGGACCA CGGTCCTCCGAAGCGAGAACAGCCCAGAAGT TGGACGAAAAGTTTCAGTGCACGCCGCGAGC CCCGACCCCCTCCACCCCGCTCCGGGCGCGG GCTCCGGCCCCCTGCCCCGCGGCTCGCCGCCGCGT CCACTGTCCGCCCGCCGGCGGGAGGAGGTGG TAGCTGGGGCTGGGGGCGGTGTCTGT
---------------	----------------------	-----	---

< SpCas9 gRNA dependent off-target genomic loci >

Site name	Protospacer	PAM	Amplicon
HEK2_site1	GAACACAATGCATAGATTGC	CGG	TCTGTAAGTGAACAATGGGTTCAAAAACAAA CAGAGAAGACAAGTAAATTAATAAGTACTGT AGTCCAGATAATCAGAGTGTGGAGAGTGAGTA AGCCAGAACACAATGCATAGATTGCCGGTAAA TAGGTTTAGATTTCATCCATTTTAAAAAATGGT GTGGGAGCATTAAATATGTATATAGTAGATAT GGAAAAATGATTCTCATAATAACTGACA
HEK2_site2	AAACATAAAGCATAGACTGC	AAA	TAGCTCAGGGAAGGAGCAGTGAGTTTGGGCAC TTGTGACAGAATAGTGGGACTATGCCAGAGAT ACACAGGAGGAGGTGGTACCTTCTAGCTCCCC CTCAAAACATAAAGCATAGACTGCAAAGTAC TCCCAAGCAGGCTGAATAACACTCGAGTACCA AAAAAATACACTCTGAAACAATACCAATAAT TTTTTTAAAAATTATATATACCATGATAGT
VEGFA_site1	GACCCCCCCCACCCCGCCCC	CGG	CTTTTTAAGGTTTCCCAATTTCATCTTCAGAAC AGTTAGAATCTACTGACTAATTCTGAGCAGAG GAATATGTGACATGAGGAGATTTCATTTCCA AAAGACCCCCCCCACCCCGCCCCCGGCTGCTG TTGCAGAGGACAAGTTGAGGAGGGCAGAGAGG ATGTGGAGAGGACAGCAGGGAGTACTGCAGCT GTCCAAGTCAGAGATTGTAGTGGCTCT
VEGFA_site2	GCCCCACCCACCCCGCCTC	TGG	GAGCCCTCCTCTCTTGCTTCTTGCTCTGACCTT GTTTGTTATTCCACAGAGAAAACCTTTCAAAAA CAGCCAGAAGAGAACATCCACGAGCTCCAC CACTGCCCCACCCACCCCGCCTCTGGCCCAT GTATTCTGCTTTCTCTTGTTTAATCTTTCCCTC TCAACAAAGAGCAACTTCTTCCTTCTAACTGC TCAGCCGAAAAGCCTTGAGATCACCC
VEGFA_site3	CGCCCTCCCCACCCCGCCTC	CGG	TCTTGCTGTACGCAGACGCCCCCATCAAGC CCACCGCCACCCCCAACCCCGTGCGGGCAG CGGTCCTCGGTGCTCTGGTCCGACGCCCTTC TGTCGCCCTCCCCACCCCGCCTCCGGCGGACA GAGCTGGGCGGGAGCTGGGACTGGTACCCCT CCCCACCTCGGAGACACCTTAACCTCAGGCT GCGATCCTCGTGCCGGGGCCCCCGCGT

VEGFA_site4	CTCCCCACCCACCCCGCCTC	AGG	CTCTGCTTCATGTTGTTGGCAGGAGGCCCTCC TTCTCATCCTTGTATCAGCTGCCTTCTCATCAC AAGATGACTATGTCCCTCTGGGCCCCATCCTCC CCTCCCCACCCACCCCGCCTCAGGCTTGAAGA GGAAAGAAGAGCGCAGGGCCAAAAAACTTT GTCCTTTAGAAAGGATCTACCTTTTAGGCTGAG AATGGCACTCTCCCCAGAAACATCC
-------------	----------------------	-----	--

< gRNA dependent RNA off-target genomic loci >

Site name	Protospacer	PAM	Amplicon
HEK3-RNA_Off-target site 1	GATACAAAAAGTTCCATTGA	GGA	TCAAGAATTACTTGTGTTTCACAGAGATAGA CTCTTTGCTTTATAGAGATTGTTGTGATTAA TATGAATATCCCAGCTTTAGAAAAGAAGTAAA CTGGATACAAAAAGTTCCATTGAGGAACAGTT ATTTACAGTATAAAAGATTTGTTTACTTTACA AAAGGCTTGTGTCTGTGTGTGTGTGTGTGTGT GTGTGTGTATTTTAAACTGACTCAG
HEK3-RNA_Off-target site 2	CAGGACTGCTGCAGACTGAG	CCA	GGAGACTCACCTCTGGCTCTTTGCCACTGTCAG AATCTGAATCTCACTGGCCCTGTGGAGTAGGG ATCCTATCTGGAGAAGTGGGAGCATGGGCTGC AGTCAGGACTGCTGCAGACTGAGCCATGTGAT GGTACGTAATGAGTTCCCTGAGGGAATGAAA CACCTCACCCTTCAAAGTCACCCTTTGGA ATTCAACACAGACACACATATCCCTTC
HEK3-RNA_Off-target site 3	AAATACATTTCTGGGCCACC	TCA	CGTGCTTGTGGACATCAGGCCTCCTGCCAGCAG TTCTTGAAGCTTCTTTTTCATTCTGCTACTCTA CCTGTATTTCTCAGTTGCAGCACTGAGTGGTCA AAATACATTTCTGGGCCACCTCAGGGAACCCA TGCATCTGCCTGGCATTAGGCAGCAGAGCCC CTGACCGTCCCCACAGGGCTCTGCCTCACGTC CTCATCTCATTGGCTGTGTAAA
RNF2-RNA_Off-target site 1	ATCACACCTCATCTTAGTCA	ACC	TACAGATATGGTAGAAAAAGTAAGAAAATTTT CACCACAAAACCAATAGTTAACTACTAACAG AGAAAGTTATACACAAAATATATCTCTCAATA CAGTGATCACACCTCATCTTAGTCAACCGACT CGATGGCCGGAGCAGAGCCTGTGGACGGGGGC TTTTGCACAGCGCTTCTCTGACTCTTGGGCTCA GCCTCCAGGTGCGCTCCCAGGGCTCTTT
RNF2-RNA_Off-target site 2	GACCAATGACTAATATGTTC	AAT	ATATGTGTACATATTAATAGGCATACACACA CATCACACACAACCAGATAAAATAAATGGTA CATACATTGCTACTATGAACACTGCACGTCCC AGACAGACCAATGACTAATATGTTCAATTA CTACTTGAATTAATACTACTTCTCTCCCCAA AATATTCTGCTTGCAGCAAACATATTCCAGTT GCTCAAGCGGTGTTGTTTCTAGGTTTCT

RNF2-RNA_Off-target site 3	AAAAATCAATGACTAAGAAC	TTA	TGAAAGCTGCAGTTTCCCTCTTTCCTATTTTTTT TTAAACACAAAAATCAATGACTAAGAACCTTA ATACTGGATGACAAATCACATCCACACTATTC GTAAATAGCCTCACAGTTAAACACATCTTAA AGACCAGTCAAATCAGTATTTACATTTTATAA ATTTCTGAAGACACCATTAGAAAAGCTTATATA TCCCTTCATTAAACAAAATAGGGAAC
----------------------------	----------------------	-----	--

< gRNA independent RNA off-target genomic loci >

Site name	Amplicon
CTNNB1	TTTGATGGAGTTGGACATGGCCATGGAACCAGACAGAAAAGCGGC TGTTAGTCACTGGCAGCAACAGTCTTACCTGGACTCTGGAATCCAT TCTGGTGCCACTACCACAGCTCCTTCTCTGAGTGGTAAAGGCAATC CTGAGGAAGAGGATGTGGATACCTCCCAAGTCCTGTATGAGTGGG AACAGGGATTTTCTCAGTCCTTCACTCAAGAACAAGTAGCTG
IP90	CTGGTTGACCAATCTGTGGTGAATAGTGGAATCTGCTCAATGAC ATGACTCCTCCTGTAAATCCTTCACGTGAAATTGAGGACCCAGAAG ACCGGAAGCCCGAGGATTGGGATGAAAGACCAAAAATCCCAGATC CAGAAGCTGTCAAGCCAGATGACTGGGATGAAGATGCCCTGCTA AGATTCCAGATGAAGAGGCCACAAAACCCGAAGGCTGGTTAGATG ATGAGCCTGAGTACGTAC
RSL1D1	TTGGCTTTCCAAATCAGTGGGTCTGACTTGAGGTCTGTGATGTGAC CCTTTTCCTCACCTGCTCAACCATTATTCACATGGACTCCATCATAT TCATTTGTAGTCATTCCCAGAGTGGCCAGTGAGGGTCTCGCTGTA TGAGAGTCGGCTACGGAATTTAGGAGAAACAGAAGTTTCTTGGCT TTCATGCTGAGCTTGTTGGTCTAAGCTTATGAG

3. Result

3-1. Protein engineering and yeast assay

< 3-1-1. Elimination of DNA-binding region >

We hypothesized that PmCDA1 itself can interact with genomic DNA without Cas9 binding to DNA leading to the gRNA-independent off-target base editing. Removing regions of PmCDA1 that would have affinity to DNA can reduce the Cas9-independent off-target. The potential binding moieties of PmCDA1 were predicted on the basis of the structure of human AID along with dsDNA (Fig. 2). The predicted binding regions were removed by truncating the protein from its termini since the corresponding amino acids for the regions are present close to the N- and C-terminal ends. We fused the truncated PmCDA1s to nCas9 at either terminus and tested their on-target activity and genome-wide non-specific mutagenic activity in yeast *Saccharomyces cerevisiae* (BY4741) cells. We adopted the CAN1 gene as a target locus, whose loss of function leads resistance to the drug canavanine. The on-target occurrence rate was calculated by counting canavanine-resistant clones. On the other hand, the LYP1 gene was employed to assess the potential of genome-wide off-target mutations by revealing resistance to S-aminoethyl-L-cysteine. These mutations are generated in a gRNA-independent fashion.

Based on the amino acid alignment of AID and PmCDA1, the potential DNA-binding moieties for PmCDA1 were located to residues 21–27 and 172–192 of the total 208 amino acids (a.a.) length of the protein, we first made a series of truncations from the C-terminal end PmCDA1 (1-201, 1-197, 1-190, 1-183, 1-179, 1-176, 1-161) (Fig. 3). To reflect the changes in their deaminase activity within the dynamic range, we intentionally omitted the uracil DNA glycosylase inhibitor (UGI) which easily increases mutation rate to the saturation level in yeast. Although it has been reported that truncated PmCDA1 (1–161) with UGI showed editing efficiency in yeast comparable to that of full-length PmCDA1 (1–208)³², our versions which are without UGI and fused to the C-terminus of nCas9 showed a significant decrease in activity as the truncation progressed (Fig. 3).

We next performed a series of truncations from the N -terminus of the 1–161 version by fusing to the N-terminus of nCas9. The N-terminus truncations of CDA1 (1–161) first showed further decreased activity, which was then recovered as truncation proceeded to 21 and 28 a.a (Fig. 4). The predicted structure of the protein indicated that simultaneous truncation of the N- and C- terminus minimizes cross-section and gives a smoother protein surface with less exposure of hydrophobic residues. Further

truncation to CDA1(30–150), which was predicted to be the smallest one with minimum exposure of the hydrophobic surface while retaining its enzymatic core domain intact showed recovered activity (Fig. 4). These results suggest that the changes in their editing activity were attributed to the conformational stability of the protein.

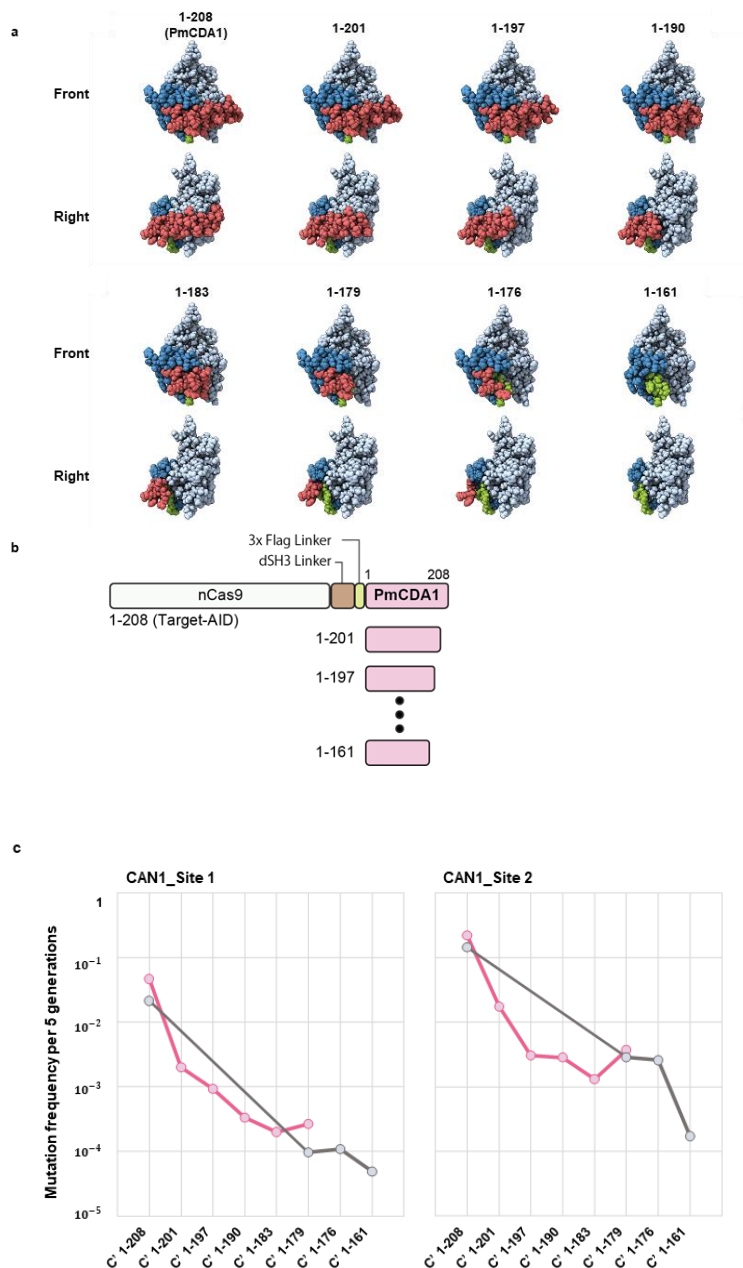


Fig. 3_Effect of truncating the C-terminus of PmCDA1. a, A series of space-filling structure predictions for C-terminally truncated PmCDA1. The non-catalytic dsDNA binding domain is shown in green (N-terminus) and red (C-terminus), respectively. The blue segments represent adjacent moieties to be truncated to smooth protein shape and minimize section. b, C-terminally truncated

Target-AID constructs tested in c. c, Trends of on-target editing efficiencies for the truncated constructs in yeast. Occurrences of canavanine resistant mutants were measured as CAN1 gene mutants. Trend lines for different data sets (pink and grey dots) are plotted for CAN1-1 and CAN1-2 target sites.

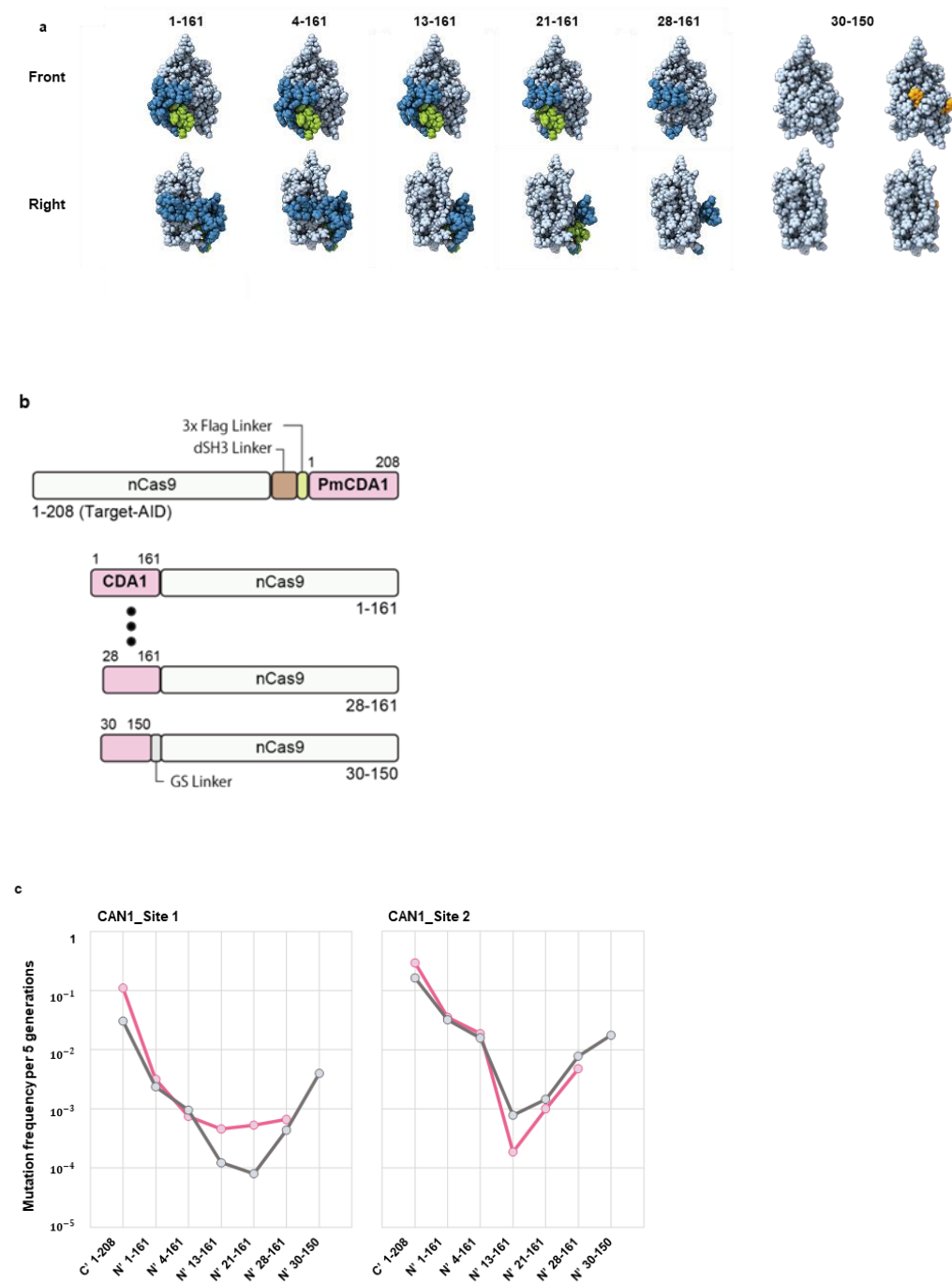


Fig. 4_Effect of truncating the both N- and C-termini of PmCDA1. a, A series of space-filling structure predictions for N- and C-terminally truncated PmCDA1. The non-catalytic dsDNA binding domain is shown in green (N-terminus) and the blue segments represent adjacent moieties to be truncated to smooth protein shape and minimize section. Potential amino acid substitution sites of 30-150 version are marked in orange. b, N- and C-terminally truncated Target-AID constructs tested in c. c, Trends of on-target editing efficiencies for the truncated constructs in yeast. Occurrences of canavanine resistant mutants were measured as CAN1 gene mutants.

< 3-1-2. Target-AID2S, 3S>

To further improve CDA1(30–150) activity, we introduced a series of mutations to the hydrophobic residues that were exposed after the truncation. Six mutations were tested in the first round and W122E was found to significantly gain activity to CDA1(30–150). Additional seven mutations were tested in combination with W122E to find W139R/Q with further improvement of the activity. Satisfyingly, the 30-150 truncated with W122E and W139Q elicited the best on-target efficiency and we named the new engineered PmCDA1 as tCDA1EQ (Fig. 5, 6). As the engineered deaminase supposedly has less affinity to DNA by itself and might be less stable than the original PmCDA1, nCas9-fusion architecture may have a greater impact on its base editing property. By combining the tCDA1EQ and to nCas9, we generated novel Target-AID2S (Small, Specific) system. Target-AID2S (PmCDA1: 121 amino acids in size) achieved around 80% recovery of on-target efficiency at site 1 and 95% recovery at site 2 in contrast to the conventional Target-AID system (PmCDA1: 208 amino acids in size).

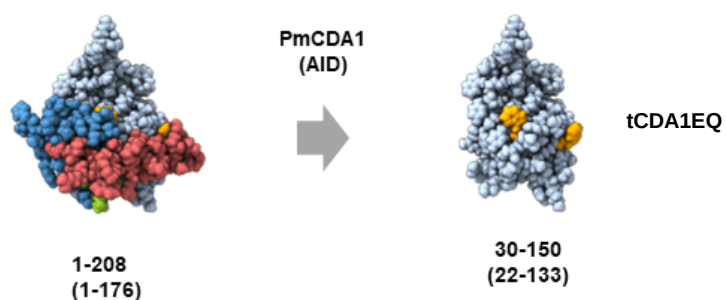


Fig. 5_Truncation of 30-150. The predicted space-filling structure of PmCDA1 before and after engineering. In addition to the direct DNA-binding sites (green and red), segments shown in blue were trimmed to minimize the protein section. The mutated amino acids (W122 and W139) are marked in orange.

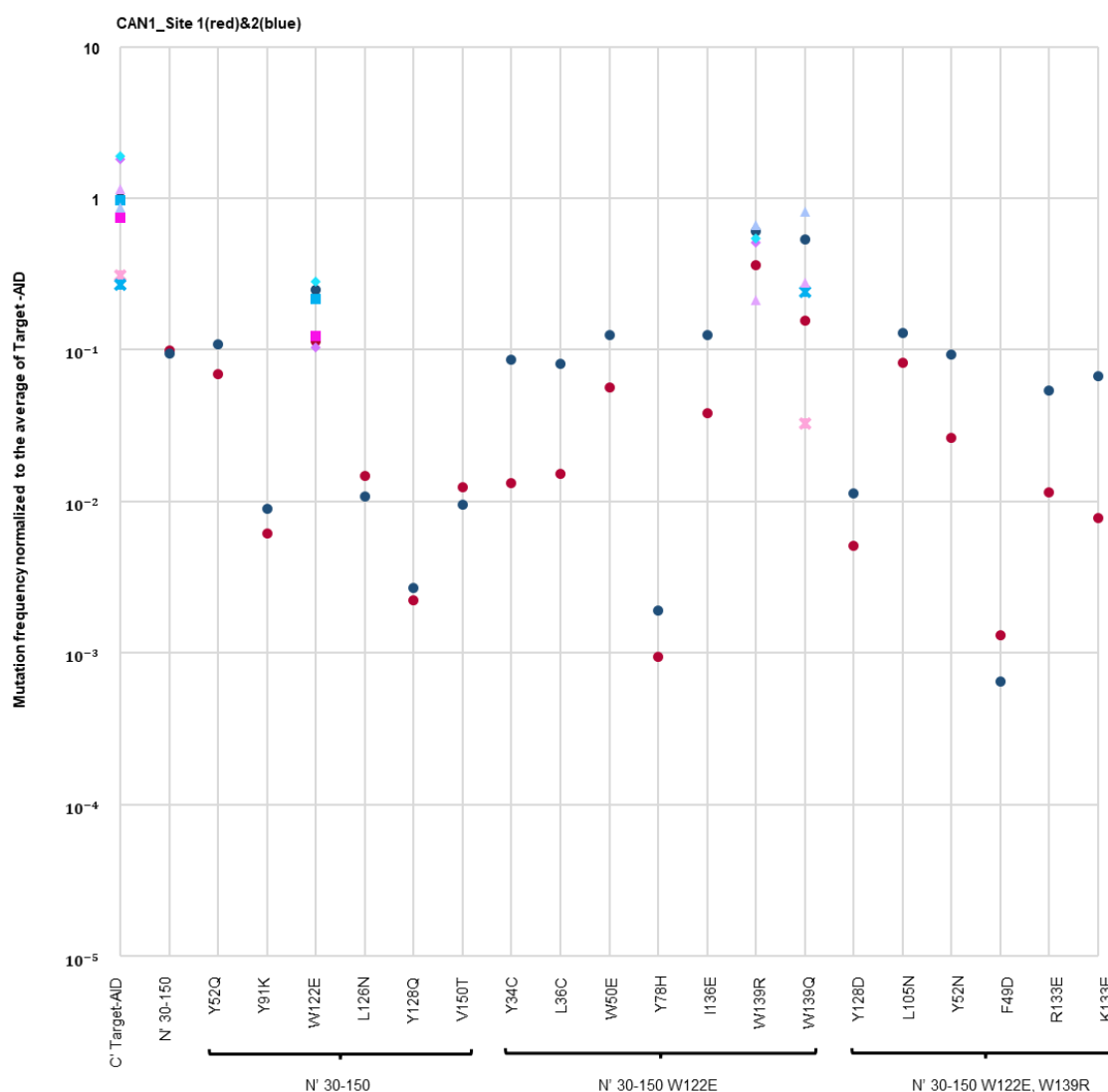
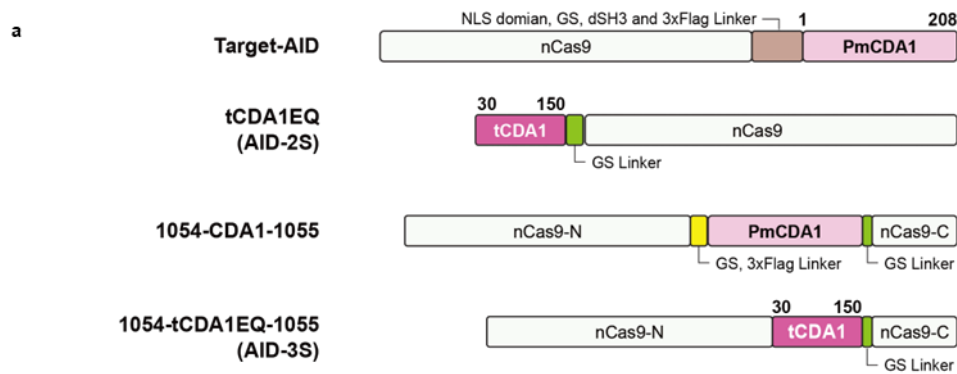


Fig. 6_Effect of amino acid substitutions in the truncated PmCDA1 (30-150). Hydrophobic residues exposed after truncation were subject to substitution with hydrophilic residues. On-target editing efficiency was measured by yeast canavanine assay and normalized to the average of Target-AID. Biological replicates were plotted for CAN1-1 (red) and CAN1-2 (blue) target sites. Different data sets are shown in different dot shapes.

The spatial interaction between proteins is an important factor that affects overall functional protein-group performance, for instance, PmCDA1 and SpCas9 protein. Thus far, PmCDA1 was designed to connect either N or C-terminal of SpCas9 protein. By exploring previous studies, we found the domain-inlaid type of constructs revealed versatile protein performance presented in APOBEC1⁴⁴. Subsequently, we questioned whether the domain-inlaid of PmCDA1 into the SpCas9 protein could alter the protein stability and overall functional performance like editing efficiency and editing window. The new guide CAN1_site 3 contains a different editing window introduced in this experiment. By investigating the relative spatial position and overall secondary structure between proteins, 4 different types of inserted PmCDA1 were established based on the prediction of their minimum functional disruption and then assayed in yeast. Intriguingly, we found N1054-CDA1 (1054-CDA1-1055) maintained the equivalent editing efficiency compare to Target-AID. Then, we applied this PmCDA1 intradomain construct to our Target-AID2S system to see if the performance could be altered or further facilitated. Surprisingly, the intradomain Target-AID3S (small, specific, and superior) we generated revealed further editing efficiency boosted 12.2-fold at on-target site 3 compare to Target-AID2S and showed equivalent efficiency to conventional Target-AID on both on-target site (Fig. 6). We reasoned that the steric limitation of domain-inlaid help hinders the stochastic movement of PmCDA1, and thus obtained superior specific-target accessibility.



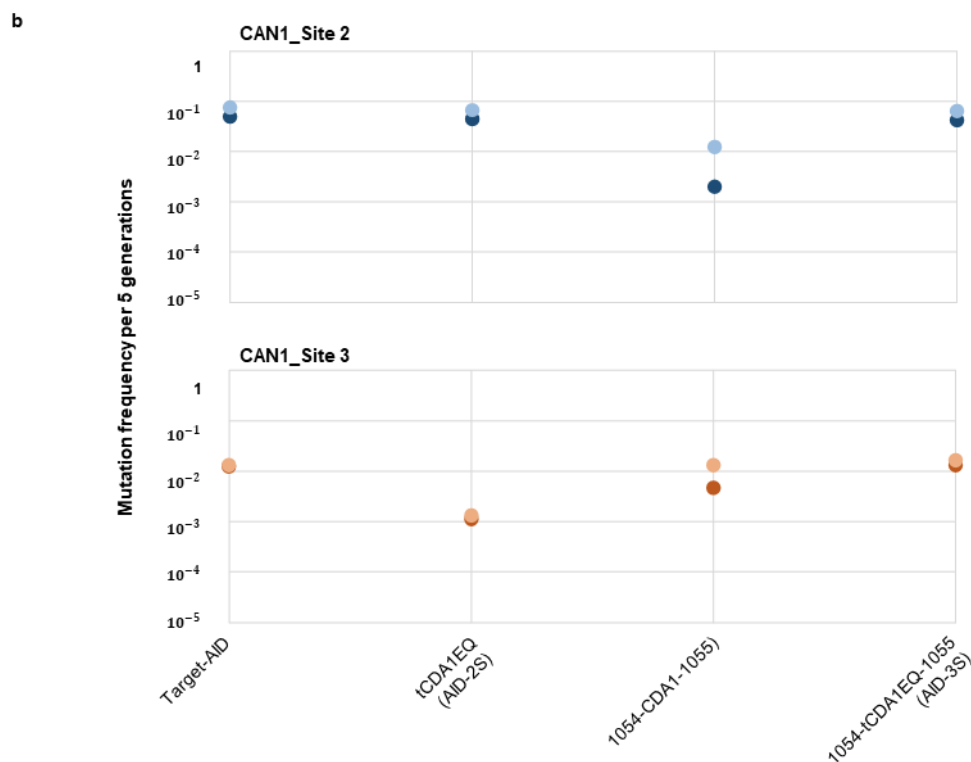


Fig. 6_ Development of domain-inlaid type Target-AID3S. a, the domain-inlaid AID-3S (1054-tCDA1EQ-1055) was developed in position of RuvC domain of Cas9. b, On-target editing efficiencies evaluated in yeast. Biological triplicates were plotted for CAN1-2 (blue dots) and CAN1-3 (orange dots) target sites.

Finally, we evaluated the overall on/off-target performance of novel Target-AID2S, 3S to conventional Target-AID system. For on-target analysis, we tested on CAN1 site 2&3. Target-AID3S revealed high on-target activity comparable to conventional Target-AID. Target-AID2S showed a comparable on-target activity on CAN1 site 2, but a decrease in CAN1 site 3. The reason for this is probably due to the target C for site 3 is closer to PAM sequence which is not preferable editing site for Target-AID2S (Fig. 7). The further editing window evaluated is needed to clarify this result.

To obtain the amenable off-target analysis data, we stuck a UGI domain to Target-AID, Target-AID2S, and Target-AID3S to boost their editing efficiency to the same level in order to omit the factor that high level off-target raised from high on-target editing efficiency. By comparing with conventional Target-AID, Both Target-AID2S and Target-AID3S revealed a remarkable deduction of genome-wide

off-target down to ~40 (site 2) and ~24 folds (site 3) respectively, whilst maintained a desirable on-target efficiency (Fig. 7). This result proved our hypothesis that the deletion of non-specific dsDNA affinity areas in PmCDA1 would help to mitigate the genome-wide off-target effect.

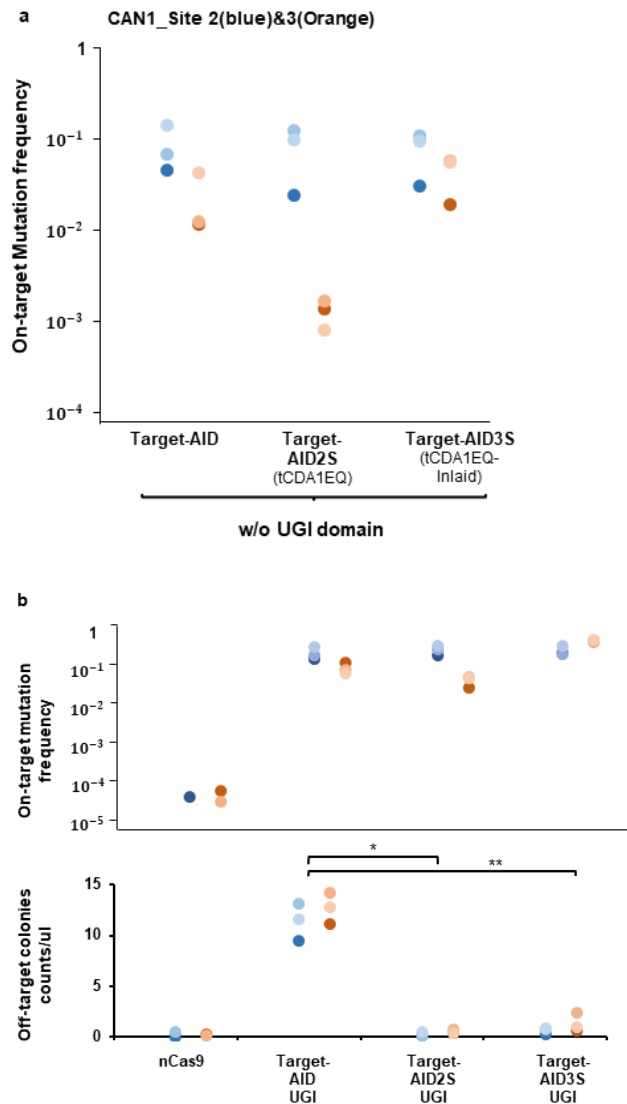


Fig. 7_ On-target and off-target profiles of the Target-AID variants assayed in yeast. a, On-target editing efficiencies for Target-AID, AID-2S, and AID-3S without UGI in yeast canavanine-resistance assay. CAN1-2 (blue dots) and CAN1-3 (orange dots) were selected as the target sites and the biological triplicates were plotted n=3. b, Occurrences of on-target mutation with UGI (canavanine resistance) and off-target mutation (thialysine resistance) were measured after induction of each construct as indicated in yeast. Biological triplicates were plotted for CAN1-2 and CAN1-3 target sites n=3. Two-tailed unpaired Student's t-test; p-value ($P^* = 1.03E-08$ and $P^{**} = 3.56E-08$).

3-2. Mammalian HEK293T cell assay

< 3-2-1. Variants and methodology for comparison >

To expand the application scope, we then evaluated our Target-AID2S and Target-AID3S in HEK293 mammalian cell and compared with YE1, YE2, R33A+K34A base editors³⁹ that reported with appreciable on-target efficiency, whilst maintained the limit off-target effect. 8 on-target sites used in previous studies were selected for evaluation (Fig. 8).

Target Sites

	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	PAM
1	G	A	A	C	A	C	A	A	A	G	C	A	T	A	G	A	C	T	G	C	
2	G	G	C	C	C	A	G	A	C	T	G	A	G	C	A	C	G	T	G	A	
3	G	G	C	A	C	T	G	C	G	G	C	T	G	G	A	G	G	T	G	G	
4	G	T	C	A	T	C	T	T	A	G	T	C	A	T	T	A	C	C	T	G	
5	G	A	C	C	C	C	C	T	C	C	A	C	C	C	C	G	C	C	T	C	
6	G	A	C	C	C	T	C	A	G	C	C	G	T	G	C	T	G	C	T	C	
7	G	G	C	C	T	C	C	G	T	A	T	C	A	C	T	C	T	C	T	G	
8	C	T	T	C	C	A	G	T	T	G	A	A	T	T	C	A	C	C	A	G	

Fig. 8_ The eight on-target sites. The selection of on-target sites is based on the editing C arranged through all position in the gRNA sequence for lower the bias for inefficiency on-target evaluation.

Analytical methods have also been developed to evaluate gRNA-independent off-target potential of the base editors. Genome-wide mutations can be thoroughly elucidated by whole genome sequencing (WGS). However, WGS is expensive, time-consuming, and has low throughput. In addition, it may not be sensitive enough for further comparative analysis of the improved base editors. Previous WGS-based studies had indicated that actively transcribed regions were prone to base editing off-target effect, because the R-loops which is formed by the exposure of single-stranded DNA by RNA transcription is a preferred substrate for the deaminases. To mimic such hot spots, localized R-loop was formed by using an orthogonal nuclease-defective CRISPR system which was co-transfected with base editing systems targeting another distant locus, and this is called R-loop assay³⁹ that report in the previous study (Fig. 9). Deep sequencing of the R-loop region allowed for rapid and sensitive comparative evaluation of gRNA-independent off target potency.

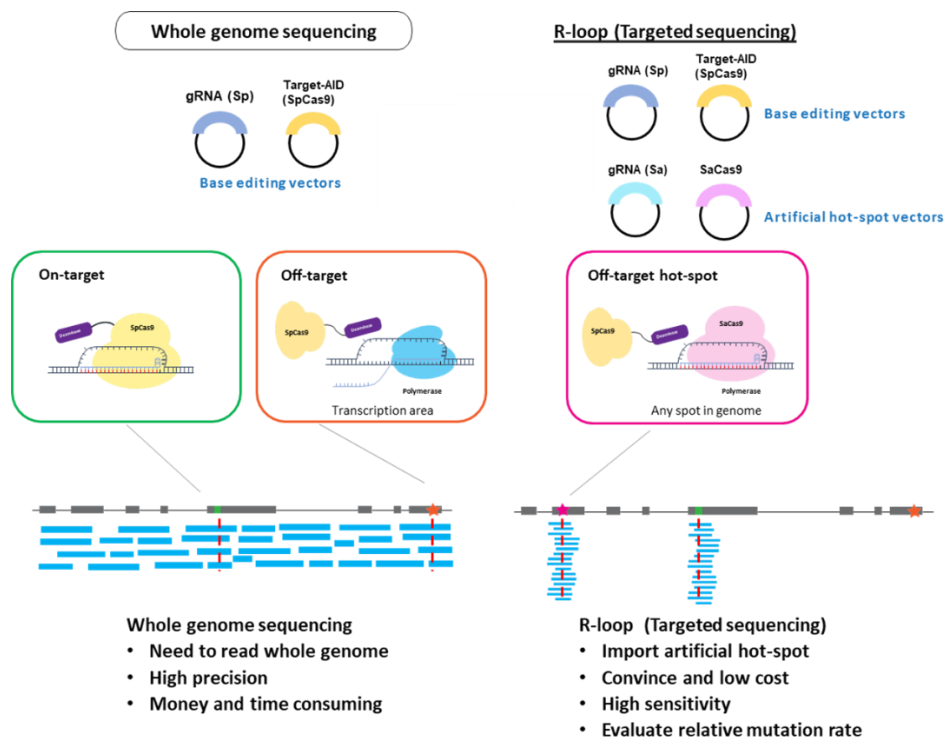
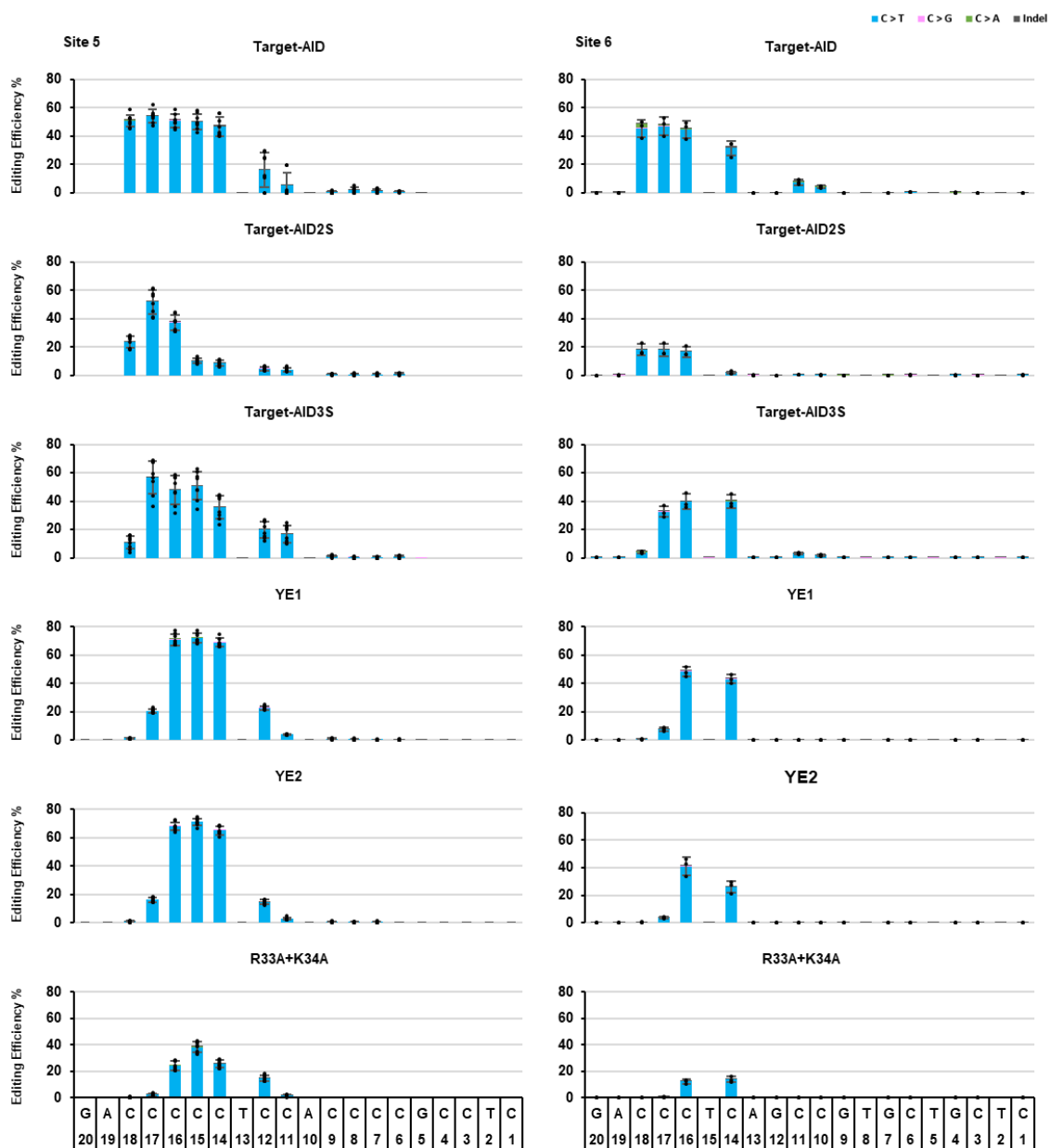


Fig. 9_ The mechanism of R-loop off-target assay. Compare to WGS, R-loop assay is cost-efficiency method to evaluated gRNA-independent off-target that is easier to conduct in lab-based research.

< 3-2-2. On-target performance >

The on-target activity presented with the 20nt on-target sequence from PAM sequence (Fig. 8, 10). Target-AID was capable to perform >20% of editing to all sites. Target-AID2S and Target-AID3S also performed well except for site 8. Interestingly, at site 1, the high C to G conversion also was observed at C-15 (25%) which indicated the UGI activity might be hindered at site 1 or at certain target context Target-AID3S may embrace the greater potential to C to G mutation rather than C to T. Both Target-AID2S and Target-AID3S presented less indel and C to A conversion relative to Target-AID. YE1 and YE2 were reported as alternate deaminase type for APOBEC1 with restricted editing window and less off-target. In this on-target assay, those two variants revealed comparable on-target performance with Target-AID3S. APOBEC1 R33A+K34A was designed as the novel platform for a lower off-target RNA editing but also exhibited reduced DNA off-target. However, it showed target-context dependent which is consistent with the find in previous study reported its displayed highly target-dependence in this study also that obtain >20% of editing only in sites 3, 4, and 7. Compare with the preference of 5' T or C context APOBEC1 CBE variants, PmCDA1 CBE variants exhibit less affinity to certain motif context (Fig. 10).



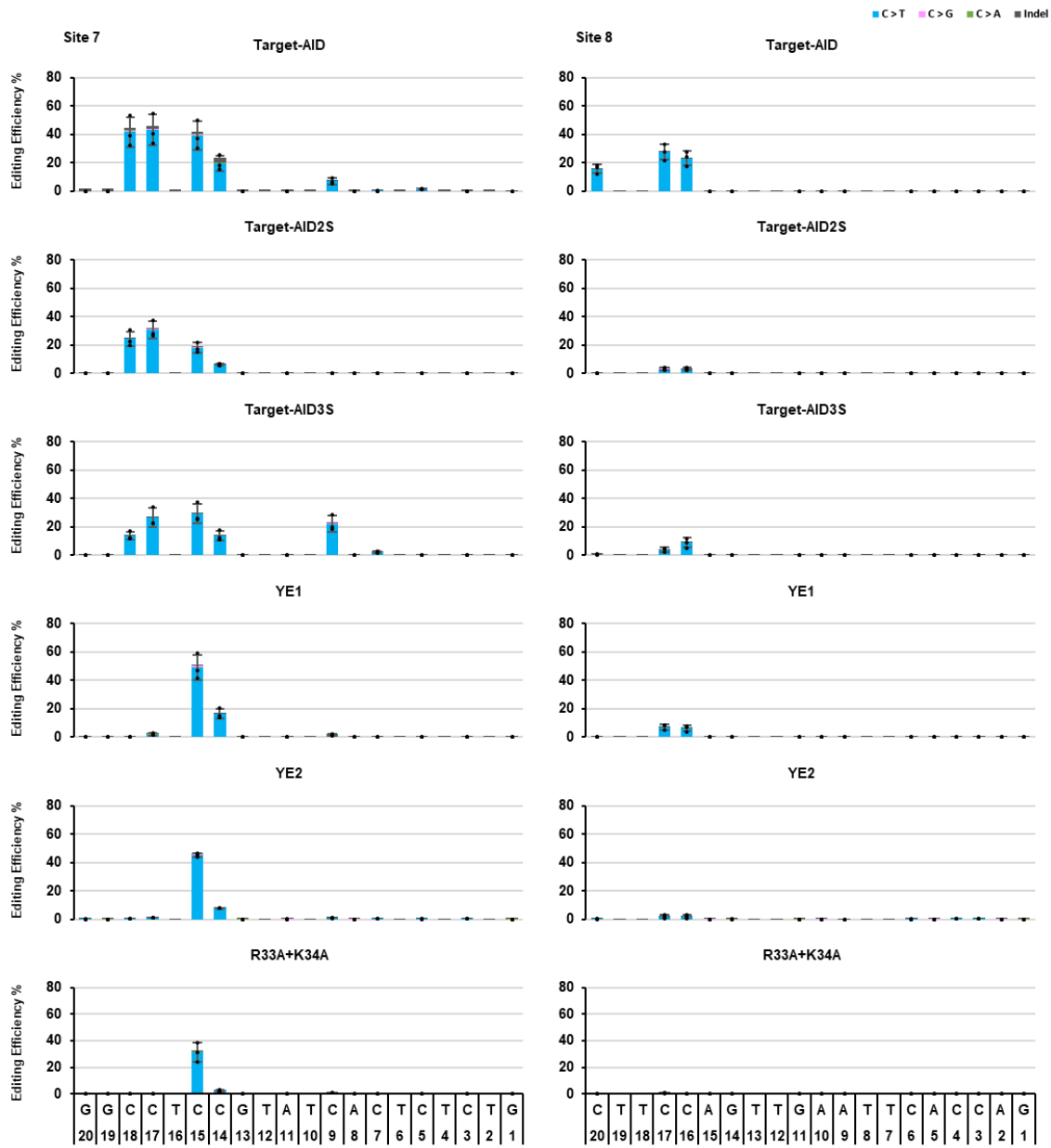


Fig 10. On-target editing profiles of the CBE variants. On-target sites (1~8) were selected and analyzed by deep sequencing. Mutation frequencies at each C position (numbered from the PAM sequence side of the target sequence toward the 5' side) are shown. Data are presented as mean values $\pm$ s.d. of C to T conversion (n=6 for site 1~4, n=3 for site 5~8).

< 3-2-3. Editing window >

After sum up the data of 8 on-target sites, we could conduct the editing window analysis for each variant (Fig. 11). Conventional Target-AID revealed high C affinity that induced widely C to T mutations from C-18 to C-14. With engineered platform Target-AID2S, the editing window showed magnificent narrowed down, meanwhile, the editing efficiency of most affinity C-18, 17 was guaranteed. This cause was thought due to protein truncation led to the weaker target affinity and mitigated deaminase activity of Target-AID2S. On the other hand, domain-inlaid type Target-AID3S unveiled the shifting of editing window toward the PAM sequence and retained two editing peaks at C-16 and C-13. This phenomenon and wider editing window detected were speculated due to insertion of PmCDA1 into nCas9 alter the spatial distance to target sequence and make it easier to approach. For those APOBEC1 derived CBE variants (YE1, YE2, and R33A+K34A), the editing window demonstrated consistent with the previous reports that most affinity at C-15.

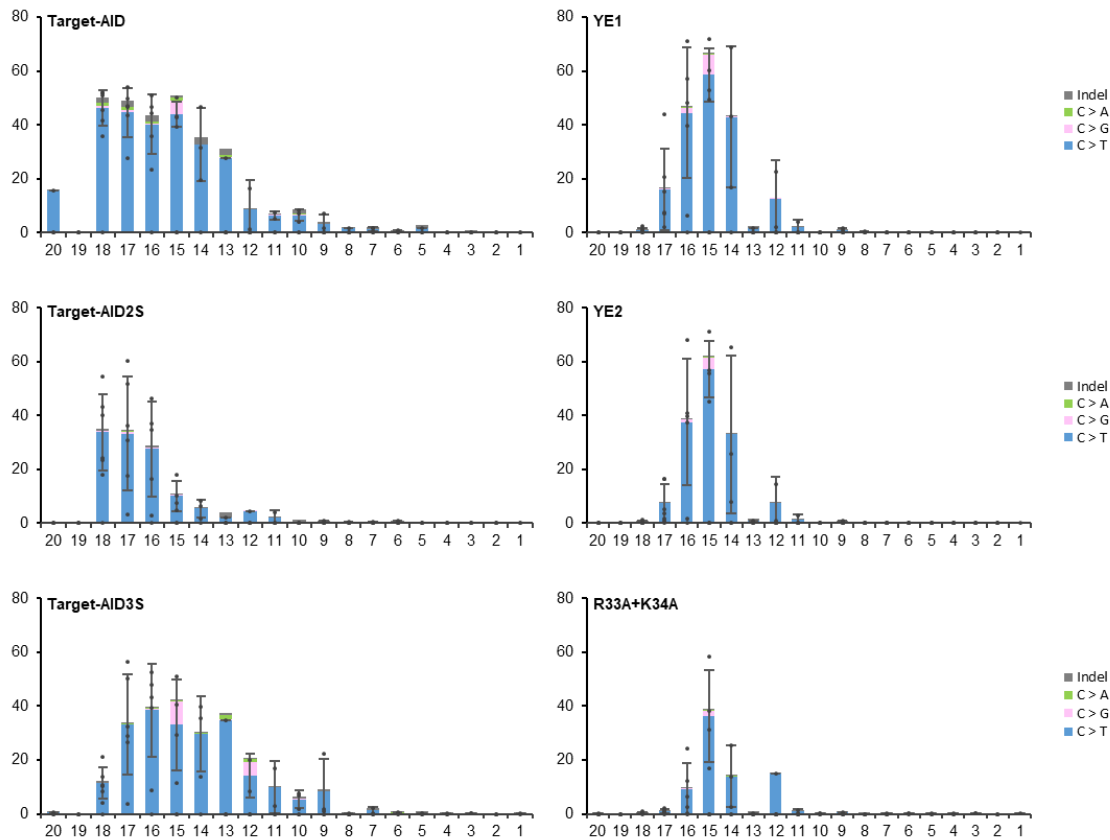


Fig. 11_ Evaluation of editing window. Each position averaged the editing efficiency cross all sites. The nucleotide C positions (numbered from the PAM sequence side of the target sequence toward the 5' side) with the averaged C to T conversion frequencies for all target sites are shown. Data are presented as mean values $\pm$ s.d (n=8). C position alignment of each site is shown in Fig. 8.

< 3-2-4. Off-target occurrence >

For correctly and precisely address off-target characterize of our newly built Target-AID2S, 3S, we conducted gRNA-in/dependent RNA/DNA off-target evaluation. We recruited YE1, YE2, R33A+K34A, BE4max for comparison.

RNA off-target

We conducted target RNA sequencing of three mRNA sites (Fig. 12, CTNNB1, RSL1D1, and IP90) which is reportedly to exhibit high gRNA-independent off-target⁴⁵ and six gRNA-dependent mRNA sites (Fig. 13, HEK3_1, 2, 3 and RNF2_1, 2, 3)⁴⁶. Both gRNA-independent and gRNA-dependent RNA-editing exhibited by BE4max as reported was significantly lowered by the all-other variants tested, while YE1 retained detectable RNA off-target at several sites. Original Target-AID and the new variants did not show significant RNA off-target, consistent with the previous report. Target-AID2S, 3S also did not revealed any RNA off-target as same as Target-AID (Fig. 12, 13).

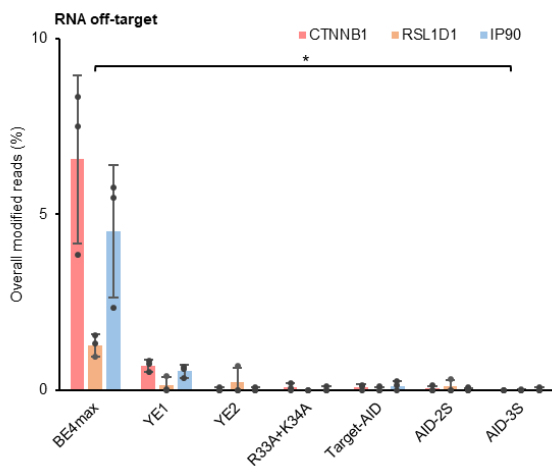


Fig. 12_RNA independent off-target. Targeted RNA sequencing of three gRNA-independent locus reported before (CTNNB1, RSL1D1, and IP90) was performed for four rAPOBEC1 CBEs and three Target-AID CBEs. Data are presented as mean values $\pm$ sd ($n = 3$). Statistically significant differences between BE4max and other CBEs were supported by two-tailed unpaired Student's t-test; p-value ($P^* = 1.98E-03 \sim 2.36E-03$ for RSL1D1 site, $1.45E-02 \sim 2.18E-02$ for IP90 site, $8.90E-03 \sim 1.33E-02$ for CTNNB1 site)

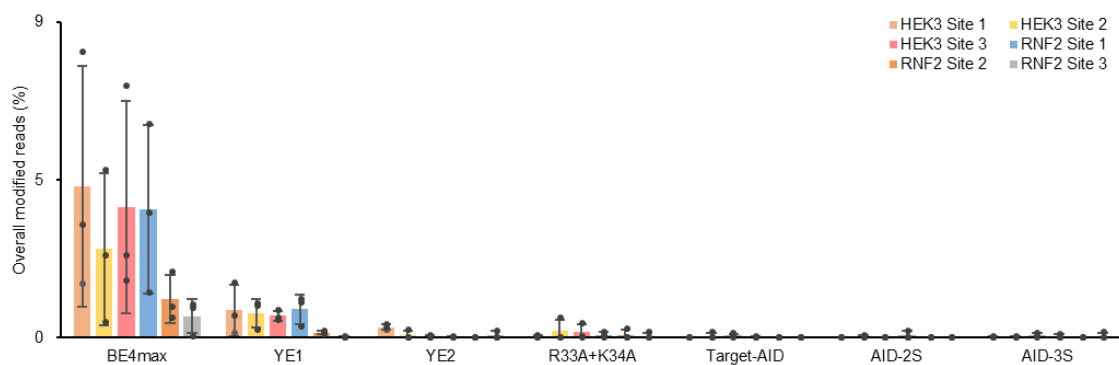
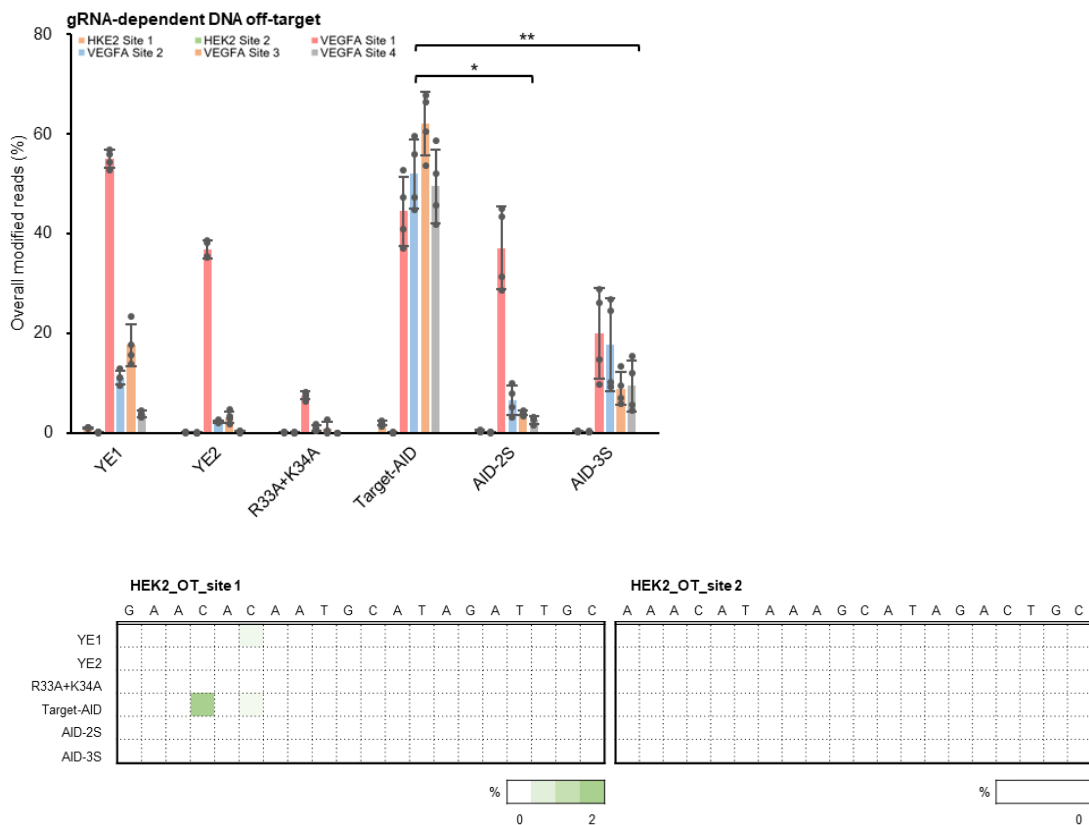


Fig. 13_RNA dependent off-target. Target RNA sequencing was conducted with six gRNA-dependent mRNA sites as reported (HEK3_1, 2, 3 and RNF2_1, 2, 3). Data are presented as mean values $\pm$ s.d. (n=3).

gRNA-dependent DNA off-target

The gRNA-dependent DNA off-target effect was also assessed by deep-sequencing of the 6 reported sites (HEK2_OT1, 2; VEGFA_OT1, 2, 3, 4)^{47,48} (Fig. 14). Target-AID2S and Target-AID3S showed substantially reduced off-target especially in HEK2 site 1 and VEGFA site 2, 3, 4. YE2 and R33A + K34A showed improved off-target editing at all the sites analyzed compared to YE1.



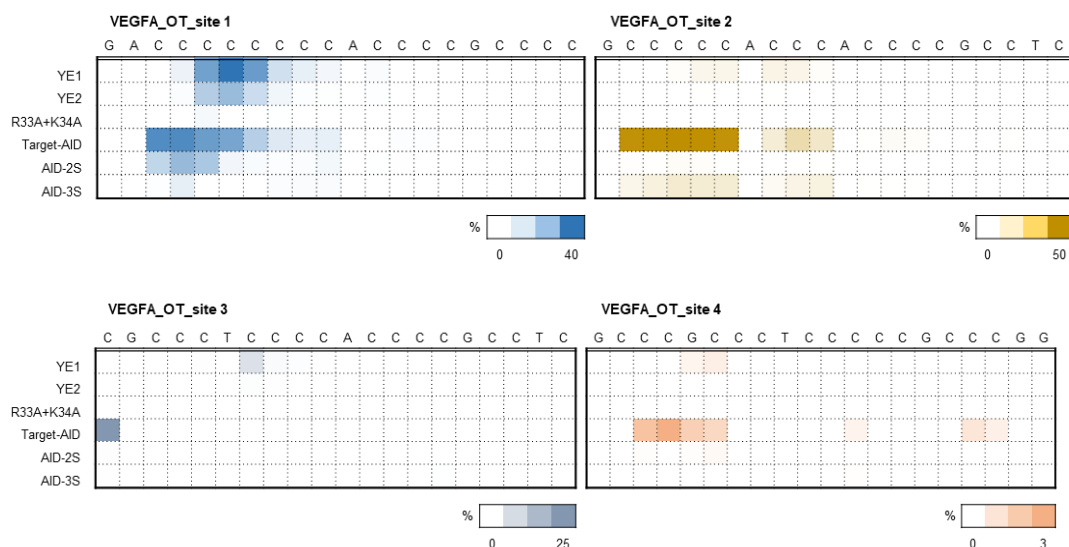
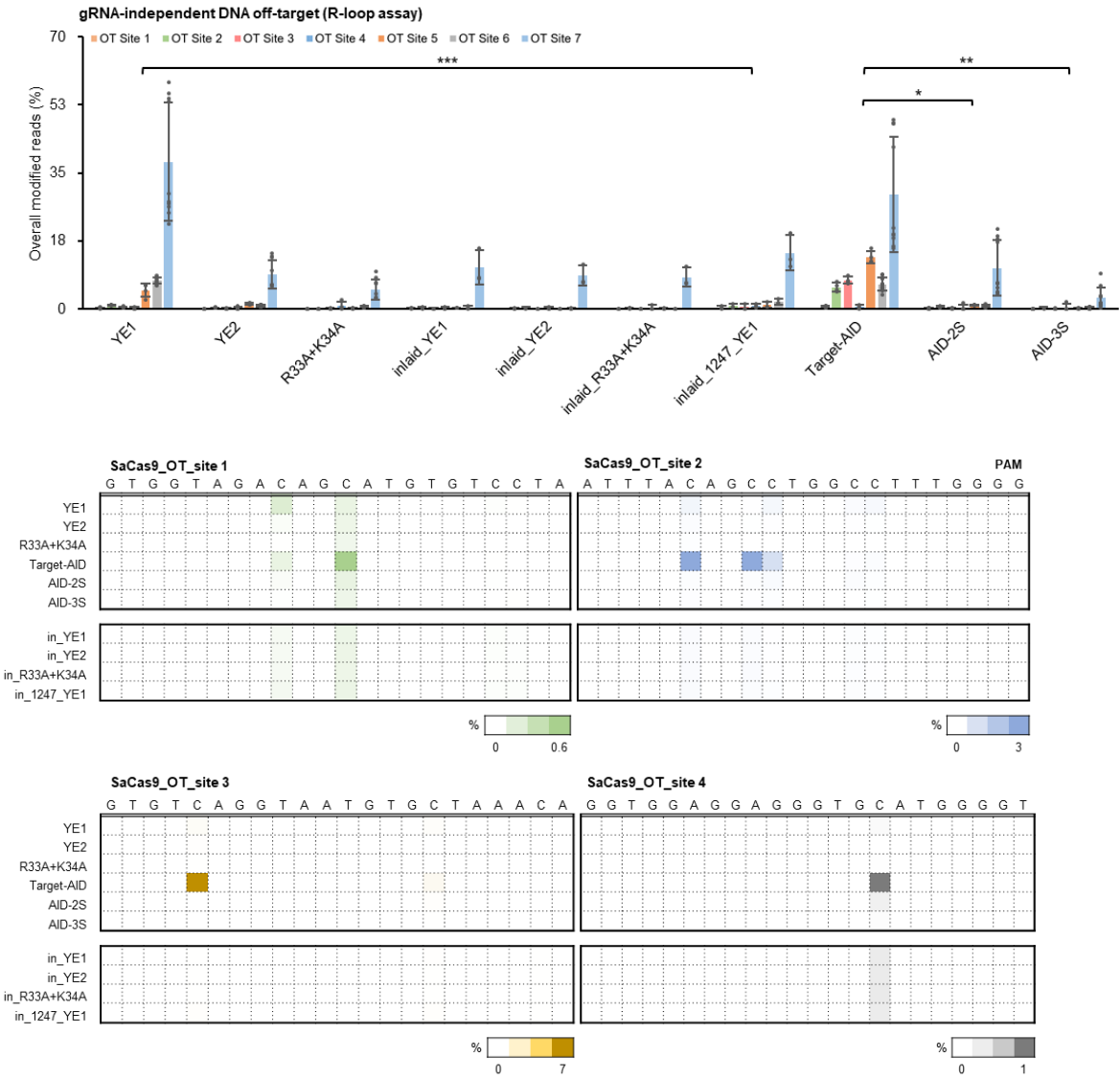


Fig. 14 gRNA-dependent DNA off-target. The two HEK2 off-target sites and four VEGFA off-target sites were analyzed by DNA amplicon deep sequencing. Target-AID2S, 3S showed significant reduction compared to Target-AID at HEK2 site 1 and VEGFA site 2, 3, and 4, supported by the two-tailed unpaired Student's t-test; p-value ($P^* = 1.79\text{E-}06 \sim 9.88\text{E-}04$ for AID2S, $P^{**} = 6.02\text{E-}06 \sim 8.13\text{E-}04$ for AID3S). The mutation frequencies at each nucleotide position are shown in heat diagram.

gRNA-independent DNA off-target:

The gRNA-independent DNA off-target effects were assessed by using orthogonal SaCas9 R-loop assay in HEK293T cells. SaCas9 off-targets site 1~6 was selected following the previous studies³⁹ and an additional site 7 (VEGFA locus) was chosen as its C-rich context may provide higher sensitivity to deamination by CBEs. Target-AID showed detectable off-target editing at all seven sites (Fig. 15), while Target-AID2S showed no detectable off-target occurrence in the site 1, 3 and significantly reduced off-target editing at site 2, 5, 6, 7, which was comparable to YE2 and R33A + K34A. YE1 showed rather higher off-target editing at site 6 and 7. Target-AID3S showed the lowest, hardly detectable off-target editing across all seven sites tested. This may be attributed to the inlaid architecture which sterically limits the access of the enzyme beyond Cas9-binding DNA strand, in addition to the eliminated DNA affinity. On average, Target-AID2S and -3S respectively exhibited approximately 4.5-folds and 13.7-folds reduction of R-loop off-target editing compared to the original Target-AID while maintaining appreciable on-target editing efficiency (Fig. 10, 15). Collectively, consistent with conclusions conducted in yeast experiment, Target-AID2S and Target-AID3S also

elicited remarkably off-target mitigation in the mammalian cell in stark contrast with canonical Target-AID which proved our hypothesis that complete removal of dsDNA affinity domain would adequately reduce the off-target effect. Furthermore, domain-inlaid also showed as an effective strategy to attenuate the off-target effect.



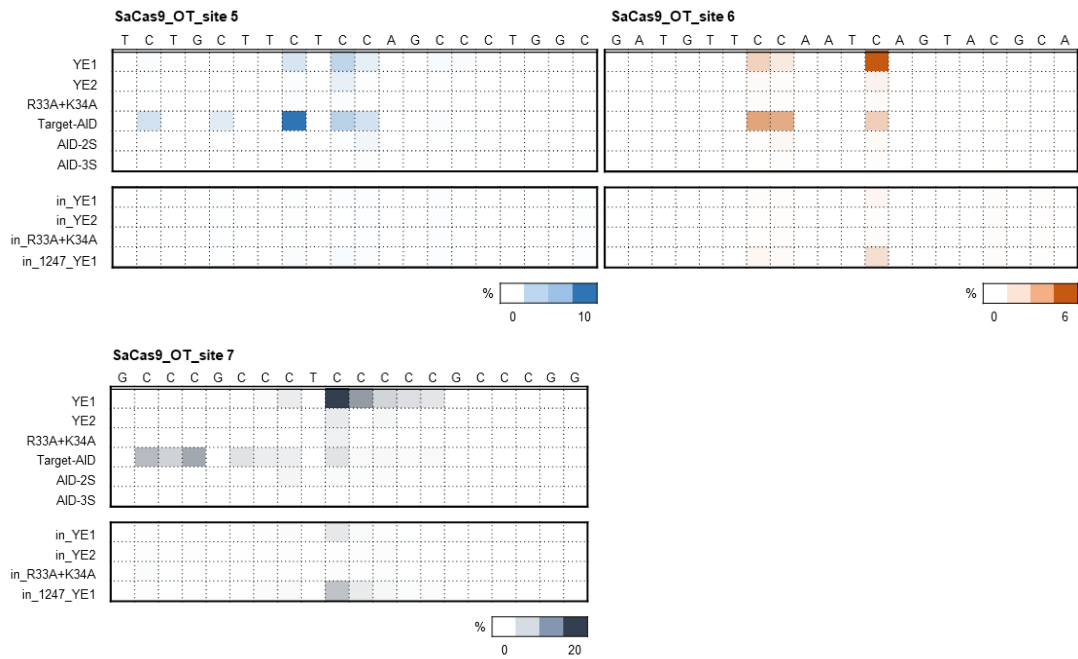


Fig. 15_gRNA-independent DNA off-target. gRNA-independent DNA off-target was compared by using seven off-target R-loop sites. Compared to Target-AID, AID-2S and -3S showed significant reduction across all sites, supported by two-tailed unpaired Student's t-test; p-value ($P^* = 3.36\text{E-}06 \sim 5.37\text{E-}03$ for AID2S, $P^{**} = 2.89\text{E-}06 \sim 8.93\text{E-}04$ for AID3S). Inlaid_1247_YE1 also showed a significant difference to YE1 at site 5, 6 and 7 ($P^{***} = 4.59\text{E-}02, 7.59\text{E-}04, 1.59\text{E-}04$). Off-target frequencies are shown as the percentage of reads containing the mutation. Data are presented as mean values $\pm$ sd ($n > 3$). The mutation frequencies at each nucleotide position are also shown in heat diagram.

Inlaid architecture investigation:

To further define the effect of domain inlaid architecture, inlaid-YE1, inlaid-YE2 and inlaid-R33A+K34A were constructed, resembling Target-AID3S architecture which splits Cas9 at 1054 a.a. position. Additionally, the 1247 a.a. position of Cas9 which is also predicted as a flexible position⁴⁴, was tested as inlaid (1247)-YE1. These rAPOBEC1 inlaid variants, however, showed overall reduction in on-target editing efficiency, especially inlaid-YE2 and inlaid-R33A+K34A by evaluating seven target sites (Fig. 16). The off-target effect of YE1 was substantially reduced by both inlaid architectures, while no or little improvement was seen for inlaid-YE2 and inlaid-R33A+K34A (Fig. 15).

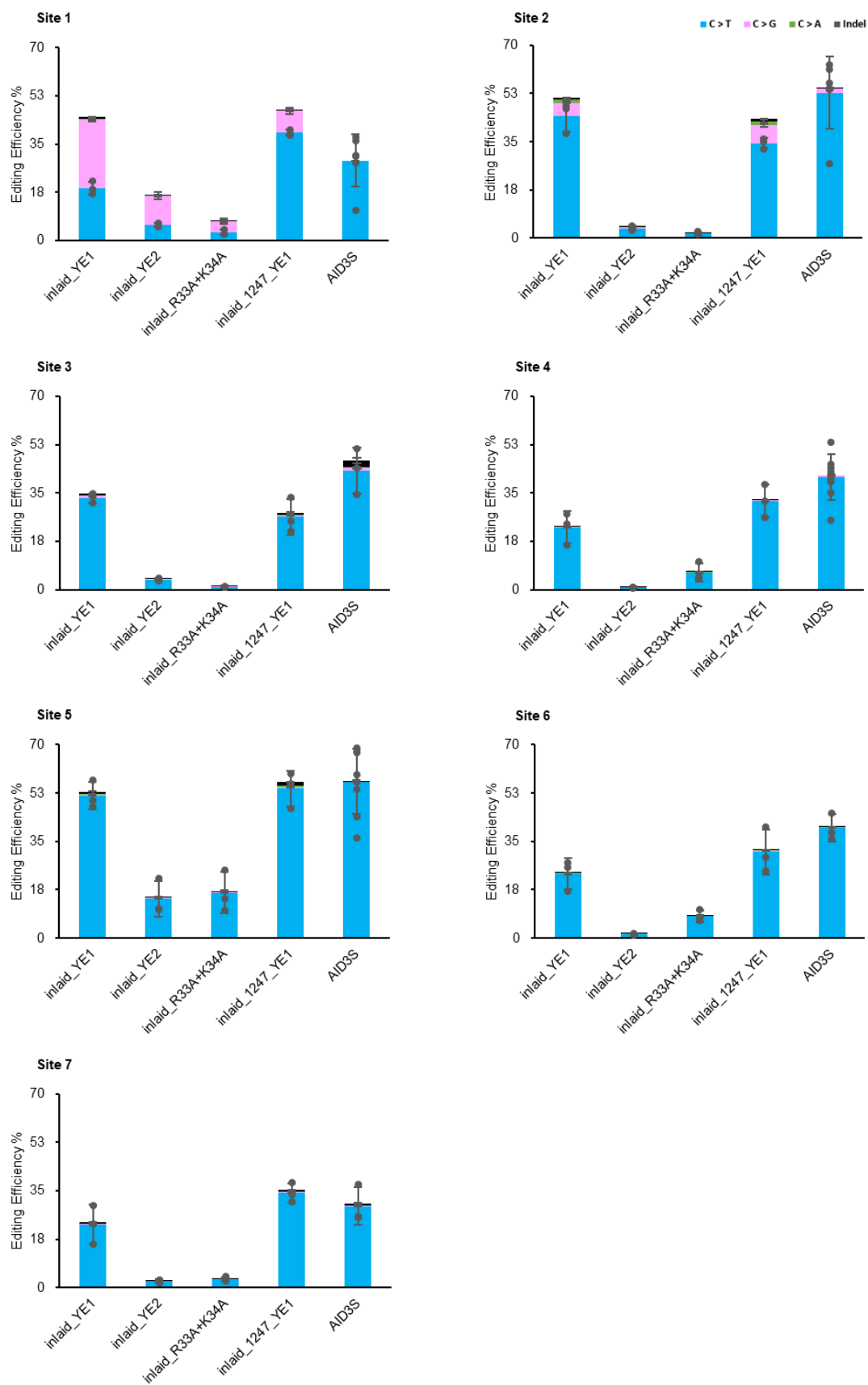


Fig. 16_ On-target editing profiles of the inlaid rAPOBEC1 CBE variants. On-target sites (1~7) were selected and analyzed by deep sequencing. The highest C to T conversion frequencies for each variant at each target are shown. Data are presented as mean values $\pm$ s.d. of C to T conversion (n=3).

< 3-2-5. On/off-target profile >

Taken together the data obtained in these assays, we plotted a landscape of candidates we identified as being contained especially promising on-target activity and minimizing genome-wide Cas9-independent off-target effect in each target site (Fig. 17).

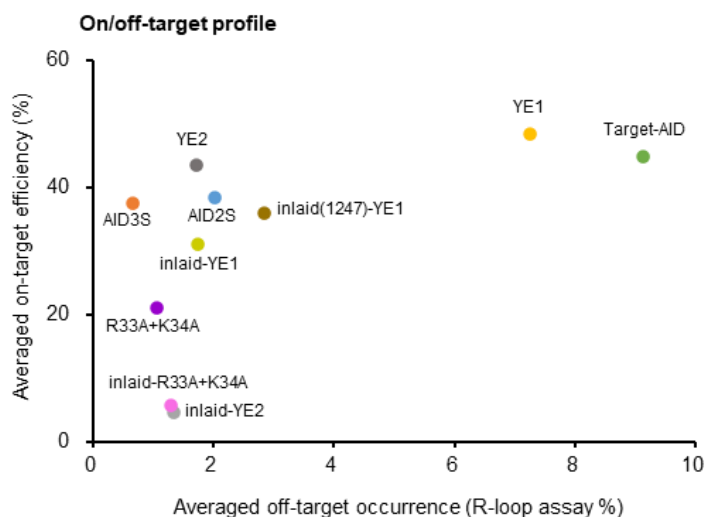


Fig. 17_ Overall on/off-target profile. On-target editing versus average off-target editing profile for all CBEs tested in this study. The y axis represents the mean on-target editing at the eight on-target, and the x axis represents the mean off-target editing at the seven orthogonal R-loop sites.

< 3-2-6. Apply into SaCas9 >

The engineered PmCDA1 (tCDA1EQ) is substantially smaller (121 a.a.) in size compared to the wild-type (208 a.a.). Smaller molecular size as a genome editing component is advantageous especially for in vivo delivery tools. Therefore, we introduced the small ortholog SaCas9 system to minimize the size of base editing system furthermore (Fig. 18). To develop SaAID3S, tCDA1EQ was inlaid into 615–616 a.a. position of nSaCas9 within the HNH domain facing to the polynucleotide-binding cleft. Small Scp1 promoter and SpA terminator were also employed to compose a total length of 4036 bp plus 332 bp of gRNA expression cassette. For comparison, the conventional form of SaCas9 version of Target-AID (SaAID) was also developed, which contains full-length PmCDA1 with linker, UGI, CMV promoter, and SV40 terminator to compose a total length of 5220 bp without gRNA cassette. To normalize transfection efficiency that may vary depending on the vector size, the transfected cells were sorted by the fluorescent signal of iRFP670 expressed from the vector backbone. At the two target sites tested, SaAID and SaAID3S showed comparable editing efficiency with differences in mutation window (Fig. 19). Relative gRNA-independent DNA off-target effect was assessed for SaAID and SaAID3S by employing four SpCas9 targets as R-loop off-target sites. SaAID showed significant off-target effect at the four R-loop sites tested, all of which were substantially reduced in SaAID3S (Fig. 20). The RNA off-target was also assessed at the three sites (CTNNB1, RSL1D1, and IP90), showing no detectable editing (Fig. 21).

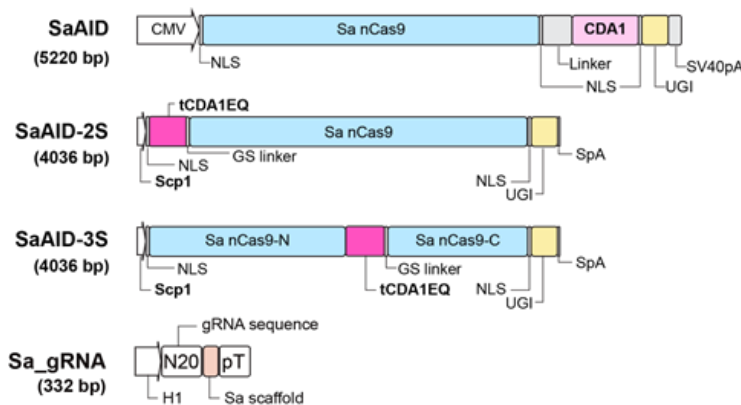


Fig. 18_ Structural design with SaCas9. Domain architectures of SaAID, SaAID2S and SaAID3S. The gRNA expression cassette was combined into each effector plasmid.

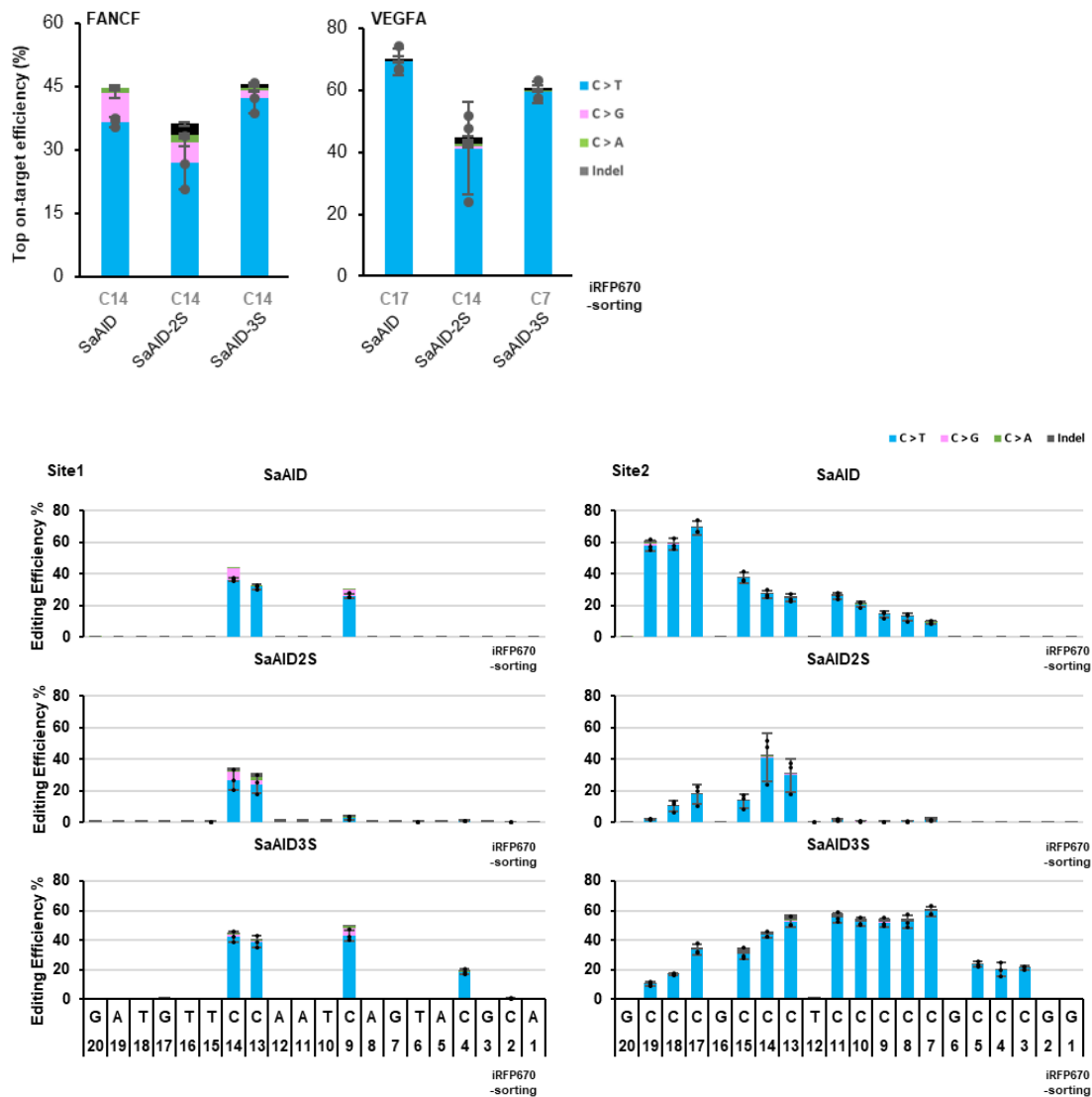


Fig. 19_ On-target editing efficiency of SaAID, SaAID2S and SaAID3S. On-target editing frequencies of SaAID, SaAID2S and SaAID3S are shown with the highest position, and data are presented as mean values $\pm$ s.d (n=3). To normalize transfection efficiencies, cells were sorted by the expression of iRFP670 from the plasmid backbone. The mutation frequencies of each nucleotide position are also shown in below.

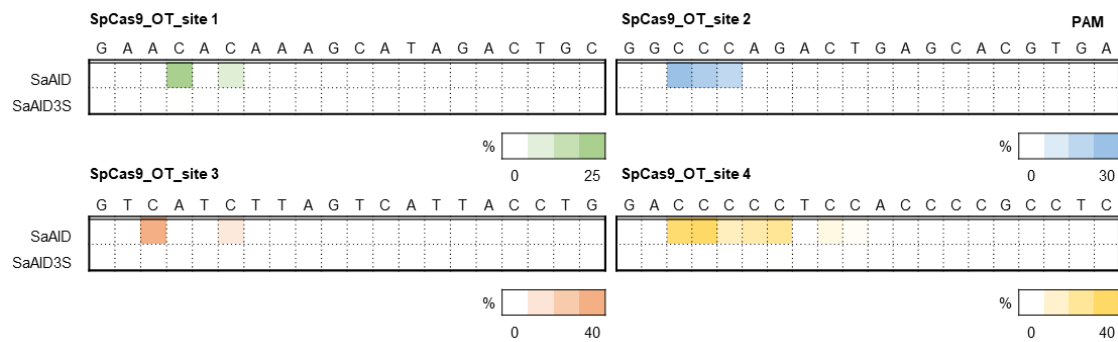


Fig. 20_ gRNA-independent DNA off-target of SaAID and SaAID3S. Heat map of DNA off-target profile in gRNA-independent fashion. gRNA-independent off-target occurrences analyzed by R-loop assay and deep sequencing with four SpCas9 target sites (HEK2, HEK3, RNF2 and VEGFA). SaAID3S revealed significant difference to SaAID cross all four sites

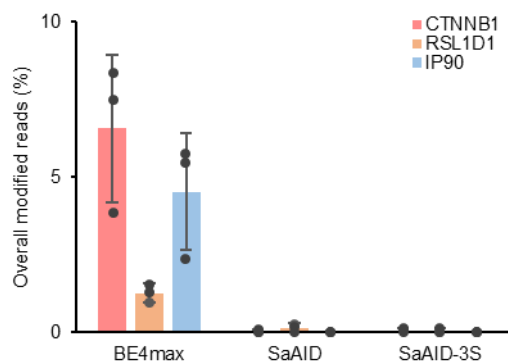


Fig. 21_ gRNA-independent RNA off-target of SaAID, SaAID2S and SaAID3S. Target RNA sequencing conducted for SaAID and SaAID3S with three gRNA-independent sites as reported (CTNNB1, RSL1D1, and IP90) in comparison with BE4max. Data are presented as mean values $\pm$ s.d. (n=3).

< 3-2-7. Load onto AAV vector >

To develop a single AAV base-editing vector, a series of viral particles were prepared with varied compositions of AAV-SaAID-2S (A2S-1 to A2S-8) and AAV-SaAID-3S (A3S-1 to A3S-7) by using Scp1, CMV, minCMV or Scp3 promoters, bGH or Short-PolyA terminators and H1 or U6 gRNA promoters (Fig. 22, 23). Their editing efficiency was evaluated in vitro by infecting HEK293T cell in comparison with the previously developed dual AAV partitioning system (DUAL-533 and DUAL-739). DUAL-533 is based on v5 *S. aureus* CBE using intein-split SaBE3.9max⁴⁹, which had been developed by splitting SaCas9 at 533 a.a. position and ligating via Npu DnaE intein. DUAL-739 is based on Npu DnaE Intein-SaKKH-BE3⁵⁰, splitting SaCas9 KKH variant at 739 a.a. position (Fig. 22). DUAL-533 retained the editing efficiency which is consistent with the previous study⁴⁹. Although no significant editing was observed for any of A3S series tested and DUAL-739, we found significant editing efficiency for the A2S series, especially those with smaller compositions (Fig. 24).

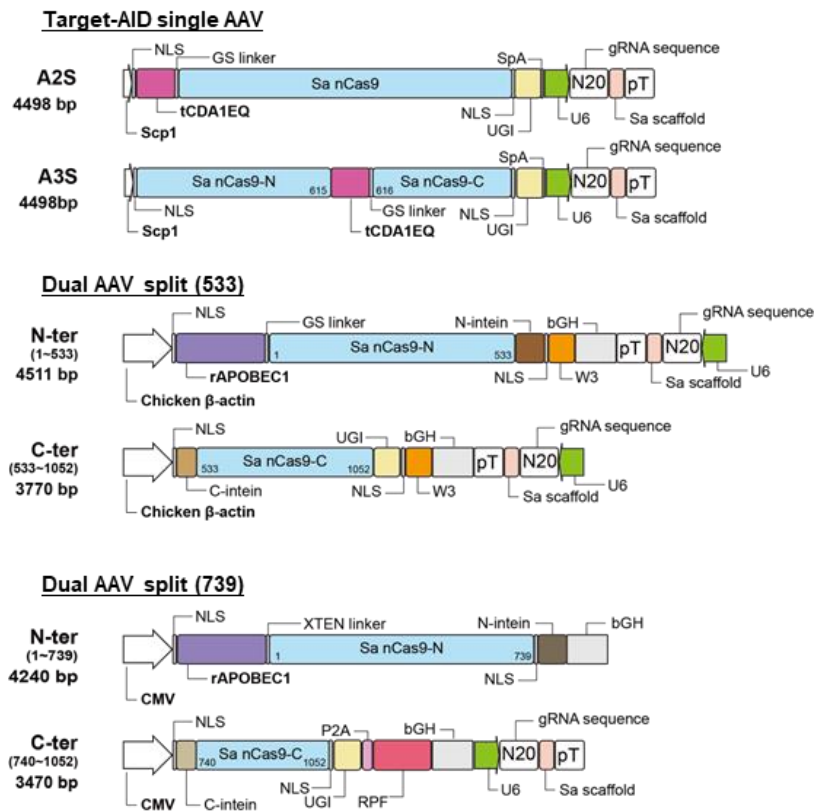


Fig. 22_ Domain architectures of AAV variants. Inside of the ITR domains is shown. The dual AAV system was introduced base on previous study.

AID2S

No	Promoter	PolyA-tail	gRNA Pro	Size(bp)	Conc (vg/ml)
A2S-1	Scp1	Short-PolyA	H1	4498	4.71E+10
A2S-2	CMV	bGH	U6	5160	4.13E+11
A2S-3	Scp1	bGH	U6	4657	2.30E+11
A2S-4	CMV	Short-PolyA	U6	5001	4.79E+11
A2S-5	Scp1	Short-PolyA	U6	4498	6.56E+11
A2S-6	minCMV	Short-PolyA	U6	4456	3.42E+11
A2S-7	Scp3	Short-PolyA	U6	4498	3.94E+11
A2S-8	Scp1	Short-PolyA	U6-flipped	4523	2.40E+11

AID3S

No	Promoter	PolyA-tail	gRNA Pro	Size(bp)	Conc (vg/ml)
A3S-1	Scp1	Short-PolyA	H1	4498	4.47E+10
A3S-2	CMV	bGH	U6	5160	6.16E+10
A3S-3	Scp1	bGH	U6	4657	6.72E+11
A3S-4	CMV	Short-PolyA	U6	5001	7.84E+11
A3S-5	Scp1	Short-PolyA	U6	4498	1.11E+12
A3S-6	minCMV	Short-PolyA	U6	4456	2.76E+12
A3S-7	Scp3	Short-PolyA	U6	4498	2.20E+12

No	Promoter	PolyA-tail	gRNA Pro	Size(bp)	Conc (vg/ml)
----	----------	------------	----------	----------	--------------

Dual AAV split (533)

rAPOBEC1_ N-ter	Chicken β -actin	bGH	U6-flipped	4511	2.91E+12
C-ter	Chicken β -actin	bGH	U6-flipped	3770	3.89E+12

Dual AAV split (739)

rAPOBEC1_ N-ter	CMV	bGH		4240	6.54E+12
C-ter	CMV	bGH	U6	3470	7.54E+10

Fig. 23_ The composition and obtained titer of AAV variants used in this study. Size is counted inside ITR domains. The concentration was measured by ddPCR.

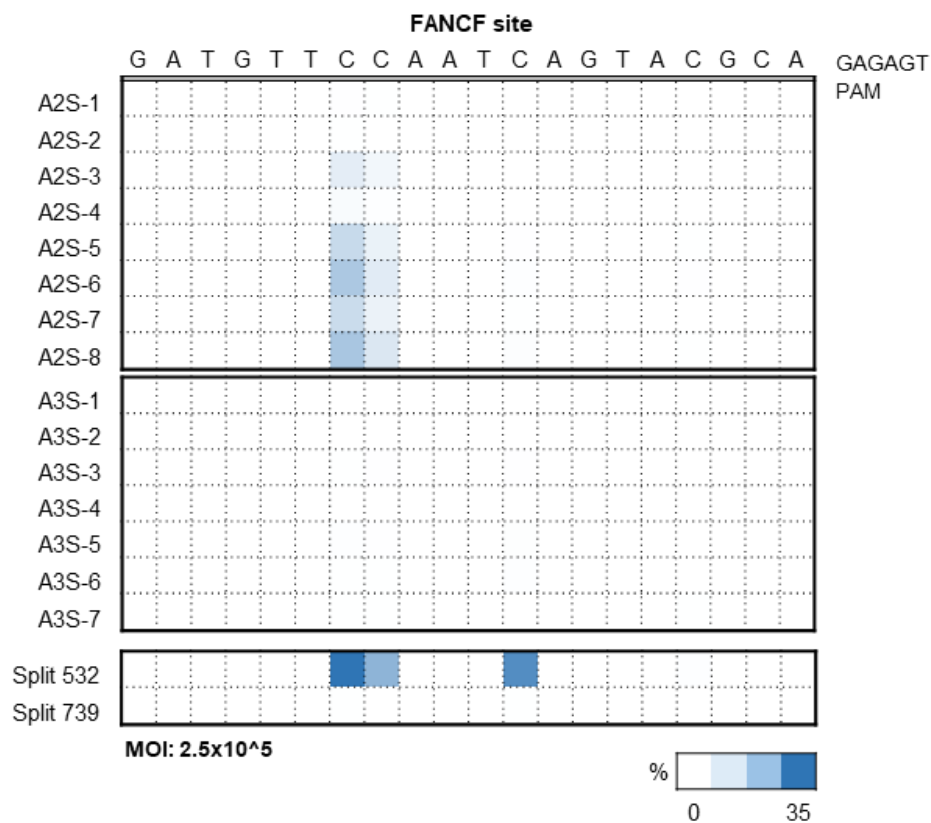


Fig. 24_ On-target editing at FANCF site. The heat map of editing was shown by AAV infection in HEK293T and analyzed by deep sequencing. The MOI: 2.5x10⁵ was selected base on the previous study. The position of each nucleotide was labeled on the top of graph.

4. Discussion

In this study, we developed high-fidelity cytosine base editors through structural engineering of the deaminase PmCDA1 to remove the potential non-specific DNA binding moiety. Although the previous studies have explored a series of C-terminal truncations of PmCDA1 to show narrowed editing window and reduced genome-wide off-targets in yeasts³², simple stepwise removal of the region led to a substantial reduction of its net deaminase activity when measured without UGI. As the UGI-fusion form is so effective in yeast that it may mask the evaluation of the net activity of the deaminase, which may not be readily applicable in other organisms. We intentionally omit UGI for strict evaluation of the enzymatic activity in yeast, then add UGI in the following off-target assay and human cell applications. Based on the predicted structure, we deliberately truncated both N- and C- termini of PmCDA1 to cleanly eliminate the DNA-binding domain and to minimize the protein section. The amino acid substitutions to lessen hydrophobicity at the exposed surface further recovered the activity, probably due to improved protein stability or folding. The obtained tCDA1EQ version demonstrated comparable on-target activity and greatly reduced off-target effect, especially in the Target-AID3S inlaid form (Fig. 10, 15, 17). Furthermore, Target-AID3S showed a wider editing window and further less R-loop off-target effect than Target-AID2S. As Target-AID2S also outperformed pre-existing base editors in on/off-target profile (Fig. 17), its narrower editing window should be useful for precise on-target editing.

YE1 was initially developed to narrow editing window⁵¹ then revealed to have reduced off-target effects for both DNA and RNA⁵². While YE2 and R33A + K34A elicited further lowered off-target effects, they performed less robust on-target editing³⁹. In this study, YE1 showed relatively high off-target editing, especially at the R-loop site 6 and 7. This might be due to the motif preference of rAPOBEC1 to 5' TC motif^{22,53}, which was clearly observed at the both R-loop off-targets (Fig. 15). YE2 and R33A + K34A also showed the same trend but much less extent, probably due to weakened substrate binding capacity. In contrast to rAPOBEC1 and other APOBEC family proteins, PmCDA1 apparently did not show such strong motif preferences nor RNA editing^{25,54,55}. By exploring whether domain inlaid type will not only affect off-target occurrence in ABE^{35,56}, but also in rAPOBEC1 CBEs, we introduced inlaid type inlaid-YE1, inlaid-YE2, inlaid-R33A + K34A, and inlaid(1247)-YE1 for comparison. As expected, the inlaid type rAPOBEC1 CBEs showed significant off-target reduction,

although their overall on-target efficiency also decreased, possibly due to steric hinderance between Cas9 and rAPOBEC1 whose molecular size is larger with additional domains. Cas9-gRNA dependent DNA off-target effect was also shown to be significantly reduced in Target-AID2S and -3S (Fig. 14). Possibly, the DNA affinity provided by deaminase may cooperate with the off-target binding of Cas9-gRNA and subsequent editing.

Minimal off-target effects and robust on-target editing are expected to have a wide range of applications from plant and microbial breeding to clinical use. The Target-AID3S concept has also been demonstrated with the SaCas9 ortholog, providing the smallest base editing system with appreciable on-target efficiency while maintained the limit off-target effect. (Fig. 18, 19, 20, 21)

In vivo gene therapy requires the delivery of genome editing components to target tissues/cells. Nonviral nanoparticles allow for DNA transport capacities in excess of 10 kbp, but the cytotoxicity and specificity of synthetic polymers are major limitations in their in vivo use. Lentivirus vector has a cargo capacity of up to 8 kb and stable transduction activity, while their genome is randomly integrated into the host chromosome, which poses risks such as inactivation of tumor suppressor genes. Adenovirus vectors carry less risk of genomic insertion and allow for larger cargos, but their application has stalled due to their high immunogenicity. AAV vectors have shown appreciable transduction efficiency and long-term transgene expression in various tissues while maintaining a low level of immunogenicity, although its cargo capacity of the expression cassette is only 4–5 kb.

We have developed single AAV vector particle of SaAID2S that performed substantial editing efficiency (Fig. 22, 23, 24). Although it did not outperform the DUAL-533 dual infection at the current version, it is to be noted that the DUAL-533 is based on BE3.9 max that employs high-off-target rAPOBEC1. Single AAV with minimized composition is also advantageous for therapeutic applications because it reduces the preparation process and the number of protein domains that may be responsible for immunogenicity. DAS-739 failed to perform editing, possibly due to target dependence because DUAL-739 employs SaCas9 KKH variant which recognizes altered PAM sequence, or by a technical problem as DUAL-739 C-ter vector exhibited low virus titer. We also did not observe editing activity in A3S series although these are in the same range of size as SaAID2S ones. This may be due to unexpected ssDNA structures or the appearance of sequences that interfered with the function of the AAV vector, because mRNA sequence optimization was done by using full length Cas9 before PmCDA1 was inlaid. Optimization of the entire sequence to avoid secondary

structure and interference, or different order with gRNA cassettes may be tested in the future study. While further exploration is needed for in vivo study, we have demonstrated the ability of a single AAV CBE packaging and editing. We hope our study could help to further expand the gene editing platform.

5. Innovation strategy research

(English preface)

5-1. Summary of innovation strategy

Title of innovation strategy: “Expansion of BioPalette's business and exploration of new business opportunities through gene therapy for hepatitis B virus using the novel small Target-AID2S, 3S technology”.

Genome editing has been researched since the 1990s, and while it is not an entirely new concept, the emergence of CRISPR in 2012 had a significant global impact, making it one of the most widely recognized technologies in the history of genetic engineering. Furthermore, based on CRISPR, the base editing technology announced in 2016 has demonstrated significant potential applications in fields such as biology and medicine. Target-AID, one of the representative technologies of base editing, was developed by the Nishida team at Kobe University, and this technology served as the basis for the establishment of the company BioPalette. In May 2017, BioPalette entered into a licensing agreement with Kobe University, including the basic patent for base editing and related patents, and received the technology transfer from Kobe University.

In 2021, building on the Target-AID technology, the Nishida team at Kobe University developed a new and compact high-precision base editing technology, Target-AID2S, 3S (inventors: Li-author, Hikoyen, Nishida). This technology, as a related technology of Target-AID, automatically underwent technology transfer to BioPalette. Meanwhile, BioPalette, through a cross-licensing agreement with Beam Therapeutics had its business scope restricted.

Analyzing these strategies, including technology strategy, intellectual property strategy, business strategy, and financial strategy, as a series of strategic cycles, aims to demonstrate the superiority of Target-AID2S, 3S in the field of gene cell therapy (hereinafter referred to as "gene therapy") and lead to negotiations for a new cross-licensing agreement with Beam Therapeutics (hereinafter referred to as "Beam"). BioPalette will then negotiate a new cross-licensing agreement with Beam. Ultimately, the company (BioPalette) aims to establish a solid intellectual property foundation in the gene therapy field targeting the Asian region, in addition to the microbiome field where it has already aggregated the intellectual property needed for the commercialization of base editing.

【Intellectual Property Strategy】

By looking through the patent situation surrounding CRISPR-Cas9 technology and base editing technology, we analyzed the industry position and technological advantages of major startups established based on each technology, while introducing patents related to the novel technologies Target-AID2S and 3S. To begin with, we introduced the patent dispute on CRISPR-Cas9 technology and the basic patent on base editing. Then, the background of the establishment of the Kobe University venture company, BioPalette, Inc. and its basic technology, Target-AID technology, as well as the cross-licensing agreement with Beam, would be introduced to show the application and position of Target-AID technology in the overall genome editing field. After that, an overview of existing genome editing technologies in the medical field in general was reviewed, and the licensing status and issues of patents related to these technologies was analyzed. Finally, the new Target-AID 2S and 3S technologies and their patents, which are not covered by the cross-licensing between BioPalette and Beam, was explained in terms of their unique technical and intellectual property advantages in various application fields, especially in the gene therapy field.

【Business Strategy】

Focusing on the business strategy of BioPalette, we analyzed the business environment and the future potential of the technology it possesses, and indicated the possibility of restructuring the cross-licensing agreement with Beam. As part of the business environment analysis, we examined existing startups (in the gene therapy fields) and their development pipelines, analyzed market segments within the industry, and presented an overall picture of BioPalette's business strategy. Then, the safety issues in gene therapy, which have received the most attention, were addressed, and the usefulness of novel technologies in the next-generation gene therapy market (the importance of Target-AID2S, 3S) was analyzed. Furthermore, we analyzed the size of the future gene therapy market and the needs of society, and discussed the future potential of the new Target-AID2S, 3S. After this business analysis, environmental analysis, and discussion of the technological future, we presented ideas for restructuring the cross-licensing agreement with Beam.

【Technology Strategy and Financial Strategy】

The technology strategy and financial strategy for the restructuring of the cross-licensing agreement were introduced. Regarding the technology strategy, hepatitis B virus (HBV) selected as the target disease and a therapeutic approach using Target-AID2S, 3S would be designed. Then, we discussed the technological advantages of Target-AID2S, 3S technology compared to Beam's base-editing technology, while evaluating it in animal cells in the in-lab environment. If we are able to achieve positive results with its HBV pipeline, it will have an advantage of BioPalette in renegotiating the cross-licensing agreement with Beam, resulting in an expansion of its business domain in the microbiome field, as well as increasing its potential for business opportunities in the human therapy field. In addition, the roadmap for the program provides an overall research plan and milestones. On the other hand, as part of the financial strategy, we evaluated the experimental cycle and period, and estimate the initial in-lab development costs and the initial investment required. Subsequently, the expertise, stability, and characteristics of the contract research organization (CRO) were analyzed to select a suitable partner for animal studies and clinical trials. For the clinical cost study, we provided general standard values within the industry.

5. イノベーション・ストラテジー研究

タイトル：「新規小型 Target-AID2S,3S 技術を用いた B 型肝炎ウイルスの遺伝子治療によるバイオパレット社の事業拡大及び新規事業の探索」

5-1. イノベーション・ストラテジー概要

ゲノム編集は 1990 年代から研究が行われており、全く新しい概念というわけではないが、2012 年の CRISPR の登場は全世界に大きな影響を及ぼし、CRISPR は遺伝子工学の歴史上において最も広く注目された技術のひとつとなっている。さらに、これをベースとして、2016 年に発表された塩基編集技術は生物学や医学などの分野への応用に大きなポテンシャルを示している。塩基編集の代表的な技術の一つである Target-AID 技術は、神戸大学の西田チームが開発したもので、この技術を基にバイオパレット社を設立した。2017 年 5 月にバイオパレット社は神戸大学と塩基編集の基本特許を含む関連特許の実施許諾契約を締結し、神戸大学からバイオパレット社への技術移転を行った。また 2021 年、Target-AID 技術をベース、神戸大学の西田チームが小型かつ高精度な新規塩基編集技術 Target-AID2S, 3S が開発され（発明者：Li-筆者、光延、西田）、この技術も Target-AID 技術の関連技術として、自動的にバイオパレット社へ技術移転した。一方、バイオパレット社は Beam Therapeutics（以下「Beam 社」と記載する）とのクロスライセンス契約により、事業分野が制限された。なお、本文章のイノベーション戦略は、アカデミアの発明者の視点から、新たな Target-AID2S, 3S 技術を活用してバイオパレット社の戦略を立案することである。技術戦略、知財戦略、事業戦略、財務戦略の 4 つの戦略に関して、これらを一連の戦略サイクルとして分析することで、Target-AID2S, 3S の遺伝子細胞治療（以下「遺伝子治療」と記載する）分野での優位性を実証し、Beam 社との新たなクロスライセンス契約の交渉につなげる。最終的には、バイオパレット社が塩基編集の事業化のために必要となる知財を既に集約しているマイクロバリオーム分野に加えて、アジア地域を対象とする遺伝子治療分野の知財基盤を確立することを目指す。

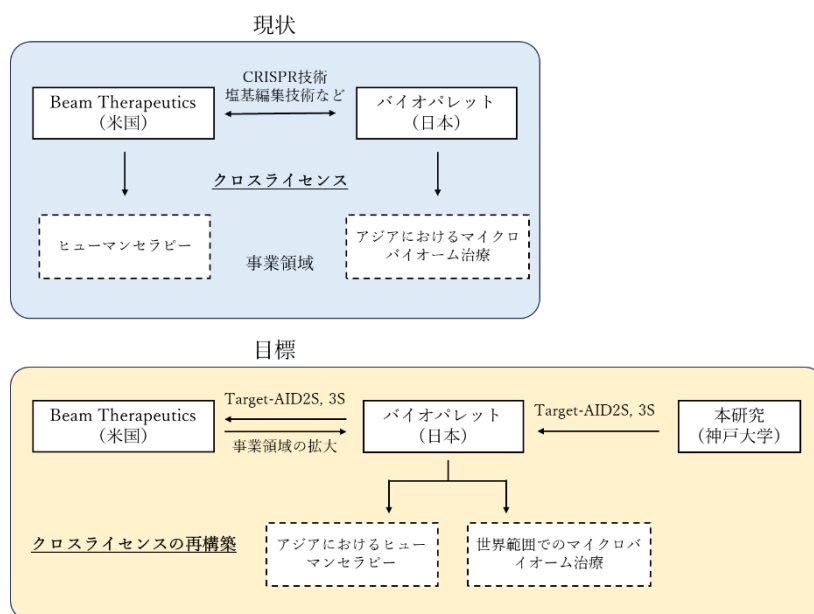


図1：イノベーション戦略のまとめ

【知財戦略】

CRISPR-Cas9 技術と塩基編集技術を巡る特許状況を通観することで、各技術をベースに設立されたスタートアップの業界での位置づけや技術的優位性を分析する一方で、新規技術 Target-AID2S, 3S に関する特許を紹介する。まず、CRISPR-Cas9 技術の特許紛争と塩基編集の基本特許を紹介する。その後、神戸大学発ベンチャーバイオパレット社の設立経緯と同社の基盤技術である Target-AID 技術、Beam 社とのクロスライセンス契約を紹介することで、ゲノム編集分野全体における Target-AID 技術の応用と立ち位置を紹介する。その後、医療分野全般における既存のゲノム編集技術を俯瞰して、それらの技術に関する特許のライセンス状況及び問題点を分析する。最後に、バイオパレット社と Beam 社のクロスライセンスでカバーされていない新規 Target-AID2S, 3S の技術及びその特許について、医療分野を中心とした様々な応用分野における独自の技術的・知財的メリットを説明し、業界内の技術的優位性を分析する。

【事業戦略】

バイオパレット社の事業戦略に着目し、事業環境及び、保有する技術の将来性を分析し、Beam 社とのクロスライセンス契約の再構築の可能性を示す。事業環境分析として、既存のスタートアップ（微生物・医療分野）とその開発パイプラインの調査を行い、業界内の市場セグメントを分析して、市場セグメントの選択を含むバイオパレット社の事業戦略の全体

像を示す。その後、最も注目されている遺伝子治療における安全性の問題を取り上げ、次世代遺伝子治療市場における新規技術の有用性（Target-AID2S, 3S の重要性）を分析する。さらに、今後の遺伝子治療の市場規模と社会ニーズの分析を行い、新型 Target-AID2S, 3S の将来性を考察する。このような事業分析、環境分析、技術将来性の考察を行った後、Beam 社とのクロスライセンス契約の再構築に関するアイデアを提示する。

【技術戦略及び財務戦略】

クロスライセンス契約の再構築に関する技術戦略と財務戦略を紹介する。技術戦略に関しては、B 型肝炎ウイルス（HBV）を標的疾患として選定し、Target-AID2S, 3S による治療手法を構築する。その後、インラボの動物細胞での評価をしつつ、Beam 社の塩基編集技術と比較し、Target-AID2S, 3S 技術の技術優位性を考察する。Beam 社も HBV のパイプラインを保有しているが、まだ初期段階である。バイオパレット社が HBV のパイプラインで良好な成果を上げることができれば、Beam 社とのクロスライセンス契約の再交渉が有利になり、結果としてマイクロバイオーム分野での事業領域の拡大と、ヒューマンセラピー分野での事業機会の獲得の可能性が高まる。また、研究のロードマップを示すことによって、全体的な研究計画とマイルストーンを提示する。一方、財務戦略としては、実験サイクル、周期を評価し、初期のインラボ開発コスト及び、必要な初期投資額を試算する。その後、臨床開発業務受託機関（CRO）の専門性、安定性、特徴を分析し、動物実験や臨床試験に適したパートナーを選定する。また、臨床コスト調査について、業界内の一般的な標準値を提示した。

5-2. ゲノム編集技術に関する知財の状況

< 5-2-1. CRISPR 基本特許における特許紛争 >

ゲノム編集の技術的インパクトと影響力の高さは、その研究や特許の数にも表れている。ゲノム編集に関する特許は世界で多く出願されていて、アメリカ、EU、中国の発表件数がトップ3になっている。それら特許の中には、ゲノム編集技術の事業化（社会実装）に必須となる基本特許と位置づけられるものも存在し、そうした基本特許を保有する数少ない組織が産業化の中心的役割を担っている（図2、図3）。

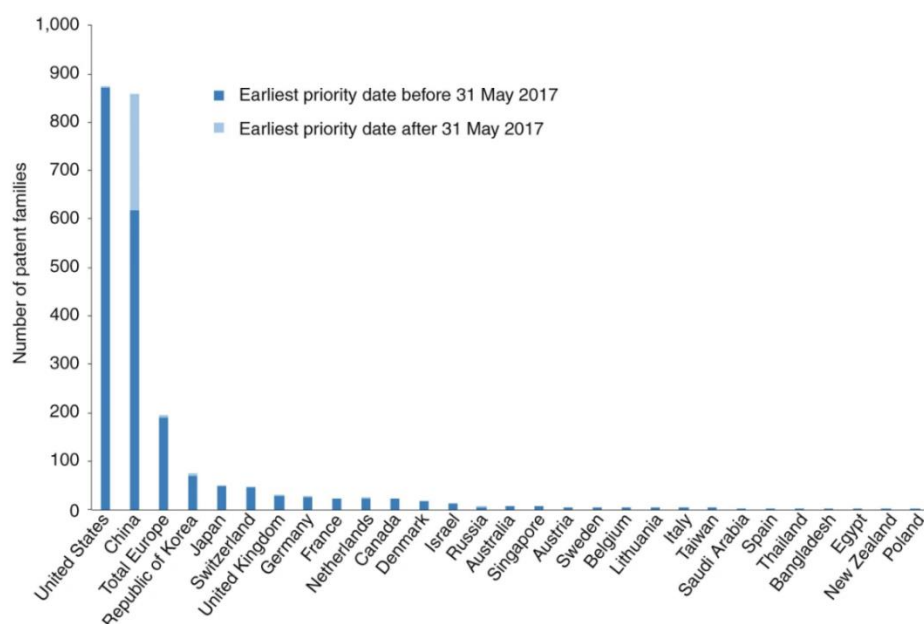


図 2：CRISPR 特許の出願件数. Ref: Worldwide CRISPR patent landscape shows strong geographical biases (2019)

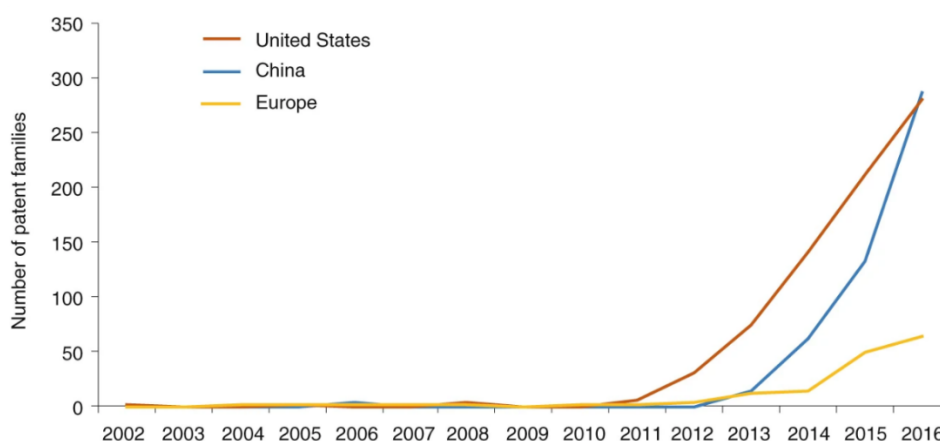


図 3：アメリカ、EU、中国の発表件数の増加. Ref: Worldwide CRISPR patent landscape shows strong geographical biases (2019)

ゲノム編集技術の市場価値の高さと社会的インパクトは、特許の争奪戦にも表れている。CRISPR に関しては、ヴィリニウス大学（リトアニア）、CVC（出願人である University of California, Berkeley、University of Vienna、Emmanuelle Charpentier の頭文字を取って、CVC と表記される）、ToolGen 社（韓国）、ブロード研究所（アメリカ）がそれぞれ出願している 4 つの特許が CRISPR 基本特許と言われている。

その中で、CRISPR 基本特許の係争は、特に CVC とブロード研究所の間で起こったものが最も注目された⁵⁷。UC Berkeley の生化学者である Jennifer Doudna と、現在マックス・プランク感染生物学研究所の Emmanuelle Charpentier が細菌由来の CRISPR が標的 DNA を切断できるという証拠を初めて発表したのは 2012 年 6 月のことで、Feng Zhang 率いるブロード研究所チームが CRISPR 技術が真核生物でも機能するという証拠を発表する 7 ヶ月前のことであった。しかし、CVC チームは最初の論文で、CRISPR が真核細胞内で機能することを示していなかった。アメリカ特許商標庁(USPTO)は 2014 年からブロード研究所に CRISPR 関連特許をいくつか発行し、2016 年に CVC チームの特許抵触の主張に基づく法廷闘争（インターフェアランス：同一の発明に対して、複数の出願がなされた場合に発明の先後を決定する手続き）を引き起こした。その結果、最初の特許審判控訴委員会（PTAB）裁判が行われ、真核生物の CRISPR 技術応用と CRISPR 技術の他の用途は、それぞれブロード研究所と CVC チームが特許を取得できる別の発明であると裁定し、複雑な評決が下されることとなった。これに不満だった CVC チームは連邦巡回控訴裁判所（CAFC）に控訴したが、CAFC は CVC チームの控訴を棄却した。その後、CVC チームは新たな主張を提出し、PTAB は 2 度目のインターフェアランス手続きの開始を宣言した。今回、PTAB は、CRISPR が真核細胞で機能するという最初の実証について、どちらのグループが最良の証拠を持っているかをより直接的に比較した。PTAB の決定は、先にこの一線を越えたという CVC チームの主張を認めず、ブロード研究所に優位な立場を与えた。その後、数年にわたる審理と連邦裁判所での数回の訴訟の末、USPTO は 2022 年 2 月、最終的にブロード研究所に特許を認める決定を下した。この決定は、CRISPR が商業的に使用された場合、CVC チームにとって巨額の損失となる可能性があることを意味する。また、すでに CVC チームから技術供与を受けている企業にも影響を与えた。

< 5-2-2. 塩基編集の基本特許 >

塩基編集においては 2 者が基本特許を有している。David Liu の所属するハーバード大学 (Base editor) と、西田敬二の所属する神戸大学 (Target-AID) である (図 4)。ハーバード大学の David Liu は 2013 年に Base editor 技術の特許を USPTO に出願し (US10167457B2)、2016 年にこの技術に関する論文を発表した。神戸大学側では、西田が Target-AID 技術について 2014 年 3 月に日本で特許出願し、1 年後の 2015 年 3 月に PCT (特許協力条約) 出願した。また、David Liu のグループと同じく 2016 年に Target-AID の技術を Science 誌で報告した。Target-AID 技術の基本特許のアメリカでの審査履歴によれば、3 年間にわたり頻繁な審査対応を行った結果として、2020 年 3 月に特許査定通知が出され、2020 年 5 月に特許が登録されている。また、欧州ではアメリカに先行して特許査定となっており、欧州内の 22 ヶ国において登録されている。その他、日本、中国、シンガポール、香港でも登録となっている (WO 2015133554)。

ゲノム編集の分野で基本特許を保有している機関は少数である。その中でも特に、ハーバード大学とマサチューセッツ工科大学を中核とするボストングループと、カリフォルニア大学グループ、この 2 つのアメリカグループがゲノム編集分野の勢力図の中心となっている。一方、神戸大学がこれら基本特許の内の一つを取得したことは、日本におけるゲノム編集の産業化の基盤形成にとって極めて重要である。

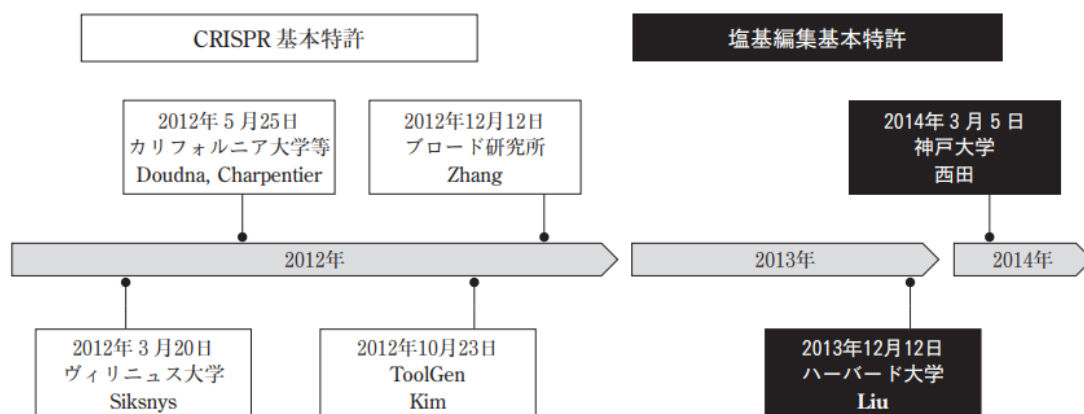


図 4：CRISPR 基本特許と塩基編集基本特許⁵⁸

<5-2-3. バイオパレット社の設立と知財の状況 >

2017年2月、西田が中心となって開発した塩基編集技術を技術シーズとして事業化を目指すバイオパレット社が設立された。バイオパレット社は株式会社科学技術アントレプレナーシップ (STE) による神戸大学発バイオベンチャー創業支援スキームで設立された初のバイオベンチャーとなり、バイオ専門のベンチャーキャピタリストで複数のバイオベンチャー企業の創業経験を持つ村瀬祥子氏が代表取締役に就任した。設立から約3ヶ月後の2017年5月、神戸大学との間で、塩基編集 Target-AID 技術に関する基本特許を含む複数の関連特許について実施許諾契約を締結し、神戸大学からバイオパレット社への技術移転が行われた⁵⁹。そして、契約内容により、神戸大学が保有している Target-AID の基本特許のみならず、将来開発される関連新技術も自動的にバイオパレット社に移転することとなる。技術移転とその内容は、すべての大学発ベンチャー企業の始まりであり、その後の成長を支える基盤として重要である。また2021年、筆者は発明者の一人として開発した Target-AID2S, 3S 技術もその後 Target-AID 技術の関連技術として、自動的にバイオパレット社に移転した⁵⁸。

バイオパレット社の重要な事業基盤となるのは神戸大学から独占実施許諾を受けた塩基編集の基本特許 (国際出願番号：PCT/JP2015/056436) である。この特許は2017年8月に日本で初めて成立した。その後、2017年10月にシンガポール、2019年6月に欧州、2019年12月に中国、2020年3月にアメリカ、2020年9月にカナダで権利化され、世界の主要国を着実にカバーしている (図5)。また、権利化されただけでなく、ほぼ全ての権利化国において分割出願が行われ、将来起こりうる事態に柔軟に対応できる強固な知的財産基盤を戦略的かつ総合的に構築している。

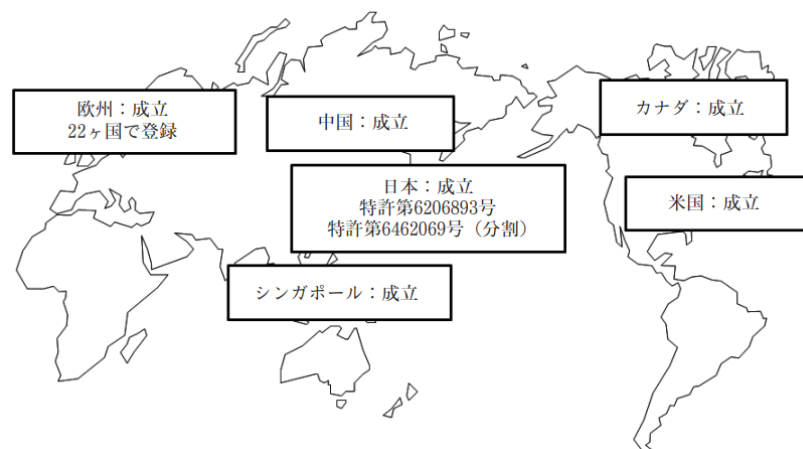


図5：塩基編集基本特許の成立国⁵⁸

< 5-2-4. 医療分野におけるゲノム編集スタートアップの概要と知財権利関係 >

ゲノム編集は、遺伝子の変異によって引き起こされる遺伝性疾患の治療、細胞ベースのがん免疫療法、他にも治療法が存在しない疾患や治療の満足度が低い疾患に対する革新的な治療法など、大いなる可能性を秘めている。その実用化は、医療分野が最も積極的に追求されている領域である。したがって、ゲノム編集技術に関連する知財を持つスタートアップ企業に注目し、それらの権利関係を分析することは、医療分野でのゲノム編集の実用性と将来性の理解に非常に役立つと考えられる。

第一世代ゲノム編集技術 ZFN の特許を有するのは Sangamo Therapeutics Inc (旧 Sangamo BioSciences Inc) である。当社は 1995 年にカリフォルニア州リッチモンドで設立された (図 6)。2018 年、Edward Rebar は Sangamo 社の副社長兼最高技術責任者に就任した。Sangamo 社は 7 回のラウンドで合計 5 億 2210 万ドルの資金を調達した。最新の資金調達は 2019 年 4 月にポスト IPO エクイティラウンドから行われ、1 億 4,550 万ドルを調達した。California Institute for Regenerative Medicine、National Institutes of Health (NIH) と Matthew Brown Companies が現在の投資家である。Sangamo 社は、ZFN の研究、開発、商業化に特化したバイオテクノロジー企業である。独自の遺伝子治療技術に基づき、ZFN 候補化合物を開発している。HIV/AIDS、血友病 A および B、ハーラー症候群 (MPS I)、ハンチントン病、ファブリー病、アルツハイマー病などの治療薬候補としてパイプラインを進めている。2019 年 2 月、Sangamo 社は、ハンター症候群の患者を対象に、DNA を永久的に改変する史上初の ZFN を用いた遺伝子治療の臨床試験を実施したことを発表した。現在、ハンター症候群以外、血友病 A の遺伝子治療もフェーズ III まで進めている。

第二世代ゲノム編集技術 TALEN の特許を有するのは Cellectis 社である。Cellectis 社は 1999 年、アンドレ・チュリカによって設立された。ゲノム工学におけるメガヌクレアーゼの使用をベースに事業を展開した (図 6)。Cellectis は 2007 年に株式を公開し、ユーロネクストでの株式公開で 2120 万ユーロを調達した。2010 年には新しいエレクトロポレーション技術を開発した Cyto Pulse 社を買収し、一年後、2011 年 1 月には、Iowa State University と University of Minnesota から TALEN ゲノム編集技術のライセンスを取得した。2014 年初頭の従業員数は約 300 名となっていた。しかし、CRISPR-Cas9 技術の出現により、Cellectis 社のメガヌクレアーゼ及び TALEN ゲノム編集技術は不経済となり、2013 年には倒産寸前まで追い込まれた。その後、同社は経営再建を図り、研究所を閉鎖して人員を削減し、研究の重点をがん免疫療法のための CAR-T 技術に移した。2014 年初頭、同社は Servier (製薬企

業)と大規模な資金調達契約を締結し、同年後半には Pfizer (製薬企業) および Allogene (製薬企業) とさらに大規模な契約を締結した。現在、CAR-T 技術をひとつの柱として研究開発を進める一方、TALEN 技術を用いた遺伝子治療薬の開発も進めている。

社名	Sangamo Therapeutics	Collectis
本社所在地	米国(カリフォルニア州リッチモンド)	フランス(パリ)
基盤技術	ZFN	TALEN
創業年	1995	1999
上場	2000年4月 NASDAQ	2015年4月 FRA
上場時の資金調達額	-	-
資金調達ラウンド数	7	7
資金調達総額	\$522.1M	\$248.2M
時価総額 (2023年 7月19日時点)	\$247.43M	111.44M EUR
従業員数 (2023年 7月19日時点)	478	234

図 6：ZFN、TALEN の医療応用を目指すバイオベンチャー

CRISPR 技術を用いて、医療応用を目指す海外の有力バイオベンチャー、5 社を紹介する。その中で、CRISPR 基本特許を有するバイオベンチャーは 3 社がある。Editas Medicine、CRISPR Therapeutics 及び、Intellia Therapeutics である。また、Beam Therapeutics は Editas Medicine とのライセンス契約により、塩基編集を対象とする CRISPR 技術の使用権限を取得した。2019 年 9 月、Prime Medicine は Beam Therapeutics とパートナーシップを結んで、契約に基づき、Beam Therapeutics は Prime Medicine に対し、ヒト疾患治療のための遺伝子編集薬を開発・商業化するために、Beam Therapeutics が管理する CRISPR 技術、送達技術およびその他の特定の技術に関する非独占的ライセンスを付与した (図 7、8)。

Editas Medicine (以下「Editas 社」と記載する) は当初、2013 年 9 月にブロード研究所の Feng Zhang、カリフォルニア大学バークレー校の Jennifer Doudna、ハーバード大学の George Church、David Liu、J. Keith Joung によって設立されたが、Doudna は 2014 年 6 月、Cas9 の知的財産に関する法的相違を理由に退職した。また、ベンチャーキャピタの Third Rock Ventures、Polaris Partners、Flagship Ventures からの資金調達を行った。Editas 社は、CRISPR ヌクレアーゼ (Cas9 と Cas12a) に関するライセンスを活用し、先天性血液疾患及びオンコロジー領域の遺伝子治療薬開発に取り組んでいる。

CRISPR Therapeutics は 2013 年、Emmanuelle Charpentier、Shaun Foy、Rodger Novak によって設立された。Charpentier はその後、2020 年のノーベル化学賞を Jennifer Doudna と共同受賞した。2015 年、CRISPR Therapeutics 社と Vertex Pharmaceuticals 社は、CRISPR-Cas9 を利用してヒト疾患の根本的なゲノム上の疾患に着目し、新たな治療法を開発することに焦点を当てた戦略的研究提携を結んだ。また、2016 年には Bayer と契約を締結し、Bayer との合併事業として Casebia Therapeutics の運営を開始した。CRISPR Therapeutics 社は先天性血液疾患、オンコロジー及び再生医療の分野で遺伝子治療薬の開発を行っている。

Intellia Therapeutics（以下「Intellia 社」と記載する）は 2014 年 5 月に設立され、Jennifer Doudna はそのサイエンス・ファウンダーの一人である。そして、資金調達に関しては、Intellia 社は Atlas Venture と Novartis 社（製薬企業）の出資を受けた。Novartis 社は CAR-T における CRISPR の応用に関心があったため、シリーズ A ラウンドに資金を提供しており、2015 年 1 月に Novartis 社と Intellia 社は、パートナーシップを締結し、Novartis 社が CAR-T プログラムに CRISPR を使用する権利を取得した。また、両社は CRISPR を用いて、血液疾患に対する遺伝子治療薬の開発に取り組んでいた。2017 年 4 月、Intellia 社は Regeneron 社（製薬企業）とパートナーシップを締結し、Regeneron 社は Intellia 社の CRISPR 技術プラットフォームを最大 10 の創薬ターゲットに使用する独占権を獲得した。

バイオペレット社と同じく塩基編集をプラットフォーム技術とする Beam 社は、ブロード研究所 David Liu 教授の塩基編集技術の医療応用を目指すバイオベンチャーとして 2017 年に設立された。2018 年 5 月、同社はシリーズ A ラウンドの資金調達として、アメリカの大手ベンチャーキャピタル 2 社から総額 8700 万ドルという多額の資金を調達した。ひとつはバイオペレット社の投資家でもあった F-Prime、もうひとつは Beam 社に投資した直後にバイオペレット社の投資家となった ARCH である。2020 年 2 月の IPO では 1 億 8000 万ドルを調達した。そして、2022 年 1 月、Beam 社は Pfizer とパートナーシップを締結し、塩基編集技術を用いた希少疾患の治療法開発に関するコラボレーションを発表した。現在、血液、免疫、HBV 及び、目疾患の遺伝子治療パイプラインを有し、開発を進めている。その中で、鎌状赤血球症及び、T-cell 編集治療が最も進んでおり、現在フェーズ I/II まで達した。

Prime Medicine（以下「Prime 社」と記載する）は同じくブロード研究所の David Liu 教授の研究室で開発された Prime editing ゲノム編集技術をベースとして立ち上げられた、遺伝子治療薬を目指すバイオベンチャーである。Prime 社は 12 人の投資家から出資を受けており、ARCH と F-prime、そして Google Venture が主要な出資者である。PE の最大の技術的課題は編集効率が低いことであり、生体内で直接編集治療（in vivo 治療）を行うことは困難であるため、編集効率のさらなる向上と最適化が必要である。そして、現在 Prime 社はマサチューセッツ州ケンブリッジに 3 拠点、マサチューセッツ州ウォータータウンに化学施設を構え、2024 年にはケンブリッジに新しい常設施設が建設される予定で、総面積は 15 万平方フィートとなる。一方、自動化、AI、RNA 技術に至るまで、企業全体の中核能力を中心に知的財産戦略を構築している。また、2022 年 10 月に IPO を成功させ、現在までに 5 億ドル以上を資金調達しており、この先、更なる急成長が予想される。

社名	Editas Medicine	CRISPR Therapeutics	Intellia Therapeutics
本社所在地	米国（マサチューセッツ州ケンブリッジ）	スイス	米国（マサチューセッツ州ケンブリッジ）
基盤技術	CRISPR	CRISPR	CRISPR
創業年	2013	2013	2014
上場	2016年 2 月 NASDAQ	2016年10月 NASDAQ	2016年 5 月 NASDAQ
上場時の資金調達額	\$94M	\$56M	\$108M
資金調達ラウンド数	6	5	5
資金調達総額	\$781.6M	\$127M	\$1.2B
時価総額 (2023年 7月19日時点)	\$710.90M	\$4.56B	\$3.89B
従業員数 (2023年 7月19日時点)	226	458	598

社名	Beam Therapeutics	Prime Medicine
本社所在地	米国（マサチューセッツ州ケンブリッジ）	米国（マサチューセッツ州ケンブリッジ）
基盤技術	塩基編集	Prime editing
創業年	2017	2019
上場	2020年 2 月 NASDAQ	2022年 10 月 NASDAQ
上場時の資金調達額	\$173M	\$175M
資金調達ラウンド数	4	2
資金調達総額	\$689M	\$315M
時価総額 (2023年 7月19日時点)	\$2.31B	\$1.22B
従業員数 (2023年 7月19日時点)	523	175

図 7：CRISPR、塩基編集の医療応用を目指すバイオベンチャー

ベンチャー企業にとって、知財係争や訴訟のリスクを減らすことは、成長戦略上、極めて重要な点である。先紹介したカリフォルニア大学 CVC チームとブロード研究所との間の CRISPR 基本特許をめぐる争いは、何年もかかり、多額の費用がかかっており、その特許技術に基づくバイオテクノロジー企業も紛争の影響を受けている。医療分野でブロード研究所の CRISPR 基本特許の独占実施権を獲得している Editas 社は、2015 年、2016 年に CRISPR 関連の特許係争のために発生した法的費用が、それぞれ 470 万ドルと 1090 万ドルであると公表した。

塩基編集技術の基本特許もハーバード大学と神戸大学がそれぞれ出願している 2 つの特許が存在する。CRISPR 基本特許の状況から、バイオパレット社が神戸大学の塩基編集技術を活用した事業開発を進めていくにあたり、ハーバード大学の塩基編集技術に関連する知的財産をどのように扱うか、その対処方針によって、事業展開に多く影響する。バイオパレット社としては、CRISPR 基本特許のような紛争を避けることが重要だと考えている。

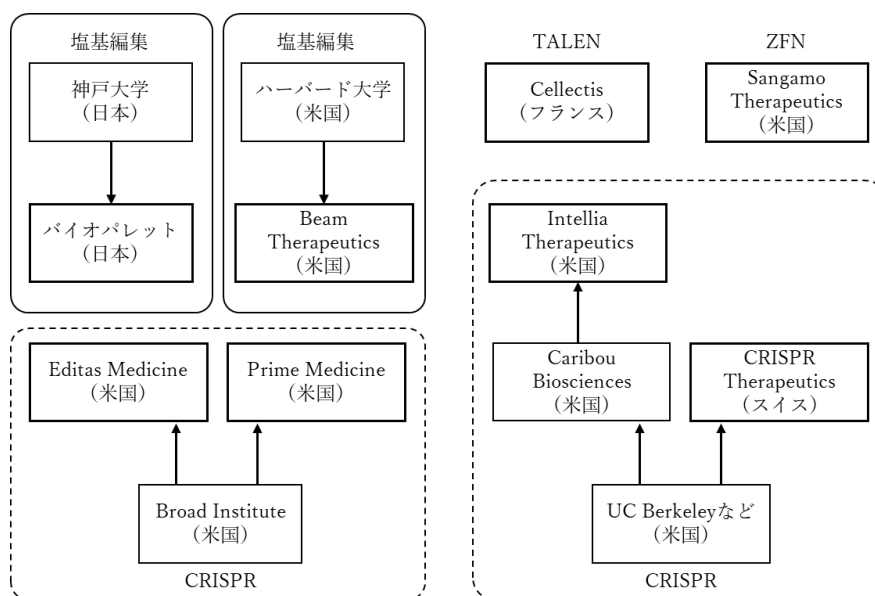


図 8：医療分野の主なプレーヤーと権利関係

<5-2-5. バイオパレット社と Beam 社のクロスライセンス >

投資家が共通している Beam 社とバイオパレット社は、お互いの間での紛争を避けるために最初から協力することを意図していた（図9）。Beam 社は、塩基編集技術を用いたヒト治療を開発するために設立されたベンチャー企業であり、ハーバード大学から実施許諾を受けているライセンスの範囲はヒト治療という事業分野に限定されており、医療以外の分野で事業を行うことは認められていなかった。一方、バイオパレット社が神戸大学から得たライセンスの範囲は特定の分野に限定されていなかった。故に、バイオパレット社の側からみた場合には、実質的には2つの選択肢があり、「医療分野全体について Beam 社へ独占的にライセンスアウトする一方向のライセンス契約」または「医療分野を 2つの領域に切り分けて、自らが選択しない領域については相手に独占的实施許諾を相互に与えるクロスライセンス契約」である⁵⁸。

結果としては、バイオパレット社は Beam 社とのライセンス形態を独占的クロスライセンス契約とする方針とした。クロスライセンスにより、バイオパレット社は Beam 社が保有する権利を含む、CRISPR 技術、塩基編集に関連する知的財産を統合し、選択した領域で安全に自社の事業を進めることができる強固な知的財産基盤を築くことができた。両社の契約交渉において最も重要な点は、それぞれの事業領域をどのように設定するかである。2019年3月に署名された独占的クロスライセンス契約において、バイオパレット社の事業領域は「マイクロバイーム治療（塩基編集による編集対象を微生物とするヒト治療）」と定義され、一方、Beam 社の事業領域は「マイクロバイーム治療を除くヒトの治療（塩基編集による編集対象をヒト細胞とするヒト治療）」と定められた。この契約は、知的財産権がアメリカの機関や企業によって制御され、日本企業の商業化が非常に難しい領域であるゲノム編集分野において、マイクロバイームというひとつの分野について事業化の基盤となる知的財産権を日本企業が独占的に確保するという画期的な契約であった。

バイオパレット社が「マイクロバイーム治療」を自社の事業領域として選択した理由は、CRISPR を用いた細菌編集の技術的な難しさにある。従来の CRISPR は DNA の二本鎖切断活性があるため微生物のゲノム改変では致死率が高いが、塩基編集技術は DNA の二本鎖切断を起こさず、細胞への毒性が低いことから微生物への応用が期待されている。故に、塩基編集技術はマイクロバイーム分野ではユニークなアプローチであり、バイオパレット社の独自性を発揮できる。

また、バイオパレット社にとって、Beam 社とのクロスライセンス契約の締結は、それぞれの事業領域の選定及び競争の回避ができ、さらには Beam 社の背後にあるブロード研究所とも良好な技術協力関係を築けるものであった。また、Beam 社は 2019 年 3 月にシリーズ B 資金調達ラウンドの完了を発表し、総額 1 億 3500 万米ドルを調達したが、これはバイオパレット社とのクロスライセンス契約締結とほぼ同時期であった。これらのことから、投資家の目から見て、バイオパレット社とのクロスライセンス契約締結は、Beam 社の今後の事業活動にとっても大きな意義があるものと推察される。

近年、ブロード研究所の基盤となる革新的な技術を活用して設立されたバイオテクノロジーのスタートアップ企業が増加していることは知られている。先述の Beam 社や Prime 社のような企業の他にも、2021 年にブロード研究所の Feng Zhang によって設立された Aera Therapeutics もその一例であり、遺伝子治療の新規デリバリーシステムに注力することによって注目を集めている。ブロード研究所は CRISPR 技術、DNA 編集、RNA 編集、デリバリーシステム、さらには AI 関連のバイオ情報技術まで幅広い分野で強力な開発チームと特許権を有している。したがって、今回のクロスライセンス契約の成立は、今後のバイオパレット社とブロード研究所の更なる協力の礎を築き、バイオパレット社の国際展開と技術の柔軟性を拡大する一助となると考えられる。

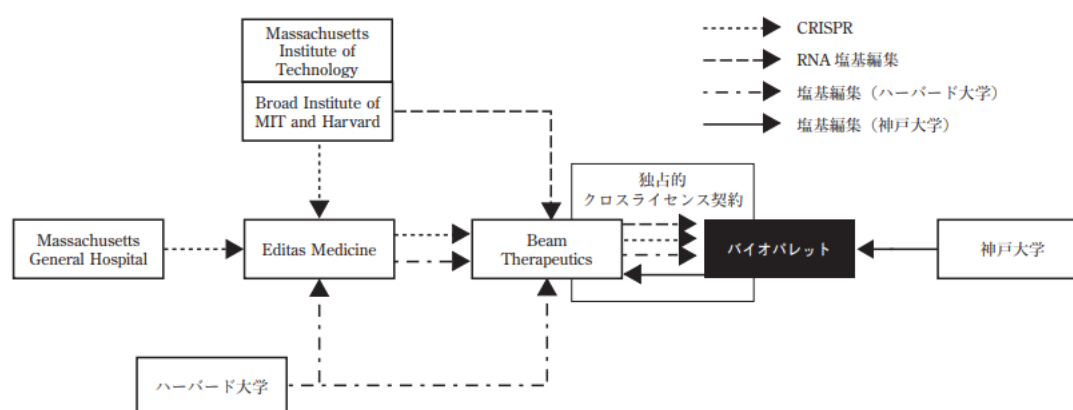


図 9：バイオパレット社と Beam 社の独占的クロスライセンス

< 5-2-6. 新たな特許出願：Target-AID2S, 3S 特許 >

近年、神戸大学は塩基編集の分野で新たなブレークスルーを起こし、Target-AID2S, 3S 技術が開発された。この技術は神戸大学が 2021 年 9 月に国際特許出願し、2022 年 3 月に国際公開された (PCT/JP2021/032689)。発明者は西田、李 (筆者)、光延で、いずれも塩基編集技術 Target-AID を発明した研究室のメンバーである。発明の具体的名称は「小型化シチジンデアミナーゼを含む二本鎖 DNA の改変用複合体」で、核酸改変モジュール、疎水性塩基置換、デアミナーゼの改変、二本鎖 DNA の改変など 30 項目が記載されている。さらに、この特許は 2022 年に哺乳類細胞で高精度の塩基編集実験に成功し、2022 年 8 月に Nature Communications 誌に掲載された。

Target-AID2S, 3S 技術は、ゲノム編集の精密性において大幅に改善でき、農水産物などの育種において、意図しない変異が生じる可能性が大幅に低減され、より安全安心な新品種の作出ができるようになり、気候変動に対応する作物や環境負荷の小さい作物、また健康効果の高い作物などが開発されると期待される。育種改良分野でのポテンシャルの高さに加え、この技術の最大の魅力は医療分野への応用である。先に紹介した CRISPR 関連のゲノム・塩基編集技術に比べ、Target-AID2S, 3S は編集の安全性が高く、副作用の発生率も低い。また、この新技術は分子サイズを大幅に制限することができ、シングルウイルス粒子に搭載することが可能となり、塩基編集技術の実用性が大幅に向上した (図 10)。したがって、この技術を保有する神戸大学やマイクロバイオーーム育種・治療パイプラインを持つバイオパレット社にとって応用価値が高いだけでなく、先に紹介した Beam 社のように医療分野を事業目標とするゲノム編集スタートアップにとっても魅力的な技術である (図 10)。

また、この技術の発明がバイオパレット社と Beam 社のクロスライセンス契約締結後であることも重要である。神戸大学との実施許諾契約にしたがって、この技術の発明と特許化により、バイオパレット社は新たな影響力のある知的財産を得た。そのことにより、今後の Beam 社や他のパートナー候補とのライセンス交渉を有利に進め、自社の持続的競争優位の向上を図る武器を手にしたと言える。

技術	Base editor (BE4)	Target-AID	Target-AID2S	Target-AID3S
所属	Beam Therapeutics	バイオパレット	バイオパレット	バイオパレット
編集効率 (8個標的による)	56.5%	52.8%	38.2%	42.2%
DNA Off-target 発生率 (ゲノムワイド)	> 15%	> 10%	< 2%	< 1%
RNA Off-target 発生率 (ゲノムワイド)	> 5%	-	-	-
安全性	低	低	高	高
分子サイズ	6490 bp	6161 bp	4525 bp	4525 bp
AAV導入手法	ダブルウイルス	ダブルウイルス	シングルウイルス	シングルウイルス

図 10：従来塩基編集技術と Target-AID2S, 3S 技術の比較

5-3. バイオパレット社の事業戦略：塩基編集の事業環境分析とクロスライセンス再構築

< 5-3-1. ヒューマンセラピー分野の事業環境分析 >

● 政治 (Politics)

技術の発展、特に CRISPR-Cas9 技術が医療分野で大きな魅力を示すにつれ、国際機関や政府は関連支援政策を立案・実施してきた。2021 年 7 月、WHO は「ヒトゲノム編集のためのガバナンスフレームワーク」を発表し、公衆衛生のためにヒトゲノム編集をより良いツールとすることを目的に、ヒューマンセラピーにおけるゲノム編集に関する勧告を提示した。WHO の新しい報告書は、登録、国際研究、医療渡航、技術に関する教育など、ヒトゲノム編集の 9 つの異なるシナリオに対する勧告を示すとともに、ヒトゲノムに関連する研究を実施、規制、監督するためのさまざまな手段を明確にするための新たな規制の枠組みを示しており、これが技術の体系的な改善につながり、すべての国において安全で効果的かつ倫理的なヒトゲノム編集を確保する一助となることが期待される。

日本では、遺伝子治療薬（再生医療等製品に分類）は、厚生労働省によって、他の医薬品とは異なる、特定の、独自の、迅速な承認の枠組みの下で規制されている（ファストトラック）。また、2013 年 11 月、再生医療安全法（RM 法）と医薬品医療機器等法（PMD 法）は、

日本における再生医療の開発と販売許可を加速させることを目的とした新たな法的枠組みを導入した。独立行政法人医薬品医療機器総合機構（PMDA）は、有効性が見込まれ、安全性が確認された予備的臨床試験に基づいて、条件付き期限付き製造販売承認を発行することができる。日本では、有効な治療法がなく、生命を脅かすような重篤な疾病を対象とする製品について、厚生労働省は革新的な製品への産業界の参入を促し、諸外国に先駆けて製造販売承認を取得することを促進するため、「SAKIGAKE」戦略を提案している。SAKIGAKE 指定された医薬品は、優先的な相談、審査期間の延長、プレミアム価格などの恩恵も受けることができる。

アメリカは常に革新的な医薬品にとって最高の環境を持っており、それに対して強い支援姿勢を持っている。アメリカでは、遺伝子治療はヒト細胞、組織、細胞・組織由来製品（HCT/P）に分類され、生物製剤評価研究センター（CBER）により医薬品および/または生物製剤として規制されている。FDA は 2016 年 3 月、REGROW 法において、日本のファストトラックの枠組みに類似した条件付き承認経路を提案した。REGROW 法は、安全性と有効性に関するより包括的でない証拠を受け入れるファストトラック承認経路を構築するため、第 III 相試験を廃止しようとするものであった。しかし、REGROW 法は、より低い製造販売承認基準によって、効果のない、潜在的に安全でない製品の市場参入を許すことになるとの懸念から、学界から反対された。しかし、アメリカ政府の革新的医薬品への支援は絶え間ない。同年 2016 年 12 月、FDA から公布した 21 世紀治療法において、重篤な疾患や生命を脅かす疾患に対するアンメット・メディカル・ニーズに対応する再生医療の開発と販売承認を奨励する目的で、再生医療先進医療（Regenerative Medicine Advanced Therapy：RMAT）指定と呼ばれる再生医療の新たな早期承認パスウェイが導入された。その後、FDA は 2022 年 3 月に、ヒトゲノム編集を伴うヒト遺伝子治療製品に関するガイダンス文書案と、CAR-T 細胞療法に関するガイダンス文書案を発表した。草案では、これらの治験用遺伝子治療製品の安全性と品質を評価するために、製品設計、製造、試験、前臨床安全性評価、臨床試験デザインなど、治験薬（IND）申請時に提供すべき情報が列挙されている。草案の定義によれば、Cas9、塩基編集、その他のゲノム編集ツールを用いた DNA の改変はすべてヒトゲノム編集とみなされる。これらの法案と関連規制の支援により、新薬の申請は迅速化され、コストを削減することができる。

- 経済 (Economy)

世界経済フォーラムの報告書によると、2022 年半ばの時点で、初期段階の研究から後期の臨床試験まで、世界全体で 2,000 以上の遺伝子治療が開発中であり、これは 2019 年の 2 倍である。世界の遺伝子治療市場は、2022 年の 53.3 億ドルから 2027 年には 198.8 億ドルに成長すると予想されている。また、Precedence Research 社によると、世界の CRISPR-Cas 遺伝子市場は 2032 年までに約 146 億 5000 万米ドルの規模に達する見込みで、2023 年から 2032 年までの間に約 19% の年間成長率 (CAGR) を記録すると予測されている (図 11)。

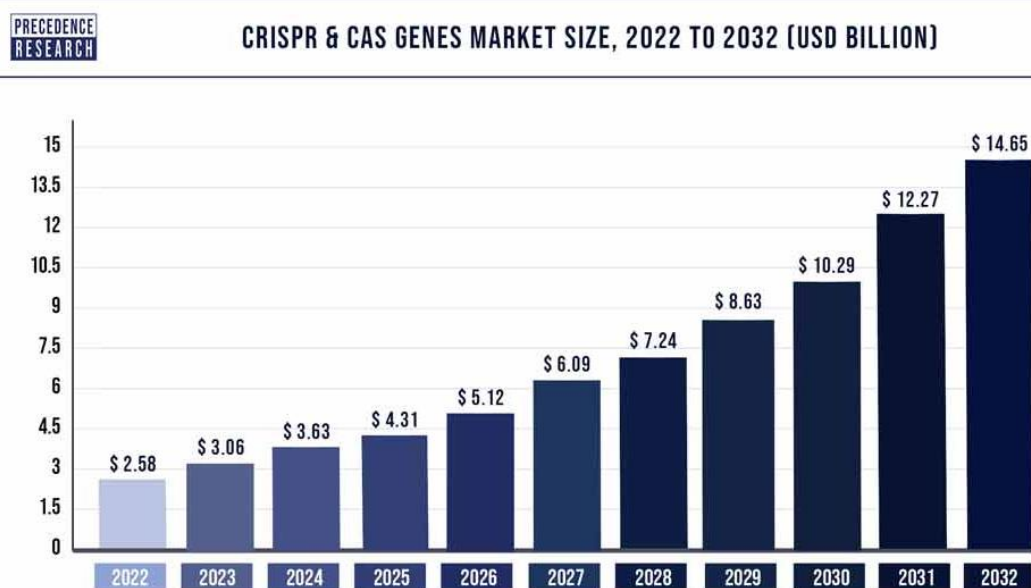


図 11：CRISPR-Cas 技術の市場と予測. Ref: CRISPR & Cas Genes Market, precedence research (2023)

Novartis、Roche、Eli Lilly、BMS などの大手製薬会社は、2017 年から 2019 年にかけて M&A 取引額が高く、M&A を通じて遺伝子細胞治療分野への参入を加速させている。現在、欧米の製薬会社トップ 10 の 90% が遺伝子治療薬を開発しており、FDA は、2025 年以降、毎年少なくとも 10 品目の遺伝子細胞治療薬が認可されると予測している。2030 年までには、60 以上の遺伝子治療薬が認可されるだろう。2021 年 11 月、Positive Sequence (Shanghai) Biotech Ltd. は約 3 億人民元の最新資金調達ラウンドを終了し、新規の塩基編集である tBE の開発と臨床応用のための製品パイプラインの開発、および特許のライセンス供与を加速させている。Arbor Biotechnologies 社は、肝臓と中枢神経系の遺伝性疾患の治療と次世代 CRISPR ゲノム編集ツールの開発プロジェクトを推進するため、2 億 1500 万ドルを調達した。

遺伝子治療はこれからの医療分野のテーマであり、その経済的価値は顕著である。特に CRISPR-Cas 技術は新たな医療革命をもたらしたとも言え、その市場価値の大きさから、経済環境は非常に整っていることがわかる。将来的には、遺伝子治療技術の進歩が医療資源の配分方法を変えるかもしれない。

● 社会 (Society)

近い未来、遺伝子治療は、嚢胞性線維症、 $\alpha 1$ 抗トリプシン欠損症、血友病、 β サラセミア、鎌状赤血球症などの特定の遺伝性疾患の予防、治療、治癒に用いられる可能性が高い。さらに、がんや HIV のような感染症など、既存の治療法では不十分な疾患に対して、新たな治療法を提供することもできる。このように、遺伝子治療は人類社会の健康にとって広範な意義を持っていることがわかる。しかし、遺伝子治療は新しい治療法であるため、その高コストが社会への普及の大きな障害となっている。

より多くの高価な遺伝子治療薬が市場に参入するにつれ、政府は予算負担を感じ始めている。特に韓国では、 ~ 0.2 百万米ドル (3 億ウォン) を超える高額医薬品を申請した患者数は、2016 年の 147 人から 2020 年には 410 人へと 3 倍に増加し、申請総額は同期間に 26 百万米ドル (370 億ウォン) から 102 百万米ドル (1470 億ウォン) へと 4 倍に増加している。これから、より多くの高価な医薬品が驚異的な成長率で市場に投入される見込みであり、それらの予算上の懸念を事前に処理することは重要な政策優先事項となるだろう。英国は、非常に希少な疾患に対する遺伝子治療法を評価する HSTP (Highly Specialized Technologies Program) を開始し、希少疾患の研究とイノベーションを促進する同時に、希少疾患に対して、より公平で公正な治療アクセスを確保している⁶⁰。また、この制度は CAR-T 細胞療法も対象としている。ドイツでは、診断関連群 (DRG) システムが運用されている遺伝子治療は、NUB1 ステータスを取得すれば、予算外の償還を受けることができる。イタリアでは、革新的医薬品として認定された遺伝子治療薬に対して、特別な基金を通じて償還給付が行われている。政府の支援により、遺伝子治療薬の社会的インパクトと実用範囲は大幅に拡大した。

● 技術 (Technology)

2010 年から 2022 年にかけて、遺伝子細胞治療薬の臨床試験総数は、年間成長率 (CAGR) 20~25% で世界中で着実に増加している (図 12)。まず、細胞治療の臨床試験の国別内訳を

見ると、2010 年代初頭には EU、英国、米国が世界の細胞治療臨床試験総数の 90%以上を実施し、支配的な地位を占めていた。しかし、同時期に年平均成長率 55%で中国が存在感を増し、最終的に 2022 年には世界のパイプラインの 21%に達するため、EU の優位性は弱まっている。同期間中、米国はその存在感を強め、2010 年にはパイプライン全体の 50%という圧倒的な地位を占めていたが、2022 年には 70%を超えるまでになった。韓国の割合は緩やかに増加しており、2010 年にはわずか 1.5%に過ぎなかったが、2022 年には 6.4%に達する。一方、遺伝子治療では、EU、英国、米国がパイプライン全体のかなりの部分を占める中、中国の台頭が目立つようになり、2022 年時点ではパイプライン全体の 17%にまで達している。国家間競争の中で、遺伝子細胞治療技術研究と臨床研究は飛躍的に拡大した。

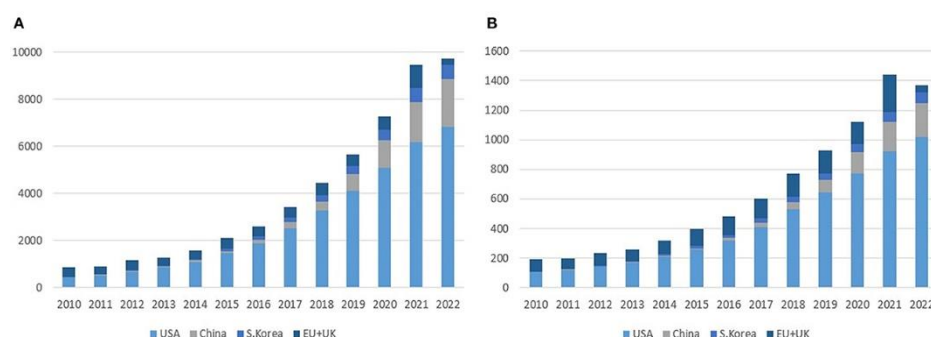


図 12：遺伝子細胞治療の臨床試験：地域別 (2010-2022) A：細胞治療、B：遺伝子治療. Ref: Cell and gene therapy regulatory, pricing, and reimbursement framework: With a focus on South Korea and the EU (2023)

近年、CRISPR-Cas9 技術をベースとした様々な派生・改良技術が開発されている。例えば、新しい超精密ゲノム編集ツールである PE、off-target の検出感度を向上させるために設計された GOTI 技術、真核生物の大分節遺伝子ノックダウンのための CRISPR-Cas3 ゲノム編集技術、複数の遺伝子座の同時編集のための Cas12a 酵素、編集効率を向上させる CAST ツールなどである。2022 年 3 月、オランダのデルフト工科大学の研究者らは、DNA の切断位置と off-target の確率を予測する新しい物理学ベースの Cas9 モデルを開発し、科学者たちがより精度の高い新たなゲノム編集ツールを開発するための基礎を築いた。

ゲノム編集技術に始まり、遺伝子導入技術、臨床応用技術に至るまで、遺伝子治療に関連する技術は日進月歩で革新が進んでいる。したがって、遺伝子治療の技術環境は非常に有利であり、近い将来、遺伝子治療が広く応用されることが予想できる。

PEST 分析を通じて、遺伝子治療によるヒューマンセラピーの外部事業環境は整っていると結論づけることができる。

<5-3-2. ヒューマンセラピー分野のパイプライン分析 >

Editas medicine

Editas 社は血液疾患及び、がんの遺伝子治療に着目している。特に、血液疾患の鎌状赤血球症 (SCD) とベータサラセミア (TDT) の創薬開発が進んでおり、このパイプラインは EDIT-301 と名付けられた。EDIT-301 はゲノム編集ヌクレアーゼである Ascas12a を用いて、ガンマグロビン遺伝子 1 および 2 のプロモーター領域を編集し、胎児ヘモグロビン (HbF) の発現を増加させることで、胎児ヘモグロビンが遺伝的に持続する自然なメカニズムを模倣し、SCD を治療する。EDIT-301 は ex vivo の手法を用いて、疾患細胞を編集したあと、患者に導入する。現在 4 名の患者に投薬され、臨床試験フェーズ I/II まで進んでいる。EDIT-301 は 4 人の患者において良好な忍容性を示し、ブスルファンによる骨髄破壊療法および自家造血幹細胞移植と同様の安全性プロファイルを示した。重篤な副作用は発生しなかったが、Cas12 に伴う off-target が発生することが確認された。また、Editas 社は、EDIT-301 以外に EDIT-101 というパイプラインを有している。EDIT-101 は、Leber congenital amaurosis 10 (LCA10、CEP290 関連網膜変性疾患) による失明の治療薬として開発中であり、米国の LCA10 患者約 1,500 人に影響を及ぼす IVS26 CEP290 変異対立遺伝子を修復するように設計されている。しかし、EDIT-101 の CRISPR-Cas9 による off-target が多く存在し、眼に関する重篤な副作用が観測され、臨床試験の評価基準を満たさなかった。故に Editas 社は EDIT-101 の開発は一時停止し、共同研究パートナーを探している。

Editas 社は、血液疾患と眼科疾患に最も投資し、注力してきたが、眼科疾患における臨床試験の結果、多くの副作用が確認され、開発を中止せざるを得なかった。これは、編集精度の向上と off-target 効果の低減が極めて重要であることを示している。一方、血液疾患を標的とする EDIT-301 はフェーズ I/II まで進捗しており、これからさらに患者導入数を増やして、検証していく方針である。また Editas 社は最近、抗がん剤開発に焦点を当てたパイプラインを他のバイオ医薬品企業と立ち上げた。

CRISPR Therapeutics

CRISPR Therapeutics 社は、血液、免疫、再生医療領域に対して開発中の医薬品をいくつか有している。その中で、Vertex Pharmaceuticals 社と共同開発中の希少な血液疾患であるベータサラセミアと鎌状赤血球症の治療薬がフェーズ III まで進んでいる。CRISPR Therapeutics 社は今年 4 月、重症鎌状赤血球症 (SCD) および輸血依存性ベータサラセミア (TDT) を対象とした治験薬 exagamglogene autotemcel (exa-cel) (旧名 CTX001™) の生物学的製剤承

認申請（BLA）の FDA への提出が完了したことを発表した。現在、進行中のフェーズ III 試験および長期追跡試験では、TDT または SCD 患者に対する治療の有効性と安全性を評価している。共有されたデータは、投与された 83 人の患者（TDT48 人、SCD35 人）すべての患者が臨床的有用性を示しており、治療法の有効性が証明された。また、安全性に関して、今までの報告では、2 人の TDT 患者が治療に関連すると考えられる重篤な副作用（Serious Adverse Event: SAE）を認めた。1 人の患者は、3 つの SAE（血球貪食性リンパ組織球症（HLH）、急性呼吸窮迫症候群および頭痛）が発生し、も 1 人は 1 つの SAE（特発性肺炎症候群）が観測された。

Editas 社と比較すると、CRISPR Therapeutics 社のパイプラインはより着実に進展している。当初の血液疾患パイプラインがフェーズ III に到達することに成功したほか、免疫・再生医療領域の開発も大きなブレイクスルーをもたらし、いずれも臨床段階に到達している。また、資金の調達や他のバイオ医薬品企業との連携も非常にスムーズに行っている。一方、臨床試験の急速な進行と対照的に、臨床データにおいては、依然として深刻な副作用（SAEs）が存在することが判明した。

Intellia Therapeutics

Intellia 社が現在保有するパイプラインも、設立当初とは比較的大きく変化している。Novartis 社は CAR-T 細胞療法における CRISPR の応用に興味を持っていたため、Intellia 社のシリーズ A ラウンドに資金を提供していた。2015 年 1 月、Novartis 社と Intellia 社は合意に達し、Novartis 社は 自社の CAR-T プログラムに CRISPR を使用する権利を取得し、両社は、ベータサラセミア（TDT）や鎌状赤血球症（SCD）を含む造血幹細胞が関与する疾患を対象として CRISPR を用いる遺伝子治療のパイプラインを立ち上げた。しかし、このプログラムは最後まで続かなかった。2022 年に発表された Intellia 社の決算によると、Novartis 社は Intellia 社の CRISPR を使用して開発した鎌状赤血球症（SCD）を対象とする ex vivo 治療プログラムを中止した。

2017 年 3 月、Intellia 社と Regeneron 社は、トランスサイレチンアミロイドーシス（ATTR）に対する CRISPR ベースの治療法を共同開発すると発表し、これが現在フェーズ I まで進んでいる（NTLA-2001）。2022 年 9 月 16 日、Intellia 社は、ATTR の単回投与治療薬として開発中の in vivo CRISPR-Cas9 ゲノム編集治療薬 NTLA-2001 の進行中のフェーズ I 臨床試験の良好な中間結果を発表した。

また、Intellia 社は、遺伝性血管性浮腫（HAE）に関するパイプラインを保有しており、現

在、フェーズ II に達している。安全性に関しては、NTLA-2002 は 3 つの投与量レベルすべてにおいて忍容性が良好であり、副作用の大部分は軽度であった。現在までに用量制限毒性、重篤な副作用、グレード 3 以上の副作用は観測していないが、軽度の副作用が頻繁に発生することが確認された。

Intellia 社は、ATTR および HAE パイプラインの臨床試験で良好なデータを達成しており、開発がスムーズに進捗していることを示している。また、臨床データにより、患者の症状はかなり緩和され、再発も抑制されたが、フェーズ II で頻繁に報告された副作用は依然として懸念事項である。

Beam Therapeutics

Beam 社も血液疾患において SCD と TDT を対象とするパイプラインを保有しており、それぞれ BEAM-101 と BEAM-102 と名付けている。特に、胎児ヘモグロビン(HbF)の発現をアップレギュレートする BEAM-101 をリードプログラムとして、開発に注力している。BEACON 試験は、重症 SCD の成人患者を対象に BEAM-101 の安全性と有効性を評価する非盲検、単群、多施設、フェーズ I/II の臨床試験である。BEACON 試験では、患者が登録されると、まず造血幹細胞を回収するための一連のプロセスが行われる。その後、細胞が編集され、自己 BEAM-101 製剤が作られる。2022 年末には BEACON 試験の最初の患者が登録された。造血幹細胞を用いた in vitro 実験では、BEAM-101 は 80% の HbS 修復率を示し、赤血球の形状も回復した。次の段階は、臨床で、BEAM-101 の効果を評価することである。また、Beam 社の報告でもうひとつ注目すべきは、ゲノム編集による DNA へのダメージである。そのデータから、塩基編集を用いることで、一般的な CRISPR-Cas9 よりも DNA のランダム突然変異、DNA の挿入、ゲノム全体への損傷が劇的に減少することがわかる。

また、Beam 社は、研究開発段階にある B 型肝炎ウイルス (HBV)、グレコ・ギン貯蔵病、スタグホーン病のパイプラインを有している。さらに、2022 年 Pfizer 社との提携により、未発表のパイプラインがいくつか加わった。中核技術である塩基編集は、従来の CRISPR-Cas9 よりも安全性が高く、DNA 損傷が少ないことが実証されている。しかし、デアミナーゼを加えるため、タンパク質全体のサイズが大きくなり、細胞内への導入が不便になるほか、デアミナーゼによる off-target 効果も副作用の可能性がある。この問題が解決されれば、塩基編集は理想的で安全な遺伝子治療のツールとなるだろう。

Prime Medicine

Prime 社は、Prime editing 技術の高い柔軟性により、多くの疾患領域をターゲットとし、豊富なパイプラインを有している。現在、Prime 社は 18 のパイプラインを有しており、そのうちの半分（9 プログラム）は「immediate」の戦略的カテゴリーに分類されている。血液疾患に関しては、鎌状赤血球症（Beam Therapeutics 社と提携）、慢性肉芽腫性疾患（GCD）、ファンコニー貧血などを対象としている。これら血液疾患を対象としたプログラムには生体外投与（ex vivo 治療）が適用されている。肝疾患領域では、ウィルソン病（肝臓の肝細胞における ATP7B H1069Q と R778L の変異をゲノム上の位置で修正する）およびグリコーゲン貯蔵病 1b を対象としており、どちらのプログラムも脂質ナノ粒子（LNP）デリバリーを応用している。眼疾患領域へのアプローチは、ロドプシン（網膜の桿体光受容体における RHOP23H 変異を本来のゲノム位置で修正する）およびアッシャー症候群に焦点を当てた網膜色素変性症プログラムを含み、いずれもアデノ随伴ウイルス（AAV）ベクターをデリバリー手段として採用している。耳疾患領域では、ギャップ結合タンパク質 $\beta 2$ （GJB2）遺伝子を標的とするアッシャー症候群 3 型や非症候群性難聴を含む。

スタート直後から 18 本のパイプラインをリリースした Prime 社は、成長の勢いがある企業と言える。さらに、同じブロード研究所の傘下である Beam 社と手を組むことで、技術基盤を拡大している。一方、Prime 社は体外だけでなく、AAV や LNP を用いた体内での遺伝子治療においてもパイプラインを有している。したがって、正確で理想的なゲノム編集を実現するためには、Prime editing 技術の効率をより高める必要がある。

総括

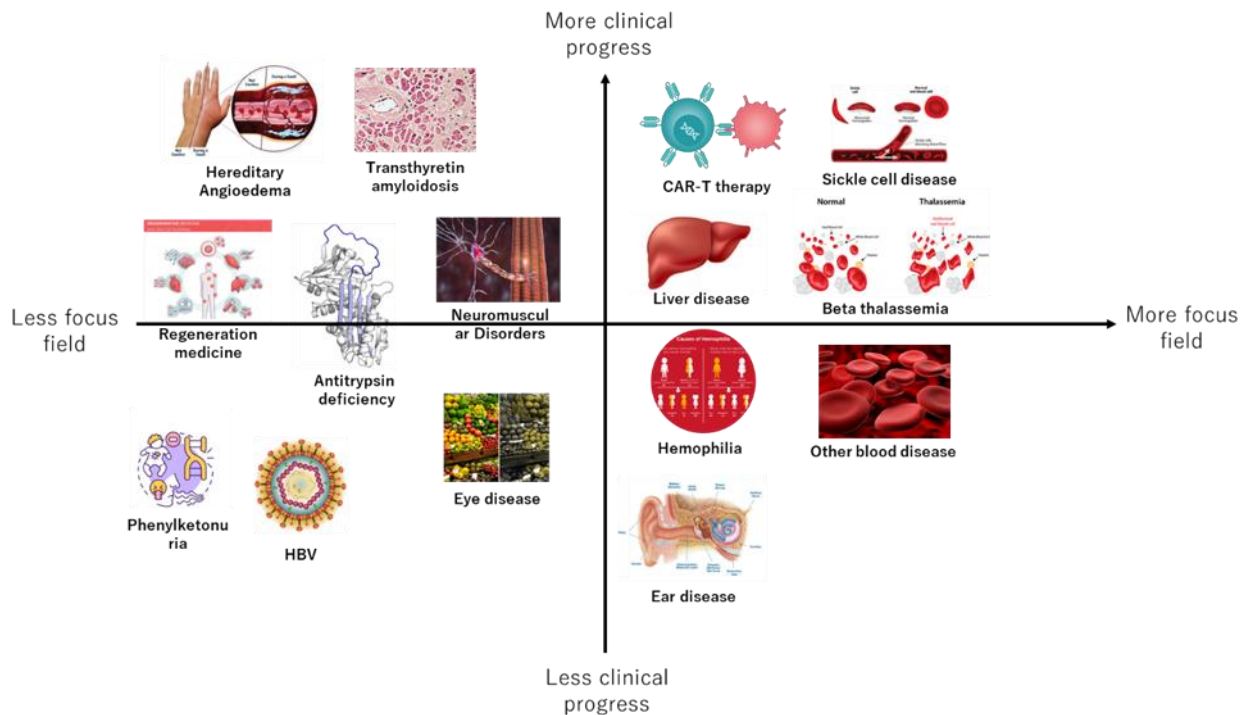


図 13：5 社のパイプラインのまとめ

多様なゲノム編集ツールが開発されたことで、遺伝子治療の応用範囲は飛躍的に拡大した。図 13 は、現在 CRISPR ベースの技術を使用している主要遺伝子治療企業 5 社をまとめたものである。このように、血液疾患が現在の開発の主軸であり、基本的に上記の 5 社すべてが関連パイプラインを有している。血液疾患は体外で編集できるため、人体を直接編集するよりもリスクははるかに低いと推測される。同様に、CAR-T 技術を向上させるためにゲノム編集を活用する場合も、体外でリスクをコントロールでき、短期的に実現しやすいと考えられる。ただし、目の疾患や耳の疾患、ウイルス感染症などの疾患は体外編集では治療できないため、開発の進捗は遅くなる。Editas 社の眼疾患パイプラインや Beam 社の耳疾患パイプラインなど、リスクを理由に中止されたパイプラインは多い。

遺伝子治療領域に参入したいバイオパレット社にとっては、体内編集の方向を選択したほうが有利だと考えられる。その理由としては、体内編集は幅広い用途に使用でき、実績のある体内編集技術は存在しないからである。しかし前述したように、体内編集は体外編集よりもはるかにリスクが高く、難しい。そこで、特許を取得した新しい Target-AID2S, 3S 技術

の重要性が前面に出てくる。この技術は現在、世界最小かつ最も正確な編集ツールであるため、体内での安全かつ効率的な編集に大きな期待が寄せられている。したがって、この技術が使えるバイオパレット社は、安全で効率的な体内編集パイプラインを確立することにおいて、大きな強みを持っている。

< 5-3-3. 遺伝子細胞治療薬における需要と市場の継続的な成長（予測） >

遺伝子細胞治療は、最初の治療が 2017 年に FDA に承認されたことで、近年、メディアの注目と投資家の資金を大きく集めている。2022 年までに FDA が承認した遺伝子細胞治療薬は 27 種類で、これはこれまでに承認された 340 種類の生物学的製剤（バイオ医薬品）の約 8%に相当する。現時点においては低分子医薬品も含めた医薬品全体のごく一部に過ぎないが、承認される真に革新的で新規な医薬品の構成を見ると、遺伝子細胞治療薬は承認総数の中でより大きな割合を占め続けている⁶¹。

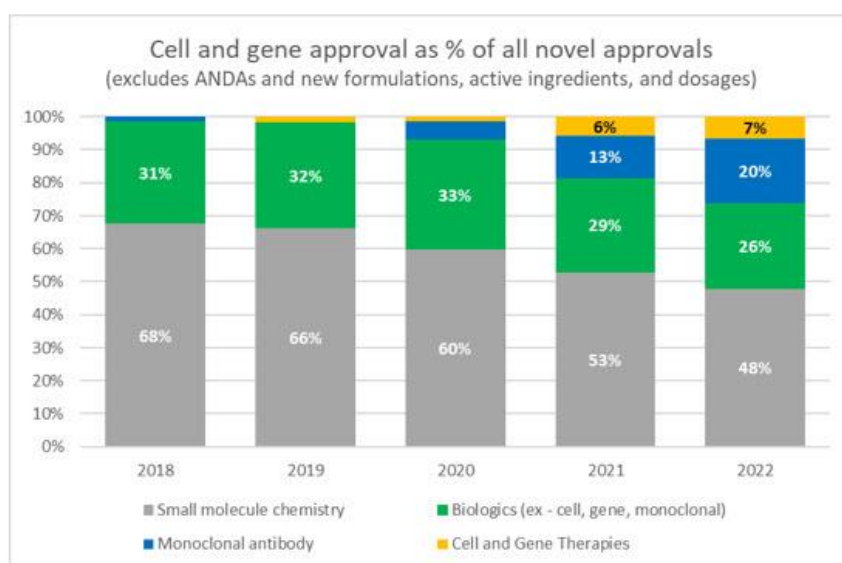


図 14：薬品種類の割合 2023's Market Outlook For Cell And Gene Therapies. Ref: 2023's Market Outlook For Cell And Gene Therapies（2023）

2022 年 12 月現在、進行中の遺伝子細胞治療の臨床試験のうち、大半は今後 6 年間で FDA への申請を予定しており、約 16%が第 3 相試験である。遺伝子細胞治療薬のパイプラインは依然として強固であり、新たな試験が開始される割合は増加している。

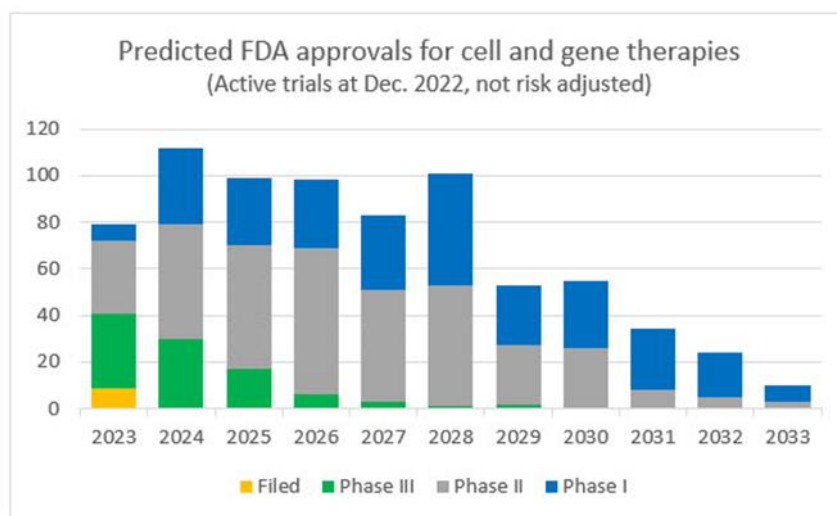


図 15：遺伝子細胞治療薬の臨床試験進行予測. Ref: 2023's Market Outlook For Cell And Gene Therapies (2023)

公衆衛生や希少疾患の治療に貢献するだけでなく、遺伝子細胞治療は次世代の医薬品の先駆けとなり、医薬品販売の原動力となるだろう。遺伝子細胞治療薬の 2022 年の世界売上高は医薬品全体の売上高の約 1% を占めるに過ぎないが、予測される成長率は驚異的である。従来型医薬品と非遺伝子細胞治療薬生物製剤の売上高は、2022 年から 2028 年にかけて CAGR (年平均成長率) 5% を達成すると予想されているが、遺伝子細胞治療薬は CAGR 46% を達成し、2028 年までに約 860 億ドルの売上高になると予測されている。

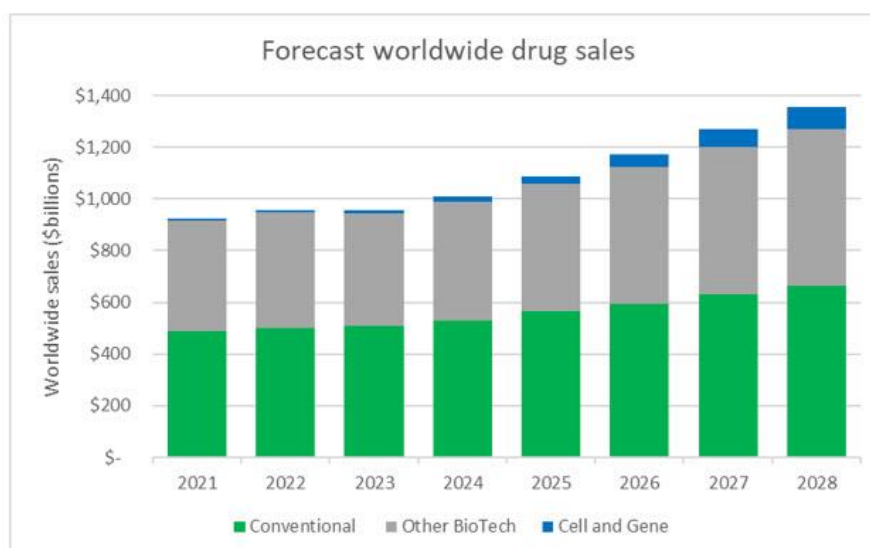


図 16：薬品別の売上予測. Ref: 2023's Market Outlook For Cell And Gene Therapies (2023)

細胞や遺伝子に投資される総額は、従来の医薬品や生物製剤に投入される金額のほんの一部に過ぎないが、このような資本の拡大がもたらす波及効果を認識することは重要である。投資家の関心はイノベーションを促進し、新興企業を奨励し、競合部門から経験豊富な人材を引き寄せる。その結果、研究室や投資ポートフォリオに成果がもたらされる。バイオテクノロジーの黎明期には、従来型医薬品間の競争が激化し、パイプラインや市場の関心が生物製剤へとシフトしていった。現在、遺伝子細胞治療がライフサイエンスのエコシステムにおいて独自のセクターとして台頭しており、より多くの人材、投資、パイプラインをさらにこれから惹きつけるだろう。

2022 年に資本市場で承認された医薬品のうち、高額な薬価となった上位 3 つが遺伝子治療薬であり、いずれも卸売価格は約 300 万ドルである。遺伝子細胞治療薬の性質そのものが、これらの医薬品の高価格を生み出している。これらの医薬品は開発コストが非常に高く、リスクが極めて高く、対象患者数が少なく、一生に一度の使用が必要な治療に取って代わろうとするものである。これらの治療法は医療分野の新時代を切り開いている同時に、保険償還制度や政府による価格設定に関する枠組みの整備も進められている。バリュー・ベース契約（患者が特定のエンドポイントを満たさなかった場合に治療費の大部分を払い戻す契約）への前払いや、医療保険やメーカーとのサブスクリプション契約を通じて、より多くの遺伝子細胞治療薬を患者が利用できるようにする取り組みが実施されている。

遺伝子細胞治療領域の臨床試験数はここ数年増加の一途をたどっているが、新規参入企業も増加している。さらに、公的・民間投資の割合が増加し、医薬品のライセンス契約数が過去最多となり、FDA の承認件数も急増した。遺伝子細胞治療薬にはまだ多くの課題があるものの、ヘルスケアとライフサイエンスの将来を切り開く上で不可欠かつ重要な役割を果たしていることは間違いない。

< 5-3-4. クロスライセンスの再構築 >

本文章を通して、遺伝子治療の安全性と off-target の問題について述べた。さらに、技術分析およびパイプライン分析を通して、新規 Target-AID2S, 3S 技術を活用することで、バイオパレット社は遺伝子治療の体内編集分野において、高い安全性を有する治療パイプラインの創出が期待できると判明した。したがって、具体的な標的を選択し、Target-AID2S, 3S 技術が安全かつ簡便な体内編集パイプラインの確立に有利であるデータを示し、Beam 社とのクロスライセンスの再構築を交渉するための戦略を立案する。新規 Target-AID2S, 3S 技術を Beam 社にライセンスアウトすることで、アジアにおけるバイオパレット社のヒューマンセラピー分野での事業化の権利、および世界におけるマイクロバイオーーム分野での事業化の権利を取得することに目標にしている。

Beam 社の興味と関心を喚起し、交渉を有利に進め、より多くの権限を得るためには、Target-AID2S, 3S 技術の生体内編集の利点を実証できる標的疾患の選択が重要となる。標的疾患の選定には2つの重要な要素があると考えられる。ひとつは標的の新規性、もうひとつは Beam 社に対する将来性である。

ゲノム編集による血液疾患や CAR-T 療法など、すでに多くの企業が開発し、臨床研究が進んでいる疾患を今回のターゲットに選定すると、標的遺伝子を設定するのは容易であるが、最終的には非常に不利になると考えられる。なぜなら、まずバイオパレット社において、現時点で遺伝子治療関連のパイプラインは存在しないため、ゼロからのスタートには多額の資金と時間的コストを要するためである。初期的な結果が出る頃には、同じ効果を持つ遺伝子治療薬がすでに市場に存在している可能性もあり、競争力そのものが大きく低下する。さらに、仮に Target-AID2S, 3S が既存の遺伝子治療薬よりも安全で効率的な結果を示したとしても、Beam 社は興味を示さないばかりか、逆に弾圧するだろう。なぜなら、成熟したパイプラインや市場から撤退する新薬の場合、多くの投資が無駄になり、同じ意味で Target-AID2S, 3S の技術を導入して開発し直すことはないからだ。したがって、Target-AID2S, 3S を新たな遺伝子治療ツールとして市場に登場させることは望んでおらず、クロスライセンスの再構築に合意することは非常に難しい。そのため、既成の成熟したパイプラインは避け、新興のパイプラインを選ぶべきである。

また、新興ターゲットの中でも、Beam 社の既存の初期パイプラインを選択し、異なる方法による安全な治療の可能性を示すことで、Target-AID2S, 3S の Beam 社への魅力が高まると思われる。なぜなら、同じターゲットを選択することは、Beam 社に一定のプレッシャー

を与えることができる。だが、まだ開発の初期段階であるため、より優れた開発パイプラインを築くために脅威となり得る新興技術を統合したいという発想が強くなり、自然と Target-AID2S, 3S 技術の重要性が Beam 社内で高まってくる。さらに、我々は異なる治療アプローチを取り、Beam 社の既存のパイプラインと組み合わせることで、開発の可能性がさらに拡大すると考えられる。このように、新規性が高く、Beam 社将来のパイプライン開発につながるターゲットを選択することで、Target-AID2S, 3S 技術に対する Beam 社の需要を最大化して、バイオパレット社がより多くの権限を獲得できる可能性を高めることができる。

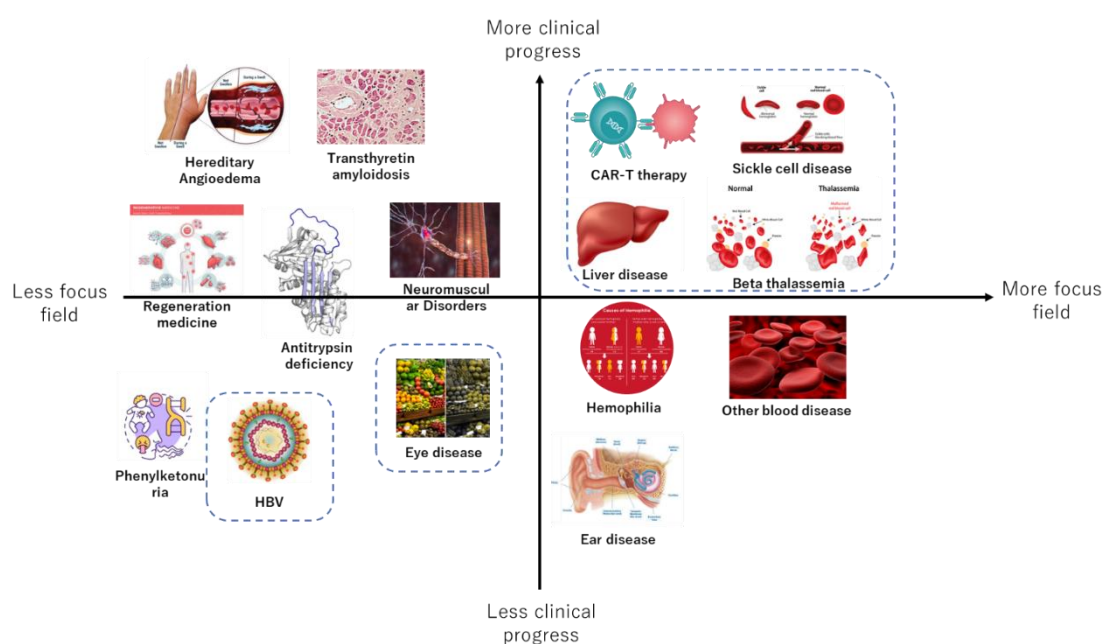


図 17：5 社のパイプラインのまとめ（Beam 社のパイプラインは青い線で囲まれる）

そのため、ターゲットとする疾患の選定に際しては、まず、図 17 の青い点線で囲んだ Beam 社の既存パイプラインに注目した。見ての通り、開発の初期段階にあるパイプラインは HBV と眼疾患（スターガルト病）である。その二つのパイプラインを詳しく分析すると、Beam 社は HBV に対するゲノム編集導入の手法として、AAV ではなく LNP を使用している。比較的成熟し、臨床実績のある AAV に比べ、LNP は新しい薬物デリバリーシステムであり、標的特異性の欠如、短い血液循環時間、生体内での不安定性など、技術的に多くの課題がまだ存在している。LNP が薬物デリバリーシステムとして初めて承認されたのは 2018 年であり、これに対して 2019 年には AAV が世界中で 250 以上の臨床試験に使用されている。しかし、Beam 社が AAV ではなく LNP を選択した重要な要因のひとつは、AAV には搭載する

遺伝子のサイズに厳しい制限があるためであると考えられる。今までに報告された BE の分子サイズでは 1 つの AAV に載せることができないため、2 つの AAV に分けて BE を導入する方法しかないが、これでは 2 倍の AAV 量が必要となり、細胞毒性も倍増し、安全性に関する懸念がある。一方、本事業で開発した新規 Target-AID2S, 3S 技術は小型に改良したため、1 つの AAV にも問題なく搭載できる。したがって、HBV 疾患については、Target-AID2S, 3S の技術を活用することで、安定、安全かつ実用性の高い新しい治療法を提供することができるため、今回の研究開発対象に最も適しているターゲットであると考ええる。

5-4. クロスライセンス再構築のための技術戦略と財務戦略

< 5-4-1. Target-AID2S, 3S×B 型肝炎ウイルス (HBV) >

HBV に対するワクチンは既に存在するが、それでも HBV は肝細胞がんの主な原因であり、毎年 80 万人以上の HBV 感染者が HBV や HDV によって引き起こされる肝がんや肝硬変で死亡していると推定されている。ペグインターフェロンによる治療はしばしば副作用を伴い、その有効性は HDV 患者では非常に限られている。ヌクレオシド系逆転写酵素阻害剤は、HBV に起因する慢性肝炎をコントロールすることはできるが、治癒させることはできないため、HBV を体内から排除する効果的な方法が臨床的に急務となっている⁶²。

本事業における具体的なアプローチとして 2 つの方法を考えている。ひとつは、Target-AID2S, 3S を用いて HBV 膜タンパク質を破壊することによって HBV ウイルスを不活性化することである。もうひとつは、細胞の NTCP 部位を編集することによって HBV ウイルス感染を防ぐことである。

HBV 膜タンパク質の破壊

HBV の外殻は、細胞由来の脂質、HBV 外膜タンパク質 (L、M、S)、HBV 表面抗原 (AGL) から構成されており、外膜タンパク質の変異により HBV 粒子の外層にある抗原を破壊することで、HBV ウイルスを不活化することができる (図 18)。そして、先行研究の調査と実験に基づき、外膜タンパク質の C147S 変異が HBV ウイルスを効果的に不活化できることを実証した (図 19)。これに加えて、我々は HBV 外膜タンパク質 L、M、S 上の他の編集部位を探索し、膜タンパク質を破壊する可能性のある部位をリストアップした。したがって、我々の第一の研究戦略は、Target-AID 2S および 3S 技術を用いてこれらの部位を編集し、HBV 外膜タンパク質を破壊することによって HBV ウイルスを不活性化することである (図 20)⁶³。

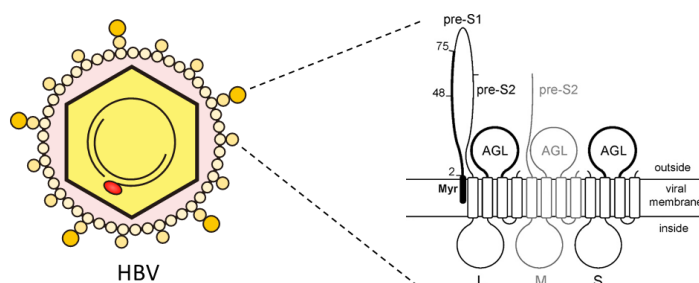


図 18：HBV ウイルスと外膜タンパク質. Ref: The Pre-S1 and Antigenic Loop Infectivity Determinants of the Hepatitis B Virus Envelope Proteins Are Functionally Independent (2009)

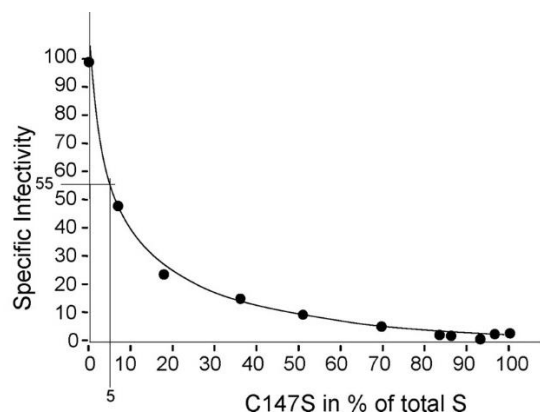


図 19：C147S 変異による HBV 感染力の低下. Ref: The Pre-S1 and Antigenic Loop Infectivity Determinants of the Hepatitis B Virus Envelope Proteins Are Functionally Independent (2009)

	Target gene	Sequence(5'-3')	Mutation	Genotype	Reason for selection
1	Pol	c cgagagctctagact c gtg g t	V391 to M or E or I/ V392 to M or E or I		Nonsense mutation in HBsAg (S-protein)
	S		W35 to Stop		
2	pol	c gctggatgtgtctc g gg c gtt	A434 to T/ A435 to T		Highly conserved regions
	S		R78 to Q or H/ R79 to H		
3	Pol	C ctatgggagtg g gc c ct c ag(T c c gtttctc T tg)	S524 to N/		Highly conserved region; 17 bp of guideRNA will accommodate all.
	S	C ctatgggagtg g gc c ct c ag(C c c gtttctc C tg)	V168 to I		
4	S-protein	c ctct T catcctgctg c tatg c c	C90 to Y/	Except genotype B	Mutated to cysteine
		c ctct G catcctgctg c tatg c c	C90 to Y/	Genotype B	
5	Pol	c atct c tt A ttggtctt c t g g	H448 to Y/ L449 to F/ L450 to F	Genotype A	Highly conserved regions
		c atct c tt G ttggtctt c t g g	H448 to Y/ L449 to F/ L450 to F	Except genotype A	
6	pol	c aaggtatgtgcccgtt g t			Highly conserved regions
	S		C107C to Y		
7	S-protein	(acc G) c agagctagactcgtg g tg g	Q30 to Stop/ P29 to L or S	Genotype A, D, E and G	Nonsense mutation in HBsAg (S-protein)
		(acc A) c agagctagactcgtg g tg g	Q30 to Stop/ P29 to L or S	Genotype B, C and F	
8	pol	t cctcttcac c tgctg c tat g	P440 to F or S or L/ L441 to F/ H442 to Y	Except type B	Highly conserved regions
	S		L84 to F		

図 20：選択したポテンシャルな標的

NTCP 遺伝子への点変異の導入

近年の研究では、タウロコール酸ナトリウム共輸送ポリペプチド (NTCP) レセプターが、HBV と HDV の両方の細胞内レセプターとして同定された。NTCP におけるホモ接合性の Arg252His 置換は、肝細胞への胆汁酸塩の取り込み障害と関連しており、この肝胆汁酸トランスポートの重要な役割が確認された⁶⁴。NTCP は染色体 14 上に位置する SLC10A1 遺伝子 (Solute Carrier family 10, member 1) にコードされている。NTCP は膜貫通タンパク質であり、細胞膜を介したナトリウムと胆汁酸の輸送に関与している。大きな HBV エンベロープタンパク質のプレ S1 ドメインの N 末端は NTCP に結合し、NTCP は肝細胞の基底側膜に主に発現し、HBV や HDV の肝細胞への侵入をサポートしている。NTCP 遺伝子のミスセンス rs2296651 変異体 (c.800C>T, S267F ; 267 位のセリンをフェニルアラニンに置換) は、膜受容体の構造を変化させることによって HBV 感染に影響を与え、その結果、肝細胞の HBV/HDV 感染感受性を低下させる可能性がある (図 21)。いくつかの研究により、S267F 変異体が HBV 感染性を低下させ、肝疾患の進行リスクを低下させることが示されている^{65,66}。したがって、我々の第二の研究戦略は、Target-AID2S, 3S 技術によって NTCP 遺伝子の S267F 編集を実現し、未感染者の HBV 感染予防を実現するとともに、既に HBV に感染している患者に対しては、S267F 編集によって病気の進行を抑制することである。

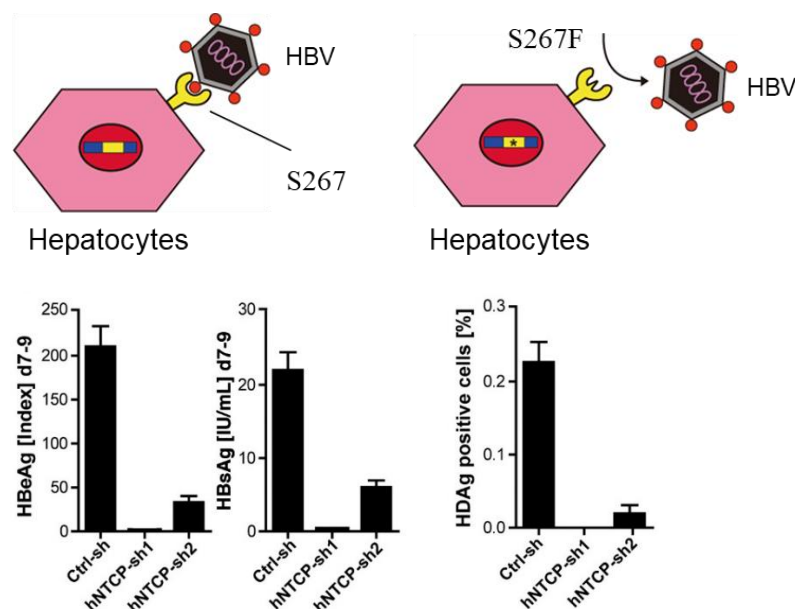


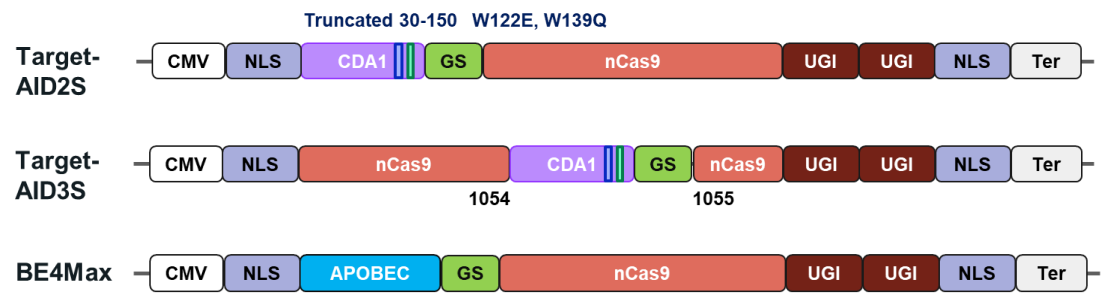
図 21 : NTCP ノックアウトによる HBV 抗原の発現量および感染細胞数の低減. Ref: Hepatitis B and D viruses exploit sodium taurocholate co-transporting polypeptide for species-specific entry into hepatocytes (2013) を基に筆者作成

< 5-4-2. Target-AID2S, 3S 技術による HBV ウイルスの失活および感染力低下実験 >

私たちはまず、Target-AID2S, 3S を用いて、HBV の 3 つの外膜遺伝子座と NTCP 遺伝子を編集のターゲットとして選択し、HEK293T 細胞での on-target の編集効率を評価した。また、Target-AID2S, 3S の比較対象として、ブロード研究所の David Liu グループの BE 編集技術 BE4 を使用した。

【実験設計】

エフェクターの構造

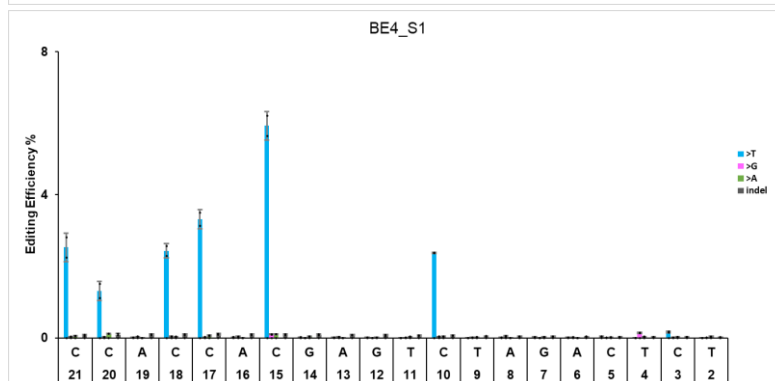
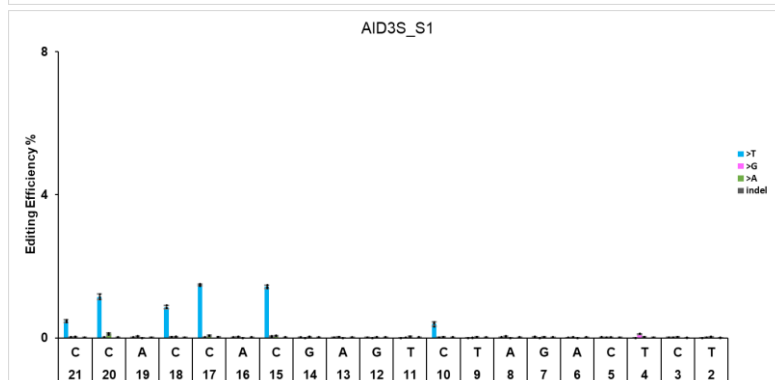
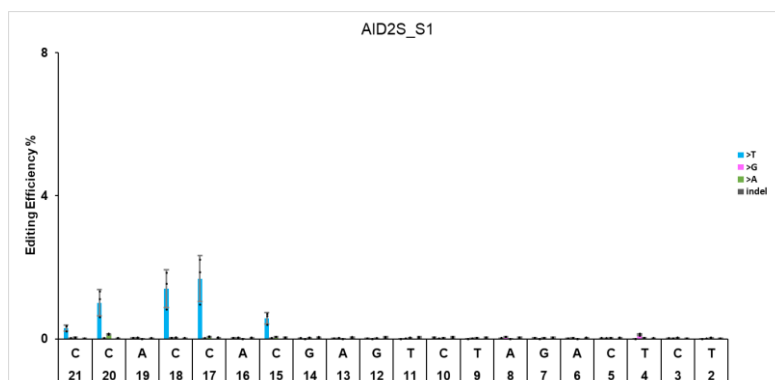


標的の配列

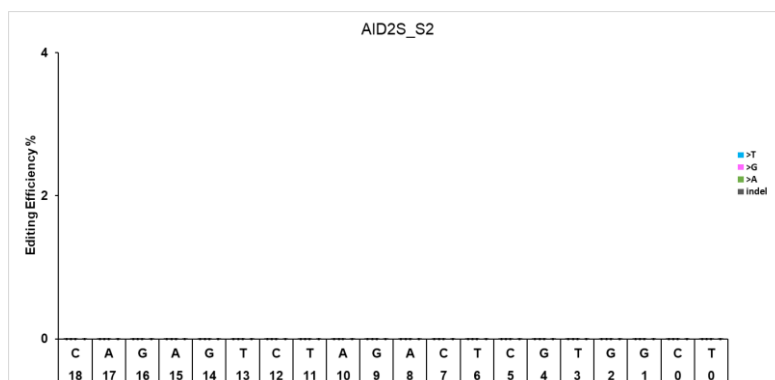
	Target	Target gene	Sequence(5'-3')	Mutation	Genotype	Reason for selection
S1	Membrane Proteins	Pol	CCGCAGAGTCTAGACTC	V391 to M or E or I/	All	Nonsense mutation in HBsAg (S-protein)
		S	GTGGT	V392 to M or E or I		
S2	Membrane Proteins	pol	CGCTGGATGTGTCTGCG	A434 to T/ A435 to T	All	Highly conserved regions
		S	GCGTT	R78 to Q or H/ R79 to H		
S3	Membrane Proteins	Pol	CCTATGGGAGTGGGCCT	S524 to N	All	Highly conserved region; 17 bp of guideRNA will
		S	CAG	V168 to I		
N	Genome	NTCP	CTGTTCCACCATCCTCA ATG	S267 to F	All	Prevent HBV binding

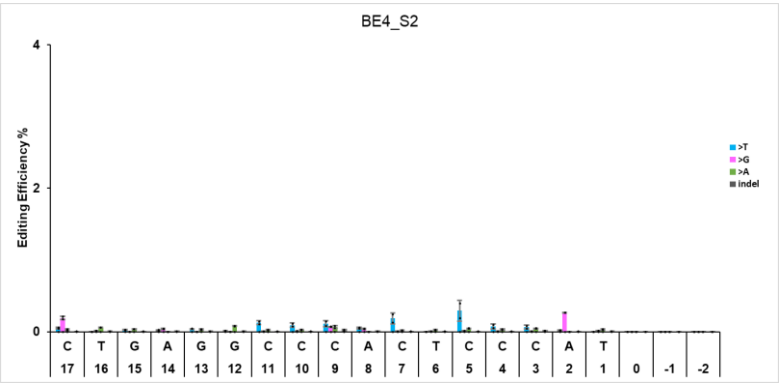
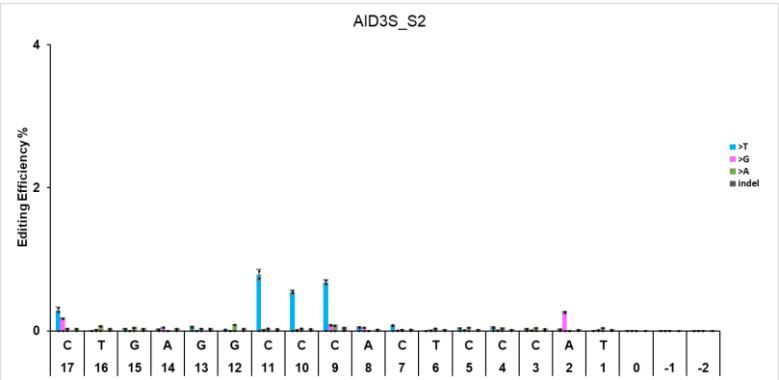
【実験結果】

HBV 膜タンパク質ターゲット：S1

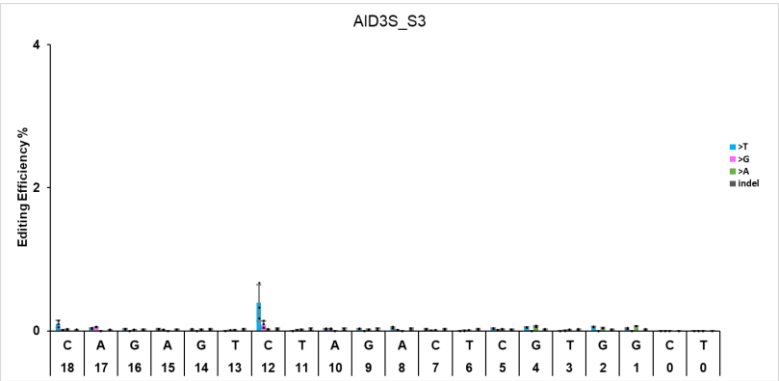
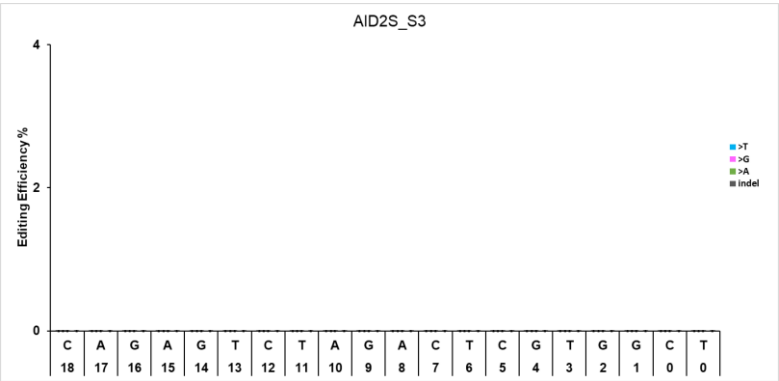


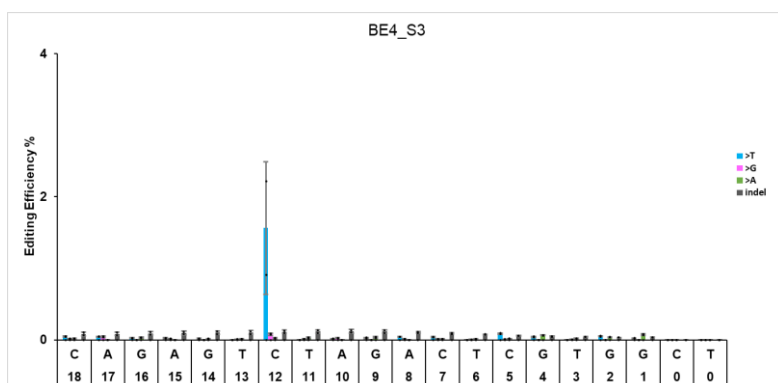
HBV 膜タンパク質ターゲット：S2





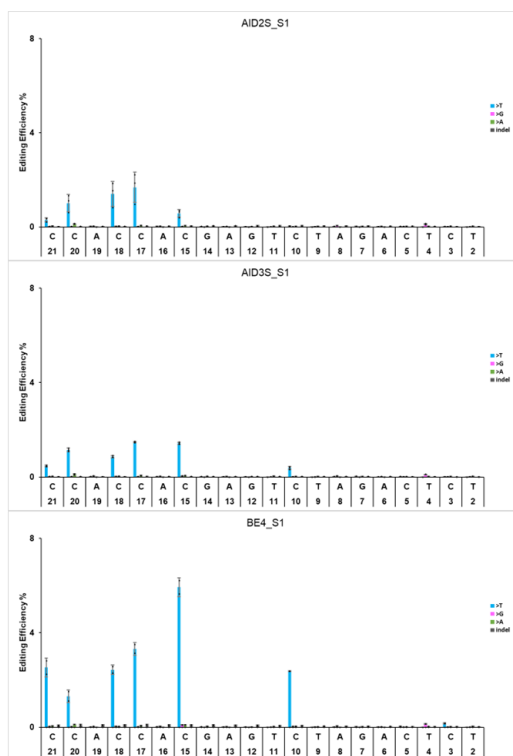
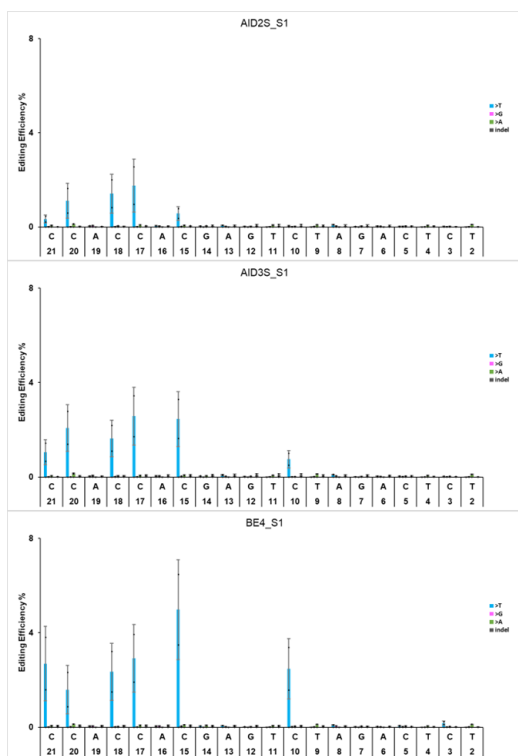
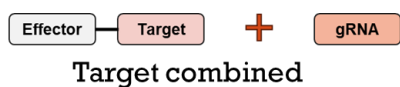
HBV 膜タンパク質ターゲット：S3





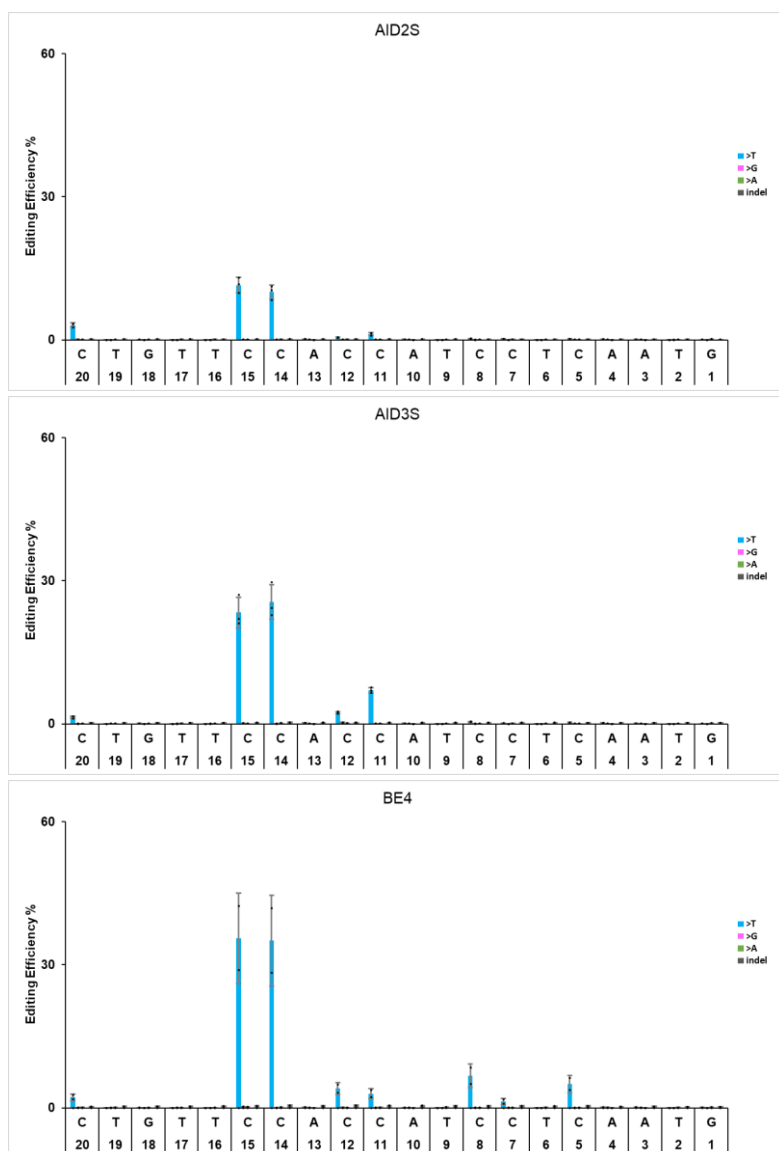
結果として、選択した3つのHBV膜タンパク質編集部位(S1、S2、S3)では、Target-AID2S、3SとBE4の全てが望ましい効率を示さず、最大編集効率はわずか6%しかなく、Target-AID2SはS2とS3では編集効率を示さなかった。その原因として、HBV膜タンパク質とゲノム編集を発現するプラスミド両方が同時に核に届く必要があるため、編集効率が低くなったと推測される。

したがって次に、標的が核に到達する確率を高めるために、HBV膜タンパク質の遺伝子情報をまとめてエフェクターに載せ、新たな試験を行った。



しかし、結果は芳しくなく、編集効率の向上は観測できなかった。また、3つのターゲットしか実験対象としていないため、これらのターゲットの編集効率自体が低い可能性もある。故に、今後はHBVの他の膜タンパク質ターゲットの選択実験を行うと同時に、新たな評価系の考察と設計を行う予定である。

人ゲノム NTCP ターゲット



ヒト NTCP 遺伝子座については、理想的な編集効率が観察された。Target-AID2S の編集効率は 15%と比較的低いが、Target-AID3S と BE4 ではほぼ同レベルの編集効率を得られ、それぞれ 30%と 35%となっている。今後、NTCP を標的とした AAV 粒子を作製し、同様の編集効率を達成できるかどうかを評価する予定である。

< 5-4-3. 研究開発のロードマップ >

今後も編集効率の最適化を進め、2024 年 5 月には実用範囲に達することを目標としている。また、AAV ウイルスベクターの実験と最適化も同時に行う予定である。その後、ウイルス毒性を評価するために、実際の HBV ウイルス粒子を導入し、細胞毒性アッセイを行う予定である。しかし、当研究室では、ウイルスを扱うためには施設面での課題があるため、神戸大学医学部との共同試験を検討している。

一方では、医療関連企業の調査を行い、適切なパートナーを選定し、共同で in vivo 動物実験を行い、最適化する。2024 年の 8 月までに、実験データの回収を完了させる。その後、解析データや資料を準備し、Beam 社との契約交渉に入る。

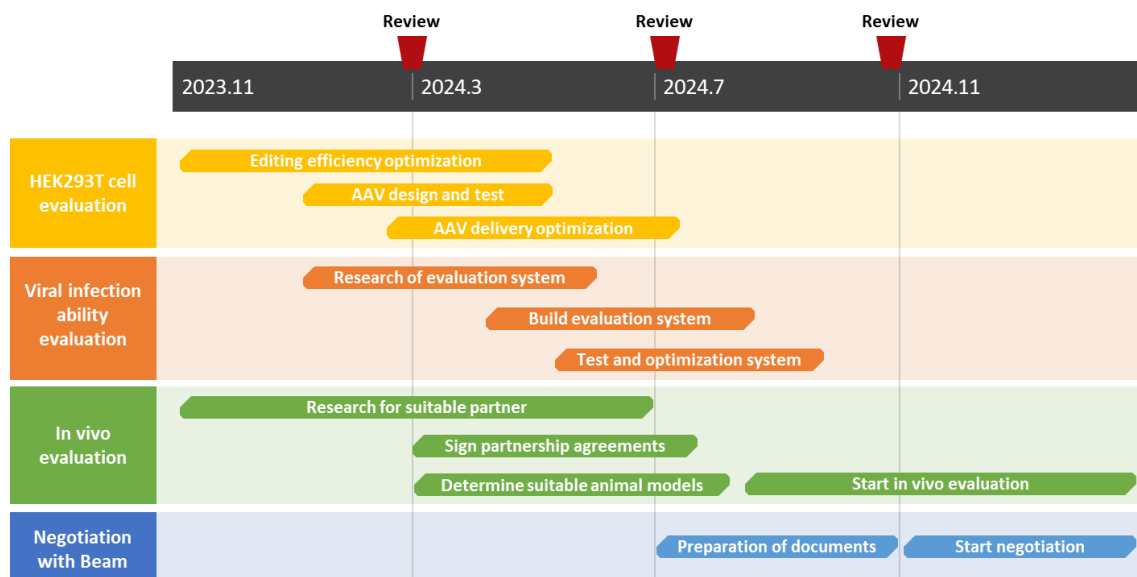


図 22：研究開発ロードマップ

<5-4-4. インラボ研究開発コストの試算 >

実験に関する消耗品をリストアップし、インラボ段階で必要な資金を試算した（2023 年 1 月～2024 年 7 月）。合計 3,370,048 円 のインラボ開発コストが必要となる。

品名	メーカー	コード	容量	数量	定価 (円)	総額 (円)
1.2mL 内ネジキ ャップクライオジ ェニックバイアル	Corning®	430487	50 / パ ック	1	38,400	38,400
アズノールディス ポコーンラージ棒	AS ONE	1-3192-01		1	8,300	8,300
プロワイプ	エリエール	S220	220 枚	1	4,968	4,968
JM109	TAKARA	9052	100 ul x 10	5	21,000	105,000
Sapphire リロード チップ 低残留 詰 換え用ユニバーサ ル 10 μ l ナチュラ ル	グライナー	771259	96 $\times$ 5 ユ ニット	5	21,120	105,600
1200 μ L エクスト ラロング 交換ブ レート	WATOSON	122-805CS	96 $\times$ 10 プレート	1	10,500	10,500
1000uL チップ(詰 め替え)大箱	WATOSON	123R- 757CS	96 本 $\times$ 5 段 $\times$ 5 パ ック	2	54,000	108,000
200uL チップ(詰 め替え)大箱	WATOSON	123R- 755CS	96 本 $\times$ 5 段 $\times$ 10 パック	2	75,000	150,000
C-Slide ディスポ ーザブル細胞計数 盤 (4 チャンネル・ 個別包装)	ワケンビーテ ック	CRCSD- NI50	50 枚	2	15,000	30,000

24 well plate	IWAKI	4820-010	20 枚	2	19,800	39,600
48 well plate	Corning®	356509	50 枚 x5	1	53,500	53,500
SensiCare ニトリ ルゴム手袋 S	MEDLINE	SEM486801	250 枚入	2	1,380	2,760
NucleoSpin® Tissue	Takara	U0952B	250 回入	1	74,200	74,200
NucleoSpin® Plasmid Transfection- grade	Takara	U0490C	250 回入	2	37,800	75,600
FastGene PlasmidMini	日本ジェネテ イクス	FG-90502	300preps	2	25,000	50,000
FastGene Gel/PCR Extraction Kit	日本ジェネテ イクス	FG-91302	300preps	2	25,000	50,000
シャーレ CL	sansei	01-002	12 枚 x50 袋	2	8,400	16,800
LB 培地 Lennox	ナカライテス ク	20066-95	500g	2	9,000	18,000
アガロース(低電気 浸透, 高ゲル強度)	ナカライテス ク	01163-05	500g	1	68,950	68,950
カネカ 簡易 DNA 抽出キット	カネカ	KN- T110005	50 tests	5	8,980	44,900
DMEM 培地	Thermo Fisher	10566024	500 mLx10	1	22,200	22,200
Opti-MEM 培地	Thermo Fisher	31985070	500 mL	2	4,000	8,000
1.5mL マイクロチ ューブ	WATSON	131-815C	1,000 本	10	3,700	37,000
サンプリングチュ ーブ 2.0ml ナチ	Thermo Scientific	508-GRD- Q	500 本 (500 本	10	1,620	16,200

ユラル			x1 パック)			
RiboMAX™ Large Scale RNA Production Systems	Promega	P1300		1	42,000	42,000
Ribo m7G Cap Analog	Promega	P1712		1	51,000	51,000
フェノール・クロロホルム・イソアミルアルコール	Merck	77619-100ML		1	13,600	13,600
エタノール	ナカライ	14713-53	3L	1	10,200	10,200
オートクレーブ廃棄バッグ L	アズワン	1-1471-03	100 枚	1	7,000	7,000
50mL チューブ	HARMONY	CFT5000	50 本 x10 袋	2	12,500	25,000
15 ml チューブ	HARMONY	CFT1500	50 本 x10 袋	2	9,000	18,000
アイソフロー	ベックマンコールター	8599600	20L	4	6,000	24,000
消毒用エタノール MIX カネイチ 5L	兼一薬品工業	8-1928-02	5Lx3 本	2	6,000	12,000
トリス - 酢酸 - EDTA 緩衝原液 50 倍 濃 縮 Nuclease and Protease tested (TAE)	ナカライテスク	32666-81	1L	1	12,000	12,000
PCR foil seal	4titude	4ti-0550	100 枚	2	22,000	44,000
PCR 8 連フラットキャップ	日本ジェネテイクス	FG-008FC	120 ストリップ / 袋	10	2,200	22,000

PCR Seal	日本ジェネテイクス	4ti-0500	100 枚	2	11,000	22,000
KOD fx neo	TOYOBO	KFX-201X5	200U x 5	2	98,000	196,000
AAVpro® Purification Kit Midi (All Serotypes)	TAKARA	6675		1	40,810	40,810
Amicon ultra-0.5 centrifugal filter unit	Merckmillipore	UFC510096	96	2	57,300	114,600
AAVpro® Titration Kit (for Real Time PCR) Ver.2	TAKARA	6233		2	79,000	158,000
AAVpro® 293T Cell Line	TAKARA	632273		1	70,000	70,000
Droplet Reader オイル	Bio-Rad	1863004	1Lx2	1	180,000	180,000
DG8 Cartridges and Gaskets	Bio-Rad	1864007		1	280,000	280,000
PCR プレート 96、分割可能, スカートなし, 250 µL, PCR clean, 青	eppendorf	0030133404	20 枚	3	21,500	64,500
ddPCR Supermix for Probe (no dUTP)	Bio-Rad	1863024	1mlx5	1	100,000	100,000
Droplet generator オイル for Probes	Bio-Rad	1863005	70mlx10	1	65,000	65,000
0.2ml Low profile Thin-walled 8 Tube & Domed Cap strips	Thermo Scientific	AB-0775	1 box	1	34,100	34,100

PhiX Control v3	Illumina	FC-110-3001	10 ul	2	-	-
Qubit dsDNA HS Assay Kit	Invitrogen	Q32854		2	39,900	79,800
0.5 ml PCR Tube with Flat Cap, Thin-Wall, Non-Sterile, MaxyClear	AXY	PCR-05-C	1000 pieces	1	5,900	5,900
Nucleobond Xtra Midi EF	TAKARA	U0420B	50 回	1	97,000	97,000
MiniSeq Mid Output Reagent Cartridge 300 cycle	Illumina	FC-420-1004		5	86,500	432,500
合計					¥3,370,048	

< 5-4-5. 臨床試験パートナーの検討 >

動物モデルでの in vivo 実験や臨床試験については、ヒューマンセラピー分野での専門的なパイプラインやリソースを持たないバイオパレット社が単独で行うことは難しい。そのため、アウトラボでさらなる開発を進めるためのビジネスパートナーとして、臨床試験関連企業との連携が必要となってくる。

開発業務受託機関（CRO）は、開発プロセス全体を通じて臨床パートナーとして、企業と肩を並べて目標やスケジュールを設定し、相互理解を深めながら、協力企業の延長線上で活動することができる。では、CRO を選ぶ際、何を重視すべきなのだろうか。以下は、最も考慮すべき点と考えられる⁶⁷。

需要と規模

臨床試験の管理には、試験デザイン、無作為化、治験薬供給管理（RTSM）、データ管理、データ収集から、システム、生物統計学、ファーマコビジュランス、薬事コンプライアンス、メディカルライティングに至るまで、多数の複雑な要素が伴う。CRO サービス市場は 2023 年までに推定 563.4 億ドル規模になり、IQVIA、LabCorp、Parexel、PRA、PPD、ICON、Syneos、Charles River、Wuxi、SGS といった有力企業がその 50%以上を占める。バイオテクノロジー

企業の第1相臨床試験の平均コストは300万ドルから500万ドルとすると、大手CROの年間収益の平均0.002%未満に相当する。つまり、「大手10社」は、より大規模で収益性の高い製薬会社の試験に引き寄せられることを好むということである。バイオテック企業の治験が必要な注目を浴びず、大手企業の要請を優先されるリスクは高い。例えば、大規模スポンサーの治験が多くのリソースを必要とする場合、小規模治験のリソースをそちらに振り向ける傾向があり、その結果、小規模企業は自社の想定するスケジュールで治験実施できない可能性が存在する。

専門性の評価

医薬品、バイオマーカー、医療機器の開発を成功させるには、綿密に計画された臨床試験が不可欠である。企業は、臨床試験の経過を注意深く計画し、予測可能性を積極的に改善し、リスクを低減し、GO/NO GOの決定を迅速に下すために協力してくれるパートナーを必要としている。一方、高品質のデザインコンサルティングを提供できるパートナーが魅力的だが、専門性が合うかどうかの考慮も重要である。結局のところ、総合的なサービスを提供する企業は、全体としてハイレベルなニーズに応えることができるかもしれないが、すべての領域にわたる特定の専門知識を持っているとは限らない。一方、より焦点を絞った企業は、その専門性を特定の分野に限定し、その分野での質の提供に全面的なリソースを傾注する。不適切な試験デザイン、不適切な投与量の選択、最適でない評価スケジュール、不適切な評価基準／マーカー、データ解析方法の問題など、あらゆる理由で試験は失敗する可能性がある。つまり、ターゲットする分野が詳しいパートナーを選択し、それに合わせた試験デザインは成功に不可欠なのである。

カルチャー・フィット

適切なカルチャー・フィット、すなわち人とミッションの適合性と一貫性は、協働作業にとって極めて重要である。それは、チームワーク、信頼、透明性、そして継続的なビジネスチャンスを意味する。オープンマインドな企業もあれば、保守的な企業も存在する。企業は、その治験がCROの使命や価値観に共鳴するかどうか、科学や顧客サービスを重視する会社であるかどうか、さらに地理的な位置などより実際の検討事項を考慮すべきである。

実績

企業は新薬の創製に特化しており、技術評価に関する経験と専門知識が求められる。CROの実績も、パートナーとして評価する上で重要な要素である。コンサルタントサービスを依頼することは、CROの経験と専門性を評価するのに最適な方法であると考えている。また、優れた臨床パートナーは、堅牢な臨床試験のデザイン、設定、実施、終了、提出方法を理解している必要がある。バイオテクノロジー企業は、CROがどれだけ多くの臨床試験に関与してきたか、どの治療領域で仕事をしてきたか、科学文献における存在感、規制当局とのつながりなども見る必要がある。

一貫性

大手CROは離職率が高いため、この点が臨床試験における迅速かつ質の高いサービスの提供にあたり、課題となることがある。新しいリソース（人材）を見つけ、教育訓練するためには一定の期間を要するため、離職率の高さは臨床試験を混乱させる可能性がある。また、新しいチームメンバーをスピードアップさせるプロセスが欠けている場合も、一貫性が損なわれる可能性がある。また、試験中にリソースの調整が必要になることもある。例えば、COVID-19の危機など、他のニーズが優先され、いくつかの試験が保留される可能性がある。したがって、トラブルに迅速に対応でき、一貫性を重視するCROを選ぶべきである。また、選択したCROパートナーが、ある試験から別の試験へのリソースの再割り当てが可能かどうか、またその経験があるかどうかを検討すべきである。

このような注意点を念頭に置いて、私たちはCRO業界に目を向け、日本で活躍しているCRO企業に注目した。前述したように、臨床試験を進めていく上で、提携先のCROと頻繁に打ち合わせを行う必要があり、臨床現場に頻繁に足を運んでデータをフィードバックしてもらい、さらなる開発につなげる必要がある。そのため、地理的に近いことも要因の一つであるが、企業文化や意思伝達の速さの面でも日本のCROの方が連携しやすいと考えられる。そのため、日本における規模が異なる7つの注目されるCROについて、調査を行った。その結果、イーピーエス社とFortrea Japan社が最も有力なパートナーと考えられる。

イーピーエス社は大規模 CRO であり、豊富な臨床試験実績がある。また、大規模企業ならではの、社員の満足度の高さが臨床開発の安定性を担保していた。最後に、新たな臨床パイプラインの開発にも力を入れており、再生医療分野における臨床試験手順の標準化・発展に大きく注力している。遺伝子治療も新興のパイプラインであるが、日本ではまだ数が少ないため、イーピーエス社にとっては新たな臨床パイプラインを開発する良い機会であり、開発の重視度が高いと推測される。

Fortrea Japan 社の最も魅力的な点は、遺伝子治療の臨床試験を専門とするコンサルタントとチームを持っていることである。これによって、技術的なレベルでのコミュニケーションが容易になるだけでなく、私たちのニーズがよりよく理解され、実現される。また、Target-AID は世界で最もよく知られたゲノム編集技術の一つであり、この技術による初の臨床試験は、Fortrea Japan 社にとって十分な興味があり、真剣に取り組んでくれると思われ、私たちの臨床試験がスムーズに行われる可能性が高まる。加えて、Fortrea Japan 社は比較的小規模であるため、バイオパレット社のようなベンチャー企業のパイプラインも重視すると思われる。また、離職率が低く、従業員満足度が高いことも、臨床開発の安定性が高いと考えられる。

また、今後事業を展開していく中で、様々な CRO と積極的に相談し、より多くの CRO とコンタクトを取るようにし、CRO との交流を通じて経験を蓄積し、最終的にパートナー先を決めていきたいと考えている。

< 5-4-6. 臨床試験コストの調査 >

各治療分野における第 1 相、第 2 相、第 3 相臨床試験の平均費用は、それぞれ約 300 万ドル、1400 万ドル、2200 万ドルである。治療分野によって、以下の表で示しているように、要する費用が少し異なってくる。

Therapeutic Area	Phase 1	Phase 2	Phase 3	Phase 1, 2, & 3 Subtotal	FDA NDA/BLA Review Phase Cost	Total
Anti-Infective	\$4.2	\$14.2	\$22.8	\$41.2	\$2.0	\$43.2
Cardiovascular	\$2.2	\$7.0	\$25.2	\$34.4	\$2.0	\$36.4
Central Nervous System	\$3.9	\$13.9	\$19.2	\$37.0	\$2.0	\$39.0
Dermatology	\$1.8	\$8.9	\$11.5	\$22.2	\$2.0	\$24.2
Endocrine	\$1.4	\$12.1	\$17.0	\$30.5	\$2.0	\$32.5
Gastrointestinal	\$2.4	\$15.8	\$14.5	\$32.7	\$2.0	\$34.7
Genitourinary System	\$3.1	\$14.6	\$17.5	\$35.2	\$2.0	\$37.2
Hematology	\$1.7	\$19.6	\$15.0	\$36.3	\$2.0	\$38.3
Immunomodulation	\$6.6	\$16.0	\$11.9	\$34.5	\$2.0	\$36.5
Oncology	\$4.5	\$11.2	\$22.1	\$37.8	\$2.0	\$39.8
Ophthalmology	\$5.3	\$13.8	\$30.7	\$49.8	\$2.0	\$51.8
Pain and Anesthesia	\$1.4	\$17.0	\$52.9	\$71.3	\$2.0	\$73.3
Respiratory System	\$5.2	\$12.2	\$23.1	\$40.5	\$2.0	\$42.5
Average	<u>\$3.4</u>	<u>\$13.6</u>	<u>\$21.8</u>	\$38.7	\$2.0	\$40.7

図 23：各疾患分野による臨床試験コスト（\$ Millions）.Ref: How Much Does a Clinical Trial Cost? SOFPROMED（2020）を基に筆者作成

臨床試験の費用は、試験の規模（患者数）、場所、臨床施設の数、治療領域、薬剤の種類、プロトコルごとに必要な特定の検査や手順など、いくつかの要素によって異なる。試験規模（患者数）は、試験フェーズと密接な関係がある。初期段階（第 1 相）の試験で必要とされる患者数は少数（平均 15～20 人）である一方、国際的な第 3 相試験では数百人（数千人）の被験者が必要とされることもある。さらに、学術的（非商業的）試験と企業主導（商業的）試験は区別すべきである。営利を目的としない学術的なスポンサーは、営利を目的とせずに臨床試験を実施することができ、臨床試験活動を賄うための資金が少なくてすむ。一方、営利目的のスポンサーは、医薬品開発プログラムを実施するための予算が多い。したがって、臨床試験費用に関する試算は容易ではないため、同じく塩基編集ベンチャー企業である Beam 社の損益計算書を参照した（図 24）。損益計算書より、2023 年第 1 四半期の研究開発

費は 9760 万ドルであり、その中で、主に BEAM-101 および BEAM-201 の製造および臨床費用、BEAM-301 および BEAM-302 の IND 申請のための試験およびアッセイ開発費用は 3,090 万であると示されている。

	2023	2022	2023	2022
License and collaboration revenue	\$ 20,116	\$ 16,652	\$ 44,324	\$ 25,084
Operating expenses:				
Research and development	97,608	74,556	197,254	139,966
General and administrative	24,656	24,062	48,146	43,309
Total operating expenses	122,264	98,618	245,400	183,275
Loss from operations	(102,148)	(81,966)	(201,076)	(158,191)
Other income (expense):				
Change in fair value of derivative liabilities	(900)	12,200	4,700	25,800
Change in fair value of non-controlling equity investments	6,148	(4,124)	(6,649)	(11,809)
Change in fair value of contingent consideration liabilities	2,171	(120)	1,875	332
Interest and other income (expense), net	11,953	2,060	21,914	2,704
Total other income (expense)	19,372	10,016	21,840	17,027
Net loss	\$ (82,776)	\$ (71,950)	\$ (179,236)	\$ (141,164)
Unrealized gain (loss) on marketable securities	(1,250)	(1,481)	415	(4,140)
Comprehensive loss	\$ (84,026)	\$ (73,431)	\$ (178,821)	\$ (145,304)
Net loss per common share, basic and diluted	\$ (1.08)	\$ (1.02)	\$ (2.41)	\$ (2.03)
Weighted-average common shares outstanding, basic and diluted	76,335,175	70,210,227	74,315,721	69,461,207

図 24 : Beam 社の 2023 年第 1 四半期損益計算書. Ref: Beam's profit and loss report for the first quarter of 2023.

また、臨床試験 CRO の見積書には通常、以下の費用項目が含まれる：規制関連業務、施設の特定と選定、施設の契約と支払い、サイトの開始と起動、施設管理、オンサイトモニタリング、医薬品安全管理、医薬品物流、生物学的サンプルの物流、医薬品物流、メディカルライティング、施設クローズアウト、プロジェクト管理、試験ファイル／文書管理、データ管理、統計、品質管理、中央 CRO/スポンサーとのコミュニケーション、パススルー費用。

臨床試験は、今後最も多くの費用を投じる必要があるところであるが、現段階では初期開発の不確実性が高いため、妥当な値を予測することは難しい。臨床コストはこれから、研究が進展し、CRO と協議する中で徐々に決定される。

6. Conclusion

In summary, we showed the engineering of deaminase PmCDA1 and Cas9 helps to create a much more precisely editing profile and allows us to generated three novel base editing systems: Target-AID2S and Target-AID3S. Then, by combining with SaCas9 and other impact components, we generated the world's smallest base editing system and demonstrated the ability of single AAV-mediate editing. The protein engineering/cell delivery approaches we accomplished in this study substantially expand the scope and capabilities of genome editing. We believe our work can contribute to creating a flexible, safe, and novel genome editing platform.

Looking back over history, we find there is a big gap between research-break-through and business-start-up, therefore a systematic analysis of what field to go into, what are the unique aspects of my business and is this field have a good market seems very important. In this paper, we focused on the technical advance of our developed technology and planned our innovation strategies through a business proposal to BioPalette due to technical license. By analyzing technology strategy, intellectual property strategy, business strategy, and financial strategy, as a series of strategic cycles, and demonstrating the potential of Target-AID2S, 3S using in HBV treatment, we believe our strategies will lead BioPalette to negotiations for a new cross-licensing agreement with Beam Therapeutics. We hope our work could inspire more researchers and help their fascinating idea not just stop at the academic level, but also bring into the innovation of business.

Reference

1. Gaj, T., Sirk, S. J., Shui, S. L. & Liu, J. Genome-editing technologies: Principles and applications. *Cold Spring Harb. Perspect. Biol.* **8**, (2016).
2. Sebastiano, V. *et al.* In situ genetic correction of the sickle cell anemia mutation in human induced pluripotent stem cells using engineered zinc finger nucleases. *Stem Cells* **29**, (2011).
3. Zhang, F. *et al.* Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat. Biotechnol.* **29**, (2011).
4. Yusa, K. *et al.* Targeted gene correction of α 1-antitrypsin deficiency in induced pluripotent stem cells. *Nature* **478**, (2011).
5. Li, H. *et al.* In vivo genome editing restores haemostasis in a mouse model of haemophilia. *Nature* **475**, (2011).
6. Soldner, F. *et al.* Generation of isogenic pluripotent stem cells differing exclusively at two early onset parkinson point mutations. *Cell* **146**, (2011).
7. Zhang, Y. *et al.* Transcription activator-like effector nucleases enable efficient plant genome engineering. *Plant Physiol.* **161**, (2013).
8. Haun, W. *et al.* Improved soybean oil quality by targeted mutagenesis of the fatty acid desaturase 2 gene family. *Plant Biotechnol. J.* **12**, (2014).
9. Hockemeyer, D. *et al.* Genetic engineering of human pluripotent cells using TALE nucleases. *Nat. Biotechnol.* **29**, (2011).
10. Wienert, B. *et al.* Editing the genome to introduce a beneficial naturally occurring mutation associated with increased fetal globin. *Nat. Commun.* **6**, (2015).
11. Nishimasu, H. *et al.* Crystal Structure of Staphylococcus aureus Cas9. *Cell* **162**, (2015).
12. Verma, A. K. *et al.* Current status and perspectives on the application of crispr/cas9 gene-editing system to develop a low-gluten, non-transgenic wheat variety. *Foods* vol. 10 (2021).
13. Saurabh, S. Genome Editing: Revolutionizing the Crop Improvement. *Plant Molecular Biology Reporter* vol. 39 (2021).
14. Hsu, P. D., Lander, E. S. & Zhang, F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* vol. 157 (2014).
15. Lyu, C. *et al.* Targeted genome engineering in human induced pluripotent stem cells from patients with hemophilia B using the CRISPR-Cas9 system. *Stem Cell Res. Ther.* **9**, (2018).
16. Shimo, T., Hosoki, K., Nakatsuji, Y., Yokota, T. & Obika, S. A novel human muscle cell model of Duchenne muscular dystrophy created by CRISPR/Cas9 and evaluation of antisense-mediated exon skipping. *J. Hum. Genet.* **63**, (2018).
17. Bjursell, M. *et al.* Therapeutic Genome Editing With CRISPR/Cas9 in a Humanized Mouse Model Ameliorates α 1-antitrypsin Deficiency Phenotype. *EBioMedicine* **29**, (2018).
18. Xu, S. *et al.* Editing aberrant splice sites efficiently restores b-globin expression in b-thalassemia.

- Blood* **133**, (2019).
19. Park, S. H. *et al.* Highly efficient editing of the β -globin gene in patient-derived hematopoietic stem and progenitor cells to treat sickle cell disease. *Nucleic Acids Res.* **47**, (2019).
 20. Ledford, H. Alternative CRISPR system could improve genome editing. *Nature* **526**, (2015).
 21. Porto, E. M., Komor, A. C., Slaymaker, I. M. & Yeo, G. W. Base editing: advances and therapeutic opportunities. *Nature Reviews Drug Discovery* vol. 19 (2020).
 22. Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A. & Liu, D. R. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533**, (2016).
 23. Roy, A., Kucukural, A. & Zhang, Y. I-TASSER: A unified platform for automated protein structure and function prediction. *Nat. Protoc.* **5**, (2010).
 24. Banno, S., Nishida, K., Arazoe, T., Mitsunobu, H. & Kondo, A. Deaminase-mediated multiplex genome editing in *Escherichia coli*. *Nat. Microbiol.* **3**, (2018).
 25. Nishida, K. *et al.* Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science (80-.).* **353**, (2016).
 26. Landrum, M. J. *et al.* ClinVar: Public archive of interpretations of clinically relevant variants. *Nucleic Acids Res.* **44**, (2016).
 27. Gaudelli, N. M. *et al.* Programmable base editing of T to G C in genomic DNA without DNA cleavage. *Nature* **551**, (2017).
 28. Zuo, E. *et al.* Cytosine base editor generates substantial off-target single-nucleotide variants in mouse embryos. *Science (80-.).* **364**, (2019).
 29. Jin, S. *et al.* Cytosine, but not adenine, base editors induce genome-wide off-target mutations in rice. *Science (80-.).* **364**, (2019).
 30. McGrath, E. *et al.* Targeting specificity of APOBEC-based cytosine base editor in human iPSCs determined by whole genome sequencing. *Nat. Commun.* **10**, (2019).
 31. Qiao, Q. *et al.* AID Recognizes Structured DNA for Class Switch Recombination. *Mol. Cell* **67**, (2017).
 32. Tan, J., Zhang, F., Karcher, D. & Bock, R. Expanding the genome-targeting scope and the site selectivity of high-precision base editors. *Nat. Commun.* **11**, (2020).
 33. Daya, S. & Berns, K. I. Gene therapy using adeno-associated virus vectors. *Clinical Microbiology Reviews* vol. 21 (2008).
 34. Kumar, N. *et al.* The development of an AAV-based crispr sacas9 genome editing system that can be delivered to neurons in vivo and regulated via doxycycline and cre-recombinase. *Front. Mol. Neurosci.* **11**, (2018).
 35. Nguyen Tran, M. T. *et al.* Engineering domain-inlaid SaCas9 adenine base editors with reduced RNA off-targets and increased on-target DNA editing. *Nat. Commun.* **11**, (2020).
 36. Chen, Y. *et al.* Development of Highly Efficient Dual-AAV Split Adenosine Base Editor for In

- Vivo Gene Therapy. *Small Methods* **4**, (2020).
37. Pettersen, E. F. *et al.* UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, (2021).
 38. Dicarlo, J. E. *et al.* Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic Acids Res.* **41**, (2013).
 39. Doman, J. L., Raguram, A., Newby, G. A. & Liu, D. R. Evaluation and minimization of Cas9-independent off-target DNA editing by cytosine base editors. *Nat. Biotechnol.* **38**, (2020).
 40. Ran, F. A. *et al.* In vivo genome editing using *Staphylococcus aureus* Cas9. *Nature* **520**, (2015).
 41. Juven-Gershon, T., Cheng, S. & Kadonaga, J. T. Rational design of a super core promoter that enhances gene expression. *Nat. Methods* **3**, (2006).
 42. Lima, S. A. *et al.* Short poly(A) tails are a conserved feature of highly expressed genes. *Nat. Struct. Mol. Biol.* **24**, (2017).
 43. Clement, K. *et al.* CRISPResso2 provides accurate and rapid genome editing sequence analysis. *Nature Biotechnology* vol. 37 (2019).
 44. Wang, Y., Zhou, L., Liu, N. & Yao, S. BE-PIGS: A base-editing tool with deaminases inlaid into cas9 pi domain significantly expanded the editing scope. *Signal Transduction and Targeted Therapy* vol. 4 (2019).
 45. Yu, Y. *et al.* Cytosine base editors with minimized unguided DNA and RNA off-target events and high on-target activity. *Nat. Commun.* **11**, (2020).
 46. Zhao, T. *et al.* Small-molecule compounds boost genome-editing efficiency of cytosine base editor. *Nucleic Acids Res.* **49**, (2021).
 47. Tsai, S. Q. *et al.* GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat. Biotechnol.* **33**, (2015).
 48. Fu, Y. *et al.* High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* **31**, (2013).
 49. Levy, J. M. *et al.* Cytosine and adenine base editing of the brain, liver, retina, heart and skeletal muscle of mice via adeno-associated viruses. *Nat. Biomed. Eng.* **4**, (2020).
 50. Villiger, L. *et al.* Treatment of a metabolic liver disease by in vivo genome base editing in adult mice. *Nat. Med.* **24**, (2018).
 51. Kim, Y. B. *et al.* Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions. *Nat. Biotechnol.* **35**, (2017).
 52. Zuo, E. *et al.* A rationally engineered cytosine base editor retains high on-target activity while reducing both DNA and RNA off-target effects. *Nat. Methods* **17**, (2020).
 53. Grünewald, J. *et al.* Transcriptome-wide off-target RNA editing induced by CRISPR-guided DNA base editors. *Nature* **569**, (2019).
 54. Grünewald, J. *et al.* CRISPR DNA base editors with reduced RNA off-target and self-editing

- activities. *Nat. Biotechnol.* **37**, (2019).
55. Lada, A. G. *et al.* Genome-Wide Mutation Avalanches Induced in Diploid Yeast Cells by a Base Analog or an APOBEC Deaminase. *PLoS Genet.* **9**, (2013).
 56. Chu, S. H. *et al.* Rationally Designed Base Editors for Precise Editing of the Sickle Cell Disease Mutation. *Cris. J.* **4**, (2021).
 57. Cohen, J. The latest round in the CRISPR patent battle has an apparent victor, but the fight continues. *Science* (80-.). (2020) doi:10.1126/science.abe7573.
 58. 山本一彦; 桑山 斉; 高 畑 豪太郎; 三宅秀昭. 「神戸大学発バイオベンチャーのイノベーション戦略」 (国民経済雑誌第226巻第1号 pp.11－34) 2022年07月.pdf.
 59. Bio Palette Co., Ltd. Official website.
 60. National Institute for Health and Care Excellence. *NICE health technology evaluations: the manual. Process and methods [PMG36]. National Institute for Health and Care Excellence* (2022).
 61. Newswire, P. R. Cell and Gene Therapy CROs Market, 2018-2030. *NY-REPORTLINKER* (2018).
 62. Abbas, Z., Memon, M. S., Mithani, H., Jafri, W. & Hamid, S. Treatment of chronic hepatitis D patients with pegylated interferon: A real-world experience. *Antivir. Ther.* **19**, (2014).
 63. Le Duff, Y., Blanchet, M. & Sureau, C. The Pre-S1 and Antigenic Loop Infectivity Determinants of the Hepatitis B Virus Envelope Proteins Are Functionally Independent. *J. Virol.* **83**, (2009).
 64. Vaz, F. M. *et al.* Sodium taurocholate cotransporting polypeptide (SLC10A1) deficiency: Conjugated hypercholanemia without a clear clinical phenotype. *Hepatology* **61**, (2015).
 65. Ni, Y. *et al.* Hepatitis B and D viruses exploit sodium taurocholate co-transporting polypeptide for species-specific entry into hepatocytes. *Gastroenterology* **146**, (2014).
 66. Yan, H. *et al.* Sodium taurocholate cotransporting polypeptide is a functional receptor for human hepatitis B and D virus. *Elife* **2012**, (2012).
 67. Nestor, M. *et al.* The Future of Contract Research Organizations in Dermatology and Aesthetic Research. *J. Clin. Aesthet. Dermatol.* **15**, (2022).

Acknowledgments

My sincere gratitude goes to professor Akihiko Kondo, professor Keiji Nishida, assistant professor Hitoshi Mitsunobu, and assistant professor Shin Yoshioka, really appreciate the supports and advice on a daily experiment. Also, I would like to express my deepest gratitude to professor Kazuhiko Yamamoto for guiding and inspiring me in the construction of my innovation strategies. Finally, I would like to thank all the others who gave me advice and encouragement in the lab or class. This work is supported by the Nishida's Lab, Graduate School of Science, Technology and Innovation, Kobe University.

Doctor Thesis, Kobe University

“Cytosine base editing systems with minimized off-target effect and molecular size”, 134 pages

Submitted on Jan, 22, 2024

The date of publication is printed in cover of repository version published in Kobe University Repository Kernel.

© Author's name (Li Ang)

All Right Reserved, 2024