universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„An Exploration of Prime Editing Approaches in Primary T Cells to Unravel the Effects of Immune Disease Associated Variants"

verfasst von / submitted by

Vinika Gandhi BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2024 / Vienna, 2024

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066 877
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Genetik und Entwicklungsbiologie
Betreut von / Supervisor:	Univ.-Prof. Mag. Dr. Pavel Kovarik

TABLE OF CONTENT

ACKOWELEGDEMENTS

My very deepest gratitude goes to my mother, whose constant encouragement and support have been the foundation of my academic journey. Her belief in me has been my greatest motivation and source of strength throughout my studies.

I am incredibly grateful to Professor Lars Steinmetz for giving me the opportunity to complete this thesis under his guidance. My time in his laboratory was not only academically enriching, but personally transformative as I was inspired by the scientific atmosphere and team spirit among colleagues.

I would especially like to thank my mentor in the lab and beyond, Dewi Moonen, whose invaluable mentorship and support were central to my growth and development. She has been a pleasure to work with and I am grateful for the insights and knowledge she has generously shared with me. Her constructive feedback and engaging discussions helped to further my scientific understanding and improve my skills. She skillfully managed to find a balance between autonomy and guidance, creating an environment fostered for self-learning and growth.

I am grateful to Annique Claringbould for her support in the final stages of my work, especially for her patience and her great expertise in navigating the bioinformatics aspects.

In addition, I am grateful to all my colleagues at the Steinmetz laboratory EMBL Heidelberg, especially Amanda Hughes, Daniel Schraivogel, Dominik Lindenhofer, Kai Fenzl, Linda Müller, and Shengdi Li, for providing a supportive environment in which I could learn and thrive.

A heartfelt appreciation goes to Laura, my neighbor and colleague at EMBL Heidelberg, whose friendship and shared passion for science made my time there even more enjoyable.

I am deeply grateful to my partner Chris for being a solid source of strength and unwavering support throughout my Master thesis. His encouragement and presence have been absolutely invaluable to me, and I am deeply grateful for his support during this time.

Finally, I would like to thank my supervisor at the University of Vienna, Professor Pavel Kovarik, for his support during my work.

ZUSAMMENFASSUNG

Diese Masterarbeit befasst sich mit Precision-Editing-Methoden für CD4-positive primäre T-Zellen, um in Zukunft gezielt Enhancer-Regionen zu editieren, die mit Immunkrankheiten in Verbindung stehen. Ziel ist es, dadurch Verständnis des genregulatorischen Netzwerks von CD4-positiven primären T-Zellen für die zukünftige Forschung zu erweitern. Es wurden verschiedene Editing-Methoden in Betracht gezogen, darunter Base-Editing, das CRISPR-Cas9-System und Prime-Editing (PE), wobei letztlich das PE aufgrund seiner Vielseitigkeit bei der Einführung von Edits und Mutationen ausgewählt wurde. PE-Konstrukte wurden mit Merkmalen zur Steigerung der Editing-Effizienz entwickelt, darunter All-in-One-Plasmide für die Transfektion in einem Schritt und ein detailliertes Design von engineerd prime editing guide RNAs epegRNAs, die für die Editing-Effizienz entscheidend sind. Die extrazelluläre Domäne von CD4-Glykoproteinen auf der Oberfläche von T-Zellen wurde als Editing-Ziel gewählt, wobei das AIO-M-Plasmid mit der MLH1dn-Sequenz aufgrund seiner Fähigkeit, die Effizienz des PE zu erhöhen, ausgewählt wurde. Als Transfektionsmethode wurde die Elektroporation gewählt, die eine erfolgreiche Transfektion von CD4-T-Zellen ermöglichte. Die GFP-Expression zeigte das Vorhandensein des Konstrukts an, wobei die mit AIO-M transfizierten Zellen eine Transfektionseffizienz von 0,7 % bis 15 % aufwiesen, und die sortierten GFP-positiven Zellen wurden einer Sequenzierung unterzogen. CRISPresso2 wurde für die bioinformatische Datenanalyse verwendet, wobei die wichtigsten Editierparameter für eine zuverlässige Quantifizierung der Edits genau festgelegt wurden. Erste Ergebnisse zeigten, dass 0,18 % Edits entdeckt wurden, was auf ein Potenzial für die Anwendung von Prime Editing hindeutet, aber es sind weitere Beweise und Versuche notwendig, um die Resultate zu bestärken. Weitere Experimente sind erforderlich, um den Nutzen des PE bei CD4-positiven primären T-Zellen zu bestätigen.

ABSTRACT

This master thesis focuses on precision editing methods for CD4-positive primary T cells, with the future aim of targeting enhancer regions associated with immune diseases. The objective is also to expand the understanding of the gene regulatory network of CD4 positive primary T cells using precision editing methods for future research. Various editing methods were considered, including base editing, the CRISPR-Cas9 system, and prime editing, ultimately selecting prime editing for its versatility in introducing edits and mutations. Prime editing constructs were designed with features to enhance editing efficiency, including all-in-one plasmids for single-step transfection and detailed design of engineerd prime editing guide RNAs (epegRNAs), crucial for editing efficiency. The extracellular domain of CD4 glycoproteins on T cell surfaces was chosen as the editing target, with the All-in-One AIO-M plasmid containing the MLH1dn sequence selected for its ability to enhance prime editing efficiency. Electroporation was chosen as the transfection method, enabling successful transfection of CD4 T cells. GFP expression indicated construct presence, with AIO-M transfected cells showing a transfection efficiency ranging from 0.7 % to 15 %, and sorted GFP-positive cells were subjected to sequencing. CRISPresso2 was utilized for bioinformatic data analysis, precisely specifying prime editing parameters for reliable edit quantification. Initial results showed 0.18 % edits detected, indicating potential for prime editing application, but further evidence is needed to strengthen the results. Further experiments are necessary to confirm the utility of prime editing in CD4-positive primary T cells.

Introduction

It is known that mutations in regulatory elements in the non-coding regions of the genome, leading to T cell dysfunction, underlie immune and autoimmune diseases. The effectiveness of the adaptive immune system depends on its ability to distinguish between foreign pathogens and self-antigens. In cases of autoimmunity, the disruption of this immune recognition process leads to abnormal immune responses directed against healthy tissue. Two main types of T cells exist: cytotoxic T cells and helper T cells. Cytotoxic T cells are characterized by the CD8 (Cluster of Differentiation 8) receptor on their membranes and eliminate infected and cancer cells. Helper T cells are characterized by the CD4 receptor on their membranes, and do not eliminate cells directly but signal to coordinate the immune response of T cells, B-cells and macrophages.[1]

This study aims to establish connections between enhancers housing immune disease-associated variants and the genes they regulate. As CD4 T cells play key roles in the immune system, understanding their functions is vital for unraveling immune responses, disease mechanisms, and devising novel therapies. To achieve this the aim is to enhance the precision editing toolkit for primary CD4+ T cells and conduct a small-scale prime editing screen targeting GWAS variants within prioritized enhancer regions. Moreover, leveraging an innovative approach called targeted perturb-seq (TAP-Seq), pioneered in the Steinmetz lab, enables concurrent analysis of transcriptome and genetic perturbations at a single-cell resolution. [2] By elucidating the impacts of disease-associated variants on gene regulation, this method holds promise for identifying potential therapeutic targets.

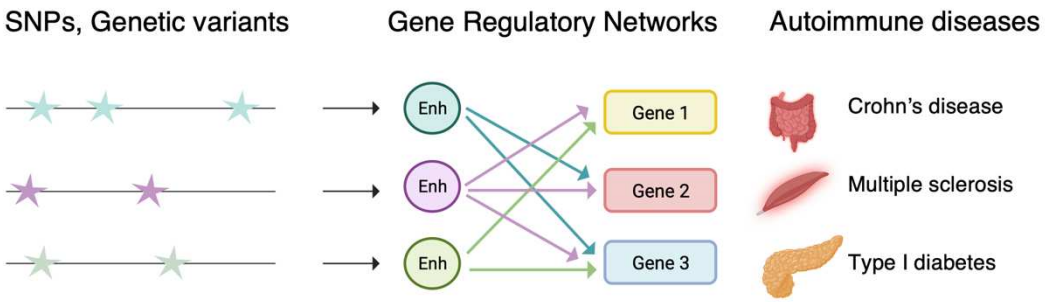


Figure 1: Single nucleotide polymorphisms (SNP) can often be found in non-coding regions of the genome at gene regulatory elements, such as enhancers. These enhancer regions can be linked to the genes they regulate. An enhancer can often regulate more than one gene. These genes may influence disease phenotype.

Over 80 autoimmune diseases are known to date including multiple sclerosis, rheumatoid arthritis, and type 1 diabetes, which are inherited and influenced by polygenic factors (Figure 1). The effectiveness of the adaptive immune system hinges on its ability to differentiate between foreign pathogens and self-antigens.[3] In cases of autoimmunity, the disruption of this immune recognition process results in abnormal immune responses targeting healthy tissues.[4] Central to the orchestration of adaptive immune responses are T cells, a class of lymphocytes crucial for defense against specific pathogens and exogenous agents. T cell activation initiates a complex

process of differentiation, which results in distinct subtypes each with specialized functions. [5,6] Optimal T cell activation is significant for a robust immune response, while excessive activity serves as a hallmark in immune-mediated pathologies.[7]

T Cell Activation

T cell precursors or thymocytes are derived from lymphoid hematopoietic stem cells that leave the bone marrow and develop in the primary lymphoid organ, the thymus. An environment with components such as cytokines and chemokines that help the thymocytes to develop into mature T cells is provided by the thymus. During the thymic differentiation process, surface molecules such as CD4 and CD8 are expressed. A positive and negative selection process then takes place resulting in the differentiation into CD4+ and CD8+ T cells. CD4, a surface glycoprotein, acts as a co-receptor alongside the T cell receptor (TCR). T cells remain in their naïve state until they encounter a specific antigen that is recognized and bound by the TCR on their cell surface. The CD4+ T cell subset comprises a portion of the lymphocyte population, because they are important in the control and clearance of infections and play a key role in the activation of B cells and CD8+ T cells. [8] During an immune response, the TCR and CD4 co-receptor interact with the antigen-major histocompatibility II complex (MHC-II complex) presented by antigen-presenting cells (APCs). Subsequently, the co-stimulatory CD28 receptor on T cells engages with CD80/CD86 molecules expressed on APCs (see Figure 2). This sequence of molecular events culminates in the production of interleukin-2 (IL-2), a critical signaling molecule that facilitates T cell differentiation. Upon activation, CD4+ T cells differentiate into specialized subtypes depending on the presence of specific cytokines.[9]

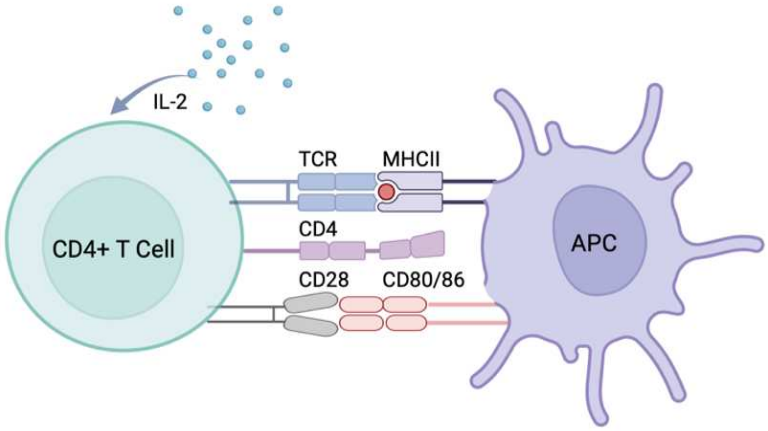


Figure 2: Several stimuli are required for T cells to be activated. The T cell receptor (TCR) interacts with the antigen presenting major histocompatibility II complex (MHC-II complex) of an antigen-presenting cell (APC). Costimulatory receptor CD28 interacts with CD80/86 on the APC. As a result of these receptor interactions, the uptake of the cytokine interleukin-2 (IL-2) takes place. The production of IL-2, which acts as a signaling molecule and survival signal, this and other cytokines determine T cell differentiation. Figure was adapted from Luckheeram *et al.* (2012).

Disease Associated Variants in Enhancer Regions

Failure to initiate an appropriate immune response due to inadequate stimuli often serves as an indicator of immune disorders. Enhancers play a key role in the regulation of genes required for T cell activation.[10,11] These cis-regulatory elements which are present in intergenic, intronic, and occasionally exon regions. They are characterized by clusters of binding sites for transcription factors and are targeted by cell-specific transcription factors, as well as chromatin modifiers and architectural proteins such as Cohesin, Condensin, and CCCTC-binding factor (CTCF).[12] 90 % of all GWAS SNPs fall in non-coding regions and are not directly associated with the coding sequence of a gene. When GWAS studies with annotations of cis-regulatory elements specific to immune cell types are analyzed around 60 % of the immune disease-associated genetic variants map to enhancers specific to immune cells. This suggests that these enhancers play an important role in the regulation of genes involved in immune function and that, when disrupted by genetic variants, they may contribute to the development of immune-related diseases.[13,14]

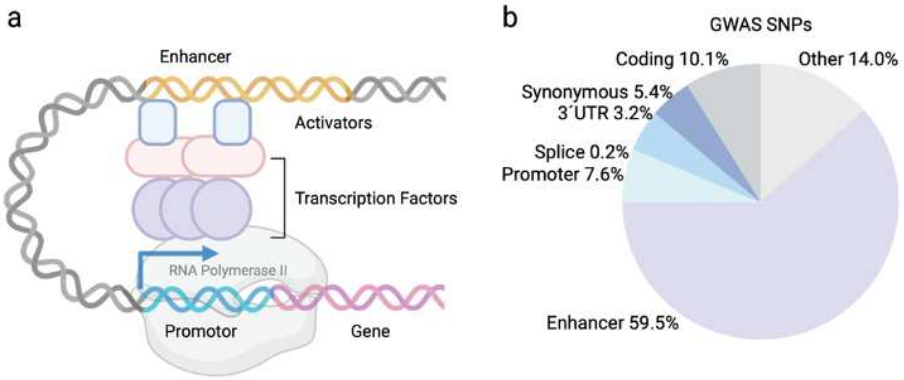


Figure 3: a) Enhancers are genetic elements that regulate transcription. These enhancers do not necessarily have to be in the proximity of the genes they regulate. Therefore, they are also referred to as distal control elements to which transcription factors can bind. Upon binding to enhancers, certain proteins induce alterations in the physical conformation of DNA. These DNA conformations are called chromatin loops and help enhancers to get in proximity of the target gene. Thus, an interaction takes place between the enhancer- associated activator and the promoter region-associated transcription factor and RNA polymerase II. This interaction of different components can initiate the alteration of gene expression. Figure was adapted from Claringbould *et al.* b) Pie chart indicating the genetic elements that GWAS SNPs associated with autoimmune disease fall into. Figure adapted from Fahr *et al.*

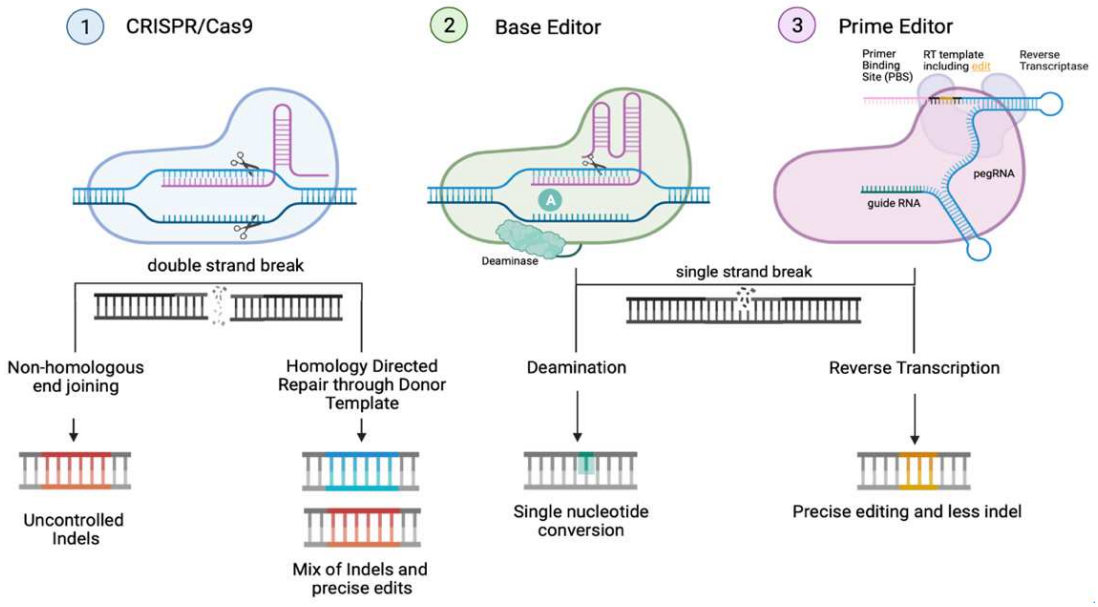
Precision Genome Editing in Primary T Cells

The Clustered regularly interspaced short palindromic repeats (CRISPR) system, originating in bacteria and archaea, defends against foreign DNA by incorporating viral DNA sequences into its genome. [15,16]Guided by CRISPR RNAs, Cas proteins cleave viral DNA upon reinfection. However, its lack of specificity underscores the need for careful design and optimization in genetic engineering and genome editing applications. The single-guide RNA (sgRNA) consists of a sequence complementary to the target DNA and a scaffold sequence that binds CRISPR associated protein 9 (Cas9), which introduces double-strand breaks (DSBs) at the target site. Repair mechanisms include non-homologous end joining (NHEJ), leading to unpredictable genetic changes, and homology-directed repair (HDR), allowing precise alterations. To improve specificity and efficiency, modified Cas9 variants like high-fidelity Cas9

and smaller Cas proteins such as Cas12 and Cas13 have been developed. [17] Additionally, optimized guide RNA (gRNA) design and computational tools for predicting off-target sites enhance target specificity in CRISPR applications. Furthermore, the development of base editing and prime editing technologies offers alternatives that enable precise modifications without inducing DSBs, reducing the reliance on error-prone repair pathways.

Base editing (BE) entails the substitution of one base with another. DNA base editors (BEs) are fusions between a catalytically impaired Cas nuclease that is only able to nick DNA and a deaminase base modification enzyme acting on single-stranded DNA (ssDNA). After binding to the target site on the DNA, base pairing between the guide RNA and the target DNA strand results in the displacement of a small segment of single-stranded DNA. The DNA bases within the single-stranded DNA bubble are modified by the enzyme deaminase. The catalytically impaired nuclease also creates a nick in the unedited DNA strand, causing the cells to repair the unedited strand using the edited strand as a template. There are two base editors: cytosine base editors (CBEs) convert C-G base pair into a T-A base pair, and adenine base editors (ABEs) convert A-T base pairs into G-C base pair. Around 58 % of human genetic variants are associated with disease come from point mutations or called single-nucleotide polymorphisms (SNP). Since BE can be used to edit at single-nucleotide level, it has the potential to be used for the correction of disease associated SNPs. One of the limitations of the base editing technique are bystander mutations of other bases beside the intended one within the editing window (Figure 4). These mutations have the potential to negatively impact gene function. [18]

One of the latest developments in gene engineering is Prime Editing (PE). PE is a versatile technique used to make specific changes in the genome of various organisms without the need for a double strand break (DSB). Compared to the previous method of base editing with PE, it is possible to overcome the restriction of only being able to introduce the four different transition mutations (C to T, A to G, T to C and G to A). Prime editing offers greater versatility compared to base editing. It enables a broader spectrum of base changes, encompassing all 12 possible substitution mutations, including the four changes that can be achieved with base editing methods. Additionally, prime editing enables the lengthening or shortening of gene sequences. Combined with recombinases, large DNA sequences of up to several kilobases in size can also be deleted or inserted at predetermined locations in the genome.[19] During the PE process a synthetic guide (sgRNA) carries the template encoding the desired edit (prime editing guide, epegRNA). The DNA is transcribed from an RNA template, which requires a reverse transcriptase. The reverse transcriptase for this method is fused to catalytically impaired SpCas9 nickase (H840A). The fusion protein is directed by an epegRNA to the target site. The nickase introduces a single strand break (SSB), creating a 3'-flap in the DNA strand displaced by the guide RNA. The nicked strand with the 3'-flap finds a primer binding site (PBS) on the epegRNA, that, upon elongation, provides the desired repair edit. Compared to CRISPR-Cas9, prime editing leads to far less indel byproducts in mammalian cell lines. with a ≤0.1 % probability for off-targets (Figure 4). [20,21]



	1) CRISPR/Cas9	2) Base Editor	3) Prime Editor
Off-target effects	Significant off-target effects	Less to no off-target effects	Less to no off-target effects
	Non-specific indels at DSB site	Bystander edits within a narrow window of 4-10 nucleotides	No bystander edits
	DNA donor template leads to plasmid integration in the genome	No DSB	No DSB
	Possible genome-wide-off-targets		High precision, because of three hybridization steps
Flexibility	Can introduce insertions, deletions, and all types of substitutions	Restricted to transition mutations (cannot introduce indel mutations)	Can introduce insertions, deletions, and all types of substitutions
	HDR pathway are restricted to the S and G2 phase of the cell cycle	Correction of larger genes >4kb	Less stringent PAM requirements
		Can introduce C > T, G > A, A > G, T > C and C > G substitutions only	Correction of larger genes > 4kb
		Editing applicable in dividing and terminally differentiated cell types	Editing applicable in dividing and terminally differentiated cell types

Figure 4: Schematic overview of three CRISPR editing methods. 1) The Cas9 enzyme that is led by a guide RNA (gRNA) to the target site and makes a double strand break (DSB). The DSB can lead to different repair pathways: non-homologous end joining (NHEJ) and homology-directed repair (HDR). The addition of a donor template increases the probability of HDR, which is the preferred method for precise editing as HDR leads to fewer uncontrolled indel events. 2) base editors require the guide-directed Cas enzyme to be fused with a deaminase. The Cas enzyme fused with deaminase can introduce a single nucleotide conversion, without the need of a DSB. Since this process does not require a DSB, repair pathways can be circumvented, and fewer off-targets occur. 3) Prime editors require prime editing guide RNAs (pegRNAs), that identify the target sequence. It contains the extended synthetic guide RNA (sgRNA) consisting of a primer binding site (PBS) and a reverse transcriptase template sequence (RT), which includes the desired edit. A Cas9 H840A, a catalytically impaired nickase, which introduces a single-strand break (SSB) and a M-MLV reverse transcriptase, that can synthesize DNA from a single-stranded RNA template are needed to introduce the desired edit. Figure adapted and content of table includes information from Kantor *et al.*

Engineering Highly Efficient and Versatile Prime Editing Constructs

Generally, Prime Editing (PE) requires the presence of multiple vectors in a cell, which may pose a potential limitation. This multiple-vector approach involves the use of up to three plasmids: one containing the prime editing enzyme (a fusion protein of Cas9 nickase and a reverse transcriptase), a prime editing guide RNA (pegRNA), and optionally, a third plasmid containing the RNA template for the HDR-mediated repair. In some cases, only a single vector is effectively delivered to the cell, while the other components may be missing. To overcome this obstacle and maximize the likelihood of edits, Adikusuma *et al.* (2021) designed an All-In-One prime editing plasmid (PEA). As the name suggests the plasmid contains all the PE elements, namely the prime editing guide RNA (pegRNA) sequence, the prime editing nuclease sequence, which is a sequence encoding for an engineered enzyme that consist of a reverse transcriptase (RT) domain fused with an RNA-targeting CRISPR-Cas9 nickase, as well as a fluorescent reporter, allowing for easier selection of the cells that successfully received the plasmid. In this work, the PEA was taken as a starting point and used for the design of an All-In-One (AIO) plasmid that further incorporates several optimized PE elements. For the All-In-One (AIO) plasmid used in this work, the optimized prime editing system "PEmax" was used. The origin of the PEmax system was the PE1 system, in which Moloney murine leukemia virus (M-MLV) RT was fused with the C-terminus of Cas9 (H840A) nickase. Through different mutations of the M-MLV RT a new PE system with an improved prime editing efficiency called PE2 was generated. Additionally, a new PE structure (PEmax) was constructed through engineering the PE2 protein, which uses a human codon-optimized RT. It has been shown in past studies that the structure and sequence length of the prime editing guide RNA (pegRNA) influences the efficiency of PE2.[22] The study indicated that a PBS length of 13-nt, and a RT template length between 10 and 16-nt is ideal for initial optimizations, it was further recommended to use a pegRNA 3' extension without the first base C.[23] During prime editing the 3’ and 5’ ends of the pegRNA are exposed and therefore not protected from exonucleolytic degradation within the cell.[25] To protect the 3-prime end, a stable pseudoknot can be incorporated. The pseudoknot evopreQ1 is one of the smallest naturally derived RNA structure motifs with a tertiary structure that was modified for better editing results and named tevopreQ1 (trimmed evopreQ, 37 nt long). To connect the PBS and RT, an additional linker, made of 8 nucleotides serving as a junction between the pegRNA and the 3' motif tevopreQ1 (trimmed evopreQ1), was designed. The linker was chosen using the pegRNA Linker Identification Tool (pegLIT) to ensure that it did not base-pair with the structure of the pegRNA.[26] For designing engineered pegRNA (epegRNA), the above criteria were considered.

It was discovered that the prime editing (PE) process can be partially restrained by DNA mismatch repair, leading to reduced editing efficiency. To counteract this, additionally constructs with a sequence encoding for the dominant negative MMR protein (MLH1dn) were designed.[24] The MLH1dn sequence results in a mutated MLH1 that can form complexes with the wild-type MLH1 protein and disrupt the normal DNA Mismatch Repair (MMR) function of the cell. The MLH1dn proteins can enhance prime edits by an average 7.7-fold and 2.0-fold compared to PE2 and PE3 systems.[24] To take advantage of this effect, all-in-one constructs containing an MLH1dn sequence (AIO-M) were developed to co-express together with the other PE elements and evaluate effects on PE efficiency in CD4+ primary T cells.

The piggyBac prime editing (PB-PE) method from Wolff *et al.* (2021) was used, to combine the piggyBac (PB) transposon system with the All-in-One (AIO) plasmid. In PB-PE, the AIO plasmid and a plasmid containing the sequence of a DNA PB transposase are introduced into the cell (Figure 5). [27] The transposase, a mobile genetic element, operates via a "cut-and-paste" mechanism, allowing it to move from one genomic location to another. PiggyBac commonly integrates into genomic TTAA sites, enhancing PE efficiency through integration at multiple sites. The PB-PE system leverages the large integration capacity of the piggyBac transposon system, allowing the insertion of genetic constructs with varying sequence lengths. [28,29]This feature enables the integration of all PE components simultaneously into multiple locations in the genome of primary CD4+ T cells, facilitating the integration of foreign DNA into a selected genome.

PB-PE works by flanking the transgene with inverted repeats (IR) and integrating it into the desired gene segment using the cut-and-paste principle of the piggyBac transposase. Through multiple integrations of the transposon in the genome, regions that enhance or suppress the expression of certain genes can be identified, aiding in the understanding of gene regulatory networks.

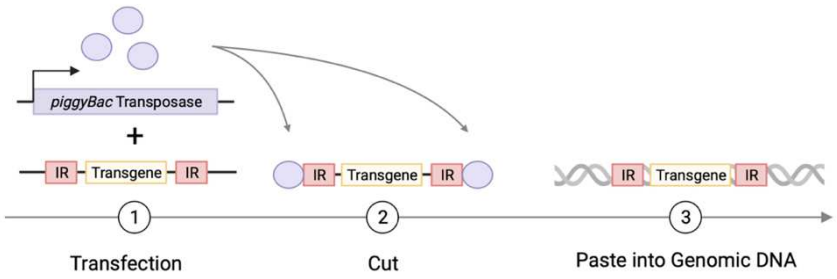


Figure 5: The mechanism of piggyBac transposition. 1) The plasmids containing the piggyBac transposase sequence and the plasmid with the transgene of interest flanked by the inverted repeats (IR) are delivered to a cell. There the transposase is expressed from the piggyBac expression vector. 2) The transposase recognizes the IR on the transgene vector plasmid and cuts the sequence. 3) As a result, the sequence of the transgene is integrated into the genomic DNA of the cell. Figure was adapted from Nelson JW *et al.* (2021).

Methods

Construction of Prime Editing Plasmids by Gibson Assembly

To obtain All-in-One (AIO) PE plasmids, the PEmax prime editor, with (AIO-M) or without (AIO) the dominant negative MMR protein (MLH1dn), under the control of the CAG promoter, was cloned into the backbone (Fragment 4). The plasmid construction process involved several steps (Figure 5) and Gibson assembly was used to assemble the fragments. See appendix The fragments and plasmid construction steps are described below. The detailed properties of the plasmids used for fragment generation of the different PE plasmids can be found under Appendix as Supplementary Table 1. The plasmids and the primers used for the Methods section of the work can be found in the Appendix as Supplementary Table 2.

Fragment 1, consisting of a CAG promotor (1712 bp), was derived by restriction digest from plasmid pDLI215. For the restriction digest the enzymes *MfeI* (NEB, #R3589) and *KpnI* (NEB, #R3142) were used and incubated at 37°C for 1 hour (Figure 5, step 1). The restriction digest reaction was prepared according to "Optimizing Restriction Endonuclease Reactions" instructions (NEB). The reaction was incubated at 37°C for 15 minutes. The resulting fragment was isolated using gel electrophoresis and purified with MinElute Gel Extraction Kit (Quiagen, #28604) according to manufacturer's instruction.

Fragment 2 measures 503 bp and contains the human U6 promoter (hU6). The fragment contains *Bbs1* sites, that were later used to introduce the engineered prime editing RNAs (epegRNAs) sequences into. It was obtained by PCR amplification using Q5® High-Fidelity DNA Polymerase (NEB, #M0491S) and primers 1175 and 1176 from plasmid pDLI301 (Figure 5, step 2). Details of Gibson assembly PCR reaction can be found in table 1.

Fragment 3, containing the Cas9 and reverse transcriptase (RT) sequence with a length of 6279 bp was obtained from pDLI142. First pDLI142 was modified by site-directed mutagenesis (SDM) to remove the *BbsI* and *BsmBI* sites using the QuikChange Multi Site-Directed Mutagenesis Kit (Agilent, #200514) according to manufacturer's instructions, using primers 1173 and 1174 respectively. The thermocycling reaction was followed by a restriction digest by adding 1 µl of *DpnI* restriction enzyme (10 U/µl) to the amplification reaction. The mixture was gently mixed by pipetting and then incubated for 1 hour to digest the parental (nonmutated) ds-DNA (Figure 5, step 1). Subsequently, the Cas9 and RT containing fragment was amplified through PCR using primers 1177 and 1178 (Figure 5, step 2).

Fragment 4 contains the Enhanced Green Fluorescent Protein (EGFP) selection marker, bovine growth hormone polyadenylation signal (bGH poly A), Ampicillin (Amp) selection marker and an origin of replication (ORI). This fragment was amplified from plasmid PEA1-GFP and the *BbsI* sites were mutated by using the Multi Site-Directed Mutagenesis Kit (Agilent, #200514). For the backbone of the AIO and AIO-M plasmids, for which fragment 4 was used, the second nicking gRNA sequence and the adeno-associated virus serotype 2 inverted terminal repeat (AAV2 ITR) were removed from

the plasmid PEA1-GFP using the *SbfI* restriction enzyme The restriction digest was performed using *SbfI-HF* (NEB, #R3642) according to manufacturer's instructions (Figure 5, step 1). Subsequently the resulting sticky ends were ligated with T4 ligase (NEB, M0202) for 10 minutes at RT. Afterwards the T4 ligase was heat inactivated at 65°C for 10 minutes and the resulting plasmid was transformed into NEB® 5-alpha Competent *E. coli* (High Efficiency) cells (NEB, C2987I) following manufacturer's instructions. Multi-Site-Directed Mutagenesis Kit (Agilent, #200514) was used according to manufacturer's instructions, using primers 1171 and 1172 to remove the *BsmBI* sites carrying point mutations (Figure 5, step 1). Finally, the fragment 4 was amplified through PCR by using the primers 1179 and 1180 and had a length of 3767 bp (Figure 5, step 2). Details of all PCR reactions used for amplification of the Gibson assembly fragments can be found in table 1. All fragments generated were purified prior to assembly using DNA Clean & Concentrator-5 kit (Zymo Research, #D4013) by following the manufacturers protocol.

To obtain the fragment 5 for AIO-M, the plasmid pDLI214 was used to amplify the fragment by PCR using the primers 1183 and 1184. Details of all PCR reactions can be found in table 1.

The resulting fragments 1-4 and 1-5 were used in two Gibson assemblies to obtain plasmids AIO and AIO-M. Each fragment concentration was measured via Nanodrop and assembled using the NEBuilder® HiFi DNA Assembly kit (NEB, #E2621) according to manufacturer's instructions. at a 1:1 ratio, 0,05 pmol per fragment (Figure 5, step 3). The reaction was incubated at 50°C for 25 min. Subsequently the ligation reaction was transformed using NEB® stable Competent *E. coli* (NEB, #C3040), according to manufacturer's instructions. From the isolated clones the plasmid was obtained by using the QIAwave Plasmid Miniprep Kit (QIAGEN, #27204)and following manufacturer's protocol.

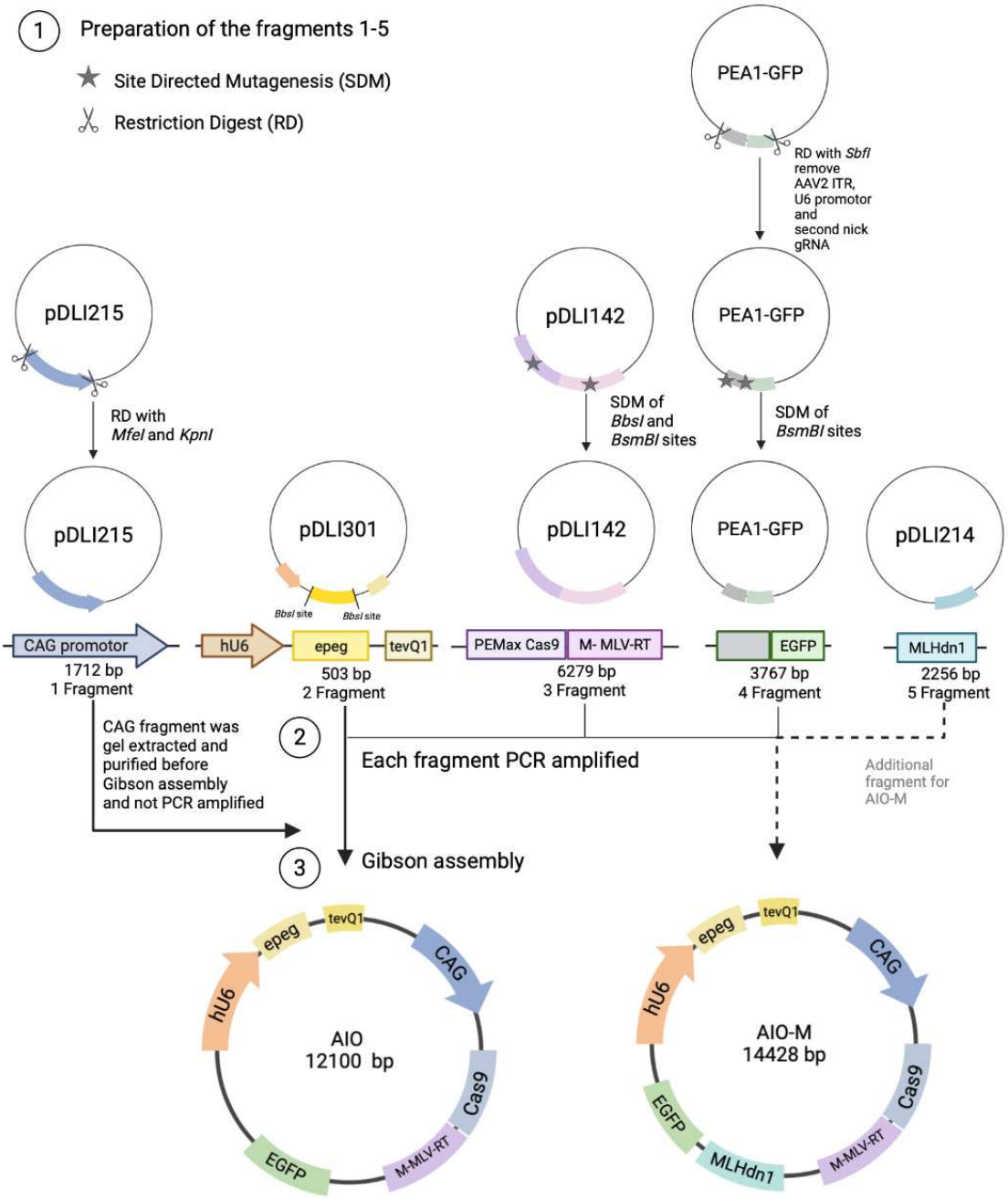


Figure 6: Schematic overview of the cloning strategy for plasmids AIO and AIO-M. 1) 4 fragments were prepared for Gibson assembly of plasmid AIO and 5 fragments for plasmid AIO-M 2) Fragment 1 was obtained from the plasmid pDLI215 by restriction digest using *MfeI* and *KpnI*. Fragment 2 was PCR amplified from plasmid pDLI301, containing the human U6 promoter, Bbs1 sites to insert the engineered prime editing RNAs (epegRNAs) and the sequence encoding tevopreQ1. For fragment 3, Cas9 and reverse transcriptase (RT) sequence originate from plasmid pDLI142. First, the RT sequence was modified by silencing the restriction sites *BsmBI* and *BbsI* by SDM. Afterwards the fragment 3 was PCR amplified and each Fragment 4 was amplified from plasmid PEA1-GFP and the Bbs1 sites were silenced it contained the Enhanced Green Fluorescent Protein (EGFP) selection marker, bovine growth hormone polyadenylation signal (bGH poly A), Ampicillin (Amp) selection marker and an origin of replication (ori). The backbone used for the plasmids AIO and AIO-M was taken from the plasmid PEA1-GFP, were the second nicking gRNA and the adeno-associated virus serotype 2 inverted terminal repeat (AAV2 ITR) were removed using the *SbfI* restriction enzyme to create sticky ends. Fragment 5 was amplified from plasmid pDLI214 caring the sequence for the dominant negative MMR protein (MLH1dn) and was only used for the AIO-M plasmid. 3) The plasmids were finally assembled by Gibson assembly.

Table 1: PCR settings for amplification of the various fragments for Gibson assembly with different lengths

			Fragments			
Step	Temperature	Time	F2	F3	F4	F5
Initial Denaturation	98°C	30 s				
20 Cycles	98°C	10 s				
Annealing temperature	*50–72°C depending on fragment	30 s	61°C	64°C	60°C	60°C
Elongation	72°C	20–30 seconds/kb	30 s	3 min	40 s	1 min
Final Extension	72°C	2 minutes				
Hold	4°C	forever				

Constructing the PiggyBac Prime Editing Plasmid with Inverted Terminal Repeats

To enable PB insertion of the prime editing cassettes, All-In-One plasmids AIO and AIO-M were modified to contain inverted repeats (IRs) flanking the prime editing cassette. One 67 bp long IR was inserted downstream of the bGH poly A signal and the other IR upstream of hU6 promoter with a length of 37bp.

First, each of the single stranded oligonucleotide pairs of the IR were phosphorylated and annealed. For each reaction 100 μM of each pair and 0,5 μl T4 PNK (NEB, #M022) was incubated at 37°C for 30 minutes, followed by 95°C for 5 min. The reaction was cooled down slowly to RT. After annealing, the oligonucleotides were diluted 1:50 in distilled water (dH_2O). The plasmids AIO and AIO-M were digested upstream of the hU6 promotor by the RE *PciI* (NEB, #R0655) and dephosphorylated in a single reaction with the following thermocycler conditions 60 min 37°C, 5 min 80°C, hold at 4°C in the presence of rSAP (NEB #M0371), according to manufacturer's instructions (Figure 6). The digested AIO and AIO-M backbones were PCR purified using the QIAquick PCR Purification Kit (50) (Qiagen, #28104) and eluted in 30 μl dH2O. The annealed left IR oligos and backbone were ligated using the T4 ligase (NEB, #M0202) at a vector to insert ratio of 1:20. Subsequently the ligation reaction was transformed into NEB® stable competent *E. coli* (NEB, #C3040I). The resulting colonies were screened for correct insertions by colony PCR using insert-specific primers (1247 and 1248) and using the PCR protocol for OneTaq® Quick-Load 2X Master Mix with Standard Buffer (NEB, #M0486). Each of the primers had a concentration of 10 pmol/ μl. The colony was picked with a pipette tip and diluted in 10 μl dH_2O. 1 μl was used for the PCR reaction. the annealing temperature was set to 47°C annealing temperature and the insert size was 110 bp.

Subsequently, this process was repeated to insert the right IR, using RE *XmaI* (NEB, #R0180) instead of *PciI* (NEB, #R0655) and using primers 1246 and 277 for colony PCR, the annealing temperature was set to 52°C and the insert size was 982 bp. To verify that the inserts did not contain mutations, they were analyzed through Sanger sequencing and correct insertions were selected, resulting in plasmids PB (PiggyBac prime editing plasmid) and PB-M (PiggyBac prime editing plasmid with MLH1dn).

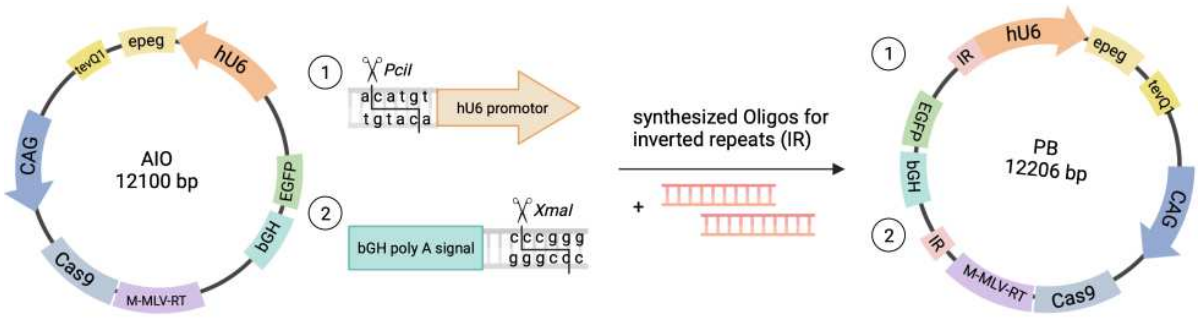


Figure 7: Example Schematic overview of the construction process for the plasmid PB. The previously prepared plasmid AIO was used for the process. First, the synthesized single stranded oligos for the IR (inverted repeats) were ligated so that they were double stranded annealed. 1) The AIO plasmid was cut with the *PciI* restriction enzyme upstream of the hU6 promotor and then first IR was ligated into the site. The plasmid was transformed into competent bacterial cells. (*E. coli*). 2) Then the plasmid with one integrated IR was cut with the restriction enzyme *XmaI* downstream of the bGH poly A signal sequence. Then the resulting PB plasmid was transformed into competent cells.

Design and Assembly of the Engineered Prime Editing Guide RNAs

To investigate the capability to prime edit in primary T cells using the constructs described above, engineered prime editing guideRNAs (epegRNAs) were designed using the computational tools “pegLIT” (pegRNA Linker Identification Tool) and “primeDesign”.[30,31] With the help of the primeDesign the Spacer, RTT and PBS parts of the epegRNA were designed. With the pegLIT tool nine different linker sequences were created. For each epegRNA, a complementary oligonucleotide pair was designed for the extension and spacer sequences. The scaffold was kept constant for each of the constructs. All fragments (extension, spacer, and scaffold) were designed with overlapping overhangs to each other to allow for assembly of the epegRNAs and ordered through Merck. The editing regions were on exon 3 and 4 of the gene CD4-201 (Figure 8 and Table 2).

Table 2 : Targeting sites on gene CD4-201 with corresponding edits.

Targets	**Edit**	**Position**	**#Exon on gene CD4-201**
C, cysteine	C41*	41	3
K, lysine	(-CC)AG, deletion	26	3
R, arginine	GA(+AC)C, insertion	81	4

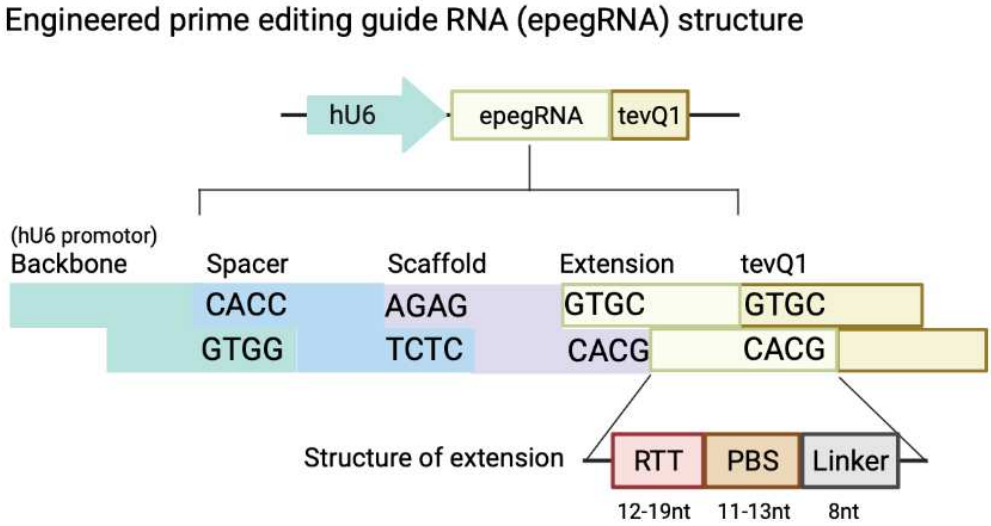


Figure 8: The structure of the engineered prime editing guide RNA (epegRNA). The epegRNA consists of a spacer, a scaffold, and the extension with complementary overhangs. The extension part of the epegRNA consists of a reverse transcriptase template (RTT) that was 12 to 19 nucleotides long, a primer binding site (PBS) that was 11 to 13 nucleotides long and a linker of 8 nucleotides.

First, each of the single stranded oligonucleotide pairs was phosphorylated and annealed. For each reaction 100 pmol of each oligo and 0,5 µl T4 PNK was incubated at 37°C for 30 minutes followed by 95°C for 5 min and slowly cooled down to 25°C. After annealing the oligonucleotides were diluted 1:250 in dH_2O. The AIO, AIO-M, PB (piggyBac prime) and PB-M (piggyBac prime with MLH cassette) plasmids were digested with the restriction enzyme (RE) *BbsI-HF* and a Shrimp alkaline phosphatase (NEB, #M0371, rSAP) was added to dephosphorylate the sticky ends. Finally, the linearized plasmids were assembled with the respective epegRNAs using T4 ligase (NEB, M0202) at a vector to insert ration of 1:20 in a separate step. A total of 36 constructs were produced carrying targeting prime editing guide RNAs. Each of the constructs was transformed into 20 µl of Stable competent *E. coli* (NEB, #C3040.) with 1 µl of the ligation reaction. The resulting colonies were screened by colony PCR for epegRNA insert using previously designed insert-specific primers 717 and 1243 using the PCR protocol for OneTaq® Quick-Load 2X Master Mix with Standard Buffer (NEB, #M0486). To verify correct assembly, selected colonies were sequenced using primers 1045 and 1243 and analyzed. A correct assembly of each construct was chosen and stocked for further use.

Isolating PBMCs from Buffy Coats

Buffy coats were collected in 50 mL Falcon tubes and diluted 1:1 with RPMI (+1mM EDTA) and mixed by inverting. Two Sepmate tubes, were carefully filled through the central hole with 15 ml Ficoll density gradient medium (Stemcell, #15460). Next, 17 ml of diluted blood was carefully pipetted down the side of the tube and centrifuged at 1200 g for 10 minutes at RT. After centrifugation, the top layer was poured off carefully. Subsequently the residual red blood cells were removed using the Red Blood Cell Lysis Buffer (Roche, #11814389001), following the manufacturer's protocol. The PBMC layer was aspirated with a Pasteur's pipette and transferred to a clean 50ml

tube. The PBMCs were washed by diluting up to 50 ml with FACS buffer (PBS + 2 % Fetal Bovine Serum FCS Thermo, #16000044 + 1mM EDTA) and spun at 350g, acceleration at 6 and brake at 6 for 5 min at RT. The supernatant was carefully discarded, ensuring not to disturb the cell pellet. A second wash-step was performed at 400 g for 5 min at RT. The supernatant was removed carefully, and the pellet was resuspended in 20ml FACS buffer. To determine the cell number, 10 μl of the cell suspension was mixed with 10 μl of Trypan blue and 10 μl of the mixture was applied to a counting slide and counted with the "Countess 2 Automated Cell Counter". Afterwards 20 μl of the resuspended PBMCs was taken and diluted in 180 μl (1:10) in FACS buffer. The 20ml PBMCs in FACS buffer were spun at 400 g for 5 min at RT. Supernatant was discarded and FACS buffer was added to obtain a final cell concentration of $50x10^6$ cells/ml.

Isolation of Human Naïve CD4+ T Cell Isolation

Naïve CD4+ T Cells were isolated from PBMCs using the EasySep TM Human Naïve CD4+ T Cell Isolation Kit II (STEMCELL Technologies). 50x106 cells/ml in FACS buffer was mixed with 50 μl of EasySep TM Human Naïve CD4+ T Cell Isolation Cocktail II per ml of sample and incubated for 5 min at RT, protocol was followed according to manufacturer's instructions with following modifications. After washing the cells once, a second washing step was added by placing the tubes in the EasySep magnet and incubating for 3 min at RT. Cells from each donor were combined in different tubes and yield was measured by counting the cells. Cells were spun at 400g for 5 min at RT and resuspended in Freeze medium (45 % complete RPMI with IL-2 added, 45 % of FBS and 10 % of DMSO) at 8x106 cells. Rapidly 1 ml of cell suspension was transferred to a cryogenic vial and placed into an isopropanol freezing container (Nalgene® Mr. Frosty) and transferred in a -80°C freezer.

Thawing, Culturing and Activating Naïve CD4+ T Cells

Isolated T cells were cultured in RPMI (RPMI 1640 Medium Thermo #11875093, L-Glutamin, HEPES, #52400025) supplemented with Sodium pyruvate (1mM, Thermo, #11360088), Fetal Bovine Serum (FBS, Thermo, #16000044 10 %, heat inactivated for 30 min at 56°C), MEM Non-Essential Amino Acids Solution (1 %, Thermo, #11140035), Penicillin-Streptomycin (1 %, Thermo, #15070063), 2-Mercaptoethanol (55μM, Thermo, 21985023) and freshly added IL-2 (1 μl per 1 ml medium). To thaw frozen T cells, a vial of frozen primary T cells was thawed quickly in a water bath at 37°C, until a small ice clump was still visible. The cells were added to a 15ml Falcon tube with 6 ml of RPMI and centrifuged at 400g for 5 min at RT. The supernatant was removed, and the cells were washed two times with 7 ml of PBS by slowly pipetting up and down. Cells were resuspended in complete RPMI (RPMI with added supplements as described above) at 1x10^6/ml and 1 mL was added to a 24-well (ThermoFisher) plate well (1x10^6 cells/1ml RPMI well). The cells were incubated for 24h in a humidified 37 °C/5 % CO2 incubator. 1 day post thawing, cells were activated with (1:1 cells:beads ratio) CD3/CD28 Dynabeads (ThermoFisher, #11131D), following manufacturer's instructions.

Electroporation of Activated Naïve CD4+ T Cells

Prior to electroporation, the concentration and quality of the plasmids used was checked by Nanodrop for a A260/280 ratio of at least 1,8. CD3/C28 Dynabeads were removed from activated cells by transferring the cells to an Eppendorf tube and placing them on a magnet. Supernatant, containing the cells, was transferred to a fresh Eppendorf tube. P3 primary 4D-Nucleofector™ Solution (Lonza, #V4XP-3032) was equilibrated to room temperature for 15 minutes prior to electroporation. A 16-well Nucleocuvette Strip (Lonza, #V4XP-3032) was pre-chilled on ice. An aliquot of 2,5ml culture medium was pre-warmed to 37 °C. A 24-well cell culture plate was prepared with 1 ml medium per well and placed in a humidified 37 °C/5 % CO2 incubator. 1μg (in 2 μl) of each plasmid and 0,4 ug (or 1μg, 0,75μg and 0,5 μg) of pmaxGFP Vector (0,4μg/2μl) were pipetted in separate 1,5 ml tubes.

Subsequently 25k cells per (for one Nucleocuvette™ Strip well) reaction were centrifuged at 200xg rpm for 10 minutes at room temperature. Next the cell pellet, resuspended in 20 μl P3 primary 4D-Nucleofector™ Solution, was added to the tubes containing the DNA and immediately transferred to the Nucleofector 16-well stripe. The Nucleofector stripe was gently taped to make sure the sample covers the bottom of the cuvette. Afterwards the Nucleofector strip was placed into the retainer of the 4D-NucleofectorTM X Unit (Lonza, AAF-1003X). Post-Nucleofection the Nucleovette strip was removed from the electroporator, and the cells were resuspended with 80 μl pre-warmed medium. The Nucleovette strip was washed with cell culture medium to remove residual cells from the strip and the additional volume was added to the same 24-well cell culture well. Immediately post electroporation, cells were reactivated with reactivation medium, containing 1 ml of RPMI and Dynabeads with a ratio of 2:1 cells:bead or 12,5 μl of ImmunoCult. 1 ml reactivation medium per sample was added to the 24-well cell culture well and placed in a humidified 37 °C/5 % CO2 incubator. After 6 hours half of the medium of the cells were removed and replaced by fresh one. It is important to do the removal step very carefully and not to disturb the cells. Cells were analyzed 48h, 72h and 9 days post electroporation using flow cytometry.

Staining of Cells and FACS Analysis

1-2x10^5 cells were centrifuged at 400g for 5 min. The supernatant was discarded, and the cells were washed with 1 ml PBS. After washing the cell pellet was resuspended in 50 μl of CD4-BV650 antibody (Biolegend, 100469) 1:100 diluted in PBS and incubated at RT for 20 min protected from light. Next, cells were washed two times with 1ml of FACS buffer (PBS + 2 %BSA +2.5mM EDTA). The cell suspension was passed through the filter of a FACS tube (Falcon® Cell Strainers, 352235). 1 drop of DRAQ7 viability dye (Biostatus, DR71000) was added to the samples and the tubes were kept on ice until analysis. The cells were analysed by using the BD FACSymphony™ A3 Cell Analyzer. For the gating strategy first, the live cells were gated to separate the live cells from cellular debris and dead cells (FSC-A against SSC-A), then the single cells were gated further (SSC-A against SSC-H), the single cells were gated further to get separate the singlets from the doublets (FSC-A against FSC-H). the next gate was set for the GFP positive live cells (FSC-A against DRAQ7), the last gate was set to distinguish CD4 positive cells from the CD4 negative ones (SSC-A against CD4)

Cell Sorting and Extraction of gDNA

The cells were sorted directly in 100 μl lysis buffer using the BD FACSAria™ Fusion Flow Cytometer, as a viability dye DRAQ7 (Biostatus, DR71000) was used. The gate was set to 405/450/50-A vs 488/530/30-A to sort for GFP positive cells. After sorting the cells in lysis buffer (from Monarch® Genomic DNA Purification Kit, NEB, #T3010S/L directly the gDNA was extracted according to the protocol for “Extraction and Purification of Genomic DNA from Cells” (NEB, #T3010).

Two-Step Library Preparation for Illumina Sequencing

PCR1 Reaction

In the initial PCR reaction, the amplification of the specific region of interest was carried out while concurrently adding adapter sequences to facilitate binding to the NGS platform. Staggered primers (Supplementary Table 2) containing the necessary adapter sequences were employed. Due to a low yield in gDNA template, this PCR was repeated using the product of the first PCR as template. The DNA concentration was measured using Qubit.

PCR2

In a second PCR reaction (PCR2), indexing barcodes were incorporated using primers that included both the sequencing primer and a distinctive barcode or index sequence unique to each sample. The index sequences play a critical role in multiplexing, as they enable the differentiation and segregation of sequence data from various samples within the same sequencing run. PCR2 amplified the DNA fragments generated in PCR1, appending the index sequences to each fragment, resulting in the creation of uniquely barcoded libraries for each sample (Supplementary Table 2).

Two-Sided Size Selection with Magnetic Beads

A two-sided selection was performed using AMPure XP Beckman Coulter magnetic beads to purify the sequencing library based on size after PCR2. This procedure facilitated the retention of desired fragments while excluding unwanted ones. AMPure XP SPRI beads from Beckman Coulter were employed at concentrations of 0.5 x and 1x beads/sample. The individual samples were processed separately. The usage of AMPure XP SPRI beads (Beckman Coulter, #A63881) from following manufacturer’s instructions, with minor adjustments made for the two-sided selection. For the adjustments dH2O was added to each sample to reach a total volume of 100 μl after which 50 μl of AMPure XP beads was added, pipette mixed and incubated for 5 minutes at RT, with intermittent mixing. Following this, the sample was placed on a magnet, specifically on the high-volume side, once the suspension cleared, the supernatant was transferred to a new tube. The 1x selection followed on the supernatant, employing 75,1 μl of 1x beads, with the beads themselves retained this time. Post-selection, the beads underwent two washes using freshly prepared 80% ethanol. ethanol was removed and the beads were air-dried for a maximum of 2 minutes. After air-drying, the beads were eluted in 30 μl of dH2O and placed on the

magnet at the low volume position. The supernatant was transferred to a new tube, and the DNA concentration was measured using a Qubit fluorometer.

Quality Control of the prepared sequencing library

The Agilent Bioanalyzer 2100 and High Sensitivity DNA Analysis kit (Agilent #5067-4626) was used to measure the size distribution of the sequencing library prior to sequencing. Samples were prepared following manufacturer instructions. Following analysis samples AIO1-M (sample 1), AIO4-M (sample 2) and AIO8-M (sample 3) were pooled together. The sample for which the cells were transfected with AIO5-M plasmid was left out due to quality issues with the sample.

Illumina Sequencing

The prepared library was submitted to the EMBL Genecore facility. A MiSeq-MICRO 150bp PE sequencing kit was used. Samples were submitted as multiplexed with dual barcodes (8bp in length), with a PhiX spike of 15 %. The average fragment length was 290 bp, and 150 cycles were run on either side. GRCh38/hg38 (Human) was used as reference genome.

NGS Data Analysis

The NGS data analysis in this study utilized CRISPResso2, a bioinformatics tool developed by the Pinello lab. The parameters used in this study included read 1 and read 2 (r1 and r2) provided as FASTQ files. The amplicon sequence (-a or --amplicon_seq) used in the experiment varied among the three samples (sample 1, sample 2, and sample 3). The minimum average quality score (-q or --min_average_read_quality) was set to 30 per read. This quality score adjustment effectively filtered out reads with quality scores below 30, reducing the likelihood of false-positive calls and ensuring the reliability of the considered base calls. Additionally, this approach maintained high-confidence data and upheld data integrity. This was followed by the spacer sgRNA sequence in RNA 5' to 3' order, excluding the PAM sequence (--prime_editing_pegRNA_spacer_seq), which is crucial for prime editing and differed between samples. The extension sequence (--prime_editing_pegRNA_extension_seq) was structured in RNA 5' to 3' order, beginning with the Reverse Transcriptase template (RTT) and followed by the Primer binding site (PBS). A crucial aspect for precise analysis was the quantification window (-qwc or --quantification_window_coordinates), which was specified by defining one base pair position to another of the intended edit. The first base in the sequence was designated as 0. The quantification window allowed for a detailed examination of genetic changes and provided a quantitative measure of the editing outcomes within that defined region. Additionally, the specific prime-edited reference (--prime_editing_override_prime_edited_ref_seq) was also input. All parameters were individually chosen for the 3 different samples analyzed (see table 5).

Table 3 Parameters used for CRISPResso2

	Sample 1, AIO1-M	Sample 2, AIO4-M	Sample 3, AIO8-M
r1	/g/steinmetz/project/Gandhi_MA/Master_part2/Data/Seq_results/GHP6J_VG_CD4AmpliSeq_23s002530-1-1_Hughes_lane11_1_sequence.txt.gz\	/g/steinmetz/project/Gandhi_MA/Master part2/Dats/5eg_results/GHP6J_VG_CD4AmpliSeq_23s002530-1-1_Hughes_lane12_1_sequence.txt.gz\	/g/steinmetz/projectGandhi_MA/Master_part2/Data/5eq_results/GHP5J_VG_CD4AmpliSeq_23s002530-1-1_Hughes_lane13_1_sequence.txt.gz\
r2	/g/steinmetz/project/Gandhi_MA/Master_part2/Dats/Seg_results/GHP6J_VG_CD4AmpliSeq_23s002530-1-1_Hughes_lane11_2_sequence.txt.gz\	/g/steinmetz/project/Gandhi_MA/Master _part2/Data/Seq_results/GHP6J_VG_CD4AmpliSeq_23s002530-1-1_Hughes_lane12_2_sequence.txt.gz\	/g/steinmetz/project/Gandhi_MA/Master _part2/Data/Seg_results/GHP6J_VG_CD4AmpliSeq_236002530-1-1_Hughes_lane13_2_sequence.txt.gz\
a	ATCITTATCTGGTTGGAGTTTTTCCAGTGGAATTGTATGCTCTTCTTCTGGGAAGCTGTACAGGTCAGTTCCACTGTATCCCCTTTTTTGCCCAGCACCACTTTCTTTCCCTGAGTGGCTGCTGG	TTTTTTGCCCAGCACCACTTTCTTTCCCTGAGTGGCTGCTGGGAGGAGCGCTAAGTGGAAAAGAMAGGAGTCAGGTCTCAATGTCGCCGCTGTCCCCCT	CCAAGCTGAATGATCGCGCTGACTCAAGAAGAAGCCTTTGGGACCAAGGAAACTTTCCCCTGATCATCAAGAATCITAAGATAGAAGACTCAGATACTTACATC
prime_editing_override_prime_edited_ref_seq	ATCTTTATCTGGTTGGAGTTTTTCCAGTGGAATTGTATGCTCTTCTTCGGGAAGCGTCTAGGTCAGTTCCACTGTATCCCCTTTTTTGCCCAGCACCACTTTCTTTCCCTGAGTGGCTGCTGG	TTTTTTGCCCAGCACCACTTTCTTTCCCTGAGTGGCTGCTGAGGAGCGCTAAGTGGAAAAGAAAGGAGTCAGGTCTCAATGTCGCCGCTGTCCCCCT	CAAGCTGAATGATCGCGCTGACTCAAGAAGAAGCCTTTGGGAACCCAAGGAAACTTTCCCCTGATCATCAAGAATCTTAAGATAGAAGACTCAGATACTTACATC
n	sample1_CD4_1_QW	sample2_CD4_4_QW	sample3_CD4_8_QW
q	30	30	30
prime_editing_pegRNA_spacer_seq	CTTCTTCTGGGMGCTGTAC	TTCCCTGAGTGGCTGCTGGG	AGAGAAGCCTTTGGGACCA
prime_editing_pegRNA_extension_seq	ACTGACCTAGACAGCTICCCAGAA	CACTTAGCGCTCCTCAGCAGCCACTC	AGTTTCCTTGGGTTCCCAGGCTT
quantification_window_coordinates	59-60	40-42	42-42
debug			
O	/g/steinmetz/project/Gandhi_MA/Master_part2/Output/Crispresso_Output	/g/steinmetz/project/Gandhi_MA/Master_part2/Output/Crispresso_Output	/g/steinmetz/project/Gandhi_MA/Master_part2/Outout/Crispresso_Output

Results

For previous studies the components used for PE were delivered to the cell on multiple plasmids. Usually this multiple-vector approach involves the use of up to three plasmids: one containing the prime editing enzyme (a fusion protein of Cas9 nickase and a reverse transcriptase), a prime editing guide RNA (pegRNA), and optionally, a third plasmid containing the RNA template for the HDR-mediated repair. However, transfection can result in the loss of a plasmid, which leads to a loss of editing events. Thus, a single plasmid could increase the transfection efficiency and potentially enhance editing efficiency. Recent advancements allowed to also improve the editing efficiency, such as combining prime editing with a piggyBac transposon system. As editing efficiency heavily depends on transfection efficiency, transfection by electroporation and lipofection were compared. The best performing protocol, electroporation, was further optimized and applied for genome editing in CD4+ T cells.

Generation of All-In-One Prime Editing Constructs

The experimental approach aimed to develop two novel plasmids for efficient prime editing in CD4+ T cells: an All-In-One plasmid (AIO) containing all PE3max components for single-step transfection, and a PiggyBac Prime Editing Plasmid (PB) combining the PE3max system with the PiggyBac transposon system to enable stable integration and enhance editing efficiency. Following methodologies outlined by Adikusuma *et al.* (2021), four distinct plasmid backbones were constructed: AIO, AIO-M with an additional MLH1dn sequence, PB, and PB-M with MLH1dn sequence. Upon expression of the domiant negative MMR protei (MLHdn1) sequnce a mutated MLH1 forms a complex with the wild-type MLH1 protein and enhances PE by disrupting the DNA Mismatch Repair (MMR). The high cargo capacity and multiple integration capability of piggyBac at TTAA sites across the genome potentially could facilitate stable transfection. [32] Integration of all prime editing components can potentially be achieved as a result. Fusion of the PE3max cassette with piggyBac generated the PB plasmid, further enhancing prime editing efficiency.

All constructs were designed to be under the control of the hU6 promoter for the expression of engineered prime editing guide RNAs (epegRNA) and the CAG promoter for the expression of Cas9 (H840), reverse transcriptase (M-MLV-RT), and the selection marker EGFP. The epegRNA inserts for a pilot experiment to assess editing efficacy were designed to target the CD4 co-receptor gene, disturbing CD4 expression and thus allowing for easy analysis of successful editing. To verify correct assembly of the plasmids, six single colonies were selected after Gibson assembly of the AIO plasmid (plasmid 545) and subsequent transformation. Isolated plasmid from these colonies was subjected to digestion with a restriction enzyme that cleaves three times. Prior to the restriction digest a virtual digest was simulated, indicating the expected presence of three differently sized fragments of 9659 bp, 2959 bp, and 1810 bp, which would indicate successful assembly. In Figure 9 a gel illustrates successful assembly in colonies GA1, GA2, GA3, GA5, and GA6, while GA4 did not assemble correctly, displaying a single band at >4000 bp. The colonies are here abbreviated as GA1-GA6, the plasmid resulting from those colonies was the AIO plasmid.

Subsequently, the AIO plasmids of which the restriction digest indicated successful assembly underwent sequencing by Plasmidsaurus for whole-genome sequencing with Oxford Nanopore Technology (ONT) to confirm correct assembly and identify any potential mutations (Figure 9b and c).

Following the successful assembly of the AIO, AIO-M, PB, and PB-M backbone plasmids, nine distinct engineerd prime editing guide RNAs (epegRNA) constructs were generated, yielding a total of 9 constructs for each plasmid (Figure 9). Sanger sequencing of the epegRNA region in the assembled plasmid was used to confirm no mutations were introduced. The sequencing results were aligned with the designed reference sequence. No mutations were detected in the epegRNA regions (Figure 10).

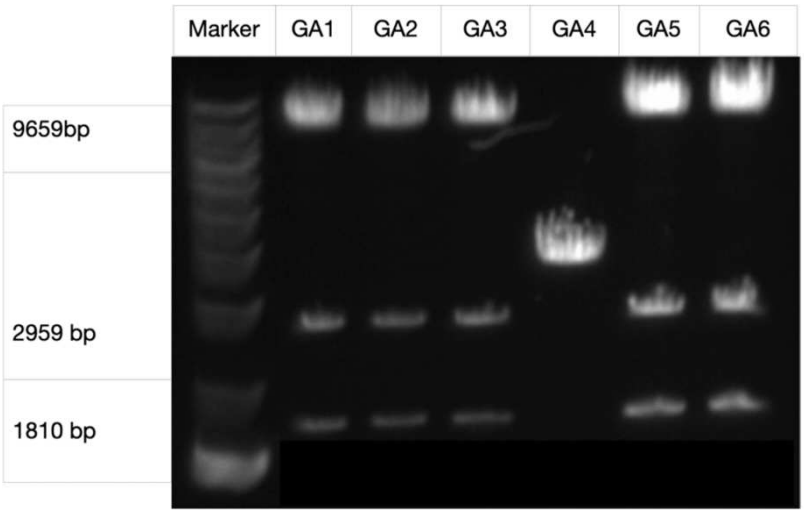


Figure 9: Shows a 0,8% Agarose gel, plasmids from 6 different colonies digested with *SbfI-HF*. Lane 1 contains a DNA ladder. Lanes GA1 to GA6 indicate the 6 different colonies. The expected band sizes were 9659 bp, 2959 bp and 1810 bp. All colonies, except for GA4, show the correct fragment sizes, indicative of successful assembly.

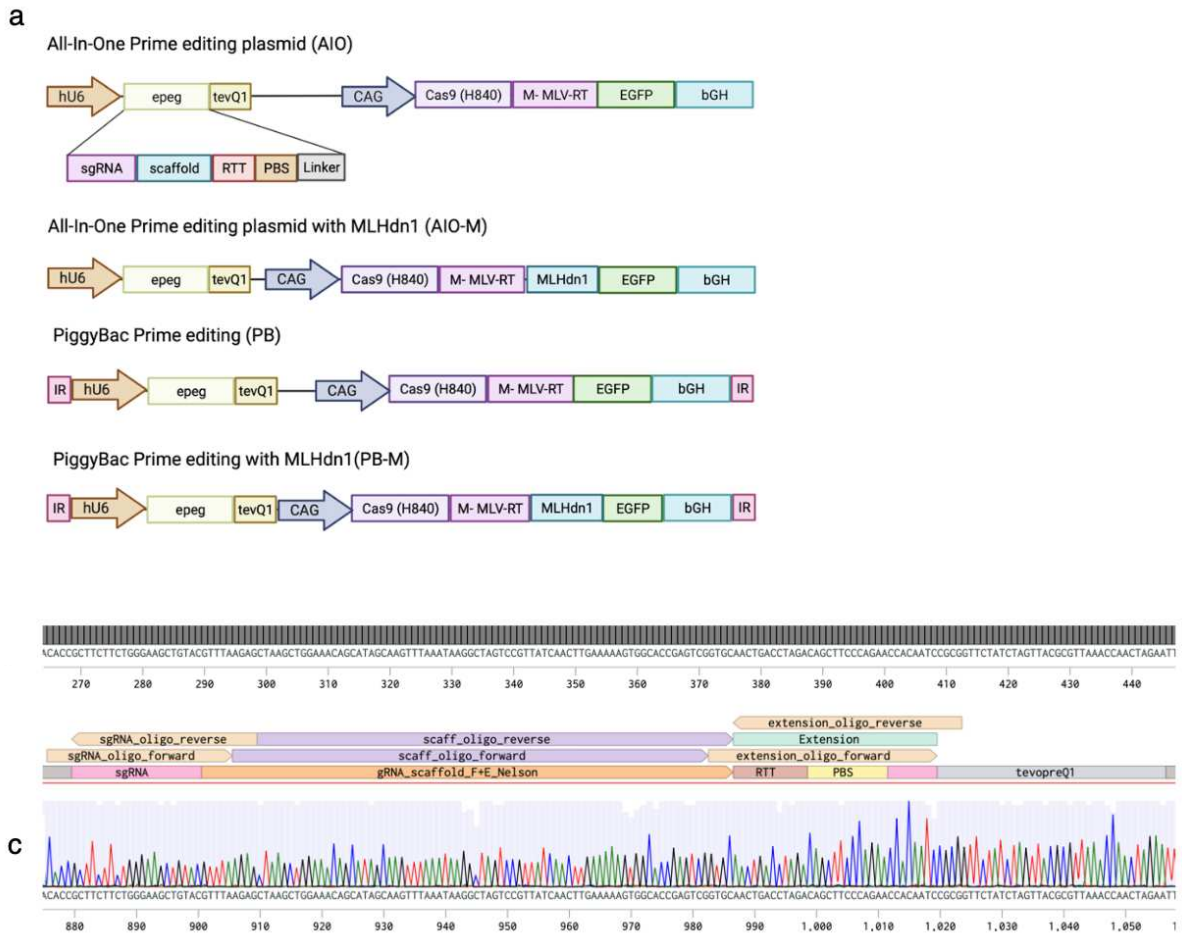


Figure 10: a) Schematic representation of the prime editing constructs. From top to bottom: the All-In-One Prime Editing Plasmid (AIO), comprising the hU6 promoter, the inserted engineered prime editing guide RNA (epegRNA), the structural motif sequence tevQ1, CAG promoter, Cas9 (H840), the reverse transcriptase sequence (M-MLV- RT), the EGFP selection marker, and the bGH polyA signal. The epegRNA consists of a synthetic guide RNA (sgRNA), a scaffold, reverse transcriptase template (RTT), the primer binding site (PBS), and the linker. The All-In-One Prime Editing Plasmid with MLHdn1 (AIO-M) consists of the same sequences as AIO, with an additional sequence for the dominant negative MMR protein (MLH1dn). The piggyBac prime editing plasmid (PB) is the AIO plasmid with inverted repeats (IR). Another plasmid is the PB plasmid with the MLHdn1 sequence (PB-M). b) representative example of the sequencing results from one of the constructs (AIO). The results are presented as chromatographs used to verify correct epegRNA inserts and provide information about sequencing quality. No mutations are detected.

Successful Transfection of CD4+ T Cell by Electroporation with All-In-One Plasmids

Following the successful assembly of all four plasmids (AIO, AIO-M, PB, and PB-M), each containing epegRNAs, the ability to induce prime editing in primary T cells through a single-step electroporation process was tested. Activated CD4+ T cells were electroporated with AIO plasmids to initiate the experiment.

All constructs were labeled with an EGFP selection marker, allowing for the assessment of transfection efficiency and expression of the construct by FACS analysis. To assess cell viability, DRAQ7 staining was performed. The small and easy to transfect pmaxGFP plasmid served as a positive control, providing a reference for maximum transfection efficiency. Due to the sensitivity of CD4+ primary T cells to environmental changes, it was expected that the electroporation transfection method will decrease cell viability. Cells that were not electroporated served as a negative control. FACS analysis was performed on days 2, 4, and 9 post-electroporation. Cells were gated for viability, single cells, and GFP positivity using both negative and positive controls (Figure 10b). It was expected that a higher percentage of pmaxGFP transfected cells would successfully transfect when compared to cells transfected with the AIO plasmid, due to the larger size of the AIO plasmid resulting in a lower transfection rate compared. Indeed, upon electroporation it was observed that T cells transfected with the pmaxGFP vector exhibited a 10-fold higher viability compared to the AIO-treated samples on day 4 (Figure 11c). On day 2, pmaxGFP transfected cells had a viability of 25 %, which decreased to 21 % on day 4, and no viable cells were observed on day 9. In contrast, the AIO-treated samples had a viability of 11 % on day 2, dropping to 2 % on day 4, and no viable cells were detected on day 9. The negative control maintained a constant viability between 94 % and 96 % over the nine days of FACS analysis. When considering the GFP positive percentage of cells as a measure of transfection efficiency, it was observed that the negative control showed no GFP positive cells, as expected. Of the pmaxGFP tranfected positive control cells 52 % was GFP positive on day 2, increasing to 53 % on day 4, and dropping to 30 % on day 9. Of the cells transfected with the AIO plasmid, 2 % was GFP positive on day 2, increasing to 9 % on day 4, and decreasing to 4 % on day 9. These results indicate that we are able to transfect cells with the AIO plasmid, reaching a maximum of 9 % positive cells. As expected, it was indeed observed that the positive control had higher transfection efficiency compared to the AIO plasmid (Supplementary Table 4). However, it is expected that cell viability can be improved, considering that viability decreases steeply after electroporation, with no viability measured at day 9 (Figure 12c).

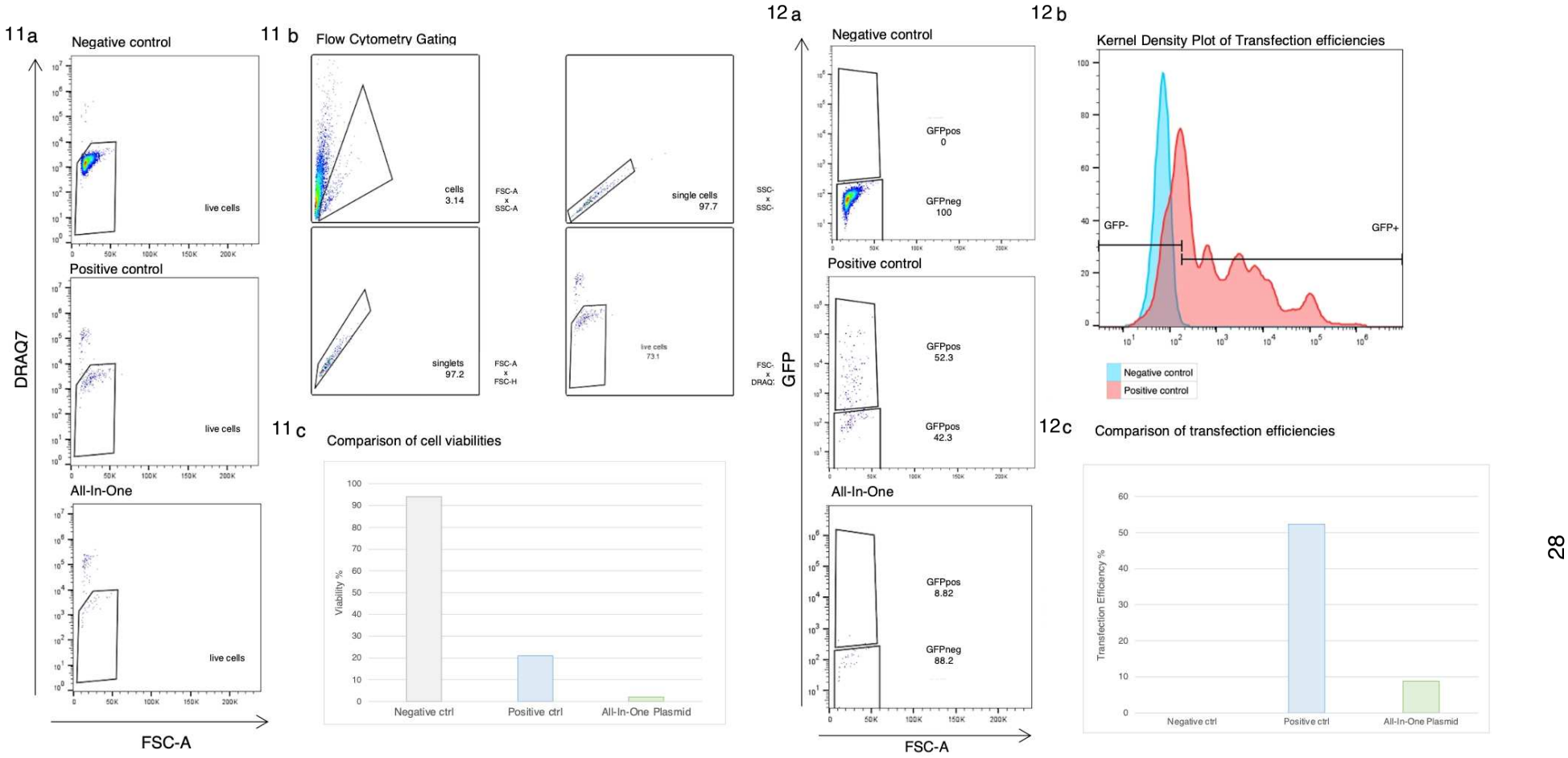


Figure 11: Viability of CD4 T cells after day 4 of electroporation 11a) Bivariate density plots in pseudo-colors on the y-axis, the DRAQ7) is plotted against the forward scatter area (FSC-A) of the negative control, the positive control (pmaxGFP vector) and AIO plasmid. 11b) Representative example of the gating strategy. 11c) Bar graph comparing the cell viabilities of the negative control, positive control and the AIO transfected cells. Negative control shows a viability of around 90 %. Positive control (pGFPmax) has a viability of 21 %, AIO transfected samples are 2 %. n=1 for all samples. 12 a) Bivariate density plots in pseudo-colors on the y-axis, the GFP is plotted against the forward scatter area (FSC-A) of the negative control, the positive control (pmaxGFP vector) and AIO plasmid. 12 b) Overlay of GFP (488-530, 30-A) expression in the negative control (blue) and pmaxGFP transfected samples (red). c) Comparison of the transfection efficiencies of the negative control, the positive control and the AIO plasmid at day 4 post electroporation. Negative control shows 0 % GFP positive cells. Positive control (pmaxGFP) shows 52 % GFP positive cells and AIO transfected contained 9 % GFP positive cells and n=1 for all samples.

Electroporation Shows a Higher Transfection Efficiency Compared to Lipofection

The aim here was to try out an alternative transfection method that could possibly increase the percentage of successfully transfected and living CD4+ T cells. Therefore, lipofection was considered as an alternative transfection. Lipofection involves the use of lipid-based transfection reagents, the reagent is mixed with the plasmid with which the cells should be transfected with and forms cationic lipids. Through ionic reaction the cationic lipids enclose the plasmid and react with the anionic cytoplasmic membrane of the cells. The cells take up the liposomes through endocytic pathway.[33] This method may cause less cellular stress and damage compared to electroporation, where cells are subjected to electric fields resulting in higher cell viability compared to electroporation. It is noteworthy that lipofection was expected to achieve a transfection efficiency of around 8.1 %, which, compared to the electroporation method yielding 60 %-70 % of transfected cells, is considerably lower.[34]

As a proof of principle of transfection efficiency compared to electroporation, the cells were initially transfected only with pmaxGFP. Transfection efficiency was assessed by analyzing GFP expression using FACS. In Figure 12a show that a transfection efficiency of 52.3 % was achieved with the electroporation method, as based on the percentage GFP positive cells. In contrast, 0 % of lipofection transfected cells were GFP positive, indicating that there was no successful transfection with the lipofection method. In Figure 12b, the gate for CD4 positivity was set within the GFP-positive population (GFPpos). It is noteworthy that electroporation showed a higher transfection efficiency compared to lipofection in this context. Based on these results, electroporation was selected as the preferred transfection method for subsequent experiments. Notably, although an equal number of cells were transfected in both electroporation and lipofection, upon analysis, fewer cells remained in the electroporated sample as compared to the lipofectamine treated sample. Likely, this is due to the negative effect of electroporation on cell viability compared to lipofection. Therefore, another experimental approach was set out to optimize cell viability upon electroporation.

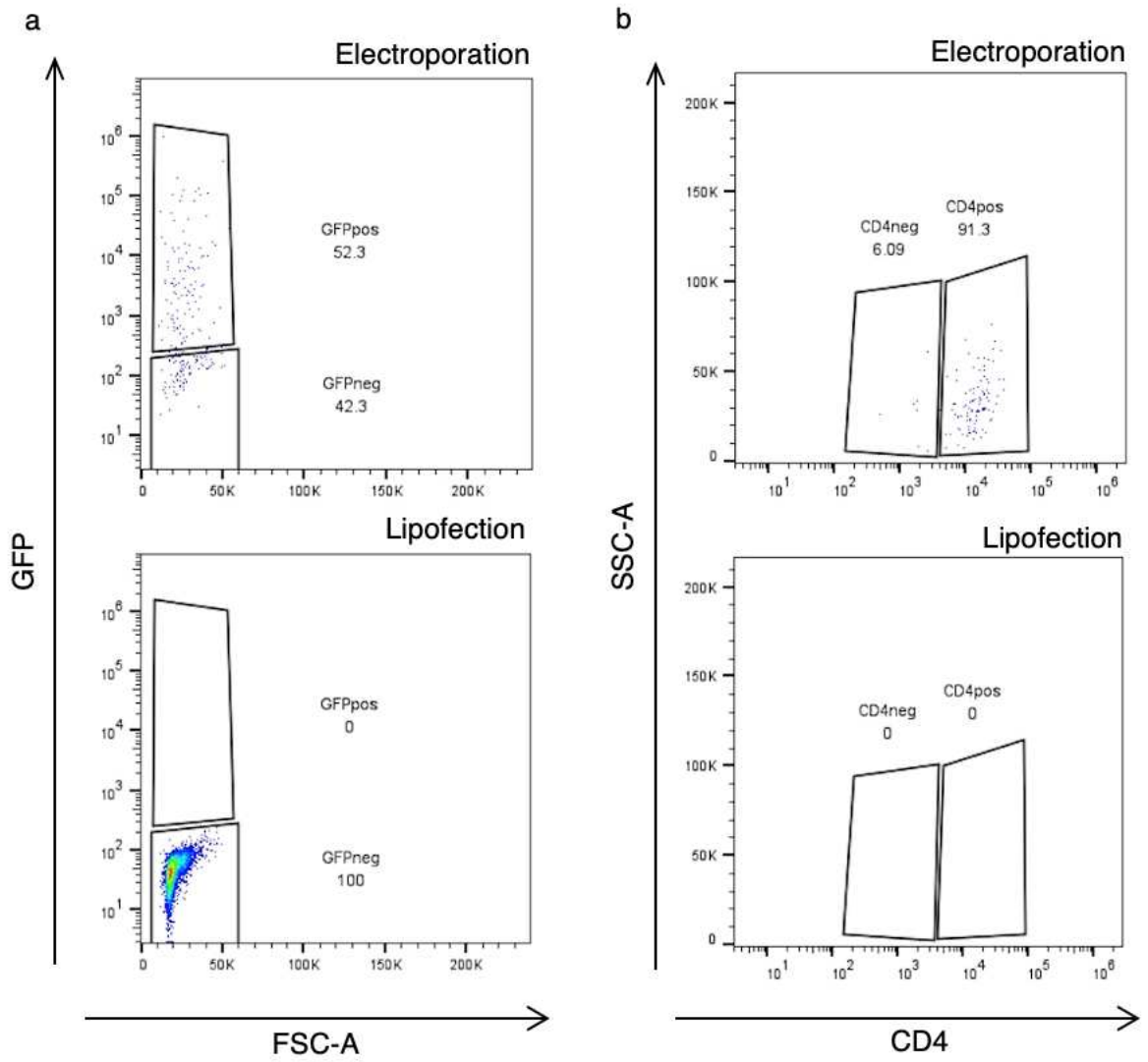


Figure 12: Comparing the Transfection efficiencies of Electroporation with Lipofection. a) Bivariate density plots in pseudo-colors on the y-axis, the GFP is plotted against the forward scatter area (FSC-A) of the cells transfected with pmaxGFP vector through electroporation and cells transfected with pmaxGFP vector through lipofection. B) Bivariate density plots in pseudo-colors on the y-axis, the side scatter area (SSC-A) plotted against CD4 of the cells transfected with pmaxGFP vector through electroporation and cells transfected with pmaxGFP vector through lipofection

Adapted Transfection Method Leads to Higher Cell Viability of the AIO treated Cells

The aim was to increase the viability of primary CD4+ T cells during and after the electroporation process, therefore adjustments had to be made to the electroporation protocol used. Because the transfection efficiency with AIO is dependent on viability, enough cells must survive over the period of the measurement days to make a statement about the success of the transfection. Furthermore, enough cells must survive and be transfected with the AIO to be able to prove whether and how many cells were successfully prime edited. In addition, the individual measurement days are a challenge for cell survival, as the preparation for the FACS analysis includes washing steps, which was be expected to result in additional cell loss. Therefore, an adjusted

protocol was developed, which included several protocol changes. Firstly, the centrifugation force (g-force) applied during the protocol was reduced from an initial value (300 g) to 200 g to minimize the mechanical stress experienced by the cells. In order to retain cell pelleting at this lowered speed, the duration of centrifugation was extended to 10 minutes. Secondly, a recovery phase after electroporation was introduced by leaving the cells in the electroporation strip for 45 minutes and adding pre-warmed medium, rather than immediately transferring cells from the electroporation strip to a culture plate containing growth medium. Additionally, the number of cells transfected in each electroporation reaction was significantly reduced, from 1 million cells to 25,000 cells per reaction. Post-electroporation, cells were maintained in a cell culture medium enriched with reactivation components in a 2:1 ratio of cells to Dynabeads, promoting proliferation. After a 6-hour incubation period, the medium was replaced to decrease the concentration of the potentially toxic electroporation solution within the cell culture medium.

The adapted protocol resulted in a higher percentage of cell viability for AIO-transfected cells, reaching 72 %. This exceeded the viability of positive control cells (64 %) transfected with the pmaxGFP plasmid (Figure 13). The cells were transfected with AIO constructs and the positive control the pmaxGFP plasmid. The positive control exhibited increasing viability over time, peaking at 64 % on day 11 (d11). In the initial method used for the positive control (pmaxGFP), a transfection efficiency of 53 % was observed on day 3 (3d), with a slight decrease to 51 % on day 4 (4d), and a further drop to 30 % on the final measurement day. In contrast, the optimized method exhibited significant improvements, with a transfection efficiency of 69 % on day 3 (3d) and peaking at 80 % on day 6 (6d), before declining to 12 % on day 13 (13d). The decrease over time can be explained as it is a transient transfection, which means that the successfully integrated AIO plasmid is lost over time through the natural cell division process. As expected, the negative controls, for which the cells were not transfected with any plasmid in both methods, exhibited a transfection efficiency of 0 %, aligning with the fact that these cells were not subjected to plasmid transfection. AIO transfected cells exhibited a transfection efficiency of 2 % on day 2 (2d), increasing to a peak of 9 % on day 4 (4d) and decreasing to 4 % on day 9 (9d), when using the initial protocol. Using the adapted protocol, transfection efficiency was measured at 0.1 % on day 3 (3d), which unexpectedly lower than initial method. The efficiency peaked at 15 % on day 6 (6d) and gradually decreased, on day 13 (13d) it was at 1 %. These results indicate that the adapted protocol improved both cell viability and transfection efficiency when compared to the initial protocol. It was therefore decided to continue with the adapted protocol in all further experiments.

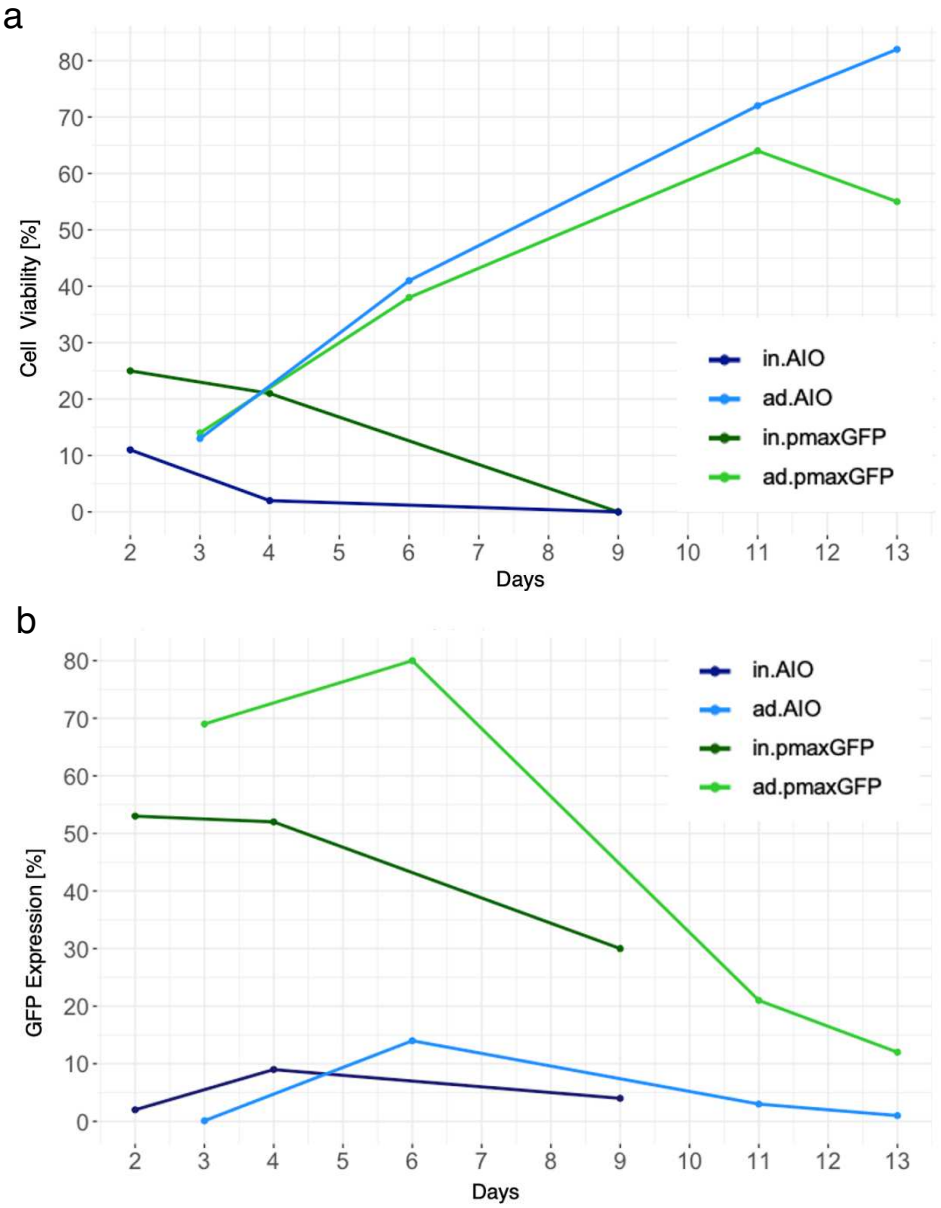


Figure 13. The cell viabilities and transfection efficiencies of the initial and adapted protocol compared. a) Line graph presents comparison of the cell viabilities of the AIO treated cells from the initial protocol (in.AIO) with the adapted protocol (ad.AIO) and the cell viabilities of the positive control where the cells were treated with the pmaxGFP from the initial protocol (in.pmaxGFP) with the adapted protocol (ad.pmaxGFP). For the initial protocol the cells were measured via FACS Analysis on day 2, day 4 and day 9 post electroporation. For the adapted protocol the cells were measured on day 3, day 6, day 11 and day 13 post electroporation. The samples processed according to the adapted protocol show a higher viability compared to the initial protocol. The samples in.AIO (11 % cell viability) and in.pmaxGFP (25 % cell viability) show the highest viability on day 3, at day 9 the cells did not survive and the CV sunk to 0 %. For the samples ad.AIO and ad.pmaxGFP the viability starts with around 15 % and increases over the next days up to 65-80 %. b) Line graph presents comparison of the transfection efficiency of the AIO treated cells from the initial protocol (in.AIO) with the adapted protocol (ad.AIO) and the transfection efficiency of the positive control where the cells were treated with the pmaxGFP from the initial protocol (in.pmaxGFP) with the adapted protocol (ad.pmaxGFP) . For the initial protocol the cells were measured via FACS Analysis on day 2, day 4 and day 9 post electroporation. For the adapted protocol the cells were measured on day 3, day 6, day 11 and day 13 post electroporation. The in.AIO samples show a GFP expression of around 10 % on day 3 compared to that the ad.AIO sample shows a GFP expression of 15 % at day 6, at the continuing measurement days the GFP expression decreases. The sample in.pmxGFP showed a GFP expression of around 50 % at day 3 and the ad.pmaxGFP showed a GFP expression of 80 % at day 6, afterwards the expression decreases.

Comparison of GFP Expression in AIO-Transfected CD4+ Primary T Cells

The aim of this experimental approach was to assess the different editing efficiencies in CD4+ primary T cells through the transfection of four distinct plasmids: AIO2 (All-in-one plasmid with epegRNA insert 2), AIO2-M (All-in-one plasmid and MLH sequence with epegRNA insert 2), AIO5 (All-in-one plasmid with epegRNA insert 5), and AIO5-M (All-in-one plasmid and MLH sequence with epegRNA insert 5). Utilizing the adapted transfection protocol, we set pit to investigate whether these plasmids exhibited variations in both transfection and editing efficiency. All plasmids incorporated a GFP sequence under the hU6 promoter, and the expression of GFP-positive cells served as a metric for transfection efficiency. Each plasmid harbored different epegRNAs targeting the CD4 co-receptor gene in T cells. The CD4 gene was chosen as a target because the main characteristic of CD4-positive primary T cells is that the glycoprotein is present in an abundant amount on their surface. Therefore, it is easy to detect when the gene expression of the CD4 gene is disrupted by prime editing, resulting in less or no CD4 surface proteins. The anti-CD4 antibody was used in the staining procedure for subsequent FACS analysis. This anti-CD4 antibody was selected since it binds to the extracellular domain, which is located outside the cell. If no CD4 protein is present on the surface because the gene expression of the corresponding gene is disrupted by prime editing, the cells should be CD4-negative because the antibody has no binding site and will be washed away during the staining process. The target sites for prime editing were also chosen to target exons of the CD4 gene that form the external domain of the CD4 protein. The external domain was chosen to to ensure that the the translation process of gene is incorrect and that the binding site for the AB is truncated with sufficient probability. This allows an assessment of the prime editing efficiency of the AIO constructs used. The selected AIO constructs were AIO2 and AIO2-M containing an epegRNA designed to induce an edit in exon 3 of the CD4 gene, leading to a premature stop codon during translation. AIO2-M included an additional MLH sequence, intended to enhance prime editing outcomes by disrupting the mismatch repair pathway. AIO5 and AIO5-M aimed to introduce a 2-base pair frameshift mutation into exon 3 of the CD4 gene, causing incorrect translation of the corresponding protein.

The cells were transfected with constructs AIO2, AIO2-M, AIO5 and AIO5-M. Analysis of GFP expression on days 3d, 6d, 11d, and 13d revealed varying GFP expression levels across different constructs (Figure 14a). GFP expression was highest at day 6, with values of 80.2 % for the pmaxGFP positive control, 12% for AIO2, 13.7 % for AIO2-M, 16.2 % for AIO5, and 15.9 % for AIO5-M. Plasmids carrying the additional MLH sequence, namely AIO2-M and AIO5-M, exhibited slightly higher GFP expression over the different measurement days. AIO2-M demonstrated a 1.7 % increase compared to AIO2, while AIO5-M (15.9 %) and AIO5 (16.2 %) showed comparable values (Figure 14b). Detailed FACS analysis on the different days highlighted the influence of the MLH1dn sequence. Although the AIO-M constructs did not show a statistically significant increase in transfection efficiencies compared to those without the MLH sequence, the decision to proceed with them for further experiments was based on the potential benefits offered by the MLH sequence. Despite similar results between constructs with and without the MLH sequence, previous studies indicate that the presence of the MLH sequence may enhance prime editing efficiencies by

inhibiting the mismatch repair pathway. Thus, selecting the AIO-M constructs for subsequent experiments is a strategic choice aimed at maximizing prime editing outcomes, pending further investigation to validate these findings.

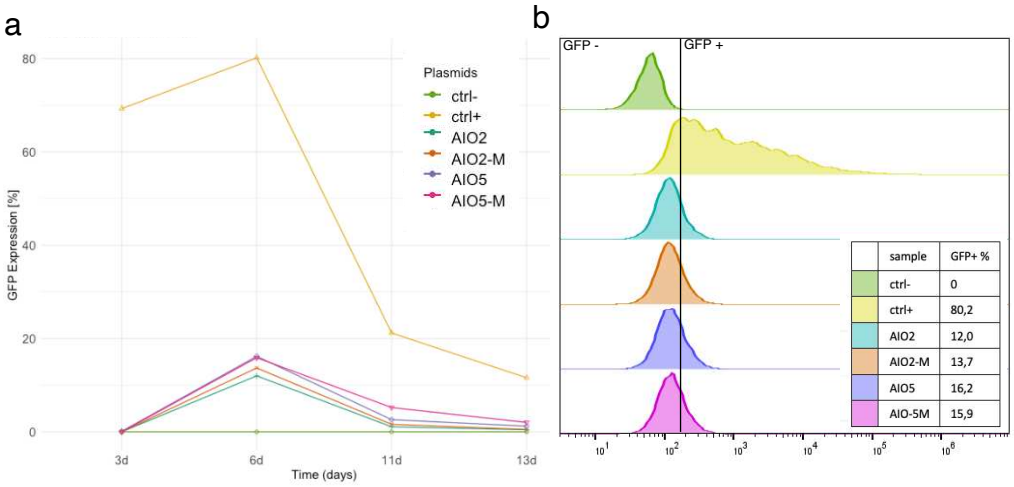


Figure 14: a line graph depicting the changes GFP expression in CD4+ primary T cells following electroporation with four different constructs (AIO2, AIO2-M, AIO5 and AIO5-) over time. The y-axis represents the percentage of GFP positive cells (%), while the y-axis indicates the days of measurement. GFP expression was quantified by FACS analysis on Day 3 (3d), Day 6 (6d), Day 11 (11d), and Day 13 (13d) after electroporation. b) histograms illustrating the GFP expression on the day with highest GFP expression; Day 6 (6d) post-transfection. Top to bottom: negative control, positive control, AIO2, AIO2-M, AIO5 and AIO5-M. The vertical line indicates the threshold used for distinguishing GFP and GFP+ cells.

The Figure 15 depicts FACS plots illustrating cells transfected with the AIO2, AIO2-M, AIO5 and AIO5-M constructs on day 6 post-electroporation, corresponding to the peak GFP expression period. Expectations were for maximal expression of the introduced construct, signifying an enhanced chance of CD4 gene editing and, subsequently, an increased likelihood of observing CD4-negative cells. No CD4-negative cells could be found, which could imply that no prime-editing has taken place in the cells, because the prime editing should inhibit the gene expression of the CD4 gene and thus prevent its anti-CD4 antibody from binding. However, it is noteworthy to mention that it is also possible that cells have undergone a prime editing event in the CD4 gene, but it takes longer until CD4 can no longer be detected. This is because editing alone does not immediately inhibit remove all the already produces CD4 glycoprotein from the surface of the cells, it takes time for the effect to become visible. That means after prime editing a cell with an intact CD4 glycoproteins, it will still be expressed on the cell surface. The conclusion was that these receptors will be replaced in time, but the turnover time of a CD4 receptor is unknown. Thus, a successful prime editing event will not result in the immediate absence of the CD4 receptor. However, to strengthen this assumption, evidence was needed and should be provided by a further experiment.

Therefore, a new analytical approach was proposed: repeating the experiment and sorting cells based on GFP expression levels on day 6 into GFP-positive and GFP-negative groups. Subsequently, detailed sequencing analysis of the GFP-positive cells will be conducted to ensure clear evidence of prime-edited cells.

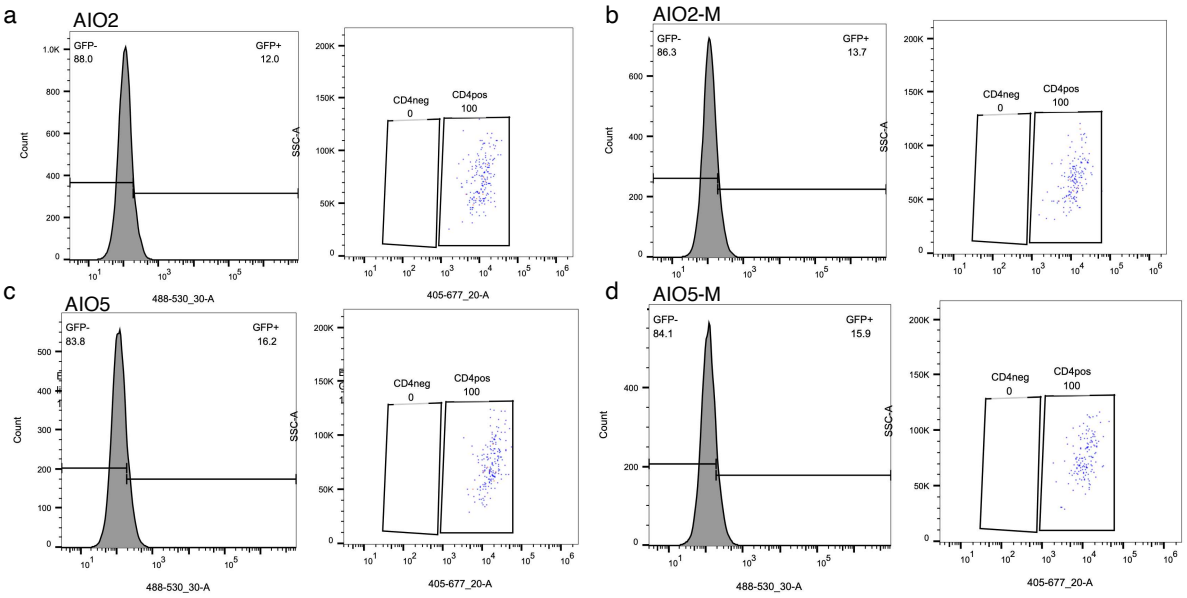


Figure 15: Comparison of the Editing Efficiencies of four different AIO plasmids. Shown is a Kernel Density Plots of GFP negative (GFP-) and GFP positive (GFP+) cells on the left side of each sample, on the right side shown the bivariant density plots in pseudo-colors shows the CD4pos (CD4 positive cells that were gated from the GFP positive population during FACS analysis. CD4 positive T cells were transfected with AIO2, AIO2-M, AIO5 and AIO5-M constructs to edit the CD4 gene. For all samples all the CD4 gates were set from GFP positive cell population. In the samples shown all the cells were CD4 positive and no CD4 negative cells were detected. Analysis was performed on day 6 post-electroporation.

Adaption of NGS Library Preparation for Small Sample Sizes

The previous experimental analyses utilized flow cytometry (FACS) to confirm successful uptake of AIO plasmids by the cells but failed to validate successful prime editing. The prime-editing constructs aimed to induce a nonsense or frame-shift mutation, resulting in a truncated CD4 glycoprotein domain or a stop codon. In either scenario, the CD4 receptor expression on the cell surface should cease. The anti-CD4 antibody used for FACS analysis targets the external domain of the CD4 glycoprotein. Consequently, successful prime editing would prevent anti-CD4 binding, leading to the identification of CD4-negative cells. However, cells with intact CD4 glycoproteins would still express them on the surface, with the assumption that they will eventually be replaced over time. Nonetheless, the turnover time of a CD4 receptor remains unknown.

Therefore, the objective of the new experimental approach was to sort GFP-positive (GFP+) cells and subject them to sequencing analysis to detect any sequence alterations indicative of prime editing at the anticipated target site in the CD4 gene.

To ascertain the editing efficiency, the subsequent step involved the sorting of GFP+ cells. Transfections were performed using the four all-in-one prime editing plasmids, namely AIO1-M, AIO4-M, AIO5-M, and AIO8-M. Constructs with the AIO-M were chosen because in previous experiments, the results showed similar transfection efficiency between constructs with and without the MLH sequence (AIO). However,

previous studies indicate that the presence of the MLH sequence may enhance prime editing efficiencies by inhibiting the mismatch repair pathway. Therefore, the selection of AIO-M constructs for this experiment was a strategic decision aimed at achieving the best results in processing until further research is conducted to validate enhancing PE capability of MLH.

Cells were transfected with the plasmid AIO1-M, targeting exon 3 at position C41, which corresponds to the 41st nucleotide and is characterized by a cytosine (C) residue. In this case, the edit introduced a nucleotide substitution. The original codon, TGT, was altered to a premature stop codon, TAG. This modification initiated a premature termination of the translation process, preventing the synthesis of a complete CD4 protein. The edit with the AIO4-M and AIO5-M construct should also occur in exon 3 at position P20, corresponding to the 20th nucleotide, and affecting a proline residue. The edit involved a deletion of two consecutive cysteines (CC) within the codon CCA. This deletion should result in the transformation of CCA to TCA, inducing a frameshift during translation. Therefore, the translation process should be prematurely terminated, leading to a truncated protein. The difference in AIO4-M and AIO5-M were the designs of the used epegRNAs, this fact is not relevant for this section of the results and will be explained further below (“Analysis and Quality Assessment of the Modified Reads and Validation of Prime Editing of each Sample"). Construct AIO8-M should target exon 4 at position D88, referring to the 88th nucleotide and an aspartic acid residue. In this case, the edit should be an insertion, specifically involving the codon GAC. This insertion introduced an extra 'A' and transformed the codon into GAA. This insertion also induced a frameshift during translation, ultimately preventing the successful completion of the translation process.

As a negative control (ctrl-) cells that were electroporated without the presence of any plasmid were used. As positive control (ctrl+), cells were electroporated with pmaxGFP. This sample served as a control for the experiment efficiency. The positive control exhibited a substantial 39.6 % GFP+ cell population (Supplementary Figure 1), which was a lower percentage compared to the previous experiments. In the previous experiment the GFP expression of the positive control was around 80 % (Figure 15). The FACS analysis of AIO1-M, AIO4-M, AIO5-M, and AIO8-M indicated proportions of only 0.7 %, 0.5 %, 0.5 %, and 0.7 % GFP+ cells, respectively. Due to the low number of GFP+ cells, only few cells for each condition could be sorted, resulting in a low genomic DNA (gDNA) yield. Standard gDNA isolation protocols typically require a minimum of 1 x 10^6 cells, with 1 x 10^4 cells being the lower threshold. Only 1610, 854, 547 and 990 cells could be sorted that were positive for respectively, AIO1-M was sorted with 1610 cells, AIO4-M with 854 cells, AIO5-M with 547 cells, and AIO8-M with 990 cells. To compensate for the suboptimal cell numbers for gDNA isolation, adjustments were made during the library preparation process. Specifically, a greater number of polymerase chain reaction (PCR) cycles was implemented. These additional PCR cycles aimed to enhance the representation of genomic material in the NGS library and mitigate the impact of lower starting cell numbers.

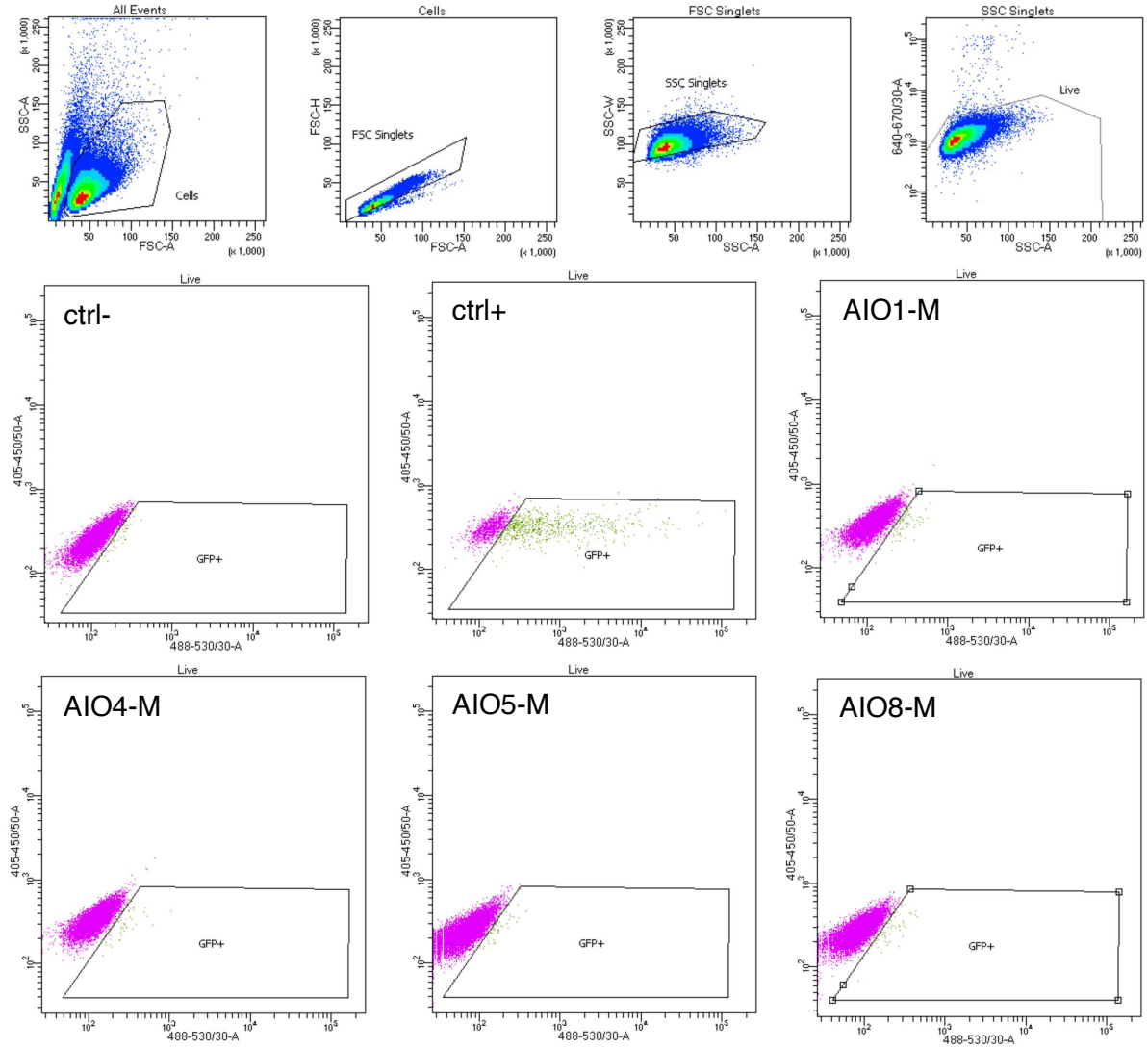


Figure 16: Sorting cells through FACS by setting gates for GFP positive (GFP+) cells. Shown on the top, density plots of the positive control (ctrl+) in pseudo-colors to represent an example of the gating strategy. Cells were sorted on day 6 post-electroporation with four different All-In-One plasmids (AIO1-M, AIO4-M, AIO5-M and AIO8-M). The plasmids carried a EGFP sequence that is expressed upon uptake by the cells and used as a indicator of successful transfection. For the negative control (ctrl-) the cells were not treated with any plasmid and only underwent he electroporation. For ctrl+ the cells were electroporated with a pmaxGFP plasmid.

Sample-Specific Insights: Validating Editing and Sequencing

The genome of CD4+ primary T cells underwent editing across various exon regions of the CD4 receptor gene. The aim was to amplify potentially prime edited DNA from CD4 positive primary T cells to obtain sufficient product for subsequent NGS sequencing. Additionally, creating a high-quality library was crucial for obtaining reliable NGS results and validating whether prime editing was successful. Four editing constructs (AIO1-M, AIO4-M, AIO5-M, and AIO8-M) were introduced into the cells via electroporation. Each construct contained a GFP sequence for subsequent cell sorting based on GFP expression. Following transfection and cell sorting, genomic DNA (gDNA) was extracted from the four distinct samples. Due to the relatively low cell

count post-sorting, a limited gDNA yield was anticipated. This expectation guided the subsequent amplicon PCR, which specifically targeted the region of interest.

To prepare these DNA fragments for Illumina sequencing, DNA adapters and unique DNA barcodes were incorporated in the subsequent steps. DNA extraction from the cells was followed by primer design specific to regions flanking the target region. These primers facilitated PCR amplification of the region of interest. Subsequently, PCR products underwent purification to remove excess primers and contaminants. Following amplification of the region of interest, size selection was performed using beads to obtain a library containing desired amplicon fragments.

For quantification and quality assessment of the size-selected library, electrophoresis using a Bioanalyzer was conducted. This provided insights into size distribution and nucleic acid fragment concentration within the samples. Electrophoresis runs were evaluated using a ladder, with distinctive peaks of 10.380 base pairs (bp) for the upper ladder and 35 bp for the lower ladder expected (Figure 17).

Subsequent analysis revealed that samples AIO1-M, AIO4-M, and AIO8-M displayed distinctive peaks aligning with expected fragment sizes, indicating suitability for downstream NGS analysis. However, sample AIO5-M exhibited less distinctive peaks, suggesting potential contamination. Given this, sample AIO5-M was omitted from library preparation to prevent data distortion. Samples AIO1-M, AIO4-M, and AIO8-M were pooled and sent for NGS sequencing (Supplementary Table 5).

Pooling was feasible due to the use of unique DNA barcodes during PCR amplification of regions of interest. These barcodes allowed for multiplexing, enabling concurrent sequencing of multiple samples. The adapters, containing short DNA sequences complementary to sequences on the Illumina sequencing flow cell, facilitated attachment of amplicons to the flow cell surface. Additionally, they played a crucial role in primer hybridization during sequencing, aiding in the formation of DNA clusters on the flow cell surface. Unique barcodes within the adapters served the purpose of sample tracking during data analysis, providing information on the source of each read. Barcodes are unique DNA sequences specific to each sample, were incorporated into adapters during library preparation. During sequencing, these barcodes were read alongside DNA fragments, enabling identification, and tracking of sequences from different samples. PCR amplification of the region of interest occurred through multiple cycles of denaturation, annealing, and extension, resulting in exponential amplification of the target DNA sequence.

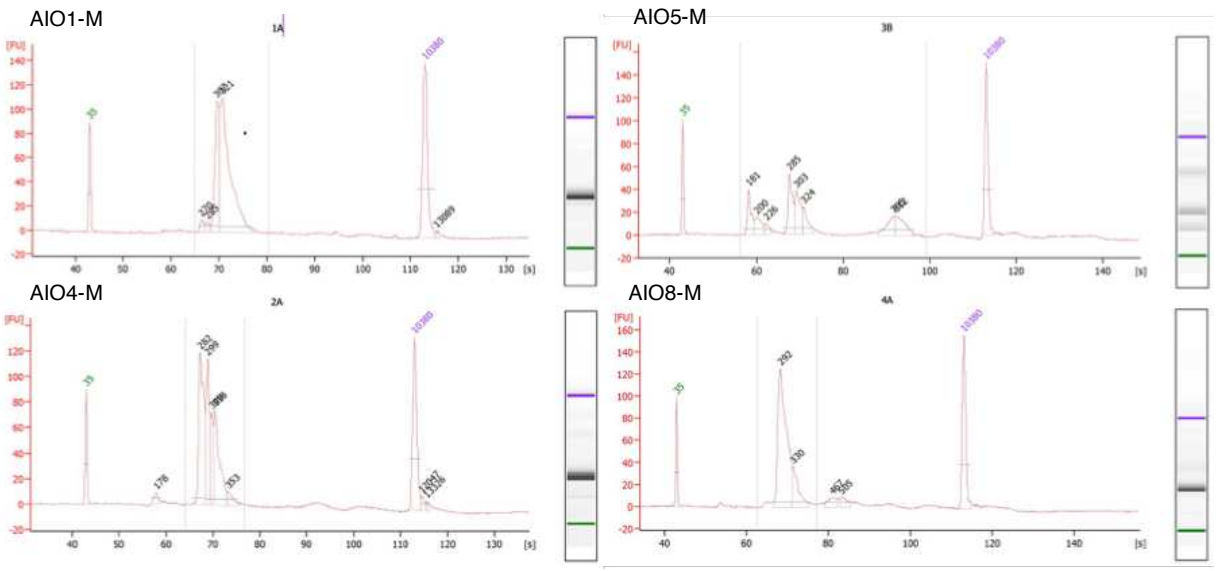


Figure 17: Electropherogram showing the size distributions of NGS libraries prepared from cells transfected with A) AIO1-M, B) AIO5-M, C) AIO4-M, D) AIO8M, each construct encodes for a different epegRNA, targeting a different site within the CD4 gene, therefore a different amplicon was amplified. The x-axis shows the seconds (s) and the y-axis shows the fluorescence units (FU). Expected amplicon sizes are 302 bp (AIO1-M), 277 bp (AIO4-M), 277 bp (AIO5-M) and 280 bp (AIO8-M).

Downstream NGS Analysis by Using CRISPResso2 Software

CRISPResso2 is a specialized software designed for analyzing genome editing experiments, particularly those involving amplicon sequencing. Its workflow includes aligning sequencing data with the GRCh38/hg38 reference genome and generating a report on editing efficiency. The software filters out low-quality reads, trims adapters, aligns reads to a reference sequence, and quantifies induced mutations (Figure 18 and Figure 19). Unlike many other analysis tools, CRISPResso2 takes into consideration the biological nuances of genome editing, such as the components of the prime editing system like the nickase and epegRNA. This ensures accurate quantification of indel events, as these factors are precisely defined in the software. Another crucial parameter in CRISPResso2 analysis was specified the quantification window, which specifies the base pair positions of the intended edit within the amplicon sequence. Therefore the algorithm is able to use the quantification window coordinates and ensures that reads precisely align with the intended edit location within the amplicon.

Analysis and Quality Assessment of the Modified Reads and Validation of Prime Editing of each Sample

Three samples, sample 1, sample 2, and sample 3, respectively containing sequencing data of the cells transfected with editing constructs AIO1-M, AIO4-M and AIO8-M. Each of the editing constructs was designed to introduce modifications in the CD4 receptor gene, affecting exon 3 or exon 4. For sample 1 (AIO1-M) cells were transfected with the plasmid AIO1-M, targeting exon 3 at position C41, which corresponds to the 41st nucleotide and is characterized by a cytosine (C) residue. In this case, the edit introduced a nucleotide substitution. The original codon, TGT, was

altered to a premature stop codon, TAG. This modification initiated a premature termination of the translation process, preventing the synthesis of a complete CD4 protein.

The analyzed NGS results showed that in sample 1 there were a total of 663,380 reads sequenced, of which 663276 Aligned to the reference (r). Among these, 661212, or 99,68 % reads were categorized as unmodified as compared to the reference. A smaller subset, 2068 reads (0,31 %), displayed modifications in relation to the reference. Within the modified reads, only one read (r) was classified as prime-edited, indicating a rare occurrence of prime editing in this sample (Figure 19). Sample 1 exhibited a substitution at the expected amplicon position. However, this substitution was identified in only one read, raising concerns about its significance as concrete evidence of a successful prime editing event. It is essential to acknowledge the possibility of a sequencing error. Furthermore, it is worth noting that the substitution resulted in an alteration from AC to GT, diverging from the anticipated CT change. Additionally, 0.02 % of the reads (103 reads) fell into the ambiguous category, signifying an inability to categorize them as clear modifications, potentially these reads could be prime edited as well, but after analyzing these reads further by aligning the reads manually to the reference gave no evidence of them being prime-edited.

The cells of sample 2 were transfected with the AIO4-M construct. For construct AIO4-M mutations should occur in exon 3 at position P20, corresponding to the 20th nucleotide, and affecting a proline residue. The edit involved a deletion of two consecutive cysteines (CC) within the codon CCA. This deletion should result in the transformation of CCA to TCA, inducing a frameshift during translation. Therefore, the translation process should be prematurely terminated, leading to a truncated protein.
The sequencing result for sample 2 encompassed a total of 1,192,050 reads, with 0.18 % (3876 reads) recorded as modified, while the majority remained unmodified (99.82 %). Eight reads were categorized as prime-edited, while one read was classified as ambiguous. The eight reads were confirmed as prime edited by the manual alignment of the individual read with the reference sequence. When visually inspecting the alignment of the 2137 reads (Figure 20) that were classified as modified, it seems that these are also prime edit results, it seems that in this case, CRISPResso2 did not correctly categorize the reads, the origin of this inability to correctly categorize the reads is unknown. Notably, the Figure 21 displays a distinct peak signifying a deletion at the precise amplicon position targeted for editing.

Construct AIO8-M, which was used for sample 3 should target exon 4 at position D88, referring to the 88th nucleotide and an aspartic acid residue. In this case, the edit should be an insertion, specifically involving the codon GAC. This insertion introduced an extra 'A' and transformed the codon into GAA. This insertion also induced a frameshift during translation, ultimately preventing the successful completion of the translation process. In the case of Sample 3, out of a total of 989,372 reads, 0.05 % displayed modifications. However, no reads were identified as prime-edited. This outcome led to the characterization of this sample as unsuccessful in the context of prime editing.

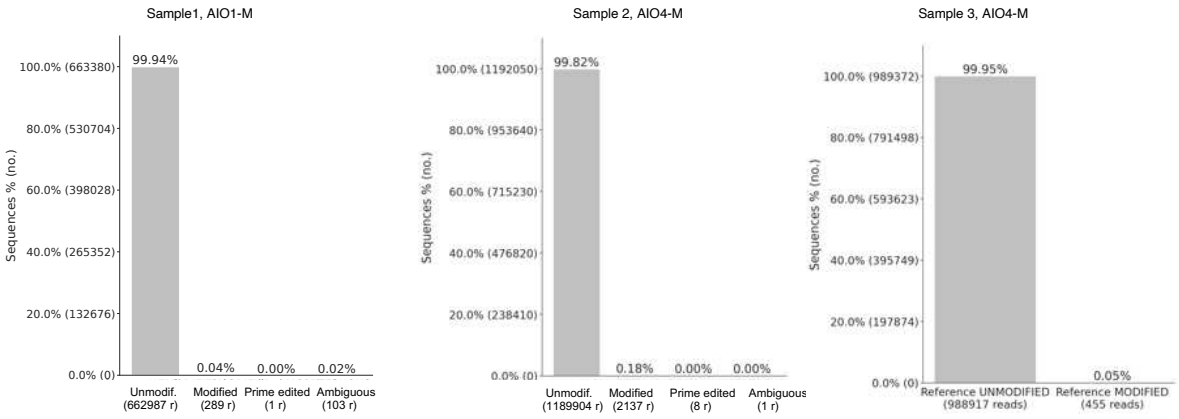


Figure 18: In this Figure, the bar graph presents the count of reads at various processing stages. The graph is organized from left to right to represent Sample 1 (AIO1-M), Sample 2(AIO4-M), and Sample 3 (AIO8-M). The y-axis illustrates the percentage of reads that remain at each processing stage, with the corresponding number of sequenced reads provided in parentheses.

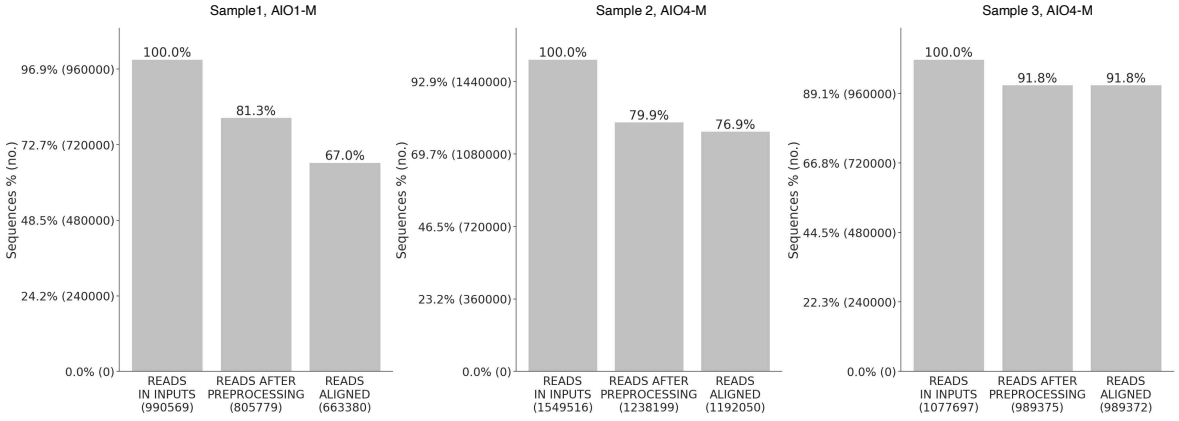


Figure 19: Bar Graph of categorized aligned reads. In each bar graph the y-axis shows the % of reads and the number of reads in the brackets, indicates the categories; unmodified (Unmodif.), "Modified", "Prime-edited" and "Ambiguous" reads (r). The category unmodified stands fort the reads, that were not edited when compared to the provided reference sequence. In the category modified all the reads fall that were edited, in the prime edited category are the once that got the specific modification that was intended with the specific AIO construct. Ambiguous reads are the ones that could not be categorized by the pipeline in any other category. For sample 1 CD4+ primary T cells were transfected with the AIO1-M construct, In sample 2, cells were transfected with AIO5-M and fin sample 3 cells were transfected with AIO4-M.

Analysis of the Modified Reads and Validation of Prime Editing

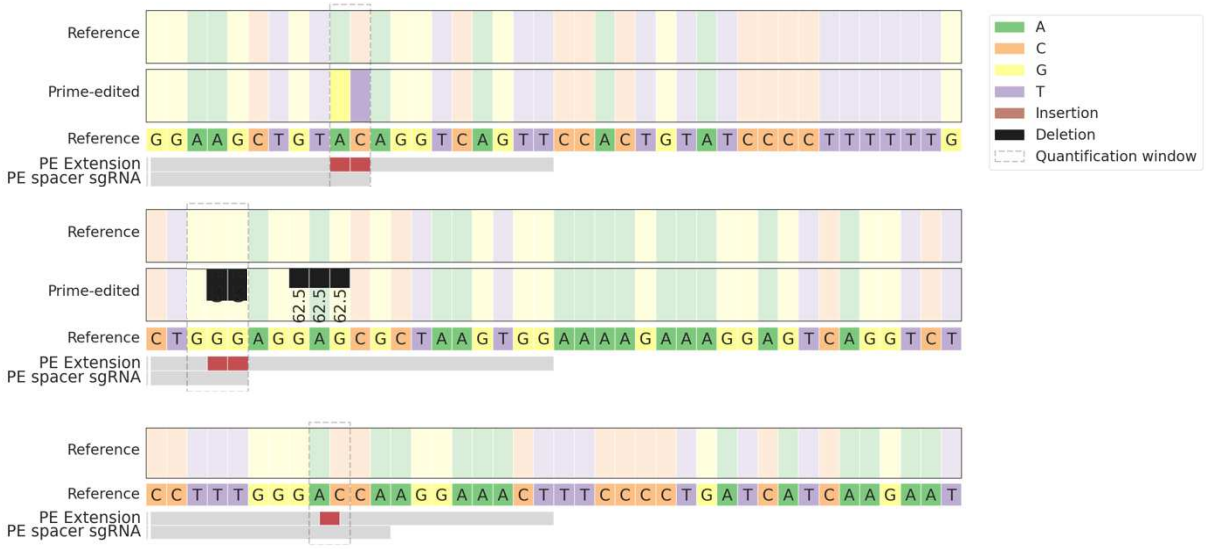


Figure 20: Distribution of nucleotides around the prime editing extension (PE Extension) (AACTGACCTAGACAGCTTCCCAGAA) from top to bottom first sample 1, PE Extension (CCACTTAGCGCTCCTCAGCAGCCACTC) sample 2 and PE Extension (AGTTTCCTTGGGTTCCCAAAGGCTT) sample 3. In this Figure, the distribution of nucleotides around the Prime Editing Extension is depicted for three distinct samples: Sample 1, Sample 2, and Sample 3. At the top of each sample, the reference sequence to which the extension aligned is presented, followed by the prime-edited version of the reference and the detailed sequence of the reference, spanning 40 bases. Following this, the PE Extension is aligned to the Reference Sequence, enabling a visualization of the expected edit location and the alignment of the Prime Editing Spacer sgRNA (PE spacer sgRNA). Dashed lines are utilized to indicate the Quantification Window (QW) applied for each sample. Sample 1: Displays a substitution of the bases AC to GT within the QW. Sample 2: Reveals the deletion of the bases GG and additional deletions. Notably, the additional deletions appear to occur in the sequence GAG downstream of the intended deletion, with the percentage of these additional deletions provided (62.5%). Sample 3: Was anticipated to contain an insertion; however, the insertion is not clearly discernible in the Figure.

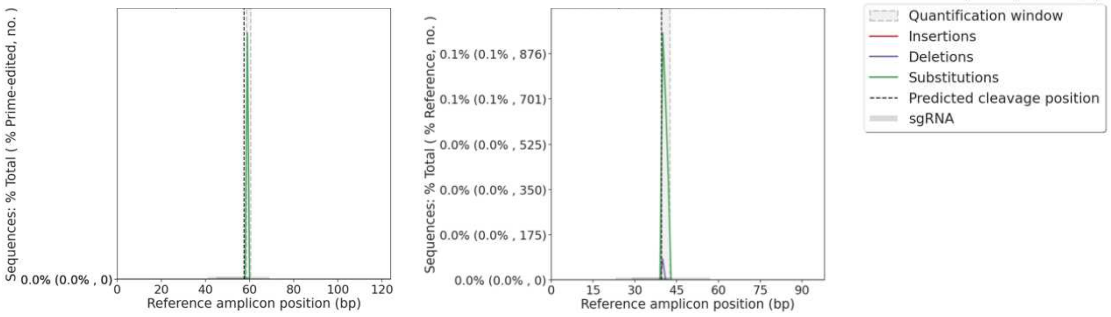


Figure 21: Frequency of Insertions, Deletions, and Substitutions Across the Entire Amplicon. This Figure illustrates the distribution of insertions (in red), deletions (in purple), and substitutions (in green) across the entire amplicon. The data includes modifications that extend beyond the quantification window. Sample 1 is depicted on the left, while Sample 2 is on the right side of the graph. The x-axis indicates the position where the edit should be located, while the y-axis represents the quantification window for the expected edit. The dashed lines on the graph delineate this quantification window. The y-axis displays both the percentage and the number of reads that exhibit modifications, including insertions, deletions, and substitutions.

Discussion

Experimental Approach and Objectives

This study seeks to link enhancers containing immune disease-related variants with the genes they control, focusing on CD4 positive primary T cells. Improving precision editing in primary CD4+ T cells and screening GWAS variants in prioritized enhancer regions are key objectives. Additionally, utilizing targeted perturb-seq (TAP-Seq) allows for analyzing transcriptome and genetic changes at a single-cell level. This approach promises to uncover how disease-associated variants affect gene regulation, offering insights into potential therapeutic targets.

Choice of the Precision Editing Method

The PEmax system, derived from the PE1 system, was utilized, employing a human codon-optimized reverse transcriptase (RT) to enhance efficiency in human cells, particularly CD4 positive primary T cells. The design of the PEmax system ensures efficient transcription and translation of the RT gene in human cells, optimizing prime editing outcomes.[22]

Comparing Transfection Methods: Electroporation versus Lipofection

The exploration of alternative transfection methods aimed to enhance the viability of CD4-positive primary T cells and ideally also the transfection efficiency for subsequent experiments. Lipofection was considered as a potential alternative to electroporation due to its hypothesized ability to induce less cellular stress and damage. Unlike electroporation, where cells are exposed to electric fields, lipofection relies on lipid-based transfection reagents to encapsulate plasmids and deliver them into cells via endocytosis. This method was anticipated to yield lower transfection efficiency, around 8.1 %, but higher cell viability compared to the 60 %-70 % transfection efficiency achieved with electroporation. Therefore, a proof-of-principle experiment was conducted using pmaxGFP to assess transfection efficiency. Flow cytometry analysis revealed a transfection efficiency of 52.3 % with electroporation, as evidenced by the percentage of GFP-positive cells. In contrast, no GFP-positive cells were observed with lipofection, indicating a lack of successful transfection.

The results demonstrated that electroporation achieved higher transfection efficiency compared to lipofection. Consequently, electroporation was selected as the preferred transfection method for subsequent experiments. Despite both methods delivering an equal number of transfected cells, fewer cells remained viable following electroporation compared to lipofection, likely due to the adverse effects of electroporation on cell viability. Therefore, efforts were directed towards optimizing cell viability post-electroporation to further enhance experimental outcomes.

In conclusion, while lipofection presents an alternative transfection approach, its lower transfection efficiency and preserved viability make electroporation the preferred method for efficient transfection of CD4-positive primary T cells in this context. Future studies may focus on refining electroporation protocols to minimize cellular stress and optimize cell viability, ultimately improving experimental outcomes. Another

transfection method could be considered, namely using the lentivirus delivery system in primary T cells. Lentivirus vectors can incorporate constructs up to 10 kbp and provide stable transgene expression.[35,36] However, it should be noted that constructs with a size of 12 to 14 kbp should be split into two parts, one part containing the nickase and the other containing the epegRNA.

Adjusted Electroporation Protocol Compared to the Initial Electroporation Protocol

The primary objective was to enhance the viability of primary CD4+ T cells during and after the electroporation process, with the aim of improving the success of transfection and prime editing efficiency using the All-in-One plasmid (AIO) constructs. The results demonstrate the successful optimization of the electroporation protocol, resulting in improved cell viability and transfection efficiency compared to the initial method.

One of the critical considerations in optimizing the electroporation protocol was to minimize cell loss during the process. The reduction in centrifugation force and the introduction of a recovery phase post-electroporation were key modifications aimed at mitigating mechanical stress and enhancing cell survival. In previous studies it was found that after Electroporation, CD4 T cells undergo dramatic morphological changes. These were characterized by wrinkled and dilated plasma membranes before recovering 1 hour later. [37] It was also shown in past studies that the increase the recovery time of the cells, means leaving them in the electroporation cuvette for longer increased the cell survival massively. [38] Another study showed that plasmids can potentially get entrapped within electropermeabilized membranes. Electropores are formed during the electroporation process through which plasmids may pass. However, there is a possibility of plasmids becoming lodged in these pores and being introduced into the cell as the membrane reseals. [41] Therefore the optimization of the protocol also incorporated an extended duration for which electroporated cells remained within the electroporation cuvettes, lasting 45 minutes post-electroporation, with the addition of a small volume of pre-warmed cell culture medium. The washing step before the electroporation process also plays a key role in cell survival rates. It is important to ensure the complete removal of PBS during this step, as PBS, combined with the RPMI medium, contains a high salt content that can elevate the medium's conductivity, potentially resulting in the overheating of the electroporation stripe and, consequently, cell damage and death.[40]

The optimized method also resulted in notable enhancements in transfection efficiency, with peak efficiency observed on day 6 before declining gradually over time. The decline in GFP expression over time is explainable because the construct used were transiently transfected, means they do not stably integrate in the cell and get lost over time due to cell division. In the initial protocol it was 9 % (day 4 after electroporation) compared to the adapted protocol the percentage of GFP positive cells increased to 15 % (on day 6) after electroporation. GFP expression following electroporation with the AIO plasmids was consistently lower when comprised to transfection with the pmaxGFP (positive control). Tendentially, this could be due to the larger size of the AIO plasmids (ranging from 12 to 14 kbp) as compared to the

pmaxGFP (3.4 kpb) plasmid that was used for the positive control. The size could impose a burden on the cells during electroporation, potentially slowing their expression relative to the smaller pmaxGFP plasmid. In past studies it was shown that larger plasmids (6 to 16 kpb) result in low cell viability and transfection efficiency compared to small plasmids (3.5 kpb) used under the same experimental conditions. In the study the decrease was shown to be linked directly to the physical plasmid size, irrespective of transgene expression. It is possible that the large size of the plasmid makes it more difficult to pass the cell membrane, resulting in lower transfection rates, because of its low mobility.[38]

In the adapted protocol the cell number that was electroporated was increased, it was shown that increased cell number results in enhanced viability and transfection efficiency.[37] The cells were reactivated after electroporation, because upon activation the cells start to expand and with that the cell number increases. That was important because as many viable cells as possible were needed that survive after electroporation and by adding the reactivation medium this could be achieved. This could be seen because in the adapted protocol the number of viable cells increased (80 % on day 13) and recovered better than in the initial protocol where it dropped to 0 %. Additionally, an increase in plasmid volume during electroporation has been linked to a higher percentage of GFP-positive cells, but in contrast leading to a reduced cell viability. [39] However, it is important to note the unexpected lower transfection efficiency observed on day 3 using the adapted protocol. This discrepancy may be attributed to various factors, including variability in experimental conditions or technical issues during the electroporation process. Further investigation is warranted to identify the specific cause and address any potential limitations associated with the adapted protocol.

Overall, the successful optimization of the electroporation protocol underscores the importance of methodological refinement in enhancing the efficiency and reproducibility of this experiment. The decision to continue with the adapted protocol for all further experiments was supported by the improvements in both cell viability and transfection efficiency observed in this study.

Comparison of the Different Constructs Transfection Efficiencies

The results aimed to evaluate the transfection and editing efficiencies of four distinct plasmids (AIO2, AIO2-M, AIO5, and AIO5-M) in CD4+ primary T cells. The adapted transfection protocol was used to investigate potential variations in both transfection and editing efficiency among these plasmids. Transfection efficiency was assessed by measuring GFP expression, which was under the control of the hU6 promoter in all plasmids. The plasmids differed in their epegRNA designs targeting the CD4 co-receptor gene in T cells. AIO2 and AIO2-M aimed to induce a premature stop codon in exon 3 of the CD4 gene, while AIO5 and AIO5-M aimed to introduce a frameshift mutation in the same exon. The choice of targeting the CD4 gene was strategic, given its abundant expression on the surface of CD4+ T cells, facilitating the detection of editing outcomes.

Analysis of GFP expression over multiple days revealed varying levels of transfection efficiency across different constructs. Interestingly, plasmids carrying the additional MLH sequence (AIO2-M and AIO5-M) exhibited slightly higher GFP expression on day 4 post electroporation compared to those without the MLH sequence. Although previous studies suggest that the MLH sequence may enhance prime editing efficiencies by inhibiting the mismatch repair pathway, the results did not conclusively demonstrate a significant improvement in transfection efficiency. However, the choice to continue with the AIO-M constructs for subsequent experiments reflects our interest in maximizing prime editing outcomes.[42]

Further investigation is warranted to validate these findings and elucidate the specific mechanisms underlying the potential benefits of the MLH sequence. Additionally, future studies may explore alternative strategies to enhance prime editing efficiencies in CD4+ primary T cells. An alternative would be to compare the PB and PB-M constructs because these constructs combine the PE system with the piggyBac transposon system. The selected PB or PB-M plasmid would have to be transfected with the transposase plasmid, but in contrast to transfection with AIO and AIO-M, one would have a stable integration at several locations in the genome, which could ensure that the transfected construct is expressed at a higher level and thus increase the transfection efficiency. This could potentially also increase the probability that more cells are edited and make it easier to compare the transfection and editing efficiencies of the different constructs. However, it would have to be ensured that the two plasmids to be transfected, PB or PB-M, are transfected with the transposase plasmid in the correct ratio, as studies have shown that the concentration of plasmids transfected during electroporation has an influence on the survival probability and transfection efficiency of the cells. [39] It should also be considered to use a different promotor for the constructs in use, here the prime editing components were expressed under the hU6 promotor. Studies have shown that the expression of a transgene is dependent on the promoter that is used.[43] The promoters that could be used for further research should be specific for CD4 positive primary T cells. For instance CMV, EF-1, hPGK and RPBSA, these are used for a different subset (CAR T cells) on T cells, but could also be used in CD4+ primary T cells and help to see if an increase in transfection and editing efficiency is visible. [44]

Evaluation of Prime Editing Efficiencies

The experimental approach aimed to confirm if prime editing in CD4+ primary T cells following transfection with various all-in-one prime editing plasmids took place. The strategy involved first sorting GFP-positive (GFP+) cells, indicative of successful transfection, isolating their genomic DNA (gDNA), preparing a library for NGS and subjecting them to sequencing analysis to detect sequence alterations suggestive of prime editing at the anticipated target sites within the CD4 gene. The constructs utilized in the experiment, include AIO1-M (sample 1), AIO4-M (sample 2), AIO5-M, and AIO8-M (sample 3), were designed to induce specific mutations in exon 3 or exon 4 of the CD4 receptor gene. Each construct aimed to introduce either a premature stop codon or a frameshift mutation, resulting in the truncation of the CD4 glycoprotein

extracellular domain. This truncation was expected to lead to the downregulation of CD4 receptor gene expression on the cell surface, thus rendering the cells CD4-negative. The anti-CD4 antibody used for FACS analysis specifically targeted the external domain of the CD4 glycoprotein, enabling the identification of CD4-negative cells. To prepare DNA fragments for Illumina sequencing, we incorporated DNA adapters and unique DNA barcodes into the workflow. Primer design specific to regions flanking the target region facilitated PCR amplification of the gene of interest followed by purification to eliminate contaminants. Subsequent size selection using beads ensured the enrichment of the library with desired amplicon fragments.

The experimental results demonstrated a substantial decrease in the proportion of GFP+ cells following transfection with the prime editing constructs compared to the positive control transfected with pmaxGFP. While the positive control exhibited a GFP+ cell population of 39.6 %, the percentages for cells transfected with AIO1-M, AIO4-M, AIO5-M, and AIO8-M were much lower, ranging from 0.5 % to 0.7 %. In comparison, the experiment using adapted electroporation had a higher percentage of GFP-expressing cells, which was 80 % for the positive control (pmaxGFP) and about 15 % for the AIO plasmids used. Therefore, the numbers of GFP-expressing cells were significantly lower than in previous experiments. One explanation for this is that the adapted electroporation used CD4-positive primary T cells from a different donor and the sorting used a different donor. It has been shown that cells taken from different donors in culture may exhibit changes that affect rates of cell loss and population doubling as well as transduction efficiency.[45]

The limited number of GFP+ cells (approximately 1000 cells) resulted in a low yield of genomic DNA (gDNA), which posed challenges for downstream analyses. Typically, 1000 cells represent the minimum requirement for most gDNA extraction kits. Following genomic DNA extraction, the concentration was measured using Qubit, yielding a concentration of 0.2 ng/μl. To address the low concentration of the samples, a concentration step was performed before proceeding with PCR amplification of the region of interest. Subsequently, adapter and barcode sequences were attached to facilitate downstream analysis. The PCR steps were also optimized by implementing size-selection of the product after the initial PCR amplification. This size-selection process allowed for the removal of unwanted DNA fragments and enrichment of the desired amplicons. Subsequent rounds of PCR amplification were performed to increase the starting material for the second PCR and to ensure sufficient representation of the target sequences. Evaluation of the size-selected library's quality and quantity was conducted using electrophoresis with a Bioanalyzer, providing insights into size distribution and nucleic acid fragment concentration. Samples AIO1-M, AIO4-M, and AIO8-M exhibited distinctive peaks aligning with expected fragment sizes, indicating their suitability for downstream NGS analysis. However, sample AIO5-M displayed less distinct peaks, suggesting potential contamination. To maintain data integrity, sample AIO5-M was excluded from library preparation. It is worth noting that the repetition of the PCR may have led to the amplification of not only the desired amplicon but also errors in the form of undesired DNA fragments. Despite size selection attempts, these fragments could not be effectively removed, potentially contributing to the presence of multiple small peaks observed in sample AIO5-M. Therefore, samples AIO1-M, AIO4-M, and AIO8-M were pooled and sequenced.

These results show the importance of optimizing PCR conditions and ensuring stringent quality control measures to minimize the presence of artifacts and contaminants in the sequencing library.

The NGS data analyses was done using the CRISPRsso2 software. In the case of sample 1, transfected with the AIO1-M construct, the intended modification involved a nucleotide substitution at position C41 of exon 3, resulting in a premature stop codon and subsequent termination of CD4 protein synthesis. However, the NGS analysis revealed a low occurrence of modifications, with only one read identified as prime-edited. This solitary instance makes the reliability of the observed substitution, particularly considering its divergence from the expected nucleotide change.

For sample 2, transfected with AIO4-M, the intended edit aimed to induce a frameshift mutation by deleting two consecutive cysteines at position P20 of exon 3. Despite a relatively higher proportion of modified reads detected, CRISPResso2 categorization issues raised doubts regarding the accuracy of the identified prime-edited reads. Manual alignment showed the presence of prime editing events, underscoring the need for careful interpretation of sequencing data and validation of results. Sample 2 showed prime editing in 8 reads, and 3876 reads were categorized as modified from 1192050 reads in total, which means that 0,18 % of reads were modified, and 0,001% reads were categorized as prime-edited. When comparing to the reference sequence confirming the deletion at the target position and intended target. It is not entirely clear why the 3876 reads did not fall under prime-edited, as the quantification window set for the probe was specific to the site being edited, there appears to be a problem in the pipeline which should be investigated further for future experiments. The investigation could include the use of known datasets or synthetic data that includes predefined editing events. With this validation step one could assess the sensitivity and specificity of the quantification window parameter of the pipeline in accurately detecting prime editing events.

In sample 3, transfected with AIO8-M targeting exon 4 at position D88, the intended insertion failed to manifest as prime-edited reads. Although modifications were detected, none were classified as prime-edited, indicating a lack of successful editing outcomes. This underscores the variability in editing efficiency observed across different constructs and target sites, highlighting the importance of comprehensive screening and validation procedures.

In order to form a meaningful conclusion for this sample, the entire experiment would have to be repeated. It is possible that the construct used for this sample was not efficient for prime editing. One assumption to which this could be attributed due to is that the epegRNA used for construct AIO8-M was designed should not disrupt the PAM sequence and leave it intact. However, recent studies have shown that the disruption of PAM sequence with the introduced edit leads to an improvement in prime editing efficiency. This suggests that the result of the study would agree, but it must be mentioned that further verification would be necessary to definitively assert this.[46]

A new experimental approach base editing could be an alternative to prime editing, but compared to prime editing, which allows all 12 types of mutations, it is limited in

the type of changes that can be made. Despite this limitation, base editing offers the possibility of making 4 single nucleotide substitutions, which may be sufficient for precise mapping of enhancer regions to disease-associated gene networks. Approximately 58 % of human disease-associated genetic variants are due to single nucleotide polymorphisms (SNPs), which are predominantly located in non-coding regions such as enhancers that regulate multiple genes and influence disease phenotypes. Base editing enables the correction of single nucleotide mutations. However, a disadvantage of this technique is the potential for concomitant mutations affecting other bases in the vicinity of the intended change.

Future studies should focus on optimizing editing parameters and validation techniques to advance the utility of prime editing as a precise genome engineering tool in primary cell models.

Conclusion

The aim was to expand the toolkit for CD4+ primary T cells and to decipher immune-associated disease variants using the prime editing method, using an All-in-One plasmid with different enhancements for PE. In order to use these results as a toolkit for further development, additional refinements could be made. Although challenging due to the small cell number, the experiment highlights the importance of rigorous validation of prime editing results and optimization of experimental protocols for future studies. The implementation of adjustments in library preparation demonstrates the flexibility required to overcome experimental limitations and maximize the yield of meaningful data. Furthermore, the observed discrepancies in GFP expression between the positive control and primary editing constructs indicate that transfection protocols need to be further optimized to improve transfection efficiency and subsequent editing results. Therefore, despite the challenges posed by the experiment in terms of cell number and transfection efficiency, the results provide valuable insights into the feasibility of prime editing in primary CD4+ T cells. Continued refinement of experimental protocols and optimization of editing constructs could potentially help achieve more robust and reproducible prime editing results in future studies. The observed variations in prime editing efficiency between samples highlight the complexity of the prime editing process and the challenges associated with consistently achieving the desired modifications. Factors such as target site accessibility, transfer efficiency, and cellular repair mechanisms can influence editing results, requiring careful optimization of experimental conditions. In addition, the shortcomings of the used analysis methods emphasize the importance of integrating multiple validation strategies to accurately confirm editing events. In summary, while the results provide insights into the performance of prime-editing constructs in primary CD4+ T cells, additional data and validations are essential to improve the efficiency and reliability of editing.

For future experimental approaches, it is recommended to consider obtaining CD4+ primary T cells from a single donor for the entire experimental procedure. Additionally, the software used for evaluation should be tested in advance against an existing or purely synthetic dataset to ensure the accuracy required to qualify the edit. In conclusion, while the study represents a significant step towards understanding prime editing in CD4+ primary T cells, continued efforts are needed to overcome existing challenges and unlock the full potential of prime editing as a precise genome engineering tool in therapeutic and research settings.

References

1. Xiong, Y. & Bosselut, R. CD4-CD8 differentiation in the thymus: connecting circuits and building memories. *Curr Opin Immunol* **24**, 139–145 (2012).
2. Schraivogel, D. *et al.* Targeted Perturb-seq enables genome-scale genetic screens in single cells. *Nat Methods* **17**, 629–635 (2020).
3. Gregersen, P. K. & Olsson, L. M. Recent Advances in the Genetics of Autoimmune Disease. *Annu Rev Immunol* **27**, 363–391 (2009).
4. Richard-Miceli, C. & Criswell, L. A. Emerging patterns of genetic overlap across autoimmune disorders. *Genome Medicine* **4**, 6 (2012).
5. Skapenko, A., Leipe, J., Lipsky, P. E. & Schulze-Koops, H. The role of the T cell in autoimmune inflammation. *Arthritis Research & Therapy* **7**, S4 (2005).
6. Frontiers | Differentiation and Transmigration of CD4 T Cells in Neuroinflammation and Autoimmunity. https://www.frontiersin.org/articles/10.3389/fimmu.2017.01695/full.
7. Alberts, B. *et al.* Helper T Cells and Lymphocyte Activation. in *Molecular Biology of the Cell. 4th edition* (Garland Science, 2002).
8. Chatzileontiadou, D. S. M., Sloane, H., Nguyen, A. T., Gras, S. & Grant, E. J. The Many Faces of CD4+ T Cells: Immunological and Structural Characteristics. *Int J Mol Sci* **22**, 73 (2020).
9. Expanding roles for CD4+ T cells in immunity to viruses | Nature Reviews Immunology. https://www.nature.com/articles/nri3152.
10. Nguyen, M. L. T., Jones, S. A., Prier, J. E. & Russ, B. E. Transcriptional Enhancers in the Regulation of T Cell Differentiation. *Front Immunol* **6**, 462 (2015).
11. Maurya, S. S. Role of Enhancers in Development and Diseases. *Epigenomes* **5**, 21 (2021).
12. Panigrahi, A. & O'Malley, B. W. Mechanisms of enhancer action: the known and the unknown. *Genome Biol* **22**, 108 (2021).
13. Suzuki, A., Guerrini, M. M. & Yamamoto, K. Functional genomics of autoimmune diseases. *Annals of the Rheumatic Diseases* **80**, 689–697 (2021).
14. Farh, K. K.-H. *et al.* Genetic and Epigenetic Fine-Mapping of Causal Autoimmune Disease Variants. *Nature* **518**, 337–343 (2015).

15. McGinn, J. & Marraffini, L. A. Molecular mechanisms of CRISPR-Cas spacer acquisition. *Nat Rev Microbiol* **17**, 7–12 (2019).
16. Wiedenheft, B., Sternberg, S. H. & Doudna, J. A. RNA-guided genetic silencing systems in bacteria and archaea. *Nature* **482**, 331–338 (2012).
17. Hillary, V. E. & Ceasar, S. A. A Review on the Mechanism and Applications of CRISPR/Cas9/Cas12/Cas13/Cas14 Proteins Utilized for Genome Engineering. *Mol Biotechnol* **65**, 311–325 (2023).
18. Rees, H. A. & Liu, D. R. Base editing: precision chemistry on the genome and transcriptome of living cells. *Nat Rev Genet* **19**, 770–788 (2018).
19. Kantor, A., McClements, M. E. & MacLaren, R. E. CRISPR-Cas9 DNA Base-Editing and Prime-Editing. *Int J Mol Sci* **21**, 6240 (2020).
20. Anzalone, A. V. *et al.* Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* **576**, 149–157 (2019).
21. Adikusuma, F. *et al.* Optimized nickase- and nuclease-based prime editing in human and mouse cells. *Nucleic Acids Res* **49**, 10785–10795 (2021).
22. Lu, C. *et al.* Prime Editing: An All-Rounder for Genome Editing. *Int J Mol Sci* **23**, 9862 (2022).
23. Hansen, S., McClements, M. E., Corydon, T. J. & MacLaren, R. E. Future Perspectives of Prime Editing for the Treatment of Inherited Retinal Diseases. *Cells* **12**, 440 (2023).
24. Chen, P. J. *et al.* Enhanced prime editing systems by manipulating cellular determinants of editing outcomes. *Cell* **184**, 5635-5652.e29 (2021).
25. Godbout, K. & Tremblay, J. P. Prime Editing for Human Gene Therapy: Where Are We Now? *Cells* **12**, 536 (2023).
26. Nelson, J. W. *et al.* Engineered pegRNAs improve prime editing efficiency. *Nat Biotechnol* **40**, 402–410 (2022).
27. Zhao, S. *et al.* PiggyBac transposon vectors: the tools of the human gene encoding. *Transl Lung Cancer Res* **5**, 120–125 (2016).
28. Karttunen, K. *et al.* Transposable elements as tissue-specific enhancers in cancers of endodermal lineage. *Nat Commun* **14**, 5313 (2023).

29. Wolff, J. H., Haldrup, J., Thomsen, E. A., Andersen, S. & Mikkelsen, J. G. piggyPrime: High-Efficacy Prime Editing in Human Cells Using piggyBac-Based DNA Transposition. *Frontiers in Genome Editing* **3**, (2021).
30. pegLIT. https://peglit.liugroup.us/.
31. PrimeDesign. https://drugthatgene.pinellolab.partners.org/.
32. Galvan, D. L. *et al.* Genome-Wide Mapping of PiggyBac Transposon Integrations in Primary Human T Cells. *J Immunother* **32**, 837–844 (2009).
33. Maurisse, R. *et al.* Comparative transfection of DNA into primary and transformed mammalian cells from different lineages. *BMC Biotechnology* **10**, 9 (2010).
34. Harris, E., Zimmerman, D., Warga, E., Bamezai, A. & Elmer, J. Nonviral gene delivery to T cells with Lipofectamine LTX. *Biotechnol Bioeng* **118**, 1693–1706 (2021).
35. Kalidasan, V. *et al.* A guide in lentiviral vector production for hard-to-transfect cells, using cardiac-derived c-kit expressing cells as a model system. *Sci Rep* **11**, 19265 (2021).
36. Bilal, M. Y., Vacaflores, A. & Houtman, J. C. D. Optimization of Methods for the Genetic Modification of Human T Cells. *Immunol Cell Biol* **93**, 896–908 (2015).
37. Zhang, M. *et al.* The impact of Nucleofection® on the activation state of primary human CD4 T cells. *J Immunol Methods* **408**, 123–131 (2014).
38. Lesueur, L. L., Mir, L. M. & André, F. M. Overcoming the Specific Toxicity of Large Plasmids Electrotransfer in Primary Cells In Vitro. *Mol Ther Nucleic Acids* **5**, e291 (2016).
39. Zhang, Z., Qiu, S., Zhang, X. & Chen, W. Optimized DNA electroporation for primary human T cell engineering. *BMC Biotechnology* **18**, 4 (2018).
40. Zu, Y., Huang, S., Liao, W.-C., Lu, Y. & Wang, S. Gold Nanoparticles Enhanced Electroporation for Mammalian Cell Transfection. *J Biomed Nanotechnol* **10**, 982–992 (2014).
41. Sukharev, S. I., Klenchin, V. A., Serov, S. M., Chernomordik, L. V. & YuA, C. Electroporation and electrophoretic DNA transfer into cells. The effect of DNA interaction with electropores. *Biophysical Journal* **63**, 1320 (1992).
42. Chen, P. J. *et al.* Enhanced prime editing systems by manipulating cellular determinants of editing outcomes. *Cell* **184**, 5635-5652.e29 (2021).

43. Rad S. M., A. H., Poudel, A., Tan, G. M. Y. & McLellan, A. D. Promoter choice: Who should drive the CAR in T cells? *PLoS One* **15**, e0232915 (2020).
44. Fang, Y. *et al.* Safety and Efficacy of an Immune Cell-Specific Chimeric Promoter in Regulating Anti-PD-1 Antibody Expression in CAR T Cells. *Molecular Therapy - Methods & Clinical Development* **19**, 14–23 (2020).
45. National Academies of Sciences, Engineering, and Medicine; Health and Medicine Division; Board on Health Sciences Policy; Forum on Regenerative Medicine. *Exploring Sources of Variability Related to the Clinical Translation of Regenerative Engineering Products: Proceedings of a Workshop*. (National Academies Press (US), Washington (DC), 2019).
46. Li, Y., Chen, J., Tsai, S. Q. & Cheng, Y. Easy-Prime: a machine learning–based prime editor design tool. *Genome Biology* **22**, 235 (2021).

Appendix and Supplementary Data

Supplementary Table 1: Abbreviations used throughout the text.

AIO	All-in-One Prime Editing Plasmid
AIO-M	All-in-One Prime Editing Plasmid with MLH1dn
MLH1dn	dominant negative MMR protein
epegRNAs	engineerd prime editing guide RNAs
AIO2	All-in-One Prime Editing Plasmid with epeg insert 2
AIO5-M	All-in-One Prime Editing Plasmid with epeg insert 5
AIO1-M	All-in-One Prime Editing Plasmid with MLH1dn with epeg insert 1
AIO2-M	All-in-One Prime Editing Plasmid with MLH1dn with epeg insert 2
AIO4-M	All-in-One Prime Editing Plasmid with MLH1dn with epeg insert 4
AIO5-M	All-in-One Prime Editing Plasmid with MLH1dn with epeg insert 5
AIO8-M	All-in-One Prime Editing Plasmid with MLH1dn with epeg insert 8
epegRNA 1	Exon 3 as target, introduce Stop codon at C41, disrupted PAM
epegRNA 2	Exon 3 as target, introduce Stop codon at C41, disrupted PAM
epegRNA 3	Exon 3 as target, introduce Stop codon at C41, intact PAM
epegRNA 4	Exon 3 as target, introduce frameshift at P20 by deleting CC, disrupted PAM
epegRNA 5	Exon 3 as target, introduce frameshift at P20 by deleting CC, disrupted PAM
epegRNA 6	Exon 3 as target, introduce frameshift at P20 by deleting CC, disrupted PAM
epegRNA 7	Exon 4 as target, introduce frameshift at D88 by adding AC, disrupted PAM
epegRNA 8	Exon 4 as target, introduce frameshift at D88 by adding AC, intact PAM
epegRNA 9	Exon 4 as target, introduce frameshift at D88 by adding AC, disrupted PAM
EGFP	Enhanced Green Fluorescent Protein
bGH poly A	bovine growth hormone polyadenylation signal
IR	Inverted Repeats
PB	PiggyBac prime editing plasmid
PB-M	PiggyBac prime editing plasmid with MLH1dn
RTT	Reverse Transcriptase Template
PBS	Primer Binding Site
pDLI215	Plasmid, contained CAG promotor, for Fragment 1
pDLI42	starting plasmid for Fragment 3

Supplementary Table 2: Abbreviations for plasmids used, their number and short description of properties.

Abbreviation	Number	Usage
PEA1-GFP	543	All in one Prime editing plasmid https://www.addgene.org/171993/
PEA1-Puro	544	All in one Prime editing plasmid https://www.addgene.org/171991/
AIO	545	All in one Prime editing plasmid started from 543, removed secondary gRNA cassette, replaced CMV promoter for CAG full, replaced Cas9 for Cas9PEMax, replaced peg casette with epeg casette, removed additional BbsmI and BbsI sites for golden gate cloning of the pegRNA
AIO-M	546	All in one Prime editing plasmid started from 545, inserted MLH1dn
PB	547	All in one Prime editing plasmid started from 545, inserted piggyBac left inverted repeat and piggyBac right inverted repeat
PB-M	548	All in one Prime editing plasmid started from 545, inserted MLH1dn
AIO1	549	All in one Prime editing plasmid started from 545, inserted pegRNA 1, targeting CD4
AIO2	550	All in one Prime editing plasmid started from 545, inserted pegRNA 2, targeting CD4
AIO3	551	All in one Prime editing plasmid started from 545, inserted pegRNA 3, targeting CD4
AIO4	552	All in one Prime editing plasmid started from 545, inserted pegRNA 4, targeting CD4
AIO5	553	All in one Prime editing plasmid started from 545, inserted pegRNA 5, targeting CD4
AIO6	554	All in one Prime editing plasmid started from 545, inserted pegRNA 6, targeting CD4
AIO7	555	All in one Prime editing plasmid started from 545, inserted pegRNA 7, targeting CD4
AIO8	556	All in one Prime editing plasmid started from 545, inserted pegRNA 8, targeting CD4
AIO9	557	All in one Prime editing plasmid started from 545, inserted pegRNA 9, targeting CD4
AIO1-M	558	All in one Prime editing plasmid started from 546, inserted pegRNA 1, targeting CD4
AIO2-M	559	All in one Prime editing plasmid started from 546, inserted pegRNA 2, targeting CD4
AIO3-M	560	All in one Prime editing plasmid started from 546, inserted pegRNA 3, targeting CD4
AIO4-M	561	All in one Prime editing plasmid started from 546, inserted pegRNA 4, targeting CD4
AIO5-M	562	All in one Prime editing plasmid started from 546, inserted pegRNA 5, targeting CD4
AIO6-M	563	All in one Prime editing plasmid started from 546, inserted pegRNA 6, targeting CD4
AIO7-M	564	All in one Prime editing plasmid started from 546, inserted pegRNA 7, targeting CD4
AIO8-M	565	All in one Prime editing plasmid started from 546, inserted pegRNA 8, targeting CD4
AIO9-M	566	All in one Prime editing plasmid started from 546, inserted pegRNA 9, targeting CD4
PB1	567	All in one Prime editing plasmid started from 547, inserted pegRNA 1, targeting CD4
PB2	568	All in one Prime editing plasmid started from 547, inserted pegRNA 2, targeting CD4
PB3	569	All in one Prime editing plasmid started from 547, inserted pegRNA 3, targeting CD4
PB4	570	All in one Prime editing plasmid started from 547, inserted pegRNA 4, targeting CD4
PB5	571	All in one Prime editing plasmid started from 547, inserted pegRNA 5, targeting CD4
PB6	572	All in one Prime editing plasmid started from 547, inserted pegRNA 6, targeting CD4
PB7	573	All in one Prime editing plasmid started from 547, inserted pegRNA 7, targeting CD4
PB8	574	All in one Prime editing plasmid started from 547, inserted pegRNA 8, targeting CD4
PB9	575	All in one Prime editing plasmid started from 547, inserted pegRNA 9, targeting CD4
PB1-M	576	All in one Prime editing plasmid started from 548, inserted pegRNA 1, targeting CD4
PB2-M	577	All in one Prime editing plasmid started from 548, inserted pegRNA 2, targeting CD4
PB3-M	578	All in one Prime editing plasmid started from 548, inserted pegRNA 3, targeting CD4
PB4-M	579	All in one Prime editing plasmid started from 548, inserted pegRNA 4, targeting CD4
PB5-M	580	All in one Prime editing plasmid started from 548, inserted pegRNA 5, targeting CD4
PB6-M	581	All in one Prime editing plasmid started from 548, inserted pegRNA 6, targeting CD4
PB7-M	582	All in one Prime editing plasmid started from 548, inserted pegRNA 7, targeting CD4
PB8-M	583	All in one Prime editing plasmid started from 548, inserted pegRNA 8, targeting CD4
PB9-M	584	All in one Prime editing plasmid started from 548, inserted pegRNA 9, targeting CD4

Supplementary Table 3: Oligos list used for all experiments.

Oligo #	Usage	Sequence
1171	used for Fragment 4, removed *BsmBI* sites FW	CAGACAAGCTGTGACCGTCTGCGGGAGCTGCATGTG TCAGAG
1172	used for Fragment 4, removed *BsmBI* sites RV	GTCATCACCGAAACGCGCGAAACGAAAGGGCCTCGT GATACGC
1173	used für Fragment 3, remove *BbsI* and *BsmBI* sites from pDLI142 FW	GCGGCGGAAGCACCCTGAACATTGAGGACGAGTATA GACTGCATG
1174	used für Fragment 3, remove *BbsI* and *BsmBI* sites from pDLI142 RV	GAAGCTGGGACCATGGAGAAGGCCCGTGGCCTATCT GTC
1175	amplification of Fragment 2 FW	CTGGCCTTTTGCTCACATGTAGGGCCTATTTCCCATG ATTCCT
1176	amplification of Fragment 2 RV	AATGACCCCGTAATTGATTACTATTAATAAC TAGTCAATTGTCGACGTTTAAACTGCT
1177	amplification of the Fragment 3, with remove *BbsI* and *BsmBI* sites from pDLI142 FW	TGCTGGTTATTGTGCTGTCTCATCATTTTG GCAAAGGTACGACAAGAAGTACAGCATCGG
1178	amplification of the Fragment 3, with remove *BbsI* and *BsmBI* sites from pDLI142 RV	TTCGTGGCCGCCGGCCTTTTAGACTCGAATTCGCTG CCG
1179	amplification of the Fragment 4, with remove *BsmBI* sites from PEA1-GFP FW	ACGGCAGCGAATTCGAGTCTAAAAGGCCGGCG
1180	amplification of the Fragment 4, with remove *BsmBI* sites from PEA1-GFP RV	AATCATGGGAAATAGGCCCTACATGTGAGCAAAAG
1181	PEA1-GFP_final_Alt_Fragment4.REV	AGGCGTATCACGAGGCCCTTGTCACAGCTTGTCTGTA AGCGG
1182	PEA1-GFP_final_Alt_Fragment5.FOR	GCTTACAGACAAGCTGTGACAAGGGCCTCGTGATAC GC
1183	amplification of the Fragment 5 from pDLI214 FW	AAAAGGAATTCGGCAGTGGAGCTACTAACTTCAGCCT GC
1184	amplification of the Fragment 5 from pDLI214 RV	AGCAGACTTCCTCTGCCCTCTCCCGATCCGTT
1187	Barcode for NGS sample 1 AIO1-M	AATGATACGGCGACCACCGAGATCTACATAGATC GCACACTCTTTCCCTACACGACGCTCTTCCG
1188	Barcode for NGS sample 2 AIO4-M	AATGATACGGCGACCACCGAGATCTACACTCTCT ATACACTCTTTCCCTACACGACGCTCTTCCG

1190	Barcode for NGS sample 3 AIO8-M	AATGATACGGCGACCACCGAGATCTACAAGA GTAGAACACTCTTTCCCTACACGACGCTCTTCCG
1199	PegRNA_Scaffold_FW	AGAGCTAAGCTGGAAACAGCATAGCAAGTTTAAATAA GG CTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTC G
1200	PegRNA_Scaffold_RV	GCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGG ACTAGCCTTATTTAAACTTGCTATGCTGTTTCCAGCTT AG
1201	epegRNA 1 single guide RNA FW used for AIO1, AIO1-M, PB1, PB1-M	CACCGCTTCTTCTGGGAAGCTGTACGTTTA
1202	epegRNA 1 single guide RNA RV used for AIO1, AIO1-M, PB1, PB1-M	CTCTTAAACGTACAGCTTCCCAGAAGAAGC
1203	epegRNA 1 extension RNA FW	GTGCAACTGACCTAGACAGCTTCCCAGAACCACAATC
1204	epegRNA 1 extension RNA RV	CGCGGATTGTGGTTCTGGGAAGCTGTCTAGGTCAGT T
1205	epegRNA 2 single guide RNA FW used for AIO2, AIO2-M, PB2, PB2-M	CACCGCTTCTTCTGGGAAGCTGTACGTTTA
1206	epegRNA 2 single guide RNA RV used for AIO2, AIO2-M, PB2, PB2-M	CTCTTAAACGTACAGCTTCCCAGAAGAAGC
1207	epegRNA 2 extension RNA FW	GTGCACAGTGGAACTGACATAGACAGCTTCCCAGCC AACCTC
1208	epegRNA 2 extension RNA RV	CGCGGAGGTTGGCTGGGAAGCTGTCTATGTCAGTTC CACTGT
1209	epegRNA 3 single guide RNA FW used for AIO3, AIO3-M, PB3, PB3-M	CACCGAATTGTATGCTCTTCTTCTGTTTA
1210	epegRNA 3 single guide RNA RV used for AIO3, AIO3-M, PB3, PB3-M	CTCTTAAACAGAAGAAGAGCATACAATTC
1211	epegRNA 3 extension RNA FW	GTGCTAGACAGCTTCCCAGAAGAAGAGCATACACCC AGACC
1212	epegRNA 3 extension RNA RV	CGCGGGTCTGGGTGTATGCTCTTCTTCTGGGAAGCT GTCTA
1213	epegRNA 4 single guide RNA FW used for AIO4, AIO4-M, PB4, PB4-M	CACCGTTCCCTGAGTGGCTGCTGGGGTTTA
1214	epegRNA 4 single guide RNA RV used for AIO4, AIO4-M, PB4, PB4-M	CTCTTAAACCCCAGCAGCCACTCAGGGAAC
1215	epegRNA 4 extension RNA FW	GTGCTCCACTTAGCGCTCCTCAGCAGCCACTCTCCTT CTC
1216	epegRNA 4 extension RNA RV	CGCGGAGAAGGAGAGTGGCTGCTGAGGAGCGCTAA GTGGA
1217	epegRNA 5 single guide RNA FW used for AIO5, AIO5-M, PB5, PB5-M	CACCGTTCCCTGAGTGGCTGCTGGGGTTTA

1218	epegRNA 5 single guide RNA RV used for AIO5, AIO5-M, PB5, PB5-M	CTCTTAAACCCCAGCAGCCACTCAGGGAAC
1219	epegRNA 5 extension RNA FW	GTGCTTAGCGCTCCTCAGCAGCCACTCAGATCCTTAA
1220	epegRNA 5 extension RNA RV	CGCGTTAAGGATCTGAGTGGCTGCTGAGGAGCGCTAA
1249	epegRNA 6 single guide RNA FW used for AIO6, AIO6-M, PB6, PB6-M	CACCGCACCACTTTCTTTCCCTGAGGTTTA
1250	epegRNA 6 single guide RNA RV used for AIO6, AIO6-M, PB6, PB6-M	CTCTTAAACCTCAGGGAAAGAAAGTGGTGC
1223	epegRNA 6 extension RNA FW	GTGCTCCTCAGCAGCCACTCAGGGAAAGAAAGTCATAAATA
1224	epegRNA 6 extension RNA RV	CGCGTATTTATGACTTTCTTTCCCTGAGTGGCTGCTGAGGA
1225	epegRNA 7 single guide RNA FW used for AIO7, AIO7-M, PB7, PB7-M	CACCGAGAAGAAGCCTTTGGGACCAGTTTA
1226	epegRNA 7 single guide RNA RV used for AIO7, AIO7-M, PB7, PB7-M	CTCTTAAACTGGTCCCAAAGGCTTCTTCTC
1227	epegRNA 7 extension RNA FW	GTGCGGAAAGTTTCCTTGGGTTCCCAAAGGCTTGAATATAA
1228	epegRNA 7 extension RNA RV	CGCGTTATATTCAAGCCTTTGGGAACCCAAGGAAACTTTCC
1229	epegRNA 8 single guide RNA FW used for AIO8, AIO8-M, PB8, PB8-M	CACCGAGAAGAAGCCTTTGGGACCAGTTTA
1230	epegRNA 8 single guide RNA RV used for AIO8, AIO8-M, PB8, PB8-M	CTCTTAAACTGGTCCCAAAGGCTTCTTCTC
1231	epegRNA 8 extension RNA FW	GTGCAGTTTCCTTGGGTTCCCAAAGGCTTAGAATAAC
1232	epegRNA 8 extension RNA RV	CGCGGTTATTCTAAGCCTTTGGGAACCCAAGGAAACT
1233	epegRNA 9 single guide RNA FW used for AIO9, AIO9-M, PB9, PB9-M	CACCGTGATCAGGGGAAAGTTTCCTGTTTA
1234	epegRNA 9 single guide RNA RV used for AIO9, AIO9-M, PB9, PB9-M	CTCTTAAACAGGAAACTTTCCCCTGATCAC
1235	epegRNA 9 extension RNA FW	GTGCTGGGAACCCAAGGAAACTTTCCCCTAACTGCTA
1236	epegRNA 9 extension RNA RV	CGCGTAGCAGTTAGGGGAAAGTTTCCTTGGGTTCCCA
1237	PiggyBac Inverted Repaet right side (betweeen *XmaI* and *PcI*) FW	CCGGCATGCGTCAATTTTACACATGATTATCTTTAACGTACGTCACAATATGATTATCTTTCTAGGG
1238	PiggyBac Inverted Repaet right side	CCGGCCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATGTGTAAAATTGACGCATG

	(betweeen *XmaI* and *PcI*) RV	
1239	PiggyBac Inverted Repaet left side (betweeen *hU6 promotor* and *PcI*) FW	CATGCCCTAGAAAGATAGTCTGCGTAAAATTGACGCATG
1240	PiggyBac Inverted Repaet left side (betweeen *hU6 promotor* and *PcI*) RV	CATGCATGCGTCAATTTTACGCAGACTATCTTTCTAGGG
717	colony PCR primer for pegRNA insert RV, colony PCR primer for PiggyBac right side RV	AATGGACTATCATATGCTTACCGT
1243	Primer for sequencing of PiggyBac Inverted Repaet right side (betweeen *XmaI* and *PcI*)	GTCGACGTTTAAACTGCTATCC
1244	Primer for sequncing of PiggyBac Inverted Repaet left side (betweeen *hU6 promotor* and *PcI*) FW	CGTCCTTTCCACAAGATATATAAAGCCAA
1245	colony PCR primer for pegRNA insert FW	AGCTGGAAACAGCATAGCAAG
1246	colony PCR primer for PiggyBac right side FW	ATGCGTCAATTTTACACATGATTATCTTT
1247	colony PCR primer for PiggyBac left side FW	CCCTAGAAAGATAGTCTGCGTA
1248	colony PCR primer for PiggyBac left side RV	TCCAATTATCTCTCTAACAGCC
1291	PCR1, adapter NGS library preparatio, sample 1 AIO1-M staggered 1 Primer FW	CTCTTTCCCTACACGACGCTCTTCCGATCTCTTTTCCACTTAGCGCTCCTC
1292	PCR1, adapter NGS library preparatio, sample 1 AIO1-M staggered 2 Primer FW	CTCTTTCCCTACACGACGCTCTTCCGATCTACTTTTCCACTTAGCGCTCCTC
1293	PCR1, adapter NGS library preparatio, sample 1 AIO1-M staggered 3 Primer FW	CTCTTTCCCTACACGACGCTCTTCCGATCTTCCTTTTCCACTTAGCGCTCCTC
1294	PCR1, adapter NGS library preparatio, sample 1 AIO1-M staggered 4 Primer FW	CTCTTTCCCTACACGACGCTCTTCCGATCTGACCTTTTCCACTTAGCGCTCCTC
1295	PCR1, adapter NGS library preparatio, sample 1 AIO1-M staggered 1 Primer RV	CTGGAGTTCAGACGTGTGCTCTTCCGATCTAGGAGCCCTGATTTCCCAGA
1296	PCR1, adapter NGS library preparatio,	CTGGAGTTCAGACGTGTGCTCTTCCGATCTCAGGAGCCCTGATTTCCCAGA

	sample 1 AIO1-M staggered 2 Primer RV	
1297	PCR1, adapter NGS library preparatio, sample 1 AIO1-M staggered 3 Primer RV	CTGGAGTTCAGACGTGTGCTCTTCCGATCTCTAGGAGCCCTGATTTCCCAGA
1298	PCR1, adapter NGS library preparatio, sample 1 AIO1-M staggered 4 Primer RV	CTGGAGTTCAGACGTGTGCTCTTCCGATCTGATAGGAGCCCTGATTTCCCAGA
1299	PCR1, adapter NGS library preparatio, sample 2 AIO4-M staggered 1 Primer FW	CTCTTTCCCTACACGACGCTCTTCCGATCTCAAAGGTGGAGGATGGGGTAG
1300	PCR1, adapter NGS library preparatio, sample 2 AIO4-M staggered 2 Primer FW	CTCTTTCCCTACACGACGCTCTTCCGATCTCCAAAGGTGGAGGATGGGGTAG
1301	PCR1, adapter NGS library preparatio, sample 2 AIO4-M staggered 3 Primer FW	CTCTTTCCCTACACGACGCTCTTCCGATCTTACAAAGGTGGAGGATGGGGTAG
1302	PCR1, adapter NGS library preparatio, sample 2 AIO4-M staggered 4 Primer FW	CTCTTTCCCTACACGACGCTCTTCCGATCTCGTCAAAGGTGGAGGATGGGGTAG
1303	PCR1, adapter NGS library preparatio, sample 2 AIO4-M staggered 1 Primer RV	CTGGAGTTCAGACGTGTGCTCTTCCGATCTGGTCAGTTCCACTGTATCCCC
1304	PCR1, adapter NGS library preparatio, sample 2 AIO4-M staggered 2 Primer RV	CTGGAGTTCAGACGTGTGCTCTTCCGATCTAGGTCAGTTCCACTGTATCCCC
1305	PCR1, adapter NGS library preparatio, sample 2 AIO4-M staggered 3 Primer RV	CTGGAGTTCAGACGTGTGCTCTTCCGATCTCAGGTCAGTTCCACTGTATCCCC
1306	PCR1, adapter NGS library preparatio, sample 2 AIO4-M staggered 4 Primer RV	CTGGAGTTCAGACGTGTGCTCTTCCGATCTCGAGGTCAGTTCCACTGTATCCCC
1307	PCR1, adapter NGS library preparatio, sample 3 AIO8-M staggered 1 Primer FW	CTCTTTCCCTACACGACGCTCTTCCGATCTCTTCTGCTCCCAGGTCCATC
1308	PCR1, adapter NGS library preparatio, sample 3 AIO8-M staggered 2 Primer FW	CTCTTTCCCTACACGACGCTCTTCCGATCTTCTTCTGCTCCCAGGTCCATC
1309	PCR1, adapter NGS library preparatio, sample 3 AIO8-M staggered 3 Primer FW	CTCTTTCCCTACACGACGCTCTTCCGATCTCACTTCTGCTCCCAGGTCCATC

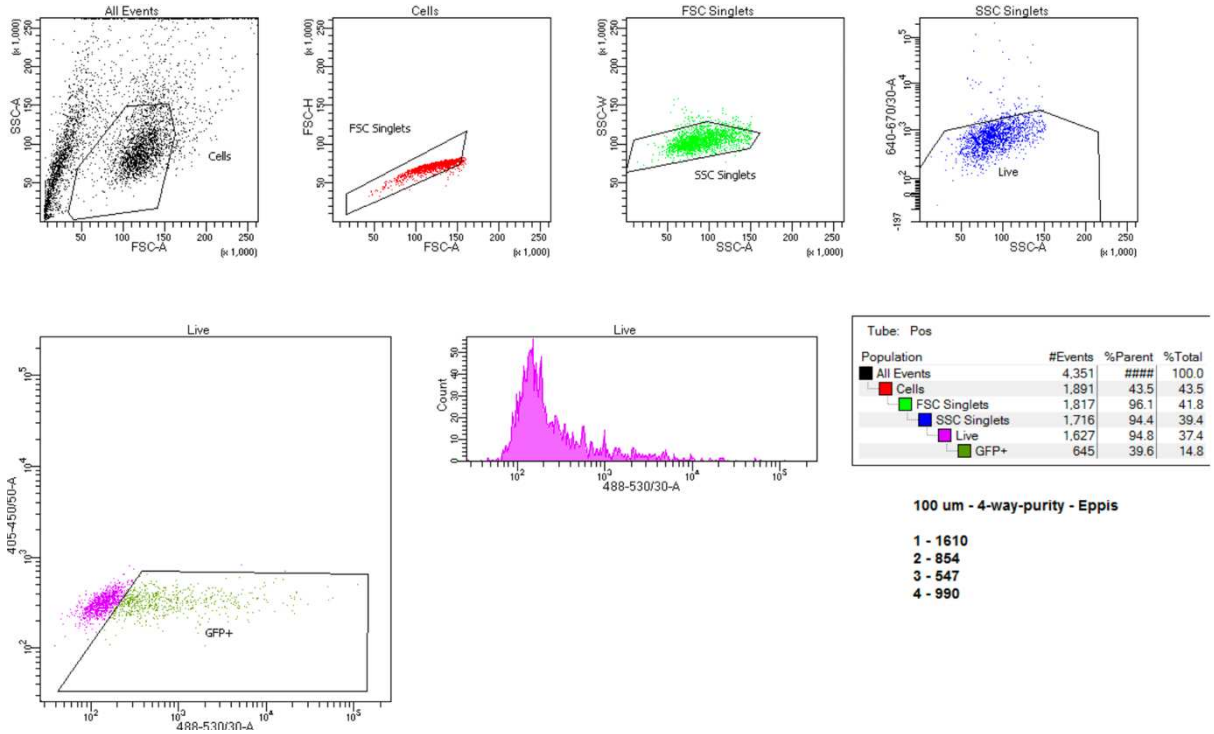
1310	PCR1, adapter NGS library preparatio, sample 3 AIO8-M staggered 4 Primer FW	CTCTTTCCCTACACGACGCTCTTCCGATCTCATCTTCTGCTCCCAGGTCCATC
1311	PCR1, adapter NGS library preparatio, sample 3 AIO8-M staggered 1 Primer RV	CTGGAGTTCAGACGTGTGCTCTTCCGATCTCTTCTGGTCCTCCACTTCACA
1312	PCR1, adapter NGS library preparatio, sample 3 AIO8-M staggered 2 Primer RV	CTGGAGTTCAGACGTGTGCTCTTCCGATCTGCTTCTGGTCCTCCACTTCACA
1313	PCR1, adapter NGS library preparatio, sample 3 AIO8-M staggered 3 Primer RV	CTGGAGTTCAGACGTGTGCTCTTCCGATCTACCTTCTGGTCCTCCACTTCACA
1314	PCR1, adapter NGS library preparatio, sample 3 AIO8-M staggered 4 Primer RV	CTGGAGTTCAGACGTGTGCTCTTCCGATCTATCCTTCTGGTCCTCCACTTCACA

Supplementary Table 4: Percentages of transfection efficiency and viability post electroporation.

Viability in %			
Day post-electroporation	Negative ctrl	Positive ctrl	AIO
2 day	95	25	11
4 day	94	21	2
9 day	96	0	0
Transfection Efficiency in %			
Day post-electroporation	Negative ctrl	Positive ctrl	AIO
2 day	0	53	2
4 day	0	52	9
9 day	0	30	4

Supplementary Table 5: Detailed Bioanalyzer Data.

	amount of cells	Expected Size [bp]	Size [bp]	Conc. [pg/µl]	Molarity [pmol/
AIO1-M	1610	302	307	121	597
AIO4-M	854	277	282	221	1191
AIO5-M	547	277	285	91	486
AIO8-M	990	280	292	391	2029



Supplementary Figure 1: Detailed results of the cell sorting process, of the positive control (pmaxGFP). Shown on the top, density plots of the positive control (ctrl+) in black, red, green and blue, the gating strategy. Cells were sorted on day 6 post-electroporation with four different All-In-One plasmids (AIO1-M (1610 cells), AIO4-M (854 cells), AIO5-M (547 cells) and AIO8-M (990 cells)). The plasmids carried a EGFP sequence that is expressed upon uptake by the cells and used as an indicator of successful transfection. For the negative control (ctrl-) the cells were not treated with any plasmid and only underwent he electroporation. For ctrl+ the cells were electroporated with a pmaxGFP plasmid.