EGFR Gene Editing via CRISPR/Cas9 and Prime Editing

Eunju Kim[1,*], Jeehoon Kim[1,*], Sehri Oh[1,*], Junsuk Lee[1,2], Changho Sohn[1,2,**], and Jinwoo Cheon[1,2,3,**]
[1]Center for Nanomedicine, Institute for Basic Science (IBS), Seoul, Republic of Korea
[2]Graduate Program of Nano Biomedical Engineering (NanoBME), Yonsei Advanced Science Institute, Yonsei University, Seoul, Republic of Korea
[3]Department of Chemistry, Yonsei University, Seoul, Republic of Korea
*These authors contributed equally to this work.

Introduction

Harnessing the power of CRISPR/Cas9 system as a gene editing tool has greatly revolutionized the field, owing to its ease of design and high sensitivity. CRISPR/Cas9 system effectively circumvents the limitations of previous gene editing tools and has paved the way for a new era of genome editing. However, more research is pointing to CRISPR/Cas9's own limitations. One of the major flaws of this technology is the relatively high frequency of off-target effects. To overcome these hurdles, a 4th generation gene editing tool called Prime Editing (PE) has been recently developed. Prime editing has three components: Cas9 nickase, reverse transcriptase, and the prime editing guide RNA (pegRNA). The pegRNA can be further broken down into the spacer sequence, scaffold, primer binding site, and reverse transcription (RT) template. The RT template already includes the desired sequence that can be transcribed by reverse transcriptase. This newly synthesized DNA replaces the original strand, yielding an extremely precise editing as opposed to the random insertion/deletion knockout generated by the CRISPR/Cas9 system. To this end, we conducted a series of studies to compare the editing efficiency of the CRISPR/Cas9 system and prime editing. The knockout of the target gene EGFR by CRISPR/Cas9 system has been verified by T7E1 assay and a plasmid reporter system as evidenced by a strong green fluorescence signal. The editing rate of prime editing, along with that of CRISPR/Cas9, has been quantified by next generation sequencing. NGS data displays a high editing frequency for the CRISPR/Cas9 system, whereas editing by prime editing was not detected. One of the possible reasons for such outcome is the lack of high-throughput experiment to optimize the pegRNA components specific to the EGFR gene.

Scheme

Fig. 1 | Overall Scheme

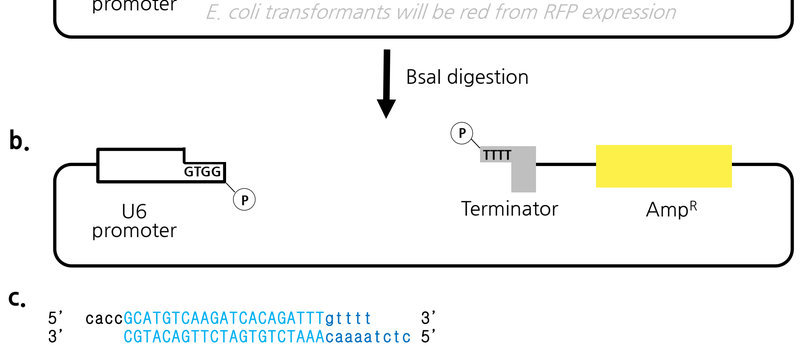


Fig. 2 | EGFR reporter system (Bae *et al.* 2019)

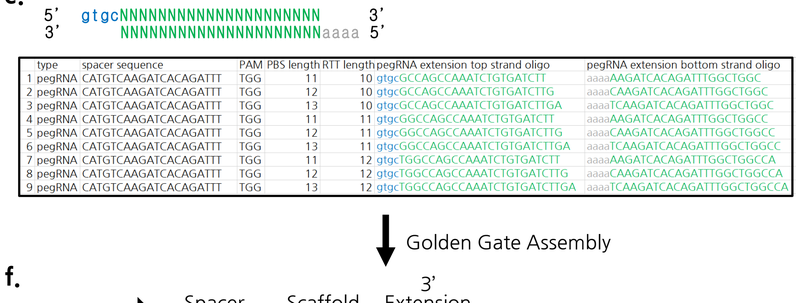


Fig. 3 | pegRNA design and Golden Gate Assembly

type	spacer sequence	PAM	PBS length	RTT length	pegRNA extension top strand oligo	pegRNA extension bottom strand oligo
1 pegRNA	CATGTCAAGATCACAGATTT	TGG	11	10	gtgcGCCAGCCAAATCTGTGATCTT	aaaaAAGATCACAGATTTGGCTGGC
2 pegRNA	CATGTCAAGATCACAGATTT	TGG	12	10	gtgcGCCAGCCAAATCTGTGATCTTG	aaaaCAAGATCACAGATTTGGCTGGC
3 pegRNA	CATGTCAAGATCACAGATTT	TGG	13	10	gtgcGCCAGCCAAATCTGTGATCTTGA	aaaaTCAAGATCACAGATTTGGCTGGC
4 pegRNA	CATGTCAAGATCACAGATTT	TGG	11	11	gtgcGGCCAGCCAAATCTGTGATCTT	aaaaAAGATCACAGATTTGGCTGGCC
5 pegRNA	CATGTCAAGATCACAGATTT	TGG	12	11	gtgcGGCCAGCCAAATCTGTGATCTTG	aaaaCAAGATCACAGATTTGGCTGGCC
6 pegRNA	CATGTCAAGATCACAGATTT	TGG	13	11	gtgcGGCCAGCCAAATCTGTGATCTTGA	aaaaTCAAGATCACAGATTTGGCTGGCC
7 pegRNA	CATGTCAAGATCACAGATTT	TGG	11	12	gtgcTGGCCAGCCAAATCTGTGATCTT	aaaaAAGATCACAGATTTGGCTGGCCA
8 pegRNA	CATGTCAAGATCACAGATTT	TGG	12	12	gtgcTGGCCAGCCAAATCTGTGATCTTG	aaaaCAAGATCACAGATTTGGCTGGCCA
9 pegRNA	CATGTCAAGATCACAGATTT	TGG	13	12	gtgcTGGCCAGCCAAATCTGTGATCTTGA	aaaaTCAAGATCACAGATTTGGCTGGCCA

Fig. 3 | a, pegRNA acceptor vector (U6-pegRNA-GG-acceptor). **b,** BsaI digested backbone. **c,** pegRNA spacer annealed oligonucleotides. **d,** pegRNA scaffold annealed oligonucleotides (5' phosphorylated). **e,** EGFR-targeting nine pegRNA 3' extension annealed oligonucleotides' sequences (Online pegRNA design platform: 'Prime Design', https://drugthatgene.pinellolab.partners.org). **f,** assembled pegRNA vector.

Methods

Fig. 4 | Next Generation Sequencing - Illumina MiniSeq

a. Library preparation for deep-sequencing

P5	Index	Rd2 SP	DNA insert	Rd2 SP	Index	P7
P5	Index1	Rd2 SP	Genomic EGFR edited via CRISPR/Cas9 system	Rd2 SP	Index1'	P7
P5	Index2	Rd2 SP	Plasmid EGFR edited via CRISPR/Cas9 system	Rd2 SP	Index2'	P7
P5	Index3	Rd2 SP	Genomic chromosome 16 target gene via Prime editing system	Rd2 SP	Index3'	P7
P5	Index4	Rd2 SP	Genomic EGFR edited via CRISPR/Cas9 system	Rd2 SP	Index4'	P7
P5	Index5	Rd2 SP	Plasmid EGFR edited via CRISPR/Cas9 system	Rd2 SP	Index5'	P7

b. Overall workflow

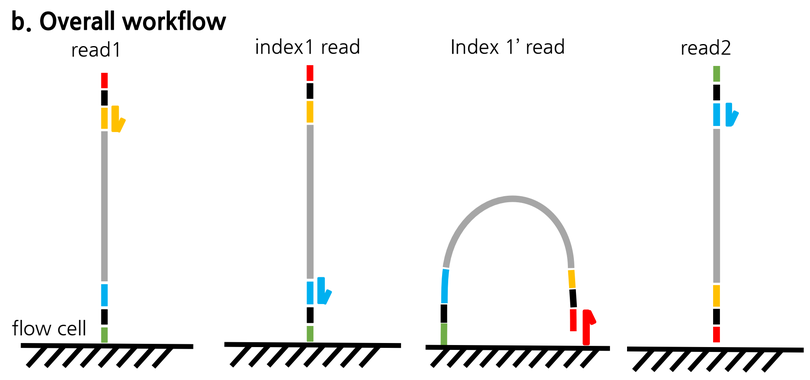
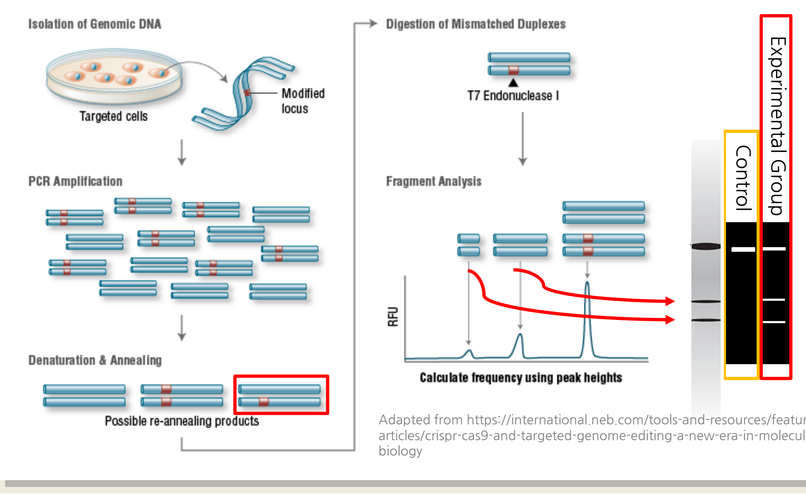


Fig. 5 | T7E1 assay



Adapted from https://international.neb.com/tools-and-resources/feature-articles/crispr-cas9-and-targeted-genome-editing-a-new-era-in-molecular-biology

Results

Fluorescence microscopy

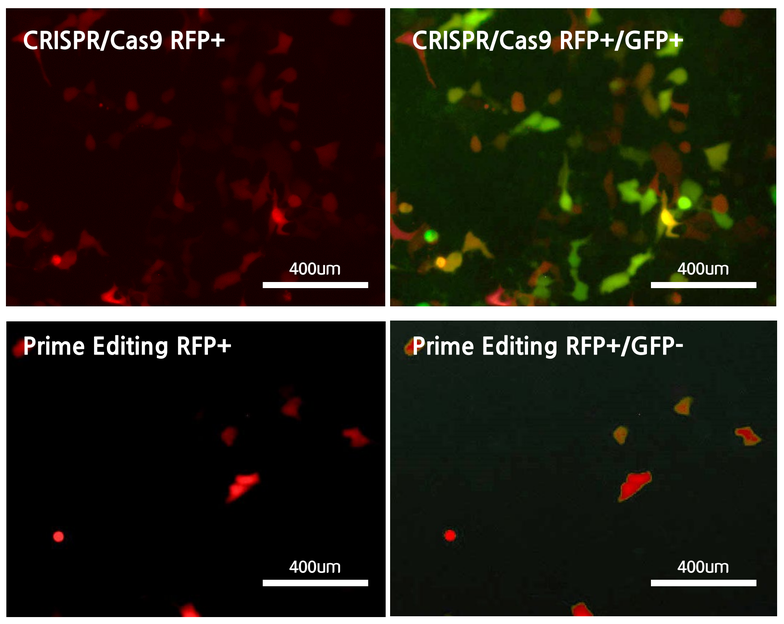


Fig. 6 | HEK293 cells transfected with reporter plasmid imaged by confocal microscope. Red indicates red fluorescence by constitutive RFP expression. Green indicates either a successful insertion of 1bp or deletion of 2bp in the target region of the reporter plasmid.

FACS (CRISPR/Cas9)

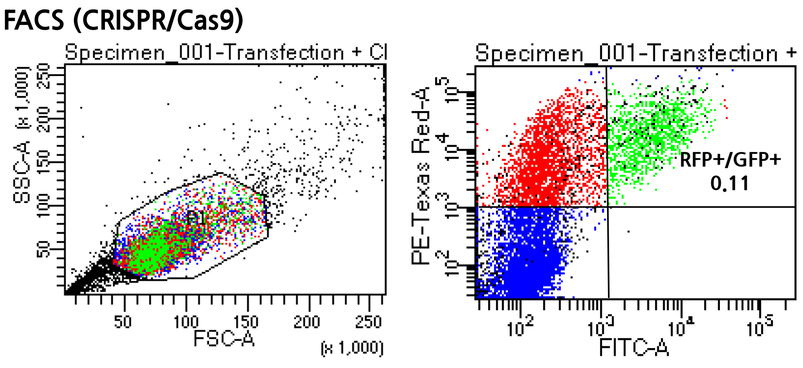


Fig. 7 | HEK293 cells sorted by FACS. Number inside the gate represents the proportion of GFP+ cells as a percentage of the total cell population.

T7E1 assay (CRISPR/Cas9)

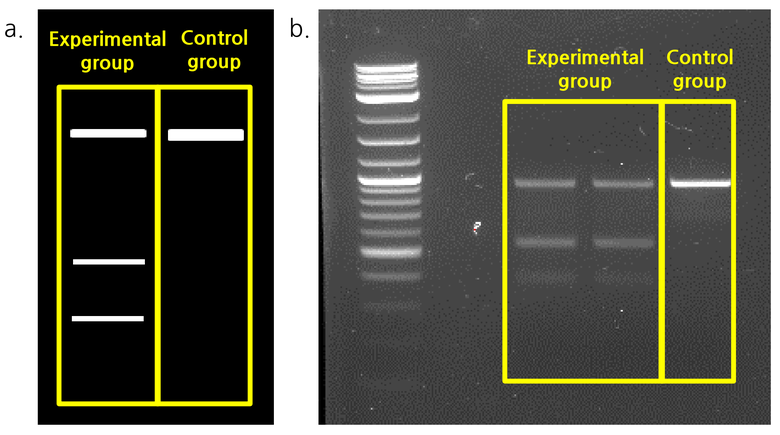


Fig. 8 | a, expected gel image result. **b,** actual gel image data.

Deep sequencing data (CRISPR/Cas9)

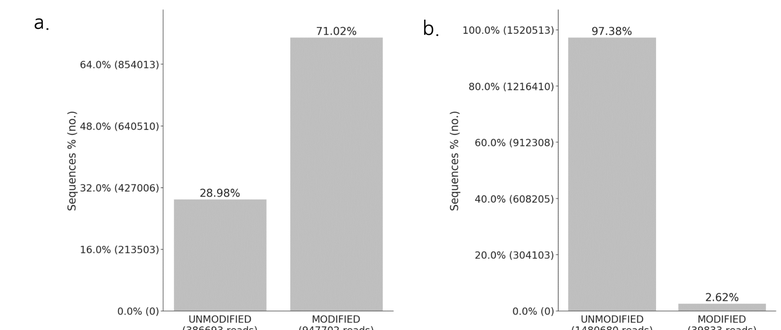


Fig. 9 | Alignment and editing frequency of reads as determined by the percentage and number of sequence reads showing unmodified and modified alleles. a, CRISPR/Cas9 editing reads on genomic DNA. **b,** CRISPR/Cas9 editing reads on reporter plasmid.

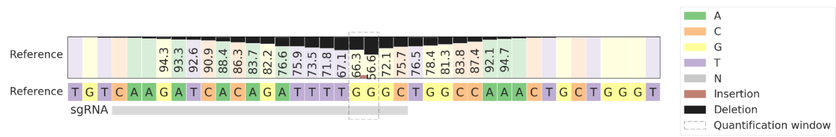


Fig. 10 | Genomic DNA nucleotide distribution around the sgRNA target sequence.

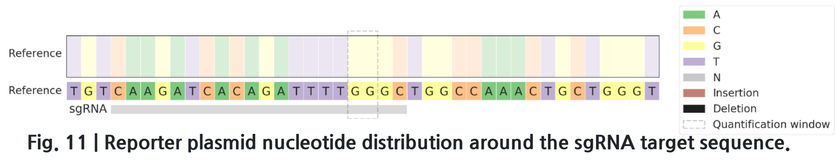
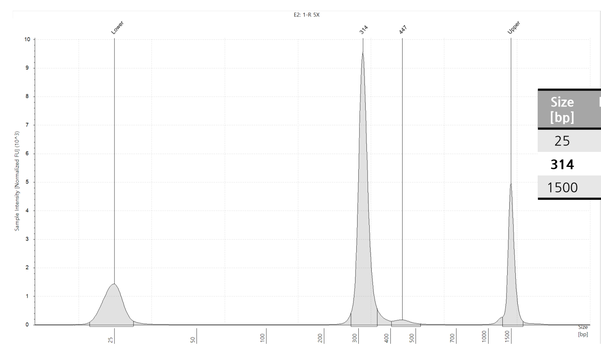


Fig. 11 | Reporter plasmid nucleotide distribution around the sgRNA target sequence.



Peak Table

Size [bp]	Peak Molarity [pmol/l]	Observations
25	365	Lower Marker
314	**91.1**	
1500	6.67	Upper Marker

Fig. 12 | 2nd PCR amplicon length quantification by automated gel electrophoresis. (TapeStation)

Discussion

We explored two types of gene editing tools: CRISPR/Cas9 and prime editing by targeting the EGFR gene. The CRISPR/Cas9 system has been tested by transfecting HEK293 cells with EGFR reporter plasmid, EGFR-targeting sgRNA plasmid, and cas9 plasmid. Similarly, prime editing was tested by transfecting the cells with EGFR reporter plasmid, EGFR-targeting pegRNA plasmid, and prime editor plasmid. We synthesized 9 candidates of novel EGFR-targeting pegRNAs by inserting each component (spacer, scaffold, PBS/RT template) to an acceptor vector through Golden Gate Assembly, after which the assembled pegRNA vector sequences were confirmed by Sanger sequencing. A total of four validation methods were used to verify the desired edits on both the genomic and plasmid levels: 1) RFP/GFP reporter, 2) T7E1 assay, 3) Sanger sequencing, and 4) next-generation sequencing. The reporter system and T7E1 assay both showed positive results for the CRISPR/Cas9 system, indicating the successful knockout of the EGFR gene by random insertion/deletion, while prime editing only yielded negative results. Sanger sequencing utilizes primers flanking the target region of the gene that are used to amplify the products by PCR, which are then integrated into a vector by blunt end cloning and submitted for sequencing. However, sanger sequencing may not be optimal for analyzing a heterogeneous group of populations since the colonies subject to sequencing may not accurately reflect the entire gene pool. As such, Sanger sequencing showed only negative results for both CRISPR/Cas9 and prime editing treated groups. NGS, on the other hand, is capable of distinguishing a rare variant in a given group of gene populations. Analysis of CRISPR/Cas9 treated groups showed 71.02% editing frequency, whereas prime editing did not yield any significant edit reads. One of the major challenges for targeting a specific gene via prime editing is designing the optimal pegRNA that can specifically edit the target gene. The 9 pegRNA designs that were used in this study was largely based on mathematical modeling proposed in a previous pegRNA library study where PBS length of 11 to 13 nt and RT length of 10 to 12 nt was recommended based on the highest average efficiency(Kim *et al.* 2020). However, this does not take into account the characteristics specific to our target gene conferred by the individual sequences such as the GC contents or the thermodynamic stability that dictates the integration of 3' flap containing the edit. The most effective way to optimize the combination of PBS and RT template length is to survey variable lengths of pegRNAs via high-throughput screening specific to the given target gene. Editing types other than deletion may also be considered to alter the gene such as insertions or substitutions to safely convert the mutant EGFR gene into a normal allele and maximize editing efficiency.

References

Anzalone, A.V., Randolph, P.B., Davis, J.R. et al. Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* 576, 149-157 (2019).
Bae T, Kim H, Kim JH, et al. Specificity Assessment of CRISPR Genome Editing of Oncogenic EGFR Point Mutation with Single-Base Differences. *Molecules* 25(1):52 (2019).
Kim, H.K., Yu, G., Park, J. et al. Predicting the efficiency of prime editing guide RNAs in human cells. *Nat Biotechnol.* 39, 198-206 (2020).