PRECISE GENOME ENGINEERING VIA PROGRAMMABLE BASE EDITING

Reported by Emily J. Geddes February 22, 2018

INTRODUCTION

Modern genome editing has been revolutionized by nuclease-based genome editing techniques, particularly the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) technology. This system harnesses the DNA binding affinity and endonuclease activity of the Cas9 protein to target a locus of interest and introduce precise genetic modifications. Cas9 generates double-stranded (ds) DNA breaks at the target locus, which can be repaired through the efficient error-prone nonhomologous end-joining (NHEJ) pathway, leading to gene disruption, or the inefficient high-fidelity homology-directed repair (HDR) pathway, leading to gene correction.[1] Applications of CRISPR-Cas9 have enabled the broad study of gene function through target gene inactivation and gene replacement; additionally, it is now being tested in clinical trials for the treatment of genetic diseases. While extensive optimization of these methods has been performed, nuclease-based genome editing has some inherent limitations. The primary challenge is that the introduction of dsDNA breaks largely leads to the formation of stochastic insertions and deletions (indels).[2] Because most genetic diseases are due to a point mutation in a target locus, rather than stochastic disruption of a gene, methods to precisely correct point mutations are highly desirable.

BASE EDITING OF C•G to T•A

The Liu lab envisioned that directly converting one DNA base to another at a specific locus in the absence of dsDNA breaks would improve the efficiency of gene correction relative to HDR without the introduction of indels. The initial experimental design of base editor 1 (BE1) incorporated catalytically deactivated Cas9 (dCas9) from *Streptococcus pyogenes*, which has mutations in its two nuclease domains and does not cleave the DNA backbone. This protein is fused to the cytidine deaminase APOBEC1 by a 16-residue XTEN linker (Fig. 1).[2] Directed by a corresponding guide RNA, dCas9 binds a genomic locus of interest, forming a RNA-DNA ternary complex. This exposes a region of approximately five nucleotides for the tethered cytidine deaminase to convert C to U, which has the base-pairing properties of T. Although base editing efficiency for BE1 was high in vitro, the efficiency within cells was significantly diminished due to uracil DNA glycosylase (UDG) activity. UDG catalyzes the removal of U from DNA in cells and initiates the base excision repair (BER) pathway. In an effort to prevent BER at the target site, uracil DNA glycosylase inhibitor (UGI) was fused onto the C-terminus of BE1. To further improve editing efficiency, the catalytic His residue at position 840 in the Cas9 HNH domain was restored (Cas9 nickase). The catalytic His selectively nicks the unedited DNA strand, directing repair pathways to preferentially resolve U•G mismatches into the desired U•A product. The combination of these modifications is known as base editor 3 (BE3), and the potential of this technology was validated by correcting two disease-related mutations in mammalian cells.

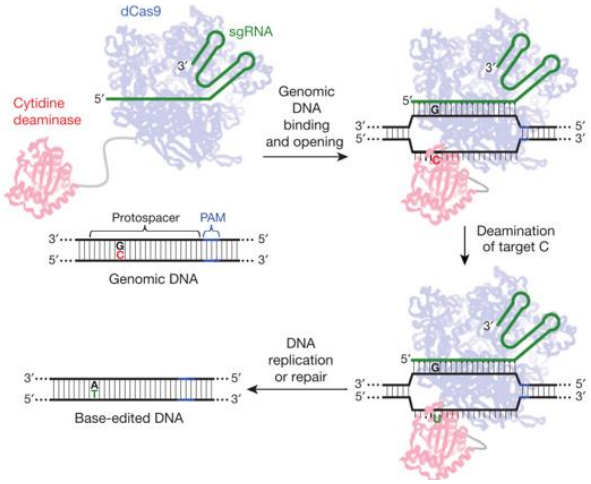

Figure 1. Schematic representation of the first-generation base-editing strategy.

A series of follow-up studies introduced several improvements to the BE3 complex. The nuclease editing window was narrowed by incorporating systematic mutations to alter the binding of the APOBEC enzyme to the DNA backbone, directing the enzyme to target 1-2 nucleotides instead of 5.[3] The number

of accessible base editing targets was broadened by using Cas9 proteins from various bacterial species to develop five additional base editors. These proteins are known to have different required binding sequences, known as protospacer adjacent motifs.[3] The incorporation of mutations known to decrease the affinity of Cas9 for DNA led to the development of a high-fidelity base editor that reduces off-target editing while maintaining on-target editing efficiency.[4] Finally, through the addition of a second UGI, base editor 4 (BE4) was generated with enhanced efficiency, product purity, and reduced indel formation.[5] These BE systems have since been widely applied in organisms such as bacteria, plants, zebrafish, mice, and human embryos.[6]

BASE EDITING OF A•T TO G•C

Until this point, all reported base editors mediated C•G to T•A conversion. The Liu group hypothesized that an analogous system could be designed in which an adenine deaminase fused to the Cas9 nickase protein could convert A to I, which is read as G by polymerases. However, no known enzymes can deaminate adenine within DNA, so they sought to evolve an adenine deaminase that would accept DNA as a substrate. TadA tRNA deaminase was selected as the target enzyme, and a bacterial selection method was developed in which defective antibiotic resistance genes were incorporated that contained point mutations at critical positions; reversion of A•T to G•C would restore antibiotic resistance (Fig. 2).[7] Seven rounds of selection were performed to identify an optimal adenine base editor (ABE). Ultimately, numerous selective and efficient late-stage ABEs were developed to convert A•T to G•C in vitro to correct pathogenic mutations and introduce disease-suppressing mutations. This development greatly expands the capabilities of base editing and the scope of pathogenic single-nucleotide polymorphisms that can be targeted precisely with genome editing.

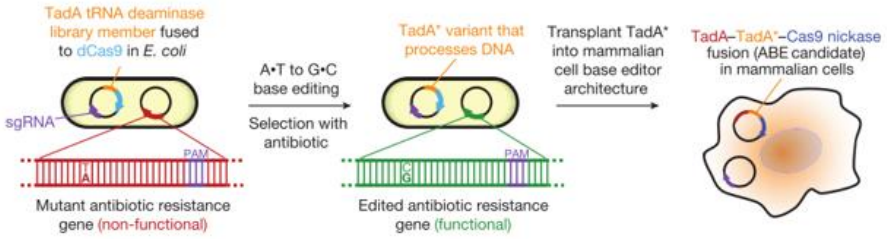

Figure 2. Schematic representation of the TadA ABE selection process.

OUTLOOK AND CHALLENGES

The requirement for dsDNA breaks with traditional genome editing methodologies had previously limited the precision and efficiency of correcting point mutations. Base editing now provides a powerful approach that circumvents these challenges, and since its introduction, significant progress has already been made towards improving and expanding the scope, specificity, efficiency, and applicability of this method. Future directions will most likely include additional ABE optimization and the development of genetic disease models. However, substantial work still needs to be done before base editing can reach its goal of clinical efficacy. Low in vivo editing efficiencies, ineffective delivery systems, and genetic mosaicism are important problems that will need to be addressed. Additionally, concerns have arisen over the potential immune system response due to the presence of Cas9 antibodies in human serum, as well as the challenge of FDA regulation of patient-specific therapies.

(1) Hess, G.T.; Tycko, J.; Yao, D.; Bassik, M.C. *Mol. Cell Rev.* **2017**, *68*, 26-43.
(2) Komor, A.C.; Kim, Y.B.; Packer, M.S.; Zuris, J.A.; Liu, D.R. *Nature* **2016**, *533*, 420-424.
(3) Kim, Y.B.; Komor, A.C.; Levy, J. M.; Packer, M.S.; Zhao, K.T.; Liu, D.R. *Nat. Biotechnol.* **2017**, 1-6.
(4) Rees, H.A.; Komor, A.C.; Yeh, W.H.; Caetano-Lopes, J.; Warman, M.; Edge, A.S.B.; Liu, D.R. *Nat. Commun.* **2017**, *8*, 1-10.
(5) Komor, A.C.; Zhao, K.T.; Packer, M.S.; Gaudelli, N.M.; Waterbury, A.L.; Koblan, L.W., Kim, Y.B.; Badran, A.H.; Liu, D.R. *Sci. Adv.* **2017**, *3*, 1-9.
(6) Kuscu, C.; Parlak, M.; Tufan, T.; Yang, J.; Szlachta, K.; Wei, X.; Mammadov, R.; Adli, M. *Nat. Methods* **2017**, *5*, 676-682.
(7) Gaudelli, N.M.; Komor, A.C.; Rees, H.A.; Packer, M.S.; Badran, A.H.; Bryson, D.I; Liu, D.R. *Nature* **2017**, *551*, 464-471.