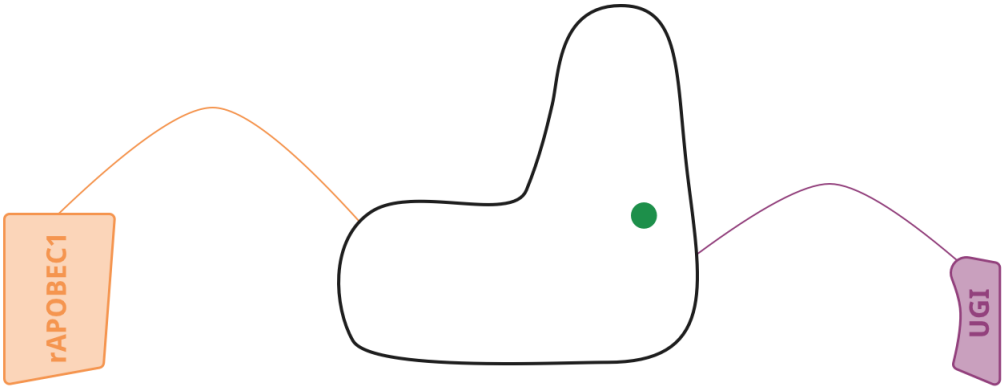
CRISPR Base Editing for Precise Genome Edits

Søren Hough
February 6, 2018



When discussing CRISPR with a general audience, it is often taken for granted that a simple combination of a guide RNA and a nuclease like Cas9 can edit any DNA locus to the precise needs of the investigator. However, we know that CRISPR is actually more like scissors, or a delete function in a word processor. Precise genomic edits to recapitulate point mutations and other clinically-relevant variants remain difficult to induce in a reliable, high-efficiency manner. Without this ability, forward screens and gain-of-function experiments are difficult to execute.

One of the ways researchers have approached this problem is by introducing homologous DNA templates with Cas9 and a guide RNA. This approach, known as CRISPR knockin, is meant to trigger homology-directed repair (HDR) or microhomology-directed repair (MMEJ) and has seen varying degrees of success depending on donor design, delivery method and other experimental factors. Desktop Genetics even developed an algorithm called CADMUS for consistent high-throughput guide and donor design for HDR-based knockin experiments.

Other methods have also emerged for improving the reliability and efficiency of CRISPR knockin experiments. Homology-independent targeted integration (HITI), which avoids the homologous recombination pathway, has been shown to induce high levels of gene insertion in zebrafish (Auer et al. 2014), human stem and somatic cells (He et al. 2016) and post-mitotic neurons *in vivo* in mice (Suzuki, Tsunekawa, Hernandez-Benitez, Wu et al. 2016). These studies suggest that HITI can dramatically outperform traditional HDR methods of CRISPR edits with 10 to 20-fold increases in efficiency (He et al. 2016).

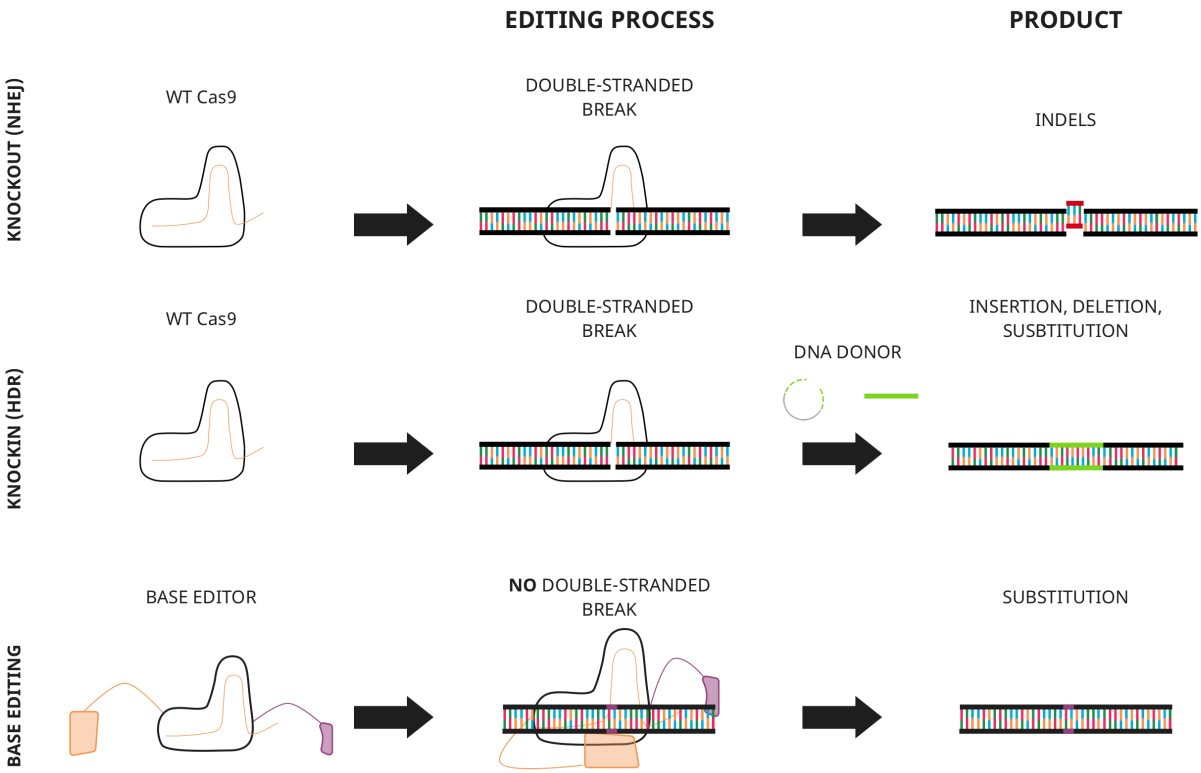


Figure 1. *Traditional CRISPR versus base editing. In general, permanent CRISPR edits via NHEJ, HDR, HITI (not pictured) and MMEJ (not pictured) involve a double-stranded break in target DNA. Base editing either nicks or does not cut the genome and precisely induces a base substitution.*

Despite these successes, the need for simplified and potentially orthogonal approaches to precise genome editing remains (Hess, Tycko, Yao et al. 2017). Enter base editing (BE), a new technique which avoids many of the hurdles of HDR and HITI. In base editing, like CRISPR interference and activation methods, the CRISPR system does not catalyze a double-stranded break in the target DNA (**Fig. 1**).

Instead, a catalytically inactive Cas9 (dCas9) or Cas9 D10A nickase (nCas9) is fused to an effector called cytidine deaminase (rAPOBEC1). This complex then chemically modifies a downstream nucleotide causing a C to T substitution, allowing for exact substitutions to introduce functionally relevant variants (Komor et al. 2016). We'll discuss some of the progress made with base editors since the concept was first introduced, as well as the modifications and future directions which might make base editing a more viable tool for basic and clinical research.

How Base Editing Works

Base editing offers an enticing alternative to standard CRISPR approaches. At the moment, it is difficult to predict the outcome of most genome editing experiments. Double-stranded breaks in DNA are often repaired to a nonrandom (but understudied) profile of insertion/deletion (indel) mutations (van Overbeek et al. 2016). Even in the presence of a

DNA donor, indels generated via nonhomologous end joining are the preferred repair product.

With base editing, a simpler and more predictable “base swap” occurs at reportedly high efficiency with little to no indel formation depending on the system used. The outcome of this process is a precise, designable point mutation. Giving investigators the ability to introduce site-specific mutations opens the door for new gain-of-function screens, functional genetics experiments and even opportunities for disease correction (Komor et al. 2016).

For base editing to work, a Cas9-rAPOBEC1 complex binds a target site guided by a guide RNA. Then, the cytidine deaminase converts a cytosine to a uracil. According to the authors, the system is most effective converting cytosines located at positions 4-8 in the protospacer region of the target (**Fig. 2**). Controlling which of these cytosines is converted remains a topic of research (Komor et al. 2016).

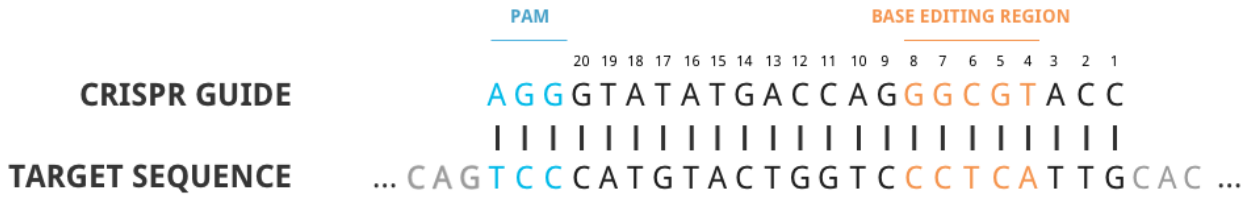


Figure 2. *Base editors modify certain cytosines. After dCas9 or nCas9 binds to a PAM/ protospacer, cytidine deaminase can modify cytosines at positions 4-8 in the protospacer. The sequence used here is an example.*

In the study reported by Komor and authors, they report three versions of their base editor. The first, BE1, is the simplest: dCas9 fused to rAPOBEC1. This complex converts a C to a U, although there is potential for the cell to revert the U back to a C to fix the G-U base pairing via mismatch repair (MMR). This explains why BE1 achieved high levels of DNA editing in a test tube (44%), but lower efficiency in eukaryotic cells (0.8-7.7%%) in the presence of endogenous repair machinery (Komor et al. 2016).

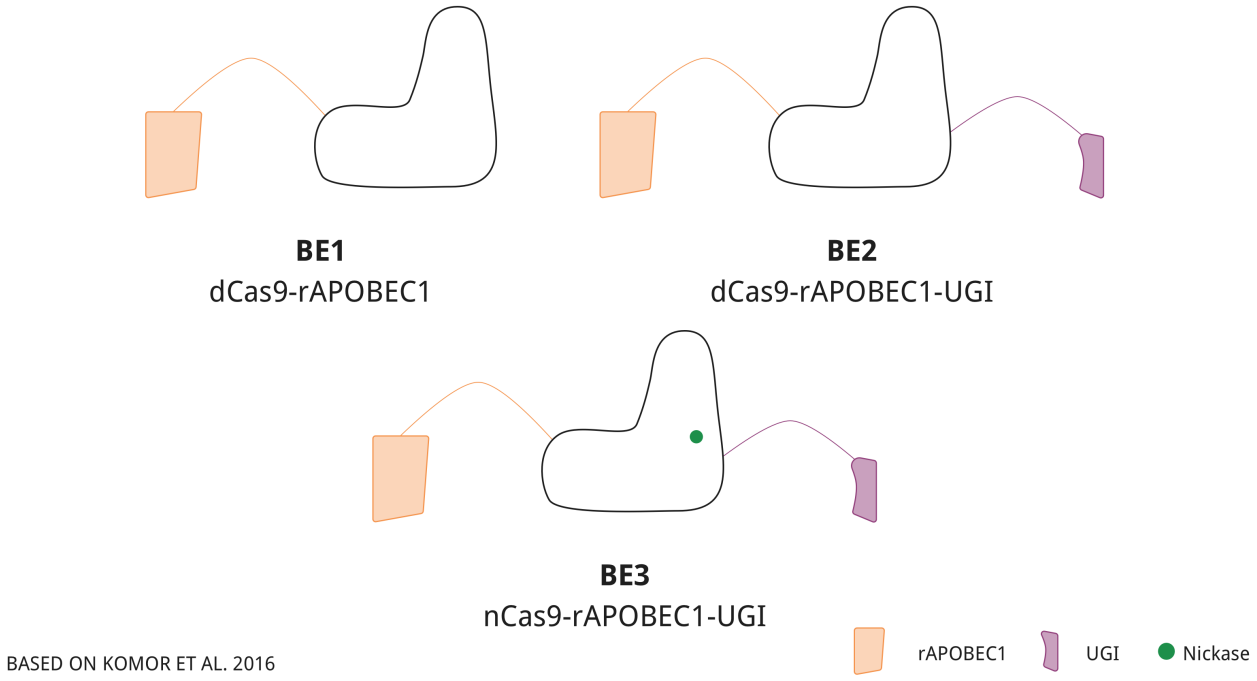


Figure 3. *Different base editing systems. The simpler BE1 system features a dCas9 fused to a cytidine deaminase. BE2 includes both cytidine deaminase and UGI to prevent converted C nucleotides from reverting. BE3 uses a nickase (nCas9) instead of dCas9 and can target both strands.*

The MMR phenomenon was accounted for with the second version of base editing, BE2, where a uracil DNA glycosylase inhibitor (UGI) was also included in the dCas9 complex to block the reversion. This led to an improved editing efficiency of 20% in cells. A third version of base editors, termed BE3, used nCas9 fused to rAPOBEC1 and UGI. This system can induce indels (unlike the dCas9 versions, which generated indels at a rate of . 1% or less), but is liable to generate higher editing efficiencies (37%) to a desired T-A final product (**Fig 3.**) (Komor et al. 2016).

Improvements of Base Editing System

Base editing systems, while still new to the CRISPR world, have been iterated on in similar ways to traditional Cas9 systems, including expanding both their range and specificity (**Table 1**). For example, the authors of the first base editing paper reported an updated version of their BE3 system with more specific editing behavior. To achieve this, they mutated nCas9 following an older report by Kleinstiver et al. (Rees et al. 2017).

The group termed their new system HF-BE3 (high-fidelity BE3) and saw a three-fold reduction in off-target editing at specific sites in VEGFA *in vitro* while generally retaining on-target editing parity. The group repeated their success when delivering HF-BE3 as RNP to zebrafish embryos and mouse inner ears. The authors noted that on-target editing efficiency was reduced *in vivo* in mice (Rees et al. 2017).

The same group also developed significant modifications to their base editors in order to expand their targeting range and precision. To do this, Kim and coauthors looked first at

expanding BE PAM targeting capabilities. To do this, thee team used both SpCas9 mutants with altered PAM recognition as well as the SpCas9 ortholog SaCas9 from *Staphylococcus aureus* (Kim et al. 2017). In addition to changing PAM specificity, the SaCas9 base editor is also attractive because it is much smaller (3.16 kb) than SpCas9 (4.1 kb) and therefore potentially easier to deliver in a viral vector for *in vivo* therapeutic development.

Base Editor Variants and Effects

NAME	DESCRIPTION	EFFECT	PAPER
BE1	dCas9-rAPOBEC1	No cut, GC to AT	Komor et al. 2016
BE2	dCas9-rAPOBEC1-UGI	No cut, prevent MMR, convert GC to AT	Komor et al. 2016
BE3	nCas9-rAPOBEC1-UGI	Nick, high efficiency edit, prevent MMR, convert GC to AT	Komor et al. 2016
HF-BE3	High Fidelity BE3	Increased specificity	Rees et al. 2017
SaBE3	Staphylococcus aureus BE3	Smaller system, altered PAM recognition	Kim et al. 2017
BE3 Variants	SaKKH-BE3, VQR-BE3, EQR-BE3, and VRER-BE3	Altered PAM recognition	Kim et al. 2017
rAPOBEC1	Modified rAPOBEC1	Narrow editing window from 5 to 1-2 nucleotides	Kim et al. 2017
eBE	Enhanced base editor (BE with free UGI)	Reduced indel rates	Wang, Xue, Yan et al. 2017
ABE	Adenine base editors	Convert AT to GC	Gaudelli et al. 2017
HF2-BE2	High-Fidelity BE2	Modify both strands, expand targeting range	Liang, Sun et al. 2017

Table 1. *Base editor variants. Standard BE1, BE2, BE3 systems have been combined with other work in engineered CRISPR nucleases to expand targeting capability, improve specificity and more.*

Kim and coworkers then looked at mutating the other central component of base editors: cytidine deaminase. The authors report that they were able to shrink the editing window of rAPOBEC1 from a potential 5-nucleotide window to 1-2 nucleotides, allowing more precise genomic edits. When Kim et al. looked at targetable mutations in the ClinVar database, they found that decreasing the window size of base editing meant a 2.5-fold increase disease variants that can be corrected without accidentally editing nearby cytosines. The authors subsequently combined the more specific cytidine deaminase with one of the Cas9 variants (Cas9-VQR) used to expand BE PAM targeting range and observed a compound effect (Kim et al. 2017).

A paper by Wang, Xue, Yan and coauthors used some of the discoveries by Komor et al. to improve on the higher-efficiency nCas9 BE3 system. The authors pointed out that although nCas9-rAPOBEC1-UGI leads to high editing efficiency rates, they also lead to unwanted (albeit minimal) indel mutations (4-12%) in some cases (Komor et al. 2016, Nishida et al. 2016). The authors hypothesized that including additional free UGI when delivering BE3 (a system termed “enhanced base editor,” or eBE) might reduce indel formation. They saw that increasing the amount of free UGI did ultimately suppress indel

rates and, as an added benefit, enhanced on-target editing rates over BE3 (Wang, Xue, Yan et al. 2017).

Another group significantly expanded the range of base editors by evolving a new protein to target previously uneditable nucleotides. Gaudelli and coauthors mutated a transfer RNA adenosine deaminase into a protein which can successfully deaminate adenine, a previously unachievable feat. Fusing this protein to dCas9, a system referred to as adenine base editors (ABEs), provided the novel function of changing A-T base pairs to G-C. The authors suggest that this opens up the potential to study another 50% of known pathogenic mutations. According to the paper, ABEs are also more efficient with minimal off-target effects compared to traditional base editing systems (Gaudelli et al. 2017).

Lab-Based Applications of Base Editing

Base editing offers another method for exploring the function of genes via gene knockout. Two papers report novel methods of introducing nonsense mutations in protein coding DNA in order to precisely ablate gene expression. Base editing for gene knockout is also permanent, unlike RNAi- or CRISPRi-based gene knockdown, making it comparable to standard CRISPR approaches. This may present an advantage over traditional CRISPR methods where the nonrandom indel profile remains difficult to predict and where the nuclease is more likely to induce off-target effects.

The other problem with generating indel mutations is that they can cause downstream frameshifts, potentially impacting the expression of other genomic elements. With base editing, the investigator can control for this by designing codon substitutions that maintain the open reading frame. The authors of both papers suggest potential for bringing their base editing methods to the scale of a genome-wide screen as an alternative to normal knockout libraries.

One of these methods, presented by Kuscu and colleagues, is called CRISPR-STOP. CRISPR-STOP is a BE3 nickase approach which targets one of four specific codons (CAG, CGA, CAA, TGG) and converts them to STOP codons (**Fig. 4**). The authors first determined all potential guide RNAs which could be used to target the codons of interest. Kuscu et al. then created and used small libraries to demonstrate the utility of this method for genetic screening. The team successfully evaluated their approach with a phenotypic readout of fluorescence (FACS sorting) and drug resistance (Kuscu et al. 2017).

Another paper by Billon, Bryant et al. describes a system they term iSTOP. In this approach, a BE3-based system is once again directed to convert one of four codons (CAA, CAG, CGA, TGG) into STOP codons (**Fig. 4**). The authors narrowed their potential targets by selecting codons with a cytosine in a restriction site; this allowed for a simple digestion-based tests for successful editing. The study also presented a database of guide RNAs which can be used to target 97-99% of genes in eight different eukaryotes (Billon, Bryant et al. 2017).

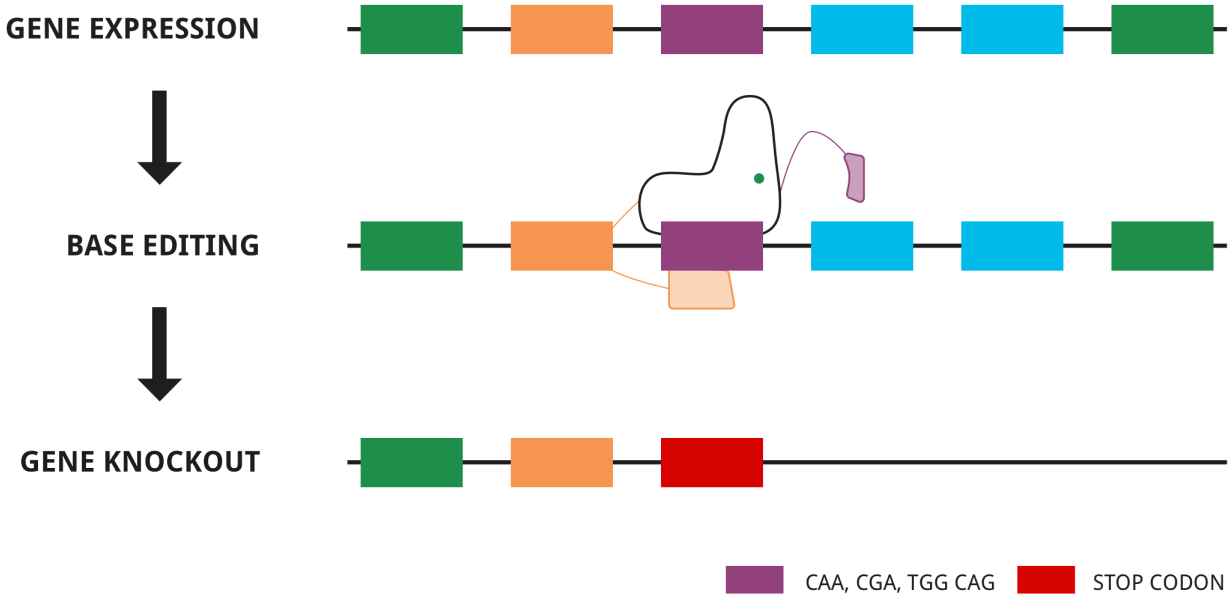


Figure 4. *Base editing to introduce stop codons. In both CRISPR-STOP and iSTOP methods, BE3 transforms one of four codons to a STOP codon, thereby functionally knocking out the target gene.*

Another intriguing variation of the CRISPR base editing strategy is CRISPR-X. Hess and coauthors discussed how the hyperactive cytidine deaminase AID (activation induced cytidine deaminase) can also be paired with dCas9. This system recruits AID via an MS2 hairpin binding site. dCas9-AID induces C>T, G>A and other mutations in a 100 bp region proximal to the binding site. These mutations can then be screened to discover novel functions or recapitulate disease phenotypes (Hess et al. 2016).

CRISPR-X and similar methods (Ma, Zhang et al. 2016) mirror the goals of "*in situ* saturation mutagenesis." In that case, a library of donor templates is used to introduce HDR-based point mutations in the target region to better understand the association between sequence and protein function. Hess at al. draw a distinction between that method and CRISPR-X by reiterating that this is a more precise approach which avoids the unpredictability of indel generation following double-stranded breaks (Hess et al. 2016). Significantly, CRISPR-X and similar AID-based methods differ from iSTOP and CRISPR-STOP in that they do not focus solely on protein-coding DNA; these protocols can provide novel insight into the noncoding genome, as well.

Bringing Base Editing *In Vivo*

In addition to the paper by Rees et al., researchers have demonstrated the utility of base editing in a variety of other *in vivo* contexts. For example, Zhang, Qin and coauthors introduced a zebrafish codon-optimized nCas9 system based on BE3 to determine whether base editing could be successful *in vivo* in their model organism. The team reported that this approach was functional in zebrafish embryos (Zhang, Qin et al. 2017).

After seeing success *in vivo*, the authors fused an nCas9 variant called nCas9-VQR to cytidine deaminase to expand targeting options (VQR recognizes NGA instead of NGG). Zhang, Qin and the team achieved up to 30% editing rates — but also unwanted indels at a rate of 35%. Using a dCas9-VQR base editing alternative, on-target editing efficiency dropped to 15% while indel rates lowered to 5% (Zhang, Qin et al. 2017).

Kim, Ryu and coworkers demonstrated that BE3 can be used to efficiently create nonsense mutations in the *Tyr* and *Dmd* genes in mouse embryos. They introduced BE3 into mice using different delivery methods (mRNA, RNP) and achieved successful mutations at a rate of ~50% in edited embryos. The group also deep sequenced algorithmically predicted off-target sites and found no evidence of nonspecific editing. The authors verified this finding with whole-genome sequencing (WGS). WGS turned up one putative off-target edit, but deep sequencing revealed this to be a false positive result (Kim, Ryu et al. 2017).

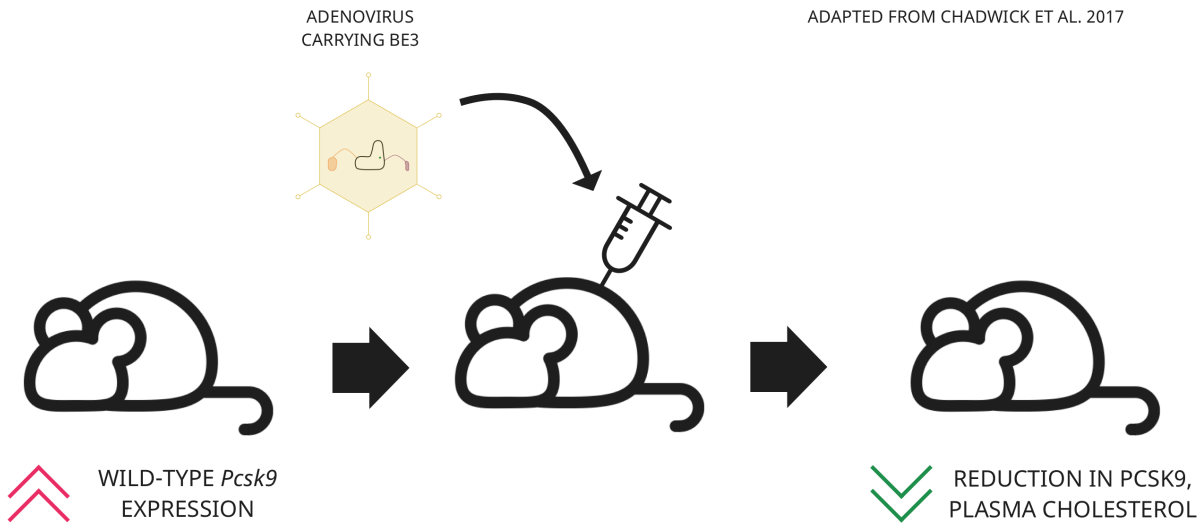


Figure 5. *Base editing in adult mice. Chadwick and coauthors use adenovirus containing BE3 to reduce Pcsk9 levels and, consequently, cholesterol in adult mice.*

Base editing in mouse zygotes was achieved by another group using “high fidelity” BE2 (HF2-BE2). Liang, Sun et al. developed HF2-BE2, a new system that can modify both strands of DNA as the nickase-based BE3 system does. The authors reported that HF2-BE2 edited at a near 100% biallelic efficiency. Perhaps more interestingly, HF2-BE2 was also able to deaminate cytosines either upstream or downstream of the protospacer region (Liang, Sun et al. 2017). This simultaneously expands the range of targetable C’s in the genome and raises questions around off-target effects.

Looking beyond germline edits, Chadwick et al. sought to demonstrate base editing in adult mice (**Fig. 5**). To do this, they delivered BE3 and a guide RNA targeting *Pcsk9* to the liver in a single adenoviral vector. Pcsk9 is a gene for which nonsense mutations have been associated with lower cholesterol levels. This makes it an attractive target for preventative gene therapy against heart disease. After monitoring levels in the mice for

four weeks, the authors reported an over 50% decrease in Pcsk9 expression and a 28% decrease in plasma cholesterol (Chadwick et al. 2017).

As with other forms of CRISPR, base editing also has the opportunity to transform the agricultural industry. Both tomato and rice have each been edited with a system the authors termed Target-AID. Target-AID was first developed by Ma, Zhang and coauthors in 2016 for mammalian cells. The version by Shimatani, Kashojiya, Takayama et al. is composed of dCas9 fused to a cytidine deaminase from *Petromyzon marinus* (a sea lamprey) called PmCDA1. To demonstrate the utility of Target-AID in crop modification, the authors used base editing to alter pesticide resistance in rice and introduce heritable mutations in tomatoes (Shimatani, Kashojiya, Takayama 2017).

Base Editing: Another Arrow In an Ever-Expanding Quiver

Base editing is positioned as a fascinating alternative to standard CRISPR systems. BE systems will continue to be iterated on as we've seen with standard SpCas9 with a consistent drive toward higher activity, specificity and targeting range. Base editing is even joining wild-type Cas13 and SaCas9 in the race to edit RNA with CRISPR nucleases, as well.

One group used a deactivated Cas13 (dCas13) fused to adenosine-to-inosine deaminase (ADAR2) which can modify adenine in messenger RNA to inosine (translationally equivalent to a G). The authors suggest that this approach (termed REPAIR) can be used to remove pathogenic mutations in RNA transcripts This poses a valuable alternative to modifying the genome, which is permanent, and could serve as a transient therapy with more nuance than CRISPRi, RNAi or Cas13-based approaches (Cox, Gootenberg, Abudayyeh et al. 2017).

Wherever base editing goes next, the potential of these novel mechanisms is exciting. Fusing activators, repressors and complex epigenetic modifiers have already borne fruit both in vitro and in vivo. Base editing comes from this lineage of dCas9-based CRISPR modification. Whatever other novel genome editing techniques may emerge in the future, it's clear that we are only just beginning to see this wave of CRISPR technology take shape.

Sources

Auer TO, Duroure K, De Cian A et al. Highly efficient CRISPR/Cas9-mediated knock-in in zebrafish by homology-independent DNA repair. Genome Res. 2014 Jan;24(1):142-53. doi: 10.1101/gr.161638.113. Epub 2013 Oct 31. PubMed PMID: 24179142.

Chadwick AC, Wang X, Musunuru K. In Vivo Base Editing of PCSK9 (Proprotein Convertase Subtilisin/Kexin Type 9) as a Therapeutic Alternative to Genome Editing. Arterioscler Thromb Vasc Biol. 2017 Sep;37(9): 1741-1747. doi: 10.1161/ATVBAHA.117.309881. Epub 2017 Jul 27. PubMed PMID: 28751571.

Cox DBT, Gootenberg JS, Abudayyeh OO et al. RNA editing with CRISPR-Cas13. Science. 2017 Nov 24;358(6366):1019-1027. doi: 10.1126/science.aaq0180. Epub 2017 Oct 25. PubMed PMID: 29070703.

Gaudelli NM, Komor AC, Rees HA et al. Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. Nature. 2017 Nov 23;551(7681):464-471. doi: 10.1038/nature24644. Epub 2017 Oct 25. PubMed PMID: 29160308.

He X, Tan C, Wang F et al. Knock-in of large reporter genes in human cells via CRISPR/Cas9-induced homology-dependent and independent DNA repair. Nucleic Acids Res. 2016 May 19;44(9):e85. doi: 10.1093/nar/gkw064. Epub 2016 Feb 4. PubMed PMID: 26850641.

Hess GT, Frésard L, Han K et al. Directed evolution using dCas9-targeted somatic hypermutation in mammalian cells. Nat Methods. 2016 Dec;13(12):1036-1042. doi: 10.1038/nmeth.4038. Epub 2016 Oct 31. PubMed PMID: 27798611.

Hess GT, Tycko J, Yao D et al. Methods and Applications of CRISPR-Mediated Base Editing in Eukaryotic Genomes. Mol Cell. 2017 Oct 5;68(1):26-43. doi: 10.1016/j.molcel.2017.09.029. Review. PubMed PMID: 28985508.

Kim K, Ryu SM, Kim ST et al. Highly efficient RNA-guided base editing in mouse embryos. Nat Biotechnol. 2017 May;35(5):435-437. doi: 10.1038/nbt.3816. Epub 2017 Feb 27. PubMed PMID: 28244995.

Kim YB, Komor AC, Levy JM et al. Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions. Nat Biotechnol. 2017 Apr;35(4):371-376. doi: 10.1038/nbt.3803. Epub 2017 Feb 13. PubMed PMID: 28191901.

Komor AC, Kim YB, Packer MS et al. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature. 2016 May 19;533(7603):420-4. doi: 10.1038/nature17946. Epub 2016 Apr 20. PubMed PMID: 27096365.

Kuscu C, Parlak M, Tufan T et al. CRISPR-STOP: gene silencing through base-editing-induced nonsense mutations. Nat Methods. 2017 Jul;14(7):710-712. doi: 10.1038/nmeth.4327. Epub 2017 Jun 5. PubMed PMID: 28581493.

Liang P, Sun H, Sun Y et al. Effective gene editing by high-fidelity base editor 2 in mouse zygotes. Protein Cell. 2017 Aug;8(8):601-611. doi: 10.1007/s13238-017-0418-2. Epub 2017 Jun 5. PubMed PMID: 28585179.

Ma Y, Zhang J, Yin W et al. Targeted AID-mediated mutagenesis (TAM) enables efficient genomic diversification in mammalian cells. Nat Methods. 2016 Dec;13(12):1029-1035. doi: 10.1038/nmeth.4027. Epub 2016 Oct 10. PubMed PMID: 27723754.

Rees HA, Komor AC, Yeh WH et al. Improving the DNA specificity and applicability of base editing through protein engineering and protein delivery. Nat Commun. 2017 Jun 6;8:15790. doi: 10.1038/ncomms15790. PubMed PMID: 28585549.

Shimatani Z, Kashojiya S, Takayama M et al. Targeted base editing in rice and tomato using a CRISPR-Cas9 cytidine deaminase fusion. Nat Biotechnol. 2017 May;35(5):441-443. doi: 10.1038/nbt.3833. Epub 2017 Mar 27. PubMed PMID: 28346401.

Suzuki K, Tsunekawa Y, Hernandez-Benitez R et al. In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration. Nature. 2016 Dec 1;540(7631):144-149. doi: 10.1038/nature20565. Epub 2016 Nov 16. PubMed PMID: 27851729.

van Overbeek M, Capurso D, Carter MM et al. DNA Repair Profiling Reveals Nonrandom Outcomes at Cas9-Mediated Breaks. Mol Cell. 2016 Aug 18;63(4):633-646. doi: 10.1016/j.molcel.2016.06.037. Epub 2016 Aug 4. PubMed PMID: 27499295.

Wang L, Xue W, Yan L et al. Enhanced base editing by co-expression of free uracil DNA glycosylase inhibitor. Cell Res. 2017 Oct;27(10):1289-1292. doi: 10.1038/cr.2017.111. Epub 2017 Aug 29. PubMed PMID: 28849781.

Zhang Y, Qin W, Lu X et al. Programmable base editing of zebrafish genome using a modified CRISPR-Cas9 system. Nat Commun. 2017 Jul 25;8(1):118. doi: 10.1038/s41467-017-00175-6. PubMed PMID: 28740134.