

Screening of CRISPR/Cas base editors to target the AMD high-risk Y402H complement factor H variant

Minh Thuan Nguyen Tran,¹ Mohd Khairul Nizam Mohd Khalid,¹ Alice Pébay,^{2,3,6} Anthony L. Cook,⁴ Helena H. Liang,² Raymond C.B. Wong,^{2,3} Jamie E. Craig,⁵ Guei-Sheung Liu,^{1,3} Sandy S. Hung,^{2,3} Alex W. Hewitt^{1,2,3}

(The last two authors jointly supervised and contributed equally to this work.)

¹Menzies Institute for Medical Research, University of Tasmania, Tasmania, Australia; ²Centre for Eye Research Australia, Royal Victorian Eye and Ear Hospital, Victoria, Australia; ³Ophthalmology, Department of Surgery, University of Melbourne, Victoria, Australia; ⁴Wicking Dementia Research and Education Centre, University of Tasmania, Hobart, TAS 7000, Australia; ⁵Department of Ophthalmology, Flinders University, Flinders Medical Centre, Bedford Park, Australia; ⁶Department of Anatomy and Neuroscience, University of Melbourne, Victoria, Australia

Purpose: To evaluate the efficacy of using a CRISPR/Cas-mediated strategy to correct a common high-risk allele that is associated with age-related macular degeneration (AMD; rs1061170; NM_000186.3:c.1204T>C; NP_000177.2:p.His402Tyr) in the complement factor H (*CFH*) gene.

Methods: A human embryonic kidney cell line (HEK293A) was engineered to contain the pathogenic risk variant for AMD (HEK293A-CFH). Several different base editor constructs (BE3, SaBE3, SaKKH-BE3, VQR-BE3, and Target-AID) and their respective single-guide RNA (sgRNA) expression cassettes targeting either the pathogenic risk variant allele in the *CFH* locus or the *LacZ* gene, as a negative control, were evaluated head-to-head for the incidence of a cytosine-to-thymine nucleotide correction. The base editor construct that showed appreciable editing activity was selected for further assessment in which the base-edited region was subjected to next-generation deep sequencing to quantify on-target and off-target editing efficacy.

Results: The tandem use of the Target-AID base editor and its respective sgRNA demonstrated a base editing efficiency of facilitating a cytosine-to-thymine nucleotide correction in 21.5% of the total sequencing reads. Additionally, the incidence of insertions and deletions (indels) was detected in only 0.15% of the sequencing reads with virtually no off-target effects evident across the top 11 predicted off-target sites containing at least one cytosine in the activity window (n = 3, pooled amplicons).

Conclusions: CRISPR-mediated base editing can be used to facilitate a permanent and stably inherited cytosine-to-thymine nucleotide correction of the rs1061170 SNP in the *CFH* gene with minimal off-target effects.

Age-related macular degeneration (AMD) is a leading cause of late-onset central vision loss affecting individuals over the age of 50 years [1]. The condition has a substantial global burden and is expected to affect up to 288 million people by the year 2040 [2]. It involves environmental, genetic, and physiologic determinants [3] that cause damage to the macula, which is an area within the eye required for sharp, central vision. Here, the retinal pigment epithelium (RPE) is important for maintaining retinal homeostasis by providing nutrients and ionic support to the apical photoreceptors while facilitating the phagocytic removal of waste products and the exchange of biomolecules with the fenestrated choroidal capillaries [4]. The basal deposition of drusen, which can occur at either the RPE or Bruch's membrane, can significantly affect photoreceptor health [5]. This has been

previously thought to occur due to the dysregulation of the alternative complement pathway, which causes inflammation of the RPE. The dysregulation of the alternative complement pathway has been shown to be associated with the extensively characterized single nucleotide polymorphism (SNP), rs1061170, in the Complement Factor H (*CFH*) gene [6].

The rs1061170 variant is characterized by a cytosine nucleotide (C) in place of a thymine nucleotide (T), which results in a tyrosine to histidine amino acid residue substitution at position 402 (Y402H) in the CFH protein. The incidence of the disease associated variant at this locus results in a 3.3- to 7.7-fold increase in AMD progression for Caucasians homozygous for the risk variant allele, and a 2.2- to 4.6-fold increase for heterozygous individuals [6]. The amino acid residue substitution principally affects the CFH protein, which is a serum glycoprotein responsible for attenuating the response of the alternative complement pathway by cleaving pro-inflammatory mediators [7, 8]. The Y402H allelic variant affects the hydrogen bonding capacity

Correspondence to: Alex Hewitt, Menzies Institute for Medical Research, University of Tasmania, Hobart 7000, Australia; Phone: +61 407359824; FAX: +61 3 6226 7704; email: hewitt.alex@gmail.com

of the CFH protein and impairs its ability to dock to the appropriate receptor on the RPE or Bruch's membrane [9, 10]. This stereochemical impairment results in an increase in the fraction of uncleaved pro-inflammatory cytokines and causes inflammation of the RPE.

The discovery of an adaptive prokaryotic immune system that requires the use of only a single endonuclease (Cas9) to mediate a DNA double-strand break (DSB) [11], allowed for the development of a programmable nuclease that could target DNA in a site-specific manner [12]. Here, the single-guide RNA (sgRNA) component of the CRISPR/Cas9 system facilitates targeted DNA editing at a user-defined nucleotide sequence [12, 13]. The binding between the Cas9 protein and the target DNA strand is dictated by its protospacer-adjacent motif (PAM) site, which is a short consensus sequence of nucleotides that occur adjacent to the target DNA sequence [12, 14].

Herein, we describe the application of a modified Clustered Regularly Interspersed Short Palindromic Repeats/CRISPR associated protein 9 (CRISPR/Cas9) system, known as 'base editing' [15], to mediate a precise single nucleotide change targeting the AMD-associated risk allele at the rs1061170 locus. Base editing mitigates the previously characterized risks associated with introducing DSBs into the genome, which can lead to imprecise insertions and deletions (indels) [16, 17]. Further, base editing has a dramatically higher editing efficiency compared to homology-directed repair (HDR), a process that has a notoriously low efficiency in somatic cells due to the competing DNA repair pathway of non-homologous end joining (NHEJ), where nucleotide insertions and deletions are typically generated [18, 19].

The CRISPR-mediated base editor fuses a cytidine deaminase, such as APOBEC3 [20], APOBEC1 [15], or human-AID (activation-induced cytidine deaminase) [21, 22], to a modified Cas9 protein to allow for site-directed cytosine-to-thymine nucleotide corrections [23]. Its nickase mechanism and the fusion of a uracil DNA glycosylase inhibitor (UGI) allows the introduction of a nucleotide correction in a stably inherited manner in both dividing and non-dividing cells [12, 24]. Although there are several variants of base editors currently available [20, 25], we selected and screened for the most well characterized variants of base editors. CRISPR/Cas base editors targeting the disease-relevant locus were screened due to their varying PAM requirements, sgRNA target sequence, and anticipated activity window. As such, base editing variants with the same mutagenicity profiles were selected to observe if the sequence-specific context of the locus has any bearing on the efficacy of the targeted correction (BE3, SaKKH-BE3, SaBE3, and VQR-BE3) [25].

Here, the PAM variant base editors (SaKKH-BE3, SaBE3, and VQR-BE3) share the same structural arrangement of BE3 in terms of the choice of the deaminating protein (APOBEC3), linker length, and terminal protein fusions, but markedly differ in the range of PAM targetable DNA substrates given their variations in the Cas9 protein component. Further, to profile the effects of the activity window on correcting the rs1061170 SNP, a base editor with a different architecture and activity window (Target-AID) was evaluated [26]. While the respective PmCDA1 and APOBEC3 proteins of Target-AID and BE3 both fulfill a deaminating role, each protein possesses a differing mutagenicity profile and processive capability [27-30]. These differences in protein function require much-needed consideration and clarification for use in the mostly cytosine-deplete sequence space surrounding the rs1061170 SNP. After the initial screening of base editors at the Y402H *CFH* locus, we selected the most active base editor for further off-target assays to explore its therapeutic potential.

METHODS

Generation of a HEK293A-CFH(p.His402) fragment cell line: To evaluate the efficacy of base editing to correct the AMD-associated rs1061170 variant, a commercial human embryonic kidney cell line (HEK293A; Invitrogen, CA) was engineered to contain the pathogenic risk variant by using a lentiviral method to knock-in a gene fragment containing exon 9 of the *CFH* gene. Briefly, the pLenti.AS2.Luci.puro vector (RNAiCore; Academia Sinica, Taipei, Taiwan) was digested using *NheI* and *EcoRI* (New England Biolabs, MA) to insert the synthetic *CFH* gene fragment (gBlock gene fragment, Integrated DNA Technologies, IA). Cell line authentication was performed using short tandem repeat DNA profiling with the following markers: AMEL, CSF1PO, D5S818, D7S820, D13S317, D16S539, D21S11, TH01, TPOX, and vWA through the Australian Genome Research Facility LTD, VIC, Australia (Appendix 1, Appendix 2, and Appendix 3). Cell line transduction and ongoing cell culture conditions were followed as previously described [31].

Modification of SpCas9 sgRNA scaffold: CHOPCHOP was used to evaluate the incidence of internal RNA interactions between the protospacer portion of the sgRNA and the sgRNA scaffold backbone [32]. The webtool RNAfold [33] was used to investigate the folding spontaneity of the unmodified and modified SpCas9 sgRNA scaffolds. The following scaffold modifications were considered: a canonical A54:U60 substitution in the critical stem loop one region and a modified stem loop two and three linker region, and an A:U nucleobase pair flip at U25 in the repeat:anti-repeat duplex of

the sgRNA scaffold [34] and an extension of the repeat:anti-repeat duplex (F-E modified scaffold; Appendix 2). Each modified scaffold contained a 5' spacer portion specific for rs1061170 and a GGG PAM site, and a 3' scaffold backbone (sgRNA2 and sgRNA2G, Appendix 2). The modified scaffolds were prepared according to the gBlocks® Gene Fragments (Integrated DNA Technologies, IA) protocols as described above.

CRISPR-base editing constructs and sgRNA cassette generation: The Target-AID construct was assembled using the HiFi DNA Assembly method (New England Biolabs, MA) where the Target-AID was subcloned from the pcDNA3.1_pCMV-nCas-PmCDA1-ugi pH1-gRNA (HPRT) plasmid (5,365 base pairs; Addgene plasmid #79620); the P2A-fragment and Blastidicin resistance gene were amplified from the LentiCas9-BLAST plasmid (Addgene Plasmid #52962); the vector backbone was amplified from pCMV-BE3 (3,361 base pairs; Addgene plasmid #73021). The individual PCR fragments of the construct were amplified using the Q5 High-Fidelity DNA polymerase (New England Biolabs) and purified using the QIAquick® PCR Purification Kit (QIAGEN, Hilden, Germany) according to the manufacturer's instructions. NEB® Stable Competent *E. coli* (New England Biolabs) was used for transformation following standard transformation protocol. Plasmid DNA was harvested using the Wizard® Plus SV Minipreps DNA Purification Systems (Promega Corporation, WI) according to the manufacturer's instructions.

For the U6-promoter driven expression of sgRNAs for the SaCas9 PAM variants, SaBE3 (Addgene Plasmid #85169) and SaKKH-BE3 (Addgene Plasmid #85170), *CFH*-locus specific sgRNAs (Appendix 2) were annealed and cloned into the digested px552-CMV-U6-SaCas9 (Addgene Plasmid #107053) construct at the SapI site. Whereas the cloning of SpCas9-specific sgRNA for the SpCas9 PAM variant, VQR-BE3 (Addgene Plasmid #85171) and sgRNA oligos (Appendix 2) were annealed and cloned into the px552-CMV-mCherry-U6-SpCas9 sgRNA scaffold (Addgene Plasmid #60958). Constructs were transformed into NEB® Stable Competent *E. coli* (New England Biolabs) and plasmid DNA was harvested using the Wizard® Plus SV Minipreps DNA Purification Systems (Promega).

HEK293A-CFH(p.His402) fragment cell line transfection with CRISPR/sgRNA constructs and cassettes: HEK293A-CFH(p.His402) fragment cells were plated to a final density of 75,000 cells per well in a 12-well plate. After one day of incubation at 37 °C and 5% CO₂, the confluency of the cells was checked to ensure it was between 40% and 70%. Then, 500 ul of 10% fetal bovine serum (FBS) media (Thermo Fisher

Scientific™, MA) was replaced and the DNA:nanoparticle complex was prepared following the manufacturer's protocol. For each condition, an aliquot containing 450 ng of sgRNA scaffold (gBlocks® Gene Fragments) or 1 µg of px552-CMV-mCherry-U6-SaCas9 or SpCas9 sgRNA scaffold (Appendix 2), and 1 µg of the CRISPR/Cas9 plasmid were prepared in 46 µl of Opti-MEM (Thermo Fisher Scientific™). Then, 3.5 µl of FuGENE® HD (Promega) was added, followed by pulse spin and vortex. The reaction aliquot was then incubated for 10 min at room temperature before adding it in a dropwise manner to each well of the 12-well plate. The HEK293A-CFH(p.His402) fragment cells were incubated at 37 °C for three days.

Following transfection of the HEK293A-CFH(p.His402) fragment cells with the CRISPR and sgRNA constructs, 20 µg/ml of blasticidin was added one day after transfection for a six-day (blasticidin) selection of the cells. Genomic DNA was harvested using the QIAamp DNA minikit (QIAGEN) from enriched cells, six days after transfection. Amplification of the exogenous *CFH* fragment tag was performed using KOD DNA polymerase (Merck Millipore, Darmstadt, Germany) with the Myc and FLAG tag primers (Appendix 2) and PCR products were purified using the Wizard® SV Gel and PCR Clean-Up System (Promega, WI). PCR samples were sent for Sanger sequencing with the Myc tag primer.

High-throughput next generation deep sequencing of base-edited gDNA samples: Nested PCR amplification reactions were performed on the extracted gel-purified PCR amplicons for the on-target base edited cells following transfection using custom primers annealing to the *CFH* gene fragment and a compatible overhang for Illumina forward and reverse adapters (Appendix 2). For the off-target assays, the top 20 scoring sgRNAs were selected using CHOPCHOP [32]. The sgRNAs were then further evaluated using the following criteria: a cytosine occurs within a -1 to +5 nucleotide position within the 20 nucleotide protospacer, whereby the PAM consensus sequence is in position +21 to +23; a compatible NGG-style PAM site must flank the 20-nucleotide protospacer; the sgRNA protospacer for the off-target site must contain no more than three mismatches compared to the on-target sgRNA sequence. Briefly, using standard conditions for Q5 High-Fidelity DNA polymerase (New England Biolabs), the on-target PCR amplicons were amplified using the following thermocycling conditions: initial denaturation at 98 °C for 30 s, cycling at 98 °C for 10 s, 60 °C for 10 s, 72 °C for 15 s, and a final extension of 72 °C for 2 min. Off-target sites were PCR-amplified using the extracted gDNA sample directly after transfection on both the *LacZ*-treated negative control and on the base editor-treated

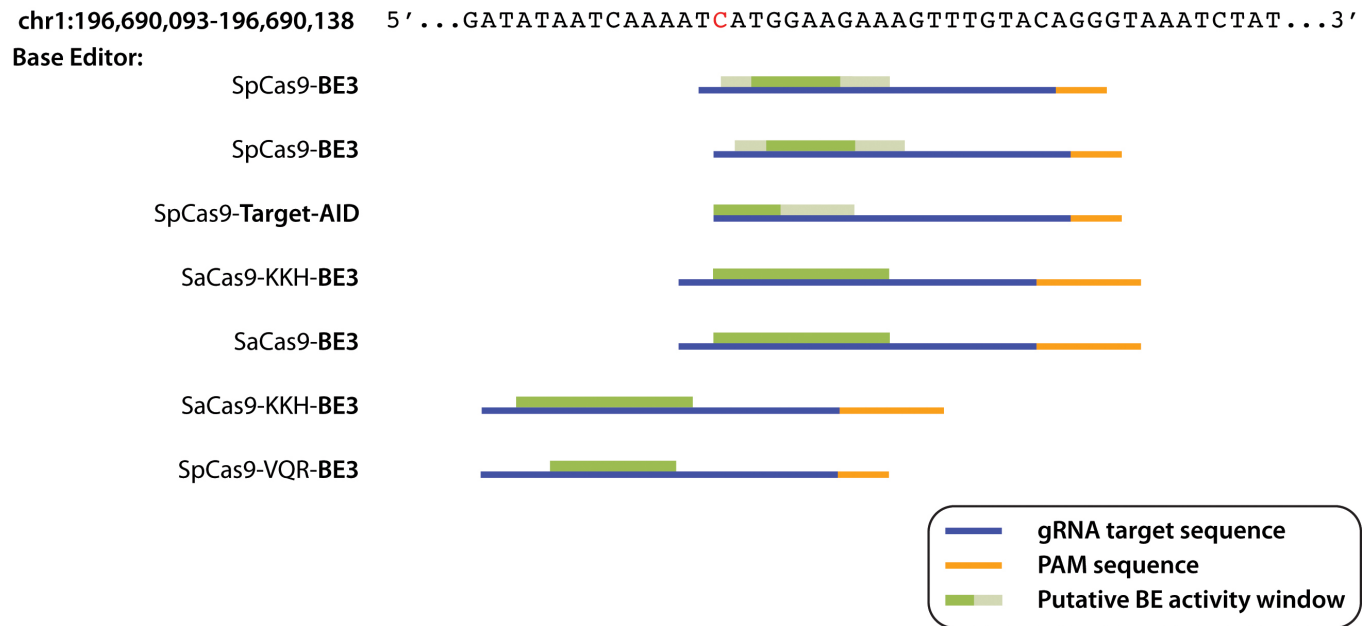


Figure 1. Activity of different base editors targeting the rs1061170 SNP. The base editor variant, sgRNA protospacer and PAM consensus nucleotide, and incidence of base editing are indicated. Base editing was determined if a peak was observed at the risk variant allele following analysis of Sanger sequencing chromatograms compared to the *LacZ* negative control (n = 3 biologic replicates performed on different days for all base editor constructs). PAM sequences and gRNA target sequences are shown in orange and blue, respectively. Putative base editor activity windows are displayed in green, where dark green indicates higher editing efficiency and light green denotes lower editing efficiency.

gDNA samples using the previously described thermocycling conditions. A secondary PCR reaction was performed on the resultant PCR amplicons following validation of the correct size using the gel purification method as previously described. Briefly, the Nextera™ DNA CD Indexes (24 Indexes, 24 Samples; Illumina, Inc., San Diego, CA) was used to barcode the previously tagged PCR amplicons containing the Illumina adaptor overhangs. The *LacZ*-negative control treated or base-edited samples each have, respectively, one on-target PCR amplicon and 11 off-target PCR amplicons. The base-edited or *LacZ*-negative control PCR amplicons were each barcoded individually according to a secondary PCR reaction (Appendix 3) [35]. The resultant PCR amplicon library was then quantified using Qubit measurement and normalized to 4 nM before being pooled into a single tube for next generation deep sequencing using the MiSeq Reagent Kits v2 (500-cycles; Illumina, Inc. CA). The matlab script, which was adapted from Gaudelli and colleagues [35], was used for determining the incidence of cytosine-to-thymine nucleotide changes in either the on-target treated sample or the off-target sites, relative to the *LacZ* negative control (Appendix 3).

RESULTS

Base editor constructs (BE3, Target-AID, SaBE3, SaKKH-BE3, and VQR-BE3) were co-transfected with their respective sgRNA expressing cassettes targeting the rs1061170 locus in the HEK293A-CFH(p.His402) fragment cell line (Figure 1). No base editing was observed on direct Sanger sequencing with the BE3, SaBE3, SaKKH-BE3, or VQR-BE3 constructs, whereas appreciable cytosine-to-thymine editing was observed with the Target-AID construct (Figure 2A; n = 3).

Next-generation deep sequencing was performed on the Target-AID base-edited cells. The median number of reads was 53,631 across the on-target and putative off-target loci (ranging from 569 to 430,308 reads). The overall editing efficiency showed that a cytosine-to-thymine nucleotide conversion occurred in 21.5% of the total sequencing reads at the target risk allele (pooled amplicons, n = 3). Nucleotide product purity for the base edited locus also revealed no unexpected nucleotide transitions or transversions beyond that of the intended C:G to T:A nucleotide transition (Figure 2B). The target protospacer revealed a minimal incidence of a G:C to A:T nucleotide transition at the G8 nucleotide. Further, the incidence of indel formation at the target region was also evaluated and revealed an indel formation frequency of only 0.15%.

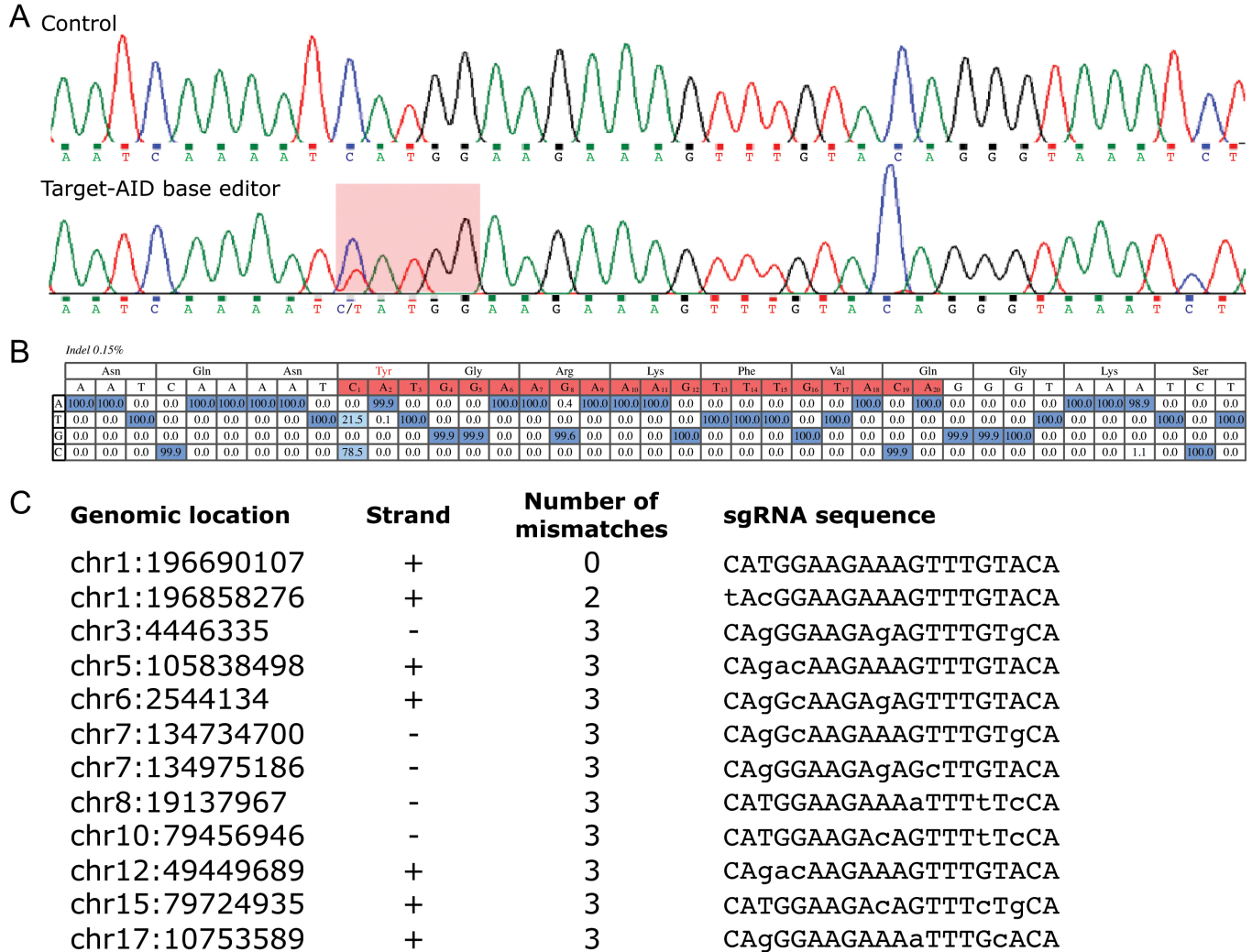


Figure 2. Editing efficiency of the Target-AID base editor at Y402H *CFH* variant (*rs1061170*). **A**: Sanger sequencing chromatograms showing the target locus. Control refers to HEK293A-*CFH*(p.His402) cells that have been treated with a combination of the Target-AID base editor with a sgRNA targeting *LacZ*. The second chromatogram displays results following transfection with the Target-AID base editor and a sgRNA specific to the *rs1061170* region. The highlighted region is the putative activity window of Target-AID showing the incidence of a C:G to T:A nucleotide correction. **B**: Base calling for targeted next-generation deep sequencing of the base-editor treated HEK293A-*CFH*(p.His402) cells. Percentages represent the fraction of sequencing reads occupied by a particular nucleotide at the specified position. The target protospacer is highlighted in red and each nucleotide is numbered 1 to 20 in which the PAM consensus sequence is considered positions +21 to +23. Deep sequencing shows that 21.5% of the target cytosines were edited into thymine nucleotides at position C1. Additionally, the percentage of indel formation is 0.15% (amplicon results of three independent biologic replicates). **C**: List of sgRNA sequences screened at the top 11 off-target sites that also contain at least one cytosine in the putative Target-AID activity window in the presence of a NGG PAM site. Each off-target site was evaluated using next generation deep sequencing and no off-target base editing was identified.

To investigate the incidence of unintended off-target effects, the top 11 putative off-target sites were profiled. Off-target sites were selected based on the criteria of one-to-three nucleotide mismatches occurring within the sgRNA protospacer relative to the target locus sgRNA, and the incidence of at least one cytosine nucleotide occurring within the expected -1 to 5 nucleotide window of the sequence space flanked by an NGG-compatible PAM consensus sequence.

Next-generation deep sequencing showed no detectable off-target effects at any of the nucleotide positions within the profiled protospacer sequence and their surrounding regions (Figure 2C).

Using the HEK293A-*CFH*(p.His402) fragment cells, we also investigated the utility of a SpCas9 scaffold enhancement. Cotransfection of a plasmid expressing SpCas9 and a sgRNA targeting the *rs1061170* SNP demonstrated clear evidence of

indel formation occurring three nucleotides upstream from the GGG PAM site ($n = 3$). Incidence of the neighboring AGG PAM motif ($n = 3$, data not shown) showed poor indel formation at this nucleotide sequence, suggesting poor Cas9:sgRNA localization. In an attempt to further optimize the editing efficiency of the CRISPR/Cas9 platform, the sgRNA component was altered with key modifications aimed at increasing the stability of Cas9:sgRNA complexation. We evaluated the effects of reducing the linker length between critical stem loop regions of the sgRNA scaffold (stem loops one and two), extending the repeat:anti-repeat duplex of the scaffold, introducing key nucleotide substitutions in the sequence space following the sgRNA protospacer to remove a putative transcription stop signal from within the sgRNA scaffold, and introducing a stabilizing nucleotide substitution into stem loop one to reduce the incidence of strong interactions adversely affecting sgRNA folding. Despite these modifications, the overall effects of modifying the sgRNA scaffold were negligible when compared to either the original sgRNA scaffold or the modified variant with the extended repeat:anti-repeat duplex. Overall cutting efficiency indicating Cas9:sgRNA localization did not show further improvements than what had already been observed (Appendix 2).

DISCUSSION

This proof-of-concept work demonstrates the feasibility of targeting the AMD-associated high Y402H *CFH* risk-variant cytosine by facilitating a cytosine-to-thymine (or guanine-to-adenine) nucleotide transition. We observed that base editing occurs in a remarkably precise manner in which only a single nucleotide correction is seen at the target cytosine, with no significant off-target deamination event occurring outside the sequence space of rs1061170. Given the characteristically low incidence of cytosines occurring within the proximal sequence space, the resultant deamination event is highly localized. Additionally, there were virtually no detectable off-target deamination events observed at the top eleven off-target sites and no appreciable indel formation was observed despite the use of a nickase variant of SpCas9.

Our work extensively explored the immediate sequence space of the rs1061170 SNP by evaluating the incidence of base editing with a plethora of PAM-diverse base editor variants and their respective sgRNAs. It is tempting to consider the four-to-five nucleotide activity window of the base editor as generous when the target is only a single nucleotide. Nonetheless, the task of selecting the most appropriate base editor can be frustrated by the spatially exhaustive exercise of re-orientating the sgRNA, the sense strand, and the antisense

strand. For example, additional parameters require careful consideration, such as the restrictive PAM site requirements, the apparent activity bias of some sgRNAs sequences, and the incidence of unintended amino acid residue substitutions in the surrounding sequence space. Here, base editing was only observed with the pairing between the Target-AID base editor construct and its respective sgRNA. No base editing was observed with other constructs despite the target cytosine occurring in several putative base editor activity windows (Figure 1). This finding is likely attributable to the fact that the optimal activity window varies across the putative window and is specific to each protospacer sequence and base editor construct [36]. For example, in a 20-nucleotide spacer, no run-off base editing at the C1 position was observed from the activity of BE3, which has a maximal activity window positioned between four and eight nucleotides wherein the PAM consensus sequence is considered to be positions 21 to 23 nucleotides. This observation strongly highlights the importance of overlapping the suggested, putative activity window of the base editor construct with its target cytosine. Although, the use of other PAM-relaxed Cas9 orthologs, such as the recently developed xCas9 or SpCas9-NG, may appear attractive, it was noted that a significant proportion of the introduced mutations, whether through random mutagenesis in the case of xCas9 or the rational design for SpCas9-NG, appeared to affect the on-target base editing efficiency of these variants at the NGG PAM sites [37, 38]. Nonetheless, other factors such as the sequence-specific nucleotide context of each sgRNA may also be a prevailing factor in determining base editing efficiency.

We evaluated two neighboring SpCas9 sgRNAs corresponding to a GGG and an AGG PAM sequence (Figure 1) and found poorer indel formation with the AGG ($n = 3$, data not shown) PAM consensus sequence relative to the GGG PAM motif, which was an observation that mirrored the apparent partiality of some nucleotide sequences for Cas9:sgRNA localization and activation [39, 40]. Chadwick and colleagues tested over 30 different sgRNAs but found that only 12 were amenable to base editing at their assayed locus [41]. Although we attempted to address the potential incidence of misfolded sgRNA scaffolds by introducing key mutations aimed at reducing internal RNA interactions and interruptions to structurally significant motifs within the scaffold, we observed no further increase in Cas9 activity than what was already evident with either the original scaffold or the F-E modified variant [42]. Prior to evaluating the incidence of base editing, each sgRNA sequence was ranked algorithmically and their scores suggested appreciable activity at the suggested sites, which we initially confirmed

by observing the frequency of indel formation at the loci [40, 43].

Future directions for this work would be to evaluate the changes in the serum levels of oxidative stress biomarkers in an in vivo system with a risk-variant encoded CFH protein profile against that of the corrected variant. Promising human-chimeric mice models have shown that some functional, photoreceptor rescue was possible when the risk variant allele was replaced with its non-risk variant form [44]. Other mouse models have further elucidated the in vivo mechanistic consequences of the Y402H mutation and its role in increasing the serum biomarkers for oxidative stress [44, 45]. Therefore, it would be reasonable to suggest that the use of the Target-AID base editor with its respective sgRNA could provide scope for measuring the in vivo recovery of photoreceptor activity.

In conclusion, we demonstrated the feasibility of using a cutting-free and potentially indel-free approach toward facilitating gene-based anticipatory therapy for a common high-risk variant that is found to be strongly associated with AMD. Appreciable base editing efficiencies were noted at the target loci in a highly precise manner with virtually no observed off-target events or indel formation. Here, the feasibility of base editing on one of the most significant risk factors for AMD progression has significant implications for the development of a potential, single-dose injectable therapy targeting dysregulated RPE cells and reducing the systemic contribution of cleavage-incompetent CFH proteins produced by the liver [49], if administered to the liver and the RPE using DNA-free approaches [50].

APPENDIX 1. MODIFICATION OF THE SGRNA SCAFFOLD BACKBONE REVEALS NEGLIGIBLE MUTATOR EFFECTS ON THE FREQUENCY OF INDEL FORMATION.

(A) A schematic overview of the sgRNA scaffold. The linker between stem loops one and two was truncated to a single U nucleotide, whereas the incidence of strong internal gRNA interactions within stem loop 1 was abrogated by introducing a canonical A54:U60 nucleotide substitution. The combination of these two mutations show robust tolerance for the scaffold to be modified. The F-E modified scaffold (d) comprising an extended repeat:anti-repeat region and tetraloop also shows comparable effects on indel formation frequency [42]. (B) The frequency of indel formation following the introduction of these mutations shows a negligible mutator effect. To access the data, click or select the words “Appendix 1.”

APPENDIX 2.

Oligonucleotide primer and gBlocks Gene Fragments sequences and used for the construction and sequencing of constructs. To access the data, click or select the words “Appendix 2.”

APPENDIX 3.

Twelve pairs of *CFH*-specific primers containing Illumina adapters were used to subject base-edited and *LacZ*-negative control treated gDNA samples to next generation deep sequencing. The chromosomal position and the reference nucleotide sequence for the Matlab base calling script is displayed [35]. To access the data, click or select the words “Appendix 3.”

ACKNOWLEDGMENTS

This work was supported by the Ophthalmic Research Institute of Australia and the Macular Disease Foundation of Australia. Financial support was also obtained from an Australian National Health and Medical Research Council (NHMRC) Centres of Research Excellence (CRE) #1023911 and Project Grant (APP1123329). JEC and AWH are supported by NHMRC Fellowships, while AP is supported by an ARC Future Fellowship. The Centre for Eye Research Australia (CERA) receives Operational Infrastructure Support from the Victorian Government. The contents of the published material are solely the responsibility of the Administering Institution, a Participating Institution or individual authors and do not reflect the views of the NHMRC. We gratefully thank Vikrant Singh for helping with the matlab scripts for base calling and indel formation.

REFERENCES

1. Resnikoff S, Pascolini D, Etya'ale D, Kocur I, Pararajasegaram R, Pokharel GP, Mariotti SP. Global data on visual impairment in the year 2002. Bull World Health Organ 2004; 82:844-51. [PMID: 15640920].
2. Wong WL, Su X, Li X, Cheung CMG, Klein R, Cheng C-Y, Wong TY. Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: a systematic review and meta-analysis. Lancet Glob Health 2014; 2:e106-16. [PMID: 25104651].
3. Guymer R. The genetics of age-related macular degeneration. Clin Exp Optom 2001; 84:182-9. [PMID: 12366313].
4. Booij JC, Baas DC, Beisekeeva J, Gorgels TGMF, Bergen AAB. The dynamic nature of Bruch's membrane. Prog Retin Eye Res 2010; 29:1-18. [PMID: 19747980].

5. Coleman HR, Chan C-C, Ferris FL III, Chew EY. Age-related macular degeneration. *Lancet* 2008; 372:1835-45. [PMID: 19027484].
6. Thakkinstian A, Han P, McEvoy M, Smith W, Hoh J, Magnusson K, Zhang K, Attia J. Systematic review and meta-analysis of the association between complement factor H Y402H polymorphisms and age-related macular degeneration. *Hum Mol Genet* 2006; 15:2784-90. [PMID: 16905558].
7. Klein RJ, Zeiss C, Chew EY, Tsai JY, Sackler RS, Haynes C, Henning AK, SanGiovanni JP, Mane SM, Mayne ST, Bracken MB, Ferris FL, Ott J, Barnstable C, Hoh J. Complement factor H polymorphism in age-related macular degeneration. *Science* 2005; 308:385-9. [PMID: 15761122].
8. Ripoche J, Day AJ, Harris TJR, Sim RB. The complete amino acid sequence of human complement factor H. *Biochem J* 1988; 249:593-602. [PMID: 2963625].
9. Ormsby RJ, Ranganathan S, Tong JC, Griggs KM, Dimasi DP, Hewitt AW, Burdon KP, Craig JE, Hoh J, Gordon DL. Functional and structural implications of the complement factor H Y402H polymorphism associated with age-related macular degeneration. *Invest Ophthalmol Vis Sci* 2008; 49:1763-70. [PMID: 18263814].
10. Clark SJ, Bishop PN, Day AJ. The Proteoglycan Glycomatrix: A Sugar Microenvironment Essential for Complement Regulation. *Front Immunol* 2013; 4:412-[PMID: 24324472].
11. Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA, Eckert MR, Vogel J, Charpentier E. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 2011; 471:602-7. [PMID: 21455174].
12. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 2012; 337:816-21. [PMID: 22745249].
13. Jinek M, East A, Cheng A, Lin S, Ma E, Doudna J. RNA-programmed genome editing in human cells. *eLife* 2013; 2:e00471-[PMID: 23386978].
14. Xu T, Li Y, Van Nostrand JD, He Z, Zhou J. Cas9-based tools for targeted genome editing and transcriptional control. *Appl Environ Microbiol* 2014; 80:1544-52. [PMID: 24389925].
15. Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 2016; 533:420-4. [PMID: 27096365].
16. Gasiunas G, Barrangou R, Horvath P, Siksnys V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci USA* 2012; 109:E2579-86. [PMID: 22949671].
17. Sapranaukas R, Gasiunas G, Fremaux C, Barrangou R, Horvath P, Siksnys V. The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res* 2011; 39:9275-82. [PMID: 21813460].
18. Christian M, Cermak T, Doyle EL, Schmidt C, Zhang F, Hummel A, Bogdanove AJ, Voytas DF. Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* 2010; 186:757-61. [PMID: 20660643].
19. Rahdar M, McMahon MA, Prakash TP, Swayze EE, Bennett CF, Cleveland DW. Synthetic CRISPR RNA-Cas9-guided genome editing in human cells. *Proc Natl Acad Sci USA* 2015; 112:E7110-7. [PMID: 26589814].
20. Komor AC, Zhao KT, Packer MS, Gaudelli NM, Waterbury AL, Koblan LW, Kim YB, Badran AH, Liu DR. Improved base excision repair inhibition and bacteriophage Mu Gam protein yields C:G-to-T:A base editors with higher efficiency and product purity. *Sci Adv* 2017; 3:4774-.
21. Hess GT, Fresard L, Han K, Lee CH, Li A, Cimprich KA, Montgomery SB, Bassik MC. Directed evolution using dCas9-targeted somatic hypermutation in mammalian cells. *Nat Methods* 2016; 13:1036-42. [PMID: 27798611].
22. Ma Y, Zhang J, Yin W, Zhang Z, Song Y, Chang X. Targeted AID-mediated mutagenesis (TAM) enables efficient genomic diversification in mammalian cells. *Nat Methods* 2016; 13:1029-35. [PMID: 27723754].
23. Sohail A, Klapacz J, Samaranyake M, Ullah A, Bhagwat AS. Human activation-induced cytidine deaminase causes transcription-dependent, strand-biased C to U deaminations. *Nucleic Acids Res* 2003; 31:2990-4. [PMID: 12799424].
24. Jore MM, Lundgren M, van Duijn E, Bultema JB, Westra ER, Waghmare SP, Wiedenheft B, Pul Ü, Wurm R, Wagner R, Beijer MR, Barendregt A, Zhou K, Snijders APL, Dickman MJ, Doudna JA, Boekema EJ, Heck AJR, van der Oost J, Brouns SJJ. Structural basis for CRISPR RNA-guided DNA recognition by Cascade. *Nat Struct Mol Biol* 2011; 18:529-36. [PMID: 21460843].
25. Kim YB, Komor AC, Levy JM, Packer MS, Zhao KT, Liu DR. Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions. *Nat Biotechnol* 2017; 35:371-6. [PMID: 28191901].
26. Nishida K, Arazoe T, Yachie N, Banno S, Kakimoto M, Tabata M, Mochizuki M, Miyabe A, Araki M, Hara KY, Shimatani Z, Kondo A. Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science* 2016; 353:Internet[PMID: 27492474].
27. Harris RS, Petersen-Mahrt SK, Neuberger MS. RNA editing enzyme APOBEC1 and some of its homologs can act as DNA mutators. *Mol Cell* 2002; 10:1247-53. [PMID: 12453430].
28. Bransteitter R, Pham P, Calabrese P, Goodman MF. Biochemical analysis of hypermutational targeting by wild type and mutant activation-induced cytidine deaminase. *J Biol Chem* 2004; 279:51612-21. [PMID: 15371439].
29. Chelico L, Pham P, Calabrese P, Goodman MF. APOBEC3G DNA deaminase acts processively 3'→5' on single-stranded DNA. *Nat Struct Mol Biol* 2006; 13:392-9. [PMID: 16622407].
30. Pham P, Bransteitter R, Petruska J, Goodman MF. Processive AID-catalyzed cytosine deamination on single-stranded DNA simulates somatic hypermutation. *Nature* 2003; 424:103-7. [PMID: 12819663].

31. Hung SSC, Chrysostomou V, Li F, Lim JKH, Wang J-H, Powell JE, Tu L, Daniszewski M, Lo C, Wong RC, Crowston JG, Pébay A, King AE, Bui BV, Liu G-S, Hewitt AW. AAV-Mediated CRISPR/Cas Gene Editing of Retinal Cells In Vivo. *Invest Ophthalmol Vis Sci* 2016; 57:3470-6. [PMID: 27367513].
32. Montague TG, Cruz JM, Gagnon JA, Church GM, Valen E. CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. *Nucleic Acids Res* 2014; 42:W401-7. [PMID: 24861617].
33. Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* 2003; 31:3406-15. [PMID: 12824337].
34. Lim Y, Bak SY, Sung K, Jeong E, Lee SH, Kim JS, Bae S, Kim SK. Structural roles of guide RNAs in the nuclease activity of Cas9 endonuclease. *Nat Commun* 2016; 7:13350. [PMID: 27804953].
35. Gaudelli NM, Komor AC, Rees HA, Packer MS, Badran AH, Bryson DI, Liu DR. Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature* 2017; 551:464-71. [PMID: 29160308].
36. Rees HA, Liu DR. Base editing: precision chemistry on the genome and transcriptome of living cells. *Nat Rev Genet* 2018; 19:Internet.
37. Nishimasu H, Shi X, Ishiguro S, Gao L, Hirano S, Okazaki S, Noda T, Abudayyeh OO, Gootenberg JS, Mori H, Oura S, Holmes B, Tanaka M, Seki M, Hirano H, Aburatani H, Ishitani R, Ikawa M, Yachie N, Zhang F, Nureki O. Engineered CRISPR-Cas9 nuclease with expanded targeting space. *Science* 2018; 361:1259-62. [PMID: 30166441].
38. Hu JH, Miller SM, Geurts MH, Tang W, Chen L, Sun N, Zeina CM, Gao X, Rees HA, Lin Z, Liu DR. Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* 2018; 556:57-63. [PMID: 29512652].
39. Xu X, Duan D, Chen S-J. CRISPR-Cas9 cleavage efficiency correlates strongly with target-sgRNA folding stability: from physical mechanism to off-target assessment. *Sci Rep* 2017; 7:143. [PMID: 28273945].
40. Doench JG, Hartenian E, Graham DB, Tothova Z, Hegde M, Smith I, Sullender M, Ebert BL, Xavier RJ, Root DE. Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat Biotechnol* 2014; 32:1262-7. [PMID: 25184501].
41. 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* [Internet] 2017; Available from: <http://dx.doi.org/10.1161/atvbaha.117.309881>.
42. Chen B, Gilbert LA, Cimini BA, Schnitzbauer J, Zhang W, Li G-W, Park J, Blackburn EH, Weissman JS, Qi LS, Huang B. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* 2013; 155:1479-91. [PMID: 24360272].
43. Doench JG, Fusi N, Sullender M, Hegde M, Vaimberg EW, Donovan KF, Smith I, Tothova Z, Wilen C, Orchard R, Virgin HW, Listgarten J, Root DE. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol* 2016; 34:184-91. [PMID: 26780180].
44. Ding JD, Kelly U, Landowski M, Toomey CB, Groelle M, Miller C, Smith SG, Klingeborn M, Singhapricha T, Jiang H, Frank MM, Bowes Rickman C. Expression of human complement factor H prevents age-related macular degeneration-like retina damage and kidney abnormalities in aged Cfh knockout mice. *Am J Pathol* 2015; 185:29-42. [PMID: 25447048].
45. Aredo B, Li T, Chen X, Zhang K, Wang CX, Gou D, Zhao B, He Y, Ufret-Vincenty RL. A chimeric Cfh transgene leads to increased retinal oxidative stress, inflammation, and accumulation of activated subretinal microglia in mice. *Invest Ophthalmol Vis Sci* 2015; 56:3427-40. [PMID: 26030099].
46. Khandhadia S, Hakobyan S, Heng LZ, Gibson J, Adams DH, Alexander GJ, Gibson JM, Martin KR, Menon G, Nash K, Sivaprasad S, Ennis S, Cree AJ, Morgan BP, Lotery AJ. Age-related macular degeneration and modification of systemic complement factor H production through liver transplantation. *Ophthalmology* 2013; 120:1612-8. [PMID: 23562165].
47. Rees HA, Komor AC, Yeh WH, Caetano-Lopes J, Warman M, Edge ASB, Liu DR. Improving the DNA specificity and applicability of base editing through protein engineering and protein delivery. *Nat Commun* 2017; 8:15790. [PMID: 28585549].

Articles are provided courtesy of Emory University and the Zhongshan Ophthalmic Center, Sun Yat-sen University, P.R. China. The print version of this article was created on 16 March 2019. This reflects all typographical corrections and errata to the article through that date. Details of any changes may be found in the online version of the article.