Journal of Genetics and Genomics xxx (xxxx) xxx

Contents lists available at ScienceDirect

Journal of Genetics and Genomics

Journal homepage: www.journals.elsevier.com/journal-of-genetics-and-genomics/

Letter to the editor

Improving the Cpf1-mediated base editing system by combining dCas9/dead sgRNA with human APOBEC3A variants

Base editor-mediated C-to-T base conversion obviates the requirements of double-strand breaks, thereby showing promise as a tool for disease modeling and gene therapy (Gaudelli et al., 2017; Rees and Liu 2018). The most actively used base editor comprises a Cas9 nickase (nCas9) with cytidine deaminase and fused uracil DNA glycosylase inhibitor at the carboxy terminus of nCas9 to inhibit uracil N-glycosylase effects (Pearl, 2000; Kunz et al., 2009; Rees and Liu 2018). This Cas9-mediated base editor (nCas9-BE) has been successfully used to induce C-to-T base conversion in various organisms, including animals and plants (Hess et al., 2017; Rees and Liu 2018). However, the editing scope of nCas9-BE is limited by its G/C-rich protospacer adjacent motif (PAM) sequences. Catalytically inactive Cpf1 from the *Lachnospiraceae* bacterium (dLbCpf1) was recently used to substitute nCas9 to fuse with rat cytosine deaminase APOBEC1, thereby developing a new base editor, named dCpf1-BE (Li et al., 2018). Similar to nCas9-BEs, dCpf1-BE induces C-to-T conversion within the editing scope. Several unique features inherited from LbCpf1 distinguish dCpf1-BE from nCas9-BEs. dCpf1-BE can recognize T-rich PAM on the 5′ side of the protospacer and allows base editing in A/T-rich regions (Zetsche et al., 2015), thereby expanding the target scope for base conversion. Moreover, dCpf1-BE uses catalytically dead LbCpf1 to bind to the genome, creating extremely low levels of indels. However, dCpf1-BE showed relatively low editing efficiency at many targets in the human genome (Li et al., 2018), indicating that its practical applications are limited. The reason for this low activity is unknown. Recent studies revealed that nucleosomes impede the access of the CRISPR (Clustered regularly interspaced short palindromic repeats) effector to the target DNA (Tsompana and Buck, 2014; Horlbeck et al., 2016; Yarrington et al., 2018). The binding of an orthogonal and catalytically dead Cas9 (dCas9) at a proximal location of target sites could alter the local chromatin structures and make inaccessible target sites available to CRISPR effectors, thereby improving gene editing efficiency (Chen et al., 2017; Liu et al., 2019a).

We assume that the efficiency of dCpf1-BE may be affected by local chromatin structures. Therefore, we designed a strategy to add dCas9 to the proximal location sides of the dCpf1-BE target site (named dCpf1-BE+) and to increase the C-to-T base editing efficiency of the dCpf1-BE system (Fig. 1A). We first tested this concept in dCpf1-BE with rat cytosine deaminase APOBEC1 by using the *THBD* gene as the target (Fig. 1B). Six sgRNAs (single guide RNA for dCas9) for the upstream sequences and five sgRNAs for the downstream sequences of the dCpf1-BE target site were designed. The distance between the sgRNA and dCpf1-BE target sites ranged from 7 bp to 120 bp (Fig. S1). Each sgRNA or sgRNA pair with dCas9 was cotransfected with dCpf1-BE and the corresponding Cpf1 sgRNA into the HEK293 cell line. The C-to-T mutation frequency was investigated by polymerase chain reaction (PCR) and Sanger sequencing and calculated using EditR software (MoriarityLab, University of Toronto) (Kluesner et al., 2018). As shown in Fig. S1, binding a single dCas9 at either side of the Cpf1 target boosted the C-to-T conversion efficiency of dCpf1-BE, which increased by 10−14% compared with those of dCpf-BE. dCas9 binding on both sides better improved mutation efficiency, which increased by 10.5−16% compared with that of dCpf-BE. The shortened distance between dCas9 sgRNA and the Cpf1-BE target site resulted in higher C-to-T frequencies. The best enhancement of C-to-T conversion efficiency was achieved by using the two nearest sgRNAs, with a distance of 7 bp (sgRNA 6, 15.6%) or 9 bp (sgRNA 7, 13.3%), which was higher than that of the Cpf-BE group (10.3%). Consistently, when the two nearest sgRNAs (sgRNA 6 and sgRNA 7) were used together, dCas9 further increased the base mutation efficacy to 18.3% (Fig. 1B). dCas9 could promote the base mutation efficiency of Cpf1-BE, and the double nearest sgRNAs proximal to dCpf1 target sites need to be used to guarantee the highest mutation rate for the dCpf1-BE system.

To validate the universality of this strategy in genome editing, ten more target sites were tested (Fig. 1C). The base editing efficiency induced by dCpf1-BE+ was considerably improved at all the tested targets (Fig. 1E and S2). The dCpf1-BE+ system achieved different editing activities for Cs in different positions. For example, when editing the *B2M* target site, the base conversion frequencies were enhanced from 23.6% to 42.3%, 14.6% to 27.6%, and 8% to 11.6% at positions C7, C8, and C11, respectively (Fig. S2D). However, with the dCpf1-BE+ system, the C-to-T editing frequency was relatively inefficient at targets containing GC-rich context in the editing window for some genes, including *APOE*, *EMX1*, and *VEGFA* (Fig. S2C, G, and J). This performance was in agreement with a previous finding that rat cytosine deaminase APOBEC1 was inefficient in GC contexts (Komor et al., 2016). Previous reports have revealed that another deaminase, human APOBEC3A (hA3A), shows higher base editing efficiency than rat APOBEC1 (Gehrke et al., 2018; Wang et al., 2018). Particularly, hA3A performed base editing in the GC contexts. We replaced rat APOBEC1 with human APOBEC3A in the Cpf1-mediated base editing system, named dCpf1-hA3A (Fig. 1D). We investigated the activity of the 10 endogenous genes described previously in the HEK293 cell line. dCpf1-hA3A showed higher base editing efficiency than dCpf1-BE at all editable Cs (Fig. 1E and S3). For instance, at the *CFTR* target site, the C-to-T base conversion efficiencies at C8, C10, and C12 were 26.3%, 19.3%, and 35.3%, respectively. For dCpf1-hA3A, there were only 4.3%, 12%, and 15%,

https://doi.org/10.1016/j.jgg.2020.07.010

Please cite this article as: M. Lian, F. Chen, X. Huang *et al.*, Improving the Cpf1-mediated base editing system by combining dCas9/dead sgRNA with human APOBEC3A variants, Journal of Genetics and Genomics, https://doi.org/10.1016/j.jgg.2020.07.010

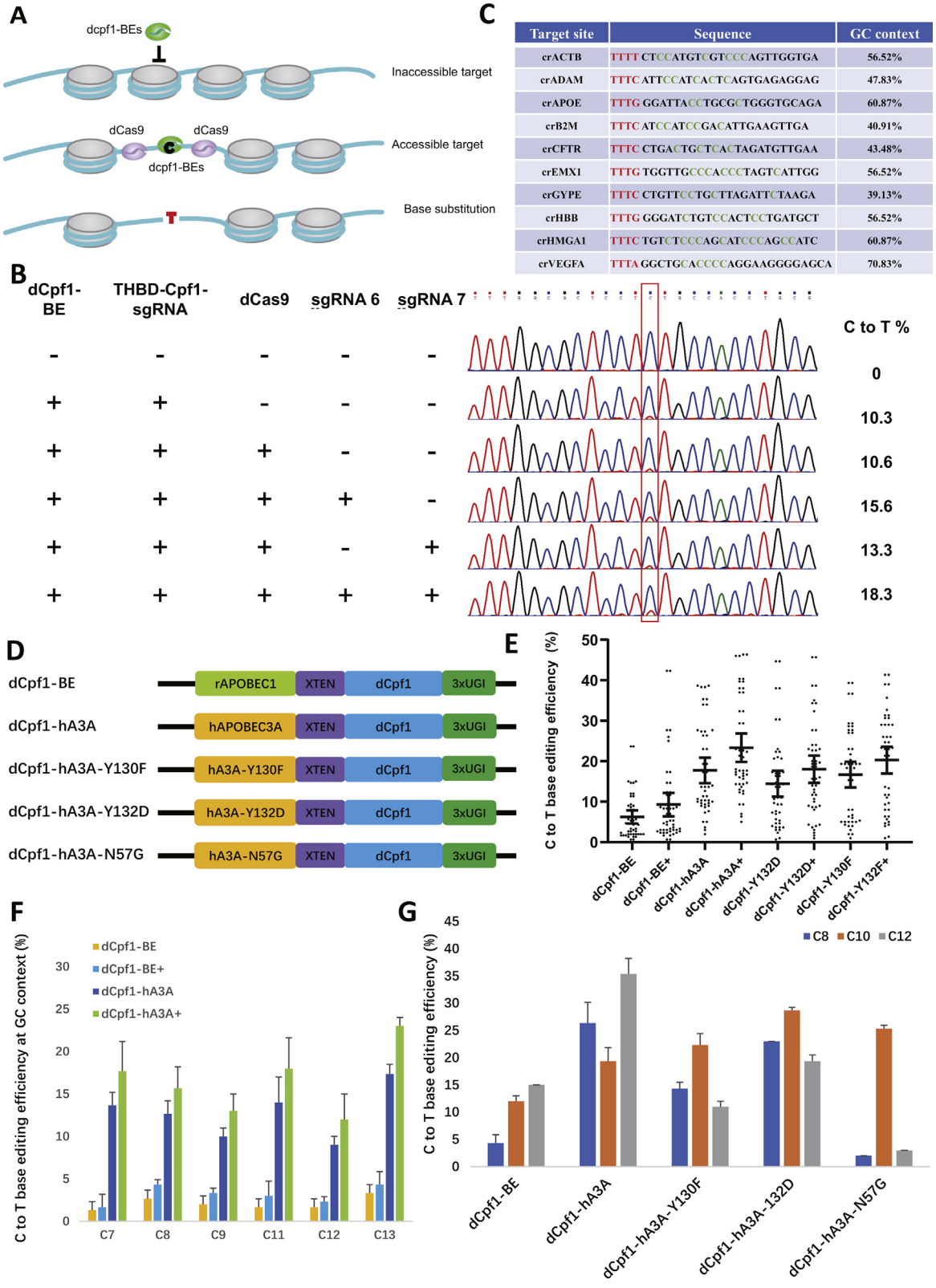

Fig. 1. dCas9 auxiliary binding and human APOBEC3A fusing could enhance the efficiency and specificity of Cpf1-mediated base editing. **A**: Schematic of the dCas9 proximity location strategy. dCpf1-BE cannot access an endogenous target in a certain chromatin context. Adding dCas9 binding to proximal locations at the dCpf1-BE target site alters the local chromatin, thereby enabling dCpf1-BE to access its target and convert C to T. **B**: Target site sequences used in this study. **C**: dCpf1-BE converts C to T efficiently under different conditions. dCas9 guided with sgRNA6 or sgRNA7 could enhance dCpf1-BE activity, and dCas9 guided with both gRNA acted synergistically. Data are representative of three independent experiments. The number of sgRNA corresponds with Fig. S1. **D**: Schematic representation of dCpf1-BE and dCpf1-hA3A fusion architecture. **E**: Representation of the general efficiencies of dCpf1-BE, dCpf1-BE+, dCpf1-hA3A, dCpf1-hA3A+, dCpf1-hA3A-Y132D, dCpf1-hA3A-Y132D+, dCpf1-hA3A-Y130F, and dCpf1-hA3A-Y130F+ at ten tested targets. Data are representative of three independent experiments. **F**: Comparison of the editing efficiency of C-to-T conversions between dCpf1-BE, dCpf1-BE+, dCpf1-hA3A, and dCpf1-hA3A+ at the *EMX1* target site. **G**: Comparison of the editing efficiency of C-to-T conversions among dCpf1-BE, dCpf1-hA3A, dCpf1-hA3A-Y130F, dCpf1-hA3A-Y132D, and dCpf1-hA3A-N57G at the *CFTR* target. The data in **F** and **G** represent the mean ± SD of three independent experiments.

Please cite this article as: M. Lian, F. Chen, X. Huang *et al.*, Improving the Cpf1-mediated base editing system by combining dCas9/dead sgRNA with human APOBEC3A variants, Journal of Genetics and Genomics, https://doi.org/10.1016/j.jgg.2020.07.010

respectively, at the corresponding Cs for dCpf1-BE (Fig. S2E and S3E). As expected, dCpf1-hA3A showed robust deaminating activity in GC contexts. For the *EMX1* target site, which contains GC contexts, dCpf1-hA3A induced higher editing efficiencies than dCpf1-BE and dCpf1-BE+ within the same editing window (Fig. 1F). Then, we added dCas9 to strengthen the base editing efficiency of the dCpf1-hA3A system (called dCpf1-hA3A+). As shown in Fig. 1E, dCpf1-hA3A+ could further enhance dCpf1-hA3A. For instance, at the *EMX1* target, the editing efficiencies of dCpf1-hA3A+ were 8.2- to 14.3-fold higher than those of dCpf1-BE (1.3% vs. 17.6% at position C7; 2.6% vs. 15.6% at position C8; 2% vs. 13% at position C9; 1.6% vs. 18% at position C11; 1.6% vs. 12% at position C12; and 3.3% vs. 23% at position C13) (Fig. 1F). We also observed a similar striking improvement in the editing efficiencies at targets with the GC context, such as *APOE, HMGA1,* and *VEGFA* (Fig. S3).

However, the editing window of hA3A was wider (approximately 10 nt at positions 6−16 in the sgRNA target site) than that of rat APOBEC1 (approximately 5 nt at positions 8−13 in the sgRNA target site). Thus, dCpf1-hA3A increased the possibility of editing unwanted bystander cytosine (Wang et al., 2018). Three engineered hA3A variants, namely, hA3A-Y130F, hA3A-Y132D, and hA3A-N57G, have been reported to overcome this drawback (Gehrke et al., 2018; Wang et al., 2018; Liu et al., 2019b). hA3A-Y130F and hA3A-Y132D could limit the editing window to that of BE3 (approximately 6 nt in human cells). In particular, hA3A-N57G showed a striking preference for the TCR (A/G) motif and reduced bystander mutation, thereby providing a highly precise base editing tool for TC contexts. We subsequently used these three engineered hA3A variants to avoid editing unwanted bystander Cs (Fig. 1D). Our results showed that dCpf1-hA3A and dCpf1-hA3A-Y130F and dCpf1-hA3A-Y132D could induce higher editing efficiencies than dCpf1-BE at all editable sites of the 10 tested targets (Fig. 1E, S4, S5). The deaminase activity of dCpf1-hA3A-N57G was high when editing Cs located in the TCR (A/G) motifs, whereas the editing efficiency was remarkably reduced in the adjacent non-TC contexts within the editing scope (Fig. 1G and S6). When dCas9 was added to a proximal location of the base editing target, the editing frequency was significantly promoted in the dCpf1-mediated base editing system with three engineered variants of APOBEC3A (Fig. 1E, S4−S6).

Off-target effect is always a major concern for any editing tool. In a Cpf1-mediated base editing system with the addition of dCas9, the off-target could be caused by Cpf1 itself, named the Cpf1-dependent off-target. In addition, dCas9 and its corresponding sgRNA in the system would yield a single-stranded DNA loop, providing a working region for cytosine deaminase (Doman et al., 2020), which could result in an extra off-target effect, named Cpf1-independent off-target.

To evaluate the specificity of the new Cpf1-mediated base editing system, we first analyzed Cpf1-dependent off-target events in 10 tested targets. Six potential off-target sites containing less than 3 mismatches were predicted using Cas-OFFinder (http://www.rgenome.net/cas-offinder/) (Bae et al., 2014) among the 10 tested targets. PCR and Sanger sequencing were performed to confirm the existence of the off-targets in cells edited by all Cpf1-mediated base editors. As shown in Fig. S7, we observed off-target events at two off-target sites, one of which was *GYPE* off-target 1, which shared exactly the same sequence as the *GYPE* target site. The other off-target site was *HBB* off-target, which contained one mismatch, but was not in the seed region. No obvious mutation was observed in the other three predicted off-targets with 3-nt mismatches. This result indicated that Cpf1-dependent off-target effects were extremely low in a Cpf1-mediated base editing system.

We then analyzed the Cpf1-independent off-target effect caused by the addition of dCas9/sgRNA through PCR and Sanger sequencing. Initially, when 20-nt sgRNA was used to guide dCas9, we observed that all five cytosine deaminases (rat APOBEC1, human APOBEC3A, and the three engineering variants of human APOBEC3A) displayed undesired C-to-T mutation, which reduced the specificity of Cpf1-mediated base editing. As shown in Fig. S8, the highest Cpf1-independent undesired C-to-T mutation rate exceeded 20% (27.6% for rat APOBEC1, 31% for hA3A, 24.6% for hA3A-Y132D, and 21.6% for hA3A-Y130F). Recently, Liu et al. (2019a) used extra dead sgRNAs with 15-nt spacer sequences, which had no capacity to mediate double-strand breaks, to alter the local accessibility and to improve Cas9 activity. Their work inspired us to speculate that the truncated spacer sequences (15 nt) could possibly reduce Cpf1-independent undesired off-target effects while still guiding dCas9 to promote Cpf1-mediated base editing frequency. To test this hypothesis, we selected 15-nt instead of 20-nt spacer sequences to overcome the off-target issues caused by dCas9/sgRNA in our system. We evaluated the effects of 15-nt sgRNA on the efficiency and off-target effects of the aforementioned 10 tested targets. As expected, the truncated sgRNA could significantly reduce dCas9/sgRNA-mediated off-target effects for all five cytosine deaminases (rat APOBEC1, human APOBEC3A, and its engineering variants.). As shown in Fig. S8, the mean value of the Cpf1-independent C-to-T mutation was reduced to one-third of that guided with a 20-bp spacer sequence. In addition, the truncated sgRNA retained the ability to guide dCas9 to promote Cpf1-mediated base editing efficacy (Fig. 1F).

In summary, we established an effective dCpf1-mediated base editing system by adding dCas9/sgRNA. In this system, the nearest double sgRNAs corresponded to dCas9 proximal to dCpf1 target sites, and hA3A engineering variants with narrowed editing windows (dCpf1-hA3A-Y130F and dCpf1-hA3A-Y132D) or TCR context preferences (dCpf1-hA3A-N57G) were used to achieve ideal mutation efficiency. The truncated sgRNA (15 bp) corresponding to dCas9 was used to overcome the off-target issues caused by dCas/gsRNA. For this system, even if the 15-nt sgRNA was used, the off-targets still existed, which most likely were Cpf1 and dCas9 independent. The most recently reported engineered APOBEC1 variant W90Y + R126E (YE1), which minimized deaminase-dependent off-target editing while retaining the substrate-targeting scope of high-activity CBEs (Doman et al., 2020), can be used to optimize our Cpf1-mediated base editing system to reduce off-target mutations.

CRediT authorship contribution statement

Han Wu, Liangxue Lai, Kepin Wang: Conceptualization, Methodology. **Meng Lian, Fangbing Chen, Xingyun Huang, Xiaozhu Zhao**: Investigation. **Shixue Gou, Nan Li, Qin Jin**: Data curation. **Han Wu, Meng Lian**: Writing - original draft preparation. **Hui Shi, Yanhui Liang, Jingke Xie**: Supervision. **Weikai Ge, Zhenpeng Zhuang, Jiaowei Wang, Yinghua Ye, Yi Yang**: Validation. **Kepin Wang**: Writing - reviewing and editing.

Acknowledgments

This work was financially supported by the National Natural Science Foundation of China (31901048), the National Key Research and Development Program of China (2017YFA0105103, 2019YFA0111500), the Strategic Priority Research Program of the Chinese Academy of Sciences (XDA16030503), Key Research & Development Program of Guangzhou Regenerative Medicine and Health Guangdong Laboratory (2018GZR110104004), the Youth Innovation Promotion Association of the Chinese Academy of Sciences (2019347), Science and Technology Planning Project of Guangdong Province, China (2019A1515012090, 2017B020231001, 2017B030314056, 2014B020225003), Science and Technology

Program of Guangzhou, China (201704030034, 202007030003, 202002030382), and Research Unit of Generation of Large Animal Disease Models, Chinese Academy of Medical Sciences (2019-I2M-5-025).

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jgg.2020.07.010.

References

Bae, S., Park, J., Kim, J.S., 2014. Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. Bioinformatics 30, 1473–1475.

Chen, F., Ding, X., Feng, Y., Seebeck, T., Jiang, Y., Davis, G.D., 2017. Targeted activation of diverse CRISPR-Cas systems for mammalian genome editing via proximal CRISPR targeting. Nat. Commun. 8, 14958.

Doman, J.L., Raguram, A., Newby, G.A., Liu, D.R., 2020. Evaluation and minimization of Cas9-independent off-target DNA editing by cytosine base editors. Nat. Biotechnol. 38, 620–628.

Gaudelli, N.M., Komor, A.C., Rees, H.A., Packer, M.S., Badran, A.H., Bryson, D.I., Liu, D.R., 2017. Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. Nature 551, 464–471.

Gehrke, J.M., Cervantes, O., Clement, M.K., Wu, Y., Zeng, J., Bauer, D.E., Pinello, L., Joung, J.K., 2018. An APOBEC3A-Cas9 base editor with minimized bystander and off-target activities. Nat. Biotechnol. 36, 977–982.

Hess, G.T., Tycko, J., Yao, D., Bassik, M.C., 2017. Methods and applications of CRISPR-mediated base editing in eukaryotic genomes. Mol. Cell 68, 26–43.

Horlbeck, M.A., Witkowsky, L.B., Guglielmi, B., Replogle, J.M., Gilbert, L.A., Villalta, J.E., Torigoe, S.E., Tjian, R., Weissman, J.S., 2016. Nucleosomes impede Cas9 access to DNA *in vivo* and *in vitro*. Elife 5, e12677.

Kluesner, M.G., Nedveck, D.A., Lahr, W.S., Garbe, J.R., Abrahante, J.E., Webber, B.R., Moriarity, B.S., 2018. EditR: a method to quantify base editing from sanger sequencing. Crispr. J. 1, 239–250.

Komor, A.C., Kim, Y.B., Packer, M.S., Zuris, J.A., Liu, D.R., 2016. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature 533, 420–424.

Kunz, C., Saito, Y., Schar, P., 2009. DNA repair in mammalian cells: mismatched repair: variations on a theme. Cell. Mol. Life Sci. 66, 1021–1038.

Li, X., Wang, Y., Liu, Y., Yang, B., Wang, X., Wei, J., Lu, Z., Zhang, Y., Wu, J., Huang, X., Yang, L., Chen, J., 2018. Base editing with a Cpf1-cytidine deaminase fusion. Nat. Biotechnol. 36, 324–327.

Liu, G., Yin, K., Zhang, Q., Gao, C., Qiu, J.L., 2019a. Modulating chromatin accessibility by transactivation and targeting proximal dsgRNAs enhances Cas9 editing efficiency *in vivo*. Genome Biol. 20, 145.

Liu, Z., Chen, S., Shan, H., Zhang, Q., Chen, M., Lai, L., Li, Z., 2019b. Efficient and precise base editing in rabbits using human APOBEC3A-nCas9 fusions. Cell Discov. 5, 31.

Pearl, L.H., 2000. Structure and function in the uracil-DNA glycosylase superfamily. Mutat. Res. 460, 165–181.

Rees, H.A., Liu, D.R., 2018. Base editing: precision chemistry on the genome and transcriptome of living cells. Nat. Rev. Genet. 19, 770–788.

Tsompana, M., Buck, M.J., 2014. Chromatin accessibility: a window into the genome. Epigenet. Chromatin 7, 33.

Wang, X., Li, J., Wang, Y., Yang, B., Wei, J., Wu, J., Wang, R., Huang, X., Chen, J., Yang, L., 2018. Efficient base editing in methylated regions with a human APOBEC3A-Cas9 fusion. Nat. Biotechnol. 36, 946–949.

Yarrington, R.M., Verma, S., Schwartz, S., Trautman, J.K., Carroll, D., 2018. Nucleosomes inhibit target cleavage by CRISPR-Cas9 *in vivo*. Proc. Natl. Acad. Sci. U.S.A. 115, 9351–9358.

Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O., Slaymaker, I.M., Makarova, K.S., Essletzbichler, P., Volz, S.E., Joung, J., van der Oost, J., Regev, A., Koonin, E.V., Zhang, F., 2015. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. Cell 163, 759–771.

Meng Lian[1]
CAS Key Laboratory of Regenerative Biology, Guangdong Provincial Key Laboratory of Stem Cell and Regenerative Medicine, Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou, 510530, China

Institute of Physical Science and Information Technology, Anhui University, Hefei, 230601, China

Fangbing Chen[1], Xingyun Huang[1], Xiaozhu Zhao, Shixue Gou, Nan Li, Qin Jin, Hui Shi, Yanhui Liang, Jingke Xie, Weikai Ge, Zhenpeng Zhuang, Jiaowei Wang
CAS Key Laboratory of Regenerative Biology, Guangdong Provincial Key Laboratory of Stem Cell and Regenerative Medicine, Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou, 510530, China

University of Chinese Academy of Sciences, Beijing, 100049, China

Yinghua Ye
CAS Key Laboratory of Regenerative Biology, Guangdong Provincial Key Laboratory of Stem Cell and Regenerative Medicine, Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou, 510530, China

Research Unit of Generation of Large Animal Disease Models, Chinese Academy of Medical Sciences (2019RU015), Guangzhou, 510530, China

Bioland Laboratory (Guangzhou Regenerative Medicine and Health Guangdong Laboratory), Guangzhou, 510005, China

Yi Yang
Key Laboratory for Major Obstetrics Diseases of Guangdong Province, Key Laboratory of Reproduction and Genetics of Guangdong Higher Education Institutes, The Third Hospital Affiliated of Guangzhou Medical University, Guangzhou, 510150, China

Kepin Wang[**], Liangxue Lai[*], Han Wu[***]
CAS Key Laboratory of Regenerative Biology, Guangdong Provincial Key Laboratory of Stem Cell and Regenerative Medicine, Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou, 510530, China

Research Unit of Generation of Large Animal Disease Models, Chinese Academy of Medical Sciences (2019RU015), Guangzhou, 510530, China

Bioland Laboratory (Guangzhou Regenerative Medicine and Health Guangdong Laboratory), Guangzhou, 510005, China

[*] Corresponding author.

[***] Corresponding author.

[**] Corresponding author.
E-mail addresses: wang_kepin@gibh.ac.cn (K. Wang), lai_liangxue@gibh.ac.cn (L. Lai), wu_han@gibh.ac.cn (H. Wu).

21 January 2020
Available online xxx

[1] These authors contributed equally to this work.