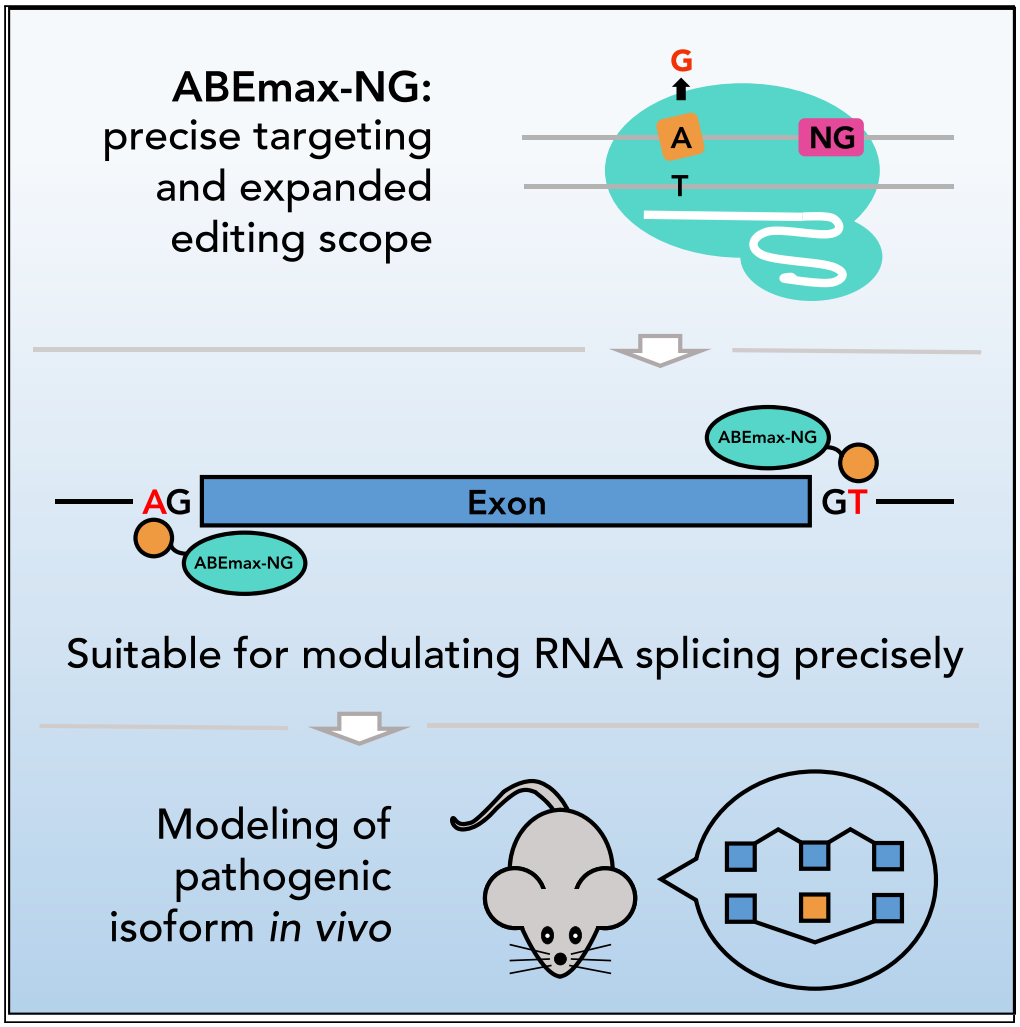
iScience

CellPress

Article

Developing ABEmax-NG with Precise Targeting and Expanded Editing Scope to Model Pathogenic Splice Site Mutations *In Vivo*



Shisheng Huang, Zhaodi Liao, Xiangyang Li, ..., Xu Ma, Qiang Sun, Xingxu Huang

nfpcc_ma@163.com (X.M.)
qsun@ion.ac.cn (Q.S.)
huangxx@shanghaitech.edu.cn (X.H.)

HIGHLIGHTS

ABEmax performs precise A•T to G•C conversion with an expanded scope

ABEmax-NG covers more splicing sites, resulting in precise RNA splicing modulation

ABEmax-NG efficiently and precisely models pathogenic RNA splicing *in vitro* and *in vivo*

Huang et al., iScience *15*, 640–648
May 31, 2019
https://doi.org/10.1016/j.isci.2019.05.008

iScience

CellPress

Article

Developing ABEmax-NG with Precise Targeting and Expanded Editing Scope to Model Pathogenic Splice Site Mutations *In Vivo*

Shisheng Huang,[1,2,7] Zhaodi Liao,[3,4,7] Xiangyang Li,[1,2,7] Zhen Liu,[3] Guanglei Li,[1,5] Jianan Li,[1,2] Zongyang Lu,[1,2] Yu Zhang,[1] Xiajun Li,[1] Xu Ma,[5,*] Qiang Sun,[3,*] and Xingxu Huang[1,6,8,*]

SUMMARY

RNA splicing is related to many human diseases; however, lack of efficient genetic approaches to modulate splicing has prevented us from dissecting their functions in human diseases. Recently developed base editors (BEs) offer a new strategy to modulate RNA splicing by converting conservative splice sites, but it is limited by the editing precision and scope. To overcome the limitations of currently available BE-based tools, we combined SpCas9-NG with ABEmax to generate a new BE, ABEmax-NG. We demonstrated that ABEmax-NG performed precise A·T to G·C conversion with an expanded scope, thus covering many more splicing sites. Taking advantage of this tool, we precisely achieved A·T to G·C conversion exactly at the splice sites. We further modeled pathogenic RNA splicing *in vitro* and *in vivo*. Taken together, we successfully generated a versatile tool suitable for precise and broad editing at the splice sites.

INTRODUCTION

As a critical biological process, RNA splicing plays many important roles in transcriptome diversity and development (Kalsotra and Cooper, 2011). Although more and more splicing isoforms have been increasingly identified with the development of new sequencing techniques, little is known about the functions of the vast majority of the splice isoforms due to lack of effective genetic approaches to modulate splicing.

Base editor (BE) systems have been recently developed that can induce single-nucleotide changes efficiently without formation of double-stranded DNA breaks (Rees and Liu, 2018). The controllability and precision of BEs make them widely used to induce or correct single point mutations (Liu et al., 2018b; Zeng et al., 2018). As splice sites are highly conserved in various species (Lim and Burge, 2001), BEs can be used to modulate RNA splicing by mutating splice sites. Indeed, recent studies have applied cytosine base editors (CBEs) to modulate multifarious RNA splicing by converting conserved guanine to adenine at either 5′ splice donor (SD) sites or 3′ splice acceptor (SA) sites (Gapinske et al., 2018; Liu et al., 2018a; Yuan et al., 2018). However, CBE-based splicing modulation is limited by the deficiency of CBE. First, CBE generates indels, although it is reduced by fusing uracil DNA glycosylase inhibitor (Wang et al., 2017). Second, CBE induces high proximal mutations, which makes it impossible to be applied to precise single-site editing (Lee et al., 2018). ABE (adenine base editor)-based splicing modulation has been successfully achieved in plant (Kang et al., 2018). Nevertheless, the requirement for the protospacer adjacent motif (PAM) to be adjacent to the target locus by current Cas9 variants limits the target sites of both ABE- and CBE-based splicing modulation. Therefore BEs suitable for splicing modulation are earnestly needed.

The application of ABE in splice site modulation in plants suggests the potential of ABE in RNA splicing modulation of animals. As ABE induces significantly few unwanted base conversions, indels, and proximal mutations (Lee et al., 2018), and a relaxed-PAM-recognized SpCas9 variant (SpCas9-NG) was developed recently (Nishimasu et al., 2018), we take advantage of them and try to combine the SpCas9-NG with ABEmax, which is the most efficient version of ABE (Koblan et al., 2018), to develop the NG-PAM-recognized adenine base editor (hereafter named ABEmax-NG). We demonstrated that ABEmax-NG has the character of precise single-base editing with expanded editing scope that is suitable for

[1]School of Life Science and Technology, ShanghaiTech University, Shanghai 201210, China

[2]University of Chinese Academy of Sciences, Beijing 100049, China

[3]Institute of Neuroscience, Chinese Academy of Sciences (CAS) Key Laboratory of Primate Neurobiology, CAS Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai 200031, China

[4]College of Life Sciences, University of Chinese Academy of Sciences, Beijing 100049, China

[5]National Research Institute for Family Planning, No. 12 Dahuishi Road, Beijing 100081, China

[6]CAS Center for Excellence in Molecular Cell Science, Shanghai Institute of Biochemistry and Cell Biology, Chinese Academy of Sciences; University of Chinese Academy of Sciences, Shanghai 200031, China

[7]These authors contributed equally

[8]Lead Contact

*Correspondence: nfpcc_ma@163.com (X.M.), qsun@ion.ac.cn (Q.S.), huangxx@shanghaitech.edu.cn (X.H.)

https://doi.org/10.1016/j.isci.2019.05.008

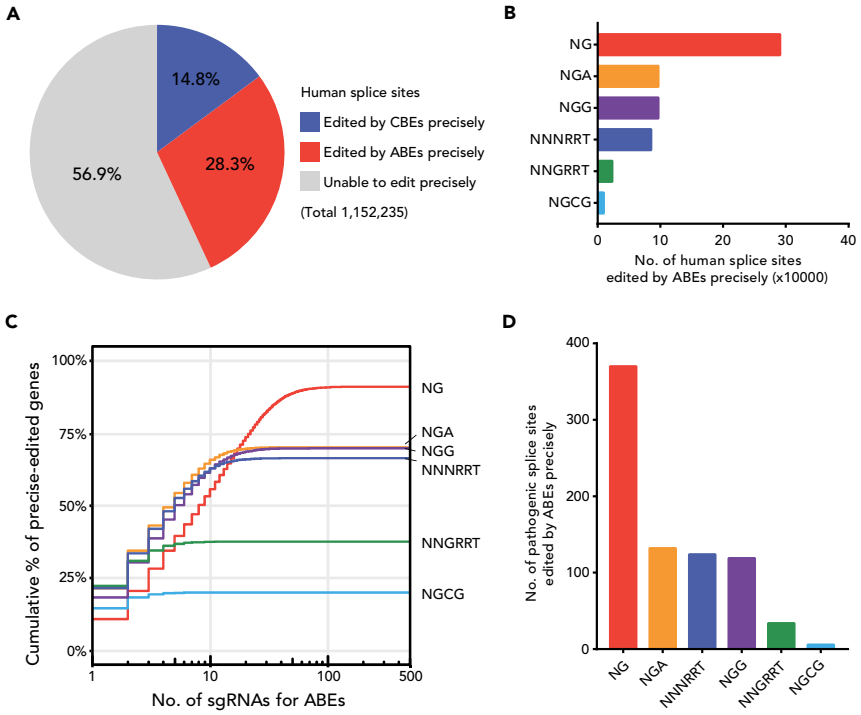


Figure 1. ABEmax-NG Has the Broadest Scope to Modulate RNA Splicing

(A) The proportion of human splice sites that can be edited precisely by CBE and ABE variants with distinct PAM specificities (NG, NGG, NGA, NGCG, NNGRRT, NNNRRT) and the corresponding editing windows. Human reference genome (hg38) and the annotation from GENCODE version 29 were used for analysis. The sgRNAs with a single target site in their editing windows were considered as precise editing sgRNAs and used for further analysis.

(B) The number of human splice sites can be edited precisely by ABEs. Distributions for distinct PAM specificities were shown.

(C) Cumulative percentage of genes whose splice sites can be edited precisely by ABEs. The horizontal axis shows the number of sgRNAs that can be designed per gene. Distributions for distinct PAM specificities were shown.

(D) The number of pathogenic human splice sites in the ClinVar database can be edited precisely by ABEs. Distributions for distinct PAM specificities were shown.

RNA splicing modulation. Furthermore, we have successfully modeled pathogenic splice site mutation *in vivo*.

RESULTS

ABEmax-NG Is a Potential Tool with the Expanded Editing Scope that Is Suitable for Modulating RNA Splicing Precisely

Given splice sites only contain highly conserved two bases in various species (Lim and Burge, 2001), BE-based splicing modulation has to be very precise. ABE is considered more precise than CBE because it induces significantly few unwanted base conversions, indels, and proximal mutations (Lee et al., 2018). Also, ABE-mediated A•T to G•C conversion can be used to mutate the SA and SD sites; we hypothesize that ABE may be more suitable for modulating splicing than CBE. To demonstrate this hypothesis, we first explored the potential of ABEmax-NG, a combination of SpCas9-NG and ABEmax, in modulating RNA splicing over the human genome by bioinformatics analysis. As there are relatively conserved guanines near the SD and SA sites (Figure S1), we found that ABEs can precisely edit about twice as many splice sites as CBEs (Figure 1A), indicating that ABEs have advantages over CBEs not only in editing precision but also in the editing scope of the target splice sites. Then we further analyzed the editing scope among different PAMs over all validated ABE variants (Rees and Liu, 2018). As expected, ABEmax-NG can cover almost all editable sites, which is far beyond other variants (Figure 1B). With this, it has the most potential single guide RNAs (sgRNAs), and meanwhile, the most targeting genes (Figure 1C), making the ABEmax-NG a flexible tool to modulate various RNA splicing reactions. As expected, ABEmax-NG is also the most efficient tool that targets most human pathogenic splice sites (Figure 1D). This

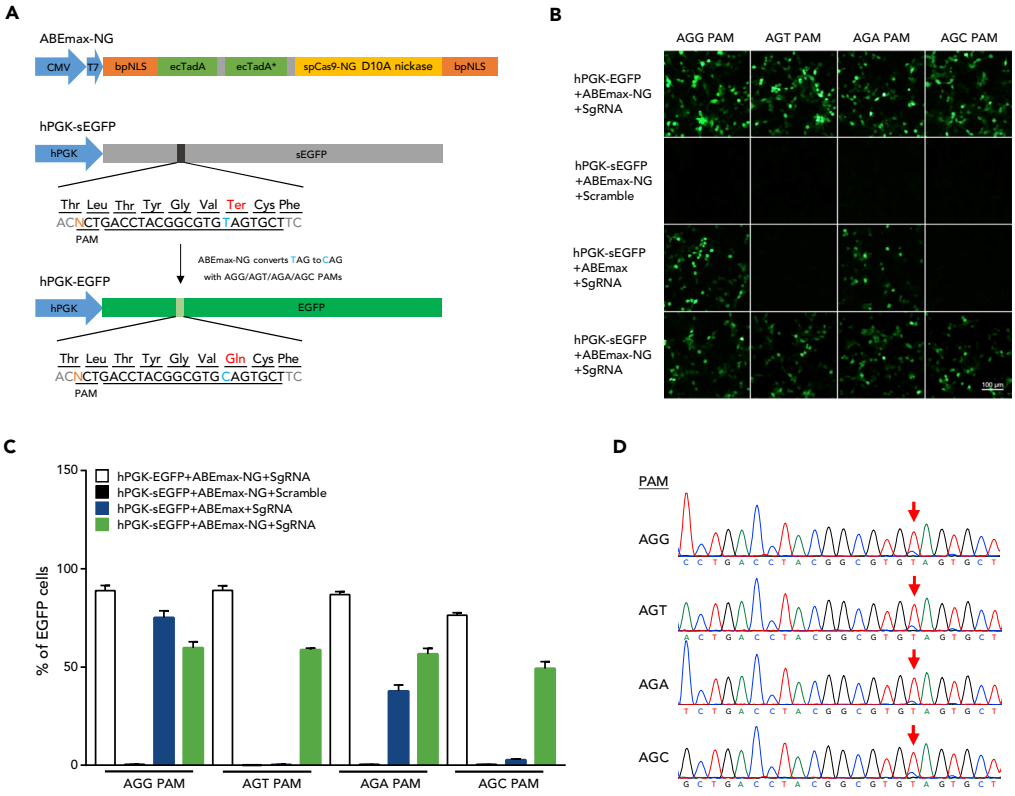


Figure 2. Verification of Versatile Base Editing of ABEmax-NG Using EGFP Reporter System

(A) ABEmax-NG construct and the schematic of the EGFP reporter system. The TAG stop codon was converted to glutamine by ABEmax-NG to restore EGFP protein production. The different third base in PAM does not change the original threonine of EGFP protein.

(B) ABEmax-NG-induced conversion of sEGFP to EGFP in HEK293FT cells. Scramble, non-targeting sgRNA served as negative control; sgRNA, sgRNA targeting sEGFP. Scale bars, 100 μm.

(C) Analysis of base editing by fluorescence-activated cell sorting. The percentage was calculated as the number of EGFP-positive cells/total transfected cells. Data are represented as the mean ± SEM (n = 3 from three independent experiments).

(D) The chromatograms of Sanger sequencing showing ABEmax-NG can induce sEGFP to EGFP conversion in different NG(N) PAM.

genome-wide analysis suggests that ABEmax-NG is a potential tool with expanded editing scope for precise modulation of RNA splicing.

The ABEmax-NG Performs Efficient Base Editing in HEK293FT Cells

To demonstrate the potential of ABEmax-NG, we first constructed it by introducing R1335V/L1111R/D1135V/G1218R/E1219F/A1322R/T1337R mutations for converting SpCas9 nickase to SpCas9-NG nickase on ABEmax (Figure 2A). Then we tested its versatility by checking different NG(N) PAM recognition of ABEmax-NG in HEK293FT cells using an enhanced green fluorescent protein (EGFP) reporter, which was developed by mutating two bases of the EGFP-coding gene, one at the third position of the Thr-63 codon, which does not change the amino acid sequence of EGFP but provides multiple PAM sequences for ABEmax-NG recognition test, and another at the first position of the Gln-69 codon, converting the Gln-69 codon (CAG) into a stop codon (TAG), thus changing EGFP to stopped-EGFP (sEGFP). This sEGFP can be restored when the stop codon is corrected by ABEmax-NG (TAG to CAG), hence the editing efficiency of ABEmax-NG in different PAM sequence can be measured as the frequency of EGFP-expressing cells (Figure 2A). Then we designed an sgRNA to target sEGFP with the mutated T at position 7 of the protospacer, counted from the PAM sequence (Figure 2A; Table S1). The plasmids expressing sgRNA and ABEmax-NG were co-transfected with the EGFP reporter system.

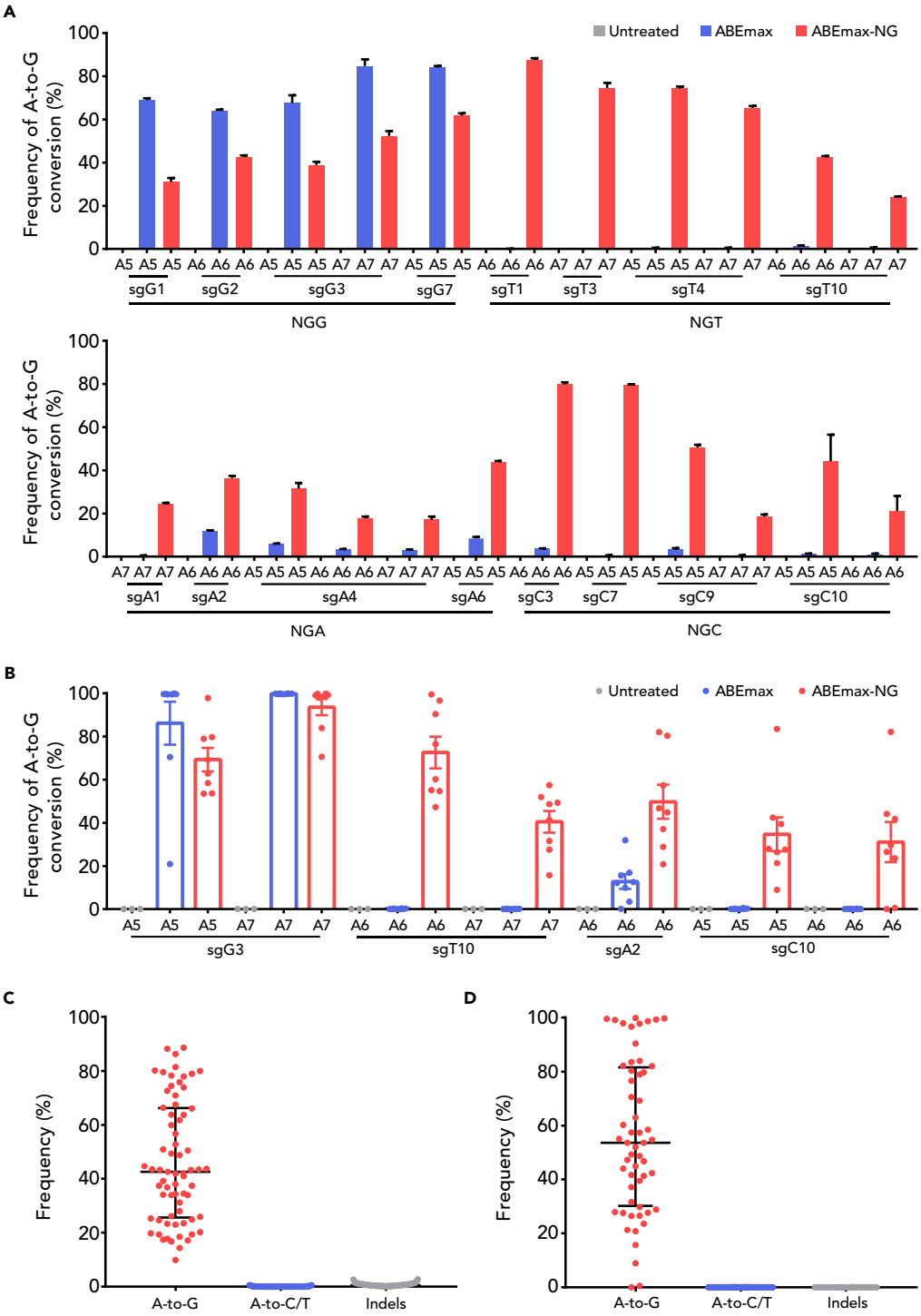


Figure 3. Efficient A-to-G Substitution in Mouse N2a Cells and Embryos by ABEmax-NG

(A) A-to-G editing by ABEmax and ABEmax-NG at 16 endogenous target sites in mouse N2a cells. The target base in the editing window was shown, counting the end distal to the protospacer adjacent motif (PAM) as position 1. Data are represented as the mean ± SEM (n = 3 from three independent experiments).

(B) A-to-G editing efficiency of ABEmax and ABEmax-NG at the four endogenous target sites in mouse embryos. Data are represented as the mean ± SEM (n = 3 for untreated, n = 8 for ABEmax, n = 8 for ABEmax-NG).

Figure 3. *Continued*

(C) Statistical analysis of the A-to-G editing frequency, unwanted base conversions, and indels induced by ABEmax-NG in (A). The median and interquartile range (IQR) are shown.

(D) Statistical analysis of the A-to-G editing frequency, unwanted base conversions, and indels induced by ABEmax-NG in (B). The median and IQR are shown.

EGFP fluorescence could be observed in the cells 48 h after transfection targeting different PAM sequences (Figure 2B). Based on the analysis by flow cytometry, ABEmax-NG can recognize all types of NG(N) PAMs efficiently. In contrast, ABEmax, as expected, only recognized the NGG PAM with high efficiency and the NGA PAM with modest efficiency (Figure 2C). Sanger sequencing results further confirmed that ABEmax-NG induces the target base editing, resulting in the conversion of sEGFP to EGFP with different NG(N) PAMs (Figure 2D). These results demonstrated that ABEmax-NG is a versatile BE with expanded editing scope.

ABEmax-NG Can Induce Precise A-to-G Base Conversions *In Vitro* and *In Vivo*

To characterize this new tool, we comprehensively examined the editing of ABEmax-NG at 40 endogenous target sites with NG(N) PAMs in mouse-derived Neuro-2a (N2a) cells (Table S1). Compared with ABEmax that only induces A-to-G conversion at the NGG sites, ABEmax-NG performed efficient editing toward all types of PAMs (Figure S2). To further analyze the ABEmax-NG-mediated editing, we next chose 16 NG(N) sites for targeted deep sequencing. As expected, ABEmax-NG induced A-to-G conversion efficiently with few unwanted base conversions (the frequency: 0%–0.54%), indels (the frequency: 0%–2.79%), and proximal mutations (the frequency: 0%–0.32%) (Figures 3A, 3C, S3A, and S3B), indicating that ABEmax-NG is a precise BE.

Then we set out to determine whether ABEmax-NG can also work well *in vivo*. To this end, four sgRNAs covering all types of NG(N) PAMs have been subjected to test experiments in a total of 64 mouse embryos as previously described (Liu et al., 2018b). As expected, efficient base editing was achieved (Figures S5A and 3B). It is noteworthy that almost no unwanted base conversion, indel, and proximal mutation were detected (Figures 3D, S3C, and S3D), confirming that ABEmax-NG acts as a precise BE *in vivo*.

ABEmax-NG Can Serve as a Useful Modulator for RNA Splicing

The precision of ABEmax-NG suggests that ABEmax-NG could be a versatile tool for RNA splicing modulation, which is only limited to two editable bases for each SD and SA site. Thus, four splicing sites corresponding to human homologous pathogenic mutations were first tested in N2a cells (Table S1). Interestingly, the targeted deep sequencing showed that all the selected splice sites were edited by ABEmax-NG efficiently and precisely (Figures S4A and S4B). As expected, the corresponding changes to RNA splicing were detected by RT-PCR amplification of the respective cDNAs (Figure S4C). These results were further confirmed by Sanger sequencing (Figure S4D). Then we tested its performance *in vivo*. As expected, ABEmax-NG also worked well in mouse embryos except one embryo harboring few indels (the frequency is 3.35%; Figures 4A and 4B). Taken together, we demonstrated that ABEmax-NG performed efficient and precise splice site editing and is suitable for modulation of RNA splicing.

The outperformance of ABEmax-NG encouraged us to model pathogenic RNA splicing in mouse. For this purpose, we focused on BBS2 gene, a member of the Bardet-Biedl syndrome (BBS) gene family. It is demonstrated that the splice site mutation of c.472-2A > G for this gene may cause human BBS, which is a developmental disorder that affects multiple systems (Innes et al., 2010). The ABEmax-NG-encoding mRNA together with sgBBS2 mRNA were co-injected into zygotes as previously (Liu et al., 2018b). A total of 40 embryos were transferred into two surrogate mothers, which generated 19 offspring (Figure S5B). Based on the results of deep sequencing, all 19 founder mice carried the expected target site mutation (Figure S6A).

It is known that the target mutation causes skipping of Exon 4 of the BBS2 gene (Figure 4C). Thus, we chose the founder mouse M12 harboring a heterozygous target mutation (Figures 4D and S6A), to further characterize the genotype and RNA splicing in different tissues (heart, liver, spleen, lung, kidney, brain, testis and intestine) at 4 weeks. As expected, all examined tissues carried the heterozygous target mutation and the isoform of the BBS2 RNA transcript skipping Exon 4 (Figures 4E, 4F, S6B, and S6C). These results demonstrated that ABEmax-NG can serve as a useful RNA splicing modulator *in vivo*.

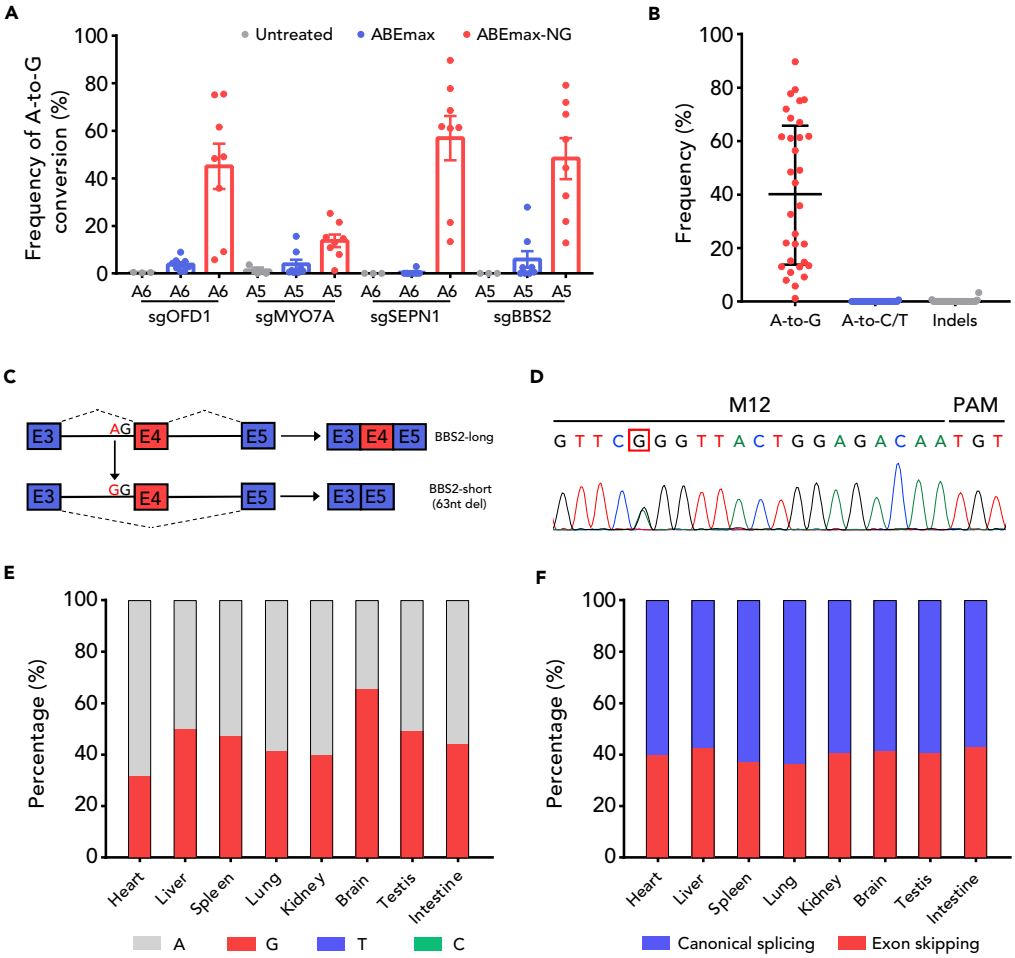


Figure 4. ABEmax-NG Mediates Pathogenic Exon Skipping *In Vivo*

(A) A-to-G editing efficiency of ABEmax and ABEmax-NG at four splice sites in mouse embryos. Data are represented as the mean ± SEM (n = 3 for untreated, n = 8 for ABEmax, n = 8 for ABEmax-NG).

(B) Statistical analysis of the A-to-G editing frequency, unwanted base conversions, and indels induced by ABEmax-NG in (A). The median and interquartile range are shown.

(C) Schematic of ABEmax-NG-induced BBS2-long to BBS2-short switch. Top panel: Exon 4 is included in normal BBS2. Bottom panel: ABEmax-NG mutates the invariant A to G at the splice acceptor site, leading to the exclusion of Exon 4.

(D) The chromatogram of Sanger sequencing of tail confirmed heterozygous genotype of founder mouse M12.

(E) The individual fraction of each base induced by ABEmax-NG in BBS2 target splice site from different tissues of founder mouse M12 at 4 weeks. Data are analyzed by deep sequencing.

(F) Quantification of the rate of exon skipping of BBS2 Exon 4 from different tissues of founder mouse M12 at 4 weeks by deep sequencing of RT-PCR amplification of the cDNA.

Off-Target Analysis by WGS

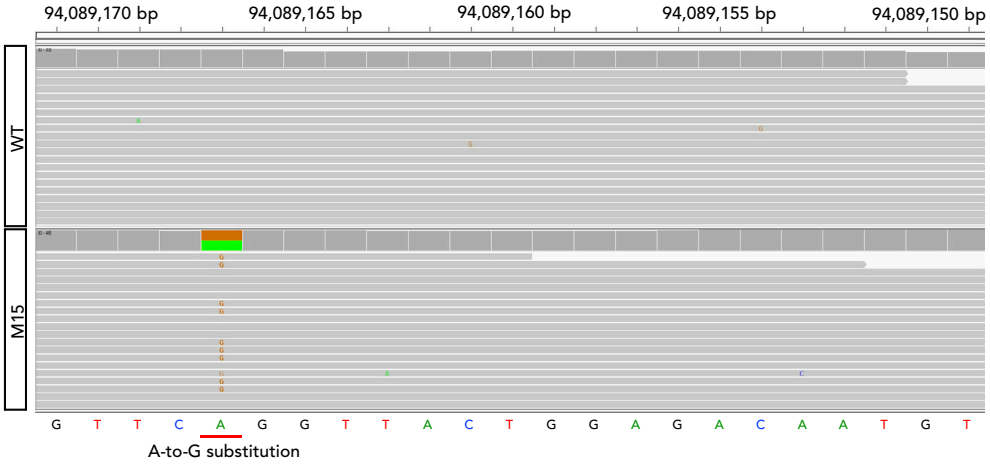
The increased off-target possibility of SpCas9-NG prompted us to comprehensively investigate the off-target effects of ABEmax-NG carefully (Nishimasu et al., 2018). We performed whole-genome sequencing (WGS) on genomic DNA samples isolated from the founder mouse M12 and a wild-type mouse (WT) at depth of 33× and 30×, respectively (Figure 5D). After filtering out mouse dbSNPs, a total of 353,868 and 330,821 single nucleotide polymorphisms were detected over the genomes of WT and M12, respectively. The mutation frequency did not seem to increase significantly in M12 (Figure 5A). The heterozygous target mutation was also confirmed by WGS (Figure 5B). Given the mismatch tolerance of SpCas9 (Cho et al., 2014), we use CasOT (Xiao et al., 2014) to explore potential off-target sites. We used the criteria that up to a mismatch of 2 bp may be included in the seed region and a mismatch of 5 bp in the non-seed region with NG PAM. Among 24,297 potential off-target sites (Figure 5C), two off-target sites that

A

	WT	M12	Uniquely assigned to M12
total SNPs	5,186,902	4,635,562	
SNPs after excluding dbSNPs	353,868	330,821	16,324
A>G SNPs	49,333	45,937	1,873
A>C SNPs	18,003	16,653	886
A>T SNPs	19,848	18,538	1,055
T>C SNPs	51,209	47,579	1,765
T>G SNPs	18,154	16,532	724
T>A SNPs	19,747	18,574	1,185

WT: Wild type

B



C

Mismatch in seed region	Mismatch in non-seed region	No. of on-/off-target sites
0	0	1
0	1-5	41
1	0-5	1438
2	0-5	22818
Total		1/24297

D

Sample	No. of reads	Mean coverage
WT	606,596,385	29.88X
M12	617,466,475	33.00X

WT: Wild type

E

Sample	On-target	Off-target
M12	1/1	2/24297

Figure 5. Whole-Genome Analysis of BBS2 Mutant (M12) and Wild-Type (WT) Mice

(A) Summary of single nucleotide polymorphism (SNP) analysis. After filtering out naturally occurring variants in the mouse SNP database, 330,821 SNPs were obtained over the M12 genome. The number of A/T conversions were shown.

(B) Confirmation of the on-target base editing by analyzing the whole-genome sequencing results of M12.

(C) Summary of on-/off-target sites. A total of 24,298 sites, including 1 on-target site and 24,297 off-target sites were analyzed.

(D) Summary of the whole-genome sequencing.

(E) Summary of off-target analysis.

are located in the intronic and intergenic regions were detected (Figures 5E, S7A, and S7B). For further demonstration, we performed Sanger sequencing of the two off-target sites, confirming the off-target mutagenesis (Figure S7C).

DISCUSSION

In this study, taking advantage of ABE with few unwanted base conversions, indels, and proximal mutations (Lee et al., 2018), and SpCas9-NG with relaxed PAM (Nishimasu et al., 2018), we introduced R1335V/L1111R/D1135V/G1218R/E1219F/A1322R/T1337R mutations into ABEmax and developed ABEmax-NG to induce A-to-G conversion with expanded editing scope and high precision. The versatility of ABEmax-NG confers it as a useful RNA splicing modulator, which is currently the most suitable tool available targeting most human pathogenic splice sites.

More and more splice isoforms have been identified with the development of sequencing technique. It is important to explore the function of the different splicing isoforms because of their potential important biological roles (Kalsotra and Cooper, 2011). The application of BE in splice site modulation opens the potential of BE in studying RNA splicing (Gapinske et al., 2018; Kang et al., 2018; Liu et al., 2018a; Yuan et al., 2018). Nevertheless, the current BE-based splicing modulation is limited. Here we found that ABEs can precisely edit about twice as many splice sites as CBEs (Figure 1A), demonstrating that ABEs have more potential for studying splicing isoform. Moreover, with the application of the relaxed-PAM-recognized SpCas9-NG, the ABEmax-NG has the broadest editing scope to modulate RNA splicing.

Considering that both SD and SA sites contain only two bases, the precision of the base editing is critical for splicing modulation. As ABE performs precise base editing (Lee et al., 2018), as expected, ABEmax-NG induced almost no unwanted base conversion and indel when targeted at the splice sites (Figures S4A, S4B, 4A, and 4B). Compared with CBE, which induced unwanted base editing around splice sites (Lee et al., 2018; Yuan et al., 2018), ABEmax-NG obviously is a better tool for modulation of RNA splicing.

In summary, we have developed ABEmax-NG, a versatile BE that is the best available tool to successfully model human pathogenic splice site mutations *in vitro* and *in vivo*.

Limitations of the Study

Although ABEmax-NG covers the majority of splice sites, there are still some isoforms that cannot be modulated.

METHODS

All methods can be found in the accompanying Transparent Methods supplemental file.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.isci.2019.05.008.

ACKNOWLEDGMENTS

We thank members of Huang lab, Sun lab, and Chen lab for helpful discussions. This work is supported by the National Key R&D Program (2016YFC0905901, 2016YFC1000307), National Natural Science Foundation of China (81830004), National Postdoctoral Program for Innovative Talents (BX201700266), and Local Grants (16JC1420200, 17JC1420103).

AUTHOR CONTRIBUTIONS

X.H., Q.S., and X.M. conceived, designed, and supervised the project. S.H., Z.L., and X.L. performed most experiments with the help of G.L., J.L., Z.L., G.Y., and Y.Z. Z.L. and X.L. provided expert technical assistance. S.H. and X.H. wrote the paper with inputs from all authors. X.L. edited the manuscript. X.H., Q.S., and X.M. managed the project.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: March 29, 2019
Revised: May 1, 2019
Accepted: May 7, 2019
Published: May 31, 2019

REFERENCES

Cho, S.W., Kim, S., Kim, Y., Kweon, J., Kim, H.S., Bae, S., and Kim, J.S. (2014). Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. Genome Res. *24*, 132–141.

Gapinske, M., Luu, A., Winter, J., Woods, W.S., Kostan, K.A., Shiva, N., Song, J.S., and Perez-Pinera, P. (2018). CRISPR-SKIP: programmable gene splicing with single base editors. Genome Biol. *19*, 107.

Innes, A.M., Boycott, K.M., Puffenberger, E.G., Redl, D., MacDonald, I.M., Chudley, A.E., Beaulieu, C., Perrier, R., Gillan, T., Wade, A., et al. (2010). A founder mutation in BBS2 is responsible for Bardet-Biedl syndrome in the Hutterite population: utility of SNP arrays in genetically heterogeneous disorders. Clin. Genet. *78*, 424–431.

Kalsotra, A., and Cooper, T.A. (2011). Functional consequences of developmentally regulated alternative splicing. Nat. Rev. Genet. *12*, 715–729.

Kang, B.C., Yun, J.Y., Kim, S.T., Shin, Y., Ryu, J., Choi, M., Woo, J.W., and Kim, J.S. (2018). Precision genome engineering through adenine base editing in plants. Nat. Plants *4*, 427–431.

Koblan, L.W., Doman, J.L., Wilson, C., Levy, J.M., Tay, T., Newby, G.A., Maianti, J.P., Raguram, A., and Liu, D.R. (2018). Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. Nat. Biotechnol. *36*, 843–846.

Lee, H.K., Willi, M., Miller, S.M., Kim, S., Liu, C., Liu, D.R., and Hennighausen, L. (2018). Targeting fidelity of adenine and cytosine base editors in mouse embryos. Nat. Commun. *9*, 4804.

Lim, L.P., and Burge, C.B. (2001). A computational analysis of sequence features involved in recognition of short introns. Proc. Natl. Acad. Sci. U S A *98*, 11193–11198.

Liu, Z., Chen, M., Chen, S., Deng, J., Song, Y., Lai, L., and Li, Z. (2018a). Highly efficient RNA-guided base editing in rabbit. Nat. Commun. *9*, 2717.

Liu, Z., Lu, Z., Yang, G., Huang, S., Li, G., Feng, S., Liu, Y., Li, J., Yu, W., Zhang, Y., et al. (2018b). Efficient generation of mouse models of human diseases via ABE- and BE-mediated base editing. Nat. Commun. *9*, 2338.

Nishimasu, H., Shi, X., Ishiguro, S., Gao, L., Hirano, S., Okazaki, S., Noda, T., Abudayyeh, O.O., Gootenberg, J.S., Mori, H., et al. (2018). Engineered CRISPR-Cas9 nuclease with expanded targeting space. Science *361*, 1259–1262.

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

Wang, L., Xue, W., Yan, L., Li, X., Wei, J., Chen, M., Wu, J., Yang, B., Yang, L., and Chen, J. (2017). Enhanced base editing by co-expression of free uracil DNA glycosylase inhibitor. Cell Res. *27*, 1289–1292.

Xiao, A., Cheng, Z., Kong, L., Zhu, Z., Lin, S., Gao, G., and Zhang, B. (2014). CasOT: a genome-wide Cas9/gRNA off-target searching tool. Bioinformatics *30*, 1180–1182.

Yuan, J., Ma, Y., Huang, T., Chen, Y., Peng, Y., Li, B., Li, J., Zhang, Y., Song, B., Sun, X., et al. (2018). Genetic modulation of RNA splicing with a CRISPR-guided cytidine deaminase. Mol. Cell *72*, 380–394.e7.

Zeng, Y., Li, J., Li, G., Huang, S., Yu, W., Zhang, Y., Chen, D., Chen, J., Liu, J., and Huang, X. (2018). Correction of the marfan syndrome pathogenic FBN1 mutation by base editing in human cells and heterozygous embryos. Mol. Ther. *26*, 2631–2637.

ISCI, Volume *15*

Supplemental Information

Developing ABEmax-NG with Precise Targeting and Expanded Editing Scope to Model Pathogenic Splice Site Mutations *In Vivo*

Shisheng Huang, Zhaodi Liao, Xiangyang Li, Zhen Liu, Guanglei Li, Jianan Li, Zongyang Lu, Yu Zhang, Xiajun Li, Xu Ma, Qiang Sun, and Xingxu Huang

Supplemental Figures and Legends

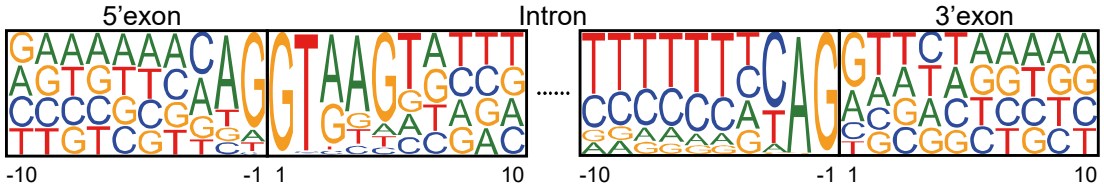


Figure S1. Human genomic sequence features in the surrounding area of splice sites, Related to Figure 1.

Human reference genome (hg38) and the annotation from GENCODE version 29 were used for analysis. The font size indicates the probability of bases at each position.

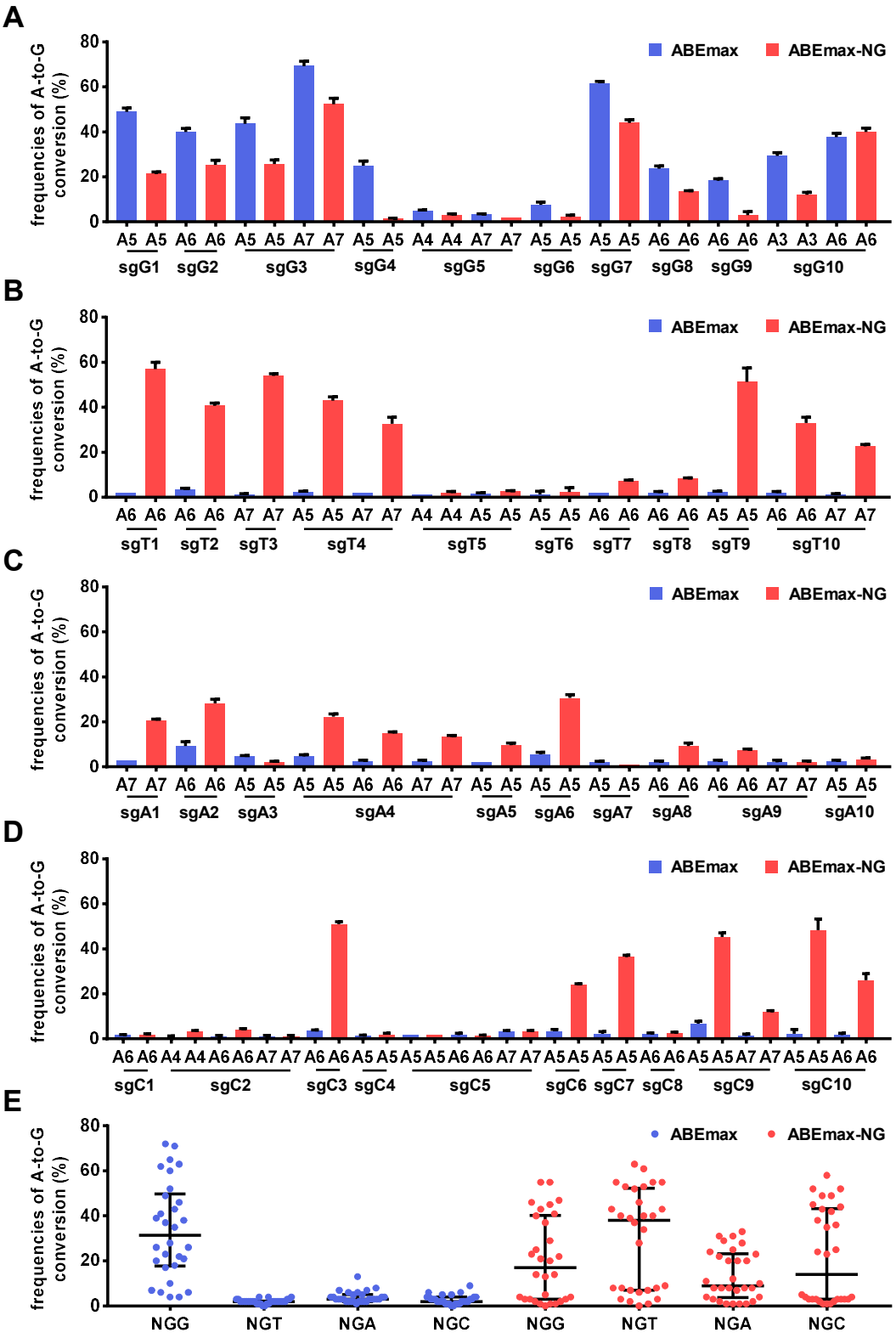


Figure S2. Analysis of ABEmax-NG-induced A-to-G substitution in mouse N2a cells, Related to Figure 3.

(A-D) A-to-G editing efficiency of ABEmax and ABEmax-NG at endogenous target sites with NGG PAM (A), NGT PAM (B), NGA PAM (C) and NGC PAM (D). Data are represented as the

mean ± s.e.m. (n = 3 from three independent experiments). The editing efficiency was calculated by EditR based on Sanger sequencing chromatograms.

(E) Statistical analysis of the A-to-G editing frequency at a total of 40 endogenous target sites in A-D. The median and interquartile range (IQR) are shown.

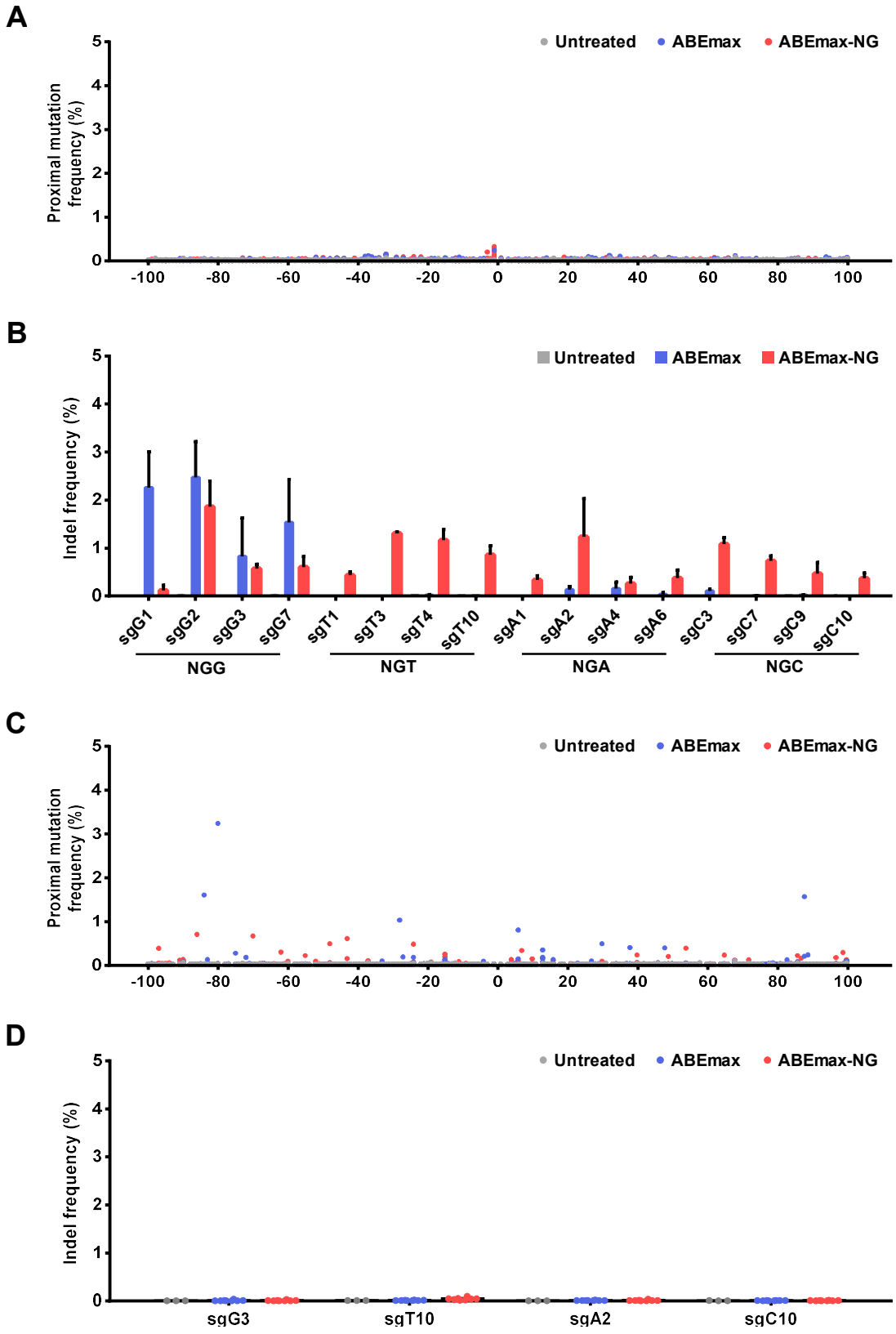


Figure S3. Analysis of proximal off-targets and indels for ABEmax and ABEmax-NG *in vitro* and *in vivo*, Related to Figure 3.

(A) Proximal off-targets of ABEmax and ABEmax-NG at 16 endogenous target sites in N2a cells were analyzed by deep sequencing. The mutation rates of A-to-G or T-to-C sites ±100 bp

surrounding the protospacer were calculated, designated the side of the protospacer distal to the PAM as negative positions, while the side that includes the PAM as positive numbers, counting against their positions relative to the protospacer.

(B) Analysis of indels at 16 endogenous target sites in N2a cells by deep sequencing. Reads containing at least 1 inserted or deleted nucleotides in the protospacer were calculated as indel-containing reads. Indel frequency was calculated as the number of indel-containing reads among the total number of mapped reads.

(C) Proximal off-targets of ABEmax and ABEmax-NG at 4 endogenous target sites in mouse embryos were analyzed by deep sequencing.

(D) Analysis of indels at 4 endogenous target sites in mouse embryos by deep sequencing.

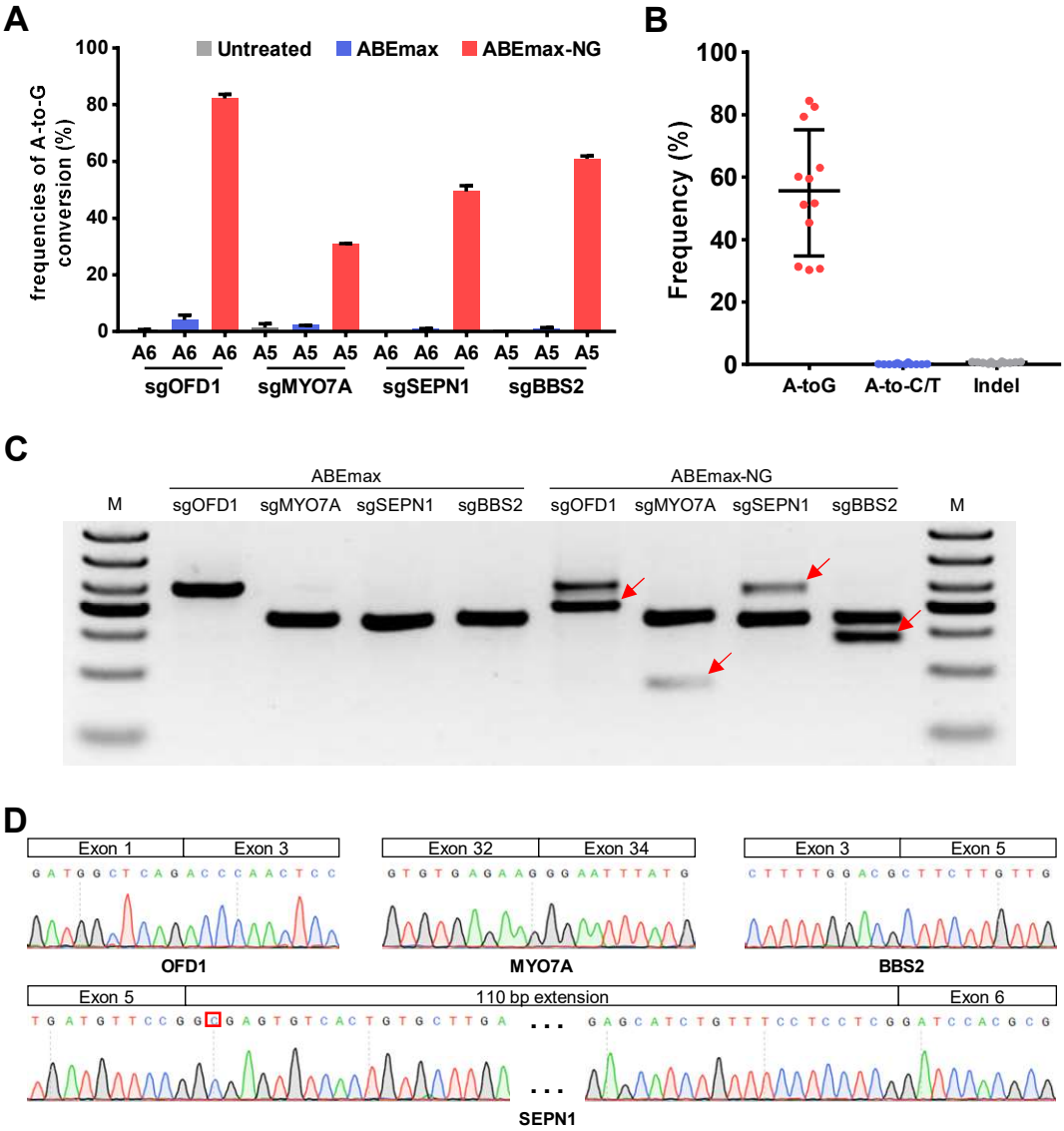


Figure S4. ABEmax-NG modulated endogenous RNA splicing in N2a cells, Related to Figure 4.

(A) A-to-G editing by ABEmax and ABEmax-NG at 4 splice sites in mouse N2a cells. Data are represented as the mean ± s.e.m. (n = 3 from three independent experiments).

(B) Statistical analysis of the A-to-G editing frequency, unwanted base conversions and indels induced by ABEmax-NG in (A). The median and interquartile range (IQR) are shown.

(C) Different RNA isoforms induced by ABEmax-NG were determined by RT-PCR. ABEmax served as the control. New splicing isoforms were highlighted by red arrowheads.

(D) Sanger sequencing chromatograms of the RT-PCR products confirmed the new splicing isoforms. Exon 2 of OFD1, exon 33 of MYO7A and exon 4 of BBS2 skipped in new splicing isoforms. The target splice site highlighted by red frame was converted from T to C in new splicing isoform of SEPN1 harboring 110 bp extension of exon 5.

A

Target gene	sgRNA	Editor	No. of examined embryos	Mutant ratio (%)		
				No. of targeted mutants[a]	No. of A-to-C/T[a]	No. of indels[a]
AKR1C19	sgG3	ABEmax	8	8(100)	0(0)	0(0)
EYA1	sgT10	ABEmax	8	0(0)	0(0)	0(0)
SIX6	sgA2	ABEmax	8	7(87)	0(0)	0(0)
BHLHA9	sgC10	ABEmax	8	0(0)	0(0)	0(0)
OFD1	sgOFD1	ABEmax	8	7(87)	0(0)	0(0)
MYO7A	sgMYO7A	ABEmax	8	5(63)	0(0)	0(0)
SEPN1	sgSEPN1	ABEmax	8	1(13)	0(0)	0(0)
BBS2	sgBBS2	ABEmax	8	4(50)	0(0)	0(0)
AKR1C19	sgG3	ABEmax-NG	8	8(100)	0(0)	0(0)
EYA1	sgT10	ABEmax-NG	8	8(100)	0(0)	0(0)
SIX6	sgA2	ABEmax-NG	8	8(100)	0(0)	0(0)
BHLHA9	sgC10	ABEmax-NG	8	8(100)	0(0)	0(0)
OFD1	sgOFD1	ABEmax-NG	8	8(100)	0(0)	0(0)
MYO7A	sgMYO7A	ABEmax-NG	8	8(100)	0(0)	0(0)
SEPN1	sgSEPN1	ABEmax-NG	8	8(100)	0(0)	0(0)
BBS2	sgBBS2	ABEmax-NG	8	8(100)	0(0)	1(13)

[a]Calculated from the number of examined embryos

B

Target gene	sgRNA	Editor	No. of transferred embryos	No. of offspring	Mutant ratio (%)			
					No. of targeted mutants[a]	No. of homozygous target mutants[a]	No. of A-to-C/T[a]	No. of indels[a]
BBS2	sgBBS2	ABEmax-NG	40	19	19(100)	0(0)	0(0)	0(0)

[a]Calculated from the number of offspring

Figure S5. Summary of the manipulation and genotyping of mouse embryos and newborn pups, Related to Figure 4.

(A) Summary of the manipulation and genotyping of mouse embryos.

(B) Summary of the manipulation and genotyping of newborn pups.

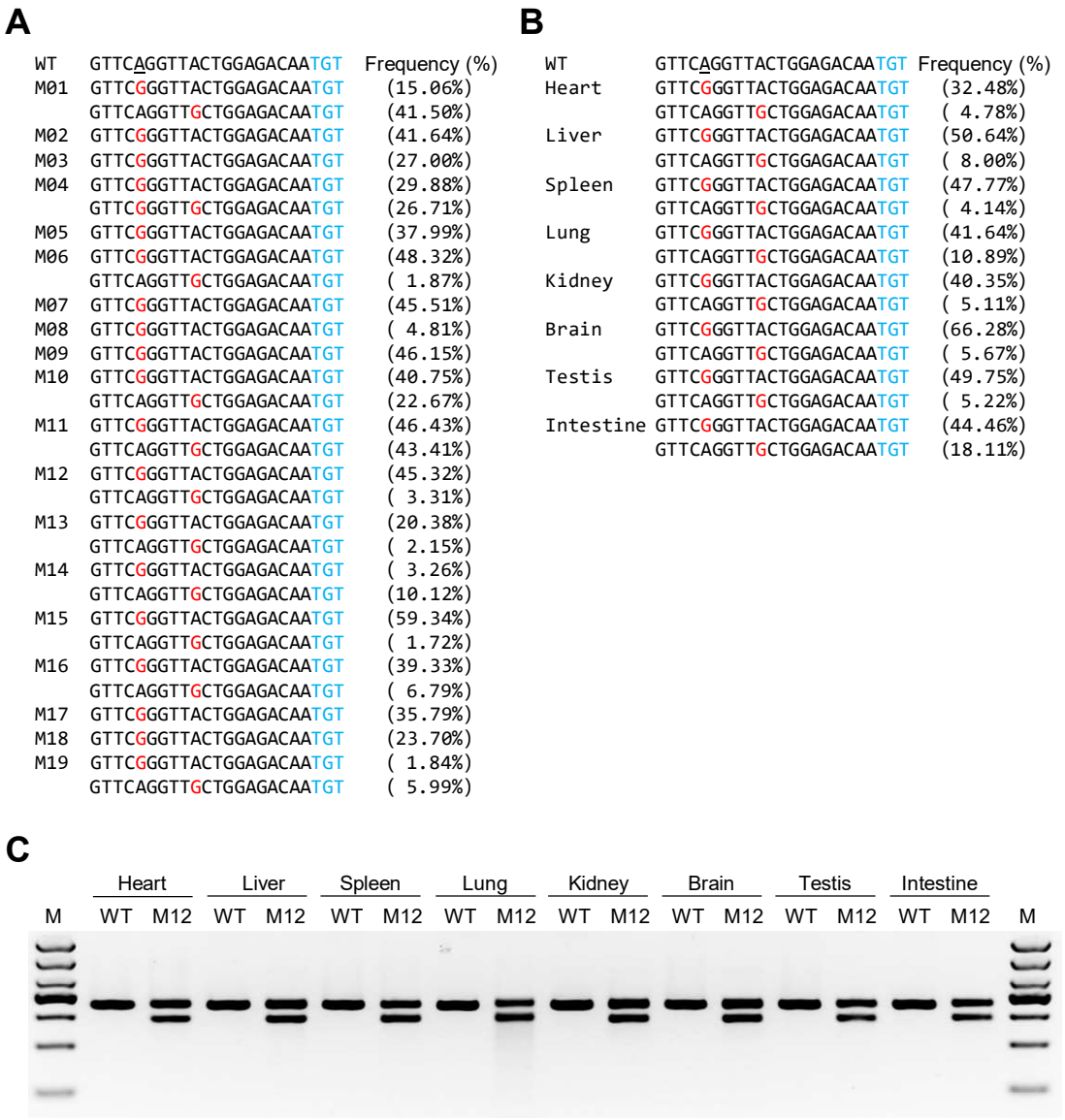


Figure S6. Verification of genotype and RNA splicing in founder mice, Related to Figure 4.

(A) Alignments of modified sequences from newborn pups after microinjection of ABEmax-NG mRNA and sgRNAs targeting at BBS2 splice site into one-cell embryos. The PAM sequences and substitutions are highlighted in blue and red, respectively. Frequency is calculated from deep sequencing.

(B) Alignments of modified sequences from different tissues of the founder mouse M12.

(C) Different RNA isoforms induced by ABEmax-NG were determined by RT-PCR from different tissues of founder mouse M12. Wild-type mouse served as the control.

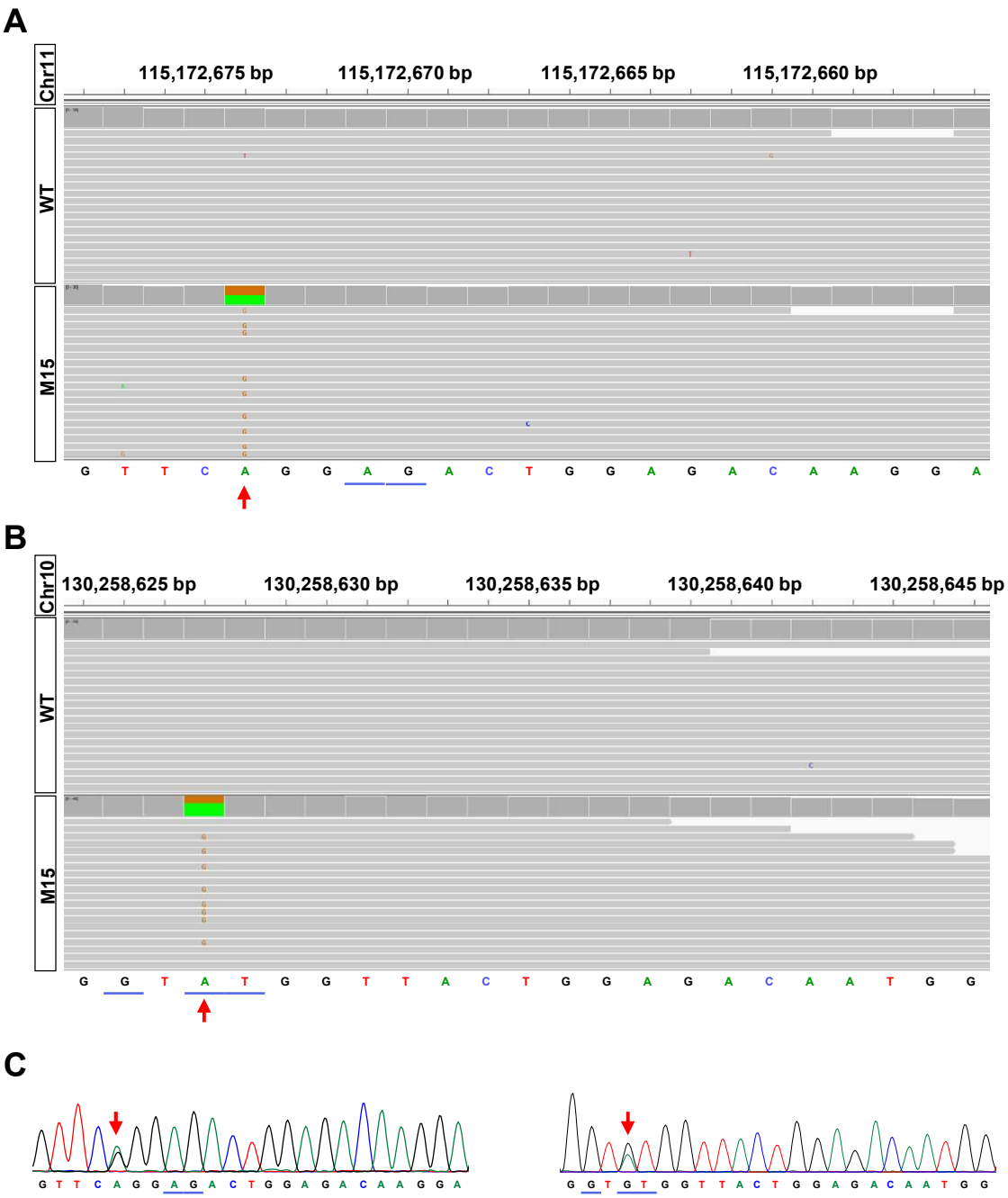


Figure S7. Verification of off-target sites, Related to Figure 5.

(A) Integrative Genomics Viewer (IGV) images showing the alignments of sequence reads at the off-target site NC_000077.6: g.115172674T>C.

(B) IGV images showing the alignments of sequence reads at the off-target site NC_000076.6: g.130258627A>G.

(C) Verification of off-target sites by Sanger sequencing chromatograms. Left: NC_000077.6: g.115172674T>C. Right: NC_000076.6: g.130258627A>G.

Blue lines represent the mismatches of sgRNA sequence and red arrows indicate the off-target sites.

Supplemental Tables

Table S1. sgRNAs used in this study, Related to Figure 2-4.

Site	Protospacer sequence	PAM	Target gene	Corresponding human genotype	Associated genetic disease
sgEGFP	AGCACTACACGCCGTAGGTC	AGN	sEGFP		
sgA1	CAATCCAGACACTGGTGGTC	AGA	CHRNE		
sgA2	CGGGCAGCGACCATAGGAAG	CGA	SIX6		
sgA3	GAGCACTCGTCGAGGTCTGC	AGA	FBN1		
sgA4	ATGGAAAGCAGACACGATAG	TGA	ITPR1		
sgA5	CAAGATGTATGGCGAGTATG	TGA	FGD1		
sgA6	GAACATGAACTCTTACGACT	CGA	TMEM67		
sgA7	TTCTATGAGCAGAAAATTAA	AGA	RNF216		
sgA8	ACCTCAGGTAATGTAGCATC	AGA	MLH1		
sgA9	GGATGAAACTATAGCGGGAT	CGA	NUP205		
sgA10	ATTCAGCTCCCGGAACATCT	CGA	TRP53		
sgC1	CTTCCAGGGGGAGCGAGGAA	AGC	COL6A1		
sgC2	TACACAACCTCACAGTCCTC	AGC	MKKS		
sgC3	CCTCTATTGTGCTGTCATGT	TGC	LMBR1		
sgC4	CGGGAGCCCGCTAGGTGGCC	AGC	MEGF8		
sgC5	TAAGAAAGTACCAAATCGAC	AGC	MTM1		
sgC6	TACCAGTCCCCTTCGCTCCC	TGC	CD207		
sgC7	CAGCAGCTCGTCCTTCACTG	CGC	NFIX		
sgC8	AGGTCAGCACTCTGACCACG	TGC	NBEAL2		
sgC9	TATTACAGAAACCAGCCCCG	AGC	DES		
sgC10	GGCTAACGTGCGGGAGCGCA	AGC	BHLHA9		
sgG1	ATTGATGTAATGGATGCAGT	GGG	NDUFS1		
sgG2	GTTTCAGAATCGAAGGGTGA	AGG	HOXD13		
sgG3	AGACATATTCCTCACTACAA	AGG	AKR1C19		
sgG4	CTTTAGCTTGACATGCAGCG	CGG	NIPBL		
sgG5	CCCACCAGCTCAAATGCAAT	GGG	SLC16A2		
sgG6	AGCCAGGTGGGCGGTTCTCT	TGG	FERMT1		
sgG7	GCGCATGGCCACTTCCTGTG	GGG	LMNA		
sgG8	AATTCACTGTAAAGCTGGAA	AGG	PTEN		
sgG9	CCCTCAGGGGTACTCTGACT	CGG	ZEB2		
sgG10	CTACTATGACCTCTATGGTG	GGG	PTPN11		
sgT1	CTTGTATCAGGACCACATGC	AGT	WNT5A		

sgT2	TCCGCAGCCGCCCCACAACC	AGT	WNT5A		
sgT3	AACGTGATGGCCATGTCGCC	TGT	SUFU		
sgT4	AGCCAGACTCTGCCGATGAC	AGT	GJA8		
sgT5	GAAAATGTTCTTGGCTGTTT	TGT	TYR		
sgT6	ATTTACCCCAAGGGAGGCCG	AGT	PLCB4		
sgT7	GGTTGATGAGCACACTGGCC	AGT	AHI1		
sgT8	GCGGGAGCGCCAGCGCACGC	AGT	TWIST1		
sgT9	CACCATGGCTCTACGGCGAC	AGT	CKAP2L		
sgT10	TTTGGAAGGAAAGTGGTATA	CGT	EYA1		
sgBBS2	GTTCAGGTTACTGGAGACAA	TGT	BBS2	NC_000016.10:g.56510923T>C	Bardet-Biedl syndrome 2
sgOFD1	CTGATACCTGAAGTGTGTCC	AGT	OFD1	NC_000023.11:g.13735348T>C	Oral-facial-digital syndrome
sgMYO7A	CCTCAGGAGGACGACCTGGC	TGA	MYO7A	NC_000011.10:g.77194352A>G	Deafness, autosomal recessive 2
sgSEPN1	CACTCACCGGAACATCACGG	TGT	SEPN1	NC_000001.11:g.25809152T>C	Eichsfeld type congenital muscular dystrophy

Table S2. Primers used for PCR amplification, genotyping and transcription, Related to Figure 2-4.

Primer name	Primer sequence
sgA1_F	AGAGCTTAGCCTGTATCACC
sgA1_R	TAGAACAATCTCTGGCAGCC
sgA10_F	CCTGTAAGTGGAGCCAGCTT
sgA10_R	AAGTCAGTTCTCGTAGGGTG
sgA2_F	GAACCCAACCCACAGCTCTT
sgA2_R	CAGCGGGAACTTCTTCCTTA
sgA3_F	CACATTGCAGCACTAGAAGC
sgA3_R	TACCATCATACAGCTCTGCC
sgA4_F	ATGTCCTGACAGATACAGGC
sgA4_R	TGGTCTGTGCTGATAGGTCA
sgA5_F	GAGTTAGGCTAGGGTTTCAC
sgA5_R	AGCCATGAGGCAGCTTTAAC
sgA6_F	AGACACACATGCAGGCAAAG
sgA6_R	AGCCAAGGAAGGTTCTGTCT
sgA7_F	GCTTGGGTTCCTAGTGAGTA
sgA7_R	GCGCATGTGGTACTAGTGAT
sgA8_F	CTTTATGGCCATGTGTGAGG
sgA8_R	GGATGCTGTTTCTTGACCCA
sgA9_F	TGGCAGTCGGATGTAAAGTG
sgA9_R	GGTGGTTCTCCTCAGTTCAT
sgBBS2_F	CCGAGGTTGCTCTTGTCTTC
sgBBS2_R	ACAGAGGCAAGGACCAGTGA
sgBBS2_RT_F	TGAACCCTGAGCTTGGCTAT
sgBBS2_RT_R	GTCTGCCACAATCTCATCTTC
sgC1_F	AGGCACAACTCTAAGCCCAA
sgC1_R	TCAGACTTGGTCAGCCTGAA
sgC10_F	TGGACCCAAATAGCTGAAGC
sgC10_R	CGGAACACTGAAGCTGGAAT
sgC2_F	CATACCTGGAGACCAGTCTT
sgC2_R	CTGAAGTCAACTGGGATTCG
sgC3_F	ATCCAGCCATCCTAGAGTGT
sgC3_R	CAATGAACGCTCATGGAGTC

sgC4_F	TCAGTACGGAGGTTCAGTGA
sgC4_R	GTCATAGGAGGTGGAGACAT
sgC5_F	CACTCACCAACCAACATGGT
sgC5_R	GGGGTAGTCTCAAGTGAGAT
sgC6_F	AGATGCTGCACCTGCAAATC
sgC6_R	GCCATATAAGACACGGAGGT
sgC7_F	GGATGAGTTCCACCCGTTTA
sgC7_R	CCGTGATGGTTAGCACAAAG
sgC8_F	ATTGCCGATCCAGCAGATGC
sgC8_R	CACCCACCAAGCCTATCGAA
sgC9_F	AAGACTGGTCCCTCTCTCTA
sgC9_R	TGTTGTTGCTGTGTAGCCTC
sgEGFP_F	TCGTTGACCGAATCACCGAC
sgEGFP_R	CCTTGAAGTCGATGCCCTTC
sgG1_F	CTGGTAGTAATCCCAACACG
sgG1_R	CCATAACACCATGGGACACA
sgG10_F	GCTGTGGGTTGCCATAGTTA
sgG10_R	TGGGCTTACAATACACTGCC
sgG2_F	GATGTGGCTTTAAACCAGCC
sgG2_R	CAATGCTTGCCTTTCTAGGC
sgG3_F	CCATTTATGCCACTCTCTTC
sgG3_R	TGTTTACCTGAACTCACTGC
sgG4_F	TGCAATTGCCGTTCGAACAA
sgG4_R	TGCCTCTTCAAAGCATAACC
sgG5_F	GCAGATAGTAGAAACGGAGA
sgG5_R	TGCCCATGTACTCTGTTTAG
sgG6_F	AATCGTGACACCTGAGCTAG
sgG6_R	GTAAGTGACAGGGGATGTAG
sgG7_F	AGACTCCAGCTTACAGAGCA
sgG7_R	AATCCAGAACCCTGTCCACT
sgG8_F	GAAGACCATAACCCACCACA
sgG8_R	CAGGGATGAGGGATACACTA
sgG9_F	AAATGTGGCAAGCGCTTCTC
sgG9_R	TCTCTTCCTCATCCCGTATC
sgMYO7A_F	CCCATGATTGCCTTGTGAAG

sgMYO7A_R	CTGTCAGGCAGAAGACATCA
sgMYO7A_RT_F	AAAGGAGGTCTTCACACCCT
sgMYO7A_RT_R	GGGCATAATTGACCACATCC
sgOFD1_F	TAAGCATCTTAGGGCTTCCG
sgOFD1_R	TCCTGCTCACTACATAGACG
sgOFD1_RT_F	AAAGCAGATGAGGATGGCTC
sgOFD1_RT_R	CCGCATCACAGCTCTCTTTA
sgSEPN1_F	CAGAGCTGACAAACCAGCTA
sgSEPN1_R	CATCCTGGTCACCAGCTAAT
sgSEPN1_RT_F	GCTATTTGTCCAACAACCGC
sgSEPN1_RT_R	CATACAGCCACTCCATGTCC
sgT1_F	TTCAAGCCCCTGAATGGCTG
sgT1_R	AGGCTGTAAAGCAGACAGCT
sgT10_F	CAGCCTACACACTGCTAATG
sgT10_R	GAGTCACCTGACAATGTCCT
sgT2_F	CCCAGCAAGATTTAGGCTTC
sgT2_R	TGCAGGTTGGGGATAAATGG
sgT3_F	GTCCCTGTTAGTGAGCAGGT
sgT3_R	TGAGGACAGCAGCACCCATA
sgT4_F	TAGTCGGCACAGATGAGGCA
sgT4_R	GACGAAGATGATCTGCAGCA
sgT5_F	ATGGGCTATGTACAAACTCC
sgT5_R	GAAGGATATCCTGGCAGGAA
sgT6_F	CCCGTGCTATCTACCTGCTT
sgT6_R	GTTATGGCTTCGTTTTGGGG
sgT7_F	CCCCAAAGTATGAGACCGAT
sgT7_R	TGGTGCATTCCCTCCATTTC
sgT8_F	AGATGATGCAGGACGTGTCC
sgT8_R	CCTGGTACAGGAAGTCGATG
sgT9_F	AACTCTCCAATCGCAGAGCC
sgT9_R	ATCAGTGAACCTGCAGAGAC
IVT-F	TCTCGCGCGTTTCGGTGATGACGG
IVT-R	AAAAAAAGCACCGACTCGGTGCCACTTTTTC

Table S3. Primers used for target deep sequencing, Related to Figure 2-4.

Primer name	Primer sequence
sgA1_HTS_F	CTCATCTCACTGGTAAGACC
sgA1_HTS_R	ATAGTCGTGCCAGTCCTAAG
sgA2_HTS_F	GTTCCAGCTGCCCATTTTGA
sgA2_HTS_R	TGGCGTGCGATTCCTTAGTA
sgA4_HTS_F	TGAAGGTCCTCTGCAGTGCT
sgA4_HTS_R	GCTCAAAGAAGCAGGTTCAT
sgA6_HTS_F	GCTAATGTGCGCTGCTTCCT
sgA6_HTS_R	CCAGAATGGAACAGAATGCACA
sgBBS2_HTS_F	GCCTTCCATTTTTGAGGCCA
sgBBS2_HTS_R	CCACACACCTCTGTCTTCCC
sgBBS2_RT_HTS_F	TCATCGGTGGAAACTGTGCT
sgBBS2_RT_HTS_R	CCAAACCGACTGCCATACAT
sgC10_HTS_F	CTCCTGCTCGGAGGCTGGCA
sgC10_HTS_R	GGATAGAGCAGTGATGCGGTGG
sgC3_HTS_F	GAGGTTAGCTCCTGTAACTTC
sgC3_HTS_R	GCGACATAAACCAGGACAAT
sgC7_HTS_F	GGATGAGTTCCACCCGTTTA
sgC7_HTS_R	CTTCTTGCCCGTGATGGTTA
sgC9_HTS_F	GACAGTTGTCATAAGAAAGGTG
sgC9_HTS_R	CACCCAGCTGGTGAAAGACA
sgG1_HTS_F	GAGGGTTTGAGGCCAACATG
sgG1_HTS_R	CCTCATTACCTCTCCAGTTCTTGTG
sgG2_HTS_F	CTTACACCAAACTGCAGCTC
sgG2_HTS_R	TAAACTGTCTGTGGCCAACC
sgG3_HTS_F	TATCCATTTAGCTCTAGAGG
sgG3_HTS_R	TGCAACTAAGAGTTCAACCT
sgG7_HTS_F	AGCATCAGAGGTTGGACAAG
sgG7_HTS_R	ATGGAGGAGCTCTTCTCCAT
sgT1_HTS_F	GCAGGTCTCTAGGTATGAAT
sgT1_HTS_R	CCTACCTATTTGCATCACCC
sgT10_HTS_F	TTTCCCTGGGTTGGATAGAG
sgT10_HTS_R	GGTAAGAAACTGCCATGGGT
sgT3_HTS_F	ATGTTGCTCTTTAGCCCCAG

sgT3_HTS_R	CTGCGTCTCACTTGTAACCA
sgT4_HTS_F	GGAAAGAGATCATCTCAGAG
sgT4_HTS_R	ATCGTAGCAGACATTCTCAC
sgT9_HTS_F	TTCCTGTTATCCCGCCCTCT
sgT9_HTS_R	TCCAGCGTCCTTAACCTAAC
sgOFD1_HTS_F	CAGGTCCACGTTGACACTAG
sgOFD1_HTS_R	TAAATGTGGAGCCACCTGCG
sgMYO7A_HTS_F	GACAGGGAGAACACAGAGTC
sgMYO7A_HTS_R	ATGTAAGTGGGCACGAGGCT
sgSEPN1_HTS_F	TCATCCATCGCCTGTTAAGC
sgSEPN1_HTS_R	CCTCTCACCAAATCTGTAGG

Transparent Methods

Genome-wide analysis. To identify editable splice sites, human reference genome (hg38) and the annotation from GENCODE version 29 were used. sgRNAs of validated CBE and ABE variants were then designed to target splice sites according to their distinct PAM specificities and corresponding editing windows. The sgRNAs with a single target site in their editing windows were considered as precise editing sgRNAs and used for further analysis. Pathogenic human splice sites were annotated by ClinVar database.

Animals. All the experiment protocols involving mice were approved by the Animal Care and Use Committee of the Institute of Neuroscience, Chinese Academy of Sciences, Shanghai, China. Mice were maintained in an Assessment and Accreditation of Laboratory Animal Care credited specific pathogen free facility under a 12 h dark-light cycle. B6D2F1 (C57BL/6 x DBA/2) and ICR mouse strains were used as embryo donors and foster mothers, respectively.

Plasmid construction. For construction of sgRNAs, oligos were synthesized, annealed and cloned into BsaI site of the sgRNA expression vector. Plasmids used include pGL3-U6-sgRNA-PGK-puromycin (Addgene, 51133), pGL3-U6-sgRNA-EGFP (Addgene, 107721), pUC57-sgRNA expression vector (Addgene, 51132), pGL3-U6-sgRNA-mCherry.

Cell culture and transfection. HEK293FT and Neuro-2a (N2a) cells were maintained in Dulbecco's Modified Eagle Medium (DMEM) (Hyclone), supplemented with 10% fetal bovine serum (FBS) (v/v) (Gemini) and 1% Penicillin Streptomycin (v/v) (Gibco). Cells were seeded on poly-D-lysine coated 12-well plates (JETBIOFIL) and transfected at approximately 70% confluence with ABEmax or ABEmax-NG expressing plasmid (1000 ng) and sgRNA-expressing plasmid (500 ng) using Lipofectamine 2000 (Life Technologies) according to the manufacturer's protocol. Puromycin was added 1 day post-transfection and was maintained in culture until untreated control cells were all died. For deep sequencing, GFP positive cells were harvested from fluorescence-activated cell sorting (FACS) 72 h after transfection. Cells were cultured at 37°C with 5% of CO_2.

Flow cytometry. Cells were harvested and subjected to flow cytometry 48 hours after transfection. sgRNAs were annealed and cloned into pGL3-U6-sgRNA-mCherry. GFP signal was detected with flow cytometry. ABEmax-NG/ABEmax, sgRNA and hPGK-sEGFP/hPGK-EGFP plasmids were transfected simultaneously. At total of 10,000 cell events were collected and analyzed using FlowJo

software.

Genomic DNA extraction and genotyping. Genomic DNA of cells was extracted using QuickExtract™ DNA Extraction Solution (Lucigen) according to manufacturer's protocols, genomic DNA of mouse was extracted by phenol-chloroform method, and genomic DNA of zygotes was amplified according to methods described below. The isolated DNA was PCR-amplified with Phanta® Max Super-Fidelity DNA Polymerase (Vazyme). Primers used were listed in Table S2.

Whole genome random amplification. After developing to blastocyst stage *in vitro*, single embryo was transferred to 200 μl tube containing 5 μl of an alkaline lysis solution (200 mM KOH/50 mM dithiothreitol). After an incubation of 10 min at 65°C, 5 μl of neutralization solution (900 mM Tris-HCl, pH 8.3/300 mM KCl/200 mM HCl) was added. The lysed and neutralized sample was added with 5 μl of a 400 μM solution of random primers (Genscript, Nanjing, China), 6 μl of 10 x PCR buffer (Takara, Dalian, China), 3 μl of a mixture of the 4 dNTPs (each at 2.5 mM), and 1 μl of Taq polymerase (Takara, Dalian, China), and brought to 60 μl with water. Fifty primer-extension cycles were carried out in a MyCycler thermo-cycler (Bio-Rad, US). Each cycle consisted of a 1 min denaturation step at 92℃, a 2 min annealing step at 37℃, a programmed ramping step of 10 sec/degree to 55℃, and a 4 min incubation at 55℃ for polymerase extension. Then the products were used as the PCR templates.

***In vitro* transcription.** In brief, ABEmax/ABEmax-NG vector was linearized by BbsI enzyme (NEB) and *in vitro* transcribed using T7 Ultra Kit (Ambion) according to the manufacturer's protocols. mRNA was purified by Mini Kit (Qiagen). sgRNA oligos were annealed into pUC57-sgRNA expression vectors with T7 promoter. Then sgRNAs were amplified and transcribed *in vitro* by MEGAshortscript Kit (Ambion). The sgRNAs were purified by MEGAclear Kit (Ambion) according to the manufacturer's protocols. Primers used for transcription *in vitro* were listed in Table S2.

Microinjection of mouse zygotes and embryo transfer. B6D2F1 female mice at 4 weeks of age were superovulated and mated with B6D2F1 male mice. Fertilized one-cell embryos were collected from the oviducts. For microinjection, mRNA mixtures containing sgRNA (50 ng/μl) and ABEmax-NG/ABEmax (100 ng/μl) were injected into the cytoplasm of zygotes in a droplet of M2 medium containing 5 μg/ml cytochalasin B (CB) using a piezo (Primetech) microinjector. The injected

zygotes were cultured in KSOM mediums at 37°C under 5% of CO_2 in air and transferred to oviducts of pseudopregnant females at 0.5 dpc.

RNA analysis. RNA was immediately extracted from cultured cells or mouse tissues by using TRIzol reagent (Invitrogen), according to the manufacturer's instructions. Complementary DNA (cDNA) was generated using the HiScript II Q RT SuperMix (Vazyme), and was PCR-amplified with Phanta® Max Super-Fidelity DNA Polymerase (Vazyme). The PCR-amplified fragments were separated by agarose gel electrophoresis. Products with abnormally size were purified from the gel and sequenced using the amplification primers.

Targeted deep sequencing. Target sites were amplified from extracted genomic DNA using Phanta® Max Super-Fidelity DNA Polymerase (Vazyme). PCR products with different barcodes were pooled together for deep sequencing on Illumina HiSeq X Ten (2 × 150 PE) at the Novogene Bioinformatics Institute, Beijing, China. Primers used for deep sequencing were listed in Table S3. The adapter pair of the pair-end reads were removed using AdapterRemoval version 2.2.2, and pair-end read alignments of 11 bp or more bases were combined into a single consensus read. All processed reads were then mapped to the target sequences using the BWA-MEM algorithm (BWA v0.7.16). For each site, the mutation rate was calculated using bam-readcount with parameters -q 20 -b 30. Indels were calculated based on reads containing at least 1 inserted or deleted nucleotide in protospacer. Indel frequency was calculated as the number of indel-containing reads/total mapped reads.

Whole genome sequencing. Whole genome sequencing of mouse genomic DNA extracted from the tail was sequenced using an Illumina HiSeq X Ten (2 × 150 PE) at the Novogene Bioinformatics Institute, Beijing, China. The WT control mouse has the same genetic background with the Founder mouse and from our previous study (SRR8263608). All cleaned reads were mapped to the mouse reference genome (GRCm38/mm10) using BWA v0.7.16 with default parameters. Sequence reads were removed for duplicates using Sambamba v0.6.7 and realigned using Genome Analysis Toolkit (GATK v3.7) IndelRealigner. Variants were identified by GATK HaplotypeCaller and the following criteria were applied to all SNPs: (1) sequencing depth (for each individual) > 1/3× and < 3×; (2) variant confidence/quality by depth > 2; (3) RMS mapping quality (MQ) > 40.0; (4) Phred-scaled P value using Fisher's exact test to detect strand bias < 60; (5) Z-score from the Wilcoxon rank sum test of Alt vs. Ref read MQs (MQRankSum) > −12.5; and (6) Z-score from the Wilcoxon rank sum

test of Alt vs. Ref read position bias (ReadPosRankSum) > −8. After filtering out variants in the SNP database (dbSNP) and also found in the wild-type genome, potential off-target sites were predicted by CasOT-1.0 considering up to 2-bp mismatch in seed region and 5-bp mismatch in non-seed region with NG PAM.

Data and Software Availability. High-throughput sequencing data will be deposited in the NCBI Sequence Read Archive database under accession code (PRJNA527206). All other data are available upon reasonable request.