

HMGN1 enhances CRISPR-directed dual-function A-to-G and C-to-G base editing

Received: 5 October 2022

Accepted: 20 April 2023

Published online: 27 April 2023

 Check for updatesChao Yang^{1,2}, Zhenzhen Ma³, Keshan Wang⁴, Xingxiao Dong⁵, Meiyu Huang⁶,
Yaqiu Li^{1,2}, Xiagu Zhu⁷, Ju Li⁸, Zhihui Cheng³, Changhao Bi^{1,2} ✉ &
Xueli Zhang^{1,2} ✉

C-to-G base editors have been successfully constructed recently, but limited work has been done on concurrent C-to-G and A-to-G base editing. In addition, there is also limited data on how chromatin-associated factors affect the base editing. Here, we test a series of chromatin-associated factors, and chromosomal protein HMGN1 was found to enhance the efficiency of both C-to-G and A-to-G base editing. By fusing HMGN1, GBE and ABE to Cas9, we develop a CRISPR-based dual-function A-to-G and C-to-G base editor (GGBE) which is capable of converting simultaneous A and C to G conversion with substantial editing efficiency. Accordingly, the HMGN1 role shown in this work and the resulting GGBE tool further broaden the genome manipulation capacity of CRISPR-directed base editors.

CRISPR-directed base editors (BEs) including a deoxynucleotide deaminase and a catalytically impaired Cas9 or Cpf1 can introduce single-base conversion. The recently developed CBE¹, ABE², and GBE/CGBE^{3–5} enable programmable C-to-T, A-to-G, and C-to-G base editing, respectively. These technologies could edit the genome sequence without inducing a DNA double-strand break or requiring donor DNA templates, with exciting prospects in the genetic therapy of mutational-associated diseases. It was reported that ABE could directly correct the pathogenic single-base mutation of nuclear lamin A in cultured fibroblasts from children with progeria⁶. And CBE was successfully applied to edit BCL11A enhancer, which prevented sickle cell phenotype and attenuated the imbalanced globin chain in erythrocytes⁷. To broaden the application of this technology, we and other labs have been working on increasing the efficiency and diversifying the editing patterns of BEs. Richter et al.⁸ greatly improved the editing efficiency of ABE through phage-assisted evolution. Koblan as well as our lab constructed and further enhanced the editing efficiency of C-to-G transition with DNA repair factors⁹ or DNA binding proteins¹⁰. However, the genome editing capacity still has room for improvement.

One base editor normally catalyzes the transition for a single type of base pair, and it is not possible to simultaneously implement multiple editors in most scenarios. However, studies reported the discovery of MNVs (multi-nucleotide variants), which were recognized as two or more nearby variants existing on the same haplotype in an individual. Significantly, the MNVs were identified as a clinically and biologically important class of genetic variation, influencing the functional interpretation of genomic data^{11,12}. To extend the editing possibilities and study MNVs-associated genetic diseases, researchers have developed dual-function base editors, which enable simultaneous C-to-T and A-to-G conversion in mammalian cells^{13–16}. In addition, an AGBE system was constructed using human APOBEC3A and TadA, which could catalyze four types of base conversions¹⁷. These reconstructed BEs broadened the capability of base editing for applications in genetic therapy or gene regulation. However, a series of MNVs with simultaneous C-to-G and A-to-G mutations remain poorly understood, and related genetic correction studies are hindered by the lack of genome editing tools. For instance, it was reported that the concurrent C-to-G and A-to-G mutation in GAA (alpha-glucosidase) resulted in glycogen storage disease type II¹⁸. Additionally, the high efficiency and

¹Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin, China. ²Key Laboratory of Systems Microbial Biotechnology, Chinese Academy of Sciences, Tianjin, China. ³College of Life Sciences, Nankai University, Tianjin, China. ⁴Department of Urology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China. ⁵School of Biological Engineering, Dalian Polytechnic University, Dalian, China. ⁶College of Life Sciences, Guangxi Normal University, Guilin, China. ⁷College of Biotechnology, Tianjin University of Science and Technology, Tianjin, China. ⁸College of Life Sciences, Tianjin Normal University, Tianjin, China. ✉e-mail: bi_ch@tib.cas.cn; zhang_xl@tib.cas.cn

specificity of base editors are indispensable for the construction of disease models and the development of genetic therapies for mutational disease. However, the recently developed ABE showed low efficiency and specificity in performing concurrent C-to-G and A-to-G¹⁷ conversions. Despite the prime editor that could introduce the concurrent C-to-G and A-to-G transition, normally the base editing could induce more efficient editing in a restricted editing range^{4,19}. Furthermore, the operation of the base editor is more convenient than that of the prime editor. Thus, it is desirable further expand the diversity and capacity of BEs with a highly specific and efficient dual-function base editor, which simultaneously catalyzes C-to-G and A-to-G conversions.

It was reported that Cas9 nuclease activity was positively associated with elevated chromatin accessibility. Ding et al.²⁰ demonstrated that Cas9-dependent deletion or insertion was improved via fusion with chromatin-modulating peptides, particularly at refractory target sites. Furthermore, Liu et al.²¹ found that transactivation modules could enhance Cas9-dependent editing efficiency, especially at nuclease-refractory target sites. Both of their functions are demonstrated to increase chromatin accessibility, which might promote Cas9 activity. More importantly, the editing efficiency of ABE and CBE was also enhanced with the inhibition of histone deacetylase to alter the chromatin state²². Collectively, these works indicate that genome editing is affected by the chromatin microenvironment. Chromatin remodelers and modifiers were reported to orchestrate the chromatin state, and thus influence the DNA repair process. It was demonstrated that BARD1 (BRCA1-associated RING domain protein 1) coordinates the binding of H2AK15ub and unmethylated H4K20 to promote homologous recombination²³, as well as corroborated the dimethylation of histone H3K36 and enhanced the repair of DNA double-strand breaks²⁴. Moreover, LSD1 (lysine-specific histone demethylase 1)-mediated histone demethylation was reported to be involved in base excision repair induced by hydrogen peroxide²⁵. Nevertheless, there were still only limited reports on how chromatin-associated factors influence base editing. Accordingly, we hypothesized that the fusion of chromatin-associated factors with BEs might increase the efficiency of base editing, which could be a promising strategy to further optimize BEs.

In this research, we engineer GBE and ABE by fusing them with a series of chromatin-associated factors, which optimizes the editing efficiency or indel frequency. Furthermore, we fuse HMGNI with GBE, ABE, and GGBE for higher editing efficiency, lower indel byproducts or a unique editing pattern. Finally, we also demonstrated the application of GGBE in correcting or creating MNVs in mammalian cells.

Results

Analysis of chromatin-associated factors for improving GBE and ABE editing

To further improve the efficiency and expand the editing scope of base editors, we intended to optimize them by integrating chromatin-associated factors and further construct a specific dual-function C-to-G and A-to-G base editor (GGBE) (Fig. 1a). Accordingly, the nascent GBE⁵ was selected for investigation of C-to-G editing, while the latest highly efficient A8e (ABE8e-V106W)⁸ with reduced off-target editing was selected to explore A-to-G editing. Then we constructed a series of engineered GBE and A8e variants fused with 16 chromatin-associated factors, including chromatin-modulating factors, histone methylation factors, histone acetylation factors, and histone ubiquitination factors. Since previous work demonstrated the GBE variants with amino-terminal fusions exhibited the highest editing efficiency^{26,27}, this fusion pattern was also employed for the selected chromatin-associated factors (Fig. S1A). HEK293T cells were transiently co-transfected with the reconstructed GBE/A8e and gRNA vectors, and the editing efficiency within the editing window as well as the indel frequency were determined via high-throughput sequencing (HTS) and CRISPResso2²⁸.

Our data revealed that several chromatin-associated factors improved the C-to-G editing efficiency at the HIRA and EMX1 loci, especially the chromatin-modulating factors and histone ubiquitination factors. Notably, the HMGNI-fused GBE (HMGNI-GBE) showed the highest editing efficiency at position C6 of the protospacer, which was also superior to SadN-GBE, an enhanced GBE variant (Fig. 1b). The HMGNI-fused A8e (HMGNI-A8e) also exhibited the highest editing efficiency, but was similar or slightly higher to the control at the HEK4 and EMX1 loci (Fig. 1c). Importantly, the indel frequency was higher for HMGNI-GBE and lower for HMGNI-A8e at the tested loci (Fig. S1B). To further query the potentially adverse effects of HMGNI fusion, the protein expression of base editors, gene expression of targeted loci, and cell viability were evaluated. The results indicated that the over-expression of HMGNI did not significantly influence the expression of the base editors, gene expression of the targeted loci, or cell viability (Fig. S1C–E).

To further analyze the role of HMGNI in optimizing GBE and A8e, we tested the HMGNI-GBE and HMGNI-A8e at more genomic loci in HEK293T cells. The data showed that HMGNI-GBE substantially increased the editing efficiency at eight genomic loci, especially at position C6 of the protospacer (Fig. 1d). Surprisingly, the increased indel frequency was not observed at all testing loci, and even decreased at the TET2-site1 (Fig. S1F). Our results also indicated that the HMGNI-A8e showed a modestly higher editing efficiency at most of the testing loci, especially at TET2-site4, with average increases of up to 37.40% (Fig. 1e). Furthermore, the indel frequency of HMGNI-A8e slightly decreased at several loci (Fig. S1G). More importantly, the HMGNI fusions with GBE and A8e also enhanced the editing yield in HeLa cells (Fig. S1H) and exhibited substantial C-to-G and A-to-G transition in primary prostate carcinoma cells (Fig. S1I).

To further support the positive effects of HMGNI and construct enhanced GBE and ABE variants, various components with diverse arrangements were tested with HMGNI. Firstly, the HMGNI was fused to highly efficient miniCGBE variant⁴, which incorporates the R33A mutation and deletes the UNG component of GBE. The data showed the HMGNI-miniCGBE enhanced the editing yield but also slightly increased the indel products at testing loci (Fig. S2A–C). We also attempted to replace the UNG component in the GBE to increase the editing yield. The Ung1 from *Saccharomyces cerevisiae*¹⁰ and the Udgx from *Mycobacterium smegmatis*^{29,30} were placed between APOBEC1 and Cas9 in the GBE system based on previously optimized arrangements⁹ (Fig. S2D). The Udgx fusion showed a higher editing frequency at several genomic targets, especially in the PAM-proximal region (Fig. S2E). Finally, given that the pioneer factors were found to promote the chromatin accessibility and thereby enabled PAM-proximal base editing for CBE and GBE in previous research²⁷, we assumed that since HMGNI also increases chromatin accessibility^{31,32}, it might similarly enable the PAM-proximal editing of A-to-G. Thus, we constructed an A8e variant where HMGNI was placed between TadA and Cas9 (HMGNI-A8e-M; Fig. S2F) similar to the pioneer factors²⁷. Our data showed that this reported A8e variant also exhibited a higher A-to-G yield of PAM-proximal adenines (Fig. S2G). Taken together, our data proved that HMGNI efficiently enhanced the C-to-G conversion and modestly improved the efficiency of A-to-G editing.

Construction of simultaneous C•G-to-G•C and A•T-to-G•C base editors

To extend the editing capacity for more applications of BEs, we speculate that the fusion of ABE and GBE could introduce a dual-functional base conversion. To test this, the GBE components including APOBEC1 and UNG were fused to A8e to construct a TABE-UNG (TadA8e-APOBEC1-Cas9-UNG; Fig. 2a). In addition, given that the UNG component might negatively affect the C-to-G transition⁹, we also constructed a TABE system by omitting the UNG component (Fig. 2a). Then the TABE-UNG and TABE were tested at two genomic loci in

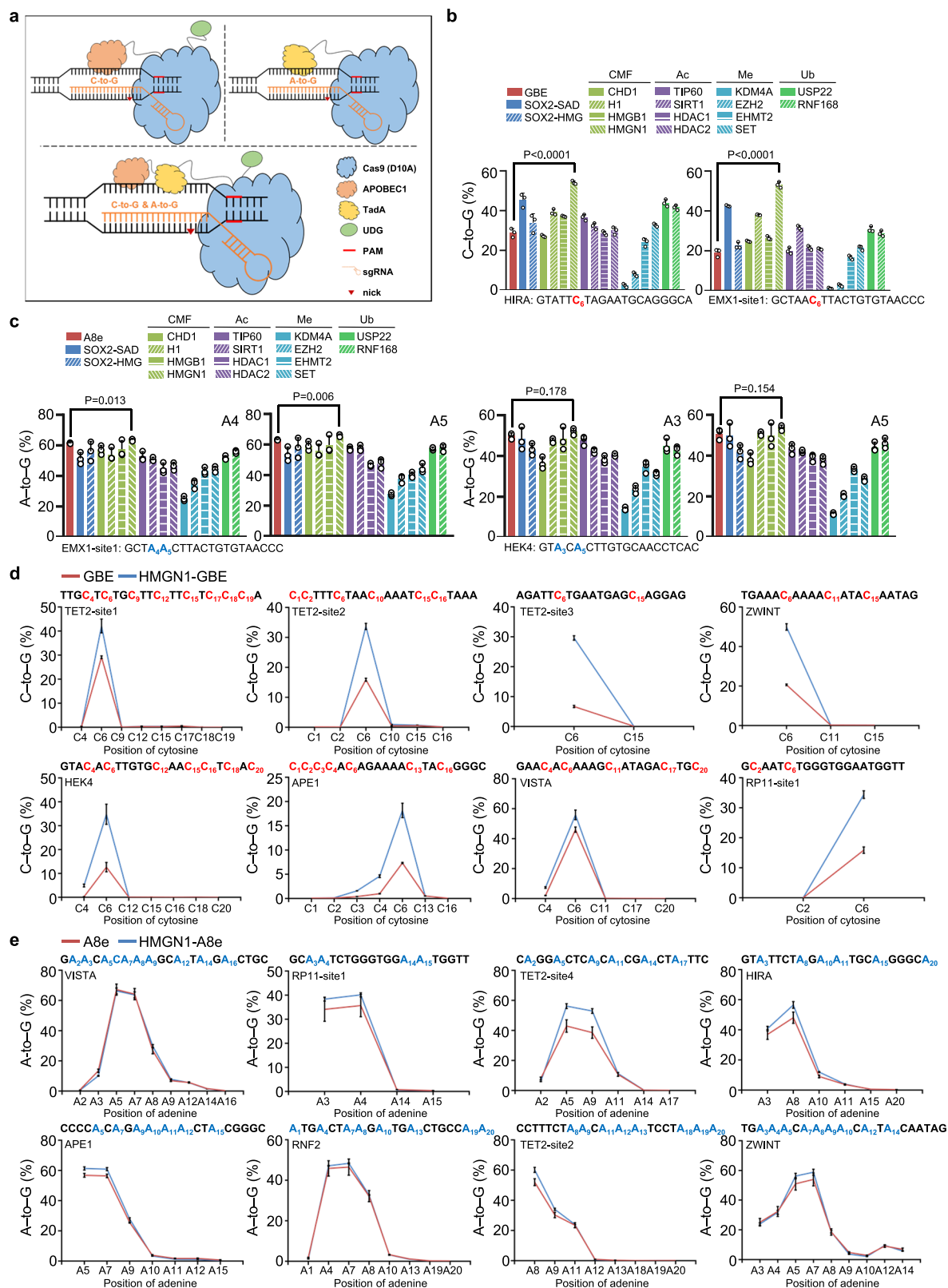


Fig. 1 | Analysis of chromatin-associated factors for improving GBE and ABE editing. **a** Schematic of GBE, ABE, and GGBE binding to DNA. **b** Base editing efficiency of C-to-G across GBE and GBE variants fused with 16 chromatin-associated factors at HIRA and EMX1-site1 loci in HEK293T cells. **c** Base editing efficiency of A-to-G across A8e and A8e variants fused with 16 chromatin-associated factors at HIRA and EMX1-site1 loci in HEK293T cells. **d** Base editing efficiency of C-to-G between GBE and HMGN1-GBE at eight genomic loci in HEK293T cells. **e** Base editing efficiency of A-to-G between A8e and HMGN1-A8e at eight genomic loci in HEK293T cells.

editing efficiency of A-to-G between A8e and HMGN1-A8e at eight genomic loci in HEK293T cells. CMF chromatin-associated factors, Ac acetylation, Me methylation, Ub ubiquitination. Mean \pm SEM (**b-e**) of all individual values of sets of $n = 3$ independent replicates are shown. All statistical analysis for samples were conducted using unpaired Student's t test (two-tailed) in GraphPad Prism 8. Source data are provided as a Source Data file.

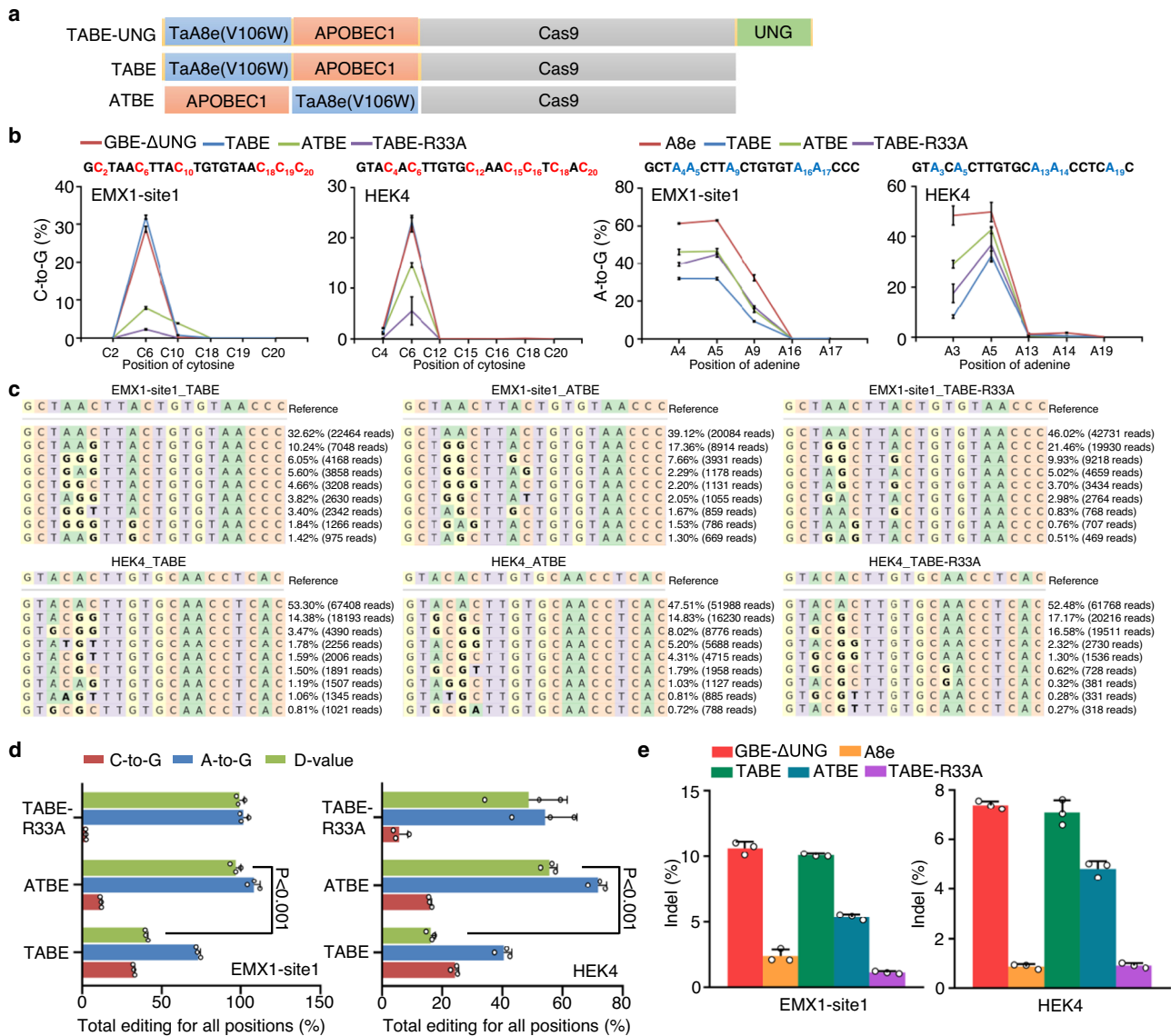


Fig. 2 | Construction of simultaneous C-to-G and A-to-G base editors. **a** Schematic of base editor variant with a fusion of APOBEC1, TadA8e, Cas9, and UNG. **b** Base editing efficiency of C-to-G (left) and A-to-G (right) across GBE-ΔUNG, A8e, Tabe, ATBE and Tabe (R33A) at EMX1-site1 and HEK4 in HEK293T cells. **c** The genotypes and reads and editing frequency of indicated genotypes across Tabe, ATBE, and Tabe (R33A) at EMX1-site1 (upper) and HEK4 (lower). **d** Comparison of

editing frequency and D-value across Tabe, ATBE, and Tabe (R33A). Total editing is the sum of editing for all positions. **e** Comparison of indel frequency across GBE-ΔUNG, A8e, Tabe, ATBE and Tabe (R33A) at EMX1-site1 (left) and HEK4 (right) in HEK293T cells. Mean ± SEM (**b**, **d**, **e**) of all individual values of sets of $n = 3$ independent replicates are shown. Source data are provided as a Source Data file.

HEK293T cells. Although our data showed that Tabe-UNG could simultaneously convert A-to-G and C-to-G, the editing efficiency was significantly lower than that of Tabe (Fig. S3A).

To potentially improve the base editing efficiency, a ATBE (APOBEC1-TadA8e-Cas9) system was established (Fig. 2a), and the APOBEC1-R33A mutant^{4,33} was also incorporated to improve the editing efficiency. To evaluate the balance of editing efficiency between A-to-G and C-to-G transition for dual functional base editors, we introduced a D-value, which is a measure of the difference in editing efficiency between A-to-G and C-to-G transition in our analysis. Despite ATBE induced a higher A-to-G conversion than Tabe (Fig. 2b, c and S3B), the D-value was significantly increased, indicating an imbalanced A-to-G and C-to-G efficiency (Fig. 2d). In addition, the APOBEC1-R33A mutant with higher C-to-G editing was introduced in Tabe, but we observed a relatively lower editing efficiency of C-to-G at two genomic loci for the Tabe-R33A (Fig. 2b, c and S3B). Further, sequencing reads indicated that the Tabe also exhibited the highest percentage of

concurrent C-to-G and A-to-G editing (Fig. S3C). Importantly, the indel frequency of Tabe system retained a similar level to GBE (Fig. 2e). Thus, we chose the Tabe system as the dual-function C-to-G and A-to-G base editor (GGBE1.0) for further investigation.

Next, we intended to optimize the GGBE by incorporating the chromosomal protein HMGNI and other UNG proteins. HMGNI and Udgx were integrated into the GGBE in different arrangements (GGBE1.1-5; Fig. 3a), and the resulting constructs were tested at two genomic loci in HEK293T cells. Our data indicated that the N-terminal fusion of HMGNI (GGBE1.1) did not significantly improve the concurrent C-to-G and A-to-G base editing (Fig. 3b-d and S3D, E) or the D-value (Fig. 3e). With the addition of Udgx between APOBEC1 and Cas9 (GGBE1.5), both the D-value and base editing efficiency was decreased (Fig. 3b-e and S3D, E). Notably, we observed a higher PAM-proximal C-to-G editing with HMGNI fusion between APOBEC1 and Cas9 (GGBE1.3), indicating an improvement of the editing window (Fig. 3b-d and S3D). Additionally, the indels of GGBE1.3 significantly

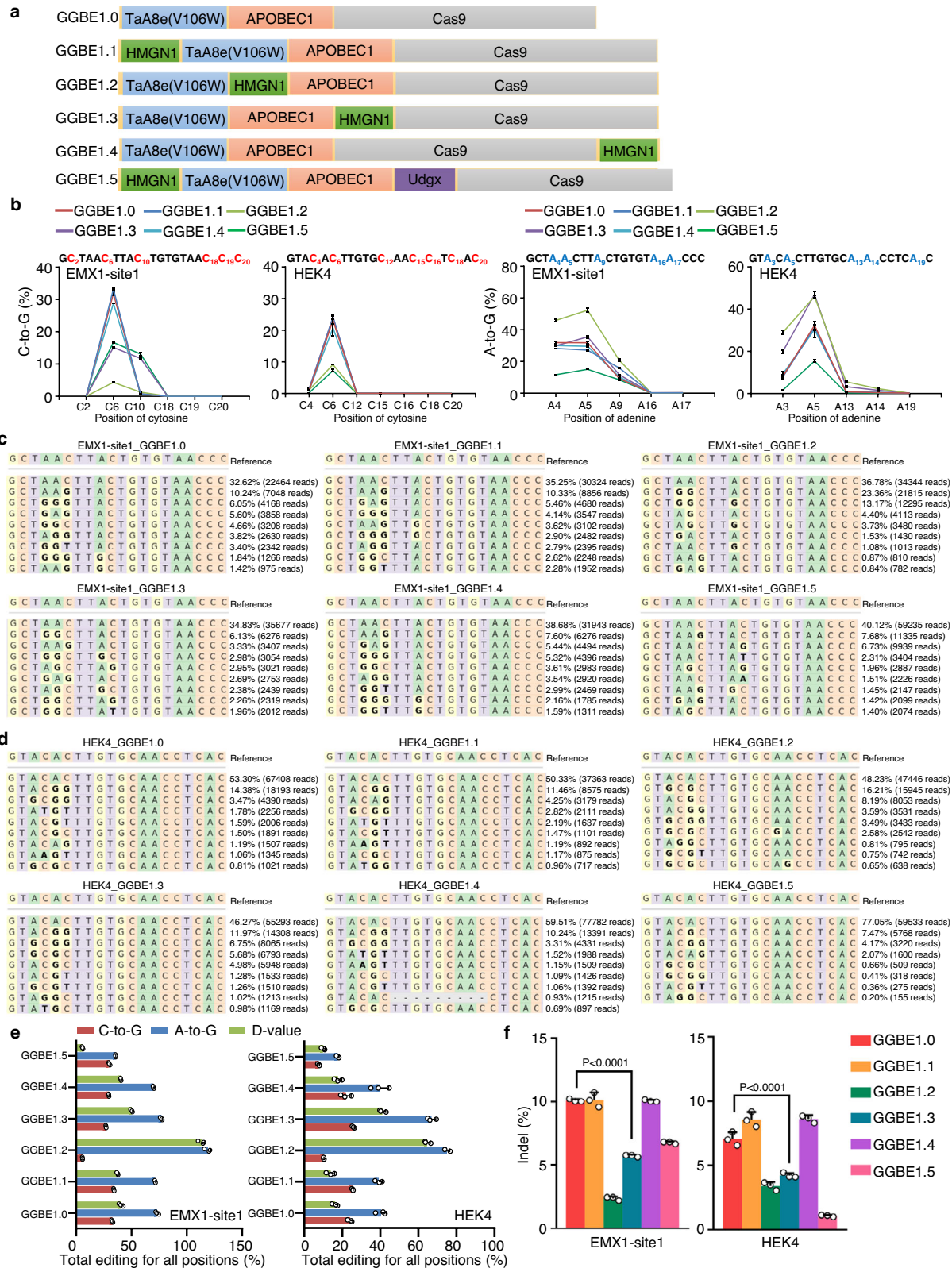


Fig. 3 | Optimization of GGBE in diverse arrangements with HMGNI. a Schematic of GGBE variants with fusion of HMGNI. **b** Base editing efficiency of C-to-G (left) and A-to-G (right) across GGBE1.0-GGBE1.5 at EMX1-site1 and HEK4 in HEK293T cells. **c** The genotypes and reads and editing frequency of indicated genotypes across GGBE1.0-GGBE1.5 at EMX1-site1. **d** The genotypes and reads and editing frequency of indicated genotypes across GGBE1.0-GGBE1.5 at HEK4. **e** Comparison of editing

frequency and D-value across GGBE1.0-GGBE1.5. Total editing is the sum of editing for all positions. **f** Comparison of indel frequency across GGBE1.0-GGBE1.5 at EMX1-site1 (left) and HEK4 (right) in HEK293T cells. Mean ± SEM (**b**, **e**, **f**) of all individual values of sets of $n = 3$ independent replicates are shown. All statistical analysis for samples were conducted using unpaired Student's t test (two-tailed) in GraphPad Prism 8. Source data are provided as a Source Data file.

decreased compared to GGBE1.0 (Fig. 3f). Considering that the GBE could also induce C-to-A/T conversion, the editing purity of cytosine was also evaluated. The data indicated that the proportion of C-to-G transition was lower in GGBE1.0 than in the original GBE at C6 of protospacer, whereas the GGBE1.3 induced a significantly higher C-to-G proportion for PAM-proximal cytosines at EMX1-site1 (Fig. S3F). Accordingly, GGBE1.0 and GGBE1.3 were selected for further verification of specific concurrent C-to-G and A-to-G base editing in subsequent analysis. Collectively, our data demonstrated that the GGBE could induce a concurrent C-to-G and A-to-G base editing, and the fusion of HMGNI could expand the editing window of GGBE.

Characterization and comparison of GGBE variants in mammalian cells

To further characterize the base editing features of GGBE and its variants, we then tested the GGBE1.0 and GGBE1.3 with another 12 genomic loci in HEK293T cells. We observed that A-to-G and C-to-G editing by GGBE1.0 and GGBE1.3 were mostly induced at positions 4–7 and 5–7 of the protospacer, respectively (Fig. 4a, b and S4A). The average editing efficiency by GGBE1.0 within the editing window were 16.52% for A-to-G conversion and 19.47% for C-to-G conversion, respectively. GGBE1.3 exhibited a moderately increased editing efficiency for both C-to-G and A-to-G transition at most tested genomic loci, especially for PAM-proximal base editing. The average editing efficiency of GGBE1.3 within the editing window were 25.94% and 21.82% for A-to-G and C-to-G conversion, respectively. For most tested gRNAs, the indel frequency of GGBE1.3 was modestly lower than that of GGBE1.0 (Fig. 4c). The percentage of concurrent C-to-G and A-to-G editing between GGBE1.0 and GGBE1.3 varied across these genomic loci, of which GGBE1.3 exhibited similar or higher simultaneous editing compared to GGBE1.0 (Fig. S4B). In addition, a higher proportion of C-to-G was also observed at several PAM-proximal cytosines in GGBE1.3, such as the PPP1R12C-site3-C9 and TET2-site2-C10 (Fig. S4C). Finally, the GGBE1.0 or GGBE1.3 were also tested in HeLa and primary prostate carcinoma cells with simultaneous C-to-G and A-to-G transition (Fig. S4D, E).

To further characterize the editing efficiency and specificity of GGBE, we compared it with the AGBE, a recently developed tool that was also reported to induce a concurrent C-to-G and A-to-G base conversion¹⁷. Considering the components and arrangements of dual base editors (Fig. 5a), the miniAGBE-4 was selected for the comparison at three genomic loci in HEK293T cells. The data showed that the A-to-G and C-to-G transition in GGBE1.0 and GGBE1.3 was obviously higher than miniAGBE-4 at RP11-site1 and HEK4 (Fig. 5b, c and S5). A significantly higher proportion of C-to-G conversion was also observed in GGBE variants compared to miniAGBE-4 within the editing window (C5–C7), especially at HEK4–C6 site (Fig. 5d). More importantly, we found that the miniAGBE-4 could induce a significantly higher indel frequency than the GGBE variants (Fig. 5e). Overall, our analysis demonstrated that the GGBE1.0 and GGBE1.3 introduced a concurrent C-to-G and A-to-G base editing events with high editing efficiency and specificity.

Off-target analysis of HMGNI-fused GBE, ABE, and GGBE in HEK293T cells

Given that the HMGNI could alter chromatin accessibility, we further evaluated the off-target effects of HMGNI-fused GBE, ABE, and GGBE. To address the gRNA-dependent off-target effects, potential off-target (OT) sites were selected for analysis using Cas-OFFinder³⁴, or based on previously reported genomic loci³⁵, after which cumulative C-to-G or A-to-G editing was calculated. Our data showed a slightly higher off-target editing was observed in HMGNI-fused base editors (Fig. S6A, B).

Next, the effect of HMGNI-fused variants on Cas9-independent off-target DNA editing was characterized. A previously developed orthogonal R-loop assay³⁶ (Fig. S6C) was employed to evaluate off-target DNA editing at three genomic loci. Briefly, HEK293T cells were

co-transfected with plasmids encoding SpBE (*Streptococcus pyogenes* base editor) variants and an on-target sgRNA, along with a catalytically inactive SaCas9 (dSaCas9) and a SaCas9 sgRNA targeting a genomic locus unrelated to the SpBE on-target site. Then the Cas9-independent off-target editing was estimated based on detected editing efficiency in these dSaCas9-generated R-loops. Our data hinted that the HMGNI-fused base editors exhibited a higher editing frequency in one of the three R-loops (Fig. S6D). Finally, to measure the extent of cellular RNA editing by these base editors, HEK293T cells were transfected with the indicated base editors, after which the C-to-N and A-to-I mutation frequencies across the transcriptome were detected. The data showed that the HMGNI-fused variants did not induce a higher alteration of RNA editing (Fig. S6E), including the editing frequency (Fig. S6F) and a number of RNA single nucleotide variants (SNVs) (Fig. S6G). Collectively, our data indicated that the HMGNI-fused base editors might induce a modestly increased gRNA-dependent and Cas9-independent off-target DNA editing, but not RNA off-target editing.

Potential application of GGBE in MNVs

Our results demonstrated that GGBE enables the efficient and concurrent conversion of C-to-G and A-to-G, thereby expanding the editing spectrum of conversion types. This expanded editing capability could be used to study and treat genetic diseases associated with MNVs, a newly identified category of genomic DNA sequence variants. Here, we identified a series of MNVs that could be rescued or created via GGBE through bioinformatics analysis (Supplementary Data 3). To further demonstrate the application value of GGBE, lentivirus-mediated knock-in³⁷ was employed to create model MNVs-associated loci on the chromosome of HEK293T cells, after which GGBE was introduced to evaluate the capacity to revert these model pathogenic MNVs. For example, Nakajo syndrome induced by the GG-to-CA mutation of PSMB8 (proteasome 20S subunit beta 8)³⁸ was selected for investigation, which was characterized by skin eruption, splenomegaly, hyper γ -globulinemia, etc³⁹. Our data indicated that the editing efficiency of targeted MNVs-associated clinical variants with GGBE1.0 ranged from 0.61% to 12.84% with simultaneous C-to-G and A-to-G editing (Fig. S7A, B). Overall, our analysis showed that the reported GGBE could be potentially employed to create cell models carrying clinically relevant MNVs, or even treat diseases by correcting these MNVs.

Discussion

While the C-to-G base editors have been constructed, a dual functional C-to-G and A-to-G base editor has not been explored in detail. Further, the research focusing on how the chromatin microenvironment influenced genome editing is limited, especially for base editing. In this work, a chromatin-associated factor HMGNI was found to improve both the base editing events of C-to-G and A-to-G. A dual deaminase-mediated base editor (GGBE) was constructed for simultaneous C-to-G and A-to-G conversion, which has the potential for studying and developing genetic therapies for MNVs.

Previous research demonstrated that several chromatin-modulating factors contributed to a similar improvement in Cas9 activity, including HMGNI, HMGB1, histone H1, and CHD1²⁰. However, we found that these four factors had disparate editing enhancement effects on the C-to-G transition. The chromosomal protein HMGNI exhibited the highest improvement for GBE, significantly higher than the previous GBE variant with pioneer factor SOX2, but the editing window did not alter. Considering that the C-to-G transition is formed via the DNA repair pattern of translesion DNA synthesis (TLS)⁴⁰, and HMGNI was reported to enhance the rate of DNA repair via reducing the compaction of chromatin structure³¹, the improvement of C-to-G was reasonable. However, HMGNI-fused A8e exhibited only a modest increase in A-to-G yield. Although HMGNI might reduce the compacted chromatin to promote Cas9 binding, the editing effects of GBE

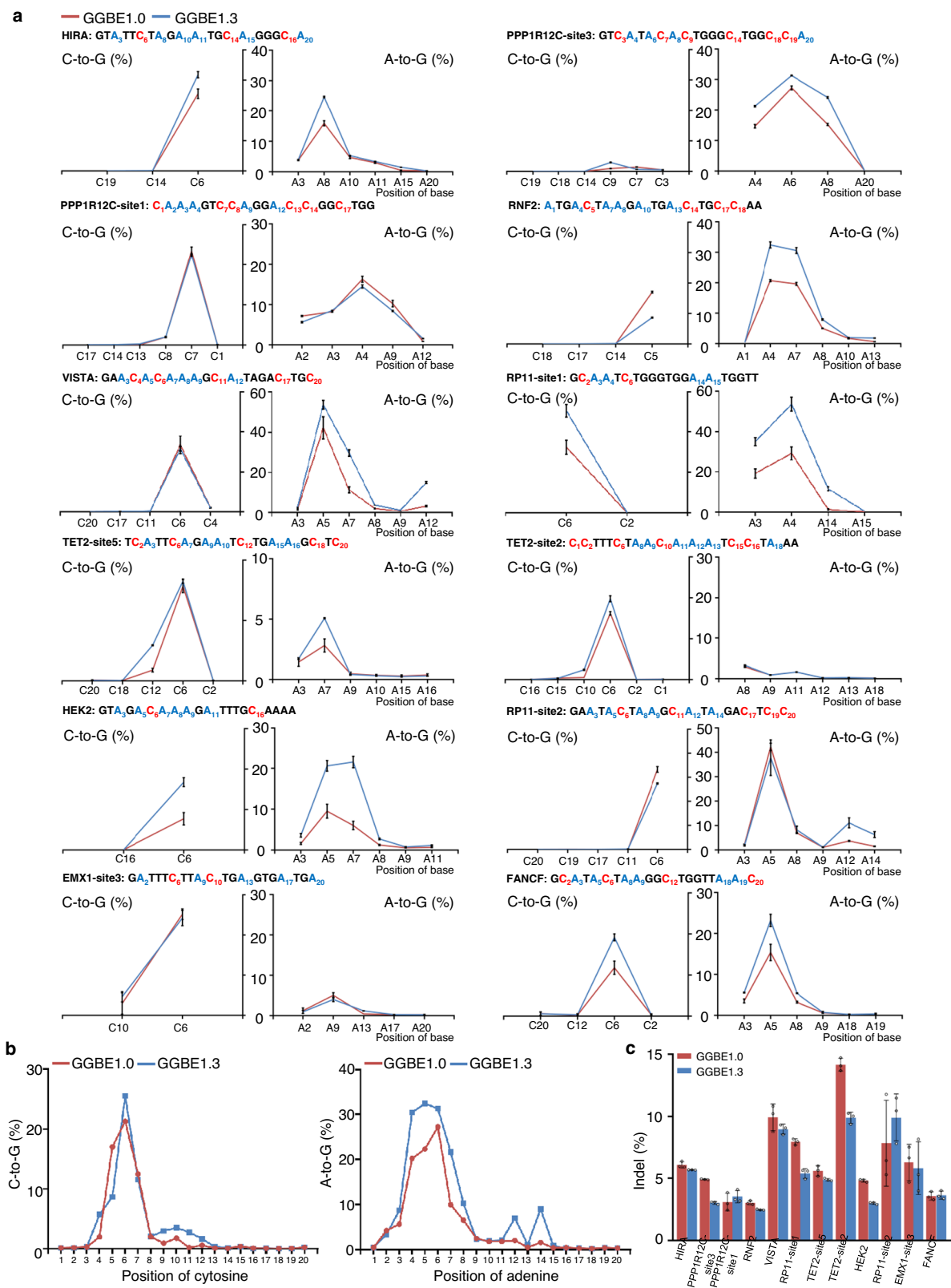


Fig. 4 | Characterization of GGBE variants in mammalian cells. a Comparison of editing efficiency between GGBE1.0 and GGBE1.3 at twelve endogenous genomic loci in HEK293T cells. **b** Average C-to-G and A-to-G base editing efficiency at C1-C20 (left) and A1-A20 (right) positions of protospacer from the twelve loci of GGBE1.0

and GGBE1.3. **c** Average indel frequency between GGBE1.0 and GGBE1.3 at 12 endogenous genomic loci in HEK293T cells. Mean \pm SEM (a, c) of all individual values of sets of $n = 3$ independent replicates are shown. Source data are provided as a Source Data file.

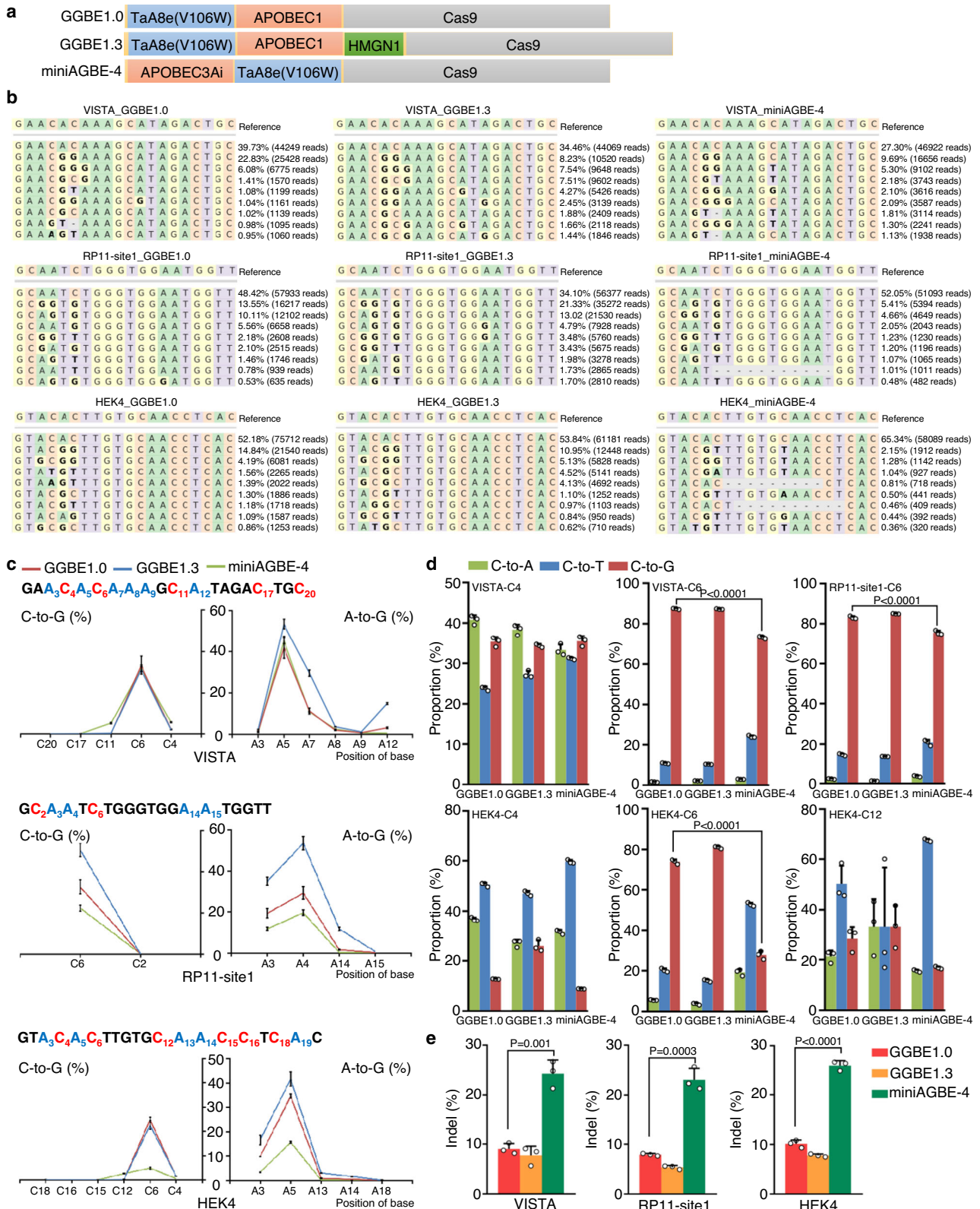


Fig. 5 | Comparison between GGBE variants and AGBE in HEK293T cells.

a Schematic of base editor variant of GGBE1.0, GGBE1.3, and mini-AGBE4. **b** The genotypes and reads and editing frequency of indicated genotypes across GGBE1.0, GGBE1.3, and mini-AGBE4 at VISTA, RP11-site1, and HEK4. **c** Comparison of editing efficiency across GGBE1.0, GGBE1.3 and mini-AGBE4 in HEK293T cells.

d Comparison of editing proportion of C-to-A/T/G across GGBE1.0, GGBE1.3 and

mini-AGBE4 in HEK293T cells. **e** Comparison of indel frequency across GGBE1.0, GGBE1.3 and mini-AGBE4 in HEK293T cells. Mean \pm SEM (**c–e**) of all individual values of sets of $n = 3$ independent replicates are shown. All statistical analyses for samples were conducted using unpaired Student’s *t* test (two-tailed) in GraphPad Prism 8. Source data are provided as a Source Data file.

and ABE are also influenced by other molecular mechanisms. While C-to-G conversion is mediated via UDG and DNA polymerase of TLS in eukaryotes⁴⁰, A-to-G conversion is recognized as the inhibition of hypoxanthine excision². Accordingly, while an open chromatin environment might facilitate the assembly of TLS-related repair factors to improve the C-to-G transition, the excision of hypoxanthine might not be influenced. However, we could not exclude other mechanisms accounting for the discrepancy between the effect of HMGNI on the GBE and A8e.

A discrepant effect on the indel products between HMGNI-fused GBE and A8e was observed. The indel products of HMGNI-GBE increased at the majority of testing sites, while that of HMGNI-A8e did not. Notably, three TET2 sites tested by us in C-to-G conversion exhibited an obvious alteration of indel products. We assumed that the sequence context or the chromatin microenvironment might influence the indel products caused by a base editor. Additionally, histone-associated factors TIP60⁴¹, USP22^{42–44}, and RNF168^{45–47}, which were observed to improve the C-to-G conversion, which was reported to be involved in DNA repair. Nevertheless, their function in base editing remains to be explored in more detail in the future. Notably, the histone methylase EZH2^{48,49}, which was reported to induce chromatin compaction, resulted in a significant decrease of C-to-G and A-to-G transition frequency, which further demonstrated the negative influence of a compacted chromatin microenvironment on base editing. More importantly, overexpression of HMGNI did not induce any negative effects on the expression of the base editor, transcription of the targeted loci, or cell viability.

Next, to further confirm the effects of HMGNI and construct improved base editor variants, diverse arrangements with HMGNI and other efficient components were attempted. We found that the HMGNI-fused miniCGBE exhibited a higher C-to-G conversion efficiency, further indicating a positive role of HMGNI. In addition, given that the human UNG was reported to negatively affect the C-to-G transition, we also replaced it with Ung1 and Udgx to construct additional GBE variants. The HMGNI-GBE-Udgx exhibited a higher editing efficiency at the majority of testing loci, especially for the PAM-proximal cytosines. This finding was consistent with previous research⁹. Finally, HMGNI-A8e-M was observed to induce a higher base transition for PAM-proximal adenines, which was similar to the fusion of pioneer factors between Cas9 and deaminase. We hypothesized that the HMGNI could improve the chromatin accessibility and also act as a long linker in this type fusion, hence increasing the binding capacity between deaminase and PAM-proximal bases. Overall, HMGNI was found to positively influence the base editing efficiency of GBE and ABE.

Although an AGBE system has been constructed recently, its editing efficiency and specificity were relatively low¹⁷. To extend the genome editing possibility and improve the specificity, a GGBE catalyzing the concurrent C-to-G and A-to-G conversion was established by the fusion of GBE and A8e. The addition of UNG or Udgx to GGBE did not increase the editing yield, suggesting that the concurrent base editing might be negatively affected by the activity of glycosylase, which was consistent with the AGBE system¹⁷. To ensure the balance of two types of base conversion, we tested several arrangements of the deaminases, and introduced a more active APOBEC1 mutant into GGBE. The TadA8e-APOBEC1-Cas9 assembly was found to exhibit the optimal base editing pattern. To potentially improve the editing activity of GGBE, several variants incorporating HMGNI were constructed. Unfortunately, the N-terminal fusion (GGBE1.1) was not found to improve the efficiency of concurrent base editing. These data indicate that simultaneous A-to-G and C-to-G base editing might potentially depend on a different molecular mechanism than single-base transition. Accordingly, the addition of HMGNI to N-terminal did not result in higher concurrent editing, despite we could not exclude the possibility that GGBE1.1 could improve the efficiency at other

genomic loci with low intrinsic chromatin accessibility. Subsequently, a GGBE variant in which HMGNI was placed between APOBEC1 and Cas9, called GGBE1.3, exhibited an improved editing window and decreased indel frequency. This fusion pattern increased the editing window possibly due to the role of a long linker and improved chromatin accessibility. We hypothesized that the decreased indel byproducts might be caused by the alteration of the editing window and also the chromatin accessibility. In addition, a comparison between GGBE and AGBE revealed that the former achieved more specific and efficient concurrent editing of C-to-G and A-to-G. Moreover, the indel frequency of AGBE was significantly higher than that of GGBE, which further demonstrated the superiority of our construct. More importantly, the HMGNI-fused base editor variants also induced slightly higher gRNA-dependent and Cas9-independent DNA off-target effects but not the RNA off-target effects. This phenomenon is in accord with the positive function of HMGNI in base editing possibly due to the increased chromatin accessibility, which might not affect RNA editing caused by deaminases. Finally, since the GGBE could induce a dual-type base conversion, it could be applied for a spectrum of scenarios, such as the correction or creation of MNVs, and mutagenesis screens, as well as for programmable installation of transcription factor-binding sites. Thus, several clinically relevant MNVs were also constructed or corrected in cellular models with substantial editing efficiencies. However, the bystanders were also observed in these MNVs, which could be further optimized using deaminases with a narrow editing windows.

In summary, the construction of HMGNI-fused GBE, ABE, and GGBE has expanded the research frontiers of BEs, and enriched the base editing toolbox in mammalian cells.

Methods

Ethics statement

The study complied with all the ethical regulations for work with patients. The study was approved by the Ethics Committee of Union Hospital, Tongji Medical College, Huazhong University of Science and Technology and written informed consent was obtained from all patients.

Cell culture, transfection, and CCK8 assay

Cell lines used in this research were purchased from the ATCC. HEK293T and HeLa cells were cultured in DMEM supplemented with 10% FBS in the humidified incubator equilibrated with 5% CO₂ at 37 °C. For transfected experiments, cells were seeded in 24-well plates (Corning, USA) and performed using polyethylenimine (Polysciences, USA) based on the manufacturer's instructions. Briefly, 600 ng of base editor and 300 ng of sgRNA-expressing plasmid were together transfected with 50 µl of Opti-MEM (Gibco, USA) containing 2.7 µl of polyethylenimine for 24 h. And then cells were replaced with fresh medium with 5 µg/ml puromycin (Merck, USA) for another 4 d. Finally, genomic DNA was extracted via QuickExtract DNA Extraction Solution (Epicentre, USA). On-target genomic regions (200 bp–300 bp) of interest were amplified by PCR for high-throughput DNA sequencing. The cell viability was evaluated via CCK8 analysis. Briefly, the HEK293T cells were seeded into 96-well plates and transfected with indicated base editors at approximately 60% confluency. Then after 48 and 72 h transfection, cell viability was measured by CCK8 reagent according to the manufacturer's instructions (DOJINDO, Japan).

Culture, transfection, and assay of human primary prostate carcinoma cell

The primary cell was dissected and cultured as previously reported⁵⁰. Briefly, fresh tissue biopsy samples were placed in media DMEM (Invitrogen, USA) with GlutaMAX (Invitrogen, USA), 100 U/ml penicillin, 100 µg/ml streptomycin (Gibco USA), Primocin 100 g/ml (InvivoGen, USA), and 10 µmol/l ROCK inhibitor (Selleck Chemical Inc.,

USA) and washed. The dissected tissue was enzymatically digested with 250 U/ml of collagenase IV (Life Technologies, USA) and TrypLE express (Gibco) in a ratio 1:2 with Collagenase IV and then centrifuged. The pellet was washed and resuspended with prostate-specific culture media composed of DMEM/F12 (Invitrogen, USA) with GlutaMAX, 100 U/ml penicillin, 100 µg/ml streptomycin (Gibco, USA), Primocin 100 Mg/ml (InvitroGen, USA), B27 (Gibco, USA), N-Acetylcysteine 1.25 mM (Sigma-Aldrich, USA), Mouse Recombinant EGF 50 ng/ml (Invitrogen, USA), Human Recombinant FGF-10 20 ng/ml (Peprotech, USA), Recombinant Human FGF-basic 1 ng/ml (Peprotech), A-83-01 500 nM (Tocris, USA), SB202190 10 µM (Sigma-Aldrich, USA), Nicotinamide 10 mM (Sigma-Aldrich, USA), (DiHydro) Testosterone 1 nM (Sigma-Aldrich, USA), PGE2 1 µM (R&D Systems, USA), Noggin conditioned media (5%) and R-spondin conditioned media (5%). The final resuspended pellet was combined with growth factor-reduced Matrigel (Corning, USA) in a 1:2 volume ratio and then pipetted onto a 24-well cell suspension culture plate.

The primary cells were transfected with indicated base editor and sgRNA plasmids using lipo3000 (GlpBio, USA). After 24 h transfection, cells were replaced with fresh medium with 1 µg/ml puromycin (Merck, USA) for another 5 d. Finally, genomic DNA was extracted via Quick-Extract DNA Extraction Solution (Epicentre, USA) and on-target genomic regions were amplified by PCR and analyzed with high-throughput DNA sequencing.

Plasmid construction

Four chromatin-modulating proteins HMGB1, HMGNI, CHD1, and Histone H1 were synthesized by AZENTA. The other chromatin-associated factors were amplified with Phusion DNA polymerase (NEB, USA) from HEK293T cDNA library. PCR products were gel purified, digested with DpnI restriction enzyme (NEB, USA), and assembled via Gibson assembly based on the manufacturer's instructions. All gRNA-expression plasmids were assembled via Golden Gate with the protospacer sequence embedded in the primers, and *RNF2* sgRNA expression plasmids were used as the template¹. The main primers are listed in Supplementary Data 1.

Strains and culture conditions

E. coli Trans5α was used as the cloning host and cultured at 37 °C in lysogeny broth (LB, 1% (w/v) tryptone, 0.5% (w/v) yeast extract, and 1% (w/v) NaCl). 100 mg/L Ampicillin (Sigma, USA) were used for screen of positive cloning.

Western blotting

The western blotting assay was performed as previously reported⁴⁸. Briefly, cellular extracts from HEK293T cells were prepared with lysis buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.5% NP-40) for 30 min at 4 °C and then denatured for 10 min at 95 °C. The cell lysates were resolved using 10% SDS-PAGE gels and transferred onto acetate cellulose membranes. For incubation, membranes were incubated with Cas9 (Beyotime Biotechnology, dilution: 1:2000, Cat:AF0123, China) or GAPDH (ABclonal, dilution: 1:10,000, Cat:AC002, USA) antibodies at 4 °C overnight followed by incubation with a secondary antibody HRP-conjugated Affinipure Goat Anti-Mouse IgG (H + L) (Proteintech, dilution: 1:5000, Cat:AS003, USA). Immunoreactive bands were visualized using western blotting luminal reagent (Millipore, USA) according to the manufacturer's recommendation.

High-throughput DNA sequencing of genomic DNA samples and data analysis

Next-generation sequencing library preparations and analysis were performed as previously reported⁵. Briefly, purified PCR fragments were treated in one reaction with End Prep Enzyme Mix for end repair, 5' phosphorylation and dA tailing, which was followed by T-A ligation to add adapters to both ends, of which PCR products were purified and

quantified. Then the sequencing was carried out on Illumina HiSeq instrument according to the manufacturer's instructions.

Analysis of amplicon sequencing data were performed using CRISPResso2 v.2.0.45 in batch mode²⁸, with window parameters set to -wc 10 -w 20. The base conversion frequency was obtained with the output file 'Nucleotide_percentage_summary.txt' and indel rates was acquired with 'CRISPRessoBatch_quantification_of_editing_frequency.txt' for base editor experiments. All genomic loci and deep sequencing oligos of sgRNA are listed in Supplementary Data 2.

RT PCR and Real-time RT PCR (qPCR)

Total cellular RNAs were isolated with SimplyP kit (Biospin, China) and used for the first strand cDNA synthesis with the Reverse Transcription System (TaKaRa, Japan). Quantitation of all gene transcripts was done by qPCR using Power SYBR Green PCR Master Mix and a Roche Roche LC96 sequence detection system with the expression of GAPDH as an internal control. The primer pairs used were: EMX1, 5'-TGTGCAT GTGCTGGCTG-3' (forward) and 5'-CTTGCCACCAAGGACTCTA-3' (reverse); HIRA, 5'-CTGGACTCTGAATGGGCTGG-3' (forward) and 5'-GGCTAGGCTCTGCCATAGG-3' (reverse); and GAPDH, 5'-CTGGGC TACTGAGCACC-3' (forward) and 5'-AAGTGGTCGTTGAGGGCAA TG-3' (reverse).

RNA-seq analysis and SNVs calling

For transcriptome analysis, 10^6 cells of each sample were collected and used for RNA extraction. The RNA-seq libraries were constructed and the high-throughput transcriptome sequencing was carried out with minia HiSeq instrument by the AZENTA company. For data analysis³³, qualified reads were mapped to the reference genome (Ensemble GRCh38) using STAR in 2-pass mode with default parameters. Then the Picard tool was used to sort and mark duplicates of the mapped BAM files, which were subject to split reads that spanned splice junctions, base recalibration, and variant calling with SplitNCigarReads, BaseRecalibrator, and HaplotypeCaller tools from GATK respectively.

The calling variants were filtered with default parameters using VariantFiltration tool from GATK. Variant loci in base editor over-expression experiments were filtered to exclude sites without high-confidence reference genotype calls in the control experiment. Base edits labeled as C-to-N comprise C-to-U/A/G edits called on the positive strand as well as G-to-A/U/C edits sourced from the negative strand. Base edits labeled as A-to-I comprise A-to-I edits called on the positive strand as well as T-to-C edits sourced from the negative strand.

Analysis of potential targets for the correction or creation of MNVs by GGBE

A list of MNVs was obtained from previous reports^{11,12}, which were screened to detect disease-correcting or disease-creating modifications enabled by GGBE. Disease-correcting or -creating conversions are defined as having targetable C and A bases with matching T and G bases in the restricted position (base position 5–7 of the protospacer). Patterns for selected disease-correcting MNV codons include GNG > ANC, GGN > ACN, NGG > NAC, GNG > CNA, GGN > CAN, and NGG > NCA; whereas patterns for disease-creating include ACN > GGN, ANC > GNG, CAN > GGN, CNA > GNG, NAC > NGG and NCA > NGG. The filtered MNVs are listed in Supplementary Data 3.

Lentivirus infection and base editing of MNVs-relevant targets

The generation of lentiviruses was conducted according to the previous reports^{51,52}. Briefly, synthetic target sequences containing pathogenic point mutations together with psPAX2 and pMD2.G, were co-transfected into the packaging cell line HEK293T at a weight ratio of 3:2:1. Viral supernatants were collected 48 h later, clarified by filtration, and concentrated by ultracentrifugation. Then the concentrated viruses were used to infect $5-10^5$ cells (20–30% confluence) in a 60-mm

dish with 5 mg/mL polybrene. Infected cells were selected by 4 $\mu\text{g ml}^{-1}$ blasticidin (Solarbio) to the culture medium. The target sequence-transduced HEK293T cells were then transfected with a mixture of plasmid encoding GGBE1.0 and targeted gRNA. After 5 days treatment with puromycin, cells were collected and the genomic DNA was subjected to deep sequencing to measure the editing efficiency of base editing.

Statistics and reproducibility

Unless otherwise noted, all data are presented as means \pm s.d. and analyzed with statistical methods from three independent experiments. The significance of the difference between the control and experiment group was calculated via student's *t* test using GraphPad Prism 8. $P < 0.05$ was considered to be statistically significant.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

HTS data generated in this study have been deposited in the NCBI Sequence Read Archive database under accession code [PRJNA946328](https://www.ncbi.nlm.nih.gov/sra/PRJNA946328). Source data are provided with this paper.

Code availability

Source code for CRISPResso2, STAR, Picard and GATK are available on github (<https://github.com/pinelloab/CRISPResso2>; <https://github.com/alexdobin/STAR>; <https://github.com/broadinstitute/picard>; <https://github.com/broadinstitute/gatk>).

References

- Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A. & Liu, D. R. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533**, 420–424 (2016).
- Gaudelli, N. M. et al. Programmable base editing of A·T to G·C in genomic DNA without DNA cleavage. *Nature* **551**, 464–471 (2017).
- Chen, L. et al. Programmable C·G to G·C genome editing with CRISPR-Cas9-directed base excision repair proteins. *Nat. Commun.* **12**, 1384 (2021).
- Kurt, I. C. et al. CRISPR C-to-G base editors for inducing targeted DNA transversions in human cells. *Nat. Biotechnol.* **39**, 41–46 (2021).
- Zhao, D. et al. Glycosylase base editors enable C-to-A and C-to-G base changes. *Nat. Biotechnol.* **39**, 35–40 (2021).
- Koblan, L. W. et al. In vivo base editing rescues Hutchinson-Gilford progeria syndrome in mice. *Nature* **589**, 608–614 (2021).
- Zeng, J. et al. Therapeutic base editing of human hematopoietic stem cells. *Nat. Med.* **26**, 535–541 (2020).
- Richter, M. F. et al. Phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity. *Nat. Biotechnol.* **38**, 883–891 (2020).
- Koblan, L. W. et al. Efficient C·G-to-G·C base editors developed using CRISPRi screens, target-library analysis, and machine learning. *Nat. Biotechnol.* **39**, 1414–1425 (2021).
- Sun, N. et al. Reconstructed glycosylase base editors GBE2.0 with enhanced C-to-G base editing efficiency and purity. *Mol. Ther. J. Am. Soc. Gene Ther.* **30**, 2452–2463 (2022).
- Wang, Q. et al. Landscape of multi-nucleotide variants in 125,748 human exomes and 15,708 genomes. *Nat. Commun.* **11**, 2539 (2020).
- Karczewski, K. J. et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **581**, 434–443 (2020).
- Grünwald, J. et al. A dual-deaminase CRISPR base editor enables concurrent adenine and cytosine editing. *Nat. Biotechnol.* **38**, 861–864 (2020).
- Sakata, R. C. et al. Base editors for simultaneous introduction of C-to-T and A-to-G mutations. *Nat. Biotechnol.* **38**, 865–869 (2020).
- Xie, J. et al. ACBE, a new base editor for simultaneous C-to-T and A-to-G substitutions in mammalian systems. *BMC Biol.* **18**, 131 (2020).
- Zhang, X. et al. Dual base editor catalyzes both cytosine and adenine base conversions in human cells. *Nat. Biotechnol.* **38**, 856–860 (2020).
- Liang, Y. et al. AGBE: a dual deaminase-mediated base editor by fusing CGBE with ABE for creating a saturated mutant population with multiple editing patterns. *Nucleic Acids Res.* **50**, 5384–5399 (2022).
- Park, H. D. et al. Three patients with glycogen storage disease type II and the mutational spectrum of GAA in Korean patients. *Ann. Clin. Lab. Sci.* **43**, 311–316 (2013).
- Anzalone, A. V. et al. Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* **576**, 149–157 (2019).
- Ding, X. et al. Improving CRISPR-Cas9 genome editing efficiency by fusion with chromatin-modulating peptides. *CRISPR J.* **2**, 51–63 (2019).
- Liu, G., Yin, K., Zhang, Q., Gao, C. & Qiu, J. L. Modulating chromatin accessibility by transactivation and targeting proximal dsRNAs enhances Cas9 editing efficiency in vivo. *Genome Biol.* **20**, 145 (2019).
- Shin, H. R. et al. Small-molecule inhibitors of histone deacetylase improve CRISPR-based adenine base editing. *Nucleic Acids Res.* **49**, 2390–2399 (2021).
- Becker, J. R. et al. BARD1 reads H2A lysine 15 ubiquitination to direct homologous recombination. *Nature* **596**, 433–437 (2021).
- Fnu, S. et al. Methylation of histone H3 lysine 36 enhances DNA repair by nonhomologous end-joining. *Proc. Natl Acad. Sci. USA* **108**, 540–545 (2011).
- Li, J., Braganza, A. & Sobol, R. W. Base excision repair facilitates a functional relationship between Guanine oxidation and histone demethylation. *Antioxid. Redox Signal.* **18**, 2429–2443 (2013).
- Dong, X. et al. Enhancing glycosylase base-editor activity by fusion to transactivation modules. *Cell Rep.* **40**, 111090 (2022).
- Yang, C. et al. Pioneer factor improves CRISPR-based C-to-G and C-to-T base editing. *Adv. Sci.* **9**, e2202957 (2022).
- Clement, K. et al. CRISPResso2 provides accurate and rapid genome editing sequence analysis. *Nat. Biotechnol.* **37**, 224–226 (2019).
- Ahn, W. C. et al. Covalent binding of uracil DNA glycosylase UdgX to abasic DNA upon uracil excision. *Nat. Chem. Biol.* **15**, 607–614 (2019).
- Tu, J., Chen, R., Yang, Y., Cao, W. & Xie, W. Suicide inactivation of the uracil DNA glycosylase UdgX by covalent complex formation. *Nat. Chem. Biol.* **15**, 615–622 (2019).
- Birger, Y. et al. Chromosomal protein HMG1 enhances the rate of DNA repair in chromatin. *EMBO J.* **22**, 1665–1675 (2003).
- Murphy, K. J. et al. HMG1 and 2 remodel core and linker histone tail domains within chromatin. *Nucleic Acids Res.* **45**, 9917–9930 (2017).
- Grünwald, J. et al. Transcriptome-wide off-target RNA editing induced by CRISPR-guided DNA base editors. *Nature* **569**, 433–437 (2019).
- Bae, S., Park, J. & Kim, J. S. Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics* **30**, 1473–1475 (2014).
- Rees, H. A., Wilson, C., Doman, J. L. & Liu, D. R. Analysis and minimization of cellular RNA editing by DNA adenine base editors. *Sci. Adv.* **5**, eaax5717 (2019).
- Doman, J. L., Raguram, A., Newby, G. A. & Liu, D. R. Evaluation and minimization of Cas9-independent off-target DNA editing by cytosine base editors. *Nat. Biotechnol.* **38**, 620–628 (2020).

37. Song, M. et al. Sequence-specific prediction of the efficiencies of adenine and cytosine base editors. *Nat. Biotechnol.* **38**, 1037–1043 (2020).
38. Ebstein, F., Poli Harlowe, M. C., Studencka-Turski, M. & Krüger, E. Contribution of the unfolded protein response (UPR) to the pathogenesis of proteasome-associated autoinflammatory syndromes (PRAAS). *Front. Immunol.* **10**, 2756 (2019).
39. Kanazawa, N. Nakajo-Nishimura syndrome: an autoinflammatory disorder showing pernio-like rashes and progressive partial lipodystrophy. *Allergol. Int.* **61**, 197–206 (2012).
40. Jiang, G. et al. Molecular mechanism of the cytosine CRISPR base editing process and the roles of translesion DNA polymerases. *ACS Synth. Biol.* **10**, 3353–3358 (2021).
41. Lashgari, A., Kougnassoukou Tchara, P. E., Lambert, J. P. & Côté, J. New insights into the DNA repair pathway choice with NuA4/TIP60. *DNA Repair* **113**, 103315 (2022).
42. Nardi, I. K., Stark, J. M., Larsen, A., Salgia, R. & Raz, D. J. USP22 interacts with PALB2 and promotes chemotherapy resistance via homologous recombination of DNA double-strand breaks. *Mol. Cancer Res.* **18**, 424–435 (2020).
43. McCann, J. J. et al. USP22 functions as an oncogenic driver in prostate cancer by regulating cell proliferation and DNA repair. *Cancer Res.* **80**, 430–443 (2020).
44. Wang, A. et al. USP22 induces cisplatin resistance in lung adenocarcinoma by regulating γ H2AX-mediated DNA damage repair and Ku70/bax-mediated apoptosis. *Front. Pharmacol.* **8**, 274 (2017).
45. Kraus, J. J. et al. RNF168-mediated localization of BARD1 recruits the BRCA1-PALB2 complex to DNA damage. *Nat. Commun.* **12**, 5016 (2021).
46. Katsuki, Y. et al. RNF168 E3 ligase participates in ubiquitin signaling and recruitment of SLX4 during DNA crosslink repair. *Cell Rep.* **37**, 109879 (2021).
47. Kim, J. J. et al. A novel reciprocal crosstalk between RNF168 and PARP1 to regulate DNA repair processes. *Mol. Cells* **41**, 799–807 (2018).
48. Grau, D. et al. Structures of monomeric and dimeric PRC2:EZH1 reveal flexible modules involved in chromatin compaction. *Nat. Commun.* **12**, 714 (2021).
49. Margueron, R. et al. Ezh1 and Ezh2 maintain repressive chromatin through different mechanisms. *Mol. Cell* **32**, 503–518 (2008).
50. Wang, K. et al. WNT5a signaling through ROR2 activates the hippo pathway to suppress YAP1 activity and tumor growth. *Cancer Res.* **83**, 1016–1030 (2023).
51. Li, B. et al. Sequence motifs and prediction model of GBE editing outcomes based on target library analysis and machine learning. *J. Genet. Genomics* **49**, 254–257 (2022).
52. Yang, C. et al. Circadian rhythm is disrupted by ZNF704 in breast carcinogenesis. *Cancer Res.* **80**, 4114–4128 (2020).

Acknowledgements

This research was financially supported by the National Natural Science Foundation of China (82203238, 32225031, 32171449), Tianjin Synthetic Biotechnology Innovation Capacity Improvement Project (TSBICIP-CXRC-034, TSBICIP-KJGG-017).

Author contributions

C.Y., J.L., Z.C., X.Z., and C.B. designed the research and wrote the manuscript. C.Y. performed the experiments studies, and the computational studies, and analyzed data. Z.M., K.W., X.D., and M.H. performed experiments. Q.L. and X.Z. performed the computational studies and analyzed data.

Competing interests

The authors declare no conflicts of interest.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-023-38193-2>.

Correspondence and requests for materials should be addressed to Changhao Bi or Xueli Zhang.

Peer review information *Nature Communications* thanks Liangxue Lai and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023