Computational and Structural Biotechnology Journal 21 (2023) 856–868

Contents lists available at ScienceDirect

Computational and Structural Biotechnology Journal

journal homepage: www.elsevier.com/locate/csbj

Dual guide RNA-mediated concurrent C&G-to-T&A and A&T-to-G&C conversions using CRISPR base editors

Yuting Zhao [a,1], Min Li [a,1], Jie Liu [b], Xiaowen Xue [a], Jingli Zhong [a], Jianxiang Lin [a], Bo Ye [c,*], Jun Chen [d,*], Yunbo Qiao [a,e,f,**]

[a] Precise Genome Engineering Center, School of Life Sciences, Guangzhou University, Guangzhou 510006, China
[b] Guangzhou Medical University, Guangzhou 511495, China
[c] Department of Thoracic Surgery, Shanghai Chest Hospital, Shanghai Jiao Tong University, Shanghai 200030, China
[d] Department of Endocrinology, Qilu Hospital, Cheeloo College of Medicine, Shandong University, Jinan 250012, China
[e] Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200125, China
[f] Shanghai Institute of Precision Medicine, Shanghai 200125, China

ARTICLE INFO

Article history:
Received 30 July 2022
Received in revised form 30 December 2022
Accepted 30 December 2022
Available online 2 January 2023

Keywords:
Dual guide RNAs (gRNAs)
Multi-nucleotide variant (MNV)
C&G-to-T&A and A&T-to-G&C conversions
TP53
SpRY

ABSTRACT

Base editing tools enable precise genome modifications, disease modeling, and promising gene therapy. However, many human genetic diseases are elicited by multi-nucleotide variants (MNVs) with heterogeneous substitutions at the same genomic locus. Based on the adenine and cytosine base editors, dual base editors that can catalyze concurrent C-to-T and A-to-G editing have been developed, while simultaneous C&G-to-T&A and A&T-to-G&C conversions on the same allele have not been achieved at the desirable site. Here we propose a strategy of combining base editors with dual guide RNAs (gRNAs) that target two overlapped neighboring loci on the opposite strands, which can induce simultaneous C&G-to-T&A and A&T-to-G&C conversions within their overlapping targeting windows. Moreover, one of the paired gRNAs is mutated to perfectly match another gRNA-edited sequence, efficiently facilitating concurrent base conversions on the same allele. To further expand the targeting scopes, PAMless SpRY Cas9-mediated base editors are combined with our optimized dual gRNAs system to induce expected concurrent base editing and to install neighboring pathogenic MNVs in TP53 in cancer cells. In addition, more complex mutation types can be achieved by integrating dual base editors and our dual gRNAs strategy. Thus, we establish a general strategy to efficiently induce MNVs in human genome, helping to dissect the functions of pathogenic MNVs with multifarious types.

1. Introduction

Along with the rapid development of high-throughput sequencing technology, diseases-associated genetic variants, including single-nucleotide variants (SNVs) and multi-nucleotide variants (MNVs), have been extensively discovered in various diseases and cancers [1–3]. SNVs, including C-to-T and A-to-G conversions, can be precisely introduced by using CRISPR-mediated cytosine and adenine base editors (CBE and ABE) [4,5], and phenotypic significance of SNVs can be dissected individually or in massively high-throughput screenings [6–8]. At present, the widely used single base editors that can efficiently introduce point mutations are BE4max and ABEmax. Based on the first generation of base editors, they fused the second UGI, optimized nuclear localization signals, and increased the linker length of APOBEC1- Cas9 nickase (Cas9n) and Cas9n-UGI respectively [9]. Recent landscape studies reveal 839,126 MNVs in metastatic solid tumors [10] and 1792,248 MNVs across the human genome from the Genome Aggregation Database (gnomAD) [3], highlighting the importance and hazardousness of MNVs during disease progression. Nevertheless, it is of great challenge to introduce MNVs into the desired genome loci and to understand the genome-wide mutational mechanisms of MNVs.

* Corresponding authors.
** Corresponding author at: Precise Genome Engineering Center, School of Life Sciences, Guangzhou University, Guangzhou 510006, China.
E-mail addresses: yebo0430@sjtu.edu.cn (B. Ye), junchen001@outlook.com (J. Chen), ybqiao@shsmu.edu.cn, ybqiao@gzhu.edu.cn (Y. Qiao).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.csbj.2022.12.055

To achieve concurrent base conversions, several groups have successfully constructed dual base editors by fusing both cytosine (APOBEC1, APOBEC3A, AID) and adenine (ecTadA/TadA* or TadA) deaminases with Cas9n, which can catalyze concurrent C-to-T and A-to-G conversions at the same target site efficiently in mammalian cells and plants [11–14]. To further expand the applications of dual C &A base editors, PAMless SpRY Cas9-mediated dual base editors have been developed to release the restriction of protospacer adjacent motif (PAM) sequences [13,15]. Importantly, dual base editors have been successfully applied in saturated mutagenesis of plant genes to generate herbicide-tolerant rice mutants [14]. In addition, simultaneous C&G-to-T&A and A&T-to-G&C conversions are also frequently needed for MNVs introducing and highly complicated mutagenesis.

In the present study, we developed a dual guide RNAs (gRNAs) system that is composed of a leading gRNA with high targeting efficiency for the original genome and a subordinated gRNA matched with the mutated DNA sequences. By combining with canonical or PAMless Cas9-mediated single or dual base editors, it can stimulate efficient C&G-to-T&A or A&T-to-G&C conversions or multifarious mutations simultaneously on the same allele in human genome. This dual gRNAs strategy could expand the toolkit and applications of base editors for functional dissection, disease modeling, and biomedical or agricultural engineering.

2. Materials and methods

2.1. Plasmid construction

BE4max, ABEmax, SpRY-BE4max, SpRY-ABE8e and CABE-RY were kind gifts from Dr. Xingxu Huang (ShanghaiTech University, China). sgRNA target sites (GNNNNNNNNNNNNNNNNGG) were chosen from the exons of the target genes. A pair of sgRNAs (sg-L and sg-R) targeting the same locus need to have overlapping targeting regions (at least 2-nt). Oligos used for sgRNA construction were synthesized, annealed, and cloned into *Bsa*I (NEB) sites of the pGL3-U6-sgRNA-mCherry and pGL3-U6-sgRNA-EGFP expression vectors with T4 DNA ligase (Vazyme, China). The primers used for sg-L and sg-R construction are listed in Supplementary Table 1.

To construct 2U6 or 2U6mut, the following oligos were designed: ATGCGTCTCGAAAC (Rev 20 N) CGGTGTTTCGTCCTTTCCACAAG, and ATGCGTCTCAACCG (For 20 N) GTTTTAGAGCTAGAAATAGCAAG; 20 N represents a pair of guide RNA sequences (For and Rev). The primers contain the cleavage site of the endonuclease in the system and can be used to amplify gRNA scaffold and U6 fragments with pUC57-U6-gRNA plasmid as a template. The oligos used for 2U6 and 2U6mut construction are listed in Supplementary Table 2–4. PCR was performed with the above oligos as primers and pUC57-U6-gRNA plasmid as templates with the following PCR procedure: (95 °C for 5 min of pre-degeneration, 35 repeated cycles (95 °C for 30 s, 58 °C for 30 s, and 72 °C for 20 s), 72 °C for 5 min for extension). After PCR amplification, the fragments containing sg-L – gRNA Scaffold – U6 – sg- were obtained. Then PCR products were inserted into the pGL3-U6-mCherry plasmid to obtain U6 - sg-L – gRNA Scaffold – U6 – sg-R- gRNA Scaffold (2u6) with Golden Gate system: BsmbI (NEB) and T4 DNA ligase (Vazyme, China) with the following procedure: 10 cycles of 37 °C for 5 min and 16 °C for 10 min, followed by 37 °C for 15 min and 80 °C for 15 min

2.2. Cell culture and transfection

HEK293T and HeLa cells were cultured in DMEM (Gibco) supplemented with 10% fetal bovine serum (Gibco) and 1% penicillin/streptomycin (Gibco), and incubated in a 5% CO_2 humidified incubator at 37 °C. The cells were seeded on 24-well plates and transfected with 500 ng base editors, 250 ng sg-L gRNA and 250 ng sg-R gRNA expression plasmids per well using EZtrans transfection reagent (Shanghai life iLAB BIO Technology) following the manufacturer's instructions. For 2U6 (or 2U6mut) system, 750 ng base editors and 250 ng 2U6 (or 2U6mut) expression plasmids were transfected per well. 96 h after transfection, the cells were collected and fluorescent-positive cells were sorted using flow cytometry for PCR amplification and sequencing.

2.3. Genomic DNA extraction and PCR amplification

About 5000–10,000 cells after flow cytometry sorting or antibiotic screening were collected and centrifuged at 12,000 rpm (13,523 g) for 5 min. Cell pellets were lysed with 25 μl lysis buffer (10 mM Tris-HCl (pH 8.0), 50 mM KCl,1.5 mM $MgCl_2$, 0.5% Nonidet P-40, 0.5% Tween-20, and 100 ug/ml proteinase K (ThermoFisher Scientific) in a PCR thermal cycler with the following procedure: 68 °C for 30 min, 16 °C for 2 min, 98 °C for 5 min 1 μl cell lysate was used as PCR template to amplify target fragments using P505 polymerase (Vazyme, China) with the following procedure in a 25 μl system: 95 °C for 5 min of pre-denaturation, 35 repeated amplification cycles (95 °C for 10 s, 58 °C for 15 s, and 72 °C for 20 s), followed by a final extension at 72 °C for 5 min. PCR products were sequenced by Sanger sequencing, The results from Sanger sequencing were uploaded to EditR (https://moriaritylab.shinyapps.io/editr_v10/) for quantifying the mutation rates. The primers used for PCR amplification and Sanger sequencing are listed in Supplementary Table 5.

2.4. Targeted deep sequencing and data analysis

PCR products were purified with AxyPre PCR Clean-Up kit (AXYGEN), and then subjected to library construction and high-throughput sequencing on an Illumina Novaseq6000 sequencing platform with PE150 mode (Novogene, China). The amplicon sequencing data were analyzed using CRISPResso2 (v.2.0.31) [16], with parameters "–base_edit –wc − 15 –fastq_output –base_editor_output–write_cleaned_report –place_report_in_output_folder". Editing efficiency was quantified from the "Quantification_window_nucleotide_percentage_table.txt" file. Indels were quantified from the "Alleles_frequency_table_around _sgRNA_* .txt". The results, including C-to-T and G-A conversion rates, A-to-G and T-to-C conversion rates, the probability of dual base mutations on the same strain of DNA, and indel rates were calculated.

2.5. Monoclonal cell construction

HeLa cells were transfected with base editors and corresponding 2U6 (3 ug base editor and 1 ug gRNA in each 6-cm dish) using EZtrans transfection reagent (Shanghai life iLAB BIO Technology), the corresponding 2U6 and primer sequences are listed in Supplementary Table 6. 24 h after transfection, puromycin (4 ug/ml; Merck) was supplemented to transfected cells expressing pGL3- 2U6–2sg-mCherry-puromycin. 48–72 h after transfection, cells were collected and sent to Sanger sequencing. To generate single clones with desired mutations, about 100 transfected cells were plated on each 96-well plate. After proliferation for 15 days, single and round clones were collected and seeded onto a 24-well plate individually, and ultimately were passaged onto 10-cm dish. After genotyping, clones with expected dual mutations were expanded for subsequent experiments and analysis.

2.6. Apoptosis detection

A total of 10,000 wild-type and mutated (from single clone) HeLa cells were seeded on 24-well plates. The cells were treated with 5 μM GPX4-IN-3 at 24 or 48 h to induce apoptosis. Cell apoptosis was measured with Annexin V-FITC Apoptosis Detection Kit (Vazyme, China). Three independent replicates were performed and presented.

2.7. RNA-Seq and analysis

HeLa cells transfected with base editors and corresponding gRNAs (3 ug base editor and 1 ug gRNA on 6-cm dish) were subjected to cell sorting and RNA extraction. A total of 500 ng RNA was subjected to bulk RNA sequencing (6 G raw data for each sample) (Novogene, China). Two single cell clones were performed, and non-transfected HeLa cell samples served as a control to analyze RNA differential expression. Briefly, sequencing reads were mapped to the human reference genome (hg38) by STAR [17] software (version 2.7.10a) and then annotated from GENCODE version v40. Raw read counts were estimated by htseq-count [18] (v2.0.1), and DESeq2 [19] (v1.34.0) was used for differential expression analysis. Heatmaps were generated using the TPM values with pheatmap (v1.0.12). Functional annotation based on gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway was performed using R/Bioconductor package clusterProfiler [19] (v4.2.2).

2.8. Statistical analysis

All data were presented as mean ± standard error from two to three individual replicates for all experiments. Data were analyzed by Student's t-test *via* GraphPad prism software 8.0.1. Results were considered significant for *p < 0.05, * *p < 0.01, * **p < 0.001, and * ** *p < 0.0001.

3. Results

3.1. Dual gRNAs strategy facilitates concurrent C&G-to-T&A or A&T-to-G&C conversions in a mini targeting window

Human genetic diseases are frequently caused by genetic variants, among which, base conversions, including C-to-T and A-to-G (G-to-A and T-to-C on the opposite strand), account for 68% of point mutations (recovered from ClinVAR database) (Fig. 1a). The majority of mutations can be introduced by using cytosine or adenine base editors [20]. Further analysis revealed 322,699 pairs of neighboring mutations within 6-nt distance from the gnomAD, including 12.59% of A&C-to-G&T and 12.49% of T&G-to-C&A conversions (Fig. 1b), which can be potentially achieved by utilizing dual base editors [11–14]. To resolve the remaining MNVs (18.5% of C&G-to-T&A and 4.49% of A&T-to-G&C mutations), we designed a pair of gRNAs (sg-L and sg-R) targeting the sense and anti-sense strands respectively, with an overlapping targeting window (base positions 4–8 of the protospacer for the base editor; at least 2-nt overlapping); the target sequences were numbered from "1" (following the flanked "CCN") to "26–32″ (followed by "NGG") (Fig. 1c). Besides, we applied BE4max and ABEmax, improved versions of BE4 and ABE base editors with higher editing efficiency [9].

To test the feasibility of paired gRNAs to induce C&G-to-T&A or A&T-to-G&C conversions on the same allele, BE4max or ABEmax were co-transfected with a pair of gRNAs into HEK293T cells (Fig. S1a) and the editing frequency was determined. As a control, base editors with a single gRNA can induce efficient base conversions within their specific targeting window in all tested sites (Fig. 1d-e; Figs. S1b-c). When a pair of gRNAs were co-transfected, BE4max induced concurrent C-to-T and G-to-A conversions in seven target sites (Fig. 1d; Fig. S1b) and ABEmax induced concurrent A-to-G and T-to-C conversions in three target sites (Fig. 1e; Fig. S1c). Moreover, simultaneous base conversions on the same DNA strand, as analyzed from the targeted deep sequencing results, were observed for both BE4max and ABEmax with high efficiency (Fig. 1d-e). Generally, the editing efficiency induced by paired gRNAs was equivalent to single gRNA at a specific target site. Because of the double nicks induced by the paired gRNAs within the target loci, higher proportions of indels were observed in most groups transfected with paired gRNAs and BE4max or ABEmax (Fig. 1f-g). However, the indel rates remained within an acceptable range. Our results demonstrate that the double gRNAs strategy is feasible for inducing simultaneous C&G-to-T&A or A&T-to-G&C in human genome.

Considering the possibility that the concurrent C&G-to-T&A or A&T-to-G&C editing may occur on different alleles, we decided to clone a pair of gRNAs into a single expression vector driven by double human U6 promoters, thereafter, named 2U6 (Fig. 2a). Apparently, the proportion of cells expressing both gRNAs ($mCherry^+$) by 2U6 was remarkably higher than the proportion of cells expressing both sg-L and sg-R ($GFP^+/mCherry^+$) in the original dual vectors, indicating that it will be easier to target the site of interest in the same cell using 2U6 vector (Fig. 2b). Compared to the original strategy, 2U6-driven gRNAs induced comparable or slightly higher editing efficiencies for both BE4max and ABEmax (Fig. 2c-f; Figs. S2a-b). Interestingly, for most tested sites, the percentages of sequencing reads containing concurrent C&G-to-T&A or A&T-to-G&C conversions in 2U6 groups were significantly higher than that in dual vector groups (Fig. 2g-h), while there was no significant increase of indel rates (Fig. 2i-j). Thus, we successfully developed a dual gRNAs strategy to achieve concurrent C&G-to-T&A or A&T-to-G&C conversions in human genome.

3.2. A modified dual gRNAs system to improve concurrent editing efficiency on the same allele

We assumed that upon the guidance of one of the paired gRNAs, the original target sequence would be base-converted by BE4max or ABEmax, and that the intermediate product would not exactly match the other gRNA, which hinders the concurrent editing on the same allele. Therefore, we defined the gRNA with higher editing efficiency as sg-L and sg-R was re-designed (sg-R mutant) to perfectly match the sequence predicted from the base conversion by sg-L (Figs. 3a and 4a). Accordingly, sg-L and sg-R mutants were constructed into 2U6 vector, named 2U6mut (Fig. 3b). Although no discernible elevation of target editing efficiency was observed on specific sites in total for both BE4max and ABEmax (Figs. 3c and 4b; Figs. S3a-b), the proportion of sequencing reads containing concurrent C&G-to-T&A or A&T-to-G&C conversions was significantly increased by using 2U6mut gRNAs (Figs. 3d and 4c), while there was a very weak effect on generating of indels (Figs. 3e and 4d). Taking Site-6 for example, BE4max and 2U6mut induced 31% of sequencing reads containing concurrent C&G-to-T&A on the same allele, while only < 1% of sequencing reads with simultaneous editing were recovered from the 2U6/BE4max group (Fig. 3 f). Consistently, a similar advantage for 2U6mut was also observed for concurrent A&T-to-G&C conversions induced by ABEmax (Fig. 4e). To test the universality of the modified dual gRNAs system in other cell types, 2U6 or 2U6mut gRNAs were co-transfected with BE4max or ABEmax base editors into Hela cells, a cervical cancer cell line. The results demonstrated that concurrent C&G-to-T&A and A&T-to-G&C conversions were induced with satisfactory efficiency by BE4max and ABEmax respectively (Fig. 4 f; Fig. S3c), especially for C&G-to-T&A simultaneous mutations, which can reach up to 80% with low indel rates (Fig. S3d). Therefore, we successfully engineered an optimized dual gRNAs system to improve concurrent editing efficiency on the same allele.

3.3. The targeting scope of dual gRNAs strategy is expanded by using SpRY-mediated base editors

According to our design, the target sequence must be flanked by "CCN" and "NGG" in the dual gRNAs system. Among all CBE- and ABE-targetable sites within human genome, only 3.44% and 3.74% of them can be potentially subjected to concurrent C&G-to-T&A or A&T-to-G&C editing respectively (Fig. S4). Thus, PAMless SpRY Cas9n was fused to cytosine and adenine base editors, named SpRY-BE4max

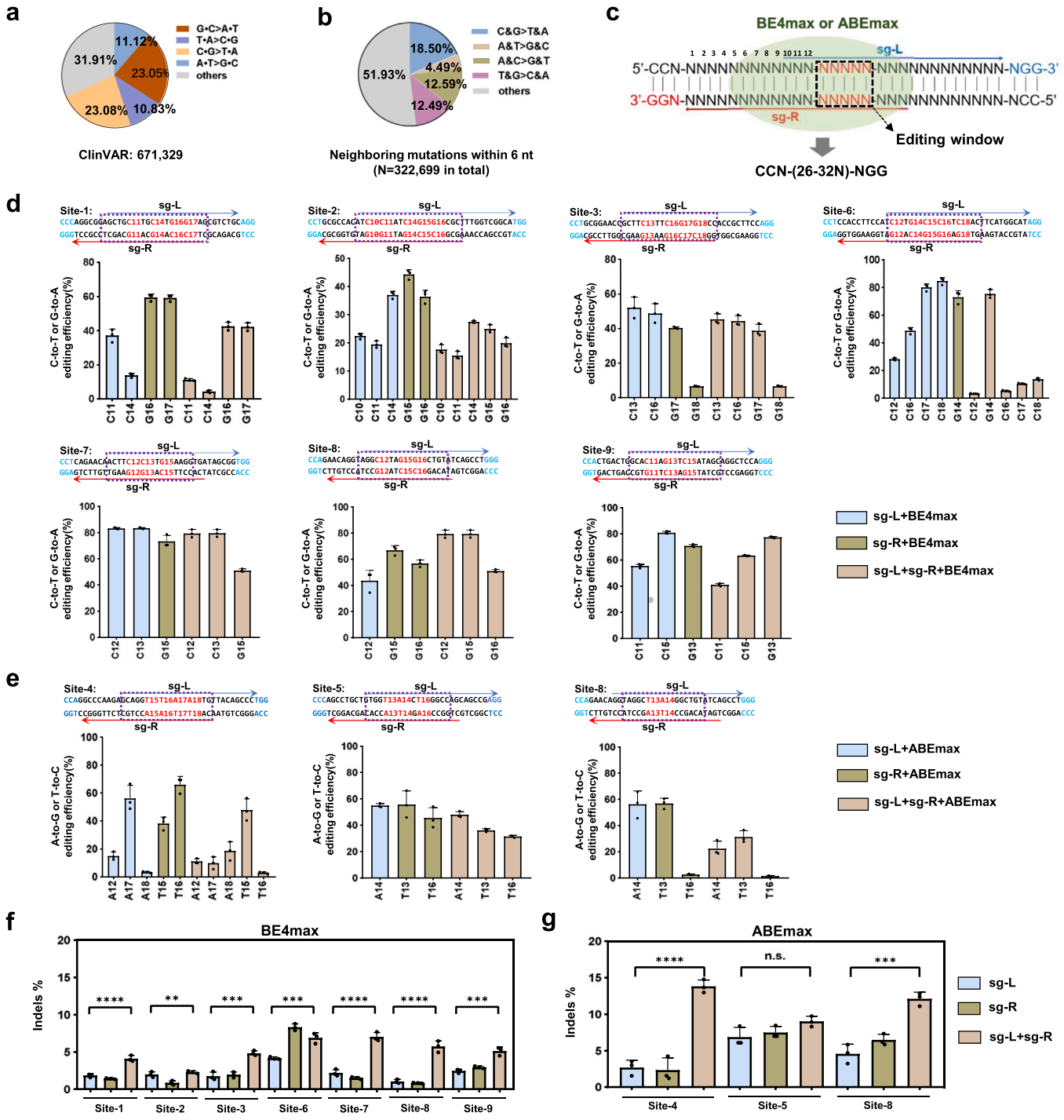

Fig. 1. The dual gRNAs strategy induces simultaneous C&G-to-T&A or A&T-to-G&C conversions. (**a**)Probability of point mutations in human genetic diseases (based on ClinVAR database). 68% of them are C-to-T and A-to-G (G-to-A and T-to-C on the opposite strand) mutations. (**b**)Probability of neighboring dual base mutations within 6-nt (raw data from gnomAD database). Four different types of dual base conversions (C&G, A&T, A&C, T&G) are indicated. (**c**) Schematic diagram of dual gRNAs editing for target sites. The sequence in the black dot box is the gRNA targeting window. The blue and red arrows represent the sg-L and sg-R sequences targeting the sense and the antisense chain, respectively. (**d**) The C-to-T or G-to-A conversion rates by single gRNA, as well as C-to-T & G-to-A dual conversions by dual gRNAs with BE4max at indicated target DNA sites in HEK293T cells. Error bars represent standard errors from three independent experiments. The edited cytosine and guanine conversions are highlighted in red. The assumed editing windows were labeled with dashed purple rectangles. (**e**) The A-to-G or T-to-C editing efficiency by single gRNA, and A-to-G & T-to-C dual-conversion efficiency by dual gRNAs with ABEmax at indicated target sites in HEK293T cells. The edited adenine and thymine are highlighted in red. The assumed editing windows were labeled with dashed purple rectangles. (**f**) Comparison of frequencies of indels generated by BE4max and ABEmax with single gRNA or double gRNAs guidance. *P*-value was calculated with unpaired, two-tailed Student's *t* test. $^{**}P < 0.01$, $^{***}P < 0.001$, $^{***}P < 0.001$, and n.s. denotes non-significant difference.

and SpRY-ABE8e (engineered from ABEmax, an upgraded version of ABE with higher editing efficiency [21]) (Fig. 5a), which can largely expand the potential targeting scope to cover up to 90% of human genome (Fig. 5b). According to the preferred PAM sequences "NRN" (R=G/A) for SpRY Cas9n, three target sites on *EMX1* were tested for SpRY-BE4max and SpRY-ABE8e respectively using 2U6 and 2U6mut gRNAs strategies. In these tested sites without "CCN" and "NGG", SpRY-BE4max and SpRY-ABE8e induced efficient C-to-T and G-to-A as well as A-to-G and T-to-C conversions within the targeting window, no matter combining with 2U6 or 2U6mut gRNAs (Fig. 5c-

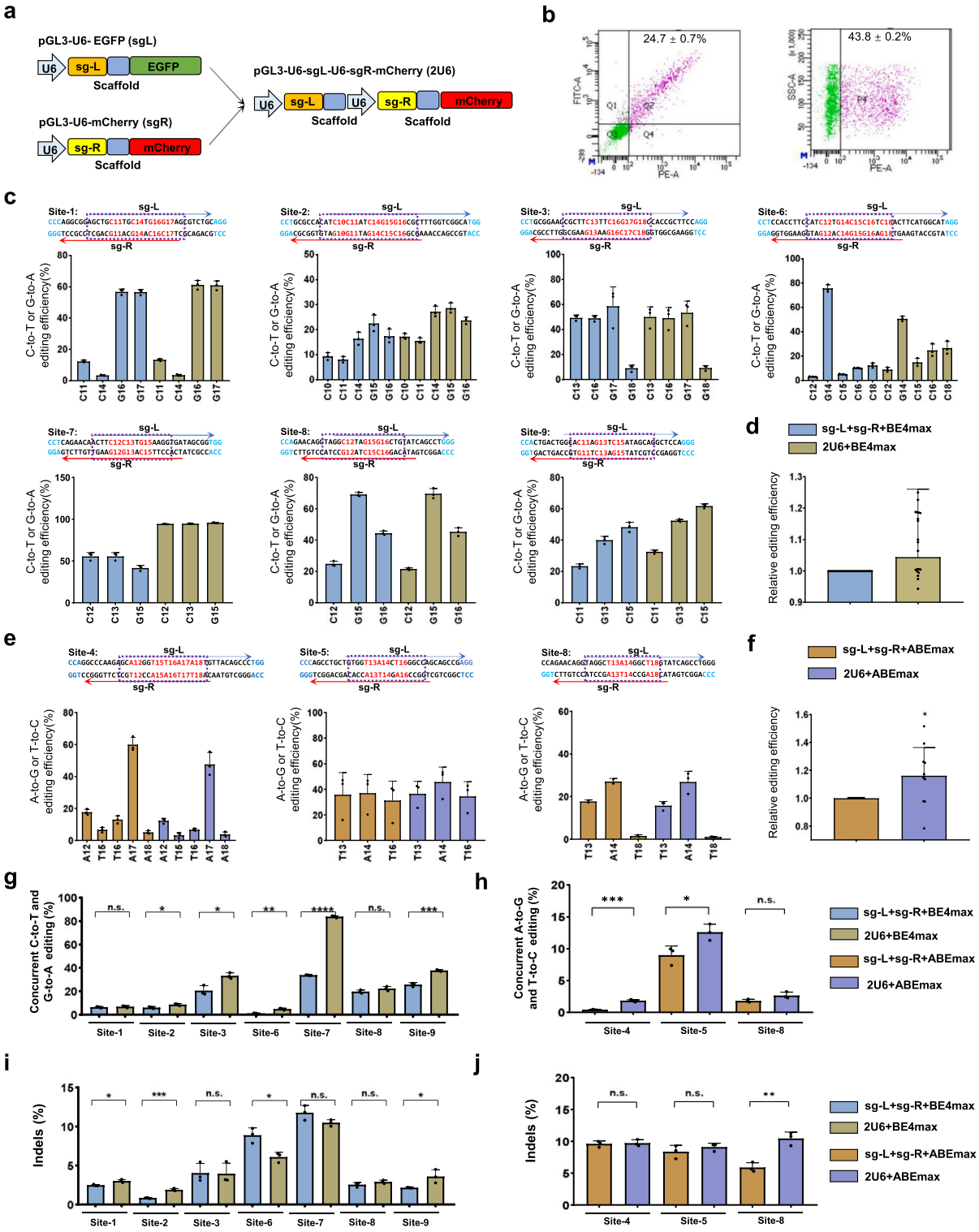

Fig. 2. Characterization of editing features induced by BE4max and ABEmax with two individual gRNAs and 2U6 dual gRNAs system. (**a**) Schematic diagram showing plasmid architectures of sg-L, sg-R, and 2U6 with two gRNAs in the same expression plasmid. (**b**) Comparison of the ratio of the positive cell population (expressing two gRNAs, purple dots) between two individual gRNAs co-transfection and 2U6 single transfection. (**c**) The C-to-T and G-to-A conversion rates induced by BE4max at each target sites were presented. The assumed editing windows were labeled with dashed purple rectangles. (**d**) The relative probability of C-to-T and G-to-A conversions at all target site were summarized. The mutation probability induced by two individual gRNAs (sg-L + sg-R) was normalized to "1". (**e**) The A-to-G or T-to-C conversion rates induced by ABEmax at each target sites were presented. The assumed editing windows were labeled with dashed purple rectangles. (**f**) The relative probability of A-to-G or T-to-C conversions at all target site were summarized. The mutation probability induced by two individual gRNAs (sg-L + sg-R) was normalized to "1". (**e-f**) The relative probability of A-to-G or T-to-C mutations by ABEmax at each target sites (e) or all target sites. The mutation probability induced by two individual gRNAs (sg-L + sg-R) was normalized to "1". (**g-h**) The efficiency of concurrent C&G-to-T&A (**g**) and A&T-to-G&C (**h**) dual conversions induced by dual two individual gRNAs and 2U6. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, and n.s. denotes non-significant difference. (i and j) The probability of indels induced by two individual gRNAs or 2U6 with BE4max (i) and ABEmax (j), Student's *t* test. $^{*}P < 0.1$, $^{**}P < 0.01$, $^{***}P < 0.001$.

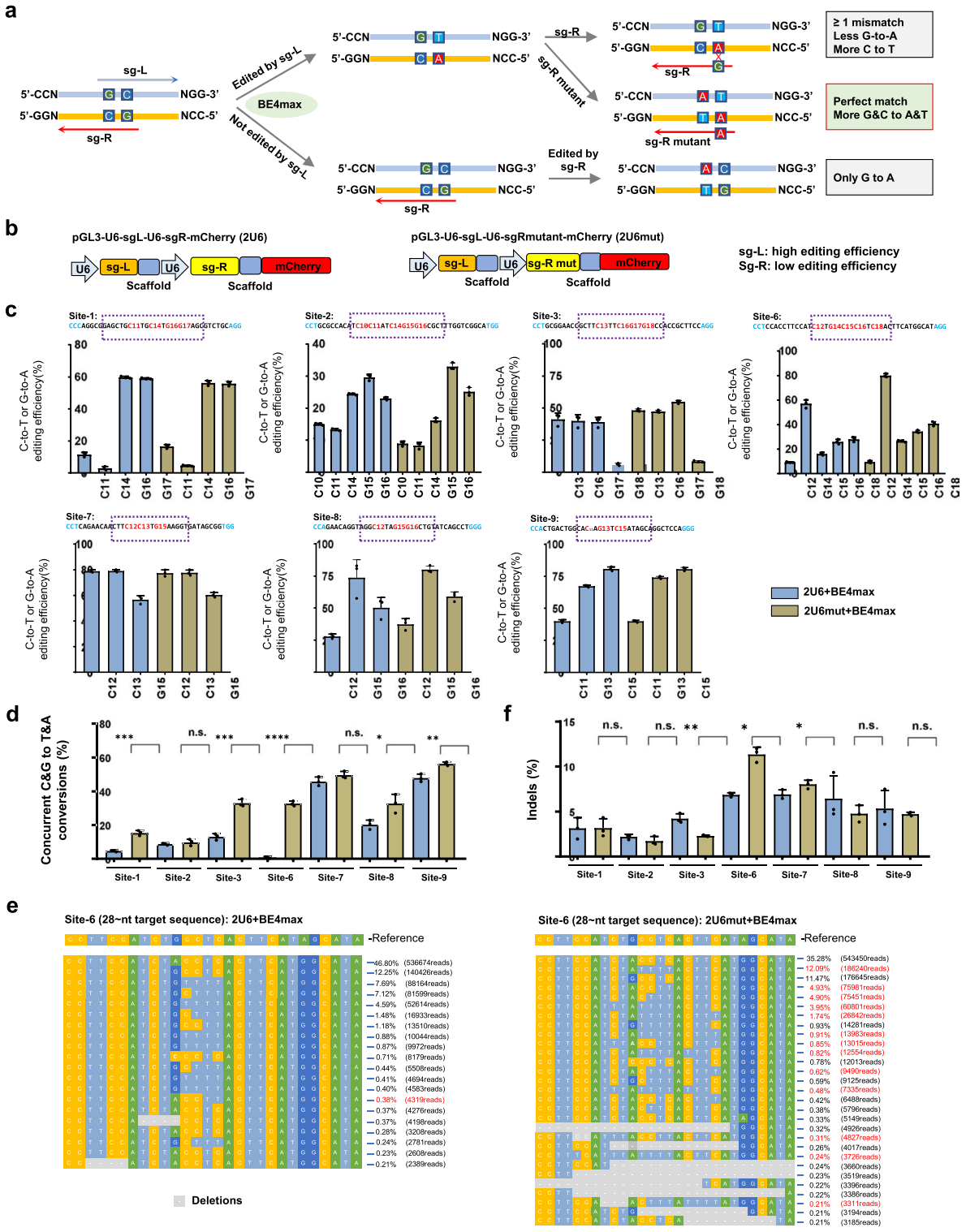

Fig. 3. Concurrent C&G-to-T&A editing efficiency on the same allele with a modified dual gRNAs system. (**a**) Schematic diagram showing the assumption of base editing efficiency with sg-L and sg-R or sg-R mutant. (**b**) Plasmid structure diagram of 2U6 containing sg-L and sg-R and 2U6mut containing sg-L and sg-R mutant. (**c**) The C-to-T or G-to-A conversion rates induced by BE4max with 2U6 or 2U6mut at indicated target sites in HEK293T cells. The PAM sequences are highlighted in blue, and the edited cytosine and guanine are highlighted in red. The assumed editing windows were labeled with dashed purple rectangles. (**d**) The concurrent C-to-T and G-to-A conversion efficiency on the same strand of DNA induced by BE4max with 2U6 or 2U6mut. $^{*}P < 0.1$, $^{**}P < 0.01$, $^{***}P < 0.001$. (**e**) The base composition of alleles with frequencies over 0.21% that result from BE4max induced on-target DNA editing with 2U6 or 2U6mut at site 6. The rows with red texts indicate the reads with concurrent C-to-T and G-to-A mutations. (**f**) Frequency of indels at each targeting site induced by BE4max with 2U6 (Blue bar) or 2U6mut (Green bar) in HEK293T cells.

d). We further analyzed the efficiency of concurrent C&G-to-T&A or A&T-to-G&C conversions at tested sites and found that the simultaneous C&G-to-T&A or A&T-to-G&C efficiency induced by 2U6mut was much higher than 2U6 gRNAs that exactly matched the original genome sequences (Fig. 5e-f), which was further verified by the proportions of reads with various base conversions (Fig. 5i-j). However, there was no significant difference between 2U6 and 2U6mut for induced indel rates (Fig. 5g-h). Thus, we propose a

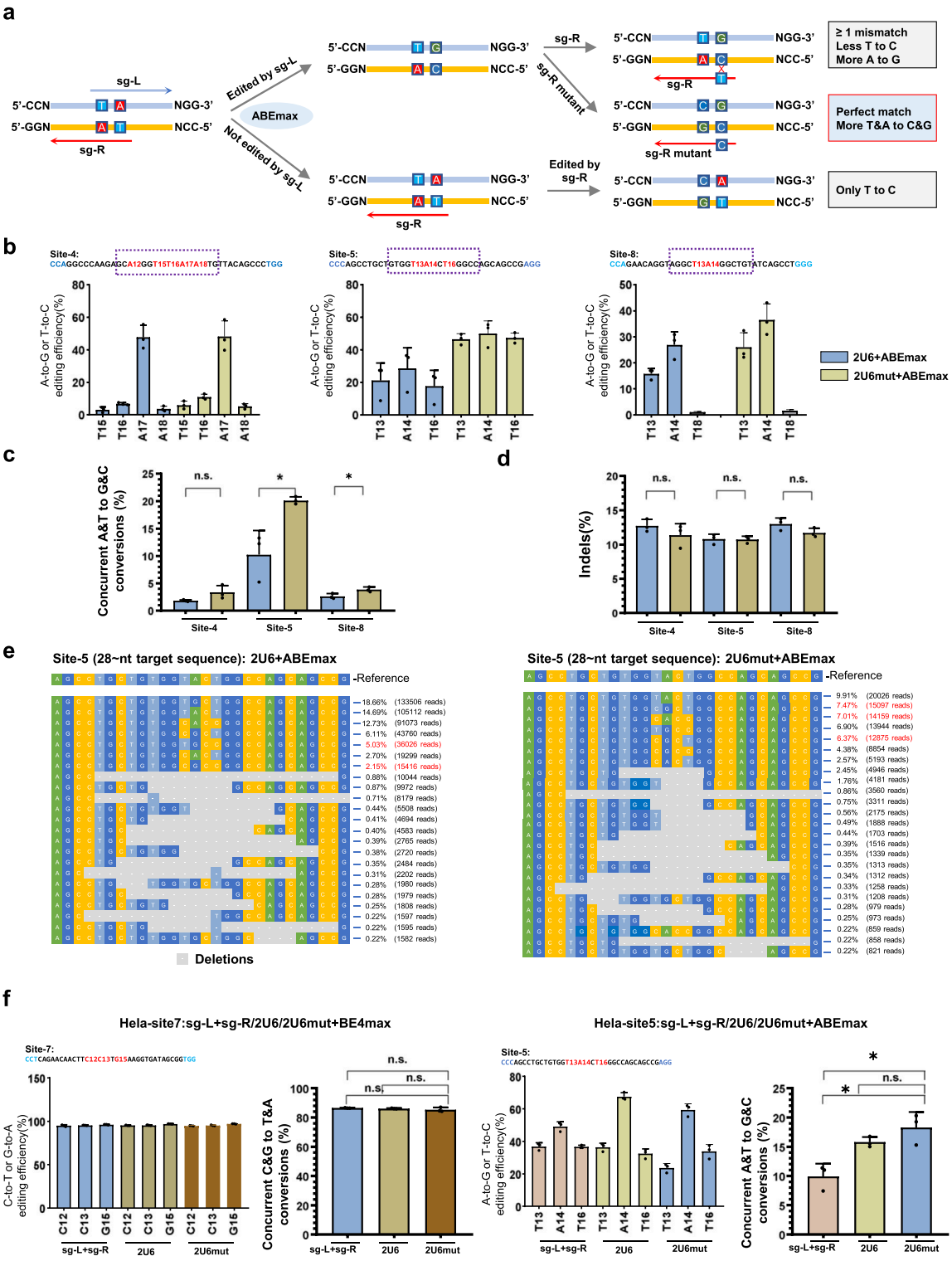

Fig. 4. Concurrent A&T-to-G&C editing efficiency on the same allele with the 2U6mut system. (**a**) Redefinition of sg-R to improve concurrent A-to-G and T-to-C editing efficiency of ABEmax. (**b**) The A-to-G or T-to-C conversion rates induced by ABEmax with 2U6 or modified 2U6mut at different target DNA sites in HEK293T cells. The PAM sequences are highlighted in blue, and the edited adenine and thymine are highlighted in red. The assumed editing windows were labeled with dashed purple rectangles. (**c**) The simultaneous A-to-G and T-to-C conversion efficiency on the same strand of DNA induced by ABEmax with 2U6 or 2U6mut. $^{*}P < 0.1$; n.s. denotes non-significant difference. (**d**) Indel rates at targeting sites induced by 2U6 (Blue bar) or 2U6mut (Green bar) with ABEmax in HEK293T cells. (**e**) The base composition of alleles with frequencies over 0.21% that result from ABEmax induced on-target DNA editing with 2U6 or 2U6mut at site 5. The rows with red texts indicate the reads with concurrent A&T-to-G&C conversions. (**f**) Concurrent C&G-to-T&A (left two histograms) or A&T-to-G&C (the right two histograms) editing efficiency on the same allele guided by the sg-L+sg-R/2U6/2U6mut systems in Hela cells. $^{*}P < 0.1$; n.s. denotes non-significant difference.

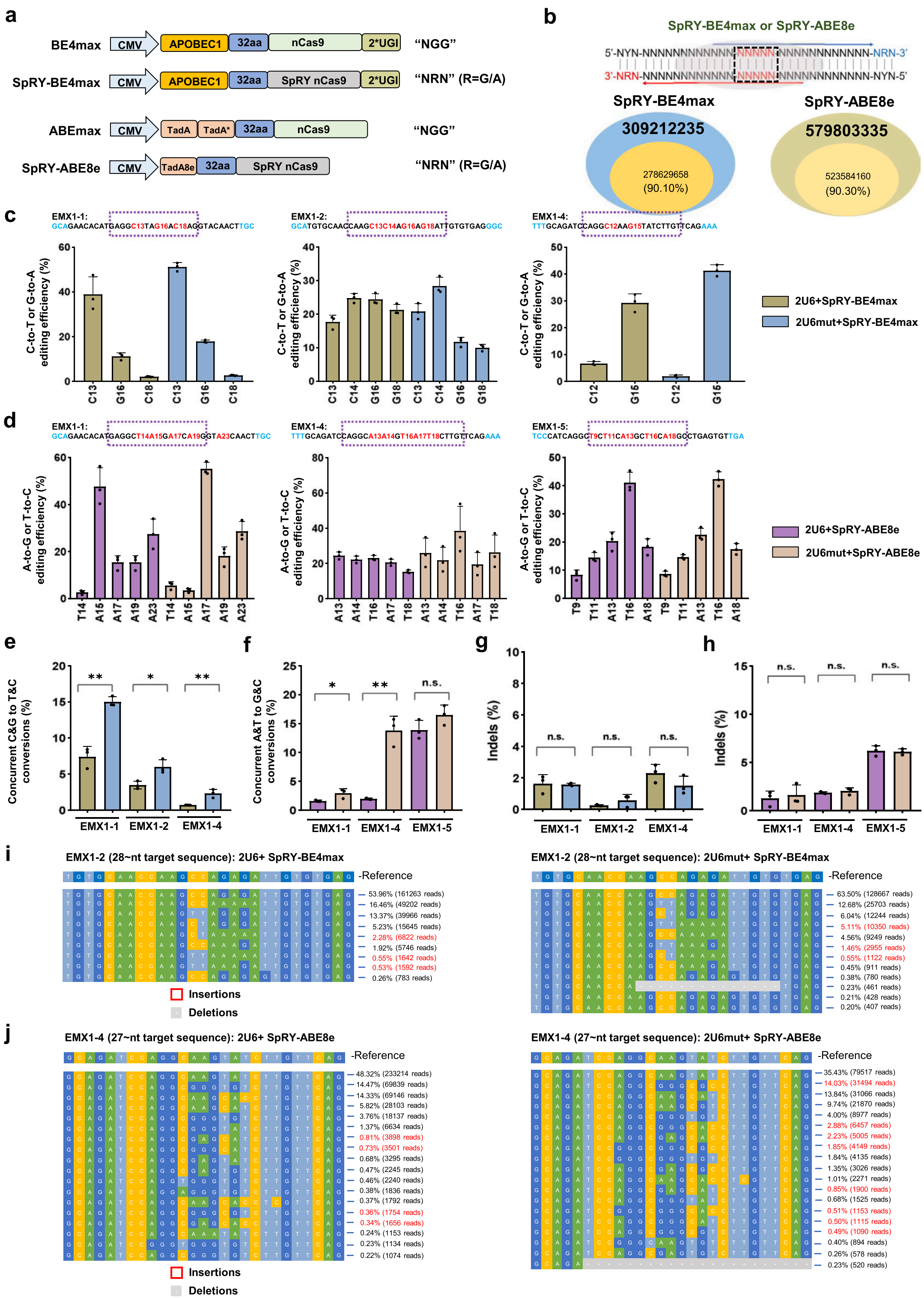

Fig. 5. The editing capability of SpRY-mediated base editors. (**a**) Schematic comparison of four base editors, including two kinds of PAMless SpRY-mediated base editors. The texts on the right represent PAMs for different base editors. (**b**) PAMless SpRY-mediated base editors were combined with dual gRNAs strategy. The entire blue and green area represents the proportion of C&G and A&T in adjacent bases of overlap region in the human genome, the yellow region represents the proportion of sequences in the human genome that conform to the editing characteristics of SpRY-mediated base editors. (**c-d**) The C-to-T or G-to-A (**c**) and A-to-G or T-to-C (**d**) conversion rates induced by SpRY-mediated base editors guided by 2U6 or 2U6mut at indicated target DNA sites in HEK293T cells. The DNA sequences represent the gRNA targets with PAM sequences highlighted in blue, and mutated bases are marked in red. The assumed editing windows were labeled with dashed purple rectangles. (**e-f**) The efficiency of concurrent C&G-to-T&A (**e**) and A&T-to- G&C(**f**) dual conversions by SpRY-BE4max and SpRY-ABE8e. (**g-h**) Indel rates at targeting sites induced by 2U6 or 2U6mut with SpRY-BE4max (**g**) and SpRY-ABE8e (**h**) in HEK293T cells. (**i-j**) The base composition of alleles with frequencies more than 0.21% induced SpRY-BE4max (**i**) and SpRY-ABE8e (**j**) at EMX1–2 (**i**) and EMX1–4 (**j**). The rows with red texts indicate the reads with concurrent C&G-to-T&A or A&T-to-G&C conversions.

feasible strategy to release PAM restrictions, expanding the targeting scope of our optimized dual gRNAs system in the human genome.

3.4. The dual gRNAs strategy combined with dual base editors enable more complex mutations

Concurrent base editing can induce concurrent base conversions and protein recoding, and dual base editors have been successfully applied in the rice plant engineering [14]. However, traditional dual base editors can only elicit C-to-T and A-to-G conversions, which might be not complex enough. We hypothesized that integrating our dual gRNAs strategy with dual base editors can catalyze more complicated mutation types. Then, four sites flanking with "CCN" and "NGG" as well as four "non-NGG" sites were tested using CABE-RY (Fig. 6a-b), a recently reported dual editor with wider editing window by fusing APOBEC3A and evolved TadA from ABE8e (simplified as TadA8e) with SpRY Cas9n [22]. As expected, relatively high targeting efficiencies were observed for C-to-T, A-to-G, G-to-A, and T-to-C conversions within a wide targeting window (10-to-23 nt starting from sg-L) (Fig. 6c). Targeted deep sequencing analysis demonstrated that sequencing reads with up to three different types of base conversions can be observed, although the percentage was relatively low; the sequencing reads with only one or two editing types accounted for the majority of collected reads (Fig. 6d-f). As exemplified in Fig. 6 g, there were 20.46% of single base conversions (C-to-T, A-to-G, G-to-A, and T-to-C), 29.1% of dual base conversions (C&A-to-T&G, G&T-to-A&C, C&G-to-T&A, A&T-to-G&C, and A&G-to-G &A), and only 0.25% of conversion with three types (G-to-A, T-to-C, and A-to-G) on site EMX1–1. It was worth noticing that the indel rates induced by dual gRNAs and dual base editors were frequently less than 5%, which was acceptable for biomedical and plant engineering, or saturation mutation-based screening (Fig. 6h-i). Therefore, we suppose that the combination of dual gRNAs strategy with dual-base editors can achieve more complex mutation types within a wide targetable window, validating the versatility of dual gRNAs system for inducing complicated mutation types.

3.5. Dual gRNAs strategy facilitates pathogenic mutations

In addition to the capacity for inducing complicated mutation types, dual gRNAs system might be capable of introducing MNVs in cell lines and animal models. *TP53* is one of the most frequently mutated genes in human cancers and its diverse mutation types and positions may be associated with the nature of mutagenic mechanisms involved in cancer etiology and other disease [23]. Among 2545 mutations in *TP53*, only 21% of them have been identified as pathogenic types (Fig. 7a-b), and the functions of the remaining 79% of mutations remain largely unknown. Particularly, some mutations have been identified as MNVs in neighboring regions [24,25]. Thus, we decided to introduce two MNVs, including C&G-to-T&A and A&T-to-G&C, into HeLa cells. To increase the concurrent mutation rates on the same allele, two pairs of 2U6mut gRNAs (sgTP53–2U6mut-1 and sgTP53–2U6mut-2) were constructed to induce L330P/Q331R and R280K/R281Q/R283Y mutations using SpRY-ABE8e and SpRY-BE4max respectively (Fig. 7c). The editing efficiencies for specific bases ranged from 38% to 70% for A-to-G and T-to-C conversions and 10–60% for C-to-T and G-to-A conversions (Fig. 7d-e).

Further analysis of the distribution of sequencing reads in sgTP53–2U6mut-1 demonstrated that as high as 39.91% of reads contained predicted base conversions to elicit L330P/Q331R mutations in *TP53*, while only 3.12% of sequencing reads possessed expected mutations without other bystander edits (Fig. S5a), indicating the necessity of optimizing the targeting accuracy of base editors in the future. To further interpret the functional significance of MNVs, single HeLa cell clones containing L330P/Q331R mutations, which are associated with Bap1 Tumor Predisposition Syndrome and Li-Fraumeni Syndrome [26], as collected in the human disease database Malacards [27], were successfully established from the sgTP53–2U6mut-1 and SpRY-ABE8e-cotransfected cells (Fig. 7e). In fact, we also tried to generate single clones containing R280K/R281Q/R283Y mutations only, but we failed to obtain it, because of its low efficiency with concurrent mutations on the same strand. Previously, p53 has been reported as an important regulator to mediate ferroptosis and cell death [28], we simply tested the effect of L330P/Q331R mutations within p53 on GPX4-IN-3 (a GPX4 inhibitor to induce ferroptosis) -induced cell death. As expected, L330P/Q331R mutation largely facilitated GPX4-IN-3-induced cell apoptosis in HeLa cells (Fig. 7 f; Fig. S5b).

To explore the potential mechanism underlying L330P/Q331R mutation-mediated cell death, two wild-type (WT) HeLa cell clones and two mutant clones were subjected to bulk RNA-seq analysis (Fig. 7 g). Surprisingly, we found that the expression of *TP53* and its target *CDKN1A* was remarkably decreased in mutant clones (Fig. 7 h). Considering the location of the L330P/Q331R neighboring a splicing donor site for exon 9, we proposed that L330P/Q331R mutations may disrupt the exon skipping rules and perturbate its splicing process, resulting in the downregulation of *TP53* transcription, which can explain the pro-apoptotic role of L330P/Q331R mutations upon GPX4-IN-3 stimulation (Fig. 7 f). Moreover, transcriptome analysis demonstrated that L330P/Q331R double mutation resulted in the downregulation of 1636 different genes and upregulation of 1537 genes (Fig. 7 g). These differentially expressed genes were mainly related to hormone metabolism, tissue development, cancer pathway, spliceosome, and transcriptional regulation (Fig. 7i-j). These data together highlight the importance to introduce pathogenic MNVs for dissecting their potential roles. Considering the well-known role of p53 in protecting cells against ferroptosis [29], it is reasonable to observe the pro-apoptotic effect of L330P/Q331R mutations here. In summary, we successfully create multiple pathogenic mutations in neighboring regions with our high-efficient dual gRNAs base editor system, which has the potential to establish disease models or to treat diseases caused by MNVs.

4. Discussion

SNVs are responsible for about half of the known disease-causing genetic variations, which can be precisely installed by cytosine or adenine base editors [20]. However, MNVs also account for a fraction of abnormalities as revealed in gnomAD database, especially for some neighboring mutations with simultaneous base conversions [3,10]. Simultaneous A&C conversions can be achieved by dual-mutation base editors, such as STEME, A&C-BEmax, SPACE, CABE, CABE-RY [11–14,22], whereas concurrent C&G and A&T conversions remain to be a challenge, at least within one step. In the present study, we combined our optimized dual guide RNA strategy with dual base editors recognizing "NGG" and "non-NGG" PAM sequences, and simultaneous C&G-to-T&A and A&T-to-G&C can be achieved, nearly without sequence restrictions, which was applied in saturation mutations and installation of pathogenic mutations.

As for dual base editors, dual C&G and A&T mutations can be stimulated by the delivery of two gRNAs separately or simultaneously. However, the proportion of desired cells and genomic products was relatively low (Figs. 1 and 2b). When concurrent base editing efficiencies were observed in Sanger sequencing results or calculated at single base, there was no significant difference between using co-transfections, 2U6, or 2U6mut strategies (Figs. 1–4). Nevertheless, when we further analyzed the distributions of deep sequencing reads, we revealed that the frequency of dual base conversions on the same DNA strand was much higher when using the 2U6mut strategy, containing a leading gRNA with higher targeting efficiency and a subsequent gRNA with mutations to match the products induced by the leading gRNA (Figs. 3–4). Roughly

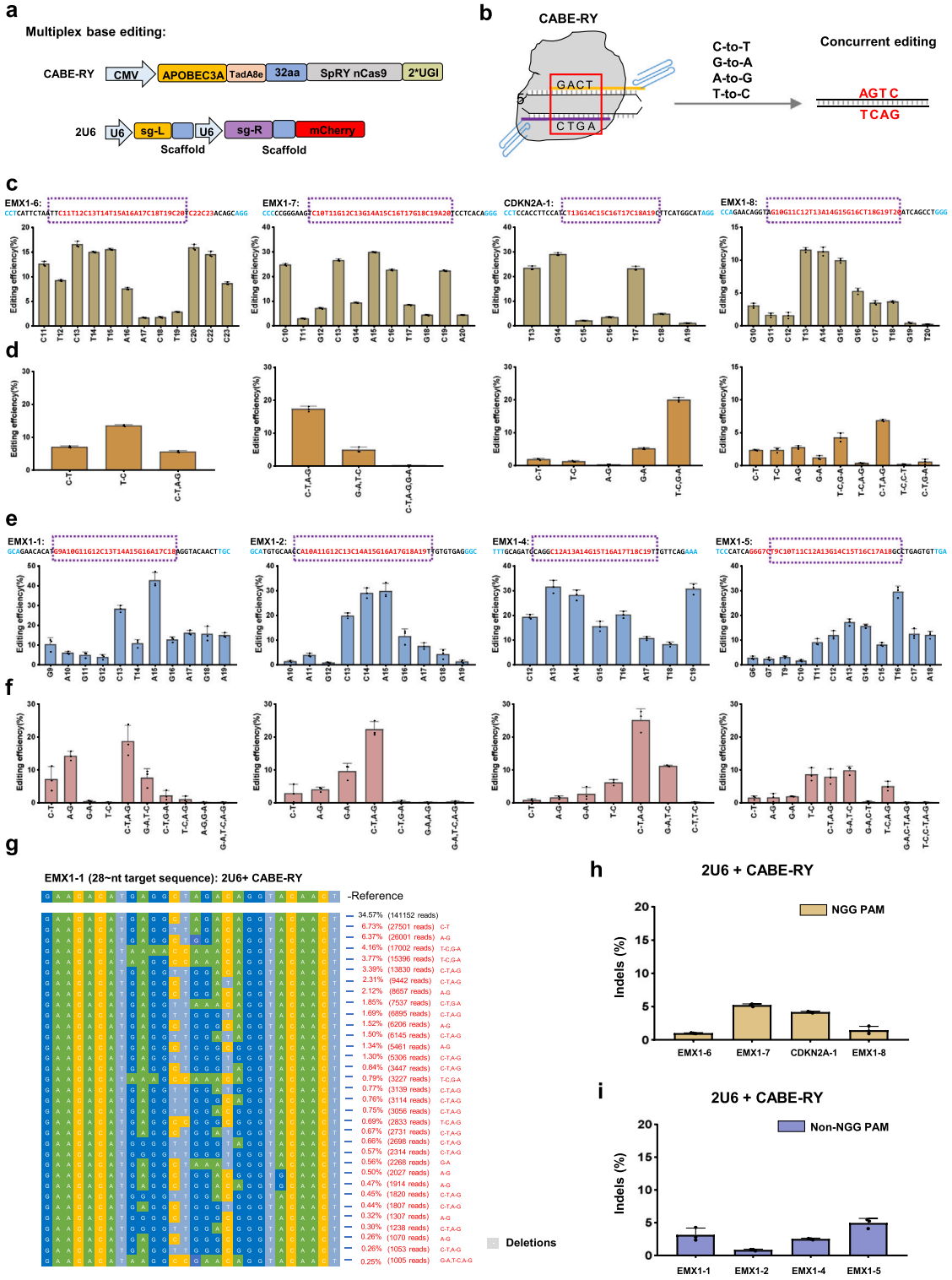

Fig. 6. Complex mutations with dual gRNAs strategy and dual base editor combinations. (**a**) Schematic diagram showing plasmid architectures of dual base editor CABE-RY and dual gRNA system 2U6. (**b**) Schematic diagram of different types of mutations produced by CABE-RY base editor with dual gRNAs strategy. (**c-d**) Summary of editing efficiencies and types induced by CABE-RY in HEK293T cells at four sites flanking with "CCN" and "NGG". The assumed editing windows were labeled with dashed purple rectangles. (**e-f**) Summary of editing efficiencies and types induced by CABE-RY in HEK293T cells at four sites flanking with "non-NGG". For analyzing the frequency of base conversions for different types in d and f, the sequencing reads with proportions of more than 0.21% were subjected to statistical analysis. The assumed editing windows were labeled with dashed purple rectangles. (**g**) The base composition of alleles with frequencies over 0.21% that result from CABE-RY-induced various on-target DNA editing at EMX1–1 site. The rows with red texts indicate the reads with concurrent C&G-to-T&A or A&T-to-G&C conversions. (**h-i**) Frequency of indels at targeting sites induced by CABE-RY in HEK293T cells at four sites flanking with "CCN" and "NGG" (**h**) and four sites flanking "non-NGG" (**i**).

comparing the induced concurrent base conversions, the C&G-to-T&A editing efficiency can reach 57% and was much higher than A&T-to-G&C mutations (< 15%) (Figs. 3–4). It seems that CBEs may cooperate better with 2U6mut gRNAs than ABEmax, and we assume that concurrent A&T-to-G&C mutations can be largely improved by using ABE8e. Compared with dual base editors, our proposed dual

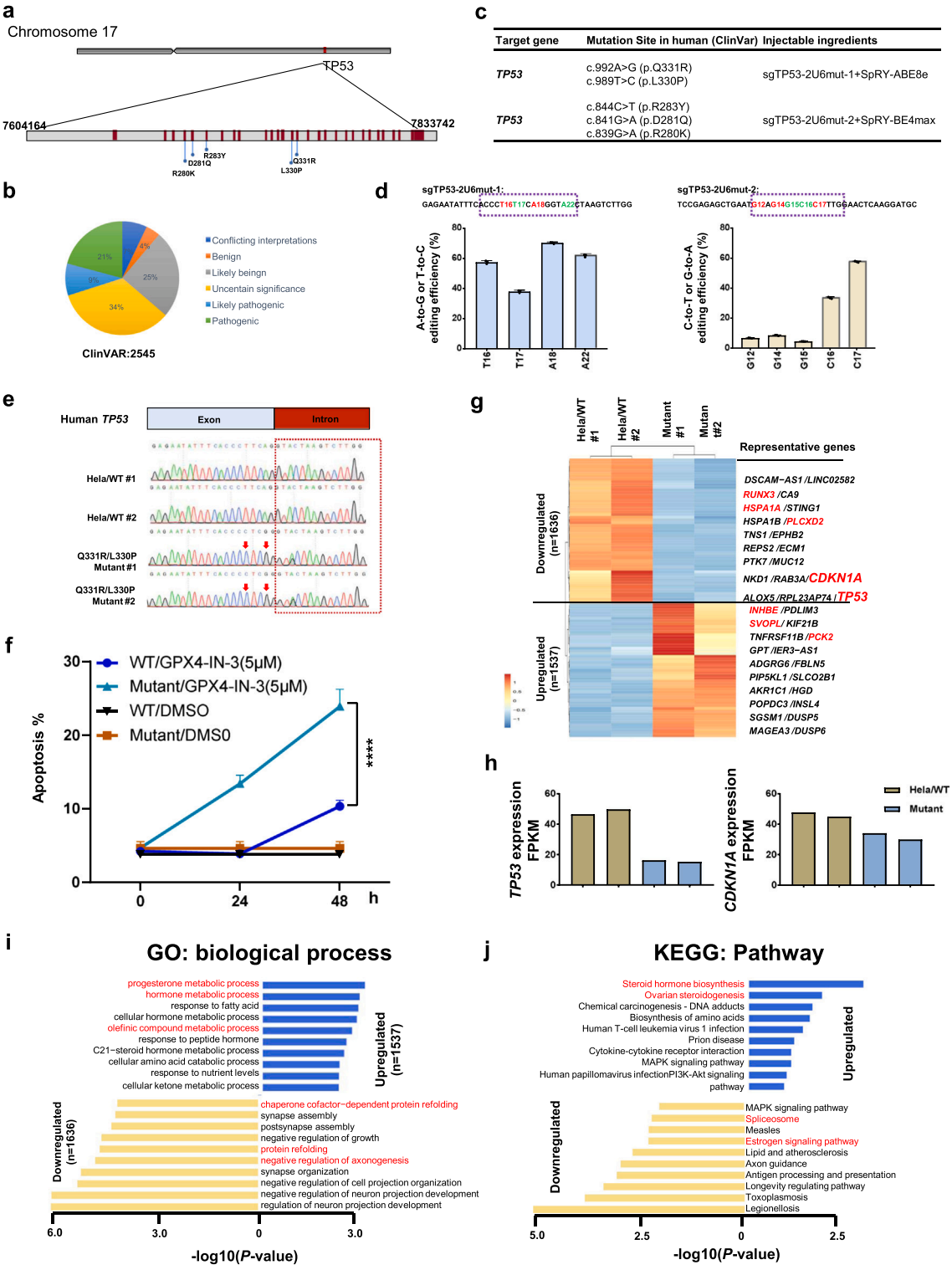

Fig. 7. Dissecting neighboring dual mutations within *TP53* with dual gRNAs system. (**a**) Schematic diagram showing *TP53* gene and the frequent mutation sites. (**b**) Summary of the currently known single nucleotide variations of *TP53* in ClinVAR disease database. (**c**) Summary of dual gRNAs and base editor information used in the present study to induce Q331R/L330P, and R283Y/D281Q/R280K mutations on *TP53* gene. (**d**) Targeted deep sequencing analysis of A-to-G or T-to-C conversions (light blue bar) as well as C-to-T or G-to-A conversions (light yellow bar) at the targeting sites. The red base represents the target mutation and the green represents the bystander edit. The assumed editing windows were labeled with dashed purple rectangles. (**e**) The chromatograms of Sanger sequencing showing examples of SpRY-ABE8e-induced simultaneous A-to-G and T-to-C conversions in Hela cells. (**f**) Q331R/L330P dual mutations promote apoptosis in HeLa cells. GPX4-IN-3 was used to induce cell apoptosis of Q331R/L330P mutant cells. Wild-type HeLa cells served as a control. (**g**) Heat map showing the differentially expressed genes in Q331R/L330P dual mutation cells compared with the wild-type HeLa cells. Representative differentially expressed genes were also presented. (**h**) *TP53* and *CDKN1A* expression levels in Q331R/L330P dual mutation cells and wild-type HeLa cells based on RNA-seq data. (**i**) GO analysis showing the enriched biological processes of the differentially expressed genes in Q331R/L330P mutant HeLa cells compared with wild-type cells. Representative differentially biological process were also presented. (**j**) KEGG analysis of the pathways enriched by the differentially expressed genes.

gRNAs strategy shows a smaller size and satisfiable efficiencies, with slightly higher but acceptable indel rates. It can act as a strong complement for inducing dual concurrent base conversions. Very similarly, twinPE also uses a single prime editor and two prime editing guide RNAs (pegRNAs) for the programmable replacement or excision of DNA sequences [30].

To further expand the targeting scope of dual gRNAs strategy, we used two SpRY Cas9-mediated base editors (SpRY-ABE8e and SpRY-BE4max) to achieve simultaneous base conversions at target sites that are not restricted by PAM sequences, broadening their applications (Fig. 5). Similarly, SpRY Cas9-mediated CABE-RY has also been applied for expanding the targeting scope of dual base editors [22]. Furthermore, dual gRNAs system has been successfully used for inducing complex mutations in human cells (Fig. 6). Although multiple mutation types were observed by combining dual gRNAs system with dual base editors, nearly no sequencing reads containing four types of mutations were identified, indicating that this combination strategy needs further optimization to obtain more saturated mutations, such as combining with some upgraded base editors [31]. Moreover, when comparing with dual base editors in the combinational editing, dual gRNAs system showed a bit recessive trait with lower concurrent C&G-to-T&A and A&T-to-G&C edits (Fig. 6). Previously, single-base editors have been applied for saturation screens [6,7], while the mutation types are largely limited. It is a promising strategy to acquire saturation screen using the combinations of dual gRNAs system and dual base editors. Importantly, we also propose a proof of concept to install pathogenic mutations into *TP53* efficiently in human cancer cells and roughly interpret the functional significance of two neighboring mutations within *TP53* by transcriptome analysis (Fig. 7), which can be broadened for disease modelling and functional dissection of MNVs *in vitro* and *in vivo*. Taken together, we propose a dual guide RNA strategy to acquire efficient concurrent C&G-to-T&A and A&T-to-G&C conversions and to expand its applications in saturation mutations and disease modeling, showing great potential for biomedical engineering and decoding disease-associated MNVs.

5. Conclusion

This work developed a dual guide RNA (gRNAs) system, which can stimulate efficient C&G-to-T&A or A&T-to-G&C conversion or multiple mutations at the same allele in the human genome by combining with canonical or PAMless Cas9-mediated single or dual base editors. Thus, we establish a general strategy to efficiently induce MNVs in human genome, helping to dissect the functions of pathogenic MNVs with multifarious types.

CRediT authorship contribution statement

Yuting Zhao (First Author): Conceptualization, Validation, Methodology, Software, Investigation, Visualization, Formal analysis, Writing – original draft; **Min Li (First Author):** Software, Data curation, Formal analysis, Visualization, Writing – original draft; **Jie Liu:** Conceptualization, Validation; **Xiaowen Xue:** Conceptualization, Validation; **Jingli Zhong:** Conceptualization, Validation; **Bo Ye (Corresponding Author):** Funding acquisition, Resources, Supervision; **Jun Chen (Corresponding Author):** Funding acquisition, Supervision, Writing – review & editing; **Yunbo Qiao (Corresponding Author):** Conceptualization, Methodology, Software, Supervision, Investigation, Formal analysis, Writing – review & editing.

Data availability

Targeted deep sequencing and RNA-seq data have been deposited into NCBI Sequence Reads Archive (SRA) database (SRA: PRJNA848093). The authors declare that all used plasmids, annotated DNA sequences, and other related data are all available from the authors upon request.

Conflict of interest

We declare that we have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported in part by Excellent Youth Foundation of Guangdong Scientific Committee (2020B1515020018; Y.Q.), National Key Research and Development Program (2018YFC1004700; Y.Q.), Shandong Excellent Young Scientists Fund Program (Overseas, 2022HWYQ-027; J.C.).

Author contributions

Y. Q., J.C., and B.Y. designed, conceived, and supervised the work. Y.Z., J.L. (Jie Liu) performed the experiments and data analysis. M.L. performed computational analysis. X.X., J.Z., and J.L. (Jianxiang Lin) helped with the experiments, data analysis, and discussion.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.csbj.2022.12.055.

References

[1] Matthijs G, Souche E, Alders M, Corveleyn A, Eck S, et al. Guidelines for diagnostic next-generation sequencing. Eur J Hum Genet 2016;24(1):2–5.
[2] Weiss MM, Van Der Zwaag B, Jongbloed JD, Vogel MJ, Bruggenwirth HT, et al. Best practice guidelines for the use of next-generation sequencing applications in genome diagnostics: a national collaborative study of Dutch genome diagnostic laboratories. Hum Mutat 2013;34(10):1313–21.
[3] Wang Q, Pierce-Hoffman E, Cummings BB, Alfoldi J, Francioli LC, et al. Landscape of multi-nucleotide variants in 125,748 human exomes and 15,708 genomes. Nat Commun 2020;11(1):2539.
[4] Gaudelli NM, Komor AC, Rees HA, Packer MS, Badran AH, et al. Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. Nature 2017;551(7681):464–71.
[5] Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature 2016;533(7603):420–4.
[6] Cuella-Martin R, Hayward SB, Fan X, Chen X, Huang JW, et al. Functional interrogation of DNA damage response variants with base editing screens. Cell 2021;184(4):1081–97. e1019.
[7] Hanna RE, Hegde M, Fagre CR, DeWeirdt PC, Sangree AK, et al. Massively parallel assessment of human variants with base editor screens. Cell 2021;184(4):1064–80. e1020.
[8] Kweon J, Jang AH, Shin HR, See JE, Lee W, et al. A CRISPR-based base-editing screen for the functional assessment of BRCA1 variants. Oncogene 2020;39(1):30–5.
[9] Koblan LW, Doman JL, Wilson C, Levy JM, Tay T, et al. Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. Nat Biotechnol 2018;36(9):843–6.
[10] Priestley P, Baber J, Priestley MP, Steeghs N, De Bruijn E, et al. Pan-cancer whole-genome analyses of metastatic solid tumours. Nature 2019;575(7781):210–6.
[11] Zhang X, Zhu B, Chen L, Xie L, Yu W, et al. Dual base editor catalyzes both cytosine and adenine base conversions in human cells. Nat Biotechnol 2020;38(7):856–60.
[12] Grunewald J, Zhou R, Lareau CA, Garcia SP, Iyer S, et al. A dual-deaminase CRISPR base editor enables concurrent adenine and cytosine editing. Nat Biotechnol 2020;38(7):861–4.
[13] Xie J, Huang X, Wang X, Gou S, Liang Y, et al. ACBE, a new base editor for simultaneous C-to-T and A-to-G substitutions in mammalian systems. BMC Biol 2020;18(1):131.
[14] Li C, Zhang R, Meng X, Chen S, Zong Y, et al. Targeted, random mutagenesis of plant genes with dual cytosine and adenine base editors. Nat Biotechnol 2020;38(7):875–82.
[15] Walton RT, Christie KA, Whittaker MN, Kleinstiver BP. Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. Science 2020;368(6488):290–6.

[16] Clement K, Rees H, Canver MC, Gehrke JM, Farouni R, et al. CRISPResso2 provides accurate and rapid genome editing sequence analysis. Nat Biotechnol 2019;37(3):224–6.
[17] Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;29(1):15–21.
[18] Putri GH, Anders S, Pyl PT, Pimanda JE, Zanini F. Analysing high-throughput sequencing data in Python with HTSeq 2.0. Bioinformatics 2022;38(10):2943–5.
[19] Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15(12):550.
[20] Rees HA, Liu DR. Base editing: precision chemistry on the genome and transcriptome of living cells. Nat Rev Genet 2018;19(12):770–88.
[21] Richter MF, Zhao KT, Eton E, Lapinaite A, Newby GA, et al. Phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity. Nat Biotechnol 2020;38(7):883–91.
[22] Tao w, Liu Q, Huang S, Wang X, Qu S, et al. CABE-RY: A PAM-flexible dual-mutation base editor for reliable modeling of multi-nucleotide variants. Mol Ther Nucleic Acids 2021;26:114–21.
[23] Olivier M, Hollstein M, Hainaut P. TP53 mutations in human cancers: origins, consequences, and clinical use. Cold Spring Harb Perspect Biol 2010;2(1):a001008.
[24] Baugh EH, Ke H, Levine AJ, Bonneau RA, Chan CS. Why are there hotspot mutations in the TP53 gene in human cancers? Cell Death Differ 2018;25(1):154–60.
[25] Petitjean A, Achatz MI, Borresen-Dale AL, Hainaut P, Olivier M. TP53 mutations in human cancers: functional selection and impact on cancer prognosis and outcomes. Oncogene 2007;26(15):2157–65.
[26] Mai PL, Malkin D, Garber JE, Schiffman JD, Weitzel JN, et al. Li-Fraumeni syndrome: report of a clinical research workshop and creation of a research consortium. Cancer Genet 2012;205(10):479–87.
[27] Rappaport N, Twik M, Plaschkes I, Nudel R, Iny Stein T, et al. MalaCards: an amalgamated human disease compendium with diverse clinical and genetic annotation and structured search. Nucleic Acids Res 2017;45(D1):D877–87.
[28] Jiang L, Kon N, Li T, Wang SJ, Su T, et al. Ferroptosis as a p53-mediated activity during tumour suppression. Nature 2015;520(7545):57–62.
[29] Dai C, Gu W. p53 post-translational modification: deregulated in tumorigenesis. Trends Mol Med 2010;16(11):528–36.
[30] Anzalone AV, Gao XD, Podracky CJ, Nelson AT, Koblan LW, et al. Programmable deletion, replacement, integration and inversion of large DNA sequences with twin prime editing. Nat Biotechnol 2021.
[31] Li J, Yu W, Huang S, Wu S, Li L, et al. Structure-guided engineering of adenine base editor with minimized RNA off-targeting activity. Nat Commun 2021;12(1):2287.