# Small-molecule activators specific to adenine base editors through blocking the canonical TGF-β pathway

**Yudong Yang[1,2,3,†], Chi Zhang[1,2,3,†], Yixuan Song[1,2,3], Yawen Li[1,2,3], Pingping Li[1,2,3], Min Huang[4], Feilong Meng[4] and Mingliang Zhang** ⓘ[1,2,3,*]

[1]Department of Histoembryology, Genetics and Developmental Biology, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China, [2]Shanghai Key Laboratory of Reproductive Medicine, Shanghai 200025, China, [3]Key Laboratory of Cell Differentiation and Apoptosis of the Chinese Ministry of Education, Shanghai 200025, China and [4]State Key Laboratory of Molecular Biology, Shanghai Institute of Biochemistry and Cell Biology, Center for Excellence in Molecular Cell Science, Chinese Academy of Sciences, Shanghai 200031, China

## ABSTRACT

**Adenine base editors (ABEs) catalyze A-to-G conversions, offering therapeutic options to treat the major class of human pathogenic single nucleotide polymorphisms (SNPs). However, robust and precise editing at diverse genome loci remains challenging. Here, using high-throughput chemical screening, we identified and validated SB505124, a selective ALK5 inhibitor, as an ABE activator. Treating cells with SB505124 enhanced on-target editing at multiple genome loci, including epigenetically refractory regions, and showed little effect on off-target conversion on the genome. Furthermore, SB505124 facilitated the editing of disease-associated genes *in vitro* and *in vivo*. Intriguingly, SB505124 served as a specific activator by selectively promoting ABE activity. Mechanistically, SB505124 promotes ABE editing, at least in part, by enhancing ABE expression and modulating DNA repair-associated genes. Our findings reveal the role of the canonical transforming growth factor-β pathway in gene editing and equip ABEs with precise chemical control.**

## INTRODUCTION

The G·C to A·T point mutation is the largest class of human pathogenic single nucleotide polymorphisms (SNPs) (1). To treat such mutations, adenine base editors (ABEs) were developed by fusing TadA, a bacteria-derived adenosine deaminase, to impaired CRISPR/Cas9 [clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated peptide 9] to enable targeted A-to-G conversion on the genome, thus presenting great therapeutic potential (2). However, achieving robust and precise editing at diverse genome loci in a controlled manner remains challenging (3). First, the mechanism for regulating ABE activity *in vivo* is largely unknown. Unlike the APOBEC family of cytidine deaminases in cytosine base editors (CBEs), TadA is absent from mammalian cells in nature. How the intrinsic factors and/or pathways in mammalian cells modulate ABE has not been reported. A better understanding of the cellular determinants for ABEs can help facilitate its evolution and application. Second, the editing efficiency of ABEs is expected to be improved. Although the current ABE variants are capable of editing at diverse genome loci, efficiency varies greatly (4). Furthermore, efficiently correcting most disease-causing SNPs remains challenging, which is particularly pertinent if the SNPs are residing in or surrounded by refractory genome regions, such as the epigenetically repressive chromatin loci that were reported to strongly suppress the activity of base editors (5,6). Therefore, robust editing is highly desired. Third, precise editing with minimized off-target activity in the genome is still challenging. Through protein evolution, current ABE variants acquired higher catalytic activity. However, this process is concomitantly coupled with more off-target editing in the genome and transcriptome, which represents a major clinical concern (7,8). Therefore, developing effective tools to boost ABEs with substantial levels of on-target, but not off-target editing, is crucial.

Chemical approaches provide rapid and reversible control for modulating diverse biological processes and are a useful tool for exploring underlying mechanisms (9–14). In this study, we performed a high-throughput chemical screen using libraries of sufficient coverage and diversity, and identified potent ABE activators. Surprisingly, a large proportion of activators target cell surface receptors. Interestingly, a spectrum of small molecules falls into the category of transforming growth factor (TGF)-β inhibitors, which has not been previously reported in modulating ABEs. Indeed,

these TGF-β inhibitors function consistently in promoting ABE editing in a dose-dependent manner. Among these small molecules, SB505124, a selective ALK5 inhibitor (15), is the most potent. Treating cells with SB505124 promoted on-target editing at multiple genome loci, including epigenetically refractory regions, greatly expanding the research scope of ABEs. Furthermore, SB505124 enhanced correction in disease-associated genes *in vitro* and *in vivo*, thus offering strong therapeutic potential. Importantly, SB505124 showed little effect on off-target editing on the genome, strengthening its potential clinical application. Intriguingly, SB505124 served as a specific booster by selectively promoting ABE editing compared with CBEs and *Streptococcus pyogenes* Cas9 (SpCas9). Mechanistically, SB505124 promotes ABE editing, at least in part, through enhancing ABE expression and modulating DNA repair-associated genes. Following the discovery of potent and specific activators, our study equips ABEs with precise chemical control and, more importantly, identifies the critical role of the canonical TGF-β pathway in gene editing.

## MATERIALS AND METHODS

### Plasmid cloning

All polymerase chain reaction (PCR) primers and oligonucleotides used for plasmid cloning are listed in Supplementary Tables S2 and S3. To make the eGFP-Stop-mCherry (ESM) reporter plasmid, DNA fragments encoding enhanced green fluorescent protein (eGFP) and mCherry, as well as the linker sequence, were amplified by PCR or synthesized and cloned into the LTR-EF1a-WPRE-LTR vector. The ABE 8.20, ABE 8e and ABE 8e-NG sequences were derived from ABE 7.10 by PCR amplification and introducing relative mutations as previously reported (7,8), and cloned into the U6-sg-CMV-ABE-NLS-bGH-CMV-eGFP vector. The guide RNAs (gRNAs) were cloned by annealing two complementary oligonucleotides encoding the target sequences and cloned into the U6-sg-CMV-ABE-NLS-bGH-CMV-eGFP vector via the BbsI site (for ABEs and SpCas9) or U6-sgRNA-Puro vector via the BsaI site (for CBEs), respectively. The gene encoding nSaCas9 was amplified by PCR from the CMV-Flag-SaCas9-NLS-bGH plasmid and cloned into the CMV-Flag-NLS-bGH vector. The D10A mutation was introduced by site-directed mutagenesis. The genes encoding Plp1 (p. A242V) and TAU (p. A152T) were amplified from mouse or human genomic DNA, respectively. Mutations, including c.725C>T for Plp1 and c.454G> A for TAU, were introduced by site-directed mutagenesis.

### Cell culture and transfection

Human embryo kidney (HEK) 293T (ATCC CRL-3216), U2OS (ATCC HTB-96), U251 (ECACC: 09063001), N2a (ATCC CCL-131), H1299 (ATCC CRL-5803) and SW1116 (ATCC CCL-233) cells were maintained in Dulbecco's modified Eagle's medium (DMEM) with 10% fetal bovine serum (FBS). K562 (ATCC CCL-243) was maintained in RPMI 1640 medium with 10% FBS. Human embryonic stem cell line H1 was maintained in mTeSR culture

medium (STEMCELL) on Matrigel- (Corning) coated culture plates. All cell lines used in this study were maintained at 37°C with 5% $CO_2$.

For cell transfection, HEK293T, U2OS, N2a, H1299 or SW1116 cells were seeded as $1.4 \times 10^5$ cells per well into 24-well plates for transfection. On the second day when grown to ~80–90% confluence, cells were transfected using Lipofectamine 3000 (ThermoFisher) with 1 μg of plasmid(s). H1, K562 and U251 were transfected using the Neon Transfection System (Invitrogen).

### Generation of the ESM cell line

The lentiviruses carrying ESM reporter genes were collected and used to infect HEK293T cells. The cells were digested to single-cell suspension and diluted into 96-well plates at one cell per well. The positive single-cell-derived clones with proper integration and expression of reporter genes were validated by sequencing, imaging and fluorescence-activated cell sorting (FACS) analysis, and selected for further experiments.

### High-throughput chemical screening

The small molecules were pre-loaded into 384-well stock plates (ABgene, 0781) individually at a concentration of 10 mM. For screening, ESM cells were seeded into 6-well plates as $8 \times 10^5$ cells per well. Twenty-four hours after seeding, cells were transfected with plasmid encoding ABE 7.10 and the stop codon targeting gRNA. Twelve hours after transfection, ESM-7.10 cells were digested as a single-cell suspension and loaded into 384-well assay plates as 5000 cells per well. A 50 nl aliquot of solution from the chemical library was instantly added to the assay plates by Mosquito HTS to a final concentration of 10 μM. Forty-eight hours after compound treatment, the intensities of eGFP and mCherry were quantified by Operetta (PerkinElmer). All images were then checked manually to exclude false positives.

### Compound treatment and DNA sequencing assay

Cells were seeded on 24-well plates and transfected with a plasmid expressing the desired editing system. Twelve hours after transfection, cells were re-plated into 96-well plates as $3 \times 10^4$ cells per well in 100 μl of culture medium. The chemicals were added individually. Seventy-two hours after treatment, the genomic DNA from cells in each well was isolated, and the target sequences were amplified for sequencing analysis. Primers used for sequencing were listed in Supplementary Tables S4 and S5.

### Western blot

HEK293T cells were lysed in RIPA lysis buffer (Beyotime, Cat. No. P0013B) supplemented with phenylmethylsulfonyl fluoride (PMSF; Beyotime, Cat. No. ST506). Total proteins were resolved on 8% sodium dodecylsulfate–polyacrylamide gel electrophoresis (SDS–PAGE) gels and transferred onto polyvinylidene fluoride (PVDF) membranes (Millipore). Membranes were blocked in 5% non-fat

dry milk in Tris-buffered saline with Tween-20 (TBST) for at least 1 h at room temperature followed by incubating with anti-SpCas9 antibody (CST, Cat. No. 14697T, 1:1000) diluted in TBST overnight at 4°C. After washing with TBST five times, membranes were incubated with peroxidase-conjugated antibodies for 1 h at room temperature. After five washes with TBST, membranes were treated with enhanced chemiluminescence (ECL, Tanon, Cat. No. 180-5001) and visualized with the Chemiluminescence Imaging System (SAGECREATION). The signal was quantified by Image J (2.1.0) and normalized to glyceraldehyde phosphate dehydrogenase (GAPDH).

### Next-generation DNA sequencing and data analysis

Next-generation sequencing (NGS) DNA amplicon libraries were established using Phusion Plus DNA Polymerase (ThermoFisher) and target site primer containing an adaptor sequence (forward: 5′-TTCCCTACACGACGCTCTTCCGATCT-3′, reverse: 5′-AGTTCAGACGTGTGCTCTTCCGATCT-3′) at the 5′ end (Supplementary Table S5). The above products were then subjected to a second-round PCR using primers containing different index sequences. The resulting amplicon libraries were mixed with an equal amount for NGS sequencing on an Illumina HiSeq platform. For sequencing data analysis, the reference sequence was set from 10 bp upstream of the protospacer to 10 bp downstream of the protospacer adjacent motif (PAM) sequence. Raw data were analyzed with the SHM pipeline (16). In brief, the paired-end reads were de-multiplexed with the fastq-multx tool from ea-utils (https://github.com/ExpressionAnalysis/ea-utils). On-target editing efficiency, unwanted base editing efficiency, indel frequency and read counts were analyzed by BE-Analyzer (17) (http://www.rgenome.net/be-analyzer/#!). Base-editing values and off-target editing were reported as a percentage of the number of reads with cytosine or adenine mutagenesis over the total aligned reads. The indel frequency of SpCas9 was analyzed by Cas-Analyzer (18) (http://www.rgenome.net/cas-analyzer/#!).

### RNA off-target analysis

Cells were seeded on 24-well plates and transfected with a plasmid expressing ABE7.10 and gRNA targeting site 18. Twelve hours after transfection, cells were re-plated into 24-well plates at $1.4 \times 10^5$ cells per well. The chemicals were added individually. Seventy-two hours after treatment, total RNA was isolated by TRNzol Universal reagent (TIANGEN, Cat. No. DP424) according to the standard protocol, then cDNA was obtained by reverse transcription. NGS libraries were amplified by Phusion Plus DNA Polymerase and sequencing was performed using the method described above. Sanger sequencing amplicons were amplified by Phusion Plus DNA Polymerase. Primers used for sequencing were listed in Supplementary Tables S4 and S5. Raw data were analyzed with the SHM pipeline (16). RNA off-target efficiency was analyzed by BE-Analyzer (17). RNA off-target editing was reported as a percentage of the number of reads with adenine mutagenesis over the total aligned reads.

Sanger sequencing data were analyzed by EditR (1.0.10) (19) (https://moriaritylab.shinyapps.io/editr_v10/).

### Small interfering RNA (siRNA)-mediated gene knockdown

All siRNAs were ordered from Dharmacon. For the siRNA-mediated gene knockdown assay, ESM cells were seeded at $3 \times 10^4$ per well in 80 μl of culture medium in 96-well plates. siRNAs were transfected with RNAiMAX Transfection Reagent (ThermoFisher) at a final concentration of 20 nM. Forty-eight hours after siRNA transfection, 200 ng of episomal plasmids encoding ABE 7.10 and sgRNA were transfected using Lipofectamine 3000 reagent according to the standard protocol. ESM cells were then treated with the indicated chemicals for 72 h, and the genomic DNA was isolated for sequencing.

### Short hairpin RNA (shRNA)-mediated gene knockdown

All shRNA plasmids were ordered from ThermoFisher (Supplementary Table S6). For the shRNA-mediated gene knockdown, ESM, HEK293T, U2OS or H1299 cells were seeded at $1.4 \times 10^5$ per well in 24-well plates and transfected with 1 μg of plasmid encoding ABE or CBE and 1 μg of shRNA-encoding plasmid with Lipofectamine 3000 (ThermoFisher). K562, U251 or H1 cells were transfected using 1 μg of plasmid encoding ABE and 1 μg of shRNA-encoding plasmid using a Neon Transfection System. Forty-eight hours later, genomic DNA was isolated for PCR and sequencing.

### Mouse experiments

Mouse experiments were performed in accordance with protocols approved by the Department of Laboratory Animal Science, Shanghai Jiao Tong University School of Medicine. Mice were housed in the specific pathogen-free (SPF) facility with free feeding. Plasmids for hydrodynamic tail vein injection were injected into 8- to 10-week-old male mice via the tail vein within 5–8 s, and the injection volume was 10% of the body weight of the mice (30 μg of plasmids per mice). SB505124 was dissolved in the solvent [5% DMSO + 40% polyethylene glycol (PEG) 300 + 5% Tween-80 + 50% ddH$_2$O] at a concentration of 2 mg/ml. Twenty-four hours after plasmid injection (day 1), SB505124 was administrated intraperitoneally daily at 10 mg/kg for four consecutive days. On day 5, the mice were sacrificed and the eGFP-positive cells were isolated by FACS sorting. A total of $5 \times 10^5$ cells were lysed using 50 μl of lysis buffer with 2% protease K, followed by incubation at 55°C for 30 min and 95°C for 10 min. The genomic DNA was isolated and the targeted sequences were amplified for Sanger sequencing. Meanwhile, $3 \times 10^6$ cells were used for total RNA isolation and quantitative real-time PCR (qRT-PCR) analysis. Primers used for qRT-PCR were listed in Supplementary Table S3.

### RNA-seq and data analysis

Total RNA was extracted with TRIzol Reagent (Invitrogen) and genomic DNA was removed using DNase I (Takara). A

3 μg aliquot of total RNA was used for library preparation with the TruSeq RNA sample preparation Kit from Illumina (San Diego, CA, USA). Paired-end RNA-seq libraries were sequenced using the Illumina NovaSeq 6000 platform (2 × 150). The mapped reads of each sample used HTseq (0.9.1) statistics to compare the read count values (20), and the differentially expressed genes (DEGs) were analyzed by DESeq2 (1.34.0) (21).

### qRT-PCR

Total RNA from cells or animal tissues was isolated using the TRNzol Universal reagent (TIANGEN, Cat. No. DP424), and was reverse transcribed into cDNA with the Reverse Transcription kit (Vazyme, R111) according to the manufacturer's instructions. The cDNA was amplified using SYBR green reagent (Vazyme, Q311) with a real-time PCR instrument (Bio-Rad) and normalized to internal control *GAPDH* or *Actb*. All primers used for qRT-PCR were listed in Supplementary Table S3.

## RESULTS

### High-throughput chemical screening identifies potent ABE activators

To identify the small-molecule activators for ABEs, we constructed a reporter plasmid encoding fluorescent proteins eGFP and mCherry, which are separated by a linker sequence containing an in-frame stop codon upstream of a PAM (Figure 1A; Supplementary Data S1). The adenine in the stop codon was placed at position 5 to the far end of PAM and could be corrected by ABEs. This reporter plasmid was lentivirally delivered into HEK293T cells and the single clones with proper integration and stable expression of reporter genes were selected for further study (Supplementary Figure S1A–C, hereafter referred to as ESM cells). Ectopic expression of ABE 7.10 and a stop codon-targeting gRNA in the ESM cells successfully converted the premature stop codon into tryptophan (TAG to TGG), leading to 27.1% of cells highly expressing both eGFP and mCherry (Supplementary Figure S1D). Conversion efficiency was ∼41% using Sanger sequencing, confirming that ESM cells are a suitable tool for imaging-based high-throughput screening (Supplementary Figure S1E).

We initiated screening using a commercial chemical library of 7647 bioactive compounds and a homemade library of 111 chemicals (Supplementary Table S1). These libraries were composed of bioactive compounds with diversified privileged structures that exhibit 'drug-like' properties, providing sufficient coverage and diversity. To initiate screening, we transfected ESM cells with an episomal plasmid expressing both ABE 7.10 and a stop codon-targeting gRNA. After 12 h, cells were seeded into 384-well plates at 5000 cells per well, and small-molecule chemicals were individually added simultaneously at 10 μM. At 48 h post-chemical treatment, intensities of eGFP and mCherry were quantified via high-content imaging, and the eGFP/mCherry ratio following treatment of individual chemicals was normalized to DMSO control (Figure 1B). Ultimately, 63 small molecules showing a high ratio (>1.5-fold) and cell viability (>70%) were identified as candidates

for ABE activators, which were then validated by a second round of screening (Figure 1C). Interestingly, a large proportion of activators (25 out of 63 candidates) targeted cell surface receptors (CSRs) (Supplementary Figure S2A). Moreover, we found that five of the 25 CSR-targeting ABE activators, i.e. SB505124, LY3200882, SD 208, SB 431542 and RepSox, which are all classified as TGF-β pathway inhibitors, were ranked on top and repeatedly promoted ABE editing in a dose-dependent manner (Figure 1H; Supplementary Figures S2A and S3A–F), strongly supporting the regulatory role of the TGF-β pathway in ABE conversion. Specifically, SB505124, a selective inhibitor of ALK5 (15), promoted ABE editing most (Figure 1C–E). Briefly, SB505124 showed the best effect at 10 μM without obvious cytotoxicity (Figure 1F; Supplementary Figure S2B, C). At 72 h post-treatment, editing kinetics reached the stationary stage (Figure 1G). NGS results confirmed that SB505124 robustly enhanced ABE editing, and no obvious increase in indel frequencies or undesired A-to-non-G conversions were observed (Supplementary Figure S4A, B). Together, using a high-throughput chemical screen, we identified and validated SB505124 as a potent activator for ABE.

Previous studies report that histone deacetylase (HDAC) inhibitors, including Romidepsin, Ricolinostat and Nexturastat A, can promote editing activity of base editors (22,23). Our screening identified that BG45, an inhibitor targeting class I HDACs, enhanced ABE editing 1.52-fold (Figure 1C; Supplementary Figure S2A). We further compared the effects of SB505124 with that of reported HDAC inhibitors. Accordingly, we observed that the editing efficiency in ESM cells was induced by 2.3-fold by SB505124, 1.5-fold by Romidepsin, 1.8-fold by Ricolinostat and 1.9-fold by Nexturastat A, respectively (Supplementary Figure S4C–E). Among these small molecules, SB505124 was the most potent activator.

SB505124 is a selective inhibitor of the TGF-β pathway by blocking ALK5-mediated activation of downstream effectors SMAD2/3 (Supplementary Figure S3A, B). To exclude the possibility of off-target (OT), we genetically knocked down ALK5 with siRNAs. We observed a 1.2-fold editing enhancement, which phenocopied SB505124 (Supplementary Figure S3G, H). For further validation, we chose additional cell lines, including U2OS, U251 and human embryonic stem cell (hESC) line H1. Intriguingly, genetic depletion of ALK5 consistently led to increased ABE editing activity. This effect has been validated at six genome loci with 19 positions, on average 2.5-fold for HEK293T, 1.8-fold for ESM, 1.3-fold for U2OS, 1.6-fold for U251 and 3.0-fold for H1, supporting evidence that inhibiting ALK5 promotes ABE editing (Supplementary Figure S3K, L). Canonically, ALK5 performs its function via phosphorylation of downstream effectors SMAD2/3. To fully characterize whether SB505124 operates via this pathway, we treated cells with ITD-1, an inhibitor that potently blocks phosphorylation of SMAD2/3 proteins. Remarkably, ITD-1 dose-dependently enhanced A-to-G conversion by 1.8-fold (Supplementary Figure S3I). However, treating cells simultaneously with SB505124 and ITD-1 did not further enhance efficiency, indicating that these two small molecules operate via the canonical TGF-β pathway (Supplementary Figure S3J). These genetic and chemical analyses collectively
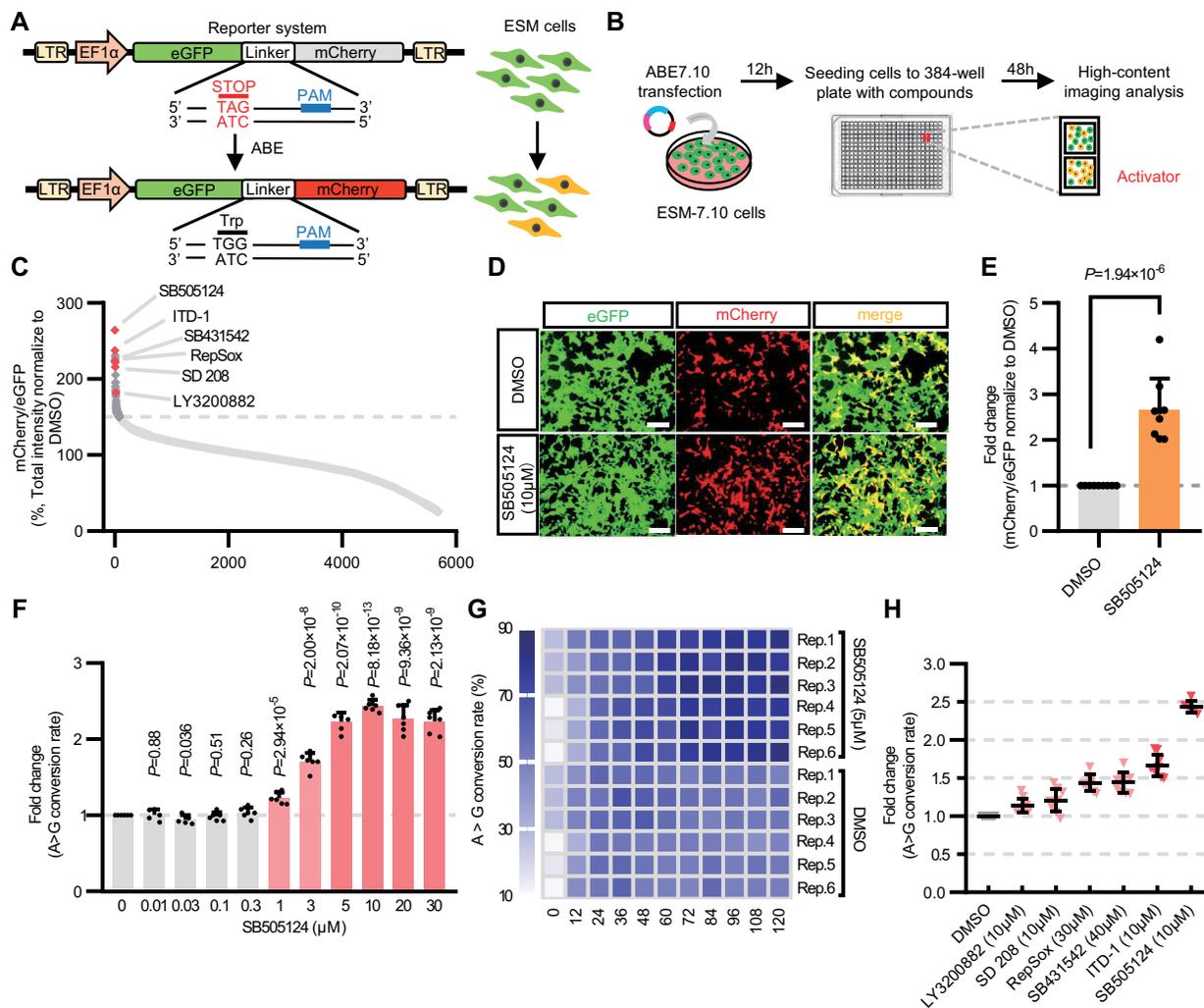
**Figure 1.** A high-throughput chemical screen to identify activators of ABE. (**A**, **B**) Schematic diagrams of the ESM reporter system and chemical screen. The reporter plasmid contains sequence encoding eGFP and mCherry in an open reading frame separated by a linker sequence with an in-frame stop codon upstream of a PAM. Ectopic expression of ABE converts the stop codon to tryptophan and leads to the expression of both eGFP and mCherry (**A**). ESM cells transfected by the ABE 7.10 system (ESM-7.10) were seeded into 384-well plates with individual compounds, and the intensities of eGFP and mCherry were quantified via high-content imaging 48 h later (**B**). (**C**) A waterfall plot shows the mCherry/eGFP ratio that was normalized to DMSO control (%) for each chemical. Chemicals exhibiting false-positive results or high cytotoxicity (>30% cell death) were removed from the diagram. Dots highlighted in red are compounds targeting the ALK5/SMAD2/3 pathway. (**D**, **E**) Representative images showing the expression of eGFP and mCherry in ESM-7.10 cells treated with DMSO or SB505124 for 48 h. Scale bar: 100 μm in (**D**). The intensities of eGFP and mCherry were quantified and the fold change of the ratio (eGFP/mCherry) for SB505124 was normalized to the DMSO control (E, $n = 9$). (**F**, **G**) Sanger sequencing results showing the efficiency of A-to-G conversion with the indicated concentrations of SB505124 48 h post-treatment (**F**, $n = 3 \times 3$) or at the indicated time points with 5 μM SB505124 (G, $n = 2 \times 3$). (**H**) Dot plots showing the efficiency of A-to-G conversion induced by the indicated small molecules that are normalized to the DMSO control ($n = 9$). Data are presented as the mean ± SD. *P*-values were determined using a two-tailed Student's *t*-test.

demonstrate SB505124 as a potent ABE activator through blocking of the canonical TGF-β pathway. As the TGF-β pathway has not previously been reported in modulating ABE editing and SB505124 is the most potent activator, we focused on SB505124 in this study.

## SB505124 promotes ABE to efficiently edit endogenous genome loci

To characterize whether SB505124 promotes ABE editing at endogenous genome loci, we transfected HEK293T cells with gRNAs targeting multiple sites across the genome (Figure 2A). These sites include those with target adenine at positions 4 to 7, the typical editing window of ABE, as well as adenines adjacent to the typical editing window that characteristically show lower editing potential. As a result, SB505124 induced two of seven sites by 20–40%, three of seven by 10–20% and two by 3–10% (Figure 2A, D; Supplementary Figure S5A). Within the typical editing window, the average editing efficiency was enhanced by 16.2, 12.1, 28 and 15.6% for positions 4 to 7, respectively (Figure 2D; Supplementary Figure S5A). Interestingly, for sites 18 and 19, editing efficiency was dramatically induced to 40–65%, equating to a 3.4-fold change (Figure 2M). Statis-
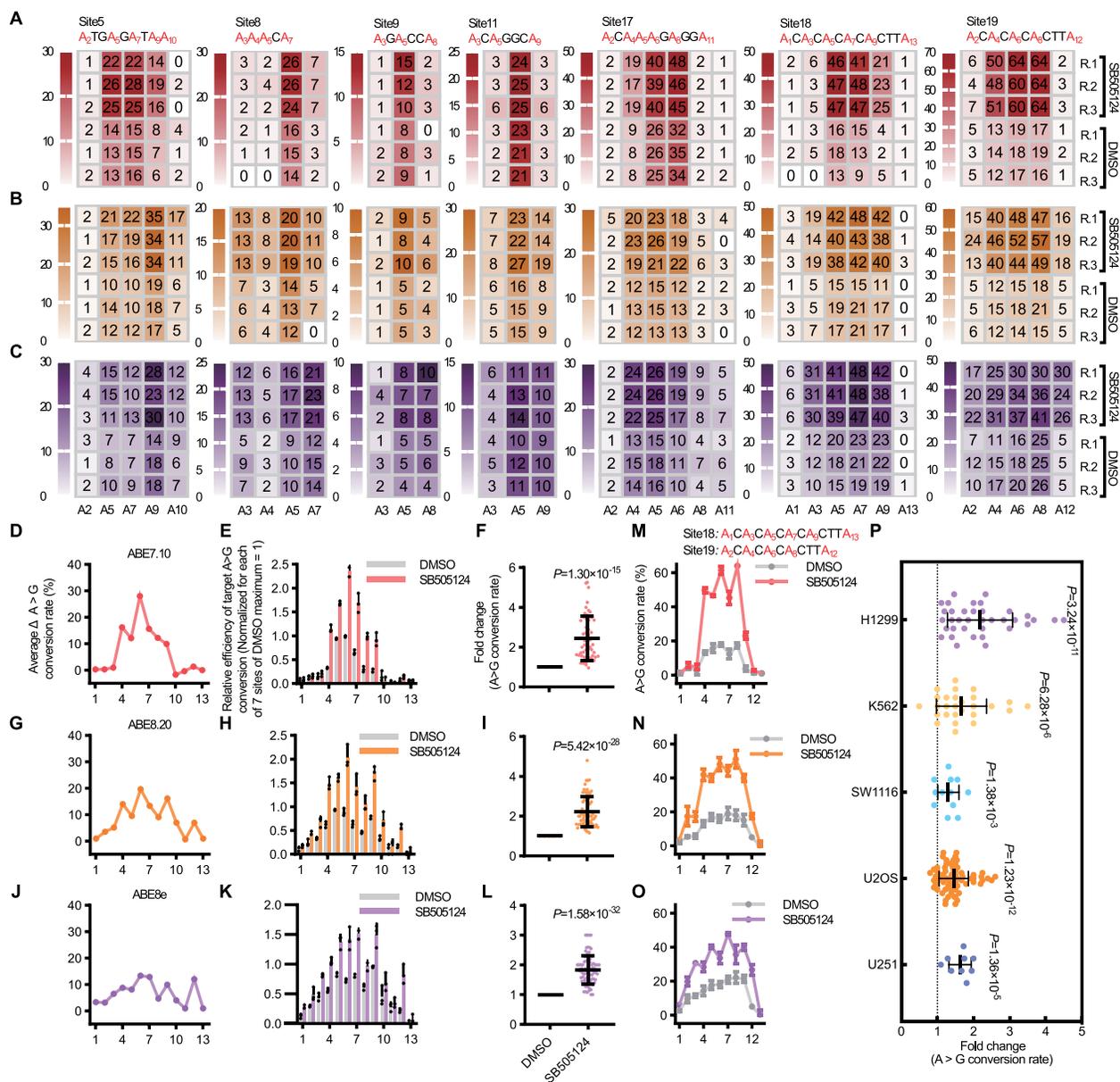
**Figure 2.** SB505124 promotes ABE editing at endogenous genome loci. (**A–C**) Heat maps showing conversion frequencies for each position of the indicated sites for ABE 7.10 (A), ABE 8.20 (B) and ABE 8e (C). Data are from three independent technical replicates. (**D–L**) Absolute changes (D, G, J) and fold changes of editing frequencies calculated for each position (E, H, K) and all sites tested (F, I, L) for ABE 7.10 (D, E, F), ABE 8.20 (G, H, I) and ABE 8e (J, K, L). Results for the SB505124 group with editing efficiency >5% were used for statistical analysis (F, I, L). Values were normalized to the maximum observed efficiency of the DMSO group at each position of the seven sites for each ABE variant. (**M–O**). Average editing frequencies at each position are shown for ABE 7.10 (M), ABE 8.20 (N) and ABE 8e (O). (**P**) A scatterplot showing the fold change of ABE conversion rates for all sites tested for the indicated cell line. Data are presented as the mean ± SD and are the results of three independent technical replicates. *P*-values were determined using a two-tailed Student's *t*-test.

tically, SB505124 promoted ABE 7.10 editing by ∼2.5-fold (Figure 2F). In addition, we examined whether SB505124 works on ABE 8.20 and ABE 8e, the other two commonly used ABEs (Figure 2B, C). Remarkably, SB505124 promoted editing at nearly all sites tested and average efficiency increased to 2.2-fold and 1.8-fold for ABE 8.20 and ABE 8e, respectively (Figure 2I, L; Supplementary Figure S5B, C). Among the three variants, ABE 8e was induced

least by SB505124 for its higher basal editing activity. To comprehensively characterize the function of SB505124 on ABE 8e, we chose five additional sites from different chromosomes, including the site proximal to the centromere, which is the condensed genome region considered difficult to edit. Sanger sequencing results confirmed that SB505124 induced ABE 8e editing for all 15 positions of five sites by an average of 1.56-fold (Supplementary Figure S5D–F). These

results demonstrate that SB505124 promotes editing activity for the three most used ABE variants.

To determine whether SB505124 modulates the ABE editing window, we chose sites 18 and 19, which cover every odd and even position, respectively, from position 1 to 13. Surprisingly, besides the target adenines within the typical editing window, we found significant editing on adenines at adjacent positions by SB505124, particularly for ABE 8.20 and ABE 8e. Briefly, SB505124 enhanced editing at positions 2 and 3, as well as positions 9 and 12 (Figure 2N, O). In sum, SB505124 dramatically promoted ABE editing across nearly all positions with substantial expansion of the editing window.

To further confirm our results, we analyzed the effect of SB505124 on ABE in five additional cell lines, i.e. H1299, K562, SW1116, U2OS and U251. As a result, we consistently observed 2.2-, 1.7-, 1.1-, 1.5- and 1.6-fold induction in these cell lines, respectively (Figure 2P; Supplementary Figure S6). These results collectively demonstrate that SB505124 has a general effect on ABE activation.

### SB505124 promotes editing at repressive genome loci

Epigenetically repressive chromatin regions, marked by high DNA methylation or repressive histone modification, are usually condensed, inaccessible and were reported to strongly suppress base editor activity (5). To examine whether SB505124 facilitates editing in these refractory regions, we first checked MSSK1, a highly methylated locus in HEK293T cells in nature (24). Adenine at position 5 is adjacent to methylated cytosines (∼75% and 100% for cytosines at positions 6 and 9, respectively). When targeted, the A-to-G conversion changed from 20.7% to 29.7% by SB505124, a 1.4-fold induction (Figure 3A, B). We further chose five additional genome loci with the methylated cytosine at different distances to the adenine to be edited (ranging from 3 to 15 nt) (24). Subsequently, conversion frequency was greatly promoted by SB505124 in all nine positions across five sites (Figure 3A, B). On average, ABE editing activity was induced 1.5-fold by SB505124 (Figure 3C, D).

We then examined whether SB505124 facilitates editing in transcriptionally repressive heterochromatin regions. We chose 35 positions marked by H3K9m3 modification within 16 heterochromatin regions from nine chromosomes. Meanwhile, we chose 19 positions within eight genome loci from eight chromosomes that were marked by H3K4m3 modification as controls for euchromatin regions (25) (Figure 3E, H). All genes located at these loci reportedly led to genetic disorders when mutated, and thus are considered promising targets for gene therapy. As a result, we found SB505124 remarkably enhanced editing by 1.49-fold in the heterochromatin regions and by 1.45-fold in euchromatin regions (Figure 3F, G, I, J). Together, these results support SB505124 promoting ABEs to efficiently edit repressive genome regions, thus expanding the research potential of ABEs.

### SB505124 promotes little off-target editing on the genome, but rather promotes off-target editing on the transcriptome

The evolution of ABEs has led to higher conversion efficiency with concomitantly more off-target editing. This effect is triggered by promiscuous ABE activity in genomic DNA or mRNA, which is considered a major clinical concern (7,8). To test whether SB505124 can promote gRNA-dependent off-target editing on the genome, we transfected HEK293T cells with gRNA targeting VEGFA3, a previously well-characterized genome locus for assessing the off-target deamination effect of ABEs (26). We analyzed the A-to-G conversion for the on-target locus, as well as the three most active OT loci. On-target editing was ∼26% and changed to 31% by SB505124 (Figure 4A). We detected a very low frequency of A-to-non-G conversion at the targeted site, which was comparable with the DMSO group (Figure 4B). Notably, NGS showed that editing frequencies at the three OT sites were not significantly different with or without SB505124 (Figure 4C, D). To more rigorously evaluate these results, we further transfected cells with gRNAs targeting HBG and EMX1 (26,27). We did not observe significant differences in editing frequency for OT sites with or without SB505124 (Supplementary Figure S7A, B). These results demonstrate that SB505124 does not promote gRNA-dependent off-target editing on the genome.

Next, we examined whether SB505124 induced gRNA-independent off-target editing on the genome. We co-transfected HEK293T cells with an *S. pyogenes*-based ABE system (nSp-ABE7.10 and gRNA targeting site 18), as well as a catalytically impaired *Staphylococcus aureus*-Cas9 system (nSaCas9 and gRNAs targeting R-loop sites 1–5). Expression of the nSaCas9 system generated R-loops at the targeted genome loci, which were the short stretch of single-stranded DNAs that served as substrates for Sp-based ABE (28) (Figure 4E). NGS results demonstrated no induction of A-to-G conversion on the five R-loops in the presence of SB505124 compared with the DMSO group (Figure 4F), suggesting that SB505124 has little effect on the off-target editing of ABE on the genome.

ABE expression can cause tens of thousands of A-to-I transitions in mRNA (29–31). To examine if SB505124 can induce mRNA off-target editing, we transfected HEK293T cells with gRNA targeting site 18 and assessed off-target deamination in previously validated hotspots (31) (Figure 4G–I). We detected 14.7% and 5.4% of A-to-I edits on the ABE_OF1 and TOPRS loci of mRNA, which were induced to 32.6% and 19.4%, respectively, by SB505124 (Figure 4G). No detectable editing for corresponding genomic DNA was observed (Figure 4H). NGS results further confirmed induction of off-target editing on mRNA at four other hotspots, ranging from 0.07% to 20.53% (Figure 4I). Collectively, SB505124 promotes little off-target editing on the genome, but rather promotes editing on the transcriptome.

### SB505124 serves as a specific activator of ABE

Given that SB505124 significantly induces ABE activity, it is crucial to know whether it modulates the activity of other Cas9-based gene editors. We chose Sp-Cas9 and CBEs as representatives for Cas9-based endonuclease and the base editor, respectively. We transfected HEK293T cells with SpCas9 and corresponding gRNAs targeting five genome loci, and analyzed the genome cut-
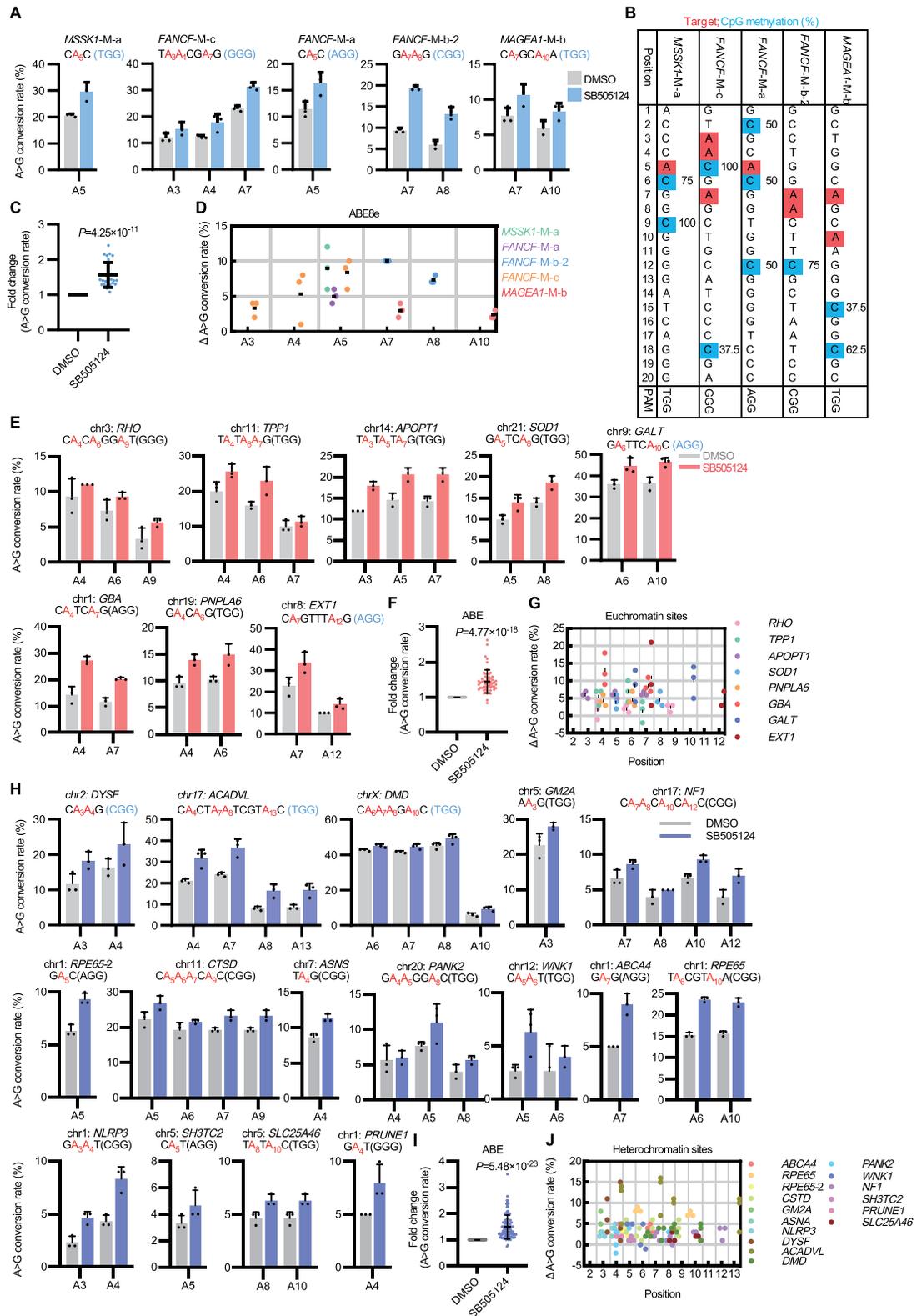
**Figure 3.** SB505124 promotes ABE editing at refractory genome loci. (**A–D**) Editing frequencies at natively methylated genome regions. Bar graphs showing the editing efficiencies at nine positions across five endogenously hypermethylated loci (A). Sequences with targeted adenines (in red), methylated cytosines (in blue) along with the methylation rate (indicated by number) and the PAM for each site (B). The fold changes (C) and absolute changes (D) for all positions tested are shown. (**E–J**) Bar graphs show the editing efficiencies of 35 positions across 16 heterochromatin loci (E) and 19 positions across eight euchromatin loci (H). Fold changes (F, I) and absolute changes (G, J) of the efficiencies for all sites tested are shown. Data are presented as the mean ± SD and are the results of three independent technical replicates. *P*-values were determined using a two-tailed Student's *t*-test.
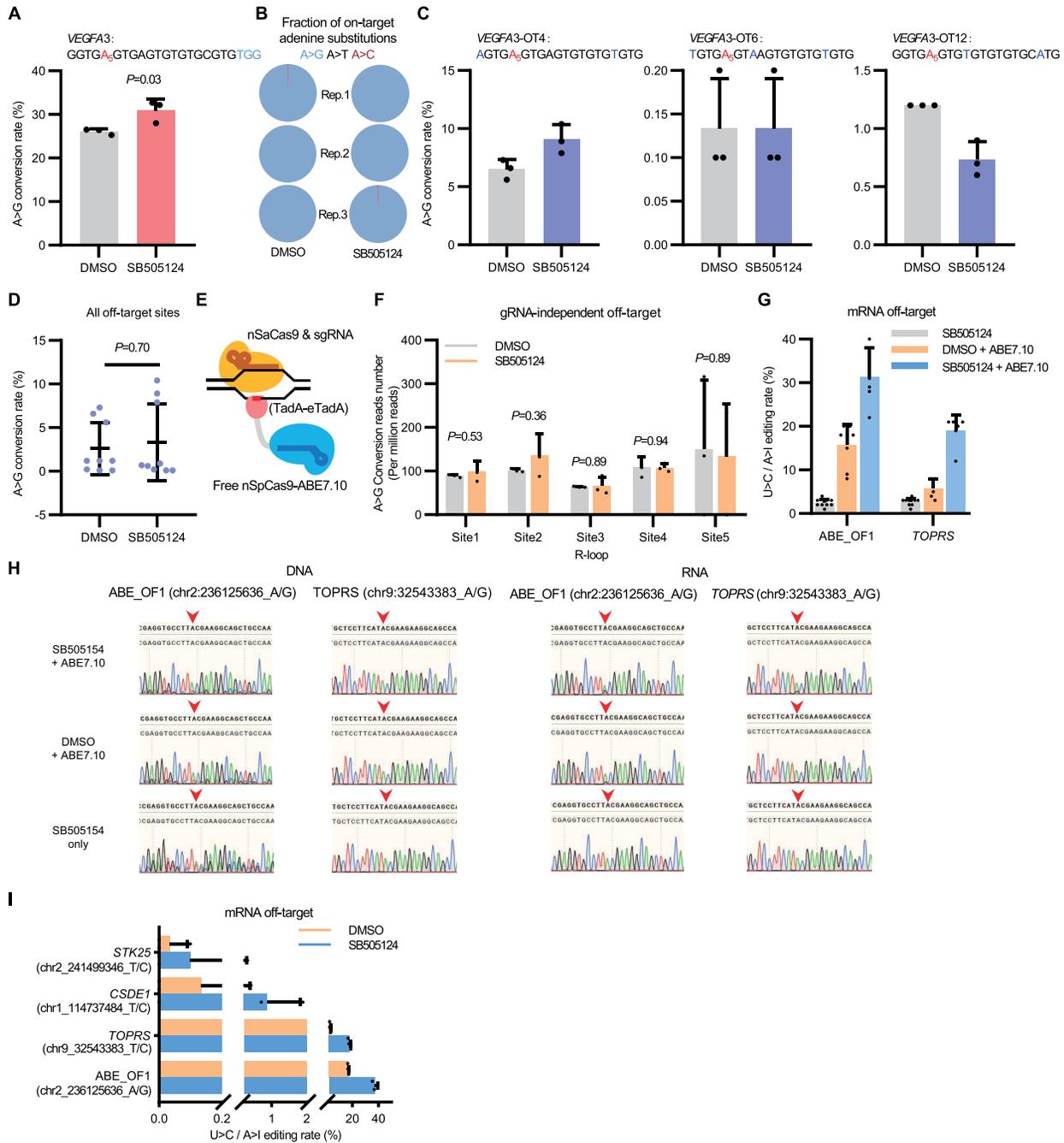
**Figure 4.** Off-target editing induced by SB505124. (**A–D**) Guide RNA-dependent DNA off-target editing was analyzed by NGS. The conversion frequencies of on-target editing at the indicated position (A), fraction of adenine substitutions (B) and off-target editing at three validated OT sites (C) for VEGFA3 are shown. The average off-target editing rate for all OT sites tested was then analyzed (D). (**E**) Schematic diagram of the R-loop assay to detect gRNA-independent DNA off-target editing. (**F**) Bar graphs showing the editing efficiencies of five R-loop regions. (**G, H**) Bar graphs showing the frequencies of A-to-I conversion rates at ABE_OF1 and TOPRS sites in mRNA transcripts (G). Sanger sequencing results show the editing on genomic loci (left panel) and their RNA transcripts (right panel) for ABE_OF1 and TOPRS sites (H). (**I**) NGS results showing the off-target editing frequencies of four mRNA transcripts by ABE 7.10 with or without SB505124. Data are presented as the mean ± SD and are the results of three technical replicates. *P*-values were determined using a two-tailed Student's *t*-test.

ting activity. For all sites tested, we detected no significant changes in editing rate at individual sites, and the average genome cutting efficiencies with SB505124 at those sites were statistically comparable with the DMSO control (Figure 5A, B). In addition, SB505124 did not induce unwanted editing, as the indel fraction showed no obvious changes (Supplementary Figure S7C), which was confirmed by NGS analysis (Figure 5C, D). These data collectively demonstrate that SB505124 does not modulate SpCas9 activity.

CBEs were developed in a similar manner to ABEs, except that the module fused to nCas9 is an APOBEC family of cytidine deaminases (24,32). To test whether SB505124 works on CBEs, we first evaluated C-to-T conversion at eight genome loci with BE3-hA3A (BE3), a potent CBE most commonly used. BE3 induced efficient C-to-T conversion at five of eight sites (frequency ranged from 10% to 30%) and inefficient editing at three sites (frequency >5%) (Figure 5E, F). However, we observed no significant changes in editing frequencies at individual sites in the presence of SB505124, and the overall conversion rate was not dramatically induced by SB505124 (Figure 5G). To more rigorously characterize this effect, we tested the editing efficiency with other CBE variants, including BE3-mA1, BE3-hA3A-Y132D and BE3-hA3A-Y130F (24). Neither editing frequency nor unwanted indels and C-to-non-T substitutions increased in the presence of SB505124 for all nine sites tested (Figure 5F, H–J). These results collectively demonstrate that SB505124 specifically induces ABE editing activity.

### SB505124 promotes the correction of disease-causing mutations by ABEs

Point mutations are the largest class of known pathogenic genetic variants, among which approximately half are G·C to A·T mutations and could be corrected by ABEs (1,2). Pelizaeus–Merzbacher disease (PMD) is an X-linked recessive disorder of the central nervous system in myelination caused by mutations involving the proteolipid protein gene 1 (PLP1). Although PLP1 gene duplication is the most common mutation, point mutations, such as a missense mutation c.725C>T (p. Ala242Val, here referred to as A242V), can result in a rarer and more severe form of PMD (33). To correct this mutation, we transfected HEK293T cells harboring the mouse Plp1 A242V mutation with ABE 8e and gRNA targeting mutated adenine. Accordingly, ABE 8e corrected A242V at a frequency of 11% and, in the presence of SB505124, was induced to 32%, a 2.9-fold increase (Figure 6A). Mutations in the microtubule-associated protein TAU gene have previously been identified in individuals at high risk of developing neurodegenerative diseases. Among these, mutation c.454G>A (p. A152T) was identified in patients diagnosed with frontotemporal spectrum disorders (34). With HEK293T cells harboring A152T mutations, we found that SB505124 induced correction from 3.3% to 8.4%, a dramatic enhancement of 2.6-fold (Figure 6B). These results collectively demonstrated that SB505124 promotes correction of disease-causing mutations in human cells.

### SB505124 facilitates *in vivo* base editing by ABE

To test whether SB505124 facilitates *in vivo* base editing, we focused on the disease-associated gene proprotein convertase subtilisin/kexin type 9 (Pcsk9), which is primarily expressed in the liver and negatively regulates the low-desity lipoprotein (LDL) receptor. Inactivating its function by disrupting Pcsk9 canonical splicing sites can reduce blood LDL levels, thus presenting a promising therapeutic target (35,36). We first validated the editing of mouse Pcsk9 in N2a cells, in which conversion efficiency of position 6 (A6) was induced from 40.7% to 44% by SB505124 (Figure 6C). To examine correction in adult mice, we performed a hydrodynamic tail vein injection of an episomal plasmid carrying ABE 7.10 and the validated gRNA (36,37) (Figure 6D). Plasmids were sustained in the liver and ABE expression was monitored by eGFP expression (Figure 6G, H). From the second day, SB505124 or the solvent was intraperitoneally injected daily. On day 5, mice were sacrificed and genomic DNA was isolated from liver cells for analysis. We found that ABE 7.10 edited Pcsk9 at A6 whereby efficiency was 37.6% (Figure 6E). With SB505124, the edit rate was 46.3% (Figure 6E). At the transcription level, we found that Pcsk9 was knocked down to 67.8% by ABE, then a further 48.3% with SB505124 (Figure 6F). Combined, these results provide strong evidence that SB505124 promotes ABE editing on the disease-associated gene *in vivo*.

### Mechanism underlying ABE activation by SB505124

To explore the underlying mechanism of ABE activation by SB505124, we first examined the expression of TadA-Cas9, as well as gRNA, in the presence of SB505124 (Figure 7A, B). Treating cells with SB505124 induced TadA-Cas9 transcription (Figure 7A). Meanwhile, an increase in gRNA in the presence of SB505124 was detected by qRT-PCR (Figure 7B). In addition, using western blot, TadA-Cas9 fusion protein was induced 1.7-fold by SB505124 (Figure 7C, D). Therefore, SB505124 activates ABE editing, albeit mildly and partially, by inducing ABE expression.

To further explore its mechanism and identify downstream targets of SB505124, we performed RNA-seq and analyzed DEGs for cells with or without SB505124 treatment. As the process of base conversion is mainly mediated by the machinery of DNA damage repair, genes for DNA base excision repair, mismatch repair (MMR) and translesion synthesis (TLS) may regulate gene editors (3,38,39). Therefore, we focused on these DNA repair-associated genes to identify downstream effector(s). By analyzing DEGs, we found that the top gene that decreased in response to SB505124 was helicase-like transcription factor (HLTF), which was also down-regulated by ITD-1 (Supplementary Figure S8A–C). HLTF, a member of the SWI/SNF family, exhibits ATPase, helicase and E3 ubiquitin ligase activities, and is involved in the DNA replication fork reversal during DNA damage repair, thereby enabling an error-free bypass of replication blocks (40,41). Notably, replication elongation proceeds faster in HLTF-deficient cells under replication stress, although it might come at the expense of genome stability (42). Recently, HLTF has also been re-
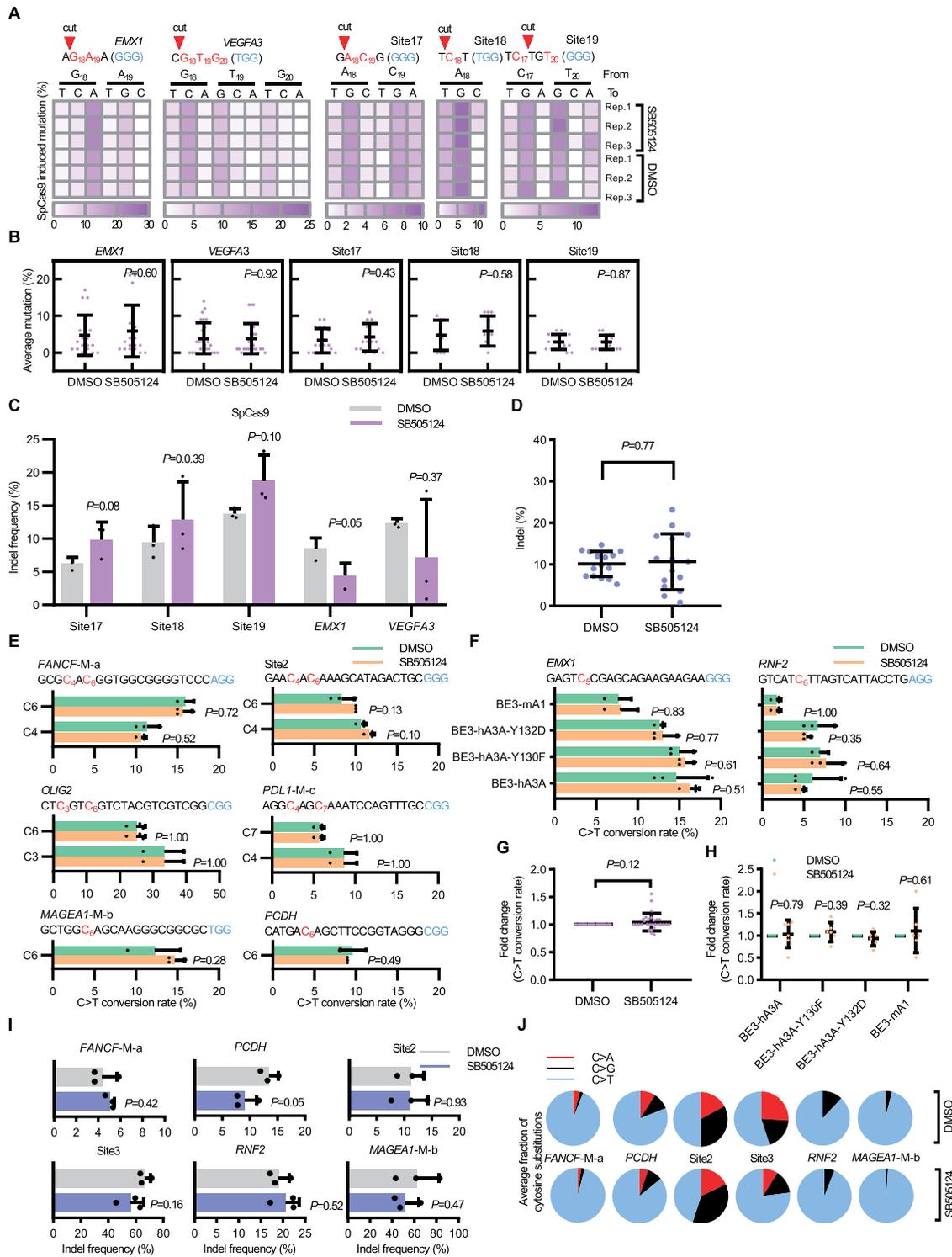
**Figure 5.** Effects of SB505124 on SpCas9 and CBEs. (**A–D**) Effect of SB505124 on SpCas9. Heat maps showing the genome cutting efficiencies at five indicated sites treated by SB505124 or DMSO ($n = 3$) (A). The average mutation rates for all sites tested are quantified and compared with the DMSO group (B). The indel rates quantified using NGS for all individual sites tested (C) and statistical analysis for all sites (D) are calculated and compared with the DMSO group. (**E**) Sanger sequencing results showing the editing efficiencies of BE3 at 10 positions at six genomic loci. (**F**) Sanger sequencing results showing the editing efficiencies of four indicated CBE variants at the EMX1 and RNF2 loci. (**G**) Statistical analysis of the normalized editing frequencies of BE3 for all six sites tested in (E). (**H**) Statistical analysis of the normalized editing frequencies of four indicated CBE variants at EMX1 and RNF2 as shown in (F). (**I**). Bar graphs showing the NGS results of the indel frequencies induced by BE3, as quantified at six indicated genome sites. (**J**) Fraction of on-target cytosine substitutions for BE3 at six indicated sites as analyzed using NGS. Data are presented as the mean ± SD and are the results of three technical replicates. *P*-values were determined using a two-tailed Student's *t*-test.
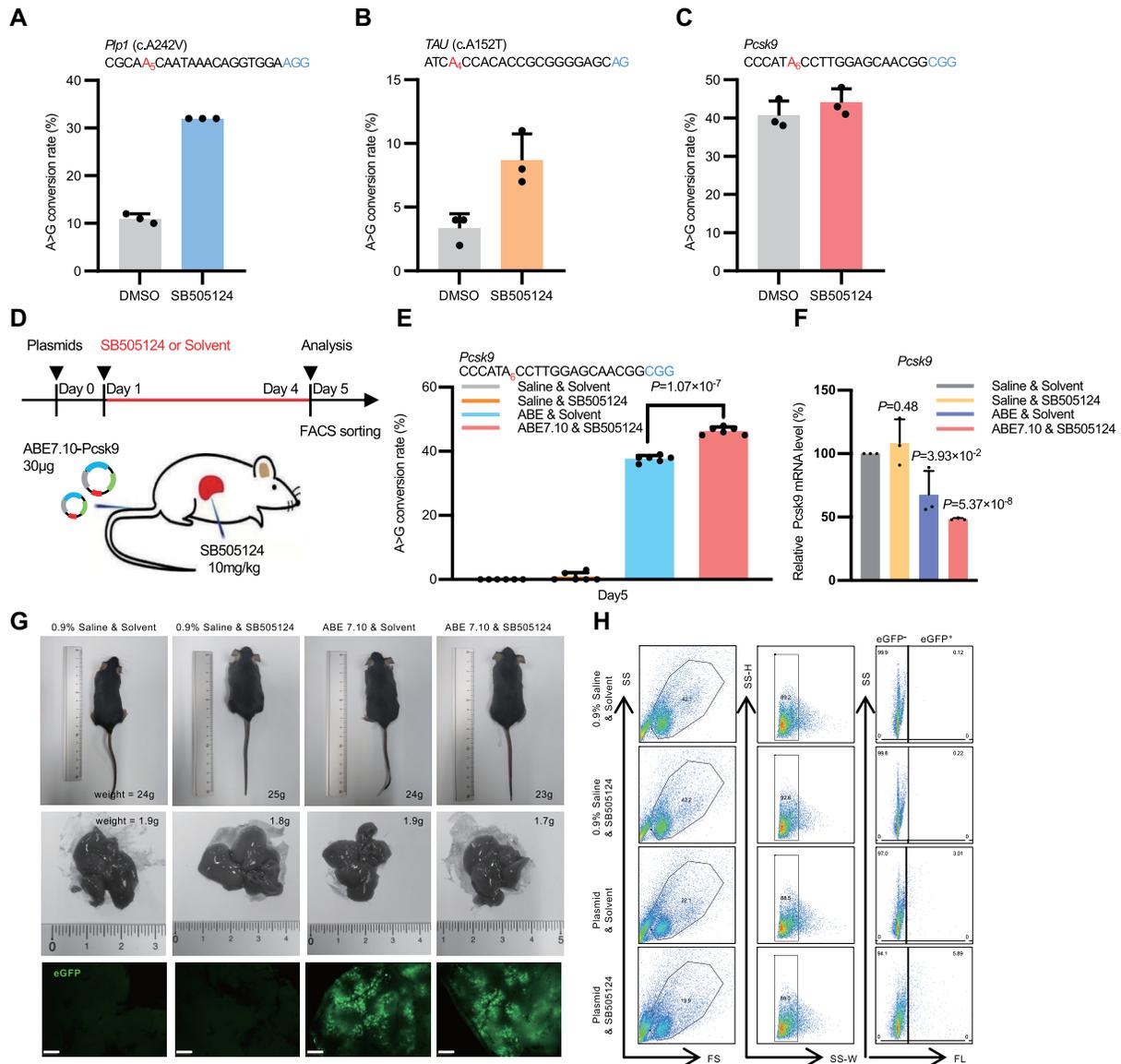
**Figure 6.** SB505124 improves ABE editing of disease-associated genes *in vivo*. (**A–C**) The conversion efficiencies of ABE 8e on Plp1 p. A242V (A), ABE 8e-NG on TAU p. A152T (B) and ABE 7.10 on Pcsk9 (C) treated with DMSO or SB505124. Data are presented as the mean ± SD and are the results of three technical replicates. *P*-values were determined using a two-tailed Student's *t*-test. (**D**) Schematic outline of mouse experiments. ABE 7.10 and validated gRNA targeting the Pcsk9 splicing site were administrated intravenously via tail vein injection, and either SB505124 or a solvent was administrated intraperitoneally on a daily basis from day 1 to day 4. Liver cells showing eGFP expression were isolated by FACS sorting by day 5. (**E**) The editing efficiencies of Pcsk9 at the indicated positions from six tissues of one mouse under the indicated conditions are shown. (**F**) The expression of Pcsk9 was determined and compared with a blank group via qRT-PCR. Three primary liver tissues were sourced from one mouse per group. *P*-values were determined using a two-tailed Student's *t*-test. (**G**) Representative photographs of C57BL/6 mice (upper panels), liver tissues (middle panels) and eGFP-expressing cells in liver tissue (bottom panels) under the indicated conditions. Scale bar: 100 μm. (**H**) FACS analysis of the eGFP-positive cells isolated from the liver tissue of C57BL/6 mice under the indicated conditions.

ported to play a role in modulating the activity of prime editor 3 (PE3), in which the knockdown of HLTF was found to improve editing efficiency (38). We found that the knockdown of HLTF by shRNAs (named shHLTF cells) repeatedly promoted ABE activity 1.4-fold (Figure 7E; Supplementary Figure S8D). This effect was further confirmed with three additional cell lines at different endogenous genome loci (Figure 7F; Supplementary Figure S8E). Importantly, treating shHLTF cells with SB505124 did not further enhance ABE editing, suggesting that they operate via the TGF-β pathway (Supplementary Figure S8F). Finally,

we analyzed whether HLTF affects CBE editing. As a result, we found that knockdown of HLTF did not affect CBE editing (Figure 7G; Supplementary Figure S8G). Interestingly, treating cells with SB505124 induced uracil *N*-glycosylase (UNG) expression in cells carrying CBE (Figure 7H; Supplementary Figure S8H). Notably, a reduced level of UNG has been reported to improve CBE editing (43,44). Indeed, UNG disruption promoted CBE editing 1.8-fold, but was less effective for ABE (Figure 7I, J; Supplementary Figure S8I, J). These data collectively suggest differential regulation patterns for distinct base editors.
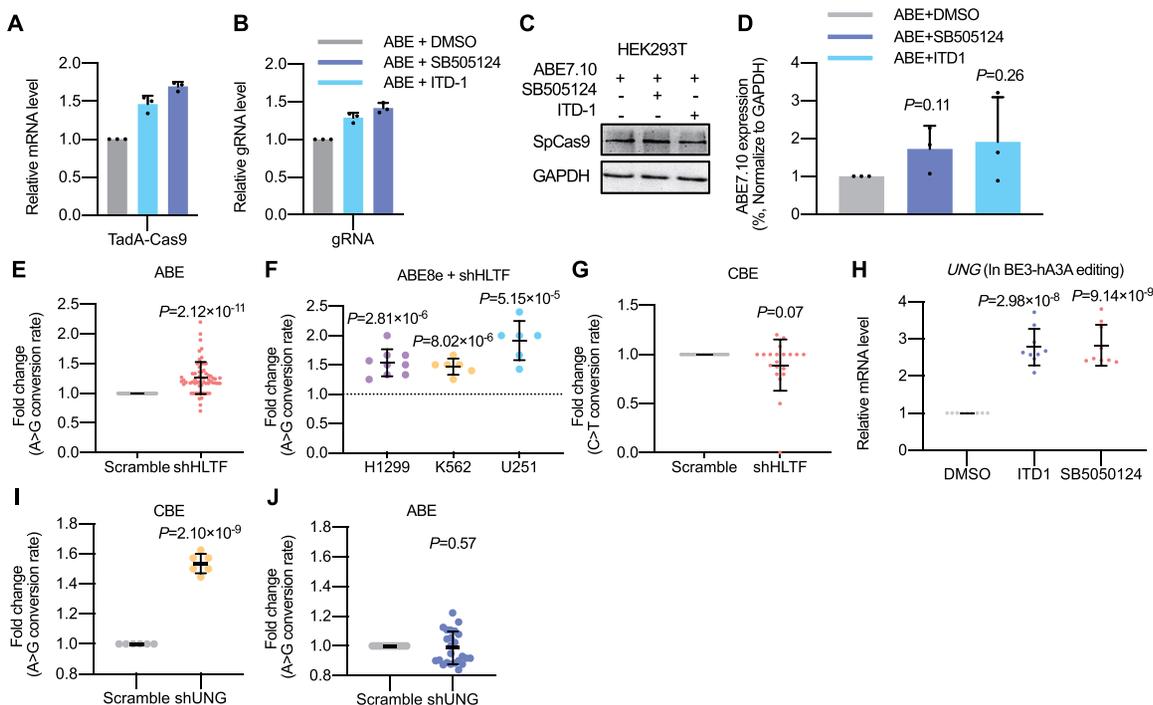
**Figure 7.** Mechanisms underlying ABE activation by SB505124. (**A, B**) qRT-PCR analysis of the transcription of TadA–Cas9 (A) and gRNA (B) under the indicated conditions. (**C, D**) Western blot analysis showing the expression of ABE under the indicated conditions in HEK293T cells (C, $n = 3$ from three biological replicates), and the quantification (D). (**E**) Fold changes of ABE editing efficiency for all target sites tested under helicase-like transcription factor (HLTF) knockdown. (**F**) ABE editing efficiency of U251, K562 and H1299 cells transfected with shRNAs targeting HLTF and normalized to those transfected with scrambled shRNAs. (**G**) Fold changes of BE3 editing efficiency at four endogenous loci when HLTF was knocked down. (**H**) qRT-PCR analysis showing UNG expression in HEK293T cells transfected with BE3 and treated by SB505124 or ITD-1, respectively ($n = 9$ from three independent biological replicates). (**I, J**) Fold changes of editing frequencies when UNG was knocked down by shRNA in cells carrying BE3 (I) or ABE (J), respectively. Data are presented as the mean ± SD and are from the results of three independent technical replicates. *P*-values were determined using a two-tailed Student's *t*-test.

## DISCUSSION

The primary application of base editors is to treat disease-associated point mutations. Therefore, achieving robust and precise editing in diverse genome regions is crucial. In this study, we screened ∼8000 small molecules and ultimately identified SB505124 as the most potent activator for promoting on-target editing of ABEs at multiple endogenous loci. SB505124 induced editing at epigenetically repressive regions, which were reported to strongly suppress the activity of base editors. In addition, SB505124 enhanced the editing of disease-associated genes *in vitro* and *in vivo*. With improved delivery technology, these activators could help facilitate *in vivo* editing, thus offering strong therapeutic potential.

The mechanism underlying ABE regulation remains elusive. In our study, we found that the canonical TGF-β pathway is involved in ABE regulation. Blocking this pathway with small-molecule inhibitors targeting ALK5 or the downstream effector SMAD2/3 greatly enhanced ABE editing activity. Further, genetic knockdown of ALK5 phenocopied that of those small molecules, confirming that the canonical TGF-β pathway is involved in regulating ABE activity. In our study, we observed a slight increase in ABE complex expression in the presence of SB505124. Given the robust activation of ABE editing by SB505124, an alternative mechanism enabling the dramatic promotion of ABE activity is possible. Recently, Liu and colleagues reported that the factors involved in MMR are important cellular determinants for PEs (38). Using RNA-seq analysis, we found that the induction of ABE editing by SB505124 may be achieved, at least in part, by down-regulating HLTF, a gene that contributes to translesion DNA synthesis (39). Interestingly, although the genetic depletion of HLTF was found to promote ABE editing, it did not exhibit any effect on CBEs. SB505124 performs differently for ABE and CBE, suggesting that certain factors can discriminate ABEs from CBEs. One recent study reported that ABE-mediated editing efficiency was distinctively higher than that of CBE in hESCs due to high UNG expression (44). In our study, we found that SB505124 induced the expression of UNG in cells carrying CBE. This may provide a mechanistic insight into the differential regulation of distinct base editors in mammalian cells. Taken together, our results suggest that the mechanisms underlying ABE activation by SB505124 involve multiple levels of biological regulation. Regardless of transcription activation, DNA replication/repair factors may play a major role in this regulation.

Precise and robust editing is the intrinsic property of gene editors and is affected by the local chromatin environment. Despite technical issues (e.g. low delivery efficiency and/or expression of gene editors in cells/tissues, such as in neu-

rons, hESCs or some other cell lines, which usually lead to poor editing efficiency), editing is affected by intrinsic cellular determinants (e.g. the expression/activity of endogenous modulators, such as UNG for CBEs and MMR genes for PEs, in specific cell types). In addition, the epigenetic status and sequence of a certain genome locus may influence the accessibility of gene editors and consequently editing efficiency (45–49). This is supported by the observation that the methylation of cytosine strongly suppresses the activity of CBEs (5) and that the efficiency of all reported gene editors varies greatly when working on different genomic loci or in different types of cells. Due to the distinct nature of chromatin/cells, the development of an efficient editing system via protein engineering/evolution and/or the discovery of effective modulators (e.g. small molecules) would improve the prospects for the future applications of gene editors.

## DATA AVAILABILITY

The next-generation sequencing data, including those for genomic DNA and the transcriptome, have been deposited in the NCBI Sequence Read Archive (https://www.ncbi.nlm.nih.gov/sra) and Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo) under accession numbers PRJNA809126, PRJNA809230 and GSE208182, respectively.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

## FUNDING

## REFERENCES

1. Landrum,M.J., Lee,J.M., Benson,M., Brown,G., Chao,C., Chitipiralla,S., Gu,B., Hart,J., Hoffman,D., Hoover,J. *et al.* (2016) ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res.*, **44**, D862–D868.

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

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

4. Koblan,L.W., Erdos,M.R., Wilson,C., Cabral,W.A., Levy,J.M., Xiong,Z.M., Tavarez,U.L., Davison,L.M., Gete,Y.G., Mao,X. *et al.* (2021) In vivo base editing rescues Hutchinson–Gilford progeria syndrome in mice. *Nature*, **589**, 608–614.

5. Nabel,C.S., Jia,H., Ye,Y., Shen,L., Goldschmidt,H.L., Stivers,J.T., Zhang,Y. and Kohli,R.M. (2012) AID/APOBEC deaminases disfavor modified cytosines implicated in DNA demethylation. *Nat. Chem. Biol.*, **8**, 751–758.

6. Arbab,M., Shen,M.W., Mok,B., Wilson,C., Matuszek,Z., Cassa,C.A. and Liu,D.R. (2020) Determinants of base editing outcomes from target library analysis and machine learning. *Cell*, **182**, 463–480.

7. Gaudelli,N.M., Lam,D.K., Rees,H.A., Sola-Esteves,N.M., Barrera,L.A., Born,D.A., Edwards,A., Gehrke,J.M., Lee,S.J., Liquori,A.J. *et al.* (2020) Directed evolution of adenine base editors with increased activity and therapeutic application. *Nat. Biotechnol.*, **38**, 892–900.

8. Richter,M.F., Zhao,K.T., Eton,E., Lapinaite,A., Newby,G.A., Thuronyi,B.W., Wilson,C., Koblan,L.W., Zeng,J., Bauer,D.E. *et al.* (2020) Phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity. *Nat. Biotechnol.*, **38**, 883–891.

9. Zhang,M., Lin,Y.H., Sun,Y.J., Zhu,S., Zheng,J., Liu,K., Cao,N., Li,K., Huang,Y. and Ding,S. (2016) Pharmacological reprogramming of fibroblasts into neural stem cells by signaling-directed transcriptional activation. *Cell Stem Cell*, **18**, 653–667.

10. Maji,B., Gangopadhyay,S.A., Lee,M., Shi,M., Wu,P., Heler,R., Mok,B., Lim,D., Siriwardena,S.U., Paul,B. *et al.* (2019) A high-throughput platform to identify small-molecule inhibitors of CRISPR-Cas9. *Cell*, **177**, 1067–1079.

11. Yu,C., Liu,Y., Ma,T., Liu,K., Xu,S., Zhang,Y., Liu,H., La Russa,M., Xie,M., Ding,S. *et al.* (2015) Small molecules enhance CRISPR genome editing in pluripotent stem cells. *Cell Stem Cell*, **16**, 142–147.

12. Ma,X., Chen,X., Jin,Y., Ge,W., Wang,W., Kong,L., Ji,J., Guo,X., Huang,J., Feng,X.H. *et al.* (2018) Small molecules promote CRISPR-Cpf1-mediated genome editing in human pluripotent stem cells. *Nat. Commun.*, **9**, 1303.

13. Maruyama,T., Dougan,S.K., Truttmann,M.C., Bilate,A.M., Ingram,J.R. and Ploegh,H.L. (2015) Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat. Biotechnol.*, **33**, 538–542.

14. Chu,V.T., Weber,T., Wefers,B., Wurst,W., Sander,S., Rajewsky,K. and Kuhn,R. (2015) Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nat. Biotechnol.*, **33**, 543–548.

15. DaCosta Byfield,S., Major,C., Laping,N.J. and Roberts,A.B. (2004) SB-505124 is a selective inhibitor of transforming growth factor-beta type I receptors ALK4, ALK5, and ALK7. *Mol. Pharmacol.*, **65**, 744–752.

16. Yeap,L.S., Hwang,J.K., Du,Z., Meyers,R.M., Meng,F.L., Jakubauskaite,A., Liu,M., Mani,V., Neuberg,D., Kepler,T.B. *et al.* (2015) Sequence-intrinsic mechanisms that target AID mutational outcomes on antibody genes. *Cell*, **163**, 1124–1137.

17. Hwang,G.H., Park,J., Lim,K., Kim,S., Yu,J., Yu,E., Kim,S.T., Eils,R., Kim,J.S. and Bae,S. (2018) Web-based design and analysis tools for CRISPR base editing. *BMC Bioinf.*, **19**, 542.

18. Park,J., Lim,K., Kim,J.S. and Bae,S. (2017) Cas-analyzer: an online tool for assessing genome editing results using NGS data. *Bioinformatics*, **33**, 286–288.

19. Kluesner,M.G., Nedveck,D.A., Lahr,W.S., Garbe,J.R., Abrahante,J.E., Webber,B.R. and Moriarity,B.S. (2018) EditR: a method to quantify base editing from Sanger sequencing. *CRISPR J*, **1**, 239–250.

20. Putri,G.H., Anders,S., Pyl,P.T., Pimanda,J.E. and Zanini,F. (2022) Analysing high-throughput sequencing data in Python with HTSeq 2.0. *Bioinformatics*, **38**, 2943–2945.

21. Love,M.I., Huber,W. and Anders,S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**, 550.

22. Shin,H.R., See,J.E., Kweon,J., Kim,H.S., Sung,G.J., Park,S., Jang,A.H., Jang,G., Choi,K.C., Kim,I. *et al.* (2021) Small-molecule inhibitors of histone deacetylase improve CRISPR-based adenine base editing. *Nucleic Acids Res.*, **49**, 2390–2399.

23. Zhao,T., Li,Q., Zhou,C., Lv,X., Liu,H., Tu,T., Tang,N., Cheng,Y., Liu,X., Liu,C. *et al.* (2021) Small-molecule compounds boost genome-editing efficiency of cytosine base editor. *Nucleic Acids Res.*, **49**, 8974–8986.

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

25. Patel,D., Patel,M., Datta,S. and Singh,U. (2019) CGGBP1 regulates CTCF occupancy at repeats. *Epigenetics Chromatin*, **12**, 57.

26. Liang,P., Xie,X., Zhi,S., Sun,H., Zhang,X., Chen,Y., Chen,Y., Xiong,Y., Ma,W., Liu,D. *et al.* (2019) Genome-wide profiling of adenine base editor specificity by EndoV-seq. *Nat. Commun.*, **10**, 67.

27. Tsai,S.Q., Zheng,Z., Nguyen,N.T., Liebers,M., Topkar,V.V., Thapar,V., Wyvekens,N., Khayter,C., Iafrate,A.J., Le,L.P. *et al.* (2015) GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat. Biotechnol.*, **33**, 187–197.

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

29. Grunewald,J., Zhou,R., Iyer,S., Lareau,C.A., Garcia,S.P., Aryee,M.J. and Joung,J.K. (2019) CRISPR DNA base editors with reduced RNA off-target and self-editing activities. *Nat. Biotechnol.*, **37**, 1041–1048.

30. Rees,H.A., Wilson,C., Doman,J.L. and Liu,D.R. (2019) Analysis and minimization of cellular RNA editing by DNA adenine base editors. *Sci. Adv.*, **5**, eaax5717.

31. Zhou,C., Sun,Y., Yan,R., Liu,Y., Zuo,E., Gu,C., Han,L., Wei,Y., Hu,X., Zeng,R. *et al.* (2019) Off-target RNA mutation induced by DNA base editing and its elimination by mutagenesis. *Nature*, **571**, 275–278.

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

33. Yamamoto,T., Nanba,E., Zhang,H., Sasaki,M., Komaki,H. and Takeshita,K. (1998) Jimpy(msd) mouse mutation and connatal Pelizaeus–Merzbacher disease. *Am. J. Med. Genet.*, **75**, 439–440.

34. Kovacs,G.G., Wohrer,A., Strobel,T., Botond,G., Attems,J. and Budka,H. (2011) Unclassifiable tauopathy associated with an A152T variation in MAPT exon 7. *Clin. Neuropathol.*, **30**, 3–10.

35. Musunuru,K., Chadwick,A.C., Mizoguchi,T., Garcia,S.P., DeNizio,J.E., Reiss,C.W., Wang,K., Iyer,S., Dutta,C., Clendaniel,V. *et al.* (2021) In vivo CRISPR base editing of PCSK9 durably lowers cholesterol in primates. *Nature*, **593**, 429–434.

36. Rothgangl,T., Dennis,M.K., Lin,P.J.C., Oka,R., Witzigmann,D., Villiger,L., Qi,W., Hruzova,M., Kissling,L., Lenggenhager,D. *et al.* (2021) In vivo adenine base editing of PCSK9 in macaques reduces LDL cholesterol levels. *Nat. Biotechnol.*, **39**, 949–957.

37. Song,C.Q., Jiang,T., Richter,M., Rhym,L.H., Koblan,L.W., Zafra,M.P., Schatoff,E.M., Doman,J.L., Cao,Y., Dow,L.E. *et al.* (2020) Adenine base editing in an adult mouse model of tyrosinaemia. *Nat. Biomed. Eng.*, **4**, 125–130.

38. Chen,P.J., Hussmann,J.A., Yan,J., Knipping,F., Ravisankar,P., Chen,P.F., Chen,C., Nelson,J.W., Newby,G.A., Sahin,M. *et al.* (2021) Enhanced prime editing systems by manipulating cellular determinants of editing outcomes. *Cell*, **184**, 5635–5652.

39. Sale,J.E. (2013) Translesion DNA synthesis and mutagenesis in eukaryotes. *Cold Spring Harb. Perspect. Biol.*, **5**, a012708.

40. Poole,L.A. and Cortez,D. (2017) Functions of SMARCAL1, ZRANB3, and HLTF in maintaining genome stability. *Crit. Rev. Biochem. Mol. Biol.*, **52**, 696–714.

41. Unk,I., Hajdu,I., Fatyol,K., Hurwitz,J., Yoon,J.H., Prakash,L., Prakash,S. and Haracska,L. (2008) Human HLTF functions as a ubiquitin ligase for proliferating cell nuclear antigen polyubiquitination. *Proc. Natl Acad. Sci. USA*, **105**, 3768–3773.

42. Kile,A.C., Chavez,D.A., Bacal,J., Eldirany,S., Korzhnev,D.M., Bezsonova,I., Eichman,B.F. and Cimprich,K.A. (2015) HLTF's ancient HIRAN domain binds 3′ DNA ends to drive replication fork reversal. *Mol. Cell*, **58**, 1090–1100.

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

44. Park,J.C., Jang,H.K., Kim,J., Han,J.H., Jung,Y., Kim,K., Bae,S. and Cha,H.J. (2022) High expression of uracil DNA glycosylase determines C to T substitution in human pluripotent stem cells. *Mol. Ther. Nucleic Acids*, **27**, 175–183.

45. Knight,S.C., Xie,L., Deng,W., Guglielmi,B., Witkowsky,L.B., Bosanac,L., Zhang,E.T., El Beheiry,M., Masson,J.B., Dahan,M. *et al.* (2015) Dynamics of CRISPR-Cas9 genome interrogation in living cells. *Science*, **350**, 823–826.

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

47. Isaac,R.S., Jiang,F., Doudna,J.A., Lim,W.A., Narlikar,G.J. and Almeida,R. (2016) Nucleosome breathing and remodeling constrain CRISPR-Cas9 function. *Elife*, **5**, e13450.

48. Verkuijl,S.A. and Rots,M.G. (2019) The influence of eukaryotic chromatin state on CRISPR-Cas9 editing efficiencies. *Curr. Opin. Biotechnol.*, **55**, 68–73.

49. Wu,X., Scott,D.A., Kriz,A.J., Chiu,A.C., Hsu,P.D., Dadon,D.B., Cheng,A.W., Trevino,A.E., Konermann,S., Chen,S. *et al.* (2014) Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nat. Biotechnol.*, **32**, 670–676.