nature biotechnology

ARTICLES

https://doi.org/10.1038/s41587-019-0178-z

Corrected: Author Correction

Programmable RNA editing by recruiting endogenous ADAR using engineered RNAs

Liang Qu[1,2,5], Zongyi Yi[1,3,5], Shiyou Zhu[1,5], Chunhui Wang[1,5], Zhongzheng Cao[1,3,5], Zhuo Zhou[1,5], Pengfei Yuan[4,5], Ying Yu[1], Feng Tian[1], Zhiheng Liu[1,3], Ying Bao[1], Yanxia Zhao[4] and Wensheng Wei[1*]

Current tools for targeted RNA editing rely on the delivery of exogenous proteins or chemically modified guide RNAs, which may lead to aberrant effector activity, delivery barrier or immunogenicity. Here, we present an approach, called leveraging endogenous ADAR for programmable editing of RNA (LEAPER), that employs short engineered ADAR-recruiting RNAs (arRNAs) to recruit native ADAR1 or ADAR2 enzymes to change a specific adenosine to inosine. We show that arRNA, delivered by a plasmid or viral vector or as a synthetic oligonucleotide, achieves editing efficiencies of up to 80%. LEAPER is highly specific, with rare global off-targets and limited editing of non-target adenosines in the target region. It is active in a broad spectrum of cell types, including multiple human primary cell types, and can restore α-L-iduronidase catalytic activity in Hurler syndrome patient-derived primary fibroblasts without evoking innate immune responses. As a single-molecule system, LEAPER enables precise, efficient RNA editing with broad applicability for therapy and basic research.

Genome editing technologies using engineered nucleases, such as zinc finger nucleases[1], transcription activator-like effector nucleases (TALENs)[2–4] and Cas proteins of CRISPR system[5–7], have been applied to manipulate the genome in a myriad of organisms. Recently, a cytidine or an adenosine deaminase has been coupled with CRISPR-Cas9 to create programmable DNA base editors[8–10], offering new opportunities for correcting disease-causing mutations. In addition to DNA editing, the ADAR adenosine deaminases have been exploited to achieve precise adenosine-to-inosine editing of RNAs. Three kinds of ADAR protein have been identified in mammals, ADAR1 (isoforms p110 and p150), ADAR2 and ADAR3 (refs. [11,12]), whose substrates are double-stranded RNAs. Inosine is believed to mimic guanosine (G) during translation[13,14]. To achieve targeted RNA editing, the ADAR protein or its catalytic domain was fused with a λN peptide[15–17], a SNAP-tag[18–22] or a Cas protein (dCas13b)[23], and a guide RNA (gRNA) was designed to recruit the chimeric ADAR protein to the specific site. Alternatively, overexpressing ADAR1 or ADAR2 proteins together with an R/G motif-bearing gRNA was also reported to enable targeted RNA editing[24–27].

All these reported nucleic acid editing methods in mammalian systems rely almost exclusively on ectopic expression of two components, an enzyme and a gRNA, although Katrekar et al.[27] have obtained preliminary evidence for editing in the absence of ADAR overexpression. Although these binary systems work efficiently in most studies, some inherited obstacles limit their broad applications, especially in therapies. Because the most effective in vivo delivery for gene therapy is through viral vectors[28], and the highly desirable adeno-associated virus vectors are limited in cargo size (~4.5 kilobases), this makes it challenging to accommodate both the protein and the gRNA[29,30]. Over-expression of ADAR1 has recently been reported to confer oncogenicity in multiple myelomas due to aberrant hyper-editing on RNAs[31] and to generate substantial global off-targeting edits[32]. In addition, ectopic expression of proteins or their domains of non-human origin runs the potential risk of eliciting immunogenicity[30,33]. Moreover, pre-existing adaptive immunity and p53-mediated DNA damage response may compromise the efficacy of the therapeutic protein, such as Cas9 (refs. [34–38]). Endogenous mechanisms for RNA editing have been harnessed by injecting pre-assembled target transcript:RNA duplex into *Xenopus* embryos[39]. Stafforst and colleagues recently reported a RNA editing method, named RESTORE, which works through recruiting endogenous ADARs using chemosynthetic antisense oligonucleotides with complex chemical modification[40]. Here, we describe an alternative approach that uses endogenous ADAR for RNA editing. We show that expressing arRNA enables efficient, precise editing of endogenous RNA and correction of pathogenic mutations.

Results

Exploiting endogenous ADAR for RNA editing. We fused the deaminase domain of the hyperactive E1008Q mutant ADAR1 ($ADAR1_{DD}$)[41] to the catalytic inactive LbuCas13 (dCas13a), a RNA-guided RNA-targeting CRISPR effector[42] (Supplementary Fig. 1a). To assess RNA editing efficiency, we constructed a surrogate reporter harboring *mCherry* and *EGFP* genes linked by 3× GGGGS-coding sequence and an in-frame UAG stop codon (Reporter-1, see Supplementary Fig. 1b). The reporter-transfected cells only expressed mCherry protein, while targeted editing on the UAG of the reporter transcript could convert the stop codon to UIG and consequently permit the downstream enhanced green fluorescent protein (EGFP) expression. Such a reporter allows us to measure the A-to-I editing efficiency through monitoring EGFP level. We then designed hU6 promoter-driven CRISPR RNAs (crRNAs) containing 5′ scaffolds subjected to Cas13a recognition and variable lengths of spacer sequences for targeting ($crRNA^{Cas13a}$, see Supplementary Table 1). The sequences complementary to the target transcripts all contain a CCA opposite to the UAG codon so as to introduce a cytidine (C) mis-pairing with

[1]Biomedical Pioneering Innovation Center, Beijing Advanced Innovation Center for Genomics, Peking-Tsinghua Center for Life Sciences, Peking University Genome Editing Research Center, State Key Laboratory of Protein and Plant Gene Research, School of Life Sciences, Peking University, Beijing, China. [2]Peking University–Tsinghua University–National Institute of Biological Sciences Joint Graduate Program, Peking University, Beijing, China. [3]Academy for Advanced Interdisciplinary Studies, Peking University, Beijing, China. [4]EdiGene Inc., Life Science Park, Changping District, Beijing, China. [5]These authors contributed equally: Liang Qu, Zongyi Yi, Shiyou Zhu, Chunhui Wang, Zhongzheng Cao, Zhuo Zhou, Pengfei Yuan. *e-mail: wswei@pku.edu.cn

the adenosine (A) (Supplementary Fig. 1b) because adenosine deamination preferentially occurs in the A-C mismatch site[13,14]. To test the optimal length of the crRNA, non-targeting or targeting crRNAs of different lengths were coexpressed with dCas13a-ADAR1$_{DD}$ in human embryonic kidney 293T (HEK293T) cells stably expressing the Reporter-1. Evidence of RNA editing indicated by the EGFP expression was observed with crRNAs containing matching sequences at least 40-nucleotides (nt) long, and the longer the crRNAs, the higher the editing efficiency (Supplementary Fig. 1c). Expression of long crRNA^{Cas13a} alone appeared sufficient to activate EGFP expression, and the coexpression of dCas13a-ADAR1$_{DD}$ rather decreased crRNA activity (Supplementary Fig. 1c,d). The EGFP expression was clearly sequence-dependent because the control RNA (Ctrl RNA) could not activate EGFP expression (Supplementary Fig. 1c,d).

Moreover, we tested a recently reported RNA editing system, termed REPAIR, which used dCas13b to direct targeted RNA editing by ADAR proteins[23]. We found that a 70-nt crRNA spacer with the dCas13b scaffold (crRNA^{dCas13b}) could also induce EGFP expression in the absence of dCas13b-ADAR fusions, while depletion of endogenous ADAR1 abrogated this effect (Supplementary Table 1 and Supplementary Fig. 1e).

As certain long engineered crRNA^{Cas13a} enabled RNA editing independent of dCas13a-ADAR1 fusions, we decided to remove the Cas13a-recruiting scaffold from the crRNA. It turned out that a 70-nt gRNA devoid of scaffold sequence induced strong EGFP expression in close to 40% of total cells harboring the Reporter-1 (Fig. 1a,b). Because endogenous ADAR proteins could edit dsRNA substrates[12], we reasoned that the long gRNAs could anneal with the target transcripts to form dsRNA substrates that in turn recruit endogenous ADAR proteins for targeted editing. We thus designated such a gRNA as arRNA.

To determine whether endogenous ADAR proteins are responsible for the above observation, we examined arRNA-mediated RNA editing in ADAR-deficient cells. Since *ADAR2* messenger RNA was barely detectable in HEK293T cells (Supplementary Fig. 2a), we generated HEK293T *ADAR1*$^{-/-}$ cells, rendering this cell line deficient in both ADAR1 and ADAR2 (Fig. 1c,d). The depletion of ADAR1 abrogated 70-nt arRNA (arRNA$_{70}$)-induced EGFP signals (Fig. 1b), demonstrating that arRNA-induced EGFP reporter expression solely depended on native ADAR1. Moreover, exogenous expression of ADAR1^{p110}, ADAR1^{p150} or ADAR2 in HEK293T *ADAR1*$^{-/-}$ cells (Fig. 1c,d) successfully rescued the expression of EGFP (Fig. 1e and Supplementary Fig. 2b). Sanger sequencing analysis on the arRNA$_{70}$-targeting region showed an A/G overlapping peak at the predicted adenosine site within UAG, indicating a significant A-to-I (G) conversion (Fig. 1f). Next-generation sequencing (NGS) further confirmed that the A-to-I conversion rate was about 13% (Fig. 1g). The quantitative PCR (qPCR) analysis showed that arRNA$_{111}$ did not perturb the expression of targeted transcripts (Supplementary Fig. 3), ruling out the possible RNA interference (RNAi) effect of the arRNA. Collectively, our data demonstrated that the arRNA can generate a significant level of editing on the targeted transcripts.

LEAPER enables RNA editing in multiple cell lines. Because the expression of endogenous ADAR proteins is a prerequisite for LEAPER-mediated RNA editing, we tested the performance of LEAPER on a panel of cell lines originated from distinct tissues, including HT29, A549, HepG2, RD, SF268, SW13 and HeLa. We first examined the endogenous expression of all three kinds of ADAR proteins. ADAR1 was highly expressed in all tested cell lines, but ADAR3 was detected only in HepG2 and HeLa cells (Supplementary Fig. 4a,b). ADAR2 was non-detectable in any cells, a result that was not due to the failure of western blotting because ADAR2 protein could be detected from ADAR2-overexpressing HEK293T cells (Supplementary Fig. 4a,b), consistent with previous reports that ADAR1 is ubiquitously expressed, while the expressions of ADAR2 and ADAR3 are restricted to certain tissues[11].

We then set out to test the editing efficiencies of a re-designed arRNA$_{71}$ targeting the Reporter-1 (Supplementary Fig. 5a and Supplementary Table 2) in these cell lines. LEAPER worked in all tested cells, albeit with varying efficiencies (Supplementary Fig. 4c). These results were in agreement with the previous report that the ADAR1/2 protein levels correlate with the RNA editing yield[43], with the exception of HepG2 and HeLa cells. The suboptimal editing efficiencies in these two lines were likely due to the abundant expressions of ADAR3 (Supplementary Fig. 4a,b), which has been reported to play an inhibitory role in RNA editing[44]. LEAPER also worked in three different cell lines of mouse origin (NIH3T3, mouse embryonic fibroblast (MEF) and B16) (Supplementary Fig. 4d), paving the way for testing its therapeutics potential through animal and disease models. Collectively, we conclude that LEAPER is a versatile tool for a wide spectrum of cell types and for different organisms.

Characterization and optimization of LEAPER. To better characterize and optimize LEAPER, we investigated the choices of nucleotide opposite to the adenosine within the UAG triplet of the targeted transcript. Reporter-1-targeting arRNA$_{71}$ showed variable editing efficiencies with a changed triplet (5′-CNA, N denotes one of A/U/C/G) opposite to the targeted UAG. The A-C mismatch resulted in the highest editing efficiency, and the A-G mismatch yielded the fewest evident edits (Supplementary Table 2 and Fig. 2a). We then tested all 16 combinations of 5′ and 3′ neighbor sites surrounding the cytidine (5′-N^{1}CN^2) of the A-C mismatch (Supplementary Table 2) and found that the 3′ neighboring adenosine was required for the efficient editing, while adenosine is the least favorable nucleotide at the 5′ site (Fig. 2b,c). We thus concluded that the CCA motif on the arRNA confers the highest editing efficiency targeting the UAG site.

Length of RNA appeared relevant to arRNA efficiency in directing the editing on the targeted transcripts (Supplementary Fig. 1c), consistent with a previous report[43]. To fully understand this effect, we tested arRNAs with variable lengths targeting two different reporter transcripts—Reporter-1 and Reporter-2 (Supplementary Fig. 5a,b and Supplementary Table 2). Based on the reporter EGFP activities, the length of arRNA correlated positively with the editing efficiency, for both reporters, peaking at 111–191-nt (Fig. 2d). Although one arRNA$_{51}$ to be appeared working, 71-nt was the minimal length for arRNA to work for both reporters (Fig. 2d).

Next, we investigated the effect of the A-C mismatch position within an arRNA on editing efficiency. We fixed the lengths of all arRNAs to 71-nt, and slid the UAG-targeting ACC triplet from 5′ to 3′ within arRNAs (Supplementary Table 2). It turned out that placing the A-C mismatch in the middle region resulted in high editing yield, and arRNAs with mismatch sites close to the 3′ end outperformed those close to the 5′ end in both reporters (Fig. 2e). For convenience, we placed the A-C mismatch at the center of arRNAs for all of our subsequent studies.

We also tested the targeting flexibility of LEAPER. For all 16 triplet combinations (5′-N^{1}AN^2) on Reporter-3 (Supplementary Fig. 5c), we fixed the arRNA length (111-nt) and ensured the preferred A-C mismatch (Fig. 2f and Supplementary Table 2). NGS results showed that all N^{1}AN^2 motifs could be edited. The UAN^2 and GAN^2 are the most and the least preferable motifs, respectively (Fig. 2f,g). Collectively, the nearest neighbor preference of the target adenosine is 5′ U > C ≈ A > G and 3′ G > C > A ≈ U (Fig. 2g).

Editing endogenous transcripts using LEAPER. Next, we examined whether LEAPER could enable effective editing on endogenous transcripts. Using arRNAs of different lengths, we targeted the UAG motifs in the transcripts of *PPIB*, *KRAS* and *SMAD4* genes and an UAC motif in *FANCC* gene transcript (Fig. 3a and Supplementary Table 2). Encouragingly, targeted adenosine sites in all four transcripts were edited by their corresponding arRNAs with all four sizes, and longer arRNAs tended to yield higher editing rates

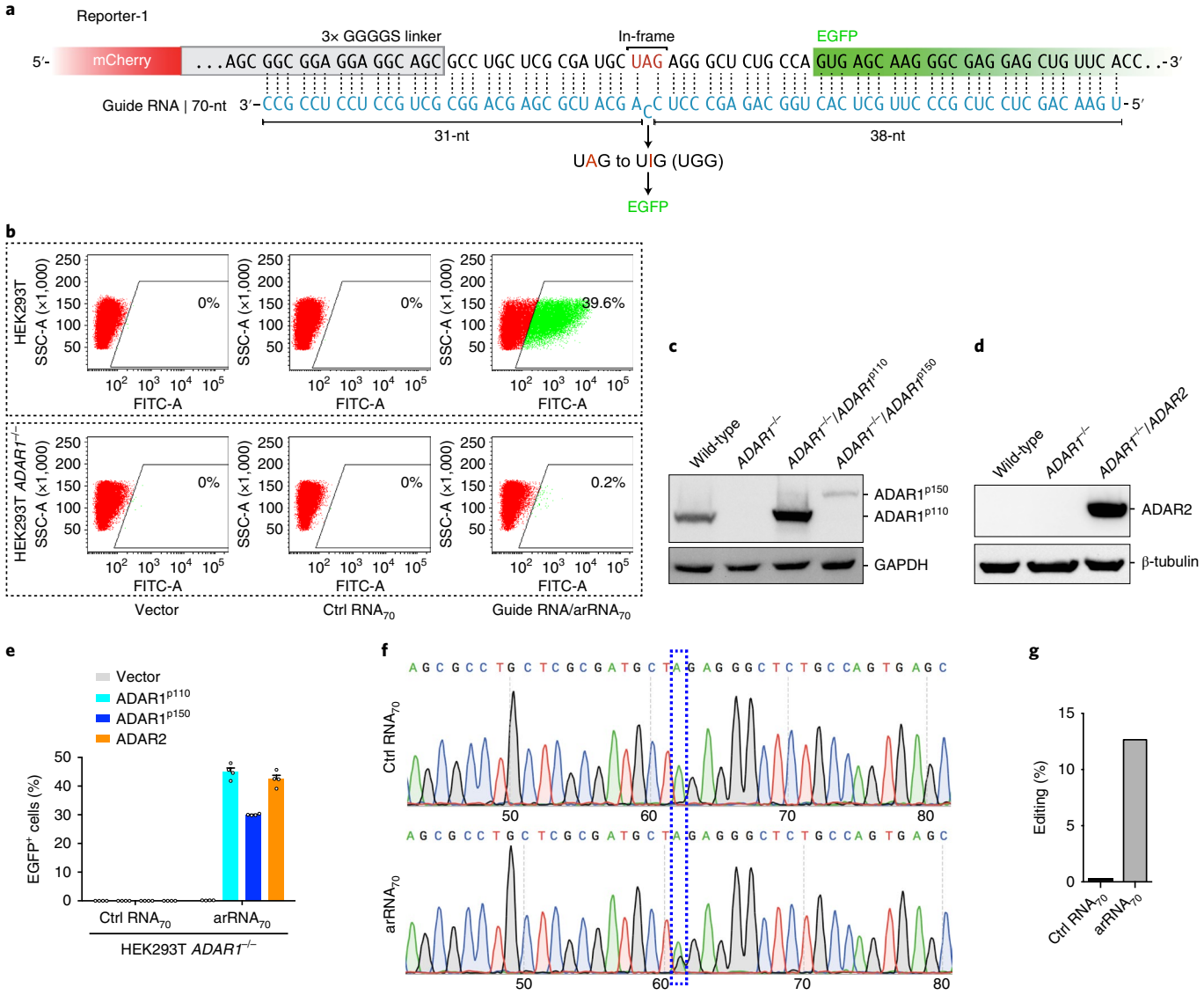

Fig. 1 | Leveraging endogenous ADAR1 protein for targeted RNA editing. a, Schematic of the Reporter-1 and the 70-nt arRNA. **b**, Representative FACS analysis of arRNA-induced EGFP expression in wild-type (HEK293T, upper) or ADAR1 knockout (HEK293T *ADAR1*$^{-/-}$, lower) cells stably expressing the Repoter-1. **c**, Western blot analysis showing expression levels of ADAR1 proteins in wild-type and HEK293T *ADAR1*$^{-/-}$ cells, as well as those in HEK293T *ADAR1*$^{-/-}$ cells transfected with ADAR1 isoforms (p110 and p150). **d**, Western blot analysis showing expression levels of ADAR2 proteins in wild-type and HEK293T *ADAR1*$^{-/-}$ cells, as well as those in HEK293T *ADAR1*$^{-/-}$ cells transfected with ADAR2. **e**, Quantification of the EGFP positive (EGFP^{+}) cells. Reporter-1 and indicated ADAR-expressing constructs were cotransfected into HEK293T *ADAR1*$^{-/-}$ cells, along with the Ctrl RNA$_{70}$ or with the targeting arRNA$_{70}$, followed by FACS analysis. EGFP^{+} percentages were normalized by transfection efficiency, which was determined by mCherry^{+}. Data are mean values ± s.e.m. $n = 4$, n represents the number of independent experiments performed in parallel. **f**, The electropherograms showing Sanger sequencing results in the Ctrl RNA$_{70}$ (upper) or the arRNA$_{70}$ (lower)-targeted region. **g**, Quantification of the A-to-I conversion rate at the targeted site by deep sequencing.

(Fig. 3b). Of note, the 151-nt arRNA^{PPIB} edited ~50% of total transcripts of *PPIB* gene (Fig. 3b). No arRNAs showed RNAi effects on their targeted transcripts (Supplementary Fig. 6a) or ultimate protein level (Supplementary Fig. 6b). Besides, LEAPER is able to achieve desirable editing rate on non-UAN sites (Fig. 3c and Supplementary Table 2), showing the flexibility of LEAPER on editing endogenous transcripts. We further tested whether LEAPER could simultaneously target multiple sites. We observed multiplex editing of both *TARDBP* and *FANCC* transcripts by coexpression of two arRNAs (Supplementary Table 2), with the efficiency even higher than those with individual arRNAs (Fig. 3d), indicating that LEAPER is multiplexable.

Because ADAR1/2 tend to promiscuously deaminate multiple adenosines in an RNA duplex[45], all adenosines on target transcripts within the arRNA coverages are likely subjected to variable levels of editing. The longer the arRNA, the higher the possibility of such off-targets. We therefore examined all adenosine sites within the arRNA covering regions in these targeted transcripts. For *PPIB* transcripts, very little off-target editing was observed throughout the sequencing window for variable sizes of arRNAs (Fig. 3e,f). However, in targeting *KRAS*, *SMAD4* and *FANCC* genes, multiple off-target edits were detected (Supplementary Fig. 7a–f). For *KRAS* in particular, 11 out of 30 adenosines underwent substantial A-to-I conversions in the sequencing window of arRNA$_{111}$ (Supplementary Fig. 7a,b).

We next attempted to minimize such off-target effects. Because an A-G mismatch suppressed editing for UAG targeting (Fig. 2a), we postulated that pairing a guanosine with a non-targeting adenosine might reduce undesirable editing. We then tested the effect

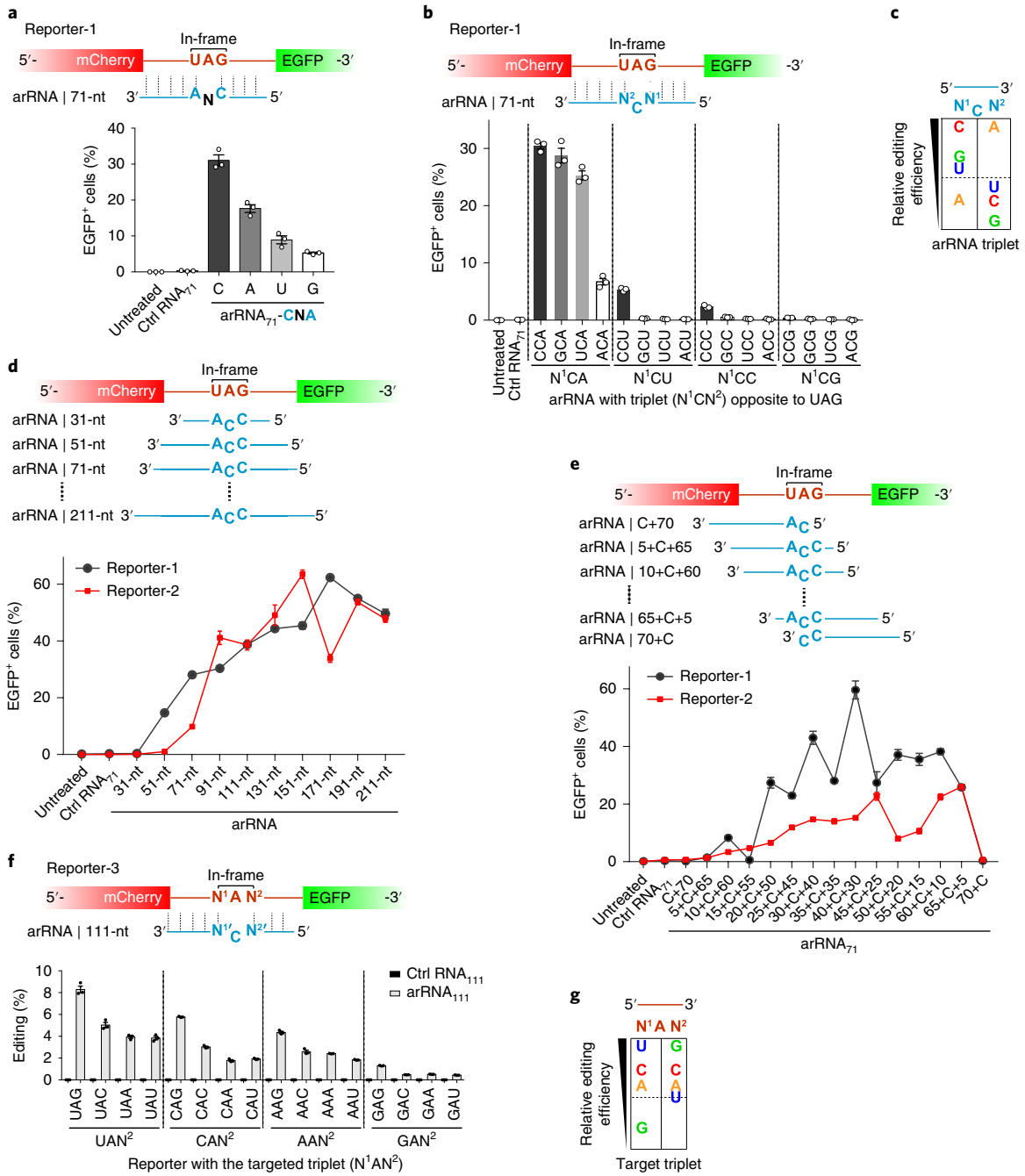

Fig. 2 | Characterization and optimization of LEAPER. a, Top, schematic of the design of arRNAs with a changed triplet (5′-CNA, N denotes A, U, C or G) opposite to the target UAG. Bottom, EGFP$^+$ percentage showing the effects of variable bases opposite to the targeted adenosine on RNA editing efficiency. **b**, Top, the design of arRNAs with changed neighboring bases flanking the cytidine in the A-C mismatch (5′-N^{1}CN^2). Bottom, the effects of 16 different combinations of N^{1}CN^2 on RNA editing efficiency. **c**, Summary of the preference of 5′ and 3′ nearest neighboring sites of the cytidine in the A-C mismatch. **d**, Top, the design of arRNAs with variable lengths. Bottom, the effect of arRNA length on RNA editing efficiency based on Reporter-1 and Reporter-2. **e**, Top, the design of arRNAs with variable A-C mismatch position. Bottom, the effect of A-C mismatch position on RNA editing efficiency based on Reporter-1 and Reporter-2. **f**, Top, the design of the triplet motifs in the reporter-3 with variable nearest neighboring bases surrounding the targeting adenosine (5′-N^{1}AN^2) and the opposite motif (5′-N^{2}CN^1) on the 111-nt arRNA (arRNA$_{111}$). Bottom, deep sequencing results showing the editing rate on targeted adenosine in the 5′-N^{1}AN^2 motif. **g**, Summary of the 5′ and 3′ base preferences of LEAPER-mediated editing at the Reporter-3. Error bars in **a**, **b**, **d**–**f** all indicate mean values ± s.e.m. $n = 3$, n represents the number of independent experiments performed in parallel.

of the A-G mismatch on adenosine in all possible triplet combinations (5′-N1AN2) as in Reporter-3 (Supplementary Fig. 5c and Supplementary Table 2). In comparison with A-C mismatch, A-G mismatch decreased the editing on adenosine in all tested triplets, except for UAG and AAG (~2%) (Fig. 3g). To further reduce editing rates at unwanted sites, we went on testing the effect of two consecutive mismatches. It turned out that the additional mismatch at the 3′ end nucleotide of the triplet opposite to either UAG or AAG abolished its corresponding adenosine editing (Fig. 3h and Supplementary Table 2). In light of these findings, we attempted to apply this rule to reduce off-targets in *KRAS* transcripts. We first designed an arRNA$_{111}$-AG6 that created A-G mismatches on all

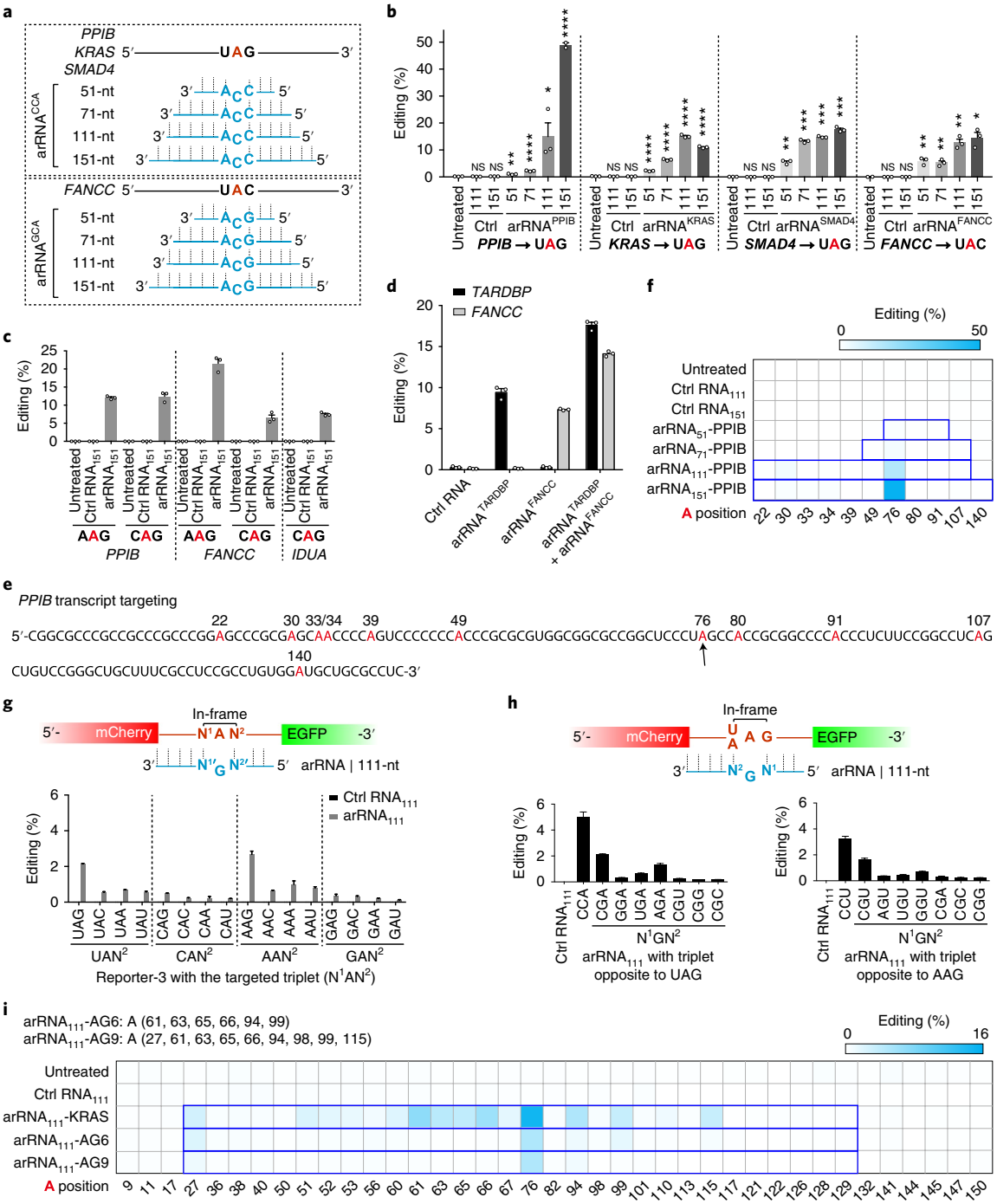

Fig. 3 | Editing endogenous transcripts with LEAPER. a, Schematic of the targeting endogenous transcripts of four disease-related genes (*PPIB*, *KRAS*, *SMAD4* and *FANCC*) and the corresponding arRNAs. **b**, Deep sequencing results showing the editing rate on targeted adenosine of the *PPIB*, *KRAS*, *SMAD4* and *FANCC* transcripts by introducing indicated lengths of arRNAs. **c**, Deep sequencing results showing the editing rate on non-UAN sites of endogenous *PPIB*, *FANCC* and *IDUA* transcripts. **d**, Multiplex editing rate by two 111-nt arRNAs. Indicated arRNAs were transfected alone or were cotransfected into the HEK293T cells. The targeted editing at the two sites was measured from cotransfected cells. **e**, Schematic of the *PPIB* transcript sequence covered by the 151-nt arRNA. The black arrow indicates the targeted adenosine. All adenosines were marked in red. **f**, Heatmap of editing rate on adenosines covered by indicated lengths of arRNAs targeting the *PPIB* gene (marked in bold frame in blue). For the 111-nt arRNA or $arRNA_{151}$-PPIB covered region, the editing rates of A22, A30, A33 and A34 were determined by RNA-seq because of the lack of effective PCR primers for amplifying this region. Otherwise the editing rate was determined by targeted deep-sequencing analysis. **g**, Top, the design of the triplet motifs in the reporter-3 with variable nearest neighboring bases surrounding the targeting adenosine (5′-N^1AN^2) and the opposite motif (5′-$N^{2\prime}GN^{1\prime}$) in the 111-nt arRNA ($arRNA_{111}$). Bottom, deep sequencing results showing the editing rate. **h**, Top, the design of arRNAs with two consecutive mismatches in the 5′-N^1GN^2 motif opposite to the 5′-UAG or the 5′-AAG motifs. Deep sequencing results showing the editing rate by an $arRNA_{111}$ with two consecutive mismatches in the 5′-N^1GN^2 motif opposite to the 5′-UAG motif (bottom left) or the 5′-AAG motif (bottom right). **i**, Heatmap of the editing rate on adenosines covered by engineered $arRNA_{111}$ variants targeting the *KRAS* gene. Data in **b**–**d**, **g**, **h** are presented as the mean ± s.e.m. $n = 3$; unpaired two-sided Student's *t*-test, $^*P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$; $^{****}P < 0.0001$; NS, not significant. Data in **f** and **i** are presented as the mean ($n = 3$). Experiments were independently performed in parallel.

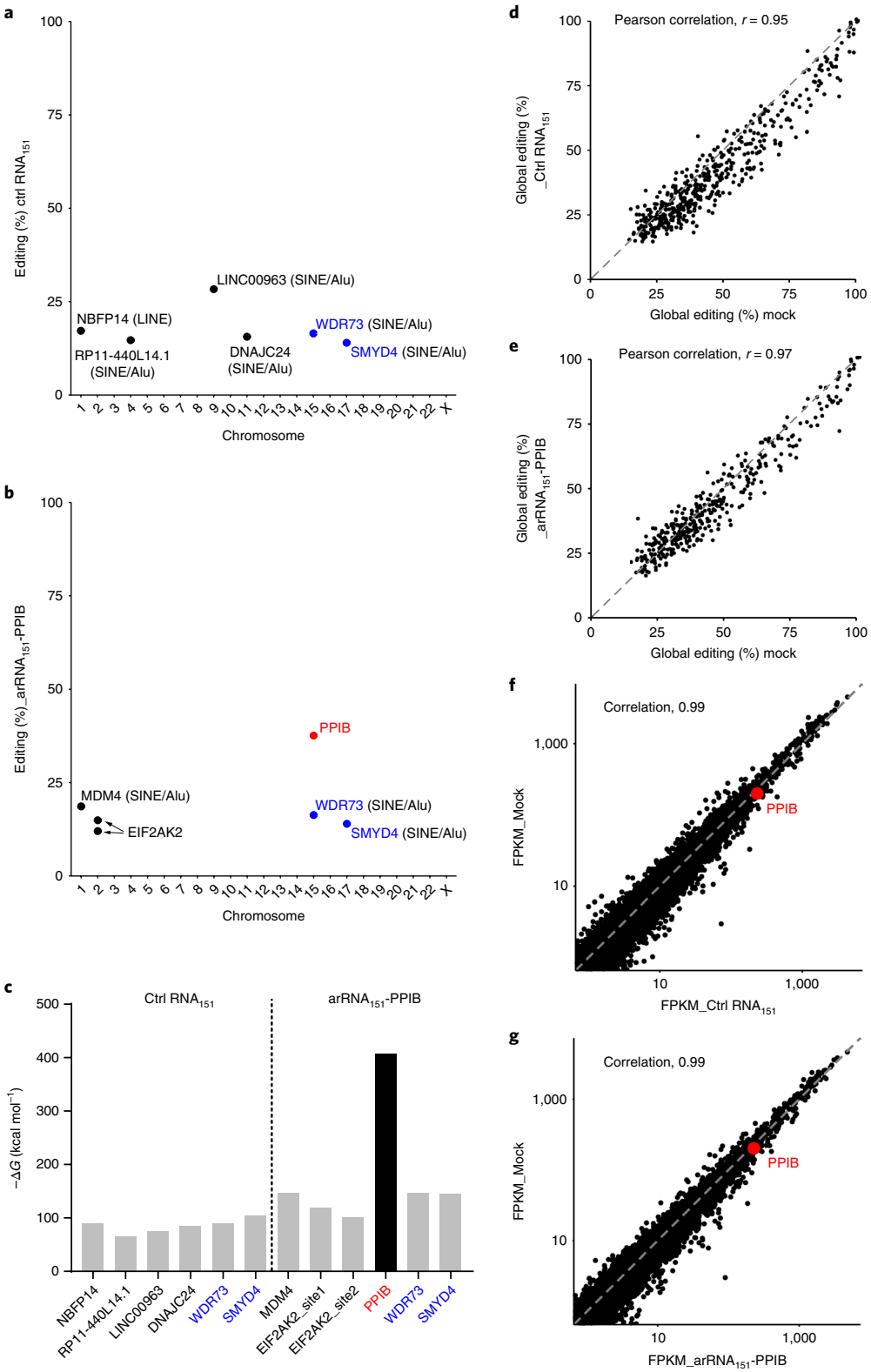

Fig. 4 | Transcriptome-wide specificity of RNA editing by LEAPER. a,**b**, Transcriptome-wide off-targeting analysis of Ctrl RNA_{151} (**a**) and $arRNA_{151}$-PPIB (**b**). The on-targeting site (PPIB) is highlighted in red. The potential off-target sites identified in both Ctrl RNA and PPIB-targeting RNA groups are labeled in blue. **c**, The predicted annealing affinity between off-target sites and the corresponding Ctrl RNA_{151} or $arRNA_{151}$-PPIB. The minimum free energy (ΔG) of dsRNA formed by off-target sites (150-nt upstream and downstream of the editing sites) and the corresponding Ctrl RNA_{151} or $arRNA_{151}$-PPIB was predicted with RNAhybrid, an online website tool. **d**,**e**, Transcriptome-wide analysis of the effects of Ctrl RNA_{151} (**d**) and $arRNA_{151}$-PPIB (**e**) on native editing sites by transcriptome-wide RNA-sequencing. Pearson's correlation coefficient analysis was used to assess the differential RNA editing rate on native editing sites. Four independent experiments were performed. **f**,**g**, Differential gene expression analysis of the effects of Ctrl RNA_{151} (**f**) and $arRNA_{151}$-PPIB (**g**) with RNA-seq data at the transcriptome level. The FPKM value was calculated with the STRINGTIE tool. Pearson's correlation coefficient analysis was used to assess the global differential gene expression.

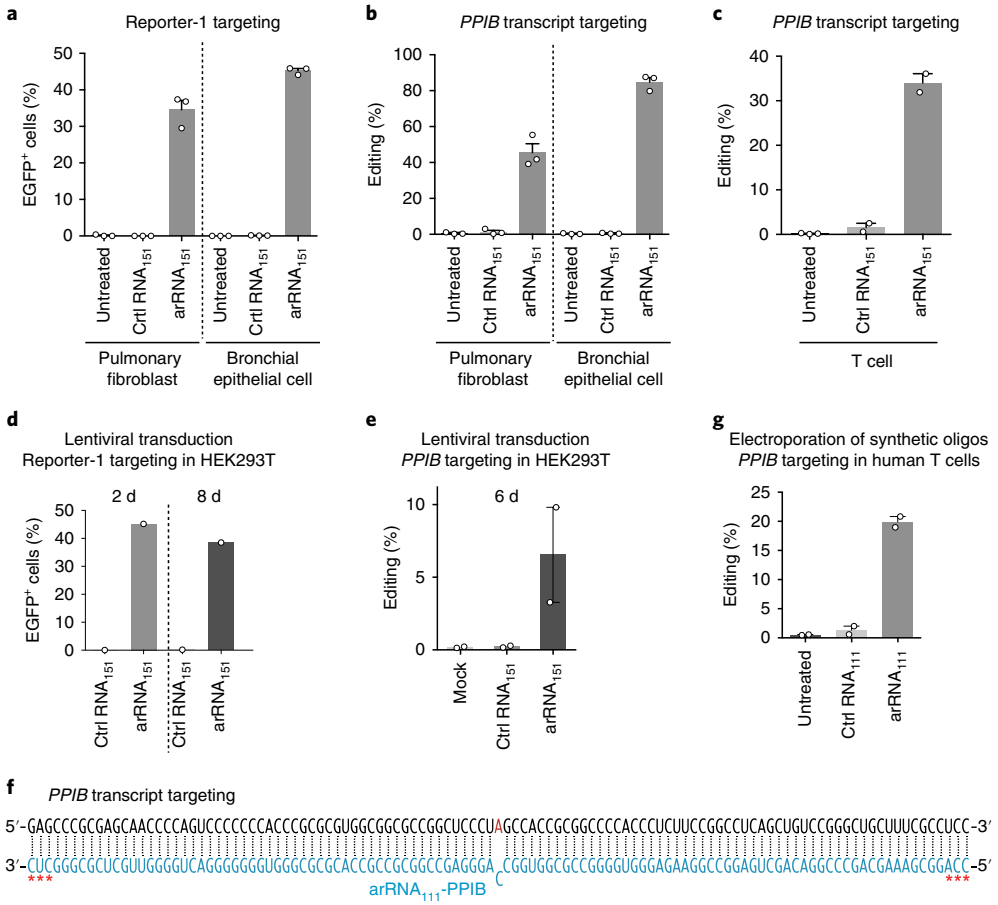

Fig. 5 | RNA editing in multiple human primary cells by LEAPER. a, Quantification of the EGFP$^+$ cells induced by LEAPER-mediated RNA editing. Human primary pulmonary fibroblasts and human primary bronchial epithelial cells were transfected with Reporter-1, along with the 151-nt control RNA (Ctrl RNA$_{151}$) or the 151-nt targeting arRNA (arRNA$_{151}$) followed by FACS analysis. **b**,**c**, Deep sequencing results showing the editing rate on *PPIB* transcripts in human primary pulmonary fibroblasts, human primary bronchial epithelial cells (**b**) and human primary T cells (**c**). **d**, Quantification of the EGFP$^+$ cells. HEK293T cells stably expressing the Reporter-1 were infected with lentivirus expressing 151-nt of Ctrl RNA or the targeting arRNA. FACS analyses were performed 2 and 8 d post infection. The ratios of EGFP$^+$ cells were normalized by lentiviral transduction efficiency (BFP$^+$ ratios). **e**, Deep sequencing results showing the editing rate on the *PPIB* transcripts on lentiviral transduction of 151-nt arRNAs into HEK293T cells. **f**, Schematic of the *PPIB* sequence and the corresponding 111-nt targeting arRNA. *(in red) represents nucleotide with 2′-*O*-methyl and phosphorothioate linkage modifications. **g**, Deep sequencing results showing the editing rate on the *PPIB* transcripts on electroporation of 111-nt synthetic arRNA oligonucleotides into human primary T cells. Data in **a**,**b** and the ntreated group **c** are presented as the mean ± s.e.m. $n = 3$. We performed on three plates of cells for each experiment. Data of Ctrl RNA$_{151}$ and arRNA$_{151}$ (**c**) are presented as the mean ± s.e.m. $n = 2$. Experiments were independently performed in parallel.

'editing-prone' motifs covered by arRNA$_{111}$ (Fig. 3i, Supplementary Fig. 7a and Supplementary Table 2), including AAU (the 61st), UAU (the 63rd), UAA (the 65th), AAA (the 66th), UAG (the 94th) and AAG (the 99th). This arRNA$_{111}$-AG6 eliminated most of the off-target editing, while maintained an on-target editing rate of ~5%. Consistent with the findings in Fig. 3g, the single A-G mismatch could not completely minimize editing in AAG motif (99th) (Fig. 3i and Supplementary Fig. 7a). We then added more mismatches on arRNA$_{111}$-AG6, including a dual mismatch (5′-CGG opposite to the targeted motif 5′-AAG), plus three additional A-G mismatches to mitigate editing on the 27th, 98th and the 115th adenosines (arRNA$_{111}$-AG9) (Supplementary Table 2). Consequently, we achieved a much improved specificity for editing, without additional loss of editing rate on the targeted site (A76) (Fig. 3i). In summary, engineered LEAPER incorporating additional rules enables efficient and more precise RNA editing on endogenous transcripts.

RNA editing specificity of LEAPER. In addition to the arRNA-covered dsRNA region, the potential off-targets may occur on other transcripts through partial base pairing of arRNA. We then performed a transcriptome-wide RNA-sequencing analysis to evaluate the global off-target effects of LEAPER. Cells were transfected with plasmids expressing Ctrl RNA$_{151}$ or arRNA$_{151}$-PPIB before being subjected to RNA-seq analysis. We identified six potential off-targets in the Ctrl RNA$_{151}$ group (Fig. 4a) and five in the arRNA$_{151}$-PPIB group (Fig. 4b), and the *PPIB* on-target rate based on NGS analysis was ~37% (Fig. 4b). Further analysis revealed that all sites, except for the two sites from *EIF2AK2* transcripts, were located in either SINE (Alu) or LINE regions (Fig. 4a,b); both of which are prone to ADAR-mediated editing[46], suggesting that these off-targets may not be derived from pairing between the target transcripts and the arRNA or control RNA. Of note, two off-targeting transcripts, *WDR73* and *SMYD4*, appeared in both groups, suggesting they are unlikely to be involved in sequence-dependent RNA editing. Indeed, minimum free energy analysis indicated that all these possible off-target transcripts failed to form a stable duplex with either Ctrl RNA$_{151}$ or arRNA$_{151}$-PPIB (Fig. 4c). To further test whether arRNA generates sequence-dependent off-targets, we selected potential off-target sites by comparing sequence similarity for both arRNA$_{151}$-PPIB and arRNA$_{111}$-FANCC. *TRAPPC12*

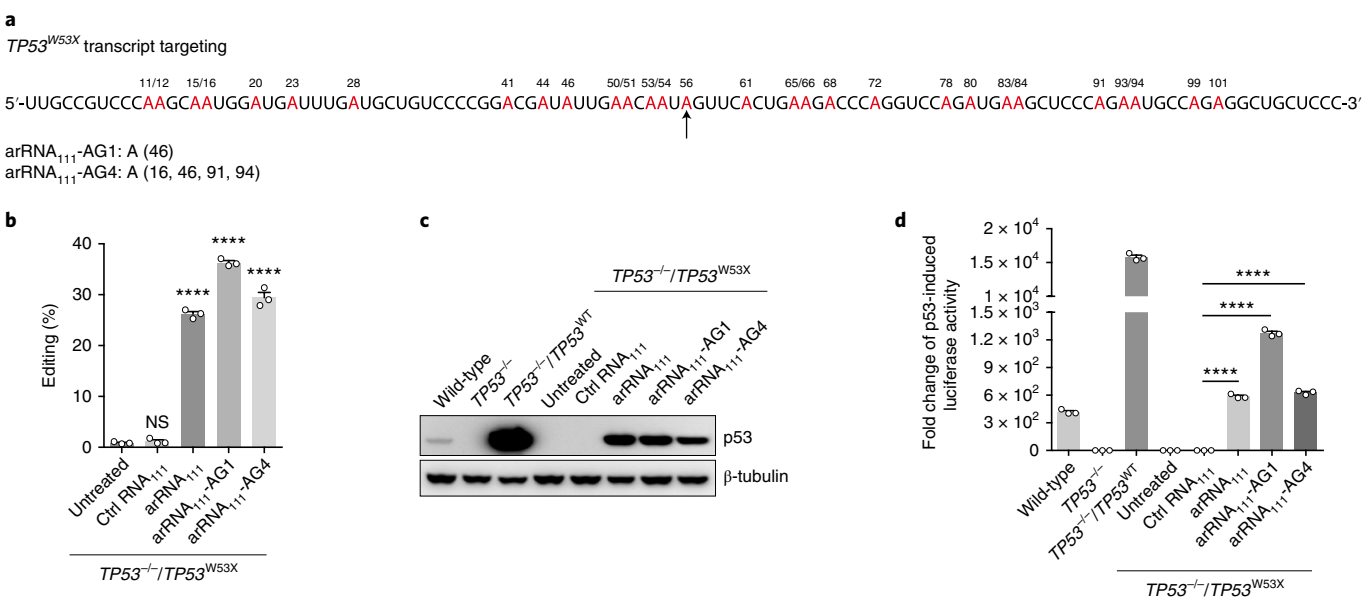

Fig. 6 | Restoration of transcriptional regulatory activity of mutant *TP53*^{W53X} by LEAPER. a, Top, schematic of the *TP53* transcript sequence covered by the 111-nt arRNA containing c.158G-to-A clinical-relevant non-sense mutation (Trp53Ter). The black arrow indicates the targeted adenosine. All adenosines were marked in red. Bottom, the design of two optimized arRNAs targeting *TP53*^{W53X} transcripts with A-G mismatch on A^{46th} for arRNA$_{111}$-AG1, and on A^{16th}, A^{46th}, A^{91th} and A^{94th} together for arRNA$_{111}$-AG4 to minimize the potential off-targets on editing-prone motifs. **b**, Deep sequencing results showing the targeted editing on *TP53*^{W53X} transcripts by arRNA$_{111}$, arRNA$_{111}$-AG1 and arRNA$_{111}$-AG4. **c**, Western blot showing the recovered production of full-length p53 protein from the *TP53*^{W53X} transcripts in the HEK293T *TP53*$^{-/-}$ cells. **d**, Detection of the transcriptional regulatory activity of restored p53 protein using a p53-Firefly-luciferase reporter system, normalized by cotransfected Renilla-luciferase vector. Data in **b**–**d** are presented as the mean ± s.e.m. $n = 3$, n represents the number of independent experiments performed in parallel; unpaired two-sided Student's t-test, $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$; $^{****}P < 0.0001$.

transcripts for arRNA$_{151}$-PPIB and three sites in the *ST3GAL1*, *OSTM1-AS1* and *EHD2* transcripts for arRNA$_{111}$-FANCC were top candidates (Supplementary Fig. 8a,b). NGS analysis revealed that no editing could be detected in any of these predicted off-target sites (Supplementary Fig. 8a,c). These results indicate that LEAPER empowers efficient editing at the targeted site, while maintaining transcriptome-wide specificity without detectable sequence-dependent off-target edits.

Safety assessment of LEAPER in mammalian cells. Because arRNAs rely on endogenous ADAR proteins for editing on-target transcripts, we wondered whether exogenous arRNAs affects native RNA editing events. Therefore, we analyzed the A-to-I RNA editing sites shared by a mock group and arRNA$_{151}$-PPIB group from the transcriptome-wide RNA-sequencing results. Neither the Ctrl RNA$_{151}$ group nor the arRNA$_{151}$-PPIB group showed a significant difference compared to the mock group (Fig. 4d,e), indicating that LEAPER had little impact on the normal A-to-I editing function of endogenous ADAR1.

To verify whether arRNA affects global gene expression, we performed differential gene expression analysis. In correlation analysis with the fragment per kilobase of exon model per million (FPKM) expression data, neither Ctrl RNA$_{151}$ nor arRNA$_{151}$-PPIB affected the global gene expression in comparison with the mock group (Fig. 4f,g). Moreover, DESeq2 analysis also revealed that there was no significant differential gene expression between the arRNA$_{151}$-PPIB group and Ctrl RNA$_{151}$ group (Supplementary Fig. 9 and Supplementary Table 3). Consistent with our previous observation (Supplementary Fig. 6a), arRNAs did not show any RNAi effect on the expression of *PPIB* (Fig. 4f,g and Supplementary Fig. 9).

Considering that the arRNA forms an RNA duplex with the target transcript and that RNA duplex might elicit an innate immune response, we investigated whether the introduction of arRNA has such an effect. To test this, we selected arRNAs targeting four gene transcripts that had proved effective. We did not observe any mRNA induction of interferon-β (IFN-β) (Supplementary Fig. 10a) or interleukin-6 (IL-6) (Supplementary Fig. 10b), which are two hallmarks of innate immune activation. As a positive control, a synthetic analog of dsRNA-poly(I:C) induced strong IFN-β and IL-6 expression (Supplementary Fig. 10a,b). LEAPER does not seem to induce immunogenicity in target cells, a feature that is important for safe therapeutics.

Corrections of pathogenic mutations by LEAPER. We next investigated whether LEAPER could be used to correct more pathogenic mutations. Aimed at clinically relevant mutations from six pathogenic genes, *COL3A1*, *BMPR2*, *AHI1*, *FANCC*, *MYBPC3* and *IL2RG*, we designed 111-nt arRNAs for each of these genes carrying corresponding pathogenic G-to-A mutations (Supplementary Fig. 11a and Supplementary Tables 2 and 4). By coexpressing arRNA/complementary DNA pairs in HEK293T cells, we identified significant amounts of target transcripts with A-to-G corrections in all tests (Supplementary Fig. 11b). Because G-to-A mutations account for nearly half of known disease-causing point mutations in humans[10,47], the A-to-G conversion by LEAPER may offer immense opportunities for therapeutics.

RNA editing in multiple human primary cells by LEAPER. To further explore its clinical use, we set out to test LEAPER in multiple human primary cells. First, we tested LEAPER in human primary pulmonary fibroblasts and human primary bronchial epithelial cells with 151-nt arRNA (Supplementary Table 2) to edit the Reporter-1 (Supplementary Fig. 5a). We found that 35–45% of EGFP^{+} cells could be obtained by LEAPER in both human primary cells (Fig. 5a). We then tested LEAPER in editing endogenous gene *PPIB* in these two primary cells and human primary T cells and found that arRNA$_{151}$-PPIB could achieve >40, >80 and >30% of editing rates in human primary pulmonary fibroblasts, primary bronchial epithelial

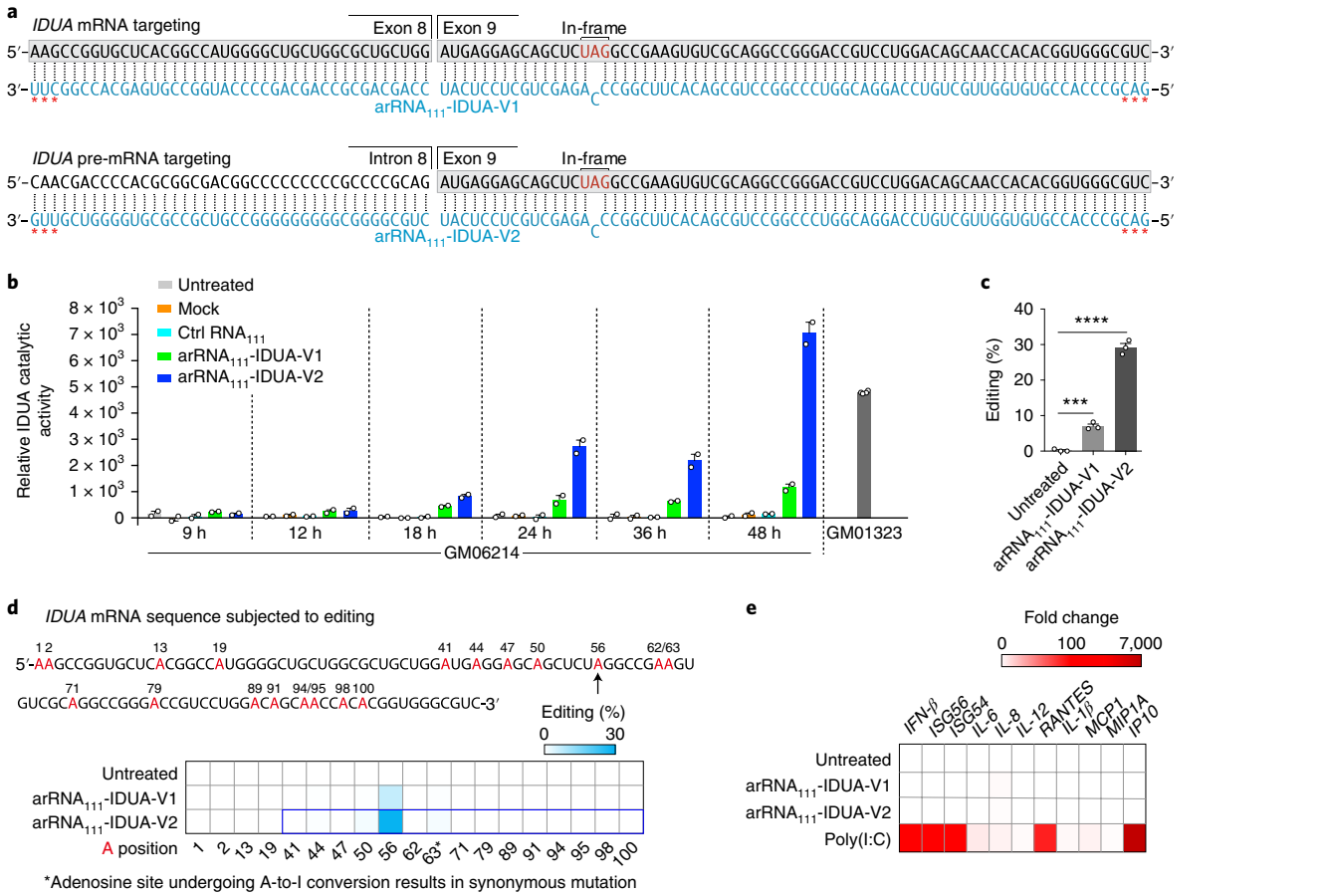

Fig. 7 | Restoration of IDUA activity in Hurler syndrome patient-derived primary fibroblast by LEAPER. a, Top, genetic information of pathogenic mutation in fibroblast GM06214 derived from Hurler syndrome patient; TGG-to-TAG mutation in exon 9, resulting in premature stop codon on gene *IDUA*; middle, schematic of the *IDUA* mature mRNA sequence of GM06214 cells (black) containing a homozygous TGG-to-TAG mutation in exon 9 of the *IDUA* gene (Trp402Ter) and the corresponding 111-nt targeting arRNA$_{111}$-IDUA-V1 (blue); bottom, schematic of the *IDUA* pre-mRNA sequence of GM06214 cells (black) and the corresponding 111-nt targeting arRNA$_{111}$-IDUA-V2 (blue). *(in red) represents nucleotides with 2′-*O*-methyl and phosphorothioate linkage modifications. **b**, Measuring the catalytic activity of IDUA with 4-methylumbelliferyl IDUA substrate at different time points. Data are presented as the mean ± s.e.m. $n = 2$. **c**, Deep sequencing results showing the targeted editing rate on *IDUA* transcripts in GM06214 cells, 48 h post electroporation. **d**, Top, schematic of the *IDUA* transcript sequence covered by the 111-nt arRNAs. The arrow indicates the targeted adenosine. All adenosines were marked in red. Bottom, a heatmap of editing rate on adenosines covered by indicated arRNAs in the *IDUA* transcript (marked in the bold frame in blue). **e**, qPCR showing the expressions of type-I interferon, interferon-stimulated genes and pro-inflammatory genes on arRNA or poly(I:C) electroporation. Data are presented as the mean ($n = 3$). Experiments were independently performed in parallel.

cells (Fig. 5b) and primary T cells (Fig. 5c), respectively. The high editing efficiency of LEAPER in human primary cells is particularly encouraging for its potential application in therapeutics.

RNA editing by clinically relevant formats of arRNAs. We then investigated whether LEAPER could be delivered by more clinically relevant methods. We first tested the effect of arRNA through lentivirus-based expression. Reporter-1-targeting arRNA$_{151}$ induced strong EGFP expression in more than 40% of total cells harboring the Reporter-1 in HEK293T cells 2 d post infection (dpi). At 8 dpi, the EGFP ratio maintained at a comparable level of ~38% (Fig. 5d and Supplementary Table 2), suggesting that LEAPER could be tailored to therapeutics that require continuous administration. For native gene editing, we delivered *PPIB*-targeting arRNA$_{151}$ through lentiviral transduction and observed over 6% of target editing at 6 dpi (Fig. 5e).

We next tested synthesized arRNA oligonucleotides and electroporation delivery method for LEAPER. The 111-nt arRNA targeting *PPIB* transcripts as well as Ctrl RNA were chemically synthesized with 2′-*O*-methyl and phosphorothioate linkage modifications at the first three and last three residues of arRNAs (Fig. 5f). After being introduced into T cells through electroporation, arRNA$_{111}$-PPIB oligos achieved ~20% of editing on *PPIB* transcripts (Fig. 5g), indicating that LEAPER holds promise for the development of oligonucleotide drugs.

Recovery of transcriptional activity of p53^{W53X} by LEAPER. Next, we studied the potential therapeutic use of LEAPER. We first targeted the tumor suppressor gene *TP53*, which is known to play a vital role in the maintenance of cellular homeostasis, but undergo frequent mutations in >50% of human cancers[48]. The c.158G-to-A mutation in *TP53* is a clinically relevant non-sense mutation (Trp53Ter), resulting in a non-functional truncated protein[48]. We designed one arRNA$_{111}$ and two alternative arRNAs (arRNA$_{111}$-AG1 and arRNA$_{111}$-AG4) (Supplementary Table 2), all targeting *TP53*^{W53X} transcripts (Fig. 6a), with the last two being designed to minimize potential off-targets. We generated HEK293T *TP53*$^{-/-}$ cell line to eliminate the effects of native p53 protein. All three forms of *TP53*^{W53X}-targeting arRNAs converted ~25–35% of *TP53*^{W53X} transcripts on the mutated adenosine site (Fig. 6b), with variable

reductions of unwanted edits for $arRNA_{111}$-AG1 and $arRNA_{111}$-AG4 (Supplementary Fig. 12). Western blot showed that $arRNA_{111}$, $arRNA_{111}$-AG1 and $arRNA_{111}$-AG4 could all rescue the production of full-length p53 protein based on the *TP53*^{W53X} transcripts in HEK293T *TP53*$^{-/-}$ cells, while the Ctrl RNA_{111} could not (Fig. 6c). Using a p53-luciferase *cis*-reporting system[49,50], we found all three versions of arRNAs could restore p53 activity in transcriptional regulation, and the optimized version $arRNA_{111}$-AG1 performed the best (Fig. 6d). In conclusion, we demonstrated that LEAPER is capable of repairing the cancer-relevant premature stop codon of *TP53* and restoring its function.

Restoration of α-L-iduronidase (IDUA) activity by LEAPER. Finally, we examined the potential of LEAPER in treating a monogenic disease—Hurler syndrome, the most severe subtype of Mucopolysaccharidosis type I (MPS I) due to the deficiency of IDUA, a lysosomal metabolic enzyme responsible for the degradation of mucopolysaccharides[51]. We chose a primary fibroblast GM06214 that was originally isolated from Hurler syndrome patient. The GM06214 cells contain a homozygous TGG-to-TAG mutation in exon 9 of the *IDUA* gene, resulting in a Trp402Ter mutation in the protein. We designed two versions of arRNAs by synthesized RNA oligonucleotides with chemical modifications, $arRNA_{111}$-IDUA-V1 and $arRNA_{111}$-IDUA-V2, targeting the mature mRNA and the pre-mRNA of *IDUA*, respectively (Fig. 7a and Supplementary Table 2). After introduction of $arRNA_{111}$-IDUA-V1 or $arRNA_{111}$-IDUA-V2 into GM06214 cells via electroporation, we measured the targeted RNA editing rates via NGS analysis and the catalytic activity of IDUA. Both $arRNA_{111}$-IDUA-V1 and $arRNA_{111}$-IDUA-V2 significantly restored the IDUA catalytic activity in *IDUA*-deficient GM06214 cells progressively with time after electroporation, and $arRNA_{111}$-IDUA-V2 performed much better than $arRNA_{111}$-IDUA-V1, while no IDUA activity could be detected in three control groups (Fig. 7b).

To further evaluate the extent to which the restored IDUA activity in GM06214 by LEAPER relieves Hurler syndrome, we examined the IDUA activity in GM01323 cells, another primary fibroblast from a patient with Scheie syndrome, a much milder subtype of MPS I than Hurler syndrome due to the remnant IDUA activity. We found that the catalytic activity of IDUA in GM06214 cells harboring $arRNA_{111}$-IDUA-V2 was higher than GM01323 cells 48 h post electroporation (Fig. 7b). Consistent with these results, NGS analysis indicated that $arRNA_{111}$-IDUA-V2 converted nearly 30% of A-to-I editing, a much higher rate than $arRNA_{111}$-IDUA-V1 (Fig. 7c). Further analysis revealed that minimal unwanted edits were detected within the arRNA-covered regions of *IDUA* transcripts (Fig. 7d). Neither $arRNA_{111}$-IDUA-V1 nor $arRNA_{111}$-IDUA-V2 induced expressions of a panel of genes involved in type-I interferon and pro-inflammatory responses (Fig. 7e). These results showed the therapeutic potential of LEAPER in targeting certain monogenetic diseases.

Discussion

In this report, we show that expression of a linear arRNA of adequate length can guide endogenous ADAR proteins to edit adenosine to inosine on targeted transcripts. LEAPER has several advantages over existing editing approaches. The small size of the arRNA molecule is reminiscent of RNAi, in which a small dsRNA induces a native mechanism for targeted RNA degradation[52], and enables delivery by a variety of viral and non-viral vehicles. Unlike RNAi, LEAPER catalyzes a precise A-to-I switch without cutting or degrading targeted transcripts (Supplementary Fig. 6a). Although the length requirement for arRNA is longer than for RNAi, arRNA neither induces immune-stimulatory effects at the cellular level (Supplementary Fig. 10 and Fig. 7e) nor affects the function of endogenous ADAR proteins (Fig. 4d,e), making it a safe strategy for RNA targeting. In contrast, ectopic expression of ADAR proteins or their catalytic domains has been reported to induce substantial global off-target edits[32] and possibly cancer[31]. Similarly, DNA base editors have been reported to generate substantial off-target single-nucleotide variants in mouse embryos, rice or human cell lines due to the expression of an effector protein[53–56]. LEAPER achieves efficient editing with rare global off-target editing (Fig. 4a,b and Supplementary Fig. 8). In addition, it may be less immunogenic than methods that require the introduction of foreign proteins.

In comparison with RESTORE[40], a recently reported RNA editing method, of which the gRNA of RESTORE is limited to chemosynthetic antisense oligonucleotides depending on complex chemical modification, arRNA of LEAPER can be generated in a variety of ways, including chemical synthesis and expression in vivo from viral or non-viral vectors (Fig. 5).

There is still room for improvement in LEAPER's efficiency and specificity. arRNA fused with an ADAR-recruiting scaffold may increase local ADAR protein concentration and consequently enhance editing yield. Ways of stabilizing arRNA or increasing its expression may further boost RNA editing efficiency. We envisage that LEAPER holds potential for broad applications in therapeutics and biomedical research.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of code and data availability and associated accession codes are available at https://doi.org/10.1038/s41587-019-0178-z.

Received: 11 April 2019; Accepted: 4 June 2019;
Published online: 15 July 2019

References

1. Porteus, M. H. & Carroll, D. Gene targeting using zinc finger nucleases. *Nat. Biotechnol.* **23**, 967–973 (2005).
2. Boch, J. et al. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **326**, 1509–1512 (2009).
3. Moscou, M. J. & Bogdanove, A. J. A simple cipher governs DNA recognition by TAL effectors. *Science* **326**, 1501 (2009).
4. Miller, J. C. et al. A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol.* **29**, 143–148 (2011).
5. Jinek, M. et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).
6. Cong, L. et al. Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013).
7. Mali, P. et al. RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826 (2013).
8. Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A. & Liu, D. R. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533**, 420–424 (2016).
9. Ma, Y. et al. Targeted AID-mediated mutagenesis (TAM) enables efficient genomic diversification in mammalian cells. *Nat. Methods* **13**, 1029–1035 (2016).
10. Gaudelli, N. M. et al. Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. *Nature* **551**, 464–471 (2017).
11. Tan, M. H. et al. Dynamic landscape and regulation of RNA editing in mammals. *Nature* **550**, 249–254 (2017).
12. Nishikura, K. Functions and regulation of RNA editing by ADAR deaminases. *Annu. Rev. Biochem.* **79**, 321–349 (2010).
13. Bass, B. L. & Weintraub, H. An unwinding activity that covalently modifies its double-stranded RNA substrate. *Cell* **55**, 1089–1098 (1988).
14. Wong, S. K., Sato, S. & Lazinski, D. W. Substrate recognition by ADAR1 and ADAR2. *RNA* **7**, 846–858 (2001).
15. Montiel-Gonzalez, M. F., Vallecillo-Viejo, I., Yudowski, G. A. & Rosenthal, J. J. Correction of mutations within the cystic fibrosis transmembrane conductance regulator by site-directed RNA editing. *Proc. Natl Acad. Sci. USA* **110**, 18285–18290 (2013).
16. Sinnamon, J. R. et al. Site-directed RNA repair of endogenous Mecp2 RNA in neurons. *Proc. Natl Acad. Sci. USA* **114**, E9395–E9402 (2017).
17. Montiel-Gonzalez, M. F., Vallecillo-Viejo, I. C. & Rosenthal, J. J. An efficient system for selectively altering genetic information within mRNAs. *Nucleic Acids Res.* **44**, e157 (2016).

18. Hanswillemenke, A., Kuzdere, T., Vogel, P., Jekely, G. & Stafforst, T. Site-directed RNA editing in vivo can be triggered by the light-driven assembly of an artificial riboprotein. *J. Am. Chem. Soc.* **137**, 15875–15881 (2015).
19. Schneider, M. F., Wettengel, J., Hoffmann, P. C. & Stafforst, T. Optimal guideRNAs for re-directing deaminase activity of hADAR1 and hADAR2 in trans. *Nucleic Acids Res.* **42**, e87 (2014).
20. Vogel, P., Hanswillemenke, A. & Stafforst, T. Switching protein localization by site-directed RNA editing under control of light. *ACS Synth. Biol.* **6**, 1642–1649 (2017).
21. Vogel, P., Schneider, M. F., Wettengel, J. & Stafforst, T. Improving site-directed RNA editing in vitro and in cell culture by chemical modification of the guideRNA. *Angew. Chemie* **53**, 6267–6271 (2014).
22. Vogel, P. et al. Efficient and precise editing of endogenous transcripts with SNAP-tagged ADARs. *Nat. Methods* **15**, 535–538 (2018).
23. Cox, D. B. T. et al. RNA editing with CRISPR-Cas13. *Science* **358**, 1019–1027 (2017).
24. Fukuda, M. et al. Construction of a guide-RNA for site-directed RNA mutagenesis utilising intracellular A-to-I RNA editing. *Sci. Rep.* **7**, 41478 (2017).
25. Wettengel, J., Reautschnig, P., Geisler, S., Kahle, P. J. & Stafforst, T. Harnessing human ADAR2 for RNA repair—recoding a PINK1 mutation rescues mitophagy. *Nucleic Acids Res.* **45**, 2797–2808 (2017).
26. Heep, M., Mach, P., Reautschnig, P., Wettengel, J. & Stafforst, T. Applying human ADAR1p110 and ADAR1p150 for site-directed RNA editing-G/C substitution stabilizes guideRNAs against editing. *Genes* **8**, E34 (2017).
27. Katrekar, D. et al. In vivo RNA editing of point mutations via RNA-guided adenosine deaminases. *Nat. Methods* **16**, 239–242 (2019).
28. Yin, H., Kauffman, K. J. & Anderson, D. G. Delivery technologies for genome editing. *Nat. Rev. Drug Discov.* **16**, 387–399 (2017).
29. Platt, R. J. et al. CRISPR-Cas9 knockin mice for genome editing and cancer modeling. *Cell* **159**, 440–455 (2014).
30. Chew, W. L. et al. A multifunctional AAV-CRISPR-Cas9 and its host response. *Nat. Methods* **13**, 868–874 (2016).
31. Teoh, P. J. et al. Aberrant hyperediting of the myeloma transcriptome by ADAR1 confers oncogenicity and is a marker of poor prognosis. *Blood* **132**, 1304–1317 (2018).
32. Vallecillo-Viejo, I. C., Liscovitch-Brauer, N., Montiel-Gonzalez, M. F., Eisenberg, E. & Rosenthal, J. J. C. Abundant off-target edits from site-directed RNA editing can be reduced by nuclear localization of the editing enzyme. *RNA Biol.* **15**, 104–114 (2018).
33. Mays, L. E. & Wilson, J. M. The complex and evolving story of T cell activation to AAV vector-encoded transgene products. *Mol. Ther.* **19**, 16–27 (2011).
34. Wagner, D. L. et al. High prevalence of *Streptococcus pyogenes* Cas9-reactive T cells within the adult human population. *Nat. Med.* **25**, 242–248 (2019).
35. Simhadri, V. L. et al. Prevalence of pre-existing antibodies to CRISPR-associated nuclease Cas9 in the USA population. *Mol. Ther. Methods Clin. Dev.* **10**, 105–112 (2018).
36. Charlesworth, C. T. et al. Identification of preexisting adaptive immunity to Cas9 proteins in humans. *Nat. Med.* **25**, 249–254 (2019).
37. Haapaniemi, E., Botla, S., Persson, J., Schmierer, B. & Taipale, J. CRISPR-Cas9 genome editing induces a p53-mediated DNA damage response. *Nat. Med.* **24**, 927–930 (2018).
38. Ihry, R. J. et al. p53 inhibits CRISPR-Cas9 engineering in human pluripotent stem cells. *Nat. Med.* **24**, 939–946 (2018).
39. Woolf, T. M., Chase, J. M. & Stinchcomb, D. T. Toward the therapeutic editing of mutated RNA sequences. *Proc. Natl Acad. Sci. USA* **92**, 8298–8302 (1995).
40. Merkle, T. et al. Precise RNA editing by recruiting endogenous ADARs with antisense oligonucleotides. *Nat. Biotechnol.* **37**, 133–138 (2019).
41. Zheng, Y., Lorenzo, C. & Beal, P. A. DNA editing in DNA/RNA hybrids by adenosine deaminases that act on RNA. *Nucleic Acids Res.* **45**, 3369–3377 (2017).
42. East-Seletsky, A. et al. Two distinct RNase activities of CRISPR-C2c2 enable guide-RNA processing and RNA detection. *Nature* **538**, 270–273 (2016).
43. Daniel, C., Widmark, A., Rigardt, D. & Ohman, M. Editing inducer elements increases A-to-I editing efficiency in the mammalian transcriptome. *Genome Biol.* **18**, 195 (2017).
44. Chen, C. X. et al. A third member of the RNA-specific adenosine deaminase gene family, ADAR3, contains both single- and double-stranded RNA binding domains. *RNA* **6**, 755–767 (2000).
45. Savva, Y. A., Rieder, L. E. & Reenan, R. A. The ADAR protein family. *Genome Biol.* **13**, 252 (2012).
46. Nishikura, K. A-to-I editing of coding and non-coding RNAs by ADARs. *Nat. Rev. Mol. Cell Biol.* **17**, 83–96 (2016).
47. Landrum, M. J. et al. ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res.* **44**, D862–D868 (2016).
48. Floquet, C., Deforges, J., Rousset, J. P. & Bidou, L. Rescue of non-sense mutated p53 tumor suppressor gene by aminoglycosides. *Nucleic Acids Res.* **39**, 3350–3362 (2011).
49. Kern, S. E. et al. Identification of p53 as a sequence-specific DNA-binding protein. *Science* **252**, 1708–1711 (1991).
50. Doubrovin, M. et al. Imaging transcriptional regulation of p53-dependent genes with positron emission tomography in vivo. *Proc. Natl Acad. Sci. USA* **98**, 9300–9305 (2001).
51. Ou, L. et al. ZFN-mediated in vivo genome editing corrects murine Hurler syndrome. *Mol. Ther.* **27**, 178–187 (2019).
52. Fire, A. et al. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806–811 (1998).
53. Zuo, E. et al. Cytosine base editor generates substantial off-target single-nucleotide variants in mouse embryos. *Science* **364**, 289–292 (2019).
54. Jin, S. et al. Cytosine, but not adenine, base editors induce genome-wide off-target mutations in rice. *Science* **364**, 292–295 (2019).
55. Kim, D., Kim, D. E., Lee, G., Cho, S. I. & Kim, J. S. Genome-wide target specificity of CRISPR RNA-guided adenine base editors. *Nat. Biotechnol.* **37**, 430–435 (2019).
56. Grunewald, J. et al. Transcriptome-wide off-target RNA editing induced by CRISPR-guided DNA base editors. *Nature* **569**, 433–437 (2019).

Acknowledgements

We acknowledge the staff of the BIOPIC High-throughput Sequencing Center (Peking University) and Genetron Health for their assistance in NGS analysis, the National Center for Protein Sciences (Beijing) and the core facilities at the School of Life Sciences (Peking University, X. Zhang, F. Wang and L. Du) for help with Fluorescence Activated Cell Sorting. We thank the High-Performance Computing Platform at Peking University for providing platforms of NGS data analysis. We thank M. Mo for technical assistance, J. Wang for providing plasmids encoding disease-relevant genes and primary cells and we also thank Z. Jiang for providing the mouse melanoma cell line B16. This project was supported by funds from Beijing Municipal Science & Technology Commission (grant no. Z181100001318009), the National Science Foundation of China (no. 31430025), Beijing Advanced Innovation Center for Genomics at Peking University and the Peking-Tsinghua Center for Life Sciences (to W.W.); the National Science Foundation of China (no. 31870893) and the National Major Science & Technology Project for Control and Prevention of Major Infectious Diseases in China (no. 2018ZX10301401, to Z.Z.) and the Beijing Nova Program (no. Z181100006218042, to P.Y.).

Author contributions

W.W. conceived and supervised the project. W.W., L.Q., Z.Y., S.Z., C.W., Z.C. and Z.Z. designed the experiments. L.Q., Z.Y., C.W., S.Z., Z.C. and P.Y. performed the experiments with the help from F.T., Y.B. and Y.Z. Y.Y. conducted all the sample preparation for NGS. Z.Y. and Z.L. performed the data analysis. L.Q., S.Z., Z.Z. and W.W. wrote the manuscript with the help of all other authors.

Competing interests

A patent has been filed relating to the data presented. W.W. is a founder andscientific adviser for EdiGene.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41587-019-0178-z.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to W.W.

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

Methods

Plasmids construction. For the three versions of dual fluorescence reporters (Reporter-1, -2 and -3), *mCherry* and *EGFP* (the start codon ATG of EGFP was deleted) coding sequences were PCR amplified, digested using BsmBI (ThermoFisher Scientific, ER0452), followed by T4 DNA ligase (NEB, M0202L)-mediated ligation with GGGGS linkers. The ligation product was subsequently inserted into the pLenti-CMV-MCS-PURO backbone.

For the dLbuCas13-ADAR$_{DD}$ (E1008Q) expressing construct, the ADAR1$_{DD}$ gene was amplified from the ADAR1^{p150} construct (a gift from J. Han's laboratory, Xiamen University). The dLbuCas13 gene was amplified by PCR from the Lbu_C2c2_R472A_H477A_R1048A_ H1053A plasmid (Addgene no. 83485). The ADAR1$_{DD}$ (hyperactive E1008Q variant) was generated by overlap-PCR and then fused to dLbuCas13. The ligation products were inserted into the pLenti-CMV-MCS-BSD backbone.

For arRNA-expressing construct, the sequences of arRNAs were synthesized and golden-gate cloned into the pLenti-sgRNA-lib 2.0 (Addgene no. 89638) backbone, and the transcription of arRNA was driven by hU6 promoter. For the ADAR-expressing constructs, the full-length ADAR1^{p110} and ADAR1^{p150} were PCR amplified from the ADAR1^{p150} construct, and the full-length ADAR2 were PCR amplified from the ADAR2 construct (a gift from J. Han's laboratory, Xiamen University). The amplified products were then cloned into the pLenti-CMV-MCS-BSD backbone. The dPspCas13b-ADAR2$_{DD}$-E488Q plasmid was purchased from Addgene (no. 103849).

For the constructs expressing genes with pathogenic mutations, full-length coding sequences of *TP53* (ordered from Vigenebio) and another six disease-relevant genes (*COL3A1*, *BMPR2*, *AHI1*, *FANCC*, *MYBPC3* and *IL2RG*, gifts from J. Wang's laboratory, Institute of Pathogen Biology, Chinese Academy of Medical Sciences) were amplified from the constructs encoding the corresponding genes with introduction of G-to-A mutations through mutagenesis PCR. The amplified products were cloned into the pLenti-CMV-MCS-mCherry backbone through the Gibson cloning method[57].

Cell culture. The HeLa and B16 cell lines came from Z. Jiang's laboratory (Peking University) and the HEK293T cell line was from C. Zhang's laboratory (Peking University). The RD cell line came from J. Wang's laboratory (Institute of Pathogen Biology, Peking Union Medical College & Chinese Academy of Medical Sciences). SF268 cell lines were from the Cell Center, Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences. A549 and SW13 cell lines were from EdiGene Inc. HepG2, HT29, NIH3T3 and MEF cell lines were maintained in our laboratory at Peking University. These mammalian cell lines were cultured in Dulbecco's Modified Eagle Medium (Corning, 10-013-CV) with 10% fetal bovine serum (FBS) (CellMax, SA201.02), additionally supplemented with 1% penicillin–streptomycin under 5% CO_2 at 37 °C. Unless otherwise described, cells were transfected with the X-tremeGENE HP DNA transfection reagent (Roche, 06366546001) according to the manufacturer's instructions.

The human primary pulmonary fibroblasts (no. 3300) and human primary bronchial epithelial cells (no. 3210) were purchased from ScienCell Research Laboratories, Inc. and were cultured in Fibroblast Medium (ScienCell, no. 2301) and Bronchial Epithelial Cell Medium (ScienCell, no. 3211), respectively. Both media were supplemented with 15% FBS (BI) and 1% penicillin–streptomycin. The primary GM06214 and GM01323 cells were ordered from Coriell Institute for Medical Research and cultured in Dulbecco's Modified Eagle Medium (Corning, 10-013-CV) with 15% FBS (BI) and 1% penicillin–streptomycin. All cells were cultured under 5% CO_2 at 37 °C.

Isolation and culture of human primary T cells. Primary human T cells were isolated from leukapheresis products from healthy human donor. Briefly, peripheral blood mononuclear cells (PBMCs) were isolated by Ficoll centrifugation (Dakewei, AS1114546), and T cells were isolated by magnetic negative selection using an EasySep Human T Cell Isolation Kit (STEMCELL, 17951) from PBMCs. After isolation, T cells were cultured in X-vivo15 medium, 10% FBS and IL2 (1,000 U ml^{-1}) and stimulated with CD3/CD28 DynaBeads (ThermoFisher, 11131D) for 2 d. Leukapheresis products from healthy donors were acquired from AllCells LLC China. All healthy donors provided informed consent.

Cell line construction. For the stable reporter cell lines, the reporter constructs (pLenti-CMV-MCS-PURO backbone) were cotransfected into HEK293T cells, together with two viral packaging plasmids, pR8.74 and pVSVG. After 72 h, the supernatant virus was collected and stored at −80 °C. The HEK293T cells were infected with lentivirus, then mCherry-positive cells were sorted via fluorescence-activated cell sorting (FACS) and cultured to select a single clone cell lines stably expressing dual fluorescence reporter system without detectable EGFP background. The HEK293T *ADAR1*$^{-/-}$ and *TP53*$^{-/-}$ cell lines were generated according to a previously reported method[58]. ADAR1-targeting single-guide RNA and PCR amplified donor DNA containing CMV-driven puromycin resistant gene were cotransfected into HEK293T cells. Then cells were treated with puromycin 7 d after transfection. Single clones were isolated from puromycin resistant cells followed by verification through sequencing and western blot.

RNA editing of endogenous or exogenous-expressed transcripts. For assessing RNA editing on the dual fluorescence reporter, HEK293T cells or HEK293T *ADAR1*$^{-/-}$ cells were seeded in six-well plates (6×10^5 cells per well). After 24 h, cells were cotransfected with 1.5 μg reporter plasmids and 1.5 μg arRNA plasmids. To examine the effect of ADAR1^{p110}, ADAR1^{p150} or ADAR2 protein expression, the editing efficiency was assayed by EGFP$^+$ ratio and deep sequencing.

HEK293T *ADAR1*$^{-/-}$ cells were seeded in 12-well plates (2.5×10^5 cells per well), then 24 h later, cells were cotransfected with 0.5 μg of reporter plasmids, 0.5 μg arRNA plasmids and 0.5 μg ADAR1/2 plasmids (pLenti backbone as control). The editing efficiency was assayed by EGFP$^+$ ratio and deep sequencing.

To assess RNA editing on endogenous mRNA transcripts, HEK293T cells were seeded in six-well plates (6×10^5 cells per well). Twenty-four hours later, cells were transfected with 3 μg of arRNA plasmids. The editing efficiency was assayed by deep sequencing.

To assess RNA editing efficiency in multiple cell lines, $8–9\times10^4$ (RD, SF268, HeLa) or 1.5×10^5 (HEK293T) cells were seeded in 12-well plates. For cells difficult to transfect, such as HT29, A549, HepG2, SW13, NIH3T3, MEF and B16, $2–2.5\times10^5$ cells were seeded in a six-well plate. Twenty-four hours later, reporters and arRNAs plasmid were cotransfected into these cells. The editing efficiency was assayed by an EGFP$^+$ ratio.

To evaluate the EGFP$^+$ ratio, at 48–72 h post transfection, cells were sorted and collected by FACS analysis. The mCherry signal was served as a fluorescent selection marker for the reporter/arRNA-expressing cells, and the percentages of EGFP$^+$/mCherry$^+$ cells were calculated as the readout for editing efficiency.

For NGS quantification of the A-to-I editing rate, at 48–72 h post transfection, cells were sorted and collected by FACS assay and were then subjected to RNA isolation (TIANGEN, DP420). Then, the total RNAs were reverse-transcribed into cDNA via PCR with reverse transcription (RT–PCR) (TIANGEN, KR103-04), and the targeted locus was PCR amplified with the corresponding primers (listed in Supplementary Table 5). The PCR products were purified for Sanger sequencing or NGS (Illumina HiSeq X Ten).

RNA editing analysis for targeted sites. For deep sequencing analysis, an index was generated using the targeted site sequence (upstream and downstream 20-nt) of arRNA covering sequences. Reads were aligned and quantified using BWA (v.0.7.10-r789). Alignment BAMs were then sorted by Samtools, and RNA editing sites were analyzed using REDitools (v.1.0.4). The parameters are as follows: -U [AG or TC] -t 8 -n 0.0 -T 6-6 -e -d -u. All the significant A-to-G conversion within the arRNA targeting region calculated by Fisher's exact test (P value < 0.05) were considered as edits by arRNA. The conversions except for targeted adenosine were off-target edits. The mutations that appeared in control and experimental groups simultaneously were considered to be due to single nucleotide polymorphism.

Transcriptome-wide RNA-sequencing analysis. The Ctrl RNA$_{151}$ or arRNA$_{151}$-PPIB-expressing plasmids with the blue fluorescent protein (BFP) expression cassette were transfected into HEK293T cells. The BFP$^+$ cells were enriched by FACS 48 h after transfection, and RNAs were purified with RNAprep Pure Micro kit (TIANGEN, DP420). The mRNA was then purified using NEBNext Poly(A) mRNA Magnetic Isolation Module (New England Biolabs, E7490), processed with the NEBNext Ultra II RNA Library Prep Kit for Illumina (New England Biolabs, E7770), followed by deep sequencing analysis using Illumina HiSeq X Ten platform (2×150-base pair paired end; 30G for each sample). To exclude nonspecific effect caused by transfection, we included the mock group in which we only treated cells with transfection reagent. Each group contained four replications.

The bioinformatics analysis pipeline followed the work by Vogel et al.[22]. The quality control of analysis was conducted by using FastQC, and quality trim was based on Cutadapt (the first 6-bp for each read were trimmed and up to 20-bp were quality trimmed). AWK scripts were used to filtered out the introduced arRNAs. After trimming, reads with lengths shorter than 90-nt were filtered out. Subsequently, the filtered reads were mapped to the reference genome (GRCh38-hg38) by STAR software[59]. We used the GATK Haplotypcaller[60] to call the variants. The raw VCF (variant call format) files generated by GATK were filtered and annotated by GATK VariantFiltration, bcftools and ANNOVAR[61]. The variants in dbSNP, 1,000 Genome[62] and EVS were filtered out. The shared variants in four replicates of each group were then selected as the RNA editing sites. The RNA editing level of the mock group was viewed as the background, and the global targets of Ctrl RNA$_{151}$ and arRNA$_{151}$-PPIB were obtained by subtracting the variants in the mock group.

To assess whether LEAPER perturbs natural editing homeostasis, we analyzed the global editing sites shared by the mock group and arRNA$_{151}$-PPIB group (or Ctrl RNA$_{151}$ group). The differential RNA editing rates at native A-to-I editing sites were assessed using Pearson's correlation coefficient analysis. Pearson correlations of the editing rate between the mock group and arRNA$_{151}$-PPIB group (or Ctrl RNA$_{151}$ group) were calculated and annotated in Fig. 6.

$$\rho(X,Y)=\frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X\sigma_Y}$$

where X means the editing rate of each site in the mock group; Y means the editing rate of each site in the Ctrl RNA$_{151}$ group (Fig. 6a) or arRNA$_{151}$-PPIB group

(Fig. 6b); σ_X is the standard deviation of X; σ_Y is the standard deviation of Y; μ_X is the mean of X; μ_Y is the mean of Y and E is the expectation.

The RNA-Seq data were analyzed for the interrogation of possible transcriptional changes induced by RNA editing events. The analysis of transcriptome-wide gene expression was performed using HISAT2 (v.2.1.0) and STRINGTIE (v.1.3.4d) software[63]. We used Cutadapt (v.1.16) and FastQC (v.0.11.8) for the quality control of the sequencing data. Then the sequencing reads were mapped to reference genome (GRCh38-hg38) using HISAT2, followed by Pearson's correlation coefficient analysis as mentioned above. The FPKM value was calculated with STRINGTIE.

We then used DESeq2 (v.1.18.1)[64] to calculate the differential gene expression with the FPKM expression data. Genes with an FPKM value less than one in all the four replicates were set at one for correction. Genes with adjusted P values less than 0.01 and $\log_2$(fold change) greater than two were filtered and regarded as significantly differentially expressed genes.

Western blot. We used the mouse monoclonal primary antibodies, respectively, against ADAR1 (Santa Cruz, sc-271854), ADAR2 (Santa Cruz, sc-390995), ADAR3 (Santa Cruz, sc-73410), p53 (Santa Cruz, sc-99), KRAS (Sigma, SAB1404011), GAPDH (Santa Cruz, sc-47724) and β-tubulin (CWBiotech, CW0098). The HRP-conjugated goat anti-mouse IgG (H + L, 115-035-003) secondary antibody was purchased from Jackson ImmunoResearch. Then, 2×10^6 cells were sorted to be lysed and an equal amount of each lysate was loaded for SDS–PAGE. Then, sample proteins were transferred onto polyvinylidene difluoride membrane (Bio-Rad Laboratories) and immunoblotted with primary antibodies against one of the ADAR enzymes (anti-ADAR1, 1:500; anti-ADAR2, 1:100; anti-ADAR3, 1:800), followed by secondary antibody incubation (1:10,000) and exposure. The β-tubulins were re-probed on the same polyvinylidene difluoride membrane after stripping of the ADAR proteins with the stripping buffer (CWBiotech, CW0056). The experiments were repeated three times. The semi-quantitative analysis was done with Image Lab software.

Cytokine expression assay. HEK293T cells were seeded on 12-well plates (2×10^5 cells per well). When approximately 70% confluent, cells were transfected with 1.5 μg of arRNA. As a positive control, 1 μg of poly(I:C) (Invitrogen, tlrl-picw) was transfected. Forty-eight hours later, cells were collected and subjected to RNA isolation (TIANGEN, DP430). Then, the total RNAs were reverse-transcribed into cDNA via RT–PCR (TIANGEN, KR103-04), and the expression of IFN-β and IL-6 were measured by qPCR (TAKARA, RR820A). The sequences of the primers are listed in the Supplementary Table 5.

Transcriptional regulatory activity assay of p53. The *TP53*^{W53X} cDNA-expressing plasmids and arRNA-expressing plasmids were cotransfected into HEK293T *TP53*$^{-/-}$ cells, together with p53-Firefly-luciferase *cis*-reporting plasmids (YRGene, VXS0446) and Renilla-luciferase plasmids (a gift from Z. Jiang's laboratory, Peking University) for detecting the transcriptional regulatory activity of p53. Then, 48 h later, the cells were collected and assayed with the Promega Dual-Glo Luciferase Assay System (Promega, E4030) according to the manufacturer's protocol. Briefly, 150 μl Dual-Glo Luciferase Reagent was added to the collected cell pellet and, 30 min later, the Firefly luminescence was measured by adding 100 μl Dual-Glo Luciferase Reagent (cell lysis) to a 96-well white plate by the Infinite M200 reader (TECAN). After 30 min, 100 μl Dual-Glo stop and Glo Reagent were sequentially added to each well to measure the Renilla luminescence and calculate the ratio of Firefly luminescence to Renilla luminescence.

Electroporation in primary cells. For arRNA-expressing plasmids electroporation in the human primary pulmonary fibroblasts or human primary bronchial epithelial cells, 20 μg plasmids were electroporated with Nucleofector 2b Device (Lonza) and Basic Nucleofector Kit (Lonza, VPI-1002), and the electroporation program was U-023. For arRNA-expressing plasmids electroporation in human primary T cells, 20 μg plasmids were electroporated into human primary T with Nucleofector 2b Device (Lonza) and Human T cell Nucleofector Kit (Lonza, VPA-1002), and the electroporation program was T-024. Forty-eight hours post electroporation, cells were sorted and collected by FACS assay and were then subjected to the following deep sequencing for targeted RNA editing assay. The electroporation efficiency was normalized according to the fluorescence marker.

For the chemosynthetic arRNA or control RNA electroporation in human primary T cells or primary GM06214 cells, RNA oligonucleotide was dissolved in 100 μl opti-MEM medium (Gibco, 31985070) with the final concentration 2 μM. Then 1×10^6 GM06214 cells or 3×10^6 T cells were resuspended with the above electroporation mixture and electroporated with Agile Pulse In Vivo device (BTX) at 450 V for 1 ms. Then the cells were transferred to warm culture medium for the following assays.

IDUA catalytic activity assay. The gathered cell pellet was resuspended and lysed with 28 μl 0.5% Triton X-100 in 1 × PBS buffer on ice for 30 min. Then 25 μl of the cell lysis was added to 25 μl 190 μM 4-methylumbelliferyl-α-L-iduronidase substrate (Cayman, 2A-19543-500), which was dissolved in 0.4 M sodium formate buffer containing 0.2% Triton X-100, pH 3.5 and incubated for 90 min at 37 °C in the dark. The catalytic reaction was quenched by adding 200 μl 0.5 M NaOH/Glycine buffer, pH 10.3 and then centrifuged for 2 min at 4 °C. The supernatant was transferred to a 96-well plate, and fluorescence was measured at 365 nm excitation wavelength and 450 nm emission wavelength with Infinite M200 reader (TECAN).

Statistics. An unpaired two-sided Student's t-test was implemented for group comparison as indicated in the figure legends. For transcriptome-wide RNA-seq data, DESeq2 (v.1.18.1) was used for analyzing statistical significance. Statistical analyses were performed with R and Prism 6 (GraphPad Software, Inc.).

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

Data availability

All data presented in this manuscript are available from the corresponding author upon reasonable request. Transcriptome-wide RNA-seq data are accessible via the NCBI Sequence Read Archive database with accession code PRJNA544353.

References

57. Gibson, D. G. et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
58. Zhou, Y., Zhang, H. & Wei, W. Simultaneous generation of multi-gene knockouts in human cells. *FEBS Lett.* **590**, 4343–4353 (2016).
59. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
60. Van der Auwera, G. A. et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr. Protoc. Bioinformatics* **43**, 11.10.1–33 (2013).
61. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010).
62. Genomes Project, C. et al. An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56–65 (2012).
63. Pertea, M., Kim, D., Pertea, G. M., Leek, J. T. & Salzberg, S. L. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* **11**, 1650–1667 (2016).
64. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

nature research

Corresponding author(s): Wensheng Wei

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	No software was used.
Data analysis	REDitools version 1.0.4 was used for targeted RNA editing analysis. FastQC, STAR, GATK Haplotypcaller, GATK VariantFiltration, bcftools, ANNOVAR, HISAT2, STRINGTIE and DESeq2 were adopted for transcriptome-wide RNA-sequencing analysis. The minimum free energy of double-stranded RNA was predicted by RNAhybrid. R script was used for ploting figure. GraphPad Prism 6 was used for basic statistical analysis and graph production.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Data availability. Data are accessible via the NCBI Sequence Read Archive database with accession code PRJNA544353, or also are available from the corresponding author upon reasonable request.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	In this study, targeted RNA editing of reporter transcripts or endogenous transcripts was done with independent experiments performed in parallel, and the numble of replicates was listed in the text and figure legends. Transcriptome-wide RNA-sequencing was done with four independent experiments performed in parallel. And the group sizes in other experiments were selected based on the prior knowledge of variation (e.g. Western blot, qPCR or FACS analysis).
Data exclusions	No data were excluded.
Replication	The numble of replications is always mentioned in text, methods and figure legends. All attempts at replication were successful.
Randomization	During FACS experiments for sorting transfected cells, all the cells were transfected and sorted via FACS according to the corresponding fluorescence maker, and untreated group and mock group (fluorescence-negative) were conducted for gating.
Blinding	No blinding was performed due to the involvement of several experimentators.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	☐ Unique biological materials
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Anti-ADAR1 antibody (Santa Cruz, sc271854); Anti-ADAR2 antibody (Santa Cruz, sc390995); Anti-ADAR3 antibody (Santa Cruz, sc73410); Anti-p53 antibody (Santa Cruz, sc-99); Anti-KRAS antibody (Sigma, SAB1404011); Anti-GAPDH antibody (Santa Cruz, sc47724); Anti-beta-tubulin antibody (CWBiotech, CW0098).
Validation	All antibodies used in this study were validated by the manufacturer, and the western blot experiments were performed according to the manufacturer's instruction. And the western blot data were provided in the manuscript.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	The HeLa and B16 cell lines were from Z. Jiang's laboratory (Peking University). And the HEK293T cell line was from C. Zhang's laboratory (Peking University). RD cell line was from J Wang's laboratory (Institute of Pathogen Biology, Peking Union Medical College & Chinese Academy of Medical Sciences). SF268 cell lines was from Cell Center, Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences. A549, SW13, HepG2, HT29, NIH3T3, and MEF cell lines were maintained in W Wei's laboratory (Peking University). Human primary pulmonary fibroblast or human primary bronchial epithelial cell were purchased from ScienCell Research Laboratories, Inc. Primary human T cells were isolated from leukapheresis products, which were purchased from AllCells LLC China.
Authentication	STR analysis was used for cell line authentication.
Mycoplasma contamination	All cells were tested negative for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Cells were transfected using the X-tremeGENE HP DNA transfection reagent (06366546001; Roche, Mannheim, German), according to the supplier's protocols. About 48 to 72 hours later, cells were digested with trypsin and collected for the following FACS according to the fluorescence maker (mCherry, EGFP or BFP).
Instrument	BD Aria SORP and BD LSRFortessa SORP
Software	BD FACSDiva
Cell population abundance	The fluorescence maker was encoded in the expression plasmids for sorting transfected cells, and about 7×10^5 cells were collected for further process.
Gating strategy	Firstly, the starting cell population were selected according to SSC-A, FSC-A, SSC-W and SSC-H gates. Then fluorescence-positive cell population were determined according to the fluorescence-negative population control (untreated group or mock group).

☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.