RNA editing with CRISPR-Cas13

David B. T. Cox1,2,3,4,†, Jonathan S. Gootenberg1,2,3,4,†, Omar O. Abudayyeh1,2,3,4,†, Brian Franklin1,2,3,4, Max J. Kellner1,2,3,4, Julia Joung1,2,3,4, Feng Zhang1,2,3,4,†

Nucleic acid editing holds promise for treating genetic disease, particularly at the RNA level, where disease-relevant sequences can be rescued to yield functional protein products. Type VI CRISPR-Cas systems contain the programmable single-effecter RNA-guided ribonuclease Cas13. We profiled type VI systems in order to engineer a Cas13 ortholog capable of robust knockingdown and demonstrated RNA editing by using catalytically inactive Cas13 (dCas13) to direct adenosine-to-inosine deaminase activity by ADAR2 (adenosine deaminase acting on RNA type 2) to transcripts in mammalian cells. This system, referred to as RNA Editing for Programmable A to I Replacement (REPAIR), which has no strict sequence constraints, can be used to edit full-length transcripts containing pathogenogenic mutations. We further engineered this system to create a high-specificity variant and minimized the system to facilitate viral delivery. REPAIR presents a promising RNA-editing platform with broad applicability for research, therapeutics, and biotechnology.

Precise nucleic acid–editing technologies are valuable for studying cellular function and as novel therapeutics. Current editing tools, based on programmable nucleases such as the prokaryotic CRISPR-associated nucleases Cas9 (1–4) or Cpf1 (5), have been widely adopted for mediating targeted DNA cleavage, which in turn drives targeted gene disruption through nonhomologous end joining (NHEJ) or precise gene editing through template-dependent homology-directed repair (HDR) (6). NHEJ uses host machineries that are active in both dividing and post-mitotic cells and provides efficient gene disruption by generating a mixture of insertion or deletion (indel) mutations that can lead to frame shifts in protein-coding genes. HDR, in contrast, is mediated by host machineries whose expression is largely limited to replicating cells. Accordingly, the development of gene-editing capabilities for post-mitotic cells remains a major challenge. DNA base editors, consisting of a fusion between Cas9 nickase and cytidine deaminase, can mediate efficient cytidine-to-uridine conversions within a target window and substantially reduce the formation of double-strand breaks–induced indels (7, 8). However, the potential of DNA base editors are limited by the requirement of Cas9 for a protospacer adjacent motif (PAM) at the editing site (9). Here, we describe the development of a precise and flexible RNA base editing technology using the type VI CRISPR-associated RNA-guided ribonuclease (RNase) Cas13 (10–13).

Cas13 enzymes have two higher eukaryotes and prokaryotes nucleotide-binding (HEPN) endorNase domains that mediate precise RNA cleavage with a preference for targets with protospacer flanking sites (PFSs) observed biochemically and in bacteria (10, 11). Three Cas13 protein families have been identified to date: Cas13a (previously known as C2c2), Cas13b, and Cas13c (12, 13). We recently reported that Cas13a enzymes can be adapted as tools for nucleic acid detection (14) as well as mammalian and plant cell RNA knockdown and transcript tracking (15), and observed that the biochemical PFS was not required for RNA interference with Cas13a (15). The programmable nature of Cas13 enzymes makes them an attractive starting point to develop tools for RNA binding and perturbation applications.

The adenosine deaminase acting on RNA (ADAR) family of enzymes mediates endogenous editing of transcripts via hydrolytic deamination of adenosine to inosine, a nucleobase that is functionally equivalent to guanosine in translation and splicing (16, 17). There are two functional human ADAR orthologs, ADAR1 and ADAR2, which consist of N-terminal double-stranded RNA–binding domains and a C-terminal catalytic deamination domain. Endogenous target sites of ADAR1 and ADAR2 contain substantial double-stranded identity, and the catalytic domains require duplexed regions for efficient editing in vitro and in vivo (18, 19). The ADAR catalytic domain is capable of deamidating target adenosines without any protein cofactors in vitro (20). ADAR1 has been found to target mainly repetitive regions, whereas ADAR2 mainly targets nonrepetitive coding regions (17). Although ADAR proteins have preferred motifs for editing that could restrict the potential flexibility of targeting, hyperactive mutants, such as ADAR2(E488Q) (21), relax sequencing constraints and increase adenosine-to-inosine editing rates. (Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.) In the mutants, other amino acids were substituted at certain locations; for example, E488Q indicates that glutamic acid at position 488 was replaced by glutamine. ADARs preferentially deaminate adenosines mispaired with cytidine bases in RNA duplexes (22), providing a promising opportunity for precise base editing. Although previous approaches have engineered targeted ADAR fusions via RNA guides (23–26), the specificity of these approaches has not been reported, and their respective targeting mechanisms rely on RNA-RNA hybridization without the assistance of protein partners that may enable target recognition and stringency.

We assessed a subset of the family of Cas13 enzymes for RNA knockdown activity in mammalian cells and identified the Cas13b ortholog from Prevotella sp. P5-125 (PspCas13b) as the most efficient and specific for mammalian cell applications. We then fused the ADAR2 deaminase domain with the E488Q mutation (ADAR2DQ) to catalytically inactive PspCas13b and demonstrated RNA editing for programmable A to I (G) replacement (REPAIR) of reporter and endogenous transcripts as well as disease-relevant mutations. Last, we used a rational mutagenesis scheme to improve the specificity of Cas13b-ADAR2DQ fusions in order to generate REPAIRv2 with more than 919-fold higher specificity.

Comprehensive characterization of Cas13 family members in mammalian cells

We previously developed Leptotrichia wadei Cas3a (LwaCas3a) for mammalian knockdown applications, but it required a monomeric superfolder green fluorescent protein (msfGFP) stabilization domain for efficient knockdown, and although the specificity was high, knockdown levels were not consistently below 50% (15). We sought to identify a more robust RNA-targeting CRISPR system by characterizing a genetically diverse set of Cas3 family members in order to assess their RNA knockdown activity in mammalian cells (Fig. 1A). We generated mammalian codon-optimized versions of multiple Cas3 proteins, including 21 orthologs of Cas3a, 15 of Cas3b, and seven of Cas3c, and cloned them into an expression vector with N- and C-terminal nuclear localization signal (NLS) sequences and a C-terminal msfGFP to enhance protein stability (table S1). To assay interference in mammalian cells, we designed a dual-reporter construct expressing the independent Gausssia (Glue) and Cupridina (Clue) luciferases under separate promoters, allowing one luciferase to function as a measure of Cas3 interference activity and the other to serve as an internal control. For each Cas3 ortholog, we designed PFS-compatible guide RNAs, using the Cas3b PFS motifs derived from an ampicillin interference assay (fig. S1, table S2, and supplementary text).
Fig. 1. Characterization of a highly active Cas13b ortholog for RNA knockdown. (A) Schematic of stereotypical Cas13 loci and corresponding CRISPR RNA (crRNA) structure. (B) Evaluation of 19 Cas13a, 15 Cas13b, and seven Cas13c orthologs for luciferase knockdown by use of two different guides. Orthologs with efficient knockdown by using both guides are labeled with their host organism name. Values are normalized to a nontargeting guide designed against the Escherichia coli LacZ transcript, with no homology to the human transcriptome. (C) PspCas13b and LwaCas13a knockdown activity (as measured by luciferase activity) by using tiling guides against Gluc. Values represent mean ± SEM. Nontargeting guide is the same as in (B). (D) PspCas13b and LwaCas13a knockdown activity (as measured by luciferase activity) by using tiling guides against Gluc. Values represent mean ± SEM. Nontargeting guide is the same as in (B). (E) Expression levels in log₂([transcripts per million (TPM)+1] values of all genes detected in RNA-seq libraries of nontargeting control (x axis) compared with Gluc-targeting condition (y axis) for LwaCas13a (red) and shRNA (black). Shown is the mean of three biological replicates. The Gluc transcript data point is labeled. Nontargeting guide is the same as in (B). (F) Expression levels in log₂([transcripts per million (TPM)+1] values of all genes detected in RNA-seq libraries of nontargeting control (x axis) compared with Gluc-targeting condition (y axis) for PspCas13b (blue) and shRNA (black). Shown is the mean of three biological replicates. The Gluc transcript data point is labeled. Nontargeting guide is the same as in (B). (G) Number of significant off-targets from Gluc knockdown for LwaCas13a, PspCas13b, and shRNA from the transcriptome-wide analysis in (E) and (F). and the 3′ H (not G) PPF from previous reports of Cas13a activity (10).

We transfected human embryonic kidney (HEK) 293FT cells with Cas13-expression, guide RNA, and reporter plasmids and then quantified levels of Cas13 expression and the targeted Gluc 48 hours later (Fig. 1B and fig. S2A). Testing two guide RNAs for each Cas13 ortholog revealed a range of activity levels, including five Cas13b orthologs with similar or increased interference across both guide RNAs relative to the recently characterized LwaCas13a (Fig. 1B), and we observed only a weak correlation between Cas13 expression and interference activity (fig. S2, B to D). We selected the top five Cas13b orthologs and the top two Cas13a orthologs for further engineering.

We next tested Cas13-mediated knockdown of Gluc without msfGFP to select orthologs that do not require stabilization domains for robust activity. We hypothesized that Cas13 activity could be affected by subcellular localization, as we previously reported for optimization of LwaCas13a (15). Therefore, we tested the interference activity of the seven selected Cas13 orthologs C-terminally fused to one of six different localization tags without msfGFP. Using the luciferase reporter assay, we identified the top three Cas13b designs with the highest level of interference activity: Cas13b from Prevotella sp. P5-125 (PspCas13b) and Cas13b from Porphyromonas gulae (PguCas13b) C-terminally fused to the HIV Rev nuclear export sequence (NES), and Cas13b from Riemerella anatipestifer (RanCas13b) C-terminally fused to the mitogen-activated protein kinase NES (fig. S3A). To further distinguish activity levels of the top orthologs, we compared the three optimized Cas13b constructs with the optimal LwaCas13a-msfGFP fusion and to short hairpin–mediated RNA (shRNA) for their ability to knock down the endogenous KRAS (V-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog) transcript by using position-matched guides (fig. S3B). We observed the highest levels of interference for PspCas13b (average knockdown, 62.9%) and thus selected this for further comparison with LwaCas13a.

To more rigorously define the activity of PspCas13b and LwaCas13a, we designed position-matched guides tiling along both Gluc and Chl transcripts and assayed their activity using our luciferase reporter assay. We tested 93 and 20 position-matched guides targeting Gluc and Chl, respectively, and found that PspCas13b had consistently increased levels of knockdown relative to LwaCas13a (average of 92.3% for PspCas13b versus 40.1% knockdown for LwaCas13a) (Fig. 1, C and D).

Specificity of Cas13 mammalian interference activity

To characterize the interference specificities of PspCas13b and LwaCas13a, we designed a plasmid library of luciferase targets containing single mismatches and double mismatches throughout the target sequence and the three flanking 5′ and 3′ base pairs (fig. S3C). We transfected HEK293FT cells with either LwaCas13a or PspCas13b, a fixed guide RNA targeting the unmodified target sequence, and the mismatched target library
corresponding to the appropriate system. We then performed targeted RNA sequencing (RNA-seq) of uncleaved transcripts in order to quantify depletion of mismatched target sequences. We found that LwaCas13a and PspCas13b had a central region that was relatively intolerant to single mismatches, extending from base pairs 12 to 26 for the PspCas13b target and 13 to 24 for the LwaCas13a target (fig. S3D). Double mismatches were even less tolerated than single mutations, with little knockdown activity observed over a larger window, extending from base pairs 12 to 29 for PspCas13b and 8 to 27 for LwaCas13a in their respective targets (fig. S3E). Additionally, because there are mismatches included in the three nucleotides flanking the 5’ and 3’ ends of the target sequence, we could assess PFS constraints on Cas3 knockdown activity. Sequencing that showed that almost all PFS combinations allowed robust knockdown, indicating that a PFS constraint for interference in mammalian cells likely does not exist for either enzyme tested. These results indicate that Cas13a and Cas13b display similar sequence constraints and sensitivities against mismatches.

We next characterized the interference specificity of PspCas13b and LwaCas13a across the mRNA fraction of the transcriptome. We performed transcriptome-wide mRNA sequencing to detect significant differentially expressed genes. LwaCas13a and PspCas13b demonstrated robust knockdown of Gluc (Fig. 1, E and F) and were highly specific compared with a position-matched shRNA, which showed hundreds of off-targets (Fig. 1G), a finding consistent with our previous characterization of LwaCas13a specificity in mammalian cells (15).

**Cas13-ADAR fusions enable targeted RNA editing**

Given that PspCas13b achieved consistent, robust, and specific knockdown of mRNA in mammalian cells, we envisioned that it could be adapted as an RNA-binding platform to recruit RNA-modifying domains, such as ADARDD, for programmable RNA editing. To engineer a PspCas13b lacking nuclease activity (dPspCas13b, referred to as dCas13b hereafter), we mutated conserved catalytic residues in the HEPN domains and observed loss of luciferase RNA knockdown (fig. S4A). We hypothesized that a dCas13b-ADARDD fusion could be recruited by a guide RNA to target adenosines, with the hybridized RNA creating the required duplex substrate for ADAR activity (Fig. 2A). To enhance target adenosine deamination rates, we introduced two additional modifications to our initial RNA editing design: We introduced a mismatched cytidine opposite the target adenosine, which has been previously reported to increase deamination frequency, and fused dCas13b with the deaminase domains of human ADAR1 or ADAR2 containing hyperactivating mutations in order to enhance catalytic activity (ADAR1DD(E1008Q) (27) or ADAR2DD(E488Q)) (27).

**Fig. 2. Engineering dCas13b-ADAR fusions for RNA editing.**

(A) Schematic of RNA editing by dCas13b-ADARDD fusion proteins. dCas13b is fused to ADARDD, which naturally deaminates adenosines to inosines in dsRNA. The crRNA specifies the target site by hybridizing to the bases surrounding the target adenosine, creating a dsRNA structure for editing and recruiting the dCas13b-ADARDD fusion. A mismatched cytidine in the crRNA opposite the target adenosine enhances the editing reaction, promoting target adenosine deamination to inosine, a base that functionally mimics guanosine in many cellular reactions. (B) Schematic of *Cypridina* luciferase W85X target and targeting guide design. Deamination of the target adenosine restores the stop codon to the wild-type tryptophan. Spacer length is the region of the guide that contains homology to the target sequence. Mismatch distance is the number of bases between the 3’ end of the spacer and the mismatched cytidine. The cytidine mismatched base is included as part of the mismatch distance calculation. (C) Quantification of luciferase activity restoration for (left) dCas13b-ADAR1DD(E1008Q) and (right) dCas13b-ADAR2DD(E488Q) with tiling guides of length 30, 50, 70, or 84 nt. All guides with even mismatch distances are tested for each guide length. Values are background-subtracted relative to a 30-nt nontargeting guide that is randomized with no sequence homology to the human transcriptome. (D) Schematic of the sequencing window in which A-to-I edits were assessed for *Cypridina* luciferase W85X. Sequencing quantification of A-to-I editing for 50-nt guides targeting *Cypridina* luciferase W85X. Blue triangle indicates the targeted adenosine. For each guide, the region of duplex RNA is outlined in red. Values represent mean ± SEM. Nontargeting guide is the same as in (C).
To test the activity of dCas13b-ADAR2DD, we generated an RNA-editing reporter on Cluc by introducing a nonsense mutation [W86X (UGG→UAG)], which could functionally be repaired to the wild-type codon through A-to-I editing (Fig. 2B) and then be detected as restoration of Cluc luminescence. We evenly tiled guides with spacers 30, 50, 70, or 84 nucleotides (nt) long across the target adenosine so as to determine the optimal guide placement and design (Fig. 2C). We found that dCas13b-ADAR1DD(e1008Q) required longer guides to repair the Cluc reporter, whereas dCas13b-ADAR2DD(e488Q) was functional with all guide lengths tested (Fig. 2C). We also found that the hyperactive E488Q mutation improved editing efficiency as wild-type ADAR2DD displayed reduced luciferase restoration (fig. S4B). From that dCas13b-ADAR1DD(e1008Q) required longer guide placement and design (Fig. 2C). We found 30, 50, 70, or 84 nucleotides (nt) long across the target site and found that both guide lengths resulted in the expected A-to-I edit, with 50-nt spacers achieving higher editing percentages (Fig. 2, D and E, and fig. S4C). We also observed that 50-nt spacers had an increased propensity for editing at non-targeted adenosines within the sequencing window, likely because of increased regions of duplexed RNA (Fig. 2E and fig. S4C).

We next targeted an endogenous gene, PPIB. We designed 50-nt spacers tiling PPIB and found that we could edit the PPIB transcript with up to 28% editing efficiency (fig. S4D). To test whether REPAIR could be further optimized, we modified the linker between dCas13b and ADAR2DD(e488Q) (fig. S4E and table S3) and found that linker choice modestly affected luciferase activity restoration. Additionally, we tested the ability of dCas13b and guide alone to mediate editing events, finding that the ADAR2DD is required for editing (fig. S5, A to D).

**Defining the sequence parameters for RNA editing**

Given that we could achieve precise RNA editing at a test site, we wanted to characterize the sequence constraints for programming the system against any RNA target in the transcriptome. Sequence constraints could arise from dCas13b-targeting limitations, such as the PFS, or from ADAR sequence preferences (28). To investigate PFS constraints on REPAIRv1, we designed a plasmid library that carries a series of four randomized nucleotides at the 5’ end of a target site on the Cluc transcript (Fig. 3A). We targeted the center adenosine within either a UAG or AAC motif and found that for both motifs, all PFSs generated detectable levels of RNA editing, with a majority of the PFSs having >50% editing at the target site (Fig. 3B). Next, we sought to determine whether the ADAR2DD in REPAIRv1 had any sequence constraints immediately flanking the targeted base, as has been reported previously for ADAR2DD (29). We tested every possible combination of 5’ and 3’ flanking nucleotides directly surrounding the target adenosine (Fig. 3C) and found that REPAIRv1 was capable of editing all motifs (Fig. 3D). Last, we analyzed whether the identity of the base opposite the target A in the spacer sequence affected editing efficiency and found that an A-C mismatch had the highest luciferase restoration, in agreement with previous reports of ADAR2 activity, with A-G, A-U, and A-A having drastically reduced REPAIRv1 activity (fig. S5E).

**Correction of disease-relevant human mutations using REPAIRv1**

to demonstrate the broad applicability of the REPAIRv1 system for RNA editing in mammalian cells, we designed REPAIRv1 guides against two disease-relevant mutations: 878G→A (APVR2 W293X) in X-linked nephrogenic diabetes insipidus and 1517G→A (FANCC W506X) in Fanconi anemia. We transfected expression constructs for cDNA of genes carrying these mutations into HEK293FT cells and tested whether REPAIRv1 could correct the mutations. Using guide RNAs containing 50-nt spacers, we were able to achieve 35% correction of APVR2 and 23% correction of FANCC (Fig. 4, A to D). We then tested the ability of REPAIRv1 to correct 34 different disease-relevant G→A mutations (table S4) and found that we were able to achieve substantial editing at 33 sites with up to 28% editing efficiency (Fig. 4E). The mutations we chose are only a fraction of the pathogenic G-to-A mutations (5739) in the ClinVar database, which also includes an additional 11,943 G-to-A variants (fig. 4F and fig. S6). Because there are no strict sequence constraints (Fig. 5), REPAIRv1 is capable of potentially editing all of these disease-relevant mutations, especially given that we observed editing regardless of the target motif (Figs. 3C and 4G).

Delivering the REPAIRv1 system to diseased cells is a prerequisite for therapeutic use, and we therefore sought to design REPAIRv1 constructs that could be packaged into therapeutically relevant viral vectors, such as adeno-associated viral (AAV) vectors. AAV vectors have a packaging limit of 4.7 kb, which cannot accommodate the large size of dCas13b-ADAR2DD (4473 base pairs (bp)) along with promoter and expression regulatory elements. To reduce the size, we tested a variety of N-terminal and C-terminal truncations of dCas13 fused to ADAR2DD(e488Q) for RNA-editing activity. We found that all C-terminal
truncations tested were still functional and able to restore luciferase signal (fig. S7), and the largest truncation, C-terminal Δ984–1090 (total size of the fusion protein, 4152 bp) was small enough to fit within the packaging limit of AAV vectors.

Transcriptome-wide specificity of REPAIRv1

Although RNA knockdown with PspCas13b was highly specific in our luciferase tilling experiments, we observed off-target adenosine editing within the guide:target duplex (Fig. 2E). To see whether this was a widespread phenomenon, we tiled an endogenous transcript, KRAS, and measured the degree of off-target editing near the target aden-

Fig. 4. Correction of disease-relevant mutations with REPAIRv1.

(A) Schematic of target and guide design for targeting AVPR2 878G→A. (B) The 878G→A mutation (indicated by blue triangle) in AVPR2 is corrected to varying levels by using REPAIRv1 with three different guide designs. For each guide, the region of duplex RNA is outlined in red. The heatmap scale bar is the same as in (B). Values represent mean ± SEM. Nontargeting guide is the same as in Fig. 2C. (C) Schematic of target and guide design for targeting FANCC 1517G→A. (D) The 1517G→A mutation (indicated by blue triangle) in FANCC is corrected to varying levels by using REPAIRv1 with three different guide designs. For each guide, the region of duplex RNA is outlined in red. The heatmap scale bar is the same as in (B). Values represent mean ± SEM. Nontargeting guide is the same as in Fig. 2C. (E) Quantification of the percent editing of 34 different disease-relevant G→A mutations selected from ClinVar by using REPAIRv1. Nontargeting guide is the same as in Fig. 2C. (F) Analysis of all the possible G→A mutations that could be corrected by using REPAIR as annotated in the ClinVar database. (G) The distribution of editing motifs for all G→A mutations in ClinVar is shown versus the editing efficiency by REPAIRv1 per motif as quantified on the Gluc transcript. Values represent mean ± SEM.

osine (Fig. 5A). We found that for KRAS, although the on-target editing rate was 23%, there were many sites around the target site that also had detectable A-to-I edits (Fig. 5B).

Because of the observed off-target editing within the guide:target duplex, we initially evaluated transcriptome-wide off-targets by performing RNA-seq on all mRNAs with 12.5x coverage. Of all the editing sites across the transcriptome, the on-target editing rate was 23%, there were many sites around the target site that also had detectable A-to-I edits (Fig. 5B). Because of the observed off-target editing within the guide:target duplex, we initially evaluated transcriptome-wide off-targets by performing RNA-seq on all mRNAs with 12.5x coverage. Of all the editing sites across the transcriptome, the on-target editing rate was 23%, there were many sites around the target site that also had detectable A-to-I edits (Fig. 5B). Because of the observed off-target editing within the guide:target duplex, we initially evaluated transcriptome-wide off-targets by performing RNA-seq on all mRNAs with 12.5x coverage. Of all the editing sites across the transcriptome, the on-target editing rate was 23%, there were many sites around the target site that also had detectable A-to-I edits (Fig. 5B).
edits, supporting the hypothesis that REPAIR off-target edits are driven by dCas13b-independent ADAR2DD(E488Q) editing events (fig. S8D).

Next, we sought to compare two RNA-guided ADAR systems that have been described previously (fig. S9A). The first uses a fusion of ADAR2DD to the small viral protein lambda N (28). Because of the guide-independent nature of the off-targets, we hypothesized that destabilizing ADAR2DD(E488Q)—RNA binding would selectively decrease off-target editing, but maintain on-target editing because of increased local concentration from dCas13b tethering of ADAR2DD(E488Q) to the target site. We mutated residues in ADAR2DD(E488Q) previously determined to contact the duplex region of the target RNA (Fig. 6A) (29). To assess efficiency and specificity, we tested 17 single mutants with both targeting and nontargeting guides, under the assumption that background luciferase restoration in the nontargeting condition would be indicative of broader off-target activity. We found that mutations at the selected residues had substantial effects on the luciferase activity for targeting and nontargeting guides (Fig. 6, A and B, and fig. S10A). A majority of mutants either significantly improved the luciferase activity for the targeting guide or increased the ratio of targeting to nontargeting guide activity, which we termed the specificity score (Fig. 6, A and B).

We selected a subset of these mutants (Fig. 6B) for transcriptome-wide specificity profiling by next-generation sequencing. As expected, off-targets measured from transcriptome-wide sequencing correlated with our specificity score (fig. S10B) for mutants. We found that with the exception of ADAR2DD(E488Q/R455E), all sequenced REPAIRv1 mutants could effectively edit the reporter transcript (Fig. 6C), with many mutants showing reduction in the number of off-targets (Fig. 6C and figs. S10C and S11). We further explored motifs surrounding off-targets for the various specificity mutants and found that REPAIRv1 and most of the engineered variants exhibited a strong 3′ G preference for their off-target edits, which is in agreement with the characterized ADAR2 motif (fig. S12A) (28).

We focused on the mutant ADAR2DD(E488Q/T375G)—because it had the highest percent editing of the four mutants with the lowest numbers of transcriptome-wide off-targets—and termed it REPAIRv2. Compared with REPAIRv1, REPAIRv2 exhibited increased specificity, with a reduction from 18,385 to 20 transcriptome-wide off-targets with high-coverage sequencing (125x coverage, 10 ng of REPAIR vector transfected) (Fig. 6D).

In the region surrounding the targeted adenine in Cluc, REPAIRv2 also had reduced off-target editing, visible in sequencing traces (Fig. 6E). In motifs derived from the off-target sites, REPAIRv1 presented a strong preference toward 3′ G but showed off-target edits for all motifs (fig. S9B); by contrast, REPAIRv2 only edited the strongest off-target motifs (fig. S12B). The distribution of edits on transcripts was heavily skewed for REPAIRv1, with highly edited genes having more than 60 edits (fig. S3A), whereas REPAIRv2 only edited one transcript (EEF1A1) multiple times (fig. S3B). REPAIRv1 off-target edits were predicted to result in numerous variants, including 1000 missense base changes (fig. S13A), with 93 events in genes related to
Fig. 6. Rational mutagenesis of ADAR2<sub>dd</sub> to improve the specificity of REPAIRv1.

(A) Quantification of luciferase signal restoration (on-target score, red boxes) by various dCas13-ADAR2<sub>dd</sub> mutants as well as their specificity score (blue boxes) plotted along a schematic of the contacts between key ADAR2 deaminase residues and the dsRNA target (the target strand is shown in gray; the nontarget strand is shown in red). All deaminase mutations were made on the dCas13-ADAR2<sub>dd</sub>(E488Q) background. The specificity score is defined as the ratio of the luciferase signal between targeting guide and nontargeting guide conditions. [Schematic of ADAR2 deaminase domain contacts with dsRNA is adapted from (20)].

(B) Quantification of luciferase signal restoration by various dCas13-ADAR2 mutants versus their specificity score. Nontargeting guide is the same as in Fig. 2C.

(C) Quantification of on-target editing and the number of significant off-targets for each dCas13-ADAR2<sub>dd</sub>(E488Q) mutant by transcriptome-wide sequencing of mRNAs. Values represent mean ± SEM. Nontargeting guide is the same as in Fig. 2C.

(D) Transcriptome-wide sites of significant RNA editing with (top) REPAIRv1 and (bottom) REPAIRv2, with a guide targeting a pretermination site in Cluc. The on-target Cluc site (254 A→I) is highlighted in orange. Ten nanograms of REPAIR vector was transfected for each condition. (E) Representative RNA-seq reads surrounding the on-target Cluc-editing site (254 A→I; blue triangle) highlighting the differences in off-target editing between (top) REPAIRv1 and (bottom) REPAIRv2. A-to-I edits are highlighted in red; sequencing errors are highlighted in blue. Gaps reflect spaces between aligned reads. Nontargeting guide is the same as in Fig. 2C. (F) RNA editing with REPAIRv1 and REPAIRv2, with guides targeting an out-of-frame UAG site in the endogenous KRAS and PPIB transcripts. The on-target editing fraction is shown as a sideways bar chart on the right for each condition row. For each guide, the region of duplex RNA is outlined in red. Values represent mean ± SEM. Nontargeting guide is the same as in Fig. 2C.
but with substantially fewer off-target editing events at both dosages (fig S15). REPAIRv2 had enhanced specificity compared with REPAIRv1 at both dosages (fig. S15B), a finding that also extended to two guides targeting distinct sites on PP1B (fig. S16, A to D). It is also worth noting that in general, the lower-dosage condition (10 ng REPAIR vector) had fewer off-targets than that of the higher dosage condition (150 ng REPAIR vector) (fig. S5).

To assess editing specificity with greater sensitivity, we sequenced the low-dosage condition (10 ng of transfected DNA) of REPAIRv1 and v2 at much higher sequencing depth (125x coverage of the transcriptome). Increased numbers of off-targets were found at higher sequencing depths corresponding to detection of rarer off-target events (fig. S17). Furthermore, we speculated that different transcriptome states could also potentially alter the number of off-targeting events. Therefore, we tested REPAIRv2 activity in the osteosarcoma U2OS cell line, observing six and seven off-targets for the targeting and non-targeting guide, respectively (fig. S18).

We targeted REPAIRv2 to endogenous genes to test whether the specificity-enhancing mutations reduced nearby edits in target transcripts while maintaining high-efficiency on-target editing. For guides targeting either KRAS or PP1B, we found that REPAIRv2 had no detectable off-target edits, unlike REPAIRv1, and could effectively edit the on-target adenosine at efficiencies of 27.3% (KRAS) or 13% (PP1B) (fig. 6F). This specificity extended to additional target sites, including regions that demonstrate high levels of background in nontargeting conditions for REPAIRv1, such as other KRAS or PP1B target sites (fig. S19). Overall, REPAIRv2 eliminated off-targets in duplexed regions around the edited adenosine and showed dramatically enhanced transcriptome-wide specificity.

**Discussion**

We show here that the RNA-guided RNA-targeting type V1-B CRISPR effector Cas13b is capable of highly efficient and specific RNA knockdown, providing the basis for improved tools for interrogating essential genes and noncoding RNA as well as controlling cellular processes at the transcript level. Catalytically inactive Cas13b (dCas13b) retains programmable RNA-binding capability, which we leveraged here by fusing dCas13b to the adenosine deaminase domain of ADAR2 to achieve precise A-to-I edits, a system we term REPAIRv1. Further engineering of the system produced REPAIRv2, which has dramatically higher specificity than previously described RNA-editing platforms (25, 29) while maintaining high levels of on-target efficacy.

Although Cas13b exhibits high fidelity, our initial results with dCas13b-ADAR2<sub>204</sub>(E488Q) fusion revealed a substantial number of off-target RNA editing events. To address this, we used a rational mutagenesis strategy to vary the ADAR2<sub>204</sub> residues that contact the RNA duplex, identifying a variant, ADAR2<sub>204</sub>(E488Q/T375G), that is capable of precise, efficient, and highly specific editing when fused to dCas13b. Editing efficiency with this variant was comparable with or better than that achieved with two currently available systems, BoxB-ADAR2<sub>204</sub>(E488Q) or ADAR2 editing. Moreover, the REPAIRv2 system created only 20 observable off-targets in the whole transcriptome, which is at least an order of magnitude better than both alternative editing technologies. Although it is possible that ADAR could deaminate adenosine bases on the DNA strand in RNA-DNA heteroduplexes (20), it is unlikely to do so in this case because Cas13b does not bind DNA efficiently and because REPAIR is cytoplasmically localized. Additionally, the lack of homology of off-target sites to the guide sequence and the strong overlap of off-targets with the ADAR2<sub>204</sub>(E488Q)-only condition suggest that off-targets are not mediated by off-target guide binding. Deeper sequencing and novel inosine enrichment methods could further refine our understanding of REPAIR specificity in the future.

The REPAIR system offers many advantages compared with other nucleic acid-editing tools. First, the exact target site can be encoded in the guide by placing a cytidine within the guide across from the desired adenosine to create a favorable A-C mismatch ideal for ADAR-editing activity. Second, Cas13 has no targeting sequence constraints, such as a PFS or PAM, and no motif preference surrounding the target adenosine, allowing any adenosine in the transcriptome to be potentially targeted with the REPAIR system. The lack of motif for ADAR editing, in contrast with previous literature, is likely due to the increased local concentration of REPAIR at the target site owing to dCas13b binding. DNA base editors can target either the sense or antisense strand, whereas the REPAIR system is limited to transcribed sequences, constraining the total number of possible editing sites. However, because of the less constrained nature of targeting with REPAIR, this system can effect more edits within ClinVar (Fig. 4C) than Cas9-DNA base editors. Third, the REPAIR system directly deaminates target adenosines to inosines and does not rely on endogenous repair pathways to generate desired editing outcomes. Therefore, REPAIR should be able to mediate efficient RNA editing even in post-mitotic cells such as neurons. Fourth, in contrast to DNA editing, RNA editing is transient and can be more easily reversed, allowing the potential for temporal control over editing events. The transient nature of RNA-mediated edits will likely be useful for treating diseases caused by temporary changes in cell state, such as local inflammation, and could also be used to treat disease by modifying the function of proteins involved in disease-related signal transduction. For instance, REPAIR editing would allow the reordering of some serine, threonine, and tyrosine residues that are the targets of kinases (fig. S20). Phosphorylation of these residues in disease-relevant proteins affects disease progression for many disorders, including Alzheimer’s disease and multiple neurodegenerative conditions (30). REPAIR might also be used to transiently or even chronically change the sequence of expressed, risk-modifying G-to-A variants so as to decrease the chance of entering a disease state for patients. For instance, REPAIR could be used to functionally mimic A-to-G alleles of IFIH1 that protect against autoimmune disorders such as type I diabetes, immunoglobulin A deficiency, psoriasis, and systemic lupus erythematosus (31, 32).

The REPAIR system provides multiple opportunities for additional engineering. Cas13b possesses pre-CRISPR-RNA (crRNA) processing activity (19), allowing for multiplex editing of multiple variants—none of which alone may not affect disease, but together might have additive effects and disease-modifying potential. Extension of our rational design approach, such as combining promising mutations and directed evolution, could further increase the specificity and efficiency of the system, while unbiased screening approaches could identify additional residues for improving REPAIR activity and specificity.

Currently, the base conversions achievable by REPAIR are limited to generating inosine from adenosine; additional fusions of dCas13 with other catalytic RNA editing domains, such as APOBEC (apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like), could enable cytidine-to-uridine editing. Additionally, mutagenesis of ADAR could relax the substrate preference to target cytidine, allowing for the enhanced specificity conferred by the duplexed RNA substrate requirement to be exploited by C-to-U editors. Adenosine-to-inosine editing on DNA substrates may also be possible with catalytically inactive DNA-targeting CRISPR effectors, such as dCas9 or dCpf1, either through formation of DNA-RNA heteroduplex targets (20) or mutagenesis of the ADAR domain.

We have demonstrated the use of the PspCas13b enzyme as both an RNA-knockdown and RNA-editing tool. The dCas13b platform for programmable RNA binding has many applications, including live transcript imaging, splicing modification, targeted localization of transcripts, pulldown of RNA-binding proteins, and epitranscriptomic modifications. We used dCas13 to create REPAIR, adding to the existing suite of nucleic acid–editing technologies. REPAIR provides a new approach for treating genetic disease or mimicking protective alleles and establishes RNA editing as a useful tool for modifying genetic function.

**REFERENCES AND NOTES**


ACKNOWLEDGMENTS

We thank R. Murthy, A. Baumann, A. Philippakis, and E. Banks for computational guidance; C. Muus for sequencing assistance; A. Smargon for gene synthesis of a Cas13b orthologs 1–8; V. Verdine for assistance with RNA sequencing library preparation; and R. Macrae, R. Belliveau, and the entire F.Z. laboratory for discussions and support. D.B.T.C. is supported by award T32GM007753 from the National Institute of General Medical Sciences. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of General Medical Sciences or the NIH. O.O.A. is supported by a Paul and Daisy Soros Fellowship and a NIH F30 National Research Service Award 1F30-CA210382. F.Z. is a New York Stem Cell Foundation–Robertson Investigator. F.Z. is supported by NIH grants 1R01-HG009761, 1R01-MH110049, and 1DP1-HL141201; the Howard Hughes Medical Institute; the New York Stem Cell Foundation; the Simons Foundation; the Paul G. Allen Family Foundation; the Valey Foundation; and J. and P. Poitras, R. Metcalfe, and D. Cheng. D.B.T.C., J.S.G., O.O.A., and F.Z. are co-inventors on U.S. provisional patent application no. 62/525,181 relating to REPAIR and other applications relating to CRISPR and CRISPR-Cas13 technology filed by the Broad Institute. Sequencing data are available at Sequence Read Archive under BioProject accession no. PRJNA414340. The authors plan to make the reagents widely available to the academic community through Addgene and to provide software tools via the F.Z. laboratory website (www.genome-engineering.org) and GitHub (github.com/fengzhanglab).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/358/6366/1019/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S20
Tables S1 to S9
References (33–38)
21 September 2017; accepted 16 October 2017
Published online 25 October 2017
10.1126/science.aag0180
RNA editing with CRISPR-Cas13
David B. T. Cox, Jonathan S. Gootenberg, Omar O. Abudayyeh, Brian Franklin, Max J. Kellner, Julia Joung and Feng Zhang

Science 358 (6366), 1019-1027.
DOI: 10.1126/science.aaq0180 Originally published online October 25, 2017

Precise transcriptome engineering
Efficient and precise RNA editing to correct disease-relevant transcripts holds great promise for treating genetic disease. Cox et al. took advantage of the ability of Cas13b, an effector from a type VI CRISPR-Cas system, to target specific RNAs directly (see the Perspective by Yang and Chen). They fused Cas13b with the ADAR2 adenosine deaminase domain and used rational protein engineering to improve the resultant enzyme. These approaches yielded an RNA knockdown and editing platform that allowed efficient and specific RNA depletion and correction in mammalian cells.

Science, this issue p. 1019; see also p. 996