MAGPiE enables precise and versatile RNA manipulation via cleavage and ligation

Long-Qi Li[1,3], Yuan-Yuan Jiang[1,2,3,*], Rui-Kang Chen[1,3], Rui-Han Zhang[1], Shuo Lin[1], Zhi-Hui Liu[1], Shuai Jin[1], Yun Yang[1], Chao-Qun Wu[1], Shu-Lin Jin[1], Gao-Li Wang[1], Ming-Yu Hou[1], Jun-Jie Gogo Liu[1,4,*]

[1]Beijing Frontier Research Center for Biological Structure, State Key Laboratory of Membrane Biology, Center for Synthetic and Systematic Biology, Tsinghua-Peking Center for Life Sciences, School of Life Sciences, Tsinghua University, Beijing 100084, China

[2]State Key Laboratory of Crop Gene Resources and Breeding, Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing, 100081, China

[3]These authors contributed equally

[4]Lead contact

*Correspondence: jiangyuanyuan@caas.cn (Y.Y.J.); junjiegogoliu@tsinghua.edu.cn (J.-J.G.L.)

ABSTRACT

RNA editing offers a promising alternative to permanent genome modification, yet most existing approaches remain largely confined to single-nucleotide substitutions. Here we present MAGPiE, a programmable RNA editing platform capable of diverse editing outcomes with limited site dependence. In MAGPiE1.0, target RNAs are site-specifically cleaved by a type III CRISPR complex and subsequently ligated by RtcB to ribozyme-generated RNA donors. Using a fluorescence reporter system, we established efficient receptor–donor ligation across multiple target sites and defined an optimal crRNA design. An optimized RtcB configuration, termed MAGPiE2.0, further enabled editing of endogenous mRNAs and noncoding RNAs with moderate efficiencies. To improve editing performance, we developed MAGPiE2.1–2.3 by introducing splint- or dCas13-mediated donor recruitment and by perturbing exoribonucleases. Receptor–donor base pairing that mimics native RtcB substrates increased editing efficiency by 2.6-fold in the MAGPiE3.0 framework and also provided the basis for an intron-targeting version. Beyond single-nucleotide substitutions, MAGPiE enables multi-nucleotide editing and programmable fragment replacement, including precise RNA-level knock-in for fluorescence tagging. Together, these results establish MAGPiE as a versatile platform that expands the scope of programmable RNA manipulation.

INTRODUCTION

DNA editing technologies such as CRISPR–Cas9 enable targeted modification of the genome, but their permanent nature raises concerns regarding long-term safety[1]. RNA manipulation offers an attractive alternative

by avoiding irreversible genomic alterations while enabling dynamic regulation and tunable levels of intervention[2,3]. Existing RNA manipulation strategies have broadly followed two major paradigms. RNA base editing harnesses endogenous or engineered deaminases to mediate defined single-nucleotide conversions, most notably A-to-I[4-9] and C-to-U[10,11] editing. More recently, U-to-Ψ modification has also been exploited to correct premature termination codons[12-15]. Besides, trans-splicing tools are based on splicing reaction between endogenous and delivered half-intron molecules and replace the mutated exons[16]. However, both approaches face inherent constraints. Current base-editing strategies remain largely restricted to single-nucleotide modifications and access only a limited subset of the twelve possible base substitutions (four transitions and eight transversions), while generally lacking the capacity for programmable insertions or deletions. Intron-mediated trans-splicing, meanwhile, is constrained by the endogenous intronic context and often exhibits limited efficiency owing to competition with native cis-splicing[17]. An ideal RNA editing platform would therefore combine broad editing versatility with high precision and minimal dependence on target-site context.

Here we present MAGPiE (MAnipulation of RNA Graft between Pairwise Ends), a programmable RNA editing platform based on a novel cut-and-replace strategy, whereby a defined RNA segment is excised and replaced with a donor RNA carrying the desired modification, in principle enabling diverse classes of RNA editing. Most cellular RNA species bear 5′-phosphate (or derivatives thereof) and 3′-hydroxyl ends. By contrast, a subset of RNA intermediates carries noncanonical 5′-OH and 2′,3′-cyclic phosphate (2′,3′-cP) termini, which can be efficiently ligated by endogenous RtcB ligase[18,19]. The prokaryote-derived type III-A CRISPR–Csm complex has been adapted for programmable RNA cleavage in mammalian cells and generates the same 5′-OH and 2′,3′-cP termini[20-22]. Although previous studies have reported ligation of Csm-cleaved RNA fragments[23,24], these approaches have been restricted to nucleotides deletion rather than the installation of user-defined sequence changes. Instead, we sought to introduce diverse kind of mutations through supplying exogenous RNA donors carrying desired sequence modifications. Self-cleaving ribozymes, including Twister, hammerhead (HH), and hepatitis delta virus (HDV) ribozymes, can likewise generate RNA fragments bearing 5′-OH and 2′,3′-cP termini and have been used for applications including functional RNA circularization[25] and large-mRNA delivery[26]. In MAGPiE, Csm from *Streptococcus thermophilus* cleaves the target transcript to generate an RNA receptor, while a self-cleaving ribozyme generates an exogenous RNA donor with a compatible terminus. Endogenous RtcB subsequently ligates the receptor and donor, thereby enabling, in principle, a broad range of sequence modifications at the target site. The name MAGPiE also draws inspiration from an ancient Chinese legend in which a bridge formed by magpies (鹊桥) reunites two separated lovers, echoing the joining of two RNA ends. In this study, we systematically evaluated MAGPiE using a split-EGFP reporter, improved its efficiency through component engineering and modulation of endogenous cellular pathways, and demonstrated its applicability to endogenous RNA manipulation in human cells.

RESULTS

Ligation between Csm- and ribozyme-cleaved products reconstitutes EGFP mRNA

We established a split-EGFP reporter system in HEK293T cells to monitor RNA fragment ligation (Fig. 1a). Two complementary configurations were designed. In MAGPiEa, Csm targets the N-terminal EGFP (Nt-EGFP) fragment to generate a 3′ end bearing 2′,3′-cP, whereas the C-terminal EGFP (Ct-EGFP) donor is generated by ribozyme (Rz) self-cleavage with a 5′-OH terminus. In MAGPiEb, the arrangement is reversed: Csm targets the Ct-EGFP fragment, while ribozyme self-cleavage generates the Nt-EGFP donor. In both configurations, endogenous RtcB ligates the receptor and donor RNAs to reconstitute full-length EGFP mRNA (Fig. 1a). We selected the RzB hammerhead (HH) ribozyme for MAGPiEa and the Twister ribozyme for MAGPiEb to produce donor RNA with flexible terminal nucleotides (Supplementary Fig. S1a).

The type III-A CRISPR–Csm complex comprises five Csm proteins (Csm1–Csm5), with multiple Csm3 endoribonuclease subunits cleaving the complementary target RNA at 6-nt intervals[27,28]. To define how this periodic cleavage pattern influences ligation, we designed five crRNAs that positioned the intended junction at each possible 6-nt-spaced cleavage site. crRNAs were named according to the position of the intended cleavage site relative to the 5′ end of the spacer (for example, N06 and N12 for MAGPiEa; C06 and C12 for MAGPiEb) (Fig. 1b). HEK293T cells were co-transfected with the split-EGFP constructs, Csm complex and crRNAs, with a both-ends Rz-fusion construct serving as the positive control (PC)[26] (Fig. 1c). Fluorescence microscopy 48 h after transfection revealed successful EGFP reconstitution in multiple groups (Fig. 1d). Flow cytometry identified N24 as the most active crRNA in MAGPiEa, yielding 45.2% EGFP-positive cells compared with 50.3% for the positive control. MAGPiEb C06, C12 and C18 produced comparable fluorescence levels (Fig. 1e; Supplementary Fig. S1b).

Next-generation sequencing (NGS) across the ligation junction corroborated the fluorescence measurements: N24 yielded the highest fraction of intended ligation products in MAGPiEa, whereas C06, C12 and C18 produced similar fractions in MAGPiEb (Fig. 1c, f; Supplementary Fig. S2). Mapping ligation junctions relative to the spacer 5′ end further revealed inherent positional preferences. In MAGPiEa, an average of 84.6% of ligation events occurred at position 24 (Fig. 1g), consistent with the reported preference of Csm for cleavage at the site farthest from the spacer 5′ end[27,28]. The predominant position-24 product corresponds to cleavage by the standard Csm complex containing four catalytically active Csm3 subunits (Fig. 1h). We also detected ligation at position 30 (purple dashed line), indicative of a Csm species containing five Csm3 subunits, consistent with the previously reported heterogeneity of Csm complexes[29]. MAGPiEb showed a broader distribution of ligation junctions, with the major peak at position 6 followed by positions 12 and 18, indicating comparable cleavage at these sites (Fig. 1g, h).

N24 and C06 achieve optimal precise ligation rate across diverse target sites

Guided by the positional preferences of ligation identified above, we designed a series of spacers targeting four randomly selected EGFP split sites (S2–S5), testing S2, S3 and S4 with MAGPiEa and S2, S3 and S5 with MAGPiEb (Fig. 2a). Flow cytometry (Supplementary Fig. S3a) showed that N24 consistently outperformed the other spacers across all sites tested with MAGPiEa, whereas C06 yielded the highest fluorescence reconstitution on average with MAGPiEb (Fig. 2b, c). The reconstitution efficiencies were broadly comparable across the different split sites (S2–S5), indicating low dependence on local sequence context.

To directly quantify ligation products, we incorporated a short sequence identical to the donor (dark green box) into the receptor as a shared primer-binding site, allowing the unedited receptor and edited products to be amplified simultaneously (Fig. 2a). Through amplicon deep sequencing, ligated products were distinguished by a unique donor-derived sequence, so that the percentage of ligated reads or precise edited reads can be calculated (Supplementary Fig. S3b). Across all targets, MAGPiEa reached a maximum ligation efficiency of 19.4% at S3-N18 and a maximum precise editing efficiency of 9.2% at S3-N24 (Fig. 2d). MAGPiEb configuration showed lower efficiency, reaching 3.0% ligation at S3-C18 and 1.2% precise editing at S3-C06 (Fig. 2e), consistent with the fluorescence measurements in Fig. 2b, c. Notably, N24 and C06 also produced the highest fractions of precisely edited products across all tested split sites, reaching 79.6% at S3-N24 and 91.5% at S3-C06, respectively (Fig. 2f; Supplementary Fig. S3c).

We next asked whether MAGPiE could support internal RNA fragment replacement through three-fragment ligation. The receptor RNA was cleaved at two positions (S2 and S5) to remove the intervening stop-codon-containing segment, while the donor was flanked by 5′ and 3′ ribozymes to generate termini compatible with ligation to both flanking receptor fragments (Fig. 2g). The NGS results showed detectable ligation products (Supplementary Fig. S4). Together, these results establish N24 and C06 as preferred crRNA configurations for MAGPiEa and MAGPiEb, respectively, demonstrate the relatively low site dependence of MAGPiE, and extend the system from two-fragment ligation to potential internal RNA fragment replacement.

MAGPiE2.0 enhances ligation efficiency and enables fragment replacement in endogenous RNAs

RtcB is broadly distributed across archaea, bacteria, algae and animals[30] and functions in RNA ligation during processes including tRNA[18,31] and XBP1 mRNA splicing[32,33], as well as RNA repair[34-37]. To assess how RtcB activity influences MAGPiE efficiency, we expressed RtcB orthologs from *Homo sapiens* (HsRtcB; eukaryotic), *Escherichia coli* (EcRtcB; bacterial) and *Pyrococcus horikoshii* (PhRtcB; archaeal), with or without an N-terminal nuclear localization signal (NLS), under either the SV40 or EF1α promoter. Flow cytometry showed that nuclear localization generally enhanced ligation across all RtcB orthologs. NLS-EcRtcB produced the strongest EGFP reconstitution, increasing fluorescence by 1.6–1.8-fold relative to the control lacking exogenous RtcB. The

SV40-driven NLS-EcRtcB configuration showed more consistent performance and was therefore adopted as MAGPiE2.0 (Fig. 3a).

We next sought to further improve EcRtcB activity using two complementary AI-guided strategies. ESM-2 was first used to identify amino-acid substitutions predicted to be more compatible with the native sequence context than the corresponding wild-type residues[38]. In parallel, structure-conditioned inverse-folding approaches[39] were applied using an AlphaFold 3-predicted structure as the template[40]: ProteinMPNN was used to propose candidate single-amino-acid substitutions[41], whereas ESM-IF was used to score all possible single-point mutations[42]. For experimental validation, we selected the highest-ranked ESM-2 candidates, along with consensus mutations independently prioritized by both ProteinMPNN and ESM-IF. Unexpectedly, all tested variants exhibited ligation efficiencies comparable to wild-type EcRtcB, with mean fluorescence intensity (MFI) values remaining within ±20% of the wild-type level (Supplementary Fig. S5). These results suggest that the specific RtcB variants generated here may have limited capacity for improvement of the ligation efficiency, in contrast to the pronounced effect of nuclear targeting.

We next evaluated MAGPiE2.0 for fragment replacement in endogenous transcripts. We selected four protein-coding RNAs (PPIB, PARK7, SMAD4 and CKB) and two nuclear non-coding RNAs (NEAT1 and XIST), and designed MAGPiEa- or MAGPiEb-mediated fragment replacement using the preferred N24 or C06 spacer configuration, respectively (Fig. 3b; Supplementary Fig. S6a). Amplicon sequencing confirmed introduction of the intended sequence modifications across these endogenous targets, including editing outcomes that are difficult to achieve with existing base-editing or trans-splicing approaches (Fig. 3c, d; Supplementary Fig. S6b). At the two most efficiently edited sites, PPIB with MAGPiEa and NEAT1 with MAGPiEb, precisely edited reads accounted for approximately 80% of the ligated products (Fig. 3d). These results extend MAGPiE from overexpressed reporter transcripts to endogenous mRNAs and non-coding RNAs and demonstrate its capacity for programmable fragment replacement in native cellular transcripts.

Optimization through engineering MAGPiE components and manipulating cellular determinants

Given that Csm-mediated target RNA cleavage has been extensively characterized and optimized[22-24], we reasoned that productive association between the donor and receptor RNAs might represent a major remaining bottleneck. We therefore first sought to increase ligation by manipulating the cleavage products (MAGPiE2.1). An MS2 stem–loop was incorporated into the RNA donor and MCP was fused to the Csm complex, with the aim of recruiting the donor to the proximity of receptor. However, none of these configurations produced a significant improvement in ligation efficiency (Supplementary Fig. S7). We next sought to juxtapose the donor and receptor RNAs while recruiting RtcB through an MS2-based splint. The splint was implemented in three formats: supplied as a free circular RNA (Trans_circ_Splint)[25], appended to the 3′ end of the crRNA (crRNA_Splint), or expressed in cis with the donor transcript and released by ribozyme self-cleavage (DonRNA_Splint) (Fig. 4a;

Supplementary Fig. S8a). Among the designs tested, Splint-2 in the DonRNA_Splint configuration increased fluorescence intensity by 2.0-fold (Fig. 4b; Supplementary Fig. S8b), possibly due to the proximity of donor and splint, increasing the likelihood of pairing. Circularization of DonRNA_Splint using the Tornado system[25] further increased its stability and raised the MFI to 2.8-fold relative to the control (Fig. 4b). Applying the same design to the endogenous NEAT1 target increased both ligation efficiency and precise editing efficiency by approximately 1.4-fold (Supplementary Fig. S8c).

Catalytically inactive Cas13 proteins have previously been engineered to promote programmable RNA–RNA interactions[43-45]. We therefore developed an orthogonal recruitment strategy by fusing a CasRx (RfxCas13d)[46] crRNA cassette to the end of the RNA donor (MAGPiE2.2). The spacer within the cassette was designed to target the receptor RNA, enabling dCas13-mediated recruitment of the donor into proximity with the receptor. In addition to inactivating the two HEPN catalytic centers through the R239A, H244A, R858A and H863A substitutions, we introduced K940A at the residue required for crRNA processing[47] to prevent cleavage upstream of the direct repeat. EcRtcB was further fused to the N terminus of dCasRx-K940A to couple donor recruitment with RNA ligation (Fig. 4c). At the S3-N24 site in the split-EGFP reporter, two spacers targeting the receptor coding sequence increased fluorescence by 1.1- and 2.2-fold, respectively, whereas a non-targeting spacer exhibited no improvement (Fig. 4d).

It has been reported that the target RNA is degraded by endogenous RNA surveillance pathways following cleavage by Cas13 or the RNA-induced silencing complex (RISC)[48]. We reasoned that the receptor and donor RNAs produced by Csm and ribozyme might be susceptible to degradation before ligation could occur. Therefore, we explored the possibility that intervention of endogenous RNA processing pathways could enhance MAGPiE efficiency (MAGPiE2.3). Candidates were first depleted by shRNA treatment, followed 48 h later by transfection of plasmids encoding the Csm complex and fluorescence reporter (Fig. 4e; Supplementary Fig. S9a). Major pathways responsible for degrading exposed RNA termini include the 3′–5′ exoribonuclease exosome, which functions in both the nucleus and cytoplasm[49,50], and the 5′–3′ exoribonucleases XRN1 and XRN2, which are predominantly cytoplasmic and nuclear, respectively[51,52]. Among the exosome components tested, depletion of EXOSC5, a key structural component of the complex, produced 5.1-fold increase in MAGPiEa and 2.3-fold in MAGPiEb (Fig. 4f). Knockdown of the nuclear exonuclease XRN2 also enhanced editing efficiency, whereas depletion of cytoplasmic XRN1 had little effect. These results suggest that MAGPiE cleavage and ligation occur predominantly in the nucleus, consistent with the nuclear localization of the Csm complex used in our system. We therefore extended this analysis to the nuclear exosome targeting (NEXT) complex (MTR4, ZCCHC8 and RBM7)[53], an exosome cofactor that recruits and delivers RNA substrates to the catalytic core[54-56]. Depletion of each NEXT component enhanced ligation efficiency in both MAGPiEa and MAGPiEb (Fig. 4f). The MFI(EGFP)/MFI(mCherry) ratio was likewise increased, excluding a nonspecific effect arising from globally elevated RNA abundance (Supplementary Fig. S9b-d). Editing of the endogenous PPIB

and NEAT1 transcripts showed a consistent trend (Supplementary Fig. S9e), demonstrating that modulation of endogenous nuclear RNA decay pathways can effectively improve MAGPiE efficiency.

Juxtaposition of RNA termini promotes MAGPiE ligation

Native RtcB substrates, including tRNA[57] and XBP1 mRNA[58], position the 5′ and 3′ termini in close spatial proximity through stem–loop structures. We therefore asked whether analogous receptor–donor pairing could facilitate intracellular ligation. To this end, we redesigned the donor in MAGPiE3.0 to form either a 16-bp or 32-bp duplex with the Csm-cleaved receptor, while retaining a 1-nt overhang at the 5′ terminus and a 6-nt overhang at the 3′ terminus to mimic the geometry of native tRNA substrates (Supplementary Fig. S10a). Introducing a 16-bp complementary region between the receptor and donor increased ligation efficiency by an average of 2.6-fold, with the highest ligation efficiency reaching 1.5% among the variants tested (Fig. 5a, b).

For editing within coding regions, such receptor–donor pairing requires site-specific donor design, for example through synonymous substitutions that strengthen base pairing without altering the encoded protein sequence. Alternatively, we redirected Csm cleavage to intronic regions, allowing the receptor–donor duplex formed after ligation to be subsequently removed by splicing. In MAGPiE3.0a, Csm was programmed to target intron 4 of the PPIB transcript, and a 5′-OH donor containing a 150-nt receptor-complementary region was supplied. Downstream of this region, the donor contained the cis-regulatory elements required for splicing, including a branch-point sequence (YNYURAY), a polypyrimidine tract and a 3′ splice site (YAG)[44], followed by the replacement exon 5 (Fig. 5c; Supplementary Fig. S10b). In MAGPiE3.0b, Csm was directed to intron 1, and the donor contained, from 5′ to 3′, a modified exon 1, a 5′ splice site (AG|GURAGU), and a 150-nt receptor-complementary region[59] terminating in a 2′,3′-cP. MAGPiE3.0a and MAGPiE3.0b both produced the intended edits at all three sites tested on PPIB pre-mRNA, with MAGPiE3.0a reaching a maximum editing efficiency of 5.8% with Donor1 (Fig. 5d).

MAGPiE enables fluorescence tagging through RNA-level knock-in

Beyond precise modification of endogenous RNAs, the capacity of MAGPiE for programmable segment replacement enables knock-in at the RNA level. As a proof of concept, we applied MAGPiE2.0 to a panel of transcripts, including H2BC11, LMNB1, TUBA1B, MAPRE3 and TPX2. Each transcript was expressed and targeted in the last codon of the coding sequence, along with a donor carrying the EGFP coding sequence lacking a start codon, such that productive ligation would generate an in-frame fusion with the target transcript (Fig. 6a; Supplementary Fig. S11a). In cells with H2BC11 (H2B Clustered Histone 11) targeted, EGFP fluorescence was readily detected in the nucleus (Fig. 6c), and precise in-frame fusion was confirmed at the RNA level (Supplementary Fig. S11b, c). Quantification across multiple target genes showed RNA knock-in efficiencies of up to approximately 4% by RT–qPCR (Fig. 6b), demonstrating the precise insertion of a functional coding sequence and highlighting the potential of MAGPiE for endogenous fluorescence tagging. Collectively,

these results extend MAGPiE beyond sequence correction and establish its capacity for diverse manipulations of native transcripts.

DISCUSSION

This study establishes MAGPiE as a programmable cut-and-replace framework that substantially broadens the accessible scope of RNA editing. Unlike deaminase-based approaches that are largely restricted to defined single-nucleotide conversions, MAGPiE can, in principle, support diverse editing outcomes, including insertions, deletions, sequence replacements and potentially non-natural modifications, with limited positional constraints. Its capacity for multi-fragment ligation further raises the possibility of replacing disease-associated RNA segments while minimizing perturbation of the surrounding sequence. MAGPiE may therefore provide opportunities in both therapeutic development and basic research. Nearly half of reported human pathogenic variants comprise insertions/deletions or structural variants[60], highlighting a potential disease space that is poorly addressed by conventional RNA base editors. More broadly, the ability to replace or insert defined RNA segments could enable applications ranging from RNA tracking and protein-domain recombination to non-coding RNA studies and functional interrogation of regulatory RNA elements (Supplementary Fig. S12).

The establishment and optimization of MAGPiE also revealed several features of its intracellular reaction environment. RtcB constructs targeted to the nucleus promoted EGFP reconstitution more efficiently than their counterparts lacking nuclear localization. Consistently, perturbation of the nuclear exosome cofactor NEXT complex and the nuclear 5′–3′ exoribonuclease XRN2 enhanced MAGPiE activity, supporting a model in which cleavage and ligation occur predominantly in the nucleus, in agreement with the nuclear localization of NLS-tagged Csm. Whether an analogous reaction can be efficiently implemented in the cytoplasm remains to be explored. More broadly, depletion of the exosome, NEXT and XRN2 enhanced ligation efficiency, indicating that endogenous RNA decay pathways compete with productive receptor–donor ligation and represent actionable determinants of MAGPiE performance. Although several RNA-processing pathways were individually interrogated here, larger-scale genetic screens, such as CRISPR interference (CRISPRi) screens, may uncover additional regulators with stronger effects[61,62]. Replacing RNAi-mediated depletion of inhibitory factors with dominant-negative variants could also simplify implementation. The splint experiments further suggest that optimal receptor–donor juxtaposition requires appropriate terminal geometry: Splint-2, which leaves both ligating termini accessible, outperformed the fully base-paired Splint-1 configuration, potentially reflecting effects on ribozyme self-cleavage or the substrate preference of RtcB.

Previous studies of Csm-mediated RNA editing achieved nucleotide deletions either in 6-nt increments[23] or at arbitrary lengths[24]. By introducing ribozyme-generated RNA donors, MAGPiE extends this framework from deletion to programmable sequence replacement and the installation of diverse modifications. Several limitations nevertheless remain. Transcript depletion following Csm cleavage is an important concern for

cleavage-and-repair-based RNA manipulation, as also observed in a conceptually related preprint released concurrently with our study[63]. Importantly, the enhancement achieved by perturbing nuclear RNA decay pathways suggests that cleavage-associated transcript loss is not merely an intrinsic limitation of the strategy, but can be partially mitigated by modulating endogenous RNA metabolism. Further optimization of these pathways may therefore improve both transcript preservation and productive ligation. Although we defined preferred spacer configurations for precise ligation, MAGPiE1.0 and MAGPiE2.0 still generate a fraction of junction products shifted by 6 nt. Strategically placed crRNA mismatches that suppress cleavage at undesired positions18, or alternative RNA-targeting effectors with more defined cleavage patterns, such as Cas7-11[64], may improve product purity. Editing efficiencies at endogenous targets also remain moderate. Combining the optimization strategies developed here, together with alternative approaches for donor recruitment and stabilization, should further improve performance. Extending MAGPiE to a broader range of disease-relevant and application-oriented transcripts will be important for defining its generalizability and ultimately realizing its potential as a versatile RNA engineering platform.

REFERENCES

1. Angelini Stewart, A., Ahrens-Nicklas, R.C., Tsai, S.Q., Musunuru, K., Giannikopoulos, P., and Clelland, C.D. (2026). Measurement and clinical interpretation of CRISPR off-targets. Nat Genet *58*, 20-27.
2. Pfeiffer, L.S., and Stafforst, T. (2023). Precision RNA base editing with engineered and endogenous effectors. Nat Biotechnol *41*, 1526-1542.
3. Yang, D., Wu, X., Yao, Y., Duan, M., Wang, X., Li, G., Guo, A., Wu, M., Liu, Y., Zheng, J., et al. (2025). An RNA editing strategy rescues gene duplication in a mouse model of MECP2 duplication syndrome and nonhuman primates. Nat Neurosci *28*, 72-83.
4. Merkle, T., Merz, S., Reautschnig, P., Blaha, A., Li, Q., Vogel, P., Wettengel, J., Li, J.B., and Stafforst, T. (2019). Precise RNA editing by recruiting endogenous ADARs with antisense oligonucleotides. Nature Biotechnology *37*, 133-138.
5. Qu, L., Yi, Z., Zhu, S., Wang, C., Cao, Z., Zhou, Z., Yuan, P., Yu, Y., Tian, F., Liu, Z., et al. (2019). Programmable RNA editing by recruiting endogenous ADAR using engineered RNAs. Nat Biotechnol *37*, 1059-1069.
6. Yi, Z., Qu, L., Tang, H., Liu, Z., Liu, Y., Tian, F., Wang, C., Zhang, X., Feng, Z., Yu, Y., et al. (2022). Engineered circular ADAR-recruiting RNAs increase the efficiency and fidelity of RNA editing in vitro and in vivo. Nat Biotechnol *40*, 946-955.
7. Katrekar, D., Yen, J., Xiang, Y., Saha, A., Meluzzi, D., Savva, Y., and Mali, P. (2022). Efficient in vitro and in vivo RNA editing via recruitment of endogenous ADARs using circular guide RNAs. Nat Biotechnol *40*, 938-945.
8. Song, D., Liu, G., Zhang, W., Ren, J., Jin, X., Sun, Y., Yi, Z., Qiu, S., Tang, H., Yi, Z., et al. (2026). RNA structure programs endogenous ADAR for precise and efficient editing. Cell *189*, 4359-4376 e4327.
9. Cox, D.B.T., Gootenberg, J.S., Abudayyeh, O.O., Franklin, B., Kellner, M.J., Joung, J., and Zhang, F. (2017). RNA editing with CRISPR-Cas13. Science *358*, 1019-1027.

10. Huang, X., Lv, J., Li, Y., Mao, S., Li, Z., Jing, Z., Sun, Y., Zhang, X., Shen, S., Wang, X., et al. (2020). Programmable C-to-U RNA editing using the human APOBEC3A deaminase. Embo j *39*, e104741.
11. Abudayyeh, O.O., Gootenberg, J.S., Franklin, B., Koob, J., Kellner, M.J., Ladha, A., Joung, J., Kirchgatterer, P., Cox, D.B.T., and Zhang, F. (2019). A cytosine deaminase for programmable single-base RNA editing. Science *365*, 382-386.
12. Song, J., Dong, L., Sun, H., Luo, N., Huang, Q., Li, K., Shen, X., Jiang, Z., Lv, Z., Peng, L., et al. (2023). CRISPR-free, programmable RNA pseudouridylation to suppress premature termination codons. Mol Cell *83*, 139-155 e139.
13. Adachi, H., Pan, Y., He, X., Chen, J.L., Klein, B., Platenburg, G., Morais, P., Boutz, P., and Yu, Y.T. (2023). Targeted pseudouridylation: An approach for suppressing nonsense mutations in disease genes. Mol Cell *83*, 637-651 e639.
14. Luo, N., Huang, Q., Dong, L., Liu, W., Song, J., Sun, H., Wu, H., Gao, Y., and Yi, C. (2025). Near-cognate tRNAs increase the efficiency and precision of pseudouridine-mediated readthrough of premature termination codons. Nat Biotechnol *43*, 114-123.
15. Liu, J., Yan, X., Wu, H., Ji, Z., Shan, Y., Wang, X., Ran, Y., Ma, Y., Li, C., Zhu, Y., et al. (2025). RNA codon expansion via programmable pseudouridine editing and decoding. Nature *643*, 1410-1420.
16. Doi, A., Delaney, C., Tanner, D., Burkhart, K., and Bell, R.D. (2024). RNA exon editing: Splicing the way to treat human diseases. Mol Ther Nucleic Acids *35*, 102311.
17. Berger, A., Maire, S., Gaillard, M.C., Sahel, J.A., Hantraye, P., and Bemelmans, A.P. (2016). mRNA trans-splicing in gene therapy for genetic diseases. Wiley Interdiscip Rev RNA *7*, 487-498.
18. Tanaka, N., Meineke, B., and Shuman, S. (2011). RtcB, a novel RNA ligase, can catalyze tRNA splicing and HAC1 mRNA splicing in vivo. J Biol Chem *286*, 30253-30257.
19. Moncan, M., Rakhsh-Khorshid, H., Eriksson, L.A., Samali, A., and Gorman, A.M. (2023). Insights into the structure and function of the RNA ligase RtcB. Cellular and Molecular Life Sciences *80*.
20. Staals, Raymond H.J., Zhu, Y., Taylor, David W., Kornfeld, Jack E., Sharma, K., Barendregt, A., Koehorst, Jasper J., Vlot, M., Neupane, N., Varossieau, K., et al. (2014). RNA Targeting by the Type III-A CRISPR-Cas Csm Complex of Thermus thermophilus. Molecular Cell *56*, 518-530.
21. Tamulaitis, G., Kazlauskiene, M., Manakova, E., Venclovas, Č., Nwokeoji, Alison O., Dickman, Mark J., Horvath, P., and Siksnys, V. (2014). Programmable RNA Shredding by the Type III-A CRISPR-Cas System of Streptococcus thermophilus. Molecular Cell *56*, 506-517.
22. Colognori, D., Trinidad, M., and Doudna, J.A. (2023). Precise transcript targeting by CRISPR-Csm complexes. Nat Biotechnol *41*, 1256-1264.
23. Nemudraia, A., Nemudryi, A., and Wiedenheft, B. (2024). Repair of CRISPR-guided RNA breaks enables site-specific RNA excision in human cells. Science *0*, eadk5518.
24. Sun, Y., Wu, Y., He, Z., Wang, Y., Hou, W., Cao, Y., Zhou, Q., and Zhang, R. (2025). Type III CRISPR-mediated flexible RNA excision with engineered guide RNAs. Mol Cell *85*, 989-998 e984.
25. Litke, J.L., and Jaffrey, S.R. (2019). Highly efficient expression of circular RNA aptamers in cells using autocatalytic transcripts. Nat Biotechnol *37*, 667-675.
26. Lindley, S.R., Subbaiah, K.C.V., Priyanka, F., Poosala, P., Ma, Y., Jalinous, L., West, J.A., Richardson, W.A., Thomas, T.N., and Anderson, D.M. (2024). Ribozyme-activated mRNA trans-ligation enables large gene delivery to treat muscular dystrophies. Science *386*, 762-767.
27. You, L., Ma, J., Wang, J., Artamonova, D., Wang, M., Liu, L., Xiang, H., Severinov, K., Zhang, X., and Wang, Y. (2019). Structure Studies of the CRISPR-Csm Complex Reveal Mechanism of Co-transcriptional Interference. Cell *176*, 239-253 e216.

28. Guo, M., Zhang, K., Zhu, Y., Pintilie, G.D., Guan, X., Li, S., Schmid, M.F., Ma, Z., Chiu, W., and Huang, Z. (2019). Coupling of ssRNA cleavage with DNase activity in type III-A CRISPR-Csm revealed by cryo-EM and biochemistry. Cell Res *29*, 305-312.
29. Liu, T.Y., Liu, J.J., Aditham, A.J., Nogales, E., and Doudna, J.A. (2019). Target preference of Type III-A CRISPR-Cas complexes at the transcription bubble. Nat Commun *10*, 3001.
30. Englert, M., Sheppard, K., Aslanian, A., Yates, J.R., 3rd, and Soll, D. (2011). Archaeal 3'-phosphate RNA splicing ligase characterization identifies the missing component in tRNA maturation. Proc Natl Acad Sci U S A *108*, 1290-1295.
31. Popow, J., Englert, M., Weitzer, S., Schleiffer, A., Mierzwa, B., Mechtler, K., Trowitzsch, S., Will, C.L., Lührmann, R., Söll, D., and Martinez, J. (2011). HSPC117 Is the Essential Subunit of a Human tRNA Splicing Ligase Complex. Science *331*, 760-764.
32. Ray, A., Zhang, S., Rentas, C., Caldwell, K.A., and Caldwell, G.A. (2014). RTCB-1 mediates neuroprotection via XBP-1 mRNA splicing in the unfolded protein response pathway. J Neurosci *34*, 16076-16085.
33. Jurkin, J., Henkel, T., Nielsen, A.F., Minnich, M., Popow, J., Kaufmann, T., Heindl, K., Hoffmann, T., Busslinger, M., and Martinez, J. (2014). The mammalian tRNA ligase complex mediates splicing of XBP1 mRNA and controls antibody secretion in plasma cells. Embo j *33*, 2922-2936.
34. Tanaka, N., and Shuman, S. (2011). RtcB is the RNA ligase component of an Escherichia coli RNA repair operon. J Biol Chem *286*, 7727-7731.
35. Burroughs, A.M., and Aravind, L. (2016). RNA damage in biological conflicts and the diversity of responding RNA repair systems. Nucleic Acids Res *44*, 8525-8555.
36. Shuman, S. (2023). RNA Repair: Hiding in Plain Sight. Annu Rev Genet *57*, 461-489.
37. Wirth, A.N., Naarmann-de Vries, I.S., Pinnen, A.M., Gopal, A., Righetti, A., Leppek, K., Dieterich, C., and Peschek, J. (2026). Eukaryotic tRNA ligases mediate RNA break repair. bioRxiv, 2026.2005.2026.727988.
38. Lin, Z., Akin, H., Rao, R., Hie, B., Zhu, Z., Lu, W., Smetanin, N., Verkuil, R., Kabeli, O., Shmueli, Y., et al. (2023). Evolutionary-scale prediction of atomic-level protein structure with a language model. Science *379*, 1123-1130.
39. Fei, H., Li, Y., Liu, Y., Wei, J., Chen, A., and Gao, C. (2025). Advancing protein evolution with inverse folding models integrating structural and evolutionary constraints. Cell *188*, 4674-4692.e4619.
40. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A.J., Bambrick, J., et al. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature *630*, 493-500.
41. Dauparas, J., Anishchenko, I., Bennett, N., Bai, H., Ragotte, R.J., Milles, L.F., Wicky, B.I.M., Courbet, A., de Haas, R.J., Bethel, N., et al. (2022). Robust deep learning–based protein sequence design using ProteinMPNN. Science *378*, 49-56.
42. Hsu, C., Verkuil, R., Liu, J., Lin, Z., Hie, B., Sercu, T., Lerer, A., and Rives, A. (2022). Learning inverse folding from millions of predicted structures. In C. Kamalika, J. Stefanie, S. Le, S. Csaba, N. Gang, and S. Sivan, eds. Proceedings of the 39th International Conference on Machine Learning. PMLR.
43. Cao, C., Li, A., Xu, C., Wu, B., Liu, J., and Liu, Y. (2023). Enhancement of protein translation by CRISPR/dCasRx coupled with SINEB2 repeat of noncoding RNAs. Nucleic Acids Res *51*, e33.
44. Fiflis, D.N., Rey, N.A., Venugopal-Lavanya, H., Sewell, B., Mitchell-Dick, A., Clements, K.N., Milo, S., Benkert, A.R., Rosales, A., Fergione, S., and Asokan, A. (2024). Repurposing CRISPR-Cas13 systems for robust mRNA trans-splicing. Nat Commun *15*, 2325.

45. Chandrasekaran, S.S., Tau, C., Fu, B.X.H., Nemeth, M., Bartie, L., Pawluk, A., Konermann, S., and Hsu, P.D. (2026). Rewriting endogenous human transcripts with dual CRISPR-guided 3′ trans-splicing. Cell Systems *17*.
46. Konermann, S., Lotfy, P., Brideau, N.J., Oki, J., Shokhirev, M.N., and Hsu, P.D. (2018). Transcriptome Engineering with RNA-Targeting Type VI-D CRISPR Effectors. Cell *173*, 665-676 e614.
47. Zhang, B., Ye, Y., Ye, W., Perculija, V., Jiang, H., Chen, Y., Li, Y., Chen, J., Lin, J., Wang, S., et al. (2019). Two HEPN domains dictate CRISPR RNA maturation and target cleavage in Cas13d. Nat Commun *10*, 2544.
48. Lima, W.F., De Hoyos, C.L., Liang, X.H., and Crooke, S.T. (2016). RNA cleavage products generated by antisense oligonucleotides and siRNAs are processed by the RNA surveillance machinery. Nucleic Acids Res *44*, 3351-3363.
49. Ogami, K., Chen, Y., and Manley, J.L. (2018). RNA surveillance by the nuclear RNA exosome: mechanisms and significance. Noncoding RNA *4*.
50. Keidel, A., Long, C.L., Iwasa, J., and Conti, E. (2025). RNA-Degrading Exosome Complexes: Molecular Mechanisms and Structural Insights. Annu Rev Cell Dev Biol *41*, 505-528.
51. Nagarajan, V.K., Jones, C.I., Newbury, S.F., and Green, P.J. (2013). XRN 5'→3' exoribonucleases: structure, mechanisms and functions. Biochim Biophys Acta *1829*, 590-603.
52. Balaratnam, S., Hoque, M.E., West, N., and Basu, S. (2022). Decay of Piwi-Interacting RNAs in Human Cells Is Primarily Mediated by 5′ to 3′ Exoribonucleases. ACS Chemical Biology *17*, 1723-1732.
53. Lubas, M., Christensen, M.S., Kristiansen, M.S., Domanski, M., Falkenby, L.G., Lykke-Andersen, S., Andersen, J.S., Dziembowski, A., and Jensen, T.H. (2011). Interaction profiling identifies the human nuclear exosome targeting complex. Mol Cell *43*, 624-637.
54. Puno, M.R., and Lima, C.D. (2022). Structural basis for RNA surveillance by the human nuclear exosome targeting (NEXT) complex. Cell *185*, 2132-2147 e2126.
55. Gerlach, P., Garland, W., Lingaraju, M., Salerno-Kochan, A., Bonneau, F., Basquin, J., Jensen, T.H., and Conti, E. (2022). Structure and regulation of the nuclear exosome targeting complex guides RNA substrates to the exosome. Mol Cell *82*, 2505-2518 e2507.
56. Garland, W., and Jensen, T.H. (2024). Nuclear sorting of short RNA polymerase II transcripts. Mol Cell *84*, 3644-3655.
57. Schmidt, C.A., Giusto, J.D., Bao, A., Hopper, A.K., and Matera, A G. (2019). Molecular determinants of metazoan tricRNA biogenesis. Nucleic Acids Research *47*, 6452-6465.
58. Peschek, J., Acosta-Alvear, D., Mendez, A.S., and Walter, P. (2015). A conformational RNA zipper promotes intron ejection during non-conventional XBP1 mRNA splicing. EMBO Rep *16*, 1688-1698.
59. Schmitt-Ulms, C., Kayabolen, A., Manero-Carranza, M., Zhou, N., Donnelly, K., Nuccio, S.P., Kato, K., Nishimasu, H., Gootenberg, J.S., and Abudayyeh, O.O. (2024). Programmable RNA writing with trans-splicing. bioRxiv, 2024.2001.2031.578223.
60. Song, J., Zhuang, Y., and Yi, C. (2024). Programmable RNA base editing via targeted modifications. Nature Chemical Biology *20*, 277-290.
61. Chen, P.J., Hussmann, J.A., Yan, J., Knipping, F., Ravisankar, P., Chen, P.F., Chen, C., Nelson, J.W., Newby, G.A., Sahin, M., et al. (2021). Enhanced prime editing systems by manipulating cellular determinants of editing outcomes. Cell *184*, 5635-5652 e5629.
62. Yan, J., Oyler-Castrillo, P., Ravisankar, P., Ward, C.C., Levesque, S., Jing, Y., Simpson, D., Zhao, A., Li, H., Yan, W., et al. (2024). Improving prime editing with an endogenous small RNA-binding protein.

Nature *628*, 639-647.

63. Colognori, D.A., Wasko, K.M., Trinidad, M.I., Zhou, Z., and Doudna, J.A. (2026). Spligation enables programmable chimeric RNA generation in living cells. bioRxiv, 2026.2003.2006.709984.

64. Kato, K., Zhou, W., Okazaki, S., Isayama, Y., Nishizawa, T., Gootenberg, J.S., Abudayyeh, O.O., and Nishimasu, H. (2022). Structure and engineering of the type III-E CRISPR-Cas7-11 effector complex. Cell *185*, 2324-2337 e2316.

MATERIALS AND METHODS

Cell culture

HEK293T cells were cultured in DMEM with high glucose (Hyclone), containing 10% (v/v) FBS (Biological Industries), 1% (v/v) penicillin-streptomycin (Hyclone) at 37°C in 5% CO_2. The cell line was kindly gifted from Caixia Gao Lab and is checked routinely for mycoplasma by PCR detection.

Plasmid construction

The Csm plasmids with mCherry or puroR were modified from pDAC627 (Addgene plasmid #195240; http://n2t.net/addgene:195240; RRID: Addgene_195240) and synthesized by Tsingke Biotechnology. The spacer region contains two *Bbs*I-recognizing sites, facilitating convenient spacer alteration through the Golden Gate method. The shRNA used in this study was driven by a U6 promoter. The spacer as well as shRNA sequences used in this study were listed in Supplementary Table S1.

The reporter plasmid was modified from pcDNA3.1 backbone: The Nt-GFP fragments were cloned into the MCS downstream of the original CMV promoter; then another CMV_promoter-Ct_GFP-bGHpA_terminator genes were inserted at *Pci*I restriction sites through Giboson Assembly (One Step Seamless Cloning Mix, CWBIO). For endogenous RNA modification assay, the RNA donor was cloned from HEK293T cDNA and verified by Sanger sequencing, also driven by the CMV promoter. For RtcB assay, RtcB derived from *E.coli* or *P. horikoshii* was synthesized by Tsingke Biotechnology, and HsRtcB was cloned from HEK293T cDNA. RtcB with or without N-NLS was inserted into the reporter plasmid under the control of SV40 or EF1α promoter. The gene sequences noted above were listed in Supplementary Table S3.

Transfection and microscopy

4.5×10^4 HEK293T cells were seeded at 48-well plate 24 h before transfection. 250 ng Csm plasmid and 125 ng reporter/donor plasmid per well were incubated with 0.75 μL LFDR transfection reagent (YALEPIC) in Opti-MEM (Gibco) following the manufacturer's instructions. Cells were passaged at 24 h post transfection for sub-confluency and collected at 48 h for flow-cytometry and RNA preparation. The fluorescence signal was also observed at 48 h with Nikon Eclipse Ts2-FL through DAPI, FITC and TRITC filter sets. Microscopy images were processed and analyzed with FIJI (ImageJ).

Flow-cytometry

HEK293T cells were dissociated with 0.25% trypsin (Hyclone), resuspended in 10% FBS-containing PBS, and washed once using PBS (Hyclone). Then Single-cell suspensions were analyzed using CytoFLEX LX (Beckman Coulter). For experiments using mCherry as a selection marker, transfected cells were gated based on the Y610-mCherry channel. Green fluorescence was measured in the B525-FITC channel. Untreated cells and cells transfected with NT Csm plasmid were used as controls.

RT-qPCR

1×10^5 HEK293T cells were collected for RNA extraction using AxyPrep RNA Miniprep Kit (Axygen), which was further reverse transcribed into cDNA using HiScript II Q RT SuperMix (Vazyme, Cat. No. R223-01). The cDNA served as templates for quantitative PCR (qPCR) to detect the RNA level. qPCR reactions using the iTaq Universal SYBR Green Supermix (Bio-Rad) on a CFX96 real-time system (Bio-Rad) were conducted with the following settings: preheating (95°C for 3 min); 40 cycles of alternating within denaturation (95°C for 5 s), annealing and extension (60°C for 30 s). Relative quantification (for detection of knock down efficiency) was performed using the ΔΔCt method with GAPDH as the internal control. Absolute quantification (for measurement of ligation efficiency) was conducted using a standard curve generated from serially diluted plasmid standards.

Next-generation sequencing and data analysis

For NGS, around 200-nt regions containing the cleavage and editing sites were amplified from cDNA. The 1st round PCR used 2× Phanta Max Master Mix (Vazyme) and primers containing adaptor sequences. Then the 2nd round PCR were performed to introduce index and P5/P7 region. Amplicons were purified by VAHTS DNA Clean Beads (Vazyme) and sequenced using the DNBSEQ-T7 PE150 platform. Primers used in the experiments above are shown in Supplementary Table S2.

Raw sequencing reads containing the target sites were processed with Trim Galore for quality filtering and assembled using FLASH2. Assembled reads with or without the 20-nt unique donor-derived sequences were separated using seqkit grep with mismatches no more than 2 (parameter: -m 2). Then the reads with donor sequence were aligned to template with intended editing by CRISPResso2 with the quantification window center set at the intended cleavage site of Csm as well as the following parameters: -w 30 --plot_window_size 30 --ignore_substitutions. The reads without donor sequence were aligned to the original receptor sequence. Total reads number equals to the sum of these two kinds of aligned reads. The ligation efficiency was calculated by the number of reads with donor divided by the number of total reads. The intended editing efficiency was calculated by the number of precise editing reads divided by the number of total reads (Supplementary Fig. S3b).

Sequence- and structure-based computational design of RtcB variants

For the sequence-based strategy, the 650M ESM-2 model was used to calculate, at each residue position, the conditional probability of each of the 20 standard amino acids given the surrounding sequence context. For every possible single-amino-acid substitution, P_{Mut} was defined as the model-assigned probability of the substituted amino acid. All candidate substitutions were ranked according to P_{Mut}, and 22 high-ranking single-amino-acid variants were selected for experimental testing.

For the structure-based strategy, the AlphaFold 3-predicted structure of EcRtcB was used as the structural

template for sequence evaluation with ProteinMPNN and ESM-IF. ProteinMPNN was used to sample 30 times at each position while keeping all other residues fixed, and the resulting sampling frequencies were used to estimate site-specific amino acid preferences. In parallel, sequences with all possible single-amino-acid substitutions were generated and evaluated using the ESM inverse-folding model. For each sequence, a structure-conditioned sequence-compatibility score was calculated and compared with that of the wild-type sequence, yielding an ESM-IF score representing the predicted change in compatibility with the input structure. Candidate variants were selected for experimental testing when they exhibited both a ProteinMPNN sampling frequency greater than 75% and an ESM-IF score greater than 0.0015.

AUTHOR CONTRIBUTIONS

L.Q.L., Y.Y.J., and J.J.G.L. conceptualized and designed the research. R.K.C. proposed and conducted the recruitment experiment. Y.Y.J., L.Q.L., R.K.C., and G.L.W. did the molecular cloning experiments. L.Q.L., R.K.C., and M.Y.H. conducted cell transfection. L.Q.L., R.H.Z. and C.Q.W. performed microscopy analysis. R.K.C., L.Q.L., and S.J. conducted flowcytometry analysis. L.Q.L., Z.H.L. and S.L.J. did the RT-qPCR assay. Y.Y.J., L.Q.L., Y.Y., and R.H.Z. performed next-generation sequencing and efficiency measurement. S.L. conducted the AI-guided optimization. Y.Y.J., L.Q.L., R.K.C. and S.J. wrote the manuscript with input from all authors. J.J.G.L. supervised the project and secured funding.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (32425030 to J.J.G.L. and 32501286 to Y.Y.J.); the National Key R&D Program (2022YFF1002801 to J.J.G.L.); the Agriculture Science and Technology Major Project (to J.J.G.L.); the New Cornerstone Science Foundation through the XPLORER PRIZE (J.J.G.L.).

COMPETING INTERESTS

The authors declare no competing interests.

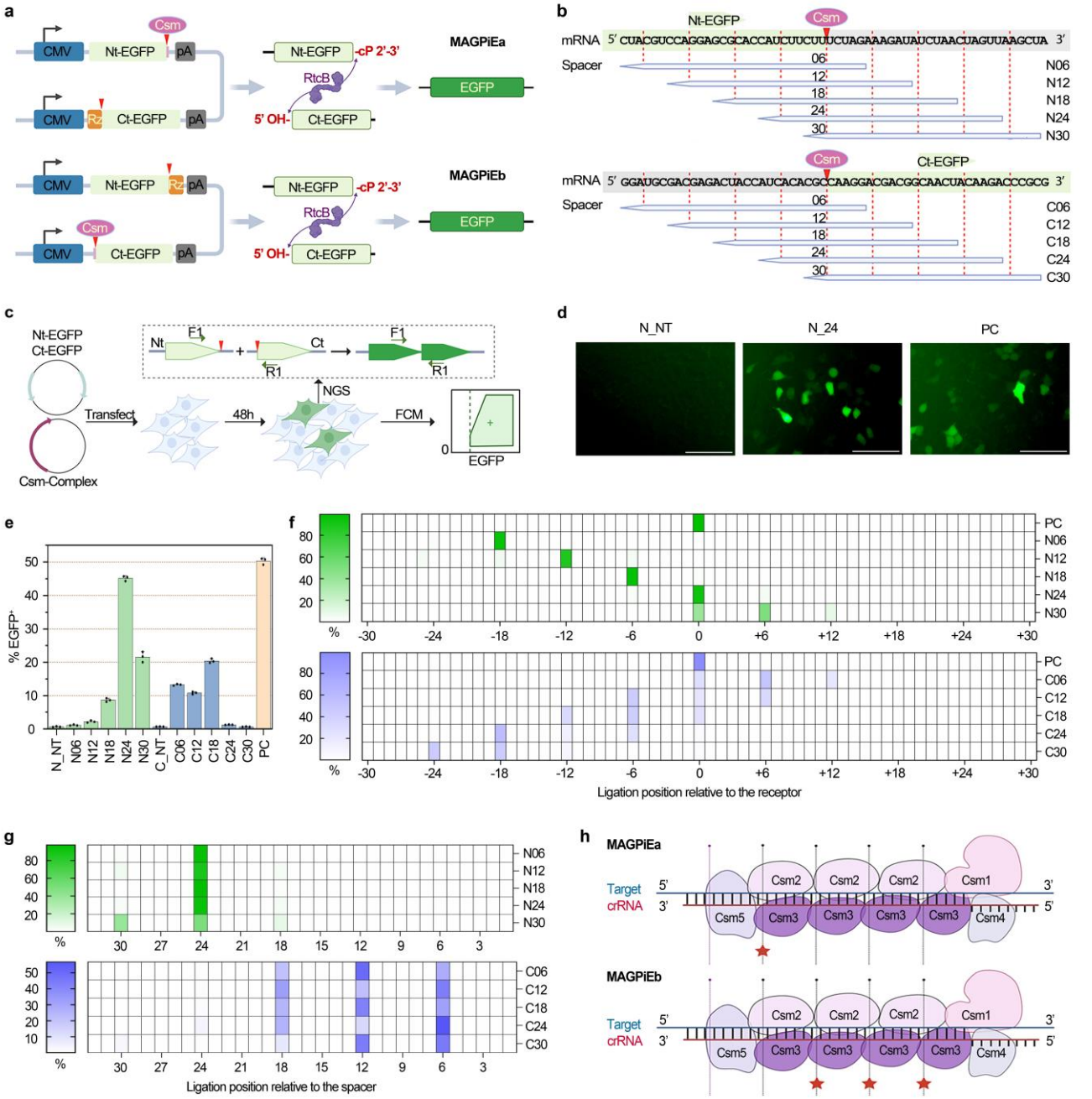


Figure 1. MAGPiE achieved cleavage and ligation of RNA receptor and donor.

(a) Schematic of the split-EGFP reporter. Two configurations are shown: in MAGPiEa, Csm targets Nt-EGFP while a ribozyme (Rz) is fused to Ct-EGFP; in MAGPiEb, Csm targets Ct-EGFP while Rz is fused to Nt-EGFP. Csm cleavage and ribozyme self-cleavage generate a 2′,3′-cyclic phosphate (cP 2’-3’) and a 5′-hydroxyl (5’ OH) terminus, respectively, which are ligated by endogenous RtcB to reconstitute full-length EGFP.

(b) crRNA design. crRNAs targeting either fragment at 6-nt intervals are named according to the expected cleavage position relative to the spacer 5′ end.

(c) Experiment workflow. Co-transfection of split-EGFP reporters, Csm complex and crRNAs, followed by flow cytometry (FCM) and next-generation sequencing (NGS).

(d) Representative fluorescence images for N24 in MAGPiEa, non-targeting control (NT), and positive control (PC). Scale bar, 100 μm.

(e) Flow cytometry analysis of EGFP-positive ($EGFP^+$) cells. PC, positive control. NT, non-target. Values and error bars represent the means and s.d. from three independent biological replicates.

(f) Distribution of ligation position relative to the receptor. Positive values indicate longer receptor fragments than those generated at the intended cleavage site, whereas negative values indicate shorter ones. Zero represents precise ligation site.

(g) Distribution of ligation position relative to the spacer 5′ end.

(h) Schematic summarizing the cleavage–ligation patterns. Red asterisks mark the major ligation sites. For MAGPiEa, a dominant ligation site is located at position 24, last cleavage site of the four-Csm3 complex as shown; the minor population of five-Csm3 complex leading to cleavage at position 30 (purple dashed line) is not depicted. For MAGPiEb, multiple ligation sites are indicated.

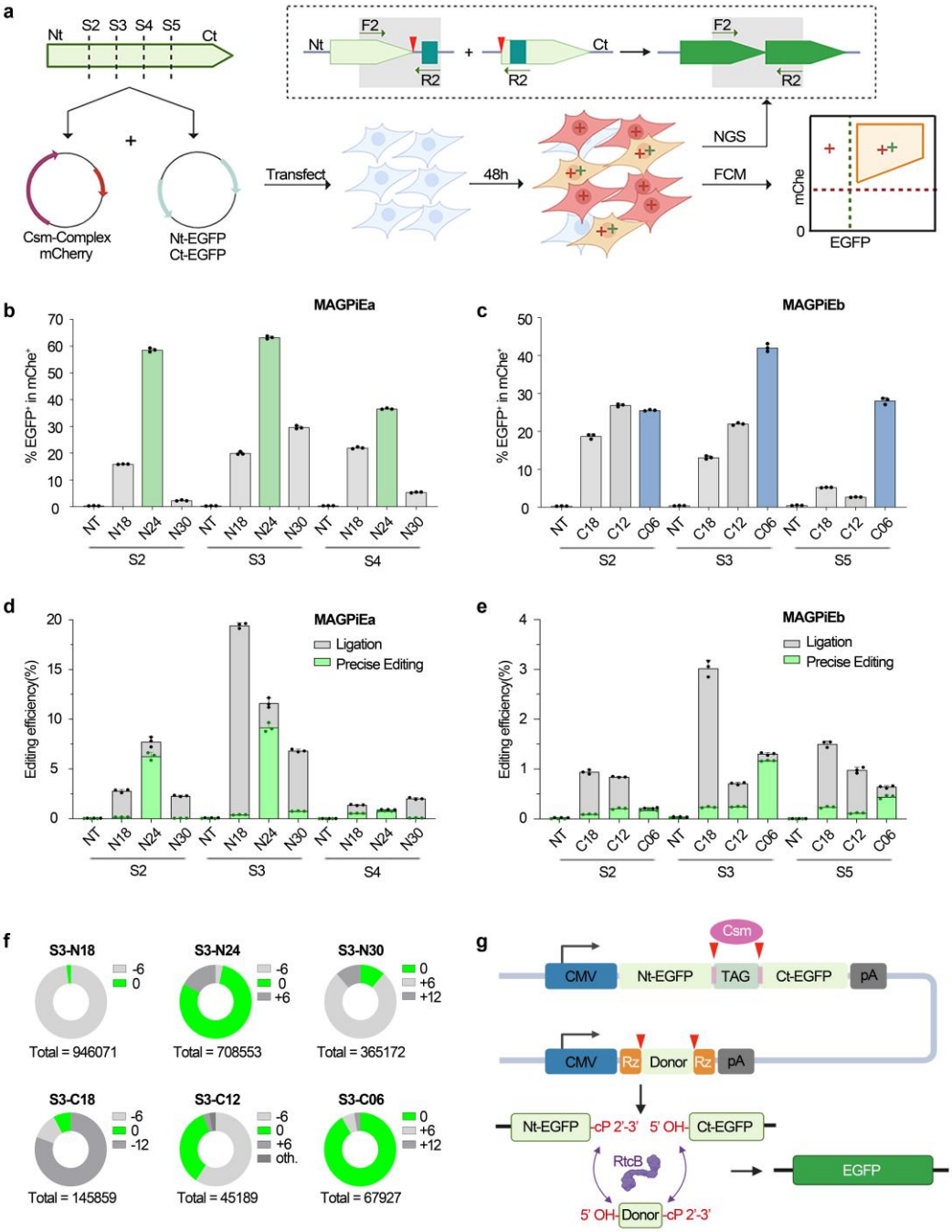


Figure 2. Verification of MAGPiE at multiple sites and attempt at three-fragment ligation.

(a) Workflow for fluorescence reconstitution assay at multiple split sites. Csm complex is co-expressed with mCherry as a selection marker.

(b and c) Flow cytometry analysis of EGFP-positive cells gated on $mCherry^+$ cells for MAGPiEa (b) and MAGPiEb (c).

(d and e) NGS quantification of ligation and precise editing efficiencies for MAGPiEa (d) and MAGPiEb (e) configurations.

(f) NGS analysis of ligation products at S3 site. Pie charts show the composition of ligation events for indicated targets.

(g) Schematic of three-fragment ligation. Cleavage at the S2 and S5 sites by Csm removes the intervening stop codon and generates two flanking RNA receptors. The RNA donor carries RzB HH and Twister ribozyme at its 5′ and 3′ ends, respectively, creating compatible termini with the two RNA receptors and enabling RtcB ligation.

Values and error bars in (b), (c), (d), and (e) are presented as the means and s.d. from three independent biological replicates.

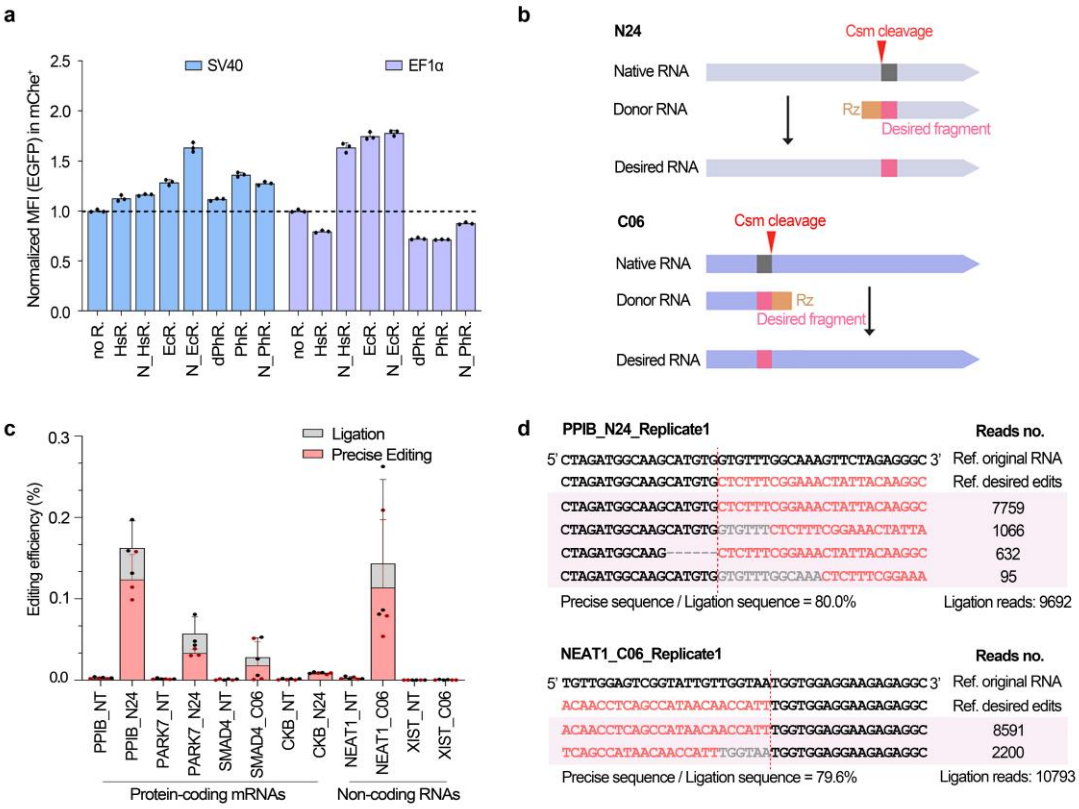


Figure 3. RtcB-optimized MAGPiE2.0 improves ligation and enables endogenous RNA replacement.

(a) Optimization of RtcB constructs. Comparison of RtcB orthologs from *Homo sapiens* (HsR.), *Escherichia coli* (EcR.) and *Pyrococcus horikoshii* (PhR.), with or without nuclear localization signal (NLS) as denoted by the "N_" prefix, driven by SV40 or EF1α promoters. dPhR. is an inactive variant of PhRtcB, harboring a C98A mutation. Mean fluorescence intensity (MFI) normalized to control is shown (no R.).

(b) Schematic of RNA fragment replacement designs. Desired fragments are shown in pink, ribozymes in orange, and fragments to be modified in gray.

(c) Quantification of fragment replacement efficiency by NGS.

(d) Representative NGS read alignments for PPIB and NEAT1. Desired fragments are shown in pink, and non-precise editing events are shown in gray.

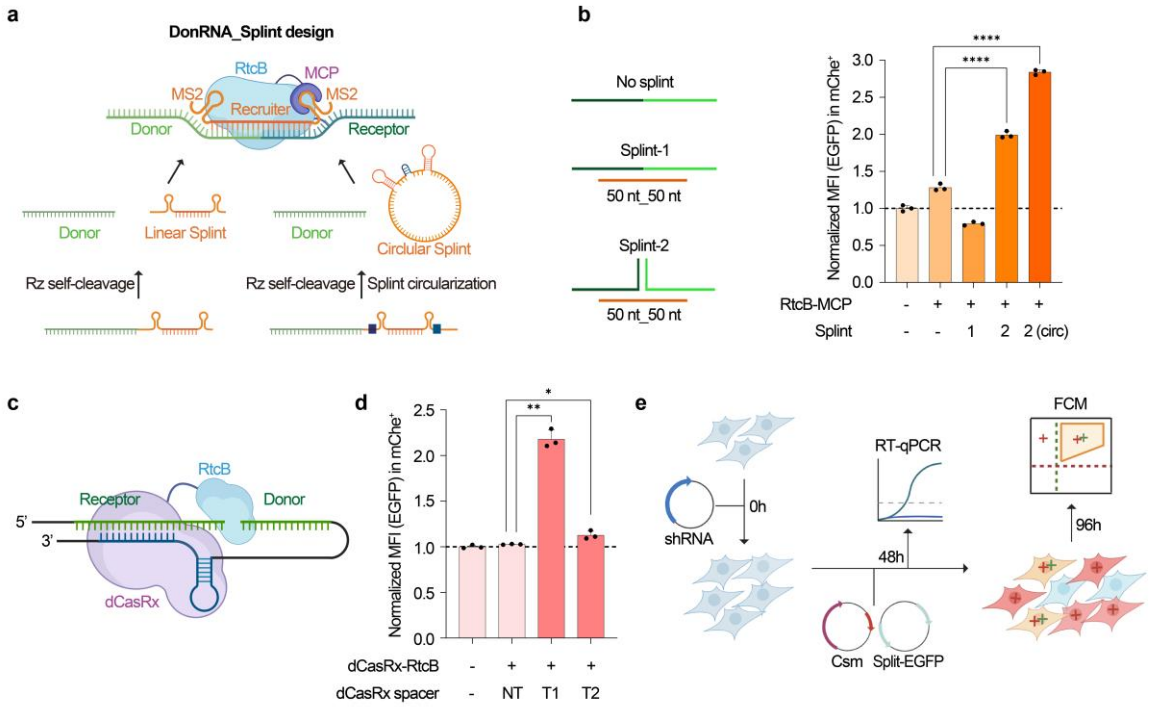


Figure 4. Engineering MAGPiE components and cellular determinants for enhanced ligation efficiency.

(a) Schematic of linear and circular DonRNA_splint strategies. The DonRNA_splint is expressed in cis with the donor transcript and released by ribozyme cleavage at the donor-distal terminus, whereas MCP-fused RtcB is recruited via the MS2-splint.

(b) Measurement of fluorescence reconstitution efficiency with different strategies. MFI normalized to control is shown.

(c) Schematic of dCasRx-RtcB mediated donor recruitment. Donor RNA is fused with crRNA cassette; dCasRx contains HEPN-inactivating and K940A mutations and is N-terminally fused with RtcB.

(d) Measurement of split-EGFP reporter activity with different spacers. The conditions include control (no spacer), non-targeting spacer (NT), and two targeting spacers (T1 and T2). MFI values are normalized to the control.

(e) Workflow for assessing the effect of exoribonucleases knockdown. shRNA-mediated knockdown of exoribonucleases was performed. qPCR was conducted at 48 h to assess knockdown efficiency. MAGPiE system was also transfected at 48 h. Flow cytometry was performed at 96 h (48 h after MAGPiE transfection).

(f) Measurement of MAGPiE efficiency upon knockdown of components in endogenous RNA decay pathways. MFI normalized to control (NC) is shown.

Values and error bars in (b), (d), and (f) are presented as the means and s.d. from three independent biological replicates. Statistical significance was determined by unpaired *t*-test with Welch's correction. *$p < 0.05$, **$p < 0.01$, ****$p < 0.0001$.

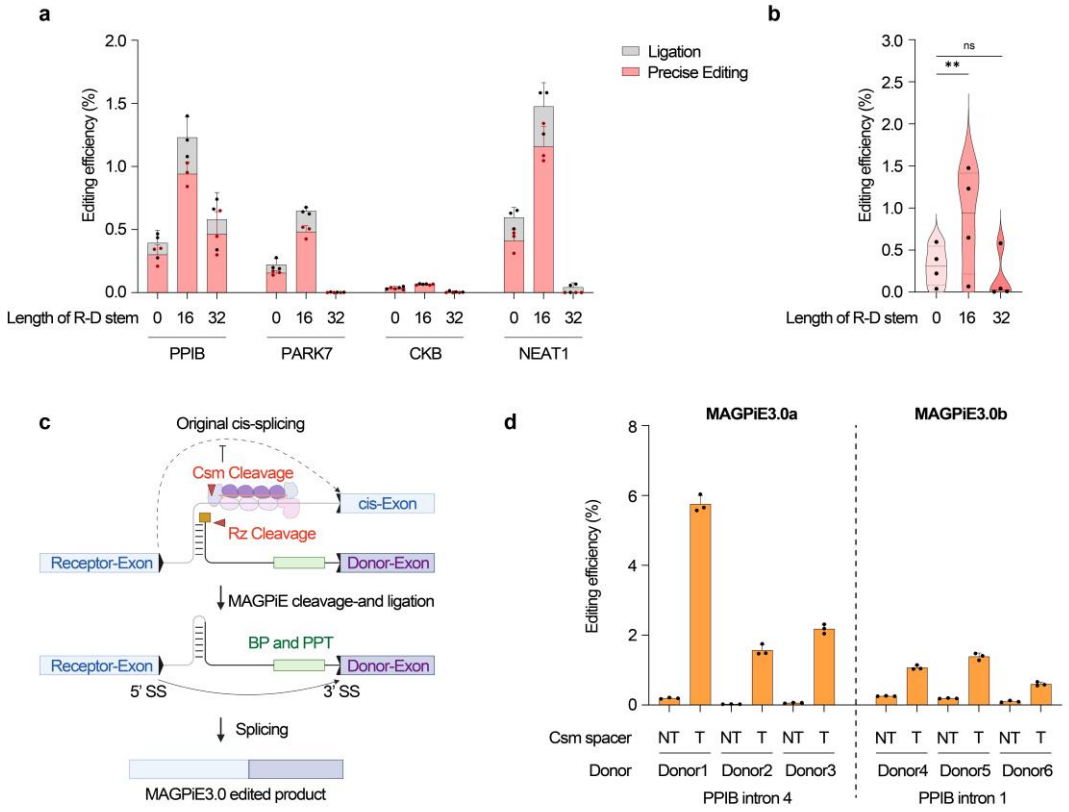


Figure 5. MAGPiE3.0-mediated ligation via stem-loop duplex and its utilization in intronic regions.

(a) Ligation efficiency of MAGPiE3.0 at four endogenous loci with designed stem-loop lengths of 0, 16, or 32 bp between the receptor and donor.

(b) Violin plot of ligation efficiencies from the four loci shown in (a), grouped by stem-loop length. Statistical significance was assessed using ratio paired *t*-test. $^{**}p < 0.01$; ns, not significant.

(c) Schematic of MAGPiE3.0-mediated ligation targeted to intronic regions. The donor RNA is designed to form a duplex with the Csm-cleaved receptor, which can be removed by the spliceosome after ligation. Key splicing elements are indicated, including the branch-point sequence (BP), polypyrimidine tract (PPT), 3′ splice site (3′SS), and 5′ splice site (5′SS).

(d) Ligation efficiency of MAGPiE3.0 targeting the intron of PPIB gene with MAGPiE3.0a and MAGPiE3.0b, each designing three donor (e.g., Donor1, Donor2, and Donor3). NT, non-target; T, target.

Values and error bars in (a) and (d) are presented as the means and s.d. from three independent biological replicates.

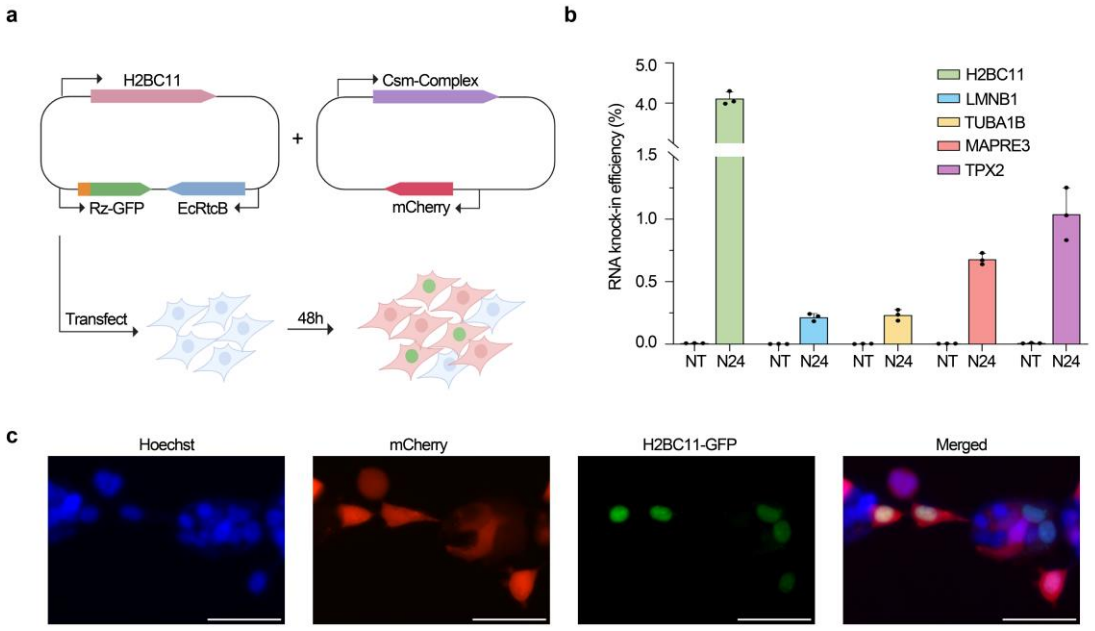


Figure 6. MAGPiE-mediated RNA-level knock-in for fluorescence tagging.

(a) Experimental workflow for RNA knock-in, as shown for H2BC11 and similarly applied to other target genes.

(b) RT-qPCR analysis showing RNA knock-in for the indicated target genes. Values are shown as ligated products relative to receptor expression with error bars representing the mean ± s.d. from three independent biological replicates.

(c) Representative fluorescence images of H2BC11-EGFP knock-in. Nuclear-localized EGFP signal indicates in-frame fusion. Scale bar, 50 μm.

Supplemental information

MAGPiE enables precise and versatile RNA manipulation via cleavage and ligation

Long-Qi Li[1,3], Yuan-Yuan Jiang[1,2,3,*], Rui-Kang Chen[1,3], Rui-Han Zhang[1], Shuo Lin[1], Zhi-Hui Liu[1], Shuai Jin[1], Yun Yang[1], Chao-Qun Wu[1], Shu-Lin Jin[1], Gao-Li Wang[1], Ming-Yu Hou[1], Jun-Jie Gogo Liu[1,4,*]

[1]Beijing Frontier Research Center for Biological Structure, State Key Laboratory of Membrane Biology, Center for Synthetic and Systematic Biology, Tsinghua-Peking Center for Life Sciences, School of Life Sciences, Tsinghua University, Beijing 100084, China

[2]State Key Laboratory of Crop Gene Resources and Breeding, Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing, 100081, China

[3]These authors contributed equally

[4]Lead contact

*Correspondence: jiangyuanyuan@caas.cn (Y.Y.J.); junjiegogoliu@tsinghua.edu.cn (J.-J.G.L.)

Supplementary Figures

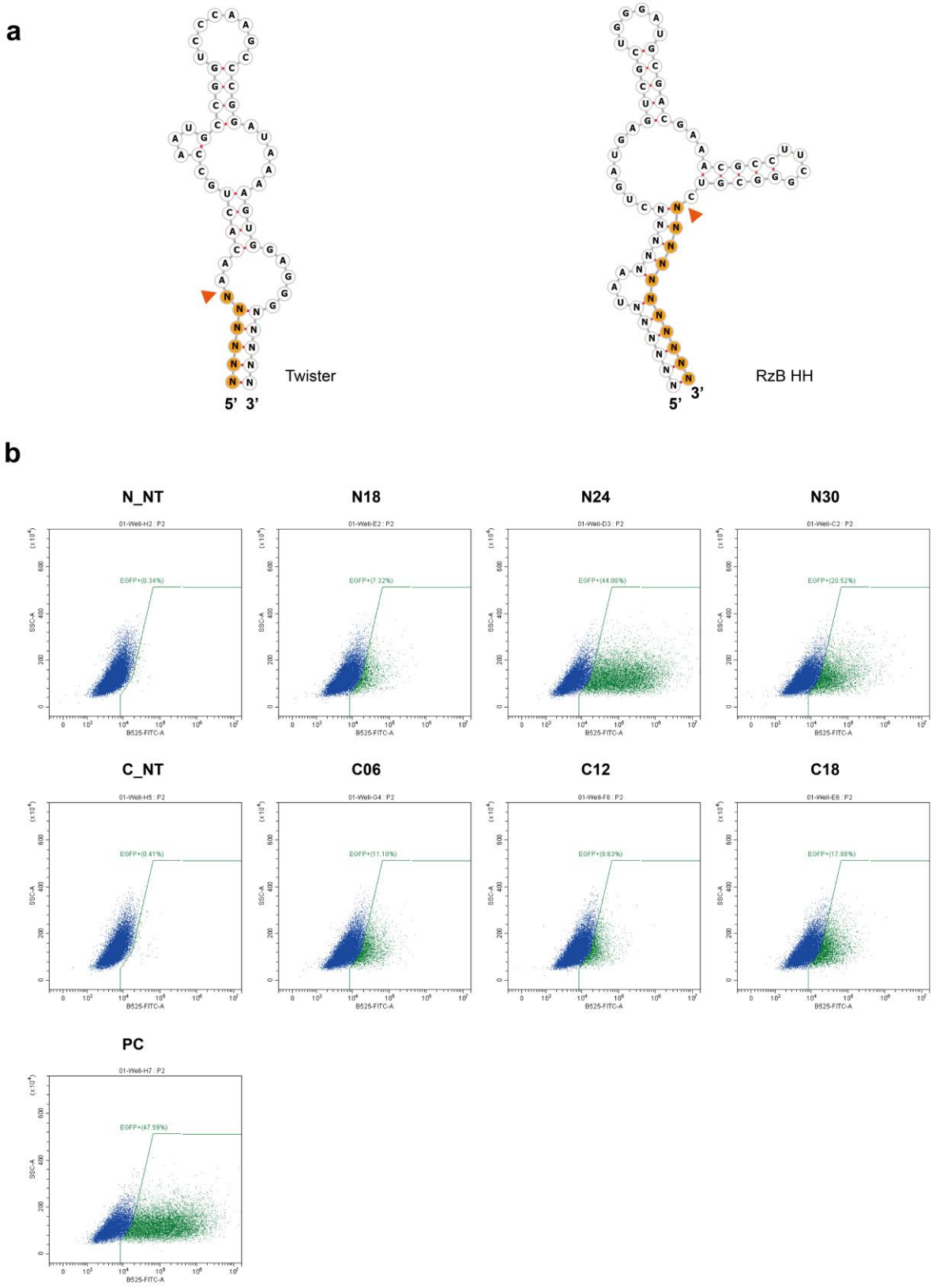


Figure S1. Secondary structure of ribozyme and representative flow cytometry gating plots of GFP reconstitution.

(a) Secondary structure of the Twister ribozyme and the RzB hammerhead (HH) ribozyme. Red arrows indicate the self-cleavage sites, and nucleotides in orange denote the donor sequence.

(b) Representative flow cytometry plots with gating to define the $EGFP^+$ cell population.

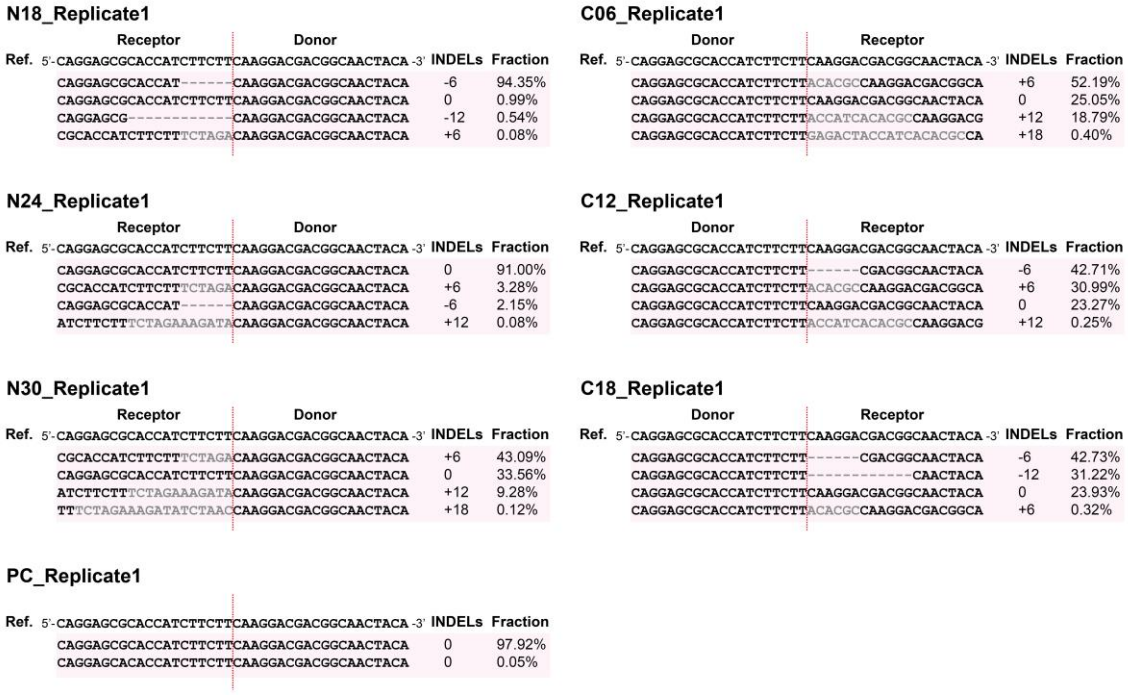


Figure S2. NGS read alignments of reconstituted EGFP mRNA.

Representative NGS read alignments. Inserted or deleted nucleotides are shown in gray, and the mutation type and frequency are indicated on the right.

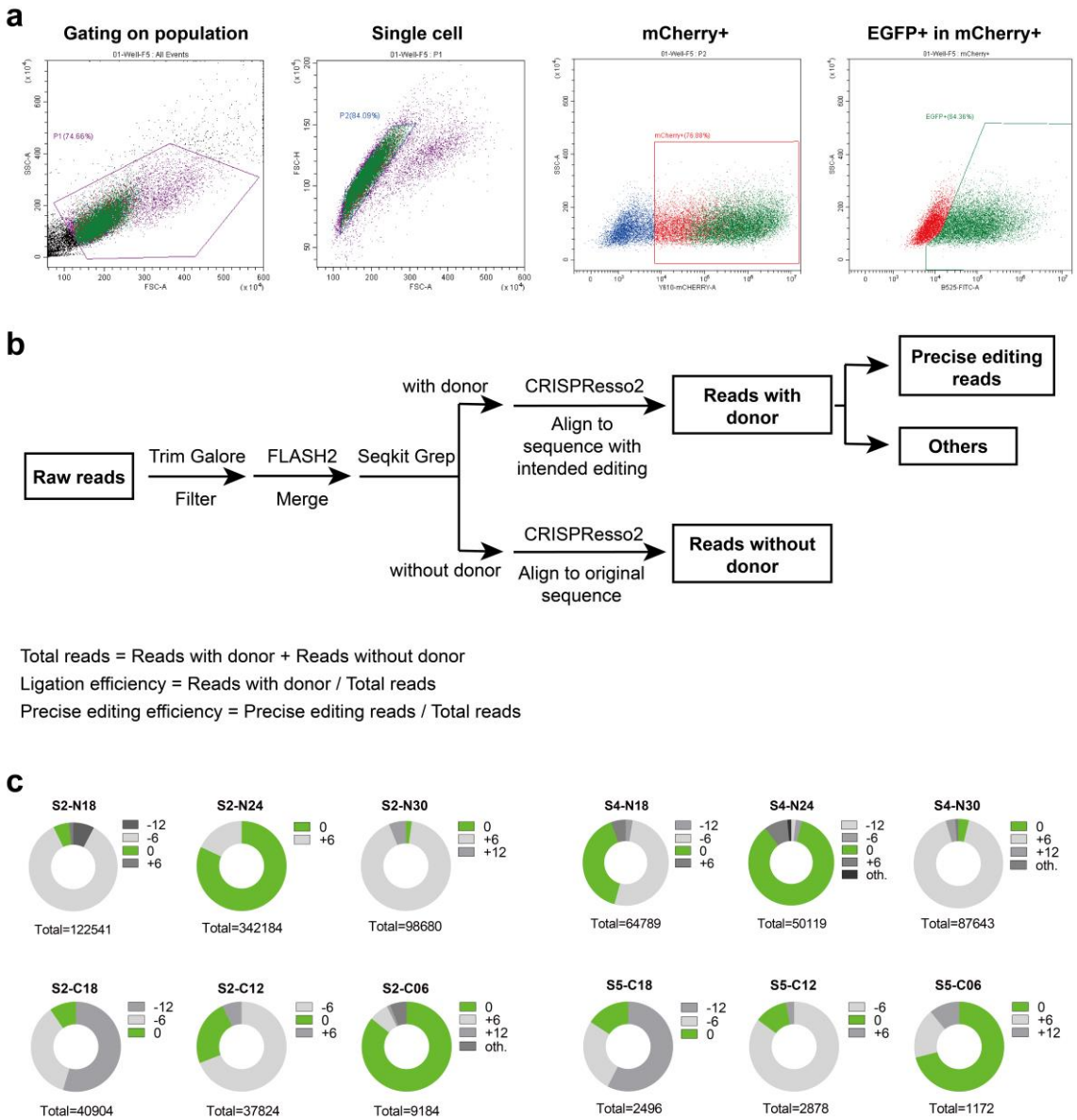
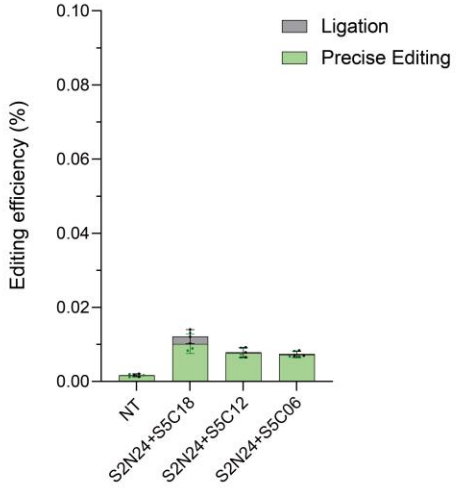


Figure S3. Dual-color flow cytometry gating plots and NGS-based analysis of ligation efficiency and types.
(a) Gating strategy to determine the percentage of EGFP⁺ in mCherry⁺ cells. Cells are initially gated for viability and singularity, and subsequently for mCherry positivity to select transfected cells. The proportion of EGFP-positive cells is quantified.
(b) Workflow of NGS analysis for quantifying ligation efficiency and precise editing efficiency.
(c) NGS analysis of ligation products at the S2, S4, and S5 site. Pie charts show the distribution of ligation events for indicated targets.

a



b

S2N24+S5C18_Replicate1

Ref. 5'-GCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCT.....CTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGG-3'	INDELs	Reads no.
GCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCT.....CTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGG	0 / 0	1093
GCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCT.....CTGGTGAACCGCATCGAGCT------CATCGACTTCAAGG	0 / -6	134
CCCGAAGGCTACGTGGTCCTCCAGGAGCGCACCATCTTCT.....CTGGTGAACCGCATCGAGCT------CATCGACTTCAAGG	+6 / -6	27
GCCATGCCCGAAGG------CCAGGAGCGCACCATCTTCT.....CTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGG	-6 / 0	3
GCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCT.....CTGGTGAACCGCATCGAGCT----------CTTCAAGG	0 / -12	2

S2N24+S5C12_Replicate1

Ref. 5'-GCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCT.....CTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGG-3'	INDELs	Reads no.
GCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCT.....CTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGG	0 / 0	1206
GCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCT.....CTGGTGAACCGCATCGAGCT------CATCGACTTCAAGG	0 / -6	44

S2N24+S5C06_Replicate1

Ref. 5'-GCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCT.....CTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGG-3'	INDELs	Reads no.
GCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCT.....CTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGG	0 / 0	920
CCCGAAGGCTACGTGGTCCTCCAGGAGCGCACCATCTTCT.....CTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGG	+6 / 0	1

Figure S4. Three-fragment ligation with both MAGPiEa and MAGPiEb.

(a) NGS-based quantification of ligation efficiency and precise editing efficiency for three-fragment ligation.

(b) Representative NGS read alignments. Inserted or deleted nucleotides are shown in gray. The mutation types at the two ligation junctions as well as the reads number are indicated on the right.

.

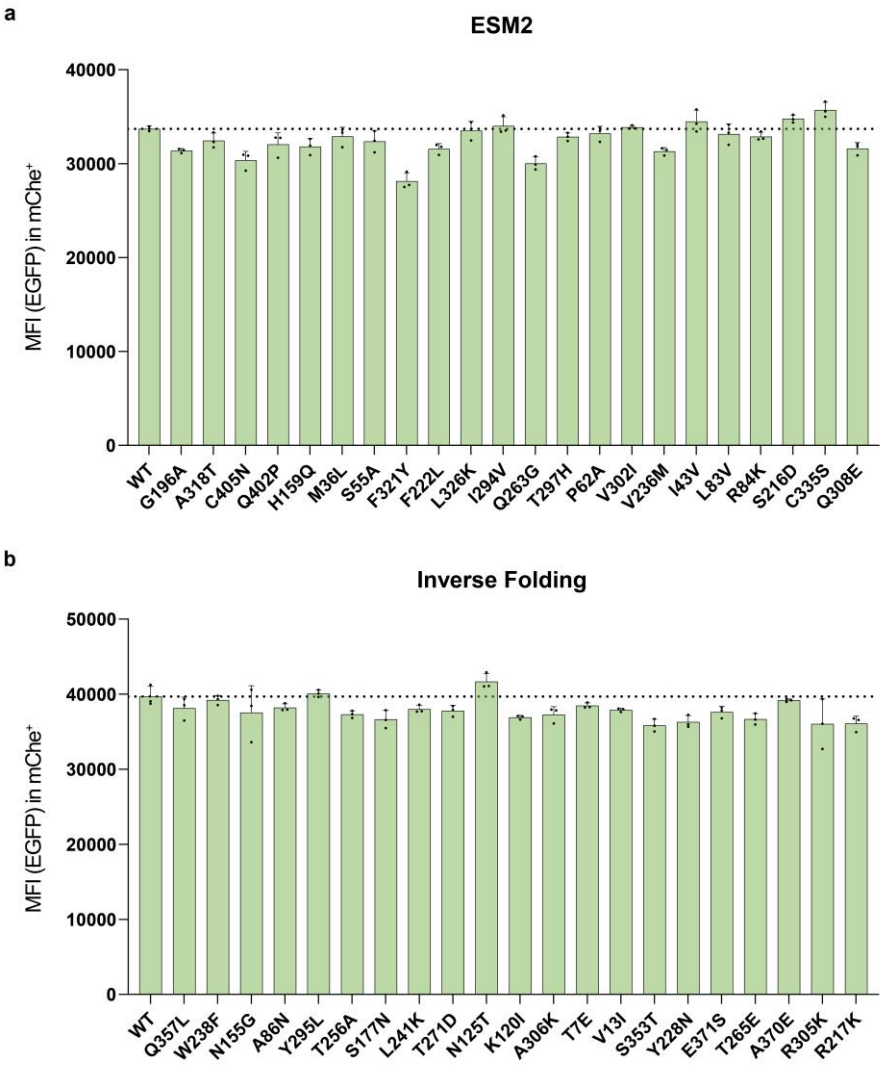


Figure S5. AI-guided optimization of RtcB variants evaluated by split-EGFP reporter.

(a) EGFP intensity of ESM-2-predicted RtcB variants in the flowcytometry assay.

(b) EGFP intensity of Inverse-folding-designed RtcB variants in the flowcytometry assay.

Data are presented as mean ± s.d. from three independent biological replicates.

Figure S6. Schematic of endogenous RNA fragment replacement and representative NGS read alignments.
(a) Schematic illustration of endogenous gene fragment replacement. The crRNA targeting sites are marked with asterisks, and the modifications are indicated below in orange. For protein-coding genes, the CDS are shown as orange solid boxes, and UTR correspond to tan hollow boxes. For ncRNAs, the exons are orange hollow boxes. Introns are represented by folded lines. The transcripts are not drawn to scale.
(b) Representative NGS results of endogenous gene fragment replacement**.** Pink regions represent successfully modified sequences, and gray regions denote insertions and deletions.

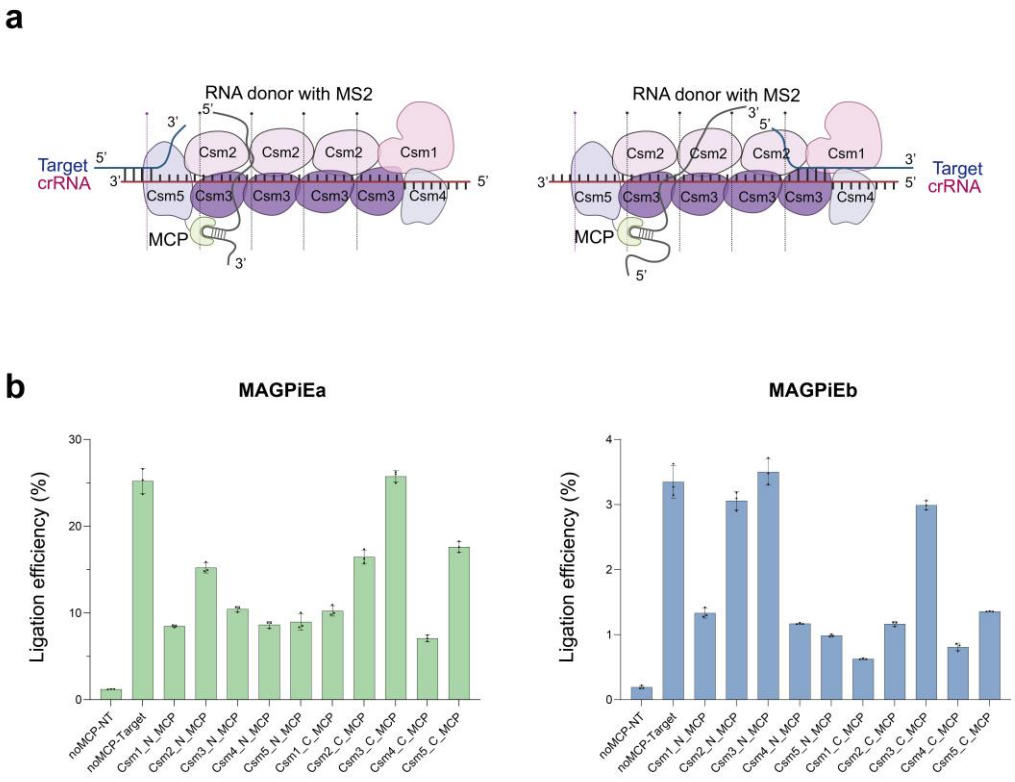


Figure S7. Recruitment of MS2-tagged donor by MCP-fused Csm.

(a) Schematic of MCP-MS2-based donor recruitment. MS2 is inserted into the UTR of the RNA donor, while MCP is individually fused to Csm1-Csm5 protein via a GS linker, facilitating recruitment of the RNA donor.

(b) Ligation efficiency was quantified by RT–qPCR. The RNA receptor and ligated products are absolutely quantified using serially diluted plasmids as standards. Ligation efficiency is defined as the ratio of the copy number of ligated products to that of the RNA receptor.

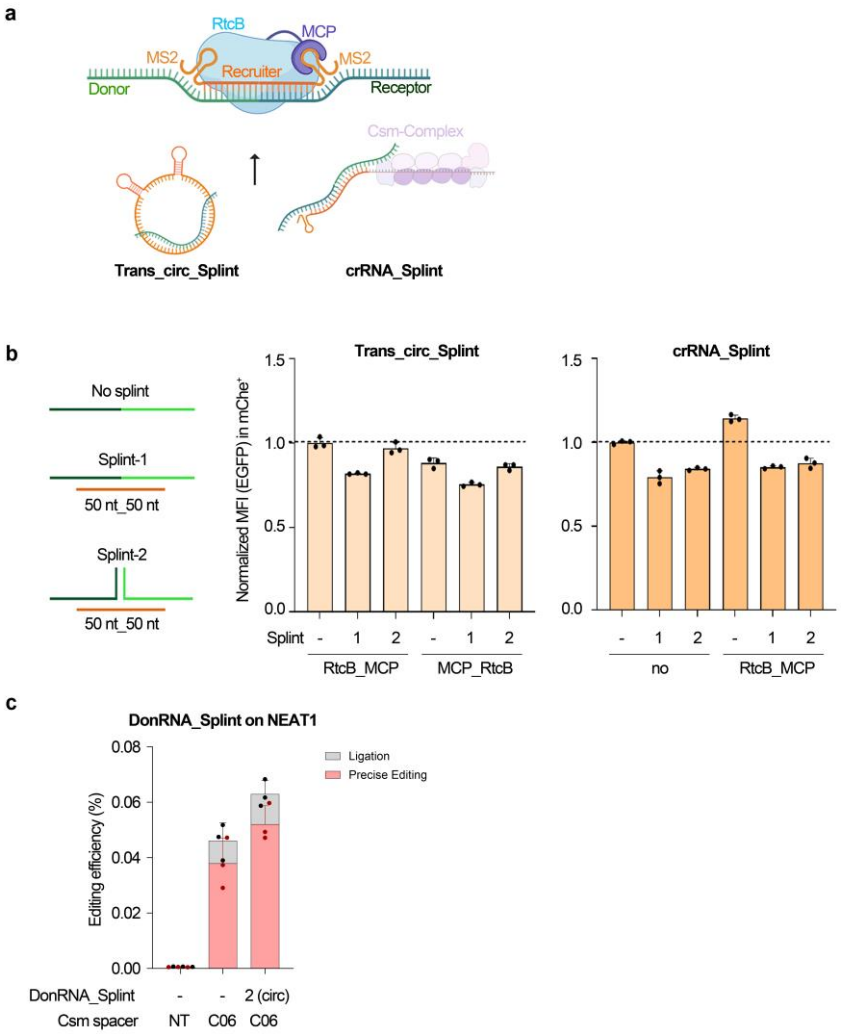


Figure S8. Splint-mediated recruitment strategies and evaluation.

(a) Schematic of two MS2-based splint formats for donor–receptor recruitment: free circular splint (Trans_circ_Splint) and splint appended to the 3′ end of crRNA (crRNA_Splint). MCP is fused to RtcB, recruited by MS2-splint.

(b) Measurement of fluorescence reconstitution efficiency with different strategies. Each strategy includes two distinct designs (Splint-1, Splint-2). MFI normalized to control is shown.

(c) Editing efficiency of MAGPiE variants at the endogenous NEAT1 locus under indicated conditions.

Values and error bars in (b) and (c) are presented as the means and s.d. from three independent biological replicates.

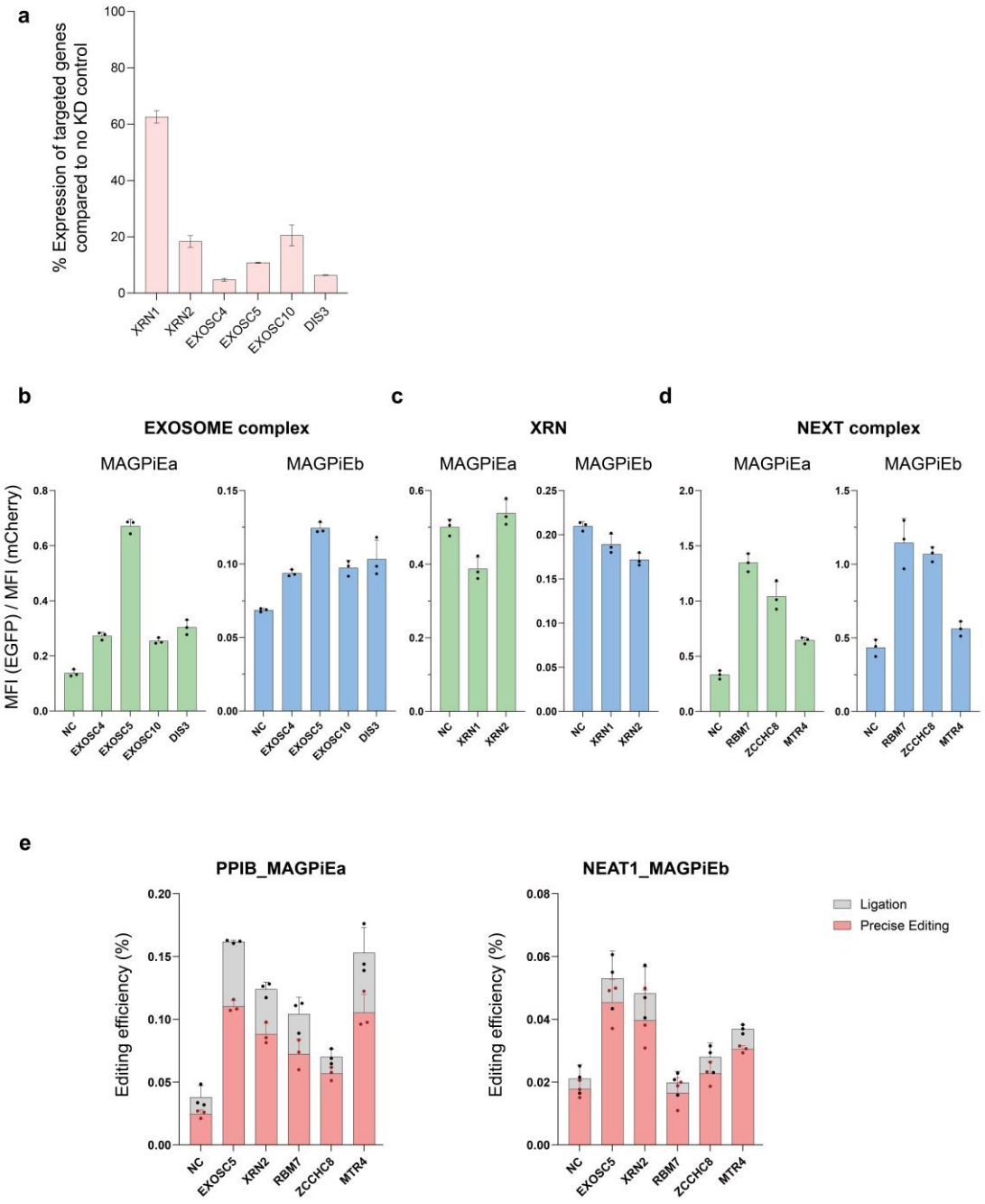


Figure S9. Validation of MAGPiE efficiency upon knockdown of components in RNA decay pathways.

(a) RNAi efficiency of shRNA treatment relative to a non-targeting shRNA control defined as $2^{-\Delta\Delta Ct}$. Relative expression of each RNA was normalized to GAPDH as internal reference.

(b-d) Ratio of MFI (EGFP) to MFI (mCherry). The fluorescence intensity of untargeted mCherry was also measured and used as an internal control to account for global effects of exonuclease knockdown on mRNA stability.

(e) Editing efficiency of MAGPiE variants at the endogenous PPIB and NEAT1 locus under indicated conditions.

All the values and error bars represent the means and s.d. from three independent biological replicates.

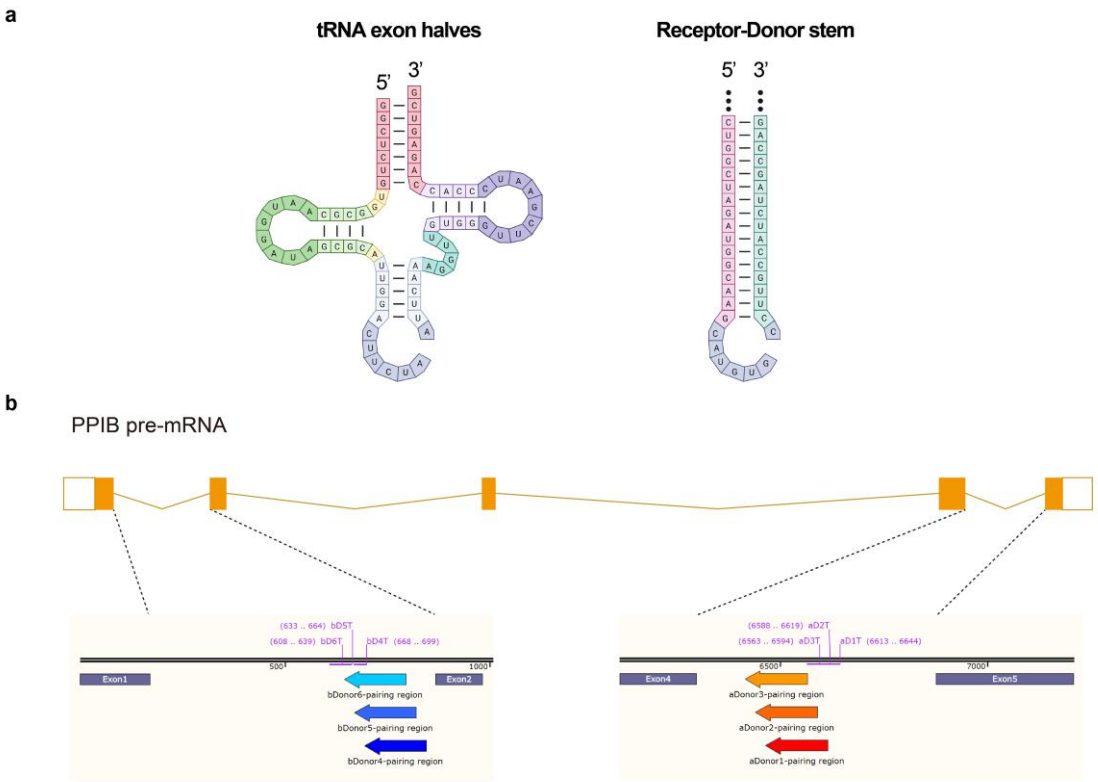


Figure S10. MAGPiE3.0 design and intronic targeting.

(a) Schematic of the receptor–donor stem-loop mimicking native tRNA exon halves. Left: tRNA exon halves. Right: receptor–donor stem-loop design. Related to Fig. 5(a) and 5(b).

(b) Design of MAGPiE3.0-mediated targeting of the PPIB pre-mRNA intron. Related to Fig. 5(d).

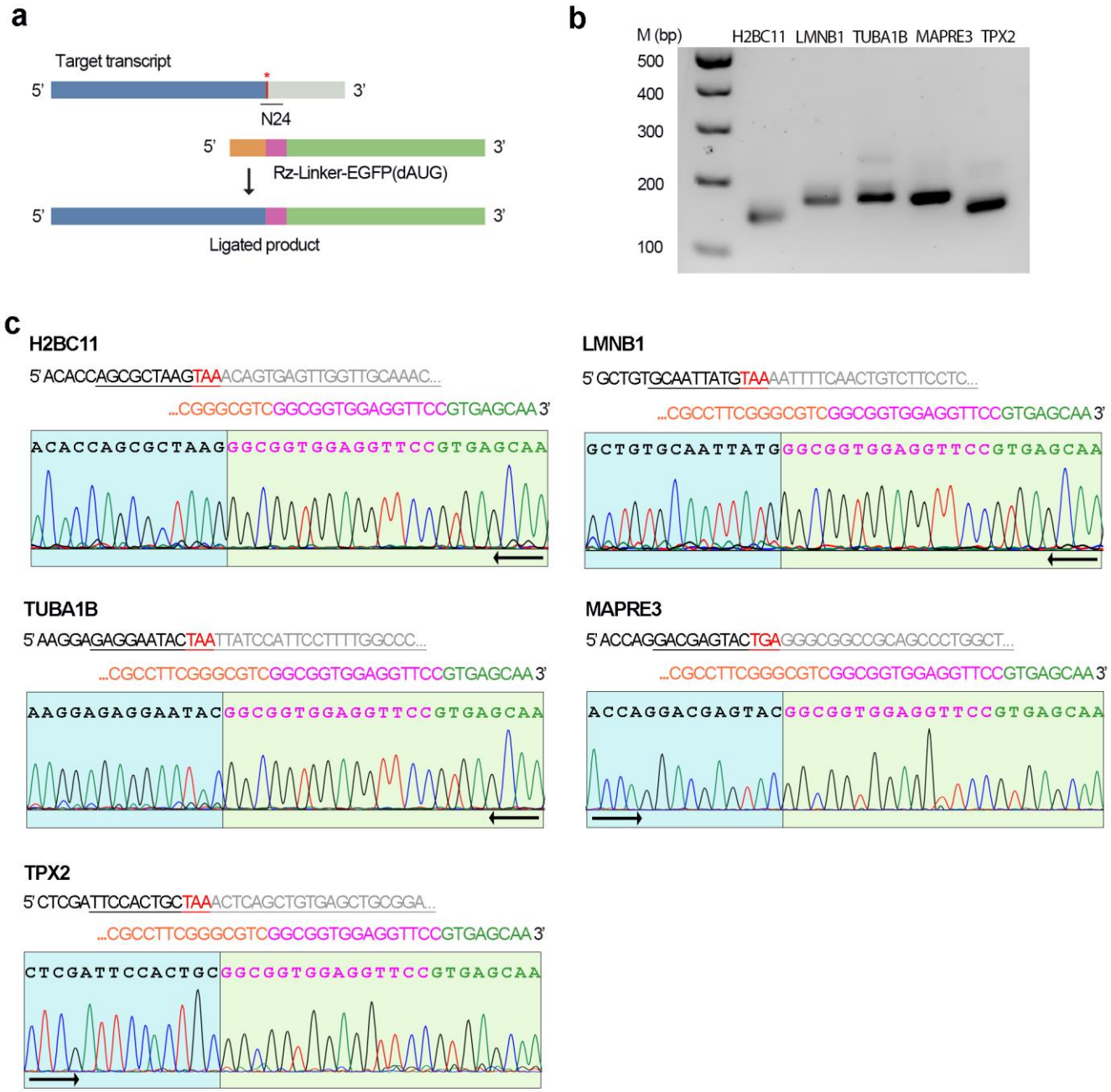


Figure S11. Characterization for RNA knock-in.

(a) Schematic of RNA knock-in targeting a native transcript. The target transcript is shown with the stop codon (indicated by a red asterisk) and the crRNA-targeted site (N24, black line). The donor consists of a self-cleaving ribozyme (Rz, orange), a flexible linker (GGGGS, pink), and EGFP lacking its start codon (green). Cleavage by Csm and the ribozyme, followed by RtcB ligation, generates the full-length product, forming an in-frame EGFP fusion.

(b) Agarose gel electrophoresis of PCR products amplified using primers flanking the ligation junction. Each lane corresponds to one indicated gene.

(c) Representative sequencing chromatograms of the ligated product. Above each chromatogram, the reference sequences of the target transcript and donor are shown. Arrows below the chromatograms indicate the sequencing direction. The stop codon is highlighted in red, and the 3′ UTR region is shown in gray. Additional labeling follows the conventions in (A).

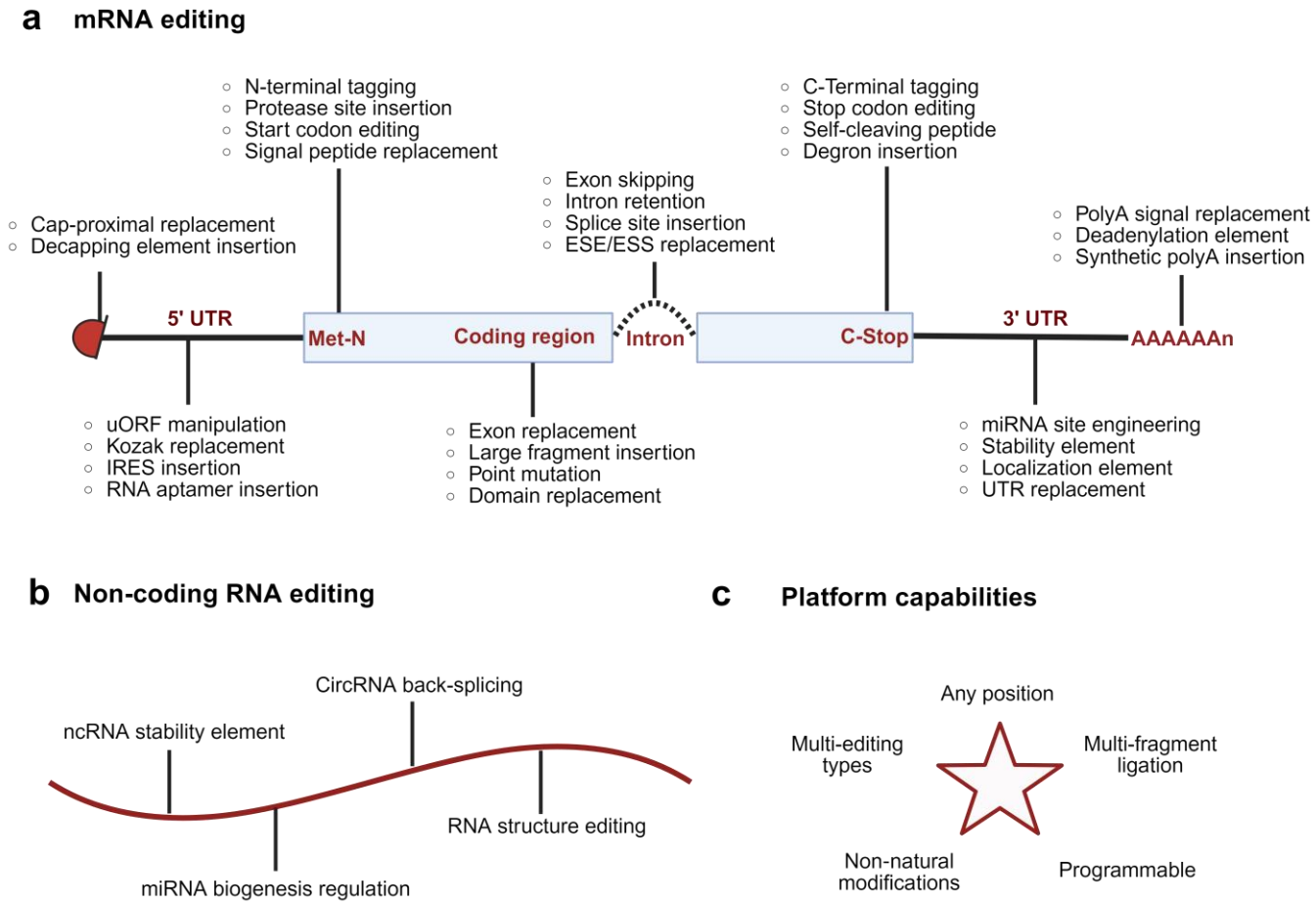


Figure S12. Summary of MAGPiE-mediated RNA editing capabilities.

(a) mRNA editing. Schematic of a representative RNA transcript (5′ cap, 5′ UTR, coding region, intron, 3′ UTR, polyA). Functional capabilities are categorized by region and positioned adjacent to the corresponding RNA elements.

(b) Non-coding RNA editing. Functional applications for non-coding RNAs, including lncRNAs, circRNAs and miRNA precursors, are summarized.

(c) Platform capabilities. A pentagon diagram illustrating core features of the MAGPiE system: any position, multi-fragment ligation, multi-editing types, non-natural modifications, as well as high programmability.