

Multi-color RNA imaging with CRISPR-Cas13b systems in living cells

Liang-Zhong Yang^{a,1}, Bao-Qing Gao^{b,1}, Youkui Huang^a, Ying Wang^b, Li Yang^{c,d}, Ling-Ling Chen^{a,d,e,*}



^a State Key Laboratory of Molecular Biology, Shanghai Key Laboratory of Molecular Andrology, CAS Center for Excellence in Molecular Cell Science, Shanghai Institute of Biochemistry and Cell Biology, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai, China

^b CAS Key Laboratory of Computational Biology, Shanghai Institute of Nutrition and Health, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai, China

^c Center for Molecular Medicine, Children's Hospital, Fudan University and Shanghai Key Laboratory of Medical Epigenetics, International Laboratory of Medical Epigenetics and Metabolism, Ministry of Science and Technology, Institutes of Biomedical Sciences, Fudan University, Shanghai, China

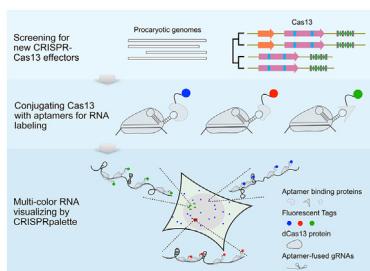
^d School of Life Science and Technology, ShanghaiTech University, Shanghai, China

^e School of Life Science, Hangzhou Institute for Advanced Study, University of Chinese Academy of Sciences, Hangzhou, China

HIGHLIGHTS

- Identification of eight catalytically dead CRISPR-(d)Cas13 proteins for RNA imaging.
- Robust tracking of RNAs with dHgm4Cas13b and dMisCas13b in living cells.
- Pre-crRNA processing null dCas13b enables RNA labelling with RNA aptamers.
- CRISPRpalette allows multi-color simultaneous labeling of RNAs in living cells.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
dCas13
Pre-crRNA processing
RNA aptamer
RNA tracking
NEAT1
GCN4

ABSTRACT

Visualizing RNA dynamics is important for understanding RNA function. Catalytically dead (d) CRISPR-Cas13 systems have been established to image and track RNAs in living cells, but efficient dCas13 for RNA imaging is still limited. Here, we analyzed metagenomic and bacterial genomic databases to comprehensively screen Cas13 homologues for their RNA labeling capabilities in living mammalian cells. Among eight previously unreported dCas13 proteins that can be used for RNA labeling, dHgm4Cas13b and dMisCas13b displayed comparable, if not higher, efficiencies to the best-known ones when targeting endogenous *MUC4* and *NEAT1_2* by single guide (g) RNAs. Further examination of the labeling robustness of different dCas13 systems using the *GCN4* repeats revealed that a minimum of 12 *GCN4* repeats was required for dHgm4Cas13b and dMisCas13b imaging at the single RNA molecule level, while >24 *GCN4* repeats were required for reported dLwaCas13a, dRfxCas13d and dPguCas13b. Importantly, by silencing pre-crRNA processing activity of dMisCas13b (ddMisCas13b) and further incorporating RNA aptamers including PP7, MS2, Pepper or BoxB to individual gRNAs, a CRISPRpalette system was developed to successfully achieve multi-color RNA visualization in living cells.

* Corresponding author. State Key Laboratory of Molecular Biology, Shanghai Key Laboratory of Molecular Andrology, CAS Center for Excellence in Molecular Cell Science, Shanghai Institute of Biochemistry and Cell Biology, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai, China.

E-mail address: linglingchen@sibcb.ac.cn (L.-L. Chen).

¹ These authors contributed equally to this work.

1. Introduction

RNA plays important roles in cells from guiding protein synthesis to acting as regulatory molecules, which are highly associated with their dynamics and subcellular localizations. To track motions of RNAs in a spatiotemporal manner, robust imaging techniques have been developed to visualize RNAs particularly in living cells, by taking advantages of fluorescent probes and RNA aptamers. On the one hand, using chemically modified fluorescent RNA or DNA probes, target RNAs containing sequences complementary to these fluorescent probes can be detected, exemplified by using molecular beacon (MB) to image RNAs that are artificially integrated with at least eight MB-recognized repeats (Chen et al., 2017). On the other hand, when target RNAs are inserted with tandem repeats of RNA aptamers with genetic manipulation, such RNAs can be lighted up by fluorescence-tagged RNA binding proteins (RBPs) or fluorophores. Using RNA aptamer-based approaches, events of RNA processing like transcription, splicing, translocation, translation, and decay have been well analyzed (Bertrand et al., 1998; Cawte et al., 2020; Chen et al., 2009, 2019; Daigle and Ellenberg, 2007; Larson et al., 2011; Wu et al., 2019). Very recently, incorporated with hybridization probes and aptamers, the CRISPR-Cas systems have been also adopted to image RNAs in living cells with different settings (Nelles et al., 2016; Wang et al., 2019; Yang et al., 2019).

A CRISPR-Cas system contains single or multiple Cas protein effectors, as well as a guide RNA (gRNA) and sometimes with a scaffold RNA (Makarova et al., 2020). The gRNA is composed of a spacer similar to an RNA probe targeting an endogenous RNA, and a direct repeat (DR) recognized by a Cas protein. Two CRISPR-Cas systems have been exploited to track RNAs in living cells: RNA-targeting Cas9 (RCas9) and CRISPR-dCas13. RCas9 is derived from the DNA-targeting CRISPR-Cas9 system but possesses capability to target RNAs for imaging (Chen et al., 2013; Cong et al., 2013; Jinek et al., 2012; O'Connell et al., 2014; Strutt et al., 2018), thus allowing to track *ACTB* mRNAs in stress granules and pathogenically expressed microsatellite repeat expansions (*MRE* RNAs) in stressed cells (Batra et al., 2017; Nelles et al., 2016). Other than RCas9, CRISPR-Cas13 is originally identified as an RNA-guided, RNA-targeting system, including Cas13a, b, c, and d families (Abudayeh et al., 2017; Konermann et al., 2018; Shmakov et al., 2015; Smargon et al., 2017). With appropriate gRNAs, EGFP-fused, catalytically dead CRISPR-(d)LwaCas13a was expressed in cells to image *ACTB* mRNAs in stress granules (Abudayeh et al., 2017). *In vitro* purified, fluorescently modified gRNAs delivered into cells that express dRfxCas13d has also enabled visualization of transcription sites of 24 × *MS2*-labeled nascent RNAs (Wang et al., 2019). In addition, with two orthogonal CRISPR-dPspCas13b and dPguCas13b, our recent study has enabled simultaneous co-labeling of the long noncoding RNA (lncRNA) *NEAT1* in parasplices and *SatIII* in nuclear stress bodies (nSBs) (Yang et al., 2019).

Despite of these progresses, most CRISPR-dCas13 systems, except dPspCas13b, displayed limited labeling capability (Yang et al., 2019), which constrains their broad application in RNA imaging, such as co-labeling distinct RNA species in a multi-color manner. To achieve this goal, here we analyzed metagenomic and bacterial genomic databases and identified 46 previously unreported CRISPR-dCas13 candidates. After mammalian codon-optimization and fused with EGFP, eight of 46 previously unreported CRISPR-dCas13 showed RNA labeling capability. Notably, two such CRISPR-dCas13 homologies, dHgm4Cas13b and dMisCas13b, displayed a comparable labeling efficiency, if not higher, to that of dPspCas13b in labeling *MUC4*, *NEAT1_2* and *GCN4* repeats. By combining pre-crRNA processing null dMisCas13b (ddMisCas13b) with RNA aptamer-fused gRNAs, we further developed a system, named CRISPRpalette, which has visualized RNAs in triple colors in living cells, expanding the application of CRISPR-Cas13 system in RNA imaging.

2. Results

2.1. Computational profiling of CRISPR-Cas13 candidates

To identify additional CRISPR-Cas13 proteins for RNA imaging, we developed a computational pipeline to systematically screen metagenomic and bacterial genome databases in a stepwise manner (Fig. 1A). First of all, given that an active Cas13 is normally accompanied by CRISPR arrays, we searched such repeat sequences in 571,097 archaeal and bacterial genomes (step 1, Fig. 1A). Among 166,441 genomic DNA sequences that contain 865,661 CRISPR arrays, we extracted all possible open reading frames (ORFs) embedded in these DNA sequences, which led to the identification of 3,222,345,443 ORFs (step 2, Fig. 1A). Next, by using previously reported Cas13 proteins (Makarova et al., 2020) as inputs, 6,084 sequences of possible Cas13-like ORFs longer than 500 amino acids (aa) were obtained within 10 kb next to the identified CRISPR arrays (step 3, Fig. 1A). Given that most of active Cas13 systems contain two ribonuclease HEPN domains (Cox et al., 2017; Konermann et al., 2018; Mahas et al., 2021; Shmakov et al., 2015; Smargon et al., 2017), 1,224 out of 6,084 Cas13-like ORFs were selected using the criterion of containing at least two Rxx...xxH motifs (length of amino acids between R and H ranges from 4 to 7) (step 4, Fig. 1A). Next, ORFs with ≤ 80% identity to reported Cas13 were collected (step 5, Fig. 1A), which were further combined to generate a list of 215 previously-undescribed Cas13 candidates (step 6, Fig. 1A, Table S2).

Based on their origins, lengths of ORFs, and low sequence similarities to reported Cas13 systems, 46 Cas13 candidate ORFs were ultimately selected to test their RNA labeling potential experimentally. Among these previously-unreported 46 ORFs, 12 belong to CRISPR-Cas13a, 33 to CRISPR-Cas13b, and one to CRISPR-Cas13c (Fig. 1B). Most of these Cas13 ORFs encode Cas13 effectors in the length between 1,000 and 1,250 aa, while the CRISPR-Cas13c homologous ORF encodes an 839 aa protein, which is much shorter than those by reported Cas13c ORFs (Fig. 1C). Further characterization showed that all 46 CRISPR-Cas13 sequences were close (within 10 kb) to CRISPR repeats containing at least four base pairs in the corresponding predicted secondary structures. Meanwhile, some (CRISPR-Cas13a and CRISPR-Cas13b) were near to sequences for WYL proteins, reported as regulators of the CRISPR-Cas13d system (Yan et al., 2018), and/or sequences of reverse transcriptase in an opposite direction (Fig. S1). In addition to ORFs, analyzing the DR sequences led to a similar phylogenetic tree cluster of these 46 ORFs, indicating that DR sequences in CRISPR arrays can be also used to bait CRISPR-Cas13 candidates (Fig. S2).

2.2. Screening CRISPR-Cas13 proteins for effective RNA labeling

To examine their capabilities for RNA imaging, we synthesized sequences individually of all 46 ORFs with mammalian codon optimization and mutations at both HEPN motifs for corresponding nuclease-dead Cas13 protein (dCas13) expression. Further, sequences of nuclear localization signal (NLS) and EGFP were fused to each dCas13 ORF sequence at its 5' or 3' end to generate the corresponding dCas13-EGFP expressing plasmid (Fig. S3A). We chose endogenous *MUC4* as the target RNA for the screening, since *MUC4* contains highly repetitive sequences in its exon2 (Chen et al., 2013; Nollet et al., 1998) (Fig. 2A) and also because *MUC4* has been successfully labeled at the single molecule level in cells by our previously defined dPspCas13b-EGFP system (Yang et al., 2019).

Here, we delivered each of 46 dCas13-EGFP expressing plasmid with the plasmid carrying the corresponding gRNAs that target *MUC4* into HeLa cells to imaging endogenous *MUC4* RNA, using dPspCas13b-EGFP and dPguCas13b-EGFP as positive controls (Yang et al., 2019). As a result, eight dCas13b-EGFP proteins showed detectable *MUC4* signals

(Fig. 2B and Fig. S3B). Among them, dHgm4Cas13b-EGFP, dMisCas13b-EGFP and dPba3Cas13b-EGFP exhibited the highest labeling efficiency, comparable to that of dPspCas13b-EGFP (Fig. 2C). Meanwhile, dBba2Cas13b-EGFP, dPba4Cas13b-EGFP, dPba5Cas13b-EGFP and dHgm6Cas13b-EGFP showed the medium labeling efficiency, similar to that of dPguCas13b-EGFP, while dPbaCas13b showed the weakest labeling efficiency (Fig. 2C). Further MUC4 RNA FISH in these dCas13-effectively labeled cells confirmed that the observed dCas13-EGFP signals were from the endogenously expressed MUC4 RNAs (Fig. 2D). We noticed that not all smFISH signals were labeled by dCas13b proteins and that smFISH signals of dCas13 labeled MUC4 presented comparably lower intensities. This was in part due to that the accessibility of RNA transcripts is occupied by dCas13 proteins because the target sequence of smFISH is the same as sgRNA used for dCas13b (Fig. 2A).

After confirming their labeling efficiencies for endogenous RNAs, we set to evaluate capabilities of these eight dCas13 proteins in single RNA molecule imaging, by using 24 or 16 repeats of GCN4 as readouts (Fig. 3A) as previously reported (Yang et al., 2019). To achieve this goal, each of eight dCas13 proteins was individually fused with $3 \times$ sfGFP in the C-terminus, two NLS in the N- and one NLS in the C-terminus for efficient RNA labeling (Fig. S4A) (Yang et al., 2019). We compared the newly identified dCas13 proteins in RNA imaging with the previously-reported ones, including dPspCas13b (Yang et al., 2019), dLwaCas13a (Abudayyeh et al., 2017), dRfxCas13d (Wang et al., 2019) and dPguCas13b (Yang et al., 2019). As shown in Fig. 3B, dHgm4Cas13, dMisCas13b and dPba3Cas13b could label both 24 and 16 repeats of GCN4 RNA at the single molecule level, similar as dPspCas13b did (Fig. 3B). In contrast, other dCas13s, including previously-reported

dLwaCas13a (Abudayyeh et al., 2017), dRfxCas13d (Wang et al., 2019) and dPguCas13b (Yang et al., 2019), failed to detect 24 or 16 repeats of GCN4 RNA with the same setting (Fig. S4B).

Extensive analysis with 12, 8 or 4 repeats of GCN4 as target RNAs further revealed that dHgm4Cas13b and dMisCas13b were the two most efficient RNA labeling systems, as shown by their efficient labeling of 12 repeats of GCN4, which were further confirmed by single molecule FISH (smFISH) results (Fig. 3C and Fig. S4C). Of note, in order to avoid the effects of dCas13b binding RNA sequences on smFISH sensitivities (Fig. 2A and D), we designed smFISH probes targeting iRFP670 that was fused to GCN4 repeats rather than the dCas13b-system targeted GCN4 repeats themselves (Fig. 3A). Strikingly, dMisCas13b could even detect 4 repeats of GCN4 (Fig. S4C), despite that not all signals were co-localized with those detected by smFISH (Fig. 3C). Interestingly, dPba3Cas13b effector showed relatively low ability to label expressed GCN4 RNA with 16 repeats in the cytoplasm; however, it could detect GCN4 transcription sites in the nucleus, confirmed by smFISH (Fig. 3C and Fig. S4C). It remains unclear how dPba3Cas13b but not dHgm4Cas13b or dMisCas13b could detect GCN4 transcription sites; but these observations together suggested the potential application of dPba3Cas13b in the study of de novo transcription of endogenous gene loci without the need of genetic manipulation. The requirement of the minimum GCN4 repeats for ectopically-expressed RNA visualization by all examined dCas13 systems has been summarized in Fig. 3D, which was largely consistent with the labeling efficiencies of the endogenous MUC4 RNA (Fig. 2).

Taken together, our screening of previously-uncharacterized Cas13 effectors has identified additional dCas13 systems with robust RNA labeling capability in living cells. Among them, dHgm4Cas13b and dMisCas13b showed the highest efficiency for RNA visualization.

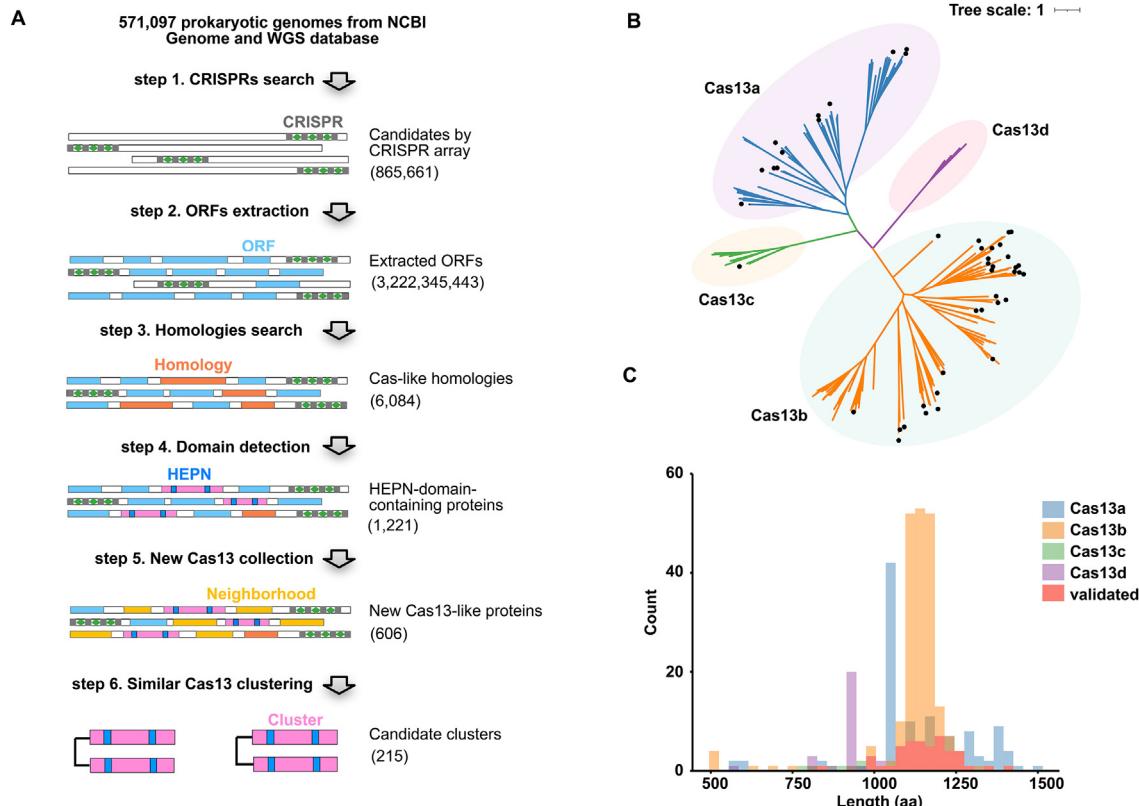


Fig. 1. Identification of previously unreported CRISPR-Cas13 proteins. (A) Bioinformatic pipeline to discover putative previously unreported CRISPR-Cas13 proteins. (B) Maximum-likelihood tree of newly identified CRISPR-Cas13 proteins and previously known Cas13a, Cas13b, Cas13c and Cas13d. Newly identified CRISPR-Cas13 proteins are marked with black dots. Cas13 sequences were aligned using MAFFT 7.487 (Katoh and Standley, 2013) and phylogenetic tree was generated from the PhyML 3.0 (Guindon et al., 2010). (C) Length distribution of newly identified and previously reported Cas13a, Cas13b, Cas13c and Cas13d.

2.3. Characteristics of gRNA sequence requirements for dCas13 labeling

Next, we attempted to inspect how different designs of gRNAs, including lengths and mismatches to targeted RNA sequences, would influence activities of dCas13 effectors in RNA imaging. Since the repeat nature of the endogenous *MUC4* (Fig. 2) and exogenous *GCN4* was not suitable for characterizing the gRNA sequence requirements, we turned to use the nuclear enriched lncRNAs, *NEAT1_2* and *SatIII*, as target RNAs in this analysis. With single gRNAs targeting 3' of *NEAT1_2* and *SatIII* respectively, all three proteins of dHgm4Cas13b-EGFP, dMisCas13b-EGFP and dPba3Cas13b-EGFP could visualize *NEAT1_2* in nuclear paraspeckles in HeLa cells and *SatIII* in nuclear stress bodies in sodium arsenite (SA)-stimulated HeLa cells (Figs. S5A and S5B). RNA smFISH

further confirmed that these dHgm4Cas13b, dMisCas13b, and dPba3-Cas13b labeled signals were indeed from endogenous *NEAT1_2* and *SatIII*, respectively (Figs. S5A and S5B). Of note, smFISH probes recognize the middle region of *NEAT1_2*, which is in the core of paraspeckles, whereas gRNAs of the dCas13b systems targeting the 3' end of *NEAT1_2*, which is in the shell of paraspeckles. These labeling approaches have clearly shown the spherical organization of *NEAT1_2* in paraspeckles (Wang et al., 2018; West et al., 2016; Yang et al., 2019). We then chose the dMisCas13b-EGFP system and the *NEAT1_2* target site for detailed characterization.

To assess the labeling specificity of dMisCas13b, we introduced single (Fig. 4A) or double nucleotide mismatches (Fig. 4B) into the 22-nt gRNA for 3' end of *NEAT1_2* (*gNEAT1_2*) and examined their influence on

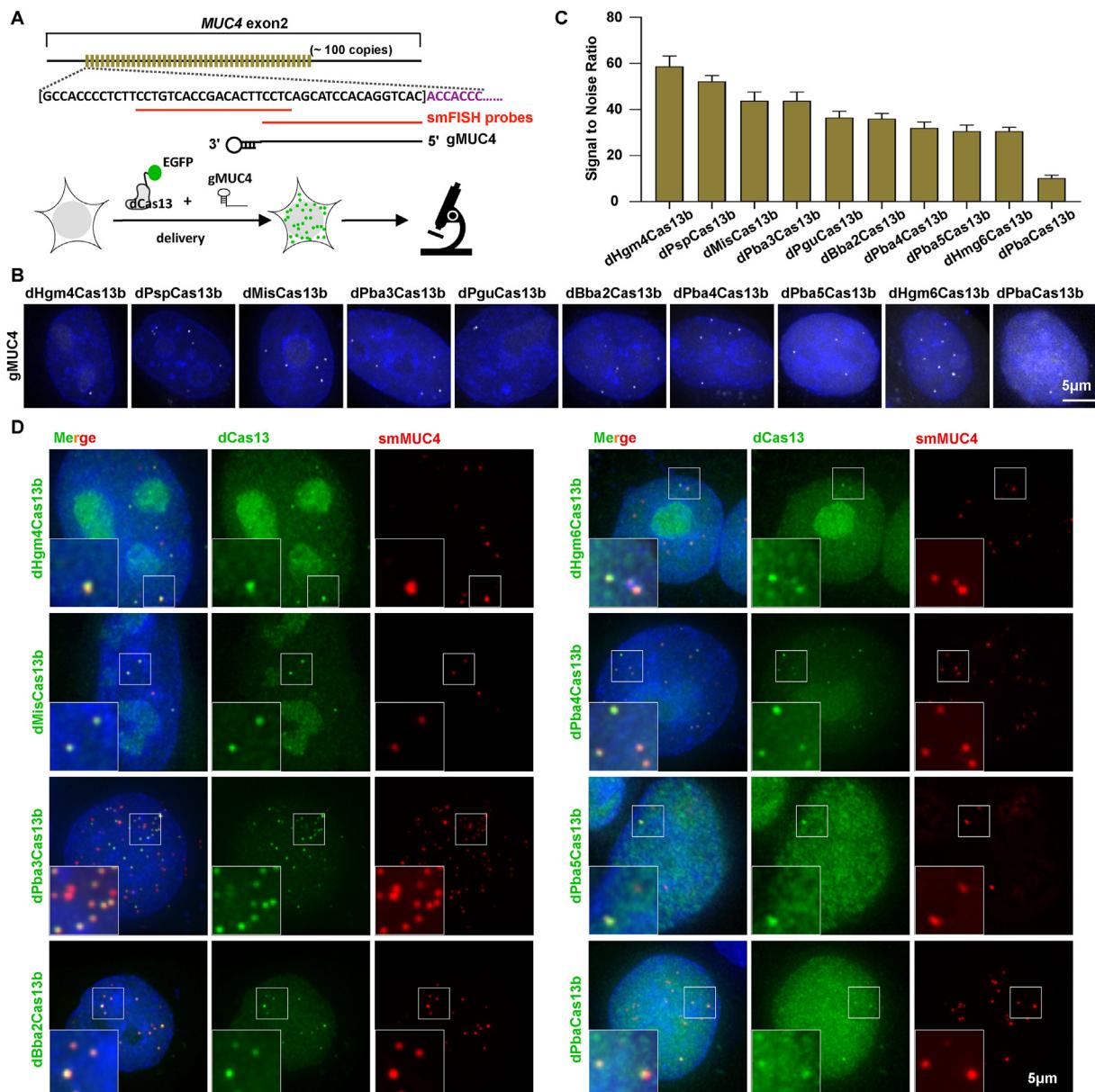
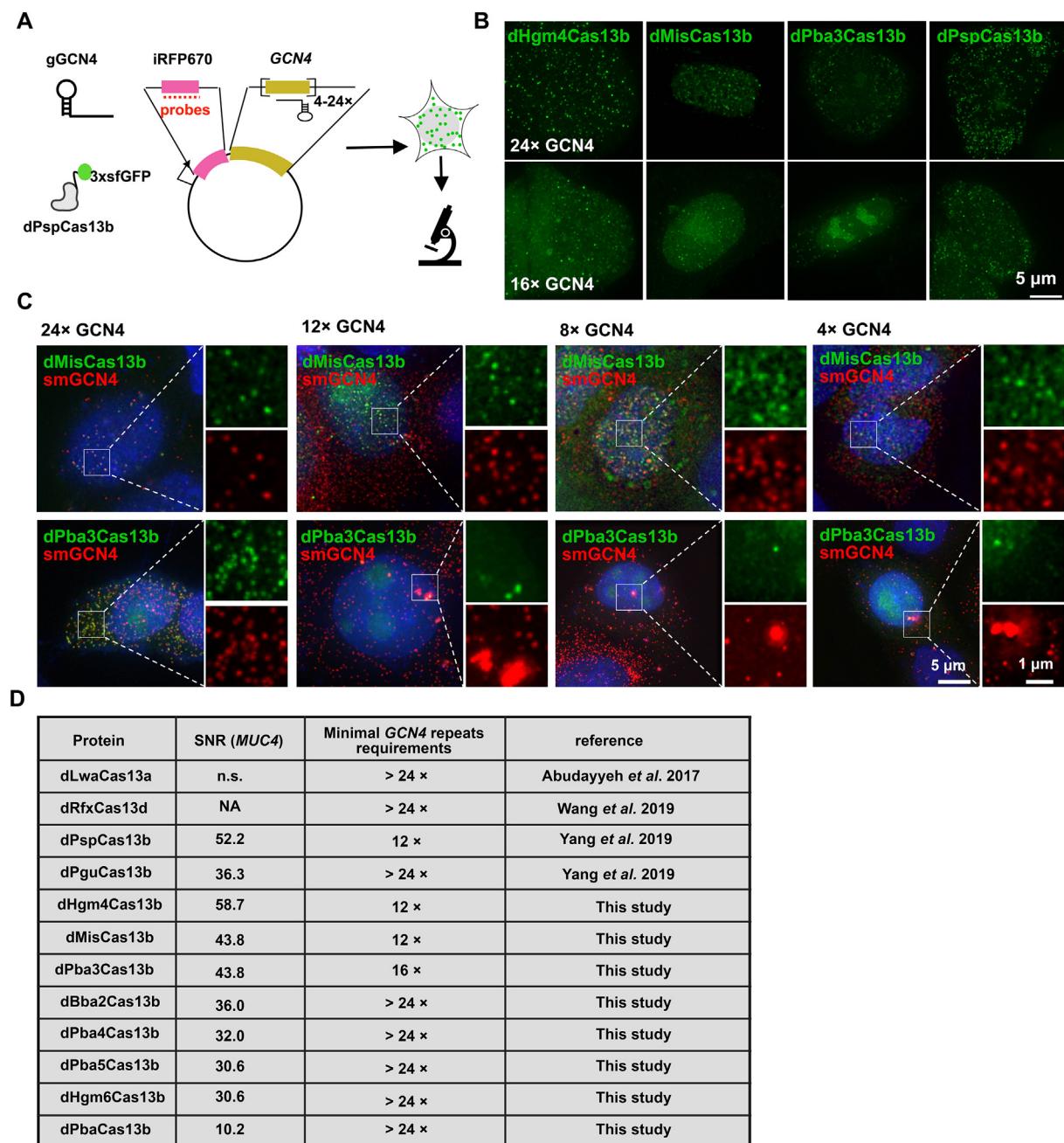


Fig. 2. Screening of new CRISPR-Cas13 proteins for effective RNA labeling. (A) Schematic of effective CRISPR-dCas13 screening for RNA imaging targeting *MUC4* RNAs. dCas13 candidates were fused by EGFP at C-terminal and delivered into HeLa cells with gRNAs targeting *MUC4* by transient transfection. 36–48 h post transfection, cells were fixed and imaged by delta-vision microscopy with a 60/1.42 NA Plan Apo oil-immersion objective. smFISH probes were shown as red lines; the gRNA in the dCas13b systems was shown as a black line with a stem-loop. Of note, they recognize the same sequences in the target RNA. (B) Representative images of *MUC4* RNAs labeled by different CRISPR-dCas13-EGFP proteins in living cells. (C) SNR (Signal to noise ratio) statistics of *MUC4* signals labeled by different CRISPR-dCas13 proteins with detectable signals. $n = 72, 99, 107, 50, 88, 78, 68, 51, 109, 41$ for each examined dCas-EGFP, respectively. Data are represented as mean \pm SEM. (D) Representative images that show the co-localization of dCas13b-EGFP labeled and smFISH-labeled *MUC4* signals. Cells transfected with individual dCas13-EGFP followed by fixation, and smFISH was performed to detect *MUC4* RNAs.

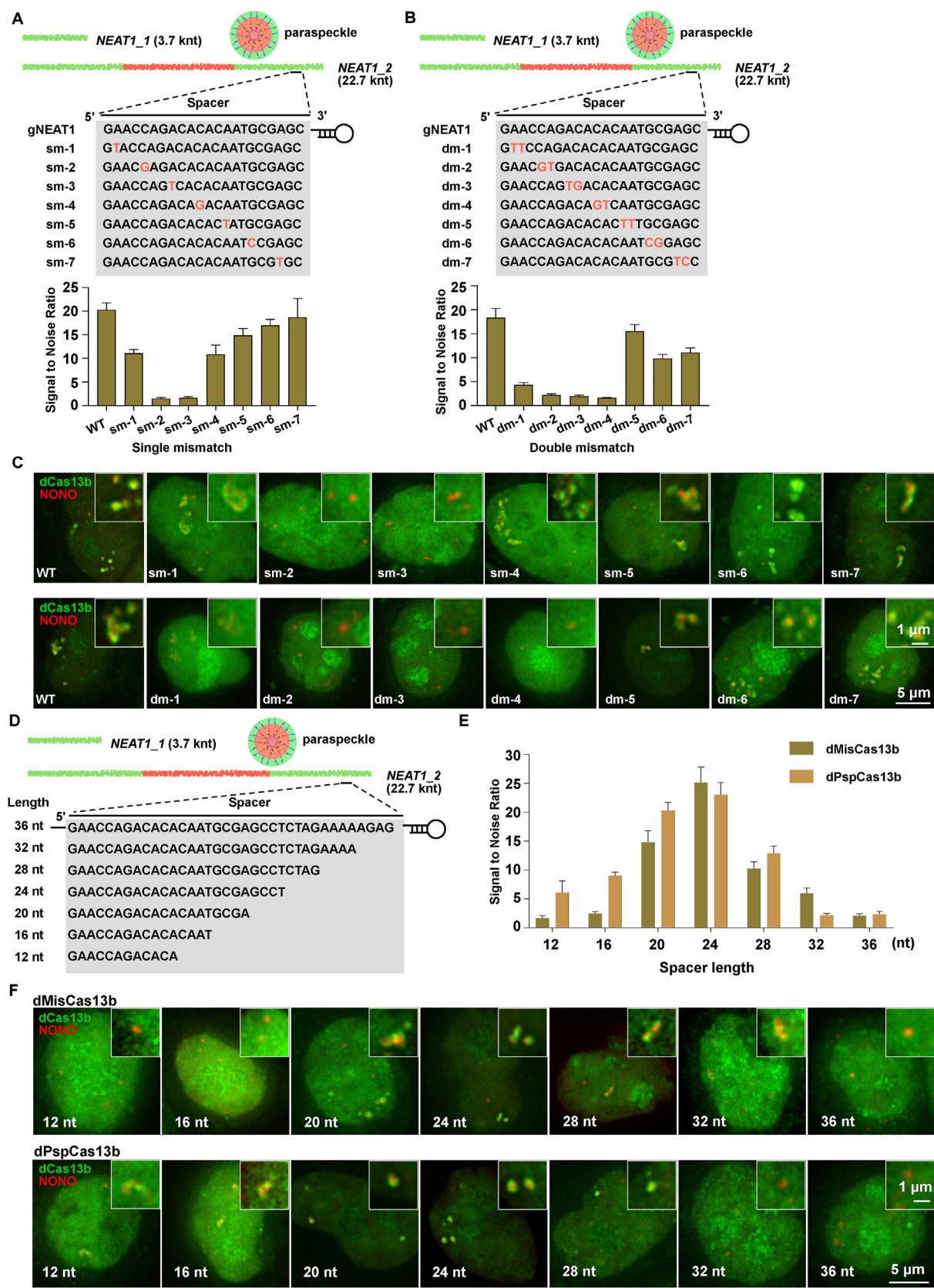


n.s., no significant signal. NA, no statistics

Fig. 3. Repeat-unit requirements for CRISPR-Cas13 proteins to visualize individual RNAs. (A) Schematic of determining the minimum *GCN4* repeats required for dCas13-3 × sfGFP to visualize RNAs with one single gRNA. smFISH probes were shown as red lines; the gRNA in the dCas13b systems was shown as a black line with a stem-loop. Of note, they recognize different sequences in the target RNA to ensure accessibility of target RNAs to probes. (B) Representative images of 24 × and 16 × *GCN4* RNAs, respectively, labeled by newly identified dCas13-3 × sfGFP with individual single gRNA, using the dPspCas13b-3 × sfGFP system as positive controls (Yang et al., 2019). (C) smFISH to confirm *GCN4*-RNA signals labeled by dMisCas13b-3 × sfGFP and dPba3Cas13b-3 × sfGFP (green), respectively. smGCN4, signals of smFISH for *GCN4* repeats, red. Note that eight *GCN4* repeats are the minimum number visualized by dCas13-3 × sfGFP proteins. (D) Summary of the sensitivity of the newly identified and previously reported CRISPR-dCas13 systems for RNA imaging, using the endogenous *MUC4* (SNR, Signal to noise ratio) and the minimum requirement of ectopically expressed *GCN4* repeats as targeted RNAs. Note that the higher SNR and the lower *GCN4* repeats labeled by dCas13-3 × sfGFP suggests the higher sensitivity of the tested dCas13 systems for RNA labeling.

guiding dMisCas13b-EGFP to *NEAT1_2* in the NONO-mRuby3 knocked-in HeLa cell line, in which NONO is a well-established paraspeckle protein localized in the core region and can be used to confirm *NEAT1* signals labeled by dCas13b-EGFP proteins (West et al., 2016; Yang et al., 2019). Consistent to the specificity of dPspCas13b, single nucleotide mismatches from 5' to the middle region were less tolerant than those in the 3' region of gRNAs (Fig. 4A and C). However, single nucleotide mismatches in the

middle region of *gNEAT1_2* abolished labeling signals by dMisCas13b-EGFP (Fig. 4A and C), but only attenuated those by dPspCas13b-EGFP (Yang et al., 2019). The discrepancy between dMisCas13b-EGFP and reported dPspCas13b-EGFP systems (Yang et al., 2019) for *NEAT1_2* labeling was even dramatic with double nucleotide mismatches in *gNEAT1_2*. Specifically, *NEAT1_2* labeling signals by dPspCas13b-EGFP could be completely abolished with double nucleotide



(caption on next page)

mismatches at positions 17 and 18 (Yang et al., 2019), while those by dMisCas13b-EGFP were all damped with double nucleotide mismatches at the 5' to middle region of *gNEAT1_2* (Fig. 4B and C). The low tolerance of gRNA mutation by the MisCas13b effector system thus suggested its high specificity for RNA imaging, and can be further adapted for multi-color RNA imaging below.

Next, we assessed the effect of gRNA spacer length on RNA labeling efficiency, different gRNAs that target the 3' of *NEAT1_2* with spacer lengths from 12 to 36 nucleotides (nt) were individually designed for dMisCas13b and the same gRNAs were also designed to parallelly compare with dPspCas13b (Fig. 4D). As shown in Fig. 4E and F, both dMisCas13b-EGFP and dPspCas13b-EGFP systems could label *NEAT1_2*, with the highest efficiency by co-transfected gRNA containing the spacer length at 24 nt, but extending or shorting the spacer length resulted in reduced labeling efficiencies. Furthermore, dMisCas13b could still label *NEAT1_2* with spacers longer than 28 nt (such as 32 nt with SNR greater than five), but dPspCas13b showed detectable labeling signals with spacers shorter than 20 nt (such as 16 or 12 nt with SNR greater than five), but not vice versa (Fig. 4E and F). Such a different spacer length requirement suggested a potential to apply dMisCas13b-EGFP and dPspCas13b-EGFP at distinct local niches to target RNAs.

2.4. Multi-color RNA imaging by engineering CRISPR-Cas13 systems

Simultaneous visualizing and tracking multiple RNAs is important for understanding motions and interactions between different RNAs in nuclear condensates (Ninomiya et al., 2020, 2021; Prasanth et al., 2005), splicing (Coulon et al., 2014; Hutchinson et al., 2007; Wan et al., 2021) and transcription (Graf et al., 2017; Xiang et al., 2014; Xing et al., 2017). With orthogonal dPspCas13b and dPguCas13b systems, dual-color RNA imaging has been achieved to visualize *NEAT1_2* and *SatII* (Yang et al., 2019). We therefore asked whether the newly identified effective dCas13, such as MisCas13b and Pba3Cas13b, -based systems could be extended to the same use of multiple color RNA imaging. Examination of the orthogonality among currently-identified dCas13b and previously-reported dPspCas13b showed their partial exchangeability with gRNAs that target *MUC4*, *NEAT1_2* or *SatII*. Briefly, dPspCas13b could bind *MUC4* and *SatII* but not *NEAT1_2* with gRNAs individually designed for Pba3Cas13b; it could also bind *NEAT1_2* and *SatII*, but not *MUC4* with gRNAs designed for MisCas13b. Differently, dMisCas13b could use all three gRNAs originally designed for dPspCas13b to target *MUC4*, *NEAT1_2* and *SatII*; but only two gRNAs for Pba3Cas13b could bind *NEAT1_2* and *SatII* rather than *MUC4*. For dPba3Cas13b, it could target *NEAT1_2* using three different gRNAs of dPspCas13b (Fig. S6). Nevertheless, these findings suggested that it is hard to achieve multiple-color RNA imaging by simply using these non-orthogonal

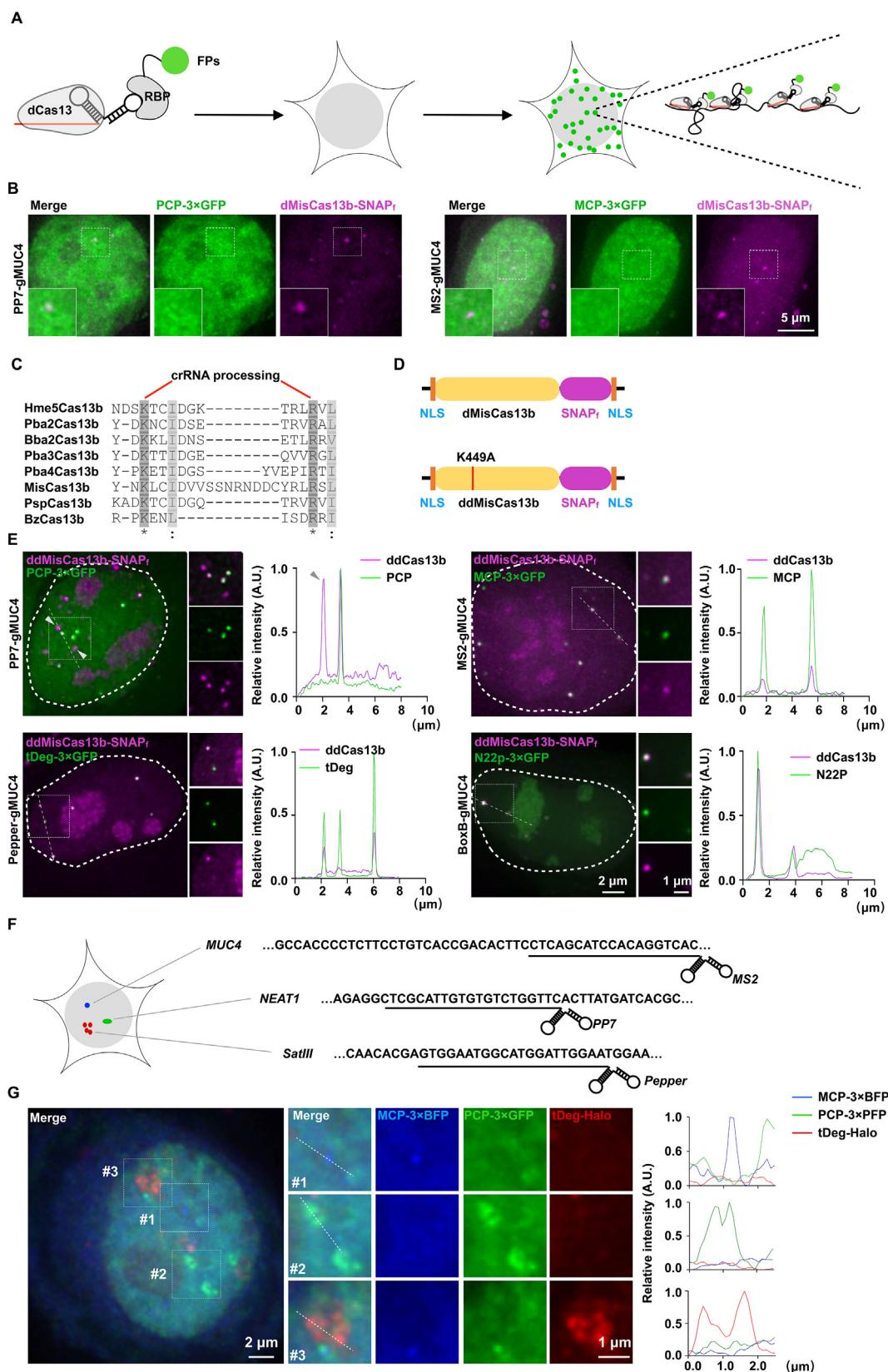
CRISPR-dCas13 systems that fused with a fluorescent protein.

Alternatively, we turned to combine CRISPR-Cas13 and RNA aptamer systems for simultaneous labeling of multiple RNAs. Various types of RNA aptamers have been developed for RNA imaging, such as MS2, BoxB, PP7 and Pepper (Bertrand et al., 1998; Daigle and Ellenberg, 2007; Larson et al., 2011; Wu et al., 2019). Genetic manipulation of tandem RNA aptamers into an RNA of interest followed by visualization of fluorescent aptamer-binding RBPs (Bertrand et al., 1998; Daigle and Ellenberg, 2007; Larson et al., 2011; Wu et al., 2019) or small molecules (Cawte et al., 2020; Chen et al., 2019; Filonov et al., 2014; Paige et al., 2011) enables multiple-color RNA labeling. Since combining CRISPR-dCas9 with RNA aptamers (CRISPRainbow) has successfully visualized multiple genomic loci (Ma et al., 2016; Shao et al., 2016), we thus set to combining CRISPR-dCas13 systems with RNA aptamers for multiple-color RNA imaging, named CRISPRpalette (Fig. 5A).

To design an efficient CRISPRpalette system, distinct RNA aptamer (PP7, MS2, Pepper, BoxB) was individually inserted at the 3' end of gRNA DR (Fig. S7A), in which a 15-nt, or 12-nt linker was required between gRNA DR and MS2 or BoxB to avoid steric hindrance for the binding of RBPs to RNA aptamers. Meanwhile, dMisCas13b protein was fused with a 22 KDa-SNAP_f tag (stained by SNAP-cell 647-SiR) (Fig. 5D), which is derived from O6-alkylguanine-DNA alkyl transferase and used as a fluorescent protein visualized by benzylguanine derivatives (Keppler et al., 2003).

To test the system, plasmids expressing PP7 or MS2 modified gRNAs targeting *MUC4*, 3 × GFP-fused aptamer binding protein and dMisCas13b-SNAP_f were co-transfected into HeLa cells. Interestingly, although *MUC4* signals were labeled by dMisCas13b-SNAP_f, these *MUC4* signals were missing in the PCP-3 × GFP or MCP-3 × GFP channel (Fig. 5B). This might be caused by the cleavage of gRNAs at the 3' of DR by dMisCas13b-SNAP_f, which released the RNA aptamer from modified gRNAs and thus impaired the *MUC4* labeling by the aptamer-modified gRNAs and PCP-3 × GFP or MCP-3 × GFP. To overcome this obstacle, we mutated the active site of dMisCas13b for pre-crRNA processing. Alignment of MisCas13b with BzCas13b, whose active sites for pre-crRNA processing were previously reported (Zhang et al., 2018), showed that the Lysine at position 449 and Arginine at position 467 were responsible for the processing of pre-crRNA arrays by MisCas13b (Fig. 5C). We thus mutated Lysine 449 to Alanine in dMisCas13b, generating the ddMisCas13b effector without its pre-crRNA processing activity (Fig. 5D). Correspondingly, when we co-transfected three plasmids individually expressing PP7 or MS2 modified gRNA that targets *MUC4*, 3 × GFP-fused aptamer binding protein and ddMisCas13b-SNAP_f, we could successfully label *MUC4* RNAs by PCP-3 × GFP or MCP-3 × GFP, largely colocalizing with ddMisCas13b-SNAP_f signals (Fig. 5E). In addition, another two aptamers, Pepper and BoxB,

Fig. 4. Characteristics of dMisCas13b and dPspCas13b in labeling RNAs. (A) Schematic of the single nucleotide gRNA-RNA mismatch within the *NEAT1_2* targeting site (top) and SNR (Signal to noise ratio) statistics (bottom) showing effects of the corresponding single nucleotide mismatch on dMisCas13b targeting efficiency. n = 19, 6, 25, 26, 12, 21, 10, 9 (cells). *NEAT1_1*, and the 5' and 3' regions of *NEAT1_2* were localized to the shell of paraspeckles (green) and the middle region of *NEAT1_2* was localized to the core of paraspeckles (red). Note that single nucleotide mismatches at the positions 5 and 8 blocked *NEAT1_2* labeling of dMisCas13b-EGFP. Data are represented as mean ± SEM. (B) Schematic of the double nucleotides gRNA-RNA mismatch position at the *NEAT1_2* targeting site (top) and SNR statistics (bottom) showing effects of the corresponding double nucleotides mismatch on dMisCas13b targeting efficiency. n = 24, 19, 21, 12, 11, 19, 22, 23 (cells). Note that double mismatches at the 5' end blocked *NEAT1_2* labeling of dMisCas13b-EGFP. See (A) in detail. (C) Representative images of CRISPR-dMisCas13b labeled *NEAT1_2* with gRNAs bearing different single nucleotide mismatches (upper) shown in (A) and double nucleotides mismatches (bottom) shown in (B) in the NONO-mRuby3 knocked-in HeLa cell line. Signals of dMisCas13b-EGFP were displayed in green, and signals of NONO-mRuby3 were shown in red. Of note, NONO is localized in the core region of paraspeckles (West et al., 2016) and dMisCas13b-labeled 3' of *NEAT1_2* is in the shell of paraspeckles. (D) Schematic of the spacer lengths from 12 nt to 36 nt of gRNAs targeting the 3' terminus of the long isoform of *NEAT1* (*NEAT1_2*). See (A) in detail. (E) SNR statistics of *NEAT1_2* signals labeled by CRISPR-dMisCas13b-EGFP and CRISPR-dPspCas13b, using single gRNA with different length. n = 7, 14, 13, 14, 12, 7, 9, 6 (cells) for dMisCas13b-EGFP (brownish green column) and n = 14, 5, 16, 23, 12, 9, 11, 7 (cells) for dPspCas13b-EGFP (brownish yellow column). Both dMisCas13b-EGFP and dPspCas13b-EGFP labeled *NEAT1_2* best with gRNAs of 24-nt spacer. It is interesting to note that dMisCas13b-EGFP displayed a higher efficiency than dPspCas13b-EGFP with a longer spacer, while dPspCas13b-EGFP displayed a higher efficiency dMisCas13b-EGFP with a shorter spacer for *NEAT1_2* imaging. Data are represented as mean ± SEM. (F) Representative images of CRISPR-dMisCas13b-EGFP (upper) and CRISPR-dPspCas13b labeled *NEAT1_2* with different lengths of gRNAs shown in (E) in the NONO-mRuby3 knocked-in HeLa cell line (Yang et al., 2019). Signals of dCas13b-EGFP were represented as green, and signals of NONO-mRuby3 were represented as red. Of note, NONO is in the core region of paraspeckles and dMisCas13b-labeled 3' of *NEAT1_2* is in the shell region of paraspeckles (West et al., 2016).



(caption on next page)

conjugated gRNAs also achieved successful *MUC4* RNA labeling (Fig. 5E). Importantly, using *NEAT1_2*, *SatIII* and 24 × *GCN4* as targets, together with the PP7 conjugated gRNAs, all these RNAs could be imaged by the CRISPRpalette system with ddMisCas13b-SNAP_f (Fig. S7B). It should be noted that signals labeled by the aptamer system and CRISPR-ddMisCas13b system did not always co-localized. For example, ddMisCas13b-SNAP_f produced some signals that did not co-localize with PCP-3 × GFP (Fig. 5E, the left upper image), likely due to the incomplete assembly of CRISPRpalette that may allow the free ddMisCas13b-SNAP_f and aptamer binding proteins to be engaged with non-specific RNAs. Further optimization of CRISPRpalette is warranted.

Nevertheless, to achieve triple-color labeling of RNAs by the CRISPRpalette system, we constructed plasmids expressing MS2-gRNA targeting *MUC4*, PP7-gRNA targeting *NEAT1_2* and Pepper-gRNA targeting *SatIII* (Fig. 5F) and delivered them with ddMisCas13b-SNAP_f, MCP-3 × BFP, PCP-3 × GFP, tDeg-Halo (stained by Janelia Fluor 549) into HeLa cells. We detected *MUC4* by MCP-BFP, and *NEAT1_2* by PCP-GFP but barely *SatIII* signals because *SatIII* RNAs were rarely expressed without stress in normal cells (Fig. S7C), showing the specificity of this CRISPRpalette system. Upon 6 h of SA treatment, *SatIII* signals labeled by tDeg-Halo appeared in *NEAT1_2* and *MUC4* double positive cells, in which each signal did not colocalize with others, further confirming the labeling specificity of the CRISPRpalette system for multiple RNAs at the same time (Fig. 5G and S7D).

3. Discussion

The CRISPR-dCas13 system displays its advantages in tracking RNAs in living cells. Among CRISPR-Cas13 subfamilies, dLwaCas13a, dPspCas13b, dPguCas13b, and dRfxCas13d are capable of imaging RNAs in living cells (Abudayyeh et al., 2017; Wang et al., 2019; Yang et al., 2019), with dPspCas13b shows the most robust activity on endogenous RNAs (Yang et al., 2019). To explore additional CRISPR-dCas13 systems with a robust labeling capacity, we screened *de novo* CRISPR-Cas13 systems (Fig. 1) and identified eight previously undescribed dCas13b proteins that are capable of labeling *MUC4* repeats in cells (Fig. 2). Among these dCas13b proteins, dHgm4Cas13b, dMisCas13b, and dPba3Cas13b showed comparable, if not higher, robustness of RNA signals than dPspCas13b (Fig. 3D). Importantly, as it has recently been shown that the high RNA labeling activity of dPspCas13b is at least in part owing to the high binding affinity to RNAs (Li et al., 2020; Tian et al., 2022), our newly identified dCas13b proteins with robust RNA labeling capacity (Figs. 2 and 3) could also be applied in other aspects of RNA research, such as RNA editing (Abudayyeh et al., 2019; Cox et al., 2017; Huang et al., 2020), RNA precipitations (Abudayyeh et al., 2017; Han et al., 2020; Yang et al., 2019; Zhang et al., 2020), and interfering with RBP binding (Konermann et al., 2018; Yao et al., 2019).

Fig. 5. Multiple-color RNA imaging by CRISPRpalette that combines double mutated CRISPR-dCas13 and RNA aptamer. (A) Overview of RNA labeling in living cells using the CRISPRpalette system. In brief, gRNAs were modified by RNA aptamers at their 3' ends. Plasmids encoding modified gRNAs, dCas13 proteins and fluorescent aptamer-binding proteins were delivered into living HeLa cells. (B) Representative images of *MUC4* RNAs labeled by CRISPR-dMisCas13b-SNAP_f using PP7 (PCP-3 × GFP), MS2 (MCP-3 × GFP) modified gRNAs. Green, RBP-3 × GFPs labeled signals. Purple, ddMisCas13b-SNAP_f labeled signals. SNAP_f was stained with SNAP-Cell 647-SiR. Note that *MUC4* can only be labeled by dMisCas13b-SNAP_f, but not by PCP-3 × GFP or MCP-3 × GFP. (C) Alignment of Cas13b orthologs showing the active site for pre-crRNA processing. Note that pre-crRNA processing sites Lysine (K) 449 and Argine (R) 467 of MisCas13b are conserved among Cas13b proteins. (D) Schematic of MisCas13b mutations fused with SNAP_f-tag for RNA imaging of the CRISPRpalette system. To obtain ddMisCas13b lacking pre-crRNA processing activity, K449 was mutated to Alanine on the dMisCas13b, which was further fused with SNAP_f-tag to indicate signals labeled by ddMisCas13b. (E) Representative images of *MUC4* RNAs labeled by CRISPR-ddMisCas13b-SNAP_f using PP7 (PCP-3 × GFP), MS2 (MCP-3 × GFP), Pepper (tDeg-3 × GFP) and BoxB (N22p-3 × GFP) modified gRNAs and line scans of fluorescent intensities of signals were indicated at the right. Green, RBP-3 × GFPs labeled signals. Purple, ddMisCas13b-SNAP_f labeled signals. SNAP_f was stained with SNAP-Cell 647-SiR. Note that *MUC4* can be both labeled by dMisCas13b-SNAP_f and different RBPs fused with 3 × GFPs. Nuclei are shown by white dotted lines. Non-colocalized signals in the upper left image (PCP-3 × GFP) are indicated by gray arrows. Non-colocalized signals in the upper left image (PCP-3 × GFP) are indicated by gray arrows. (F) Schematic of labeling *SatIII*, *NEAT1_2*, and *MUC4* in triple colors in living cells using CRISPRpalette. The single gRNA targeting *SatIII* was modified with *Pepper*; the single gRNA targeting *NEAT1_2* was modified with *PP7*, and the single gRNA targeting *MUC4* was modified with *MS2*. Of note, HeLa cells were treated with Sodium Arsenate (SA) for 6 h (h) to allow *SatIII* expression prior to imaging. (G) A representative image of *MUC4*, *NEAT1_2*, and *SatIII* RNAs labeled simultaneously by the CRISPRpalette system with gRNAs shown in (F) 6 h post SA treatment. Blue, MCP-3 × BFP labeled signals. Green, PCP-3 × GFP labeled signals. Red, tDeg-Halo labeled signals. Enlarged images in squares with the relative fluorescence intensities of the dotted lines are shown.

Additional characteristics of the gRNA spacer length and nucleotide mismatches have shown that the dMisCas13b system shows similar but distinct property when compared to the well-established dPspCas13b in RNA labeling (Yang et al., 2019) (Fig. 4). These features resulted in the non-orthogonal gRNA recognition of these two Cas13b systems or among other identified effective dCas13 proteins, such as dPba3Cas13b (Fig. S6). To overcome the non-orthogonal difficulty of newly identified dCas13b in multiple-RNA tracking, we designed RNA aptamers conjugated to 3' ends of gRNAs (Fig. 5A and Fig. S7A Fig. S7), and impaired the pre-crRNA processing activity of dMisCas13b to achieve the robust and multiple RNA labeling at the same time (Fig. 5C and D), named CRISPRpalette. This CRISPRpalette system has enabled the aptamer-conjugated gRNAs to be recognized by fluorescent RBPs (Fig. 5E–G and S7B–S7D), thereby extending multi-color RNA imaging in living cells. Of note, it has been challenging to label multiple RNAs at the same time. We have achieved simultaneous labeling of RNAs in three colors in living cells (Fig. 5G and Fig. S7D), thereby providing a convenient system for multiple-color RNA labeling without genetic manipulation. More RNA targets could in principle be visualized at the same time with additional RNA aptamers, as for multi-color genome labeling by CRISPRRainbow (Ma et al., 2016).

In addition, we noticed a tendency of most dCas13 proteins to concentrate in the nucleolus (for example, Fig. S3B), and these concentrated signals could become more obvious when designed gRNAs were less effective (Fig. 4C and F). Similar observations have been also reported in DNA imaging by CRISPR-dCas9 systems (Chen et al., 2013, 2016). As the site for ribosome biogenesis, the nucleolus is enriched with rRNAs, which could be nonspecific targets for RBPs, thereby trapping dCas13 proteins into the nucleolus, as shown in previous studies (Frottin et al., 2019; Riback et al., 2020). This indeed could be largely overcome by using CRISPRpalette, in which fluorescent RNA aptamer binding proteins, instead of dCas13b proteins, were used as the visualization signals, thus avoiding the detection of unwanted nucleolar signals (Fig. 5E–G and S7B–S7D).

Together, the identification of eight previously undescribed dCas13b proteins for RNA labeling and the combination of RNA-aptamer-conjugated gRNAs has broadened the utility of CRISPR-dCas13 systems in imaging RNAs.

4. Materials and methods

4.1. Computational pipeline for Cas13 discovery

All bacterial and archaeal genomes were downloaded from NCBI Genome and WGS databases in November 2021. PILER-CR (Edgar, 2007) was used to identify CRISPR repeats. ORFs(>500aa) within 10 kb next to identified CRISPR arrays were detected by ORFfinder(-g 11). Multiple

sequence alignments of reported known Cas13 were used by PSI-BLAST (-num_iterations 3 -evalue 0.1)(Altschul et al., 1997) to search against all extracted ORFs for Cas13 homologies identification. The HEPN domains in candidate Cas13 homologues were detected by HHpred (default)(-Soding, 2005; Soding et al., 2005) and CD-search (-evalue 10)(Marchler-Bauer et al., 2017). And two HEPN motifs were retrieved by regular expression matching in HEPN-domain-containing ORFs. To determine the ancillaries around Cas13, 10 ORFs in proximity to Cas13-like ORFs were extracted. All reported Cas and ancillary proteins were subjected to PSI-BLAST(-num_iterations 3 -evalue 1e-5) (Altschul et al., 1997) to search against the Neighborhoods. In order to obtain new Cas13-like proteins, the list of candidates was narrowed down by deleting loci containing neighboring Cas proteins belonging to other type CRISPR-Cas systems or similar to reported Cas13 (>80% identity and coverage). The highly similar sequences were clustered by MMseq2 (Steinegger and Soding, 2017) with parameters -min-seq-id 0.8 -c 0.8, resulting in final candidate Cas13 clusters.

4.2. Reagents and cells

The following reagents were purchased from the indicated companies. Lipofectamine 3000 (Thermo); Ribonucleoside Vanadyl Complex (NEB); VECTASHIELD Antifade Mounting Medium (Vector Lab); Cyanine 3-dUTP (Enzo life); DAPI (Invitrogen); Triton X-100 (ABCONE); Agarose (ABCONE); DPBS (Gibco); Paraformaldehyde (Sigma-Aldrich); Sodium arsenate (Sigma-Aldrich); Bovine Serum Albumin (ABCONE); Hieff Clone™ One Step Cloning Kit (Yeasen), and StarPrep Gel Extraction Kit StarPrep (Genstar); DMEM (Gibco); Opti-MEM (Gibco), Fetal Bovine Serum (FBS) (Gibco), SNAP-Cell 647-SiR (NEB), dye Halotag-549 was a gift of Hanhui Ma lab.

HeLa cells were purchased from the American Type Culture Collection (ATCC; <http://www.atcc.org>).

4.3. Plasmid construction

pHAGE-PCP-3 × GFP, pHAGE-MCP-3 × GFP, pHAGE-N22p-3 × GFP and pHAGE-MCP-3 × BFP were gifted from the Dr. Hanhui Ma's lab. To construct pHAGE-dCas13-EGFP-2 × NLS-IRES-puro plasmids, all newly identified Cas13 proteins were mammalian codon optimized and synthesized by Tsingke Biotechnology and Sangon Biotechnology. Both two HEPN motifs of each Cas13 proteins were mutated and nuclease dead (d) Cas13 proteins were inserted into pHAGE-dPspCas13b-EGFP-2 × NLS-IRES-puro by replacing the dPspCas13b sequence (Yang et al., 2019). In detail, pHAGE-dPspCas13b-EGFP-2 × NLS-IRES-puro was digested by NheI and XbaI to delete dPspCas13b and the linearized sequence was fused with other dCas13 sequences using the one-step clone method (Yeasen). To construct pHAGE-dCas13-3 × sfGFP-3 × NLS-IRES-puro, the sequence of fluorescent proteins 3 × sfGFP with NLS were cloned into pHAGE-dCas13-EGFP-2 × NLS-IRES-puro by replacing EGFP using one-step clone method (Yeasen). To construct pHAGE-dMisCas13b-SNAP_f-2 × NLS-IRES-puro, the sequence of SNAP_f was cloned into pHAGE-dCas13-EGFP-2 × NLS-IRES-puro by replacing EGFP using one-step clone method (Yeasen). To construct pHAGE-ddMisCas13b-SNAP_f-2 × NLS-IRES-puro, dMisCas13b was further mutated at K449 and cloned into pHAGE-dMisCas13b-SNAP_f-2 × NLS-IRES-puro by replacing dMisCas13b using one-step clone method (Yeasen).

To construct pHAGE-tDeg-Halo, the sequence of Halo-tag was cloned into pHAGE-PCP-3 × GFP to replace PCP-3 × GFP and then sequence of tDeg was cloned into C- terminus using the one-step clone method (Yeasen). To construct pHAGE-tDeg-3 × GFP, 3 × GFP was cut from pHAGE-N22p-3 × GFP and inserted into pHAGE-tDeg-Halo to replace Halo-tag using restriction enzyme BamHI and XhoI.

To construct plasmids expressing mRNA with different repeated sequences, 4 × GCN4, 8 × GCN4, 12 × GCN4, 16 × GCN4 and 24 × GCN4 were inserted into pmiRFP670-c1 plasmid using restriction enzyme

HindIII and BamHI.

The sequences of backbone plasmids for expressing gRNAs are the same as previously reported (Abudayeh et al., 2017; Cox et al., 2017; Konermann et al., 2018). All gRNA sequences were listed in Supplemental Table S1.

Associated plasmids are deposited to Addgene (<https://www.addgene.org>) and BRICS (<http://www.bricks.ac.cn/plasmid/?columnId=15>).

4.4. Plasmid transfection for live cell imaging and native RNA immunoprecipitation

To transfet plasmids for live cell imaging and native RNA immunoprecipitation, Lipofectamine™ 3000 (Invitrogen) transfection was performed with the general ratio of 3 reagents as 1 µg plasmids, 2 µL P3000 and 2 µL lipo3000 in each 12-well plate at 70% cell confluence. After 16–18 h transfection, cells were passage into 35 mm no. 1.5 glass-bottomed dishes (Cellvis) for live cell imaging or passage followed by native RNA immunoprecipitation assays. For each transfection, usage of each plasmid was listed as below:

Assay	Plasmid usage	Related to Figures
Labeling single RNAs by dCas13s	0.2 µg dCas13-EGFP + 0.8 µg gMUC4	Fig. 2B-D , Fig. S3A , Fig. 4
Repeats labeling test and optimizations of fluorescent proteins for cytoplasmic RNA labeling	0.2 µg dCas13-EGFP + 0.6 µg gRNA + 0.2 µg mRuby3-N × GCN4 (N = 4,8,12,16,24)	Fig. 3B and C , S4B, S4C
CRISPRpalette system to label single RNAs	0.2 µg dCas13-SNAP _f + 0.6 µg gRNA + 0.2 µg RBP-FPs	Fig. 5 B, 5E , S7B
Multi-color labeling by CRISPRpalette system	0.2 µg ddMisCas13b-SNAP _f + 0.6 µg PP7-gNEAT1 + 0.6 µg MS2-gMUC4 + 0.6 µg Pepper-gSatIII + 0.2 µg PCP-3 × GFP + 0.2 µg MCP-3 × BFP + 0.2 µg tDeg-Halo	Fig. 5G and S5C

4.5. Live cell imaging procedure

Transfected cells (transfected or not) were cultured on 35 mm no. 1.5 glass-bottomed dishes (Cellvis). Cells were washed once with PBS and the medium was replaced by FluoroBrite DMEM (Gibco) supplemented with 10% FBS and placed back in the incubator for 1 h followed by Widefield microscopy imaging. For cells transfected with SNAP_ftag or Halotag-fused proteins, the medium was replaced with medium containing 30 nM SNAP-Cell 647-SiR (NEB) or 6 nM Halotag-549 one half an hour before imaging.

4.6. Widefield microscopy procedure

All widefield microscopy images were performed on a DeltaVision Elite imaging system (GE Healthcare) equipped with a 60 × /1.42 NA Plan Apo oil-immersion objective, or a 100 × /1.40 NA Plan Apo oil-immersion objective (Olympus), as well as the CoolSnap HQ2 camera (Photometrics) equipped with the live cell imaging environment control system (Live Cell Instrument). Raw data of all presented figures were deconvolved by softWoRx 6.5 using the enhanced ratio method.

4.7. Single molecule RNA fluorescent in situ hybridization (smFISH) and proteins visualization

All smFISH probes were designed via Stellaris Probe Designer and labeled with Cy3 or Cy5 on the 3' ends (Supplemental Table 1). RNA FISH was carried out as described before (Raj and Tyagi, 2010). Briefly, cells were fixed with 4% PFA for 15 min and washed with DPBS for 3 × 5 min, followed by permeabilization with 0.5% Triton X-100 for 5 min and

washed with DPBS for 3×5 min. Cells were incubated in 10% formamide/2 × SSC for 10 min at room temperature followed by hybridization at 37 °C for 16 h. After hybridization, samples were mounted in VECTASHIELD antifade mounting medium (Vector Lab).

4.8. Signal-to-noise ratio (SNR) and line scan analysis

SNR was defined as the ratio of the intensity of a fluorescent signal and the power of background noise. Steps of SNR calculation were described previously (Yang et al., 2019). To obtain SNR for *NEAT1*, spot colocalized with NONO was selected first, then a circle with diameter of 3–4 μm and with center of the spot (exclude the spot that are colocalized with NONO) was selected as background followed by calculating with following formula: $\text{SNR} = P_{\text{signal}}/P_{\text{background}} = (\text{Max intensity of spots signal} - \text{Mean intensity of background GFP spot})/\text{Std. dev. of the background signal}$.

In line scan analysis, a line was drawn by ImageJ's "line tool" through the region of interest (ROI) and pixel intensity along the line was plotted using the "Plot Profile" tool. Pixel intensity data were then normalized to the maximum in the data set and then used to generate curves of relative intensity using GraphPad Prism 8. In Fig. 5E, 8-μm line was used and in Fig. 5G, 2.5-μm line was used.

All the fluorescence imaging data were analyzed by Fiji/Image J.

4.9. Statistical analysis

Data were analyzed by GraphPad Prism 8. Error bars represented as standard error of the mean (SEM) from data in at least 5 cells and these have been stated in the corresponding legends. See methods and figure legends for details.

Author contributions

L.-L.C. and L.-Z. Y conceived the project. L.-Z. Y designed and performed experiments with the help of Y. H; B.-Q. G. and Y. W performed computational analyses, supervised by L.Y. L.-Z. Y., L.Y. and L.-L.C. wrote the manuscript with the input from all authors. L.-L.C. supervised the project.

Data availability

Raw data of key experiments can be accessed on Mendeley Data: https://data.mendeley.com/datasets/hpp7wvvr4k/draft?_a=25764d35-717d-4901-9b03-29988a2830f5.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

L.-Z.Y., B. G, Y. H, S. L, Y. W, L. Y., Y.-F. X., and L.-L.C. are named as inventors on patents related to these CRISPR-Cas13 enzymes held by CAS Center for Excellence in Molecular Cell Science.

Acknowledgements

We thank Chuyan Qin and Yichun Xiong for help in experiments and in analyses, and other members in Chen and Yang laboratories for discussion. This work was supported by the CAS Project for Young Scientists in Basic Research, China (YSBR-009), National key research and development program, China (2021YFA1100203), the Shanghai Municipal Commission for Science and Technology (20JC1410300), the National Natural Science Foundation of China (NSFC) (31821004 and 31725009), the CAS Center for Excellence in Molecular Cell Science (CEMCS) (2020DF03), and the HHMI International Program (55008728) to L.-L.C.; L.-L.C. acknowledges support from the Xplorer Prize.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.cellin.2022.100044>.

References

- Abudayyeh, O. O., Gootenberg, J. S., Essletzbichler, P., Han, S., Joung, J., Belanto, J. J., Verdine, V., Cox, D. B. T., Kellner, M. J., Regev, A., et al. (2017). RNA targeting with CRISPR-Cas13. *Nature*, *550*, 280. +-.
- Abudayyeh, O. O., Gootenberg, J. S., Franklin, B., Koob, J., Kellner, M. J., Ladha, A., Joung, J., Kirchgatterer, P., Cox, D. B. T., & Zhang, F. (2019). A cytosine deaminase for programmable single-base RNA editing. *Science*, *365*, 382.
- Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J. H., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, *25*, 3389–3402.
- Batra, R., Nelles, D. A., Pirie, E., Blue, S. M., Marina, R. J., Wang, H., Chaim, I. A., Thomas, J. D., Zhang, N., Nguyen, V., et al. (2017). Elimination of toxic microsatellite repeat expansion RNA by RNA-targeting Cas9. *Cell*, *170*, 899. +-.
- Bertrand, E., Chartrand, P., Schaefer, M., Shenoy, S. M., Singer, R. H., & Long, R. M. (1998). Localization of ASH1 mRNAs in living yeast. *Mol. Cell*, *2*, 437–445.
- Cawte, A. D., Unrau, P. J., & Rueda, D. S. (2020). Live cell imaging of single RNA molecules with fluorogenic Mango II arrays. *Nat. Commun.*, *11*.
- Chen, B., Gilbert, L. A., Cimini, B. A., Schnitzbauer, J., Zhang, W., Li, G. W., Park, J., Blackburn, E. H., Weissman, J. S., Qi, L. S., et al. (2013). Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell*, *155*, 1479–1491.
- Chen, B., Hu, J., Almeida, R., Liu, H., Balakrishnan, S., Covill-Cooke, C., Lim, W. A., & Huang, B. (2016). Expanding the CRISPR imaging toolset with *Staphylococcus aureus* Cas9 for simultaneous imaging of multiple genomic loci. *Nucleic Acids Res.*, *44*, e75.
- Chen, J., Nikolaitchik, O., Singh, J., Wright, A., Bencsics, C. E., Coffin, J. M., Ni, N., Lockett, S., Pathak, V. K., & Hu, W. S. (2009). High efficiency of HIV-1 genomic RNA packaging and heterozygote formation revealed by single virion analysis. *Proc. Natl. Acad. Sci. U. S. A.*, *106*, 13535–13540.
- Chen, M. M., Ma, Z., Wu, X. T., Mao, S. Q., Yang, Y. T., Tan, J., Krueger, C. J., & Chen, A. K. (2017). A molecular beacon-based approach for live-cell imaging of RNA transcripts with minimal target engineering at the single-molecule level. *Sci. Rep.-UK*, *7*.
- Chen, X. J., Zhang, D. S., Su, N., Bao, B. K., Xie, X., Zuo, F. T., Yang, L. P., Wang, H., Jiang, L., Lin, Q. N., et al. (2019). Visualizing RNA dynamics in live cells with bright and stable fluorescent RNAs. *Nat. Biotechnol.*, *37*, 1287.
- Cong, L., Ran, F. A., Cox, D., Lin, S. L., Barretto, R., Habib, N., Hsu, P. D., Wu, X. B., Jiang, W. Y., Marraffini, L. A., et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science*, *339*, 819–823.
- Coulon, A., Ferguson, M. L., de Turris, V., Palangat, M., Chow, C. C., & Larson, D. R. (2014). Kinetic competition during the transcription cycle results in stochastic RNA processing. *Elife*, *3*.
- Cox, D. B. T., Gootenberg, J. S., Abudayyeh, O. O., Franklin, B., Kellner, M. J., Joung, J., & Zhang, F. (2017). RNA editing with CRISPR-Cas13. *Science*, *358*, 1019–1027.
- Daigle, N., & Ellenberg, J. (2007). LambdaN-GFP: an RNA reporter system for live-cell imaging. *Nat. Methods*, *4*, 633–636.
- Edgar, R. C. (2007). PILER-CR: fast and accurate identification of CRISPR repeats. *BMC Bioinf.*, *8*.
- Filonov, G. S., Moon, J. D., Svensen, N., & Jaffrey, S. R. (2014). Broccoli: rapid selection of an RNA mimic of green fluorescent protein by fluorescence-based selection and directed evolution. *J. Am. Chem. Soc.*, *136*, 16299–16308.
- Frottin, F., Schueder, F., Tiwary, S., Gupta, R., Korner, R., Schlüchthaerle, T., Cox, J., Jungmann, R., Hartl, F. U., & Hipp, M. S. (2019). The nucleolus functions as a phase-separated protein quality control compartment. *Science*, *365*, 342.
- Graf, M., Bonetti, D., Lockhart, A., Serhal, K., Kellner, V., Maicher, A., Jolivet, P., Teixeira, M. T., & Luke, B. (2017). Telomere length determines TERRA and R-loop regulation through the cell cycle. *Cell*, *170*, 72.
- Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.*, *59*, 307–321.
- Han, S., Zhao, B. S., Myers, S. A., Carr, S. A., He, C., & Ting, A. Y. (2020). RNA-protein interaction mapping via MS2-or Cas13-based APEX targeting. *P. Natl. Acad. Sci. USA*, *117*, 22068–22079.
- Huang, X. X., Lv, J. J., Li, Y. Q., Mao, S. S., Li, Z. F., Jing, Z. Y., Sun, Y. D., Zhang, X. M., Shen, S. X., Wang, X. X., et al. (2020). Programmable C-to-U RNA editing using the human APOBEC3A deaminase. *EMBO J.*, *39*.
- Hutchinson, J. N., Ensminger, A. W., Clemson, C. M., Lynch, C. R., Lawrence, J. B., & Chess, A. (2007). A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. *BMC Genom.*, *8*.
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, *337*, 816–821.
- Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.*, *30*, 772–780.
- Keppler, A., Gendreizig, S., Gronemeyer, T., Pick, H., Vogel, H., & Johnsson, K. (2003). A general method for the covalent labeling of fusion proteins with small molecules in vivo. *Nat. Biotechnol.*, *21*, 86–89.
- Konermann, S., Lotfy, P., Brudeau, N. J., Oki, J., Shokhirev, M. N., & Hsu, P. D. (2018). Transcriptome engineering with RNA-targeting type VI-D CRISPR effectors. *Cell*, *173*, 665.

- Larson, D. R., Zenklusen, D., Wu, B., Chao, J. A., & Singer, R. H. (2011). Real-time observation of transcription initiation and elongation on an endogenous yeast gene. *Science*, 332, 475–478.
- Li, J. X., Chen, Z. J., Chen, F., Xie, G. Y., Ling, Y. Y., Peng, Y. X., Lin, Y., Luo, N., Chiang, C. M., & Wang, H. S. (2020). Targeted mRNA demethylation using an engineered dCas13b-ALKBH5 fusion protein. *Nucleic Acids Res.*, 48, 5684–5694.
- Ma, H. H., Tu, L. C., Naseri, A., Huisman, M., Zhang, S. J., Grunwald, D., & Pederson, T. (2016). Multiplexed labeling of genomic loci with dCas9 and engineered sgRNAs using CRISPRainbow. *Nat. Biotechnol.*, 34, 528–530.
- Mahas, A., Wang, Q. C., Marsic, T., & Mahfouz, M. M. (2021). A novel miniature CRISPR-cas13 system for SARS-CoV-2 diagnostics. *ACS Synth. Biol.*, 10, 2541–2551.
- Makarova, K. S., Wolf, Y. I., Iranzo, J., Shmakov, S. A., Alkhnbashi, O. S., Brouns, S. J. J., Charpentier, E., Cheng, D., Haft, D. H., Horvath, P., et al. (2020). Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants. *Nat. Rev. Microbiol.*, 18, 67–83.
- Marchler-Bauer, A., Bo, Y., Han, L. Y., He, J. E., Lanczycki, C. J., Lu, S. N., Chitsaz, F., Derbyshire, M. K., Geer, R. C., Gonzales, N. R., et al. (2017). CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res.*, 45, D200–D203.
- Nelles, D. A., Fang, M. Y., O'Connell, M. R., Xu, J. L., Markmiller, S. J., Doudna, J. A., & Yeo, G. W. (2016). Programmable RNA tracking in live cells with CRISPR/Cas9. *Cell*, 165, 488–496.
- Ninomiya, K., Adachi, S., Natsume, T., Iwakiri, J., Terai, G., Asai, K., & Hirose, T. (2020). LncRNA-dependent nuclear stress bodies promote intron retention through SR protein phosphorylation. *EMBO J.*, 39.
- Ninomiya, K., Iwakiri, J., Aly, M. K., Sakaguchi, Y., Adachi, S., Natsume, T., Terai, G., Asai, K., Suzuki, T., & Hirose, T. (2021). m(6)A modification of HSATIII lncRNAs regulates temperature-dependent splicing. *EMBO J.*, 40, Article e107976.
- Nollet, S., Moniaux, N., Maury, J., Petitprez, D., Degand, P., LAine, A., Porchet, N., & Aubert, J. P. (1998). Human mucin gene MUC4: organization of its 5'-region and polymorphism of its central tandem repeat array. *Biochem. J.*, 332, 739–748.
- O'Connell, M. R., Oakes, B. L., Sternberg, S. H., East-Seletsky, A., Kaplan, M., & Doudna, J. A. (2014). Programmable RNA recognition and cleavage by CRISPR/Cas9. *Nature*, 516, 263–266.
- Paige, J. S., Wu, K. Y., & Jaffrey, S. R. (2011). RNA mimics of green fluorescent protein. *Science*, 333, 642–646.
- Prasanth, K. V., Prasanth, S. G., Xuan, Z. Y., Hearn, S., Freier, S. M., Bennett, C. F., Zhang, M. Q., & Spector, D. L. (2005). Regulating gene expression through RNA nuclear retention. *Cell*, 123, 249–263.
- Raj, A., & Tyagi, S. (2010). Detection of individual endogenous RNA transcripts in situ using multiple singly labeled probes. *Methods Enzymol.*, 472, 365–386.
- Riback, J. A., Zhu, L., Ferrolino, M. C., Tolbert, M., Mitrea, D. M., Sanders, D. W., Wei, M. T., Kriwacki, R. W., & Brangwynne, C. P. (2020). Composition-dependent thermodynamics of intracellular phase separation. *Nature*, 581, 209.
- Shao, S. P., Zhang, W. W., Hu, H., Xue, B. X., Qin, J. S., Sun, C. Y., Sun, Y. A., Wei, W. S., & Sun, Y. J. (2016). Long-term dual-color tracking of genomic loci by modified sgRNAs of the CRISPR/Cas9 system. *Nucleic Acids Res.*, 44.
- Shmakov, S., Abudayyeh, O. O., Makarova, K. S., Wolf, Y. I., Gootenberg, J. S., Semenova, E., Minakhin, L., Joung, J., Konermann, S., Severinov, K., et al. (2015). Discovery and functional characterization of diverse class 2 CRISPR-cas systems. *Mol. Cell*, 60, 385–397.
- Smargon, A. A., Cox, D. B. T., Pyzocha, N. K., Zheng, K. J., Slaymaker, I. M., Gootenberg, J. S., Abudayyeh, O. A., Essletzbichler, P., Shmakov, S., Makarova, K. S., et al. (2017). Cas13b is a type VI-B CRISPR-associated RNA-guided RNase differentially regulated by accessory proteins Csx27 and Csx28. *Mol. Cell*, 65, 618.
- Soding, J. (2005). Protein homology detection by HMM-HMM comparison (vol 21, pg 951, 2005). *Bioinformatics*, 21, 2144–2144.
- Soding, J., Biegert, A., & Lupas, A. N. (2005). The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.*, 33, W244–W248.
- Steinegger, M., & Soding, J. (2017). MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat. Biotechnol.*, 35, 1026–1028.
- Strutt, S. C., Torrez, R. M., Kaya, E., Negrete, O. A., & Doudna, J. A. (2018). RNA-dependent RNA targeting by CRISPR-Cas9. *Elife*, 7.
- Tian, S., Zhang, B., He, Y., Sun, Z., Li, J., Li, Y., , Chen, W., et al. (2022). CRISPR-iPAS: a novel dCAS13-based method for alternative polyadenylation interference. *Nucleic Acids Res.*, 50(5), e26.
- Wan, Y. H., Anastasaki, D. G., Rodriguez, J., Palangat, M., Gudla, P., Zaki, G., Tandon, M., Pegeraro, G., Chow, C. C., Hafner, M., et al. (2021). Dynamic imaging of nascent RNA reveals general principles of transcription dynamics and stochastic splice site selection. *Cell*, 184, 2878.
- Wang, H., Nakamura, M., Abbott, T. R., Zhao, D. H., Luo, K. W., Yu, C., Nguyen, C. M., Lo, A., Daley, T. P., La Russa, M., et al. (2019). CRISPR-mediated live imaging of genome editing and transcription. *Science*, 365, 1301.
- Wang, Y., Hu, S. B., Wang, M. R., Yao, R. W., Wu, D., Yang, L., & Chen, L. L. (2018). Genome-wide screening of NEAT1 regulators reveals cross-regulation between paraspeckles and mitochondria. *Nat. Cell Biol.*, 20, 1145.
- West, J. A., Mito, M., Kurosa, S., Takumi, T., Tanegashima, C., Chujo, T., Yanaka, K., Kingston, R. E., Hirose, T., Bond, C., et al. (2016). Structural, super-resolution microscopy analysis of paraspeckle nuclear body organization. *J. Cell Biol.*, 214, 817–830.
- Wu, J. H., Zaccara, S., Khuperkar, D., Kim, H., Tanenbaum, M. E., & Jaffrey, S. R. (2019). Live imaging of mRNA using RNA-stabilized fluorogenic proteins. *Nat. Methods*, 16, 862.
- Xiang, J. F., Yin, Q. F., Chen, T., Zhang, Y., Zhang, X. O., Wu, Z., Zhang, S. F., Wang, H. B., Ge, J. H., Lu, X. H., et al. (2014). Human colorectal cancer-specific CCAT1-L lncRNA regulates long-range chromatin interactions at the MYC locus (vol 24, pg 513, 2014). *Cell Res.*, 24, 1150, 1150.
- Xing, Y. H., Yao, R. W., Zhang, Y., Guo, C. J., Jiang, S., Xu, G., Dong, R., Yang, L., & Chen, L. L. (2017). SLERT regulates DDX21 rings associated with pol I transcription. *Cell*, 169, 664–678.
- Yan, W. X., Chong, S. R., Zhang, H. B., Makarova, K. S., Koonin, E. V., Cheng, D. R., & Scott, D. A. (2018). Cas13d is a compact RNA-targeting type VI CRISPR effector positively modulated by a WYL-domain-containing accessory protein. *Mol. Cell*, 70, 327.
- Yang, L. Z., Wang, Y., Li, S. Q., Yao, R. W., Luan, P. F., Wu, H., Carmichael, G. G., & Chen, L. L. (2019). Dynamic imaging of RNA in living cells by CRISPR-cas13 systems. *Mol. Cell*, 76, 981.
- Yao, R. W., Xu, G., Wang, Y., Shan, L., Luan, P. F., Wang, Y., Wu, M., Yang, L. Z., Xing, Y. H., Yang, L., et al. (2019). Nascent pre-rRNA sorting via phase separation drives the assembly of dense fibrillar components in the human nucleolus. *Mol. Cell*, 76, 767.
- Zhang, B., Ye, W. W., Ye, Y. M., Zhou, H., Saeed, A. F. U. H., Chen, J., Lin, J. Y., Perculija, V., Chen, Q., Chen, C. J., et al. (2018). Structural insights into Cas13b-guided CRISPR RNA maturation and recognition. *Cell Res.*, 28, 1198–1201.
- Zhang, Z. H., Sun, W. P., Shi, T. Z., Lu, P. F., Zhuang, M., & Liu, J. L. (2020). Capturing RNA-protein interaction via CRUIS. *Nucleic Acids Res.*, 48.