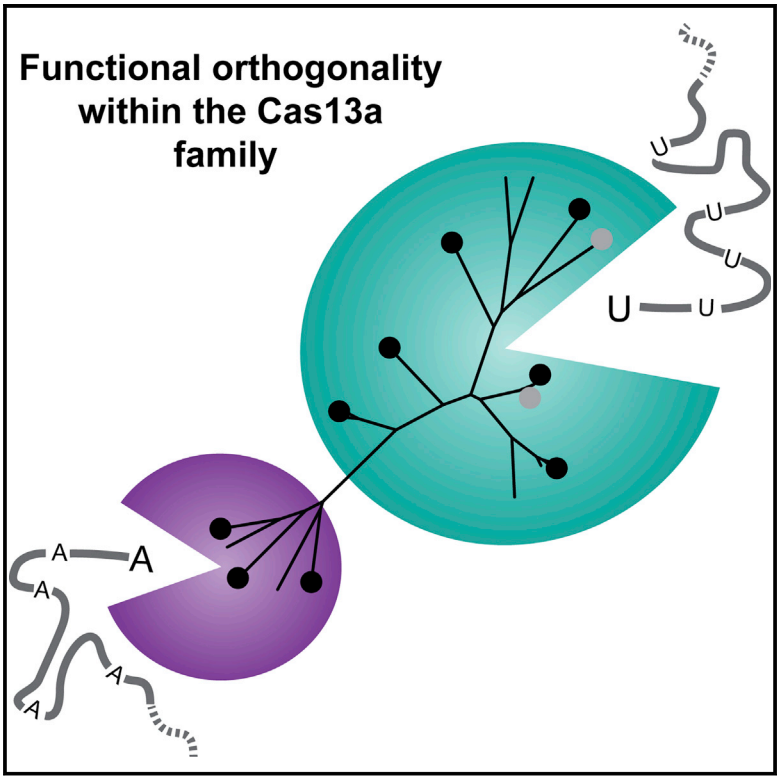
Article

Molecular Cell

RNA Targeting by Functionally Orthogonal Type VI-A CRISPR-Cas Enzymes

Graphical Abstract



Authors

Alexandra East-Seletsky, Mitchell R. O'Connell, David Burstein, Gavin J. Knott, Jennifer A. Doudna

Correspondence

doudna@berkeley.edu

In Brief

East-Seletsky et al. demonstrate that the CRISPR-Cas type VI-A protein family contains surprising functional variability. They define two enzyme subfamilies that use orthogonal guide RNAs and cleave with distinct substrate preferences.

Highlights

- Cas13a family contains two distinct subfamilies
- Cas13a subfamilies possess distinct nuclease substrate preferences
- Cas13a subfamilies also use orthogonal crRNAs
- crRNA processing is not essential for ssRNA targeting by Cas13a

East-Seletsky et al., 2017, Molecular Cell 66, 373–383
May 4, 2017 © 2017 Elsevier Inc.
http://dx.doi.org/10.1016/j.molcel.2017.04.008

CellPress

Molecular Cell
Article

RNA Targeting by Functionally Orthogonal Type VI-A CRISPR-Cas Enzymes

Alexandra East-Seletsky,[1] Mitchell R. O'Connell,[1] David Burstein,[2] Gavin J. Knott,[1] and Jennifer A. Doudna[1,3,4,5,6,7,*]
[1]Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720, USA
[2]Department of Earth and Planetary Sciences, University of California, Berkeley, CA 94720, USA
[3]Department of Chemistry, University of California, Berkeley, CA 94720, USA
[4]Howard Hughes Medical Institute, University of California, Berkeley, CA 94720, USA
[5]Innovative Genomics Institute, University of California, Berkeley, CA 94720, USA
[6]MBIB Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
[7]Lead Contact
*Correspondence: doudna@berkeley.edu
http://dx.doi.org/10.1016/j.molcel.2017.04.008

SUMMARY

CRISPR adaptive immunity pathways protect prokaryotic cells against foreign nucleic acids using CRISPR RNA (crRNA)-guided nucleases. In type VI-A CRISPR-Cas systems, the signature protein Cas13a (formerly C2c2) contains two separate ribonuclease activities that catalyze crRNA maturation and ssRNA degradation. The Cas13a protein family occurs across different bacterial phyla and varies widely in both protein sequence and corresponding crRNA sequence conservation. Although grouped phylogenetically together, we show that the Cas13a enzyme family comprises two distinct functional groups that recognize orthogonal sets of crRNAs and possess different ssRNA cleavage specificities. These functional distinctions could not be bioinformatically predicted, suggesting more subtle co-evolution of Cas13a enzymes. Additionally, we find that Cas13a pre-crRNA processing is not essential for ssRNA cleavage, although it enhances ssRNA targeting for crRNAs encoded internally within the CRISPR array. We define two Cas13a protein subfamilies that can operate in parallel for RNA detection both in bacteria and for diagnostic applications.

INTRODUCTION

Competition for survival within microbial communities drives the evolution of diverse pathways for antiviral defense (Dy et al., 2014). Unique among such defense mechanisms, CRISPR-Cas (clustered *r*egularly *i*nterspaced *s*hort *p*alindromic *r*epeats, CRISPR *a*ssociated) systems provide adaptive immunity by means of CRISPR genomic loci that contain integrated viral DNA sequences (spacers) flanked by conserved palindromic sequences (repeats) (Charpentier et al., 2015; Mohanraju et al., 2016; Wright et al., 2016). CRISPR RNAs (crRNAs) expressed from these loci assemble with Cas proteins to form surveillance complexes that recognize and cleave nucleic acids matching the virally derived segment of the crRNA. Three steps ensure ongoing adaptive immunity: (1) integration of new virus spacer sequences into the CRISPR locus, (2) Cas protein expression and crRNA biogenesis, and (3) surveillance complex assembly and target interference (Barrangou et al., 2007; Brouns et al., 2008; Garneau et al., 2010).

Although double-stranded DNA (dsDNA) is the typical CRISPR-Cas target, type III and VI CRISPR-Cas systems instead recognize single-stranded RNA (ssRNA). Unlike the DNA-targeting CRISPR effector complexes or the type III CRISPR effector complexes, which catalyze site-specific cleavage of DNA or RNA sequences specifically matching the crRNA, respectively (reviewed in Charpentier et al., 2015; Mohanraju et al., 2016; Wright et al., 2016), the type VI systems catalyze complete degradation of any ssRNA present upon activation (Abudayyeh et al., 2016; East-Seletsky et al., 2016; Smargon et al., 2017). The signature protein of type VI-A CRISPR-Cas systems, Cas13a (formerly C2c2), is a dual nuclease responsible for both crRNA maturation and RNA-activated ssRNA cleavage (East-Seletsky et al., 2016). Cas13a binds to precursor crRNA (pre-crRNA) transcripts and cleaves them within the repeat region to produce mature crRNAs. This cleavage reaction is presumed to be single-turnover, resulting in a functional Cas13a:crRNA surveillance complex. Binding to a ssRNA with sequence complementarity to the crRNA activates Cas13a for *trans*-ssRNA cleavage, potentially triggering cell death or dormancy of the host organism (Abudayyeh et al., 2016; East-Seletsky et al., 2016; Liu et al., 2017). While previous work demonstrated that these two enzymatic activities of Cas13a are chemically distinct and occur within separate active sites (East-Seletsky et al., 2016; Liu et al., 2017), the interdependence, if any, between active site substrate specificities is unknown.

Multiple CRISPR-Cas loci, whether from redundant or distinct subtypes, commonly co-exist within host genomes, but the relationships between these systems are unclear. Maturation of crRNAs from different CRISPR loci is sometimes catalyzed by the same processing enzyme (Nickel et al., 2013; Niewoehner et al., 2014), while other organisms require distinct enzymes for each crRNA repeat species (Scholz et al., 2013; Sokolowski et al., 2014). Once generated, crRNAs can sometimes assemble

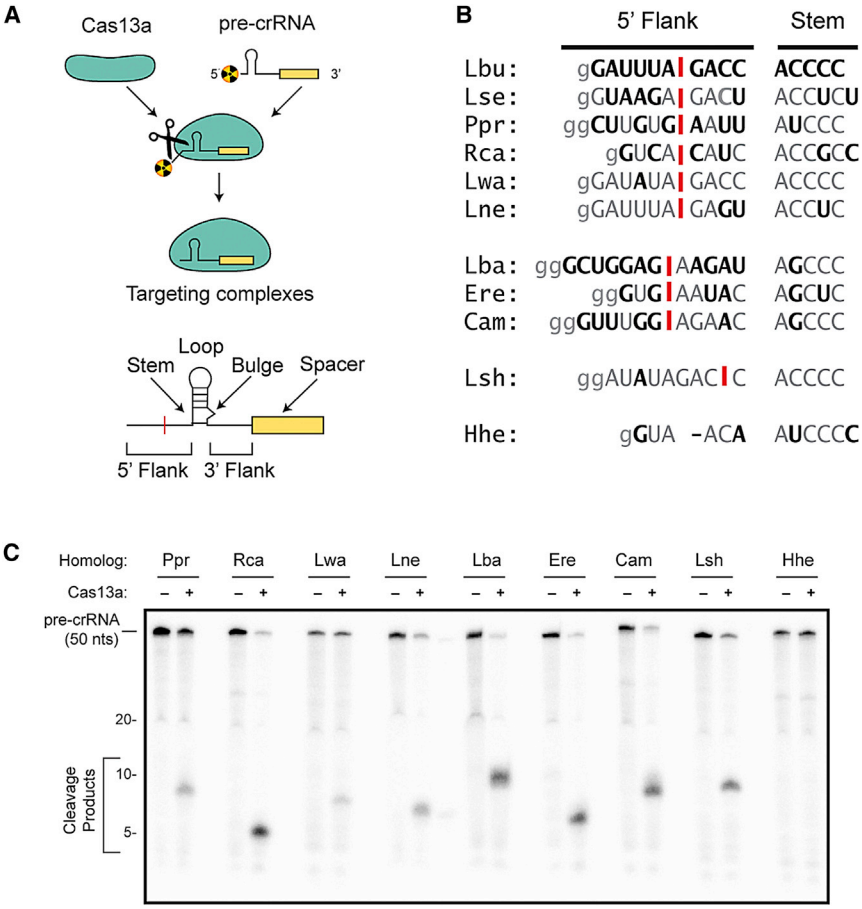


Figure 1. Pre-crRNA Processing Is Broadly Conserved within the Cas13a Protein Family

(A) Schematic of crRNA processing reaction catalyzed by Cas13a. Pre-crRNA substrates are cleaved by Cas13a to generate mature crRNAs. Below, a schematic of a pre-crRNA highlighting important functional features.

(B) Alignment of the 5′ portion of CRISPR repeat sequences from the studied type VI CRISPR systems highlighting the pre-crRNA cleavage site. Mapped cleavage sites are shown as red bars. Deviations from the Lbu crRNA-repeat sequence are noted in black text. Lowercase "g"s were required for transcription purposes and are not part of the native crRNA repeat sequences. The full CRISPR repeat sequence is diagrammed in Figure S1B. Cleavage sites were mapped in previous studies for LbuCas13a, LshCas13a, and LseCas13a (East-Seletsky et al., 2016).

(C) Representative gel of Cas13a-mediated pre-crRNA cleavage by nine Cas13a homologs after 60 min incubation with 5′-radiolabeled pre-crRNA substrates. Cleavage by LbuCas13a has been previously demonstrated (East-Seletsky et al., 2016) and can be observed in Figure 2.

into targeting complexes that include Cas proteins from other loci, but the functional consequences of such CRISPR-loci crosstalk are not known (Staals et al., 2014; Zhang et al., 2012). Cas13a provides a unique opportunity to study pathway interdependence within a CRISPR subtype for both pre-crRNA processing and target cleavage since the same protein catalyzes both activities. Furthermore, understanding the extent of orthogonality between distinct type VI-A CRISPR-Cas systems (i.e., are there subsets of type VI-A systems that recognize orthogonal sets of crRNAs and/or possess different ssRNA cleavage specificities) will advance the development of these proteins as biotechnological tools.

CRISPR proteins are classified according to sequence and phylogenetic conservation in attempts to simplify the broad diversity found within these microbial defense systems (Makarova et al., 2011, 2015; Shmakov et al., 2017). In turn, representative operons are investigated, with the subsequent conclusions applied across the entire subtype protein family. But it is unclear how appropriate these generalizations are, given the extremely fast evolution of CRISPR enzymes driven by bacteria-phage interactions. Within the type VI-A protein family, homologs only share 15%–20% residue identity and exist within diverse host genomes. This suggested that we could investigate divergent Cas13a homologs, which presumably each contain two well-defined nuclease activities, to test the functional diversity within a phylogenetically determined CRISPR-Cas subtype. Working with ten Cas13a homologs that span the major branches of the Cas13a family, we find that both bioinformatic analyses and biochemical results define two functional enzyme subfamilies within the type VI-A subtype. Orthogonal pre-crRNA and ssRNA-target cleaving activities distinguish these groups, such that each enzyme subfamily recognizes distinct crRNAs yielding targeting complexes with subfamily-distinct ssRNA-degradation patterns. Furthermore, these results revealed that pre-crRNA processing is not essential for ssRNA targeting in type VI systems, despite strong conservation of this function within the Cas13a family. Instead, pre-crRNA processing enhances *trans*-ssRNA targeting by liberating crRNAs from the long pre-crRNA transcript, leading to a revised model for type VI-A CRISPR-Cas systems. The two orthogonal subfamilies of Cas13a enable potential applications in RNA detection or imaging for the parallel detection of distinct RNA species.

RESULTS

Most Cas13a Homologs Possess pre-crRNA Processing Activity

To explore the functional diversity of Cas13a proteins, we compared the pre-crRNA processing activities of ten homologs from across the protein family tree for their capacity to produce mature crRNAs from cognate pre-crRNAs (Figures 1A and S1). Similar to the three homologs studied previously (LbuCas13a, LseCas13a, and LshCas13a) (East-Seletsky et al., 2016), seven additional Cas13a enzymes possess crRNA maturation activity (Figures 1B and 1C). Only one of the 11 Cas13a proteins tested

to date exhibited no detectable cleavage of its cognate pre-crRNA across a wide range of assay conditions (HheCas13a) (Figures 1 and S1). Of the homologs that processed their native crRNAs, all but LshCas13a cleaved at the phosphodiester bond four or five nucleotides upstream of the conserved crRNA-repeat hairpin (Figures 1B and 1C). These results show strong conservation of Cas13a-mediated crRNA biogenesis activity within type VI-A CRISPR-Cas systems.

A Conserved crRNA Maturation Center within Most Cas13a Enzymes

Previous studies have implicated two distinct regions of Cas13a as responsible for pre-crRNA processing. The general Cas13a protein architecture established by the LshCas13a crystal structure consists of an N-terminal domain and two HEPN (*h*igher *e*ukaryotes and *p*rokaryotes *n*ucleotide-binding) domains separated by two helical domains (Figure 2A) (Liu et al., 2017). For LbuCas13a, mutation of a single residue (R1079) within the HEPN2 domain was sufficient to substantially reduce pre-crRNA processing activity (East-Seletsky et al., 2016). By contrast, mutations at two positions located in the helical 1 domain, R438 and K441, were shown to diminish pre-crRNA cleavage by LshCas13a (Liu et al., 2017). While both of these regions may be involved in pre-crRNA processing, it is unclear whether processing by LbuCas13a is inhibited by helical 1 domain mutations and which domain is primarily responsible for crRNA maturation across the Cas13a protein family.

First we examined the conservation of both the helical 1 and HEPN2 domains across 19 Cas13a homologs. Previously reported alignments (Liu et al., 2017; Shmakov et al., 2015) conflict in the helical 1 domain region, suggesting high ambiguity in the relationship between homologs in this domain. We observed minimal conservation within our alignment of the helical 1 domain implicated in pre-crRNA processing, while in contrast, we observed consistent levels of conservation across the HEPN2 domain (Figures 2B, 2C, and S2; Data S1). The only pre-crRNA processing-defective homolog, HheCas13a, maintains a majority of the conserved charged residues throughout both domains, suggesting that other parts of the protein or the repeat sequence may be preventing pre-crRNA cleavage. Among the homologs used in this study, LshCas13a is the most divergent across the HEPN2 domain, potentially explaining the alternative catalytic domain and atypical cleavage site selection by this homolog.

Next, we tested whether the helical 1 domain residues critical for LshCas13a pre-crRNA processing affect LbuCas13a's pre-crRNA processing activity. Due to discrepancies between alignments presented in the previous published studies, we tested four residues (E299, K310, R311, and N314) for their role in pre-crRNA cleavage (Figure 2B) (East-Seletsky et al., 2016; Liu et al., 2017; Shmakov et al., 2015). Mutation of these residues to alanine revealed a range of impacts on pre-crRNA processing efficiencies: N314A significantly reduced the cleavage rate, R311A minimally impaired activity, and E299A and K310 had no effect on pre-crRNA processing (Figure 2D). In parallel, we performed mutagenesis within the HEPN2 domain of LbuCas13a. Alanine substitutions at R1072 and K1082 significantly reduced pre-crRNA cleavage, while other mutations in the same region (D1078A, K1080A, and K1087A) had minimal impacts on pre-crRNA processing (Figure 2E). These results suggest that the HEPN2 and to a lesser extent the helical 1 domains play significant roles in crRNA biogenesis for LbuCas13a, although the residues directly responsible for catalyzing hydrolysis remain unknown. The difference between the regions implicated in pre-crRNA processing across LbuCas13a (this study) and LshCas13a (Liu et al., 2017) do not necessarily contradict each other, as the 5′ terminus of the crRNA is held between the HEPN2 and helical 1 domains in the LshCas13a structure (Liu et al., 2017). Residues from both domains might play pivotal roles in pre-crRNA processing by stabilizing substrate binding, promoting proper substrate orientation, and/or catalyzing hydrolysis.

The lack of conservation within the HEPN2 domain of LshCas13a, its putative pre-crRNA processing active site, and its atypical pre-crRNA cleavage site led us to hypothesize that LshCas13a may utilize a different region within the helical 1 domain to catalyze pre-crRNA processing. In the absence of a three-dimensional structure of a pre-crRNA-bound Cas13a homolog, we tested this hypothesis by mapping the LshCas13a cleavage sites on non-cognate pre-crRNAs (Figure 2F). LshCas13a was able to process pre-crRNAs from LwaCas13a and LbuCas13a, generating a shifted cleavage site one nucleotide from the predicted hairpin base. In concordance with this observation, processing of the Lsh pre-crRNA by LbuCas13a, LwaCas13a, and PprCas13a occurs at the standard four-nucleotide interval from the repeat stem, differing from the cognate LshCas13a site. This supports our previous observations that the distinct LshCas13a processing site depends on the protein architecture, not the pre-crRNA sequence, and that LshCas13a is an outlier within the Cas13a tree with regard to pre-crRNA processing.

Cas13a Enzymes Initiate ssRNA Cleavage at Either Uridines or Adenosines

Previous studies established that Cas13a:crRNA complexes recognize and bind complementary ssRNA targets, hereby referred to as ssRNA activators, to trigger general RNase activity at exposed uridine residues (Figure 3A) (Abudayyeh et al., 2016; East-Seletsky et al., 2016). We wondered if the panel of homologs within this study retained the non-specific degradation activity previously demonstrated by LbuCas13a and LshCas13a and if the uridine preference within the HEPN active site is universal within the family. To systematically test general RNase activity, we monitored the ability of a ternary complex comprising Cas13a:crRNA with a bound ssRNA activator to degrade a *trans*-ssRNA target (Figure 3A). We detected *trans*-ssRNA cleavage activity for 8 of the 10 homologs over the course of 1 hr (Figures 3B and S3A). Most notable of the Cas13a homologs active for *trans*-ssRNA cleavage is HheCas13a, which possesses no detectable pre-crRNA processing activity, yet catalyzed complete degradation of substrates guided by its cognate pre-crRNA. While the previously characterized homologs LshCas13a (Abudayyeh et al., 2016) and LbuCas13a (East-Seletsky et al., 2016) both exhibit a preference for uridine 5′ to the scissile bond, products of different lengths generated by the other homologs suggest that different active site nucleotide preferences may exist within this protein family (Figure S3B).

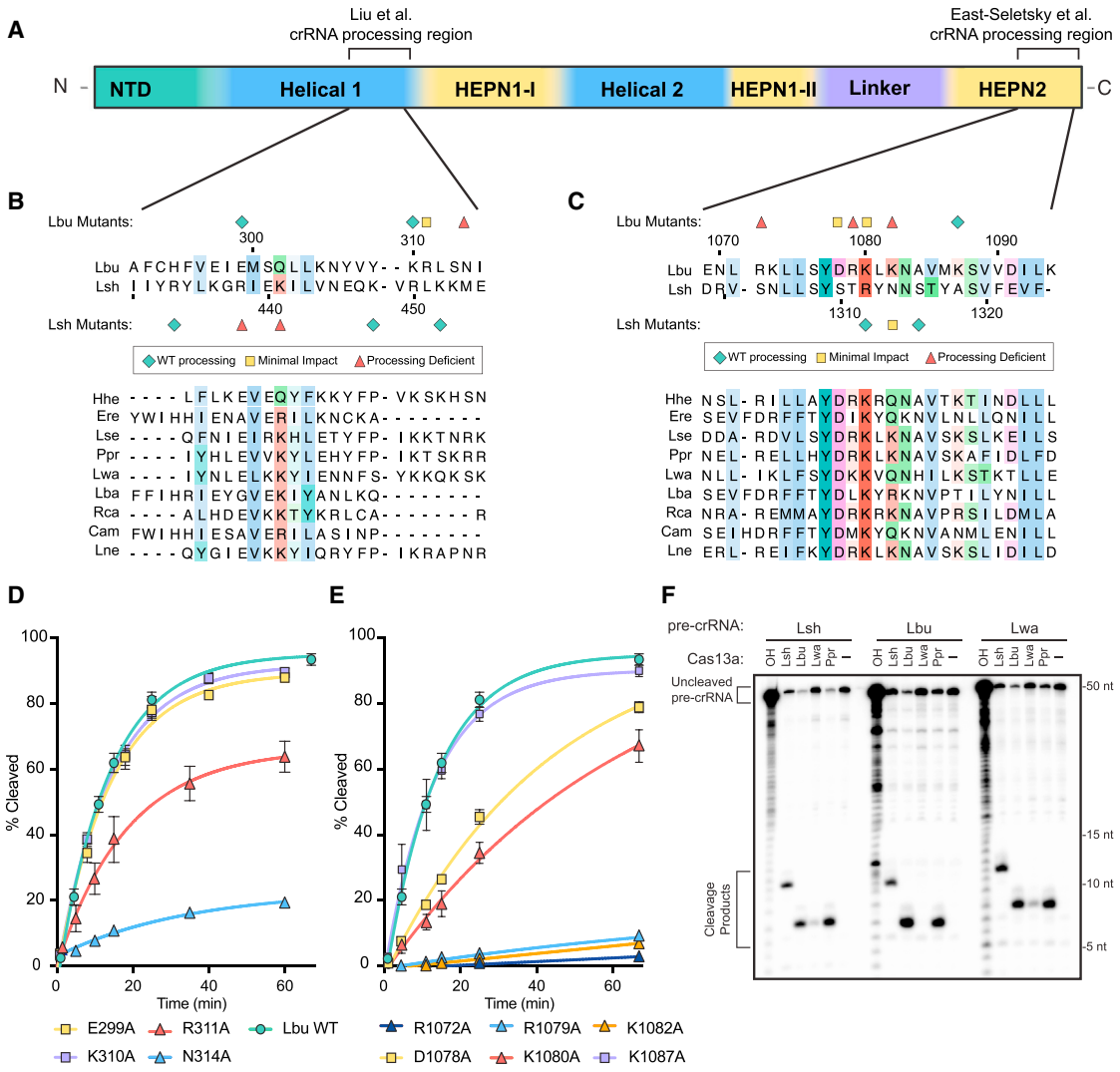


Figure 2. Identification of Residues Important for Pre-crRNA Cleavage by LbuCas13a

(A) General Cas13a domain organization schematic with annotations based off a LshCas13a crystal structure (Liu et al., 2017).

(B and C) Multiple-sequence amino acid alignment of (B) the local region in Cas13a's helical 1 domain implicated in pre-crRNA processing by studies on LshCas13a and (C) the region within in Cas13a's HEPN2 domain implicated in pre-crRNA processing by studies on LbuCas13a. See Figure S2 for full family tree alignment of these regions and Data S1 for a complete protein alignment. Residues whose mutation severely affects pre-crRNA processing are marked by red triangles, minimal impacts on processing by yellow squares, and residues whose mutation did not affect pre-crRNA processing are marked with teal diamonds. Symbols above the LbuCas13a sequences correspond to mutations made to LbuCas13a, and symbols below the LshCas13a sequence correspond to mutations made to LshCas13a by Liu et al., 2017. Coloration of the matrix alignment denotes residue conservation using the ClustalX scheme, with darker hues indicating the strength of the conservation.

(D and E) Pre-crRNA processing under single turnover conditions was measured for mutants in (D) the helical 1 domain and (E) the HEPN2 domain. Quantified data were fitted with single-exponential decays with calculated pseudo-first-order rate constants (k_{obs}) (mean ± SD, n = 3) as follows: Lbu WT, 0.074 ± 0.003 min^{-1}; E299A, 0.071 ± 0.005 min^{-1}; K310A, 0.071 ± 0.003 min^{-1}; R311A, 0.054 ± 0.007 min^{-1}; N314A, 0.029 ± 0.008 min^{-1}; R1079A, 0.009 ± 0.007 min^{-1}; D1078A, 0.023 ± 0.002 min^{-1}; K1080A, 0.016 ± 0.004 min^{-1}; and K1087A, 0.076 ± 0.007 min^{-1}; while R1072A and K1082A could not be fitted.

(F) Representative gel of pre-crRNA processing of LshCas13a, LwaCas13a, and LbuCas13a pre-crRNAs by LbuCas13a, LshCas13a, LwaCas13a, and PprCas13a proteins using standard conditions. Hydrolysis ladders for each pre-cRNA substrate demonstrate subtle differences in the migration of these fragments from differing sequences.

To further probe the *trans*-ssRNA cleavage nucleotide preferences of the Cas13a homologs, we measured the *trans*-cleavage capacity of these enzymes using 5-mer homopolymers of A, C, G, and U as substrates. Six homologs were able to cleave these short substrates (Figures 3C and S3C; Table S2). Four of the homologs, LbuCas13a, LwaCas13a HheCas13a, and PprCas13a, exhibited preferred cleavage of the homouridine substrate, although secondary preferences were observed for the homologs with the highest activities (Figures 3C and S3D). In contrast, LbaCas13a and EreCas13a preferred

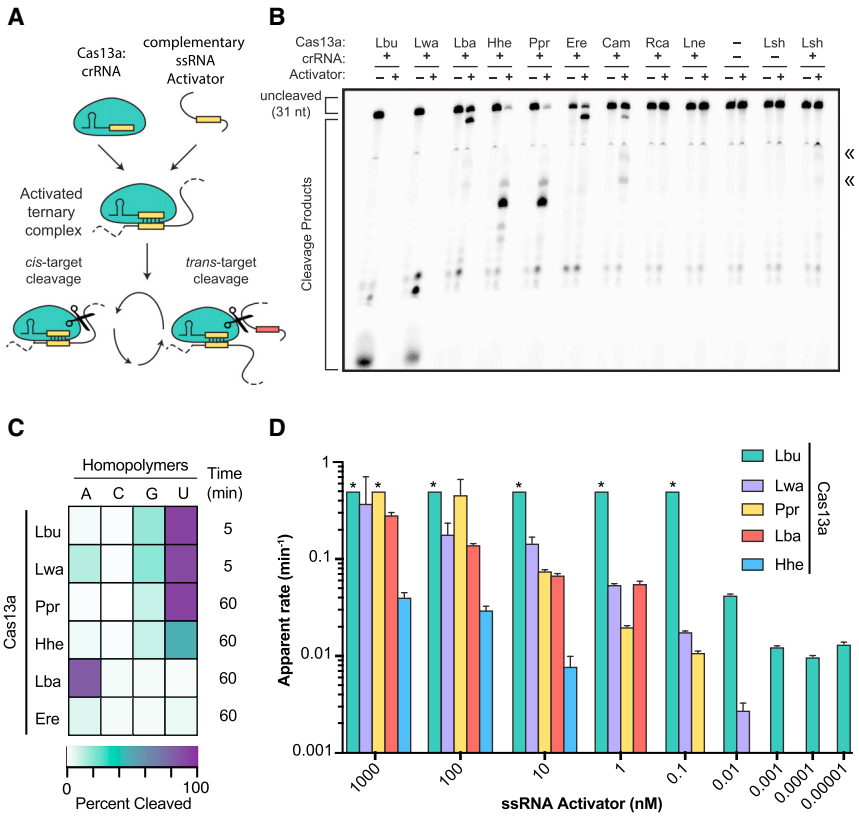


Figure 3. Members of the Cas13a Protein Family Cleave ssRNA with a Range of Efficiencies

(A) Schematic of ssRNA-targeting by Cas13a. For simplicity, *trans*-ssRNA cleavage was the focus of this study.

(B) Representative gel of Cas13a-mediated *trans*-ssRNA cleavage by all ten homologs after 60 min incubation. Cas13a:crRNA complexes were formed as described in STAR★Methods using mature crRNA products with a final RNP complex concentration of 50 nM. < 1 nM radiolabeled *trans*-ssRNA target was added to initiate reaction in the presence and absence of 50 nM unlabeled, crRNA-complementary ssRNA activator. Weak *trans*-ssRNA cleavage activity was observed by LshCas13a with product bands noted by double arrows to the right.

(C) Heatmap reporting Cas13a-catalyzed *trans*-ssRNA cleavage percentages for each 5-mer homopolymer ssRNA substrate, for six different Cas13a homologs. Assay conditions were identical to (B), except LbuCas13a and LwaCas13a, which were incubated for 5 min instead of 60 min (n = 3, values with associated errors presented in Table S3).

(D) Apparent cleavage rates of a fluorescent ssRNA reporter by five homologs across a range of ssRNA activator concentrations. Cas13a:crRNA complexes were pre-incubated at a 2:1 ratio, with a final active complex concentration of 50 nM. Complementary ssRNA activator and fluorescent ssRNA cleavage reporter were added to initiate reactions. Normalized reporter signal curve time courses were fitted with single-exponential decays, and the apparent rates are plotted (n = 3). Some conditions plateaued before the first measured time point; therefore their rates are minimally assumed to be 0.5 min^{-1} and are labeled with a * in the chart.

homo-adenosine in agreement with biochemically mapped cleavage sites on longer targets (Figure S3B). Identical product generation from these long substrates by CamCas13a is consistent with adenosine preference by this clade of the Cas13a family tree (Figures 3, S1, and S3).

One notable difference between these enzymes was the rate at which *trans*-ssRNA cleavage reaches saturation under the tested conditions of equimolar ssRNA activator and Cas13a:crRNA interference complex. These enzymatic differences could have dramatic effects on Cas13a's biological role, so we next aimed to quantify the variance in *trans*-ssRNA cleavage within the Cas13a homolog family. We developed a high-throughput screen utilizing a short fluorescent ssRNA reporter for RNA cleavage to account for both ssRNA activator binding and *trans*-ssRNA cleavage, the two core properties of Cas13a enzymes that contribute to total enzymatic output. To interrogate the sensitivity of each Cas13a homolog, decreasing amounts of complementary ssRNA activator were added to initiate the reaction, and the apparent rate of fluorescent ssRNA reporter cleavage was calculated from each of the resulting time courses. While the calculated rates are a convolution of the ssRNA activator binding affinity and the catalytic turnover rate for each of the enzymes, they give a relative measure of cleavage activity that is comparable across homologs.

Five homologs (LbuCas13a, LwaCas13a, LbaCas13a, HheCas13a, and PprCas13a) demonstrated sufficiently detectable cleavage activity within this assay for reproducible analysis. Of these five homologs, LbuCas13a exhibited the most sensitivity, with detectable reporter cleavage in the presence of only 10 fM complementary activator (Figure 3D). Only two homologs, LwaCas13a and PprCas13a, displayed enough activity to detect the activator in the picomolar range, with sensitivities of 10 pM and 100 pM, respectively. LbaCas13a and HheCas13a were much less sensitive, only becoming active at nanomolar levels of reporter, which is close to equimolar relative to the Cas13a complex. Since this assay relies on a substantial number of *trans*-cleavage events to produce detectable fluorescence, we can assume that the three homologs unable to produce detectable signal above background despite similar cleavage site preferences (EreCas13a, CamCas13a, and LshCas13a) possess even poorer complementary target sensitivities. The remarkably broad range in sensitivities ($\sim 10^7$-fold) suggests a diverse capacity of Cas13a enzymes to protect a host organism from foreign RNA.

CRISPR Repeat Sequence Determines Non-cognate Pre-crRNA Processing

The dual activities of Cas13a provide an opportunity to study the interdependence of pre-crRNA processing and targeting

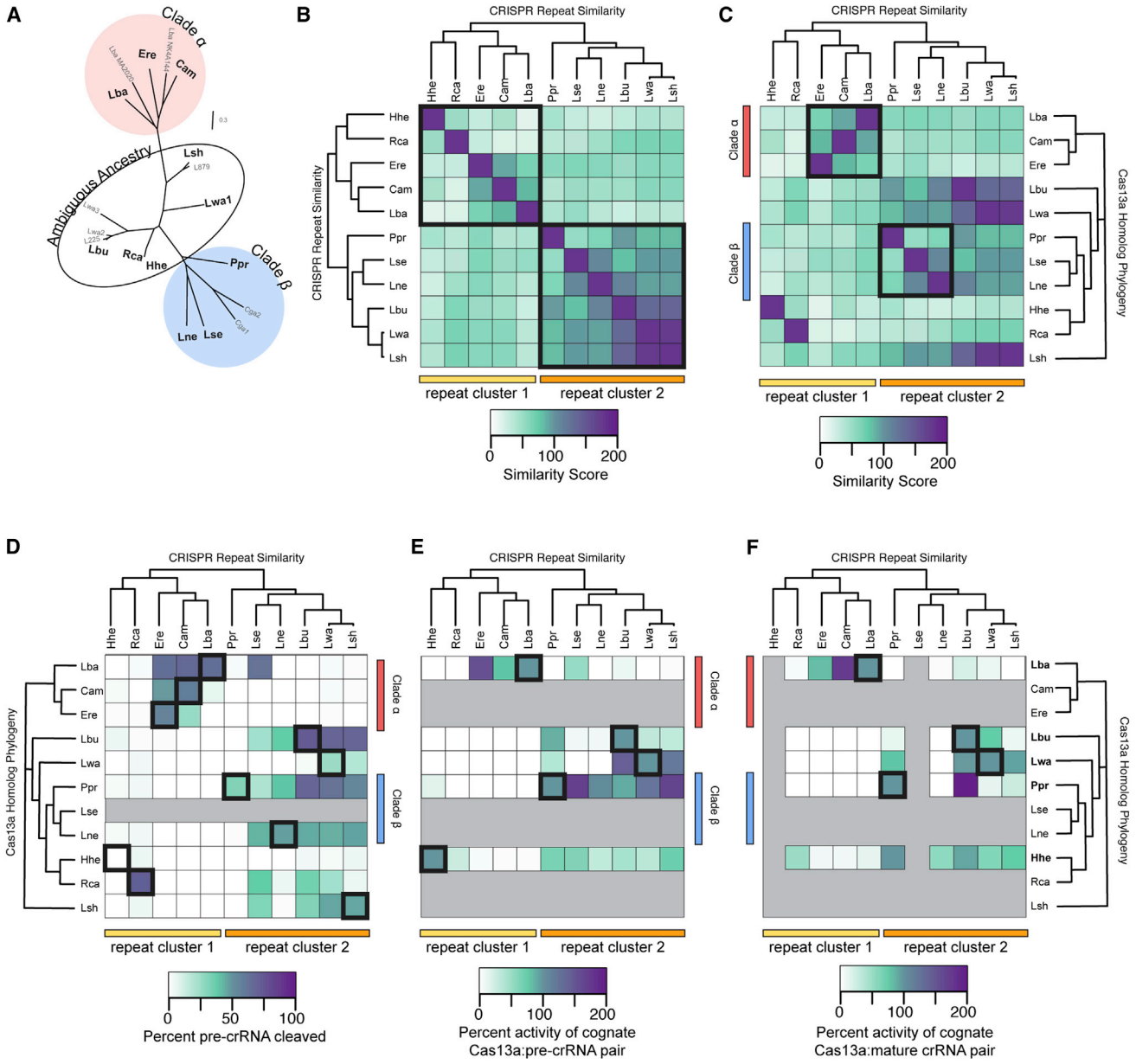


Figure 4. crRNA Exchangeability within the Cas13a Family

(A) Maximum-likelihood phylogenetic tree of all Cas13a family members. Homologs used in this study are bolded and clades are highlighted. Bootstrapped values are located in previous work (East-Seletsky et al., 2016).

(B) Symmetrical similarity score matrix for CRISPR repeats from homologs used in this study. Rows and columns are ordered by CRISPR repeat clustering.

(C) Asymmetrical similarity score matrix for CRISPR repeats from homologs used in this study. The same pairwise scores are presented here as in (B), except the rows are reordered to correspond to the Cas13a phylogenetic tree of the subset of homologs used in this study. Bootstrap values for this smaller tree are in Figure S1A.

(D–F) Functional activity matrix for (D) pre-crRNA processing by non-cognate proteins, (E) *trans*-ssRNA cleavage directed by pre-crRNAs, and (F) *trans*-ssRNA cleavage directed by mature crRNAs. Processing assays were performed using standard conditions, and 60 min reaction endpoints were analyzed. *trans*-ssRNA cleavage assays were performed using the fluorescent ssRNA reporter assay with fitted initial rates. ssRNA activator concentrations were as follows: LbuCas13a,100 pM; LwaCas13a,100 pM; PprCas13a,100 nM; LbaCas13a,10 nM; HheCas13a,100 nM. Initial rates were fit across three replicates to account for differences in fluorescence plateau values and normalized to each Cas13a:crRNA cognate pair. See Tables S4–S6 for numerical values and associated errors (n = 3).

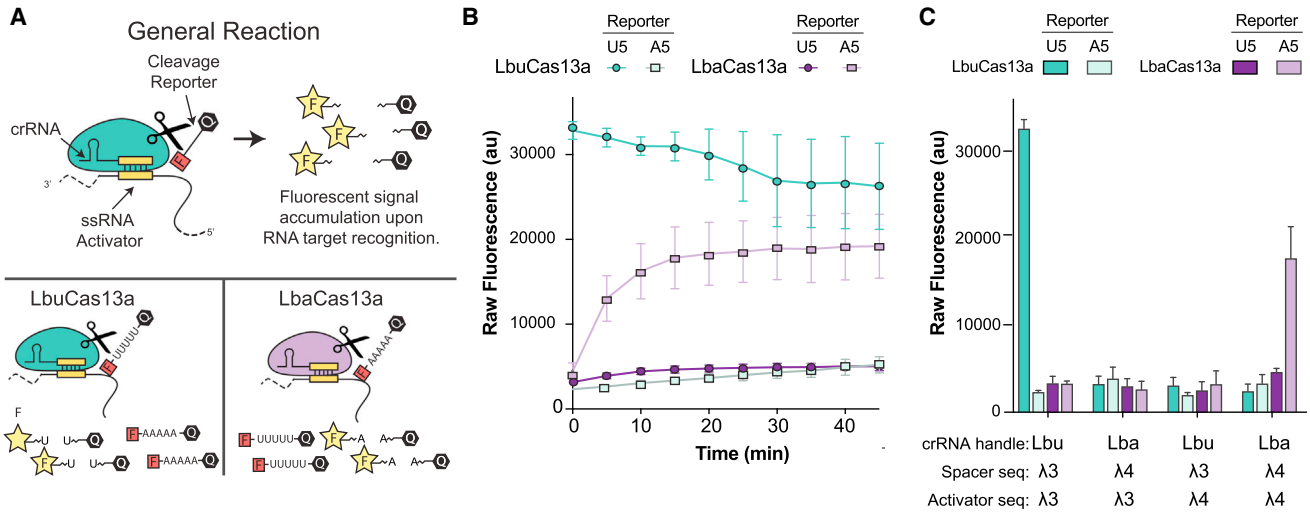


Figure 5. Functional Validation of Orthogonal Cas13a Subfamilies for RNA Detection

(A) Schematic of the RNA detection assay modified to use fluorescent homopolymer ssRNA reporter substrates to assay *trans*-ssRNA cleavage activation by either LbuCas13a or LbaCas13a.

(B) Time course of raw fluorescence measurements generated by homopolymer reporters incubated with either LbuCas13a: Lbu-crRNA: 10 pM ssRNA activator or LbaCas13a: Lba-crRNA: 1 nM ssRNA activator (mean ± SD, n = 3).

(C) Raw fluorescence measurements generated by the fluorescent homopolymer ssRNA reporters across a panel of crRNA, ssRNA activator, and Cas13a protein combinations (mean ± SD, n = 3).

between distinct type VI-A CRISPR-Cas operons. To determine the substrate requirements for both activities, we asked to what extent different homologs can recognize non-cognate crRNAs for guide processing and targeting. Initially, we hoped to predict bioinformatically the likely crRNA exchangeability through phylogenetic analysis of the Cas13a family and crRNA similarities. This analysis suggested that two distinct clades of homologs exist, termed alpha and beta for clarity (Stamatakis, 2014) (Figure 4A). Five of our purified homologs exist outside of these clades with ambiguous ancestral relationships, leading us to wonder if the pre-crRNA (CRISPR repeat) sequence might dictate functional orthogonality, or incompatibility of these crRNA sequences to facilitate *trans*-ssRNA cleavage when paired with a non-cognate Cas13a protein. Due to the short and structured content of the CRISPR repeats, we used a pairwise sequence alignment score matrix to build a hierarchical clustering relationship between the CRISPR repeats to score the variation across the family (Burstein et al., 2016). Surprisingly, this analysis pointed to the existence of two crRNA clusters, overlapping but distinct from the protein clades determined by the amino acid sequences (Figures 4B and 4C). While cluster 1 crRNAs correlate well with a subset of the alpha-clade proteins, and all the beta-clade-associated crRNAs are within cluster 2, the homologs with ambiguous phylogenetic relationships are split across the two clusters (Figure 4C).

Unable to easily predict Cas13a:crRNA orthogonality using bioinformatic analysis alone, we tested the extent of functional exchangeability between non-cognate crRNAs for both processing and *trans*-ssRNA target cleavage by each of the Cas13a homologs (Figures 4D–4F). For pre-crRNA processing, we found that the crRNA clusters defined by the pairwise sequence comparisons predicted their ability to be processed by their associated Cas13a proteins (Figure 4D). For example, pre-crRNAs from cluster 1 are only processed by the proteins of clade alpha and vice versa for cluster 2 and clade beta. In contrast, the protein classification is less predictive, as most of the ambiguously classified proteins could process sequences from repeat cluster 2, independent of where their repeat sequences were clustered.

Three homologs (PprCas13a, HheCas13a, and Rca13a) were pre-crRNA processing outliers with respect to their position within the crRNA clusters. HheCas13a was unable to cleave any non-cognate pre-crRNAs, and conversely, no homolog, including HheCas13a, processed the Hhe pre-crRNA. This suggests that the inability of HheCas13a to process its cognate pre-crRNA reflects not just the divergent repeat sequence (Figures 1 and S1), but also the loss of pre-crRNA processing activity within the protein. In contrast, the deviating activity of PprCas13a and RcaCas13a is explained by crRNA repeat sequence divergence. PprCas13a and RcaCas13a process their own crRNAs, yet both also process the crRNA repeat cluster 2 non-cognate sequences. The PprCas13a crRNA repeat sequence differs from the other cluster 2 crRNA repeats across the 5′ flanking region cleavage site, suggesting greater substrate flexibility for pre-crRNA processing by PprCas13a. Similarly, a distinguishing sequence feature of the RcaCas13a crRNA is an extended six-base-pair stem-loop relative to the standard five-base-pair stem-loop present in crRNAs of the rest of the family. It is worth noting that the positional substitution tolerance within each crRNA repeat for Cas13a pre-crRNA processing is consistent with our previous mutation studies of the LbuCas13a:crRNA complex and recent structural insights obtained from the LshCas13a:crRNA complex (East-Seletsky et al., 2016; Liu et al., 2017). Overall, these results suggest that the sequence of a type VI-A CRISPR repeat dictates

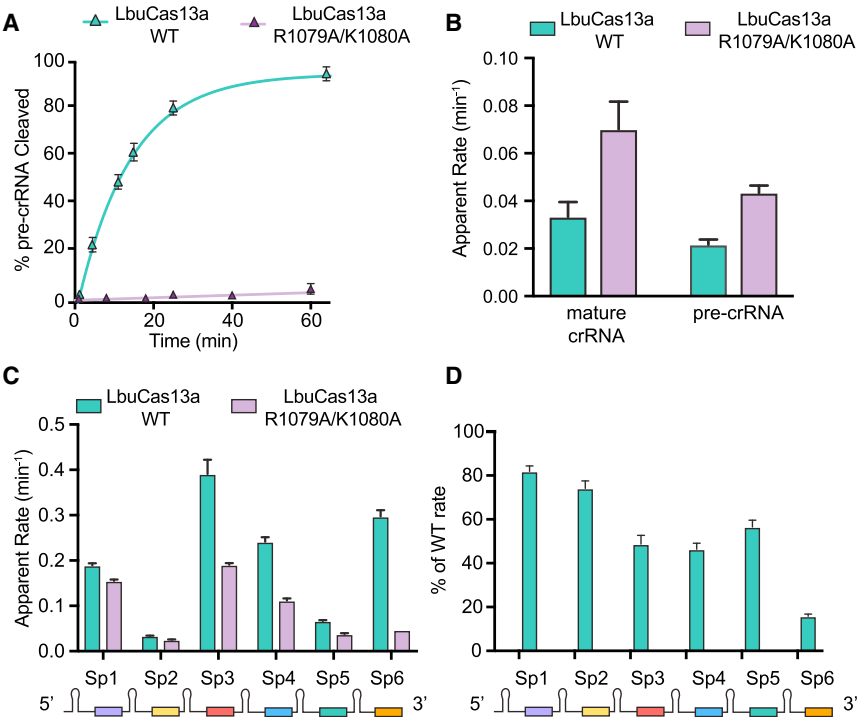


Figure 6. Deciphering the Role of crRNA Array Processing for LbuCas13a

(A) Quantified time course data of pre-crRNA processing assays for R1079A/K1080A mutant compared to WT LbuCas13a. Quantified data were fitted to single-exponential decays resulting in pseudo-first-order rate constant (k_{obs}) (mean ± SD, n = 3) for LbuCas13a WT of 0.074 ± 0.003 min^{-1}, while the R1079A/K1080A mutant could not be fit with sufficient confidence to yield a rate constant.

(B) Apparent rate of fluorescent reporter by LbuCas13a WT and R1079A/K1080A processing inactive mutant as directed by pre-crRNA and mature crRNAs. Cas13a:RNA complexes were pre-incubated for 60 min at a 1:1 ratio, and then 10 pM of activator and 150 nM reporter were added to initiate reaction (mean ± SD, n = 3).

(C) Apparent rates of fluorescent ssRNA reporter cleavage by 300 nM WT LbuCas13a or R1079A/K1080A pre-crRNA processing inactive mutant as directed by 50 nM of a CRISPR array containing six crRNA repeat-spacers. Each bar group represents the addition of 100 pM of a distinct ssRNA activator sequence complementary to schematized positions within the CRISPR array indicated below each bar group. Each rate is fitted from data from three biological replicates, and the standard deviation of the rate is depicted. Mutant protein rate is statistically different from the WT LbuCas13a for all spacer positions (one-sided t test: $p < 0.001$).

(D) Data from (C) depicted as a percentage of WT LbuCAs13a activity demonstrating the positional effect of the decreased *trans*-ssRNA targeting efficiencies by the pre-crRNA processing inactive mutant. Processing of the array is shown in Figure S5.

its capacity for pre-crRNA processing by the Cas13a family. However, homologs that evolved in the presence of divergent repeats (PprCas13a and RcaCas13a) retain the capacity to process other cluster 2 sequences.

Two Subfamilies of Functionally Orthogonal Cas13a Enzymes

Next we wondered whether the pre-crRNA processing exchangeability clusters defined in Figure 4D were competent for directing *trans*-ssRNA cleavage by non-cognate Cas13a homologs. To study Cas13a-mediated *trans*-ssRNA cleavage directed by non-cognate pre-crRNAs and mature crRNAs, we modified the previously described fluorescence assay and limited our analysis to the five homologs that exhibited significant cleavage activity in previous ssRNA-cleavage experiments (see Figure 3C). Broadly, these results mirrored the pre-crRNA processing results, with the crRNA repeat cluster identity determining functional groups, but with some striking contrasts consistent with processing and targeting being independent enzymatic activities (Figures 4E and 4F). For instance, the Ppr pre- and mature crRNA can direct ssRNA cleavage by non-cognate proteins LwaCas13a and LbuCas13a, despite their inability to process these pre-crRNAs. Another surprise is the promiscuity of HheCas13a, which is directed by all cluster 2 pre- and mature crRNAs for *trans*-ssRNA cleavage, despite lacking pre-crRNA processing activity with any of these guides. This suggests that crRNA maturation is not required for *trans*-ssRNA cleavage, an observation in agreement with our previous finding that LbuCas13a pre-crRNA processing-deficient mutants possess unaffected *trans*-ssRNA cleavage capacity (East-Seletsky et al., 2016).

Comparison of crRNA exchangeability for both pre-crRNA processing and *trans*-ssRNA cleavage defines two functionally orthogonal subfamilies within the Cas13a protein family. The first group (i.e., LbuCas13a) has wide promiscuity for both pre-crRNA processing and *trans*-ssRNA cleavage directed by crRNAs from across the protein family. This polyphyletic group also shows a preference for uridine within the *trans*-ssRNA cleavage active site. The second group (i.e., LbaCas13a), defined both by crRNA and protein sequences, has a distinct crRNA exchangeability profile and preferentially cleaves at adenosines during *trans*-ssRNA target cleavage. To test orthogonal reactivities of these Cas13a subfamilies, we first verified that LbuCas13a and LbaCas13a cleaved homo-A or homo-U ssRNA reporters, respectively (Figures 5A and 5B). While the different probes generated similar amounts of fluorescent signal, it should be noted that substantially different quantities of ssRNA activator were added due the differential sensitivities of the homologs (10 pM versus 1 nM for LbuCas13a and LbaCas13a, respectively). To verify orthogonality, a panel of control reactions with all possible non-cognate combinations between crRNA, activator, and reporter were tested, with no substantial signal detected except for the cognate combinations (Figure 5C). Taken together, these results define distinct Cas13a homologs that can function in parallel within the same system.

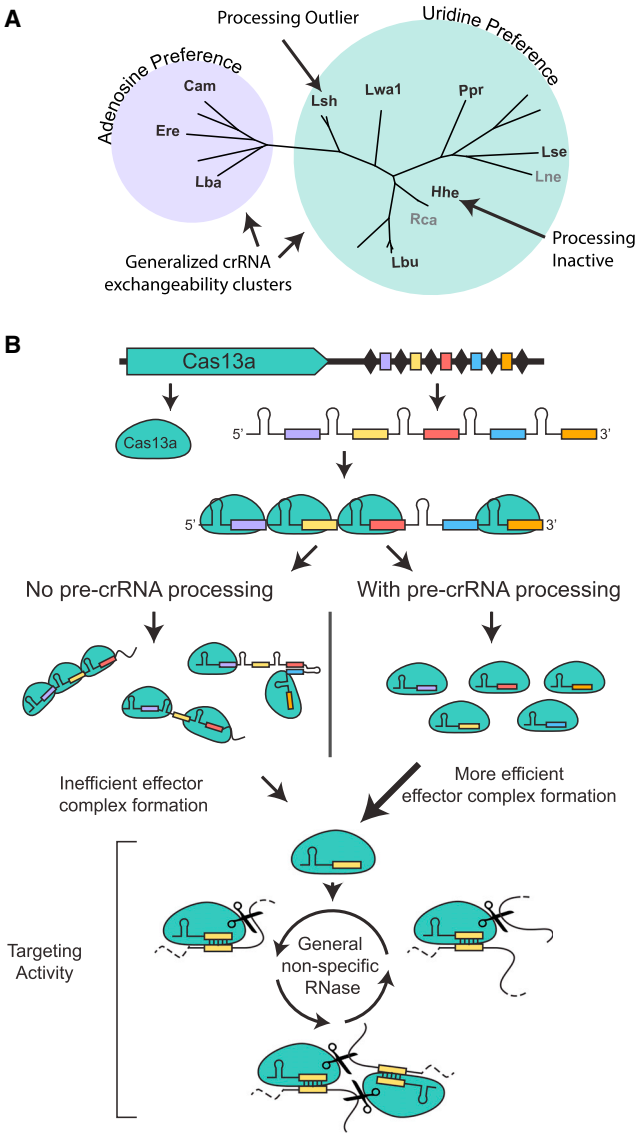


Figure 7. A Revised Model for Type VI-A CRISPR-Cas System

(A) Graphical summary of key findings in this study. Homologs used in this study are indicated with abbreviations in bold, with *trans*-ssRNA cleavage inactive homologs depicted in gray. Colored circles highlight the two orthogonal Cas13a enzyme groups, as defined by their generalized crRNA exchangeability and *trans*-ssRNA cleavage substrate nucleotide preference. (B) Revised model of type VI-A CRISPR-Cas systems incorporating the interplay between crRNA maturation and RNA targeting efficiency.

Pre-crRNA Processing Enhances Targeting Efficiencies within the Context of a CRISPR Array

One puzzling finding of this study is the lack of a stringent requirement for mature crRNA to trigger the subsequent *trans*-ssRNA target cleavage reaction by Cas13a. Additionally, processing-deficient mutants of LbuCas13a maintain similar efficiencies of *trans*-ssRNA cleavage, even when directed by a pre-crRNA instead of a mature crRNA (Figures 6 and S4). This led us to hypothesize that the role of pre-crRNA processing within type VI CRISPR loci is not necessary for ssRNA targeting but instead serves to liberate each crRNA from the confines of a long CRISPR array transcript. We wondered whether pre-crRNA processing might relieve RNA folding constraints and potential steric hindrance of neighboring Cas13a:crRNA-spacer species during crRNA loading and/or ssRNA targeting. To test this, we compared the efficiency of *trans*-ssRNA cleavage directed by a CRISPR array using either wild-type (WT) LbuCas13a or a pre-crRNA processing-inactive mutant.

Since all processing-defective single point mutants of LbuCas13a (Figure 2) retained low levels of pre-crRNA processing activity, we created a double mutant (R1079A/K1080A) that possessed no detectable processing activity, yet retained *trans*-ssRNA cleavage efficiencies similar to or greater than WT LbuCas13a (Figures 6A and 6B). Next, this mutant and the WT LbuCas13a enzyme were tested for ssRNA cleavage in the presence of a target RNA and a small CRISPR array transcript consisting of six distinct repeat-spacer units (Figures 6C and 6D). The rate of *trans*-ssRNA cleavage by the crRNA-processing inactive mutant was significantly reduced for all spacer sequences within the array (one-sided t test: $p < 0.001$ for all pairs). The reduced activity compared to WT LbuCas13a is more pronounced with each successive spacer within the array, with the last spacer directing cleavage with a rate that is only 15% of that catalyzed by the WT enzyme. This finding suggests that while pre-crRNA processing is not necessary for targeting, it enhances activity by liberating crRNAs from the CRISPR array, leading to a revised model incorporating the interplay between crRNA-processing and ssRNA-targeting within type VI-A CRISPR-Cas systems (Figure 7).

DISCUSSION

Cas13a is a unique dual ribonuclease that catalyzes two of the three steps of type VI-A CRISPR-Cas-based adaptive immunity, enabling both crRNA maturation and target RNA destruction by the same enzyme. Based on the large divergence in protein sequences and corresponding CRISPR arrays within the Cas13a enzyme family, it was bioinformatically unclear if the entire family maintained consistent functionality. Biochemically, however, analysis of ten Cas13a homologs defined two subfamilies of enzymes that recognize distinct crRNA sequences and possess different nucleotide cleavage preferences during *trans*-ssRNA degradation. These data also showed that despite being conserved in most Cas13a enzymes, pre-crRNA processing activity is not strictly required for *trans*-ssRNA cleavage. In addition, we found that the kinetics of Cas13a-catalyzed ssRNA cleavage range over seven orders of magnitude, raising the possibility that these enzymes can operate in both regulatory and cell-destruction pathways.

Beyond sharing respective sets of crRNAs, the two orthogonal subfamilies of Cas13a homologs defined in this study cleave ssRNAs preferentially at Us versus As. The evolutionary basis for these substrate preferences, which pertain to both independent active sites of the enzyme, are not yet known. It is possible that ancestral versions of these enzymes evolved in response to nucleotide composition of the host and phage transcriptomes, resulting in differing active site molecular architectures. We

anticipate that other type VI CRISPR-Cas enzymes, such as Cas13b, may also include subfamilies with divergent nucleotide preferences. Structural studies will be required to elucidate the molecular basis for *trans*-ssRNA substrate recognition to rationalize these preferences and to potentially engineer alternative cleavage activities.

The observation that pre-crRNAs can support ssRNA targeting suggests that, unlike other CRISPR-Cas systems, crRNA maturation is not essential for assembly into functional surveillance complexes. Alternative mechanisms of generating mature crRNAs have been observed (Zhang et al., 2013), but in the absence of processing, these CRISPR-Cas complexes are not competent for phage defense (Brouns et al., 2008; Deltcheva et al., 2011; Hatoum-Aslan et al., 2014; Maier et al., 2015; Semenova et al., 2015). In contrast, our data suggest that type VI systems can employ pre-crRNAs, even within the context of a CRISPR array, to trigger ssRNA degradation. Current evidence suggests that Cas12a (formerly Cpf1), the only other known dual CRISPR-Cas nuclease and a component of type V systems, also contains separate active sites for crRNA maturation and target degradation (Fonfara et al., 2016; Zetsche et al., 2017). Whether Cas12a enzymes can employ pre-crRNAs for target DNA recognition and cleavage, and whether type V systems are consistent with our revised type VI model is unknown.

The identification of two orthogonal subfamilies of RNA-guided RNA targeting enzymes could enable diverse multiplexed applications of Cas13a. Harnessing orthogonal Cas13a homologs with distinct crRNA specificities within the same application may enable RNA detection or in vivo imaging of distinct RNA species in parallel, expanding the utility of type VI-A systems. Notably, representatives of these enzyme subfamilies have not yet been found to co-exist within a single host genome. Only one strain of *Leptotrichia wadei* contains multiple type VI-A CRISPR-Cas operons, all of which appear to be from one subfamily of Cas13a.

We were surprised to find that sensitivities of ssRNA targeting by the Cas13a family vary by up to $\sim 10^7$-fold. Although the in vivo consequences of this diversity are unknown, we wonder whether type VI-A CRISPR-Cas systems orchestrate RNA-triggered responses ranging from cell death to subtle modulation of transcript levels. The recent characterization of a type VI-B CRISPR-Cas system that contains a Cas13b inhibitor within the *cas* operon provides tantalizing evidence that these systems may display much more dynamic behaviors within their respective host organisms (Smargon et al., 2017). Overall, our work demonstrates that phylogenetic classification of CRISPR proteins may conceal functional diversity within biotechnologically relevant enzyme families.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Strains
- METHOD DETAILS
  - Phylogenetic and repeat conservation analysis
  - Cas13a protein expression and purification
  - *In-vitro* RNA transcription
  - Radiolabeled ssRNA nuclease assays
  - Fluorescent ssRNA nuclease assays
  - crRNA filter-binding assays
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Radiolabeled ssRNA nuclease assays
  - Fluorescent ssRNA nuclease assays
  - Radiolabeled filter binding assays
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes five figures, six tables, and one data file and can be found with this article online at http://dx.doi.org/10.1016/j.molcel.2017.04.008.

AUTHOR CONTRIBUTIONS

Conceptualization, A.E.-S., M.R.O., and J.A.D.; Methodology, A.E.-S., M.R.O., D.B., and G.J.K.; Investigation, A.E.-S. with assistance from M.R.O.; Writing – Original Draft, A.E.-S.; Writing – Review & Editing, M.R.O., D.B., G.J.K., A.E.-S., and J.A.D.; Funding Acquisition, J.A.D.

ACKNOWLEDGMENTS

We thank the QB3 MacroLab for assistance with cloning of Cas13a constructs, Nan Ma and Kaihong Zhou for technical assistance, S.C. Knight for contributions during initial project conception, Tina Liu and Addison Wright for critical reading of the manuscript, and members of the Doudna laboratory for helpful discussions. M.R.O. is a recipient of a C.J. Martin Overseas Early Career Fellowship from the National Health and Medical Research Council of Australia. This work was supported in part by a Frontiers Science award from the Paul Allen Institute to J.A.D. and by the National Science Foundation (MCB-1244557 to J.A.D.), and J.A.D. is an Investigator of the Howard Hughes Medical Institute. J.A.D. is a co-founder of Caribou Biosciences, Inc., Editas Medicine, and Intellia Therapeutics. The authors have submitted a patent related to this work.

Received: February 21, 2017
Revised: March 21, 2017
Accepted: April 10, 2017
Published: May 4, 2017

REFERENCES

Abudayyeh, O.O., Gootenberg, J.S., Konermann, S., Joung, J., Slaymaker, I.M., Cox, D.B.T., Shmakov, S., Makarova, K.S., Semenova, E., Minakhin, L., et al. (2016). C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. Science *353*, aaf5573–aaf5573.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., and Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. Science *315*, 1709–1712.

Brouns, S.J.J., Jore, M.M., Lundgren, M., Westra, E.R., Slijkhuis, R.J.H., Snijders, A.P.L., Dickman, M.J., Makarova, K.S., Koonin, E.V., and van der Oost, J. (2008). Small CRISPR RNAs guide antiviral defense in prokaryotes. Science *321*, 960–964.

Burstein, D., Sun, C.L., Brown, C.T., Sharon, I., Anantharaman, K., Probst, A.J., Thomas, B.C., and Banfield, J.F. (2016). Major bacterial lineages are essentially devoid of CRISPR-Cas viral defence systems. Nat. Commun. *7*, 10613.

Charpentier, E., Richter, H., van der Oost, J., and White, M.F. (2015). Biogenesis pathways of RNA guides in archaeal and bacterial CRISPR-Cas adaptive immunity. FEMS Microbiol. Rev. *39*, 428–441.

Deltcheva, E., Chylinski, K., Sharma, C.M., Gonzales, K., Chao, Y., Pirzada, Z.A., Eckert, M.R., Vogel, J., and Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature *471*, 602–607.

Dy, R.L., Richter, C., Salmond, G.P.C., and Fineran, P.C. (2014). Remarkable Mechanisms in Microbes to Resist Phage Infections. Annu Rev Virol *1*, 307–331.

East-Seletsky, A., O'Connell, M.R., Knight, S.C., Burstein, D., Cate, J.H.D., Tjian, R., and Doudna, J.A. (2016). Two distinct RNase activities of CRISPR-C2c2 enable guide-RNA processing and RNA detection. Nature *538*, 270–273.

Fonfara, I., Richter, H., Bratovič, M., Le Rhun, A., and Charpentier, E. (2016). The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. Nature *532*, 517–521.

Garneau, J.E., Dupuis, M.-È., Villion, M., Romero, D.A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadán, A.H., and Moineau, S. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature *468*, 67–71.

Hatoum-Aslan, A., Maniv, I., Samai, P., and Marraffini, L.A. (2014). Genetic characterization of antiplasmid immunity through a type III-A CRISPR-Cas system. J. Bacteriol. *196*, 310–317.

Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. *30*, 772–780.

Liu, L., Li, X., Wang, J., Wang, M., Chen, P., Yin, M., Li, J., Sheng, G., and Wang, Y. (2017). Two Distant Catalytic Sites Are Responsible for C2c2 RNase Activities. Cell *168*, 121–134.e12.

Maier, L.-K., Stachler, A.-E., Saunders, S.J., Backofen, R., and Marchfelder, A. (2015). An active immune defense with a minimal CRISPR (clustered regularly interspaced short palindromic repeats) RNA and without the Cas6 protein. J. Biol. Chem. *290*, 4192–4201.

Makarova, K.S., Haft, D.H., Barrangou, R., Brouns, S.J.J., Charpentier, E., Horvath, P., Moineau, S., Mojica, F.J.M., Wolf, Y.I., Yakunin, A.F., et al. (2011). Evolution and classification of the CRISPR-Cas systems. Nat. Rev. Microbiol. *9*, 467–477.

Makarova, K.S., Wolf, Y.I., Alkhnbashi, O.S., Costa, F., Shah, S.A., Saunders, S.J., Barrangou, R., Brouns, S.J.J., Charpentier, E., Haft, D.H., et al. (2015). An updated evolutionary classification of CRISPR-Cas systems. Nat. Rev. Microbiol. *13*, 722–736.

McWilliam, H., Li, W., Uludag, M., Squizzato, S., Park, Y.M., Buso, N., Cowley, A.P., and Lopez, R. (2013). Analysis Tool Web Services from the EMBL-EBI. Nucleic Acids Res. *41*, W597–W600.

Mohanraju, P., Makarova, K.S., Zetsche, B., Zhang, F., Koonin, E.V., and van der Oost, J. (2016). Diverse evolutionary roots and mechanistic variations of the CRISPR-Cas systems. Science *353*, aad5147–aad5147.

Nickel, L., Weidenbach, K., Jäger, D., Backofen, R., Lange, S.J., Heidrich, N., and Schmitz, R.A. (2013). Two CRISPR-Cas systems in Methanosarcina mazei strain Gö1 display common processing features despite belonging to different types I and III. RNA Biol. *10*, 779–791.

Niewoehner, O., Jinek, M., and Doudna, J.A. (2014). Evolution of CRISPR RNA recognition and processing by Cas6 endonucleases. Nucleic Acids Res. *42*, 1341–1353.

Scholz, I., Lange, S.J., Hein, S., Hess, W.R., and Backofen, R. (2013). CRISPR-Cas systems in the cyanobacterium Synechocystis sp. PCC6803 exhibit distinct processing pathways involving at least two Cas6 and a Cmr2 protein. PLoS ONE *8*, e56470.

Semenova, E., Kuznedelov, K., Datsenko, K.A., Boudry, P.M., Savitskaya, E.E., Medvedeva, S., Beloglazova, N., Logacheva, M., Yakunin, A.F., and Severinov, K. (2015). The Cas6e ribonuclease is not required for interference and adaptation by the E. coli type I-E CRISPR-Cas system. Nucleic Acids Res. *43*, 6049–6061.

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.

Shmakov, S., Smargon, A., Scott, D., Cox, D., Pyzocha, N., Yan, W., Abudayyeh, O.O., Gootenberg, J.S., Makarova, K.S., Wolf, Y.I., et al. (2017). Diversity and evolution of class 2 CRISPR-Cas systems. Nat. Rev. Microbiol. *15*, 169–182.

Smargon, A.A., Cox, D.B.T., Pyzocha, N.K., Zheng, K., 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*, 618–630.e7.

Sokolowski, R.D., Graham, S., and White, M.F. (2014). Cas6 specificity and CRISPR RNA loading in a complex CRISPR-Cas system. Nucleic Acids Res. *42*, 6532–6541.

Staals, R.H.J., Zhu, Y., Taylor, D.W., Kornfeld, J.E., Sharma, K., Barendregt, A., Koehorst, J.J., Vlot, M., Neupane, N., Varossieau, K., et al. (2014). RNA targeting by the type III-A CRISPR-Cas Csm complex of Thermus thermophilus. Mol. Cell *56*, 518–530.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics *30*, 1312–1313.

Sternberg, S.H., Haurwitz, R.E., and Doudna, J.A. (2012). Mechanism of substrate selection by a highly specific CRISPR endoribonuclease. RNA *18*, 661–672.

Wright, A.V., Nuñez, J.K., and Doudna, J.A. (2016). Biology and Applications of CRISPR Systems: Harnessing Nature's Toolbox for Genome Engineering. Cell *164*, 29–44.

Zetsche, B., Heidenreich, M., Mohanraju, P., Fedorova, I., Kneppers, J., DeGennaro, E.M., Winblad, N., Choudhury, S.R., Abudayyeh, O.O., Gootenberg, J.S., et al. (2017). Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. Nat. Biotechnol. *35*, 31–34.

Zhang, J., Rouillon, C., Kerou, M., Reeks, J., Brugger, K., Graham, S., Reimann, J., Cannone, G., Liu, H., Albers, S.-V., et al. (2012). Structure and mechanism of the CMR complex for CRISPR-mediated antiviral immunity. Mol. Cell *45*, 303–313.

Zhang, Y., Heidrich, N., Ampattu, B.J., Gunderson, C.W., Seifert, H.S., Schoen, C., Vogel, J., and Sontheimer, E.J. (2013). Processing-independent CRISPR RNAs limit natural transformation in Neisseria meningitidis. Mol. Cell *50*, 488–503.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and Virus Strains		
Escherichia coli BL21(DE3)Star	QB3 Macrolab	N/A
Escherichia coli Rosetta2	QB3 Macrolab	N/A
Chemicals, Peptides, and Recombinant Proteins		
Complete-EDTA-free protease inhibitor	Sigma-Aldrich	11873580001
Tobacco Etch Virus (TEV) protease	J. Doudna Lab	N/A
Phusion Polymerase	NEB	M530L
ATP, [γ-32P]- 3000Ci/mmol	Perkin Elmer	BLU002A001MC
T4 PNK	NEB	M0236S
T7 Polymerase	J. Doudna Lab	N/A
Deposited Data		
Raw Imaging Files	This study, Mendeley Data	http://dx.doi.org/10.17632/4gpp43hpy2.1
Oligonucleotides		
ssDNA oligos (see Table S2 for sequences)	IDT	N/A
ssRNA olgios (see Table S2 for sequences)	IDT	N/A
Six-mer CRISPR array	GeneArt (ThermoFisher)	Custom order
RNaseAlert Substrate	IDT	11-04-03-03
Recombinant DNA		
p2CT-His-MBP-Lbu_C2c2_WT	East-Seletsky et al., 2016	Addgene #83482
p2CT-His-MBP-Lsh_C2c2_WT	East-Seletsky et al., 2016	Addgene #83487
p2CT-His-MBP-Lbu_C2c2_R1079A	East-Seletsky et al., 2016	Addgene #83494
p2CT-His-MBP-Lbu_C2c2_E299A	This study	Addgene #91855
p2CT-His-MBP-Lbu_C2c2_K310A	This study	Addgene #91856
p2CT-His-MBP-Lbu_C2c2_R311A	This study	Addgene #91857
p2CT-His-MBP-Lbu_C2c2_N314A	This study	Addgene #91858
p2CT-His-MBP-Lbu_C2c2_R1072A	This study	Addgene #91859
p2CT-His-MBP-Lbu_C2c2_D1078A	This study	Addgene #91860
p2CT-His-MBP-Lbu_C2c2_K1080A	This study	Addgene #91862
p2CT-His-MBP-Lbu_C2c2_K1082A	This study	Addgene #91863
p2CT-His-MBP-Lbu_C2c2_R1087A	This study	Addgene #91864
p2CT-His-MBP-Lwa_Cas13a_WT	This study	Addgene #91865
p2CT-His-MBP-Lne_Cas13a_WT	This study	Addgene #91866
p2CT-His-MBP-Lba_Cas13a_WT	This study	Addgene #91867
p2CT-His-MBP-Ere_Cas13a_WT	This study	Addgene #91868
p2CT-His-MBP-Cam_Cas13a_WT	This study	Addgene #91869
p2CT-His-MBP-Rca_Cas13a_WT	This study	Addgene #91870
p2CT-His-MBP-Hhe_Cas13a_WT	This study	Addgene #91871
p2CT-His-MBP-Ppr_Cas13a_WT	This study	Addgene #91872
Software and Algorithms		
R Studio- Open source edition	Rstudio	https://www.rstudio.com/
Prism 7 for Mac	GraphPad Software	https://www.graphpad.com/scientific-software/prism/
ImageQuant TL	GE Healthcare	N/A
EMBOSS Needle	EMBL–EBI	http://www.ebi.ac.uk/Tools/psa/emboss_needle/nucleotide.html

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for reagents should be directed to Lead Contact Jennifer Doudna (doudna@berkeley.edu). All plasmids are deposited with Addgene.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Strains

Expression vectors were transformed into Rosetta2 DE3 or BL21 *E. coli* cells grown in 2xYT broth at 37°C, induced at mid-log phase with 0.5 mM IPTG, and then transferred to 16°C for overnight expression.

METHOD DETAILS

Phylogenetic and repeat conservation analysis

Cas13a maximum-likelihood phylogenies were computed as previously described (East-Seletsky et al., 2016) using RAxML (Stamatakis, 2014) with the PROTGAMMALG evolutionary model and 100 bootstraps. Protein clades alpha and beta were defined as branch points with bootstrap values greater than 90, suggesting high confidence in having a common ancestor. The remaining proteins were labeled as ambiguous ancestry, as the phylogenetic relationships between them were low confidence, reflected in bootstrap values less than 90. Sequences were aligned by MAFFT with the 'einsi' method (Katoh and Standley, 2013). Candidates were selected to represent each of the major branches of the Cas13a protein tree. Alignments were performed for all non-redundant homologs (Figure S2 and Supporting File 1) and candidate proteins. Comparison of the CRISPR-RNA (crRNA) repeats was carried out by calculating pairwise similarity scores using the Needleman-Wunsch algorithm through the Needle tool on EMBL-EBI (McWilliam et al., 2013). Hierarchical clustering of CRISPR crRNA was performed in R using the similarity score matrix.

Cas13a protein expression and purification

Expression vectors for protein purification were assembled using synthetic gBlocks ordered from Integrated DNA Technologies (IDT). The codon-optimized Cas13a genomic sequences were *N*-terminally tagged with a His_6-MBP-TEV cleavage site sequence, with expression driven by a T7 promoter. Mutant proteins were cloned via site-directed mutagenesis of WT Cas13a constructs. Purification of all homologs was based off of a previously published protocol (East-Seletsky et al., 2016). Briefly, expression vectors were transformed into Rosetta2 DE3 or BL21 *E. coli* cells grown in 2xYT broth at 37°C, induced at mid-log phase with 0.5 mM IPTG, and then transferred to 16°C for overnight expression. Cell pellets were resuspended in lysis buffer (50 mM Tris-Cl pH 7.0, 500 mM NaCl, 5% glycerol, 1 mM TCEP, 0.5 mM PMSF, and EDTA-free protease inhibitor (Roche)), lysed by sonication, and clarified by centrifugation at 15,000 g. Soluble His_6-MBP-TEV-Cas13a was isolated over metal ion affinity chromatography, and in order to cleave off the His_6-MBP tag, the protein-containing eluate was incubated with TEV protease at 4°C overnight while dialyzing into ion exchange buffer (50 mM Tris-Cl pH 7.0, 250 mM KCl, 5% glycerol, 1 mM TCEP). Cleaved protein was loaded onto a HiTrap SP column (GE Healthcare) and eluted over a linear KCl (0.25-1.5M) gradient. Cas13a containing fractions were pooled, concentrated, and further purified via size-exclusion chromatography on a S200 column (GE Healthcare) in gel filtration buffer (20 mM Tris-Cl pH 7.0, 200 mM KCl, 5% glycerol, 1 mM TCEP) and were subsequently stored at −80°C. All homologs were purified using this protocol except LwaCas13a which was bound to a HiTrap Heparin column instead of a SP column, and the size-exclusion chromatography step was omitted due to sufficient purity of the sample post ion-exchange. All expression plasmids are deposited with Addgene.

In-vitro RNA transcription

All pre-crRNAs, mature crRNAs, and targets were transcribed in vitro using previously described methods (East-Seletsky et al., 2016; Sternberg et al., 2012). Briefly, all substrates were transcribed off a single-stranded DNA oligonucleotide template (IDT), except for mature crRNAs requiring a non-GR 5′ terminus. For these mature crRNAs, T7 polymerase templates containing a Hammerhead Ribozyme sequence immediately upstream of the mature crRNA sequence were generated using overlap PCR, and then purified for use as the template for T7 transcription (see Table S2 for sequences). Lbu six-mer CRISPR array in vitro transcription template was synthesized by GeneArt (Thermofisher) as a plasmid. The T7 promoter-CRISPR array region was PCR amplified and purified prior to use as the template for T7 transcription. All transcribed RNAs were purified using 15% Urea-PAGE, except for the array which was purified using 6% Urea-PAGE. All RNAs were subsequently treated with calf alkaline phosphatase to remove 5′ phosphates. Radiolabeling was performed as previously described (East-Seletsky et al., 2016; Sternberg et al., 2012). A, C, G and U homopolymers, and fluorescently-labeled RNA reporters for *trans*-ssRNA cleavage were synthesized by IDT. Homopolymers were purified using 25% Urea-PAGE after radiolabeling to reduce substrate heterogeneity.

Radiolabeled ssRNA nuclease assays

pre-crRNA processing assays were performed at 37°C in RNA processing buffer (20 mM HEPES-Na pH 6.8, 50 mM KCl, 5 mM $MgCl_2$, 10 μg/mL BSA, 100 μg/mL tRNA, 0.01% Igepal CA-630 and 5% glycerol) with a 100-fold molar excess of Cas13a relative to 5′-labeled pre-crRNA (final concentrations of 100 nM and < 1 nM, respectively). Unless otherwise indicated, reactions were

quenched after 60 min with 1.5X RNA loading dye (100% formamide, 0.025% bromophenol blue, and 200 μg mL^{-1} heparin). After quenching, reactions were denatured at 95°C for 5 min prior to resolving by 15% denaturing PAGE (0.5x TBE buffer). Target cleavage assays were performed at 37°C in cleavage buffer (20 mM HEPES-Na pH 6.8, 50 mM KCl, 5 mM $MgCl_2$, and 5% glycerol). Generally, Cas13a:crRNA complex formation was performed in cleavage buffer, at a molar ratio of 2:1 protein to crRNA at 37°C for 60 min, prior to adding 5′-labeled target and/or other non-radiolabeled RNA target substrates. Unless otherwise indicated, final concentrations were 100 nM Cas13a, 50 nM crRNA or pre-crRNA, 50 nM crRNA-complementary target ssRNA (henceforth referred to as 'activator') and < 1 nM *trans*-ssRNA target.

Fluorescent ssRNA nuclease assays

Cas13a:crRNA complexes were assembled in cleavage buffer, as described above. 150 nM of RNase Alert reporter (IDT) and various final concentrations (0-1 μM) of ssRNA-activator were added to initiate the reaction. Notably these reactions are in the absence of competitor tRNA or total RNA compared to our previous work (East-Seletsky et al., 2016) to more accurately measure *trans*-cleavage activity. These reactions were incubated in a fluorescence plate reader (Tecan Infinite Pro F2000) for up to 120 min at 37°C with fluorescence measurements taken every 5 min (λ_{ex}: 485 nm; λ_{em}: 535 nm). For fluorescent homopolymer ssRNA reporter studies, Cas13a:crRNA complexes were pre-incubated at 37°C for 60 min using standard conditions. Activator ssRNA and 200 nM fluorescent ssRNA reporter were added to initiate the reaction immediately before placing reaction in plate reader. For Lbu- and Lba-Cas13a containing samples with fluorescent homopolymer ssRNA reporters, 10 pM and 1 nM activator was used, respectively. For pre-crRNA array experiments, 300 nM Cas13a was first incubated with 50 nM pre-crRNA array for 1 hr in cleavage buffer to enable binding and processing of the array. 100 pM of each ssRNA activator was added along with 150 nM of RNase Alert reporter (IDT) to initiate the reaction, in biological triplicate for each spacer sequence. Apparent rates were calculated using one single-exponential decay using Prism7 (GraphPad) and calculated rates are plotted with their associated standard deviations.

crRNA filter-binding assays

Filter binding assays were carried out as previously described (East-Seletsky et al., 2016). Briefly Cas13a and radiolabeled crRNA were incubated for 1 hr at 37°C in RNA processing buffer (20 mM HEPES-Na pH 6.8, 50 mM KCl, 5 mM $MgCl_2$, 10 μg/mL BSA, 100 μg/mL yeast tRNA, 0.01% Igepal CA-630 and 5% glycerol). Tufryn, Protran and Hybond-N+ were assembled onto a dot-blot apparatus in the order listed above. The membranes were washed twice with 50 μL Equilibration Buffer (20 mM HEPES-Na pH 6.8, 50 mM KCl, 5 mM $MgCl_2$ and 5% glycerol) before the sample was applied to the membranes. Membranes were again washed with 50 μL Equilibration Buffer, dried and visualized by phosphorimaging.

QUANTIFICATION AND STATISTICAL ANALYSIS

Radiolabeled ssRNA nuclease assays

All bands were visualized by phosphorimaging (Typhoon, GE Healthcare) and quantified with ImageQuant (GE Healthcare). For pre-crRNA processing, the percent cleavage was determined as the ratio of the product band intensity to the total intensity of both the product and uncleaved pre-crRNA bands and normalized for background within each measured substrate. For *trans*-ssRNA cleavage reactions, the percentage cleavage was determined as the ratio of all fragments smaller than the target to the total intensity within the lane and normalized for background within each substrate. These data were subsequently fit to a single-exponential decay using Prism7 (GraphPad) and cleavage rates are reported in figure legends.

Fluorescent ssRNA nuclease assays

Background-corrected fluorescence values were obtained by subtracting fluorescence values obtained from reactions carried out in the absence of target ssRNA activator. For determining homolog sensitivities and array processing effects, background corrected values were fit to a single-exponential decay using Prism7 (GraphPad) and the calculated rates were plotted with their associated standard deviations from n = 3. For comparing non-cognate crRNA directed *trans*-ssRNA cleavage, initial reaction rates were instead calculated due to discrepancies in fluorescence plateau values across the dataset. Rates were then scaled relative to the cognate crRNA to normalize rates across the homologs. See Tables S4-6 for normalized values.

Radiolabeled filter binding assays

Data were quantified with ImageQuant TL Software (GE Healthcare) and fit to a binding isotherm using Prism (GraphPad Software). All experiments were carried out in triplicate. Dissociation constants and associated errors are reported in the Figure legends. See also Data S1.

DATA AND SOFTWARE AVAILABILITY

Raw phosphorimaging files and alignment file for Cas13a family are deposited on Mendeley Data (http://dx.doi.org/10.17632/4gpp43hpy2.1).