Article

A Variety of Dicer Substrates in Human and *C. elegans*

Agnieszka Rybak-Wolf,[1,3] Marvin Jens,[1,3] Yasuhiro Murakawa,[2,3] Margareta Herzog,[1] Markus Landthaler,[2,*] and Nikolaus Rajewsky[1,*]
[1]Laboratory for Systems Biology of Gene Regulatory Elements
[2]Laboratory for RNA Biology and Posttranscriptional Regulation
Berlin Institute for Medical Systems Biology, Max-Delbrück-Center for Molecular Medicine, Robert-Rössle-Strasse 10, 13125 Berlin, Germany
[3]Co-first author
*Correspondence: markus.landthaler@mdc-berlin.de (M.L.), rajewsky@mdc-berlin.de (N.R.)
http://dx.doi.org/10.1016/j.cell.2014.10.040

SUMMARY

The endoribonuclease Dicer is known for its central role in the biogenesis of eukaryotic small RNAs/microRNAs. Despite its importance, Dicer target transcripts have not been directly mapped. Here, we apply biochemical methods to human cells and *C. elegans* and identify thousands of Dicer-binding sites. We find known and hundreds of additional miRNAs with high sensitivity and specificity. We also report structural RNAs, promoter RNAs, and mitochondrial transcripts as Dicer targets. Interestingly, most Dicer-binding sites reside on mRNAs/lncRNAs and are not significantly processed into small RNAs. These passive sites typically harbor small, Dicer-bound hairpins within intact transcripts and generally stabilize target expression. We show that passive sites can sequester Dicer and reduce microRNA expression. mRNAs with passive sites were in human and worm significantly associated with processing-body/granule function. Together, we provide the first transcriptome-wide map of Dicer targets and suggest conserved binding modes and functions outside of the miRNA pathway.

INTRODUCTION

Genes are subject to posttranscriptional regulation by small RNAs (sRNA) and RNA-binding proteins (RBPs). Over the past years, many regulatory sRNAs have been discovered (Bartel, 2009; Kim et al., 2009). Most miRNAs are generated from primary transcripts that undergo two distinct steps of processing. First, DROSHA and its partner DGCR8 release ~70 nt stem-loop precursor miRNAs (pre-miRNAs) inside of the nucleus (Lee et al., 2003). Alternatively, "mirtrons" are derived from introns in a splicing-machinery-dependent and Drosha-independent fashion (Okamura et al., 2007; Ruby et al., 2007). The pre-miRNA hairpins are further cut by Dicer (Hutvágner et al., 2001). The physical distance between the Dicer PAZ and 5′ pocket and the RNase III domains functions as a molecular "ruler" to control the product size

(Lau et al., 2012; Park et al., 2011). The resulting ~22 nt miRNA/miRNA* duplexes have a 2 nt overhang at the 3′ ends, which is a characteristic signature of Dicer processing (Bartel, 2009). The duplexes are subsequently handed over to Argonaute (AGO) proteins, and one of the two strands is selected as the mature miRNA to form active RNA-induced silencing complex (RISC), which represses the expression of target genes (Bartel, 2009).

Dicer binding sites are typically indirectly inferred by mapping sRNA sequencing data to transcripts. Two adjacent read stacks suggest a pre-miRNA that was cleaved by Dicer. A more variable 3′ than 5′ end is also typical for miRNAs (Chiang et al., 2010). By computationally scoring these and other hallmarks of Dicer processing, Dicer targets are inferred (Friedländer et al., 2012). These methods invariably make assumptions about how Dicer binds and processes substrates.

sRNA sequencing revealed diverse sRNA classes, including miRNAs that are processed from tRNAs and snoRNAs (Castellano and Stebbing, 2013; Ender et al., 2008; Friedländer et al., 2014). Currently, more than 1,800 pre-miRNAs in human and 200 in *C. elegans* are listed in miRBase (Kozomara and Griffiths-Jones, 2014). However, it remains challenging to distinguish miRNAs from fragments of other transcripts based on sRNA-sequencing data alone (Kozomara and Griffiths-Jones, 2014). There are also other types of sRNAs such as endogenous short interfering RNAs (endo-siRNAs) originating from long double-stranded RNAs (dsRNAs) generated by transposable elements or convergent transcription units (Kim et al., 2009). Also sRNAs from transcript starts and ends have been described (Seila et al., 2008; Valen et al., 2011; Zamudio et al., 2014). Recent findings highlight that primary sequence, RNA structure, size, and position of stem and loop, the accessibility of 3′/5′ ends, and cofactors are all relevant for Dicer substrate recognition and/or cleavage (Feng et al., 2012; Fukunaga et al., 2012; Gu et al., 2012; Lau et al., 2012; Park et al., 2011; Tian et al., 2014). Furthermore, in vitro studies (Feng et al., 2012) and electron microscopy (Taylor et al., 2013) show that affinity and cleavage efficiency can be uncoupled properties of the bound substrate. In *C. elegans*, Dicer binds the lncRNA *rncs-1* in vitro but does not process it to sRNA, leading to sequestration of Dicer and inhibition of its function (Hellwig and Bass, 2008).

Other reports likewise point to functions of Dicer that are not connected to sRNA production. In human cells, Dicer is present
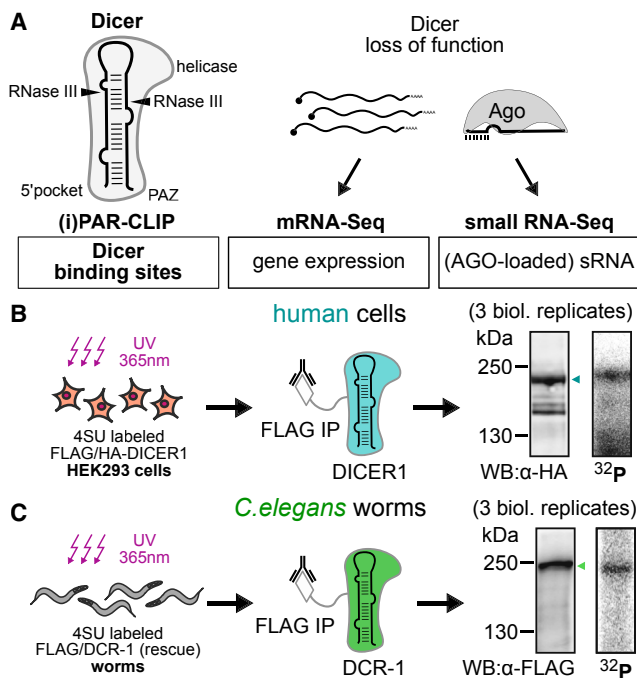
Cell

Figure 1. PAR-CLIP Identifies Transcriptome-wide Binding of Dicer
(A) Outline of experimental setup. PAR-CLIP/iPAR-CLIP of Dicer was performed in HEK293 cells and *C. elegans*. Expressed sRNA was profiled by 5′P (*C. elegans*) or AGO2/3-IP (human) deep sequencing. The impact of Dicer and Drosha knockdowns on mRNA levels was monitored by mRNA-seq, on sRNA expression by sRNA-seq.
(B and C) Overview of PAR-CLIP experiments in HEK293 cells (B) and in *C. elegans* (C). The right panel shows the phospho-image of SDS-PAGE-resolved, radiolabeled Dicer-RNA complex, immunoprecipitated from 4SU-labeled and crosslinked HEK293 cells (B) or worms (C). Arrows indicate the specific band of RNAs crosslinked to Dicer protein, confirmed by western blot (left).

in the nucleus (Doyle et al., 2013; Sinkkonen et al., 2010) and interacts with nuclear pore complexes (Ando et al., 2011). In *C. elegans*, Dicer localizes to germline P granules and is required for their formation (Beshore et al., 2011). Dicer is also present in chromatoid bodies of mammalian germ cells (Kotaja et al., 2006). If Dicer binding and RNA cleavage are largely uncoupled, many interactions may be undetectable with current methods.

Here, we present an in-depth identification and analysis of direct Dicer-binding sites in human embryonic kidney (HEK) 293 cells and *C. elegans* by applying PAR-CLIP/iPAR-CLIP (Hafner et al., 2010; Jungkamp et al., 2011) (Figure 1). Because mammals and nematodes both have a single full-length Dicer gene, assumed to process long dsRNAs and miRNAs, we were able to compare human and worm modes of Dicer targeting, highlighting core functions conserved during 550 million years of evolution. To measure the enzymatic activity of Dicer on its binding sites, we performed sRNA sequencing. In human cells, we deeply sequenced AGO-loaded sRNAs, enabling us to identify hundreds of additional miRNAs.

We also identified in both human and worm thousands of stably bound Dicer target sites that were not processed into sRNAs

("passive" sites). In *C. elegans*, the top passively bound transcript was the lncRNA *rncs-1*, confirming previous in vitro data (Hellwig and Bass, 2008) and indicating that passive sites can function by sequestering Dicer. Dicer knockdown experiments in both human cells and worms suggest that passive sites, in contrast to sites that emit sRNAs, are generally stabilized by Dicer. Furthermore, we systematically compared passive Dicer targets across human and worm and discovered that messages encoding P body/granule components are statistically significantly enriched in both species.

In summary, our data reveal hundreds of additional miRNAs, non-canonical substrates for dicing, and a novel, passive mode of Dicer binding. Moreover, we identify a large number of stably bound "passive" sites inside of many mRNAs in both human and worm. We present multiple lines of evidence that these sites are bound by Dicer nearby the loop of small hairpin structures within the intact host RNA, generally stabilizing its expression. We discuss several scenarios for the function of passive sites.

RESULTS

PAR-CLIP Identifies Transcriptome-wide Binding of Dicer in Human Cells and *C. elegans*
To identify direct in vivo Dicer targets transcriptome wide, we performed PAR-CLIP (Hafner et al., 2010) in human HEK293 cells and iPAR-CLIP (Jungkamp et al., 2011) in young adult *C. elegans* worms (Figures 1B and 1C). We used HEK293 cells expressing FLAG/HA-tagged DICER1 protein (Figure S1A available online), whereas in *C. elegans*, we used a 3×FLAG/10×His-tagged DCR-1 rescue strain (Figure S1B). For HEK293 cells and worm, we combined three independent replicates each (see Figures S1C–S1F for reproducibility). After computational quality filtering, we identified ∼8,500 and ∼2,500 reproduced Dicer-binding sites in HEK293 cells and *C. elegans*, respectively. The crosslinking introduced specific nucleotide mutations and was not biased toward highly expressed genes (Figures S1G and S1H).

Dicer PAR-CLIP Recovers Known miRNAs
We first investigated canonical miRNAs, exemplified by the human "oncomir" cluster miR-17–92 (Figure 2A) and the miR-35–41 cluster in *C. elegans* (Figure 2B). All PAR-CLIP experiments reproducibly mapped Dicer binding (Figure 2A and 2B). PAR-CLIP reads and crosslink-induced T-to-C mutations cover the entire pre-miRNA stem-loop structure. AGO-bound sRNA reads align to both sides of the stem and form precise double stacks, representing the "mature" and "star" miRNA. Additional ∼11 nt of base-paired RNA beyond the pre-miRNA reflect the requirements for Drosha processing (Han et al., 2006) of miR-17 and miR-40 (Figures 2A and 2B, inserts).

In contrast to sRNA sequencing, Dicer PAR-CLIP abundantly yielded reads spanning the loop region of miRNAs. This was further enhanced when the RNase T1 treatment was omitted (Figure S2A and S2B). Thus, our PAR-CLIP assay faithfully recovers Dicer binding to intact miRNA precursors.

To our surprise, Dicer binds not only pre-miRNAs. We mapped Dicer-binding sites to GENCODE, WormBase in *C. elegans*, and miRBase annotations. Only 321 binding sites map to 316 individual miRNAs in HEK293 cells and 100 binding

sites to 92 miRNAs in *C. elegans*. The majority of Dicer-binding sites map to a wide range of RNA transcripts, including not only other structural RNAs, such as tRNAs and snoRNAs, but also exonic and intronic mRNA regions (Figures 2C and 2D and Table S1). This is also reflected in the distribution of PAR-CLIP reads (Figures S2C and S2D).

Dicer-Binding Sites Explain the Majority of Expressed sRNA

We asked how many of the expressed sRNAs map to Dicer-binding sites. We sequenced AGO-loaded sRNAs in HEK293 cells by immunoprecipitation of stably expressed FLAG/HA-tagged AGO2 and AGO3. The two data sets were highly correlated (Figure S2E) and subsequently merged (referred to as AGOIP). AGOIP read counts were also highly correlated to normal 5′-monophosphate-specific (5′-P) sRNA sequencing, indicating that the majority of expressed sRNAs are indeed loaded into AGO (Figure S2F). Of the aggregate number of AGOIP reads from 18–26 nt length, 95% originate from human Dicer-binding sites, comparable to the 93% explained by miR-Base (Figure 2E).

C. elegans expresses at least 27 Argonautes (Gu et al., 2009). We therefore relied on conventional sRNA sequencing. Here, 61% of 18–26 nt reads originate from nematode Dicer-binding sites, comparable to 60% explained by miRBase (Figure 2F). The previously published PAR-CLIP data sets of single-stranded RNA-binding proteins, human ELAVL1 (Lebedeva et al., 2011) and *C.elegans* GLD-1 (Jungkamp et al., 2011), were analyzed in the same manner as negative controls and explained less than 1% of the expressed sRNA. Worm-specific, triphosphory-lated endo-siRNAs were sequenced independently and were found to be largely unassociated with Dicer-binding sites, consistent with previous results on their Dicer-independent production (Gu et al., 2009) (Figure S2G).

Dicer PAR-CLIP Identifies sRNA-Generating Loci with High Sensitivity

Because the pool of expressed sRNAs can be dominated by a few highly expressed miRNAs, we investigated the sensitivity of the Dicer PAR-CLIP as a function of sRNA expression. We considered all clusters of AGOIP reads that uniquely align to the genome as potential sites of sRNA production and selected those with a given minimal sRNA output. We then asked what fraction of these stacks overlap Dicer-binding sites, known miRNAs, or control sites (Figures 2G and 2H). Although Dicer-binding sites more often overlap with regions of low and interme-diate sRNA expression than the miRBase annotation, the gap disappears for highly expressed sRNAs. Thus, although compa-rable fractions of total sRNA can be explained by either miRBase or Dicer binding, Dicer PAR-CLIP offers enhanced sensitivity for lowly expressed sRNAs.

We conclude that Dicer PAR-CLIP identifies loci of sRNA pro-duction with high sensitivity and that the majority of AGO-loaded or 5′-P sRNAs (at least 95% in HEK cells and at least 61% in *C. elegans*) are probably Dicer products. Furthermore, the con-trol data sets for ELAVL1 and GLD-1 demonstrate that our anal-ysis of HEK293 and worm PAR-CLIP data has a very low rate (<1%) of falsely reporting sites that generate sRNAs.

Dicer PAR-CLIP Discovers Hundreds of Additional miRNAs

Next, we searched for previously unidentified miRNAs in the Dicer PAR-CLIP data. We identified 1,978 human DICER1-binding sites with AGOIP read counts above background level (Figure S3A). Of these, 1,678 are not listed in miRBase (Figure S3A). We ran miRDeep2 (Friedländer et al., 2012) to predict additional miRNAs from the AGOIP data independently and subsequently inter-sected these with DICER1-binding sites. This analysis resulted in 212 (not present in miRBase) miRNA candidates supported by Dicer PAR-CLIP (Table S2). However, we also found many Dicer-bound loci with a clean "double-stack" profile of AGOIP reads, the signature of processed pre-miRNAs, which were only called by miRDeep2 after deactivating the hairpin scoring (Table S2). This scoring is meant to assess how likely the hairpin fold of a sequence occurred by chance, which reduces false positives but also sensitivity. We successfully validated Dicer-dependent pro-cessing for three out of three such miRNAs by an in vitro process-ing assay (Figure S3B), demonstrating that the PAR-CLIP evidence for in vivo interactions allows us to relax the miRDeep2 filtering. In total, we report 367 new miRNA candidates predomi-nantly originating from introns (like known miRNAs), but also from a variety of other sources including 5′ UTRs (Figures 3A and 3B).

In *C. elegans*, only two sites were identified as putative addi-tional miRNAs (Figures S3C and S3D), which is not surprising considering the very deep sRNA profiling and smaller genome (Shi et al., 2013).

Additionally Identified miRNAs Are Dicer Dependent and Interact with mRNAs

To validate the miRNA candidates, not present in miRBase, we compared sRNA read counts from control cells to Dicer- or Dro-sha-depleted cells (Figures S3E–S3G). Our 5′-P sequencing data allowed us to quantify expression changes of 52 miRNA candi-dates, which were specifically and highly significantly downregu-lated upon Dicer knockdown (Figures 3C, S3H, and S3I).

Next, we screened publicly available AGO CLIP data for chimeric reads (Grosswendt et al., 2014) that would support direct miRNA:target interactions. For 20 of our miRNA candi-dates, chimeric reads with target 3′ UTRs were found (Table S2). Although some indicate binding through seed complemen-tarity, others show noncanonical miRNA:target interactions. For example, the miRNA derived from the 5′ UTR of glutamate-ammonia ligase (GLUL) appears to bind to the 3′ UTR of methyl-transferase-like 8 (METTL8) by base-pairing within the central region of the miRNA (Figure 3E).

We further validated four out of four candidates by northern analysis. Dicer dependence is reflected either by reduction of mature miRNA or by accumulation of pre-miRNA in Dicer-depleted cells (Figure 3D). Finally, we employed a luciferase reporter with perfect matches to miRNA candidates. By this assay, three out of six candidates can regulate targets in vivo (Figure 3F). In sum-mary, multiple independent assays support expression, Dicer de-pendency, and functionality of the additionally identified miRNAs.

Dicer Binds and Cleaves Structural RNAs

Known and additional miRNAs still account for only a fraction of the Dicer targets with sRNA output. We call these sites "active"
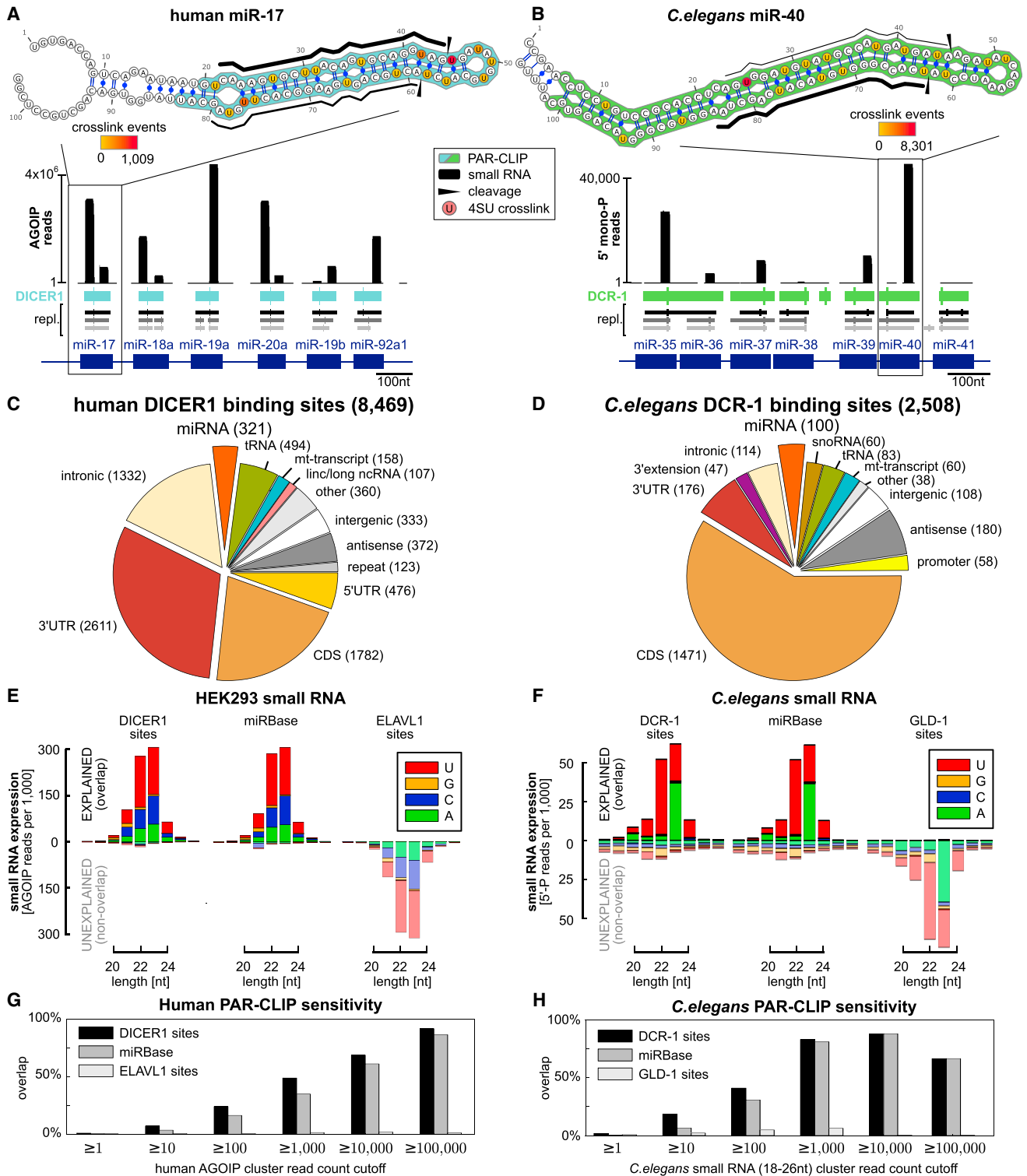
A human miR-17

crosslink events
0   1,009

4×10⁶
AGOIP reads
1

DICER1
repl.

miR-17 miR-18a miR-19a miR-20a miR-19b miR-92a1

100nt

PAR-CLIP
small RNA
cleavage
4SU crosslink

B *C.elegans* miR-40

crosslink events
0   8,301

40,000
5' mono-P reads
1

DCR-1
repl.

miR-35 miR-36 miR-37 miR-38 miR-39 miR-40 miR-41

100nt

C human DICER1 binding sites (8,469)

miRNA (321)
tRNA (494)
mt-transcript (158)
linc/long ncRNA (107)
other (360)
intronic (1332)
intergenic (333)
antisense (372)
repeat (123)
5'UTR (476)
3'UTR (2611)
CDS (1782)

D *C.elegans* DCR-1 binding sites (2,508)

miRNA (100)
snoRNA (60)
intronic (114)
tRNA (83)
3'extension (47)
mt-transcript (60)
3'UTR (176)
other (38)
intergenic (108)
antisense (180)
promoter (58)
CDS (1471)

E HEK293 small RNA

small RNA expression [AGOIP reads per 1,000]

DICER1 sites   miRBase   ELAVL1 sites

EXPLAINED (overlap)
300
150
0
150
300
UNEXPLAINED (non-overlap)

U
G
C
A

20 22 24
length [nt]

F *C.elegans* small RNA

small RNA expression [5'-P reads per 1,000]

DCR-1 sites   miRBase   GLD-1 sites

EXPLAINED (overlap)
50
25
0
25
50
UNEXPLAINED (non-overlap)

U
G
C
A

20 22 24
length [nt]

G Human PAR-CLIP sensitivity

100%
overlap
50%
0%

DICER1 sites
miRBase
ELAVL1 sites

≥1 ≥10 ≥100 ≥1,000 ≥10,000 ≥100,000
human AGOIP cluster read count cutoff

H *C.elegans* PAR-CLIP sensitivity

100%
overlap
50%
0%

DCR-1 sites
miRBase
GLD-1 sites

≥1 ≥10 ≥100 ≥1,000 ≥10,000 ≥100,000
C.elegans small RNA (18-26nt) cluster read count cutoff

Figure 2. PAR-CLIP Recovers Known miRNAs and Explains the Majority of Expressed sRNA
(A and B) Dicer PAR-CLIP data for miRNA clusters: human mir-17–92 (A) and *C.elegans* mir-35–41 (B). PAR-CLIP-binding sites from three independent experiments are indicated as gray boxes. Consensus clusters in blue (human) or green (nematode). Human AGO2/3-IP sRNA reads and *C. elegans* 5′-P sRNA read coverage in black. The top insert shows details of the folded pre-miRNA secondary structure. PAR-CLIP read coverage in blue (human) and green (*C. elegans*). Crosslinks in yellow to red, proportional to induced nucleotide conversion frequency. sRNA coverage as black lines, with width indicating expression level.

(legend continued on next page)

(at least 10 AGOIP or sRNA reads) or "highly active" (human: >1,000 AGOIP reads, *C. elegans*: >100 sRNA reads) (Figures S4A and S4B). Figure 4A shows active Dicer binding to the H/ACA box snoRNA SNORA33, the 3′ splice site of an alternative exon of Drosha mRNA, and the 5′ UTR of glutamate-ammonia ligase (GLUL) (Figure S4C shows greater detail). All of these sites give rise to ample sRNAs, comparable to a medium-expressed miRNA. The *C. elegans* examples of atypical substrates include the C/D box snoRNA F30H5.4 and Y-RNA *yrn-1* (Figure S4D), where cleavage may generate different RNA fragments of ∼50 nt length. Cleavage appears to frequently occur only on one side of the stem, suggesting that Dicer may primarily participate in its degradation. Expression of Dicer-bound structural RNAs indeed increased upon Dicer depletion, as shown by qPCR for tRNAs, snoRNAs, vault RNAs, or Y-RNAs in human (Figure S4E) and *C.elegans* (Figures S4F and S4G), independent of whether sRNA profiles match a clean double-stack profile that would resemble pre-miRNA-like processing.

Many Abundant sRNAs Are Processed with Low Precision

To study sRNA production for these substrates, we first computationally scanned binding sites for hairpin structures that could explain Dicer binding. These structures set the frame of reference to count sRNAs that align to both hairpin arms, to assess possible preferences for 5′ or 3′ arms, and to score the fidelity of sRNA processing by recording 5′ variability of read stacks (Figure 4B and Experimental Procedures). The in silico hairpin folds are supported by PARS data, which captures secondary structures in vivo (Wan et al., 2014; Figure S4H).

The resulting picture reveals a large dynamic range (six orders of magnitude) of sRNA expression generated by diverse substrates (Figures 4C and S4I). The most abundant sRNAs originate from known miRNAs, whereas Dicer binding to mRNAs, on average (but with interesting exceptions like the 5′ UTRs of GLUL or FLCN), gives rise to very little or no detectable sRNA. Expression of sRNAs from mRNAs did not correlate with mRNA expression (Figure S4J).

As reported before (Chiang et al., 2010), the more abundant sRNA preferentially arises from the 5′ arm of the hairpin. This holds for known miRNAs and becomes more pronounced for the entire set of active sites (Figure 4D). Interestingly, expression level of sRNAs does not automatically correlate with processing fidelity (Figure 4E). Although known miRNA precursors are not only abundant, but also cut with the highest precision, the additional miRNAs and a subset of bound tRNAs and snoRNAs identified here are processed with comparable precision despite lower sRNA expression. On the other hand, abundant sRNAs from mitochondria are apparently produced with very low precision (Figures 4F and S4N).

There are examples of Dicer-bound structural RNAs, mostly tRNAs, which show clean stacks of AGO-loaded RNAs, but additional and more diverse sRNA profiles in 5′ P sequencing (Figure 4F). Furthermore, upon Dicer knockdown, the sRNA output from most tRNAs is not substantially reduced (Figure S4K), whereas mitochondrial sRNAs even increase in abundance (Figures 4G and S4K). Taken together, Dicer binding to structural RNAs influences RNA stability, but concomitant sRNA production does not automatically indicate that dicing is also the source of the observed sRNA. Rather, abundant RNA species may be subject to multiple, different mechanisms of decay, including hydrolysis.

Mitochondrial Transcripts Are Regulated by Human and *C. elegans* Dicer

Because many Dicer-binding sites map to mitochondrial transcripts but display low fidelity of sRNA production, we examined the expression changes of mitochondrial transcripts upon Dicer depletion by qPCR. We reproducibly observed increased mitochondrial transcript levels after 2 days of DICER1 depletion in HEK293 cells (Figure 4H). The effect is even more prominent after 3 days (Figure S4L), which was confirmed by northern analysis (Figure S4M) and was similarly observed in DCR-1 RNAi worms (Figures S4O and S4P).

Dicer Is Associated with sRNAs Derived from Transcription Start Sites

As described before (Seila et al., 2008; Valen et al., 2011; Zamudio et al., 2014), some sRNAs originate from transcription start sites (TSS). We observe concomitant DICER1 binding around TSS of protein-coding genes (Figure S4Q), suggesting that promoter-associated sRNAs from both strands are, at least in part, DICER1 associated. In contrast, the AGOIP peak upstream of transcription termination and polyadenylation sites (PAS) does not coincide with Dicer binding (Figure S4R). This is consistent with recently published data describing PAS-associated sRNA as Dicer independent (Valen et al., 2011).

Reproducible Dicer Binding without Detectable sRNA Production

Although we were able to associate 86% of human Dicer-binding sites and 72% of *C. elegans* sites with local stem-loop folds, the majority of these Dicer-bound hairpins (73.4% and 84.5%) appeared to not give rise to detectable sRNA. These 5,349 human and 1,510 *C. elegans* "passive" sites are particularly enriched in coding sequences and 3′ UTRs (Figures 5A and 5B).

Although we already demand that all binding sites are supported by at least two out of three independent PAR-CLIP replicates, we wanted to ensure that the passive sites indeed represent stable interactions with mRNAs. To this end, we performed independent Dicer immunoprecipitations without crosslinking or RNase treatment and assayed the bound RNAs by semiquantitative RT-PCR. In each case, the amplicons were set outside of the identified Dicer-binding sites, and the RT

(C and D) Annotation breakdown of binding sites in human (C) and *C. elegans* (D).

(E and F) Barplots of size and 5′ nt distribution of sRNA reads. Top half, left to right: sRNA originating from Dicer-binding sites, miRBase miRNAs (pos. control), and ELAVL1/GLD1-binding sites (negative control). Bottom half: sRNA from remaining transcribed parts of the genome. (E) human, (F) *C. elegans*.

(G and H) Barplot showing the fraction of sRNA read stacks that overlap with Dicer-binding sites (black), miRBase miRNAs (dark gray), or ELAVL1/GLD-1-binding sites as controls (bright gray), as a function of minimal sRNA read count in human (G) and *C. elegans* (H).
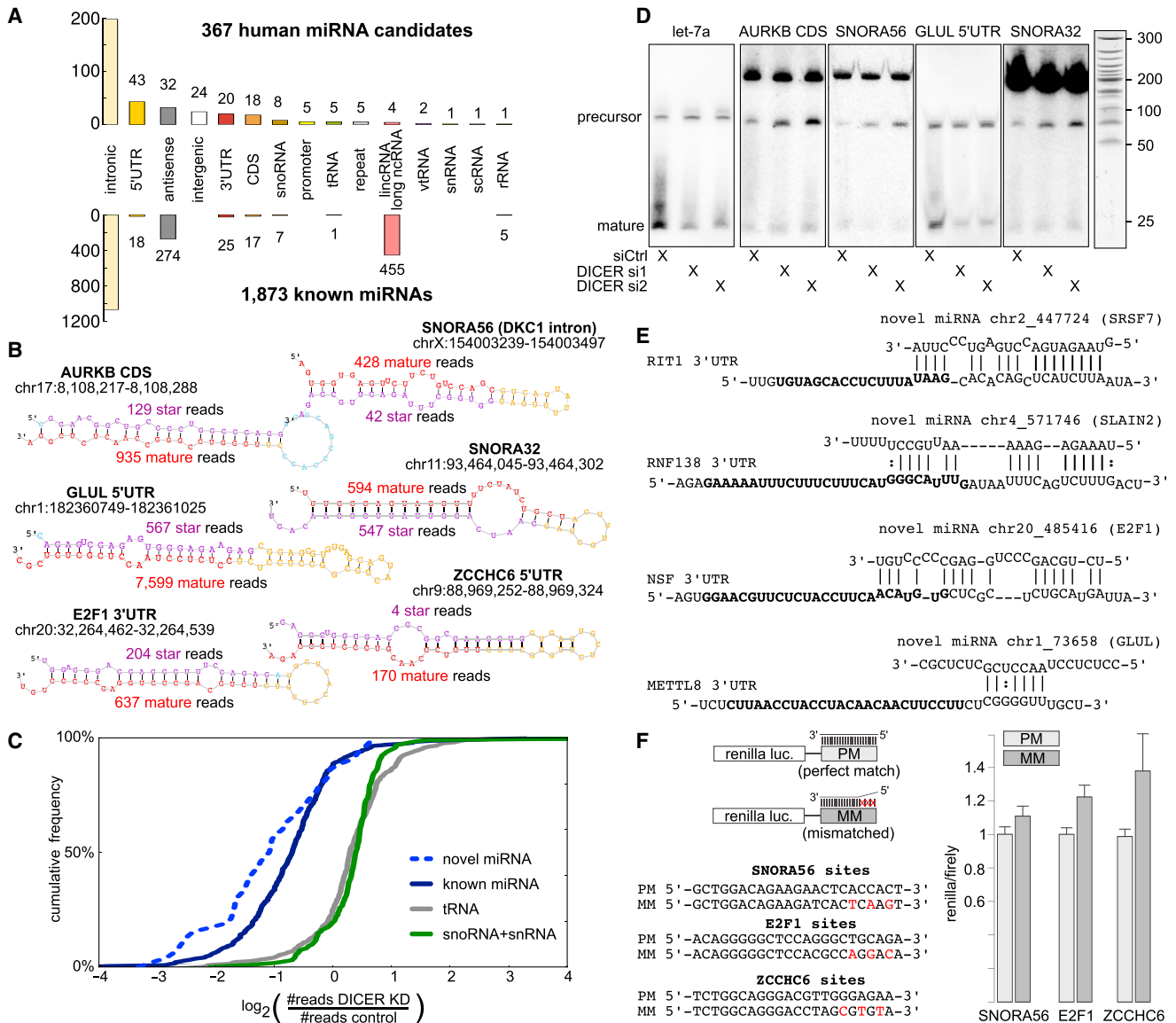
Figure 3. Identification and Validation of miRNA Targets

(A) Annotation breakdown of identified (top) and known miRBase (bottom) miRNAs.

(B) Examples of identified miRNAs of different human genomic origin. Reduced miRDeep2 plots show the precursor hairpin structure and the coverage of mature (red), star (violet), and loop (yellow) sequences.

(C) Cumulative distribution of changes in small RNA expression upon Dicer knockdown (siRNA 2). The distribution of log2 fold changes of small RNA read counts of tRNAs (gray), snoRNAs, and snRNAs (green), known miRNAs (dark blue), and miRNAs identified in this analysis (blue, dashed). Known (p < 1 ×10^{-88}, Mann-Whitney U, n = 512) and newly identified (p < 1×10^{-19}, MWU, n = 52) are significantly more downregulated after Dicer depletion than tRNAs (n = 513).

(D) Northern blots of human let-7a (pos. control) and four miRNAs validate the ∼22 nt mature product and either reduction of mature (let-7a, GLUL) or increase of precursor (AURKB, SNORA56, SNORA32) upon Dicer-depletion.

(E) Chimeric reads from AGO-CLIPs, containing miRNAs (top) and target fragments (below, fragments are extended; original chimeric part is bold) support in vivo interaction. Vertical bar: Watson-Crick pairing; colon: G-U wobble. Dashes indicate bulged nucleotides in the paired sequence.

(F) Luciferase reporter assay with perfectly matching (PM) or mismatched (MM) binding sites controlling Renilla expression. Barplot: Ratios of Renilla to firefly (control). Error bars indicate SD.

reaction was oligo-dT primed to specifically test for binding to intact mRNAs. By this assay, nine out of ten passive Dicer-binding events were validated in both human and worm (Figures 5C, 5D, S5A, and S5B), also with an antibody against endogenous DICER1 (Figure S5A).

In *C. elegans*, the most strongly bound passive target of DCR-1 in vivo is rncs-1 (Figure 5E). This lncRNA forms a long double strand that was previously described to bind DCR-1 with high affinity in vitro while being resistant to dicing due to secondary structures at the ends of the long stem (Hellwig and

Bass, 2008). The absence of abundant sRNAs from *rncs-1* argues for the absence of efficient dicing, consistent with the proposed role of *rncs-1* as a stable competitor for other Dicer substrates (Hellwig and Bass, 2008).

The transcript with the largest number of passive DICER1-binding sites in human is the mRNA encoding DICER1 itself. It is covered with 39 PAR-CLIP clusters, but like rncs-1 in worm, these sites do not emit sRNAs (Figure 5F). Another strongly bound example is a binding site in the 3′ UTR of vesicle-associated membrane protein-associated protein B/C (VAPB) (Figure 5G). The sites constitute imperfect stem-loop structures, similar to pre-miRNAs, but with larger bulges and lacking the distinct 11 nt stem segment that would be required for Drosha processing.

Passive Sites Can Interfere with Catalytic Dicer Activity

As *rncs-1* expression was shown to inhibit the catalytic activity of DCR-1 via sequestration of DCR-1 (Hellwig and Bass, 2008), we tested whether this is also possible with human passive sites.

We expressed four mRNA transcripts containing passive hairpins of different origin (DICER1, SLC2A1, TARBP2, VAPB) to high levels (in the range of GAPDH) in HEK293 cells (Figure S5C). Expression of four out of four passive hairpins reproducibly reduced the levels of endogenously expressed miRNAs by 20%–30% in comparison to constructs bearing mutations that disrupt the secondary structure (Figure 5H), suggesting that these interactions are sufficiently strong to functionally sequester Dicer protein in vivo. Dicer protein levels were not affected (Figure S5C). Taken together, our data show that Dicer stably and reproducibly binds to specific stem-loop structures inside of intact mRNAs without dicing.

Passive Dicer Targets Are Functionally Linked to RNA Granules

Mammalian and *C. elegans* Dicer is essential for the germline (Knight and Bass, 2001; Murchison et al., 2007). A hallmark of animal germ cells are RNA granules, perinuclear aggregates of RNA and protein. In *C. elegans*, Dicer is present in and required for assembly of these granules (Figure S5D and Beshore et al., 2011). As Dicer is a large protein with many protein interactors and passive binding to mRNAs does not involve RNase activity, we hypothesized that Dicer may additionally play a role in aggregating bound mRNAs to RNA granules. We intersected lists of genes known to be required for the proper formation of RNA granules with Dicer targets. To qualify as a passive Dicer target, we conservatively demand that a gene must not contain a single Dicer-binding site with sRNA output exceeding background levels (Experimental Procedures). We find that both human (p < 0.015 Fisher's exact test) and worm (p < 0.003 for P bodies, p < 2.3 × 10^{-9} for P granules) passive target transcripts of Dicer are statistically significantly enriched for genes associated with granules (Updike et al., 2011) (Figures 5I and 5J). For many of these genes, it has been demonstrated that both, their mRNAs and proteins, are localized to granules (Schisa et al., 2001). Of note, the enrichment of P-body-associated genes among Dicer targets in HEK293 cells is almost entirely due to passive targets and becomes even more significant for the most strongly bound passive targets with most crosslink events. In contrast, active

targets are not enriched for granules. The association of passive Dicer binding with RNA granules suggests a role in mRNA localization. Indeed, our FLAG-tagged DCR-1 protein recapitulates the granular localization of wild-type DCR-1 in the *C. elegans* germline (Figures 5K and S5E and Beshore et al., 2011).

Human passive targets of Dicer are also significantly (p < 2.8 × 10^{-5}) enriched for 3′ UTR targets of Staufen (Ricci et al., 2014), presumably due to the presence of double-stranded structures in their common targets. However, whereas Staufen binds to long stems (Ricci et al., 2014), Dicer binding in HEK293 cells seems restricted to the tip of the structure and requires contact to the loop region of hairpins (Figure S6B).

Homologous, Passive Targets in Human and *C. elegans*

We find 82 genes with homology between human and worm to be passive Dicer targets in both species. An interesting example is the 3′ UTR of the germline helicase CGH-1 (Figure 5L): the human ortholog DDX6 is also bound by DICER1 via its 3′ UTR (Figure 5M). Although the comparison of an animal and a human cell line across ∼550 Mya of evolution does not warrant to judge conservation of function, we find the prominent appearance of ancient genes linked to RNA granules and the germline intriguing and report passive, homologous targets in Table S3.

### Active and Passive Sites Differ in Secondary Structure

To characterize the differences between active and passive sites, we grouped Dicer sites by their activity (sRNA output level). Taking the center of the loop as a point of reference, we computed average profiles of structure, crosslinking, and other features as a function of position within the stem loop.

Known miRNAs have a characteristic fold with an unpaired loop region, flanked by paired bases within the stem of the pre-miRNA. Importantly, separated by a bulge, the stem extends for another ∼11 nt beyond the pre-miRNA boundaries, which allows processing by DROSHA/DGCR8 (Han et al., 2006). This profile is observed for active sites (Figures 6A and 6B), including the miRNA candidates identified in our analysis, and is less pronounced for sites with lower activity. Among the active sites, base-pairing stability correlates with sRNA output, which, however, appears to be optimal for intermediate folding energies (Figure S6C). In contrast, base-pairing within passive sites decays steadily to background levels, beyond ∼15 bp of stem. Figures S6D–S6G show additional profiles of G/C content and bulges.

Passive Sites Preferentially Crosslink within the Loop, Active Sites at the Ends of Precursors

If passive sites are indeed within intact mRNA transcripts, in contrast to pre-miRNA-like hairpins, no 5′ and 3′ ends should be accessible to Dicer. This difference could be reflected by the frequency of crosslink-induced transitions along the stem-loop structure because crosslinking requires close spatial proximity (Hafner et al., 2010).

Crosslinks between Dicer and HEK293-expressed miRNAs peak in three distinct regions: within the loop and at the 5′ and 3′ ends of the pre-miRNA (Figure 6C). The loop can interact with the Dicer helicase domain (Taylor et al., 2013), whereas

A H/ACA box SNORA33 (human)

B binding site with flanking sequence
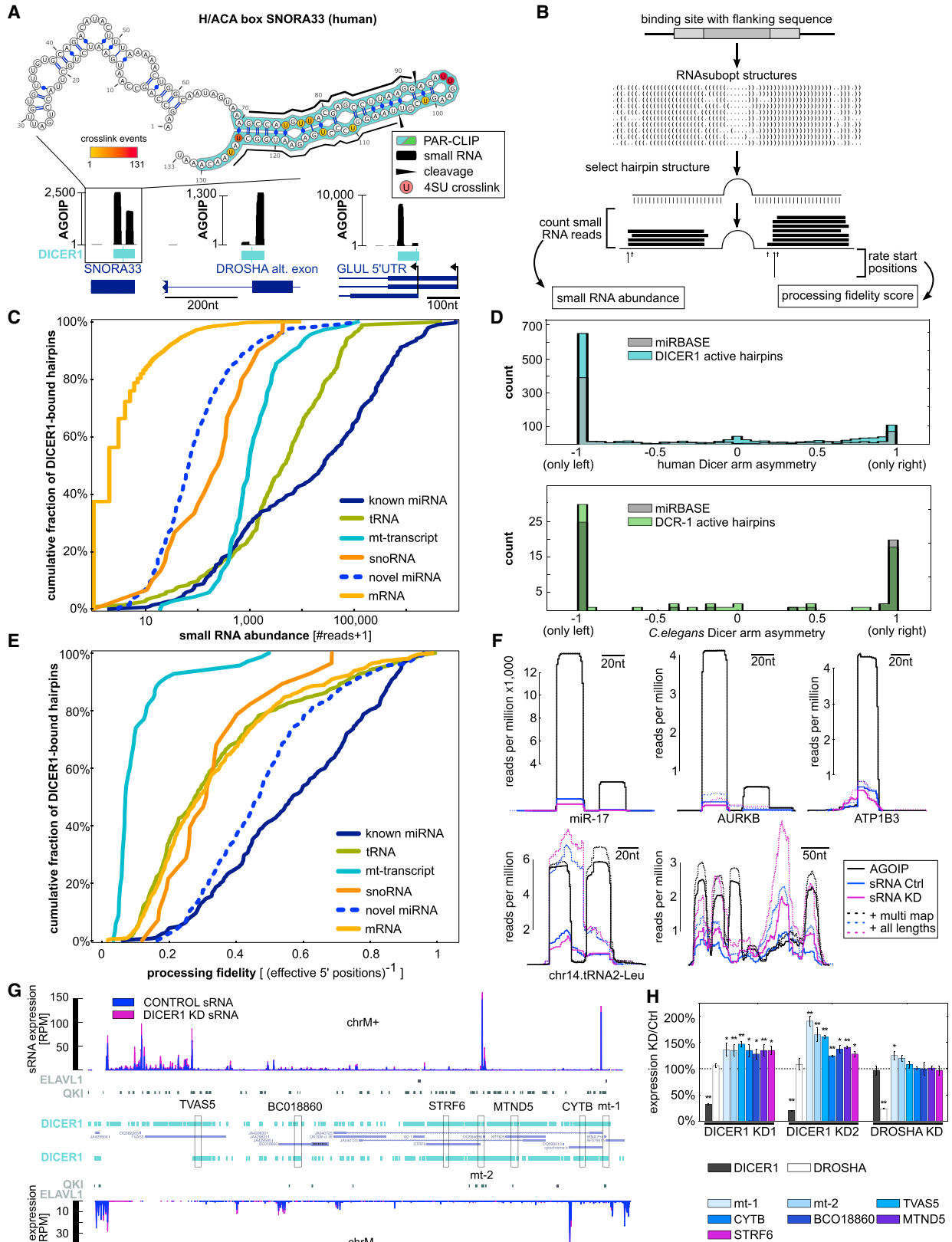
(legend on next page)

the ends are coordinated by the 5′ pocket and PAZ domain, which is required for "setting the ruler" and efficient dicing (MacRae et al., 2007; Park et al., 2011). This pattern is observed for all active sites, supporting that active sites are bound as pre-miRNA-like structures cleaved out of larger transcripts. In contrast, passive sites lack the peaks at precursor-size distance (Figure 6C). In *C. elegans,* the same difference between active and passive sites at the precursor termini is observed. However, active sites display much less crosslinking within the loop region (Figure 6D), consistent with the dispensability of the helicase domain for miRNA production in *C. elegans* (Welker et al., 2011).

Passive Sites Are Not Engaged by DGCR8 and Are Not Cleaved

For humans, we additionally interrogated DGCR8 HITS-CLIP data (Macias et al., 2012) and 5′-monophosphate sequencing of endonucleolytic RNA cleavage products (Karginov et al., 2010). Active sites and known miRNAs display prominent shoulders of DGCR8 HITS-CLIP signal, flanking the pre-miRNA and indicating processing by the Drosha/DGCR8 complex. This signal is absent from passive sites (Figure 6E). Moreover, binding of the Dicer-interacting proteins PACT and TARBP2 (Goodarzi et al., 2014) was only detected for loci with sRNA output (Figures S6H–S6J). Consistently, only active sites, but not passive sites, are supported by cleavage products aligning to either the pre-miRNA itself or the 3′ fragments of the primary transcript.

We conclude that the passive Dicer-binding sites observed in human and nematode represent local stem-loop structures, which structurally differ from typical miRNAs and are consequently not cleaved by Drosha or Dicer. Rather, they constitute structural elements within intact transcripts.

Functional Consequences of Dicer Binding and Dicing

To assess the significance and biological consequences of widespread Dicer binding to mRNAs, we performed Dicer knockdowns in HEK293 cells and *C. elegans*, followed by mRNA sequencing. To limit indirect effects due to reduced miRNA production, we chose the earliest time point at which Dicer protein levels were significantly reduced while miRNA levels were stable. As an additional control, we knocked down Drosha (Figures S7A and S7C–S7F).

In *C. elegans*, miRNA expression changes during development, and Dicer depletion causes pleiotropic effects, including sterility. Our best attempt to study the impact of Dicer binding on mRNAs therefore consists of choosing a developmental stage when Dicer can be significantly depleted (Figures S7B, S7G, and S7H) while the animals are viable. We used the *fem-1 (hc17)* strain (inducible sterility) to obtain sterile control animals.

As the sequencing-derived log fold changes correlated well with qPCR measurements on independent biological replicates (Figure S7M), we compared the distribution of changes between different groups of transcripts. Overall, the observed changes at early time points are small. However, in both HEK293 cells and *C. elegans*, we observed statistically significant effects on Dicer target transcripts. mRNAs containing active Dicer-binding sites were significantly stabilized upon Dicer depletion, whereas passive targets overall were more destabilized in both worms and cells (see Figure S7N for mRNA decay). The most strongly bound passive targets, as well as the homologous passive targets of DICER1 and DCR-1, show the strongest effects (Figures 7A, S7I, and 7B).

In contrast, Drosha depletion had almost no effect on passive targets in HEK293 cells (Figures S7J and S7K), ruling out that the observed regulation stems from global loss of miRNAs. In *C. elegans*, where results strongly depended on the developmental stage, Drosha depletion stabilized both passive and active targets (Figure S7L).

DISCUSSION

The rapid increase of sequencing data has led to the discovery of many sRNAs that have been linked to Dicer activity. However, sRNA data capture only the endpoint of a cascade of processing events on a background of degradation products, and the indirect identification of Dicer binding from sRNA requires assumptions about miRNA-like processing. Here, we biochemically identified thousands of in vivo Dicer-binding sites in a human cell line and *C. elegans.*

Although we recovered known miRNAs with high sensitivity, these account for only a fraction of our data. We identified and validated many additional miRNAs. Most of these are lowly expressed and would not have been detected without direct evidence for Dicer binding.

Figure 4. Dicer-Binding Sites in Different RNA Substrates Display a Wide Dynamic Range of sRNA Expression and Processing Fidelity

(A–F) Dicer binding to noncanonical structural RNA: snoRNA SNORA33, insert shows structure as in Figure 2A, 3′ splice site of DROSHA mRNA, 5′ UTR of GLUL mRNA in HEK293 cells. Dark blue: exon/intron structure of bound transcript; light blue: Dicer-binding site, AGOIP coverage in black.

(B) Schematic of the local hairpin search. Binding sites (dark gray) are padded with flanking sequence (light gray) and are folded with RNAsubopt. The best hairpin is selected and intersected with aligned AGOIP/sRNA reads. Numbers and positions of alignments used to score output, left/right arm asymmetry, and fidelity of sRNA processing are shown.

(C) Cumulative frequency of AGOIP read counts for different categories of Dicer-binding sites in HEK293 cells.
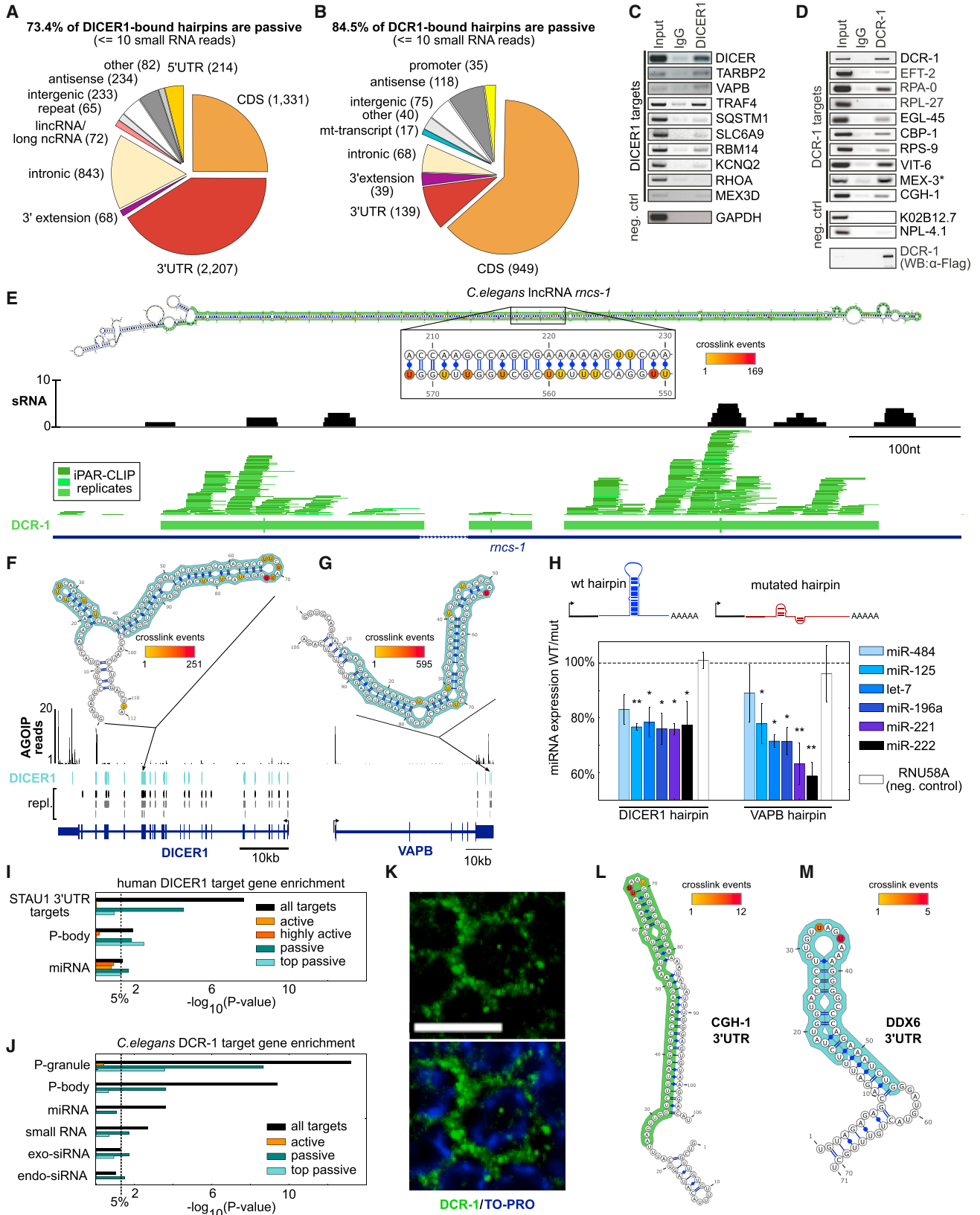
(D) Small RNA production is biased toward the left arm. Histogram of sRNA read coverage asymmetry across all sRNA precursors in human (top) and *C. elegans* (bottom). Dicer-bound active hairpins (bright blue and green, respectively) and miRBase annotated human and *C. elegans* pre-miRNAs (dark gray).

(E) Cumulative frequency of AGOIP read stack fidelity (0 for randomly scattered reads, 1 for a single 5′ start position for all reads) for different categories of Dicer-binding sites in HEK293 cells.

(F) Examples of AGO-loaded small RNA sources. Black: AGOIP reads, blue: 5′P sRNA control, purple: 5′P sRNA DICER1 KD2. Uniquely aligning reads 18–25 nt. Dashed lines: all reads.

(G) Dicer (blue boxes) and neg. control QKI and ELAVL1 (gray boxes) binding along mitochondrial transcripts (middle, blue) from the plus strand (top) or minus strand (bottom). sRNA output in blue (control) and purple (DICER1 KD with siRNA2). Amplicons for qPCR are marked by rectangles.

(H) Barplot with qPCR changes of mitochondrial transcripts upon Dicer KD. Error bars represent SEM. *p < 5%, **p < 1%, two-sided, two-sample t test on delta Ct, unequal variance, n = 3.

A **73.4% of DICER1-bound hairpins are passive**
(<= 10 small RNA reads)

other (82) 5'UTR (214)
antisense (234)
intergenic (233)
repeat (65)
lincRNA/
long ncRNA (72)
CDS (1,331)
intronic (843)
3' extension (68)
3'UTR (2,207)

B **84.5% of DCR1-bound hairpins are passive**
(<= 10 small RNA reads)

promoter (35)
antisense (118)
intergenic (75)
other (40)
mt-transcript (17)
intronic (68)
3'extension (39)
3'UTR (139)
CDS (949)

C
Input | IgG | DICER1

DICER1 targets:
DICER
TARBP2
VAPB
TRAF4
SQSTM1
SLC6A9
RBM14
KCNQ2
RHOA
MEX3D

neg. ctrl: GAPDH

D
Input | IgG | DCR-1

DCR-1 targets:
DCR-1
EFT-2
RPA-0
RPL-27
EGL-45
CBP-1
RPS-9
VIT-6
MEX-3*
CGH-1

neg. ctrl:
K02B12.7
NPL-4.1
DCR-1 (WB:α-Flag)

E *C.elegans* lncRNA *rncs-1*

crosslink events 1 — 169

sRNA 10 / 0

100nt

iPAR-CLIP replicates

DCR-1

rncs-1

F
crosslink events 1 — 251
AGOIP reads 20 / 1
DICER1
repl.
DICER1
10kb

G
crosslink events 1 — 595
VAPB
10kb

H
wt hairpin   mutated hairpin
AAAAA   AAAAA

miRNA expression WT/mut
100%
80%
60%
DICER1 hairpin   VAPB hairpin

miR-484
miR-125
let-7
miR-196a
miR-221
miR-222
RNU58A (neg. control)

I human DICER1 target gene enrichment
STAU1 3'UTR targets
P-body
miRNA
all targets
active
highly active
passive
top passive
5%
2   6   10
-log$_{10}$(P-value)

J *C.elegans* DCR-1 target gene enrichment
P-granule
P-body
miRNA
small RNA
exo-siRNA
endo-siRNA
all targets
active
passive
top passive
5%
2   6   10
-log$_{10}$(P-value)

K
DCR-1/**TO-PRO**

L
crosslink events 1 — 12
CGH-1 3'UTR

M
crosslink events 1 — 5
DDX6 3'UTR

(legend on next page)

Furthermore, the intersection of Dicer binding, sRNA sequencing, and AGOIP allowed us to investigate Dicer interactions with structural RNAs. We found not only unusual patterns of dicing, but also Dicer-independent, AGO-loaded sRNAs, presumably generated by other modes of processing or degradation. As Dicer-bound structural RNAs accumulated upon Dicer depletion (Figures S3F and S3G), regardless of the fidelity of their processing, this opens the possibility that, rather than producing *trans*-acting miRNAs, Dicer may contribute to or interfere with the degradation of many structural RNAs. Dicer is required for germline maintenance and early embryogenesis (Knight and Bass, 2001; Murchison et al., 2007), and its expression changes during differentiation. This observation could, in part, explain why the pool of tRNAs differs between proliferating and differentiated cells (Gingold et al., 2014).

To our surprise, depletion of Dicer, on average, increased the levels of mature mRNAs hosting active sites, suggesting that there can be some form of feedback to DROSHA/DGCR8 processing. Components of the sRNA pathway are known to auto- and cross-regulate each other (Bennasser et al., 2011; Han et al., 2009), and many, including three AGOs, TARBP2, and DROSHA, are also targeted by DICER1. Indeed, Dicer and Drosha levels were not independent in perturbation experiments. However, at the early time point in Figures 7A and 7B, Drosha levels were unaffected.

Based on studies in nematodes, it has been suggested that long dsRNA represents the archetypal substrate for Dicer, linked to its supposed ancient role in defense against dsRNA viruses or genome-invading elements (Sarkies and Miska, 2013). However, it appears that binding of human, somatic Dicer requires contact to the loop region, presumably utilizing the helicase domain, as we observe prominent crosslinking to hairpin loops. Indeed, we find almost no long dsRNA binding in HEK293 cells (Figures S4S, S6A, and S6B), in line with other findings that somatic mammalian cells lack endo-siRNA biogenesis (Nejepinska et al., 2012).

Our unbiased look at Dicer-binding sites pictures a continuum of sRNA expression from diverse substrates and surprisingly encompasses a large set of non-diced, passive binding sites that were previously undetectable. This class of passive binding sites, predominantly residing in mRNAs, can be stably bound

without endonuclease activity. This observation agrees with previous reports indicating that Dicer is not able to efficiently process RNA if no free ends are available (Fukunaga et al., 2014; Zhang et al., 2002) and explains the resistance of the lncRNA *rncs-1* against dicing (Hellwig and Bass, 2008).

Our data argue that passive sites are not transient interactions and that they can stabilize expression of their host transcripts. We note that these effects were strongest for mRNAs that were most strongly crosslinked or are also passive targets in the other species.

What could be the function of passive Dicer sites besides generally stabilizing target RNAs? First, we have shown that high expression of passive sites can sequester Dicer protein from other target transcripts and can interfere with its catalytic activity, analogous to lncRNA *rncs-1* in *C. elegans*. In these experiments, we expressed a few thousand passive sites per cell. It is not trivial to estimate the total copy number of all passive sites, but according to our data, a high number of passive sites may be naturally present. These sites could have a buffering function and may compete with active sites for Dicer binding. Second, the proteins encoded by passively bound mRNAs can immediately interact with Dicer, as they are translated. This would be a mechanism for efficient buildup of protein complexes. In support of this idea, several mRNAs with passive sites both in human and worm encode known Dicer interactors (DRH-1, LIN-41, TARBP2, DHX9). Third, many passively bound mRNAs encode RNA-granule-associated genes, and we show that Dicer protein itself is granularly localized in the *C. elegans* germline. Taken together with the observation that Dicer is required for RNP granule formation (Figure S5D and Beshore et al., 2011), this indicates that passive binding may be important for RNA localization and assembly of RNA-protein complexes. This finding could explain the diverse although overall weak effects observed upon DICER1 knockdown in HEK293 cells and suggests that germ cells, forming prominent RNP structures, are a more suitable system to investigate this aspect of Dicer biology. Along the same line, the *fem1* strain of *C. elegans* forms large and prominent P granules in its arrested oocytes (Schisa et al., 2001), and the observed changes of passive targets upon DCR-1 depletion were more pronounced in these animals than in the wild-type (data not shown). Independently supporting

Figure 5. Most mRNA Sites Are Not Substrates for Dicer Cleavage

(A and B) Pie charts of genomic annotation of passive binding sites for human (A) and *C. elegans* (B).

(C and D) Validation of catalytically passive Dicer targets by RIP RT-PCR. RT-PCR on RNA from Dicer IP validated nine out of ten tested in HEK293 cells (C) and nine out of ten in *C. elegans* (D). RT-PCR for abundant GAPDH mRNA served as negative controls for human. Mouse IgG IP and RT-PCR for three abundant transcripts: K02B12.7, NPL4.1 as negative control in *C. elegans*. RT-PCR reaction primed with oligo-dT primer and amplicon set outside of PAR-CLIP-binding sites. *MEX-3 has low PAR-CLIP coverage, is excluded from the strict consensus list, and was included here for comparison.

(E) *C. elegans* highly double-stranded lncRNA *rncs-1* is densely bound, but not efficiently cleaved by DCR-1. Top: secondary structure as in Figures 1D and 1E with zoomed region to show crosslinks. Bottom: sense (antisense) aligned distinct PAR-CLIP reads are indicated as green lines above (below) the gene structure. Shades of green for replicates. Thick blue lines: exons, green boxes: consensus binding sites. sRNA read coverage of *rncs-1* locus in black.

(F and G) Examples of passive Dicer-binding sites in HEK293 cells: DICER1 mRNA (F), 3′ UTR of VAPB mRNA (G). Dicer binding sites: blue boxes. AGOIP sRNA reads in black. Top inserts show predicted folding of Dicer-bound hairpins overlaid with PAR-CLIP data as in Figures 2A and 2B.

(H) Sponge assay. Barplots with the ratio of miRNA levels in cells expressing passive hairpins relative to mutated hairpins. Error bars represent SEM. *p < 5%, **p < 1%, two-sided, two-sample t test, unequal variance, n = 3.

(I and J) Barplots of -log10 p value (one-sided Fisher's exact test) for target enrichment of human (I) and *C. elegans* (J) Dicer. Vertical line demarcates p = 5%.

(K) FLAG-tagged DCR-1 expression in *C. elegans* germline RNP granules. Immunohistochemistry was performed on extruded gonads from adult worms using monoclonal anti-FLAG (green) antibody. Overview and higher-magnification view of adult germline are provided (Figure S5E). Cell nuclei were counterstained with TO-PRO (blue). Scale bar, 10 μm.

(L and M) Passive binding sites in the homologous targets conserved germline helicase-1 (CGH-1, *C. elegans*) and DEAD box RNA-helicase 6 (DDX6, human).
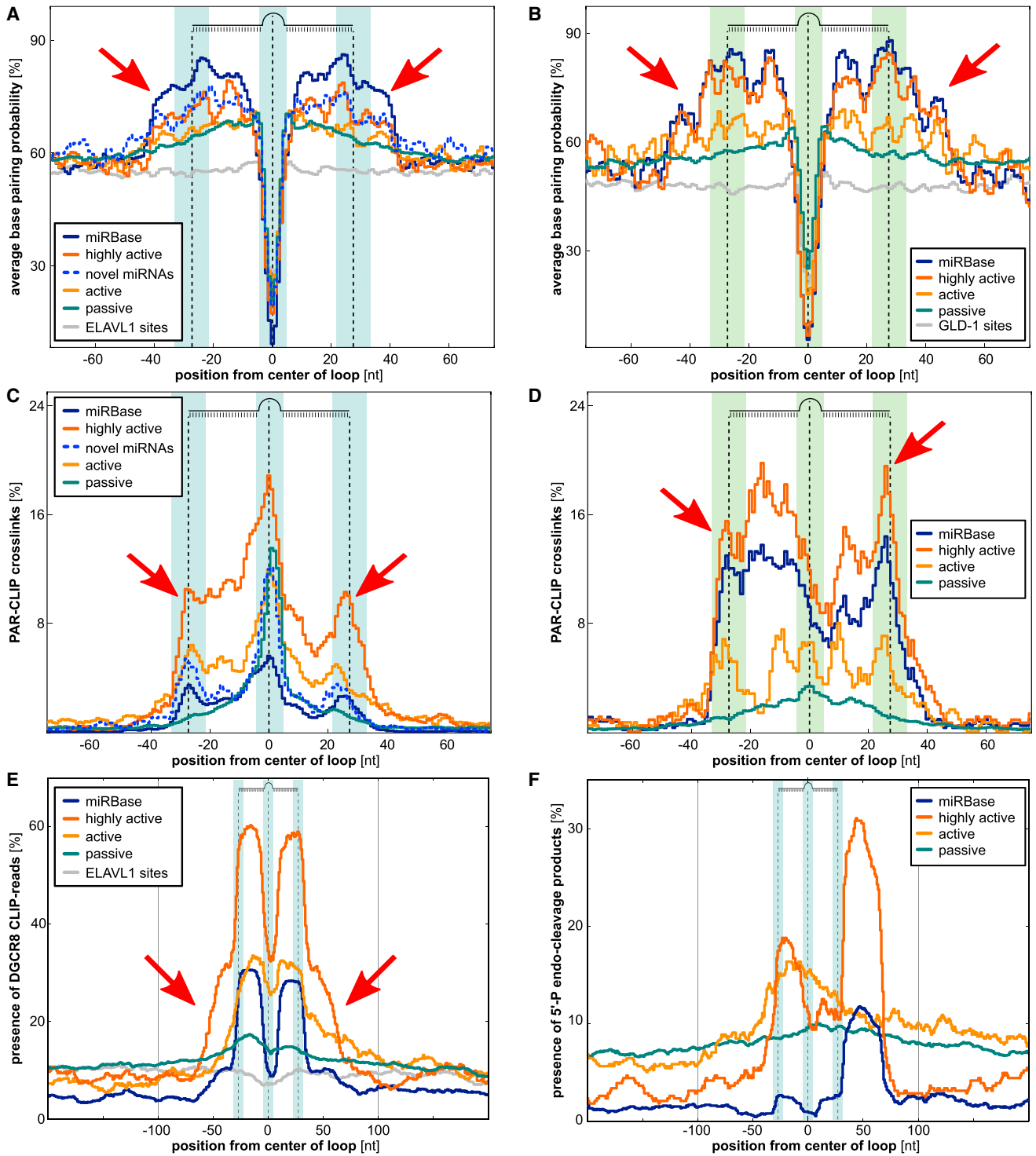
Figure 6. Secondary Structure and Dicer Crosslinking as a Function of sRNA Expression

(A and B) Average base-pairing probability at each nucleotide position in fixed windows around the loop center for human DICER1 (A) and *C.elegans* DCR-1-binding sites (B). Catalytic activity levels correspond to sRNA read count: highly active >1,000 (>100), active > 10, passive < 10 for human (*C. elegans*). ELAVL1/GLD-1 (single-stranded) binding sites as negative, miRBase miRNAs as positive control. Arrows point at the additional 11 nt of stem, characteristic for DROSHA/DGCR8 substrates.

(C and D) Average occurrence of crosslink mutations at each nucleotide position in fixed windows around the loop for human DICER1 (C) and *C. elegans* DCR-1 (D) bound hairpins. Activity levels as above. Arrows point at the peaks of crosslinking at the 5′ and 3′ ends of precursor structures, absent from passive sites.

(legend continued on next page)
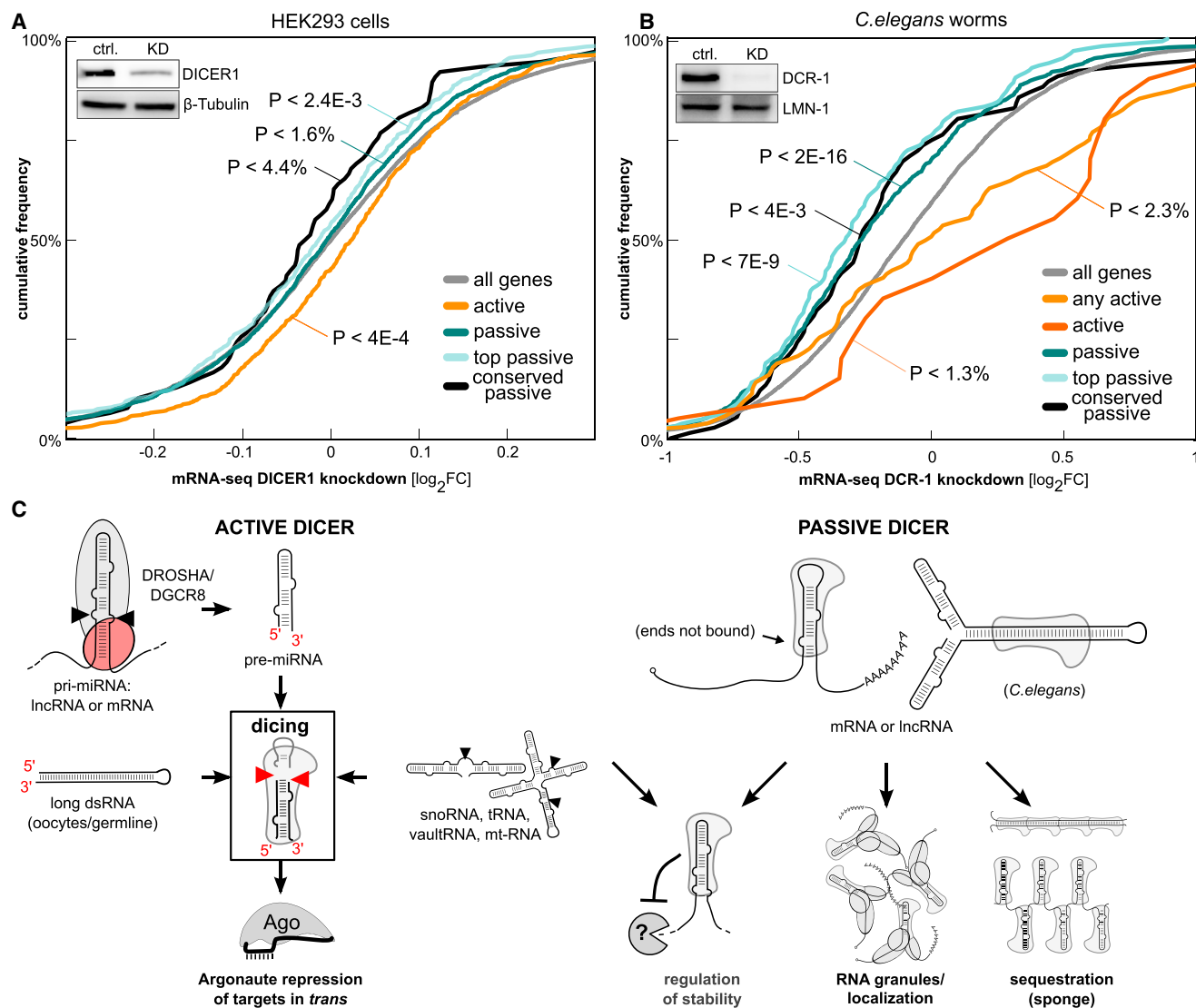
Figure 7. Functional Consequences of Dicer Binding and Dicing

(A and B) Cumulative frequencies of mRNA $\log_2$ fold changes for all detected genes (gray), catalytically active (orange/red), and passive targets (blue, most strongly bound: bright blue), as well as a conserved set of Dicer targets (black) in HEK293 cells (A) and *C. elegans* (B). Inserts: western blot validation of Dicer knockdowns. HEK293: active p = 3.95×10^{-4} (n = 753), passive p = 1.586×10^{-2} (n = 2326), top passive p = 2.349×10^{-3} (n = 456), conserved passive p = 4.337% (n = 78), all genes: n = 10,007. *C. elegans*: any active p = 2.287×10^{-2} (n = 107), active p = 1.233% (n = 20), passive p = 1.695×10^{-16} (n = 804), top passive p = 7.412×10^{-9} (n = 165), conserved passive p = 4.269×10^{-3} (n = 75), all genes: n = 8501. All p values by double-sided Mann-Whitney U.

(C) Model of Dicer-RNA interactions and function.

the idea that passive Dicer binding may influence mRNA localization, Dicer was recently found to be a shuttling protein (Doyle et al., 2013), and both human and *C. elegans* Dicer interact with nuclear pore complexes (Ando et al., 2011).

In summary, we mapped transcriptome-wide Dicer binding in human and *C. elegans*, identified and validated numerous additional miRNAs, disentangled the contribution of Dicer to the degradation of many structural RNAs, and unveiled passive binding of Dicer with possible functions outside of sRNA pathways (Figure 7C). Also, the microprocessor components DGCR8/Pasha and DROSHA have recently been shown to have miRNA-independent cellular functions (Gromak et al.,

(E) Average presence of at least one DGCR8 HITS-CLIP (Macias et al., 2012) read for each nucleotide position in fixed windows around the loop for human DICER1-bound hairpins. Activity levels as above. Arrows point at the additional shoulders at the base, outside of pre-miRNAs, absent from passive sites.

(F) Average presence of at least one endocleaved 5′-P bearing sequencing read (Karginov et al., 2010) for each nucleotide position in fixed windows around the loop for human DICER1-bound hairpins. Drosha/DGCR8 processing produces one peak starting at the 5′ end of the pre-miRNA and one at the remaining 3′ fragment. Both are absent from passive sites.

2013; Macias et al., 2012). Altogether, this indicates that the versatile functions of these ancient RBPs may have been overshadowed by a focus on miRNAs and may need to be carefully re-examined.

EXPERIMENTAL PROCEDURES

Standard molecular biology techniques and lists of reagents as well as computational analyses are described in the Extended Experimental Procedures.

PAR-CLIP and iPAR-CLIP HEK293 Cells
PAR-CLIP on HEK293 cells stably expressing FLAG/HA-DICER1 was performed as described (Hafner et al., 2010). In vivo PAR-CLIP was performed on DCR-1::FLAG rescue strain (BB92; dcr-1(ok247)III,uuEx18) as described previously (Jungkamp et al., 2011).

AGO-Associated sRNA Cloning and Sequencing
Immunoprecipitation of FLAG/HA-tagged AGO2/3 cells was performed with FLAG magnetic beads. RNA was isolated, ligated to 3′ and 5′ adapters, reverse transcribed, and PCR amplified and sequenced.

sRNA Sequencing
sRNA sequencing was performed from 5 μg total RNA according to the standard Illumina sRNA library preparation protocol. For 5′ triphosphate sRNA sequencing, RNA was treated with 5′ polyphosphatase before library preparation.

ACCESSION NUMBERS

Sequencing data have been deposited in the GEO database under GSE55333.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, seven figures, and four tables and can be found with this article online at http://dx.doi.org/10.1016/j.cell.2014.10.040.

AUTHOR CONTRIBUTIONS

A.R.-W., M.J., and Y.M. contributed equally to this work. A.R.-W. and Y.M. designed and performed experiments, supervised by M.L. and N.R. A.R.-W. performed the C. elegans DCR-1 PAR-CLIP and the majority of validation experiments, assisted by M.H. Y.M. performed all of the human PAR-CLIPs, AGO-IPs, and validation experiments. M.J. carried out all of the computational analyses, supervised by N.R. N.R, M.L, M.J, A.R.-W., and Y.M. wrote the paper.

REFERENCES

Ando, Y., Tomaru, Y., Morinaga, A., Burroughs, A.M., Kawaji, H., Kubosaki, A., Kimura, R., Tagata, M., Ino, Y., Hirano, H., et al. (2011). Nuclear pore complex protein mediated nuclear localization of dicer protein in human cells. PLoS ONE 6, e23385.

Bartel, D.P. (2009). MicroRNAs: target recognition and regulatory functions. Cell 136, 215–233.

Bennasser, Y., Chable-Bessia, C., Triboulet, R., Gibbings, D., Gwizdek, C., Dargemont, C., Kremer, E.J., Voinnet, O., and Benkirane, M. (2011). Competition for XPO5 binding between Dicer mRNA, pre-miRNA and viral RNA regulates human Dicer levels. Nat. Struct. Mol. Biol. 18, 323–327.

Beshore, E.L., McEwen, T.J., Jud, M.C., Marshall, J.K., Schisa, J.A., and Bennett, K.L. (2011). C. elegans Dicer interacts with the P-granule component GLH-1 and both regulate germline RNPs. Dev. Biol. 350, 370–381.

Castellano, L., and Stebbing, J. (2013). Deep sequencing of small RNAs identifies canonical and non-canonical miRNA and endogenous siRNAs in mammalian somatic tissues. Nucleic Acids Res. 41, 3339–3351.

Chiang, H.R., Schoenfeld, L.W., Ruby, J.G., Auyeung, V.C., Spies, N., Baek, D., Johnston, W.K., Russ, C., Luo, S., Babiarz, J.E., et al. (2010). Mammalian microRNAs: experimental evaluation of novel and previously annotated genes. Genes Dev. 24, 992–1009.

Doyle, M., Badertscher, L., Jaskiewicz, L., Güttinger, S., Jurado, S., Hugenschmidt, T., Kutay, U., and Filipowicz, W. (2013). The double-stranded RNA binding domain of human Dicer functions as a nuclear localization signal. RNA 19, 1238–1252.

Ender, C., Krek, A., Friedländer, M.R., Beitzinger, M., Weinmann, L., Chen, W., Pfeffer, S., Rajewsky, N., and Meister, G. (2008). A human snoRNA with microRNA-like functions. Mol. Cell 32, 519–528.

Feng, Y., Zhang, X., Graves, P., and Zeng, Y. (2012). A comprehensive analysis of precursor microRNA cleavage by human Dicer. RNA 18, 2083–2092.

Friedländer, M.R., Mackowiak, S.D., Li, N., Chen, W., and Rajewsky, N. (2012). miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. Nucleic Acids Res. 40, 37–52.

Friedländer, M.R., Lizano, E., Houben, A.J., Bezdan, D., Báñez-Coronel, M., Kudla, G., Mateu-Huertas, E., Kagerbauer, B., González, J., Chen, K.C., et al. (2014). Evidence for the biogenesis of more than 1,000 novel human microRNAs. Genome Biol. 15, R57.

Fukunaga, R., Han, B.W., Hung, J.H., Xu, J., Weng, Z., and Zamore, P.D. (2012). Dicer partner proteins tune the length of mature miRNAs in flies and mammals. Cell 151, 533–546.

Fukunaga, R., Colpan, C., Han, B.W., and Zamore, P.D. (2014). Inorganic phosphate blocks binding of pre-miRNA to Dicer-2 via its PAZ domain. EMBO J. 33, 371–384.

Gingold, H., Tehler, D., Christoffersen, N.R., Nielsen, M.M., Asmar, F., Kooistra, S.M., Christophersen, N.S., Christensen, L.L., Borre, M., Sørensen, K.D., et al. (2014). A dual program for translation regulation in cellular proliferation and differentiation. Cell 158, 1281–1292.

Goodarzi, H., Zhang, S., Buss, C.G., Fish, L., Tavazoie, S., and Tavazoie, S.F. (2014). Metastasis-suppressor transcript destabilization through TARBP2 binding of mRNA hairpins. Nature 513, 256–260.

Gromak, N., Dienstbier, M., Macias, S., Plass, M., Eyras, E., Cáceres, J.F., and Proudfoot, N.J. (2013). Drosha regulates gene expression independently of RNA cleavage function. Cell Rep. 5, 1499–1510.

Grosswendt, S., Filipchyk, A., Manzano, M., Klironomos, F., Schilling, M., Herzog, M., Gottwein, E., and Rajewsky, N. (2014). Unambiguous identification of miRNA:target site interactions by different types of ligation reactions. Mol. Cell 54, 1042–1054.

Gu, W., Shirayama, M., Conte, D., Jr., Vasale, J., Batista, P.J., Claycomb, J.M., Moresco, J.J., Youngman, E.M., Keys, J., Stoltz, M.J., et al. (2009). Distinct argonaute-mediated 22G-RNA pathways direct genome surveillance in the C. elegans germline. Mol. Cell 36, 231–244.

Gu, S., Jin, L., Zhang, Y., Huang, Y., Zhang, F., Valdmanis, P.N., and Kay, M.A. (2012). The loop position of shRNAs and pre-miRNAs is critical for the accuracy of dicer processing in vivo. Cell 151, 900–911.

Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, M., Jr., Jungkamp, A.C., Munschauer, M., et al. (2010). Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. Cell 141, 129–141.

Han, J., Lee, Y., Yeom, K.H., Nam, J.W., Heo, I., Rhee, J.K., Sohn, S.Y., Cho, Y., Zhang, B.T., and Kim, V.N. (2006). Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. Cell *125*, 887–901.

Han, J., Pedersen, J.S., Kwon, S.C., Belair, C.D., Kim, Y.K., Yeom, K.H., Yang, W.Y., Haussler, D., Blelloch, R., and Kim, V.N. (2009). Posttranscriptional crossregulation between Drosha and DGCR8. Cell *136*, 75–84.

Hellwig, S., and Bass, B.L. (2008). A starvation-induced noncoding RNA modulates expression of Dicer-regulated genes. Proc. Natl. Acad. Sci. USA *105*, 12897–12902.

Hutvágner, G., McLachlan, J., Pasquinelli, A.E., Bálint, E., Tuschl, T., and Zamore, P.D. (2001). A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. Science *293*, 834–838.

Jungkamp, A.C., Stoeckius, M., Mecenas, D., Grün, D., Mastrobuoni, G., Kempa, S., and Rajewsky, N. (2011). In vivo and transcriptome-wide identification of RNA binding protein target sites. Mol. Cell *44*, 828–840.

Karginov, F.V., Cheloufi, S., Chong, M.M., Stark, A., Smith, A.D., and Hannon, G.J. (2010). Diverse endonucleolytic cleavage sites in the mammalian transcriptome depend upon microRNAs, Drosha, and additional nucleases. Mol. Cell *38*, 781–788.

Kim, V.N., Han, J., and Siomi, M.C. (2009). Biogenesis of small RNAs in animals. Nat. Rev. Mol. Cell Biol. *10*, 126–139.

Knight, S.W., and Bass, B.L. (2001). A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in Caenorhabditis elegans. Science *293*, 2269–2271.

Kotaja, N., Bhattacharyya, S.N., Jaskiewicz, L., Kimmins, S., Parvinen, M., Filipowicz, W., and Sassone-Corsi, P. (2006). The chromatoid body of male germ cells: similarity with processing bodies and presence of Dicer and microRNA pathway components. Proc. Natl. Acad. Sci. USA *103*, 2647–2652.

Kozomara, A., and Griffiths-Jones, S. (2014). miRBase: annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res. *42* (Database issue), D68–D73.

Lau, P.W., Guiley, K.Z., De, N., Potter, C.S., Carragher, B., and MacRae, I.J. (2012). The molecular architecture of human Dicer. Nat. Struct. Mol. Biol. *19*, 436–440.

Lebedeva, S., Jens, M., Theil, K., Schwanhäusser, B., Selbach, M., Landthaler, M., and Rajewsky, N. (2011). Transcriptome-wide analysis of regulatory interactions of the RNA-binding protein HuR. Mol. Cell *43*, 340–352.

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Rådmark, O., Kim, S., and Kim, V.N. (2003). The nuclear RNase III Drosha initiates microRNA processing. Nature *425*, 415–419.

Macias, S., Plass, M., Stajuda, A., Michlewski, G., Eyras, E., and Cáceres, J.F. (2012). DGCR8 HITS-CLIP reveals novel functions for the Microprocessor. Nat. Struct. Mol. Biol. *19*, 760–766.

MacRae, I.J., Zhou, K., and Doudna, J.A. (2007). Structural determinants of RNA recognition and cleavage by Dicer. Nat. Struct. Mol. Biol. *14*, 934–940.

Murchison, E.P., Stein, P., Xuan, Z., Pan, H., Zhang, M.Q., Schultz, R.M., and Hannon, G.J. (2007). Critical roles for Dicer in the female germline. Genes Dev. *21*, 682–693.

Nejepinska, J., Malik, R., Filkowski, J., Flemr, M., Filipowicz, W., and Svoboda, P. (2012). dsRNA expression in the mouse elicits RNAi in oocytes and low adenosine deamination in somatic cells. Nucleic Acids Res. *40*, 399–413.

Okamura, K., Hagen, J.W., Duan, H., Tyler, D.M., and Lai, E.C. (2007). The mirtron pathway generates microRNA-class regulatory RNAs in Drosophila. Cell *130*, 89–100.

Park, J.E., Heo, I., Tian, Y., Simanshu, D.K., Chang, H., Jee, D., Patel, D.J., and Kim, V.N. (2011). Dicer recognizes the 5′ end of RNA for efficient and accurate processing. Nature *475*, 201–205.

Ricci, E.P., Kucukural, A., Cenik, C., Mercier, B.C., Singh, G., Heyer, E.E., Ashar-Patel, A., Peng, L., and Moore, M.J. (2014). Staufen1 senses overall transcript secondary structure to regulate translation. Nat. Struct. Mol. Biol. *21*, 26–35.

Ruby, J.G., Jan, C.H., and Bartel, D.P. (2007). Intronic microRNA precursors that bypass Drosha processing. Nature *448*, 83–86.

Sarkies, P., and Miska, E.A. (2013). RNAi pathways in the recognition of foreign RNA: antiviral responses and host-parasite interactions in nematodes. Biochem. Soc. Trans. *41*, 876–880.

Schisa, J.A., Pitt, J.N., and Priess, J.R. (2001). Analysis of RNA associated with P granules in germ cells of C. elegans adults. Development *128*, 1287–1298.

Seila, A.C., Calabrese, J.M., Levine, S.S., Yeo, G.W., Rahl, P.B., Flynn, R.A., Young, R.A., and Sharp, P.A. (2008). Divergent transcription from active promoters. Science *322*, 1849–1851.

Shi, Z., Montgomery, T.A., Qi, Y., and Ruvkun, G. (2013). High-throughput sequencing reveals extraordinary fluidity of miRNA, piRNA, and siRNA pathways in nematodes. Genome Res. *23*, 497–508.

Sinkkonen, L., Hugenschmidt, T., Filipowicz, W., and Svoboda, P. (2010). Dicer is associated with ribosomal DNA chromatin in mammalian cells. PLoS ONE *5*, e12175.

Taylor, D.W., Ma, E., Shigematsu, H., Cianfrocco, M.A., Noland, C.L., Nagayama, K., Nogales, E., Doudna, J.A., and Wang, H.W. (2013). Substrate-specific structural rearrangements of human Dicer. Nat. Struct. Mol. Biol. *20*, 662–670.

Tian, Y., Simanshu, D.K., Ma, J.B., Park, J.E., Heo, I., Kim, V.N., and Patel, D.J. (2014). A phosphate-binding pocket within the platform-PAZ-connector helix cassette of human Dicer. Mol. Cell *53*, 606–616.

Updike, D.L., Hachey, S.J., Kreher, J., and Strome, S. (2011). P granules extend the nuclear pore complex environment in the C. elegans germ line. J. Cell Biol. *192*, 939–948.

Valen, E., Preker, P., Andersen, P.R., Zhao, X., Chen, Y., Ender, C., Dueck, A., Meister, G., Sandelin, A., and Jensen, T.H. (2011). Biogenic mechanisms and utilization of small RNAs derived from human protein-coding genes. Nat. Struct. Mol. Biol. *18*, 1075–1082.

Wan, Y., Qu, K., Zhang, Q.C., Flynn, R.A., Manor, O., Ouyang, Z., Zhang, J., Spitale, R.C., Snyder, M.P., Segal, E., and Chang, H.Y. (2014). Landscape and variation of RNA secondary structure across the human transcriptome. Nature *505*, 706–709.

Welker, N.C., Maity, T.S., Ye, X., Aruscavage, P.J., Krauchuk, A.A., Liu, Q., and Bass, B.L. (2011). Dicer's helicase domain discriminates dsRNA termini to promote an altered reaction mode. Mol. Cell *41*, 589–599.

Zamudio, J.R., Kelly, T.J., and Sharp, P.A. (2014). Argonaute-bound small RNAs from promoter-proximal RNA polymerase II. Cell *156*, 920–934.

Zhang, H., Kolb, F.A., Brondani, V., Billy, E., and Filipowicz, W. (2002). Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP. EMBO J. *21*, 5875–5885.