

# Mapping the miRNA interactome by cross-linking ligation and sequencing of hybrids (CLASH)

Aleksandra Helwak & David Tollervey

Wellcome Trust Centre for Cell Biology, The University of Edinburgh, Edinburgh, UK. Correspondence should be addressed to A.H. (olahlwak@yahoo.com) or D.T. (d.tollervey@ed.ac.uk).

Published online 27 February 2014; doi:10.1038/nprot.2014.043

**RNA-RNA interactions have critical roles in many cellular processes, but studying them is difficult and laborious. Here we describe an experimental procedure, termed cross-linking ligation and sequencing of hybrids (CLASH), which allows high-throughput identification of sites of RNA-RNA interaction. During CLASH, a tagged bait protein is UV-cross-linked in cell cultures to stabilize RNA interactions, and it is purified under denaturing conditions. RNAs associated with the bait protein are partially truncated, and the ends of RNA duplexes are ligated together. After linker addition, cDNA library preparation and high-throughput sequencing, the ligated duplexes give rise to chimeric cDNAs, which unambiguously identify RNA-RNA interaction sites independent of bioinformatic predictions. This protocol is optimized for studying miRNA targets bound by Argonaute (AGO) proteins, but it should be easily adapted for other RNA-binding proteins and classes of RNA. The protocol requires ~5 d to complete, excluding the time required for high-throughput sequencing and bioinformatic analyses.**

## INTRODUCTION

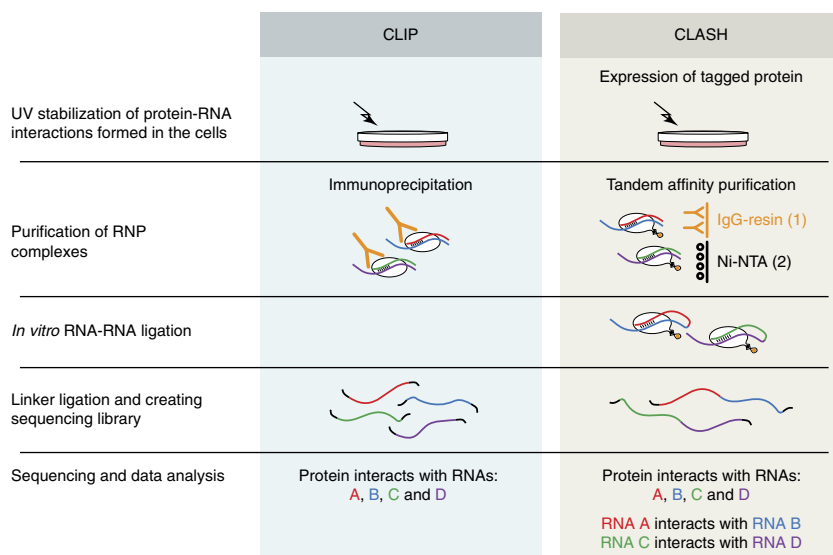
The crucial role of RNA interactions in many cellular processes, including translation and splicing, has long been well established. However, interest in the identification of targets for RNA interactions has been increased by a slew of recent discoveries, including the regulatory potential of miRNAs and long noncoding RNAs (lncRNAs) and the direct participation of RNA in transcription regulation (for recent reviews see refs. 1–3). Moreover, advances in sequencing technology have highlighted the startling complexity of the transcriptome and revealed ever-increasing numbers of different RNA species. Deep sequencing has revealed that most of the eukaryotic genome is transcribed<sup>4</sup>, but the biological functions of the majority of transcripts remain unclear.

Many RNA species function via interactions with proteins and/or other RNA molecules. A key advance in high-throughput analyses of ribonucleoprotein (RNP) complexes was the development of the CLIP technique (cross-linking and immunoprecipitation) by the Darnell laboratory<sup>5</sup>. CLIP is based on the stabilization of protein-RNA complexes in their cellular environment by UV cross-linking, immunoprecipitation of RNPs and the isolation and sequencing of bound RNAs. Many modifications of CLIP have subsequently been reported, including HITS-CLIP (high-throughput sequencing of RNA isolated by CLIP)<sup>6</sup>, PAR-CLIP (photoactivatable-ribonucleoside-enhanced CLIP)<sup>7</sup> and iCLIP (individual-nucleotide resolution UV CLIP)<sup>8</sup>. Originally applied to the neuronal splicing factor Nova, CLIP-based techniques have allowed the characterization of the RNA targets for many proteins. Although all these techniques directly reveal only protein-RNA interactions, the data can greatly aid bioinformatic predictions of RNA-RNA interactions. The first experimentally supported high-throughput prediction of miRNA targets was achieved by HITS-CLIP analyses of the AGO family of miRNA-binding proteins<sup>9</sup>.

The CLASH technique described here is the first reported high-throughput method for the direct identification of RNA-RNA interactions. It was derived from a further modified version of CLIP called CRAC (UV cross-linking and analysis of cDNAs)<sup>10</sup>. In CRAC, the immunoprecipitation step is replaced by purification

of tandem tagged proteins on an affinity resin. This eliminates the necessity for high-affinity antibodies and allows for very stringent purification of RNPs under denaturing conditions. This ensures a low recovery of background RNA and allows the analysis of individual proteins within very stable complexes (e.g., pre-ribosomes) that would resist the semidenaturing conditions used in CLIP approaches. In the course of CRAC experiments performed on the proteins binding to small nucleolar RNA (snoRNA)-rRNA duplexes, we observed that a small proportion of the RNA-RNA hybrids were ligated together and gave rise to chimeric cDNAs in the sequencing data (Fig. 1). CLASH was initially developed as a bioinformatic pipeline for the analysis of these chimeras, and it allowed the direct identification of interacting RNA molecules and their structures<sup>11</sup>. We subsequently enhanced the efficiency of chimera recovery in CLASH analyses by modifying the experimental protocol, with a crucial additional ligation step allowing enrichment for RNA-RNA chimeras. CLASH allows the identification of RNA-RNA interactions without bioinformatic predictions, and we recently applied it to the characterization of the human miRNA interactome<sup>12</sup>.

miRNAs are short noncoding RNAs that are involved in the post-transcriptional regulation of gene expression. Bound by AGO proteins, they direct effector RNP complexes to the target RNAs by means of complementarity. Each miRNA can bind many targets and each target can be regulated by multiple miRNAs, thus creating a huge network of interactions and mutual dependencies between RNA molecules. The patterns of miRNA-target mRNA base-pairing have been extensively studied, but many known interactions match poorly with the established canonical rules (for a review, see ref. 13). Defining more comprehensive miRNA targeting rules requires an experimental method that is independent of pre-existing knowledge and assumptions. By using CLASH for this purpose, we identified more than 18,000 miRNA-mRNA interactions, revealing a broad picture of the miRNA interactome in HEK 293 cells<sup>12</sup>. This protocol is specifically adjusted for studying miRNA targets bound by AGO proteins. It should, however, be readily adjusted for other RNA-binding proteins and types of RNA, as discussed further below.



**Figure 1** | Comparison of the principles of the CLIP and CLASH protocols.

**Overview of the procedure**

Living cells expressing PTH (protein A–tobacco etch virus (TEV) cleavage site–6×His)-tagged AGO are UV-irradiated (254 nm) to covalently couple existing protein-RNA complexes. Cells are lysed, PTH-tagged AGO is stringently purified under denaturing conditions, protein-bound RNA-RNA duplexes are trimmed by RNases and interacting RNA strands are ligated together to form chimeric guide–target molecules (Fig. 2). All AGO-interacting RNAs are subsequently isolated and used to produce a cDNA library ready for Illumina sequencing.

Bioinformatic analyses of CLASH data provides two types of information: precise AGO-binding sites on RNAs (similar to information obtained from CRAC or CLIP methods) and RNA-RNA hybrids that are formed within the AGO RNA-binding pocket (specific to the CLASH technique).

**Advantages of the method**

CLASH is currently the only technique that allows for direct mapping of RNA-RNA interactions on a large scale. It has various key features:

- (i) It uses UV irradiation that specifically stabilizes direct protein-RNA interactions<sup>14</sup>. As it is applied to living cells, recovered complexes should represent a snapshot of physiological interactions. Only brief irradiation is required (~1 min), offering the future possibility of time-resolved studies and kinetic analyses.
- (ii) Protein-RNA complexes are purified stringently under denaturing conditions, which results in low background; more details on the background recovery in the method are provided in ref. 12.
- (iii) Several samples can be prepared in parallel, as the procedure is streamlined by the use of the small spin columns for enzymatic reactions instead of tubes.
- (iv) The CLASH method is flexible, and it should be easily adapted to the analysis of other RNA-RNA interactions.

**Limitations and future improvements**

Stringent purification of protein-RNA complexes is a crucial step of the protocol. CLASH, therefore, strictly depends on the presence of a tagged protein in the cells. To date, CLASH has been applied only in established cell cultures, but this limitation may be overcome by using viral vectors for primary cell or organ cultures or knock-in of tagged constructs for *in vivo* studies on animals. To allow more physiological studies, we and our collaborators are preparing a genetically modified mouse in which endogenous AGO2 is replaced with a tagged protein (FLAG-6×His).

Another limitation of the current protocol arises from the low efficiency of RNA-RNA ligation. The number of chimeric reads in the sequencing data is quite variable and usually lower than 2% of all sequencing reads. A good cDNA library obtained with the current protocol is expected to yield ~15,000 unique miRNA-

mRNA interaction sites. This is a substantial number, but it does not saturate miRNA interactions in cells. A reliable comparison of alterations in miRNA targets under changing physiological conditions would require more complete coverage. Obtaining even higher numbers of RNA-RNA hybrids from each sample should therefore be a priority in the next optimization stages.

**Applications and future uses of the method**

AGO-CLASH is the method of choice for looking for new phenomena in miRNA biology: observation of new patterns of target binding and identification of novel targets or RNA regulators. Additional modifications to increase the efficiency of RNA-RNA ligation (as discussed below) would facilitate a range of applications, such as studying the dynamics of miRNA interactions in changing physiological conditions, the influences of diverse miRNA modifications on targeting efficiency or comparing targets of various miRNA family members.

As CLASH uses a simple concept of creating chimeric RNAs from two interacting RNA molecules, it should be easily applied to studies of various biological processes that involve ternary complexes, e.g., RNA-binding protein and two interacting RNA molecules. Depending on the initial phosphorylation state or modifications of the ends of interacting RNA molecules, different orders of the enzymatic reactions should be considered (see ‘Preparing RNA ends for ligation’ below). In some cases, linker-mediated ligation may give better results than a direct guide:target RNA ligation strategy.

In the case of miRNAs, the interacting RNA molecules form a base-paired duplex that is buried within the AGO protein. However, this structural arrangement is not essential, as interactions between snoRNAs and their rRNA targets can be identified by CLASH using proteins that bind adjacent to the RNA duplex<sup>11</sup>. CLASH is particularly suited to analyses in which one protein binds to a wide variety of RNA-RNA duplexes. Possible future applications include the identification of targets for the

many lncRNAs, and interactions between miRNAs and lncRNAs have already been identified in this way<sup>12</sup>. Pre-mRNA packaging factors or hnRNP proteins might be suitable as bait proteins for these CLASH analyses. The large data sets generated are also likely to be of value in training future biophysical computational models of miRNA-mRNA interactions.

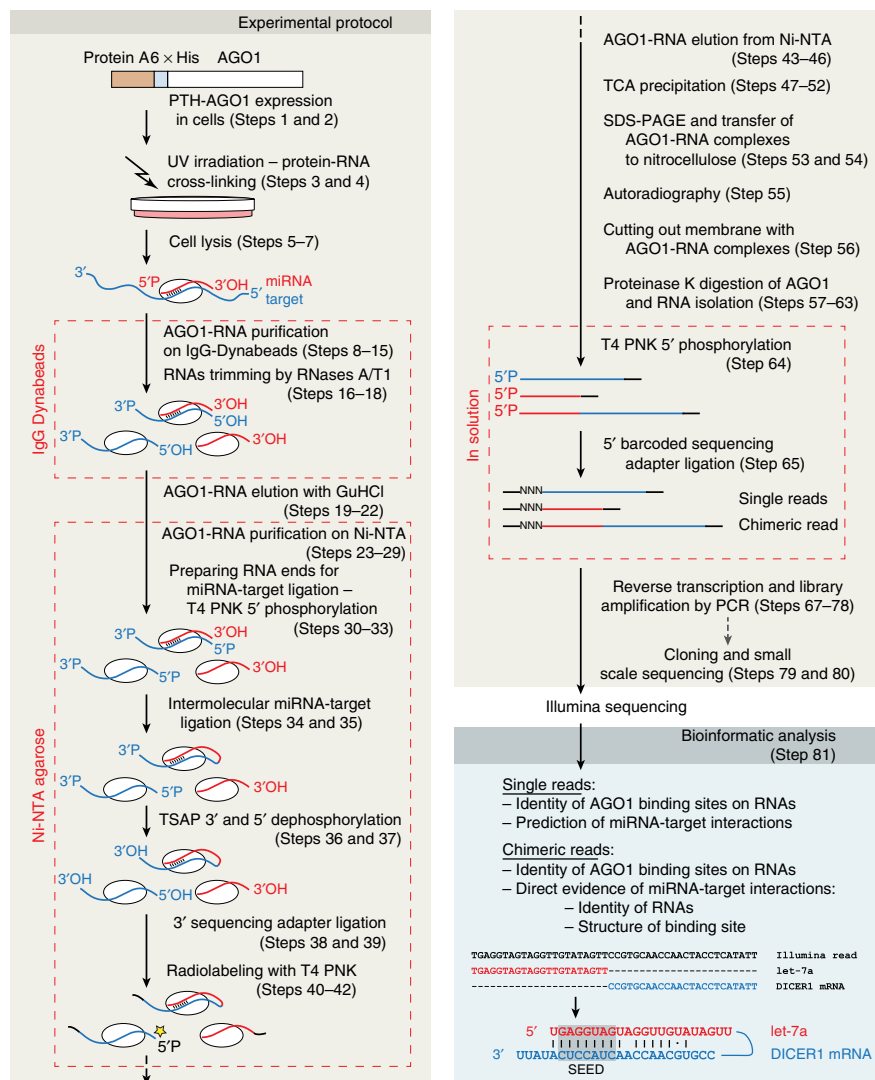
**Experimental design**

**Initial preparation steps: tagging of AGO1.** In ref. 12, we used the AGO1 protein with a tripartite tag (PTH) added at the N terminus. This consists of (i) two immunoglobulin-binding Z domains from *Staphylococcus aureus* protein A; (ii) a TEV protease cleavage site; and (iii) a 6×His-tag (Fig. 2). However, in the protocol, we do not include a TEV cleavage step for the following reasons. First, we observed that 6×His-AGO1 generated by the TEV cleavage is retained by the remaining protein A-IgG-Dynabeads complex. As endogenous AGO1 does not bind IgG-Dynabeads, this suggests that this retention is caused by a nonspecific binding of AGO1 to the cleaved protein A region of the tag, which is retained on the column. Second, performing RNA-RNA ligation directly after TEV cleavage, without AGO1 denaturation and additional stringent purification on Ni-NTA, substantially reduced ligation efficiency and increased the background of the method. Instead, we elute AGO1-RNA complexes from IgG-Dynabeads using denaturing conditions.

We believe that the only part of the PTH tag crucial for the success of the CLASH method is the 6×His-tag, and the first purification step (purification on IgG-Dynabeads, Steps 8–15) can be performed using any good efficiency tag in place of Protein A. A FLAG-PreScission cleavage site-6×His tag was successfully tested in the laboratory for proteins other than AGO1. Good antibodies should also be suitable for the first purification step, but this has not yet been confirmed.

**Initial preparation steps: cell culture.** The first stage of the CLASH technique, which is not described in detail in this protocol, is the preparation of cells expressing the tagged protein. In our experiments, we used the Flp-In T-REx 293 cell line stably transfected with pcDNA5/FRT/TO vector (Flp recombination target site, tetracycline-regulated TetO<sub>2</sub> promoter), expressing PTH-tagged human AGO1 upon induction with doxycycline (described in detail in ref. 12). However, transient transfection of 293 cells and possibly other cell lines is also a potential option.

**Purification of AGO complexes (Steps 8–15 and Steps 23–29).** Stringent purification of the AGO-RNA complexes before the



**Figure 2** | Overview of the experimental steps in the CLASH technique and the outcome of the basic bioinformatic analysis. Adapted from *Cell* (ref. 12). Copyright 2013, with permission from Elsevier.

RNA-RNA ligation step is very important for the success of the CLASH procedure. Protein-RNA cross-linking is not an efficient process, and we expect that almost all of the recovered interactions originate from complexes where one RNA strand is cross-linked to AGO and the other RNA molecule is maintained by RNA-RNA base-pairing. Therefore, one of the main concerns is to what extent the RNA-RNA interactions are stable throughout the procedure and how much background arises from random RNA-RNA interactions formed after cell lysis. For that purpose, we prepared a control experiment with mixed human and yeast lysates, looking for the inter-species chimeric sequences (described in more detail in ref. 12). For this protocol, the background appeared to be <2%; however, for the alternative protocol with less stringent purification (no guanidine hydrochloride wash before RNA-RNA ligation), it was close to 10%.

To minimize the background, we used highly denaturing 6 M guanidine hydrochloride in this protocol, possibly sacrificing the numbers of identified RNA-RNA interactions and creating bias



## PROTOCOL

toward more stable ones. However, a higher recovery of RNA-RNA interactions at the cost of reduced specificity might be achieved by careful optimization of the Ni-NTA wash conditions, if this was desired for other applications.

**RNase digestion (Steps 16–18).** A crucial step in the CLASH procedure is fragmentation of protein-bound RNAs with RNases. In our study, we used a mixture of RNase A and T1; however, other RNases are probably also suitable, and they have been used successfully in the related CLIP and PAR-CLIP techniques<sup>15</sup>. Overdigestion with RNase can generate very short RNA fragments, which are less readily mapped unambiguously to the genome or transcriptome database after cloning and sequencing, and it can lead to RNA cleavage within protein-binding sites and to biased results<sup>15</sup>. Conversely, long chimeric cDNAs may not be fully sequenced, potentially leading to one fragment failing to be mapped. Moreover, the non-base-paired loop region between two ligated RNA molecules is usually shorter than 10 nt (ref. 12), suggesting that fragments with short single-stranded overhangs might be preferentially ligated. Suboptimal RNase treatment can also influence other steps in the procedure, e.g., long RNAs can bind directly to the nitrocellulose.

Commercial RNase preparations are very concentrated, and it can be difficult to obtain a reproducible dilution directly from the original stock. It is therefore convenient to prepare a larger volume of prediluted stock (we use 1:20 dilution in water) before optimization of final RNase concentration and incubation time; do not change it between experiments.

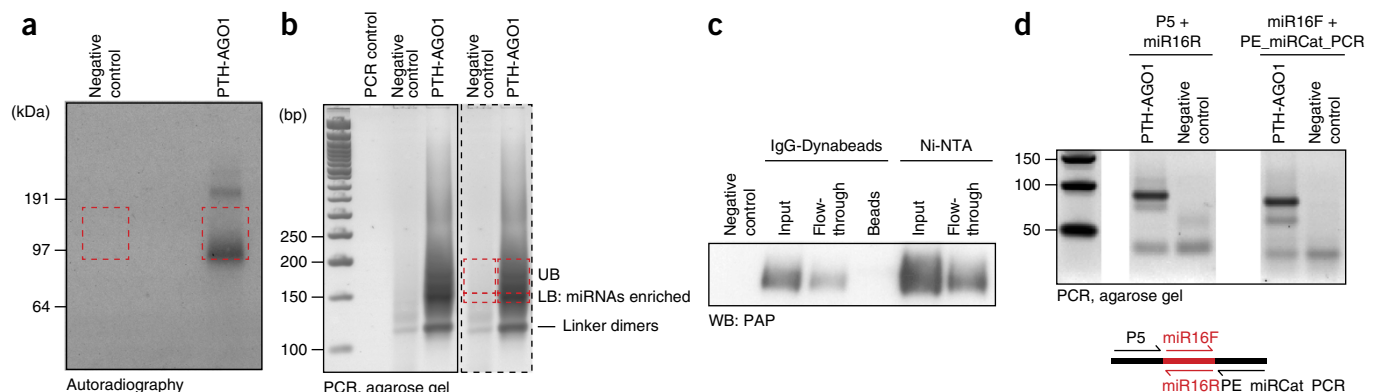
**Oligonucleotide design.** The 5' adapters for cDNA cloning carried a 5' blocking group (inverted dideoxythymidine; invddT) followed by four DNA residues, an RNA sequence complementary to primers present on Illumina sequencing plates, three random ribonucleotides to allow PCR duplicates to be identified and a 5–7 ribonucleotide bar code to allow multiple samples to be sequenced together. The L5 adapters and P5 PCR primer described here are generally applicable to various bait proteins and RNA types. However, any other adapter-primer set used

successfully in similar experiments and Illumina sequencing could probably be used.

**Preparing RNA ends for ligation (Steps 30–33).** In the AGO-CLASH protocol, we take advantage of the short length of miRNAs that tend not to be trimmed by RNases and therefore retain their natural 3'-OH ends. These miRNAs do not have to be enzymatically modified for RNA-RNA ligation (the expected phosphorylation state of RNA molecules along the CLASH protocol is depicted in Fig. 2).

Trimming of longer RNA molecules by either RNase A or T1 leaves ends (3'P and 5'OH) that are incompatible with ligation. The order of enzymatic repair reactions should therefore be reversed (5' and 3' thermosensitive alkaline phosphatase (TSAP) dephosphorylation first, then 5' T4 polynucleotide kinase (PNK) phosphorylation). Short RNAs with unusual end modifications may require the inclusion of additional enzymatic steps. Note that, although this protocol is predicted to yield only chimeras with the miRNA sequence positioned 5' to the mRNA-derived sequence (Fig. 2), the opposite orientation was also recovered at lower frequency.

**CLASH controls.** In CLASH experiments, three kinds of controls are useful: (i) a negative control sample prepared from the cell line without tagged AGO protein. This is an important control that indicates the amount of RNA background. This control sample should give little, if any, signal by autoradiography (Fig. 3a), and it should not give rise to much PCR product (Fig. 3b). However, this is not a very good negative control for the analysis of sequencing data. (ii) A useful negative control for the bioinformatic analysis is a sample prepared from a cell line that carries a different tagged protein, preferentially one that is involved in a different biological process. This can give information on the common, unspecific RNA contaminants. (iii) A control sample prepared with mixed lysates of different species (see 'Purification of AGO1 complexes' above and ref. 12 for more details). This is an especially useful sample when modifying the experimental protocol described here, as it gives information on the recovery of unspecific RNA-RNA complexes that do not originate from physiological interactions.



**Figure 3** | Examples of gel analyses at different steps in the CLASH protocol. (a) SDS-PAGE purification of the AGO1 protein, visualized by autoradiography of the radioactively labeled, and cross-linked RNA (Step 55). Red dashed lines indicate the regions excised for elution of the RNA-protein complexes. (b) MetaPhor agarose separation of the PCR-amplified DNA library ready for sequencing (Step 76). Gel in the dashed box is a fragment of the original scan (solid box) repeated for clarity of the picture. Red dashed lines indicate the regions excised for preferential recovery of miRNAs (LB) and chimeric cDNAs and targets (UB). (c) Recovery of the tagged AGO1 on IgG-Dynabeads and Ni-NTA (Step 25); western blot with peroxidase-anti-peroxidase (WB: PAP) soluble complex antibody recognizing Protein A tag; beads: residual AGO1 after elution; eluate fraction from IgG-Dynabeads is coinciding with Ni-NTA input. (d) MetaPhor agarose separation of the control RT-PCR products generated with miR-16-specific primers.



**Bioinformatic analysis (Step 81).** Basic bioinformatic analysis of CLASH data involves the following: (i) pre-processing of Illumina sequence reads (i.e., quality control and linker trimming); (ii) mapping reads to the transcriptomic or genomic database; (iii) identifying the chimeric reads defined as two high-quality hits to distinct regions of the database; and (iv) determining the

RNA-RNA interaction sites and base-pairing by the RNA secondary structure prediction algorithms.

The bioinformatic pipeline for analysis of the AGO1 CLASH data was described in ref. 12. More detailed descriptions of the bioinformatics, further discussion of variants in the analyses and their influence on its outcome are described in ref. 16.

## MATERIALS

### REAGENTS

**▲ CRITICAL** If you are planning to use a different supplier than that suggested, check that the enzymes are not His-tagged (information that is not usually included in the enzyme information sheet). Be aware that commercial enzymes are frequently contaminated with RNA derived from the organism used for recombinant protein production. Although it is possible to remove contaminating RNAs by guanidine hydrochloride washes on the Ni-NTA beads, enzymatic reactions performed in solution can be substantially affected.

**▲ CRITICAL** All the reagents used in Steps 17–71 should be RNase free. Diethylpyrocarbonate (DEPC) treatment is not required.

- PTH-AGO1-Flp-InT-REx 293 cells created from the Flp-In T-REx 293 cell line (Life Technologies, cat. no. R780-07) stably transfected with pcDNA5/FTR/TO vector (Life Technologies, V6520-20) expressing PTH-tagged human AGO1 (described in detail in ref. 12)
- DMEM with 4.5 mg per liter glucose (Life Technologies, cat. no. 41966)
- FBS (Sigma-Aldrich, cat. no. F2442)
- Dulbecco's PBS (DPBS; Life Technologies, cat. no. 14190)
- Doxycycline (Sigma-Aldrich, cat. no. D9891) **! CAUTION** This reagent is harmful if swallowed.
- Peroxidase-anti-peroxidase soluble complex antibody (Sigma-Aldrich, cat. no. P1291)
- Dynabeads M-270 epoxy (Life Technologies, cat. no. 14301)
- IgG antibody from rabbit serum (Sigma-Aldrich, cat. no. I5006)
- rATP, 100 mM (Promega, cat. no. E6011). Prepare aliquots and store them at  $-20^{\circ}\text{C}$  for at least 1 year **▲ CRITICAL** Avoid repeated thawing and freezing.
- ATP, 10 mM (New England BioLabs, supplied with T4 RNA ligase 1). Store it at  $-20^{\circ}\text{C}$ , or for long-term storage (more than 1 month), store it at  $-70^{\circ}\text{C}$ .
- BSA (Sigma-Aldrich, cat. no. A3294) dissolved in water at 10 mg ml<sup>-1</sup>
- Guanidine hydrochloride (Sigma-Aldrich, cat. no. G4505) **! CAUTION** This reagent is harmful if swallowed.
- Ni-nitrilotriacetic acid (NTA) Superflow beads (Qiagen, cat. no. 30410)
- RNasin (Promega, cat. no. N2111)
- PEG 8000 (Sigma-Aldrich, cat. no. P1458)
- miRCat-33 conversion oligos pack (Illumina-compatible 3' adapter and reverse transcription (RT) primer; Integrated DNA Technologies, cat. no. 51-01-13-10); stock concentration = 10  $\mu\text{M}$  in water. After dissolving, prepare aliquots and store them at  $-80^{\circ}\text{C}$  for at least 1 year **▲ CRITICAL** Always use fresh aliquots.
- Illumina-compatible L5 adapters (r stands for ribonucleotide, rN indicates a random ribonucleotide): 5'-invddT-ACACrGrArCrGrCrUrCrUrCrCrGrArUrCrUr-rNrNrN-bar code-3', list of bar codes used: -L5Aa: 5'-rUrArArGrC-3'OH; L5Ab: 5'-rArUrUrArGrC-3'OH; L5Ac: 5'-rGrCrGrCrArGrC-3'OH; L5Cc: 5'-rArCrTrCrArGrC-3'OH; L5Cd: 5'-rGrArCrTrArGrC-3'OH (Integrated DNA Technologies, custom order); stock concentration = 100  $\mu\text{M}$  in water. Prepare aliquots after dissolving. They can be stored at  $-80^{\circ}\text{C}$  for at least 1 year **▲ CRITICAL** Always use fresh aliquots of the linker.
- P5 PCR primer: 5'-AATGATACGGCACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTCCGATCT-3' (Integrated DNA Technologies, custom order), stock concentration = 10  $\mu\text{M}$  in water. After dissolving, prepare aliquots and store them at  $-20^{\circ}\text{C}$  for at least 3 years
- PE\_miRCat\_PCR primer: 5'-CAAGCAGAAGACGGCATAACGAGATCGGTCTGGCATTCTCTGGCCTTGGCACCCGAGAATTCC-3' (Integrated DNA Technologies, custom order), stock concentration = 10  $\mu\text{M}$  in water. After dissolving, prepare aliquots and store them at  $-20^{\circ}\text{C}$  for at least 3 years

- miR-16-specific PCR primers: miR-16F: 5'-GCAGCACGTAAATATTGGCG-3'; miR-16R: 5'-GCCAATATTACGTGCTGCTA-3' (Integrated DNA Technologies, custom order)
- <sup>32</sup>P- $\gamma$ -ATP (PerkinElmer, 6000 Ci mmol<sup>-1</sup>, NEG502Z) **! CAUTION** This is a radioactive material, and it is especially dangerous when ingested or inhaled. Use protective clothing and Plexiglas shielding, and be sure to monitor for surface contamination.
- Hybond-C Extra membrane (GE Healthcare, cat. no. RPN303E)
- Kodak BioMax MS autoradiography film (Kodak, cat. no. 8222648)
- MetaPhor agarose (Lonza, cat. no. 50180)
- SYBRSafe (Life Technologies, S33102)
- cOmplete Protease inhibitors, EDTA free (Roche Applied Science, cat. no. 11873580001)
- RNase-IT (Agilent, cat. no. 400720) diluted 1:20 with water; it can be stored at  $4^{\circ}\text{C}$  for at least 2 years
- SeeBlue Plus2 pre-stained standard (Life Technologies, cat. no. LC5925)
- NuPAGE LDS sample buffer 4 $\times$  (Life Technologies, cat. no. N0007)
- NuPAGE 4–12% (wt/vol) polyacrylamide Bis-Tris gels (Life Technologies, cat. no. NP0335)
- NuPAGE SDS-MOPS running buffer (Life Technologies, cat. no. NP0001)
- NuPAGE transfer buffer (Life Technologies, cat. no. NP00061)
- GlycoBlue (Life Technologies, cat. no. AM9515)
- MinElute PCR purification kit (Qiagen, cat. no. 28004)
- MinElute Gel extraction kit (Qiagen, cat. no. 28604)
- GeneRuler 50-bp DNA ladder (Thermo Scientific, cat. no. SM0371)
- Qubit dsDNA HS assay kit (Life Technologies, cat. no. Q232854)
- DNA loading dye, 6 $\times$  (Thermo Scientific, cat. no. R0611)
- T4 PNK, T4 polynucleotide kinase (New England BioLabs, cat. no. M0201L)
- T4 RNA ligase 1 (New England BioLabs, cat. no. M0204L)
- T4 RNA ligase reaction buffer, 10 $\times$  (supplied with T4 RNA ligase 1)
- T4 RNA ligase 2 truncated, K227Q (New England BioLabs, cat. no. M0351L)
- Thermosensitive alkaline phosphatase (TSAP; Promega, cat. no. M9910)
- Proteinase K (Roche Applied Science, cat. no. 03115836001)
- SuperScript III reverse transcriptase (Life Technologies, cat. no. 18080-044)
- First-strand buffer, 5 $\times$  (supplied with SuperScript III reverse transcriptase)
- DTT, 0.1 M (supplied with SuperScript III reverse transcriptase)
- RNase H (New England BioLabs, cat. no. M0297L)
- TaKaRa long and accurate (LA) Taq (Clontech, cat. no. RR002M)
- LA PCR buffer II, 10 $\times$  (Mg<sup>2+</sup> plus) supplied with TaKaRa LA Taq
- dNTPs, 2.5 mM (supplied with TaKaRa LA Taq). Prepare aliquots and store them at  $-20^{\circ}\text{C}$  for at least 1 year **▲ CRITICAL** Always use fresh aliquots.
- TOPO TA Cloning kit for sequencing, with One Shot TOP10 chemically competent *Escherichia coli* (Life Technologies, cat. no. K4575-40)
- Phenol (Sigma-Aldrich, cat. no. P4557, used without equilibration buffer) **! CAUTION** Phenol is toxic on inhalation, on contact with skin or if swallowed; it causes severe skin burns and eye damage. Use a hood, protective clothing, eye protection and gloves.
- Tris (Life Technologies, cat. no. 15504-020)
- HCl (Thermo Fisher Scientific, cat. no. 10000180)
- NaCl (Thermo Fisher Scientific, cat. no. 10326390)
- NP-40 (Roche Applied Science, cat. no. 11754599001)
- EDTA (Thermo Fisher Scientific, cat. no. 10213570)
- Glycerol (Thermo Fisher Scientific, cat. no. 10336040)
- $\beta$ -Mercaptoethanol (BME; Sigma-Aldrich, cat. no. M3148) **! CAUTION** This reagent is toxic if swallowed or if inhaled; it can be fatal on contact with skin, and it causes serious eye damage. Use a hood, protective clothing, eye protection and gloves. Avoid release to the environment.
- MgCl<sub>2</sub> (Sigma-Aldrich, cat. no. M8266)

## PROTOCOL

- Trichloroacetic acid, 100% (TCA; Sigma-Aldrich, cat. no. 91228)  
**! CAUTION** This reagent causes severe skin burns and eye damage. Wear protective clothing, eye protection and gloves.
- Acetone (Thermo Fisher Scientific, cat. no. 10162180)
- Imidazole (Sigma-Aldrich, cat. no. I2399)
- SDS (Sigma-Aldrich, cat. no. L4390)
- Methanol (Thermo Fisher Scientific, cat. no. 11976961)
- Ethanol (Hayman, cat. no. AR100-X)
- Boric acid (Thermo Fisher Scientific, cat. no. 10263370)
- Sodium acetate (Thermo Fisher Scientific, cat. no. 10122350)
- Chloroform (Thermo Fisher Scientific, cat. no. 10293850)
- Isoamyl alcohol (Sigma-Aldrich, cat. no. 19392)
- Liquid nitrogen

### EQUIPMENT

- Humidified 37 °C, 5% CO<sub>2</sub> incubator
- Cell culture hood
- Cell culture dishes, 150 mm (Thermo Scientific, cat. no. 157150)
- Crosslinker, Stratalinker 1800 (Stratagene) with UV bulbs,  $\lambda = 254$  nm (Ushio, cat. no. 3000016)
- Cell scrapers
- Filter units for buffer sterilization (Thermo Scientific, rapid-flow disposable filter units, CN membrane)
- Magnetic rack for 15-ml conical tubes (Life Technologies, Dynal MPC-15)
- Magnetic rack for microcentrifuge tubes (Life Technologies, cat. no. CS15000)
- Rotating wheel for microcentrifuge tubes
- Spin Columns with snap caps (Thermo Scientific, cat. no. 69725)
- Cooling plate for microcentrifuge tubes
- Thermoblock with shaking (Eppendorf Thermomixer comfort)
- Refrigerated centrifuge for conical tubes (Sorvall Legend RT)
- Refrigerated benchtop centrifuge (Eppendorf, cat. no. 5417R)
- SDS-PAGE tank XCell SureLock mini-cell (Life Technologies, cat. no. EI0001)
- Mini Trans-Blot electrophoretic transfer cell (Bio-Rad, cat. no. 170-3930)
- Power supply unit
- Phosphorescent rulers for autoradiography (Sigma-Aldrich, cat. no. R8133)
- Thermocycler (Bio-Rad, DNA Engine)
- Agarose gel electrophoresis cell
- Gel scanner (FLA-5100)
- Qubit fluorometer (Life Technologies, cat. no. Q32857)
- Film developer
- Vortex
- Filter tips (Starlab)
- Microcentrifuge tubes, 1.5 ml
- Conical tubes, 15 ml
- Disposable pipettes
- Pipette aid (e.g., Pipetboy)

- Radioactivity monitor
- Transparency film (e.g., 3 M CG6000)
- Scalpels
- Luria broth (LB) agar plates with ampicillin for bacterial selection

### REAGENT SETUP

**Phenol-chloroform-isoamyl (PCI) alcohol solution** Mix the reagents in volume proportions of 25:24:1. Store the solution at 4 °C until the color changes. Discard it if the color of the mixture changes.

**TBE buffer, 10×** TBE buffer (10×) is 890 mM Tris, 890 mM boric acid and 20 mM EDTA. Store it at room temperature (~25 °C). **▲ CRITICAL** All buffers listed below should be filter-sterilized before use, and, unless stated otherwise, they can be stored at 4 °C for at least 1 year.

**Lysis buffer** Mix 50 mM Tris-HCl (pH 7.8), 300 mM NaCl, 1% (vol/vol) NP-40 (use 50% stock), 5 mM EDTA (pH 8.0) and 10% (vol/vol) glycerol (use 50% stock) in deionized water. Add 5 mM BME and protease inhibitors just before use.

**Low-salt (LS)-IgG buffer** Mix 50 mM Tris-HCl (pH 7.8), 300 mM NaCl, 0.5% (vol/vol) NP-40, 2.5% (vol/vol) glycerol and 5 mM MgCl<sub>2</sub> in deionized water. Add 5 mM BME just before use.

**High-salt (HS)-IgG buffer** Mix 50 mM Tris-HCl (pH 7.8), 800 mM NaCl, 0.5% (vol/vol) NP-40, 2.5% (vol/vol) glycerol and 10 mM MgCl<sub>2</sub> in deionized water. Add 5 mM BME just before use.

**PNK-western blotting (WB) buffer** Mix 50 mM Tris-HCl (pH 7.8), 50 mM NaCl, 0.5% NP-40 (vol/vol) and 10 mM MgCl<sub>2</sub> in deionized water. Add 5 mM BME just before use.

**PNK buffer, 5×** Mix 250 mM Tris-HCl (pH 7.5), 250 mM NaCl, 2.5% (vol/vol) NP-40, 50 mM MgCl<sub>2</sub> and 50 mM BME. After filtering, divide the buffer into small portions and keep them frozen at -20 °C.

**Ni-WBI (wash) buffer** Mix 50 mM Tris-HCl (pH 7.8), 300 mM NaCl, 0.1% (vol/vol) NP-40, 10 mM imidazole (pH 8.0) and 6 M guanidine hydrochloride in deionized water. Add 5 mM BME just before use. Protect the buffer from light.

**Ni-WBII (wash buffer)** Mix 50 mM Tris-HCl (pH 7.8), 300 mM NaCl, 0.1% (vol/vol) NP-40 and 10 mM imidazole (pH 8.0) in deionized water. Add 5 mM BME just before use. Protect the buffer from light.

**Ni-elution buffer (EB)** Mix 50 mM Tris-HCl (pH 7.8), 50 mM NaCl, 0.1% (vol/vol) NP-40 and 150 mM imidazole (pH 8.0) in deionized water. Add 5 mM BME just before use. Protect the buffer from light.

**Proteinase K buffer** Mix 50 mM Tris-HCl (pH 7.8), 50 mM NaCl, 0.1% (vol/vol) NP-40, 10 mM imidazole (pH 8.0), 1% (wt/vol) SDS (use 10% stock), 5 mM EDTA (pH 8.0) and 5 mM BME in deionized water. Protect the buffer from light.

**IgG-coated Dynabeads** IgG-coated Dynabeads are prepared according to the protocol published in ref. 17.

### PROCEDURE

#### Cell culture ● TIMING 48 h

1| Seed PTH-AG01-Flp-In T-REx 293 or control Flp-In T-REx 293 cells onto 150-mm dishes at ~20% confluency and grow them overnight in DMEM with 10% (vol/vol) FBS.

2| The next day, induce the cells for PTH-AG01 production with doxycycline (final concentration 0.5  $\mu\text{g ml}^{-1}$ ) for 36 h. Use four plates per CLASH sample.

#### UV cross-linking ● TIMING ~10 min per CLASH sample

3| At this stage, the cells should be ~90% confluent. Briefly wash the growing cells once with room-temperature DPBS.

4| Place the culture plates without lids on an ice tray and into the Stratalinker. Irradiate at  $\lambda = 254$  nm with a power setting of 400 mJ cm<sup>-2</sup>.

**▲ CRITICAL STEP** Proceed with one plate at a time to reduce the time between DPBS washing and cell lysis, thus limiting cell exposure to changed conditions.

**Cell lysis ● TIMING 1 h**

5| Directly after cross-linking, lyse the cells on the dish by adding 2.5 ml of ice-cold lysis buffer; incubate the cells for 5 min on ice.

6| When all the dishes are ready, scrape the cells and collect the pooled lysates into 15-ml conical tubes.

7| Incubate the samples for a further 10 min on ice, and then centrifuge them at 6,400g for 30 min at 4 °C. Prepare the WB sample ('IgG-Dynabeads input') by boiling 15 µl of lysate with 5 µl of NuPAGE LDS sample buffer for 5 min; store it at -20 °C until further use at Step 25.

■ **PAUSE POINT** Cell lysates can either be used directly or divided into aliquots, frozen in liquid nitrogen and stored at -80 °C for up to 6 months.

**IgG-Dynabeads loading and washing ● TIMING 90 min**

8| Wash the IgG-coated Dynabeads three times with 5 ml of ice-cold DPBS.

▲ **CRITICAL STEP** Washing the beads directly before loading is essential for the removal of IgG released from the beads during storage. Free IgG competes with IgG-coated beads, which decreases the AGO-binding efficiency.

9| Add 20 mg of IgG-Dynabeads to each sample of freshly prepared or thawed lysates from Step 7. Incubate the samples with rotation for 40 min at 4 °C.

▲ **CRITICAL STEP** Frozen lysates should be thawed slowly on ice.

10| Place the conical tubes in the magnetic rack for ~30 s until the beads settle and the supernatant becomes clear. Carefully remove the supernatant. Prepare the WB sample ('IgG-Dynabeads flow through') by boiling 15 µl of supernatant with 5 µl of NuPAGE LDS sample buffer for 5 min; store it at -20 °C until further use at Step 25. Discard the remaining supernatant.

11| Wash the beads briefly by resuspending them in 10 ml of ice-cold LS-IgG buffer, allow the beads to pellet in the magnetic rack and discard the supernatant.

12| Wash the beads twice with ice-cold HS-IgG buffer, this time incubating them for 5 min in the buffer with rotation at 4 °C. Pellet the beads and discard the supernatant.

13| Wash the beads once briefly with 10 ml of ice-cold LS-IgG buffer. Pellet the beads and discard the supernatant.

14| Wash the beads once briefly with 10 ml of ice-cold PNK-WB buffer. Pellet the beads and discard the supernatant.

15| Gently resuspend the beads with a pipette in 1 ml of ice-cold PNK-WB buffer, and transfer them to 1.5-ml microcentrifuge tubes. Pellet the beads in the magnetic rack and discard the supernatant.

**RNase digestion on Dynabeads ● TIMING 15 min**

16| Prepare the RNase digestion buffer by adding 1 µl of diluted RNase-IT (0.5 U) to 500 µl of prewarmed PNK-WB buffer per sample.

▲ **CRITICAL STEP** This step might require optimization; more details are given in the Experimental design section.

17| Add prepared RNase digestion buffer from Step 16 to the IgG-Dynabeads from Step 15 and incubate for 7 min at 20 °C.

▲ **CRITICAL STEP** During the digestion, tap the samples gently every 30 s to prevent Dynabeads from settling.

18| Place the samples briefly on ice to slow down the reaction, and then settle the beads on the magnetic rack. Remove the RNase digestion buffer completely and discard it.

**Elution from IgG-Dynabeads ● TIMING 40 min**

19| Add 500 µl of Ni-WBI buffer to the beads and incubate the mixture for 10 min at room temperature with rotation.

20| Pellet the beads and collect the supernatant.

21| Repeat the elution twice more, once with 500 µl and once with 1 ml of Ni-WBI buffer. Pool all the eluate fractions together. Prepare the WB sample ('Ni-NTA input') as follows: 10 µl eluate + 90 µl water + 100 µl 10% (wt/vol) TCA, incubate it for 40 min on ice and then proceed as in Steps 48-52. Store the sample at -20 °C until further use at Step 25.

## PROTOCOL

22| Wash the Dynabeads with 1 ml of PNK-WB and resuspend the beads in 1 ml of PNK-WB. Prepare the WB sample ('IgG-Dynabeads beads') by mixing 15  $\mu\text{l}$  of IgG-Dynabeads suspension + 135  $\mu\text{l}$  water + 50  $\mu\text{l}$  NuPAGE LDS sample buffer, boil it for 5 min and store it at  $-20\text{ }^{\circ}\text{C}$  until further use at Step 25. Use 20  $\mu\text{l}$  per gel lane.

### Ni-NTA agarose loading and washing ● TIMING 3 h, plus 6 h for WB (optional)

23| Wash Ni-NTA beads twice with Ni-WBI buffer and prepare a 50% bead suspension in Ni-WBI buffer for all the samples together. Add 80  $\mu\text{l}$  of prepared Ni-NTA suspension to each eluate from Step 21 and incubate for 2 h at  $4\text{ }^{\circ}\text{C}$  with rotation.

24| Spin the beads at 1,000g for 10 s at  $4\text{ }^{\circ}\text{C}$  and gently aspirate the supernatant. Prepare the WB sample ('Ni-NTA flow through') by using 10  $\mu\text{l}$  of supernatant, as described in Step 21.

25| Perform WB (optional) to estimate AGO recovery during purification steps, by using WB samples prepared in Steps 7, 10, 21, 22 and 24. SDS-PAGE and transfer procedures are described in Steps 53 and 54. For the tagged AGO detection, use peroxidase-anti-peroxidase soluble complex antibody (prepare a 1:10,000 dilution in 5% nonfat dry milk in DPBS (wt/vol) and incubate the membrane for 1 h at room temperature with rotation; detect peroxidase coupled antibody using enhanced chemiluminescence (ECL) solution).

26| Wash the beads twice with 1 ml of ice-cold Ni-WBI for 10 min at  $4\text{ }^{\circ}\text{C}$  with rotation.

27| Transfer the beads to spin columns. Keep the columns in 1.5-ml microcentrifuge tubes.

28| Wash the beads twice with 0.75 ml of ice-cold Ni-WBII, and discard the supernatant. The same wash collection tubes can be used multiple times throughout the procedure.

▲ **CRITICAL STEP** All washes are carried out under gravity flow. Close the lid to get rid of the remaining buffer.

▲ **CRITICAL STEP** It is very important to carefully rinse the column with Ni-WBII. The Ni-WBI used in Step 26 contains guanidine hydrochloride, and any remains can influence the subsequent enzymatic reactions.

29| Wash the beads three times with 0.75 ml of ice-cold PNK-WB buffer and discard the supernatant. Dry the beads by closing the lid of the column, and subsequently close the bottom of the column with the supplied snap cap.

### RNA phosphorylation with T4 PNK ● TIMING 3 h

30| To each column, add 80  $\mu\text{l}$  of PNK reaction mixture, prepared in bulk for all the samples together; the volumes per sample are given below. Incubate the columns for 2.5 h at  $20\text{ }^{\circ}\text{C}$ .

Reagent	Volume ( $\mu\text{l}$ ) per reaction	Final concentration
Water	57	
PNK buffer, 5 $\times$	16	
rATP, 100 mM	0.8	1 mM
RNasin	2	1 U $\mu\text{l}^{-1}$
T4 PNK	4	0.5 U $\mu\text{l}^{-1}$
Mix total volume	80	

31| Remove the snap cap and wash the beads twice with 0.5 ml of ice-cold buffer Ni-WBI; discard the washes.

▲ **CRITICAL STEP** To prevent quick dripping from the column, open the lid first, then open the snap cap and quickly place the column in the microcentrifuge collection tube.

▲ **CRITICAL STEP** To limit the column contact area with Ni-WBI, add the buffer directly to the bottom of the column. Ni-WBI has to be rinsed off well, as it contains guanidine hydrochloride, which can affect the subsequent enzymatic reactions.

32| Wash the beads twice with 0.75 ml of ice-cold buffer Ni-WBII and discard the washes.

33| Wash the beads 3 times with 0.75 ml of ice-cold buffer PNK-WB and discard the washes.



**Inter-molecular RNA-RNA ligation ● TIMING overnight**

**34|** Close the bottom of the column with a snap cap, and add 160  $\mu\text{l}$  of ligation mixture prepared in bulk for all the samples together; the volumes per sample are given below. Incubate the samples overnight at 16 °C with gentle rotation.

Reagent	Volume ( $\mu\text{l}$ ) per reaction	Final concentration
Water	118.4	
PNK buffer, 5 $\times$	32	
rATP, 100 mM	1.6	1 mM
RNasin	4	1 U $\mu\text{l}^{-1}$
T4 RNA ligase 1	4	0.25 U $\mu\text{l}^{-1}$
Mix total volume	160	

**35|** Wash the beads as in Steps 31–33.

**Dephosphorylation with TSAP ● TIMING 1 h**

**36|** Close the bottom of the column with a snap cap, and add 80  $\mu\text{l}$  of reaction mixture prepared in bulk for all the samples together; the volumes per sample are given below. Incubate the samples for 45 min at 20 °C.

Reagent	Volume ( $\mu\text{l}$ ) per reaction	Final concentration
Water	54	
PNK buffer, 5 $\times$	16	
RNasin	2	1 U $\mu\text{l}^{-1}$
TSAP	8	0.1 U $\mu\text{l}^{-1}$
Mix total volume	80	

**37|** Wash the beads as in Steps 31–33.

**3'-end adapter ligation on beads ● TIMING 6.5 h**

**38|** Close the bottom of the column with a snap cap, and add 80  $\mu\text{l}$  of ligation mixture prepared in bulk for all the samples together; the volumes per sample are given below. Incubate the samples for 6 h at 16 °C.

Reagent	Volume ( $\mu\text{l}$ ) per reaction	Final concentration
Water	18	
PNK buffer, 5 $\times$	16	
PEG 8000, 25% (wt/vol)	32	10%
miRCat-33 3'-linker, 10 $\mu\text{M}$	8	1 $\mu\text{M}$
RNasin	2	1 U $\mu\text{l}^{-1}$
T4 RNA ligase 2, truncated, K227Q	4	10 U $\mu\text{l}^{-1}$
Mix total volume	80	

**39|** Wash the beads as in Steps 31–33.

## PROTOCOL

### Radioactive labeling of RNA ● TIMING 1 h 15 min

40| Close the bottom of the column with a snap cap, and add 80  $\mu\text{l}$  of reaction mixture prepared in bulk for all the samples together; the volumes per sample are given below. Incubate the samples for 30 min at 37 °C.

Reagent	Volume ( $\mu\text{l}$ ) per reaction	Final concentration
Water	55	
PNK buffer, 5×	16	
$^{32}\text{P}$ - $\gamma$ -ATP	3	
RNasin	2	1 U $\mu\text{l}^{-1}$
T4 PNK	4	0.5 U $\mu\text{l}^{-1}$
Mix total volume	80	

**! CAUTION** Always be careful when you are working with radioactive material. Consult local and institutional rules for detailed guidelines.

41| Wash the beads extensively with Ni-WBI; each wash volume is 0.5 ml.

▲ **CRITICAL STEP** Wash the beads multiple times at this stage to remove free  $^{32}\text{P}$ - $\gamma$ -ATP; reduce the risk of exposure to radioactivity and contamination. Monitor the radioactivity in the flow-through and repeat washes until this falls to ~10 counts per minute (c.p.m.).

42| Wash the beads as in Steps 31–33.

### Elution of PTH-AGO-RNA complexes from Ni-NTA beads ● TIMING 30 min

43| Add 200  $\mu\text{l}$  of Ni-EB buffer to the beads, close the snap cap and incubate in thermoblock setup for 5 min at room temperature with gentle mixing.

44| Collect the supernatant.

45| Repeat the elution twice more, once with 200  $\mu\text{l}$  and once with 600  $\mu\text{l}$  of Ni-EB buffer.

46| Pool all the eluate fractions together (total volume ~1 ml).

### TCA precipitation ● TIMING 2 h

47| Add 2  $\mu\text{g}$  of BSA and 200  $\mu\text{l}$  of 100% TCA and incubate it for 40 min on ice.

48| Centrifuge the samples in a tabletop centrifuge at 20,000g for 30 min at 4 °C.

49| Discard the supernatant.

**! CAUTION** Note that at this stage the supernatant can still be radioactive, as it contains unincorporated  $^{32}\text{P}$ - $\gamma$ -ATP.

50| Wash the pellets containing protein-RNA complexes twice with 1 ml of ice-cold acetone. Vortex the samples well during each wash, and spin the pellets with maximum speed (20,000g) for 10 min at 4 °C. Discard the supernatants.

51| Dry the pellets at room temperature (~5 min; or at 37 °C, but do not overdry them).

52| Resuspend the pellets in 15  $\mu\text{l}$  of water, add 5  $\mu\text{l}$  of 4 × NuPAGE LDS sample buffer and heat the samples for 10 min at 65 °C.

▲ **CRITICAL STEP** Protein pellets after TCA precipitation are easier to resuspend in water than in prediluted 1× NuPAGE LDS sample buffer.

▲ **CRITICAL STEP** The NuPAGE buffer system is quite resistant to small changes in the pH of the samples, but carefully observe the color of the samples in the NuPAGE LDS sample buffer. If it changes, add 1  $\mu\text{l}$ , or more, of 1 M Tris (pH 7.8) until the color comes back to normal.

■ **PAUSE POINT** Prepared samples can be frozen at  $-20\text{ }^{\circ}\text{C}$  for a few days.

? **TROUBLESHOOTING**

**SDS-PAGE and transfer of protein-RNA complexes** ● **TIMING 4.5 h**

53| Thaw the samples, separate the proteins and prestained protein standard on the 4–12% (wt/vol) polyacrylamide Bis-Tris NuPAGE gel with NuPAGE SDS-MOPS running buffer; run the samples at a constant voltage (150 V) until the dye reaches the bottom of the gel (~30 min).

▲ **CRITICAL STEP** It is important to use NuPAGE SDS-MOPS buffer as, unlike homemade buffers, it does not change its pH substantially during the run. An increase in buffer pH can cause RNA degradation.

▲ **CRITICAL STEP** To reduce the risk of sample cross-contamination, separate neighboring samples with empty lanes.

? **TROUBLESHOOTING**

54| Transfer gel-separated protein-RNA complexes to the nitrocellulose membrane in the cooled wet-transfer tank by using NuPage transfer buffer supplemented with 10% (vol/vol) methanol for 2 h at constant voltage (100 V).

55| Air-dry the nitrocellulose membrane for 30 s, wrap it in cling film, attach the phosphorescent ruler and expose it on the autoradiography film for 1 h (or longer, as needed).

▲ **CRITICAL STEP** Attaching the phosphorescent ruler to the cling film before exposition is crucial, as it is the only way to precisely align the nitrocellulose membrane with the exposed autoradiography film.

■ **PAUSE POINT** The nitrocellulose membrane can be frozen at  $-20\text{ }^{\circ}\text{C}$  for at least a week.

? **TROUBLESHOOTING**

**RNA purification** ● **TIMING 5 h to overnight, plus 4 h**

56| Carefully align the nitrocellulose membrane with the developed autoradiography film by using the signal from the phosphorescent ruler. Cut out the radioactive bands with disposable scalpels according to the instructions included in **Figure 3a**.

57| Place the cut-out membrane pieces in separate 1.5-ml microcentrifuge tubes. Add 400  $\mu\text{l}$  of proteinase K buffer and 100  $\mu\text{g}$  of proteinase K per tube; incubate the samples in the thermoblock setup for gentle mixing for 2 h at  $55\text{ }^{\circ}\text{C}$ .

▲ **CRITICAL STEP** After that time, check the membrane for remaining radioactivity, which should be low in comparison with the solution.

? **TROUBLESHOOTING**

58| Discard the membrane and perform PCI extraction of RNA from solution. Add 50  $\mu\text{l}$  of 3 M sodium acetate (pH 5.5), and 500  $\mu\text{l}$  of PCI, and then vortex the samples vigorously for 30 s.

! **CAUTION** Phenol is toxic on inhalation, on contact with skin or if swallowed; it causes severe skin burns and eye damage. Use a hood, protective clothing, eye protection and gloves.

59| Centrifuge the samples at maximum speed (20,000g) in a tabletop centrifuge for 5 min at room temperature. Immediately collect the upper aqueous phase (~350  $\mu\text{l}$ ). Check the lower organic phase for remaining radioactivity, which should be low in comparison with the aqueous phase.

▲ **CRITICAL STEP** Be very careful not to aspirate the lower organic fraction.

? **TROUBLESHOOTING**

60| Add 1  $\mu\text{l}$  (15  $\mu\text{g}$ ) of GlycoBlue and 1 ml of ethanol to the aqueous phase from Step 59. Incubate the samples at  $-80\text{ }^{\circ}\text{C}$  for 30 min.

▲ **CRITICAL STEP** Extended incubation time at  $-80\text{ }^{\circ}\text{C}$  causes adverse co-precipitation of sodium acetate. Alternatively, precipitate RNA overnight at  $-20\text{ }^{\circ}\text{C}$ .

61| Centrifuge the samples at maximum speed (20,000g) for 30 min at  $4\text{ }^{\circ}\text{C}$ . Discard the supernatant.

? **TROUBLESHOOTING**

62| Wash the RNA-containing pellets twice with 0.75 ml of ice-cold 70% (vol/vol) ethanol (in water), and vortex the samples well. After each wash, centrifuge the samples for 10 min with the speed and temperature settings as in Step 61. Discard the supernatant.

## PROTOCOL

63| Air-dry the pellets (~5 min, alternatively at 37 °C, do not overdry the pellets).

### ? TROUBLESHOOTING

#### PNK phosphorylation and 5' adapter ligation ● TIMING 7 h to overnight, plus 30 min

64| Resuspend each pellet in 15 µl of phosphorylation mixture prepared in bulk for all the samples together; the volumes per sample are given below. Incubate the samples for 30 min at 37 °C.

Reagent	Volume (µl) per reaction	Final concentration
T4 RNA ligase reaction buffer, 10×	1.5	
ATP, 10 mM	1.5	1 mM
T4 PNK	1	0.67 U µl <sup>-1</sup>
Water	11	
Mix total volume	15	

65| Add 5 µl of 5' adapter ligation mixture, prepared for each sample as follows. Incubate the samples for 6 h (or overnight) at 16 °C.

Reagent	Volume (µl) per reaction	Final concentration (volume = 20 µl)
Water	2	
T4 RNA ligase reaction buffer, 10×	0.5	
ATP, 10 mM	0.5	1 mM
Barcoded 5' adapter	1	5 µM
RNA ligase 1	1	0.5 U µl <sup>-1</sup>
Mix total volume	5	

#### RNA purification and concentration ● TIMING variable: 2.5 h to overnight, plus 2 h

66| For easier handling, increase each sample volume to 400 µl with water. Proceed with PCI extraction and ethanol precipitation as in Steps 58–63.

■ PAUSE POINT Dried RNA pellets can be stored indefinitely at –20 °C.

#### Reverse transcription ● TIMING 2.5 h

67| Resuspend each pellet in 13 µl of RT mix I prepared in bulk for all the samples together; the volumes per sample are given below. Incubate the samples for 3 min at 80 °C, and then immediately place the samples on ice for 5 min.

Reagent	Volume (µl) per reaction	Final concentration (volume = 20 µl)
Water	8	
dNTPs, 2.5 mM	4	0.5 mM
miRCat-33 primer, 10 µM	1	0.5 µM
Mix total volume	13	

68| To each sample, add 6 µl of RT mix II prepared in bulk for all the samples together; the volumes per sample are given below. Incubate the samples for 3 min at 50 °C.



Reagent	Volume ( $\mu\text{l}$ ) per reaction	Final concentration (volume = 20 $\mu\text{l}$ )
First strand buffer, 5 $\times$	4	
DTT, 0.1 M	1	5 mM
RNasin	1	2 U $\mu\text{l}^{-1}$
Mix total volume	6	

**69** | Add 1  $\mu\text{l}$  of Superscript III reverse transcriptase, and incubate the samples for 60 min at 50  $^{\circ}\text{C}$ .

**70** | To inactivate reverse transcriptase, incubate the samples for 15 min at 65  $^{\circ}\text{C}$ .

**71** | To degrade template RNA, add 2  $\mu\text{l}$  of RNase H and incubate the samples for 30 min at 37  $^{\circ}\text{C}$ .

**■ PAUSE POINT** Prepared cDNA can be stored indefinitely at  $-20^{\circ}\text{C}$ .

**PCR amplification of the library ● TIMING 1 h**

**72** | To PCR-amplify the library, prepare 200  $\mu\text{l}$  of PCR mixture for each sample as follows, and divide it into four PCR tubes for amplification.

Reagent	Volume ( $\mu\text{l}$ ) per reaction	Final concentration
Water	140	
LA buffer, 10 $\times$	20	
dNTP, 2.5 mM	20	0.25 mM
Primer P5, 10 $\mu\text{M}$	4	0.2 $\mu\text{M}$
Primer PE_miRCat_PCR, 10 $\mu\text{M}$	4	0.2 $\mu\text{M}$
Template cDNA from Step 71	10	
TaKaRa LA Taq	2	0.05 U $\mu\text{l}^{-1}$
Mix total volume	200	

For amplification, use the following program:

Step	Temperature ( $^{\circ}\text{C}$ )	Time	Number of cycles
Initial denaturation	95	2 min	1
Denaturation	98	20 s	
Annealing	52	20 s	19–24
Extension	68	20 s	
Final extension	72	5 min	1

**▲ CRITICAL STEP** Limit the number of PCR cycles to reduce the proportion of PCR duplicates in the sequencing result.

**▲ CRITICAL STEP** To confirm the quality of the library, a small-scale PCR amplification can initially be performed with 1  $\mu\text{l}$  of template. However, in this case, before high-throughput sequencing, increase the sequencing depth by PCR amplification of the entire cDNA library. It is also informative to set up an additional PCR with miRNA-specific primers (see ANTICIPATED RESULTS section).

## PROTOCOL

■ **PAUSE POINT** Samples can be frozen and stored for up to a year at  $-20\text{ }^{\circ}\text{C}$ . However, freezing is not recommended if library quality will be checked by cloning and small-scale sequencing (in Step 79). In this case, follow the PROCEDURE continuously to Step 77 or 79.

### Size selection of the DNA fragments in the library ● **TIMING 4 h**

73| Prepare the 3% (wt/vol) MetaPhor agarose gel in TBE buffer with SYBR Safe according to the manufacturer's instructions. The gel can be prepared in advance (i.e., 1 d before) and stored at  $4\text{ }^{\circ}\text{C}$  wrapped in cling film.

74| Concentrate the DNA library from Step 72 by using the MinElute PCR purification kit according to the manufacturer's instructions. Elute the sample with  $20\text{ }\mu\text{l}$  of buffer EB provided with the kit. Add  $4\text{ }\mu\text{l}$  of  $6\times$  DNA loading dye to each sample.

75| Place the gel tank on ice, load the samples and the 50-bp GeneRuler DNA ladder and run the gel at constant voltage (80 V) for  $\sim 2\text{ h}$  until bromophenol blue reaches the edge of the gel.

▲ **CRITICAL STEP** Unless your samples carry bar-coded adapters, separate neighboring samples with empty lanes to reduce the risk of cross-contamination.

#### ? **TROUBLESHOOTING**

76| Scan the gel on the gel scanner and print the scan in its original size. Place the print underneath the transparency film and gel on its top. Carefully overlay the gel with the scan and cut-out gel slices with DNA fragments of desired length according to the instructions included in **Figure 3b** and in the ANTICIPATED RESULTS section.

▲ **CRITICAL STEP** This is an important step in the library preparation, as it allows enrichment of chimeric reads versus miRNA reads.

#### ? **TROUBLESHOOTING**

77| Isolate DNA fragments from the gel by using the MinElute Gel extraction kit according to the manufacturer's instructions. Elute the sample with  $20\text{ }\mu\text{l}$  of buffer EB from the kit.

▲ **CRITICAL STEP** To increase the cloning efficiency, dissolve the agarose slice in QG buffer (part of the kit) at room temperature, instead of  $50\text{ }^{\circ}\text{C}$  suggested by the kit manufacturer.

■ **PAUSE POINT** It is possible to freeze the sample at  $-20\text{ }^{\circ}\text{C}$  at this point and come back to the cloning procedure the next day. However, this can affect the cloning efficiency.

78| Measure the concentration of the cDNA library by using Qubit Fluorometer and Qubit dsDNA HS assay kit according to the manufacturer's instructions. The library is ready for high-throughput sequencing.

### Cloning of the cDNA library and small-scale sequencing (optional) ● **TIMING variable**

79| To check the quality of the cDNA library by small-scale sequencing, set up cloning reaction by using the TOPO-TA cloning kit according to the manufacturer's instructions. Use  $2\text{ }\mu\text{l}$  of the library from Step 78 per  $6\text{-}\mu\text{l}$  cloning mixture. Transform the Top10-competent bacteria that are supplied with the kit, and grow them on the LB agar plates with ampicillin.

▲ **CRITICAL STEP** Handle the Top10 bacteria supplied with the TOPO-TA cloning kit with utmost care, as they are extremely sensitive.

#### ? **TROUBLESHOOTING**

80| Send the colonies or purified plasmids for Sanger sequencing.

### Bioinformatic analysis ● **TIMING $\sim 5\text{ h}$ for basic analysis**

81| For the description of the analysis, see the Experimental design ('Bioinformatic analysis') section and ref. 12. A recently developed software package for CLASH data analysis is available and described in detail in ref. 16.

#### ? **TROUBLESHOOTING**

#### ? **TROUBLESHOOTING**

Troubleshooting advice can be found in **Table 1**.

**TABLE 1** | Troubleshooting table.

Step	Problem	Possible reason	Solution
52	Slight change in the color of the NuPAGE LDS sample buffer	Low pH of the sample caused by the remaining TCA	Add 1 $\mu$ l of 1 M Tris buffer (pH 7.8)
53	Uneven migration of the samples in the gel	Low pH of the sample caused by the remaining TCA	Carefully observe the color of the samples from Step 52. If the problem persists, add a further wash with ice-cold acetone
55	Additional bands on the autoradiography film; radioactive signal in the negative control at the level of tagged AGO	Cross-contamination of the samples	Repeat the experiment, taking care not to mix the samples
	Additional bands on the autoradiography film at different levels than tagged AGO	Background proteins bound to the beads	Extend the wash in Step 26
	Radioactive smear in the whole lanes of the gel, visible at autoradiography film	Unspecific RNA is bound to AGO-RNA complexes	Extend the wash in Step 26; add more RNase-IT
	Low signal on the autoradiography film	Low expression of tagged-AGO	Optimize conditions of tagged-AGO expression; expose the gel longer
57	High radioactive signal on the nitrocellulose membrane	Low binding efficiency to IgG-Dynabeads	Check binding efficiency by western blotting (ANTICIPATED RESULTS); coat the Dynabeads once more and wash them well directly before the experiment
		Inefficient cross-linking	Optimize cross-linking conditions
		RNA degradation	Ensure that you work in RNase-free conditions
59	Substantial radioactivity in the organic phase	Inefficient proteinase K treatment; remains of protein bind to RNA to prevent it from entering the aqueous phase	Take a fresh proteinase K aliquot; add more proteinase K and/or incubate samples longer. See also Troubleshooting for Step 57
61	Pellet is loose or floating in supernatant	Traces of organic phase are aspirated during PCI extraction	Take more care in collecting the upper aqueous phase
63	RNA pellets are difficult to resuspend	Pellets are overdried	Take more care in drying pellets
75	Samples diffuse from wells during loading	Samples contain leftover ethanol from Step 74	Incubate the remaining samples for a few min at 50 °C, and add more loading dye
76	No PCR products are observed despite strong signal on the autoradiography film	Problems with enzymatic reactions	Replace enzymes and check buffers
		Problems with ligation of sequencing adapters	After resuspending in water, divide the adapters into small aliquots; use fresh aliquot for each experiment

(continued)

## PROTOCOL

**TABLE 1** | Troubleshooting table (continued).

Step	Problem	Possible reason	Solution
	Unique PCR product of approximate size of 120 bp	Only PCR-amplified linker dimers are visible on the gel; problems with enzymatic reactions	See Troubleshooting for Step 76
	PCR products <140 bp migrate in one relatively sharp band	Excessive RNase digestion in Step 17	Optimize RNase digestion conditions. See Experimental design ('RNase digestion') section
79	No bacterial colonies are seen on the plates	Unsuccessful TOPO-TA cloning	Check if your polymerase leaves 3' A-overhangs in the PCR products; if not, add A-overhangs after amplification according to the protocol supplied with the TOPO-TA cloning kit
		Top10 bacteria do not survive transformation	During transformation, be very gentle; do not pipette bacteria and do not shake them after the heat-shock before spreading them on the plate
81	Sequencing returns multiple similar sequences that cannot be mapped to the database	RNA contamination coming from commercial enzymes (frequently bacterial rRNA)	Change enzyme supplier
		Problem with 3'-end adapter ligation	miRCat-33 primer used for reverse transcription is complementary to bacterial rRNA. Therefore, a problem with 3'-end adapter ligation results in enrichment of cDNA/PCR product in bacterial rRNA
	No reads map to miRNA in the sequencing result	Major experimental problems	Repeat the experiment

### ● TIMING

Steps 1 and 2, cell culture: 48 h

Steps 3 and 4, UV cross-linking: ~10 min per CLASH sample

Steps 5–7, cell lysis: 1 h

Steps 8–15, IgG-Dynabeads loading and washing: 90 min

Steps 16–18, RNase digestion on Dynabeads: 15 min

Steps 19–22, elution from IgG-Dynabeads: 40 min

Steps 23–29, Ni-NTA agarose loading and washing: 3 h plus 6 h for WB (optional)

Steps 30–33, RNA phosphorylation with T4 PNK: 3 h

Steps 34 and 35, intermolecular RNA-RNA ligation: overnight

Steps 36 and 37, dephosphorylation with TSAP: 1 h

Steps 38 and 39, 3'-end adapter ligation on beads: 6.5 h

Steps 40–42, radioactive labeling of RNA: 1 h 15 min

Steps 43–46, elution of PTH-Ago1-RNA complexes from Ni-NTA beads: 30 min

Steps 47–52, TCA precipitation: 2 h

Steps 53–55, SDS-PAGE and transfer of protein-RNA complexes: 4.5 h

Steps 56–63, RNA purification: variable, 5 h to overnight, plus 4 h

Steps 64 and 65, PNK phosphorylation and 5' adapter ligation: variable; 7 h to overnight, plus 30 min

Step 66, RNA purification and concentration: variable, 2.5 h to overnight, plus 2 h

Steps 67–71, reverse transcription: 2.5 h

Step 72, PCR amplification of the library: 1 h

Steps 73–78, size selection of the DNA fragments in the library: 4 h

Steps 79 and 80 (optional), cloning of the cDNA library and small-scale sequencing: variable

Step 81, bioinformatic analysis: ~5 h for basic analysis



## ANTICIPATED RESULTS

The CLASH procedure is long and contains relatively few steps at which progress can be controlled. It is therefore useful to collect the samples from all the purification fractions for clues that may help future troubleshooting.

### Step 7

Before starting the CLASH procedure, it is worth determining the expression level of the tagged AGO protein by WB. The aim should be to express the tagged protein at close to endogenous levels. Overexpression risks generating spurious binding sites. Note that the protein A tag should be cleaved in order to compare the amount of tagged and endogenous proteins, as it will bind to antibodies used for WB.

### Step 25

By comparing the amount of tagged AGO in the cell lysate with AGO remaining in the flow-through from IgG-Dynabeads loading, it is possible to estimate the Dynabeads' IgG-coating efficiency and the percentage of protein bound by the IgG-Dynabeads relative to the cell lysate input. Independently of the amount of the IgG beads used per sample, we routinely observe ~30% protein loss at this stage (**Fig. 3c**). At the stage of Ni-NTA loading, we also observe up to 25% loss in the amount of AGO (**Fig. 3c**).

### Step 50

After TCA precipitation, higher radioactivity of the samples than the negative control prepared from the cell line without tagged AGO protein is expected.

### Step 56

The strength of the signal obtained by autoradiography can vary markedly from experiment to experiment. However, with successful procedure, a clear signal should be visible within 1 h. There should be a single main band in each sample lane corresponding to cross-linked AGO-RNA complexes at the level of the orange band in the prestained SeeBlue Plus 2 standard. An additional band is frequently visible above this, corresponding to AGO-RNA complexes. The negative control prepared from the cell line without tagged AGO protein should not give clear radioactive bands, but some signal in the whole lane may be visible after an overnight exposure because of background RNA.

### Step 76

A typical agarose gel picture is presented in **Figure 3b**. A fast migrating band of ~120 bp is frequently visible, both in the samples and negative controls. This corresponds to the ligated adapter dimers, and its intensity varies markedly between experiments. It is expected that there will be a broad band, relatively sharper at the bottom and smeared at the top, 20 bp above that. The lower part of the band mostly corresponds to cloned miRNAs, whereas the upper, smeared part is enriched for targets and chimeric sequences. This upper band should appear only in the cross-linked samples. Separately isolating the DNA fragments of different sizes allows enrichment for the target and chimeric sequences in the sequence data. We generally use threefold more of the upper band (UB: targets + chimeras) than the lower band (LB: miRNAs). In case of doubts, preparing additional control PCR reactions with primers that amplify a specific miRNA can confirm that the cross-linking and linker ligations were successful. For studies on HEK 293 cells, we used primers specific for the highly expressed miR-16; a typical gel picture is shown in **Figure 3d**.

### Step 78

Because the amount of cDNA in the library depends on multiple factors, it is difficult to give any expected concentration. To have enough material for the small-scale and high-throughput sequencing without creating unnecessary PCR duplicates, adjust the number of PCR cycles to obtain ~20–80 ng of DNA.

### Step 80

A small-scale sequencing analysis should provide the following information: (i) types of RNAs present in the sequencing data: expect many miRNAs for AGO-CLASH; (ii) the average length of cDNA inserts, which reflects the effectiveness of RNase treatment; and (iii) the presence of bacterial rRNA contamination.

**ACKNOWLEDGMENTS** We thank R. Holmes, G. Kudla and A. Tuck for their helpful comments on the manuscript. This work was supported by Wellcome Trust funding (no. 077248). Work in the Wellcome Trust Centre for Cell Biology is supported by Wellcome Trust core funding (no. 092076).

**AUTHOR CONTRIBUTIONS** A.H. designed the experiments, conducted the experimental work and wrote the manuscript; D.T. supervised the study and wrote the manuscript.

**COMPETING FINANCIAL INTERESTS** The authors declare no competing financial interests.

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

1. Fabian, M.R., Sonenberg, N. & Filipowicz, W. Regulation of mRNA translation and stability by microRNAs. *Ann. Rev. Biochem.* **79**, 351–379 (2010).

2. Hu, W., Alvarez-Dominguez, J.R. & Lodish, H.F. Regulation of mammalian cell differentiation by long non-coding RNAs. *EMBO Rep.* **13**, 971–983 (2012).
3. Guil, S. & Esteller, M. *Cis*-acting noncoding RNAs: friends and foes. *Nat. Struct. Mol. Biol.* **19**, 1068–1075 (2012).
4. Djebali, S. *et al.* Landscape of transcription in human cells. *Nature* **489**, 101–108 (2012).
5. Ule, J. *et al.* CLIP identifies Nova-regulated RNA networks in the brain. *Science* **302**, 1212–1215 (2003).
6. Licatalosi, D.D. *et al.* HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature* **456**, 464–469 (2008).
7. Hafner, M. *et al.* Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* **141**, 129–141 (2010).
8. König, J. *et al.* iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nat. Struct. Mol. Biol.* **17**, 909–915 (2010).
9. Chi, S.W., Zang, J.B., Mele, A. & Darnell, R.B. Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature* **460**, 479–486 (2009).
10. Granneman, S., Kudla, G., Petfalski, E. & Tollervey, D. Identification of protein binding sites on U3 snoRNA and pre-rRNA by UV cross-linking and high-throughput analysis of cDNAs. *Proc. Natl. Acad. Sci. USA* **106**, 9613–9618 (2009).
11. Kudla, G., Granneman, S., Hahn, D., Beggs, J.D. & Tollervey, D. Cross-linking, ligation, and sequencing of hybrids reveals RNA-RNA interactions in yeast. *Proc. Natl. Acad. Sci. USA* **108**, 10010–10015 (2011).
12. Helwak, A., Kudla, G., Dudnakova, T. & Tollervey, D. Mapping the human miRNA interactome by CLASH reveals frequent noncanonical binding. *Cell* **153**, 654–665 (2013).
13. Pasquinelli, A.E. MicroRNAs and their targets: recognition, regulation and an emerging reciprocal relationship. *Nat. Rev. Genet.* **13**, 271–282 (2012).
14. Greenberg, J.R. Ultraviolet light-induced crosslinking of mRNA to proteins. *Nucleic Acids Res.* **6**, 715–732 (1979).
15. Kishore, S. *et al.* A quantitative analysis of CLIP methods for identifying binding sites of RNA-binding proteins. *Nat. Methods* **8**, 559–564 (2011).
16. Travis, A.J., Moody, J., Helwak, A., Tollervey, D. & Kudla, G. Hyb: a bioinformatics pipeline for the analysis of CLASH (crosslinking, ligation and sequencing of hybrids) data. *Methods* doi:10.1016/j.ymeth.2013.10.015 (2013).
17. Oeffinger, M. *et al.* Comprehensive analysis of diverse ribonucleoprotein complexes. *Nat. Methods* **4**, 951–956 (2007).

