Supporting Information

HyPR-MS for multiplexed discovery of MALAT1, NEAT1, and NORAD lncRNA protein interactomes

Michele Spiniello[1,6], Rachel A. Knoener[1,6], Maisie I. Steinbrink[1,2], Bing Yang[3], Anthony J. Cesnik[1], Katherine E. Buxton[1], Mark Scalf[1], David F. Jarrard[2,3,4], Lloyd M. Smith[1,5,*]

[1] Department of Chemistry; University of Wisconsin – Madison; Madison, Wisconsin 53706; United States
[2] Molecular and Environmental Toxicology; University of Wisconsin – Madison; Madison, Wisconsin 53706; United States
[3] Department of Urology; University of Wisconsin School of Medicine and Public Health; Madison, Wisconsin 53705; United States
[4] Carbone Comprehensive Cancer Center; University of Wisconsin – Madison; Madison, Wisconsin 53792; United States
[5] Genome Center of Wisconsin; University of Wisconsin – Madison; Madison, Wisconsin 53706; United States
[6]These authors contributed equally
*Correspondence: smith@chem.wisc.edu; Tel. 608-263-2594

Supporting Information:

HyPR-MS Protocol

Probe Design

1. Using available databases for lncRNAs (such as https://lncipedia.org/db/search), find the sequence of each target
2. Load the sequence into the Mfold database (http://unafold.rna.albany.edu/?q=mfold/RNA-Folding-Form)
    a. Set the folding temperature to 37□ (or whatever your desired hybridization temperature will be)
    b. Leave all default parameters as they are
3. From the resulting single-strand frequency data, select regions with high values along several regions of the target RNA

    Note: When studying specific isoforms, select only from regions that are specific to that isoform

    Note: In general, we have found that 2-3 regions for capture is enough for an RNA less than 10 kb in length
4. Find the complementary and reverse sequence (http://www.bioinformatics.org/sms/rev_comp.htm) to the chosen regions of the RNA target; this will be the sequence of the capture oligonucleotide (CO)
5. Use the OligoAnalyzer 3.1 from IDT (https://eu.idtdna.com/calc/analyzer) or other publicly available software to define necessary parameters for your CO conditions
    a. Salt concentrations should be set to 375 mM (or the optimal concentration for hybridization)
    b. Find the range of CO concentrations that are possible to obtain the ideal melting temperature (Tm) for your experiment

        Note: The 'ideal' CO Tm is 20-28.5□ higher than the hybridization temperature (in this case, 56.8-65.5□). The optimal concentration of CO to use within this range will be tested in subsequent small-scale experiments.

        Note: It is generally possible to meet the Tm by choosing a CO length of 20-30 nucleotides (shorter than this will not bind specifically to you target), finding single-stranded regions of the target RNA that have low GC content, using a lower probe concentration, or modifying the salt concentration.
6. After the desired CO sequence is found and the Tm range is met, add an 8 nt toehold sequence to the 5' end of the CO.

    Note: This toehold should not affect the Tm for COs because it is not complementary to the target RNA, but it is important to include the toehold sequence when checking the formation of hairpins and self-dimerization (this can be done following the OligoAnalyzer 3.1 data analysis tips (https://eu.idtdna.com/pages/support/technical-vault/faq/application/oligo-characteristics/how-do-i-use-the-oligoanalyzer-tool-to-analyze-possible-hairpins-and-dimers-formed-by-my-oligo-)

Note: It may be necessary to slightly shift COs to avoid formation of hairpins below the melting temperature, even if the single-stranded frequency from Mfold is not as high after this shift.

a. Check the specificity of the CO to your target using NCBI BLAST

7. Create release oligonucleotides (ROs) that are reverse and complementary to the sequence of the COs for each target
8. Once all COs and ROs have been designed, use the OligoAnalyzer 3.1 software to check for heterodimer formation. Modify COs and ROs as necessary to reduce heterodimerization.

Small and Large Scale HyPR-MS Protocol

Materials

Reagents (all materials used prior to qPCR should be RNAse free)

Ice
Cells, 1% formaldehyde-treated for 10 minutes (small-scale: 5 x 10^5 per condition; large-scale: 1 x 10^8 per bioreplicate)
Ribonucleoside Vanadyl Complex (RVC) 200 mM
Dithiothreitol, solid
RNasin Plus Ribonuclease Inhibitor (Promega, Cat. No. N261B)
Halt™ Protease Inhibitor Cocktail (100X) (Thermo Fisher Scientific, Cat. No. 78430)
Lithium Chloride (LiCl) 8M
Tris pH 7.5 4M
Lithium Dodecyl Sulfate, solid (LiDS)
Triton X-100 (10% w/v)
Nuclease Free Water, not DEPC treated
Capture oligonucleotides, biotinylated (at least one for each target and scrambled control, designed through IDT)
Sera-Mag magnetic streptavidin microparticles (GE Healthcare, Cat. No. 30152105010150)
Release oligonucleotides (complementary to each capture oligonucleotide, designed through IDT)
Tris pH 8.0, 10 mM
EDTA pH 8.0, 0.1 mM
Proteinase K
TRI Reagent
Chloroform (100%)
GlycoBlue™ Coprecipitant (Invitrogen, Cat. No. AM9515)
Ethanol (100% and 75%)
Reverse Transcription Kit (Thermo Fisher Scientific, Cat. No. 4368814)
LightCycler 480 Probes Master qPCR Kit (Roche, Cat. No. 04707494001)
qPCR assays for several regions along each target (designed through IDT)
Human Genomic DNA, mixed (Promega, Cat. No. G304A)
Urea, solid
Deoxycholic Acid, solid (DCA)
HPLC water
Iodoacetamide, solid

Ammonium bicarbonate (200 mM and 50 mM)
Sequencing Grade Modified Trypsin, lyophilized (Promega, Cat. No. V5111)
Ethyl Acetate
Trifluoroacetic Acid (TFA)

Equipment
Benchtop rocker
Benchtop vortex mixer
Mini-centrifuge
Model 550 Sonic Dismembrator (Fisher Scientific)
Benchtop centrifuge
-80□ freezer
-20□ freezer
Nanodrop UV-Vis Spectrophotometer
Incubator (37□)
Thermocycler
Roche LightCycler® 480 System
Magnetic separation racks for 2 mL tubes
Magnetic separation racks for 50 mL falcon tubes
Pipets
RNAse-free pipet tips
2 mL tubes, low-bind RNAse-free
50 mL conical falcon tubes, low-bind
50ml conical falcon tubes, low-bind RNAse-free
200 uL tubes RNAse-free
RNase*Zap*™ RNase Decontamination Wipes (Thermo Fisher Scientific, Cat. No. AM9786)
Hot water bath (37□)
Parafilm
Bucket centrifuge
Centrifugal filter units (50 kb filters)
Savant Speed Vac Concentrator

Method

Each condition in the small-scale experiment requires 5 x 10^5 cells crosslinked with 1% formaldehyde for 10 minutes. An additional 40 uL aliquot is required to use as an input control. Several conditions can be tested at the same time during each experiment. Proteinase K treatment, RNA extraction, and RT-qPCR will be performed to measure capture efficiency (CE), selectivity (CS), release efficiency (RE) and fold enrichment (FE) (see main text). Based on these parameters, the best conditions and necessary number of cells to use in large-scale experiments will be determined.

Each large-scale experiment will be optimized based on results from small-scale experiments and scaled accordingly to meet the requirements for mass spectrometric analysis. The bulk of

each large-scale experimental sample will be subjected to protein purification, trypsin digestion and protein analysis, but small aliquots can be taken throughout and treated as the small-scale samples for quality assessment.

Lysis

1. Start the lysis procedure, using the following calculations to determine reagent volumes to add
    a. Desired number of cells per uL lysate = 5000 cells
    b. Pellet volume = 20 uL
    c. Final volume = (total number of cells/5000) uL
    d. RVC (200 mM) = [(Final volume*12.5)/200] uL
    e. DTT (1 M) = [(Final volume*12.5)/1000] uL
    f. RNAsin (40 U/mL) = [(Final volume/1000)*(125)/40)] uL
    g. Protease Inhibitor (100X) = [(Final volume*1.25)/100] uL
    h. LiCl (8 M) = [(Final volume*468.8)/8000]
    i. Tris (4 M) = [(Final volume*62.5)/4000] uL
    j. LiDS (10%) = [(Final volume*1.25)/10] uL
    k. Triton X-100 (10%) = [(Final volume*1.25)/10] uL
    l. Nuclease Free Water = remaining uL to achieve the final volume
2. Combine the above reagents with the cell pellet (placed in ice) and leave on ice for 10 minutes, vortexing periodically to break the pellet
3. Transfer the solution to a 2 mL low-bind tube (small-scale) or 50 mL falcon (large-scale).
4. Sonicate the lysate on ice to achieve an average chromatin fragmentation size of approximately 6kb (see Supplemental Information for gel image)
    Note: This was achieved using a Fisher Scientific Model 550 Sonic Dismembrator pulsing 4 seconds on, 4 seconds off for 12 seconds at intensity 3 with a microtip (small-scale) or 24 seconds at intensity 4 with a full probe (large-scale). Times and intensities may vary between sonicators.
5. Centrifuge at 4C, 1800 RPM for 2 minutes
6. Transfer the supernatant to a new low-bind tube and discard the pellet
7. For each tested small-scale condition, add 100 uL of sonicated lysate (5×10^5) to a separate low-bind RNAse free tube. For large-scale, transfer the lysate to a new low-bind 50 mL falcon. Discard pellets and place all tubes on ice.
    a. At this point, reserve a 40 uL sample of lysate as an input control for capture quality assessment or mass spectrometry analysis.
    b. Store any unused sonicated lysate at -80C for future experiments.

Hybridization

1. Add water to achieve 1.25x the lysate volume
2. Add COs to fall within the range of Tm values established through IDT (small-scale), or an already-established optimal CO concentration (large-scale).
    Note: If COs must be diluted to add the required amount, dilute in water
3. Add the calculated amounts of CO to each condition and hybridize by rotating for 3 hours in an incubator at 37C

Note: Treat the 40 uL input sample with the same temperature conditions as the other tubes. In this way, any degradation brought about by heat will be consistent across all samples and controls.

Buffer and Bead Preparation

1. During the 3 hour hybridization period, prepare the wash and release buffers in RNAse-free tubes

Wash Buffer	Final Conc.	Volume (uL)		**Release Buffer**	Final Conc.	Volume (uL)
LiCl	375 mM	46.9		LiCl	375 mM	46.9
Tris	50 mM	12.5		Tris	50 mM	12.5
LiDS	0.2%	20		LiDS	0.1%	10
Triton X-100	0.2%	20		Triton X-100	0.1%	10
H_2O		900.6		H_2O		920.6

2. 30 minutes before the end of hybridization, add streptavidin-modified magnetic beads to a 2 mL low-bind tube (small-scale) or a 50ml low-bind falcon (large-scale)
    a. Use 2.2 uL of beads per picomole CO added for hybridization
    Note: It is important not to perform this step too far before hybridization is complete, or else the beads may spend too much time at room temperature and become less effective at capture
3. Put the tube of beads on the magnetic tube rack and wait 3-4 minutes to allow beads to leave the solution
4. Throw out the supernatant
5. Add wash buffer to the tube and resuspend the beads
    a. Volume of wash buffer should be 5 times the volume of beads
6. Place the tube of beads on the shaker for 5 minutes
7. Put the tube with the beads on the magnetic tube rack for 3-4 minutes, then throw out the supernatant
8. Repeat steps 5-7 for a second wash
9. Resuspend the beads in their original volume with wash buffer

Bead Coupling

1. At the end of the hybridization period, add the required amount of beads to each sample
    Note: No beads are required for the 40 uL input sample, but it should still be heated and cooled in the same way as the other samples.
2. Return all samples and the input to the incubator at 37C for 1 hour

Elution

1. During the 1 hour incubation, prepare the ROs. For each target, use 100 times more RO than was added of its respective CO.
2. After the 1 hour incubation, put each condition on the magnetic tube rack for 3-4 minutes, and discard the supernatant
3. Resuspend the beads in wash buffer
    a. Volume of wash buffer should be 5 times the volume of beads for each condition
4. Place samples on a shaker for 15 minutes at 37C
5. Put tubes with beads on the magnetic tube rack for 3-4 minutes, then throw out the supernatant
6. Resuspend the beads in release buffer
    a. Volume of release buffer should be 3 times the volume of beads for each condition
7. Leave samples at room temperature for 5 minutes without shaking
8. Put tubes with beads on the magnetic tube rack for 3-4 minutes, then throw out the supernatant.
9. Resuspend the beads in release buffer
    a. Volume of release buffer should be 3 times the volume of beads for each condition
10. Add the required amount of RO for the first target

    Note: Because wash steps are performed between each target release, it may be desirable to release the target with the fewest number of copies first to avoid excess target loss.
11. Place the samples on a shaker for 30 minutes at room temperature to facilitate release

    Note: If the volume is too small for a benchtop shaker to mix thoroughly, use a foam tube holder attached to a vortex and shake on low
12. Put the tubes with beads on the magnetic tube rack for 3-4 minutes, then transfer the supernatant to a new tube and put in ice. This is the first captured sample.

    Note: In small-scale experiments this entire sample will be subjected to Proteinase K treatment, RNA extraction and RT-qPCR for capture quality assessment. In large-scale a 40 uL aliquot will be taken from each captured sample to treat as the small-scale samples while the remaining captured sample will be stored overnight prior to eFASP protein purification.
13. Resuspend the beads in release buffer
    a. Volume of release buffer should be 3 times the volume of beads for each condition
14. Put the tubes with beads on the magnetic tube rack for 3-4 minutes, then throw out the supernatant
15. Resuspend the beads in release buffer
    a. Volume of release buffer should be 3 times the volume of beads for each condition
16. Repeat steps 10-15 to release each target and scrambled control sequentially

    Note: To test the elution efficiency, resuspend the beads in 3 times their original volume of release buffer after the final release step. Use no more than a 50 uL aliquot and check for the presence of each target RNA on the beads via

RT-qPCR (release efficiency).

Proteinase K Treatment and RNA Purification

The entirety of each input, captured, and various control samples taken from each small-scale, or their corresponding aliquots from large-scale, will be treated with Proteinase K and extracted. For the bulk of each large-scale captured sample, skip to the eFASP procedure.

1. Add Proteinase K and $CaCl_2$ to each aliquot to achieve a final concentration of 200 mM for each based on a 200 uL final volume.
2. To make reagent concentrations equal based on the respective buffers each sample was resuspended in, add the necessary reagents (LiCl, Tris, LiDS, etc.) to the tubes based on a 200 uL final volume.
3. Bring the resulting solutions to 200 uL with nuclease-free water, seal with parafilm, and digest overnight at 37□ in a water bath
4. Begin RNA extraction by adding 500 uL Trizol to each sample
5. Vortex to resuspend the RNA and leave at room temperature for 5 minutes, vortexing periodically
6. Add 100 uL chloroform to each sample and shake vigorously to combine
7. Leave at room temperature for 5 minutes
8. Centrifuge at 4□, 12000 RCF for 10 minutes
9. While the samples are centrifuging, prepare new tubes for each sample containing 1.5 uL glycoblue in each
10. After centrifugation, take the upper layer of solution for each sample and add it to the tube containing glycoblue, making sure to note the volume
11. Add ethanol to these samples to achieve a final ethanol concentration of 75% in each
12. Store all samples at -20□ overnight (10+ hours)
13. Centrifuge samples at 4□, max speed for 15 minutes
14. Discard the supernatant and resuspend the pellet in 750 uL of 75% ethanol
15. Centrifuge samples at room temperature, max speed for 8 minutes
16. Discard the supernatant, removing as much liquid as possible without disturbing the pellet, and let tubes dry for 5 minutes
17. Resuspend pellets in water (15-25 uL) and measure RNA concentrations with a NanoDrop
18. Use approximately 1 ug RNA from each sample (up to 10 uL) for RT-qPCR

    Note: To estimate the amount of DNA captured during the experiment, treat 1 ug of the input with reverse transcriptase (+) and another 1 ug without reverse Transcriptase (-). Any signal measured via qPCR from the second sample is genomic DNA and can be subtracted from the first for a more accurate measurement of capture efficiency.

    Note: If qPCR assays have not been used before, follow the IDT PrimeTime qPCR Assay Resuspension Protocol and bring the stock concentration to 40x using IDTE buffer (10 mM Tris pH 8.0, 0.1 mM EDTA pH 8.0)

eFASP Procedure

1. Add solid urea to each of the captured, control, and small input samples to bring to a final concentration of 8M
   Note: The solid urea will increase the final volume, sometimes by several mL. Take this into account when measuring final urea concentration.
2. Add enough DCA to achieve a final concentration of 2%
3. Vortex to mix
4. Take 50 kb filters and tubes for all samples and label them; 1 for the lysate sample and 4 for each captured/control sample
5. Put 450 uL of each solution into their respective filter
6. Centrifuge at room temperature, 14000 RCF for 10 minutes
7. Discard the eluent
8. Repeat steps 5-7 until all the liquid from each sample has been filtered
   Note: If at any step during the eFASP procedure the liquid does not completely flow through the filter into the collection tube after centrifugation, repeat the centrifugation step
9. During this time, create buffers for the rest of the eFASP procedure

Exchange Buffer	Final Buffer Concentration	Volume
Urea (solid)	8 M	24 g
DCA (solid)	0.10%	50 mg
Add water to final volume:		50 mL

Reducing Buffer	Final Buffer Concentration	Volume
Urea (solid)	8 M	4.8 g
DTT	20 mM	200 uL
Add water to final volume:		10 mL

Alkylation Buffer	Final Buffer Concentration	Volume
Urea (solid)	8 M	24 g
Iodoacetamide (solid)	50 mM	0.462 g
Ammonium Bicarbonate	50 mM	12.5 mL

(200 mM)		
Add water to final volume:		50 mL

Digestion Buffer	Final Buffer Concentration	Volume
Urea (solid)	1 M	3 g
Ammonium Bicarbonate (200 mM)	50 mM	12.5 mL
DCA (solid)	0.10%	50 mg
Add water to final volume		50 mL

10. After all the samples have been filtered, add 400 uL exchange buffer to each filter
11. Centrifuge at room temperature, 14000 RCF for 10 minutes
12. Discard the eluent
13. Repeat steps 10-13 two more times
14. Add 200 uL reducing buffer to each filter and incubate on the benchtop at room temperature for 30 minutes
15. Centrifuge samples at room temperature, 14000 RCF for 10 minutes
16. Discard the eluent
17. Add 200 uL alkylation buffer to each filter and incubate in the dark at room temperature for 30 minutes
18. Centrifuge samples at room temperature, 14000 RCF for 10 minutes
19. Discard the eluent
20. Add 400 uL digestion buffer to each filter
21. Centrifuge samples at room temperature, 14000 RCF for 10 minutes
22. Discard the eluent
23. Repeat steps 20-22 two more times
24. Move the filters into a clean collection tube
25. Add 100 uL digestion buffer and 1 ug trypsin to each filter
26. Close and wrap each tube with parafilm, incubate overnight at 37C without rotation
27. After incubation, remove parafilm and centrifuge samples at room temperature, 14000 RCF for 10 minutes

    Note: This has eluted the proteins - do not discard the eluent!
28. Add 50 uL of 50 mM ammonium bicarbonate
29. Centrifuge samples at room temperature, 14000 RCF for 10 minutes
30. Transfer eluent to new snap-top tubes, noting the volume of each
31. Add 200 uL ethyl acetate to each tube
32. Add TFA to bring each sample to 0.5%
33. Shake manually for 1 minute, then centrifuge at room temperature, 15700 RCF for 2 minutes

34. Discard the top layer
35. Add 200 uL ethyl acetate to each tube and shake manually for 1 minute
36. Centrifuge at room temperature, 15700 RCF for 2 minutes
37. Discard the top layer
38. Repeat steps 35-37 once more
39. Place all tubes (open) in a Speed Vac and spin until samples are completely dry and store at -80C until ZipTip can be performed

Table S1: HyPR-MS Advances the Capabilities of *In Vivo* RNA Capture Technologies.

(In reference to the Results section and Table 1; A=Advantage, L=Limitation)

	CHART (West et al., 2014)	**ChIRP** (Chu et al., 2015)	**RAP-MS** (McHugh et al., 2015)	**HyPR-MS**
Capture Oligonucleotide Design	**RNAse H Assay** A: Able to design the capture oligonucleotide to hybridize to the optimal location of the target RNA using empirical evidence of where the single stranded regions are. L: Expensive and laborious. Using only one oligonucleotide for capture can result in a bias towards capturing fragments of the RNA that are close to the site of hybridization.	**Full RNA Tiling** A: Avoid location bias caused from using only one capture oligonucleotide by using many oligonucleotides that hybridize along the full length of the target RNA. L: Expensive. Unable to isolate different isoforms of the RNA because the full sequence is hybridized to the oligonulceotides.	**Full RNA Tiling** A: Avoid location bias caused from using only one capture oligonucleotide by using many oligonucleotides that hybridize along the full length of the target RNA. L: Expensive. Unable to isolate different isoforms of the RNA because the full sequence is hybridized to the oligonulceotides.	**Secondary-structure prediction software** A: Mfold software (Zuker et al., 2003) uses minimum free energy estimations to predict nucleotide regions with the highest probability to be single stranded. 2-3 oligonucleotides complementary to these regions produces sufficient capture efficiency along the length of the target RNA. Faster and less expensive than RNAse H Assay and Full RNA Tiling. Allows for isolation of isoforms. Reduced location bias than using only one oligonucleotide for capture. L: Secondary structure is only a prediction, not empirically determined.
Crosslinking Conditions	**3% formaldehyde, 30 min.** A: Identification of direct and indirect protein-RNA interactions critical to RNA processing (Schmitz et al., 2016). Applied with success in many applications such as PICh and HyCCAPP. L: Formaldehyde may cause single- and double-stranded breaks in nucleic acids (Grafstrom et al., 1984, Kondo et al., 1985) resulting in the target RNA potentially being fragmented. This is problematic if only one oligonucleotide is used for capture.	**3% formaldehyde, 30 min.** A: Identification of direct and indirect protein-RNA interactions critical to RNA processing (Schmitz et al., 2016). Applied with success in many applications such as PICh and HyCCAPP. L: Formaldehyde may cause single- and double-stranded breaks in nucleic acids (Grafstrom et al., 1984, Kondo et al., 1985) resulting in the target RNA potentially being fragmented. This is not problematic for Full RNA Tiling capture strategies.	**UV-crosslinking** A: Crosslinks only direct RNA-protein interactions which could be desirable in some applications. Milder solubilization and less target fragmentation risk. L: For discovery and understanding of RNA interactomes it is beneficial to find indirect interactors as their impact on processing is significant. (Schmitz et al., 2016). Low efficiency of crosslinking which could correspond to excessive false negatives.	**1% formaldehyde, 10 min.** A: Same advantages as those for CHART and ChIRP however the limited exposure to formaldehyde here reduces the risk of RNA fragmentation and also reduces the amount of sonication necessary for nucleic acid solubilization downstream which also reduces RNA fragmentation. This is critical for the use of only a few capture oligonucleotides to capture the full RNA sequence without using the Full RNA Tiling capture strategy.

	CHART (West et al., 2014)	**ChIRP** (Chu et al., 2015)	**RAP-MS** (McHugh et al., 2015)	**HyPR-MS**
Lysate Solubilization and DNA Fragment Size	**Sonication; 2-10 kilobases** A: Sonication fragments and solubilizes the DNA so that lncRNAs that are associated with the chromatin can be isolated from the full genome. L: Excessive sonication can cause target RNA fragmentation (see above).	**Sonication; 200-500 bases** A: Sonication fragments and solubilizes the DNA so that lncRNAs that are associated with the chromatin can be isolated from the full genome. L: Excessive sonication can cause target RNA fragmentation (see above).	**Sonication; unspecified** A: Sonication fragments and solubilizes the DNA so that lncRNAs that are associated with the chromatin can be isolated from the full genome. L: Excessive sonication can cause target RNA fragmentation (see above).	**Sonication; about 6 kilobases** A: Sonication fragments and solubilizes the DNA so that lncRNAs that are associated with the chromatin can be isolated from the full genome. Milder sonication is sufficient for HyPR-MS, because of reduced crosslinking, thus preserving the target RNA integrity. L: Excessive sonication can cause target RNA fragmentation (see above).
Elution Strategy	**RNAse H** L: only one RNA target per 10^8 cells.	**Heat and Biotin Elution** L: only one RNA target per 10^8 cells.	**RNAse and DNAse Digestion** L: only one RNA target per 10^8 cells.	**Toehold-Mediated Release (Kennedy-Darling et al., 2014)** Allows for isolation of several RNA targets from the same 10^8 cells. Dramatically decreasing the time, cost, and background variability between each RNA target captured. See Figure 1B for mechanism.

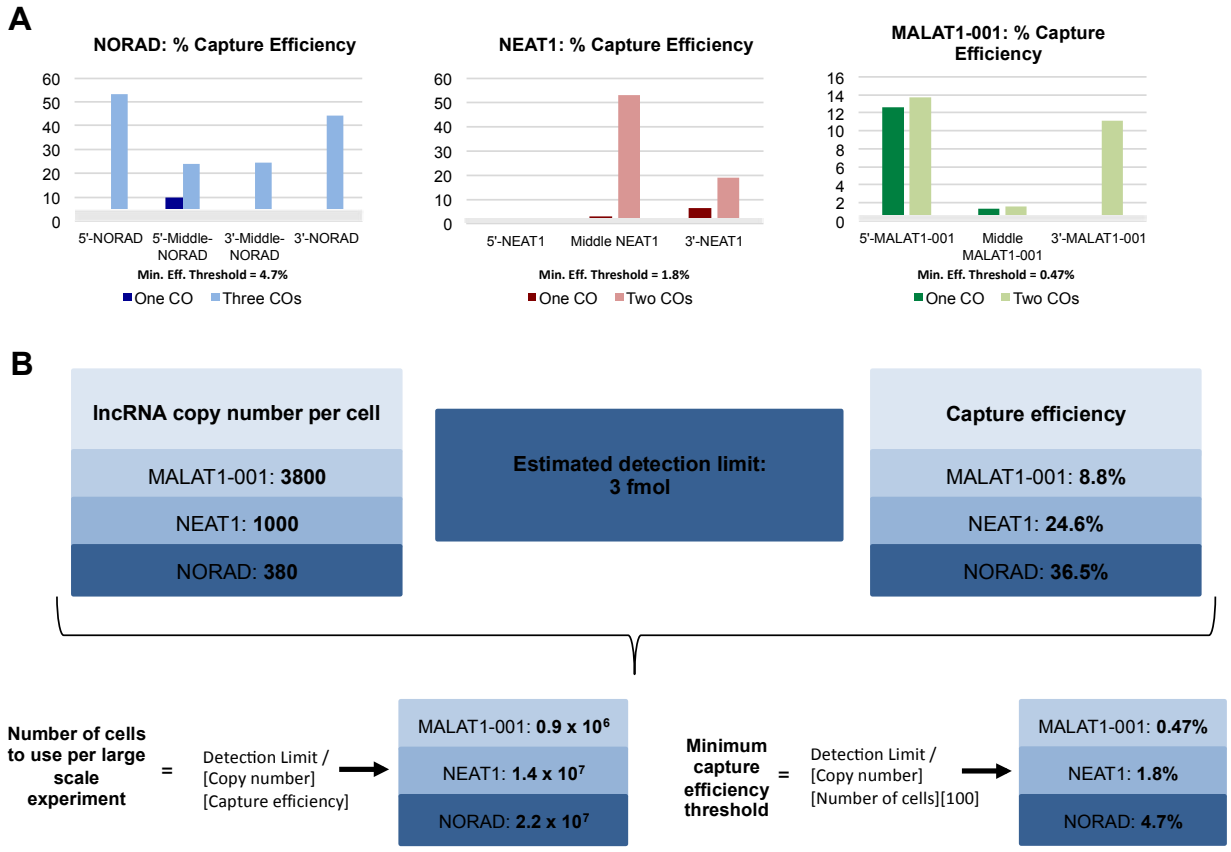

Figure S1: Determination of capture efficiencies and cell numbers required for protein detection. (In reference to the Results section and Figure 2A and 2B) (A) Two COs and three RT-qPCR assays were designed for NEAT1 and MALAT1 capture. Three COs and four RT-qPCR assays were designed for NORAD capture. (Tables S2 and S3). The efficiency of capture was assessed for captures using multiple COs (All listed in Tables S2) and for captures using only one CO (5'-NORAD, 5'-NEAT, and 5'MALAT1-001 COs were used for the captures for "One CO" efficiency calculations. Capture efficiency is defined here as the percentage of the target lncRNA transcripts that were captured and released from the beads relative to the amount available for capture in the lysate sample. The amount of target in each sample was determined using multiple qPCR assays (Table S3) that amplified various regions along each target lncRNA: near the 5'-end, towards the middle, and near the 3'-end (Figure 2A). The results here show that when using only one CO per RNA target the qPCR amplification declines at regions further from the location of the CO. The data shows that using multiple capture oligonucleotides reduces this effect and that the capture efficiencies at each location along the RNA exceed the calculated minimum efficiency thresholds for each target. For assessment of the variability in capture efficiencies, see Results, Capture Efficiency (B) The number of cells necessary to obtain sufficient protein for one capture experiment was determined using the estimated copy number of each lncRNA transcript, the estimated detection limit for a protein in a complex biological mixture, and the capture efficiency of each lncRNA as determined from small scale experiments. Then, using the determined cell number, a minimum capture efficiency threshold was calculated for each target. This threshold was used to ensure that the actual capture efficiencies calculated from each qPCR measured region of the large-scale capture experiments is sufficient to detect by MS even low abundance proteins. These efficiencies are indicated in the graphs in panel A as grey boxes. The multiple-CO captures provide sufficient copies of each target lncRNA to

surpass the capture efficiency threshold established above, while still allowing the isolation of specific isoforms of a given lncRNA.

Table S2: Capture oligonucleotide (CO) and release oligonucleotide (RO) sequences and amounts used. (In reference to the Results section and Figure 2A).

Target	Capture Oligonucleotide Sequence 5'-3' (toehold in red)	Amount of CO used (pmol)	Release Oligonucleotide Sequence 5'-3' (toehold in red)	Amount of RO used (nmol)
5`-MALAT1 - 001	TCGTATCTCAATATTTTCA TTTTCTATCTTGTTTCTAT	75	ATAGAAACAAGATAGAAA ATGAAAATATTGAGATAC GA	7.5
3`-MALAT1 - 001	TCGTATCTCACCCAGCAT TACAGTTCTT	75	AAGAACTGTAATGCTGG GTGAGATACGA	7.5
5`-NEAT1	TCGTATCTAAGTTATTTC ATCAGGCTAAGAA	357	TTCTTAGCCTGATGAAAT AACTTAGATACGA	35.7
3`-NEAT1	TCGTATCTCAAACATTCA AATTAAAAATACAATAAAT	178.5	ATTTATTGTATTTTTAATT TGAATGTTTGAGATACGA	17.9
5`-NORAD	TCGTATCTAGCTTTTTCA TATTATATACACAG	178.5	CTGTGTATATAATATGAA AAAGCTAGATACGA	17.9
NORAD Middle	TCGTATCTAATGGTTATA TCTGATAGTGTCTT	178.5	AAGACACTATCAGATATA ACCATTAGATACGA	17.9
3`-NORAD	CAACTGTCACATCAATGG CTATCAAAATGTAAATAT GG	178.5	CCATATTTACATTTTGATA GCCATTGATGTGACAGTT G	17.9
Scrambled	CAACTGTCGCGTCTTTAT TTAGTTTACTCTTGATTGT T	75	AACAATCAAGAGTAAACT AAATAAAGACGCGACAGT TG	7.5

Table S3: Primer and probe sequences for qPCR assays. (In reference to Results section and Figure 2A).

Target	Forward 5'-3'	Probe 5'-3'	Reverse 5'-3'
5`-MALAT1-001	CAGGATTCCAGGAACCAGTG	CTAGGACTGAGGAGCAAGCGAGC	TTCCTATCTTCACCACGAACTG
MALAT1-001 Middle	GAACGAATGTAACTTTAAGGCAGG	TCCAGGCACATGGCAATAGAGGC	GATCATAATCTCCCACCTGTCTAAG
3`-MALAT1-001	ACGTATTGTTTTCTCAGGTTTTGC	AAAGATGCTGGTGGTTGGCACTC	GATTTGAACCCCGTCCTGG
5`-NEAT1	GCCTCCGGTCATACTAGTTTTG	CCTTGTAGATGGAGCTTGCAGATGGA	AGGTGGGTAGGTGAGAGG
NEAT1-Middle	CACCTAAAATCAGTTTGGAAAACAAG	CTCTCCCCACAATCCCCATCCC	AGGTGGGTAGGTGAGAGG
3`-NEAT1	CACCTAAAATCAGTTTGGAAAACAAG	CTCTCCCCACAATCCCCATCCC	ACATGTAGTAAAGGCACCTCG
5`-NORAD	TGGCTGTGCCCAGACCTT	CCACGGCCGCCATTAGTC	CAGCGAACCTCTCTTTCCCACCC
5`/Middle-NORAD	CACGTTTGTTAAGTGGGTTAGATG	ACATGGAGCTGGAAGACCTGAGAAG	AATATGACCAGTCTAGCATAGAACC
3`/Middle-NORAD	GACACGTGCCTATATCCATCAG	CCTTCCAACTCCTCTCCACCACC	CTTCTAAATACGAACATTCTGGTCTAG
3`-NORAD	TTGTTAAGCCACCTCTGAGC	TGCCAACCTAATGAACAAGTCCTGACA	CCTGTATAATTCCTTCTGCCCC

Table S4: qPCR amplification parameters. (In reference to Methods section)

Program	Cycles	Temp (□)	Acquisition Mode	Analysis Mode	Hold	Ramp Rate (□/s)
Pre-Incubation	1	95	None	None	8 min	4.4
Amplification	45	95	None	Quantification	10 sec	4.4
	45	57	None	Quantification	20 sec	2.2
	45	72	Single	Quantification	1 sec	4.4
Cool	1	40	None	None	10 sec	2.2

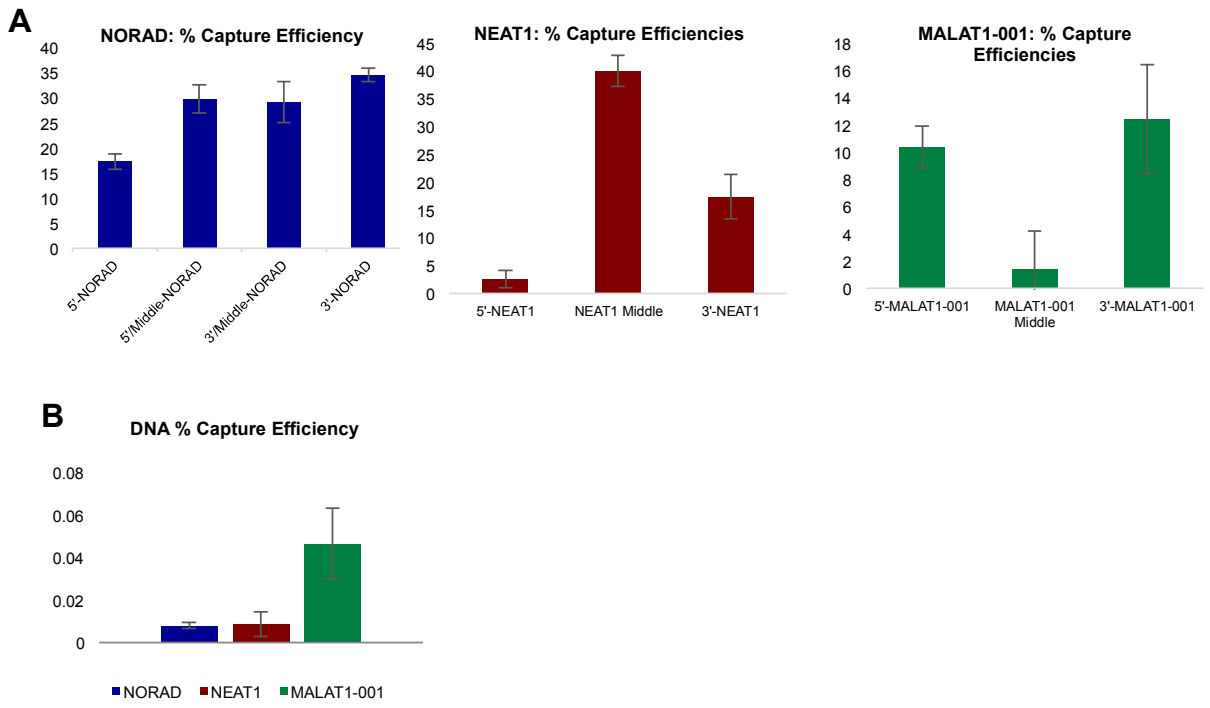

Figure S2. (In reference to Figure 2) (A) Capture efficiencies (CE) at each qPCR region for each target lncRNA. There are differences in the CE measurement at different regions of the same lncRNA. These differences could be due to the secondary structure of the RNA or the protein occupancy at the capture region of the lncRNA, resulting in fewer captured molecules of the target RNA, or the variable efficiencies of the RT-qPCR assays in complex samples. It is also an artifact of the experimental design; up to a certain point, increases in CO concentration during hybridization can increase the CE (data not shown), and this must be optimized for each target lncRNA. However, the volume of streptavidin-coated beads needed, and thus a large portion of the cost of the experiment, also increases with increasing CO concentration. For this reason, the concentration of each CO added to the capture experiment was determined to ensure that the CE would surpass the calculated CE threshold (Figure S1B), but not necessarily to maximize the CE for each region (see Table S2 for concentrations used). (B) DNA capture efficiencies (CE). Each CO can hybridize not only the target lncRNA, but also its corresponding DNA sequence, meaning any proteins interacting with that region of the genome could be falsely identified as RNA interactors. To ensure this had minimal effect on the overall protein content of each capture sample, the DNA CE was found by measuring the number of DNA copies corresponding to the target RNA sequence in the captured sample and comparing it to the number of copies in the lysate. Based on DNA CE values that are consistently lower than 0.1%, the small total DNA copy number (two copies per cell), and the mass spectrometry sensitivity (approximately 3 femtomole, or 1.8×10^9 copies), any contamination by DNA-interacting proteins is not able to affect the specificity of the RNA captures.