



iDRiP for the systematic discovery of proteins bound directly to noncoding RNA

Hsueh-Ping Chu^{1,2,3}✉, Anand Minajigi^{1,2,6}, Yunfei Chen³, Robert Morris^{4,5}, Chia-Yu Guh³, Yu-Hung Hsieh³, Myriam Boukhali^{4,5}, Wilhelm Haas^{4,5} and Jeannie T. Lee^{1,2}✉

More than 90% of the human genome is transcribed into noncoding RNAs, but their functional characterization has lagged behind. A major bottleneck in the understanding of their functions and mechanisms has been a dearth of systematic methods for identifying interacting protein partners. There now exist several methods, including identification of direct RNA interacting proteins (iDRiP), chromatin isolation by RNA purification (ChIRP), and RNA antisense purification, each previously applied towards identifying a proteome for the prototype noncoding RNA, Xist. iDRiP has recently been modified to successfully identify proteomes for two additional noncoding RNAs of interest, TERRA and U1 RNA. Here we describe the modified protocol in detail, highlighting technical differences that facilitate capture of various noncoding RNAs. The protocol can be applied to short and long RNAs in both cultured cells and tissues, and requires ~1 week from start to finish. Here we also perform a comparative analysis between iDRiP and ChIRP. We obtain partially overlapping profiles, but find that iDRiP yields a greater number of specific proteins and fewer mitochondrial contaminants. With an increasing number of essential long noncoding RNAs being described, robust RNA-centric protein capture methods are critical for the probing of noncoding RNA function and mechanism.

Introduction

The Human Genome and ENCODE Projects have shown that the protein-coding potential of the mammalian genome is <2% and that >90% of the genome is transcribed^{1,2}. While a growing number of noncoding RNA (ncRNA) is being shown to have important cellular roles during development and disease pathogenesis^{3–8}, the purpose of the pervasive transcription and the function of its RNA output have remained under active debate. With the possibility of performing systematic CRISPR-Cas9 genome editing and advances in structural characterization^{9–12}, a functional dissection of ncRNA is now more realistic than ever. However, mechanistic insight will ultimately depend on understanding the interplay between the ncRNA and associated regulatory factors. An understanding of the ncRNA mechanism has lagged behind because of challenges presented by a dearth of approachable methods to map RNA-centered interactomes and to identify high-confidence interacting protein partners.

Several *in vivo* RNA proteomic methods have recently been developed, including identification of direct RNA interacting proteins (iDRiP)^{13,14}, RNA antisense purification (RAP)¹⁵, chromatin isolation by RNA purification (ChIRP)¹⁶, capture hybridization analysis of RNA targets (CHART)^{17,18}, biotinylated RNA binding proteins (BioRBP)¹⁹ and CRISPR-assisted RNA–protein interaction detection (CARPID)²⁰. Many were initially tested on the RNA gold standard, Xist—the long ncRNA responsible for establishing global silencing of the inactive X-chromosome^{5,21} (Table 1). The methods have produced an overlapping set of interacting proteins, in addition to factors that are unique to each method^{14–16,20}. RNA-proteomic methods are still far from routine, and whether each can be generally applied to all ncRNAs remains unclear. Also lacking is a comparative analysis between methods, as well as between various RNA species to determine the applicability of the *in vivo* method to RNAs not involved in dosage compensation.

Here we describe a detailed protocol for iDRiP¹⁴ demonstrated to be successful for two other ncRNAs: (i) TERRA, a universal telomeric transcript that scaffolds factors critical for gene expression and for protection of chromosomal ends^{13,22–25}; and (ii) U1, a splicing factor with well-established interacting proteins that thereby serves as a positive control. We then conduct a performance analysis

¹Department of Molecular Biology, Massachusetts General Hospital, Boston, MA, USA. ²Department of Genetics, Harvard Medical School, Boston, MA, USA. ³Institute of Molecular and Cellular Biology, National Taiwan University, Taipei, Taiwan ROC. ⁴Massachusetts General Hospital Cancer Center, Charlestown, Boston, MA, USA. ⁵Department of Medicine, Harvard Medical School, Boston, MA, USA. ⁶Present address: Eisai, Inc., Andover, MA, USA. ✉e-mail: cchu2017@ntu.edu.tw; lee@molbio.mgh.harvard.edu

Table 1 | Comparison of methods for RNA-based capture of protein targets in vivo

Method	Cross-linking	Method of capture	Advances	Examples of tested RNAs	Ref
iDRiP-MS	UV irradiation (254 nm)	Use of a limited number of biotinylated antisense DNA capture probes (20–25 mers) to enhance capture specificity	Identification of directly interacting proteins to RNA of interest	Xist, TERRA, U1	13,14
ChIRP-MS	Formaldehyde	Use of extensive tiling biotinylated antisense DNA probes (20–25 mers)	Identification of proteins in proximity to RNA of interest	Xist, U1, U2 SARS-CoV19	16,53
RAP-MS	UV irradiation (254 nm)	Use of long tiling biotinylated antisense DNA probes (90 mers)	Identification of proteins in proximity to RNA of interest	Xist, NORAD, SARS-CoV19	15,54,55
CHART-MS	Formaldehyde	Use of limited biotinylated antisense DNA probes (20–25 mers)	Identification of proteins in proximity to RNA of interest	MALAT1, NEAT1	56
CARPID	No cross-linking	Use of dCas9Rx-biotin ligase fusion protein with gRNA to target RNA of interest. In vivo labeling of associated proteins with biotin	Identification of proteins in proximity to RNA of interest	Xist, MALAT1, DANCR	20
BioRBP	No cross-linking	Use of biotin ligase (BirA*)-PP7 coat protein fusion, with PP7-tagging of RNA of interest. In vivo labeling of associated protein with biotin	Identification of proteins in proximity to RNA of interest	Tsix	19

relative to another method, ChIRP¹⁶. The effectiveness of iDRiP for the highly repetitive and tightly chromatin-bound TERRA RNA is of interest, and modifications we incorporated could be useful for the capturing of protein interactomes for other repetitive RNAs.

Applications of the method

iDRiP is an RNA-centric method to systematically identify proteins that directly interact with an RNA of interest from cultured cells^{13,14} or tissues²⁶. It does not rely on an antibody-based immunoprecipitation procedure to capture RNA–protein complexes. In principle, the protocol can be applied to short or long RNAs, chromatin-bound or soluble RNAs, and nuclear or cytoplasmic RNAs. Previously, we had successfully applied this protocol to ncRNAs including the splicing factor U1 (164 nt), the chromatin-bound ncRNA Xist (17 kb) and the telomeric transcript TERRA (heterogeneous size, hundreds to >9 kb). The iDRiP protocol enabled us to isolate and characterize a large number of interacting proteins^{13,14}. Because iDRiP employs shorter oligos (20–25 mer) than other methods (Table 1), it more easily accesses RNA–protein complexes than a method such as RAP¹⁵, which uses considerably longer oligo probes.

Limitations and alternative modifications

Our method has been tested mostly on nuclear ncRNAs, including the soluble U1 RNA and the chromatin-bound Xist and TERRA RNAs^{13,14}. It should in principle also be applicable to cytoplasmic RNAs, though we have not explicitly tested them. We have also not tested iDRiP for RNAs <100 nt such as microRNAs and PIWI-interacting RNAs (piRNAs) (21–35 nt). Because iDRiP probe sizes are typically ~20–25 mer, it is theoretically possible to design probes for small RNAs as well. Modified DNA oligo probes such as locked nucleic acid chemistry with phosphorothioate backbones could be used, as locked nucleic acids generally have a higher affinity for nucleic acids, and phosphorothioate backbones lend greater stability than unmodified oligonucleotide probes. Additionally, iDRiP may be best suited to RNAs of moderate-to-high abundance. Xist RNA is present at only 100–300 copies per cell²⁷, while TERRA and U1 are some of the most abundantly expressed transcripts in embryonic stem (ES) cells¹³. Capturing very low abundance RNAs may be more challenging, as more cells would have to be used and the signal-to-noise ratio may be diminished for very-low-abundance RNAs (<100 copies/cell) (see Troubleshooting). Enzymatic methods, such as the recently reported BioRBP¹⁹ and CARPID²⁰, may therefore be more suitable for low-abundance transcripts, though one disadvantage of these alternatives is that they identify both directly as well as indirectly interacting proteins (see ‘Considerations for selecting an RNA proteomics method’).

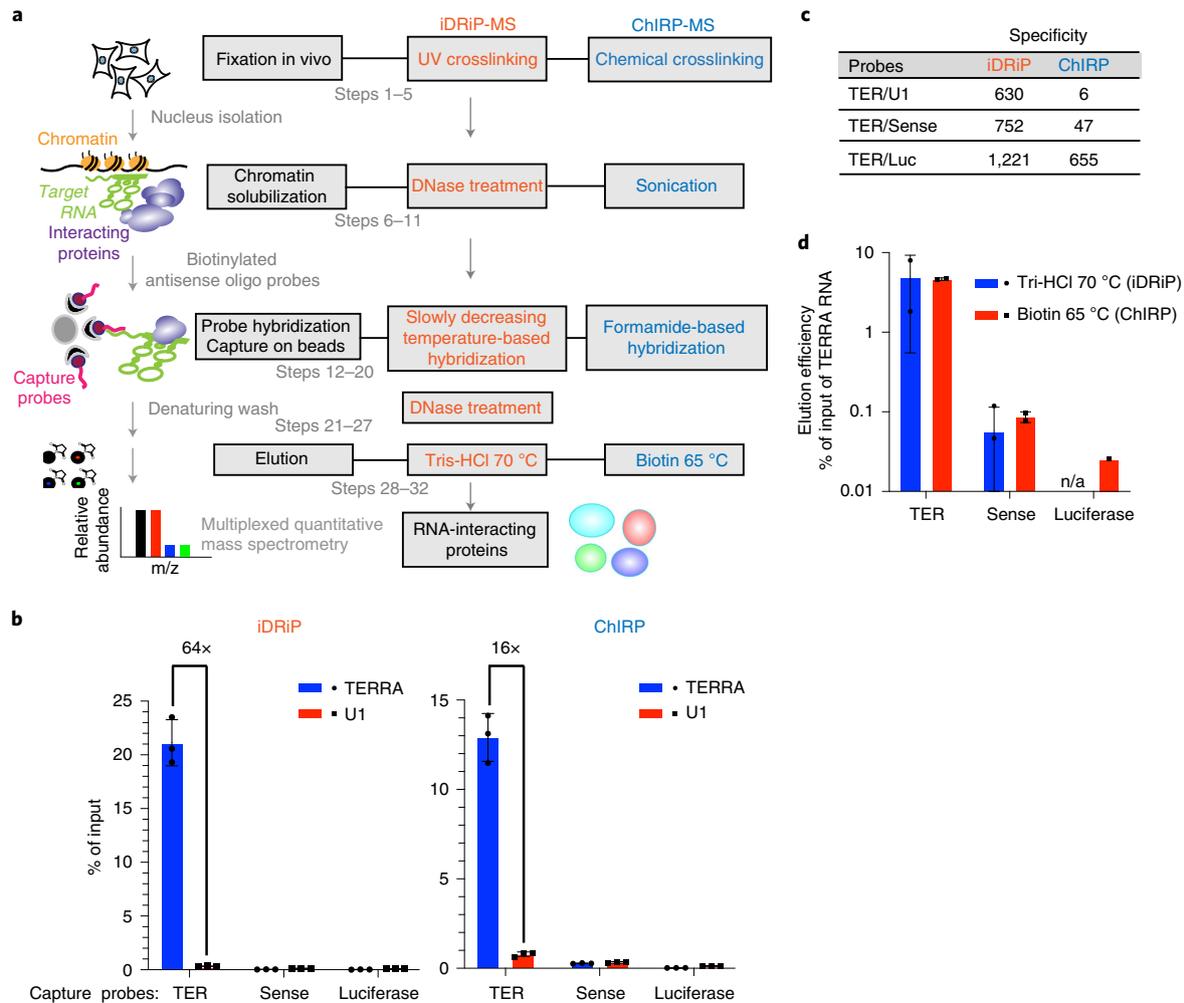


Fig. 1 | Overview of iDRiP and ChIRP. a, Workflows of iDRiP and ChIRP. **b**, The efficiency and specificity of TERRA recovery by iDRiP versus ChIRP using mES cells, expressed as % of input after RT-qPCR to detect PAR-TERRA, a TERRA transcript¹³ derived from sex chromosome, and U1 RNA. Capture probes for TERRA (TER) and the control probes (sense and luciferase) are indicated. TERRA-iDRiP and U1-iDRiP are from published data¹³. Mean ± SD are shown. **c**, Specificity as determined by fold enrichment of TERRA after normalization to various negative controls as indicated. **d**, Elution efficiencies of iDRiP and ChIRP expressed as % of input. Mean ± SD are shown. n/a, nondetectable.

Comparison with other methods

The current technologies using RNA-centric methodologies to identify RNA interacting proteins in vivo include iDRiP^{13,14}, RAP¹⁵, ChIRP¹⁶, CHART^{17,18}, BioRBP¹⁹ and CARPID²⁰ (Table 1). Many methods were initially tested on Xist RNA, with the methods generating a partially overlapping set of interacting proteins. Here we focus on a comparative analysis between iDRiP and ChIRP, as the two methods both yielded dozens of proteins for Xist, whereas RAP yielded only a dozen or so. iDRiP and ChIRP differ in a number of fundamental ways (Fig. 1a). Although both techniques probe in vivo interactions under physiological conditions, substantial differences exist in how RNA–protein interactions are fixed, how chromatin is solubilized, how capture probes are annealed and, finally, how captured RNA–protein complexes are eluted. At the beginning of both protocols, ES cells are cross-linked to preserve in situ RNA–protein interactions. iDRiP utilizes UV light (at 254 nm) to photo-cross-link RNA to proteins, whereas ChIRP utilizes formaldehyde to chemically cross-link the macromolecules. Because UV cross-linking is more efficient for capturing RNA–protein interactions than DNA–protein or protein–protein interactions, this photo-cross-linking preferentially identifies direct RNA–protein interactions²⁸. Furthermore, because UV light at 254 nm is a short-range cross-linker (‘zero-length’ scale), the RNA–protein interactions tend to be direct interactions. By contrast, the formaldehyde cross-linking employed during ChIRP fixes much larger macromolecular networks

and typically results in identification of RNA–protein interactions that are both direct and indirect, potentially brought together by multiple intermediary factors. They could, for instance, include RNAs and proteins brought together by other protein subunits, RNA–DNA hybrids that interact with proteins, or RNA–chromatin–protein interactions.

Next, ES nuclei are isolated, and chromatin is solubilized prior to isolation of RNA–protein complexes (Fig. 1a). In iDRiP, nuclei are treated with DNase I (Extended Data Fig. 1a) to release the large networks of chromatin that would be prohibitive to the isolation of discrete RNA–protein–chromatin complexes. In ChIRP, vigorous sonication is used to achieve the same end. iDRiP was designed to avoid sonication, due to concern about the physically disruptive action that could potentially fragment and destroy large RNAs, such as Xist and TERRA, and their associated protein complexes. Later, biotinylated antisense oligonucleotides (ASO) of 20–25 DNA-based nucleotides are used to capture RNA-specific complexes. In ChIRP, a formamide-based buffer is applied to enable ASO hybridization at a lower temperature of 37 °C, as formamide lowers the melting temperature (T_m) for RNA–DNA duplexes. To elute, excess biotin is used to compete away and thereby release bound RNA–protein complexes. By contrast, iDRiP applies a non-denaturing hybridization buffer, so that ASO annealing has to be performed at a higher temperature of 65 °C, followed by slow cooling to promote ASO–RNA hybridization. To elute, iDRiP employs heated washes to 70 °C and disruptive detergents to denature the hybridized ASO–RNA interaction and wash away RNA–protein interactions that have not been photo-cross-linked.

Capture efficiency and specificity, as well as signal-to-noise ratio, are the most important factors that determine whether the protocol will successfully identify proteins directly interacting with the RNA. Using iDRiP, we recently identified an extensive proteome for TERRA¹³, a long ncRNA transcribed across telomeric ends and comprising hundreds to thousands of repeats of the telomeric sequence UUAGGG^{22,24,25}. At the same time that we performed iDRiP, we also performed ChIRP for TERRA in mouse ES (mES) cells. The two methods were carried out under similar conditions with regard to cell culture and mES cell passage numbers. Most importantly, the capture probes used were identical. (Note: although ChIRP was performed around the same time, we only published the TERRA-iDRiP results in the 2017 paper¹³ because the performance characteristics of TERRA-iDRiP were more desirable and the method yielded more directly interacting proteins.)

Here we carried out a comparative analysis of the two resulting proteomes. We had designed capture probes targeting UUAGGG repeats and several negative control probes, including sense (TERRA-reverse complement), luciferase (nonmammalian) and the mammalian splicing RNA U1. After TERRA capture, qRT-PCR showed that recovery of TERRA RNA using TERRA capture probes was good for both iDRiP (>20% of input) and ChIRP (~13% of input) (Fig. 1b). However, TERRA-capture iDRiP yielded a more favorable signal-to-noise ratio, with a 64-fold enrichment of TERRA over U1 RNA, in comparison with the 16-fold enrichment using ChIRP (Fig. 1b). This difference suggested that iDRiP capture gives a slightly greater yield and specificity for TERRA. Greater specificity of iDRiP was further supported by examining the relative enrichment of TERRA probe capture over control probe capture (Fig. 1c). TERRA was enriched over 1,000-fold in TERRA-iDRiP compared with Luc-iDRiP capture (Fig. 1c). On the other hand, elution methods for ChIRP (excess biotin) and iDRiP (heat denaturation) were similarly efficient (Fig. 1d).

To obtain a quantitative proteomic analysis, we performed quantitative mass spectrometry (MS) using isobaric labeling methods²⁹ and compared the enrichments over negative controls—TERRA-sense and luciferase—for ChIRP relative to our previously published iDRiP profiles¹³. Captured proteins were quantitatively compared by multiplexed MS-based proteomics using tandem-mass-tag (TMT) reagents and MS3-based quantification^{30,31}. We averaged two biological replicates and reported proteins that are twofold enriched over luciferase controls. The correlation of two biological replicates of ChIRP shows a higher correlation in the global dataset than that of iDRiP (Extended Data Fig. 1a,b). Gene ontology (GO) analysis for enriched hits was performed using the Database for Annotation, Visualization and Integrated Discovery (DAVID) bioinformatic resources³². The number of twofold enriched targets is higher in ChIRP than in iDRiP, whereas the number of DAVID significant Biological Process gene sets (false discovery rate (FDR) <25%) is higher in iDRiP than in ChIRP (Extended Data Fig. 1c). There are 134 enriched proteins for iDRiP-MS¹³ and 205 enriched proteins for ChIRP-MS (Fig. 2a; Supplementary Tables 1,2). Among them, 27 proteins were shared (Fig. 2a, left panel). Among the top 30 hits, 11 proteins were shared (Fig. 2a, right panel). Common hits included shelterin proteins and factors involved in alternative lengthening of telomeres, chromatin modification and cell cycle regulation (Fig. 2b,c). The overall correlation between the iDRiP and ChIRP datasets was low (Pearson's $r = 0.13$, Fig. 2d, left panel). However, when considering only

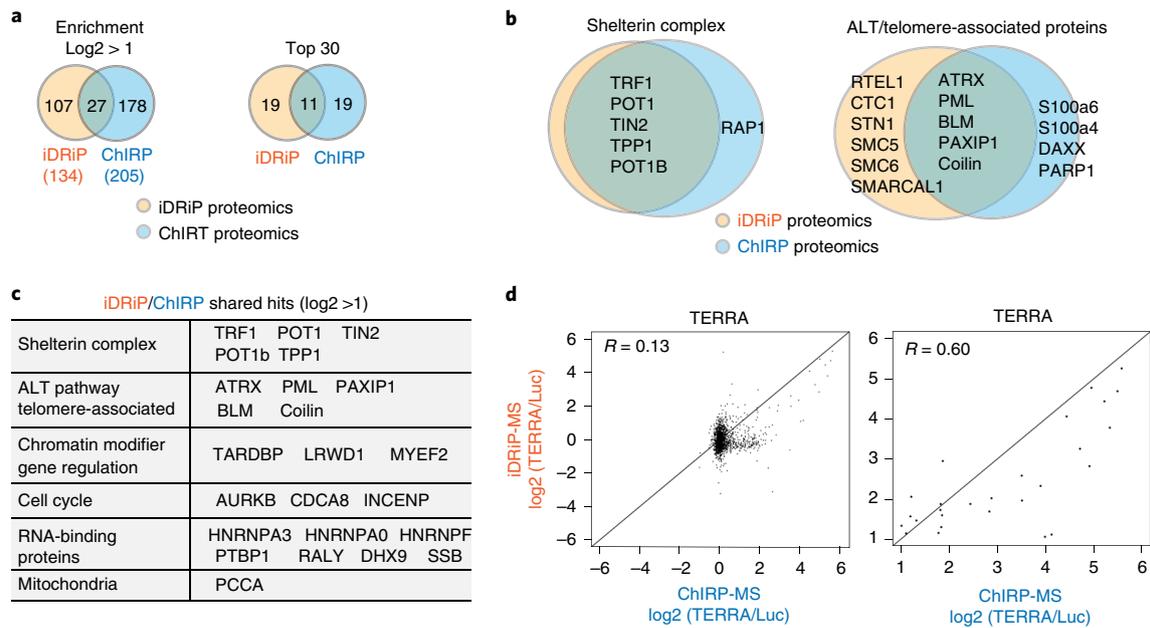
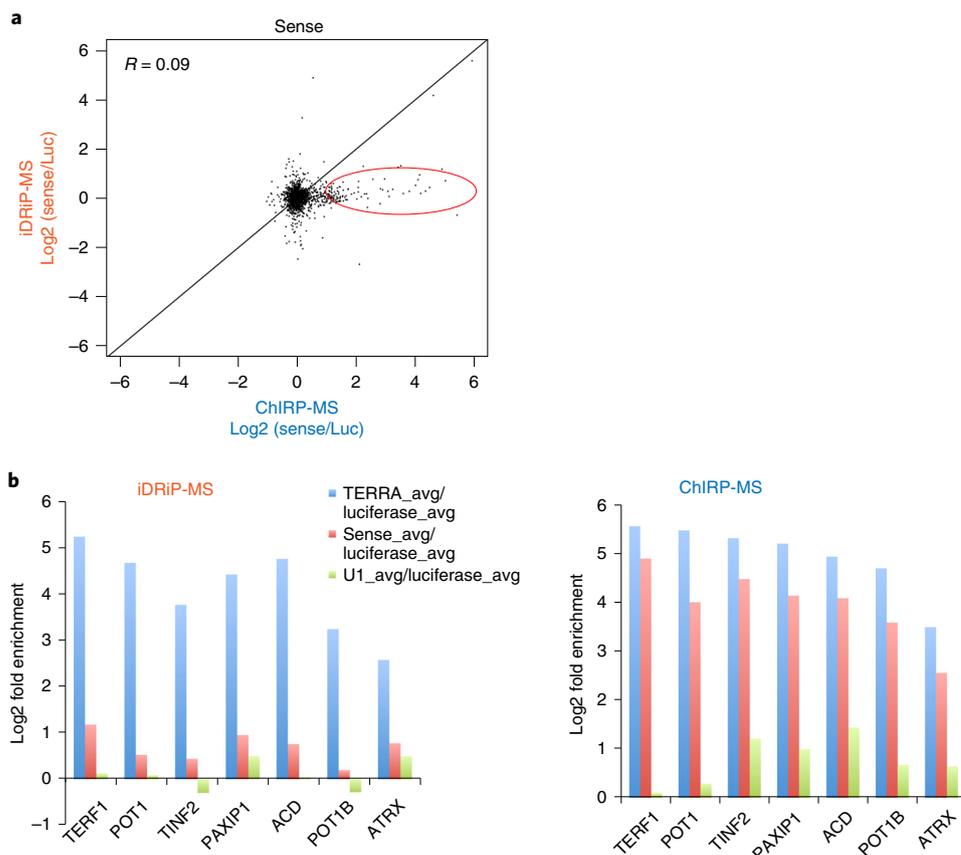


Fig. 2 | Characterization of the TERRA proteome as obtained by iDRiP versus ChIRP. a, Venn diagrams indicating the number of enriched proteins by iDRiP and ChIRP, along with the number of proteins that overlap. Candidates shown are enriched by $\text{log}_2(\text{TERRA/luciferase}) > 1$ (left). The top 30 candidates are shown (right). **b**, Shelterin complex and proteins related to alternative lengthening of telomeres (ALT) are enriched within the TERRA proteome by both methods. **c**, The shared TERRA proteome as identified by iDRiP and ChIRP. **d**, Scatterplots showing correlation between TERRA hits by iDRiP versus ChIRP. Proteomic counts were normalized using luciferase. Log_2 enrichment ratios are shown. Left: correlation using data of all proteins obtained from MS. Right: correlation using data of proteins that are enriched over twofold obtained from MS.

the enriched subsets ($\text{log}_2 > 1$), the correlation was high (Pearson's $r = 0.63$, Fig. 2d, right panel). Thus, in spite of fundamental methodological differences between iDRiP and ChIRP, the top hits were shared.

There were nevertheless subsets of proteins unique to each dataset. For example, some telomere-associated proteins were only enriched in the iDRiP dataset, including CTC1 and STN1 of the CST complex, as well as RTEL1 (Fig. 2b, Supplementary Tables 1, 2). Notably, the sense control returned far fewer enriched proteins by iDRiP than by ChIRP (Fig. 3a). A number of telomere-associated proteins were enriched in the ChIRP sense control but not in the iDRiP counterpart (Fig. 3b, Extended Data Fig. 2), possibly because ChIRP pulled down indirect interactions mediated via the telomeric chromatin. Significantly, GO enrichment analysis revealed that mitochondrial proteins were major contaminants in the TERRA-ChIRP experiment, but not in the TERRA-iDRiP experiment (Fig. 3c). Consistent with their being contaminants rather than physiological interactors, the same mitochondrial proteins were also enriched in the U1-ChIRP samples (Extended Data Fig. 3). Thus, the ChIRP methodology isolates more false positives originating outside of the nucleus such as mitochondria, which should not come in contact with telomeres under physiological circumstances.

We previously biochemically and functionally validated the iDRiP candidate, ATRX, as a major TERRA-interacting protein¹³. By ChIRP, ATRX was also enriched with TERRA. However, it was not enriched relative to the sense control and would have likely been eliminated as an interacting factor from the ChIRP TERRA proteome. Promyelocytic leukemia protein (PML)—a tumor suppressor protein frequently mutated in promyelocytic leukemia and that was recently shown to interact with telomeres³³—was also identified as a hit by both iDRiP and ChIRP (Fig. 2b, Supplementary Table 1). However, by ChIRP, PML was not enriched relative to the sense control and would not have been considered an enriched factor had we performed ChIRP alone (Supplementary Table 2). Using immuno-RNA fluorescent in situ hybridization (FISH) to visualize the localization of RNA and proteins, we observed that TERRA RNA showed 60–70% colocalization with both ATRX and PML (Fig. 4a), consistent with a TERRA–PML interaction. We further validated TERRA interacting candidates by UV light cross-linking RNA immunoprecipitation (UV-RIP) (Fig. 4b,c). EZH2, EHMT1, POT1 and aurora B are highly associated with TERRA, as shown in RNA slot blot (Fig. 4b) and qRT-PCR (Fig. 4c). For TERRA, iDRiP generally yielded greater specificity and a superior signal-to-noise ratio relative to sense, luciferase and U1 controls (Fig. 1c, Extended Data Fig. 2).



c **TERRA-iDRIP-MS**

GO term	Genes	Count	%	P-value	Benjamini
Nucleus		109	82.6	7.2E-36	1.2E-33
Chromosome		31	23.5	1.5E-25	1.3E-23
Nucleoplasm		57	43.2	2.0E-23	1.2E-21
Chromosome, telomeric region		16	12.1	7.7E-19	3.3E-17
Nuclear chromosome, telomeric region		18	13.6	7.0E-8	2.4E-16
Nucleolus		29	22.0	1.2E-2	3.4E-11
Chromosome, centromeric region		12	9.1	2.3E-9	5.7E-8
Nuclear telomere cap complex		5	3.8	6.0E-8	1.3E-6
Condensed nuclear chromosome, centromeric region		6	4.5	1.2E-7	2.3E-6
Intracellular ribonucleoprotein complex		14	10.6	2.0E-7	3.4E-6

TERRA-ChIRP-MS

GO term	Genes	Count	%	P-value	Benjamini
Mitochondrion		103	50.7	2.7E-55	5.7E-53
Mitochondrial matrix		31	15.3	2.2E-27	2.4E-25
Mitochondrial inner membrane		23	11.3	5.8E-11	4.1E-9
Nuclear telomere cap complex		6	3.0	2.0E-9	1.0E-7
Nuclear chromosome, telomeric region		12	5.9	6.0E-8	2.6E-6
Telosome		5	2.5	3.3E-7	1.2E-5
Chromosome, telomeric region		9	4.4	6.6E-7	2.0E-5
Nucleus		94	46.3	7.6E-7	2.0E-5
Intracellular ribonucleoprotein complex		16	7.9	9.0E-7	2.1E-5
Myelin sheath		12	5.9	4.4E-6	9.4E-5

Fig. 3 | iDRiP shows overall greater specificity for TERRA RNA. **a**, Scatterplot comparing log₂ enrichment of Sense-iDRiP and Sense-ChIRP results normalized to luciferase. Red circle indicates that many telomere-associated proteins are enriched by Sense-ChIRP but not by Sense-iDRiP. Pearson's *r* is shown. **b**, Captures by sense probes reveal differences between the iDRiP and ChIRP methods. **c**, GO cellular component enrichment analysis (generated using DAVID tools, GO_CC_term) for iDRiP and ChIRP datasets. Multiple mitochondrial contaminants are seen in the ChIRP dataset.

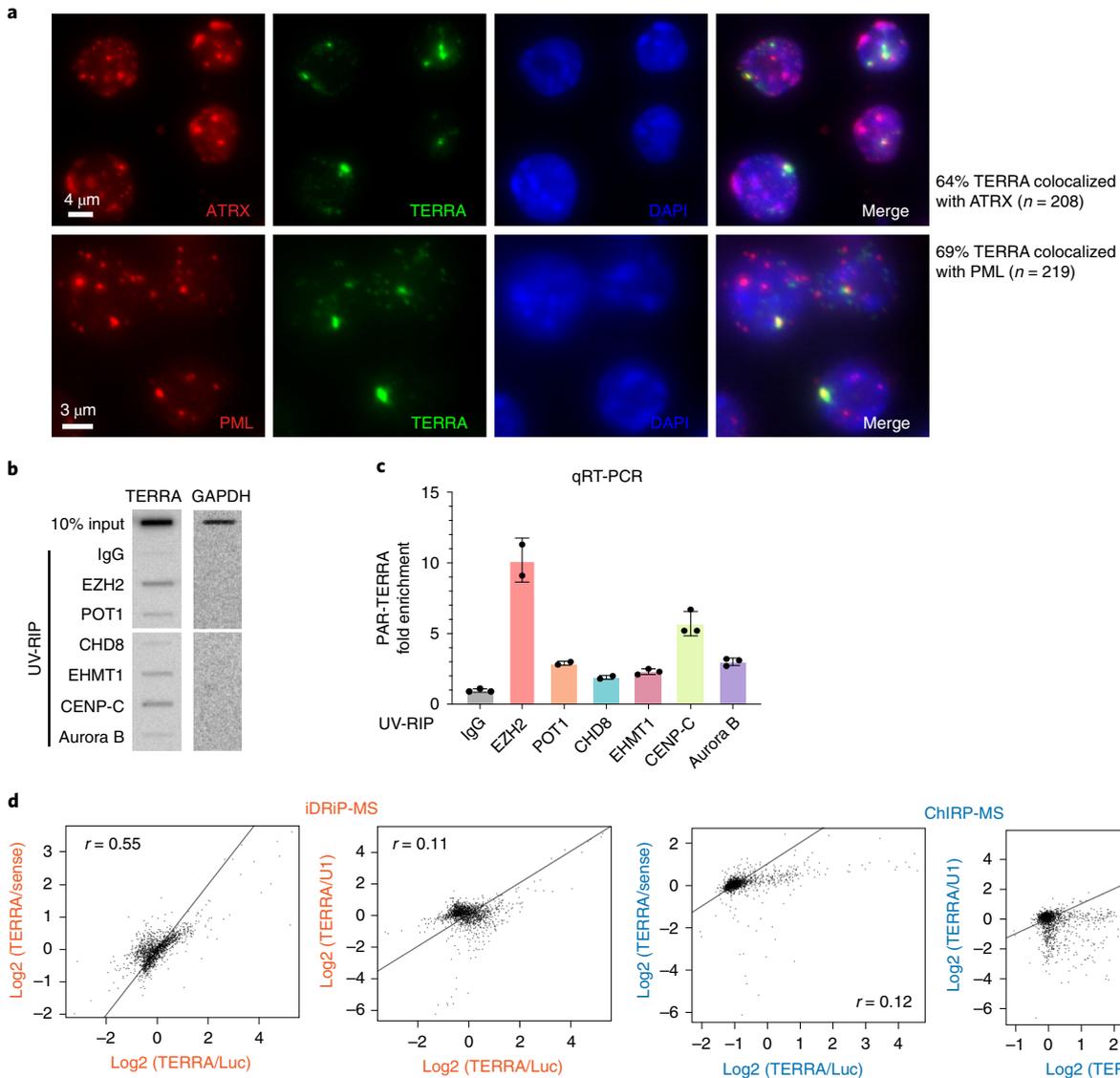


Fig. 4 | Validation of the TERRA proteome as obtained by iDRiP. **a**, Representative immuno-RNA-FISH staining showing that TERRA (green) is colocalized with ATRX (red, top panel) and PML (red, bottom panel) in mES cells. **b**, Slot blot shows the TERRA RNA enrichment after UV-RIP for EZH2, POT1, CHD8, EHMT1, CENP-C and Aurora B. Control: GAPDH RNA. See Supplementary Methods for experimental details. **c**, RT-qPCR shows that PAR-TERRA, a TERRA transcript¹³ derived from sex chromosome, was enriched after UV-RIP using antibodies against EZH2, POT1, CHD8, EHMT1, CENP-C and Aurora B. PAR-TERRA enrichment was normalized to the IgG capture. Error bars, SD. **d**, Scatterplots comparing log₂ enrichment of hits normalized to either Sense, U1 or luciferase control. iDRiP, left panels. ChIRP, right panels. Pearson's *r* is shown.

Further analysis using scatterplots demonstrated a high degree of correlation between iDRiP TERRA hits normalized to sense control versus luciferase control (Fig. 4d, left panel; Pearson's *r* = 0.55), but the same was not the case for ChIRP TERRA hits (Fig. 4d, right panels). Together, these data suggest that, at least for TERRA, iDRiP provides greater specificity for RNA proteomic analysis.

In comparing various normalization methods, we observed that while normalization of iDRiP data to sense and luciferase controls showed high correlation (Fig. 4d, left panel), normalization to U1

displayed a different pattern that suggested potential commonalities between TERRA and U1 proteomes and that therefore suggested that U1 would not be an appropriate negative control (Fig. 4d, right panel). U1 RNA is a component of the spliceosome, which comprises a large number of well-established small nuclear RNAs and protein complexes, including numerous U1 small nuclear ribonucleoproteins (snRNPs), known as U1 RNA direct interacting proteins^{34,35}.

Thus, to further test iDRiP and ChIRP methodologies, we performed RNA proteomics using the two methods in parallel to compare the recovery of known U1-interacting proteins (Fig. 5a,b). Both methods recovered U1 snRNPs, but with vastly different efficiencies. Specific U1 snRNPs were highly enriched by iDRiP (20- to 35-fold), with nearly all ranking among the top 33 hits among a total of 180 enriched proteins (Fig. 5b, Extended Data Fig. 4c, Supplementary Table 3). By contrast, the same proteins were enriched only several folds by ChIRP and were ranked low on the enrichment list (rank 48 to 1,133) (Fig. 5b, Extended Data Fig. 4d, Supplementary Table 4). ChIRP yielded a much larger enriched list of proteins (436 total). Nevertheless, iDRiP and ChIRP datasets showed 98 proteins in common (Fig. 5c). GO analysis revealed that proteins involved in mRNA processing, RNA splicing and spliceosomal snRNP assembly are enriched in both U1-iDRiP and U1-ChIRP (Fig. 5d and Extended Data Fig. 3a). To further analyze the specificity of the RNA proteomics, we examined the representation of other snRNPs—associated with the functionally related U2 (which exists in a particle with U1), and the biochemically and functionally separate U4, U5 and U6^{36,37}. Significantly, by ChIRP, almost all spliceosomal proteins are enriched, including snRNPs known to be associated with the U4/U6/U5 particle (Fig. 5e, red stars). In contrast, U1-iDRiP shows enrichment of most U1/U2 snRNPs but not of U4/U6/U5 snRNPs (Fig. 5d). Thus, iDRiP yielded greater specificity for snRNPs associated with the U1/U2 particle, whereas ChIRP appeared to isolate general splicing factors of both the U1/U2 and the U4/U6/U5 complexes.

Considerations for selecting an RNA proteomics method

Here, by investigating two well-established ncRNAs, TERRA and U1, we have performed a side-by-side analysis of two RNA proteomic methods. Our results show that both methods are able to capture bona fide RNA-interacting proteins. However, at least for TERRA and U1, the iDRiP method yielded a greater number of specific proteins and fewer contaminants (e.g., from mitochondria). The enrichment of mitochondrial proteins also suggests that nuclear isolation/extraction is less efficient in ChIRP. An advantage of ChIRP, however, is that indirectly interacting proteins could be identified. Because it employs a strong chemical cross-linking step, ChIRP could in principle be performed with less starting material. Analysis of the expression of genes encoding protein targets (twofold enriched) from ChIRP and iDRiP shows that the targets are not biased toward highly expressed RNA transcripts. In fact, the targets are clustered significantly near the low end of the transcriptome (Extended Data Fig. 3c), suggesting that both methods are able to capture low-abundance proteins. Our experience with TERRA and U1 here is similar to what has been reported for Xist, the RNA responsible for X-chromosome silencing^{14,16}. iDRiP identified a range of nuclear factors, including both low- (CTC1, OBFC1) and high-abundance epigenetic factors (SPEN, CTCF, cohesin, ATRX, HNRNPK).

Current limitations for iDRiP that would require further optimization are small RNA targets, such as microRNAs and piRNAs, in the 20–35 nt size range. Because capture probes are in the 20–25 nt size range, it is theoretically possible to design one to two probes to pull down small RNAs. Another current limitation is the capture of very-low-abundance RNAs. Enzymatic methods may be more suitable for low-abundance, sub-stoichiometric interactions. Such methods include the recently developed BioRBP¹⁹ and CARPID²⁰, both based on proximity protein labeling using biotinylation. Indeed, these methods have been tested on two dosage compensation RNAs, Tsix and Xist, respectively, both of moderate abundance^{27,38}. CARPID analysis of Xist revealed a large set of proteins (73) (ref. 20), among which a fraction overlapped with targets identified by our iDRiP analysis. Shared proteins include the experimentally validated cohesin subunits, RAD21, SMC1A¹⁴, ATRX³⁹ and BRG1⁴⁰. CARPID additionally identified ISWI chromatin remodeler, SNF2L, and the transcription coactivator, TAF15²⁰—supporting the idea that Xist RNA recruits repressors as well as repels activators^{14,40}. On the other hand, CARPID did not identify SPEN, which was identified by iDRiP, RAP and ChIRP. Technical differences between the CARPID and iDRiP methods likely account for proteins that are unique to each dataset. Using Tsix RNA as bait, BioRBP also identified a large number of high-confidence interactors, including the mRNA decapping enzyme DCP1A and the architectural glue protein CTCF¹⁹. The study then validated the proteins through functional

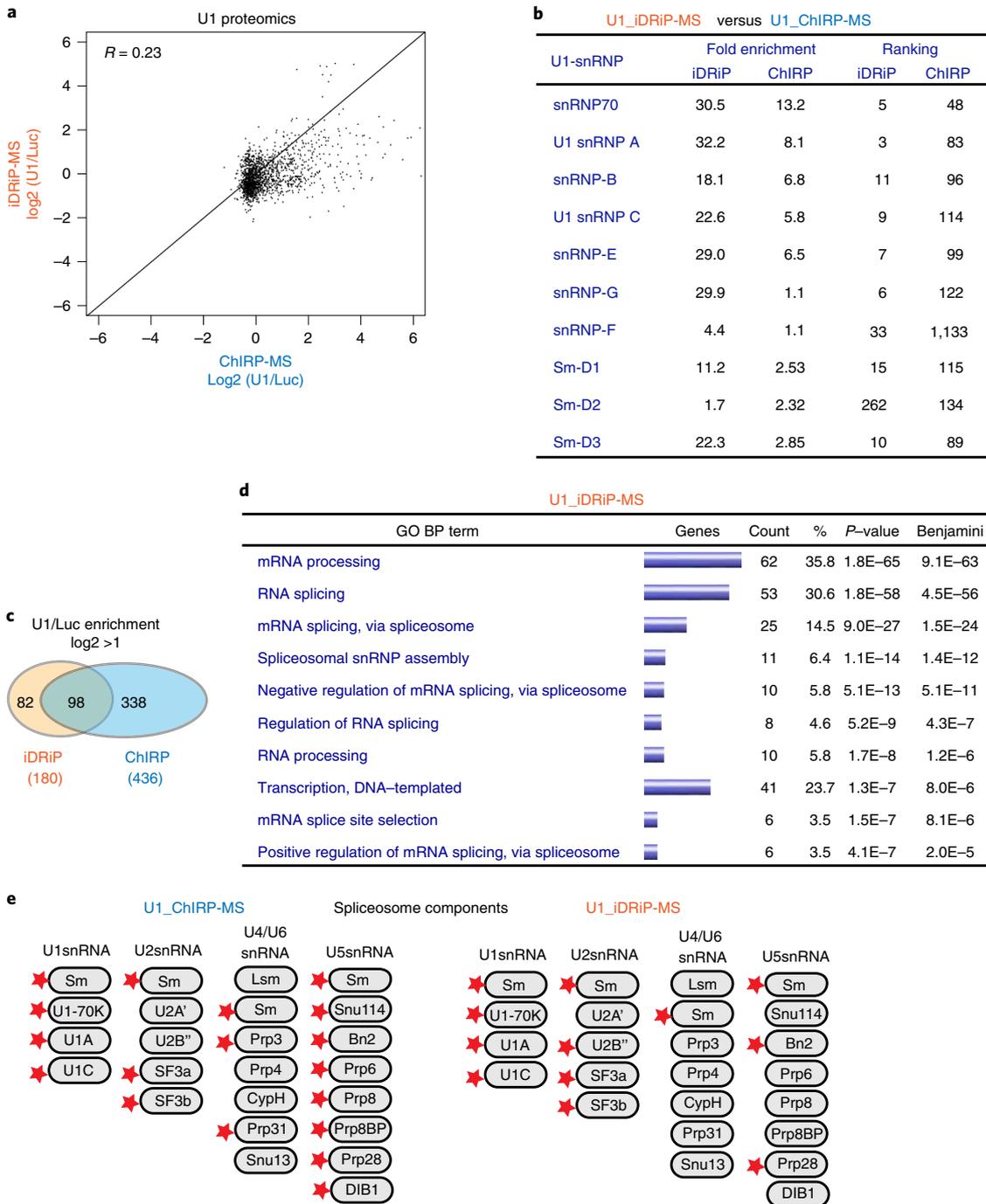


Fig. 5 | Differential U1 proteomes obtained by iDRiP versus ChIRP. **a**, Scatterplots comparing log₂ enrichment of U1-iDRiP and U1-ChIRP proteomics normalized to luciferase control from mES cells. Pearson's *r* is shown. **b**, U1 snRNPs were enriched by both methods. The fold-enrichment and relative rankings are shown. **c**, Venn diagram indicating the number of enriched proteins by iDRiP and ChIRP, along with the number of proteins that overlap. Candidates shown are enriched by log₂ (U1/luciferase) >1. The total number of enriched proteins is shown in parentheses. **d**, GO biological process (GO_BP_term) enrichment analysis for U1 interactors identified by iDRiP. **e**, Components of the spliceosome enriched by U1-ChIRP versus U1-iDRiP. Red stars indicate enriched components.

characterization, demonstrating a role for both in the formation of X-X pairs and breaking the epigenetic symmetry of two X chromosomes to determine allelic choice.

Given the pros and cons of various RNA proteomic methods, we do not suggest adopting a single method alone. Rather, we advocate for using various orthogonal methods to arrive at a complete view of the interacting proteome for any given RNA. Among the available methods discussed, only iDRiP

(and RAP) specifically targets direct interactors—made possible by the short-length cross-linker, UV. Whereas iDRiP enriches for these directly interacting proteins and is best used on moderate-to-high-abundance RNAs, the enzymatic methods (BioRBP and CARPID) may be used for lower-abundance transcripts, substoichiometric or transient interactors, and when proteins that do not directly interact and/or are within a large protein scaffold may be of interest. CHART and ChIRP would also offer alternative methods based on formaldehyde cross-linking and yield proteins that do not directly interact with RNA. Information gained from orthogonal methods could then be evaluated, and overlapping hits might be considered the highest-confidence hits, though nonoverlapping hits could arguably also be of interest owing to the unique experimental conditions of each method. Ultimately, a combination of a large family of RNA-proteomic methods including iDRiP, BioRBP, CARPID, CHART^{17,18}, ChIRP and RAP is beneficial to have in the RNA biologist's armamentarium.

Experimental design

Overview

To characterize RNA direct-interacting proteins *in vivo*, UV irradiation is applied to cross-link RNA and protein complexes. To minimize background due to DNA-bound proteins, iDRiP includes a critical DNase I treatment before elution of complexes, for purposes of fragmenting chromatin and eliminating proteins bound to single-stranded DNA that is complementary to the ASO. The washing steps are under a denaturation condition containing 1% sodium dodecyl sulfate (SDS) to eliminate noncovalent interaction through protein–protein interaction. This protocol utilizes biotinylated ASO of 20–25 DNA-based nucleotides to capture RNA of interest. The DNA probes are more stable than RNA probes and can be synthesized easily and cost-effectively. Both a negative control (luciferase) and a positive control (U1) are strongly recommended in the same run for proper normalization and ensuring the quality of iDRiP. Eluted proteins are analyzed by quantitative MS (using isobaric tags for relative and absolute quantitation)²⁹, followed by normalization with the negative control to obtain the enrichment of RNA-interacting proteins. The workflow of iDRiP-MS is shown in Fig. 1a.

Cell types and numbers

We suggest using cell lines or cell types that express high levels of the RNA target of interest. The abundance could be tested by qRT-PCR using 1–5 µg total RNA for cDNA synthesis. In general, when PCR efficiency is optimized, a mean Cq >30 for the RNA target would suggest a low-abundance RNA. We suggest choosing cell lines or cell types in which the RNA's Cq <26. The Cq of Xist is ~23 from mouse embryonic fibroblasts (MEFs). The Cq of PAR-TERRA, which is TERRA RNA derived from sex chromosomes, is 24–26 from mES cells. The Cq of U1 is ~21 from mES cells. This protocol has been shown to work for mES cells, mouse germ cells²⁶, immortalized MEFs and human U2OS cancer cells.

UV energy

UV is used to directly cross-link RNA to interacting proteins. The optimal UV energy is expected to vary slightly from cell line to cell line and RNA to RNA. When cells are grown in a monolayer, 200–400 mJ/cm² of 254 nm UV is recommended. Cells grown in multiple layers or tissues may require higher UV irradiation regimens 0.8–4.0 J/cm²⁴¹. Higher UV energy could damage RNA and reduce the number of accessible sites for probe targeting. For protein-enriched RNA complexes or RNAs with enriched secondary structures, substantially lower UV energy is recommended.

Modifications

There are some modifications of TERRA iDRiP compared with the original iDRiP that might be helpful for users investigating other RNAs with repetitive content and tight chromatin binding, as well as applying the method to different types of cells. First, UV cross-linking efficiency could differ depending on the sample. We recommend a slightly lower dose for cells grown in monolayer and a higher dose for multilayer tissues or cells prepared in suspension. For Xist-iDRiP, 200 mJ/cm² UV was used for MEFs grown in monolayer. For TERRA-iDRiP, 400 mJ/cm² UV was used for mES cells that are usually grown in a colony. Although ES cells are generally trypsinized prior to irradiation, cells may adhere together in suspension, which would require adjustments in irradiation parameters. Second, for TERRA iDRiP, we heated lysates to 65 °C prior to probe hybridization to increase solubility of chromatin-bound complexes, and insoluble cellular material was removed by centrifugation prior to RNA capture by probe, and then the lysate was pre-cleared using beads to reduce

the background noise. This is critical because insoluble material may adhere to beads and increase the background noise. When lysates were heated to 65 °C, TERRA recovery was increased twofold, as determined by RT-qPCR comparing heated with unheated samples. Third, the most critical step is the change in hybridization conditions for TERRA. We anticipated that the highly repetitive nature of TERRA and its high abundance of telomeric DNA could be an issue for probe specificity. Thus, whereas Xist capture hybridization was carried out at 37 °C overnight, TERRA was first denatured at 65 °C and then gradually cooled to 37 °C over 1–2 h to enhance capture specificity. We also avoided overnight incubations for TERRA to minimize background that could arise from longer incubations. Finally, we added an extra wash step for TERRA iDRiP, starting with a 300 mM LiCl wash (as for Xist iDRiP) and following with a lower-salt wash in 150 mM NaCl. For all capture experiments, we generally recommend using U1 as a positive control in side-by-side comparisons. A detailed protocol with RNA-specific modifications is provided in the Procedure.

Probe design

iDRiP utilizes 20–25 mer antisense oligos to capture a specific RNA. Tiling probes across the full-length RNA could be used to capture RNA as is done for RAP¹⁵, though we recommend against this approach owing to the likelihood of nonspecific RNA pulldown and decreasing the signal-to-noise ratio, resulting in pulldown of nonspecific interacting proteins. For example, we successfully captured a comprehensive interactome of >100 proteins for 17 kb Xist using nine oligo probes¹⁴. For U1, we only use one antisense oligo for iDRiP. In contrast, a tiling probe approach taken using RAP produced only a dozen abundant proteins¹⁵. Key to using fewer capture probes, however, is the ability to maintain an intact target RNA during the protocol. We recommend designing one probe per 500 nt and testing at least two sets of probes for capture efficiency before doing a large-scale iDRiP. For RNA >10 kb, one probe for 2–3 kb is recommended. We found the ideal probe length to be 20–25 nt. Optimal melting temperature for oligo probes is ~55–60 °C using the National Center for Biotechnology Information (NCBI) Primer-Blast tool. It is best to avoid T-rich sequences (>4 consecutive Ts) and GC-rich sequences (>5 continuous GCs). NCBI-Blastn should be applied to avoid high homology to other expressed sequences, which would be detrimental to specificity.

Quantitative proteomics to identify RNA-specific protein interactors

An accurate comparison of the amount of RNA co-isolated proteins with proteins in control samples is essential in identifying RNA-specific protein interactors. MS is the most important tool for unbiased quantitative mapping of proteins, and several MS methods are enabling such an analysis. The methods range from label-free quantification⁴² to the use of stable heavy isotopes for generating internal standards by either metabolically⁴³ or chemically⁴³ incorporating the stable isotopes into the proteins. An elegant way of using chemically incorporated stable isotopes for MS-based protein quantification is the use of isobaric reagents. These probes allow multiplexed quantifications through simultaneous quantification of multiple samples⁴⁴. Currently, up to 16 samples can be simultaneously quantified using this strategy⁴⁵, and further extensions of the multiplexing capacity are in development⁴⁶.

We have used multiplexed proteomics with isobaric labeling to identify RNA-specific proteins isolated using the iDRiP and ChIRP methods, but it should be noted that other quantitative proteomics methods are also suitable to do this analysis. In brief, we have used TMT10-plex reagents for barcoding the samples⁴⁷ and an MS3-based data acquisition method on an Orbitrap Fusion mass spectrometer (Thermo Fisher Scientific) for data acquisition. MS3-based data acquisition enables the generation of high-accuracy and high-precision quantitative data when using isobaric labels for quantitative proteomics^{30,31,48}. A full description of the proteomics sample preparation would go beyond the scope of this protocol, and we, therefore, point to two other protocol papers describing the steps from the protein digestions to the MS data acquisition in full detail^{29,49}.

Materials

Biological materials

- UV-cross-linked 1.5×10^8 cells: mES cells (*Mus musculus* × *Mus castaneus* hybrid ES cell line EL 16.7, female 40,XX⁵⁰, RRID: [CVCL_A6UI](#)) **! CAUTION** All cell lines should be regularly checked for mycoplasma contamination **▲ CRITICAL** We recommend using at least 10 million cells for iDRiP capture. The number of cells may be adjusted for different RNA targets.

Reagents

- Isopropanol (Sigma-Aldrich, cat. no. 278475)
- Ethanol (Sigma-Aldrich, cat. no. E7023)
- Protease inhibitors, cOmplete, EDTA-free protease inhibitor cocktail (Merck, cat. no. 4693132001)
- MyOne streptavidin C1 beads (Thermo Fisher Scientific, cat. no. 65001)
- RNase H (New England Biolabs, cat. no. M0297L)
- RNase A (Thermo Fisher Scientific, cat. no. 12091021)
- Proteinase K (Sigma-Aldrich, cat. no. 03115844001)
- SuperScript IV reverse transcriptase (Thermo Fisher Scientific, cat. no. 18090200)
- TRIzol (Thermo Fisher Scientific, cat. no. 15596018)
- Yeast RNA (Thermo Fisher Scientific, cat. no. AM7118)
- SUPERase•In RNase Inhibitor (Thermo Fisher Scientific, cat. no. AM2694)
- TURBO DNase (Thermo Fisher Scientific, cat. no. AM2239)
- Diethyl pyrocarbonate (DEPC; Sigma-Aldrich, cat. no. D5758)
- Dithiothreitol (DTT, BioShop, cat. no. DTT001)
- Phenylmethyl sulfonyl fluoride (PMSF, Sigma-Aldrich, cat. no.10837091001)
- Nonidet P-40 substitute (BioShop, cat. no. NON505)
- Dulbecco's phosphate-buffered saline (DPBS, Thermo Fisher Scientific, cat. no. 141902350)
- 1,4-Piperazinediethanesulfonic acid (PIPES; Sigma-Aldrich, cat. no. P6757)
- Sodium lauroyl sarcosine (Sigma-Aldrich, cat. no. L9150)
- Sodium deoxycholate (Sigma-Aldrich, cat. no. D6750)
- Ethylene glycol tetraacetic acid (EGTA; Sigma-Aldrich, cat. no. E4378)
- EDTA (Amresco, cat. no. 0105-1KG)
- LiCl (Sigma-Aldrich, cat. no. L9650)
- Triton X-100 (Acros Organics, cat. no. 9002-93-1)
- SDS (BioShop, cat. no. SDS001)
- Chloroform (Sigma-Aldrich, cat. no. C0549-1QT)
- Acid phenol/chloroform (Thermo Fisher Scientific, cat. no. AM9722)
- GlycoBlue (Thermo Fisher Scientific, cat. no. AM9516)
- 3' Biotin-TEG luciferase probe: 5'-GTATCCCTGGAAGATGGAAG-3'/3biotinTEG/ (Integrated DNA Technologies)
- 3' Biotin-TEG U1 probe sequence: 5'-CTCCCCTGCCAGGTAAGTAT-3'/3biotinTEG/ (Integrated DNA Technologies)
- 3' Biotin-TEG TERRA probe sequence: 5'-TAACCCTAACCCCTAACCCCTA-3'/3biotinTEG/ (Integrated DNA Technologies)
- 3' Biotin-TEG sense probe sequence: 5'-TTAGGGTTAGGGTTAGGGTT-3'/3biotinTEG/ (Integrated DNA Technologies)
- TMT10plex Isobaric Label Reagent Set (Thermo Fisher Scientific, cat. no. 90111). Alternatively, TMTduplex Isobaric Label Reagent Set (cat. no. 90065), TMTsixplex Isobaric Label Reagent Set (cat. no. 90061), TMT10plex Isobaric Label Reagent Set plus TMT11-131C Label Reagent (cat. no. A34808) or TMTpro 16plex Label Reagent Set (cat. no. A44521) can be used according to the number of samples

Equipment

- DNA LoBind microcentrifuge tubes, 1.5 ml, PCR clean, colorless (Eppendorf, cat. no. 022431021)
- Pipettes (Rainin, cat. no. 17014393/ 17014391/ 17014382, or equivalent)
- Pipette filter tip (SSIbio, cat. no. 4137NSF/4237NAF/4237NSF/4347NSF, or equivalent)
- 15 ml centrifuge tubes (Thermo Fisher Scientific, cat. no. NUC339650)
- Benchtop microcentrifuge (Thermo Fisher Scientific, Heraeus Fresco 21, cat. no. 75002555, or equivalent)
- Heat block, for 1.5 ml microcentrifuge tubes (VWR, cat. no. 12621-088, or equivalent)
- Water bath (YIHDER, model no. BU-420, or equivalent)
- Ice bucket (VWR, cat. no. 10146-290, or equivalent)
- Vortex mixer, Vortex-Genie 2 G560 (Scientific Industries, cat. no. SI-0236)
- DynaMag-2 magnetic separation stand, for 1.5 ml tubes (Thermo Fisher Scientific, cat. no. 12321D)
- Hybridization oven with a rotating wheel (Hybaid, Mini Oven MK II, cat. no. E4639, or equivalent, temperature could reach to 65 °C)
- Real-time PCR detection system (Bio-Rad, model no. 1855195, or equivalent)

- qPCR plates, 96-wel (Bio-Rad, model no. HSP9601 or equivalent suitable for real-time PCR detection system)
- Tandem mass spectrometer (Thermo Fisher Scientific, Orbitrap Fusion, Q-Exactive, Bruker Impact II or similar models)

Software

- NCBI Primer-Blast (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>)
- David Genome (<https://david.ncifcrf.gov/>)

Reagent setup

CSKT buffer

Component	Amount	Final concentration
PIPES	3.02 g	10 mM
NaCl	5.84 g	100 mM
MgCl ₂ •6H ₂ O	0.61 g	3 mM
Sucrose	102.7 g	0.3 M
Triton X-100	5 ml	0.5% (vol/vol)
DEPC-treated water	to 1 L	
Total	1 L	N/A

- Prepare 500 ml of DEPC-treated water in a bottle, and add the powders listed above to the solution.
- Adjust solution to pH 6.8–7.0 using 10 N NaOH
- Add DEPC-treated water to 1 L
- Add 5 ml of Triton X-100 to the solution
- Filter through a 0.2 µm sterile filter
- Store at 4 °C in dark bottles (or covered with foil) up to several months
- Freshly add 1× protease inhibitor cocktail before use

DNase I digestion buffer

Make up fresh as outlined below.

Component	Amount	Final concentration
1 M Tris pH7.5	500 µl	50 mM
10% of NP40	500 µl	0.5% (vol/vol)
10% sodium lauroyl sarcosinate	100 µl	0.1% (wt/vol)
DEPC-treated water	8.9 ml	
Add fresh:		
Protease inhibitor (EDTA-free)	1 tablet	1×
SUPERase•In (20 µl per 1 ml of lysate)	200 µl	1×
Turbo DNase I (100 µl per 1 ml of lysate)	1 ml	
Total	10 ml	N/A

0.5 M LiCl lysis buffer

Make up fresh as outlined below.

Component	Amount	Final concentration
1 M Tris pH 7.5	500 µl	50 mM
10% of NP40	500 µl	0.5% (vol/vol)
5 M LiCl	1 ml	0.5 M
Table continued		

(continued)		
Component	Amount	Final concentration
0.5 M EDTA	400 μ l	20 mM
0.5 M EGTA	400 μ l	20 mM
DEPC-treated water	6.1 ml	
Add fresh:		
10% sodium lauroyl sarcosinate	1 ml	1% (wt/vol)
10% sodium deoxycholate	100 μ l	0.1% (wt/vol)
Total	10 ml	N/A

2 \times binding buffer

Prepare as outlined below. Can be stored at room temperature (RT, 20–27°C) for months.

Component	Amount	Final concentration
1 M Tris-HCl (pH 7.5)	100 μ l	10 mM
500 mM EDTA	20 μ l	1 mM
5 M NaCl	4 ml	2 M
DEPC-treated water	5.9 ml	
Total	10 ml	N/A

Wash buffer 1

Make up fresh as outlined below.

Component	Amount	Final concentration
1 M Tris-HCl pH 7.5	2.5 ml	50 mM
5 M LiCl	3 ml	0.3 M
0.5 M EDTA	500 μ l	5 mM
10% SDS	5 ml	1% (wt/vol)
10% NP40	2.5 ml	0.5% (vol/vol)
DEPC-treated water	36.2 ml	
Add fresh:		
1 M DTT	50 μ l	1 mM
200 mM PMSF	250 μ l	1 mM
Protease inhibitors, EDTA-free	1 tablet	1 \times
Total	50 ml	N/A

Wash buffer 2

Make up fresh as outlined below.

Component	Amount	Final concentration
10% of SDS	1 ml	1% (wt/vol)
5 M NaCl	300 μ l	150 mM
0.5 M EDTA	100 μ l	5 mM
DEPC-treated water	8.6 ml	N/A
Add fresh:		
1 M DTT	10 μ l	1 mM
200 mM PMSF	50 μ l	1 mM
Total	10 ml	N/A

Elution buffer

Prepare as outlined below. Can be stored at RT for months.

Component	Amount	Final concentration
1 M Tris-HCl pH 7.5	0.1 ml	10 mM
0.5 M EDTA pH 8.0 (optional)	20 µl	1 mM
10% Triton X-100	50 µl	0.05% (vol/vol)
DEPC water	9.83 ml	
Total	10 ml	N/A

PK buffer

Prepare as outlined below. Can be stored at RT for months.

Component	Amount	Final concentration
5 M NaCl	200 µl	100 mM
1 M Tris-HCl pH 7.0	100 µl	10 mM
0.5 M EDTA pH 8.0	20 µl	1 mM
10% SDS	500 µl	0.5% (wt/vol)
DEPC-treated water	9.18 ml	
Total	10 ml	N/A

DEPC water

Component	Amount	Final concentration
Distilled water	1,000 ml	
DEPC	1 ml	
Total	1,000 ml	N/A

Autoclave distilled water first, then mix thoroughly with DEPC and leave it in chemical hood overnight. Autoclave again. Can be stored at RT for months.

Procedure

UV-cross-linking and preparation of the nuclei ● Timing ~2 h

- Starting with cells grown to 80% confluence in a 15 cm² plate, rinse cells with cold DPBS once and irradiate cells (with 5 ml of cold (0–8 °C) DPBS covering the 15 cm² plate) at 200–400 mJ/cm². For MEFs, we use 200 mJ/cm². For mES cells, we use 400 mJ/cm². Irradiate cells at 254 nm without a lid. **▲ CRITICAL STEP** Trypsinize cells when using ES cells, cells growing in spherical colonies or tissue before UV irradiation. For mES cells, trypsinize the cells grown in the 15 cm² plate, spin at 800g at RT for 5 min and replace the medium with 5 ml of cold PBS, and plate cells onto the plate for irradiation. Determine cell number before irradiation.
- Collect cells in a 15 ml centrifuge tube. Add 5 ml of cold DPBS to wash the plate, and collect the remaining cells from the plate. Spin at 800g for 3 min at 4 °C. Remove and discard the supernatant. Note: when preparing a large-scale experiment, pool 1.5 × 10⁸ cells in a 50 ml tube for CSKT treatment. **▲ CRITICAL STEP** If RNA of interest is in the cytoplasm, omit the CSKT treatment for nuclear isolation (Step 3).
- Resuspend the pellet in 10 ml of ice-cold CSKT (for 1.5 × 10⁸ cell pellet) buffer with protease inhibitors for nuclear extraction, and keep it on the rocker for 3–10 min at 4 °C. Spin at 800g at 4 °C for 3 min. Remove and discard the supernatant. Note: add more CSKT as needed to accommodate the number of cells.

▲ CRITICAL STEP This step is to reduce background noise from the cytoplasm. The incubation time of CSKT treatment is dependent on the cell type. We recommend 3 min for mES cells, and 10 min for MEFs. For testing the time of CSKT treatment, the efficiency of nuclei extraction could be checked by looking at cell morphology under a light microscope. Place some CSKT-treated cells onto a slide to check the efficiency. The CSKT treatment should be stopped when the plasma membrane is disrupted but the nuclear membrane remains intact.

- Resuspend the pellet (for $1.5\text{--}5 \times 10^7$ cells) with 1 ml of cold DPBS, and transfer to 1.5 ml tube. Spin at 800g at 4 °C for 3 min.
- Aspirate all the liquid on top of the pellet. Flash freeze cell pellets in liquid nitrogen, and store at -80 °C until use.

■ PAUSE POINT Cells could be stored at -80 °C for several months.

? TROUBLESHOOTING

Preparation of lysate ● Timing ~1.5 h

- Thaw nuclear pellets on ice. For 1.5×10^8 cells, resuspend the nuclear pellets in 3 ml of DNase I digestion buffer supplemented with 300 μ l of Turbo DNaseI enzyme (2 U/ μ l), 60 μ l SUPERase•In and 1 \times protease inhibitors (the total volume will reach to 3.5 ml). Separate the 3.5 ml lysate equally into 4 \times 1.5 ml microtubes. Mix well, and incubate at 37 °C for 20 min in a water bath. Note: longer incubations may help the solubility but may also increase the degradation of RNA and proteins. Pipette the lysates and invert microtubes several times during the incubation. Be careful to avoid water contamination.

▲ CRITICAL STEP This step is required for capturing nuclear RNAs. If DNase does not work well, the background noise will increase. To check the efficiency of DNase treatment, DNA can be purified using cell lysates from Step 7 by phenol/chloroform extraction. DNA size after DNase treatment should be ~100–200 bp, the size of mononucleosomes (Extended Data Fig. 1d).

▲ CRITICAL STEP Cell (nucleus) clumping is usually a problem. Separating 3.5 ml lysate into 4 \times 1.5 ml microtubes will avoid cell clumping during this step and increase the efficiency of DNase treatment. Incomplete digestion of genomic DNA may result in large RNA–chromatin complexes that could be nonspecifically pulled down by beads. The background noise such as abundant nuclear RNAs could be detected after iDRiP by RT-qPCR.

▲ CRITICAL This step is optional for cytosolic RNAs.

? TROUBLESHOOTING

- Combine 3.5 ml lysate from Step 6 with the reagents listed below in a 15 ml tube. Mix well, vortex and continue incubation at 37 °C for 5 min in the water bath, then put lysate on ice. For 3.5 ml of initial lysate, combine reagents as follows to reach the final concentrations listed in the table below (the total volume will reach 5 ml).

Component	Amount	Final concentration
0.5 M EGTA	200 μ l	20 mM
0.5 M EDTA	200 μ l	20 mM
5 M LiCl	500 μ l	0.5 M
10% sodium lauroyl sarcosinate	470 μ l	1%
10% sodium deoxycholate	50 μ l	0.1%
DEPC-treated water	70 μ l	
Total	1,490 μ l	

- After mixing, separate the 5 ml of lysate into five microtubes (1.5 ml). (Note: at this point, the lysate may still have small clumps. It is okay to continue to the following steps.) Spin down at highest speed (15,000g) at 4 °C for 2 min. Save the supernatant on ice.
- Resuspend the pellet in each tube in 600 μ l of 0.5 M LiCl lysis buffer (Note: 600 μ l are used for the pellet from 1 ml lysate. Total volume will be 600 μ l \times 5 tubes = 3 ml), leave on ice for 10 min and then incubate at 65°C for 5 min (mix frequently by inverting the tubes) in a heat block. Immediately spin down (15,000g at RT, 1 min). Save the supernatant, and leave it on ice for several minutes. Combine with the previous supernatant (from Step 8). The final volume of lysate after this step will reach ~8 ml.

▲ CRITICAL STEP Insoluble cellular material is removed by these two centrifugation steps prior to RNA capture by probes. Insoluble material may adhere to beads and increase the background noise.

We also found that heating the insoluble materials from the pellet to 65 °C can increase the solubility of chromatin-bound RNA–protein complexes.

! CAUTION When collecting supernatant, it is important to avoid collecting the pellet (insoluble materials), otherwise the background noise will increase.

- 10 Collect 1% (vol/vol) input. (It can vary from 1% to 10%). Freeze samples at –80 °C.
- 11 Use 3–5 ml of lysate per capture for the following steps. Separate lysate into several 1.5 ml microtubes. Perform the hybridization in 1.5 ml microtubes.

▲ CRITICAL STEP To test the capture efficiency of iDRiP, a small-scale (300 µl lysate) iDRiP could be done prior to a large-scale experiment (3 ml lysate).

Preclear lysate ● Timing ~0.5 h

- 12 Wash 50 µl Dynabeads (MyOne streptavidin C1 beads) with 1 ml of DEPC-treated water twice in a 1.5 ml microtube by placing on a magnet to collect the beads. Suspend beads in an equal volume of 0.5 M LiCl lysis buffer (i.e., take 50 µl beads from the original vial and resuspend 50 µl of lysis buffer).
- 13 Add 50 µl of beads (50 µl for 3 ml of lysate, ~16.3 µl beads for 1 ml of lysate) for precleaning, and incubate beads with lysates at RT for 20 min with gentle rotation. Collect beads on a magnet, transfer the lysate to new tubes and discard beads.

Conjugate oligo probes to beads ● Timing ~0.5 h

▲ CRITICAL DNA Oligos should be 20–25 mers. See ‘Reagents’ for validated sequences.

- 14 Wash with 300 µl of Dynabeads with 1 ml of DEPC water twice.
- 15 Resuspend beads in 300 µl of 1× binding buffer.
- 16 Add 150 pmoles of probes (the total amount) to beads, and incubate at RT for 20 min with gentle rotation.
- 17 Wash beads with 1 ml of 1× binding buffer twice.
- 18 Resuspend beads in 300 µl of 0.5 M LiCl lysis buffer.

Hybridization ● Timing ~3 h

▲ CRITICAL One of the most critical factors for a successful iDRiP experiment is the hybridization condition used to capture the ncRNA of interest. We anticipated that the highly repetitive nature of TERRA and its high abundance in mES cells could be an issue for probe specificity. Thus, whereas Xist capture hybridization is carried out at 37 °C overnight, TERRA is first denatured at 65 °C and then gradually cooled to 37 °C over 1–2 h to enhance capture specificity.

▲ CRITICAL Hybridization of capture probe to RNA target is done in a hybridization chamber/oven (used for Southern or northern blotting), where the temperature can be gradually decreased from 65 °C to 37 °C over 30–60 min. This aids in optimizing the hybridization temperature of different RNA-targeting probes between 65 °C and 37 °C.

- 19 Prewarm lysates (from Step 13) and 300 µl of probe-conjugated beads (from Step 18) to 65 °C for 5 min. Mix lysates with the beads, and immediately incubate the mixture at 65 °C (~100 µl beads for 1 ml of lysate in a 1.5 ml microtube).

▲ CRITICAL STEP It is important to bring up the temperatures of probes and lysates individually before hybridization to prevent nonspecific binding at a lower temperature.

- 20 Incubate at 65 °C for 15 min, then reduce the temperature slowly to 37 °C and continue the hybridization for 1 h in the hybridization oven.

▲ CRITICAL STEP It is important to reduce the temperature to 37 °C slowly to obtain the optimal hybridization efficiency. Suddenly dropping the temperature is not recommended. The easy way to slowly reduce the temperature is to set the oven temperature to 37 °C (the original temperature is 65 °C) and keep the hybridization oven closed.

! CAUTION Do not let the hybridization temperature exceed 65 °C; higher temperatures will disrupt the biotin–streptavidin interaction.

? TROUBLESHOOTING

Washing ● Timing ~1.5 h

- 21 Prewarm wash buffer 1 (see ‘Reagent setup’) to 37 °C.
- 22 Capture the beads at 37 °C on the magnet. Remove the supernatant containing unbound materials, and add wash buffer 1 to resuspend the beads. Each wash constitutes 1 ml wash buffer/1.5 ml tube under rotation (on wheel, or 3 ml wash buffer for 3 ml of lysate in 15 ml tube) each for 5 min at 37 °C

- 23 Wash beads three times.
- 24 Rinse beads once at RT in DNase I buffer (plus 0.2 M LiCl) (1 ml DNase I buffer/1.5 ml tube) to clean out the SDS and EDTA and EGTA, which can inhibit the DNase activity. Resuspend the beads in 300 μ l DNase I buffer (for 3 ml of lysate) with 0.2 M LiCl + 10 μ l TURBO DNase + 4 μ l SUPERase•In + protease inhibitor. Incubate at 37 °C for 10 min in the water bath.
- 25 Add wash buffer 1, and mix briefly before magnetic separation.
- 26 Perform two more wash (total five washes), each for 5 min at 37 °C.
- 27 Perform an additional wash with 1 ml of 150 mM of NaCl wash buffer/1.5 ml tube (wash buffer 2) at 37 °C for 5 min.
▲ CRITICAL STEP The lower-salt washing step (150 mM NaCl) is important for improving the signal-to-noise ratio. However, it may reduce the capture efficiency if incubation at 37 °C exceeds 5 min.

Elution ● Timing ~0.5 h

- ▲ CRITICAL** We have tried several methods for elution, such as excess biotin, 95 °C formamide and 70 °C low-salt solution. The RNA recovery rates did not change. We found that 0.05% (vol/vol) Triton X-100 is required for the low-salt solution method at 70 °C to efficiently elute complexes from beads. Other detergents such as 0.05% (vol/vol) NP40 may also work.
- 28 Resuspend the beads in 600 μ l elution buffer (for 3 ml of original lysate), and incubate the sample at 70 °C for 5 min, then immediately place it on the magnet (~200 μ l elution buffer is used for 1 ml of lysate).
 - 29 Collect the supernatant in a separate tube. Put the supernatant on the magnet to remove residual beads by transferring supernatant again to a new tube. Keep the elution on ice.
 - 30 For every MS experiment, quantification of the RNA of interest is necessary. Set aside 10% (vol/vol) elute for quantification of RNA pulldown by RT-qPCR.
 - 31 For submission of samples for MS, snap freeze the elution and store at –80 °C.
 - 32 For RNA extraction, take 2–10% (vol/vol) of elution (usually 6–12 μ l is enough), and add 90 μ l of PK buffer (with 5 μ l of 20 mg/ml proteinase K in 100 μ l buffer) for 15 min at 55 °C (preheat PK buffer at 55 °C for 10 min in advance). Then, add 1 ml of Trizol. Process the input sample from Step 10 (use between 0.2% and 2% of input; usually 6 μ l is enough) the same way by treating with proteinase K followed by Trizol extraction.
■ PAUSE POINT The eluted proteins and RNA (in Trizol) can be stored at –80 °C for several months.

Trizol RNA extraction ● Timing ~1 d

- 33 Thaw out RNA in Trizol, and pipette the solution several times at RT. Add 200 μ l of chloroform, and shake tubes by hand. Incubate samples at RT for 5 min.
! CAUTION Please conduct the procedures using Trizol and chloroform under a chemical hood.
- 34 Spin at 14,000g for 15 min at 4 °C.
- 35 Take the upper (aqueous) phase, and transfer to a new 1.5 ml tube.
- 36 Add 500 μ l of acid phenol/chloroform to the aqueous phase, shake and incubate at RT for 5 min.
- 37 Spin down at 14,000g for 10 min at 4 °C.
- 38 Transfer the upper phase to a new 1.5 ml tube.
- 39 Add 3 μ l of glycoBlue (15 mg/ml). Mix well.
- 40 Add 0.5 mL of isopropyl alcohol, and mix well to precipitate RNA. Incubate samples at –20 °C overnight. (Note: RNA can be stored in isopropyl alcohol at –20 °C for several months.)
- 41 Spin at 14,000g for 15 min at 4 °C.
- 42 Wash with 1 ml of 75% (vol/vol) EtOH.
- 43 Spin at 14,000g for 5 min. Carefully remove supernatant.
- 44 Air dry for 5–10 min.
- 45 Dissolve RNA in 30 μ l of DEPC-treated water, and immediately put RNA samples on ice.

cDNA synthesis

- 46 Prepare RNA-primer mix by combining the following components with each RNA sample in PCR reaction tubes. Add an equal volume of RNA from each sample for cDNA synthesis.

Component	Volume
50 μ M random hexamers	1 μ l
10 mM dNTP mix (10 mM each)	1 μ l
RNA sample (10 pg to 5 μ g total RNA)	5 μ l (or up to 11 μ l)
DEPC-treated or nuclease-free water	to 13 μ l

- 47 Mix and briefly centrifuge the components.
- 48 Heat the RNA-primer mix at 65 °C for 5 min, and keep on ice for at least 1 min.
- 49 Prepare reverse transcription reaction mix with the following components in a reaction tube (per RNA sample).

Component	Volume
5 \times SSIV buffer	4 μ l
100 mM DTT	1 μ l
RNaseOUT recombinant RNase inhibitor	1 μ l
SuperScript IV reverse transcriptase (200 U/ μ l)	1 μ l

- 50 Mix well and briefly centrifuge the contents.
- 51 Add the reverse transcription reaction mix to the annealed RNA sample from Step 49.
- 52 Incubate the mixture at the following temperatures using a PCR machine:

Temperature	Time
25 °C	10 min
50 °C	50 min
85 °C	5 min

- 53 (Optional) Add 1 μ l *E. coli* RNase H to remove RNA, and incubate at 37 °C for 20 min.
- 54 The samples are ready for RT-qPCR immediately, or they can be stored at -20 °C.

Calculating enrichment of the target RNA by RT-qPCR

- 55 Before the RT-qPCR experiment, prepare tenfold serial dilutions from the input sample. The serial dilutions of cDNA samples are used for generating a standard curve for calculating the concentrations of RNA-captured samples.
- 56 To define the RNA capture efficiency of iDRiP by RT-qPCR experiment, we use ‘percentage of the input’ as the unit of measurement. This analysis represents the amount of captured RNA relative to the amount of starting material from the input sample.
- 57 The following table is an example for setting up an RT-qPCR standard curve. The concentration of captured RNA can be calculated using the CFX Manager for qPCR data analysis provided by Bio-Rad. Both the target RNA and the reference RNA need a standard curve. The reference RNA such as U1, U6 or GAPDH can be used. Set up qPCR reactions according to the manufacturer’s manual.

Sample	Cq	Starting quantity (percent of input)
100% input		100
10% input		10
1% input		1
0.1% input		0.1
0.01% input		0.01
0.001% input		0.001

- 58 Calculate capture efficiency, enrichment of the target RNA and signal-to-noise ratio using the formulas listed below:

The capture efficiency (presented as % of input) = the amount of captured RNA/total starting RNA.
The enrichment of target RNA (specificity) = the amount of target RNA after target probe capture (% of input)/the amount of target RNA after control probe capture (% of input).

The fold enrichment of target RNA compared with the reference RNA (signal-to-noise ratio) = the amount of target RNA after capture (% of input)/the amount of reference RNA after capture (% of input).

Troubleshooting

Troubleshooting advice is provided in Table 2.

Step	Problem	Possible reason	Solution
1	Protein amount is low after capture	The RNA of interest is low in abundance	Some RNAs are expressed at low levels. As a result, the amount of captured proteins could be very low too, and it may be difficult to see specific protein bands in an SDS gel by silver staining. As long as some specific protein bands can be seen after the positive control U1-iDRiP capture (Extended Data Fig. 1e), you can go ahead with MS. Quantitative MS is necessary to eliminate the false positive candidates from abundant proteins. Including a positive and negative control in the same experiment is necessary to ensure the quality of iDRiP-MS. Increase total material (cell number) for iDRiP. Use cells that express the RNA of interest at higher levels, or overexpress the target RNA.
6	DNase treatment is not efficient	Cross-linking is not sufficient Cell aggregation and clumping	Increase the UV cross-linking dose. Use a pipette to resuspend cells frequently during the incubation or increase the time for DNase treatment to 30–45 min.
6–32	RNA quality is low. This could be detected by RT-qPCR after capture	RNA is degraded	Each buffer should be prepared using DEPC-treated water. Wear a mask if needed to avoid RNase contamination of samples.
20	Signal-to-noise ratio is low	Hybridization temperature does not drop to 37 °C gradually, or hybridization oven temperature is higher than 65 °C. High temperature (>80 °C) may disrupt the biotin-streptavidin interaction	Adjust the temperature to 37 °C without opening the oven, and do not open the oven during hybridization, which may cause temperature fluctuation.
55–58	Capture efficiency is low	Target sequences are inaccessible to probes because they are masked by proteins	Redesign probe sequences. We recommend designing at least two sets of probes to test capture efficiency at a small scale prior to performing the large-scale prep.

Timing

- Steps 1–5, UV cross-linking: 2 h
- Steps 6–11, preparation of lysate and first DNase treatment: 1.5 h
- Steps 12–18, preclean lysate and conjugate oligo probes to beads: 1 h
- Steps 19–20, hybridization: 3 h
- Steps 21–27, washing and second DNase treatment: 1.5 h
- Steps 28–32, elution: 0.5 h
- Steps 33–45, RNA extraction: 1.5 d
- Steps 46–54, cDNA synthesis: 1.5 h
- Steps 55–58, calculating enrichment of the target RNA by RT-qPCR: 3 h

Anticipated results

Calculating enrichment of the target RNA

After iDRiP capture, quantitative RT-PCR is used to check the recovery and specificity. Above 10% of target RNA can be recovered compared with the input. The signal-to-noise ratio, which compares the target RNA to other abundant RNAs such as U1 RNA or GAPDH, is very important. We usually obtain at least 50-fold enrichment (Fig. 1b). For the specificity of capture, the relative enrichment of the target RNA is calculated by comparing the amount of target RNA between the specific capture and the negative control captures. As shown in Fig. 1c, TERRA iDRiP yielded an over 1,000-fold enrichment compared with Luc-iDRiP capture. The high specificity is critical because proteins captured by the control probe are used for the normalization of the specific protein enrichment after quantitative MS.

The interaction of RNA and protein can be further validated by UV-RIP¹⁴ or immuno-RNA-FISH staining¹³ (Fig. 4a,b).

Statistical analysis

The intensity of each protein analyzed by quantitative MS (TMT) is considered as protein abundance after capture. After normalization with luciferase control, the log₂ value of each protein abundance (target/control) represents the enrichment over control. Probability density using log₂ values was plotted by fitting the data to a normal distribution (Extended Data Fig. 4a,b). We provide two different methods to obtain *P* values. (1) if only one data set is available, *P* values could be calculated by the fitting curve method. *P* value and FDR were calculated in R 3.6.2 for U1-ChIRP and U1-iDRiP (Supplementary Tables 5 and 6). A histogram of probability density was plotted using *hist(x = abundance ratio, probability = TRUE)*, and datapoints on the histogram were extracted using the *density()* function and were fitted to a normal distribution by the nonlinear least squares method using the *nls()* function in R. The Gaussian formula used in this step was $f(x) = a \cdot e^{-\frac{(x-b)^2}{2c^2}}$, and the parameters *a*, *b* and *c* were determined by fitting the datapoints when finding a model. *P* values were calculated using the original log₂ abundance ratio as *x*, and the FDR was obtained using *P* values adjusted by Benjamini–Hochberg method. (2) If there are biological replicates, *P* values can be calculated using the rank product method⁵¹. RankProd 2.0⁵², an R package, was applied to U1-iDRiP and U1-ChIRP analysis (Extended Data Fig. 4c,d, Supplementary Table 7, 8)

Data availability

All mass spectrometer RAW files in Supplementary Tables 1–4 can be accessed through the MassIVE data repository (massive.ucsd.edu) under the accession number [MSV000086552](https://massive.ucsd.edu/MSV000086552).

References

1. The ENCODE Project Consortium. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* **447**, 799–816 (2007).
2. Kung, J. T., Colognori, D. & Lee, J. T. Long noncoding RNAs: past, present, and future. *Genetics* **193**, 651–669 (2013).
3. Kapranov, P. et al. The majority of total nuclear-encoded non-ribosomal RNA in a human cell is ‘dark matter’ un-annotated RNA. *BMC Biol.* **8**, 149 (2010).
4. Rinn, J. L. & Chang, H. Y. Genome regulation by long noncoding RNAs. *Annu. Rev. Biochem.* **81**, 145–166 (2012).
5. Lee, J. T. & Bartolomei, M. S. X-inactivation, imprinting, and long noncoding RNAs in health and disease. *Cell* **152**, 1308–1323 (2013).
6. Tay, Y., Rinn, J. & Pandolfi, P. P. The multilayered complexity of ceRNA crosstalk and competition. *Nature* **505**, 344–352 (2014).
7. Anderson, K. M. et al. Transcription of the non-coding RNA upperhand controls Hand2 expression and heart development. *Nature* **539**, 433–436 (2016).
8. Kotzin, J. J. et al. The long non-coding RNA Morrbid regulates Bim and short-lived myeloid cell lifespan. *Nature* **537**, 239–243 (2016).
9. Chillon, I. & Marcia, M. The molecular structure of long non-coding RNAs: emerging patterns and functional implications. *Crit. Rev. Biochem. Mol. Biol.* **55**, 662–690 (2020).
10. Uroda, T. et al. Conserved pseudoknots in lncRNA MEG3 are essential for stimulation of the p53 pathway. *Mol. Cell* **75**, 982–995 e9 (2019).
11. Uroda, T. et al. Visualizing the functional 3D shape and topography of long noncoding RNAs by single-particle atomic force microscopy and in-solution hydrodynamic techniques. *Nat. Protoc.* **15**, 2107–2139 (2020).

12. Kim, D. N. et al. Zinc-finger protein CNBP alters the 3-D structure of lncRNA Braveheart in solution. *Nat. Commun.* **11**, 148 (2020).
13. Chu, H. P. et al. TERRA RNA antagonizes ATRX and protects telomeres. *Cell* **170**, 86–101 e16 (2017).
14. Minajigi, A. et al. Chromosomes. A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. *Science* **349** (2015).
15. McHugh, C. A. et al. The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature* **521**, 232–236 (2015).
16. Chu, C. et al. Systematic discovery of Xist RNA binding proteins. *Cell* **161**, 404–416 (2015).
17. Simon, M. D. et al. High-resolution Xist binding maps reveal two-step spreading during X-chromosome inactivation. *Nature* **504**, 465–469 (2013).
18. Simon, M. D. Capture hybridization analysis of RNA targets (CHART). *Curr. Protoc. Mol. Biol.* **Chapter 21**, Unit 21 25 (2013).
19. Aeby, E. et al. Decapping enzyme 1A breaks X-chromosome symmetry by controlling Tsix elongation and RNA turnover. *Nat. Cell Biol.* **22**, 1116–1129 (2020).
20. Yi, W. et al. CRISPR-assisted detection of RNA-protein interactions in living cells. *Nat. Methods* **17**, 685–688 (2020).
21. Distèche, C. M. Dosage compensation of the sex chromosomes. *Annu. Rev. Genet.* **46**, 537–560 (2012).
22. Azzalin, C. M., Reichenbach, P., Khoriauli, L., Giulotto, E. & Lingner, J. Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. *Science* **318**, 798–801 (2007).
23. Azzalin, C. M. & Lingner, J. Telomere functions grounding on TERRA firma. *Trends Cell Biol.* **25**, 29–36 (2015).
24. Zhang, L.-F. et al. Telomeric RNAs mark sex chromosomes in stem cells. *Genetics* **182**, 685 (2009).
25. Schoeftner, S. & Blasco, M. A. Developmentally regulated transcription of mammalian telomeres by DNA-dependent RNA polymerase II. *Nat. Cell Biol.* **10**, 228–236 (2008).
26. Brieno-Enriquez, M. A., Moak, S. L., Abud-Flores, A. & Cohen, P. E. Characterization of telomeric repeat-containing RNA (TERRA) localization and protein interactions in primordial germ cells of the mousedagger. *Biol. Reprod.* **100**, 950–962 (2019).
27. Sunwoo, H., Wu, J. Y. & Lee, J. T. The Xist RNA-PRC2 complex at 20-nm resolution reveals a low Xist stoichiometry and suggests a hit-and-run mechanism in mouse cells. *Proc. Natl Acad. Sci. USA* **112**, E4216–E4225 (2015).
28. Ule, J. et al. CLIP identifies Nova-regulated RNA networks in the brain. *Science* **302**, 1212–1215 (2003).
29. Edwards, A. & Haas, W. Multiplexed quantitative proteomics for high-throughput comprehensive proteome comparisons of human cell lines. *Methods Mol. Biol.* **1394**, 1–13 (2016).
30. Ting, L., Rad, R., Gygi, S. P. & Haas, W. MS3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics. *Nat. Methods* **8**, 937–940 (2011).
31. McAlister, G. C. et al. MultiNotch MS3 enables accurate, sensitive, and multiplexed detection of differential expression across cancer cell line proteomes. *Anal. Chem.* **86**, 7150–7158 (2014).
32. Huang, D. W. et al. Extracting biological meaning from large gene lists with DAVID. *Curr. Protoc. Bioinformatics*. <https://doi.org/10.1002/0471250953.bi1311s27> (2009).
33. Chang, F. T. et al. PML bodies provide an important platform for the maintenance of telomeric chromatin integrity in embryonic stem cells. *Nucleic Acids Res.* **41**, 4447–4458 (2013).
34. Seraphin, B. & Rosbash, M. Identification of functional U1 snRNA-pre-mRNA complexes committed to spliceosome assembly and splicing. *Cell* **59**, 349–358 (1989).
35. Pomeranz Krummel, D. A., Oubridge, C., Leung, A. K., Li, J. & Nagai, K. Crystal structure of human spliceosomal U1 snRNP at 5.5 Å resolution. *Nature* **458**, 475–480 (2009).
36. Konarska, M. M. & Sharp, P. A. Association of U2, U4, U5, and U6 small nuclear ribonucleoproteins in a spliceosome-type complex in absence of precursor RNA. *Proc. Natl Acad. Sci. USA* **85**, 5459–5462 (1988).
37. Hall, K. B. & Konarska, M. M. The 5' splice site consensus RNA oligonucleotide induces assembly of U2/U4/U5/U6 small nuclear ribonucleoprotein complexes. *Proc. Natl Acad. Sci. USA* **89**, 10969–10973 (1992).
38. Shibata, S. & Lee, J. T. Characterization and quantitation of differential Tsix transcripts: implications for Tsix function. *Hum. Mol. Genet.* **12**, 125–136 (2003).
39. Sarma, K. et al. ATRX directs binding of PRC2 to Xist RNA and Polycomb targets. *Cell* **159**, 869–883 (2014).
40. Jegu, T. et al. Xist RNA antagonizes the SWI/SNF chromatin remodeler BRG1 on the inactive X chromosome. *Nat. Struct. Mol. Biol.* **26**, 96–109 (2019).
41. Sysoev, V. O. et al. Global changes of the RNA-bound proteome during the maternal-to-zygotic transition in *Drosophila*. *Nat. Commun.* **7**, 12128 (2016).
42. Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat. Protoc.* **11**, 2301–2319 (2016).
43. Ong, S. E. et al. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteom.* **1**, 376–386 (2002).
44. Rauniyar, N. & Yates, J. R. Isobaric labeling-based relative quantification in shotgun proteomics. *J. Proteome Res.* **13**, 5293–5309 (2014).
45. Li, J. M. et al. TMTpro reagents: a set of isobaric labeling mass tags enables simultaneous proteome-wide measurements across 16 samples. *Nat. Methods* **17**, 399–404 (2020).
46. Braun, C. R. et al. Generation of multiple reporter ions from a single isobaric reagent increases multiplexing capacity for quantitative proteomics. *Anal. Chem.* **87**, 9855–9863 (2015).

47. McAlister, G. C. et al. Increasing the multiplexing capacity of TMTs using reporter ion isotopologues with isobaric masses. *Anal. Chem.* **84**, 7469–7478 (2012).
48. Lapek, J. D. et al. Detection of dysregulated protein-association networks by high-throughput proteomics predicts cancer vulnerabilities. *Nat. Biotechnol.* **35**, 983–989 (2017).
49. Kreuzer, J., Edwards, A. & Haas, W. Multiplexed quantitative phosphoproteomics of cell line and tissue samples. *Post-Transl. Modif. That Modulate Enzym. Act.* **626**, 41–65 (2019).
50. Lee, J. T. & Lu, N. Targeted mutagenesis of Tsix leads to nonrandom X inactivation. *Cell* **99**, 47–57 (1999).
51. Breitling, R., Armengaud, P., Amtmann, A. & Herzyk, P. Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. *FEBS Lett.* **573**, 83–92 (2004).
52. Del Carratore, F. et al. RankProd 2.0: a refactored bioconductor package for detecting differentially expressed features in molecular profiling datasets. *Bioinformatics* **33**, 2774–2775 (2017).
53. Flynn, R.A. et al. Systematic discovery and functional interrogation of SARS-CoV-2 viral RNA-host protein interactions during infection. Preprint at *bioRxiv* <https://doi.org/10.1101/2020.10.06.327445> (2020).
54. Munschauer, M. et al. The NORAD lncRNA assembles a topoisomerase complex critical for genome stability. *Nature* **561**, 132–136 (2018).
55. Schmidt, N. et al. The SARS-CoV-2 RNA–protein interactome in infected human cells. *Nat. Microbiol.* **6**, 339–353 (2021).
56. West, J. A. et al. The long noncoding RNAs NEAT1 and MALAT1 bind active chromatin sites. *Mol. Cell* **55**, 791–802 (2014).

Acknowledgements

This work was supported by grants from the National Institutes of Health (R37-GM58839 and RO1-HD097665) and the Howard Hughes Medical Institute to J.T.L., and by a Ministry of Science and Technology grant, MOST 108-2628-B-002-050, to H.-P.C.

Author contributions

H.-P.C. and J.T.L. conceived the project and wrote the manuscript; H.-P.C. and A.M. designed the iDRiP protocol. H.-P.C. performed iDRiP-MS, ChIRP-MS and staining experiments. Y.C. conducted the UV-RiP experiments. R.M., C.-Y.G., Y.-H.H., R.M., M.B., W.H. and J.T.L. analyzed and interpreted proteomic data. All authors contributed extensively to the work presented and prepared the manuscript.

Competing interests

J.T.L. is a scientific adviser to Translate Bio and Fulcrum Therapeutics as well as an adviser to Skyhawk Therapeutics.

Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s41596-021-00555-9>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41596-021-00555-9>.

Correspondence and requests for materials should be addressed to H.-P.C. or J.T.L.

Peer review information *Nature Protocols* thanks Philipp G. Maass, Claire Rougeulle and the other, anonymous reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 9 October 2020; Accepted: 13 April 2021;

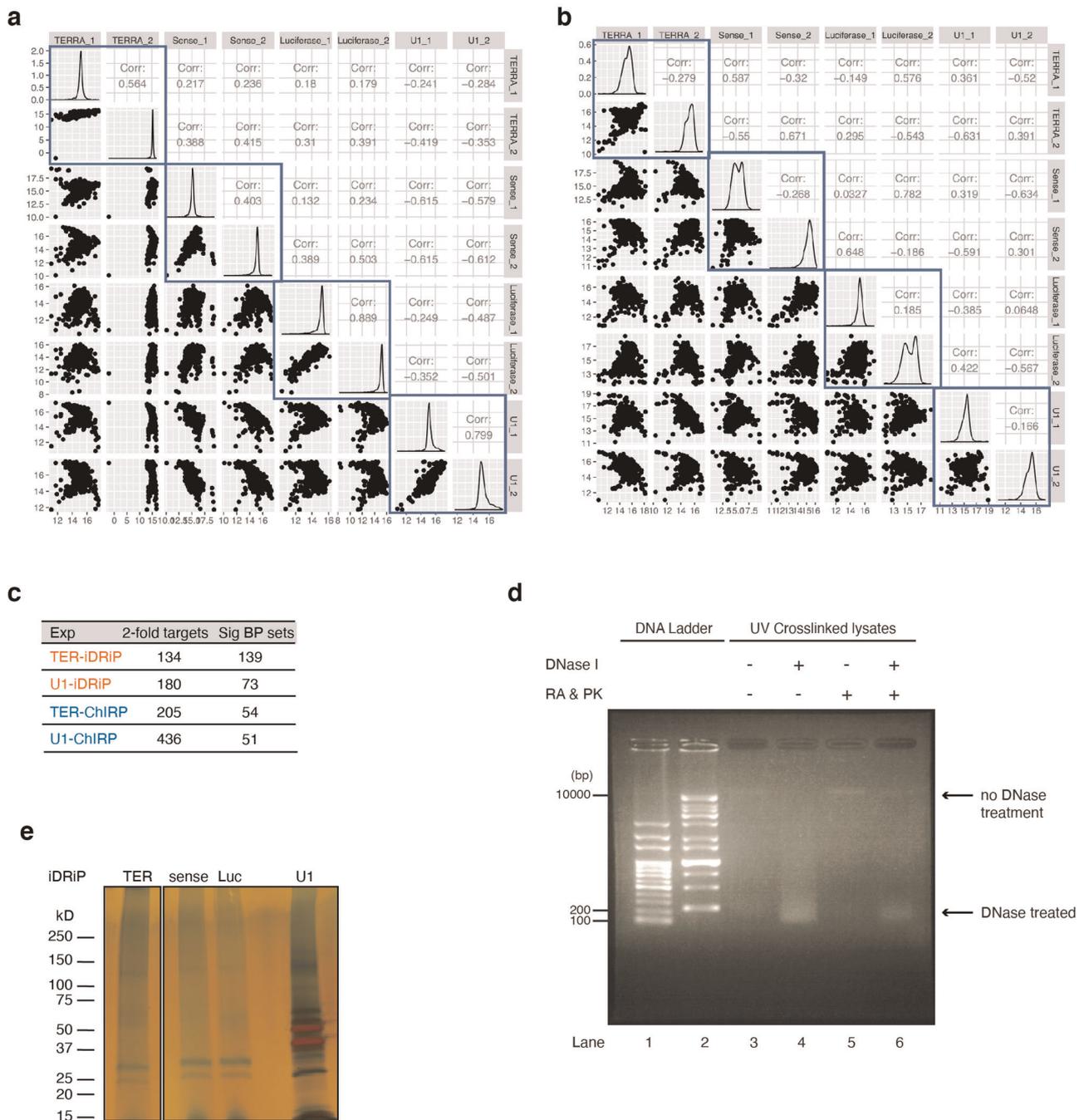
Published online: 9 June 2021

Related links

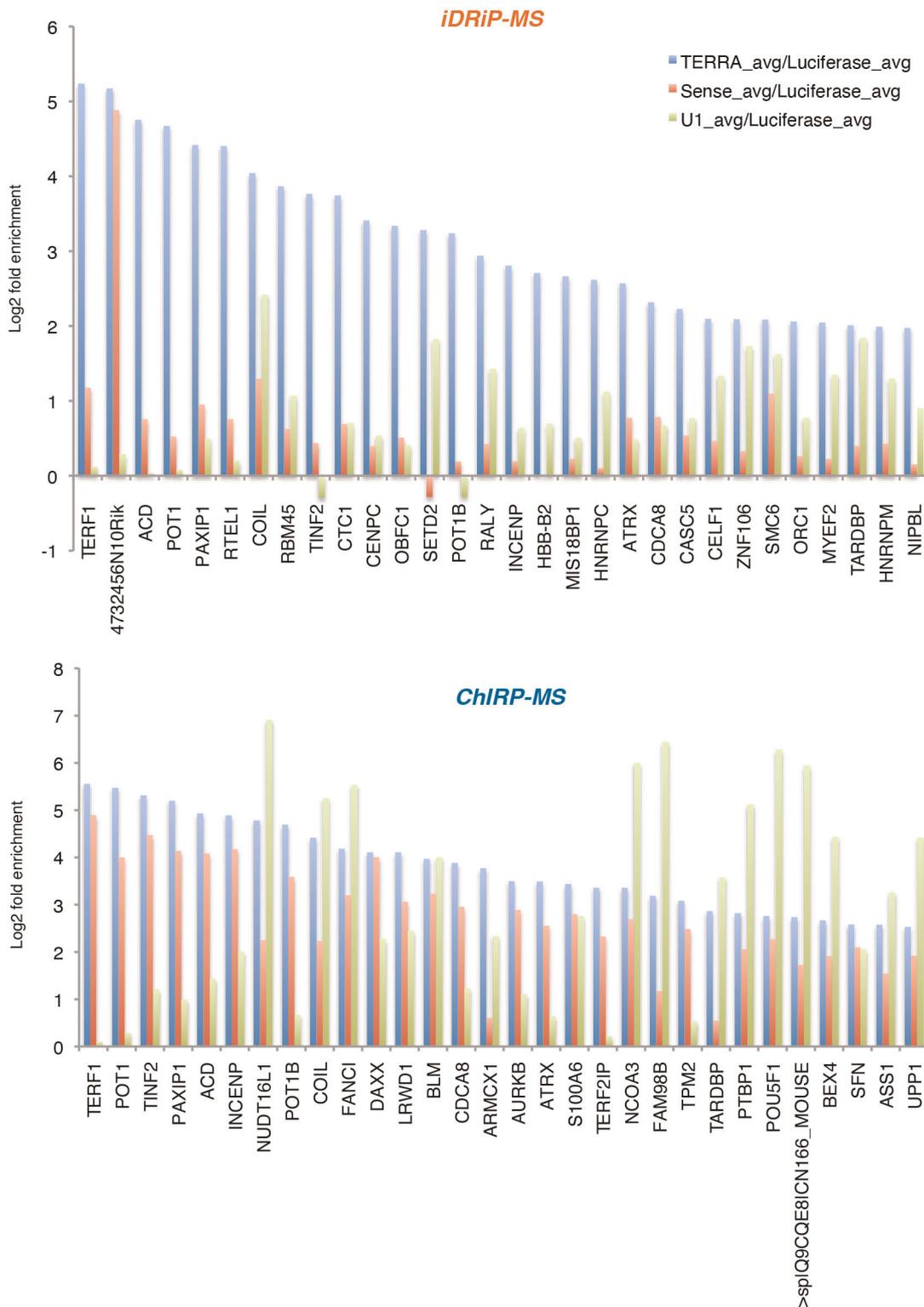
Key references using this protocol

Minajigi, A. et al. *Science* **349**, aab2276 (2015): <https://doi.org/10.1126/science.aab2276>

Chu, H., et al. *Cell* **170**, 8–9 (2017): <https://doi.org/10.1016/j.cell.2017.06.017>



Extended Data Fig. 1 | Comparison of the results obtained with the iDRiP and ChIRP methods. a, Scatterplots comparing various iDRiP-MS datasets. Each dot represents the protein abundance (log₂ intensity) obtained from quantitative MS. Replicate 1 versus replicate 2 shown in blue boxes. Pearson correlation *r* is indicated. **b**, Scatterplots comparing various ChIRP-MS datasets. Each dot represents the protein abundance (log₂ intensity) obtained from quantitative MS. Replicate 1 versus replicate 2 shown in blue boxes. Pearson correlation *r* is indicated. **c**, The number of twofold enriched targets (over luciferase control) and the number of significant biological processes (BP) sets from DAVID bioinformatics resources with FDR < 25% are indicated in each experiment. **d**, Agarose gel of DNase-treated iDRiP samples. UV-cross-linked samples were treated with or without DNase. DNA was purified by phenol/chloroform extraction. Lane 1, 100 bp marker. Lane 2, 1 kb marker. Lane 3, no DNase treatment. Lane 4, DNase treatment for 20 min. Lane 5, no DNase treatment (+RNase A, +proteinase K). Lane 6, DNase treatment (+RNase A, +proteinase K). **e**, Silver staining of the eluted proteins after iDRiP (using various probes: TERRA, sense, Luc and U1) in a polyacrylamide gel. The eluted proteins from U1-iDRiP show some specific bands compared with the negative control iDRiP (Luc).



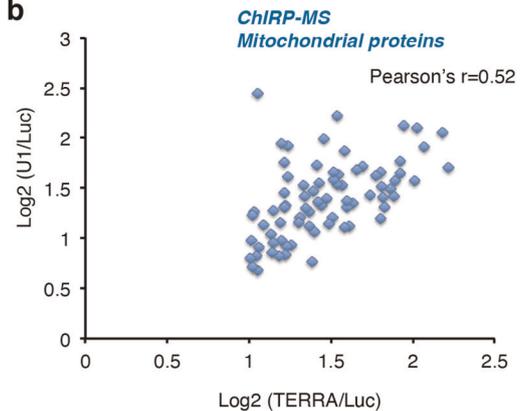
Extended Data Fig. 2 | Top TERRA interacting candidates revealed by iDRiP or ChIRP. iDRiP proteomics (top panel) shows less background in controls (sense and U1) than ChIRP proteomics (bottom panel). Log2 enrichment of proteins captured by iDRiP or ChIRP (using various probes: TERRA, sense and U1) normalized using luciferase as a control.

a

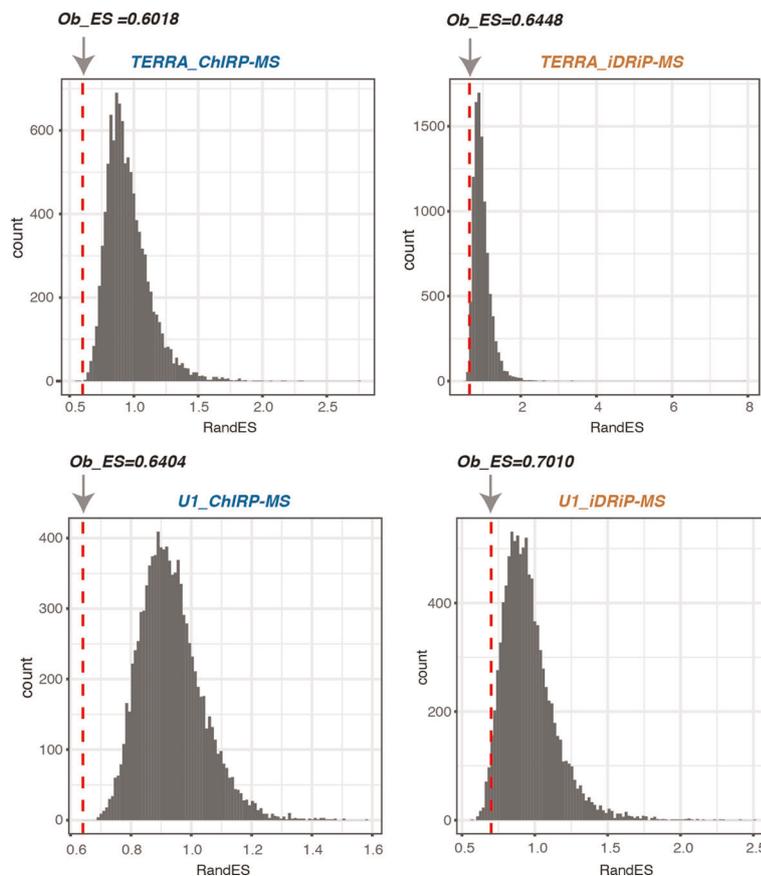
U1_ChIRP-MS

GO_BP_Term	Genes	Count	%	P-Value	Benjamini
mRNA processing		128	30.0	3.4E-130	4.6E-127
RNA splicing		107	25.1	8.0E-114	5.3E-111
mRNA splicing, via spliceosome		47	11.0	5.1E-47	2.3E-44
spliceosomal snRNP assembly		15	3.5	3.4E-17	1.1E-14
regulation of alternative mRNA splicing, via spliceosome		17	4.0	4.6E-17	1.2E-14
negative regulation of mRNA splicing, via spliceosome		14	3.3	8.5E-16	2.0E-13
RNA processing		20	4.7	1.0E-15	1.9E-13
mRNA transport		17	4.0	6.9E-11	1.1E-8
spliceosomal complex assembly		10	2.3	7.9E-11	1.2E-8
transcription, DNA-templated		85	19.9	1.3E-10	1.7E-8

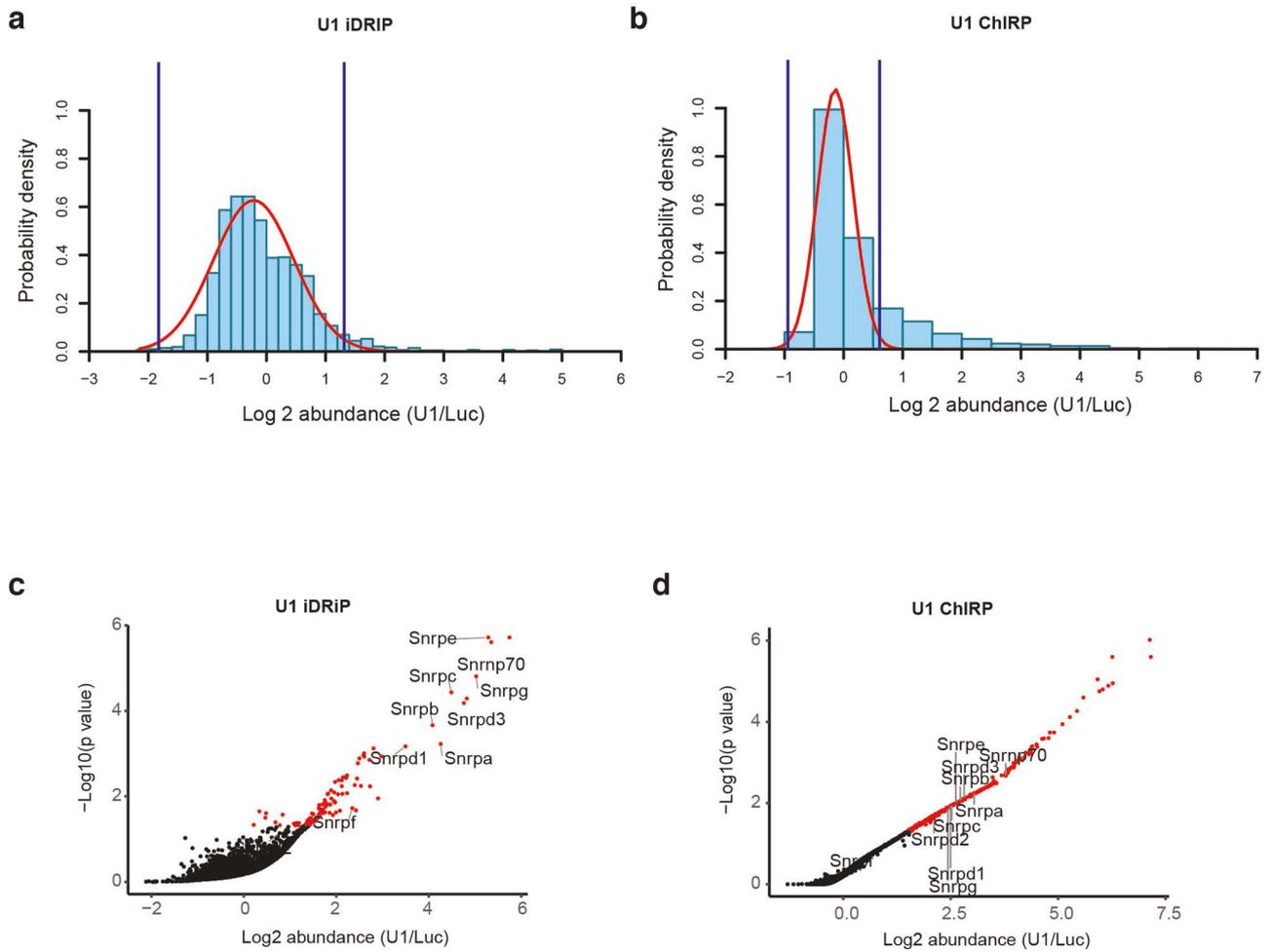
b



c



Extended Data Fig. 3 | U1 protein interactome identified by ChIRP-MS. **a**, GO biological process (GO_BP_term) enrichment analysis for U1-ChIRP interactome. **b**, Scatterplot showing log₂ enrichment of mitochondrial proteins found by TERRA-ChIRP versus U1-ChIRP. Pearson's r is shown. **c**, The ranked list of expression data for all genes from RNA-seq data¹³ was sorted from high expression to low expression. The target genes positions in the ranked list (FPKM value) are used to calculate an enrichment score for the entire set of targets (observed data). To see if the observed value is random noise, the position of gene labels was randomly ordered relative to the ranked FPKM list 10,000 times. The observed enrichment score (Ob_ES) is labeled with a red dotted line. The histogram is the distribution of randomized enrichment scores (RandES). The results show that in general the protein targets from iDRiP and ChIRP are clustered significantly near the low end of the transcriptome (1- P -value < 0.05).



Extended Data Fig. 4 | Statistics of U1-iDRiP and U1-ChIRP proteomics. **a**, Histogram of the probability density of the log₂ abundance for U1-iDRiP-MS. Red line, fitted using the Gaussian formula. Blue lines, cutoffs for $P < 0.05$. **b**, Histogram of the probability density of the log₂ abundance for U1-ChIRP-MS. Red line, fitted using the Gaussian formula. Blue lines, cutoffs for $P < 0.05$. **c**, Scatter plot of log₂ abundance versus $-\log_{10}(P \text{ value})$ for U1-iDRiP-MS. U1 snRNP genes are indicated. Red dots, genes with $P < 0.05$. **d**, Scatterplot of log₂ abundance versus $-\log_{10}(P \text{ value})$ for U1-ChIRP-MS. U1 snRNP genes are indicated. Red dots, genes with $P < 0.05$.