



HAL
open science

A Modified Cross-Linking Analysis of cDNAs (CRAC) Protocol for Detecting RNA–Protein Interactions and Transcription at Single-Nucleotide Resolution

Drice Challal, Jessie Colin, Tommaso Villa, Domenico Libri

► **To cite this version:**

Drice Challal, Jessie Colin, Tommaso Villa, Domenico Libri. A Modified Cross-Linking Analysis of cDNAs (CRAC) Protocol for Detecting RNA–Protein Interactions and Transcription at Single-Nucleotide Resolution. *Yeast Functional Genomics*, 2477, Springer US, pp.35 - 55, 2022, *Methods in Molecular Biology*, 10.1007/978-1-0716-2257-5_3. hal-03824813

HAL Id: hal-03824813

<https://hal.science/hal-03824813v1>

Submitted on 21 Oct 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Modified Cross-Linking Analysis of cDNAs (CRAC) Protocol for Detecting RNA–Protein Interactions and Transcription at Single-Nucleotide Resolution 2 3 4

Drice Challal, Jessie Colin, Tommaso Villa, and Domenico Libri 5

Abstract 6

Detecting protein–RNA interactions in vivo is essential for deciphering many important cellular pathways. Several methods have been described for this purpose, among which cross-linking analysis of cDNA, CRAC. This method relies on a first step of UV cross-linking of living yeast cells and several subsequent steps of purification of the protein–RNA complexes, some of which under denaturing condition. Without altering the general principle of the method, we have modified and improved the protocol, with the specific aim of sequencing the nascent RNA isolated from transcription complexes and generate high-resolution and directional transcription maps. 7 8 9 10 11 12 13

Keywords RNA–protein interaction in vivo, Genome-wide RNAPII mapping, Transcription, CRAC (cross-linking analysis of cDNAs) 14 15 16

1 Introduction 17

Because of differences in the stability of RNA molecules, the steady state abundance of transcripts is often a poor indicator of transcriptional activity. Many RNAs that are not, or barely, detected in wild type cells derive from genes that are transcribed with comparable intensity to genes producing abundant steady state transcripts. More accurate measures of transcription levels are generally based on the detection of RNA polymerase associated to its template, presumably during RNA production. Detection of RNA polymerase by chromatin immunoprecipitation (ChIP) is a method that does not provide a directional output and lacks resolution. Several approaches have been developed based on the sequencing of the nascent RNA that provide higher resolution mapping data. One 18 AU1 19 20 21 22 23 24 25 26 27 28 29

Drice Challal, Jessie Colin and Tommaso Villa contributed equally with all other contributors.

such method, nascent transcript sequencing (NETseq) [1] relies on the purification of the RNA polymerase II (RNAPII) transcription elongation complex under native condition and the sequencing of the associated RNA. Mapping of the 3'-OH of recovered nonpolyadenylated reads is a generally reliable indicator of RNAPII position. However, native purification conditions are not very efficient in eliminating nonnascent or contaminating RNAs of various sources. Indeed, cotranscriptional processing RNA intermediates associated with the elongation complex are also detected by this method and can be misinterpreted as indicators of RNAPII genomic position. For instance, the presence of reads derived from the 5'-exon splicing intermediate generates a strong signal immediately before the 5'-splice site [1] (see also below), which can be mis-interpreted as a sign of RNAPII pausing.

A similar method, RNAPII PAR-CLIP (photoactivable ribonucleoside-enhanced cross-linking and precipitation), requires the previous incorporation of a photoactivable ribonucleoside analogue (4-thio-uracil, 4tU) into the nascent RNA that is thereby efficiently cross-linked to the Rpb2 subunit of RNAPII. Thanks to the formation of a covalent RNA-protein adduct, purification steps can be conducted under denaturing conditions, which limits copurification of associated, nonnascent RNA molecules. The extent to which the addition of the toxic 4tU [2] can alter the physiology of the cell and the transcription process has not been fully investigated. Other methods also rely on the incorporation of 4tU in RNA and the purification of the newly transcribed molecule [3, 4]; however, the extent to which 4tU-containing RNAs are nascent (and therefore signal the position of RNAPII) as opposed to post-transcriptional is debatable, particularly so in the compact and largely intron-less genome of *S. cerevisiae*. Indeed, the length of many genes is below the number of nucleotides that RNAPII can transcribe in the timeframe of the labeling pulse. CRAC [5] is analogous to PAR-CLIP in that the protein of interest is cross-linked to the bound RNA by exposure of the cells to short wavelength UV light for a short time lapse in the absence of 4tU. Cross-linking is likely less efficient in this condition than when photoactivable nucleotide analogs are used, but there is less concern about the impact of the latter on transcription. As with PAR-CLIP, the subsequent purification includes denaturing steps, which decreases the amount of associated but not cross-linked RNAs. The position of the cross-link can also be identified by this method because transcribing through the cross-link site during the reverse transcription step induces microdeletions or misincorporation in the cDNA that can be easily identified and mapped.

Here, we describe a modified protocol (Fig. 1) that we have extensively used to generate high resolution and directional transcription maps [6, 7] but also to analyze genome-wide the association of RNA binding proteins with their RNA targets [8–10]. To

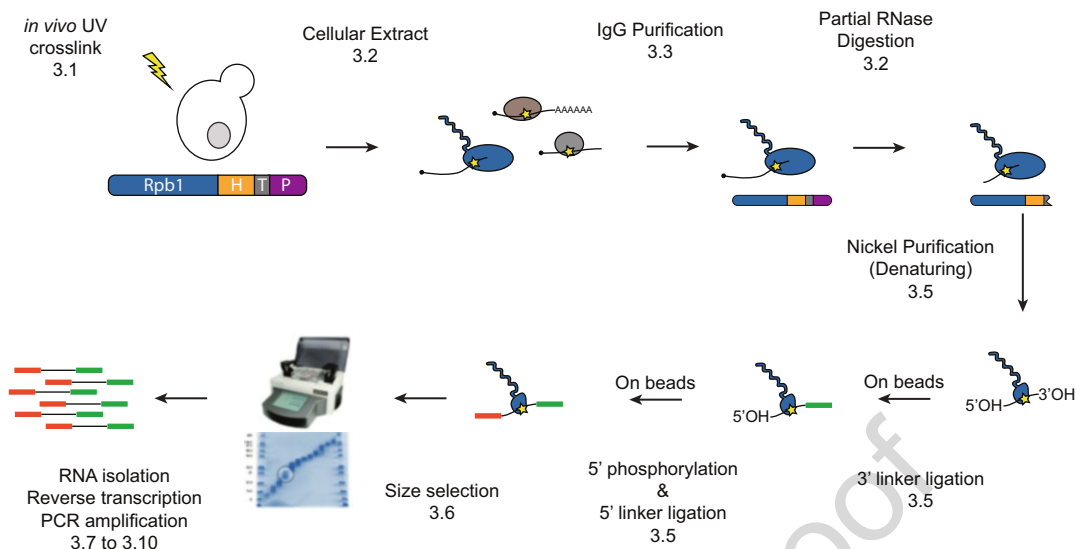
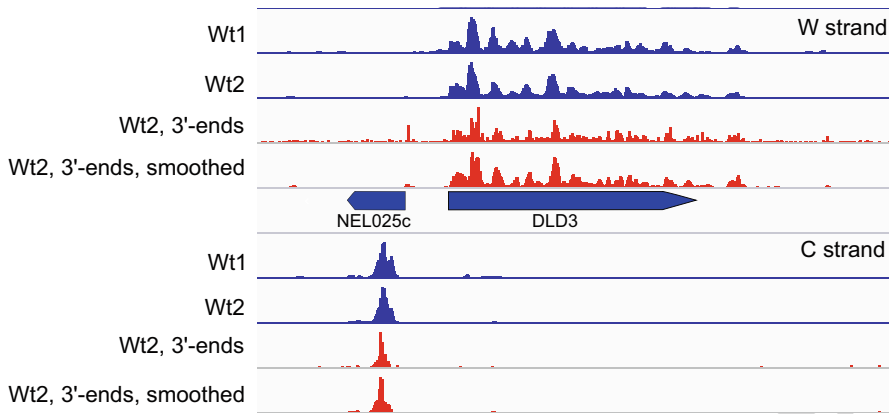


Fig. 1 Overview of the CRAC procedure. Living cells are UV cross-linked and the protein of interest (Rpb1 for transcription analysis) is purified under native conditions. A limited RNase digestion allows reduction of the size of the cross-linked RNA fragment. A second purification under denaturing conditions is performed and linkers are added to the cross-linked RNA while the complex is still attached to the beads, which allows extensive washing between each step. After elution, the RNA–protein complex is further subjected to a denaturing, size-selection purification using the GELFREE 8100 system. The RNA from the complex is finally isolated, reverse-transcribed, and the library prepared by PCR amplification is sequenced using Illumina technology. Each step of the procedure is numbered (Subheadings 3.1–3.10) and detailed in the main text

improve detection of the 3'-end of each read (defining the last 78
 position of RNAPII) we modified the original adapters so that 79
 sequencing occurs from the 3' end. Because some proteins bind 80
 polyadenylated reads, we added an RNase H/oligo-dT degradation 81
 step to limit the size of poly(A) tails that might be present on a 82
 subset of reads. Polyadenylated reads can still be detected by 83
 sequencing as oligo-dT directed RNase H cleavage cannot fully 84
 degrade the poly(A) tail. We also modified the purification protocol 85
 to improve yields, and eliminated the original radioactive labeling 86
 of the protein-bound RNA. Our protocol provides a very robust 87
 and high-resolution detection of RNAPII distribution (Fig. 2a) 88
 which is further improved by mapping only the 3'-end of each 89
 read, eventually with a smoothing treatment of the signal 90
 (Fig. 2a). Metagene analyses obtained using RNAPII CRAC are 91
 similar to the ones obtained using NET-seq (Fig. 2b). Differences 92
 are observed in the detection of the 5'-exon splicing intermediate 93
 (data not shown) and also at the very position of poly(A) tail 94
 addition (Fig. 2b) most likely due to the detection in NET-seq of 95
 a 3'-end cleavage intermediate that has not yet been polyadenylated 96
 and remains associated to the transcription elongation complex. 97
 Differences with RNAPII PAR-CLIP data are more marked as the 98

A



B

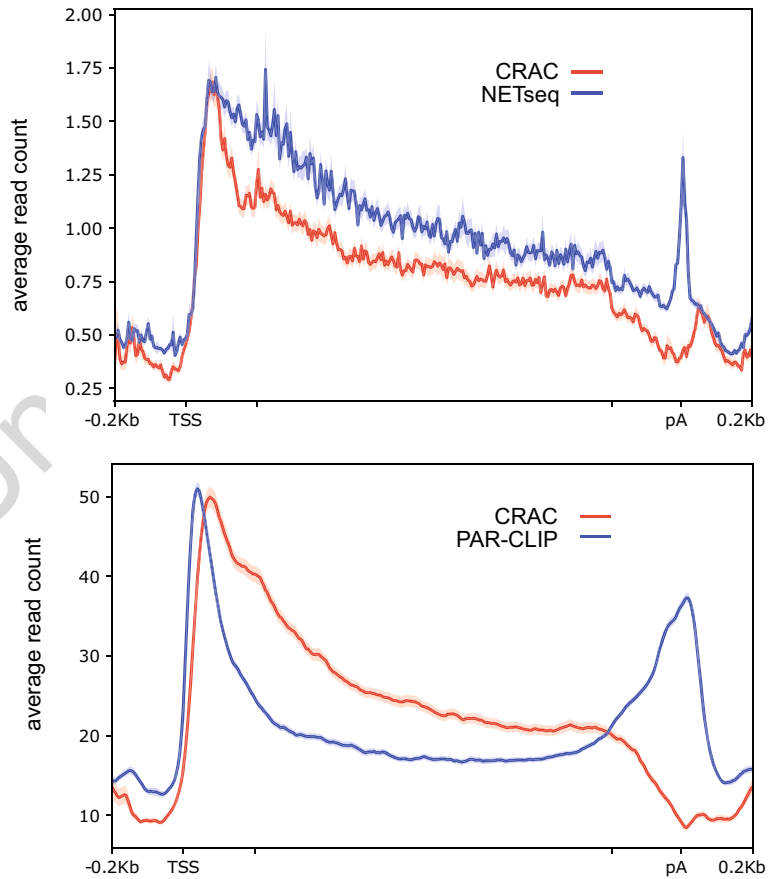


Fig. 2 (a) Snapshots showing the distribution of the RNAPII CRAC signal at the *DLD3-NEL025C* (SUT503-CUT542) locus. Wt1 and Wt2 tracks (blue) represent two independent experiments with wild type cells

latter method reveals high signals in the 3'-end of many genes (Fig. 2b). The reasons for this difference between NET-seq and CRAC on one side and PAR-CLIP on the other side is unclear. The similarity of NET-seq and CRAC signals tends to exclude the occurrence of strong RNAPII accumulation in the 3'-end of genes, but some peculiarities of PAR-CLIP RNAPII might reveal events that escape detection by the other methods.

For RNAPII CRAC, we mainly used cells expressing a C-terminally tagged Rpb1, but very similar results have been obtained with C-terminally tagged Rpb2. Interestingly, tagged Rpb3 did not cross-link efficiently to the nascent RNA in our conditions, which testifies to the specificity of the method.

2 Materials

Solutions are prepared with autoclaved deionized water and can be stored at 4 °C. Protease inhibitors and beta-mercaptoethanol are freshly added to solutions before usage. RNase-free plastic equipment and reagents are used.

2.1 Yeast Strain and Media

1. Yeast strain expressing HTP (6×His, TEV cleavage site, Protein A)-tagged protein of interest (Rpb1 or Rpb2 for RNAPII CRAC).
2. CSM-TRP media.

2.2 Reagents and Enzymes

1. Complete EDTA-free protease inhibitor cocktail tablets.
2. Pefabloc SC-Protease-Inhibitor [AEBSF].
3. DNase I recombinant, RNase-free.
4. Dynabeads M-280 Tosylactivated.
5. IgG from rabbit serum.
6. RNase H.
7. Recombinant GST-TEV protease.
8. RNase-IT Ribonuclease Cocktail.

Fig. 2 (continued) performed by different operators. In red, only the 3'-end of reads from Wt2 are shown, as raw and smoothed signals as indicated. **(b)** Metagenome analysis to compare RNAPII signals on all yeast genes for CRAC and NET-seq (top) or CRAC and PAR-CLIP (bottom). In the top panel only the 3'-ends of the CRAC reads have been plotted for homogeneity with NET-seq data. Note that the peak observed at the 3' of genes in the NET-seq profile is constituted by reads ending precisely at the polyadenylation site, and presumably represent molecules that have been cleaved but not yet polyadenylated and are associated at the RNAPII elongation complex. This is distinct from the diffuse peak observed in the PAR-CLIP profile in a similar position. The gray shadowed areas represent nonscaled regions of 200 nt around the anchoring point. NET-seq data from ref. [1]. PAR-CLIP data from ref. [15]

9. Guanidine hydrochloride.	128
10. Imidazole.	129
11. 0.3 M NaCl.	130
12. Ni-NTA Agarose.	131
13. Spin columns—Snap Cap.	132
14. RNaseOUT Recombinant Ribonuclease Inhibitor.	133
15. Quick CIP.	134
16. T4 RNA Ligase 2, truncated KQ.	135
17. T4 polynucleotide kinase.	136
18. T4 RNA Ligase 1 (ssRNA ligase).	137
19. Vivacon 500.	138
20. SilverQuest™ Silver Staining Kit.	139
21. Amersham™ Quick Stain Protein Labeling Kit.	140
22. Rabbit Peroxidase Anti-Peroxidase Soluble Complex.	141
23. Mouse monoclonal anti-polyHistidine (HIS-1).	142
24. Proteinase K, recombinant, PCR grade.	143
25. Ethanol.	144
26. 3 M NaOAc pH 5.5.	145
27. Phenol–chloroform–isoamyl alcohol 25:24:1, Saturated with 10 mM Tris–HCl, pH 8.0, 1 mM EDTA.	146
28. Chloroform.	148
29. 5PRIME Phase Lock Gel Heavy, 2 mL tubes.	149
30. GlycoBlue Coprecipitant.	150
31. SuperScript IV Reverse Transcriptase.	151
32. RNase H.	152
33. LightCycler FastStart DNA Master SYBR Green I.	153
34. LightCycler capillaries.	154
35. LA Taq.	155
36. Low Molecular Weight DNA Ladder.	156
37. Exonuclease I.	157
38. Gelfree 8100 5% Tris Acetate Cartridge Kit.	158
39. NucleoSpin Gel and PCR cleanup.	159
40. Qubit dsDNA HS Assay Kit.	160
2.3 Buffers	
1. TN150 buffer: 50 mM Tris–HCl pH 7.8, 150 mM NaCl, 0.1% NP-40, 5 mM beta-mercaptoethanol.	161
2. TN1000 buffer: 50 mM Tris–HCl pH 7.8, 1 M NaCl, 0.1% NP-40, 5 mM beta-mercaptoethanol.	162
	163
	164

	3. Wash buffer I: 6.0 M guanidine hydrochloride, 50 mM Tris-HCl pH 7.8, 1 M NaCl, 0.1% NP-40, 10 mM imidazole, 5 mM beta-mercaptoethanol.	165 166 167
	4. Wash buffer I-1 M: 6.0 M guanidine hydrochloride, 50 mM Tris-HCl pH 7.8, 1 M NaCl, 0.1% NP-40, 10 mM imidazole, 5 mM beta-mercaptoethanol.	168 169 170
	5. 1× PNK buffer: 50 mM Tris-HCl pH 7.8, 10 mM MgCl ₂ , 0.1% NP-40, 5 mM beta-mercaptoethanol.	171 172
	6. 5× PNK buffer: 250 mM Tris-HCl pH 7.8, 50 mM MgCl ₂ , 50 mM beta-mercaptoethanol.	173 174
	7. Wash buffer II: 50 mM Tris-HCl pH 7.8, 50 mM NaCl, 0.1% NP-40, 10 mM imidazole, 5 mM beta-mercaptoethanol.	175 176
	8. Elution buffer: 50 mM Tris-HCl pH 7.8, 50 mM NaCl, 0.1% NP-40, 150 mM imidazole, 5 mM beta-mercaptoethanol.	177 178
	9. Commercial ultrapure PCR-grade H ₂ O (Roche).	179
2.4 Equipment	1. Megatron W5 UV cross-linking unit (UVO3 Ltd).	180
	2. Mixer Mill MM 400 (Retsch).	181
	3. Bioruptor Standard (Diagenode).	182
	4. GELFREE 8100 Fractionation Station (Expedeon).	183
	5. Qubit Fluorometer (Thermo Fisher Scientific, Invitrogen).	184
	6. LightCycler 2.0 (Roche).	185
2.5 Oligonucleotides	Oligonucleotides were designed to sequence from the 3'-end of RNA. 3' linkers are 5'-riboadenylated (rApp), allowing for efficient and specific ligation to RNA substrates by T4 RNA ligase 2 in the absence of ATP, and are 3'-blocked (ddC). A three nucleotides barcode (underlined, lowercase) is included for multiplexing and six random nucleotides are used for eliminating PCR duplicates. The 5' linker contains an inverted dideoxythymidine at its 5'-end to prevent self-ligation and protect from degradation. For the reverse transcriptase step, the original primer described in Bohnsack et al. [5] was modified as we noted a possible recognition of endogenous rRNA.	186 187 188 189 190 191 192 193 194 195 196 197
2.5.1 3' Linkers	L3-6N-gag /5rApp/GCT <u>ctc</u> NNNNNNAGATCGGAA GAGCGTCGTGTAGGGAAAGAGTGT/3ddC/ L3-6N-acu /5rApp/GCT <u>agt</u> NNNNNNAGATCGGAA GAGCGTCGTGTAGGGAAAGAGTGT/3ddC L3-6N-ugc /5rApp/GCT <u>gca</u> NNNNNNAGATCGGAA GAGCGTCGTGTAGGGAAAGAGTGT/3ddC/ L3-6N-cua /5rApp/GCT <u>tag</u> NNNNNNAGATCGGAA GAGCGTCGTGTAGGGAAAGAGTGT/3ddC/ L3-6N-cgu /5rApp/GCT <u>acg</u> NNNNNNAGATCGGAA GAGCGTCGTGTAGGGAAAGAGTGT/3ddC/	198 199 200 201 202 203 204 205 206 207

	L3-6N-guc /5rApp/GCTgac NNNNNNAGATCGGAA	208
	GAGCGTCGTGTAGGGAAAGAGTGT/3ddC/	209
	L3-6N-uca /5rApp/GCTtga NNNNNNAGATCGGAA	210
	GAGCGTCGTGTAGGGAAAGAGTGT/3ddC/	211
	L3-6N-agu /5rApp/GCTact NNNNNNAGATCGGAA	212
	GAGCGTCGTGTAGGGAAAGAGTGT/3ddC/	213
		214
2.5.2 5' Linker	L5miRCat /5InvddT/	215
	CTTGrGrCrArCrCrCrGrArGrArArUrUrCrCrA-3.	216
		217
2.5.3 RT and PCR	RT L3-2 5-ACACTCTTTCCCTACACGACGCTCTTCCG-3	218
Oligonucleotides	P5_3'	219
	5- AATGATACGGCGACCACCGAGATCTA	220
	CACTCTTTCCCTACACGACGCTCTTCCGATCT-3	221
	miRCat_PCR2 5- CAAGCAGAAGACGGCATAACGA	222
	gatcCTTGGCACCCGAGAAT-3	223
		224
2.6 Software and Algorithms	The cracalyze pipeline is used for demultiplexing, mapping and normalizing reads using <i>S. pombe</i> spike-in. The PeakCcall suite [6] is used for extracting 5' or 3' read positions, mismatches and deletion, for smoothing the data and for peak calling.	225 226 227 228

3 Methods 229

3.1 Yeast Growth and UV Cross-Linking	1. Grow HTP-tagged strains in 2 L of CSM-TRP medium (<i>see</i> Notes 1 and 2) to OD ₆₀₀ = 0.6 at 30 °C.	230 231
	2. UV-irradiate cells at 254 nm for 50 s (<i>see</i> Note 3) using the Megatron W5 UV cross-linking unit.	232 233
	3. Harvest cells by centrifugation in 1 L bottles, 5 min at ~9000 × <i>g</i> at 4 °C.	234 235
	4. Wash 2 L pellets in 40 mL cold PBS and transfer in 50 mL conical tube, spin down cells in tabletop centrifuge for 5 min at 3200 × <i>g</i> at 4 °C, and weigh pellets.	236 237 238
3.2 Preparation of Cell Lysate	1. Resuspend pellets in 2.4 mL/g of cells of TN150 buffer supplemented with cComplete EDTA-free Protease Inhibitor Cocktail and 2 mM AEBSF (<i>see</i> Note 4).	239 240 241
	2. Flash-freeze the suspension in droplets. Cell “pop corn” can be indefinitely stored at –80 °C.	242AU2 243
	3. Mechanically break cells through 5 cycles of 3 min at 20 Hz in a Mixer Mill MM400 and store resulting cell powders at –80 °C until use (<i>see</i> Note 5).	244 245 246
	4. Thaw cell powders by immersing tubes in cold water and measure lysate volume (<i>see</i> Note 6).	247 248

5. Solubilize chromatin of lysates with DNase I (165 U/g of cells) 249
in the presence of 10 mM MnCl₂ for 1 h in Infors shaker at 250
16 °C and 180 rpm. This step can be replaced by sonication for 251
15 min on a water/ice mixture (45 s on–45 s off) and is 252
required for releasing proteins associated with chromatin. 253
6. Preclear solubilized lysate by centrifuging for 20 min at 254
4600 × *g* at 4 °C in tabletop centrifuge (*see Note 7*). 255
7. Clarify cell lysate by centrifuging for 20 min at 20,000 × *g* at 256
4 °C in SS-34 rotor. 257
8. Transfer cleared lysates to 15 mL tubes on ice by decanting/ 258
pipetting avoiding carryover of lipidic layer and measure vol- 259
ume (*see Note 8*). Save a small amount (~50 μL) for western 260
blot analysis of Input fraction. Use same amount of lysate for all 261
samples. 262

**3.3 Affinity
Purification of the
Complex Using
IgG Beads**

1. Wash necessary (*see Note 9*) amount of 50 mg/mL rabbit 263
IgG-coupled M-280 tosylactivated Dynabeads per sample 264
twice with 1 mL–2 mL TN150 buffer and once more with 265
1–2 mL of TN150 buffer supplemented with cOmplete 266
EDTA-free Protease Inhibitor Cocktail, using a magnetic 267
device. 268
2. Resuspend beads in 100 μL/sample of TN150 buffer supple- 269
mented with protease inhibitors. 270
3. Add extracts to beads and rotate in 15 mL tubes at 4 °C for 2 h. 271
4. Collect beads using the magnet and save a small amount of the 272
supernatant for western blot analysis of flow-through fraction. 273
5. Wash beads three times with 10 mL TN1000 supplemented 274
with protease inhibitor for 2–5 min, and once with 10 mL 275
TN150 (*see Note 10*). 276
6. Transfer beads to 1.5 mL tubes using TN150, remove the 277
buffer and resuspend in 600 μL TN150. 278
7. Elute His-tagged protein–RNA complexes from IgG beads by 279
cleaving the TEV site with the addition of 5 μL (20–25 U) of 280
homemade GST-TEV protease to the 600 μL of TN150 and 281
incubation for 2 h at 18 °C with 1000 rpm shaking, 1 min on– 282 AU3
20 s off in a Thermomixer. Here, an optional step of RNase H/ 283
oligo-dT cleavage can be performed to eliminate poly(A) tails 284
from polyadenylated reads. The reaction is carried out in the 285
same buffer and at the same time as TEV elution (*see Note 11*). 286
8. Collect beads and transfer eluates to fresh tubes (*see Note 12*). 287

**3.4 Partial RNase
Digestion**

1. Treat eluate with 0.05–0.2 U (1 μL of a 1/200 to 1/50 288
dilution; *see Note 13*) RNase-IT RNase cocktail for 5 min at 289
37 °C to shorten the protein-bound RNA. Vortex briefly after 290

	addition of the enzyme, move to Thermomixer and start counting time. Stagger when handling multiple samples.	291 292
	2. Quench the RNase reaction by quickly transferring 550 μ L to tubes containing 400 mg guanidine hydrochloride and vortex to dissolve. Save leftover RNase reaction (\sim 30–60 μ L) for western blot analysis of TEV Eluate fraction.	293 294 295 296
	3. Equilibrate the solution (typically \sim 830 μ L) for nickel affinity purification by adjusting to 0.3 M NaCl and 10–15 mM imidazole pH 7.9 (typically, 45 μ L of 3 M NaCl and 10 μ L of 1 M imidazole are added) (<i>see Note 14</i>).	297 298 299 300
3.5 Affinity Purification of the Complex Using Nickel Beads and on-Beads Linker Ligation		
3.5.1 Binding of Protein-RNA Complexes		
	1. Transfer 100 μ L (for each sample) of Ni-NTA agarose nickel beads slurry into Snap Cap spin columns lodged in their collection tubes and pack beads by centrifuging for 1 min at $1000 \times g$.	301 302 303 304
	2. Wash beads three times with 700 μ L Wash buffer I supplemented with protease inhibitors (<i>see Note 15</i>).	305 306
	3. Apply equilibrated samples to Ni-NTA agarose nickel beads in bottom-capped Snap Cap spin columns lodged in their collection tubes, close top of column, Parafilm around, and rotate at 4 $^{\circ}$ C overnight.	307 308 309 310
	4. Uncap top and bottom of Snap Cap spin columns and discard flowthrough (<i>see Note 16</i>) by centrifuging for 1 min at $1000 \times g$ max.	311 312 313
	5. Wash beads twice with 700 μ L Wash buffer I-1 M and then three times with 700 μ L 1 \times PNK buffer, each time by inverting top-capped Snap Cap spin columns lodged in their collection tubes and centrifuging for 1 min at $1000 \times g$.	314 315 316 317 318
3.5.2 On-Beads 3'-End Dephosphorylation (Optional)	The partial RNase digestion leads to the formation of phosphorylated RNA 3'-ends. A dephosphorylation step is generally required for the subsequent ligation of the 3' linker. However, when sequencing the nascent RNAs, the 3'-OH end in the catalytic center of the RNAPII enzyme is protected from the action of RNases. We have indeed observed identical results with and without phosphatase treatment for RNAPII CRAC.	319 320 321 322 323 324 325
	1. Add 80 μ L per sample of CIP mix (16 μ L 5 \times PNK buffer, 5 μ L (25 U) Quick CIP, 2 μ L (80 U) RNaseOUT, 57 μ L H ₂ O) on bottom-capped spin column, Parafilm top of column (<i>see Note 17</i>) and incubate at 37 $^{\circ}$ C for 30 min at 900 rpm in Thermomixer.	326 327 328 329 330
	2. Wash beads once with 400 μ L Wash buffer I-1 M and then three times with 400 μ L 1 \times PNK buffer, each time by inverting top-capped Snap Cap spin columns lodged in their collection tubes and centrifuging for 1 min at $1000 \times g$.	331 332 333 334

3.5.3	<i>On-Beads 3' Linker Ligation</i>	<ol style="list-style-type: none"> 1. Add 72 μL per sample of T4 RNA ligase 2 mix (16 μL 5\times PNK buffer, 26.7 μL 50% PEG 8000, 4 μL (800 U) T4 RNA ligase 2 truncated KQ, 2 μL (80 U) RNaseOUT, 23.3 μL H₂O) on bottom-capped spin column. 2. Add a different adaptor to each sample by adding 8 μL of 10 μM solution of 3' linker (1 μM final concentration), Parafilm top of column and incubate at 25 °C for 5–6 h at 900 rpm, 30 s on–1.5 min off in Thermomixer. 3. Wash beads once with 400 μL Wash buffer I-1 M and then three times with 400 μL 1\times PNK buffer, each time by inverting top-capped Snap Cap spin columns lodged in their collection tubes and centrifuging for 1 min at 1000 $\times g$. 	<p>336 337 338 339 340 341 342 343 344 345 346 347 348</p>
3.5.4	<i>On-Beads 5'-End Phosphorylation</i>	<p>RNase treatment leaves 5'-OH ends on RNA fragments. These must be phosphorylated to allow the addition of the 5' linker by T4 RNA ligase (<i>see</i> Subheading 3.5.5).</p> <ol style="list-style-type: none"> 1. Add 80 μL per sample of T4 PNK mix (16 μL 5\times PNK buffer, 2 μL (20 U) T4 PNK, 1 μL 100 μM ATP (1.25 μM final concentration), 2 μL (80 U) RNaseOUT, 59 μL H₂O) on bottom-capped spin column, Parafilm top of column and incubate at 37 °C for 45 min at 900 rpm in a Thermomixer (without collector tubes). 2. Wash beads twice with 400 μL Wash buffer I-1 M and then three times with 400 μL 1\times PNK buffer, each time by inverting top-capped Snap Cap spin columns lodged in their collection tubes and centrifuging for 1 min at 1000 $\times g$. 	<p>349 350 351 352 353 354 355 356 357 358 359 360 361 362</p>
3.5.5	<i>On-Beads 5' Linker Ligation</i>	<ol style="list-style-type: none"> 1. Add 80 μL per sample of T4 RNA ligase 1 mix (16 μL 5\times PNK buffer, 8 μL 10 mM ATP (1 mM final concentration), 4 μL (40 U) T4 RNA ligase 1, 1 μL 100 μM L5 miRCat 5' linker (1.25 μM final concentration), 2 μL (80 U) RNaseOUT, 49 μL H₂O) on bottom-capped spin column, Parafilm top of column and incubate at 16 °C overnight at 700 rpm, 30 s on–1.5 min off in Thermomixer (without collector tubes). 2. Wash beads three times with 400 μL Wash buffer II, each time by inverting top-capped Snap Cap spin columns lodged in their collection tubes and centrifuging for 1 min at 1000 $\times g$. 	<p>363 364 365 366 367 368 369 370 371 372 373</p>
3.5.6	<i>Purification of Protein–RNA Complexes</i>	<ol style="list-style-type: none"> 1. Move Snap Cap spin columns to clean 1.5 mL Eppendorf tubes and add 200 μL Elution buffer on beads. Incubate on bench 5 min with occasional gentle hand-shaking, then centrifuge for 1 min at 1000 $\times g$. 2. Repeat this step by keeping the Snap Cap spin columns in the same 1.5 mL tube and adding further 200 μL Elution buffer on beads. Centrifuge for 1 min at 1000 $\times g$. Save a small amount (about 10 μL) for silver or Cy5 stain analysis of Ni-NTA Eluate 	<p>374 375 376 377 378 379 380 381</p>

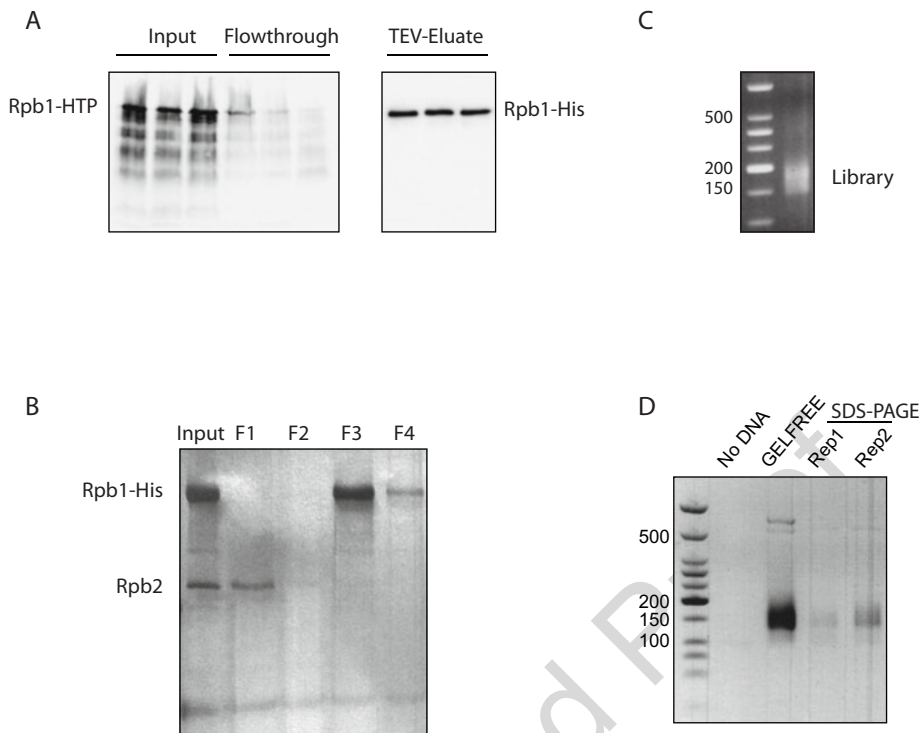


Fig. 3 Representative examples of typical controls to verify the efficiency of the protein–RNA complex purification and library construction. **(a)** Western blot to verify the efficiency of the first purification step on the IgG column. Shown are proteins from three independent preparations probed with an anti-His antibody. Input (total extract), flowthrough and TEV-eluate fractions are analyzed as indicated. **(b)** Four purification fractions from GELFREE 8100 (F1 to F4) are analyzed by PAGE and silver staining together with a fraction of the input (eluate from Subheading 3.5.6, see main text). Using this fractionation procedure Rpb1-6×His is typically found in fractions 3 and 4. The faster migrating band present in the input and in fraction F1 is most likely a small amount of Rpb2 that survived the first denaturing purification step. **(c)** Analysis of the cDNA library average size after PCR amplification using LA Taq polymerase (25 cycles, analyzed on 2% agarose gel stained with Ethidium bromide, see Subheading 3.9, step 2). The typical profile is a smear of fragments in the 150–200 bp size range. **(d)** Comparison of typical yields of cDNA library using either the original procedure (lanes SDS-PAGE, two independent replicates, Rep1 and Rep2) or the GELFREE 8100 system (lane GELFREE). Equivalent amount of TEV eluate were used for the two purification methods. Libraries were analyzed as in C

fraction (*see Note 18*) or western blot. Keep eluates on ice 382
during protein analysis before next step. 383

3. The original CRAC procedure includes a radioactive labeling 384
step of the RNA and a size selection step based on SDS-PAGE 385
and transfer on nitrocellulose membrane [5]. We found this 386
step to be time-consuming and the recovery of the RNA from 387
the membrane to be a major source of material loss. This step 388
was replaced by a different method to recover the protein– 389
RNA complex that does not use radioactivity and allows com- 390
bining the samples for a common elution (Figs. 1 and 3). We 391
presently use an apparatus called GELFREE 8100 that allows 392

fractionation of complex protein samples. The eluted samples 393
 from step 2 are combined and concentrated using a Vivacon 394
 500 ultrafiltration spin column. The concentrated, mixed sam- 395
 ple is applied to several GELFREE cassette lanes (typically 2–3 396
 lanes for 6 samples total) using manufacturer conditions for the 397
 approximate size of the protein–RNA complex to be purified. 398
 5% gel cassettes have been used for Rpb1 isolation. The eluted 399
 fractions are analyzed by PAGE as described in **Note 18** for 400
 identifying the fraction containing the protein of interest. Nev- 401
 ertheless, this purification step has been skipped without a 402
 major increase in the level of background for some proteins 403
 (*see Note 19*). 404

3.6 RNA Isolation

1. Process fractions containing protein–RNA complexes. Add 406
 SDS to 0.5% final concentration, and treat with 100 μg of 407
 proteinase K at 50 °C for 1 to 2 h at 800 rpm, 1 min on– 408
 1 min off in Thermomixer. 409
2. Extract proteinase K-digested fractions twice with phenol– 410
 chloroform–isoamyl alcohol (25:24:1) and once with chloro- 411
 form by vortexing (or mixing, *see Note 20*) and centrifuging 412
 for 5–10 min at 16,000 $\times g$ at 10 °C, each time being careful to 413
 avoid the interphase. 414
3. Precipitate RNA by adding NaOAc pH 5.5 to 0.3 M final, 1 μL 415
 GlycoBlue carrier, 2.5 volumes ethanol (or 1 volume Isopro- 416
 panol) to the aqueous phase 2 h to overnight at –20 °C. 417
4. Pellet RNA by centrifuging for 30 min at 16,000 $\times g$ at 4 °C, 418
 carefully remove supernatant, wash once with 0.5 mL–1 mL 419
 70% ethanol, centrifuge for 10 min at 16,000 $\times g$ at 4 °C, 420
 carefully remove supernatant, speed-vac dry pellet for 5 min 421
 without heating. 422

3.7 Synthesis of the cDNA Library

1. Resuspend the RNA pellet in 5 μL water at room temperature 423
 for 5 min at 900 rpm in Thermomixer and wash tube walls by 424
 pipetting (*see Note 21*). 425
2. Add 8 μL RT Mix I (2.6 μL 5 \times SSIV buffer, 0.65 μL 100 mM 426
 DTT, 1 μL 10 μM RT L3–2 primer, 3.75 μL H₂O) and 427
 incubate sample at 70 °C for 10 min in Thermomixer to anneal 428
 RT primer. 429
3. Set temperature to 50 °C and leave sample to continue anneal- 430
 ing in Thermomixer until it reaches temperature. 431
4. Briefly spin sample, then add 7 μL RT Mix II (1.4 μL 5 \times SSIV 432
 buffer, 0.35 μL 100 mM DTT, 2 μL 5 mM dNTP mix, 1 μL 433
 RNaseOUT, 1 μL SuperScript IV Reverse Transcriptase, 434
 1.25 μL H₂O) and incubate sample at 50 °C for 1 h in 435
 Thermomixer. 436
5. Inactivate reaction by incubating sample at 80 °C for 10 min. 437

6. Add 1 μL (5 U) RNase H and incubate cDNA sample at 37 °C for 30 min in Thermomixer in order to digest residual RNA.
7. Mix cDNA from the different GELFREE lanes (or different purification fractions) and dilute the sample two-fold (*see Note 22*).

3.8 Amplification of the cDNA Library

The cDNA library is amplified by PCR for sequencing. We calculate the minimum number of PCR cycles required to obtain sufficient material for Illumina NextSeq500 sequencing (currently 10 μL of a 0.2 ng/ μL solution, roughly 2 nM, assuming an average size of 150 bp for the library fragments). This is done in two steps: (1) estimation of the absolute concentration of the starting cDNA library and (2) total amplification of the library and purification of the fragments. A step for analyzing the size distribution of the library is also included.

3.8.1 Estimate of the Minimal Number of Cycles for Amplifying the cDNA Library

1. Mix 2 μL of cDNA sample or of appropriate dilutions of the reference library with 8 μL qPCR mix (0.8 μL 25 mM MgCl_2 , 0.134 μL 1a, 0.857 μL 1b, 1 μL 10 μM miRCat_PCR#2 primer (1 μM final concentration), 1 μL 10 μM P5-3' primer (1 μM final concentration), 4.2 μL H_2O) in capillaries. Spin capped capillaries 1 min at 400 $\times g$. Run qPCR reactions with Library Amplification program: 5 min 95 °C—(10 s 95 °C—40 s 60 °C—30 s 72 °C) 30 \times .
2. Estimate the number of cycles necessary for La Taq amplification of cDNA sample to obtain appropriately concentrated library for sequencing. Defining the minimal amount of PCR cycles is important for limiting the number of PCR duplicates and avoiding amplification biases. We first estimate the absolute concentration of the cDNA library before amplification by qPCR, comparing it to a standard of similar size and known concentration that can be amplified with the same primers. We find that using a capillary LightCycler qPCR apparatus gives the most reliable result, but other methods can be used, for instance Droplet Digital PCR.

We amplify 2 μL of the cDNA library in duplicate in a same reaction as a serial dilution of the standard. This gives a Ct that corresponds to the starting concentration of the library (CtStart). Using the standard for measuring PCR efficiency we also estimate the Ct number that is expected for a cDNA library that would be at the expected final concentration (CtFinal). CtFinal—CtStart is the number of PCR cycles that would be required to bring the starting cDNA library to the expected concentration. This is clarified in the example shown in Fig. 4.

From these data the PCR efficiency for this couple of oligos can be calculated (in this case $10^{-0.2658} = 1.84$) and, by extrapolation,

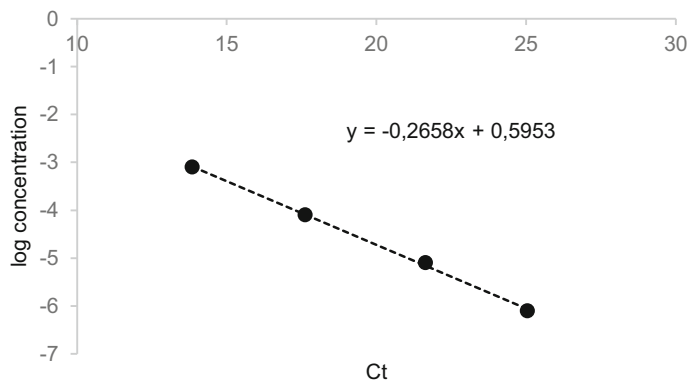


Fig. 4 Example of the estimation of the Ct number that is expected for a cDNA library that would be at the expected final concentration

Table 1
Caption

t.1
AU4

Standard concentration (ng/ μ L)	Log (concentration)	Mean qPCR values
0.0008	-3.10	13.85
0.00008	-4.10	17.62
0.000008	-5.10	21.64
0.0000008	-6.10	25.04

t.2
t.3
t.4
t.5
t.6

the Ct expected for a solution at 0.2 ng/ μ L (CtFinal, here 4.87). 483
 Our test sample has given, in the same reaction, an average Ct 484
 (CtStart) of 15.20. 10.33 cycles (i.e., 15.20–4.87) would be 485
 required for the final PCR reaction (Table 1). However, because 486
 the latter produces large volumes that are concentrated in the final 487
 step, a concentration factor should be considered for determining 488
 the number of cycles required. The final PCR is performed using a 489
 tabletop apparatus in 25 μ L reactions. In our hands, PCR amplifi- 490
 cation using La Taq polymerase (TaKaRa) in 25 μ L reactions has 491
 very comparable efficiency as qPCR using capillary LightCycler, but 492
 this should be carefully verified and considered for the final 493
 calculation. 494

The total volume of the final PCR reaction depends on that of 495
 the starting cDNA library. For a 40 μ L RT reaction (*see* Subheading 496
 3.7) we use 8 \times 25 μ L reactions (5 μ L of cDNA mix per PCR 497
 reaction). We generally set up a separate RT reaction per each 498
 protein–RNA eluate and using GELFREE we typically obtain 499
 between 1 and 3 independent size fractionation eluates (*see* Note 500
 19). Therefore, the total volume of the starting cDNA library is 501
 between 40 and 120 μ L, and the total volume of the final PCR 502
 reaction will be between 200 and 600 μ L. In the final step (*see* step 503

6 of Subheading 3.9) samples are concentrated to a final volume of 40 μL . If we consider a volume of 400 μL for the final PCR reaction of our example, this corresponds to a ten-fold concentration. Because of this final concentration factor, less PCR cycles are required to obtain the required concentration for sequencing. This “gain” in PCR cycles depends on the PCR efficiency that in our hands is very similar to the one measured by qPCR for this couple of oligos. Therefore:

- Vol. PCR: 400 μL .
- Vol. elution: 40 μL .
- Enrichment: 10.
- “gain” in PCR cycles (PCR efficiency 1.84): 3.76.

In our example, the minimal number of PCR cycles required is therefore $10.33 - 3.76 = 6.57$. A couple of additional amplification cycles are typically added to this number to account for possible errors in the estimated PCR efficiency and to obtain excess material for quality control of the library by the sequencing facility when needed.

The library concentration in the final PCR reaction is verified by qPCR (LightCycler) before purification. The Ct of this test reaction should be inferior to the expected Ct for a 0.2 ng/ μL reaction (4.87 in our example) plus the cycles corresponding to the concentration factor (3.76), that is, 8.63 for our example.

3.8.2 Analysis of the Size Profile of cDNA Library

1. Mix 2 μL of cDNA sample or H_2O control with 23 μL PCR mix (2.5 μL $10\times \text{Mg}^{2+}$ plus LA PCR buffer, 2 μL 2.5 mM dNTP mix (0.2 mM final concentration), 1 μL 10 μM miR-Cat_PCR#2 primer (0.4 μM final concentration), 1 μL 10 μM P5-3' primer (0.4 μM final concentration), 0.5 μL (2.5 U) LA Taq, 16 μL H_2O). Amplify for 25 cycles with the following program: 2 min 95 $^\circ\text{C}$ - (30 s 95 $^\circ\text{C}$ - 45 s 58 $^\circ\text{C}$ - 1 min 72 $^\circ\text{C}$) $25\times$ - 5 min 72 $^\circ\text{C}$.
2. Control LA Taq amplification on 2% agarose by running 5 μL of sample and control reactions with a low molecular weight DNA ladder. An optimal profile appears as a smeary population spreading above 150 bp and up to 200 bp without defined bands. A band of roughly 100 nt would indicate the presence of primer dimers or very small inserts.

3.9 Final PCR and Library Purification

1. Based on total cDNA sample volume, set up N separate 25 μL LA Taq PCR reactions and amplify for calculated number of cycles (typically 8–11). Mix 5 μL of cDNA sample (*see Note 23*) with 20 μL PCR mix (2.5 μL $10\times \text{Mg}^{2+}$ plus LA PCR buffer, 2 μL 2.5 mM dNTP mix (0.2 mM final concentration), 1 μL 10 μM miRCat_PCR#2 primer (0.4 μM final

- concentration), 1 μL 10 μM P5–3' primer (0.4 μM final concentration), 0.5 μL (2.5 U) LA Taq, 13 μL H_2O). Amplify with the following program: 2 min 95 °C - (30 s 95 °C - 45 s 58 °C - 1 min 72 °C)_{Nx} - 5 min 72 °C. 549
2. Pool LA Taq PCR reactions in one tube and keep on ice during the subsequent qPCR test. 553
 3. Test CRAC library amplification by capillary qPCR. Mix 2 μL of CRAC library or H_2O control with 8 μL qPCR mix as in **step 1** of Subheading 3.8.1. This should give a number of cycles at least equal to calculated theoretical number (*see* Subheading 3.9, **step 1** and **Note 24**). 555
 4. Digest primers in CRAC library sample with Exonuclease I. Add 2.5 μL (50 U) per 200 μL PCR reaction and incubate at 37 °C for at least 1 h in Thermomixer. 560
 5. Inactivate reaction by incubating sample at 80 °C for 20 min. 563
 6. Purify CRAC library with NucleoSpin Gel and PCR cleanup. Load sample several times on the same unique column if needed. Dry columns with two 1 min and 30 s centrifugations, and a 3 min incubation of the open column at 45 °C. Recover CRAC library with two successive 20 μL elutions with water or elution buffer, each with incubation at room temperature for 5 min, followed by centrifugation for 1–1.5 min. 564
 7. Quantify CRAC library with a Qubit Fluorometer and Qubit dsDNA HS Assay Kit by reading 4 μL of sample. Typical final concentration ranges from 0.5 to 2 ng/ μL . 571
 8. Store CRAC library at -20 °C until ready for sequencing. 574

3.10 Data Analysis

Analysis of the data obtained is based on the pyCRACutility suite [11] modified as described [6, 7]. Briefly, CRAC samples are demultiplexed using the pyBarcodeFilter script and subsequently, the 5' linker (Subheading 2.5.2) is clipped with Cutadapt [10, 12]. The resulting insert is quality trimmed from the 3' end using Trimmomatic rolling mean clipping [13] (window size = 5, minimum quality = 25). At this stage we use the pyCRAC script pyFastqDuplicateRemover to collapse PCR duplicates using a 6 nucleotides random tag that is part of the 3' linker (Subheading 2.5.1). During demultiplexing, pyBarcodeFilter retains this information in the header of each sequence. Reads containing more than 70% of Ts are removed. These are common source of artefactual signals due to the unspecific mapping of stretches of As (sequencing is from the 3'-end of cDNAs). The resulting sequences are reverse complemented with Fastx reverse complement (part of the fastx toolkit, http://hannonlab.cshl.edu/fastx_toolkit/) and mapped to the required genome with bowtie2 (using “-N 1” option) [14]. Unmapped reads are searched for the presence of poly (A) tails and polyadenylated reads are mapped independently. 575

When the *S. pombe* spike-in is present in the sample, reads that do not map on *S. cerevisiae* are mapped on the *S. pombe* genome after filtering for reads shorter than 40 nt to avoid cross mapping of *S. cerevisiae* reads on *S. pombe*. The amount of mapped spike-in reads is used for normalization purposes. The output (directional wig and bigwig files) provides the raw mapped read counts, the read counts normalized to the total number of mapped reads per each sample and the read counts normalized to the *S. pombe* spike-in mapped reads.

Because the whole read is considered for the read counts, the resolution of the signal depends on the average size of the reads. To increase signal resolution, we also generated tools to only map the read 3'-end (which defines the last position of RNAPII during transcription or backtracking after RNA cleavage) or its 5'-end. Cross-linking induces deletions or misincorporations during reverse transcription of the RNA, which can be used as proxy for the position of protein–RNA contact. This information can also be used to distinguish bona fide signals from background due to highly abundant RNA molecules (generally tRNAs and rRNAs). We have generated tools to extract reads containing mismatches or short deletions (the latter appear to be more frequent), as well as smoothing procedures. These scripts are available upon request.

4 Notes

1. Other appropriate drop-outs can be used to grow cells. However, tryptophan should be excluded from growth medium as it quenches UV light, thus reducing cross-linking efficiency [5].
2. *S. pombe* cells used to spike-in experiments can be grown in lower volumes of EMM + xAA (2–5 amino acids are added) minimal medium, added in the desired proportion to *S. cerevisiae* cultures if cross-linked together (usually 0.5%–1% of *S. cerevisiae* cells) or cross-linked separately for lysate mixing.
3. Optimal UV cross-linking time should be determined empirically: increased cross-linking will reduce the efficiency of the reverse transcription step and might induce DNA damage response. In our hands 50 s is best for RNAPII.
4. For unstable proteins, lysis buffer is supplemented with benzamide at 2.5 mM final concentration.
5. Frequency can be as low as 10 Hz and be increased up to 20 Hz. Higher frequency will give better lysis but might be detrimental to some proteins. Preparation of extracts can be alternatively done with a Carver Press, in which case cell pellets resuspension volumes need to be adjusted to Carver Press cylinders' capacity and extracts need to be processed

- immediately. Initial TN150 lysis buffer can be supplemented with 5%–10% glycerol (or even substituted with classic IP lysis buffers). Glycerol must be included if extracts are prepared using Carver Press.
6. Thawing of lysates can be done in Infors shaker set at 16 °C and 180 rpm or under lukewarm water (risky for unstable proteins).
 7. Preclearing of lysates after chromatin solubilization can be omitted and lysates subjected only to high-speed centrifugation. However, preclearing helps avoiding excessive lipidic carryover.
 8. A Bradford assay can be carried out on cleared lysates before the first immunoprecipitation to control for protein concentration. This can be tricky due to the residual presence of lipids and pigments.
 9. 100 µL (for average abundance proteins) to 250 µL (for high abundance proteins such as Rpb1) M-280 tosylactivated Dynabeads coupled with rabbit IgG (5 mg–15 mg of beads per samples).
 10. We have observed that cOmplete protease inhibitor inhibits to some extent TEV activity. For this reason, the final wash of Dynabeads as well as the elution step are performed in TN150 without protease inhibitor.
 11. A fraction of the reads associated to some proteins are polyadenylated, which, if the poly(A) tail is too long, might prevent reading the sequence of the fragment (sequencing is done from the 3'-end). For this reason, an optional RNase H/oligo-dT degradation step can be performed in order to partially digest poly(A) tails. The same 600 µL of TN150 in which TEV digestion is performed are in this case supplemented with 0.4 µM poly-dT oligonucleotide, 3 nM MgCl₂, and 2 µL (10 U) RNase H.
 12. If needed, dried beads can be stored at –20 °C for western blot analysis.
 13. Conditions for partial RNase digestion should be determined empirically and both enzyme concentration and digestion time can vary. For RNAPII CRAC, 0.2 U of RNase-IT cocktail are used.
 14. Imidazole must be present at roughly 10–20 mM concentration for specific binding to Ni-NTA agarose beads.
 15. When working with Snap Cap spin columns, leaving the top cap open when centrifuging will prevent foam formation.
 16. The flowthrough of Ni-NTA columns cannot be used for western blot analysis as it contains guanidine hydrochloride, which affect PAGE running.

17. For each enzymatic step, for which the reaction volume is low, be sure to cap only the bottom of the tube and close the top with Parafilm in order not to squeeze out the reaction mix.
18. Side-by-side analysis of TEV and Ni-NTA Eluates (or at least of Ni-NTA Eluates if analysis of the first immunoprecipitation has been carried out separately) should be performed to ensure successful elution and estimate eventual loss of material along the procedure, as well as general profile (degradation fragments, contaminants, etc.). This is particularly important if samples are subsequently mixed, as this is the last point in the procedure to check samples individually. Fractions obtained after GELFREE purification are also checked in the same way. Analysis can be done by Silver Staining, or by Quick Stain Cy5 (staining efficiency may vary according to proteins, protein solution composition, and relative amounts), and/or by western blot. Samples for Quick Stain Cy5 analysis are treated according to manufacturer's protocol and can be stored at -20°C . SDS to 0.1% should be added to samples from GST-TEV and Ni-NTA elution steps before Cy5 staining.
19. Unfortunately, GELFREE is being discontinued by Abcam. We are studying alternative solutions based on similar principles (e.g., elution from PAGE or other fractionation methods).
20. Phase Lock Gel Heavy tubes can be used for second phenol-chloroform-isoamyl alcohol (25:24:1) and subsequent chloroform extractions (in the same tube if volumes are compatible). In this case do not vortex samples (see also manufacturer's instructions).
21. Resuspension of RNA pellets in H_2O rather than in RT Mix I prevents foaming up of the sample due to detergents in $5\times$ SSIV buffer. Alternatively, this can be done by directly adding $13\ \mu\text{L}$ of RT Mix I (in this case prepared with $8.75\ \mu\text{L}$ H_2O).
22. Sample two-fold dilution is expected to reduce the impact of the RT buffer on the subsequent amplification steps and improve accuracy due to the pipetting of larger volumes.
23. Use of $5\ \mu\text{L}$ of cDNA sample in $25\ \mu\text{L}$ represents a scale up of the qPCR reaction for determining the concentration of the library. This is important to approach similar conditions of amplification. We use $25\ \mu\text{L}$ reaction volumes to minimize thermal inertia and independent reactions to minimize PCR artefacts that might arise from the stochastic early amplification of fragments.
24. If qPCR test of CRAC library results in significantly different number of cycles than expected, additional LA Taq PCR amplification cycles can be performed, provided the sample has not been treated with Exonuclease I.

727 **References**

- 729 1. Churchman LS, Weissman JS (2011) Nascent
730 transcript sequencing visualizes transcription at
731 nucleotide resolution. *Nature* 469:368–373
- 732 2. Burger K, Mühl B, Kellner M et al (2013)
733 4-thiouridine inhibits rRNA synthesis and
734 causes a nucleolar stress response. *RNA Biol*
735 10:1623–1630
- 736 3. Baejen C, Andreani J, Torkler P et al (2017)
737 Genome-wide analysis of RNA polymerase II
738 termination at protein-coding genes. *Mol Cell*
739 66:38–49.e6
- 740 4. Schwalb B, Michel M, Zacher B et al (2016)
741 TT-seq maps the human transient transcrip-
742 tome. *Science* 352:1225–1228
- 743 5. Bohnsack MT, Tollervey D, Granneman S
744 (2012) Identification of RNA helicase target
745 sites by UV cross-linking and analysis of
746 cDNA. *Meth Enzymol* 511:275–288
- 747 6. Challal D, Barucco M, Kubik S et al (2018)
748 General regulatory factors control the fidelity
749 of transcription by restricting non-coding and
750 ectopic initiation. *Mol Cell* 72:955–969.e7
- 751 7. Candelli T, Challal D, Briand J-B et al (2018)
752 High-resolution transcription maps reveal the
753 widespread impact of roadblock termination in
754 yeast. *EMBO J* 37:e97490
- 755 8. Luciano P, Jeon J, El-Kaoutari A et al (2017)
756 Binding to RNA regulates set1 function. *Cell*
757 *Discov* 3:17040
9. Villa T, Barucco M, Martin-Niclos M-J et al 758
(2020) Degradation of non-coding RNAs pro- 759
motes recycling of termination factors at sites 760
of transcription. *Cell Rep* 32:107942 761
10. Han Z, Jasnovidova O, Haidara N et al (2020) 762
Termination of non-coding transcription in 763
yeast relies on both an RNA pol II CTD inter- 764
action domain and a CTD-mimicking region in 765
Sen1. *EMBO J* 39:e101548 766
11. Webb S, Hector RD, Kudla G et al (2014) 767
PAR-CLIP data indicate that Nrd1-Nab3- 768
dependent transcription termination regulates 769
expression of hundreds of protein coding genes 770
in yeast. *Genome Biol* 15:R8 771
12. Martin M (2011) Cutadapt removes adapter 772
sequences from high-throughput sequencing 773
reads. *EMBnet J* 17(1). [https://doi.org/10.](https://doi.org/10.14806/ej.17.1.200) 774
[14806/ej.17.1.200](https://doi.org/10.14806/ej.17.1.200) 775
13. Bolger AM, Lohse M, Usadel B (2014) Trim- 776
momatic: a flexible trimmer for illumina 777
sequence data. *Bioinformatics* 30:2114–2120 778
14. Langmead B, Salzberg SL (2012) Fast gapped- 779
read alignment with bowtie 2. *Nat Methods* 9: 780
357–359 781
15. Schaugency P, Merran J, Corden JL (2014) 782
Genome-wide mapping of yeast RNA polymer- 783
ase II termination. *PLoS Genet* 10:e1004632 784