In situ dissection of RNA functional subunits by domain-specific chromatin isolation by RNA purification (dChIRP)

Jeffrey J. Quinn[1,2] and Howard Y. Chang[1*]

[1] Howard Hughes Medical Institute and Program in Epithelial Biology, Stanford University School of Medicine, Stanford CA 94305, USA
[2] Department of Bioengineering, Stanford University Schools of Medicine and Engineering, Stanford CA 94305, USA
* Corresponding author

Running head: RNA dissection by dChIRP

Summary

Here we describe domain-specific chromatin isolation by RNA purification (dChIRP), a technique for dissecting the functional domains of a target RNA *in situ*. For an RNA of interest, dChIRP can identify domain-level intramolecular and intermolecular RNA-RNA, RNA-protein, and RNA-DNA interactions, and maps the RNA's genomic binding sites with higher precision than domain-agnostic methods. We illustrate how this technique has been applied to the roX1 lncRNA to resolve its domain-level architecture, discover its protein- and chromatin-interacting domains, and map its occupancy on the X chromosome.

Keywords: RNA, long noncoding RNA, ChIRP, chromatin, RNA-binding protein, RNA domains

1. Introduction

With the advent of deeper RNA sequencing and improved epigenomic tools, the repertoire of noncoding RNAs has exploded with thousands of uncharacterized, potentially functional long noncoding RNAs (lncRNAs; *reviewed in* Rinn and Chang, 2012). Only a small fraction of identified lncRNAs have been mechanistically characterized, but those studied have been shown to participate in diverse processes and exhibit precise tissue- and disease-specific regulation (Cabili et al., 2011). A common hallmark of many lncRNAs is that they interact with the chromatin, thereby acting at the interface between the genome, chromatin, and transcriptional machinery. It has been proposed that lncRNAs may possess functional modular domains that coordinate such varied nucleic acid and protein interactions (Guttman and Rinn, 2012). To date, testing this hypothesis has been hindered by a relative scarcity in RNA-centric technologies for systematically studying lncRNA functions. Existing methods for identifying lncRNA domains are effective, yet limited; for example, testing the function of mutant transgenic lncRNAs with deleted or disrupted domains is laborious and suffers from drawbacks due to non-native RNA expression. Next-generation sequencing-enabled techniques such as cross-linking immunoprecipitation (CLIP) and proximity ligation can determine sites of

RNA-protein and RNA-RNA interactions, respectively; however both techniques are protein-centric in approach (Ule et al., 2005; Helwak et al., 2013).

Here we describe domain-specific chromatin isolation by RNA purification (dChIRP), a method for dissecting RNA functional domains *in situ.* This technique leverages the purification strategy of the traditional ChIRP method, which has been used to map the genomic binding sites of chromatin-associated lncRNAs towards studying mechanisms of dosage compensation, epigenetic silencing, and immunoregulation (Chu et al., 2011; Colak et al., 2014; Li et al., 2013). In dChIRP, biotinylated antisense oligos are designed to target specific regions of a target RNA (**Fig.1A**). Cells are cross-linked and chromatin is prepared by solubilizing and shearing the nucleic acids with sonication; domain-specific oligo pools are added to recover specific domains of the target RNA and affinity purified, along with any co-purified biomolecules that associate with the target RNA (**Fig.1B**). Co-recovered RNA, proteins, and DNA are extracted and analyzed by quantitative polymerase chain reaction (qPCR), immunoblotting, or sequencing (**Fig.1C**). From these results it is possible to determine the target RNA's intramolecular topology, interacting RNA species, protein- or chromatin- binding RNA domains, and the RNA's genomic binding sites. Thus, dChIRP can map the RNA-, DNA-, and protein-interacting domains of a target RNA in one *in situ* experiment.

dChIRP has been used to dissect the functional domains of roX1, a lncRNA involved in dosage compensation in *Drosophila* (Quinn et al, 2014; *reviewed in* Conrad & Akhtar, 2012). As an example of the output of a typical dChIRP experiment, we present the results of roX1 dChIRP.

2. Materials

Equipment

Centrifuge
Table-top mini centrifuge
Heat block, preheated to 95ºC
Vortex
Nanodrop spectrophotometer (Thermo Scientific)
Hybridization oven with rotator, preheated to 37ºC
Handheld motorized douncer with 1.5mL tube disposable probe tips (e.g. Argos, VWR)
Bioruptor (Diagenode) or E-Series Focused Ultrasonicator (Covaris)
1.5mL and 15mL DynaMag-2 magnetic strip (Life Technologies)
Quantitative PCR thermocycler (e.g. Roche LightCycler)

Reagents and consumables

Phosphate Buffered Saline (PBS)
16% w/v formaldehyde (Pierce) and/or 25% glutaraldehyde (Sigma) (*see* **Note 1**)
1.25M glycine
Complete Protease Inhibitor (PI; Roche) stock (50x): 1 pellet dissolved in 1mL nuclease-free H_2O (store at -20ºC)
PMSF (pheylmethylsulfonyl fluoride) stock (100x): 100mM PMSF in isopropanol (store at -20ºC)
SUPERase-In RNase Inhibitor (20U/µL, Ambion): use as a 200x stock (store at -20ºC)

Liquid nitrogen or dry ice
RNeasy Mini Columns and QIAquick PCR Purification Kit (Qiagen)
TRIzol Reagent (Life Technologies)
Phenol:chloroform:isoamyl alcohol, 25:24:1 (Life Technologies)
Phase-lock gel heavy tubes (5 PRIME)
GlycoBlue glycogen, 15 mg/mL (Life Technologies)
100% ethanol
100% isopropanol
Chloroform
3M sodium acetate (NaOAc)
Dynabeads MyOne Streptavidin C1 magnetic beads (Life Technologies)
Turbo DNase DNA-free kit (Ambion)
Sequencing library prep kit (e.g. NEBNext ChIP-seq kit)

Buffers

Nuclear Lysis Buffer: 50mM Tris-HCl pH 7.0, 10mM EDTA, 1% SDS, 1x PI, 1x PMSF, 1x SUPERase-In. (Unsupplemented Nulcear Lysis Buffer may be stored at room temperature without PI, PMSF, and SUPERase-In. Always add PI, PMSF, and SUPERase-In fresh.)

Swelling Buffer: 100mM Tris-HCl pH7.0, 10mM KOAc, 15mM MgOAc, 1% NP-40, 1x PMSF, 1x PI, 1x SUPERase-In. (Always prepare fresh.)

Proteinase K Buffer: 100mM NaCl, 10mM Tris-HCl pH7.0, 1mM EDTA, 0.5% SDS, 1mg/mL Proteinase K (Ambion). (Unsupplemented Proteinase K Buffer may be stored at room temperature without Proteinase K. Always add Proteinase K fresh.)

Hybridization Buffer: 750mM NaCl, 1% SDS, 50mM Tris-HCl pH 7.0, 1mM EDTA, 15% formamide, 1x PI, 1x PMSF, 1x SUPERase-In. (Always prepare fresh.)

Wash Buffer: 2x SSC (NaCl and sodium citrate), 0.5% SDS, 1x PMSF. (Always prepare fresh.)

DNA Elution Buffer: 50mM $NaHCO_3$, 1% SDS, 100µg/mL RNase A, 0.1U/µL RNase H. (Always prepare fresh.)

Protein Elution Buffer (1x): 95µL Laemmli Sample Buffer (Bio-Rad), 100µL nuclease-free H_2O, 5µL ß-mercaptoethanol. (Always prepare fresh.)

Protein Elution Buffer (2x): 95µL Laemmli Sample Buffer (Bio-Rad), 5µL ß-mercaptoethanol. (Always prepare fresh.)

3. Methods

The ideal RNA target for dChIRP is (i) localized to the nucleus and/or contacting chromatin, (ii) abundant, and (iii) of sufficient length for domain-specific recovery and shearing.

It is first important to determine where the RNA is localized, using techniques such as RNA fluorescent in situ hybridization (FISH) or cellular fractionation. If the RNA is present in the nucleus or in the chromatin fraction, it is wise to apply DNA-sensitive modalities of dChIRP (e.g. qPCR or sequencing) to detect the genomic binding sites of the RNA. If the RNA is outside the nucleus, qPCR and sequencing will likely not be productive.

Second, abundant RNAs are better targets than rare transcripts since their higher concentration increases the likelihood of recovering the RNA and its interaction partners. If the RNA target is rare, it may be necessary to perform dChIRP from cells ectopically overexpressing the RNA. Results from ectopic overexpression should be interpreted with caution, as they may suffer from complications due to non-physiological RNA concentration, RNA misfolding, RNA instability, and exotic locus of transcription.

Lastly, since the domain-specific recovery of dChIRP relies on RNA fragmentation by sonication, the RNA must be of sufficient length to be sheared. The lower limit of resolution for a dChIRP experiment is determined by the RNA shearing, which is ~200-500bp by sonication. As such, RNAs of at least 1kb are suitable for dissection by dChIRP. For example, dChIRP of roX1 RNA (~3.7kb) identified six distinct RNA domains (Quinn et al., 2014), but dChIRP of the smaller roX2 RNA (~0.6kb) failed to identify distinct domains (data not shown).

3.1 Design antisense DNA oligos and oligo pools

3.1.1 Access the Stellaris single-molecule RNA FISH probe designer from Biosearch Technologies (www.biosearchtech.com). Input the RNA sequence into the program. Mask repeats or regions with extensive homology to the reference genome. Allow a minimum spacing length of 2 nucleotides. The program will design a series of non-overlapping 20-mer oligos that tile the RNA of interest, avoiding regions with high occurrence in the reference genome. If the RNA is too long for the program, divide the RNA and submit each segment individually or allow a greater number of maximum probes.

3.1.2 Export the sequences of the designed oligos and check for homology between two or more oligos using a linear alignment or motif discovery program (such as MEME, Bailey et al., 1994). If two or more oligos share significant sequence homology (sense or antisense), they must be excluded, as these oligos can cross-hybridize to multiple sites of the target RNA. For this reason, it is important to avoid targeting repetitive RNA elements.

3.1.3 Divide the target RNA into domains of at least 500bp. The domains can be determined by even subdivision of the RNA or based on existing evidence of functional domains (e.g. conservation, genetic, or biochemical evidence). For example, the domains of roX1 used for dChIRP were devised based on individual-nucleotide resolution CLIP experiments, which suggested that roX1 contained three protein interaction domains (**Fig.2A**; Ilik et al., 2013)

3.1.4 Within each domain, select an equal number of oligos; these are the dChIRP oligo pools. Order antisense DNA oligos with a 3'-biotin-TEG modification.

3.1.5 Label the oligos according to their position on the RNA (e.g. RNA-1, RNA-2, etc.). Dilute to a concentration of 100µM.

3.1.6 Prepare dChIRP oligo pools by adding an equal volume of the 100µM DNA oligos to a new tube. Vortex to mix and store at -20ºC. Oligos will register ~600ng/µL on a Nanodrop using the ssDNA setting.

3.1.7 For dChIRP-sequencing experiments, prepare even-odd oligo pools. For each RNA domain, split the oligos into two alternating pools: even and odd (see **3.10.1** and Chu et al., 2011).

3.2 Collect and cross-link biological material and prepare chromatin

3.2.1 Grow a relevant cell type in tissue culture that expresses the RNA of interest. When the cultures reach confluence, harvest the cells and pool into 50mL tubes. Centrifuge cells at 1000rpm for 5min to pellet, aspirate media, and resuspend pellet in PBS. Count the cells and redistribute such that each 50mL tube contains 40 million cells, which is typically sufficient material for two ChIRP samples (*see* **Note 2**). Repeat spin and aspirate PBS from pellet.

3.2.2 Prepare 45mL 1.1% formaldehyde in PBS. Resuspend cell pellet in 5mL PBS. Add 45mL 1.1% formaldehyde in PBS, and invert to mix (*see* **Note 1**). Cross-link cells by incubating at room temperature for 10min with gentle end-to-end shaking.

3.2.3 Quench formaldehyde with 5mL 1.25M glycine and incubate at room temperature for 5min with gentle end-to-end shaking.

3.2.4 Centrifuge cells at ~3000rpm for 10min at 4ºC. Aspirate supernatant, and wash with 50mL cold PBS. Repeat spin and aspirate PBS from cross-linked cell pellet.

3.2.5 Resuspend pellet well in 2mL Swelling Buffer, split into 500µL aliquots in 1.5mL tubes, and incubate on ice 10min. Dounce for 2sec with a hand-held douncer fitted with disposable probes that fit 1.5mL tubes. This disrupts the cell membrane and enriches for cell nuclei. In a bench-top mini centrifuge, spin tubes at 2500rcf for 5min at 4ºC to pellet. Carefully aspirate supernatant.

3.2.6 Resuspend nuclei in a total of 2.5mL PBS and pool in a 50mL tube. Add 22.5mL 3.3% formaldehyde in PBS and invert to mix. Cross-link cells by incubating at room temperature for 30min with gentle end-to-end shaking.

3.2.7 Quench formaldehyde with 2.5mL 1.25M glycine and incubate at room temperature for 5min with gentle end-to-end shaking.

3.2.8 Centrifuge nuclei at ~3000rpm for 10min at 4ºC. Aspirate supernatant, and wash with 50mL cold PBS. Repeat spin and aspirate PBS from cross-linked nuclear pellet. Resuspend in 2mL PBS and prepare 1mL aliquots in 1.5mL tubes. Spin in a bench-top mini centrifuge at 2500rcf to pellet, and aspirate PBS. Pelleted, cross-linked nuclei can be flash-frozen in liquid nitrogen or dry ice and stored at -80ºC, or further processed for nuclear lysis and sonication (see **3.2.9**).

3.2.9 Resuspend cross-linked nuclei (pellet should have a mass of approximately 100mg) in 2mL freshly supplemented Nuclear Lysis Buffer. Break up the pellet by pipetting up and down until fully resuspended and smooth.

3.2.10 Solubilize and shear the biological material by sonication (e.g. using a Bioruptor or Covaris sonicator) in a 4ºC waterbath (*see* **Note 3**). Sonicate the material until the bulk of nucleic acids are sheared to 150-600bp (*see* **Note 4**).

3.2.11 Pre-clear the sonicated chromatin by spinning at maximum speed in a bench-top mini centrifuge for 10min at 4ºC. Remove the supernatant and place in a fresh 1.5mL tube. The insoluble pellet should be small; discard. The sonicated chromatin can be flash-frozen in liquid nitrogen or dry ice and stored at -80ºC, or used immediately for dChIRP (see **3.3**).

3.3 Hybridize oligos and purify RNA targets

3.3.1 At room temperature, thaw 1mL of chromatin for each sample (*see* **Note 5**). Thaw 100µM ChIRP oligo pools.

3.3.2 In addition to the 1mL of chromatin for the dChIRP experiment, set aside 10µL of chromatin for RNA Input, 10µL of chromatin for DNA Input, and 50µL of chromatin for Protein Input. These input samples are used to later quantify RNA, protein, and DNA recovery.

3.3.3 Prepare 2mL of Hybridization Buffer for each 1mL of chromatin.

3.3.4 Transfer 1mL chromatin to a 15mL Falcon tube. Add 2mL Hybridization Buffer and 1µL 100µM oligo pool. Invert to mix.

3.3.5 Gently rotate end-to-end at 37ºC for 4hrs in a hybridization oven (*see* **Note 6**).

3.3.6 With 30min remaining for the hybridization, prepare the magnetic beads. Resuspend the stock solution of beads. Use 100µL beads for each 1mL of chromatin. Add each 100µL of beads to a new 1.5mL tube, place on magnetic strip for 1min to separate beads, and carefully aspirate the buffer. Remove from magnetic strip and wash in 1mL Nuclear Lysis Buffer. Repeat for a total of three washes. Remove all buffer. Resuspend in 100µL Nuclear Lysis Buffer and pool together.

3.3.7 When hybridization is complete, add 100µL washed beads to the 15mL tubes. Invert to mix. Return to hybridization oven with gentle end-to-end shaking at 37ºC for 30min.

3.3.8 Meanwhile, prepare fresh Wash Buffer (5mL per 1mL of chromatin). Prewarm to 37ºC before use (*see* **Note 7**).

3.3.9 When bead binding is complete, remove 15mL tubes, collect briefly by short centrifugation, and separate beads on a 15mL tube-capacity magnetic strip for 2min. Aspirate buffer, and remove 15mL tube from magnetic strip. Add 1mL Wash Buffer and pipette up and down to completely resuspend beads. Transfer to clean, labeled 1.5mL tubes and gently rotate end-to-end at 37ºC for 5min. Repeat for a total of five washes, using 1.5mL tube-capacity magnetic strip and collecting sample by short centrifugation between each wash.

3.3.10 For the last wash, remove all buffer and resuspend the beads well in 1mL of Wash Buffer. Remove 10µL beads (1% total sample) and place in a new tube. This sample will be used for analyzing the RNA fraction (*see* **3.4**). The remaining ~1mL of sample will be used for analyzing either the DNA (*see* **3.5**) or Protein fraction (*see* **3.6**). Place all samples on magnetic strip to separate beads and remove all buffer.

3.4 Isolate RNA fraction

3.4.1 Take 10µL RNA Input (*see* **3.3.2**) and RNA dChIRP sample (beads from 10µL, *see* **3.3.10**). Add 90µL and 100µL Proteinase K Buffer (supplemented with fresh Proteinase K) to the Input and dChIRP sample beads, respectively. Incubate at 50ºC for 45min with shaking.

3.4.2 Collect the sample by short centrifugation. Heat samples at 95ºC for 10min to denature the streptavidin beads. Adding a cap clamp will prevent the tubes from popping open.

3.4.3 Cool the sample on ice briefly. Add 500µL TRIzol and vortex vigorously for 10sec. Incubate at room temperature for 10min. Add 100µL chloroform and vortex vigorously for 10sec. Spin at maximum speed on a bench-top mini centrifuge for 15min at 4ºC. The aqueous phase (clear) will separate from the organic (pink), possibly with a white interface.

3.4.4 Carefully remove the aqueous phase (~400µL), avoiding the organic phase and interface, and add to a new labeled tube.

3.4.5 Add 1µL GlycoBlue and vortex. Add 400µL (1 volume) of isopropanol and vortex. Precipitate RNA overnight at -20ºC.

3.4.6 The next day, carefully place the tubes in a bench-top mini centrifuge and spin the precipitated RNA at maximum speed for 1hr at 4ºC. The RNA and glycogen will form a small blue pellet. Carefully aspirate the supernatant and wash in 1mL 70% ethanol. Vortex briefly, and spin for 5min. Carefully remove all residual ethanol and let the pellet air dry for 5min.

3.4.7 Add 26µL nuclease-free H_2O and fully dissolve by vortexing and/or heating at 37ºC. Add 3µL 10X TURBO DNase Buffer and 1µL TURBO DNase. Incubate at 37ºC for 30min.

3.4.8 Add 70µL nuclease-free H_2O to bring to 100µL, and mix. Purify RNA on RNeasy mini column, as per manufacturer's protocol. Elute in 50µL nuclease-free H_2O. Store purified dChIRP RNA at -80ºC or proceed to analysis by RT-qPCR (*see* **3.7**).

3.5 Isolate Protein fraction

3.5.1 Take 50µL Protein Input (*see* **3.3.2**) and beads from ~1mL dChIRP sample (see **3.3.10**). Add 50µL 2x Protein Elution Buffer to Protein Input and mix; resuspend dChIRP sample beads in 50µL 1x Protein Elution Buffer.

3.5.2 Heat Protein Input and samples at 95ºC for 15min. Adding a cap clamp will prevent the tubes from popping open.

3.5.3 Cool tubes on ice briefly. Collect sample by short centrifugation. Separate beads from dChIRP samples on magnetic strip. Remove eluate and add to new, labeled tube.

3.5.4 Resuspend the dChIRP sample beads in 50µL 1x Protein Elution Buffer. Heat Input, dChIRP protein eluate, and dChIRP sample beads at 95ºC for 15min more.

3.5.5 Repeat Step **3.5.3**, pooling the first and second eluates for each dChIRP sample. This results in 100µL Protein Input and 100µL dChIRP Protein samples. Store the purified protein at -20ºC or proceed to analysis by immunoblotting (*see* **3.8**).

3.6 Isolate DNA fraction

3.6.1 Prepare 300µL fresh DNA Elution Buffer per DNA sample, supplemented with RNases.

3.6.2 Take 10µL DNA Input (*see* **3.3.2**) and beads from ~1mL dChIRP sample (*see* **3.3.10**). Add 140µL DNA Elution Buffer to Input and mix; resuspend dChIRP sample beads in 150µL DNA Elution Buffer. Incubate at 37ºC for 30min with gentle shaking.

3.6.3 Separate beads from dChIRP samples on magnetic strip. Collect eluate and add to new, labeled tubes.

3.6.4 Add 150µL DNA Elution Buffer to Input and beads. Repeat incubation at 37ºC for 30min with gentle shaking.

3.6.5 Separate beads from dChIRP samples on magnetic strip. Collect eluate and pool with first eluate. This results in 300µL Input and 300µL dChIRP DNA samples.

3.6.6 Add 15µL Proteinase K to each tube and incubate at 50ºC for 45min.

3.6.7 Pre-spin phase-lock columns at max speed on a bench-top mini centrifuge. Transfer DNA samples to tubes, add 300µL phenol:chloroform:isoamyl alcohol, and vortex vigorously for 5min. Spin at max speed on a bench-top mini centrifuge for 5min at 4ºC. Remove aqueous phase (~300µL, it will separate above the gel plug) and add to new labeled tube.

3.6.8 Add 3µL GlycoBlue and 30µL 3M NaOAc, and vortex to mix. Add 900µL 100% ethanol and vortex to mix. Precipitate DNA overnight at -20ºC.

3.6.9 The next day, spin the precipitated DNA at maximum speed on a bench-top mini centrifuge for 1hr at 4ºC. The DNA and glycogen will form a small blue pellet. Carefully aspirate the supernatant and wash in 1mL 70% ethanol. Vortex briefly, and spin for 5min. Carefully remove all residual ethanol and let the pellet air dry for 5min.

3.6.10 Add 50µL nuclease-free H_2O and fully dissolve the pellet by vortexing and/or heating at 37ºC. Store purified dChIRP DNA at -20ºC or proceed to analysis by qPCR (*see* **3.9**) or sequencing library preparation (*see* **3.10**).

3.7 Analyze co-recovered RNA by RT-qPCR

3.7.1 Design RT-qPCR primers to analyze RNA recovery and co-recovery by dChIRP. Pick RT-qPCR primers that produce an amplicon within each targeted domain of the RNA of interest. Additionally, if other RNA species are expected to interact with the RNA of interest, design RT-qPCR primers against those as well. As a negative control, design primers against an abundant, non-target RNA, such as GAPDH mRNA. (*see* **Note 8**)

3.7.2 Order the RT-qPCR primers and test them for linear and specific amplification of the intended amplicon. (*see* **Note 9**)

3.7.3 With validated RT-qPCR primers against the target RNA domains and a negative control RNA, perform RT-qPCR using the dChIRP RNA samples and Input. Produce a standard curve dilution of the Input (100%, 10%, 1%) against which the dChIRP samples can be absolutely quantified. For a given dChIRP RNA sample, the most abundant RNA fragment identified should be the targeted domain of the RNA of interest. The negative control RNA

should be significantly de-enriched (**Fig.2B**). RNA co-recovery can be used to interpret RNA domain-level architecture (**Fig.2C**).

3.8 Analyze co-recovered protein by immunoblotting

3.8.1 Interaction between the RNA's domains and suspected RNA-binding proteins can be interrogated by immunoblotting the dChIRP-recovered protein fraction. Select primary antibodies against suspected protein partners and validate using the Input fraction; antibodies against a negative control protein, such as actin or tubulin, can be used to demonstrate the selectivity of protein co-recovery (*see* **Note 10**).

3.8.2 Thaw the dChIRP-recovered protein samples and input at room temperature. If a precipitate forms, heat the samples to 65ºC for 5min, vortex, and repeat until dissolved. Run the samples on PAGE electrophoresis using a protein ladder to resolve sizes. Transfer the gel to a nitrocellulose membrane, and perform washing, blocking, incubation with primary antibody, and incubation with secondary antibody according to specifications for the given primary/secondary antibodies used. Develop and image protein bands (e.g. by chemi-luminescence, if using HRP-conjugated secondary antibody).

3.8.3 Protein recovery is semi-quantitatively indicative of RNA domain-specific protein interaction. See **Fig.2D** for examples of domain-specific protein interaction resolved by dChIRP.

3.9 Analyze co-recovered DNA by qPCR

3.9.1 Design qPCR primers to analyze DNA co-recovery by dChIRP from **3.6.10**. Pick qPCR primers that produce a single amplicon within known or suspected interaction sites in the genome. As a negative control, design primers against genomic regions known or suspected to not interact with the RNA of interest. (*see* **Note 11**)

3.9.2 Order the qPCR primers and test them for linear and specific amplification of the intended amplicon. (*see* **Note 12**)

3.9.3 With validated qPCR primers against genomic loci, perform qPCR using the dChIRP DNA samples and Input. Produce a standard curve dilution of the Input (100%, 10%, 1%) against which the dChIRP samples can be absolutely quantified. The negative control DNA locus should be significantly de-enriched (**Fig.2E**).

3.10 Analyze co-recovered DNA by sequencing

3.10.1 For dChIRP-sequencing experiments, it is necessary to use two independent oligo pools (e.g. even and odd) per RNA domain. This is because each oligo in a given pool may hybridize to off-target RNAs or genomic loci with sufficient complementarity. By taking two independent measurements of genomic occupancy, these oligo-specific noises can be eliminated to produce a "merged" dataset with low noise. As such, each dChIRP-sequencing experiment requires two oligo pools per targeted domain, as well as an input (*see* **3.1.7**). It is further useful to use negative

controls in dChIRP-sequencing experiments, such as sense oligos, oligos against a negative control RNA (e.g. *LacZ*), dChIRP in RNase-treated chromatin, or dChIRP in a cell line that does not express the target RNA.

3.10.2 Take recovered DNA (*see* **3.6.10**). Follow the manufacturer's manual for DNA sequencing library preparation (e.g. NEBNext ChIP-seq) for end repair, dA-tailing, adapter ligation and barcoding, size selection, amplification, and purification. Quantify libraries by standard methods (e.g. qPCR, Bioanalyzer, absorbance), and submit for high-throughput sequencing (e.g. 50-cycle single-end sequencing).

3.10.3 When sequencing is complete, run the ChIRP-seq processing pipeline, starting from raw .fastq-files (available for download at: changlab.stanford.edu/protocols). The pipeline separately maps reads from the even, odd, and input samples; the mapped reads from each sample are normalized to the total number of reads, the even and odd samples are merged, and peaks are called using MACS (Bailey et al., 1994). Export the resulting bigwig files and visualize the sequencing tracks and called peaks in a genome browser application (e.g. UCSC Genome Broswer, genome.ucsc.edu; **Fig.2F**). Further genome-wide bioinformatics analyses can be performed from these data sets, such as gene ontology analysis or motif discovery.

4. Notes

1. Depending on the nature of the RNA-chromatin interaction, optimizing the cross-linking agent or conditions can improve RNA, protein, and DNA recovery by dChIRP. The two-step, 1%+3% formaldehyde conditions discussed here were optimized for dChIRP of the roX RNAs. As an alternative cross-linker, glutaraldehyde is a five-carbon agent that is not thermally reversible, making longer and more stable cross-links; however glutaraldehyde cross-linked proteins are not amenable to electrophoresis.
2. 20 million cells are typically sufficient for one dChIRP sample. This number will vary depending on the RNA copy number, cell size, and genome size.
3. We have had success with both Bioruptor (Diagenode) and E-Series Focused Ultrasonicator (Covaris), both of which are coupled to waterbaths that keep the sample temperature at 4ºC. It is important that the sample not heat during shearing, as this can reverse formaldehyde cross-links and disrupt biomolecular interactions. Due to the automation of the Covaris sonicator, chromatin shearing is more consistent sample-to-sample and the throughput is higher.
4. Sonication time depends on the cross-linking agent used, the cell type, and the sonicator and may take several minutes to several hours. Titrate sonication time using one aliquot of cross-linked nuclei, withdrawing 10µL lysate at regular intervals for several hours. Add 90µL Proteinase K Buffer supplemented with Proteinase K and incubate at 50ºC for 45min. Extract DNA using a PCR purification kit (e.g. Qiagen) and elute in 20µL nuclease-free H_2O. Check DNA sizes on a 1% agarose gel. The bulk of the nucleic acids should be between 150-600bp. Select a shearing time that produces such fragments. Sonication time is

important to optimize, as this process shears the RNA fragments and thus allows the domain-specific recovery of the target RNA. This is also important for shearing the DNA to sequence-amenable fragments.

5. 1mL of chromatin is sufficient for one dChIRP sample (i.e. one pool of oligos) for isolating RNA and protein fractions *or* RNA and DNA fractions. To isolate RNA, protein, *and* DNA fractions, use 2mL of chromatin per sample.
6. Ensure that the caps are fastened tightly, as the SDS in the Hybridization Buffer can cause leaking.
7. Add PMSF immediately prior to washes, as PMSF has a short half-life in aqueous solution.
8. RT-qPCR amplicons should be between 50-150bp in length, so as to be below the fragment size for RNA shearing. Ideally, amplicons should fall within each region of the RNA that is tiled by oligos within an oligo pool (rather than between them). Design 2-3 primer pairs for each region, as many RT-qPCR primer pairs will fail to amplify the target amplicon specifically and/or linearly.
9. Discard RT-qPCR primers that produce more than one product. Validate the RT-qPCR primers using a dilution series (100%, 10%, 1%) of the Input RNA and confirm linear amplification by calculating the $\Delta C_{t,100\text{-}10}$ and $\Delta C_{t,10\text{-}1}$. For optimal primers, $\Delta C_t = 3.322$.
10. Glutaraldehyde cross-linked protein fractions are not amenable to PAGE, but can be used for analysis by immuno-dot-blot (*see* Chu et al., 2011).
11. qPCR amplicons should be between 50-150bp in length, so as to be below the fragment size for DNA shearing. These sites can be determined from previous ChIRP-seq or ChIP-seq data or known binding sites.
12. Discard qPCR primers that produce more than one product. Validate the qPCR primers using a dilution series (100%, 10%, 1%) of the Input DNA and confirm linear amplification by calculating the $\Delta C_{t,100\text{-}10}$ and $\Delta C_{t,10\text{-}1}$. For optimal primers, $\Delta C_t = 3.322$.

5. References

Bailey, T.L. and Elkan, C. (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc Int Conf Intell Syst Mol Biol* 2, 28–36.

Cabili, M.N., Trapnell, C., Goff, L., Koziol, M., Tazon-Vega, B., Regev, A., and Rinn, J.L. (2011) Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev* 25, 1915-1927.

Chu, C., Zhong, F. L., Artandi, S. E. and Chang, H. Y. (2011) Genomic Maps of Long Noncoding RNA Occupancy Reveal Principles of RNA-Chromatin Interactions. *Mol Cell* 44, 1-12.

Chu, C., Quinn, J.J., and Chang, H. Y. (2012) Chromatin Isolation by RNA Purification (ChIRP). *J Visualized Experiments* 61, e3912.

Colak, D., Zaninovic, N., Cohen, M.S., Rosenwaks, Z., Yang, W.Y., Gerhardt, J., Disney, M.D., and Jaffrey, S.R. (2014) Promoter-bound trinucleotide repeat mRNA drives epigenetic silencing in fragile X syndrome. *Science* 343, 1002–1005.

Conrad, T. and Akhtar, A. (2012) Dosage compensation in *Drosophila melanogaster*: epigenetic fine-tuning of chromosome-wide transcription. *Nat Rev Genet* 13, 123–134.

Guttman, M. and Rinn, J.L. (2012) Modular regulatory principles of large non-coding RNAs. *Nature* 482, 339-346.

Helwak, A., Kudla, G., Dudnakova, T. and Tollervey, D. (2013) Mapping the human miRNA interactome by CLASH reveals frequent non-canonical binding. *Cell* 153, 654–665.

Ilik, I.A., Quinn, J.J., Georgiev, P., Tavares-Cadete, F., Maticzka, D., Toscano, S., Wan, Y., Spitale, R.C., Luscombe, N., Backofen, R., Chang, H.Y., and Akhtar, A. (2013) Tandem stem-loops in

roX RNAs act together to mediate X chromosome dosage compensation in *Drosophila*. Mol Cell 51, 156-173.

Quinn, J. J., Ilik, I. A., Qu, K., Georgiev, P., Chu, C., Ahktar, A., and Chang, H. Y. (2014) Revealing long noncoding RNA architecture and functions using domain-specific chromatin isolation by RNA purification. *Nat Biotech* X, xxx.

Rinn, J.L., Chang, H.Y. (2012) Genome regulation by long noncoding RNAs. *Annu Rev* 81, 145-166.

Ule, J., Jensen, K., Mele, A. and Darnell, R.B. (2005) CLIP: a method for identifying protein- RNA interaction sites in living cells. *Methods* 37, 376–386.

Li, Z., Chao, T.-C., Chang, K.-Y., Lin, N., Patil, V.S., Shimizu, C., Head, S.R., Burns, J.C., and Rana, T.M. (2013) The long noncoding RNA *THRIL* regulates TNFa expression through its interaction with hnRNPL. *Proc Natl Acad Sci U S A* 111, 1002-1007.

Figure 1. dChIRP workflow

(A) dChIRP oligo pool design strategy. Biotinylated antisense oligos are designed to tile specific domains of the target RNA.

(B) dChIRP chromatin preparation and affinity purification method. Nuclei are cross-linked, preserving native nucleic acid-protein interactions. The nuclei are then solubilized and sheared by sonication to fragment nucleic acids. Next, pools of domain-specific antisense oligos are added to the sheared chromatin, allowed to hybridize to the RNA target under stringent conditions, and affinity purified on streptavidin-beads.

(C) dChIRP outputs. The RNA, protein, and DNA fractions from the co-purified samples are extracted and analyzed, revealing target RNA-RNA, RNA-protein, and RNA-DNA interactions. Figure adapted from Quinn et al., 2014.

Figure 2. Examples of dChIRP results

(A) Oligo pool design strategy for roX1. roX1 RNA was divided into six domains, and twelve antisense oligos were designed against each domain. The domains determined from previous individual-nucleotide resolution CLIP experiments, which suggested that domains D1, D2, and D3 contact interacting proteins whereas domains U1, U2, and U3 do not (Ilik et al., 2013).

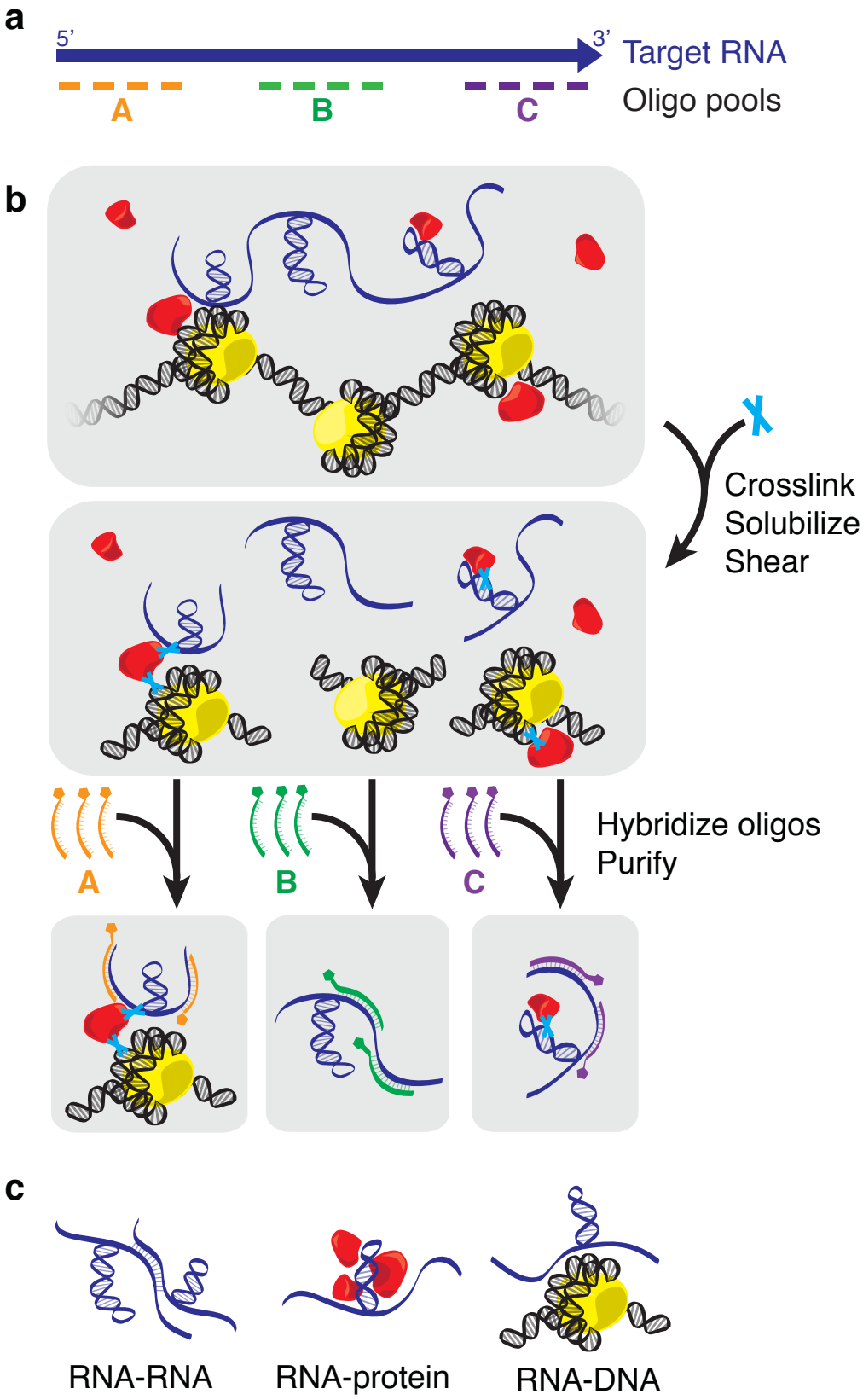
(B) roX1-specific RNA purification by dChIRP. The RNA fraction from dChIRP was extracted and analyzed by RT-qPCR, using primers against each of the six domains of roX1 and a negative control mRNA, GAPDH. roX1 dChIRP specifically enriches for roX1 RNA while de-enriching for GAPDH. ChIRP using LacZ oligos fails to enrich for roX1 RNA over GAPDH.

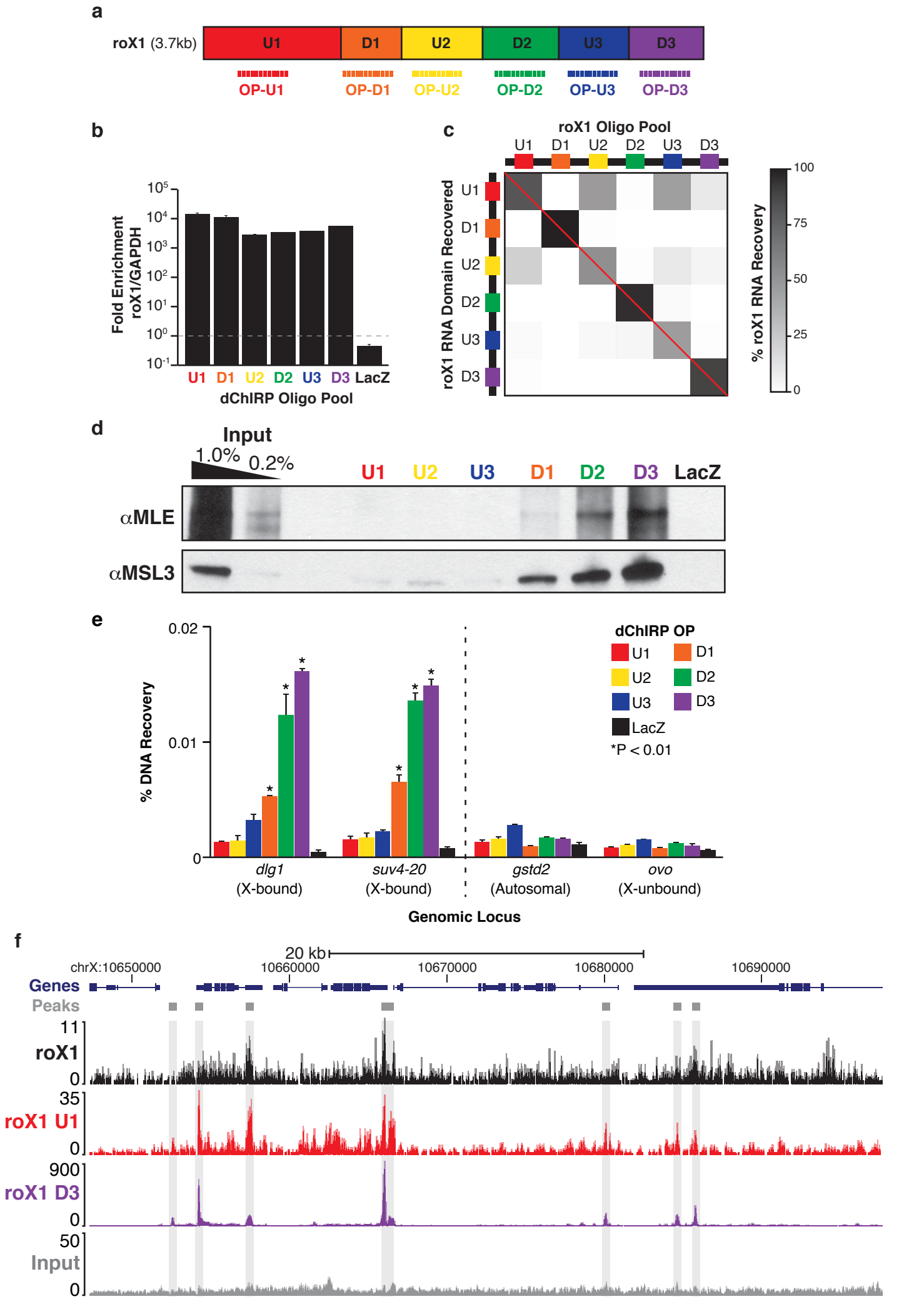
(C) roX1 RNA domain co-recovery. RNA domain recovery was predominantly domain-specific (along the red diagonal); however dChIRP of the U domains also co-recovered other U domains (signal off the red diagonal). This U domain co-recovery suggests that the U domains interact in the native ribonucleoprotein complex, and that the D domains are independent.

(D) roX1 dChIRP protein co-recovery. The protein fraction was extracted from roX1 dChIRP samples and subjected to polyacrylamide gel electrophoresis and immunoblotting with antibodies against MLE and MSL3, proteins known to interact with roX1. dChIRP of domains D1, D2, and D3 recovered MLE and MSL3, providing further evidence that these domains exclusively interact with protein. No actin was co-recovered (*see* Quinn et al., 2014; data not shown).

(E) roX1 dChIRP DNA co-recovery. The DNA fraction was extracted from roX1 dChIRP samples and analyzed by qPCR using primers against known genomic binding sites of the dosage compensation complex (*dlg1* and *suv4-20*) and two negative control loci (*gstd2* and *ovo*). roX1 dChIRP of domains D1, D2, and D3 significantly enrich for *dlg1* and *suv4-20* DNA over negative control loci, suggesting that these domains are closely associated with the chromatin at dosage compensated loci.

(F) roX1 genomic binding sites revealed by dChIRP-sequencing at a representative window on the X chromosome. Traditional ChIRP-sequencing of roX1 (black), dChIRP-sequencing of roX1 domains U1 (red) and D3 (purple), and input DNA (gray) genomic browser tracks illustrate the binding pattern of roX1 RNA on the X chromosome. The improvement of sequencing signal achievable by dChIRP is illustrated here by the increased signal in the dChIRP-sequencing tracks relative to traditional ChIRP. Furthermore, dChIRP of domain D3 has the highest signal because this domain is the most closely associated with chromatin. Figure adapted from Quinn et al., 2014.


a
5'
3'
Target RNA
A
B
C
Oligo pools
b
Crosslink
Solubilize
Shear
A
B
C
Hybridize oligos
Purify
c
RNA-RNA
RNA-protein
RNA-DNA


a
roX1 (3.7kb)
U1
D1
U2
D2
U3
D3
OP-U1
OP-D1
OP-U2
OP-D2
OP-U3
OP-D3
b
Fold Enrichment roX1/GAPDH
dChIRP Oligo Pool
U1 D1 U2 D2 U3 D3 LacZ
c
roX1 Oligo Pool
roX1 RNA Domain Recovered
% roX1 RNA Recovery
d
Input
1.0% 0.2%
αMLE
αMSL3
e
% DNA Recovery
dChIRP OP
*P < 0.01
dlg1 (X-bound)
suv4-20 (X-bound)
gstd2 (Autosomal)
ovo (X-unbound)
Genomic Locus
f
20 kb
chrX:10650000
10660000
10670000
10680000
10690000
Genes
Peaks
roX1
roX1 U1
roX1 D3
Input