Development of Oligonucleotide-directed proximity-interactome MAPping (O-MAP), for characterizing RNA-protein interactions and higher order subnuclear architecture *in situ*

Ashley Tsue

A dissertation

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2023

Reading Committee:

David Shechner, Chair

Shao-En Ong

Ning Zheng

Program Authorized to Offer Degree:

Pharmacology

University of Washington

ABSTRACT

Development of Oligonucleotide-directed proximity-interactome MAPping (O-MAP), for characterizing RNA-protein interactions and higher order subnuclear architecture *in situ*

Ashley Tsue

Chair of the Supervisory Committee:
David Shechner
Department of Pharmacology

RNA-protein interactions underscore a broad array of regulatory mechanisms throughout biology, and dysregulation of these interactions contributes to a broad range of human pathologies such as neurodegenerative disorders and many cancers[1–3]. Recent studies have discovered thousands of RNA-binding proteins that lack canonical RNA binding domains[4,5]. Additionally, the human transcriptome harbors tens of thousands of novel RNAs that can be differentially regulated in disease[6]. RNA-protein complexes, termed ribonucleoproteins (RNPs), can also serve as structural scaffolds for reorganizing local subcellular environments and assembling subnuclear bodies[7]. A classic example of RNA-scaffolded structures is the nucleolus, a subnuclear body that serves as the site of ribosome biogenesis, cell cycle regulation, and stress responses[8–10]. Nucleolar assembly is initiated de novo after each cell division—nucleated by pre-ribosomal RNA (pre-rRNA) and other nucleolar factors that are preserved during the disassembly of the nucleolus at the start of mitosis[11]. This formation is thought to be driven by RNA-protein interactions that phase separate into the nucleolar tripartite structure[12]. Other RNAs, such as X inactive specific transcript (XIST) and MALAT1, are thought to have an architectural role in organizing subnuclear Barr bodies and nuclear speckles for gene regulation. However, studying specific RNA and their protein interactomes is challenging by conventional methods. Most approaches for analyzing the composition of a subcellular structure rely on the biochemical purification of that structure[13] or enriching for interacting proteins using antisense oligonucleotides of the target RNA bound to a resin[14]. These approaches are either impossible for smaller structures or are plagued with nonspecific binding and low specificity after exposure to crude lysate[15]. To address these problems, recent work has utilized proximity labeling as a way to study specific RNA-protein interactomes within the subcellular context; however, few methods exist to specifically target and enrich for RNA scaffolded subnuclear bodies.

Utilizing existing RNA-fluorescence in situ hybridization (FISH) technologies[16] and proximity labeling, this thesis focuses on the development of Oligonucleotide-directed proximity-interactome (O-MAP) to target specific RNAs within their subcellular context and probe for interacting proteins[17]. This technique targets antisense oligonucleotides to a RNA of interest in fixed cells and recruits a proximity labeling enzyme, horse radish peroxidase (HRP), to the target RNA enabling the promiscuous biotinylation of nearby proteins. These biotinylated proteins can be isolated by streptavidin enrichment and analyzed by downstream omics such as mass spectrometry (MS). Using the 47S pre-ribosomal RNA, long noncoding RNA XIST and 7SK small nuclear RNA as models, O-MAP induces precise biotinylation of target RNA and nearby proteins that can then be systematically analyzed by mass spectrometry. O-MAP-MS at these RNA targets characterized these RNA scaffolded compartments and interacting proteomes, highlighting classes of proteins involved in nucleolar biology and nuclear speckles. Without the need for genetic manipulation, O-MAP is both easily portable to other cell lines, organoids, tissues as well as RNA targets of varying expression and lengths. Additionally, O-MAP can probe differentially-regulated proteomes in disease-relevant states including pancreatic cancer and cellular stress. These studies demonstrate the versatility and specificity of O-MAP as well as its potential to provide new characterizations in RNA interactome biology.

ACKNOWLEDGEMENTS

I would like to thank my PhD advisor, David, for his support and guidance over my graduate career. The experiences, mentorship and thoughtful opportunities he has given me has allowed me to grow in this environment and in developing the edge of RNA technologies. I also want to thank the other current and past members of our Shechner lab including Jessican Gianopulos, Chris Hsu, Caitlin Anderson, Diana Lei, and Hayden Gizinski. These wonderful lab mates have created a great and growing community for our lab. I also want to thank the members of my thesis committee over the years including John Scott, Ning Zheng, Brian Beliveau, Shao-En Ong, and Richard Gardner for their guidance and contributions to shaping O-MAP. A special thank you to the members of the Pharmacology Department, especially Debbie Bale for helping me through navigating my graduate school experience. A huge thank you to my graduate school friends, especially our year Kimya, Nicole, Simar and Melissa, for the beautiful memories throughout graduate school. Thank you to all of my family, Debbie, Terry, Grandpa Ted, Grandma Frances, Grandpa Norman, and Trevor, for your unconditional love and support throughout this time. They are always present to offer guidance and help through each chapter of my life. I want to also thank my family friends Dan Welch, the Uhlig family, and the Ruelles who are always a steadfast source of support and who have walked alongside our family for many years. I am truly grateful for all of our adventures and experiences that color my life. To my closest friends, Thomas Dudey, Dylan Altschuler, Broderic Bender and Shane Wilson, thank you for your continued support even from afar. I especially want to thank Evan Kania for taking on the challenge of developing O-MAP and his continued support throughout graduate school. I am truly grateful for your drive, experience and optimistic deposition that drives our lab and our experiences in graduate school. I sincerely thankful for your friendship.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 2 FIGURES

CHAPTER 3 FIGURES

PREFACE

Portions of the text, figures and data from this dissertation were reprinted from the following works under fair use and/or with permission CC-BY License:

1. Tsue, A. F., Kania E. E. *et al.* Oligonucleotide-directed proximity-interactome mapping (O-MAP): A unified method for discovering RNA-interacting proteins, transcripts and genomic loci in situ. *bioRxiv* 2023.01.19.524825 (2023).

1.1 INTRODUCTION

RNA carries out a broad range of biological functions from translating genetic information to regulating genes and organizing higher order structures. Aside from messenger RNA, other RNA classes are required for synthesizing functional gene products including ribosomal RNA and transfer RNAs. Furthermore, classes of noncoding RNAs extend beyond synthesizing proteins and are involved in regulating gene expression and cellular responses to stress[18]. These RNA are not acting alone, and they form transient and spatially regulated interactions with other RNA, proteins and DNA. On a larger level, these interactions integrate function with the assembly of higher order structures within the nucleus primarily to regulate transcription[4,19]. Recent studies support this architectural role of RNA and its interactions through the process of liquid-liquid phase separation (LLPS) to form non-membrane bound compartments that nucleate reversible and dynamic condensates[19–21].

A classical example of a liquid-liquid phase separated body is the nucleolus, the site of ribosome biogenesis. This membraneless subnuclear body nucleates around pre ribosomal RNA to form a tripartite structure, with each compartment dictating different roles of RNA polymerase I transcription, processing, and the assembly of ribosome RNA and protein complexes[22]. Due to its liquid-liquid phase separated nature, the nucleolus can also have dynamic responses to environmental stimuli such as heat shock or DNA damage[23–25]. Other intracellular condensates that regulate transcription and cellular stress responses include transcriptional condensates, MALAT1, NEAT1 and XIST[26,27]. Currently, there are limited techniques that can be used to characterize these RNA scaffolded intracellular condensates within the cell, and most of these membraneless condensates cannot be biochemically fractionated or affinity purified. Thus, there is a need to develop RNA-centric technologies to elucidate these higher order RNA-protein condensates within their subcellular context.

1.2 RNA SCAFFOLDED INTERACTIONS

Phase separation

Cellular bodies can play an essential role in compartmentalizing biochemical reactions using phase separation. Examples of these biomolecular condensates include nucleoli, Cajal bodies and PML nuclear bodies[28]. These cellular compartments lack physical membranes like the lipid bilayer present in the endoplasmic reticulum. They stably maintain structure, assembly, and diverse composition while still

being able to rapidly fuse to similar bodies and exchange with the surrounding environment[29,30]. These protein, RNA and DNA maintain in solution until they reach their solubility limit and phase separate creating a chemical equilibrium between compartments of different chemical properties[19]. Dysregulation of LLPS can lead to aberrant condensate and amyloid formation causing neurodegeneration and cancer[31].

Regulation of the biomolecular condensate composition and formation is still being studied. Recent work supports the idea of RNA serving as a scaffold, wherein resident molecules (e.g. DNA, RNA, and proteins) drive the formation of a higher order structure that can transiently recruit other components and form a phase separated body capable of biological function. For example, the NEAT1 non-coding RNA dictates the formation of paraspeckles in the nucleus to retain adenosine-to-inosine edited mRNAs[32,33]. Additionally, processing bodies (P-bodies) are scaffolded around mRNAs in complex with proteins associated with translation repression and mRNA decay intermediates[34]. Recruitment of interacting partners in these condensates to the scaffold may occur through specific binding interactions, multivalency, general electrostatic properties and intrinsically disordered regions (IDR)[35,36]. However, studying the composition of these bimolecular condensates is difficult within the native cellular context because they cannot be biochemically purified and their components can rapidly exchange outside of the condensate.

RNA and nuclear organization

The concept of scaffolding and recruitment in biomolecules condensates also underscores the role of RNAs in regulating higher order organization and nuclear architecture. For instance, in mammals, the X-inactive specific transcript (XIST) is a long non-coding RNA that forms a repressive nuclear compartment essential for silencing and repositioning the X chromosome to the periphery of the nucleus in females[37]. In this X chromosome inactivation (XCI), XIST first localizes to the genomic sites near its own transcription locus before coating the X chromosome for silencing through the recruitment of chromatin-modifying proteins, transcriptional silencers and other RNA-binding proteins[38,39]. XIST foci remain in these local protein compartments that undergo a rapid exchange of XIST-interacting proteins facilitated by the protein SPEN, which ultimately supports protein crowding that silences a large number of genes across the chromosome[38].

Other noncoding RNAs play important roles in gene regulation through nuclear organization. Studies have found that the long noncoding RNA MALAT1 is an essential structural component of paraspeckles, influencing alternative splicing through phosphorylation regulation and splicing factor localization[40]. Additionally, small nuclear 7SK RNA is enriched in nuclear speckles along with pre-mRNA splicing factors to regulate RNA polymerase II (RNA pol II) transcription by displacing P-TEFb proteins from transcription sites[41,42]. Downregulation of 7SK leads to mislocalization of the nuclear speckle components[43]. These RNA are able to recruit other biomolecules to spatially regulate and reorganize local neighborhoods within the nucleus.

Nucleolus structure and regulation

Present in all eukaryotic cells, the nucleolus has an important role in many cellular functions including ribosome biogenesis, stress response, and cancer development[44]. This membraneless organelle is the site of ribosome biogenesis, producing about 2 million ribosomes during 15 hours utilizing 80% of cellular ATP[45]. Briefly, the initiation of ribosomal DNA (rDNA) transcription occurs at the boundaries of the fibrillar center (FC) and dense fibrillar component (DFC) before the pre-ribosomal RNA (pre-rRNA) is processed in the DFC and granular component (GC). This tripartite structure of the nucleolus is disrupted upon interruption of rDNA transcription, especially during the eukaryotic cell cycle. Nucleoli are dissembled at the beginning of mitosis (i.e. prophase a) and then reassembled around the recruitment of pre-rRNAs and other processing factors at the end of mitosis[11,46]. Along with in vitro assays[12,47], this natural formation of the membranelles organelle nucleated by pre-ribosomal RNA and other nucleolar components contribute to the idea of the nucleolus as a phase-separated body[22].

Aside from cell cycle regulation and ribosome production, the nucleolus also has a dynamic role in stress response. Upon transcriptional stress or DNA damage induced by ionizing radiation (IR) or RNA polymerase I inhibition, nucleolar segregation occurs to disrupt the tripartite structure often forming nucleolar caps[48]. Nucleolar disruption causes the release of nucleolar p53 and other ribosomal proteins to bind to MDM2, activating a p53-dependent cellular stress response that can lead to cell cycle arrest and apoptosis[49]. In some cases of DNA damage and heat shock, translocation to the nucleolus is another regulatory response to cellular stress. For example in the case of heat stress, not only stress-responsive proteins (e.g. p53 and PML body proteins) but also unfolded proteins can be recruited into the nucleolus

to form intranucleolar bodies (INBs) and even amyloid bodies (A-bodies) that expand the nucleolus[22,50]. The dynamic nucleolar stress response have similar characteristics that demonstrate an accumulation of proteins and RNA that do not belong to normal nucleolar components[25]. Thus, the nucleolus, especially upon disruption of the tripartite structure from stress responses, becomes more difficult if not impossible to characterize by classical biochemical fractionation and protein immunoprecipitation. New technologies are needed that can probe nucleolar structure and its disrupted states within the cellular context.

1.2 RNA-CENTRIC TECHNOLOGIES

RNA function expands beyond its role as a protein precursor to participate in a broad range of functions in biomolecular condensates and regulating gene expression. RNAs dynamically interact with other biomolecules—especially proteins—to form ribonucleoproteins (RNP) complexes. Some RNA-binding proteins (RBP) possess RNA-binding domains such as the RNA recognition motif (RRM), hnRNP K homology domain, or DEAD box helicase domain[4]. However, studies capturing mRNA binding proteins through crosslinking and oligo(dt) enrichment coupled to proteomics—RNA interactome capture (RIC)[51]—determined that about half of these mRNA interact with proteins that lack conventional RNA binding domains and have connections to other biological processes including metabolic stress, cell cycle and development[5,52,53]. While techniques, such as RIC and Orthogonal Organic Phase Separation (OOPS)[54], attempt to identify RBPs through broad capture of RNAs, these techniques do not focus individual RNAs and their interactome. Protein-centric technologies and co-immunoprecipitation only work for known RBPs and highlights the challenge of characterizing unknown RBPs interacting with specific RNAs to regulate its function.

In vitro RNA-centric technologies

The field has developed RNA-centric technologies to target specific RNA and identify other interacting partners. *In vitro* methods focus on studying RNA and proteins outside of the cellular context. For example, antisense oligonucleotides complementing a RNA of interest, often end-labeled with a biotin[55], can be bound to resin and exposed to crude lysate to enrich RBPs for proteomic analysis. Similarly, target RNA can be expressed with an aptamer tag, such as the S1 biotin aptamer, or stem-loops that bind the CRISPR Cys4 endoribonuclease, or bacteriophage MS2 coat protein, to enrich for RBPs bound to a resin[56,57]. However these approaches' use of crude lysate make them prone to

nonspecific associations with other proteins that may not occur due to subcellular localization. These methods also focus on characterizing RBPs that are directly interacting with RNA and are especially biased toward the detection of abundant and highly specific interacting proteins[58].

***In cellulo* RNA-centric technologies**

In vivo approaches attempt to characterize RNA-protein interactions within the cellular environment, usually incorporating cross-linking. For instance, RNA affinity purification (RAP)[14] uses UV cross-linking before exposing lysate to long (>100 nt) antisense oligonucleotide probes. Other techniques, such chromatin isolation by RNA purification (ChIRP)[59] and capture hybridization analysis of RNA targets (CHART)[60], target specific RNAs within a crosslinked cell and are captured later by enrichment and MS analysis. While these methods imply more of a preservation of subcellular environments, these methods rely on high inputs of cell numbers of up to a billion cells and can be complicated when targeting lowly expressed transcripts[58]. Also, the interacting proteomes from several RAP experiments targeting XIST have limited overlap[61–63]. While crosslinking may improve the signal of larger RBP complexes, and even proteins in an extended neighborhood of the target RNA, these approaches can still be improved upon to increase specificity and reproducibility.

1.4 LIVE CELL PROXIMITY LABELING TECHNOLOGIES

Proximity labeling has been introduced as an approach to map the interactome of living cells. Briefly, this technique relies on an enzyme, such as peroxidase or biotin ligase[64], to promiscuously biotinylate nearby proteins, and to a lower extent, RNA[65] and DNA. These biotinylated biopolymers can then be isolated using streptavidin enrichment before being analyzed by downstream proteomics and sequencing. Many different biotinylating enzymes have been engineered to decrease label time and target biopolymers more efficiently including ascorbate peroxidases (e.g. APEX) and biotin ligases (e.g. BioID, miniTurbo, and TurboID). Through imaging, the biotinylation radius of these techniques range from 1 to 10 nm[66,67] , enabling the capture of even weak or transient interactions without crosslinking. Proximity labeling also allows for the enrichment of subcellular neighborhoods around targets of interest. In protein-centric technologies, these biotinylating enzymes can be transgenically expressed to target proteins or organelles in live cells. Their interacting RNAs and proximal proteins can then be analyzed by MS and

sequencing in techniques such as APEX-RIP[68] and APEX-seq[65]. However, there are fewer RNA-centric methods addressing specific RNA interactomes.

Recent technologies have been developed to identify protein partners of a RNA of interest within live cells. These techniques have focused on aptamer tags and CRISPR to guide biotinylating enzymes to RNAs. In RNA-protein interaction detection (RaPID)[69], the target RNA motif is flanked by BoxB stem loops that bind to the RaPID protein comprised of BirA* fused to a λN peptide that recognizes the BoxB stem loop. The biotinylated proteins are then analyzed by mass spectrometry to yield nearby proteins interacting with this RNA motif that is expressed in cells. Another aptamer based technique, RNA BioID, relies on expressing target RNA with a MS2 tag, then recruiting BirA* fused to a MS2 coat binding protein (MCP) for biotinylating nearby proteins analyzed by MS[70]. CRISPR-based RNA-United Interacting System (CRUIS)[71] and CRISPR-assisted RNA-protein interaction detection method (CARPID)[72] use dcas13 and CRISPR-CasRx, respectively, engineered to proximity labeling enzymes to biotinylate interacting proteins in living cells. These methods require careful optimization of both the engineered CRISPR proteins and guide RNAs to achieve signal over noise. Controlling transgenic construct localization and expression can lead to poor signal and contradicting RNA-protein interactomes[73]. Furthermore, these techniques have been used on endogenously abundant and overexpressed RNAs, limiting the range of specific RNA targets and their interactomes.

1.5 IN SITU PROXIMITY LABELING TECHNOLOGIES

To address the limitations of expressing transgenic constructs in live cells, recent studies have focused on *in situ* labeling approaches that target enzymes *post facto* in chemically fixed cells. From a specific protein perspective, techniques can use primary and secondary antibodies conjugated to promiscuous biotinylating enzymes to identify protein interactomes; these techniques include biotinylation by antibody recognition (BAR)[74], tyramide signal amplification-sequencing/mass spectrometry[75], and selective proteomic proximity labeling assay using tyramide (SPPLAT)[76]. These technologies use crosslinked cells omitting the need for transgenic constructs and rely on specific target antibodies to recruit the biotinylating enzyme. Using classical tyramide signal amplification, these techniques can also increase and optimize biotinylation signal by titrating labeling times, assuming the antibody is specific.

Toward RNA-centric *in situ* methods, a recent method hybridization-proximity (HyPro)[77] labeling utilizes digoxigenin-labeled (DIG) antisense oligonucleotides to target RNA. A specialized HyPro enzyme with DIG binding protein conjugated to APEX2 is then recruited to biotinylate nearby biopolymers that can be identified in sequencing and MS analysis. However, this method utilizes custom synthesized protein reagents that are difficult to optimize and might influence the specificity of biotinylation. Thus there remains a need for straightforward RNA-centric technologies that can reproducibly target diversely expressed RNAs and specifically target their interactomes within the subcellular context.

1.6 ELUCIDATING RNA-PROTEIN INTERACTIONS USING PROXIMITY LABELING *IN SITU*

This thesis work focuses on developing proximity labeling technologies that can be used to interrogate interacting proteins of target RNAs *in situ*. In Chapter 2, I describe a new technology, Oligonucleotide-directed proximity-interactome mapping (O-MAP), that targets specific RNA and their interactomes *in situ* with optimizable signal amplification. A fluorescence *in situ* hybridization (FISH) technique called signal amplification by exchange reaction (SABER)[16], allows for the programmable amplification of FISH signal for target RNAs in different cell and tissues systems. Combining this technique with classic tyramide signal amplification, allows for the biotinylation of local neighborhoods around target RNAs that can be enriched and characterized by downstream proteomics and sequencing. I show that O-MAP can be applied and optimized to different RNA targets and in diverse biological specimen types, including cells, organoids, and tissue. Moreover, I explore and characterize the nucleolar proteome by targeting the 47s pre-ribosomal RNA. I characterize the 7sk proteome, which supports its relationship to nuclear speckles and regulation of transcription elongation. I show the possibility of expanding O-MAP to versatile applications in different systems to identify RNA-interactomes, and thus advancing our capability of characterizing higher order subnuclear RNA scaffolded structures. Ultimately, this technique can help decode how RNA can regulate gene expression, splicing, and other complex biological processes especially within the context of cancer and disease.

CHAPTER 2: OLIGONUCLEOTIDE-DIRECTED PROXIMITY-INTERACTOME MAPPING (O-MAP): A UNIFIED METHOD FOR DISCOVERING RNA-INTERACTING PROTEINS *in situ*

2.1 ABSTRACT

RNA dynamically interact with proteins to form ribonucleoprotein complexes responsible for regulating a broad swath of biopolymer function and fate. To address challenges in specifically probing RNA and their interactomes, we introduce Oligonucleotide-mediated proximity-interactome MAPping (O-MAP); a versatile method that characterizes biomolecules near a target RNA *in situ*. O-MAP uses programmable oligonucleotide probes to deliver proximity biotinylating enzymes to target RNA, enabling the enrichment of nearby molecules using streptavidin pulldown. O-MAP induces exceptionally precise RNA-localized *in situ* biotinylation, and unlike alternative methods it enables straightforward optimization of its targeting accuracy. Using the 47S pre-ribosomal RNA, 7sk small nuclear RNA, and long noncoding RNA *Xist* as models, we develop O-MAP workflows for unbiased discovery of RNA-proximal proteins. This revealed a characterization of nucleolar compartments and 7sk contained nuclear speckle components. O-MAP is portable to cultured cells, organoids, and tissues, and to RNAs of various lengths, abundances, and sequence composition. And, O-MAP requires no genetic manipulation and uses exclusively off-the-shelf parts. We therefore anticipate its application to a broad array of RNA phenomena.

2.2 INTRODUCTION

In the context of the cell, very little RNA is naked[4,54,78]. Direct binding interactions with other biomolecules (proteins, RNAs, genomic loci) regulate all aspects of an RNA's lifecycle, including biogenesis, localization, turnover, and protein-coding or noncoding functions[79,80]. Moreover, higher-order interactions between transcripts and their local microenvironment are critical for organizing subcellular

architecture and compartmentalization[81,82]. In humans, for example, RNAs are central determinants of chromatin folding[83–85], and they nucleate and scaffold a host of biomolecular condensates that collectively control cellular metabolic, epigenetic, and stress-signaling pathways[81,86]. Yet, most of these critical structures have eluded detailed molecular characterization, due to a lack of methods for elucidating RNA interactions at compartment-level (nm–μm) distance[86].

Most state-of-the-art RNA interaction-discovery approaches use biotinylated antisense oligonucleotides to pull down a target RNA and its molecular partners from cell lysates[58,86–88]. While powerful, these approaches are limited in several key regards. Because it is challenging to optimize the specificity of the oligo probe pool, these approaches are limited by off-target RNA capture[89]. Moreover, enriching RNA from crude lysates can be plagued by artifactual interactions with abundant, nonspecific RNA-binding proteins, leading to false positives[15]. Overcoming this experimental background often requires large input masses (~10^8 cells[88]), especially for low-abundance RNAs. Finally, because these techniques probe an RNA *ex vivo,* they cannot capture higher-order interactions that depend upon intact subcellular structure.

We hypothesized that many of these obstacles could be addressed using *in situ* proximity-biotinylation[90]. In this approach, promiscuous biotinylating enzymes (*e.g.* peroxidases like Horseradish Peroxidase, HRP, and APEX)[69,91,92] are localized to a subcellular site of interest, typically by transgenic expression[90]. Upon addition of substrate, these enzymes generate reactive biotin species that diffuse outward and covalently tag nearby molecules *in situ*, enabling these molecules to be enriched by simple streptavidin pulldown and identified by mass spectrometry or high-throughput sequencing[13,65,93]. Yet, while proximity-biotinylation has emerged as a powerful tool for protein-targeted interaction discovery, applying the technique to RNAs remains challenging. Unlike proteins, transcripts cannot be genetically fused to the biotinylating enzyme, and hence most strategies seek to engineer artificial complexes between the enzyme and its target RNA[69,94,95]. But these transgenic approaches often produce substantial pools of mislocalized or unbound biotinylating enzymes, resulting in nonspecific background labeling that can blur the experimental signal[96]. These methods also require complex cell engineering to simultaneously overexpress multiple components, often including the RNA itself[94], making them challenging to apply to low-abundance RNAs, and altogether impossible in difficult-to-engineer lines,

organisms, and clinical samples. While a handful of methods (*e.g.* RNA-TRAP and HyPro)[77,97], potentially overcome this barrier by targeting individual transcripts without transgenic manipulation, these approaches rely on custom reagents that are cumbersome to synthesize, vet, and optimize[77,97]

To overcome these challenges, we here present Oligonucleotide-directed proximity-interactome mapping (O-MAP), a straightforward and flexible method for applying proximity-labeling to individual RNA targets in genetically unmodified samples. O-MAP utilizes the same peroxidase/tyramide chemistry used in APEX-based proximity-omics approaches[68,98], but it relies on programmable oligonucleotide probes, rather than transgenic expression, to deploy biotinylating enzymes to endogenous target RNAs (**Fig. 1a**). This approach enables precise RNA-targeting that can be easily optimized and experimentally validated, overcoming a chief limitation of oligo-pulldown based approaches. Building on this, we demonstrate O-MAP mass-spectrometry (O-MAP-MS) workflows for unbiased, compartment-wide RNA-interaction discovery. This technique requires markedly fewer (20–100-fold) cells than established methods, and use only inexpensive, off-the-shelf parts and standard manipulations. Using this workflow, we discover new interactions within the nucleolus and 7sk nuclear speckles—RNA dependent compartments that are difficult to isolate biochemically. Finally, we demonstrate that O-MAP can be readily applied to cell lines, organoids, and tissues, as well as to diverse classes of target RNAs, suggesting its utility as a broad-use RNA interaction-discovery tool.

2.3 RESULTS

O-MAP Design and Implementation

Recent advancements in RNA-Fluorescent *In Situ* Hybridization (RNA-FISH) oligonucleotide probe design have dramatically improved the method's sensitivity, flexibility, and specificity, enabling nearly any transcript to be imaged with high precision[99–101]. In developing O-MAP, we aimed to leverage these advancements to deploy the proximity-labeling enzyme HRP, rather than fluorophores, to a target RNA. As our primary model system, we chose the human 47S pre-ribosomal RNA, the long noncoding RNA that scaffolds the nucleolus[8,9]. To specifically target the 47S precursor (and avoid mature ribosomes) we selected an established RNA-FISH probe set[102] against ITS1, a "transcribed spacer" domain that is degraded during ribosome biogenesis and which never departs from the nucleolus (**Fig. 1 Fig. supplement 1**)[103]. We used these probes to explore different HRP recruitment strategies, by

appending them with functional modules that can be targeted by HRP conjugates and scoring the resulting nucleolar biotinylation by neutravidin staining (**Fig. 1 Fig. supplement 2 and 3a-b**). Several of these strategies used small-molecule haptens (*e.g.* Digoxigenin) to recruit an HRP-conjugated hapten-binding protein—similar to classical Tyramide Signal Amplification (TSA)-FISH and related interaction-discovery methods[75,77,93,97,104]. Yet, while these strategies induced prominent nucleolar biotinylation, they also exhibited variable off-target labeling that frequently rivaled the nucleolar signal (**Fig. 1 Fig. supplement 1b-e**).

In contrast, our optimized O-MAP strategy relies exclusively on programable oligonucleotide hybridization to precisely target HRP to an RNA of interest (**Fig. 1a**). In O-MAP, cells or tissues are chemically fixed with formaldehyde, and pools of oligo probes are annealed to the target RNA. These probes are chemically unmodified, but are appended with universal "landing pad" sequences originally developed for the FISH techniques Oligopaint[105] and SABER[16]. In a subsequent hybridization step, these landing pads recruit a common secondary oligo that is conjugated to HRP[106]. Upon addition of biotin-tyramide and hydrogen peroxide, HRP generates short-lived, highly reactive phenoxyl radicals that diffuse outward and pervasively biotinylate molecules near (~10 nm) the target RNA[13], enabling their enrichment.

Unlike hapten-based HRP-recruitment strategies (**Fig. 1 Fig. supplement 3**), 47S-targeted O-MAP induced prominent and precise nucleolar biotinylation with nearly undetectable background (**Fig. 1b**). Likewise, O-MAP targeting *7SK*—a small noncoding RNA thought to reside in nuclear speckles[43]—produced exclusively nucleoplasmic labeling. For each RNA, similar results were observed using multiple distinct probe sets, including probes designed using the OligoMiner Pipeline[101] to have different hybridization parameters **(Fig. 2- Fig. supplement 2**). Importantly, omitting any component of the O-MAP workflow (primary oligos, the oligo-HRP conjugate, HRP substrates), or using scrambled primary probes, completely ablated biotinylation (**Fig. 1b**). Moreover, 47S-targeted O-MAP appeared markedly more spatially refined than the analogous genetically encoded approach—nucleolar-targeted APEX2[65,107]—which generated substantial off-target, nucleoplasmic biotinylation (**Fig. 1c**).

O-MAP's landing pad design also enables a straightforward strategy for optimizing the specificity of its probe pool, thus overcoming a longstanding challenge of oligo pulldown-based approaches[58,86–88] **(Fig. 1d**). Derived from a standard RNA-FISH method[100], our O-MAP Probe Validation Assay first groups

probes into alternating sub-pools of probes divided equally and tiling the RNA of interest. Each sub-pool is outfitted with a distinct landing pad that recruits either an HRP-conjugated or fluorescent secondary oligo, enabling O-MAP and RNA-FISH to be performed simultaneously. Off-targeting (*i.e.* non-colocalizing) probes are then readily identified and eliminated. We applied this assay to our 47S- and 7SK-targeting probe sets, and in both cases observed strong overlap between the O-MAP and FISH signals, further underscoring O-MAP's precision (**Fig. 1e**). We then tested this assay on a less abundant RNA with a more confined localization. For this we chose *Xist*, the long noncoding RNA (lncRNA) that drives X-chromosome inactivation, and which coats the inactive X-chromosome (Xi)[108]. As predicted, *Xist* O-MAP induced strong biotinylation in a single prominent nuclear punctum, and which exclusively co-localized with the *Xist* RNA-FISH signal (**Fig. 1e**). Importantly, O-MAP's use of formaldehyde crosslinking and formamide denaturation did not appear to limit the recovery and enrichment of biotinylated proteins, as demonstrated with both 47S and 7SK (**Fig. 1f**). Taken with results above, this suggests that O-MAP generates exceptionally precise, RNA-targeted *in situ* biotinylation compatible with recovery and downstream analysis. We therefore sought to develop a tool for discovering proteins interacting with a target RNA (**O-MAP-MS**). We also took advantage of the flexibility of O-MAP to extend it to different organisms and targets.

O-MAP is portable across specimen types and target RNAs

A chief limitation of genetically encoded proximity-biotinylation methods is that they require the generation of a custom-built transgenic line for each new cell culture model, tissue, or organism under interrogation[90]. Because O-MAP doesn't require genetic manipulation to program its spatial targeting, it may enable rapid parallelization. We demonstrated this using O-MAP to profile the 47S pre-rRNA across a panel of cultured mammalian cell lines. In all cases, we used the same hybridization- and *in situ*-labeling conditions optimized in HeLa cells, and analogous (species-specific) 47S-targeting probes (**Fig. 1b**). This revealed O-MAP to be remarkably modular, catalyzing precise, nucleolar-targeted biotinylation in every line tested (ASPC1, SUIT2, 8988t, Panc3.27, A375, HEK 293T, U2OS, and MEFs, **Fig. 2a and Fig. 2 Fig. Supplement 1**). Notably, we observed robust labeling irrespective of nucleolar volume, density, or morphology—factors that can limit affinity-purification- and fractionation-based approaches.

This modularity can enable us to perform parallelized discovery of nucleolar interactions across multiple cell lines, which would be cumbersome by established methods.

Next, we tested if O-MAP might be portable to specimens for which transgenic approaches would be more challenging or intractable, using patient-derived PDA tumor organoids[109] and fixed mouse embryo tissue sections as models. Although each sample-type required modest re-optimization of hybridization- and *in situ*-labeling conditions, we were able to generate robust and specific 47S-targeted O-MAP in both (**Fig. 2b,c**). This demonstrates that O-MAP is extensible beyond two-dimensional cell culture, and implies its potential application for RNA-interaction discovery in clinically relevant contexts[109].

We next examined O-MAP's portability to different RNAs, by targeting an array of transcripts with diverse lengths, expression levels, sequence composition, biogenesis pathways, and localization. For each new target, probes were designed using the OligoMiner pipeline[101], and the optimal *in situ* biotinylation time was determined empirically via a labeling time course. Encouragingly, in each case, O-MAP yielded prominent *in situ* biotinylation that recapitulated the target's known subcellular localization. Our Probe Validation Assay (**Fig. 1d**) further confirmed the RNA-targeting accuracy of the probe pool: in all cases *in situ* biotinylation and RNA-FISH signals exhibited a high degree of overlap (**Fig. 2d**). This was observed with highly abundant transcripts like the chromatin-regulatory lncRNAs *MALAT1* and with low-abundance RNAs like *Firre*. We note that obtaining robust O-MAP at low-abundance targets required large primary probe pools (100–500 oligos) to bolster the number of HRP molecules recruited to the target transcript, as well as longer labeling times (60–120 minutes).

Mapping RNA-proximal proteomes with O-MAP-MS

To develop our O-MAP-MS pipeline our primary test case was the nucleolus—the subnuclear organelle that compartmentalizes and controls ribosome biogenesis—in HeLa cells[8,9]. Our approach emulated the ratiometric quantification strategy used in APEX-based proximity-proteomics, which enhances precision by measuring the relative abundances of peptides within a target compartment and the adjoining subcellular space[98] (**Fig. 3a**). In parallel experiments, we used 47S O-MAP to label nucleoli, 7SK-targeting probe sets (**Fig. 2- Fig. supplement 2**) as proxies for the neighboring nucleoplasmic compartment, and scrambled probes to model nonspecific background biotinylation. After O-MAP labeling, biotinylated proteins were enriched under denaturing conditions, and protein abundances were

measured by Tandem Mass Tag (TMTPro) quantitative mass spectrometry[110]. Owing to the reduced background labeling and tyramide amplification, our initial proof-of-principle O-MAP-MS experiments required 20–100-fold less starting material (5.5 million cells per replicate) than nucleolar biochemical fractionation[111] or the oligo pulldown method ChIRP–MS[88].

Yet, even with this small input, O-MAP-MS was able to accurately capture targeted subcellular compartments at considerable depth (1954 total proteins detected) and reproducibility (Pearson's correlations ranging from 0.77–0.99 between replicates; **Fig. 3. Fig. supplement 1**). Stereotypic nucleolar markers (*e.g.*, NPM1, NCL), were markedly enriched by 47S O-MAP (**Fig. 3b***, left*), while 7SK O-MAP highly enriched both classical components of the 7SK RNP (*e.g.* CDK9; Cyclin T1[112]), and recently discovered interactors like BAF155 and BAF160A [113](**Fig. 3b**, *right*). To expand on this, we examined more extensive sets of nucleolar and nucleoplasmic marker proteins, using the Human Protein Atlas (HPA)[114] to define markers exclusively localized within each compartment (38 nucleolar and 305 nucleoplasmic proteins observed in our data), and a cohort of proteins bilocalized between the two (155 in our data). As demonstrated in (**Fig. 3c,d**), we observed striking enrichment of resident nucleolar proteins exclusively from 47S O-MAP-MS (95% of markers, average of 5.1-fold enrichment), and converse enrichment of resident nucleoplasmic proteins exclusively from 7SK O-MAP-MS (98% of markers, average of 2.7–fold enrichment; $p = 6.83 \times 10^{-41}$, Fisher's Exact Test). Bilocalized proteins exhibited a multi-modal distribution, with 64% being more strongly enriched by 47S- than 7SK-O-MAP (average 2.1-fold enrichment, **Fig. 3c**).

To further assess O-MAP's spatial precision, we performed hypothesis-unbiased analysis of the putative nucleolar (47S-proximal) and nucleoplasmic (7SK-proximal) proteomes identified by our data. We used Receiver-Operating Characteristic (ROC) analysis to determine the optimal threshold separating these two groups ($\log_2$(47S/7SK) = 0.523, **Fig. 3 Fig. Supplementary 2-3**), thereby defining candidate lists of 258 O-MAP-Nucleolar and 1396 O-MAP-Nucleoplasmic proteins. Gene Ontology (GO)[115] and Gene Set Enrichment Analysis (GSEA)[116] of candidate nucleolar proteins revealed substantial enrichment of factors driving the transcription and processing of pre-ribosomal RNA, ribosome assembly, and nucleolar architecture, as expected (**Fig 3e,** *top* and **Fig. 3 Fig. Supplementary 4**). Likewise, the putative nucleoplasmic proteome was highly enriched in factors involved in chromatin organization, mRNA

biogenesis and DNA metabolism (**Fig. 2e***, bottom*). We furthermore examined if O-MAP might reveal more targeted analysis into the 7SK-proximal subnuclear compartment. 7SK has been suggested to accumulate in Nuclear Speckles—membraneless nuclear bodies with putative roles in mRNA biogenesis and processing[43]. Remarkably, of the 117 established speckle factors observed in our data, nearly all were strongly enriched by 7SK O-MAP-MS, relative to either 47S O-MAP or scramble controls (97% and 79% enriched, respectively, **Fig. 3f**). We used ROC analysis to define an optimal threshold cutoff for the speckle factors, noting that these factors were significantly more enriched from the broader nucleoplasmic proteome when comparing 7SK O-MAP to scrambled controls, rather than to 47S O-MAP (**Fig 3. Fig. Supplementary 2**). This analysis revealed a list of 510 candidate 7SK-proximal proteins, which were highly enriched in factors involved in mRNA biogenesis, especially pre-mRNA splicing (**Fig. 3g** and **Fig. 3 Supplementary 3-4**). This further reinforces 7SK's proposed localization to nuclear speckles, and these bodies' putative role as splicing hubs[43,117].

We next explored the temporal dynamics of O-MAP labeling. We reasoned that, during a labeling time course, proteins within a target compartment would be biotinylated with distinct kinetics from those adjoining it, allowing us to identify richer sets of compartment-specific factors by classifying common temporal profiles. To test this, we performed 47S-targeted O-MAP at times ranging from 1 second to 100 minutes, using a single 7SK time point as a normalization control (**Fig. 4a**). From K-medoid (k=12) clustering, four clusters were significantly enriched for HPA-defined nucleolar proteins (Fisher's exact test < 0.05; **Fig. 4 Fig. Supplement 1**), generating an expanded list of 313 candidate nucleolar proteins (**Fig. 2b**). Recovery of these proteins appeared to spike after one minute of *in situ* biotinylation and plateau by 10 minutes (**Fig. 4d,e**). Encouragingly, 241 (77%) proteins in this cluster have established roles in nucleolar organization or ribosome biogenesis, as annotated by the GO, HPA and Uniprot databases [118](**Fig. 4c**). Manual curation of the remaining factors revealed 31 (10%) that likely also play roles in these processes, but were misannotated (**Fig. 4c**). This demonstrates the potential power of this temporal profiling approach, which would be intractable by live-cell proximity-labeling due to the confounding effects of diffusion[119]. Taken with the above, these data strongly support O-MAP's ability to discover RNA-proximal proteomes with high precision.

2.4 Discussion

In this work, we have developed a nearly universal method for applying proximity-labeling to RNA, by co-opting the mechanics of RNA-FISH to precisely deploy biotinylating enzymes to a target transcript without genetic modification. This approach, O-MAP, supports a robust toolkit for elucidating the proteins (O-MAP-MS), transcripts (O-MAP-Seq), and genomic loci (O-MAP-ChIP) near an RNA of interest. We believe this toolkit holds considerable advantages over established RNA interaction-discovery methods, offering superior precision, biological context, flexibility, ease of use, and cost. Moreover, we anticipate that O-MAP's ability to probe higher-order interactions within a transcript's subcellular "neighborhood" will enable unprecedented analysis of RNA-mediated compartmentalization[81,82], a powerful new approach for spatial cell biology.

The most effective RNA interaction-discovery methods use antisense oligonucleotides to pull down a target RNA *ex vivo*[58,86–88], or transgenic manipulation to localize proximity-labeling enzymes to that RNA in live cells[69,95,120]. Compared to oligo pulldown-based approaches (*e.g.* ChIRP, ChART, RAP)[14,86,88], O-MAP offers more precise optimization, higher sensitivity, and broader biological context. A chief limitation of these pulldown-based techniques is their inability to optimize the targeting precision of the capture-oligo pool[86,89], leading to off-targeting noise. O-MAP overcomes this using a straightforward validation assay that rapidly identifies and eliminates off-targeting probes (**Figs. 1d**–**e**, and **2**). Furthermore, oligo pulldown techniques capture their target RNA in crude cell lysates, where artifactual interactions with abundant, promiscuous RNA-binding proteins can confound analysis[15]. Overcoming this issue typically eliminates all but the target RNA's most direct interactors[86], thus eliminating higher-order biological context. Moreover, the inefficiency of target capture, typically necessitate large inputs (>10^8 cells per replicate)[86]. In contrast, because O-MAP enzymatically amplifies the number of biotins per oligo probe, and enables interacting molecules to be directly captured (rather than co-purified by pulldown), it is markedly more efficient. Of note, our complete proteomic analysis of HeLa nucleoli (**Fig. 3**)—including all experimental replicates and controls—collectively required fewer cells than would a single replicate of ChIRP-MS[88]. This higher efficiency, and the ease with which O-MAP is deployed to different specimens (**Fig. 2**), could enable parallelized analysis across tissues, cell types, or other experimental variables, which would be unwieldly using oligo-capture methods. And unlike those methods, proximity-biotinylation

approaches like O-MAP are uniquely suited to probe dynamic, higher-order interactions that are too transient or fragile to survive a pulldown.

Compared to RNA-targeted live-cell proximity-labeling methods, O-MAP is more spatially precise, more applicable across transcripts and specimens, and is simpler to design and execute. Live cell approaches (*e.g.,* RapID, MS2/Cas13-APEX, CRUIS) assemble artificial complexes between the labeling enzyme and its RNA target, by transgenically overexpressing these components fused to localization sequences or binding motifs[58,71,95,120]. Although such transgenic approaches have had some success, they must contend with substantial background labeling from unbound enzymes[95,120] (**Fig. 1c**), and with artifacts from overexpressing the target RNA outside of its native context[121]. These issues are particularly problematic with low-abundance transcripts. In contrast, O-MAP doesn't require transgenic manipulation to control enzyme localization. The sequences, abundances, and hybridization parameters of its HRP-targeting oligos can be explicitly controlled to minimize off-targeting[16], and residual, mislocalized probes can be removed by stringent washing. This strategy enabled endogenous RNAs—even low abundance transcripts—to be biotinylated with nearly undetectable background (**Figs. 1b** and **2**). It furthermore enables O-MAP analysis in contexts where genetic manipulation would be challenging or altogether impossible, including pre-fixed tissue samples (**Fig. 2c**). Applying O-MAP in samples like these, such as clinical isolates, may provide a powerful avenue for biomarker and therapeutic target discovery.

O-MAP is conceptually similar to other methods that use affinity reagents, rather than transgenic expression, to localize proximity-labeling enzymes. Several established 'omics tools (*e.g.,* EMARS, SPPLAT, BAR, TSA-Seq, and TSA-MS[75,76,93,122,123] use antibodies to deploy HRP to a target protein *in situ*, essentially adapting an immunofluorescence workflow for novel interaction discovery. A similar tactic underlies RNA-TRAP[97] and HyPro[77], which use Digoxigenin (DIG)-modified oligos to recruit HRP to a target RNA—either via an anti-DIG antibody-conjugate[97] or linked to a custom-made DIG-binding protein[77]. In our experience, these hapten-recruitment strategies are markedly noisier than O-MAP's "landing-pad" approach (**Fig. 1 Fig. Supplement 3**), and their lack of modularity precludes strategies like our Probe Validation Assay (**Figs. 1d–e** and **2**), a vital optimization tool. Moreover, unlike the current implementation of HyPro, O-MAP is compatible with quantitative tandem mass-tag-based proteomics (**Fig. 3**) and can be expanded into other downstream ChIP-like genomic interaction discovery (*data not*

shown), and can be used with conventional fixed tissues (**Fig. 2**). O-MAP is also considerably cheaper, using inexpensive, chemically unmodified primary oligos, in lieu of DIG-conjugated probes. And critically, O-MAP uses entirely off-the-shelf parts, and doesn't require the synthesis of custom proteins.

Our work also illustrates O-MAP's unique ability to discover compartment-level interactions that are opaque to current methods. For example, while 7SK's direct binding partners are well established[112,113], our O-MAP-MS analysis represents, to our knowledge, the first proteomic characterization of this enigmatic RNA's subnuclear compartment (**Fig 3f,g**). Extending this analysis to other lines, and complementing it with O-MAP-MS and even sequencing, will enable an unprecedented molecular characterization of nucleolar architectural remodeling during oncogenesis[124].

We anticipate that O-MAP will be applicable to nearly any target RNA, and in any biological setting, for which RNA-FISH can be performed. However, we note some of the technique's current limitations. O-MAP will likely be challenging with difficult FISH targets (*e.g.,* small, low-abundance RNAs), requiring more advanced probe designs to amplify signal[16]. Moreover, formaldehyde fixation—though a mainstay of high-resolution imaging—can induce aberrations that perturb the local subcellular structure[125], and which might be improved using alternate fixation strategies. Finally, like all peroxidase-based proximity-labeling strategies, O-MAP's labeling chemistry identifies all molecules within its labeling radius, including both the target's direct and nearby interactors. Future development of O-MAP's probe design and labeling chemistry will enable more precise control over its labeling radius, allowing "shells" of RNA-interactions to be probed independently. Regardless, in its current implementation we believe that O-MAP's ability to elucidate RNA interactions *in situ* overcomes a technical roadblock that has long challenged the RNA community. Given the extraordinarily broad scope of cellular functions performed by RNA[126], this advancement will catalyze fundamental discoveries into countless biological phenomena.

2.5 Methods

Cultured cells, tissue sections and organoids

HeLa cells, female mouse embryonic fibroblasts (fMEFs, *a gift from Dr. C. Disteche, UW*), HEK 293T, Patski, SUIT2, and U-2 OS cells were cultured in High Glucose DMEM with Pyruvate (Thermo Fisher; 11995073), supplemented with 10% (v/v) Fetal Bovine Serum (FBS, Thermo Fisher; 26140079), 100 units/mL penicillin and 100 µg/mL streptomycin (Thermo Fisher; 15140122), and 1x GlutaMAX

(Thermo Fisher; 35050061). 8988T, A375, ASPC1, and Panc 3.27 cells were cultured in RPMI 1640 media (Thermo Fisher; 11875093), supplemented with 10% (v/v), FBS, 100 units/mL penicillin, and 100 µg/mL streptomycin. In all cases, cells were maintained at 37°C, under 5% CO_2. Cell lines were authenticated by STR testing (ATCC), when possible.

For imaging experiments, cells were cultured in two-well Lab-Tek borosilicate glass #1.0 chambers (Thermo Fisher; 155380). To improve HEK293T adherence, chambers were treated with gelatin (0.5% (w/v), in water, Sigma; G7765) for 30 minutes at 37°C, prior to plating. For biochemistry, proteomic, and high-throughput sequencing experiments, cells were cultured in six-well plates. When necessary, material from multiple wells was harvested and merged into a single lysate, as described below.

Human Pancreatic Ductal Adenocarcinoma organoids (**Fig. 2**) were prepared as described. Mouse tissue sections (**Fig. 2c**, *a gift from Dr. E. Nichols, UW*) were prepared from day E13.5 embryos (C57BL/6J wild type mice; Jackson Laboratory) by drop fixing in 4% (v/v) formaldehyde, followed by sucrose equilibration, and thin-sectioning in OCT compound (VWR 25608-930. Cryosections (approximately 10 µm) were prepared at the UW Histology and Imaging Core and stored on glass slides at –80°C until use.

O-MAP probe design and synthesis

Probes targeting the human 47S pre-rRNA ITS1 domain were used from previously vetted probe sets[127]. All other probes were designed using OligoMiner pipeline using the following settings. The blockParse script was run using the settings: -l 30 -L 37 -t 42 -T 47 -s 390 -F 40. Bowtie2 was used with settings: -U --no-hd -t -k 100 --very-sensitive-local -S. The outputClean script was run with the -u argument; the structureCheck script was run with the settings: -F 40 -s 390 -m dna1998 -T 42. K-mer filtering was performed in Jellyfish version 2.2.10, using a Jellyfish file for the corresponding genome (human or mouse), and using the kmerFilter function with the -m 18 and -k 1 arguments. Jellyfish files were generated for each genome assembly (hg38 for human; mm10 for mouse), using a hash size set to the appropriate size of the genome assembly[128]. For example, the command -s 3300M -m 18 -o hg38_18.jf --out-counter-len 1 -L 2 hg38.fa was used to generate the 18mer dictionary for hg38. For most targets, all probes that passed this final filtering step—typically 10–150 probes per target—were used.

For the O-MAP Probe Validation Assay (**Fig. 1d**), probes were divided into sub-pools along the length of the target RNA, appended on their 3´ termini with SABER1 or SABER2 "Landing-pad" sequences[16]. Once probes were validated, the complete pool was reformulated, appended with SABER1 and used for O-MAP alone.

Probe sets consisting of fewer than 20 probes were ordered as individual oligos (Sigma; 0.025–0.05 µg synthesis scale, standard desalting), and further purified from preparative polyacrylamide gels, as described previously[129]. Purified oligos were resuspended in nuclease-free water, quantified by UV-vis spectroscopy, pooled to a final aggregate concentration of 20 µM and stored at –20°C. Probe pools requiring more than 20 oligos were purchased as oPools (IDT; 50 pmol/oligo scale, unmodified), and dissolved to approximately 100 µM in nuclease free-water. 20 µL were desalted using the Oligo Clean & Concentrator Kit (Zymo Research; D4060), following the manufacturer's instructions. Pools were quantified by UV-vis spectroscopy, diluted to 5 µm and stored at –20°C. Fluorophore-conjugated secondary probe used in RNA-FISH ("SABER2–AF647") was purchased from IDT (100 nmol scale; HPLC purification), resuspended to 100 µM in nuclease-free water, and stored in a light-tight container at –20°C. The HRP-conjugated secondary probe ("SABER1–HRP") was purchased from Bio-Synthesis, resuspended to 10 µM in resuspension buffer (10mM NaH_2PO_4, 150 mM NaCl, pH 7.2), allotted into 10 µL single-use aliquots, frozen and stored at –20°C.

In some cases, the RNA-FISH signal was amplified by extending the FISH probe subpool with concatemers of additional "SABER2" Landing-pads. These were enzymatically added via the Polymerase Exchange Reaction (PER), essentially as described[16]. Briefly, pooled probes (5 µM, aggregate) were incubated with 0.5 µM Template hairpin and 0.1 µM Clean.G DNA hairpin (IDT), in 10 mM $MgSO_4$, 0.6 mM each ATP, CTP, and TTP, 4 U *Bst* 2.0 DNA Polymerase (NEB; M0537), 1x PBS, in a final volume of 50 µL. Reactions were incubated at 37°C for two hours in a thermocycler, heat-inactivated at 80°C for 20 minutes, and chilled on ice. The resulting PER-extended oligos were then purified with OligoClean and Concentrator Kits, eluting into nuclease-free water, and their length was examined on denaturing 10% Polyacrylamide TBE-Urea gels, stained with SYBR-Gold (Thermo Fisher; S11494).

O-MAP core protocol

The following protocol was used for omics-scale O-MAP, using cells grown in six-well dishes ($3.5x10^5$ cells/well; plated one day prior to harvest). For imaging-only experiments, cells were plated at $7x10^4$ cells per well, in two-chamber LabTeks. In all cases, RNase-free reagents and manipulations were used throughout.

The core O-MAP workflow is split over two days. The first day comprises fixation, permeabilization, peroxidase inactivation, and primary probe hybridization; the second day comprises secondary probe hybridization, an optional endogenous biotin blocking step, and *in situ* biotinylation. Thereafter, the protocol varies depending on the endpoint assay—imaging, western blotting, and proteomics.

O-MAP Day 1. All manipulations were performed at room temperature, unless noted. Cells were washed briefly with 1x Ca- and Mg-free DPBS (Thermo Fisher; 14190250) and fixed with freshly prepared 2% (v/v) formaldehyde (Electron Microscopy Sciences; 15710) in 1x PBS (Sigma; 6506), for 10 minutes without agitation. The formaldehyde solution was aspirated and the crosslinking reaction terminated by two washes with 250mM glycine in 1x PBS, five minutes each, with gentle rocking (3 rpm on a platform rocker). Cells were briefly washed with DPBS, permeabilized with 0.5% (v/v) triton-X 100 in PBS (10 min; gentle rocking), and washed three times with DPBS. Next, to inactivate endogenous peroxidases, samples were treated with 0.5% (v/v) H_2O_2, in 1x PBS, for 10 minutes with gentle rocking, and washed twice with PBS. Samples were then equilibrated in Formamide Wash Buffer (10–40% (v/v) deionized formamide (Thermo Fisher; AM9342); 2x SSC (Thermo Fisher; AM9765); 0.1% (v/v) Tween-20), for five minutes with gentle rocking. The formamide concentration was matched to the primary probe hybridization mix, as determined by the binding parameters of the primary probe. This buffer was then aspirated, and each sample was treated with 115 µL of Probe Mix (0.1 µM primary oligo probe pool, in 1x Hybridization Buffer: 10–40% deionized formamide; 2x SSC; 0.1% (v/v) Tween-20; 10% (w/v) dextran sulfate (SIGMA; D8906); in nuclease-free water) and this mix was gently spread over the sample by covering with a clean, 30 mm diameter #1.5 thickness glass cover slip (Bioptechs; 30-1313-03192). A 2x SSC-soaked kimwipe was placed between the wells to maintain humidity during hybridization. Plates were then sealed with Parafilm and incubated without agitation for 8 hours at 37°C or 42°C, depending on the probe set

O-MAP Day 2. Following primary hybridization, coverslips were removed and cells were washed three times with pre-warmed 30% Formamide Wash Buffer, 10 minutes per wash, at 37°C with gentle rocking. For imaging experiments, this was followed by a blocking step to mask endogenous biotinylated proteins, as described below (*see* "O-MAP Imaging"). In all cases, subsequent manipulations were carried out in the dark, to avoid photooxidation of the HRP conjugate. Each well was treated with 115 µL O-MAP Secondary Probe Mix (100 nM SABER1–HRP oligo, in 30% Formamide Hybridization Buffer), and covered with a clean coverslip. Samples were incubated for 1 hour at 37°C, with gentle rocking. Coverslips were then removed, and samples were washed four times with PBST (0.1% (v/v) Tween-20 in 1x PBS), 15 minutes per wash, with gentle rocking. Buffer was aspirated, and *in situ* biotinylation induced by addition of 1 mL Labeling Solution (0.8 µM biotinyl-tyramide (Sigma; SML2135), 1 mM H_2O_2, 1x PBST), and incubation at room temperature. Labeling times varied between RNA targets—ranging from 1 second (**Fig. 4a,** *top left*) to 120 minutes (**Fig. 4a**, *bottom right*)—and were determined empirically using the O-MAP Imaging assays described below. In all cases, biotinylation was halted by addition of Sodium Azide and Sodium Ascorbate (10 mM each, final) in 1x PBST, for three washes of five minutes each.

O-MAP imaging

For imaging experiments, the background signal from endogenous biotinylated proteins was blocked after the secondary probe hybridization step. Briefly, samples were washed three times in 1x PBST and incubated in pre-blocking solution (1% (w/v) nuclease-free BSA (VWR; 97061-420) in 1x PBST) at room temperature for 30 min with gentle rocking. Samples were then blocked with 1 mL of Neutravidin Blocking Solution (10 µg/mL neutravidin (Thermo Fisher; 31000), 1% (w/v) nuclease-free BSA, in 1x PBST) for 15 min with gentle rocking at room temperature, and washed three times with PBST. To saturate free streptavidin binding sites, samples were next treated with 10 µg/mL D-biotin (Thermo Fisher; B20656) in 1x PBST, for 15 minutes with gentle rocking, followed by three washes with room temperature PBST. Thereafter, *in situ* biotinylation and quenching proceeded as described above, using 50 µL volumes for primary and secondary hybridization buffers. After biotinylation and quenching, samples were stained with 1 mL 1 µg/mL neutravidin-DyLight 550 conjugate (Thermo Fisher; 84606), in 1% BSA pre-blocking solution, for one hour at room temperature with gentle rocking, followed by three

washes with 1x PBST. Samples were counterstained with DAPI (5 µg/mL, in 1x PBST) and imaged immediately, or were mounted in Vectashield (Vector Labs; H-1900-10) and stored at 4°C.

O-MAP-MS

O-MAP-labeled cells (approximately $5.5X10^6$ cells per replicate; three replicates per experimental condition) were harvested by scraping into 1x PBST, supplemented with 10 mM Sodium Azide. After pelleting by centrifugation at 800 x *g* for 10 minutes at 4°C, remaining buffer was aspirated and the pellets were flash frozen in liquid nitrogen and stored at –80°C until use. All subsequent steps were performed at room temperature, unless noted. Cell pellets were lysed in 800 µL of MS Lysis Buffer (4% (w/v) SDS in 1x PBS, with 10 mM Sodium Azide, 1x Halt EDTA-Free Protease Inhibitor Cocktail) for five minutes.

Samples were then sonicated using a Branson Digital Sonifier 250 outfitted with a double stepped microtip (Emerson Industrial Automation) at 10–12 Watts for 30 seconds (0.7 s on; 1.3 s off) for one cycle. Lysates were clarified by centrifugation at 15,000 x *g* for 10 minutes and soluble protein was quantified using the Pierce BCA Protein Assay Kit (Thermo Fisher; 23225). For each sample, 300 µg of protein was diluted to 1% SDS by the addition of three volumes Dilution Buffer (1x PBS, supplemented with 10 mM Sodium Azide and 1x Protease Inhibitors). Protein samples were then reduced with TCEP (Thermo Fisher; 77720, 10 mM final concentration) for 60 minutes with gentle rotation. To alkylate free thiol groups, samples were treated with iodoacetamide (Sigma; I1149, 20 mM final concentration) and rotated for 60 minutes at room temperature, before quenching by addition of DTT (5 mM final) and incubation for 15 minutes.

For streptavidin pulldown, to each sample was added 100 µL Pierce Streptavidin Magnetic Bead slurry (Thermo Fisher; 88817) that had been equilibrated in Diluted Lysis Buffer (1% (w/v) SDS, 1x PBS, 10 mM Sodium Azide, 1x Protease inhibitors). Samples were rocked end over end for two hours at room temperature, and streptavidin beads were then washed with the following buffers (5 minutes per wash; rocking end over end at room temperature): (1–2) Two washes in Diluted Lysis Buffer, (3) 1x PBS (4–5) Two washes in KCl Buffer (1 M KCl in 1xPBS), (6–7)Two washes in Urea Buffer (2 M Urea in 1xPBS), (8–9) Two washes in 200 mM EPPS (pH 8.5). Beads were then resuspended in 15 µL 200 mM EPPS (pH 8.5).

Bound proteins were eluted by on-bead proteolytic digestion, as follows. Lysyl endopeptidase (Lys-C, Fujifilm Wako; 121-05063) was added at a ratio of 1 µg of enzyme per 100 µg of input protein, and samples were incubated for three hours at 37°C, with shaking at 500 rpm. Trypsin (Thermo Fisher; 90057) was then added at a ratio of 1 µg of enzyme per 100 µg of input protein, and digestion continued overnight at 37°C, 500 rpm shaking.

O-MAP-MS LC-MS/MS Sample Preparation

Eluted peptides were labeled with TMTpro reagents using established protocols[110]. Briefly, eluted peptides were supplemented with acetonitrile to a final concentration of 30% (v/v), in 200 mM EPPS buffer (pH 8.5). TMTpro reagents in 100% anhydrous acetonitrile were then added to each sample at approximately a 2.5:1 (w/w) excess. The labeling reaction was allowed to continue for 1.5 hours, quenched with 5% hydroxylamine, and the labeled peptides were mixed. Pooled peptides were then dried by vacuum centrifugation. Dried, labeled peptides were resuspended in 100 µl of (5% acetonitrile, 1% formic acid) and cleaned using in-house assembled stage-tips[130]. Pooled peptides were eluted in (70% acetonitrile, 1% formic acid). Eluates were then dried to completion and stored at -80°C until analyzed by LC-MS/MS.

O-MAP-MS data acquisition and analysis

Pooled, labeled peptides were resuspended in (5% acetonitrile, 2% formic acid) and eluted over an in-house pulled 25 cm C18 column (Accucore, Thermo Fisher Scientific) throughout a 180 minute gradient from (6% acetonitrile, 0.125% formic acid) to (32% acetonitrile, 0.125% formic acid). Peptides were analyzed using an SPS-MS3 method on a Thermo Fisher Orbitrap Eclipse to quantify TMTpro reporter ions. Briefly, the duty cycle consisted of three FAIMS (FAIMSpro, Thermo Fisher Scientific) mobility regions at Compensation Voltages (CV)=-40/-60/-80V. At each CV the following were collected within a duty cycle: an MS1 scan (R=120,000, MaxIT=50ms), six MS2 scans (Ion trap, Turbo scan speed, MaxIT=50ms, AGC=200%, CID NCE = 35%), and six SPS-MS3 scans (R=50,000, MaxIT=86ms, HCD NCE = 45%, AGC = 400%). A single dynamic exclusion of 90s was used across all CVs.

Resulting spectra were analyzed using the Comet search algorithm, searched against a full human protein database with forward and reverse protein sequences (Uniprot 10/2020). Precursor monoisotopic peaks were estimated using the Monocle package[131]. Peptides and proteins were filtered to

a 1% false discovery rate using the rules of parsimony and protein picking[132]. Protein quantification was done using signal-to-noise estimates of reporter ions and these data were processed and plotted using the R statistical programming language. Receiver-Operating Characteristic (ROC) analyses were performed in Microsoft Excel, as described previously[68,98] Gene Ontology analysis was performed using MetaScape[115], and GSEA[116].

Streptavidin blotting

O-MAP labeled samples were lysed, sonicated, and clarified as described for O-MAP-MS. Samples were quantified by BCA, supplemented with 2x Laemmli loading buffer and heated to 95°C for 10 minutes. Samples, standardized by protein mass, were loaded and separated on 10% SDS-PAGE gels, transferred onto PVDF membranes and stained with Ponceau S (Sigma P7170). Membranes were blocked with 5% (w/v) powdered milk in TBS-T (20 mM Tris, pH 7.5, 150 mM NaCl, 0.1% (v/v) Tween-20) for one hour with rocking at room temperature. Membranes were washed three times in TBS-T, 5 minutes per wash, and blotted with streptavidin–HRP conjugate (Thermo Fisher S911, diluted 1:20,000 in TBS-T, supplemented with 5% (w/v) BSA) overnight at 4°C. Blots were washed three times for five minutes in TBS-T, developed using the SuperSignal West Pico PLUS Chemiluminescent Kit (Thermo Fisher 34580), and imaged.

RNA-FISH, O-MAP Probe-Validation assays, and immunofluorescence

For RNA-FISH, cells were fixed in formaldehyde, permeabilized with Triton-X 100, equilibrated in formamide wash buffer, and hybridized to primary probes as described above (*see: O-MAP Day 1*), but omitting the peroxidase inactivation step. Thereafter, samples were washed three times in 30% Formamide Wash Buffer (five minutes per wash; room temperature) and incubated with 50 µL FISH Secondary Probe Mix (100 nM SABER2–AF647 oligo (Supplementary Table 1), in 30% Formamide Hybridization Buffer), and covered with a clean coverslip in a hybridization chamber. Samples were incubated for 1 hour at 37°C in the dark. Samples were washed three times with PBST (five minutes per wash), counterstained with DAPI solution (5 µg/mL, in 1x PBST) and either imaged immediately, or stored sealed in vectashield at 4°C.

For combined O-MAP and RNA-FISH experiments (**Figs. 1d–e, 2**), including the Probe-Validation Assay (**Fig. 1d**), RNA-FISH signal was dramatically diminished when HRP- and Fluorophore-conjugated

secondary oligos were hybridized simultaneously, presumably due to fluorophore damage during the *in situ* biotinylation reaction. To avoid this, O-MAP and RNA-FISH were performed sequentially. O-MAP was performed first, as described above (*see: O-MAP Core Protocol*), including the endogenous biotin blocking step (*see: O-MAP Imaging*). After quenching the O-MAP biotinylation reaction, samples were washed three times in 30% Formamide Wash Buffer (5 minutes per wash; room temperature), and incubated with 50 µL FISH Secondary Probe Mix (100 nM SABER2–AF647, in 30% Formamide Hybridization Buffer) for 1 hour at 37°C. Samples were washed four times with PBST (five minutes per wash), stained with neutravidin–fluorophore conjugate, counterstained with DAPI solution for two minutes, and either imaged immediately in 1x PBST or mounted in vectashield.

When combined with immunofluorescence (**Figs. 1b, 2a–b**), O-MAP or RNA-FISH were performed prior to immunostaining. Cells were subjected to the RNA-FISH or O-MAP pipelines, as described above. Following secondary probe hybridization (RNA-FISH) or *in situ* biotinylation and quenching (O-MAP), cells were washed three times in 1x PBST, and then blocked with 5% (w/v) nuclease-free BSA (VWR 0332) in 1x PBST for one hour at room temperature. Samples were then incubated with rabbit anti-*NPM1* (Thermo Fisher; PA517742, used at 1:100 dilution) or mouse anti-*FBL* (Thermo Fisher; MA1-22000; used at 1:100 dilution), in 1% (w/v) BSA, 1x PBST, for one hour at room temperature with gentle rocking. Samples were washed four times with 1x PBST and then incubated with either AlexaFluor 488-conjugated anti-rabbit (Thermo Fisher; A32731TR, 1:1000 dilution), or AlexaFluor 488-conjugated anti-mouse (Thermo Fisher; A32723TR 1:1000 dilution), with neutravidin-DyLight 550 conjugate (Thermo Fisher; 84606, 1:1000 dilution) as appropriate, in 1% (w/v) nuclease-free BSA, 1x PBST, for one hour at room temperature. Samples were washed four times with 1X PBST (15 minutes per wash), counterstained with DAPI for two minutes, and either imaged immediately in 1x PBST or mounted in vectashield.

Fluorescence widefield microscopy was performed on a Leica DM IL, equipped with a HC Fluotar 100x oil immersion objective with a 1.32 numerical aperture and planar correction (Leica; 11506527), a white LED light source (Leica; EL6000) and a DFC365 FX digital camera (Leica; 11547004). The following filter cubes were used: Texas Red (Leica TX2 ET; 11504180; used with Dylight-550 conjugates), Cy5 (Leica Y5 ET; 11504181, used for Alexafluor-647), GFP (Leica GFP ET; 11504174,

used for Alexa Fluor-488), and DAPI (Leica DAPI ET; 11504204). Illumination intensity was adjusted using the light source manual control; acquisition times ranged from 40–2000 ms, as controlled by the Leica LASX software.

Fluorescence confocal microscopy was performed on a Leica SP8X microscope (UW Keck Imaging Center), outfitted with a HC CS2 63x oil immersion objective, with 1.40 numerical aperture with both planar and apochromatic correction. The average voxel size was 0.06 x 0.06 x 0.346 µm. Samples were illuminated using a 470–670nm tunable White Light Laser system, with a typical laser power of 0.1% for DAPI, 3% for 550 nm, and 30% for 647 nm. Gain and offset settings were adjusted to avoid pixel saturation. Images were line-averaged twice, with an average pixel dwell time of 1.58 µs. A bit-depth of 8 or 16 was used and zoom factor between 1-3 was used for all images.

Image processing

Images were processed using Fiji[133] and ImageJ[134], and multicolor overlays were made using the screen setting in Adobe Photoshop[16]. Most confocal images are maximum projections of z-stacks; the remainder correspond to single z-slices. Brightness and contrast were adjusted for display purposes using Fiji and ImageJ or Adobe Photoshop. In all cases, contrast adjustment was applied to improve signal visibility, by changing the minimum (black) and maximum (white) values only. Automated despeckling was applied when necessary (*e.g.* in RNA-FISH images with weak, diffuse speckling in between cells) to reduce residual background signal.

2.6 Acknowledgements

This chapter contains experiments and work that was reprinted and adapted Tsue, A. F., Kania E. E. *et al.* Oligonucleotide-directed proximity-interactome mapping (O-MAP): A unified method for discovering RNA-interacting proteins, transcripts and genomic loci in situ. *bioRxiv* 2023.01.19.524825 (2023). Thank you to E. Kania and D. Shechner on the O-MAP development team and writing contributions; J. Gianopulos, C. Hsu, C. Anderson, and J. Cabarrus for general technical assistance; R. Akilesh, P. Valdmanis, and S. Smukowski for the kind donation of materials; N. Peters, K. Collins, M. Soruco, E. Hacisüleyman, and A. Fenix for imaging assistance; S. Attar, E. Nichols, for help with cell-based models; H-T. Lau, M. G. Golkowski, S-E. Ong, R. Fields, C. McGann, and D. Schweppe for mass-

spec assistance; Y. Sancak, R. Gardner, and D. Shechner lab members for thoughtful discussions and critiques.

2.7 Figures
Figure 1

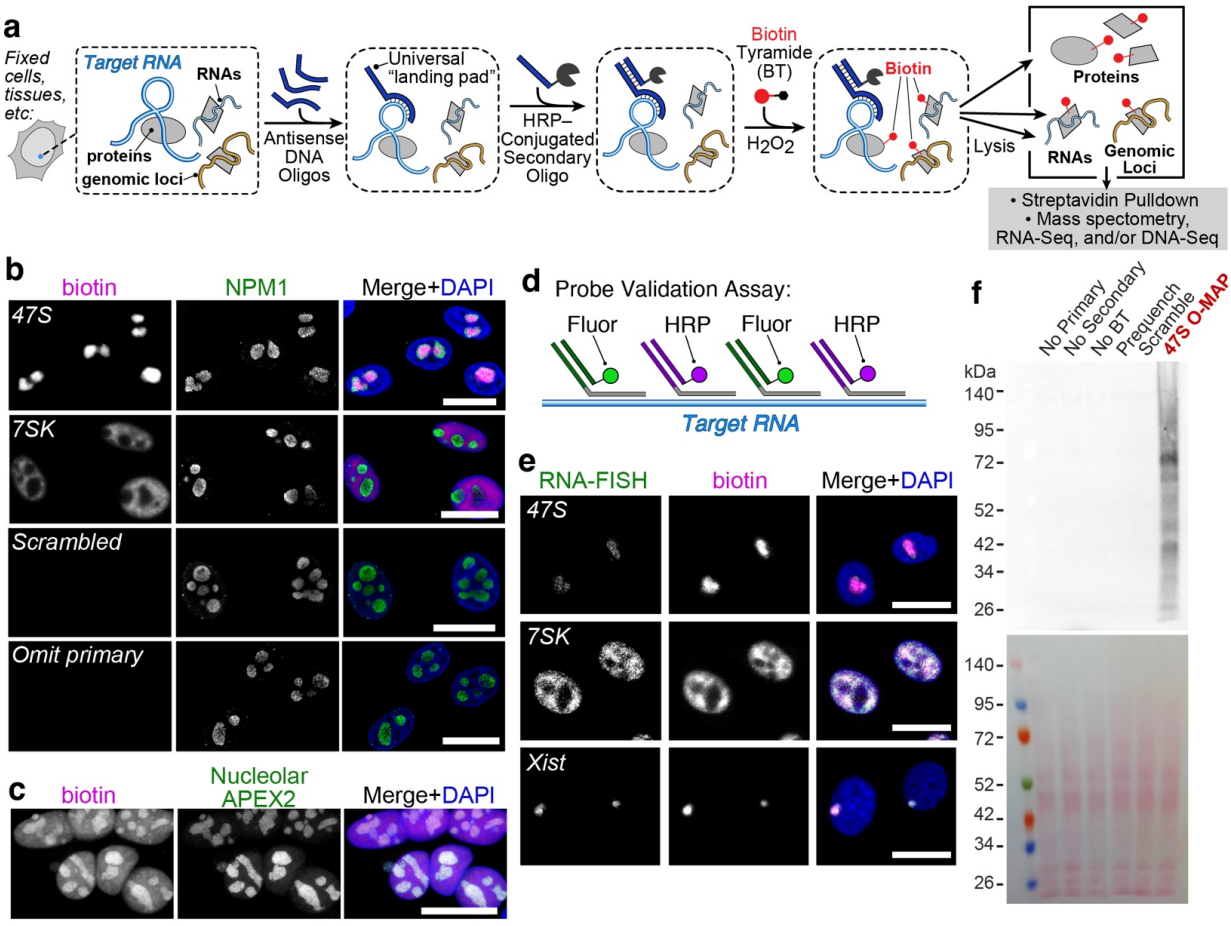


Figure 1. O-MAP Design and Implementation. **a,** Overview of O-MAP. Specimens are chemically fixed, and pools of antisense DNA probes are hybridized to the target RNA. These probes recruit a common, HRP-conjugated secondary probe that catalyzes *in situ* proximity-biotinylation. **b,** O-MAP enables precise RNA-targeted proximity labeling, enabling interaction-discovery. *In situ* biotinylation imaged by neutravidin staining; *NPM1* by immunofluorescence. Note nucleolar biotinylation in 47S O-MAP, nucleoplasmic biotinylation in 7SK O-MAP. **c,** Nucleolar-targeted APEX2 exhibits substantial off-target, nucleoplasmic labeling[65]. **d,** O-MAP Probe Validation Assay. Primary probes are split into sub-pools that enable O-MAP and RNA-FISH to be performed simultaneously. Lack of co-localization suggests probe off-targeting. **e,** Probe Validation Assays on 47S-pre-rRNA and 7SK, in HeLa cells, and *Xist*, in mouse Patski cells. **f,** Recovery of O-MAP-biotinylated proteins. *Top:* Streptavidin-HRP blot of whole cell lysates. Note ladder of biotinylated proteins from 47S O-MAP *Bottom:* Ponceau stain. All scale bars: 20 μm.

Figure 1 Figure Supplement 1

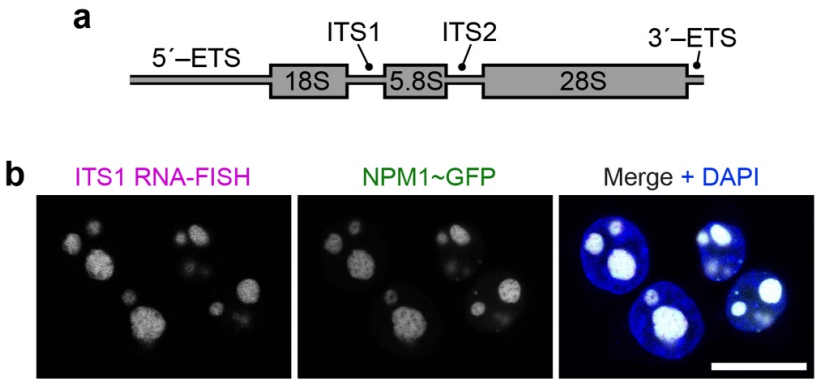


Figure 1- Figure Supplement 1. RNA-FISH validation of the 47S pre-rRNA-targeting probe set. a, Schematic of the 47S pre-rRNA. During ribosome biogenesis, the 18S, 5.8S, and 28S domains are processed and incorporated into mature ribosomes, while the 5´– and 3´–External Transcribed Spacers (5´–ETS and 3´–ETS) and Internal Transcribed Spacers (ITS1 and ITS2) are cleaved from the precursor transcript and degraded within the nucleolus. The 47S pre-rRNA probe set used in the initial stages of O-MAP development target ITS1. **b,** validation of this probe set. HEK293T cells were transiently transfected with GFP-tagged *NPM1*, a nucleolar marker, and subjected to conventional RNA-FISH. Note conspicuous overlap between the *NPM1*~GFP and ITS1 RNA-FISH signals. Scale bar: 20 μm.

Figure 1 Figure Supplement 2

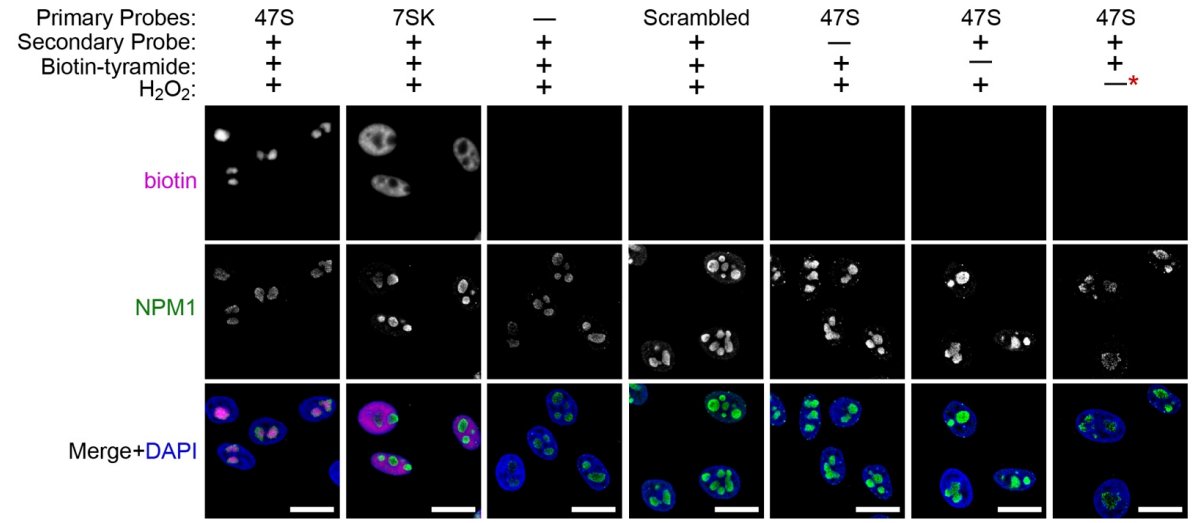


Figure 1- Figure Supplement 2. More O-MAP controls. O-MAP and negative control experiments were performed in HeLa cells, as indicated. Biotin was imaged using a fluorescent neutravidin conjugate; *NPM1* via immunofluorescence. Note that omitting any component of the O-MAP pipeline ablated biotinylation signal. The 47S-O-MAP, 7SK-O-MAP, omit primary and scrambled primary conditions (*left four columns*) are the same images presented in (**Fig. 1b**). In the "omit-H_2O_2" condition (*far right, marked* *), cells were pre-quenched with sodium azide and ascorbic acid prior to the addition of biotin-phenol and H_2O_2. Simply removing H_2O_2 from the O-MAP protocol still resulted in targeted *in situ* biotinylation (*i.e.* nucleolar labeling using 47S probes), presumably due to photoactivation of HRP. Scale bars, 20 μm.

Figure 1 Figure supplement 3

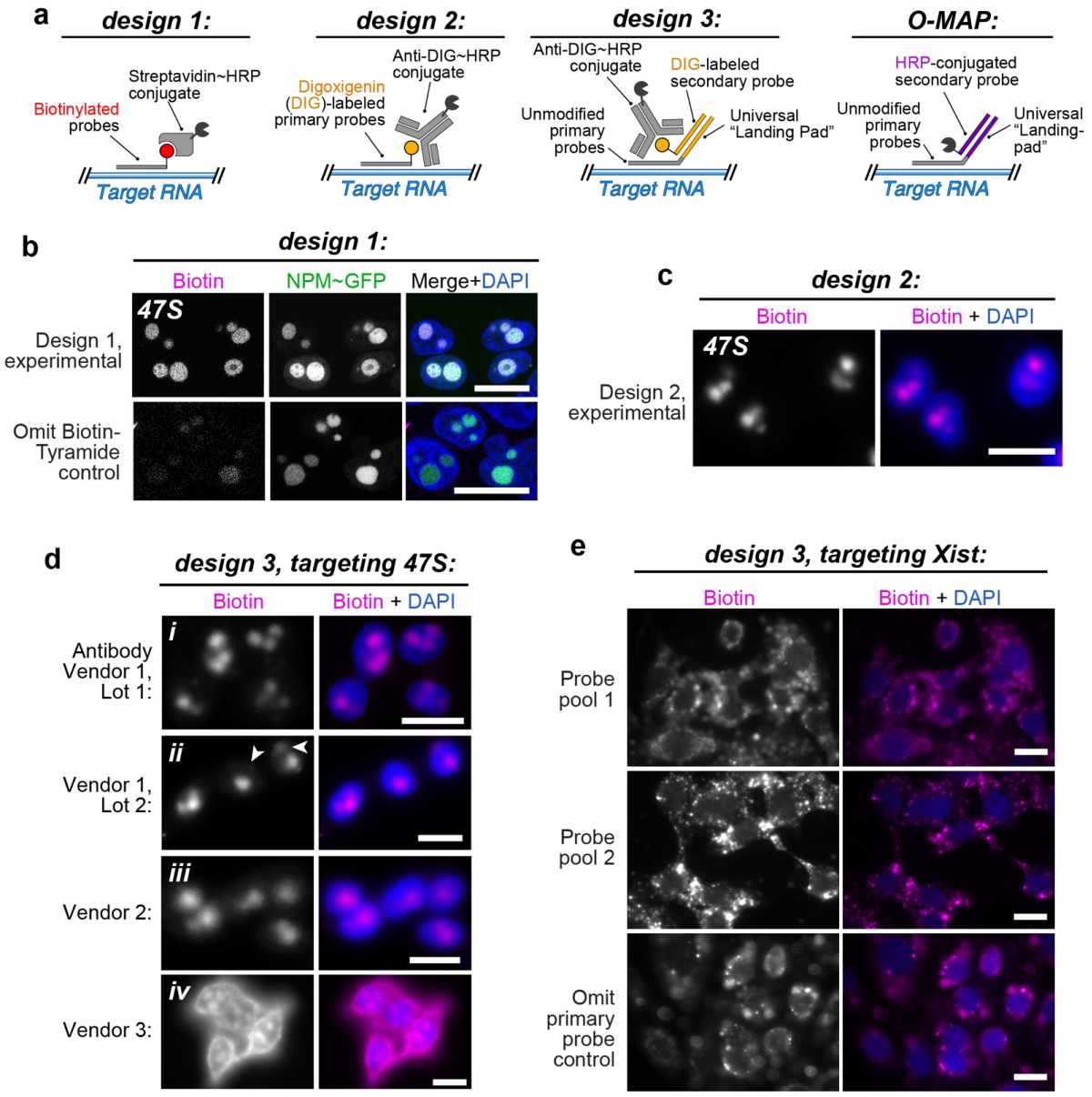


Figure 1- Figure supplement 3. Overview and limitations of preliminary O-MAP designs. a, summary of RNA-targeted HRP-recruitment strategies tested. Design 1 uses biotinylated primary probes to recruit a streptavidin-HRP conjugate. Designs 2 and 3 use Digoxigenin (DIG)-labeled primary or secondary probes to recruit an HRP–conjugated anti-DIG antibody. Our final O-MAP design, which uses HRP-conjugated oligo probes, is shown for comparison (*see also*, **Fig. 1a**). The same anti-DIG antibody is used in designs 2 and 3; the same "universal landing pad" sequences are used in designs 3 and 4. **b–e**, limitations of Designs 1–3. In all cases biotin was visualized by staining with a fluorescent neutravidin conjugate. **b,** Design 1 was disfavored because *in situ* biotinylation cannot be unambiguously distinguished from biotinylated primary probes. HeLa cells over-expressing NPM1~eGFP were probed using the same 47S-targeting probes as in (**Fig. 1**) and the main text, appended on their 3′–termini with biotin. Note nucleolar biotin signal even in the absence of biotin-tyramide (*bottom panels*). We anticipated that this background signal would be especially problematic with low-abundance target RNAs. **c,** Design 2—analogous to HyPro (PMID: 35457249)—was sometimes capable of producing well-resolved nucleolar-targeted biotinylation.

HeLa cells are shown. This approach was eventually disfavored due to antibody irreproducibility issues described below, and because the high cost of DIG-labeled oligos would limit its use with low-abundance transcripts, which can require dozens to hundreds of probes. **d,** Design 3 overcomes the oligo cost issue but still suffers from antibody background binding and irreproducibility. HeLa cells were probed with the same 47S-targeting primary probe sets used in the main text (appended with the same "landing pad" modules), a DIG-labeled secondary oligo, and four different lots or vendors of commercial HRP-conjugated antibodies. In some cases, we observed well-resolved RNA-targeted biotinylation (*panel i*), though other lots from the same vendor exhibited off-target labeling (*panel ii, arrows*). Regents from other vendors exhibited varying degrees of spatial blurring (*panel iii*), or conspicuous off-target biotinylation that rivaled or exceeded the target signal (*panel iv*). Compare these results to (**Fig. 1b**). **d,** Design 3 is particularly problematic with lower-abundance RNA targets. Patski cells were probed with the same *Xist*-targeting, landing-pad-extended probes as used in the main text, divided into two sub-pools. Anti-DIG-HRP was from Vendor 1. Note that all conditions—both probe sub-pools and the omit-primary negative control–induced substantial off-target biotinylation. Compare these results to (**Fig. 1e**). All scale bars, 20 μm.

Figure 2.

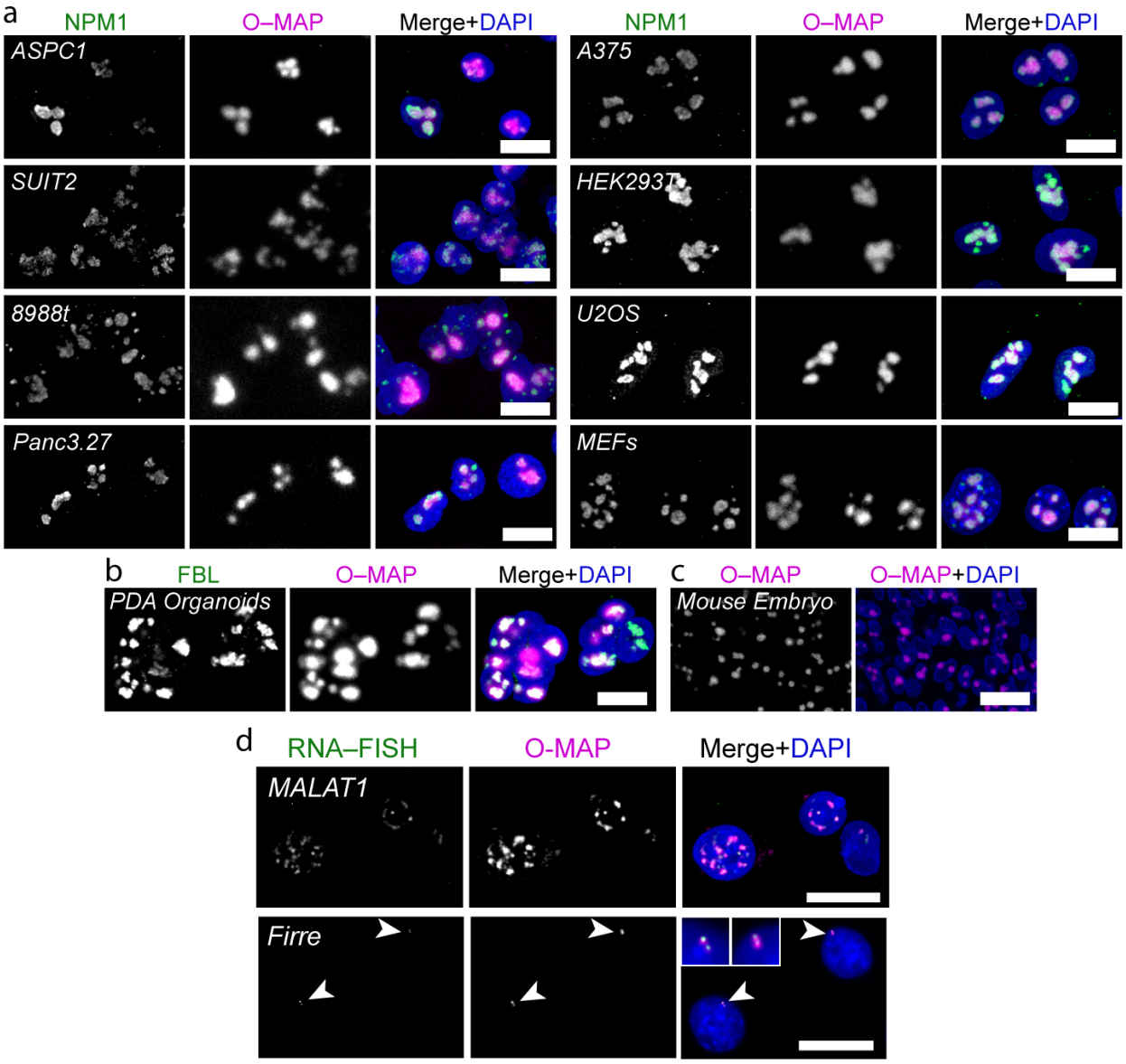


Figure 2. **O-MAP is readily ported across specimen types**. **a,** 47S-O-MAP in cultured mammalian cell lines. NPM1 immunofluorescence denotes nucleoli. All human-derived lines used the same probe set and hybridization conditions as HeLa cells (**Fig. 1**); MEFs used an analogous mouse-targeting probe set. **b,** 47S O-MAP in human patient-derived pancreatic ductal adenocarcinoma (PDA) organoids. *FBL* immunofluorescence denotes nucleoli. **c,** 47S O-MAP in cryo-preserved mouse tissue slices. **d,** O-MAP Probe Validation Assay (**Fig. 1d**) applied to a compendium of target transcripts. Note conspicuous overlap between RNA–FISH (*green*) and O-MAP (*magenta*) signals. Images from HeLa (*MALAT1*) and Patski (*Firre*). *Firre* is expressed from a single-copy transgene; all other targets are endogenous transcripts. Insets show zoomed-in sections of the same images, to highlight signal overlap. All scale bars: 20 μm.

Figure 2- Figure supplement 1:

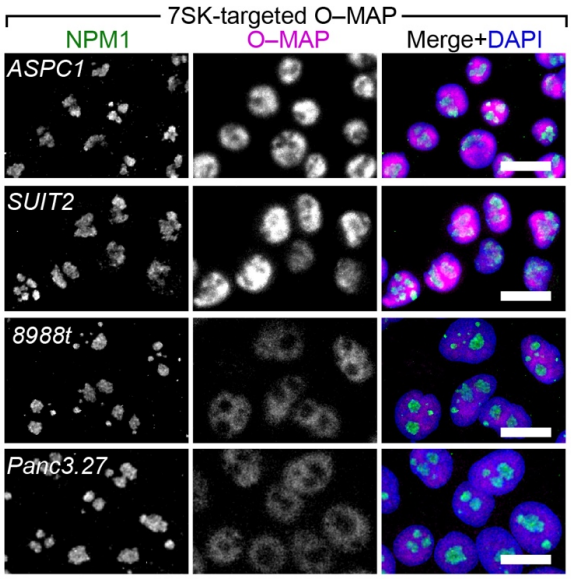


Figure 2- Figure supplement 1: ***7SK O-MAP in cultured Pancreatic Ductal Adenocarcinoma (PDA) cell lines.*** *O-MAP was visualized using a fluorescent neutravidin conjugate;* NPM1 *by immunofluorescence,* as in (**Fig. 2**)*. Scale bars: 20 μm.*

Figure 2- Figure supplement 2:

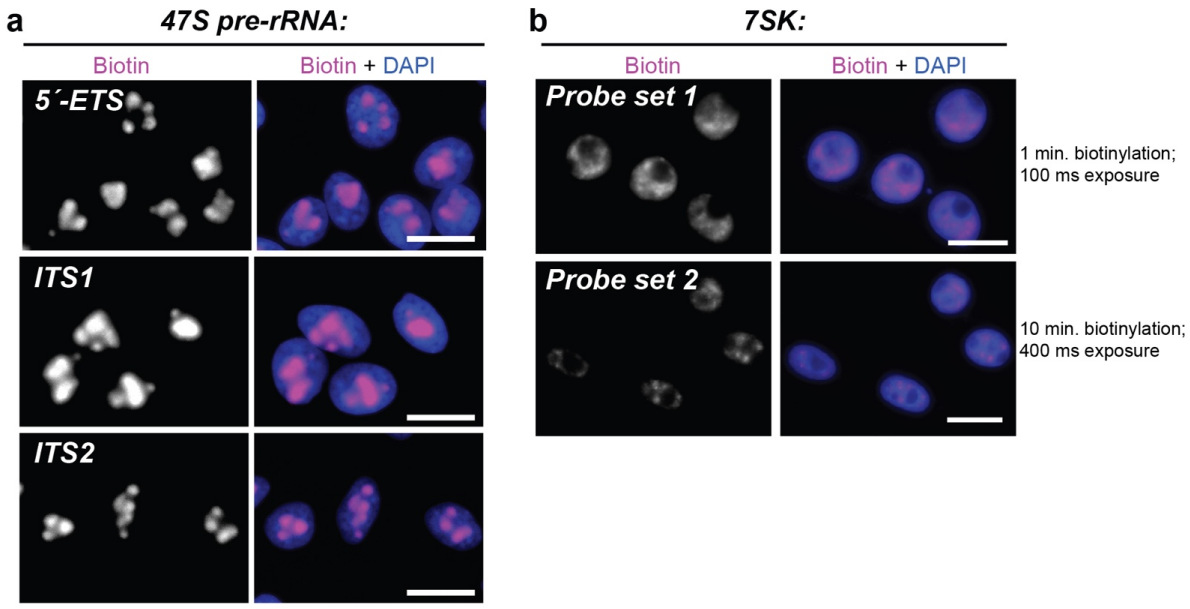


Figure 2- Figure Supplement 2: Reproducibility of O-MAP labeling across primary probe sets. a, Targeting the 47S pre-rRNA. Probe sets targeting the 5´–ETS, ITS1, and ITS2 Transcribed Spacer domains produced similar patterns of nucleolar *in situ* biotinylation. Probes targeting the 3´–ETS yielded no signal (*data not shown*). **b,** Targeting 7SK. Each set targets the entirety of the 7SK transcript, but were designed to have different hybridization parameters. Note that *in situ* biotinylation and exposure times differed between 7SK probe sets, as indicated (*right*). Biotin was visualized by staining with fluorescent neutravidin, in HeLa cells. Scale bar: 20 μm.

Figure 3.

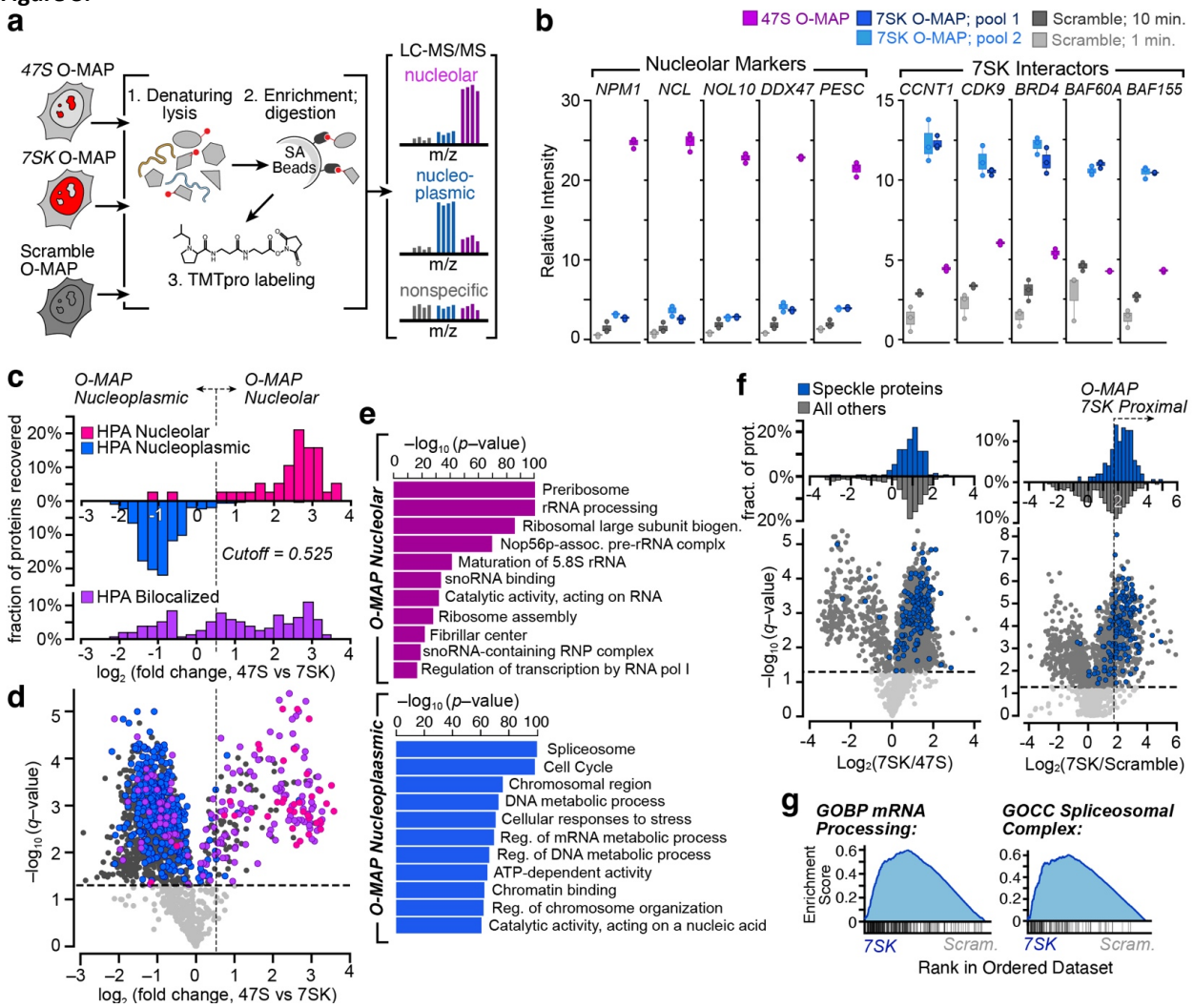


Figure 3. O-MAP-MS for probing RNA-proximal proteomes. a, Strategy for characterizing the HeLa nucleolar proteome. Parallel O-MAP experiments targeting the 47S (nucleolar), 7SK (nucleoplasmic) or using scrambled probes (background) were processed as indicated. Samples were quantified by TMTPro mass-spectrometry; each replicate (n = 3) was labeled with a unique mass-tag. **b–g**, analysis of a single-shot 47S/7SK O-MAP-MS experiment, using 5.5x10^6 cells per replicate. **b,** Recovery of known nucleolar proteins (*left*) and 7SK interactors (*right*). Two different probe sets, requiring different labeling times, were used in 7SK experiments; these were time-matched to scrambled-probe controls. **c–d,** Global enrichment of known nucleolar (*magenta*), nucleoplasmic (*blue*), and bi-localized (*purple*) marker proteins, defined by the Human Protein Atlas (HPA). Dotted line denotes the optimal threshold separating the compartments, determined by Receiver-Operating Characteristic (ROC) analysis (**Supplementary Fig. 2**). **c,** Histograms plotting the enrichment of each sub-compartmental proteome. Only proteins with q-value ≤ 0.05 are shown.Note conspicuous separation between nucleolar and nucleoplasmic markers, and multi-modal distribution of bi-localized proteins. **d,** Volcano plot showing all data. **e,** Gene Ontology (GO) analysis for the O-MAP-nucleolar (*top*) and O-MAP-Nucleoplasmic (*bottom*) proteomes, defined from ROC analysis. The top eleven most enriched terms are shown. **f,** 7SK O-MAP enriches the Nuclear Speckle proteome. Histograms and volcano plots as in **c,d**, showing enrichment of speckle proteins in both 7SK-vs-47S and 7SK-vs-Scramble comparisons. (*continued on next page)*

Figure 3. (*continued*) The latter comparison more clearly distinguished speckle proteins from the broader nucleoplasm and was used in ROC analysis to determine the optimal threshold cutoff (**Supplementary Fig 3**). **g,** Gene Set Enrichment Analysis (GSEA) of the "O-MAP 7SK-proximal" proteome reveals a striking

enrichment in RNA splicing-relevant Gene Ontology Biological Process (GOBP) and Cellular Component (GOCC) terms.

Figure 3 Figure Supplement 1

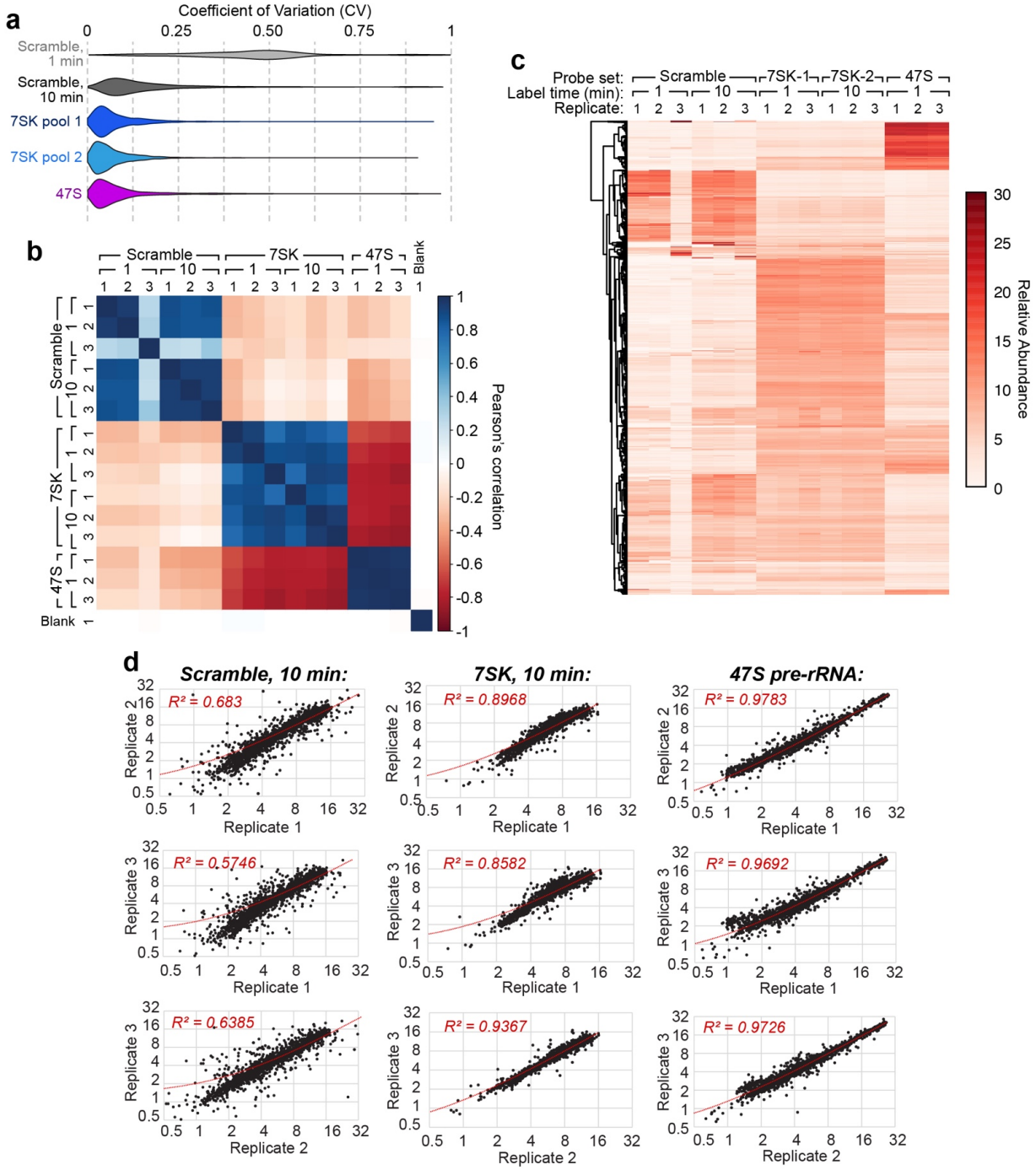


Figure 3- Figure supplement 1: Reproducibility of O-MAP-MS. a, Average coefficients of Variation (CV) between biological replicates (*n* = 3). Conventionally, CV's below 0.2 are considered acceptable. **b,** Correlation matrix between all samples. Note high positive correlations between biological replicates, highlighting O-MAP's reproducibility, and negative correlations between different sample types (*e.g.* 47S vs 7SK 1 minute), highlighting the unique composition of the proteome captured for each target. **c,** Heatmap of all proteomic data, clustered by unsupervised hierarchical clustering, showing defined clades of 47S-proximal and 7SK-proximal proteomes. Note that the 7SK 1-minute labeling condition used probe set 1; the 10-minute condition used probe set 2.

d, Scatter plots comparing scaled protein abundances for individual biological replicates for the Scramble negative control (10 minutes' labeling time, *left*), 7SK probe set two (*center*), and 47S (*right*). Data are plotted on log2 scale.

Figure 3 Figure Supplement 2

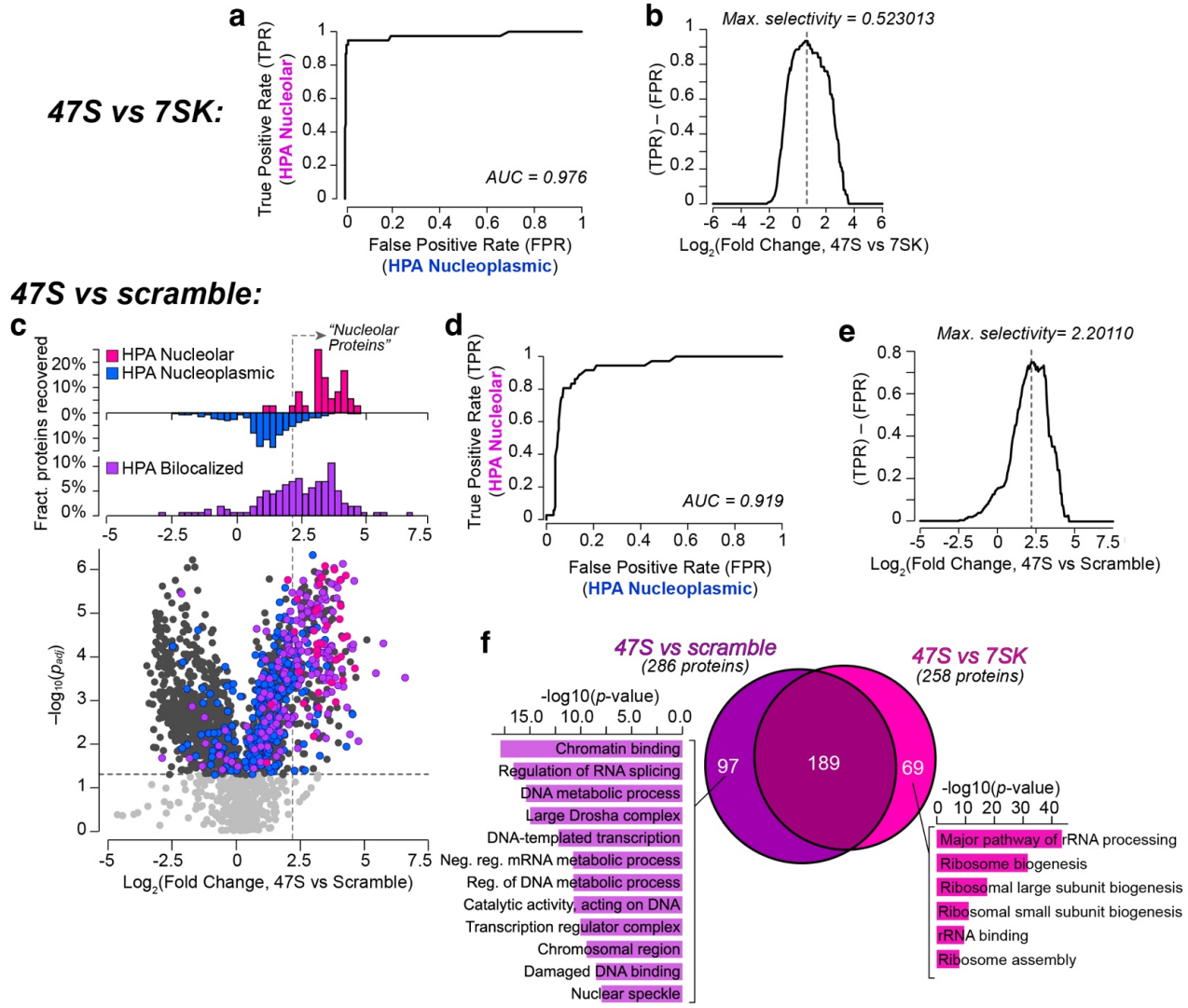


Figure 3- Figure Supplement 2: Probing the nucleolar proteome with 47S O-MAP-MS. a, Receiving-Operator Characteristic (ROC) analysis of the 47S vs 7SK O-MAP-MS experiment. True Positive and False Positive proteins were defined using lists of exclusively nucleolar and exclusively nucleoplasmic proteins, respectively, as reported by the Human Protein Atlas (HPA). An Area Under the Curve (AUC) of nearly 1.0 suggests strong and highly sensitive selectivity for nucleolar proteins over the nucleoplasmic proteome. **b**, These data were used to derive an optimal Log_2(fold change, 47S/7SK) cutoff value of 0.523, and to define a putative list of 258 O-MAP core nucleolar proteins, as described in (**Figure 2**). **c–e**, parallel analysis using (47S/Scramble controls), instead of (47S/7SK). **c,** Volcano plot and histograms of showing the enrichment of HPA-nucleolar, HPA-Nucleoplasmic, and HPA-bilocalized proteins, using the same protein marker reference lists as in (**a–b**), and (**Fig. 2c**,**d**). **d,** ROC analysis of the (47S/Scramble) data demonstrates slightly lower sensitivity than that of the (47S/7SK) analysis, though still exceptionally. **e**, As in (**b**), these data were used to determine an optimal Log_2(fold change, 47S/Scramble) cutoff value of 2.201, defining a putative cohort of 286 O-MAP core nucleolar proteins. **f,** the putative nucleolar proteomes derived from the (47S/7SK) and (47S/Scramble) ROC analyses show considerable overlap (66–73%). Outliers were used for Gene Ontology (GO)-term analysis. Factors uniquely captured by the (47S/7SK) analysis were highly enriched for ribosome biogenesis factors, while those unique to the (47S/Scramble) analysis were enriched

for nucleoplasmic functions. This suggests that the (47S/7SK) comparison more precisely captures the nucleolar proteome.

Figure 3 Figure Supplement 3

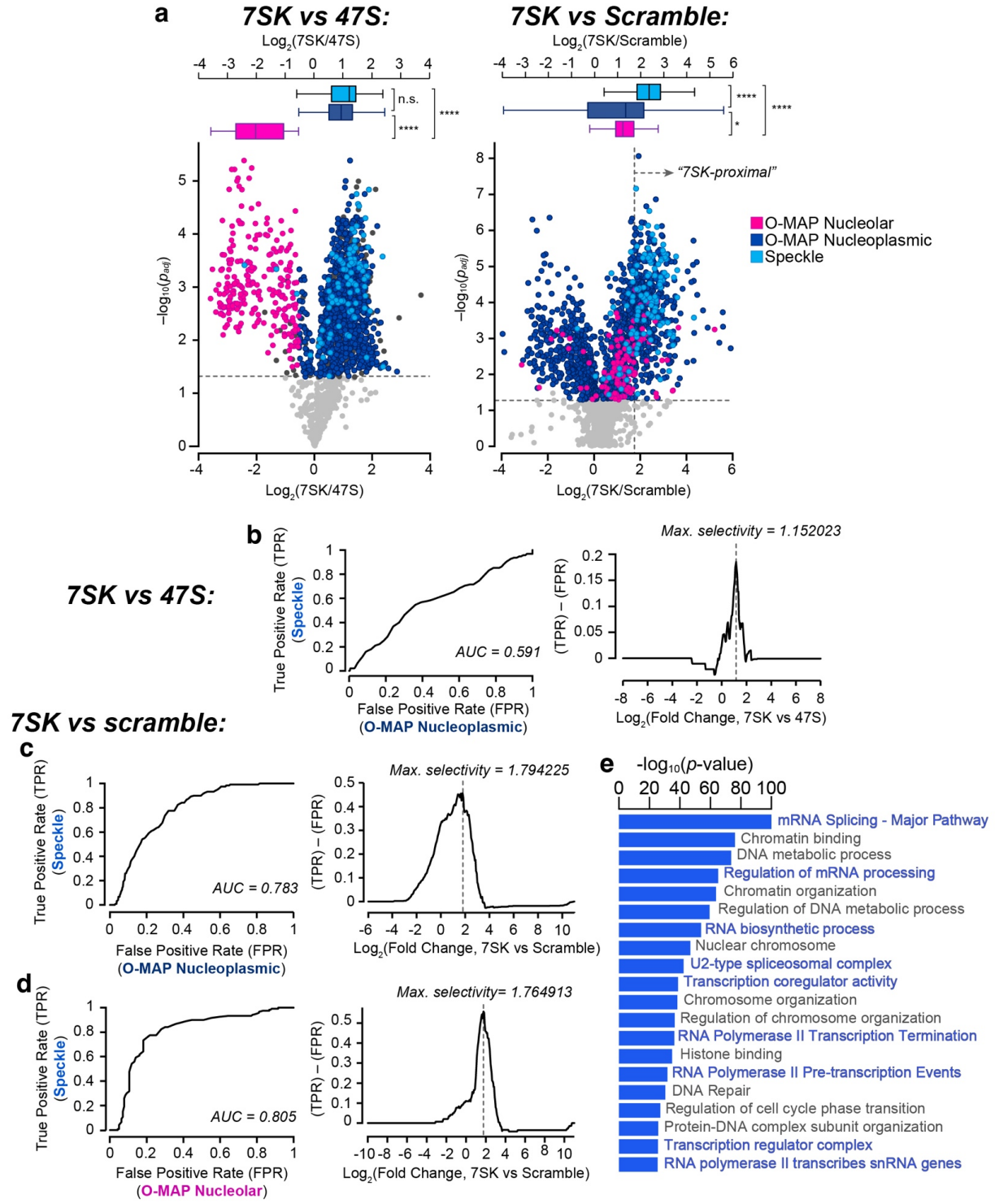


Figure 3-Figure Supplement 3: Probing the 7SK-proximal proteome with O-MAP-MS. a, Identifying the optimal basis of comparison for probing the 7SK-proximal compartment: (7SK/47S) or (7SK/Scramble). For each comparison, volcano plots demonstrate the enrichment of the nucleolar- and nucleoplasmic-proteomes, as derived from O-MAP analysis (**Fig. 3** and **Fig. 3. Fig. Supplement 2**), and of known Nuclear Speckle proteins, markers of the 7SK-proximal compartment. Box-whisker plots (*above*) summarize the

distribution of each protein group (significance testing: two-tailed, heteroscedastic Student's t-tests; n.s: not significant, * p=0.05; **** $p < 1x10^{-5}$). In each comparison, speckle proteins are significantly enriched relative to the nucleolar proteome, but in the (7SK/47S) comparison (*left*) these proteins are indistinguishable from the broader nucleoplasmic proteome (p = 0.14). In contrast, in the (7SK/Scramble) comparison (*right*) Speckle proteins are significantly enriched relative to the nucleoplasmic outgroup (p= $5x10^{-10}$). This suggests that (7SK/Scramble) is more selective for 7SK-proximal interactors over the broader nucleoplasmic proteome. This is corroborated by Receiver-Operating Characteristic (ROC) analysis, as described below, which was used to define the "7SK-proximal" cutoff in the righthand panel. **b,** ROC analysis of (7SK/47S), using speckle proteins as true positives and O-MAP-nucleolar proteins as false positives. An area under the curve (AUC) of approaching 0.5 suggests a nearly complete absence of signal. **c,** ROC analysis of (7SK/Scramble), using the same True Positive and False Positive lists as in (**b**), suggests a strong and specific separation between the protein populations, and an optimal Log_2(fold change) cutoff of ≥ 1.794. **d**, A nearly identical result is obtained when using O-MAP-nucleolar proteins as False Positives. These cutoffs define a cohort of 510 putative 7SK-proximal proteins. **e,** Gene Ontology (GO)-term analysis on this cohort is highly enriched for biological processes involved in pre-mRNA biogenesis, highlighted in blue.

Figure 3 Figure Supplement 4

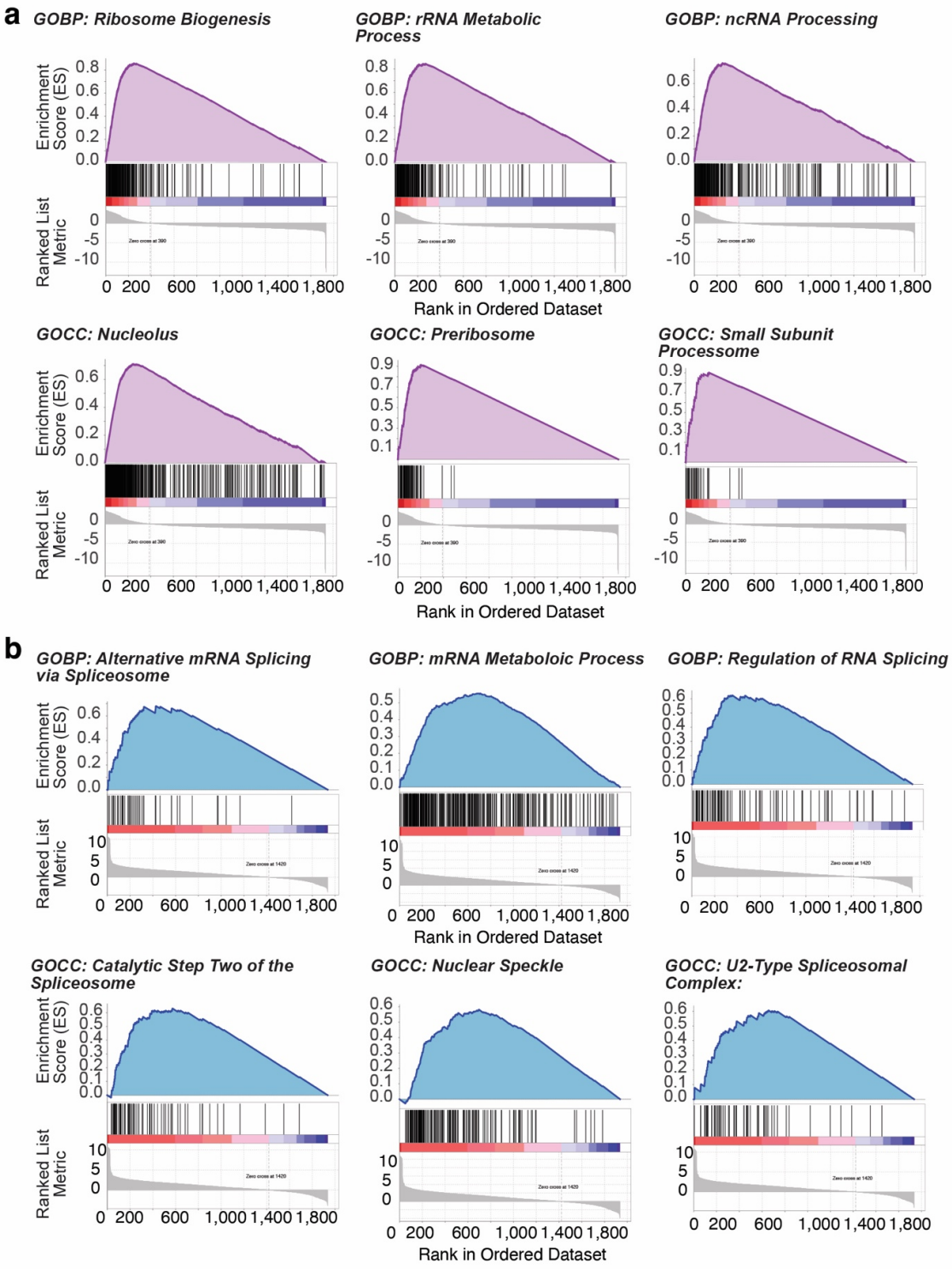
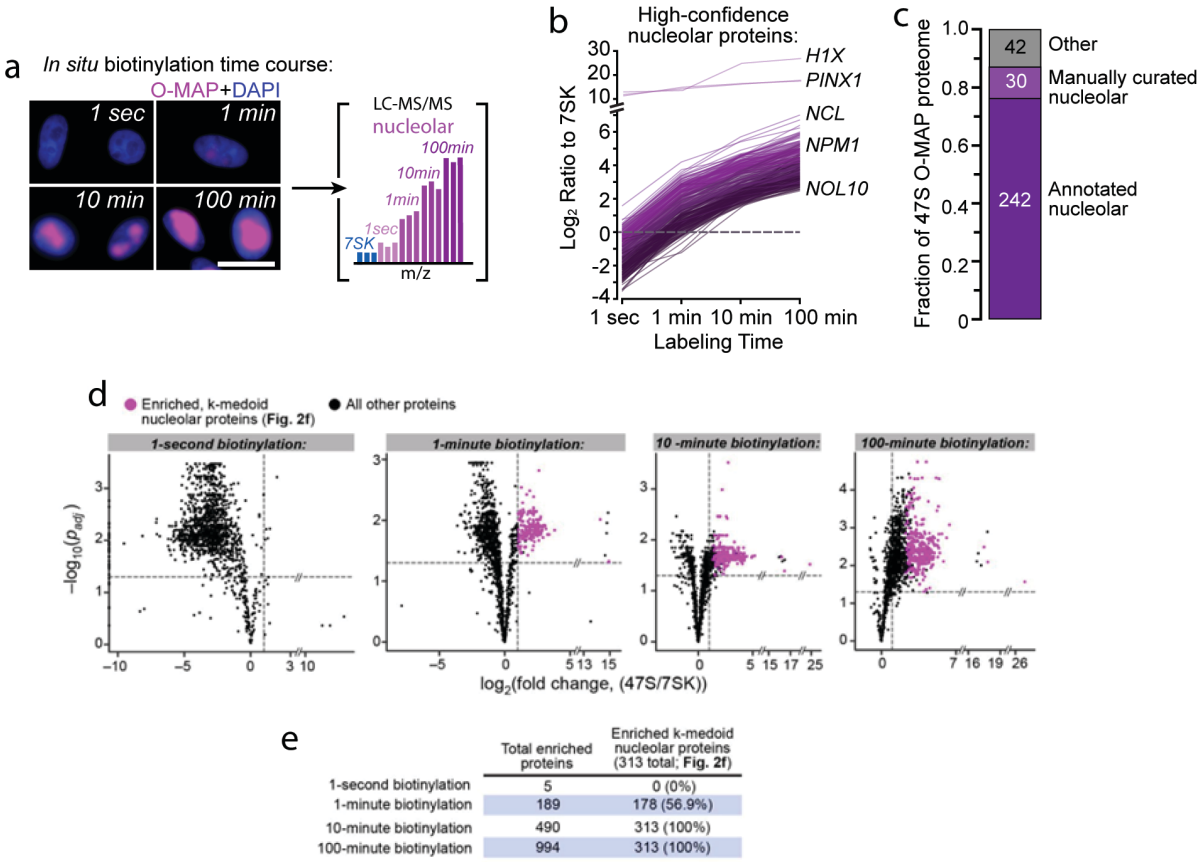


Figure 3-Figure Supplement 4: Representative highly ranked Gene Set Enrichment Analysis (GSEA) results. a, GSEA results comparing 47S and 7SK O-MAP-MS. The same ranked gene list was used for all queries. Top Gene Ontology Biological Process (GOBP) and Cellular Component (GOCC) terms are shown. For all, left-ranking genes (*red*) were more highly enriched from 47S O-MAP than 7SK O-MAP; those ranked toward the right (*blue*) were more enriched from 7SK O-MAP. **b,** GSEA results comparing

7SK O-MAP--MS to Scrambled controls. The same gene list was used for all queries. Leftmost genes were 7SK-enriched; rightmost were Scramble-enriched.

Figure 4



e

	Total enriched proteins	Enriched k-medoid nucleolar proteins (313 total; **Fig. 2f**)
1-second biotinylation	5	0 (0%)
1-minute biotinylation	189	178 (56.9%)
10-minute biotinylation	490	313 (100%)
100-minute biotinylation	994	313 (100%)

Figure 4. Coverage of the nucleolar proteome during the 47S O-MAP labeling time course. . **a–c**, Higher coverage of the nucleolar proteome using an O-MAP *in situ* biotinylation time course. **a,** Approach. 47S O-MAP was performed in parallel for the indicated times, enriched, and TMTPro-labeled as above (*n* = 3). A single 7SK probe set and time point was used for normalization. *In situ* biotinylation visualized by neutravidin staining. Scale bars: 20 μm. **b,** k-medoid clustering yields a clade of 313 high-confidence nucleolar proteins. Notable marker genes are indicated. **c,** nearly all (87%) members of the nucleolar medoid group corresponds to annotated or manually curated nucleolar proteins. **d,** Volcano plots for each 47S O-MAP labeling point, calculated relative to the 7SK/10-minute label condition. Enrichment of nucleolar proteins derived from our k-medoid analysis (**Fig. 3e–g; Fig. 3 Fig. Supplemental 4**) are highlighted in pink. **e,** Table summarizing the recovery of the nucleolar proteome (defined either by our first-pass analysis, **Fig. 3b–d**, or k-medoid analysis, **Fig. 4 Fig. Supplemental 1**), at each labeling time point. Note that coverage appears to plateau at 10 minutes.

Figure 4 Figure Supplement 1

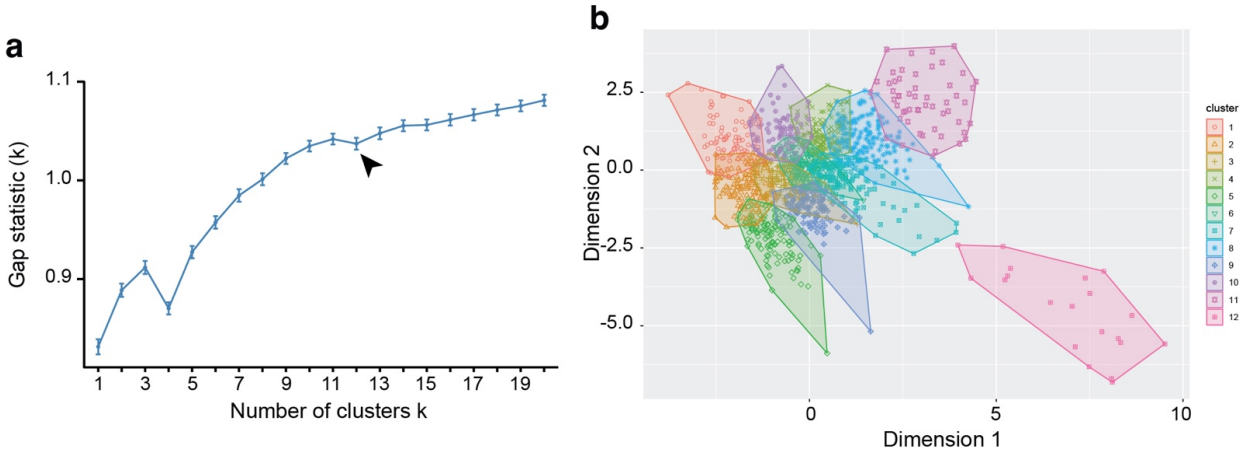


Figure 4- Figure Supplement 1: k-medoid clustering. a., Determining the optimal number of clusters. Ultimately, 12 clusters were chosen, given the dip in experimental signal (*arrowhead*). **b,** Principle component analysis (PCA) of the twelve k-medoid clusters.

CHAPTER 3: O-MAP: FUTURE APPLICATIONS IN CANCER AND NUCLEOLAR BIOLOGY
3.1 NUCLEOLAR BIOLOGY AND O-MAP

Nucleolar O-MAP and stress responses

The nucleolus operates primarily as a site for ribosome biogenesis but also harbors many dynamic roles in the cell cycle, DNA replication and repair, nuclear architecture, cellular stress responses, and a significant role in cancer[135]. During the mammalian cell cycle, nucleoli rapidly disassemble and reform around clusters of rDNA to form subnuclear compartments partitioned by ribosomal transcription and processing[11]. Driven by ribosome biogenesis, the structure of these membraneless nucleoli is scaffolded by the 47S pre-ribosomal RNA (47S pre-rRNA) that are dynamically interacting with RNA-binding proteins (RBPs) and chromatin loci. Under pathological or stress conditions, these ribonucleoprotein (RNP) interactions drastically change, and macroscopically, the nucleolus structure and morphology is reorganized[136]. However, there is a poor understanding of the significance of these RNP changes as well as the molecular mechanisms that drive nucleolar adaptive responses[137].

The dynamic responses of nucleoli to external stress is thought to be driven by the nucleolar role as a liquid-liquid phase separated body. For instance, throughout the mammalian cell cycle, nucleoli disassemble and reassemble around Nucleolar Organizing Regions (NORs) and Prenucleolar Bodies (PNBs) that contain many ribosome biogenesis factors[11,138]. Nucleolar assembly depends on the dynamic recruitment of existing pre-ribosomal RNA (pre-rRNA) and other nucleolar RNPs to sites to begin active rDNA transcription and form the classic nucleolar tripartite structure.

Relating to its dynamic LLPS nature, nucleoli possess dramatic cellular stress responses, and they undergo proteomic and structural changes upon exposure to heat stress[136]. Current technologies study end point analysis of stressed nucleolar proteomes through biochemical fractionation, or by pulling down abundant nucleolar proteins such as nucleophosmin (*NPM1*)[25,139]. From these studies, extranucleolar proteins, RNA, stress-responsive proteins (e.g. HSP70) are translocated and detained within the nucleolus, expanding the organelle and forming intranucleolar macromolecular collections resembling amyloid bodies[25,140]. Additionally, the nucleolus drastically changes morphology upon other types of environmental stress such as transcription inhibition and DNA damage. The nucleolar tripartite structure formed from nucleolar RNP interactions and liquid-liquid phase separation is disrupted to form different morphologies such as nucleolar caps[141] that DNA damage repair proteins. Yet studying these dynamic

responses, their mechanisms driving protein translocation, and overall changes in nucleolar morphology are difficult to study due to limitations in technology where disrupted nucleoli are challenging to isolate.

To address this, proximity labeling technologies can be used to deeply characterize the temporal dynamics of nucleoli remodeling, the recruitment and immobilization of proteins, and the protein quality control machinery recruited to the nucleolus (e.g. Hsp70) required for the nucleolar heat stress, DNA damage or even cancer remodeling and response to treatment. The technology featured in this thesis, Oligonucleotide directed proximity interactome MAPping (O-MAP), has shown a deep characterization of nucleoli and provides a possible tool for studying make up of disrupted nucleoli in response to stress. **Figure 3.1a** demonstrates preliminary experiments of nucleolar O-MAP in HeLa cells that are treated with heat shock, demonstrating specific nucleolar signal. This labeling supports the ability of O-MAP biotinylating enzymes to label within crowded, protein aggregated nucleoli[25]. O-MAP is also able to specifically label nucleoli disrupted by pharmacological treatment by 5,6-dichloro-1-beta-D-ribofuranosylbenzimidazole (DRB), a potent inhibitor of CDK7 and CDK9 and pol II transcription[142] (**Fig. 3.1b**). These imaging experiments support the ability of O-MAP to be used in pharmacologically disrupted states, which would be difficult in the case of live-cell proximity labeling[107] that is dependent on transgenic expression and localization of biotinylating enzymes. Since cells are crosslinked, O-MAP can be further expanded to a temporal analysis studying the dynamic nuclear heat shock response as well as disrupted nucleolar states after drug treatments.

Figure 3.1.

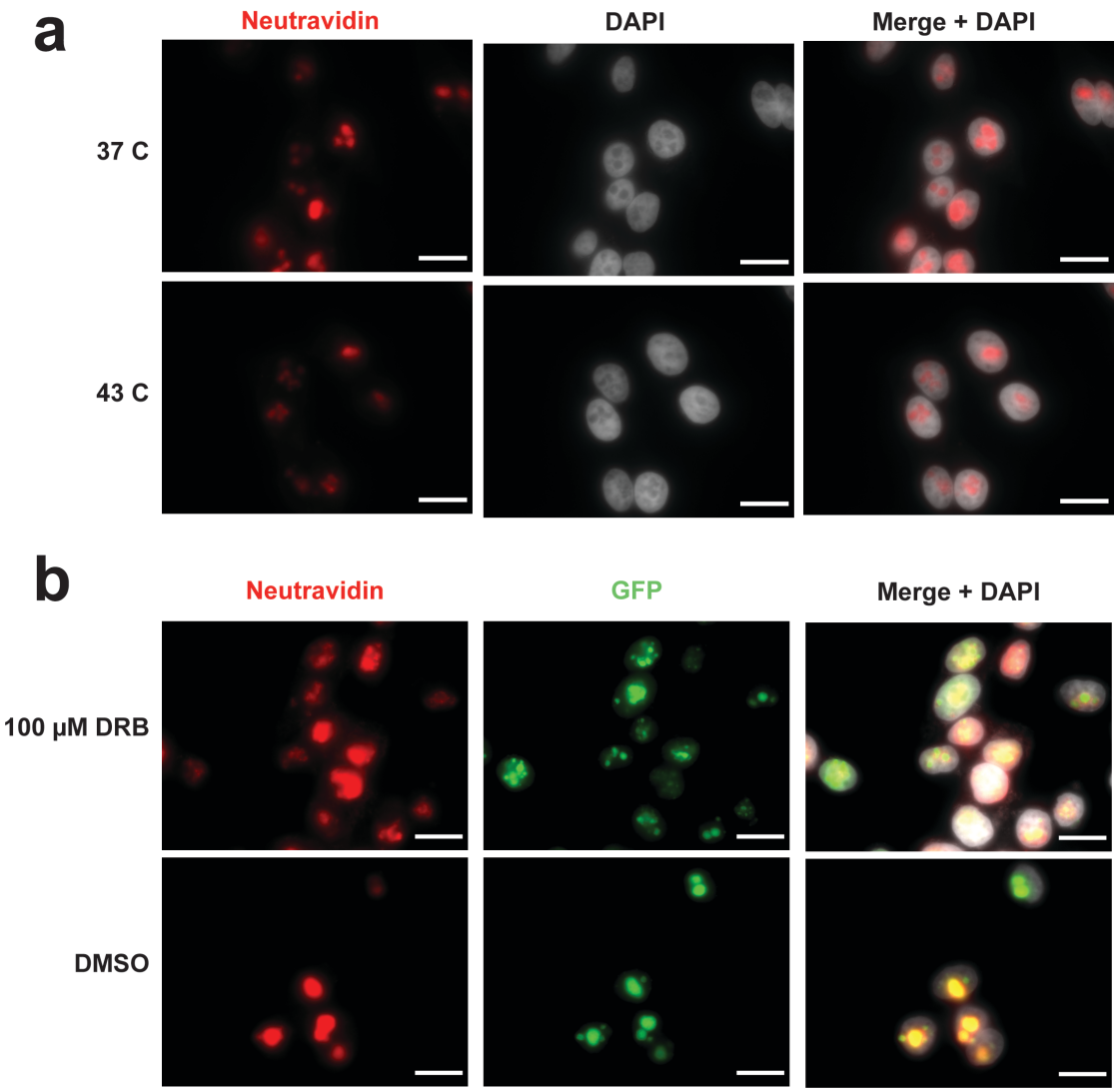


Figure 3.1. O-MAP Imaging Assay in nucleolar perturbed states. a, HeLa cells treated with 43C heat shock for 2 hours are targeted with nucleolar O-MAP and stained for biotinylated proteins. **b**, HeLa expressing nucleophosmin-GFP are treated with 100 μM DRB for 90 minutes. Nucleolar O-MAP is performed and biotinylated species are shown with neutravidin staining, while the nucleolar marker (NPM) is shown in green. Scale bars: 20 μm

O-MAP and Nucleoli in Cancer

The nucleolus also regulates tumor suppressors (e.g. p53 and c-myc) and is commonly implicated in cancer, historically serving as a biomarker for aggressive cancer increasing in both size and number[44]. This remains the case for pancreatic ductal adenocarcinoma (PDAC), the third leading cause of cancer death in the United States[143]. Based on RNA profiling, PDAC has been characterized into two main subtypes—classical and basal, of which basal is the more aggressive subtype that is recalcitrant to treatment[144]. Recent work by the Kugel lab has determined that THZ1—a small molecule inhibitor of RNA transcription and CDK7—has anti-proliferative activity specifically in the basal subtype (S. Kugel and J. Gianopulos, personal communication). Broad RNA-sequencing following THZ1 treatment in other cancers[145] supports THZ1 as influencing a DNA damage response that affects and disrupts the nucleolus[146,147]. Identifying the factors involved in the basal sensitivity to THZ1 as well as the classical resistance to treatment could lead to the identification of biomarkers and potential avenues for new combinatorial treatment with existing chemotherapeutics. However, it is difficult to biochemically fractionate the disrupted nucleolar structures after THZ1 treatment.

To address this, proximity labeling technologies can be used to deeply characterize the temporal dynamics of nucleolar remodeling in PDAC, response to THZ1 treatment and potential DNA damage nucleolar responses. This thesis displays the recently developed technology, Oligonucleotide directed proximity interactome MAPping (O-MAP), showing a deep characterization of nucleoli and provides a possible tool for studying the makeup of disrupted nucleoli in environmental stress and cancer[17]. In this way, disrupted nucleolar structure may be able to be characterized while maintaining a high spatial and temporal resolution upon perturbation. Upon treatment with THZ1, Panc 3.27 cells—a basal subtype specific cell line—showed a nucleolar disruption morphology, and a redistribution of pre-rRNA (**Fig 2a**). The pre-rRNA O-MAP imaging supports the formation of nucleolar necklaces around the nucleoli. Staining for different compartments of the nucleolus NPM (GC) and fibrillarin (DFC) show the formation of nucleolar caps and a disruption of the nucleolar tripartite structure (**Fig 2b,c**). Using OMAP-MS, the components of the nucleolar disrupted structure of THZ1 sensitive cells can be characterized, shedding light on potential proteins involved in THZ1 sensitivity for basal subtypes (as well as classical resistance).

Figure 3.2.

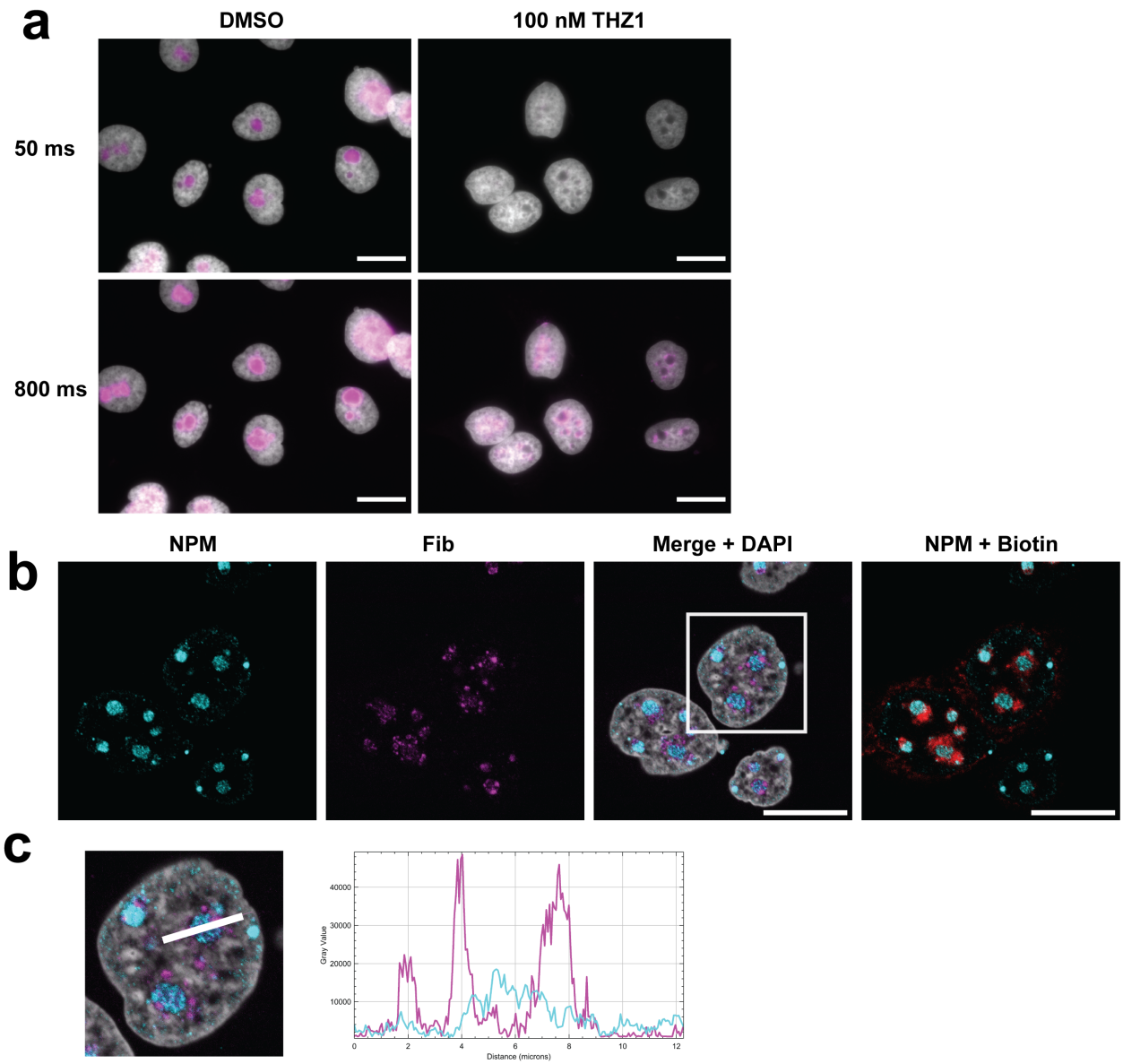


Figure 3.2. Disrupted nucleolar morphology and nucleolar cap formation after PDAC THZ1 treatment. a, Panc 3.27 cells incubated with DMSO or 100 nM THZ1 for 2 hours at 37C. Exposure imaging times from 50 ms and 800 ms (overexposed). **b,** Panc 3.27 cells incubated with 100 nM THZ1 for 2 hours were stained with IF for nucleophosmin (cyan), fibrillarin (magenta) and DAPI (grey). O-MAP was performed and biotin is shown in red with cyan NPM (left). White rectangle represents the close up image shown in **c,** with a line and corresponding line plot for both the NPM (cyan) and Fib (magenta). Scale bar 20 μm.

3.2: O-MAP AS A TOOL FOR STUDYING RNA INTERACTOMES

RNAs and their dynamic interactions with other biopolymers are increasingly important in understanding fundamental biological processes, the formation of biological condensates, and revealing the molecular bases of disease. The challenge of studying target RNAs and their interacting proteins on a subcellular level becomes even more apparent when the RNA and its interactome are organized into subnuclear bodies such as nuclear speckles, p-bodies, paraspeckles, and perinucleolar compartments[19]. In this thesis, I attempt to address the limitations of current RNA-centric technologies involving specificity, RNA abundance, and genetic manipulation by introducing O-MAP as a way to study a variety of RNA interactomes while maintaining subcellular context. I demonstrate the versatility of O-MAP to extend to RNAs of different expression levels. I also show the modularity of O-MAP in different cell systems and organisms. Using the 47S pre ribosomal RNA and 7sk, I characterize the proteomes of the nucleolus and parts of nuclear speckles where 7sk is involved in sequestering PTEFb before transcription elongation. Ongoing work on O-MAP includes the characterization of the XIST proteome and the comparison to existing literature[61,62]. Ultimately, the work presented here in this thesis introduces O-MAP as a new tool to discover RNA interactomes within the subcellular context and thus uncovering new opportunities for characterizing RNA biology, cancer and pharmacologically disrupted cells.

3.3 THE FUTURE OF *IN SITU* PROXIMITY LABELING AND RNA INTERACTOME CAPTURE

Towards the development of O-MAP sensitivity, lowly expressed RNAs, such as FIRRE, should be characterized using MS analysis. Additional O-MAP studies can test the boundaries of lowly expressed RNAs (e.g. <500 copies per cell) as well as mRNAs. Ongoing work includes lower expressed RNA OMAP-MS characterization on enhancer of zeste homolog 2 (EZH2)[148] mRNA. From a technical level, O-MAP is performed post crosslinking so additional biotinylation labeling time will not affect the state of the cell but can increase the signal strength of biotinylated material. However during the streptavidin enrichment, it becomes important to consider limiting the noise from endogenous biotinylated proteins and endogenous peroxidase activity. Future iterations of O-MAP can perhaps include a thiol-cleavable biotin[149] to increase signal while reducing these artifacts. Transcript length becomes a limiting factor in this initial design of O-MAP as no less than three oligonucleotide probes for abundant RNAs have been tested. Less abundant and short RNAs, such as some small nucleolar RNAs, can only be

targeted by a single probe and would require long labeling times as well as a considerable number of cells for streptavidin enrichment. Most RNA targets in O-MAP have shown strong biotinylation signal with 50-100 probes. To address the low number of probes and low signal, longer primer exchange reaction (PER) concatemers can be used by increasing the number of binding sites for the biotinylation enzyme conjugated oligonucleotide in SABER[16]. The limitations of O-MAP might then mirror the capacity to target certain RNAs within the RNA-FISH field. Additional steps to accommodate cytosolic localized RNAs could be a methanol or saponin permeabilization to reduce the total loss of cytoplasmic proteins from harsher triton permeabilization[150].

Within this thesis, I have shown that O-MAP can specifically label nucleoli in frozen tissues from mouse embryo. However, O-MAP-MS has yet to be applied to these systems, but other technologies, such as BAR[151], indicate that tissues can be analyzed by MS analysis after fixation and proximity labeling. O-MAP would potentially be disrupted in the case of formalin-fixed paraffin-embedded (FFPE) tissue due to the extreme fixation that often requires stringent techniques to recover material as well as the loss of nucleic acids from the preserved tissue. Additionally, O-MAP may be able to be used on other challenging models, such as organoids and stem cells to study disease relevant models.

Although this work provides new characterization of RNA and its local interacting proteomes, O-MAP can be extended to sequencing to study interacting DNA and RNA localized to the target RNA[17]. The combinational analysis of nearby chromosome regions, transcription factors, and target transcripts can lead develop a multidimensional dissection of gene regulation. The multiple component analysis could also be used to describe the components of subnuclear bodies and what drives phase separation. This analysis can be extended to disrupted structures, as long as the target RNA is present. For example, the nucleolus changes morphology upon environmental stress for heat shock and DNA damage, forming a dense nucleolar core or nucleolar caps, respectively. The components of these as well as the components required for nucleolar assembly after mitosis can be studied using the same 47S pre-rRNA targeted O-MAP in these systems. Ultimately, proximity labeling can be utilized to studying RNA driven biology and the formation of higher order structures.

REFERENCES

1. Osborne, R. J. & Thornton, C. A. RNA-dominant diseases. *Hum. Mol. Genet.* **15**, R162–R169 (2006).
2. Lukong, K. E., Chang, K., Khandjian, E. W. & Richard, S. RNA-binding proteins in human genetic disease. *Trends Genet.* **24**, 416–425 (2008).
3. Pereira, B., Billaud, M. & Almeida, R. RNA-Binding Proteins in Cancer: Old Players and New Actors. *Trends in Cancer* **3**, 506–528 (2017).
4. Hentze, M. W., Castello, A., Schwarzl, T. & Preiss, T. A brave new world of RNA-binding proteins. *Nat. Rev. Mol. Cell Biol.* **19**, 327–341 (2018).
5. Castello, A. *et al.* Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell* **149**, 1393–1406 (2012).
6. Delás, M. J. & Hannon, G. J. lncRNAs in development and disease: from functions to mechanisms. *Open Biol.* **7**, (2017).
7. Melé, M. & Rinn, J. L. 'Cat's Cradling' the 3D Genome by the Act of LncRNA Transcription. *Mol. Cell* **62**, 657–664 (2016).
8. Pederson, T. The Nucleolus. *Cold Spring Harb. Perspect. Biol.* **3**, 1–15 (2011).
9. Farley, K. I., Surovtseva, Y., Merkel, J. & Baserga, S. J. Determinants of mammalian nucleolar architecture. *Chromosoma* **124**, 323–31 (2015).
10. Quinodoz, S. A. *et al.* Higher-Order Inter-chromosomal Hubs Shape 3D Genome Organization in the Nucleus. *Cell* **174**, 744-757.e24 (2018).
11. Hernandez-Verdun, D. Assembly and disassembly of the nucleolus during the cell cycle. *Nucleus* **2**, 189–94 (2011).
12. Feric, M. *et al.* Coexisting Liquid Phases Underlie Nucleolar Subcompartments. *Cell* **165**, 1686–1697 (2016).
13. Rhee, H.-W. *et al.* Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. *Science* **339**, 1328–1331 (2013).
14. McHugh, C. A. & Guttman, M. RAP-MS: A method to identify proteins that interact directly with a specific RNA molecule in cells. in *Methods in Molecular Biology* **1649**, 473–488 (Humana Press Inc., 2018).
15. Mili, S. & Steitz, J. A. Evidence for reassociation of RNA-binding proteins after cell lysis: implications for the interpretation of immunoprecipitation analyses. *RNA* **10**, 1692–4 (2004).
16. Kishi, J. Y. *et al.* SABER amplifies FISH: enhanced multiplexed imaging of RNA and DNA in cells and tissues. *Nat. Methods* **16**, 533–544 (2019).
17. Tsue, A. F. *et al.* Oligonucleotide-directed proximity-interactome mapping (O-MAP): A unified method for discovering RNA-interacting proteins, transcripts and genomic loci in situ. *bioRxiv* 2023.01.19.524825 (2023). doi:10.1101/2023.01.19.524825
18. Statello, L., Guo, C.-J., Chen, L.-L. & Huarte, M. Gene regulation by long non-coding RNAs and its biological functions. *Nat. Rev. Mol. Cell Biol. 2020 222* **22**, 96–118 (2020).
19. Banani, S. F., Lee, H. O., Hyman, A. A. & Rosen, M. K. Biomolecular condensates: Organizers of cellular biochemistry. *Nature Reviews Molecular Cell Biology* **18**, 285–298 (2017).
20. Poudyal, R. R., Sieg, J. P., Portz, B., Keating, C. D. & Bevilacqua, P. C. RNA sequence and structure control assembly and function of RNA condensates. *RNA* **27**, 1589–1601 (2021).
21. Jain, A. & Vale, R. D. RNA phase transitions in repeat expansion disorders. *Nature* **546**, 243–247 (2017).
22. Latonen, L. Phase-to-Phase With Nucleoli – Stress Responses, Protein Aggregation and Novel Roles of RNA. *Front. Cell. Neurosci.* **0**, 151 (2019).
23. Yang, K., Yang, J. & Yi, J. Nucleolar Stress: hallmarks, sensing mechanism and diseases. *Cell Stress* **2**, 125 (2018).
24. Gai, X. *et al.* Pre-ribosomal RNA reorganizes DNA damage repair factors in nucleus during meiotic prophase and DNA damage response. *Cell Res.* **32**, 254–268 (2022).
25. Frottin, F. *et al.* The nucleolus functions as a phase-separated protein quality control compartment. *Science (80-.).* **365**, 342–347 (2019).
26. Jiang, S., Fagman, J. B., Chen, C., Alberti, S. & Liu, B. Protein phase separation and its role in tumorigenesis. doi:10.7554/eLife.60264
27. Rinn, J. & Guttman, M. RNA and dynamic nuclear organization: Long noncoding RNAs may

function as organizing factors that shape the cell nucleus. *Science (80-.).* **345**, 1240–1241 (2014).
28. Wang, B. *et al.* Liquid–liquid phase separation in human health and diseases. *Signal Transduct. Target. Ther. 2021 61* **6**, 1–16 (2021).
29. Dundr, M. *et al.* In vivo kinetics of Cajal body components. *J. Cell Biol.* **164**, 831–842 (2004).
30. Tong, X. *et al.* Liquid–liquid phase separation in tumor biology. *Signal Transduct. Target. Ther. 2022 71* **7**, 1–22 (2022).
31. Taniue, K. & Akimitsu, N. Aberrant phase separation and cancer. *FEBS J.* **289**, 17–39 (2022).
32. Yamazaki, T. *et al.* Functional Domains of NEAT1 Architectural lncRNA Induce Paraspeckle Assembly through Phase Separation. *Mol. Cell* **70**, 1038-1053.e7 (2018).
33. Wang, C. *et al.* Stress Induces Dynamic, Cytotoxicity-Antagonizing TDP-43 Nuclear Bodies via Paraspeckle LncRNA NEAT1-Mediated Liquid-Liquid Phase Separation. *Mol. Cell* **79**, 443-458.e7 (2020).
34. Luo, Y., Na, Z. & Slavoff, S. A. P-Bodies: Composition, Properties, and Functions. *Biochemistry* **57**, 2424 (2018).
35. Nott, T. J. *et al.* Phase Transition of a Disordered Nuage Protein Generates Environmentally Responsive Membraneless Organelles. *Mol. Cell* **57**, 936–947 (2015).
36. Li, P. *et al.* Phase transitions in the assembly of multivalent signalling proteins. *Nat. 2012 4837389* **483**, 336–340 (2012).
37. Penny, G. D., Kay, G. F., Sheardown, S. A., Rastan, S. & Brockdorff, N. Requirement for Xist in X chromosome inactivation. *Nature* **379**, 131–137 (1996).
38. Markaki, Y. *et al.* Xist nucleates local protein gradients to propagate silencing across the X chromosome. *Cell* **184**, 6174-6192.e32 (2021).
39. Giorgetti, L. *et al.* Structural organization of the inactive X chromosome in the mouse. *Nature* **535**, 575–579 (2016).
40. Galganski, L., Urbanek, M. O. & Krzyzosiak, W. J. Nuclear speckles: molecular organization, biological function and role in disease. *Nucleic Acids Res.* **45**, 10350–10368 (2017).
41. Barrandon, C., Bonnet, F., Nguyen, V. T., Labas, V. & Bensaude, O. The transcription-dependent dissociation of P-TEFb-HEXIM1-7SK RNA relies upon formation of hnRNP-7SK RNA complexes. *Mol. Cell. Biol.* **27**, 6996–7006 (2007).
42. Peterlin, B. M., Brogie, J. E. & Price, D. H. 7SK snRNA: A noncoding RNA that plays a major role in regulating eukaryotic transcription. *Wiley Interdisciplinary Reviews: RNA* **3**, 92–103 (2012).
43. Prasanth, K. V. *et al.* Nuclear organization and dynamics of 7SK RNA in regulating gene expression. *Mol. Biol. Cell* **21**, 4184–4196 (2010).
44. Stępiński, D. The nucleolus, an ally, and an enemy of cancer cells. *Histochem. Cell Biol.* **150**, 607 (2018).
45. Schmidt, E. V. The role of c-myc in cellular growth control. *Oncogene 1999 1819* **18**, 2988–2996 (1999).
46. Dousset, T. *et al.* Initiation of Nucleolar Assembly Is Independent of RNA Polymerase I Transcription. *Mol. Biol. Cell* **11**, 2705 (2000).
47. Mitrea, D. M. *et al.* Nucleophosmin integrates within the nucleolus via multi-modal interactions with proteins displaying R-rich linear motifs and rRNA. *Elife* **5**, (2016).
48. Latonen, L. Phase-to-phase with nucleoli - Stress responses, protein aggregation and novel roles of RNA. *Front. Cell. Neurosci.* **13**, 151 (2019).
49. Rubbi, C. P. & Milner, J. Disruption of the nucleolus mediates stabilization of p53 in response to DNA damage and other stresses. *EMBO J.* **22**, 6068 (2003).
50. Audas, T. E. *et al.* Adaptation to Stressors by Systemic Protein Amyloidogenesis. *Dev. Cell* **39**, 155–168 (2016).
51. Baltz, A. G. *et al.* The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. *Mol. Cell* **46**, 674–690 (2012).
52. Castello, A. *et al.* System-wide identification of RNA-binding proteins by interactome capture. *Nat. Protoc. 2013 83* **8**, 491–500 (2013).
53. Baltz, A. G. *et al.* The mRNA-Bound Proteome and Its Global Occupancy Profile on Protein-Coding Transcripts. *Mol. Cell* **46**, 674–690 (2012).
54. Queiroz, R. M. L. *et al.* Comprehensive quantitation of RNA-protein interaction dynamics by orthogonal organic phase separation (OOPS). *Nat. Biotechnol.* **37**, 169 (2019).
55. Zheng, X. *et al.* Detecting RNA-Protein Interaction Using End-Labeled Biotinylated RNA

Oligonucleotides and Immunoblotting. *Methods Mol. Biol.* **1421**, 35–44 (2016).
56. Lee, H. Y. *et al.* RNA-protein analysis using a conditional CRISPR nuclease. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 5416–5421 (2013).
57. Leppek, K. & Stoecklin, G. An optimized streptavidin-binding RNA aptamer for purification of ribonucleoprotein complexes identifies novel ARE-binding proteins. *Nucleic Acids Res.* **42**, (2014).
58. Ramanathan, M., Porter, D. F. & Khavari, P. A. Methods to study RNA-protein interactions. *Nat. Methods* **16**, 225–234 (2019).
59. Chu, C. & Chang, H. Y. Understanding RNA-chromatin interactions using chromatin isolation by RNA purification (ChIRP). *Methods Mol. Biol.* **1480**, 115–123 (2016).
60. Simon, M. D. Capture hybridization analysis of RNA targets (CHART). *Curr. Protoc. Mol. Biol.* (2013). doi:10.1002/0471142727.mb2125s101
61. Minajigi, A. *et al.* A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. *Science* **349**, 1DUIMMY (2015).
62. Chu, C. *et al.* Systematic discovery of Xist RNA binding proteins. *Cell* **161**, 404–416 (2015).
63. McHugh, C. A. *et al.* The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nat. 2015 5217551* **521**, 232–236 (2015).
64. Roux, K. J., Kim, D. I., Raida, M. & Burke, B. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J. Cell Biol.* **196**, 801 (2012).
65. Fazal, F. M. *et al.* Atlas of Subcellular RNA Localization Revealed by APEX-Seq. *Cell* **178**, 473-490.e26 (2019).
66. Kim, D. I. *et al.* Probing nuclear pore complex architecture with proximity-dependent biotinylation. *Proc. Natl. Acad. Sci. U. S. A.* **111**, (2014).
67. Martell, J. D. *et al.* Engineered ascorbate peroxidase as a genetically encoded reporter for electron microscopy. *Nat. Biotechnol.* **30**, 1143–1148 (2012).
68. Kaewsapsak, P., Shechner, D. M., Mallard, W., Rinn, J. L. & Ting, A. Y. Live-cell mapping of organelle-associated RNAs via proximity biotinylation combined with protein-RNA crosslinking. *Elife* **6**, (2017).
69. Ramanathan, M. *et al.* RNA–protein interaction detection in living cells. *Nat. Methods 2018 153* **15**, 207–212 (2018).
70. Mukherjee, J., Franz-Wachtel, M., Maček, B. & Jansen, R.-P. RNA Interactome Identification via RNA-BioID in Mouse Embryonic Fibroblasts. *Bio-protocol* **10**, (2020).
71. Zhang, Z. *et al.* Capturing RNA-protein interaction via CRUIS. *Nucleic Acids Res.* **48**, (2020).
72. Yi, W. *et al.* CRISPR-assisted detection of RNA-protein interactions in living cells. *Nat. Methods* **17**, 685–688 (2020).
73. Weissinger, R., Heinold, L., Akram, S., Jansen, R. P. & Hermesh, O. RNA Proximity Labeling: A New Detection Tool for RNA–Protein Interactions. *Molecules* **26**, (2021).
74. Bar, D. Z. *et al.* Biotinylation by antibody recognition—a method for proximity labeling. *Nat. Methods 2017 152* **15**, 127–133 (2017).
75. Dopie, J., Sweredoski, M. J., Moradian, A. & Belmont, A. S. Tyramide signal amplification mass spectrometry (TSA-MS) ratio identifies nuclear speckle proteins. *J. Cell Biol.* **219**, (2020).
76. Rees, J. S., Li, X. W., Perrett, S., Lilley, K. S. & Jackson, A. P. Selective Proteomic Proximity Labeling Assay Using Tyramide (SPPLAT): A Quantitative Method for the Proteomic Analysis of Localized Membrane-Bound Protein Clusters. *Curr. Protoc. Protein Sci.* **80**, 19.27.1-19.27.18 (2015).
77. Yap, K., Chung, T. H. & Makeyev, E. V. Hybridization-proximity labeling reveals spatially ordered interactions of nuclear RNA compartments. *Mol. Cell* **82**, 463-478.e11 (2022).
78. Schueler, M. *et al.* Differential protein occupancy profiling of the mRNA transcriptome. *Genome Biol.* **15**, (2014).
79. Mitchell, S. F. & Parker, R. Principles and properties of eukaryotic mRNPs. *Mol. Cell* **54**, 547–558 (2014).
80. Müller-Mcnicoll, M. & Neugebauer, K. M. How cells get the message: dynamic assembly and function of mRNA-protein complexes. *Nat. Rev. Genet.* **14**, 275–287 (2013).
81. Bhat, P., Honson, D. & Guttman, M. Nuclear compartmentalization as a mechanism of quantitative control of gene expression. *Nat. Rev. Mol. Cell Biol.* **22**, 653–670 (2021).
82. Roden, C. & Gladfelter, A. S. RNA contributions to the form and function of biomolecular condensates. *Nat. Rev. Mol. Cell Biol.* **22**, 183–195 (2021).

83. Thakur, J. & Henikoff, S. Architectural RNA in chromatin organization. *Biochem. Soc. Trans.* **48**, 1967–1978 (2020).
84. Nickerson, J. A., Krochmalnic, G., Wan, K. M. & Penman, S. Chromatin architecture and nuclear RNA. *Proc. Natl. Acad. Sci. U. S. A.* **86**, 177 (1989).
85. Nozawa, R. S. & Gilbert, N. RNA: Nuclear Glue for Folding the Genome. *Trends Cell Biol.* **29**, 201–211 (2019).
86. Simon, M. D. & Machyna, M. Principles and Practices of Hybridization Capture Experiments to Study Long Noncoding RNAs That Act on Chromatin. *Cold Spring Harb. Perspect. Biol.* **11**, (2019).
87. McDonel, P. & Guttman, M. Approaches for Understanding the Mechanisms of Long Noncoding RNA Regulation of Gene Expression. *Cold Spring Harb. Perspect. Biol.* **11**, (2019).
88. Chu, C. & Chang, H. Y. ChIRP-MS: RNA-Directed Proteomic Discovery. *Methods Mol. Biol.* **1861**, 37–45 (2018).
89. Simon, M. D. *et al.* The genomic binding sites of a noncoding RNA. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 20497–20502 (2011).
90. Qin, W., Cho, K. F., Cavanagh, P. E. & Ting, A. Y. Deciphering molecular interactions by proximity labeling. *Nat. Methods 2021 182* **18**, 133–143 (2021).
91. Kaewsapsak, P., Shechner, D. M., Mallard, W., Rinn, J. L. & Ting, A. Y. Live-cell mapping of organelle-associated RNAs via proximity biotinylation combined with protein-RNA crosslinking. *Elife* **6**, e29224 (2017).
92. Engel, K. L. *et al.* Analysis of subcellular transcriptomes by RNA proximity labeling with Halo-seq. *Nucleic Acids Res.* **50**, E24 (2022).
93. Chen, Y. *et al.* Mapping 3D genome organization relative to nuclear compartments using TSA-Seq as a cytological ruler. *J. Cell Biol.* **217**, 4025–4048 (2018).
94. Lam, S. S. *et al.* Directed evolution of APEX2 for electron microscopy and proximity labeling. *Nat. Methods* **12**, 51–54 (2015).
95. Han, S. *et al.* RNA-protein interaction mapping via MS2- or Cas13-based APEX targeting. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 22068–22079 (2020).
96. Lobingier, B. T. *et al.* An Approach to Spatiotemporally Resolve Protein Interaction Networks in Living Cells. *Cell* **169**, 350-360.e12 (2017).
97. Carter, D., Chakalova, L., Osborne, C. S., Dai, Y. feng & Fraser, P. Long-range chromatin regulatory interactions in vivo. *Nat. Genet.* **32**, 623–626 (2002).
98. Hung, V. *et al.* Proteomic Mapping of the Human Mitochondrial Intermembrane Space in Live Cells via Ratiometric APEX Tagging. *Mol. Cell* **55**, 332–341 (2014).
99. Raj, A., van den Bogaard, P., Rifkin, S. A., van Oudenaarden, A. & Tyagi, S. Imaging individual mRNA molecules using multiple singly labeled probes. *Nat. Methods* **5**, 877–879 (2008).
100. Raj, A. & Rinn, J. L. Illuminating Genomic Dark Matter with RNA Imaging. *Cold Spring Harb. Perspect. Biol.* **11**, (2019).
101. Beliveau, B. J. *et al.* OligoMiner provides a rapid, flexible environment for the design of genome-scale oligonucleotide in situ hybridization probes. doi:10.1073/pnas.1714530115
102. Padovan-Merhar, O. *et al.* Single Mammalian Cells Compensate for Differences in Cellular Volume and DNA Copy Number through Independent Global Transcriptional Mechanisms. *Mol. Cell* **58**, 339–352 (2015).
103. Henras, A. K., Plisson-Chastang, C., O'Donohue, M. F., Chakraborty, A. & Gleizes, P. E. An overview of pre-ribosomal RNA processing in eukaryotes. *Wiley Interdiscip. Rev. RNA* **6**, 225–242 (2015).
104. Raap, A. K. *et al.* Ultra-sensitive FISH using peroxidase-mediated deposition of biotin- or fluorochrome tyramides. *Hum. Mol. Genet.* **4**, 529–534 (1995).
105. Beliveau, B. J. *et al.* Versatile design and synthesis platform for visualizing genomes with Oligopaint FISH probes. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 21301–21306 (2012).
106. Wendeberg, A. Fluorescence In Situ Hybridization for the Identification of Environmental Microbes Fluorescence In Situ Hybridization for the Identification of Environmental Microbes. **2010**, 1–9 (2014).
107. Qin, W., Myers, S. A., Carey, D. K., Carr, S. A. & Ting, A. Y. Spatiotemporally-resolved mapping of RNA binding proteins via functional proximity labeling reveals a mitochondrial mRNA anchor promoting stress recovery. *Nat. Commun. 2021 121* **12**, 1–19 (2021).

108. Plath, K., Mlynarczyk-Evans, S., Nusinow, D. A. & Panning, B. Xist RNA and the mechanism of X chromosome inactivation. *Annu. Rev. Genet.* **36**, 233–278 (2002).
109. Boj, S. F. *et al.* Organoid models of human and mouse ductal pancreatic cancer. *Cell* **160**, 324–338 (2015).
110. Li, J. *et al.* TMTpro reagents: a set of isobaric labeling mass tags enables simultaneous proteome-wide measurements across 16 samples. *Nat. Methods 2020 174* **17**, 399–404 (2020).
111. Andersen, J. S. *et al.* Directed proteomic analysis of the human nucleolus. *Curr. Biol.* **12**, 1–11 (2002).
112. Diribarne, G. & Bensaude, O. 7SK RNA, a non-coding RNA regulating P-TEFb, a general transcription factor. *RNA Biol.* **6**, 122–128 (2009).
113. Flynn, R. A. *et al.* 7SK-BAF axis controls pervasive transcription at enhancers. *Nat. Struct. Mol. Biol.* **23**, 231–8 (2016).
114. Uhlen, M. *et al.* Towards a knowledge-based Human Protein Atlas. *Nat. Biotechnol.* **28**, 1248–1250 (2010).
115. Zhou, Y. *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **10**, (2019).
116. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15545–15550 (2005).
117. Spector, D. L. & Lamond, A. I. Nuclear Speckles. *Cold Spring Harb. Perspect. Biol.* **3**, 1–12 (2011).
118. Bateman, A. *et al.* UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res.* **49**, D480–D489 (2021).
119. Jan, C. H., Williams, C. C. & Weissman, J. S. Principles of ER cotranslational translocation revealed by proximity-specific ribosome profiling. *Science (80-.).* **346**, 1257521–1257521 (2014).
120. Han, Y. *et al.* Directed Evolution of Split APEX2 Peroxidase. *ACS Chem. Biol.* **14**, 619–635 (2019).
121. Bassett, A. R. *et al.* Considerations when investigating lncRNA function in vivo. *Elife* **3**, 1–14 (2014).
122. Jiang, S. *et al.* A proteomics approach to the cell-surface interactome using the enzyme-mediated activation of radical sources reaction. *Proteomics* **12**, 54–62 (2012).
123. Bar, D. Z. *et al.* Biotinylation by antibody recognition—a method for proximity labeling. *Nat. Methods* **15**, 127–133 (2017).
124. Derenzini, M. *et al.* Nucleolar function and size in cancer cells. *Am. J. Pathol.* **152**, 1291 (1998).
125. Schnell, U., Dijk, F., Sjollema, K. A. & Giepmans, B. N. G. Immunolabeling artifacts and the need for live-cell imaging. *Nat. Methods* **9**, 152–158 (2012).
126. Cech, T. R. & Steitz, J. A. The noncoding RNA revolution - Trashing old rules to forge new ones. *Cell* **157**, 77–94 (2014).
127. Padovan-Merhar, O. *et al.* Single Mammalian Cells Compensate for Differences in Cellular Volume and DNA Copy Number through Independent Global Transcriptional Mechanisms. *Mol. Cell* **58**, 339–352 (2015).
128. Beliveau, B. J. *et al.* OligoMiner provides a rapid, flexible environment for the design of genome-scale oligonucleotide in situ hybridization probes. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E2183–E2192 (2018).
129. Shechner, D. M. & Bartel, D. P. The structural basis of RNA-catalyzed RNA polymerization. *Nat. Struct. Mol. Biol.* **18**, 1036 (2011).
130. Rappsilber, J., Ishihama, Y. & Mann, M. Stop and Go Extraction Tips for Matrix-Assisted Laser Desorption/Ionization, Nanoelectrospray, and LC/MS Sample Pretreatment in Proteomics. *Anal. Chem.* **75**, 663–670 (2002).
131. Eng, J. K., Jahan, T. A. & Hoopmann, M. R. Comet: an open-source MS/MS sequence database search tool. *Proteomics* **13**, 22–24 (2013).
132. Schweppe, D. K. *et al.* Full-Featured, Real-Time Database Searching Platform Enables Fast and Accurate Multiplexed Quantitative Proteomics. *J. Proteome Res.* **19**, 2026–2034 (2020).
133. Schindelin, J. *et al.* Fiji: An open-source platform for biological-image analysis. *Nature Methods* **9**, 676–682 (2012).
134. Rueden, C. T. *et al.* ImageJ2: ImageJ for the next generation of scientific image data. *BMC*

Bioinformatics **18**, (2017).
135. Boisvert, F. M., Van Koningsbruggen, S., Navascués, J. & Lamond, A. I. The multifunctional nucleolus. *Nature Reviews Molecular Cell Biology* **8**, 574–585 (2007).
136. Boulon, S., Westman, B. J., Hutten, S., Boisvert, F.-M. & Lamond, A. I. The Nucleolus under Stress. *Mol. Cell* **40**, 216–227 (2010).
137. Orsolic, I. *et al.* The relationship between the nucleolus and cancer: Current evidence and emerging paradigms. *Seminars in Cancer Biology* **37**–**38**, 36–50 (2016).
138. Caudron-Herger, M., Pankert, T. & Rippe, K. Regulation of nucleolus assembly by non-coding RNA polymerase II transcripts. *Nucleus* **7**, 308–318 (2016).
139. Moore, H. M. *et al.* Quantitative proteomics and dynamic imaging of the nucleolus reveal distinct responses to UV and ionizing radiation. *Mol. Cell. Proteomics* **10**, (2011).
140. Latonen, L. Phase-to-phase with nucleoli - Stress responses, protein aggregation and novel roles of RNA. *Frontiers in Cellular Neuroscience* **13**, (2019).
141. Lindström, M. S. *et al.* Nucleolus as an emerging hub in maintenance of genome stability and cancer pathogenesis. *Oncogene* **37**, 2351–2366 (2018).
142. Turinetto, V. *et al.* The cyclin-dependent kinase inhibitor 5, 6-dichloro-1-beta-D-ribofuranosylbenzimidazole induces nongenotoxic, DNA replication-independent apoptosis of normal and leukemic cells, regardless of their p53 status. *BMC Cancer* **9**, 281 (2009).
143. Rahib, L. *et al.* Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States. *Cancer Res.* **74**, 2913–2921 (2014).
144. Raphael, B. J. *et al.* Integrated Genomic Characterization of Pancreatic Ductal Adenocarcinoma. *Cancer Cell* **32**, 185-203.e13 (2017).
145. Greenall, S. A. *et al.* Cyclin-dependent kinase 7 is a therapeutic target in high-grade glioma. *Oncog. 2017 65* **6**, e336–e336 (2017).
146. Zeng, S. *et al.* CDK7 inhibition augments response to multidrug chemotherapy in pancreatic cancer. *J. Exp. Clin. Cancer Res.* **41**, 1–18 (2022).
147. Lu, P. *et al.* THZ1 reveals CDK7-dependent transcriptional addictions in pancreatic cancer. *Oncogene 2019 3820* **38**, 3932–3945 (2019).
148. Duan, R., Du, W. & Guo, W. EZH2: A novel target for cancer treatment. *J. Hematol. Oncol.* **13**, 1–12 (2020).
149. Li, H., Frankenfield, A. M., Houston, R., Sekine, S. & Hao, L. Thiol-Cleavable Biotin for Chemical and Enzymatic Biotinylation and Its Application to Mitochondrial TurboID Proteomics. *J. Am. Soc. Mass Spectrom.* **32**, 2358–2365 (2021).
150. Ly, T. *et al.* Proteomic analysis of cell cycle progression in asynchronous cultures, including mitotic subphases, using PRIMMUS. *Elife* **6**, (2017).
151. Bar, D. Z. & Collins, F. S. Antibody-Driven Proximity Labeling in Fixed Tissues. *Methods Mol. Biol.* **2008**, 73–81 (2019).