

ANNALS OF THE NEW YORK ACADEMY OF SCIENCES

Special Issue: *Keystone Symposia Reports*

Concise Original Report

Noncoding RNAs: biology and applications—a Keystone Symposia report

Jennifer Cable,¹ Edith Heard,^{2,3} Tetsuro Hirose,^{4,5} Kannanganattu V. Prasanth,⁶ Ling-Ling Chen,^{7,8,9} Jonathan E. Henninger,¹⁰ Sofia A. Quinodoz,¹¹ David L. Spector,¹² Sarah D. Diermeier,¹³ Allison M. Porman,¹⁴ Dhiraj Kumar,¹⁵ Mark W. Feinberg,¹⁶ Xiaohua Shen,¹⁷ Juan Pablo Unfried,¹⁸ Rory Johnson,^{19,20} Chun-Kan Chen,^{21,22} Jeremy E. Wilusz,²³ Adelheid Lempradl,²⁴ Sean E. McGeary,^{10,25} Lamia Wahba,^{22,26} Anna Marie Pyle,^{27,28} Amanda E. Hargrove,²⁹ Matthew D. Simon,³⁰ Marco Marcia,³¹ Róża K. Przanowska,³² Howard Y. Chang,^{21,33} Samie R. Jaffrey,³⁴ Lydia M. Contreras,³⁵ Qi Chen,³⁶ Junchao Shi,³⁶ Joshua T. Mendell,³⁷ Lin He,³⁸ Erwei Song,³⁹ John L. Rinn,⁴⁰ Mukesh Kumar Lalwani,⁴¹ Murat Can Kalem,⁴² Edward B. Chuong,⁴³ Lynne E. Maquat,⁴⁴ and Xuhang Liu⁴⁵

¹PhD Science Writer, New York, New York. ²European Molecular Biology Laboratory (EMBL), Heidelberg, Heidelberg, Germany. ³Collège de France, Paris, France. ⁴Graduate School of Frontier Biosciences, Osaka University, Suita, Japan. ⁵Institute for Genetic Medicine, Hokkaido University, Sapporo, Japan. ⁶Department of Cell and Developmental Biology, Cancer Center at Illinois, University of Illinois at Urbana-Champaign, Urbana, Illinois. ⁷State Key Laboratory of Molecular Biology, Shanghai Key Laboratory of Molecular Andrology, CAS Center for Excellence in Molecular Cell Science, Shanghai Institute of Biochemistry and Cell Biology, University of the Chinese Academy of Sciences, Shanghai, China. ⁸School of Life Science and Technology, ShanghaiTech University, Shanghai, China. ⁹School of Life Sciences, Hangzhou Institute for Advanced Study, University of the Chinese Academy of Sciences, Hangzhou, China. ¹⁰Whitehead Institute for Biomedical Research, Cambridge, Massachusetts. ¹¹Department of Chemical and Biological Engineering, Princeton University, Princeton, New Jersey. ¹²Cold Spring Harbor Laboratory, Cold Spring Harbor and Genetics Program, Stony Brook University, Stony Brook, New York. ¹³Department of Biochemistry, University of Otago, Dunedin, New Zealand. ¹⁴Biochemistry and Molecular Genetics Department, University of Colorado, Anschutz Medical Campus, Aurora, Colorado. ¹⁵Department of Cancer Biology, The University of Texas MD Anderson Cancer Center, Houston, Texas. ¹⁶Cardiovascular Division, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts. ¹⁷Tsinghua-Peking Joint Center for Life Sciences, School of Medicine and School of Life Sciences, Tsinghua University, Beijing, China. ¹⁸Center for Applied Medical Research (CIMA), Department of Gene Therapy and Regulation of Gene Expression, Universidad de Navarra (UNAV), Pamplona, Spain. ¹⁹Department of Medical Oncology, Inselspital, Bern University Hospital; and Department for BioMedical Research University of Bern, Bern, Switzerland. ²⁰School of Biology and Environmental Science and Conway Institute for Biomolecular and Biomedical Research, University College Dublin, Dublin, Ireland. ²¹Center for Personal Dynamic Regulomes, Stanford University, Stanford, California. ²²Department of Genetics, Stanford University School of Medicine, Stanford, California. ²³Department of Biochemistry and Biophysics, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania. ²⁴Department of Metabolism and Nutritional Programming, Van Andel Research Institute, Grand Rapids, Michigan. ²⁵Howard Hughes Medical Institute and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts. ²⁶Department of Pathology, Stanford University School of Medicine, Stanford, California. ²⁷Department of Genetics, Yale School of Medicine, New Haven, Connecticut. ²⁸Connecticut and Howard Hughes Medical Institute, Chevy Chase, Maryland. ²⁹Department of Chemistry, Duke University, Durham, North Carolina. ³⁰Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut. ³¹European Molecular Biology Laboratory (EMBL) Grenoble, Grenoble, France. ³²Department of Biochemistry and Molecular Genetics, University of Virginia School of Medicine, Charlottesville, Virginia. ³³Howard Hughes Medical Institute, Stanford University, Stanford, California. ³⁴Department of Pharmacology, Weill Medical College of Cornell University, New York, New York. ³⁵McKetta Department of Chemical Engineering, University of Texas at Austin, Austin, Texas. ³⁶Division of Biomedical Sciences, School of Medicine, University of California, Riverside, Riverside, California. ³⁷Department of Molecular Biology, Harold C. Simmons Comprehensive Cancer Center, Hamon Center for Regenerative Science and Medicine; and Howard Hughes Medical Institute, University of Texas Southwestern Medical Center, Dallas, Texas. ³⁸Division of Cellular and Developmental Biology, Molecular and Cell Biology Department, University of California at Berkeley, Berkeley, California. ³⁹Guangdong Provincial Key Laboratory of Malignant

doi: 10.1111/nyas.14713

Tumor Epigenetics and Gene Regulation, Medical Research Center and Breast Tumor Center, Sun Yat-sen Memorial Hospital, Sun Yat-sen University; Bioland Laboratory; Program of Molecular Medicine, Zhongshan School of Medicine, Sun Yat-sen University; and Fountain-Valley Institute for Life Sciences, Guangzhou Institute of Biomedicine and Health, Chinese Academy of Sciences Guangzhou, Guangzhou, China. ⁴⁰Department of Biochemistry, BioFrontiers Institute, and Howard Hughes Medical Institute, University of Colorado Boulder, Boulder, Colorado. ⁴¹Queens Medical Research Institute, BHF Centre for Cardiovascular Sciences, University of Edinburgh, Scotland, United Kingdom. ⁴²Department of Microbiology and Immunology, Witebsky Center for Microbial Pathogenesis and Immunology, Jacobs School of Medicine and Biomedical Sciences, University at Buffalo, SUNY, Buffalo, New York. ⁴³Department of Molecular, Cellular, and Developmental Biology and BioFrontiers Institute, University of Colorado Boulder, Boulder, Colorado. ⁴⁴Department of Biochemistry and Biophysics, School of Medicine and Dentistry and Center for RNA Biology, University of Rochester, Rochester, New York. ⁴⁵Laboratory of Systems Cancer Biology, The Rockefeller University, New York, New York

Address for correspondence: annals@nyas.org

The human transcriptome contains many types of noncoding RNAs, which rival the number of protein-coding species. From long noncoding RNAs (lncRNAs) that are over 200 nucleotides long to piwi-interacting RNAs (piRNAs) of only 20 nucleotides, noncoding RNAs play important roles in regulating transcription, epigenetic modifications, translation, and cell signaling. Roles for noncoding RNAs in disease mechanisms are also being uncovered, and several species have been identified as potential drug targets. On May 11–14, 2021, the Keystone eSymposium “Noncoding RNAs: Biology and Applications” brought together researchers working in RNA biology, structure, and technologies to accelerate both the understanding of RNA basic biology and the translation of those findings into clinical applications.

Keywords: circRNA; noncoding RNA; lncRNA; nuclear condensate; phase transition; piRNA; transcription; Xist

KEYSTONE SYMPOSIA
Accelerating Life Science Discovery

SCAN FOR MORE DETAILS



Text
NonCodingRNA21
to
+1 970.236.4705
to receive additional details

Introduction

Noncoding RNAs are a diverse range of RNA species whose physiological and pathological functions are just beginning to be understood. Long noncoding RNAs (lncRNAs) are RNA transcripts of greater than 200 nucleotides that often contain many features of protein-coding RNAs. Approximately 30,000 lncRNAs have been identified in the human genome, putting them on par with the number of protein-coding RNAs. The most well-studied lncRNAs localize to the nucleus, where they can play key roles in the formation of nuclear condensates like the nucleolus and affect the transcription machinery to regulate gene expression. Dysregulation of several lncRNAs has been implicated in dis-

eases, including cancer and cardiovascular disease (CVD), thus opening the potential for new therapeutic targets.¹

Short noncoding RNA species include microRNA (miRNA), small RNA, piwi-interacting RNA (piRNA), and circular RNA (circRNA), as well as more obscure RNAs like tRNA-derived small RNA (tsRNA), rRNA-derived small RNA (rsRNA), and Y RNA-derived small RNA (ysRNA). These RNAs play roles in mRNA translation, alternative splicing events, RNA editing, and RNA silencing.² Traditional sequencing techniques often miss small RNAs. As new technologies are developed, the prevalence and importance of these small RNA species are beginning to come to light.

On May 11–14, 2021, researchers in noncoding RNA met virtually for the Keystone eSymposium Non-Coding RNAs: Biology and Applications. Speakers highlighted the mechanisms and roles of diverse types of noncoding RNA species in areas like transcription, translation, gene activation and silencing, and protein and cell signaling regulation. In addition, several speakers expounded on the potential for noncoding RNAs to act as drug targets in diseases like cancer and CVD, while others described the structure/function relationships for lncRNAs, long thought to be mostly unstructured. Finally, speakers discussed the emergence of new technologies to systematically and comprehensively identify and map noncoding RNAs.

Keynote address: the multiple roles of Xist RNA

Edith Heard from the European Molecular Biology Laboratory presented the keynote address on elucidating the role of Xist RNA in X chromosome inactivation (XCI). In almost all mammals, one of the two X chromosomes is randomly silenced during development. This silencing is epigenetically propagated through cell divisions and enables dosage compensation between the sexes.³

Xist is an lncRNA of approximately 15,000 nucleotides in the mouse, encoded by the gene *Xist* on the X chromosome. Xist initiates XCI by coating the chromosome from which it is expressed, triggering chromosome-wide gene silencing. Xist contains several conserved regions, including the A-repeat, which is essential for chromosome-wide gene silencing; the B-repeat, which recruits Polycomb repressive complexes to modify epigenetic marks on the X chromosome; and a few other dispersed regions, which collectively help Xist coat and reorganize the X chromosome. The role of Xist is believed to be primarily in initiating XCI and helping to establish early maintenance.³

Xist initiates gene silencing by associating with sites along the X chromosome that are close in 3D space to Xist. Within a few hours, the entire chromosome is coated with Xist.⁴ Heard's group and others have shown that Xist RNA induces massive changes in the 3D organization of the chromosome, characterized by the loss of topologically associated domains and the formation of two poorly structured megadomains.⁵ As genes are silenced, they relocate into the Xist-coated territory, while a small

number of genes that escape inactivation remain located outside.^{6,7}

One possible explanation for how Xist induces gene silencing could be that it physically excludes the transcriptional machinery. This is consistent with observations that soon after Xist coating, the inactive X chromosome is depleted of transcriptional machinery factors like RNA Pol II, splicing factors, and transcription factors. Heard's group used single-particle tracking to visualize Xist and RNA Pol II molecules in live embryonic stem cells (ESCs) and showed that RNA Pol II was able to freely diffuse in and out of the Xist domain, indicating that the early Xist compartment does not present a physical barrier to the transcriptional machinery.⁸

To understand the mechanism through which Xist mediates X-chromosome inactivation, Heard's group conducted a comprehensive analysis of the time- and space-dependent epigenetic changes in the inactive X chromosome upon Xist coating using native ChIP-seq, RNA-seq/TT-seq, and allele-specific analysis. They showed that the two earliest epigenomic events are loss of H3K27ac at promoters and enhancers due to the activity of HDAC3 and accumulation of H2AK119 ubiquitination (catalyzed by PRC1). H2AK119ub1 initially accumulates at intergenic regions and spreads into genic regions after gene silencing.⁹

Studies by Heard and others have identified binding partners of Xist.^{10–15} One of the most important Xist-binding partners that has been identified in multiple screens and pull-down assays is SPEN. SPEN is a large RNA-binding protein (RBP) involved in chromatin regulation. The SPEN SPOC domain associates with nuclear co-repressors (NCoR/SMRT) and recruits HDAC3, which is necessary for the early deacetylation observed upon Xist coating. Heard's group recently showed that SPEN is essential for initiating X-chromosome inactivation. Depleting SPEN abolished X-chromosome gene silencing, while knocking out SPEN in mice phenocopied *Xist* knockout. Further investigation revealed that the SPOC domain of SPEN is critical for gene silencing and for dampening the expression of genes that escape XCI. Live and fixed-cell imaging showed that SPEN accumulates on the X chromosome simultaneously with Xist. While SPEN is not required for Xist to coat the chromosome, it is required for gene silencing. Mass

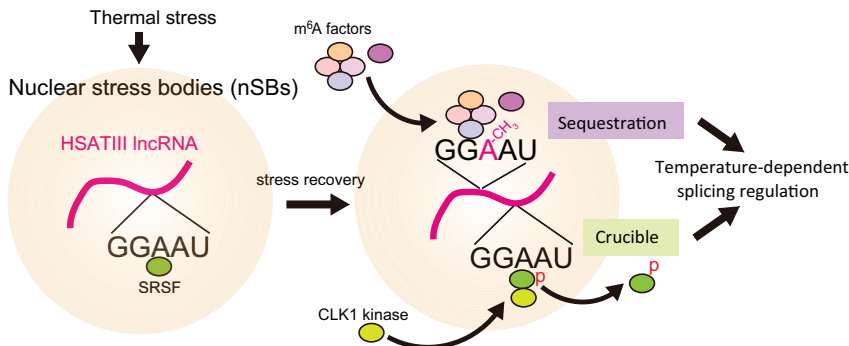


Figure 1. HSATIII arcRNAs form nSBs upon thermal stress that sequester proteins involved in intron splicing, including SRSF and SAFB. Upon recovery from thermal stress, additional factors are recruited to the nSBs, including CLK1.

spectrometry analysis revealed that the SPOC domain acts as a platform for multiple protein complexes, including the RNA Pol II transcription machinery.¹⁶

Heard put forth a model for the role of SPEN and Xist in gene silencing in which Xist's A-repeat recruits SPEN to the X chromosome, where it is targeted to transcriptionally active promoters and enhancers, likely via interactions between the SPOC domain and RNA Pol II. There, SPEN triggers gene silencing, and then SPEN disengages from chromatin, even though it continues to be recruited by Xist RNA.

Regulation of long noncoding RNA in nuclear condensates

Most noncoding RNAs retained in the nucleus are retained in membraneless nuclear bodies, like the nucleolus, nuclear speckles, and promyelocytic leukemia bodies. Recent studies have indicated that these noncoding RNAs can play an important role as regulatory or structural molecules in these nuclear domains.

Tetsuro Hirose from Osaka University presented work on delineating the functions of architectural noncoding RNAs (arcRNA), which act as a structural scaffold for membraneless nuclear bodies by sequestering IDR-RBPs and inducing phase separation.^{17,18} These membraneless bodies may have many functions, including acting as reaction crucibles for biochemical reactions, sequestering proteins, and acting as an organization hub for chromosomes.¹⁹ Hirose focused on the role of temperature-induced nuclear stress bodies (nSBs) in regulating RNA splicing. They showed that

HSATIII arcRNAs form nSBs upon thermal stress that sequester proteins involved in intron splicing, including SRSF and SAFB.²⁰ Upon recovery from thermal stress, additional factors are recruited to the nSBs, including CLK1 (Fig. 1), which phosphorylates SRSF and allows it to leave the nSB and repress intron splicing.²¹ In a separate pathway, HSAT III induces intron retention via m6A writer and reader proteins. nSBs sequester the m6A writer complex, which methylates HSATIII, leading to sequestration of the nuclear m6A reader, YTHDC1. Sequestration of the m6A writer and reader represses m6A modification of pre-mRNAs, thereby reducing m6A-dependent splicing during stress recovery.²² Hirose's work shows that nSBs can control temperature-dependent intron splicing through two distinct RNP complexes.

Noncoding RNAs in the nucleolus

Another membraneless nuclear body in which noncoding RNAs play an important role is the nucleolus. The nucleolus is the site of ribosomal RNA (rRNA) gene transcription, processing, and assembly with ribosomal proteins. It is organized around nucleolus organizer regions (NORs), which consist of tandem repeats of ribosomal genes. In humans, each NOR contains ~40 rDNA repeats that form head-to-tail tandem arrays. Noncoding RNAs are central to the nucleolar structure and function. The nucleolus consists of three major compartments: the fibrillar center (FC), which is the site of rRNA transcription; the dense fibrillar component (DFC), the site of early rRNA processing; and the granular component, the site of late rRNA processing.

Ling-Ling Chen from the Chinese Academy of Sciences presented work on RNA regulation in the nucleolus. Chen focused on sno-lncRNA that enhances pre-rRNA transcription (SLERT). SLERT is an approximately 700-nucleotide-long snoRNA found in the nucleolus that enhances pre-rRNA transcription.²³ Knocking out *SLERT* reduces pre-rRNA and ribosome synthesis and subsequently reduced cell proliferation.²³ SLERT interacts with the DEAD box helicase DDX21, which forms a shell-like structure around RNA Pol I.²³ In the nucleolus, FCs and DFCs are essential morphologically distinct subregions of mammalian cell nucleoli for rDNA transcription and pre-rRNA processing.²⁴ DDX21 shells both control FC/DFC size and suppress rDNA transcription. Chen showed that SLERT promotes Pol I transcription by changing the conformation of the DDX21 shells. In brief, DEAD box RNA helicase can adopt an open or closed conformation. In the open conformation, DDX21 can form dense hypermultimerized clusters that impair the liquidity and size of the FC/DFC units in the nucleolus. In the closed conformation, DDX21 forms loose hypomultimerized clusters that enable liquidity in the FC/DFC. SLERT keeps DDX21 in the closed state via an RNA chaperone-like mechanism at a substantially low stoichiometric ratio, thereby promoting hypomultimerization and maintaining the FC/DFC and preventing DDX21 from hijacking rDNA.²⁵

Kannanganattu V. Prasanth from the University of Illinois presented unpublished data on a family of novel nucleolus-enriched noncoding RNAs. Individual members of this class of ncRNAs associated with specific NOR-containing chromosomes and modulated rDNA gene expression. Several recent studies on the nucleolus-enriched ncRNAs, including the work from the Ling-ling Chen and Prasanth labs, have highlighted the importance of ncRNAs in rDNA gene expression and nucleolus organization.^{25–32}

Sofia A. Quinodoz in Clifford Brangwynne's lab at Princeton University presented work done in Mitchell Guttman's lab at Caltech to comprehensively map the location of all RNAs in the nucleus. Quinodoz developed SPRITE, an iterative split-and-pool tagging method that measures multiway RNA and DNA interactions in 3D space.^{33,34} Quin-

odoz recently developed a new version of SPRITE (SPRITE 2.0), which enables the detection of all classes of RNA, including low abundance noncoding RNAs and nascent pre-RNAs. They showed that SPRITE 2.0 captures known RNA–DNA interactions as well as known RNA–RNA interactions and global RNA bodies. For example, SPRITE 2.0 identified many RNA hubs via RNA interactions, corresponding to various nuclear bodies, such as the nucleolus, speckles, and Cajal bodies. Quinodoz has used SPRITE to map the 3D organization of RNA and DNA in these bodies. During the talk, Quinodoz focused on the organization of RNA and DNA in the nucleolus, though they have also applied the method to nuclear speckles, histone locus body RNAs, and Cajal body-associated RNAs. The results reveal several general principles for the organization of RNA hubs in which nascent RNAs remain near their DNA, multiple DNA loci come together in 3D space, and diffusible regulatory ncRNAs come to the hubs. Quinodoz has also used SPRITE to map the localization of over 600 lncRNAs, showing that the majority associate with chromatin and localize around their loci.^{33,34}

Noncoding RNAs as therapeutic targets

As the functional roles of noncoding RNAs become better appreciated, several have been implicated in diseases, such as cancer and CVD.^{35,36} Several talks focused on the potential to targeted misregulated noncoding RNAs to treat disease.

David L. Spector from Cold Spring Harbor Laboratory presented work on identifying and targeting lncRNAs upregulated in breast cancer that might play a role as therapeutic targets. LncRNA misregulation has been associated with various cancer types, both as prometastatic factors and tumor suppressors.³⁵ **Sarah D. Diermeier**, a former postdoc in Spector's lab, now at the University of Otago, performed an RNA-seq screen in mouse models of HER2⁺ and luminal B breast cancer to identify lncRNAs involved in breast cancer progression. They prioritized 30 lncRNAs, dubbed mammary tumor-associated RNAs (MaTARs), as promising clinically relevant targets due to the location of their genes in the genome and the lack of sequence homology with other RNA species.³⁷

Spector focused on *MaTAR25*, an intergenic lncRNA gene that localizes to the nucleus and is

upregulated in several breast cancer models but not expressed in normal tissues.³⁸ Both *in vitro* and *in vivo* assays show that loss of *MaTAR25* affects tumor formation, proliferation, invasion, and metastatic ability. Spector showed that *MaTAR25* influences cell migration by upregulating expression of tensin1, a protein found in focal adhesions.^{38,39} While high tensin1 levels had been implicated in colorectal cancer, it had not been previously associated with breast cancer. Knocking out *MaTAR25* resulted in disassembly of actin filaments and focal adhesions as well as loss of microvilli, which are critical for cell migration and invasion. Spector showed that *MaTAR25* regulates tensin1 expression via association with PURB and the *tensin1* promoter. Spector's group has identified the human ortholog of *MaTAR25*, named *LINC01271*.³⁸ *LINC01271* is overexpressed in primary breast tumors and lung metastases, with increased expression correlating with poor survival in breast cancer patients.⁴⁰ Spector's group is evaluating *LINC01271* as a potential therapeutic target that can be manipulated by an antisense-based approach to impact metastatic breast cancer progression.

Diermeier discussed another *MaTAR* identified in their RNA-seq screen, *MaTAR17*. In mammary tumor organoids, knockdown of *MaTAR17* via antisense oligonucleotides (ASOs) developed in collaboration with IONIS reduced organoid branching as a result of reduced cell proliferation and invasiveness.^{37,41} Previous work has shown that, unlike other lncRNAs, which are typically associated with one tumor type, hMATAR17 is overexpressed in several types of cancer, including head and neck, colorectal, breast, and lung cancer.⁴² Diermeier presented unpublished data elucidating the mechanisms by which *MaTAR17* drives breast and colorectal tumor progression and investigating the therapeutic potential of anti-*MaTAR17* ASOs for breast cancer. The group is also exploring the utility of *MaTAR17* as a predictive or prognostic biomarker for tracking disease progression and treatment response.

Allison M. Porman (Swain) from Aaron Johnson's lab at the University of Colorado, Anschutz Medical Campus, showed how N6-methyladenosine (m6A) modification in the lncRNA *HOTAIR* regulates its function in breast cancer. *HOTAIR* promotes metastasis in breast cancer cells by

reprogramming the chromatin state. While this was originally believed to occur via interactions with PRC2, subsequent studies found that PRC2 is not required for *HOTAIR*-mediated transcriptional repression.^{43,44} Porman identified an m6A modification site in *HOTAIR* that is required for *HOTAIR*-mediated cell proliferation and invasion in breast cancer cells. They showed that m6A at this single site enables *HOTAIR* to interact with the m6A reader YTHDC1, which is required for *HOTAIR* to associate with chromatin and repress tumor suppressor genes (Fig. 2). Porman's work demonstrates how a single site of RNA modification within an lncRNA can play a major role in cell phenotype.⁴⁵

Dhiraj Kumar from the University of Texas MD Anderson Cancer Center presented unpublished work on how another lncRNA, *Malat1*, which has been identified as a mediator of metastatic reactivation,⁴⁶ promotes metastatic reactivation in breast cancer by regulating the immune microenvironment at the metastatic site.

Mark W. Feinberg from Brigham and Women's Hospital/Harvard Medical School focused on the role of lncRNAs in atherosclerosis. While roles for lncRNAs have been identified in different aspects of CVD,^{36,47} a direct role in vessels is poorly understood. RNA-seq studies of blood vessels in mice fed a high-cholesterol diet identified *SNHG12* as an lncRNA that decreases with atheroprogession and increases with atheroregression (Fig. 3). While knocking down *SNHG12* increased plaque burden in mice, the lesions showed no differences in proinflammatory signaling pathways compared with wild-type lesions. Feinberg showed that reduced expression of *SNHG12* increases atherosclerosis independent of lipid profile or lesional inflammation via DNA damage and vascular senescence. *SNHG12* acts to control DNA damage by stabilizing the interaction between DNA-PK and Ku70/Ku80. Knocking down *SNHG12* increased DNA damage, impaired nonhomologous end-joining (NHEJ), and increased cell senescence. These effects can increase endothelial permeability to low-density lipoprotein.⁴⁸ Feinberg's work builds upon increasing evidence suggesting that senescence can independently promote atherosclerosis.^{49,50} Their group has also identified and worked to understand

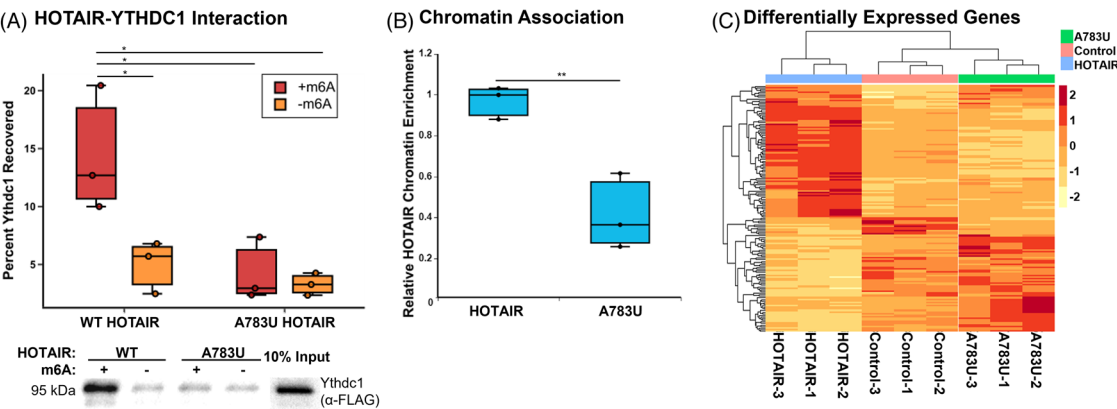


Figure 2. An m6A modification site in HOTAIR required for HOTAIR-mediated cell proliferation and invasion in breast cancer cells.

the mechanism for other lncRNAs involved in atherosclerosis.^{51–54}

Mechanisms of long noncoding RNA function

Transcription by RNA Pol II occurs via a series of sequential events, including promoter binding, transcription initiation, and transcription elongation. Each of these events is quite inefficient, such that only 1% of polymerases that bind to a promoter produce an mRNA transcript. Many lncRNAs have been shown to modulate transcription at sites near their genes, though the exact mechanisms are not clear.⁵⁵

Jonathan E. Henninger from Richard Young’s lab at the Whitehead Institute for Biomedical Research presented a model in which these nascent RNAs regulate Pol II transcription via a dynamic feedback loop. During transcription, short RNAs are produced from enhancer and promoter elements that do not have a clear function.⁵⁶ Henninger’s model is based on a type of phase separation called complex coacervation, whereby oppositely charged polymers form condensates through electrostatic interactions. In the model, during initial stages of transcription, electrostatic interactions between negatively charged RNA and positively charged proteins balance each other and form condensates. As transcription proceeds, negatively charged RNAs accumulate, the repulsive electrostatic interactions dissolve the condensate, and the system returns to a diffuse phase. The model predicts that

nascent RNAs initially stimulate and then arrest transcription, which was validated both *in vitro* and *in vivo*. Given that this electrostatic-based model is dependent on RNA size and not sequence, Henninger believes that it may explain the role of many lncRNAs, which often show little sequence conservation but are primarily located near sites of transcription, in regulating gene expression of nearby genes.⁵⁷

Several lines of evidence suggest that there are key regulatory events that stabilize RNA Pol II binding for transcription elongation;^{57–59} yet the molecular link between RNA and the transcription machinery has not been elucidated.

Xiaohua Shen from Tsinghua University presented a phase-separation mechanism by which noncoding RNAs act with RBPs to aid in the rate-limiting step of polymerase condensate formation to promote polymerase engagement and transcription. Shen showed that RBPs are prevalent among chromatin-bound proteins and are susceptible to liquid–liquid phase condensation on chromatin. Many RBPs colocalize with RNA Pol II at sites of transcription, and the degree of RBP binding to a promoter positively correlated with the level of mRNA expression, suggesting that RBPs might act collaboratively to regulate transcription. She showed the paraspeckle protein PSPC1 inhibits the RNA-induced premature release of Pol II and makes use of RNA as multivalent molecules to enhance the formation of transcription condensates and subsequent phosphorylation and release of Pol

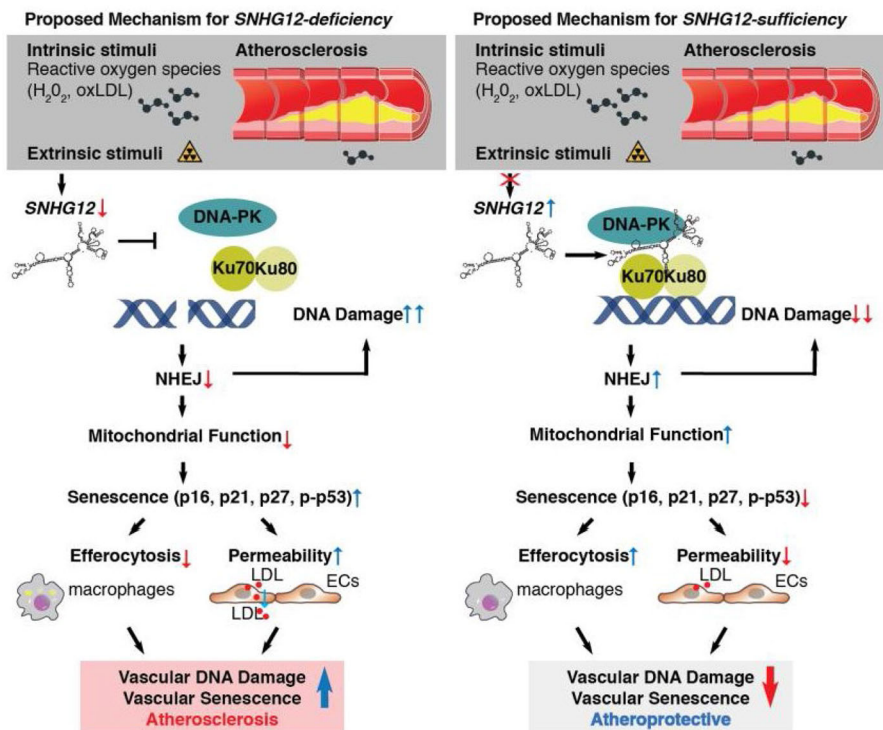


Figure 3. Proposed mechanism of lncRNA SNHG12 regulation of atherosclerosis through a DNA-PK-mediated DNA damage response in the vascular endothelium.

II. This synergistic interplay enhances polymerase engagement and activity via the RNA-binding and phase-separation activities of PSPC1. In ESCs, auxin-induced acute degradation of PSPC1 leads to genome-wide defects in Pol II binding and nascent transcription. Shen put forth a model in which RBPs stabilize the interaction between RNA Pol II and the transcription site via phase separation. During transcription, RNA Pol II produces nascent transcripts, which can evict it from the gene promoter before phosphorylation of the C-terminal domain, which is required for transcription elongation. The nascent transcripts can also recruit RBPs, which, once reaching a critical threshold, create an RBP-rich condensate that occludes RNA Pol II eviction, allowing it to be phosphorylated, and to proceed to the elongation phase.⁶⁰ The interplay between noncoding transcripts and RBPs around the transcription sites regulates dynamic assembly of polymerase clusters in cells, expanding the horizon in understanding transcription regulation beyond classic transcription and epigenetic factors.

One of the barriers to understanding the role of lncRNAs on gene transcription is the lack of explicit definitions of the types of regulatory lncRNAs. Some lncRNAs regulate gene expression in *cis*, meaning that their gene targets are located nearby on the same chromosome, while some regulated gene expression in *trans*, meaning that their targets are on a different chromosome or distally on the same chromosome. **Rory Johnson** from University College Dublin presented a new tool to identify *cis*-regulatory lncRNAs. Johnson's group is developing tools to classify lncRNAs as either *cis* or *trans* acting to determine the prevalence of each modality and to define general rules for how *cis*-acting lncRNAs work.

Joshua T. Mendell from Howard Hughes Medical Institute, University of Texas Southwestern, discussed how *NORAD* regulates Pumilio (PUM) activity. The lncRNA *NORAD* is a ubiquitously expressed cytoplasmic noncoding RNA that is activated by DNA damage.^{61–66} In contrast to most lncRNAs, it contains several repeat units, some

of which form conserved secondary structures, while some consist of Pumilio recognition elements (PREs) and bind PUM proteins. PUM represses mRNA translation by binding to PREs in mRNA transcripts and accelerating their degradation.

Given that the number of PUM target mRNAs vastly outnumbers the number of PUM-binding sites on *NORAD*, Mendell's group is interested in how *NORAD* can outcompete PUM target mRNA-binding sites to effectively inhibit PUM activity. Mendell showed that *NORAD* and PUM colocalize in punctate foci in the cytoplasm, which they termed *NORAD*-Pumilio (NP) bodies. These foci disperse in the absence of *NORAD*. PUM contains several features associated with the ability to form condensates, including large intrinsically disordered regions (IDRs) and the ability to bind RNA. They proposed that *NORAD* sequesters PUM in NP bodies, which are formed through *NORAD*-induced PUM phase separation. *In vitro* studies showed that at physiologic concentrations, *NORAD* drives phase separation of PUM. Mendell showed that *NORAD* can recruit PUM into the droplet phase not only via direct binding to PREs but also via PUM–PUM IDR interactions. This allows *NORAD* to outcompete other PRE-containing RNAs, as it can sequester approximately three times as many PUM proteins as it would be able to via PRE binding alone. Moreover, each PRE in *NORAD* can recruit PUM ~40 times more efficiently than an isolated PRE in a target mRNA. Therefore, in the cell, the effective number of *NORAD* PREs is comparable to the total number of PUM-binding sites on target mRNAs. Mendell also showed that sequestration of PUM in phase-separated condensates is important for regulating PUM activity and maintaining genomic stability in mammalian cells.⁶³

Juan Pablo Unfried from Puri Fortes's lab at the University of Navarra discussed their work on lncRNAs associated with hepatocellular carcinoma (HCC). Unfried previously published a list of lncRNAs that are highly, preferentially expressed in cancer, including HCC, that may serve as attractive drug targets. Several of these are associated with clinical outcomes, including survival, suggesting that they may play a role in disease pathology.⁶⁷ Unfried presented data on one such lncRNA, *NIHCOLE*.⁶⁸ Depletion of *NIHCOLE* from HCC

cells led to impaired proliferation, increased apoptosis, and DNA damage accumulation due to a specific decrease in the activity of the NHEJ pathway of DNA double-stranded break (DSB) repair. DNA damage induction further decreased the growth of *NIHCOLE*-depleted cells. Authors found that *NIHCOLE* associates with DSB markers and recruits several molecules of the Ku70/80 heterodimer. Band-shift analyses showed that structural domains within *NIHCOLE* support multimeric repair complexes, including the ligation complex formed by XRCC4 and DNA Ligase IV. This allows *NIHCOLE* to promote the ligation efficiency in NHEJ reconstitution assays by serving as a scaffold and facilitator of the NHEJ pathway. Collectively, they showed that *NIHCOLE*'s role in DSB repair confers an advantage to HCC cells, uncovering a novel cancer vulnerability that could be targeted with therapeutic purposes.

Enhancer RNAs, circular RNAs, and small RNAs

Several talks focused on the regulatory roles of small RNA species, such as small nuclear RNA (snRNA), piRNA, miRNA, and circRNA.

Jeremy E. Wilusz from the University of Pennsylvania showed how Integrator (Int), a multisubunit complex that associates with RNA Pol II, is regulated.⁶⁹ Integrator contains an endonuclease, IntS11, that cleaves nascent snRNA transcripts, which mature into small nuclear ribonucleoproteins (snRNPs) and catalyze splicing.⁷⁰ Wilusz's lab has shown that Int can also regulate protein-coding genes by cleaving nascent mRNA transcripts, triggering premature transcription termination that results in short, nonfunctional transcripts that are degraded.^{71,72} While these two roles for Int have similar characteristics molecularly—they both involve recruitment of Int by RNA Pol II, endonuclease-mediated cleavage, and subsequent release of an RNA transcript—they seem to have highly distinct regulatory needs. When Int cleaves at an snRNA gene, it results in a functional transcript; therefore, constitutive Int activity is required for the production of functional snRNPs. Conversely, Int activity at a protein-coding gene abolishes the production of functional mRNA, indicating that Int activity must be regulated so that protein-coding genes can be expressed under the appropriate

cellular conditions. Wilusz presented unpublished data elucidating how Int is differentially regulated at snRNA genes versus at protein-coding genes.

The piRNA pathway is the largest and most diverse class of small RNAs. Its most well-understood function is as a defense mechanism against transposons to protect germ cells against the threat of mobile genetic elements. piRNAs bind to and cleave mRNAs, thus negatively impacting translation. In the nucleus, this can also lead to epigenetic changes and transposon silencing. While many piRNAs match to transposons, many do not.^{73–75}

Lamia Wahba from Andrew Fire's lab at Stanford University presented work on understanding the function of nontransposon piRNA. Many possible mechanisms for nontransposon piRNAs have been proposed, including transposon defense through imperfect complementarity, massive RNA elimination, and translational activation or repression. Wahba presented unpublished data in *C. elegans*, in which over 99% of the 16,000 piRNAs do not map to transposons, to elucidate the mechanism of how loss of the piRNA pathway results in gradual sterility across generations.

Adelheid Lempradl from the Van Andel Institute discussed how the piRNA pathway can play a role on inheritance of metabolic phenotype. Several studies indicate that effects of a parent's environment can be inherited in their offspring. For example, Lempradl had previously established a model for diet-induced intergenerational metabolic reprogramming in *Drosophila* in which fathers fed a high-sugar diet produced offspring with a higher susceptibility of diet-induced obesity.⁷⁶ Other studies have observed similar phenotypic inheritance patterns with metabolic disorders and trauma.^{77,78} Lempradl has investigated the role of the piRNA pathway in mediating this intergenerational heritability. Most research on the piRNA pathway in *Drosophila* has been in females. Lempradl showed that the piRNA pathway is also present in mature sperm and is necessary for intergenerational inheritance of paternal metabolic state. They identified piwi-bound small RNAs in sperm and provided evidence that these RNAs repress target gene expression in offspring.⁷⁹

Sean E. McGeary from David Bartel's group at the Whitehead Institute presented work on identifying the rules that govern miRNA binding to target mRNAs. miRNAs mediate repression of targets by binding to sites within the 3' UTRs of target mRNAs. Binding occurs via pairing between two key regions: the seed region and the 3'-pairing region.⁸⁰ Previous studies to identify the rules governing pairing to the miRNA 3' end were based on transcriptome-wide analyses of 11 miRNA transfections, which may have caused miRNA-specific differences in 3'-pairing preferences to go undetected. Several studies have shown that there might be more miRNA-specific differences that govern 3'-pairing rules.^{81,82} McGeary used RNA bind-n-seq to more precisely measure the binding affinity of thousands of 3'-pairing architectures by varying the length of the 3' pairing, the starting position of the 3' pairing, the offset of the 3' pairing with respect to the seed pairing, and the nature of the seed pairing. Analyses of these measurements revealed two distinct binding modes shared by multiple miRNAs and allowed them to identify optimal 3'-pairing architectures for different miRNAs. McGeary has also developed a mathematical model-based framework to understand the relationship between 3'-pairing length, position, and offset.⁸³

Chun-Kan Chen from Howard Chang's lab at Stanford University discussed the elements that drive translation of circRNA. Circular RNA is generated by spliceosome-mediated back-splicing forming a head-to-tail closed loop.^{84–86} Previous work has shown that circRNAs can act as sponges for miRNA and RBPs.^{87–92} However, circRNAs can also serve as templates for translation.^{93–99} Translation of linear mRNA is initiated by ribosomal recognition of the 5' cap. Circular RNA has no 5' cap. Instead, the ribosome recognizes internal ribosomal entry sites (IRES) that mimic the structure of the 5' cap.^{96,98,100–106} Given that most of the research on translation has focused on linear RNA, much less is known about the RNA elements that are required for circRNA translation. Chen has systematically identified IRES sequences in circRNAs (Fig. 4). Using these, they were able to identify potential protein-coding circRNAs and showed how the resulting peptides can regulate cell function (see <https://doi.org/10.1016/j.molcel.2021.07.042>).

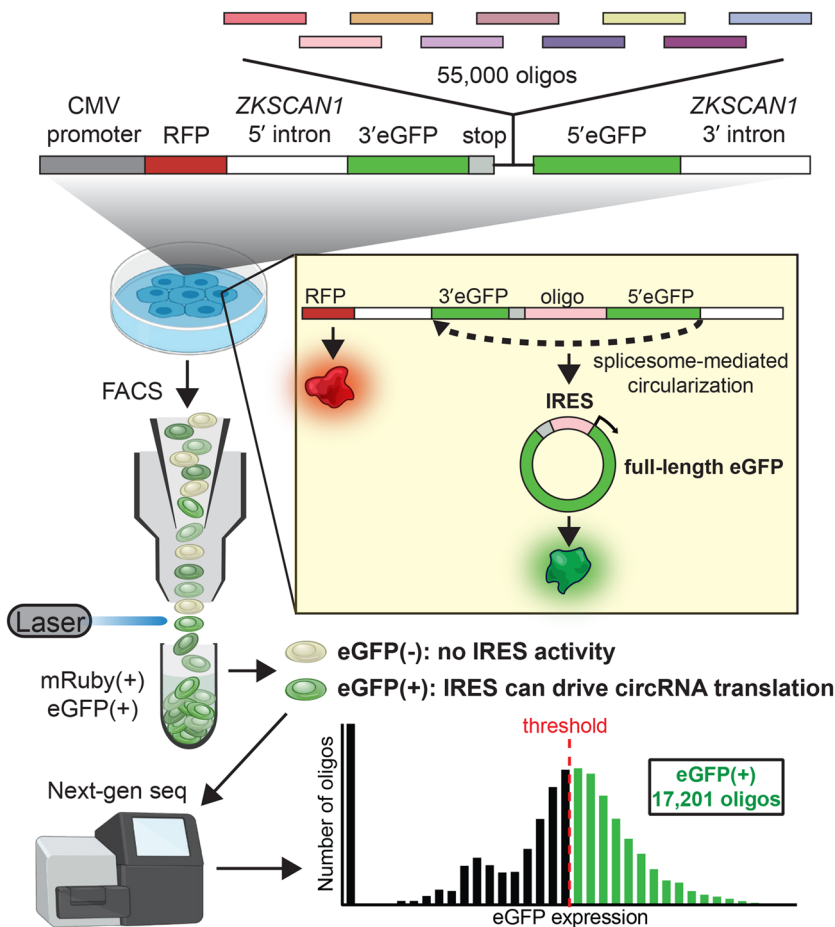


Figure 4. IRES sequences identify potential protein-coding circRNAs.

RNA structure and function

Anna Marie Pyle from Yale University discussed the importance of structure in lncRNAs. Many lncRNAs do not have specific structural elements, but it is becoming increasingly clear that some of them have modules of functional structure. Given that one of their key roles is to increase the effective RNA Pol II concentration at a locus, many believe that their sequence and structure is largely irrelevant. Pyle’s group was the first to solve a high-resolution structure of a group II self-splicing intron,¹⁰⁷ which was largely believed to lack any tertiary structure owing to its lack of sequence conservation. Similar assumptions have been made about lncRNAs. Pyle’s group has conducted secondary structure analyses on the lncRNA *HOTAIR*. As noted by Swain previously during

the meeting, *HOTAIR* is differentially expressed along the body axis and overexpressed in cancer.¹⁰⁸ Pyle showed that *HOTAIR* contains four distinct, modular secondary structural domains. Covariance analysis identified patterns of phylogenetic covariance and conservation in some of the stem regions.¹⁰⁹ While lncRNAs are often noted for their lack of sequence conservation, Pyle stressed that experimentally determined secondary structures can be instrumental in identifying regions of sequence conservation and covariance by providing a framework for organizing elements of primary sequence in space. Pyle’s group had adapted the tool R-Scape¹¹⁰ to make it more robust at identifying conservation and covariation in longer RNAs. Their tool was released as R-Scape APC-RAFS.¹¹¹ Pyle also discussed unpublished structural studies on another lncRNA, *Pnky*, as well

as a new high-throughput approach their lab is developing to rapidly detect tightly folded tertiary motifs.

Amanda E. Hargrove from Duke University discussed work on targeting lncRNAs with small molecules. One challenge to developing RNA-targeting small molecules is the relative paucity of high-resolution structures. Hargrove's group has been working to optimize selective RNA-binding small molecules in the absence of such structures. Toward this goal, they have developed a database, R-BIND, that contains approximately 150 bioactive ligands that target non-rRNA.^{112,113} Hargrove focused on their group's work in scaffold-based library design, which has revealed insights on the factors that drive RNA-binding affinity and specificity. On the basis of data showing that dimethyl amiloride is a weak, selective binder of HIV-1 TAR RNA,¹¹⁴ Hargrove's group synthesized and screened a library of amiloride (AMR)-based small molecules and identified a molecule with 100-fold increased binding affinity. They developed a model whereby simple molecular properties, such as molecular weight and charges, can be used to predict which ligands selectively bind to HIV-1 TAR RNA.^{115,116} Hargrove's group has expanded upon this model to make it more generalizable for a range of RNAs and scaffolds. They hope that general rules can guide prediction of ligand structure–activity relationships in the absence of high-resolution structures. In a similar vein, screening AMR compounds for binding to an enterovirus RNA identified a small molecule that selectively inhibited viral translation and replication.¹¹⁷ Finally, Hargrove discussed work on a separate scaffold, DPF, to target the triple helix of *MALAT-1*, the stability of which is important to prevent degradation.¹¹⁸ DPF-based compounds bind to the triple helix with a range of affinities and selectivities. The data revealed the importance of shape, with rod-like compounds displaying higher affinity and selectivity.¹¹⁹ Interestingly, the degree of destabilization did not correlate with binding affinity.¹²⁰ Hargrove's group is screening additional small molecules to identify stabilizing compounds and tease out which features are important for de/stabilization, selectivity, and affinity.

Matthew D. Simon, from Yale University, presented work on understanding the role of promoter prox-

imal pausing in gene regulation. During the initial stages of transcription, RNA Pol II pauses after transcribing approximately 40 nucleotides. This pause allows cells to either release the paused RNA Pol II and continue transcription or to terminate transcription. Simon presented data on determining the rates of transcription termination and RNA Pol II release at promoter proximal pause sites throughout the genome and demonstrated how various perturbations can affect these rates to regulate transcription.¹²¹

Marco Marcia from EMBL Grenoble presented structural-functional studies for the lncRNA MEG3. Marcia was involved in developing a representation of group II intron splicing using nearly 20 high-resolution structures and molecular dynamics to capture catalytic RNA in the act of splicing.^{122–124} They plan to similarly visualize the activity of lncRNAs using structural studies. MEG3 is involved in tumor progression via regulating the expression of p53.^{125–127} *In vitro* and *in vivo* probing analyses of MEG3 revealed five secondary structure domains. Using a gene reporter assay of p53 activation, Marcia's group demonstrated that two of these domains, domain 2 (D2) and domain 3 (D3), are essential for MEG3-mediated p53 activation. They also identified elements within these domains that contribute to p53 stimulation. Some of these regions are highly conserved, although conservation was not picked up via other structure-free methods. Marcia focused on an invariant loop within D2 that is required for p53 activation. Regions within D3 are complementary to this loop, suggesting the potential for a long-range tertiary interaction (Fig. 5). Indeed, introducing compensatory mutations between these two regions rescued the activity of the single-point mutations, suggesting that a long-range tertiary interaction between these sites does control MEG3-mediated p53 expression. Consistent with this idea, atomic force microscopy studies showed that MEG3 forms a distinct, compact 3D structure that functionally inactive mutants cannot adopt.^{128,129}

Róża K. Przanowska from Anindya Dutta's lab at the University of Virginia presented work on the structure–function relationship in the lncRNA MUNC. MUNC has two isoforms, MUNC spliced and MUNC genomic, both of which are involved

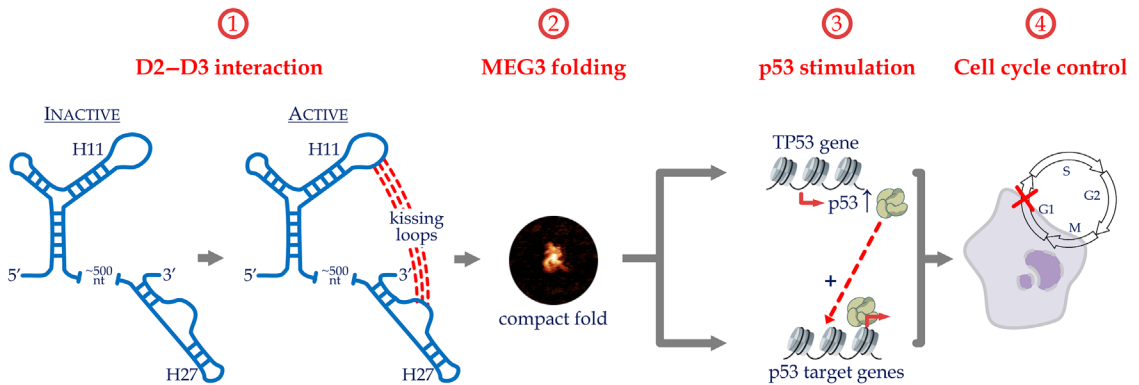


Figure 5. An invariant long range tertiary interaction between domains 2 (D2) and 3 (D3) of MEG3 is required for p53 activation mediated by this lncRNA.

in muscle satellite cell differentiation.¹³⁰ RNA-seq analysis of myoblasts overexpressing either isoform revealed that, while both isoforms are myogenic, they regulate different sets of genes, and MUNC spliced is a stronger promyogenic factor. SHAPE-MaP revealed six secondary structural elements present in both isoforms that were not predicted by folding algorithms.¹³¹ The data also revealed that MUNC spliced is more stable structurally, whereas MUNC genomic exists in a more dynamic unfolded state. Przanowska used the SHAPE-MaP-predicted secondary structure to delete and mutate various structural regions and determine which regions were responsible for MUNC function. They identified distinct structural domains responsible for regulating different genes and for binding to different genomic sites and cohesin complex.¹³²

RNA technologies

Howard Y. Chang from Stanford University discussed work in understanding the role of the X chromosome and Xist in sex-biased immunity. Several lines of evidence support a sex bias in immunity, including particularly the preponderance of autoimmune diseases among women¹³³ and the improved outcomes in women to COVID-19.¹³⁴ As Heard discussed in the keynote address, Xist is an lncRNA that coats and silences the X chromosome. While long thought to be primarily involved in establishing X-chromosome inactivation, Chang showed that a subset of X-linked genes requires ongoing Xist silencing in adult female B cells, including *TLR7*, a gene involved in innate immunity, in adult female B cells. Genes

that reactivate upon Xist silencing have low levels of methylation, which is known to maintain X-chromosome inactivation. Chang showed that Xist binds to many factors in a cell type-dependent fashion. In B cells, interaction between Xist and TRIM28, which enforces RNA Pol II pausing, is key to maintaining X-chromosome inactivation. Chang put forth a model in which TRIM28 binds to target genes and interacts with Xist. This interaction enhances the ability of TRIM28 to pause RNA Pol II, thereby blocking gene expression. Chang also showed single-cell RNA-seq data of B cells from female COVID-19 patients that demonstrate a population of atypical B memory cells (ABC) that have high levels of Xist escape. Chang proposed that expansion of these ABCs increases the reactivity to self-antigens but can also lead to a more effective immune response against some pathogens, such as SARS-CoV-2.¹³⁵

Samie R. Jaffrey from Weill Medical College of Cornell University discussed work on developing a system to express circRNA in mammalian cells. Expressing RNA at high levels in mammalian cells is a major challenge because of their high susceptibility to exonucleases. Jaffrey's group had previously leveraged an endogenous pathway that generates circRNAs, which are resistant to exonucleases, from a tRNA intron.¹³⁶ The resulting circRNA was stable in mammalian cells but only present at low levels. To optimize the production of circRNA, Jaffrey's group inserted RNA sequences into a self-splicing bacterial Twister ribozyme that generates 3' and 5' ends that are recognized by the tRNA ligase that

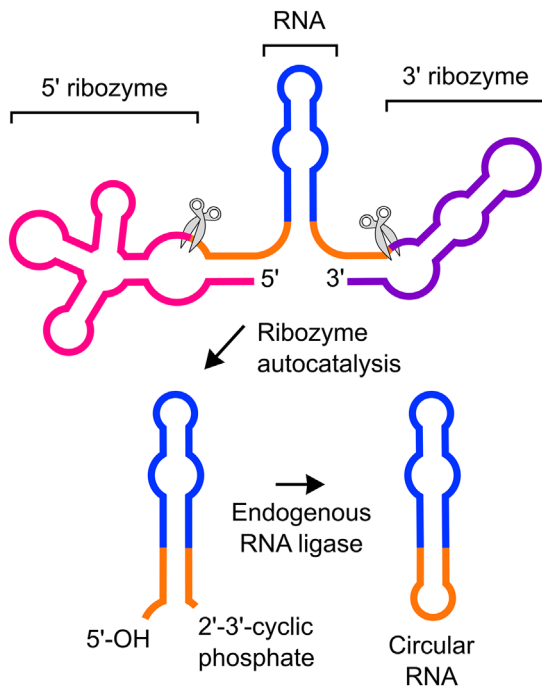


Figure 6. Tornado generates circRNA species that are highly stable and expressed at high levels in mammalian cells.

creates circRNA (Fig. 6). The resulting system, dubbed Tornado, generates circRNA species that are highly stable and expressed at high levels in mammalian cells.¹³⁷ Jaffrey showed how the Tornado system can be used to express engineered RNA-based aptamers that can act as metabolite sensors, interact with endogenous proteins to affect signaling pathways, act as decoys to block the activity of RBPs, or be translated into protein. Jaffrey hopes that the Tornado expression system will expand the use of RNA as tools to control mammalian cell function and potentially to be used as a new therapeutic modality.

Given that high-resolution structures of RNA species are difficult to obtain, other technologies are needed that can characterize RNA structural elements.

Lydia M. Contreras from the University of Texas at Austin presented work on quantifying the accessibility within RNA species to potentially characterize functionally important areas. Contreras noted that accessibility can give insights into which areas are more likely to mediate interactions with binding partners (Fig. 7). Their group has developed

a high-throughput method that measures regional RNA accessibility *in vivo* (and, more recently, *in vitro*)¹³⁸ to understand interfaces available for binding interactions that may play a role in regulatory processes.^{139,140} In this system, an RNA probe designed to target a region (9–16 nucleotides) of a target RNA is incorporated into a riboregulator, which leads to GFP expression upon RNA binding.^{139–141} A high-throughput version for use in *E. coli* was also developed that uses RNA transcript length as the output and that can probe hundreds of molecules at a time.¹⁴⁰ Contreras showed how RNA accessibility can be assessed under various conditions within cells to pinpoint regions important for regulation and/or to understand how intracellular factors affect RNA accessibility.^{140,142} For example, by using accessibility as a proxy for CsrA occupancy (low accessibility = high occupancy), the CsrB sRNA was rationally reengineered for modular activity.¹⁴² Accessibility insights for other sRNAs, namely, extreme accessibility found indicative of target binding, have supported the rational selection of predicted mRNA targets for experimental follow-up.¹⁴³ They are currently modifying the method for detection of long-range tertiary interactions within RNA.

Small RNA species are often underestimated in RNA-seq studies; this is, in part, because RNA modifications can block reverse transcription, and/or unique termini features can occlude detection.

Qi Chen from the University of California Riverside presented a new RNA-seq method that detects small RNAs often missed by traditional sequencing, dubbed Panoramic RNA Display by Overcoming RNA Modification Aborted Sequencing (PANDORA-seq). In PANDORA-seq, samples are treated with enzymes that eliminate reverse transcriptase-blocking modification and convert noncanonical termini so that they can be sequenced.¹⁴⁴ Conducting PANDORA-seq on multiple human and mouse tissues/cells revealed an RNA landscape containing enriched tsRNA and rsRNA species that had not previously been identified. Additional data suggest that these small RNAs may be functionally relevant. Chen showed that tsRNA and rsRNA expression is tissue specific and that they can regulate the mouse embryonic stem cell differentiation process. To annotate and analyze the different types of small ncRNAs,

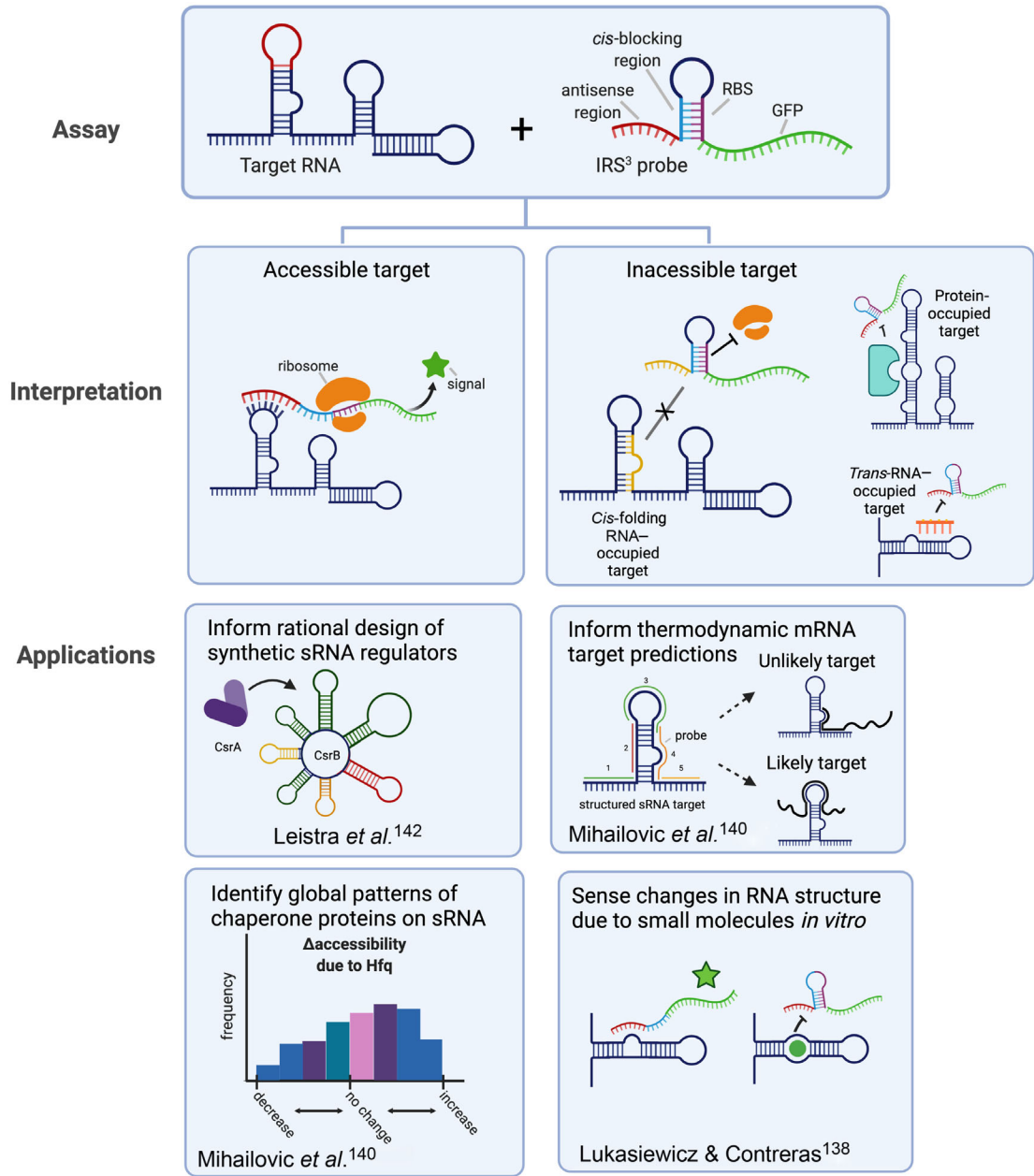


Figure 7. Probing accessible subregion within RNA. Models of understanding regional RNA accessibility and interfaces available for binding interactions that may play roles in regulatory processes.

Junchao Shi, from Chen’s group, developed the software *SPORTS*.¹⁴⁵ Shi showed that *SPORTS* performs admirably, both in precision and sensitivity, compared with other annotating software at identifying miRNA.¹⁴⁶ Moreover, *SPORTS* can also annotate other RNA species, including piRNA,

tsRNA, and rsRNA. *SPORTS* showed that tsRNAs and rsRNAs are derived from different regions within tRNAs and rRNAs, respectively. As an example of translational application, *SPORTS* was used to identify a small RNA-based signature, TRY-RNA, that consists of tsRNA, rsRNA, and

ysRNA and can distinguish between control, lung cancer, and tuberculosis subjects, demonstrating the potential for these types of tools in lung cancer screening and potentially other diseases.¹⁴⁷

Noncoding RNA in development and disease

Lin He from the University of California at Berkeley discussed how retrotransposons can affect cell differentiation during preimplantation development. The transcriptome is highly enriched with retrotransposons during preimplantation development,^{148,149} and thus these mobile elements may play an important developmental role during this stage. For example, *MERVL*, a family of LTR retrotransposons, is an important marker of totipotency.^{150–153} RNA-seq studies of various mammals showed that retrotransposons exhibit dynamic expression during preimplantation development. He showed that retrotransposon-derived promoters can affect host gene regulation, generating gene isoforms with an altered ORF. One such example is the MT2B2 promoter for *Cdk2ap1*, which is strongly induced in morulae. Activation of MT2B2 shifts the transcription start site of *Cdk2ap1*, resulting in an N-terminally truncated protein isoform. This alternative promoter is essential for preimplantation development; deleting it in mice resulted in partial embryonic lethality. Interestingly, while MT2B2 retrotransposons are mouse specific, the resultant MT2B2-induced *Cdk2ap1* isoform is evolutionarily conserved, albeit with species-specific expression regulation. The level of the N-terminally truncated *Cdk2ap1* inversely correlated with development timing in mammalian preimplantation embryos. The high level of MT2B2-dependent *Cdk2ap1* isoform in mice is correlated with a rapid progression of its preimplantation development. He put forth a model in which specific retrotransposon promoters can yield novel gene isoforms during preimplantation embryos that are essential for proper, species-specific development.¹⁵⁴

Erwei Song from Sun Yat-sen University discussed how the lncRNA *IRENA* affects macrophage phenotype in response to chemotherapy. Tumor-associated macrophages (TAMs) are the most abundant infiltrating inflammatory cells in various cancers.¹⁵⁵ In treatment-naïve breast cancer

patients, TAMs exhibit an immunosuppressive phenotype, which may drive out T cells and promote tumor progression.¹⁵⁶ Song showed that chemotherapy treatment changes the phenotype of TAMs toward an IFN-activated, proinflammatory phenotype, which enhances antitumor immunity but also promotes resistance to chemotherapy. TAMs isolated from patients after receiving neoadjuvant chemotherapy were able to enhance chemotaxis and tumor-specific cytotoxicity of CD8⁺ T cells, whereas TAMs isolated before chemotherapy were not. In mice, macrophage deficiency reduced chemotherapy-induced infiltration of CD8⁺ T cells but enhanced response to chemotherapy. These two effects of postchemotherapy TAMs—increased immune response and chemoresistance—were mediated by distinct mechanisms. Song showed that chemoresistance was mediated by NF-κB activation, while antitumor immunity was mediated by JAK1–STAT1 signaling. Song showed that the lncRNA *IRENA* can activate the NF-κB pathway following chemotherapy via its interaction with PKR. Knocking down *IRENA* abrogated the chemoresistance effect of macrophages in coculture with tumor cells without affecting the antitumor immunity effects on T cells. Song showed that *IRENA* activates NF-κB by interacting with PKR. In a mice breast cancer model, *IRENA* knockout improves response to chemotherapy and abrogates NF-κB but not JAK1–STAT1 signaling. In human breast cancer tissue, *IRENA* expression in TAMs correlates with chemoresistance and poorer survival.¹⁵⁷

John L. Rinn from the University of Colorado Boulder discussed the importance of splicing in lncRNA function. Rinn focused on the lncRNA *Tug1* and TERT mRNA. Both RNA species often retain some of their introns. Rinn showed that splicing is important for subcellular localization. Specifically, transcripts that maintain their introns localize to the nucleus, whereas spliced transcripts localize to the cytoplasm. Rinn showed that the localization pattern of *Tug1* is important for its function. Blocking *Tug1* splicing with a nuclear-localized modified ASO, thiomorpholino (TMO), depleted the cytoplasmic *Tug1* pool, and shifted localization to the nucleus, which had a negative impact on cell viability. In the case of TERT, the mRNA normally localizes to the nucleus, which would preclude

its translation. Rinn showed that retention of two introns (intron 11 and 14) maintains TERT mRNA nuclear localization. During mitosis, the nuclear membrane breaks down, TERT mRNA enters the cytoplasm, and intron 11 is spliced. Blocking intron 11 splicing with TMOs also reduced cell viability. Rinn's work suggests that splicing is important for lncRNA localization and function.¹⁵⁸

Mukesh Kumar Lalwani from Andrew Baker's lab at the University of Edinburgh presented unpublished work on the role of the lncRNA *CARMN* in heart regeneration. After injury, the heart suffers from loss of cardiomyocytes, which are replaced by noncontractile fibroblasts, which can lead to contractile dysfunction and heart failure. While adult mammals cannot induce cardiomyocyte proliferation, neonates and lower vertebrate species can regenerate and proliferate cardiomyocytes after injury.¹⁵⁹ Several lncRNAs are induced upon injury in cardiomyocytes.^{160,161} Lalwani focused on the impact of *CARMN*. Baker's group previously showed that loss of *CARMN* can enhance vascular smooth muscle cell proliferation, migration, inflammation, and plasticity, and accelerate atherosclerosis in mice.¹⁶² In addition, differential expression of *CARMN* isoforms in cardiac progenitor cells affects differentiation at different stages of development. In the fetal heart, *CARMN* isoforms promote differentiation into cardiomyocytes, while in the adult heart, *CARMN* isoforms promote differentiation to smooth muscle cells.^{163,164} Lalwani is investigating the role of *CARMN* in cardiomyocyte proliferation and cardiac regeneration in zebrafish.

Murat Can Kalem from John Panepinto's lab at the State University of New York, University at Buffalo, presented unpublished work on whether a protein post-translational modification called arginine methylation in fungi. Kalem showed that arginine methylation by Rmt5 can regulate lncRNA expression in a *Cryptococcus neoformans* infection. *C. neoformans* is an environmental fungus and a human pathogen that can cause cryptococcal meningitis in immunocompromised people. RNA-seq data have identified approximately 3000 differentially expressed genes in *C. neoformans* at the site of human meningitis, of which approximately 500 are lncRNAs.¹⁶⁵ Gene regulation is

important for *C. neoformans* to adapt to different microenvironments during infection, such as elevated temperature. One of the key factors that enables *C. neoformans* to grow at elevated temperature is the arginine methyltransferase Rmt5. Rmt5 regulates a group of lncRNAs that are also differentially expressed during infection. Arginine methylation can potentially affect protein-protein and protein-RNA interactions. Given the importance of Rmt5 in *C. neoformans* survival and expression of lncRNAs at sites of infection, Kalem is investigating whether there is a link between arginine Rmt5 target proteins and lncRNA function.

Repetitive RNAs in gene regulation

Transposable elements (TEs) are highly repetitive DNA regions that represent over half of the genome. While most TEs are not protein coding, they have been shown to be an abundant source of various kinds of noncoding regulatory sequences, such as promoters, enhancers, noncoding RNAs, and short RNAs.¹⁶⁶

Edward B. Chuong from the University of Colorado Boulder discussed how retrotransposons can regulate innate immunity. They focused on the endogenous retrovirus (ERV) *MER41*, of which approximately 9000 copies are present in primate genomes. *MER41* contains a pair of STAT1-binding motifs that can interact with STAT1 and impact the expression of nearby genes. In particular, Chuong showed that *MER41* elements exhibit IFN-inducible regulatory activity in macrophages. Upon IFN stimulation, STAT1 can bind to *MER41.AIM2*, which is located near the inflammasome component absent in melanoma 2 (*AIM2*) and induce *AIM2* expression. Chuong also showed that *MER41* elements can impact the expression of several other genes, indicating that ERVs have been co-opted to act as IFN-stimulated enhancers for a number of antiviral and inflammatory genes.¹⁶⁷ Their group is continuing to study how TEs influence the evolution of immune regulation and variation and how they impact regulatory elements in disease. Toward that goal, Chuong presented unpublished work on a separate EVR, *LTR10*, and its role in cancer. TE activation is common in cancer cells.^{168–172} Chuong's group is working to understand how the elements

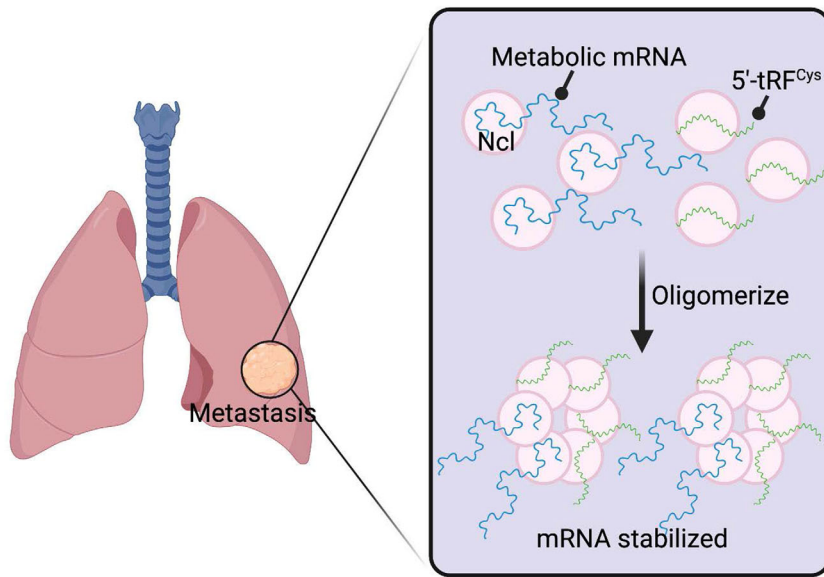


Figure 8. Interactions between nucleolin and 5'-tRFCys enhance nucleolin binding to a subset of transcripts, including several metabolic enzymes, and enhance their expression.

may cause gene dysregulation that can contribute to disease.

Lynne E. Maquat from the University of Rochester Medical Center described how loss of the fragile X retardation protein (FMRP) upregulates nonsense-mediated mRNA decay (NMD), potentially contributing to the developmental and cognitive defects observed in individuals with fragile X syndrome. Approximately 5–10% of mRNAs are degraded by NMD, which allows the cell to regulate mRNA concentration and adapt to environmental changes. For example, Maquat's group has shown that NMD competes with a separate mRNA decay pathway, Staufen-mediated mRNA decay (SMD), due to the fact that both pathways require the protein UPF1.¹⁷³ This competition contributes to myogenesis, as decreased levels of the NMD factor UPF2 enable UPF1 to interact with the SMD factor Staufen and increase SMD efficiency. Down-regulation of SMD targets, such as *PAX3*, and upregulation of NMD targets, such as *myogenin*, ultimately promote myogenesis.¹⁴² More recently, Maquat's lab has shown that UPF1 interacts directly with FMRP. This interaction promotes and/or stabilizes FMRP binding to NMD targets and thus suppresses NMD activity. In the brain, many NMD targets encode proteins important for axon guid-

ance and synaptic transmission. Therefore, in the absence of FMRP, such as in fragile X syndrome and depending on the stage of neural differentiation, NMD is hyperactivated, and NMD targets are destabilized, which negatively affects neurogenesis. Maquat showed that fibroblasts and induced pluripotent stem cells (iPSCs) from subjects with fragile X syndrome manifest hyperactivated NMD. The differentiation of these iPSCs to neurons manifests deficits in neurogenesis that are partially normalized upon administration of an NMD inhibitor.¹⁴³

Xuhang Liu from Sohail Tavazoie's lab at the Rockefeller University showed how a tRNA fragment can impact metastasis. Small RNA-seq analysis of mouse breast cancer cells with differing metastatic capacity identified 5'-tRFCys as one of the most upregulated tRFs in highly metastatic cells. High expression of 5'-tRFCys also correlates with poor clinical outcomes in breast cancer patients and is increased in breast tumors compared with normal breast tissues. Liu showed that interactions between nucleolin and 5'-tRFCys enhance nucleolin binding to a subset of transcripts, including several metabolic enzymes, and enhance their expression (Fig. 8). Liu put forth a model in which breast cancer cells contain a subset of

prometastatic transcripts that bind to nucleolin and form nucleolin monomers. In the absence of 5'-tRF^{Cys}, these monomers are not able to efficiently oligomerize, the complexes disassemble, and cells have limited metastatic potential. However, upregulation of 5'-tRF^{Cys} promotes efficient oligomerization of nucleolin-based complexes, which stabilizes prometastatic transcripts and increases the metastatic potential of cells.¹⁷⁴

Acknowledgments

R.K.P. was supported, by a Predoctoral Fellowship from the American Heart Association (18PRE33990261), a Wagner Fellowship from the University of Virginia, and F99/K00 NCI Predoctoral to Postdoctoral Fellow Transition Award (F99CA253732).

J.E.W. was supported by NIH Grant R35-GM119735.

X.L. was supported by a Bristol Myers Squibb Postdoctoral Fellowship.

S.F.T. was supported, in part, by a Faculty Scholar grant from the Howard Hughes Medical Institute, and by the DOD Collaborative Scholars and Innovators Award (W81XWH-12-1-0301), the Pershing Square Sohn Cancer Research Alliance Award, the Breast Cancer Research Foundation Award, the Emerald Foundation, NIH Grant 5R01CA215491-05, and the Black Family Metastasis Center.

Work in K.V.P. the laboratory is supported by grants from NIH (R21AG065748 and R01GM132458), Cancer Center at Illinois seed grant, the Prairie Dragon Paddlers, and the National Science Foundation (EAGER; MCB1723008).

Competing interests

J.T.M. is a scientific advisor for Ribometrix and Circ Bio. S.T. is a cofounder and shareholder of Rgenix and a member of its scientific advisory board.

H.Y.C. is a co-founder of Accent Therapeutics, Boundless Bio, and Circ Bio, and an advisor to 10x Genomics, Arsenal Biosciences, and Spring Discovery.

References

- Chowdhary, A., V. Satagopam & R. Schneider. 2021. Long non-coding RNAs: mechanisms, experimental, and computational approaches in identification, characterization, and their biomarker potential in cancer. *Front. Genet.* **12**: 649619.
- Mattick, J.S. & I.V. Makunin. 2006. Non-coding RNA. *Hum. Mol. Genet.* **15**(Spec No 1): R17–R29.
- Dossin, F. & E. Heard. 2021. The molecular and nuclear dynamics of X-chromosome inactivation. *Cold Spring Harb. Perspect. Biol.* a040196.
- Engreitz, J.M., A. Pandya-Jones, P. McDonel, *et al.* 2013. The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. *Science* **341**: 1237973.
- Giorgetti, L., B.R. Lajoie, A.C. Carter, *et al.* 2016. Structural organization of the inactive X chromosome in the mouse. *Nature* **535**: 575–579.
- Chaumeil, J., P. Le Baccon, A. Wutz, *et al.* 2006. A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced. *Genes Dev.* **20**: 2223–2237.
- Chow, J.C., C. Ciaudo, M.J. Fazzari, *et al.* 2010. LINE-1 activity in facultative heterochromatin formation during X chromosome inactivation. *Cell* **141**: 956–969.
- Collombet, S., I. Rall, C. Dugast-Darzacq, *et al.* 2021. RNA polymerase II depletion from the inactive X chromosome territory is not mediated by physical compartmentalization. *bioRxiv* 2021.03.26.437188.
- Żylicz, J.J., A. Bousard, K. Zumer, *et al.* 2019. The implication of early chromatin changes in X chromosome inactivation. *Cell* **176**: 182–197.e23.
- Chu, C., Q.C. Zhang, S.T. da Rocha, *et al.* 2015. Systematic discovery of Xist RNA binding proteins. *Cell* **161**: 404–416.
- McHugh, C.A., C.-K. Chen, A. Chow, *et al.* 2015. The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature* **521**: 232–236.
- Chen, C.-K., M. Blanco, C. Jackson, *et al.* 2016. Xist recruits the X chromosome to the nuclear lamina to enable chromosome-wide silencing. *Science* **354**: 468–472.
- Minajigi, A., J. Froberg, C. Wei, *et al.* 2015. Chromosomes. A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. *Science* **349**: 10.1126/science.aab2276 aab2276.
- Moindrot, B., A. Cerase, H. Coker, *et al.* 2015. A pooled shRNA screen identifies Rbm15, Spen, and Wtap as factors required for Xist RNA-mediated silencing. *Cell Rep.* **12**: 562–572.
- Monfort, A., G. Di Minin, A. Postlmayr, *et al.* 2015. Identification of Spen as a crucial factor for Xist function through forward genetic screening in haploid embryonic stem cells. *Cell Rep.* **12**: 554–561.
- Dossin, F., I. Pinheiro, J.J. Żylicz, *et al.* 2020. SPEN integrates transcriptional and epigenetic control of X-inactivation. *Nature* **578**: 455–460.
- Yamazaki, T., S. Souquere, T. Chujo, *et al.* 2018. Functional domains of NEAT1 architectural lncRNA induce paraspeckle assembly through phase separation. *Mol. Cell* **70**: 1038–1053.e7.
- Yamazaki, T., S. Nakagawa & T. Hirose. 2019. Architectural RNAs for membraneless nuclear body formation. *Cold Spring Harb. Symp. Quant. Biol.* **84**: 227–237.

19. Shin, Y. & C.P. Brangwynne. 2017. Liquid phase condensation in cell physiology and disease. *Science* **357**: eaaf4382.
20. Biamonti, G. 2004. Nuclear stress bodies: a heterochromatin affair? *Nat. Rev. Mol. Cell Biol.* **5**: 493–498.
21. Ninomiya, K., S. Adachi, T. Natsume, *et al.* 2020. LncRNA-dependent nuclear stress bodies promote intron retention through SR protein phosphorylation. *EMBO J.* **39**: e102729.
22. Ninomiya, K., J. Iwakiri, M.K. Aly, *et al.* 2021. m⁶ A modification of HSATIII lncRNAs regulates temperature-dependent splicing. *EMBO J.* **40**: e107976.
23. Xing, Y.-H., R.-W. Yao, Y. Zhang, *et al.* 2017. SLERT regulates DDX21 rings associated with Pol I transcription. *Cell* **169**: 664–678.e16.
24. Yao, R.-W., G. Xu, Y. Wang, *et al.* 2019. Nascent pre-rRNA sorting via phase separation drives the assembly of dense fibrillar components in the human nucleolus. *Mol. Cell* **76**: 767–783.e11.
25. Wu, M., G. Xu, C. Han, *et al.* 2021. lncRNA SLERT controls phase separation of FC/DFCs to facilitate Pol I transcription. *Science* **373**: 547–555.
26. Wang, X., X. Hu, W. Song, *et al.* 2021. Mutual dependency between lncRNA LETN and protein NPM1 in controlling the nucleolar structure and functions sustaining cell proliferation. *Cell Res.* **31**: 664–683.
27. Caudron-Herger, M., T. Pankert, J. Seiler, *et al.* 2015. Alu element-containing RNAs maintain nucleolar structure and function. *EMBO J.* **34**: 2758–2774.
28. Li, D., J. Zhang, M. Wang, *et al.* 2018. Activity dependent LoNA regulates translation by coordinating rRNA transcription and methylation. *Nat. Commun.* **9**: 1726.
29. Yap, K., S. Mukhina, G. Zhang, *et al.* 2018. A short tandem repeat-enriched RNA assembles a nuclear compartment to control alternative splicing and promote cell survival. *Mol. Cell* **72**: 525–540.e13.
30. Audas, T.E., M.D. Jacob & S. Lee. 2012. Immobilization of proteins in the nucleolus by ribosomal intergenic spacer noncoding RNA. *Mol. Cell* **45**: 147–157.
31. Bierhoff, H., K. Schmitz, F. Maass, *et al.* 2010. Noncoding transcripts in sense and antisense orientation regulate the epigenetic state of ribosomal RNA genes. *Cold Spring Harb. Symp. Quant. Biol.* **75**: 357–364.
32. Hao, Q. & K.V. Prasanth. 2021. Regulatory roles of nucleolus organizer region-derived long non-coding RNAs. *Mamm. Genome*. <https://doi.org/10.1007/s00335-021-09906-z>.
33. Quinodoz, S.A., N. Ollikainen, B. Tabak, *et al.* 2018. Higher-order inter-chromosomal hubs shape 3D genome organization in the nucleus. *Cell* **174**: 744–757.e24.
34. Quinodoz, S.A., P. Bhat, N. Ollikainen, *et al.* 2020. RNA promotes the formation of spatial compartments in the nucleus. *bioRxiv* 2020.08.25.267435.
35. Balas, M.M. & A.M. Johnson. 2018. Exploring the mechanisms behind long noncoding RNAs and cancer. *Non-Coding RNA Res.* **3**: 108–117.
36. Haemmig, S. & M.W. Feinberg. 2017. Targeting lncRNAs in cardiovascular disease: options and expeditions. *Circ. Res.* **120**: 620–623.
37. Diermeier, S.D., K.-C. Chang, S.M. Freier, *et al.* 2016. Mammary tumor-associated RNAs impact tumor cell proliferation, invasion, and migration. *Cell Rep.* **17**: 261–274.
38. Chang, K.-C., S.D. Diermeier, A.T. Yu, *et al.* 2020. MaTAR25 lncRNA regulates the Tensin1 gene to impact breast cancer progression. *Nat. Commun.* **11**: 6438.
39. Chen, H. & S.H. Lo. 2003. Regulation of tensin-promoted cell migration by its focal adhesion binding and Src homology domain 2. *Biochem. J.* **370**: 1039–1045.
40. Li, J., L. Han, P. Roebuck, *et al.* 2015. TANRIC: an interactive open platform to explore the function of lncRNAs in cancer. *Cancer Res.* **75**: 3728–3737.
41. Diermeier, S.D. & D.L. Spector. 2017. Antisense oligonucleotide-mediated knockdown in mammary tumor organoids. *Bio Protoc.* **7**: e2511.
42. Yan, X., Z. Hu, Y. Feng, *et al.* 2015. Comprehensive genomic characterization of long non-coding RNAs across human cancers. *Cancer Cell* **28**: 529–540.
43. Gupta, R.A., N. Shah, K.C. Wang, *et al.* 2010. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* **464**: 1071–1076.
44. Portoso, M., R. Ragazzini, Ž. Brenčič, *et al.* 2017. PRC2 is dispensable for HOTAIR-mediated transcriptional repression. *EMBO J.* **36**: 981–994.
45. Porman, A.M., J.T. Roberts, E.D. Duncan, *et al.* 2021. A single N6-methyladenosine site in lncRNA HOTAIR regulates its function in breast cancer cells. *bioRxiv* 2020.06.08.140954.
46. Gao, H., G. Chakraborty, A.P. Lee-Lim, *et al.* 2014. Forward genetic screens in mice uncover mediators and suppressors of metastatic reactivation. *Proc. Natl. Acad. Sci. USA* **111**: 16532–16537.
47. Haemmig, S., V. Simion & M.W. Feinberg. 2018. Long non-coding RNAs in vascular inflammation. *Front. Cardiovasc. Med.* **5**: 22.
48. Haemmig, S., D. Yang, X. Sun, *et al.* 2020. Long noncoding RNA SNHG12 integrates a DNA-PK-mediated DNA damage response and vascular senescence. *Sci. Transl. Med.* **12**: eaaw1868.
49. Childs, B.G., D.J. Baker, T. Wijshake, *et al.* 2016. Senescent intimal foam cells are deleterious at all stages of atherosclerosis. *Science* **354**: 472–477.
50. Wang, J.C. & M. Bennett. 2012. Aging and atherosclerosis: mechanisms, functional consequences, and potential therapeutics for cellular senescence. *Circ. Res.* **111**: 245–259.
51. Simion, V., H. Zhou, S. Haemmig, *et al.* 2020. A macrophage-specific lncRNA regulates apoptosis and atherosclerosis by tethering HuR in the nucleus. *Nat. Commun.* **11**: 6135.
52. Simion, V., H. Zhou, J.B. Pierce, *et al.* 2020. lncRNA VINAS regulates atherosclerosis by modulating NF-κB and MAPK signaling. *JCI Insight* **5**: 140627.
53. Zhou, H., V. Simion, J.B. Pierce, *et al.* 2021. lncRNA-MAP3K4 regulates vascular inflammation through the p38 MAPK signaling pathway and cis-modulation of MAP3K4. *FASEB J.* **35**: e21133.
54. Ni, H., S. Haemmig, Y. Deng, *et al.* 2021. A smooth muscle cell-enriched long noncoding RNA regulates cell plasticity and atherosclerosis by interacting with serum

- response factor. *Arterioscler. Thromb. Vasc. Biol.* **41**: 2399–2416.
55. Yin, Y., J.Y. Lu, X. Zhang, *et al.* 2020. U1 snRNP regulates chromatin retention of noncoding RNAs. *Nature* **580**: 147–150.
 56. Seila, A.C., J.M. Calabrese, S.S. Levine, *et al.* 2008. Divergent transcription from active promoters. *Science* **322**: 1849–1851.
 57. Henninger, J.E., O. Oksuz, K. Shrinivas, *et al.* 2021. RNA-mediated feedback control of transcriptional condensates. *Cell* **184**: 207–225.e24.
 58. Cisse, I.I., I. Izeddin, S.Z. Causse, *et al.* 2013. Real-time dynamics of RNA polymerase II clustering in live human cells. *Science* **341**: 664–667.
 59. Hnisz, D., K. Shrinivas, R.A. Young, *et al.* 2017. A phase separation model for transcriptional control. *Cell* **169**: 13–23.
 60. Shao, W., X. Bi, B. Gao, *et al.* 2021. Phase separation of RNA-binding protein promotes polymerase engagement and transcription. *bioRxiv* 2021.03.26.436939.
 61. Lee, S., F. Kopp, T.-C. Chang, *et al.* 2016. Noncoding RNA NORAD regulates genomic stability by sequestering PUMILIO proteins. *Cell* **164**: 69–80.
 62. Kopp, F., M.M. Elguindy, M.E. Yalvac, *et al.* 2019. PUMILIO hyperactivity drives premature aging of Norad-deficient mice. *eLife* **8**: e42650.
 63. Elguindy, M.M., F. Kopp, M. Goodarzi, *et al.* 2019. PUMILIO, but not RBMX, binding is required for regulation of genomic stability by noncoding RNA NORAD. *eLife* **8**: e48625.
 64. Tichon, A., N. Gil, Y. Lubelsky, *et al.* 2016. A conserved abundant cytoplasmic long noncoding RNA modulates repression by Pumilio proteins in human cells. *Nat. Commun.* **7**: 12209.
 65. Tichon, A., R.B.-T. Perry, L. Stojic, *et al.* 2018. SAM68 is required for regulation of Pumilio by the NORAD long noncoding RNA. *Genes Dev.* **32**: 70–78.
 66. Munschauer, M., C.T. Nguyen, K. Sirokman, *et al.* 2018. The NORAD lncRNA assembles a topoisomerase complex critical for genome stability. *Nature* **561**: 132–136.
 67. Unfried, J.P., G. Serrano, B. Suárez, *et al.* 2019. Identification of coding and long noncoding RNAs differentially expressed in tumors and preferentially expressed in healthy tissues. *Cancer Res.* **79**: 5167–5180.
 68. Unfried, J.P., M. Marín-Baquero, Á. Rivera-Calzada, *et al.* 2021. Long noncoding RNA NIHCOLE promotes ligation efficiency of DNA double-strand breaks in hepatocellular carcinoma. *Cancer Res.* **81**: 4910–4925.
 69. Mendoza-Figueroa, M.S., D.C. Tatomer & J.E. Wilusz. 2020. The integrator complex in transcription and development. *Trends Biochem. Sci.* **45**: 923–934.
 70. Baillat, D., M.-A. Hakimi, A.M. Nääs, *et al.* 2005. Integrator, a multiprotein mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA polymerase II. *Cell* **123**: 265–276.
 71. Tatomer, D.C., N.D. Elrod, D. Liang, *et al.* 2019. The integrator complex cleaves nascent mRNAs to attenuate transcription. *Genes Dev.* **33**: 1525–1538.
 72. Elrod, N.D., T. Henriques, K.-L. Huang, *et al.* 2019. The integrator complex attenuates promoter-proximal transcription at protein-coding genes. *Mol. Cell* **76**: 738–752.e7.
 73. Mani, S.R. & C.E. Juliano. 2013. Untangling the web: the diverse functions of the PIWI/piRNA pathway. *Mol. Reprod. Dev.* **80**: 632–664.
 74. Williams, Z., P. Morozov, A. Mihailovic, *et al.* 2015. Discovery and characterization of piRNAs in the human fetal ovary. *Cell Rep.* **13**: 854–863.
 75. Özata, D.M., T. Yu, H. Mou, *et al.* 2020. Evolutionarily conserved pachytene piRNA loci are highly divergent among modern humans. *Nat. Ecol. Evol.* **4**: 156–168.
 76. Öst, A., A. Lempradl, E. Casas, *et al.* 2014. Paternal diet defines offspring chromatin state and intergenerational obesity. *Cell* **159**: 1352–1364.
 77. Carone, B.R., L. Fauquier, N. Habib, *et al.* 2010. Paternally induced transgenerational environmental reprogramming of metabolic gene expression in mammals. *Cell* **143**: 1084–1096.
 78. Gapp, K., A. Jawaid, P. Sarkies, *et al.* 2014. Implication of sperm RNAs in transgenerational inheritance of the effects of early trauma in mice. *Nat. Neurosci.* **17**: 667–669.
 79. Lempradl, A., U. Kugelberg, M. Ionomou, *et al.* 2021. Intergenerational metabolic priming by sperm piRNAs. *bioRxiv* 2021.03.29.436592.
 80. Grimson, A., K.K.-H. Farh, W.K. Johnston, *et al.* 2007. MicroRNA targeting specificity in mammals: determinants beyond seed pairing. *Mol. Cell* **27**: 91–105.
 81. Sheu-Gruttadauria, J., P. Pawlica, S.M. Klum, *et al.* 2019. Structural basis for target-directed microRNA degradation. *Mol. Cell* **75**: 1243–1255.e7.
 82. Sheu-Gruttadauria, J., Y. Xiao, L.F. Gebert, *et al.* 2019. Beyond the seed: structural basis for supplementary microRNA targeting by human Argonaute2. *EMBO J.* **38**: e101153.
 83. McGeary, S.E., N. Bisaria & D.P. Bartel. 2021. Pairing to the microRNA 3' region occurs through two alternative binding modes, with affinity shaped by nucleotide identity as well as pairing position. *bioRxiv* 2021.04.13.439700.
 84. Vicens, Q. & E. Westhof. 2014. Biogenesis of circular RNAs. *Cell* **159**: 13–14.
 85. Chen, L. & G. Shan. 2015. Circular RNAs remain peculiarly unclear in biogenesis and function. *Sci. China Life Sci.* **58**: 616–618.
 86. Chen, I., C.-Y. Chen & T.-J. Chuang. 2015. Biogenesis, identification, and function of exonic circular RNAs. *Wiley Interdiscip. Rev. RNA* **6**: 563–579.
 87. Hansen, T.B., T.I. Jensen, B.H. Clausen, *et al.* 2013. Natural RNA circles function as efficient microRNA sponges. *Nature* **495**: 384–388.
 88. Memczak, S., M. Jens, A. Elefsinioti, *et al.* 2013. Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* **495**: 333–338.
 89. Xu, H., S. Guo, W. Li, *et al.* 2015. The circular RNA Cdr1as, via miR-7 and its targets, regulates insulin transcription and secretion in islet cells. *Sci. Rep.* **5**: 12453.

90. Ashwal-Fluss, R., M. Meyer, N.R. Pamudurti, *et al.* 2014. circRNA biogenesis competes with pre-mRNA splicing. *Mol. Cell* **56**: 55–66.
91. Kulcheski, F.R., A.P. Christoff & R. Margis. 2016. Circular RNAs are miRNA sponges and can be used as a new class of biomarker. *J. Biotechnol.* **238**: 42–51.
92. Zheng, Q., C. Bao, W. Guo, *et al.* 2016. Circular RNA profiling reveals an abundant circHIPK3 that regulates cell growth by sponging multiple miRNAs. *Nat. Commun.* **7**: 11215.
93. Yang, Y., X. Gao, M. Zhang, *et al.* 2018. Novel role of FBXW7 circular RNA in repressing glioma tumorigenesis. *J. Natl. Cancer Inst.* **110**: 304–315.
94. Zhang, M., K. Zhao, X. Xu, *et al.* 2018. A peptide encoded by circular form of LINC-PINT suppresses oncogenic transcriptional elongation in glioblastoma. *Nat. Commun.* **9**: 4475.
95. Zhang, M., N. Huang, X. Yang, *et al.* 2018. A novel protein encoded by the circular form of the SHPRH gene suppresses glioma tumorigenesis. *Oncogene* **37**: 1805–1814.
96. Legnini, I., G. Di Timoteo, F. Rossi, *et al.* 2017. Circ-ZNF609 is a circular RNA that can be translated and functions in myogenesis. *Mol. Cell* **66**: 22–37.e9.
97. Zheng, X., L. Chen, Y. Zhou, *et al.* 2019. A novel protein encoded by a circular RNA circPPP1R12A promotes tumor pathogenesis and metastasis of colon cancer via Hippo-YAP signaling. *Mol. Cancer* **18**: 47.
98. Pamudurti, N.R., O. Bartok, M. Jens, *et al.* 2017. Translation of circRNAs. *Mol. Cell* **66**: 9–21.e7.
99. Liang, W.-C., C.-W. Wong, P.-P. Liang, *et al.* 2019. Translation of the circular RNA circ β -catenin promotes liver cancer cell growth through activation of the Wnt pathway. *Genome Biol.* **20**: 84.
100. Granados-Riveron, J.T. & G. Aquino-Jarquin. 2016. The complexity of the translation ability of circRNAs. *Biochim. Biophys. Acta* **1859**: 1245–1251.
101. Schneider, T., L.-H. Hung, S. Schreiner, *et al.* 2016. CircRNA–protein complexes: IMP3 protein component defines subfamily of circRNPs. *Sci. Rep.* **6**: 31313.
102. Li, L.-J., R.-X. Leng, Y.-G. Fan, *et al.* 2017. Translation of noncoding RNAs: focus on lncRNAs, pri-miRNAs, and circRNAs. *Exp. Cell Res.* **361**: 1–8.
103. Wang, Y. & Z. Wang. 2015. Efficient backsplicing produces translatable circular mRNAs. *RNA* **21**: 172–179.
104. Chen, X., P. Han, T. Zhou, *et al.* 2016. circRNADb: a comprehensive database for human circular RNAs with protein-coding annotations. *Sci. Rep.* **6**: 34985.
105. Jang, S.K., H.G. Kräusslich, M.J. Nicklin, *et al.* 1988. A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during *in vitro* translation. *J. Virol.* **62**: 2636–2643.
106. Pelletier, J. & N. Sonenberg. 1988. Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature* **334**: 320–325.
107. Toor, N., K.S. Keating, S.D. Taylor, *et al.* 2008. Crystal structure of a self-spliced group II intron. *Science* **320**: 77–82.
108. Rinn, J.L., M. Kertesz, J.K. Wang, *et al.* 2007. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* **129**: 1311–1323.
109. Somarowthu, S., M. Legiewicz, I. Chillón, *et al.* 2015. HOTAIR forms an intricate and modular secondary structure. *Mol. Cell* **58**: 353–361.
110. Rivas, E., J. Clements & S.R. Eddy. 2017. A statistical test for conserved RNA structure shows lack of evidence for structure in lncRNAs. *Nat. Methods* **14**: 45–48.
111. Tavares, R.C.A., A.M. Pyle & S. Somarowthu. 2019. Phylogenetic analysis with improved parameters reveals conservation in lncRNA structures. *J. Mol. Biol.* **431**: 1592–1603.
112. Morgan, B.S., B.G. Sanaba, A. Donlic, *et al.* 2019. R-BIND: an interactive database for exploring and developing RNA-targeted chemical probes. *ACS Chem. Biol.* **14**: 2691–2700.
113. Morgan, B.S., J.E. Forte, R.N. Culver, *et al.* 2017. Discovery of key physicochemical, structural, and spatial properties of RNA-targeted bioactive ligands. *Angew. Chem. Int. Ed. Engl.* **56**: 13498–13502.
114. Stelzer, A.C., A.T. Frank, J.D. Kratz, *et al.* 2011. Discovery of selective bioactive small molecules by targeting an RNA dynamic ensemble. *Nat. Chem. Biol.* **7**: 553–559.
115. Patwardhan, N.N., L.R. Ganser, G.J. Kapral, *et al.* 2017. Amiloride as a new RNA-binding scaffold with activity against HIV-1 TAR. *MedChemComm* **8**: 1022–1036.
116. Patwardhan, N.N., Z. Cai, A. Umuhire Juru, *et al.* 2019. Driving factors in amiloride recognition of HIV RNA targets. *Org. Biomol. Chem.* **17**: 9313–9320.
117. Davila-Calderon, J., N.N. Patwardhan, L.-Y. Chiu, *et al.* 2020. IRES-targeting small molecule inhibits enterovirus 71 replication via allosteric stabilization of a ternary complex. *Nat. Commun.* **11**: 4775.
118. Brown, J.A., D. Bulkley, J. Wang, *et al.* 2014. Structural insights into the stabilization of MALAT1 noncoding RNA by a bipartite triple helix. *Nat. Struct. Mol. Biol.* **21**: 633–640.
119. Donlic, A., B.S. Morgan, J.L. Xu, *et al.* 2018. Discovery of small molecule ligands for MALAT1 by tuning an RNA-binding scaffold. *Angew. Chem. Int. Ed. Engl.* **57**: 13242–13247.
120. Donlic, A., M. Zafferani, G. Padroni, *et al.* 2020. Regulation of MALAT1 triple helix stability and *in vitro* degradation by diphenylfurans. *Nucleic Acids Res.* **48**: 7653–7664.
121. Zimmer, J.T., N.A. Rosa-Mercado, D. Canzio, *et al.* 2021. STL-seq reveals pause-release and termination kinetics for promoter-proximal paused RNA polymerase II transcripts. *Mol. Cell*. <https://doi.org/10.1016/j.molcel.2021.08.019>.
122. Marcia, M. & A.M. Pyle. 2012. Visualizing group II intron catalysis through the stages of splicing. *Cell* **151**: 497–507.
123. Zhao, C., K.R. Rajashankar, M. Marcia, *et al.* 2015. Crystal structure of group II intron domain 1 reveals a template for RNA assembly. *Nat. Chem. Biol.* **11**: 967–972.

124. Manigrasso, J., I. Chillón, V. Genna, *et al.* 2020. Visualizing group II intron dynamics between the first and second steps of splicing. *Nat. Commun.* **11**: 2837.
125. Chunharojrith, P., Y. Nakayama, X. Jiang, *et al.* 2015. Tumor suppression by MEG3 lncRNA in a human pituitary tumor derived cell line. *Mol. Cell. Endocrinol.* **416**: 27–35.
126. Zhang, X., K. Rice, Y. Wang, *et al.* 2010. Maternally expressed gene 3 (MEG3) noncoding ribonucleic acid: isoform structure, expression, and functions. *Endocrinology* **151**: 939–947.
127. Zhu, J., S. Liu, F. Ye, *et al.* 2015. Long noncoding RNA MEG3 interacts with p53 protein and regulates partial p53 target genes in hepatoma cells. *PLoS One* **10**: e0139790.
128. Uroda, T., E. Anastasakou, A. Rossi, *et al.* 2019. Conserved pseudoknots in lncRNA MEG3 are essential for stimulation of the p53 pathway. *Mol. Cell* **75**: 982–995.e9.
129. Uroda, T., I. Chillón, P. Annibale, *et al.* 2020. Visualizing the functional 3D shape and topography of long noncoding RNAs by single-particle atomic force microscopy and in-solution hydrodynamic techniques. *Nat. Protoc.* **15**: 2107–2139.
130. Mueller, A.C., M.A. Cichewicz, B.K. Dey, *et al.* 2015. MUNC, a long noncoding RNA that facilitates the function of MyoD in skeletal myogenesis. *Mol. Cell. Biol.* **35**: 498–513.
131. Cichewicz, M.A., M. Kiran, R.K. Przanowska, *et al.* 2018. MUNC, an enhancer RNA upstream from the MYOD gene, induces a subgroup of myogenic transcripts in trans independently of MyoD. *Mol. Cell. Biol.* **38**: e00655–17.
132. Przanowska, R.K., C.A. Weidmann, S. Saha, *et al.* 2021. Distinct MUNC lncRNA structural domains regulate transcription of different promyogenic factors. *bioRxiv* 2021.06.22.449443.
133. Invernizzi, P. & M.E. Gershwin. 2009. The genetics of human autoimmune disease. *J. Autoimmun.* **33**: 290–299.
134. Peckham, H., N.M. de Grujter, C. Raine, *et al.* 2020. Male sex identified by global COVID-19 meta-analysis as a risk factor for death and ICU admission. *Nat. Commun.* **11**: 6317.
135. Yu, B., Y. Qi, R. Li, *et al.* 2021. B cell-specific XIST complex enforces X-inactivation and restrains atypical B cells. *Cell* **184**: 1790–1803.e17.
136. Lu, Z., G.S. Filonov, J.J. Noto, *et al.* 2015. Metazoan tRNA introns generate stable circular RNAs *in vivo*. *RNA* **21**: 1554–1565.
137. Litke, J.L. & S.R. Jaffrey. 2021. Trans ligation of RNAs to generate hybrid circular RNAs using highly efficient autocatalytic transcripts. *Methods*. <https://doi.org/10.1016/j.jymeth.2021.05.009>.
138. Lukasiewicz, A.J. & L.M. Contreras. 2020. Antisense probing of dynamic RNA structures. *Methods* **183**: 76–83.
139. Sowa, S.W., J. Vazquez-Anderson, C.A. Clark, *et al.* 2015. Exploiting post-transcriptional regulation to probe RNA structures *in vivo* via fluorescence. *Nucleic Acids Res.* **43**: e13.
140. Mihailovic, M.K., J. Vazquez-Anderson, Y. Li, *et al.* 2018. High-throughput *in vivo* mapping of RNA accessible interfaces to identify functional sRNA binding sites. *Nat. Commun.* **9**: 4084.
141. Isaacs, F.J., D.J. Dwyer, C. Ding, *et al.* 2004. Engineered riboregulators enable post-transcriptional control of gene expression. *Nat. Biotechnol.* **22**: 841–847.
142. Leistra, A.N., P. Amador, A. Buvanendiran, *et al.* 2017. Rational modular RNA engineering based on *in vivo* profiling of structural accessibility. *ACS Synth. Biol.* **6**: 2228–2240.
143. Mihailovic, M.K., A.M. Ekdahl, A. Chen, *et al.* 2021. Uncovering transcriptional regulators and targets of sRNAs using an integrative data-mining approach: HNS-Regulated RseX as a case study. *Front. Cell. Infect. Microbiol.* **11**: 696533.
144. Shi, J., Y. Zhang, D. Tan, *et al.* 2021. PANDORA-seq expands the repertoire of regulatory small RNAs by overcoming RNA modifications. *Nat. Cell Biol.* **23**: 424–436.
145. Shi, J., E.-A. Ko, K.M. Sanders, *et al.* 2018. SPORTS1.0: a tool for annotating and profiling non-coding RNAs optimized for rRNA- and tRNA-derived small RNAs. *Genomics Proteomics Bioinformatics* **16**: 144–151.
146. Di Bella, S., A. La Ferlita, G. Carapezza, *et al.* 2020. A benchmarking of pipelines for detecting ncRNAs from RNA-Seq data. *Brief. Bioinform.* **21**: 1987–1998.
147. Gu, W., J. Shi, H. Liu, *et al.* 2020. Peripheral blood non-canonical small non-coding RNAs as novel biomarkers in lung cancer. *Mol. Cancer* **19**: 159.
148. Peaston, A.E., A.V. Evsikov, J.H. Graber, *et al.* 2004. Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. *Dev. Cell* **7**: 597–606.
149. Kigami, D., N. Minami, H. Takayama, *et al.* 2003. MuERV-L is one of the earliest transcribed genes in mouse one-cell embryos. *Biol. Reprod.* **68**: 651–654.
150. Macfarlan, T.S., W.D. Gifford, S. Driscoll, *et al.* 2012. Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature* **487**: 57–63.
151. Macfarlan, T.S., W.D. Gifford, S. Agarwal, *et al.* 2011. Endogenous retroviruses and neighboring genes are coordinately repressed by LSD1/KDM1A. *Genes Dev.* **25**: 594–607.
152. Ishiuchi, T., R. Enriquez-Gasca, E. Mizutani, *et al.* 2015. Early embryonic-like cells are induced by downregulating replication-dependent chromatin assembly. *Nat. Struct. Mol. Biol.* **22**: 662–671.
153. Choi, Y.J., C.-P. Lin, D. Risso, *et al.* 2017. Deficiency of microRNA miR-34a expands cell fate potential in pluripotent stem cells. *Science* **355**: eaag1927.
154. Modzelewski, A., W. Shao, J. Chen, *et al.* 2021. A species-specific retrotransposon drives a conserved Cdk2ap1 isoform essential for preimplantation development. *Cell* **184**: 5541–5558.e22. <https://doi.org/10.1016/j.cell.2021.09.021>.
155. Gentles, A.J., A.M. Newman, C.L. Liu, *et al.* 2015. The prognostic landscape of genes and infiltrating immune cells across human cancers. *Nat. Med.* **21**: 938–945.

156. Sica, A., P. Larghi, A. Mancino, *et al.* 2008. Macrophage polarization in tumour progression. *Semin. Cancer Biol.* **18**: 349–355.
157. Liu, J., L. Lao, J. Chen, *et al.* 2021. The IRENA lncRNA converts chemotherapy-polarized tumor-suppressing macrophages to tumor-promoting phenotypes in breast cancer. *Nat. Cancer* **2**: 457–473.
158. Dumbović, G., U. Braunschweig, H.K. Langner, *et al.* 2021. Nuclear compartmentalization of TERT mRNA and TUG1 lncRNA is driven by intron retention. *Nat. Commun.* **12**: 3308.
159. Xin, M., E.N. Olson & R. Bassel-Duby. 2013. Mending broken hearts: cardiac development as a basis for adult heart regeneration and repair. *Nat. Rev. Mol. Cell Biol.* **14**: 529–541.
160. Chen, J., Q. Hu, B.-F. Zhang, *et al.* 2019. Long noncoding RNA UCA1 inhibits ischaemia/reperfusion injury induced cardiomyocytes apoptosis via suppression of endoplasmic reticulum stress. *Genes Genomics* **41**: 803–810.
161. Ponnusamy, M., F. Liu, Y.-H. Zhang, *et al.* 2019. Long noncoding RNA CPR (cardiomyocyte proliferation regulator) regulates cardiomyocyte proliferation and cardiac repair. *Circulation* **139**: 2668–2684.
162. Vacante, F., J. Rodor, M.K. Lalwani, *et al.* 2021. CARMN loss regulates smooth muscle cells and accelerates atherosclerosis in mice. *Circ. Res.* **128**: 1258–1275.
163. Ounzain, S., R. Micheletti, C. Arnan, *et al.* 2015. CARMEN, a human super enhancer-associated long noncoding RNA controlling cardiac specification, differentiation and homeostasis. *J. Mol. Cell. Cardiol.* **89**: 98–112.
164. Plaisance, I., S. Perruchoud, M. Fernandez-Tenorio, *et al.* 2016. Cardiomyocyte lineage specification in adult human cardiac precursor cells via modulation of enhancer-associated long noncoding RNA expression. *JACC Basic Transl. Sci.* **1**: 472–493.
165. Chen, Y., D.L. Toffaletti, J.L. Tenor, *et al.* 2014. The *Cryptococcus neoformans* transcriptome at the site of human meningitis. *mBio* **5**: e01087–01013.
166. Misiak, B., L. Ricceri & M.M. Sasiadek. 2019. Transposable elements and their epigenetic regulation in mental disorders: current evidence in the field. *Front. Genet.* <https://doi.org/10.3389/fgene.2019.00580>.
167. Chuong, E.B., N.C. Elde & C. Feschotte. 2016. Regulatory evolution of innate immunity through co-option of endogenous retroviruses. *Science* **351**: 1083–1087.
168. Deniz, Ö., M. Ahmed, C.D. Todd, *et al.* 2020. Endogenous retroviruses are a source of enhancers with oncogenic potential in acute myeloid leukaemia. *Nat. Commun.* **11**: 3506.
169. Jang, H.S., N.M. Shah, A.Y. Du, *et al.* 2019. Transposable elements drive widespread expression of oncogenes in human cancers. *Nat. Genet.* **51**: 611–617.
170. Edginton-White, B., P. Cauchy, S.A. Assi, *et al.* 2019. Global long terminal repeat activation participates in establishing the unique gene expression programme of classical Hodgkin lymphoma. *Leukemia* **33**: 1463–1474.
171. Babaian, A. & D.L. Mager. 2016. Endogenous retroviral promoter exaptation in human cancer. *Mob. DNA* **7**: 24.
172. Lamprecht, B., K. Walter, S. Kreher, *et al.* 2010. Derepression of an endogenous long terminal repeat activates the CSF1R proto-oncogene in human lymphoma. *Nat. Med.* **16**: 571–579, 1p following 579.
173. Kim, Y.K., L. Furic, L. Desgroseillers, *et al.* 2005. Mammalian Staufen1 recruits Upf1 to specific mRNA 3'UTRs so as to elicit mRNA decay. *Cell* **120**: 195–208.
174. Liu, X., H. Alwaseem, H. Molina, *et al.* 2021. A prometastatic tRNA fragment drives nucleolin oligomerization and stabilization of bound metabolic mRNAs. *bioRxiv* 2021.04.26.441477.