RNA Society

OF NORTH CAROLINA

Symposium on RNA Biology XV: RNA Tool and Target

NOVEMBER 14-15, 2024
DUKE UNIVERSITY

Symposium on RNA Biology XV: RNA Tool and Target

NOVEMBER 14-15, 2024
DUKE UNIVERSITY

ORGANIZING COMMITTEE

Stacy Horner
Duke, Integrative Immunobiology

Kate Meyer
Duke, Biochemistry

Bruce Sullenger
Duke, Surgery

Marcos Morgan
NIEHS, Reproductive & Developmental Biology

Traci Hall
NIEHS, Epigenetics & Stem Cell Biology

Dan Dominguez
UNC, Pharmacology

Jimena Guidice
UNC, Cell Biology and Physiology

Qi Zhang
UNC, Biochemistry and Biophysics

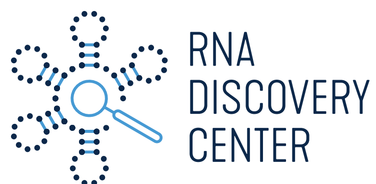
Mauro Calabrese
UNC, Pharmacology

CONTENTS

Sponsors 3
Meeting Agenda 4
Invited Speaker Abstracts 7
Selected Speaker Abstracts 10
Poster Abstracts 18

THANK YOU TO OUR SPONSORS

GOLD



SILVER

BRONZE

MEETING AGENDA

Thursday, November 14th

12:00 pm
Registration and arrival

12:25 pm
Welcome and opening remarks: Stacy Horner and Kate Meyer, Duke University

Session I: RNA Regulation and Structure
Session Chair: Qi Zhang, UNC-Chapel Hill

12:30 pm
Keynote Lecture: Mitch Guttman, Cal Tech
“RNA promotes the formation of spatial compartments throughout the nucleus”

1:10 pm
Joseph Yesselman, University of Nebraska
“High-throughput determination of RNA tertiary contact thermodynamics by quantitative DMS chemical mapping”

1:40 pm
Claire Fleurisson, UNC-Chapel Hill (15 minutes + 5 min Q&A)
“Contributions of RNA Structure Near Splice Sites to the Splicing Regulatory Code”

2:00 pm
Zoe Wright, NIEHS (15 minutes + 5 min Q&A)
“Structural and Biochemical Analyses Reveal that Spontaneous Base Flipping Helps Drive SARS-CoV-2 Protein Nsp15’s Preferences in dsRNA Substrates”

2:20 pm – 2:45 pm
Break

Session II: RNA:Protein Interactions and RNA Modifications
Session Chair: Traci Hall, NIEHS

2:45 pm
Siggy Nachtergaele, Yale University
“Exploring the roles of RNA modifications in drug-resistant glioblastoma”

3:15 pm
Kotaro Nakanishi, Ohio State University
“Gene silencing by cityRNAs”

3:45 pm
Kayla Mason, UNC-Chapel Hill (15 minutes + 5 min Q&A)
"The role of DDX3X in the initiation of cardiac sex-differential protein expression"

4:05 pm
Matt Tegowski, Duke University (15 minutes + 5 min Q&A)
"Single-cell m6A profiling in the mouse cortex reveals cell type-specific methylation profiles as well as changes during aging and stress"

The poster session and dinner will be on the 6th floor.

4:30 pm – 6:00 pm
Poster Session, Hors d'oeuvres and bar service.

6:00 pm
Dinner

Friday, November 15th

10:00 am
Breakfast and networking

Session III: RNA Targeting and Therapeutics
Session Chair: Bruce Sullenger, Duke

10:30 am
Amanda Hargrove, University of Toronto
"Strategies to modulate the conformation and function of RNA with small molecules"

11:00 am
Wilton Snead, Duke University (15 minutes + 5 min Q&A)
"A multicomponent, out-of-equilibrium logic controls the assembly of nuclear paraspeckles"

11:20 am
Edgar Faison, UNC-Chapel Hill (15 minutes + 5 min Q&A)
"Chemical Visualization of Nascent RNA Folding with Co-Transcriptional TMO-MaP"

11:40 am
Diana Dou, Duke University (15 minutes + 5 min Q&A)
"A Complex issue: When 2 X's need to Xist"

12:00 pm – 1:00 pm
Lunch

Session IV: RNA Processing, Localization, and Translation

Session Chair: Mauro Calabrese, UNC-Chapel Hill

1:00 pm
Keynote Lecture: Geraldine Seydoux, Johns Hopkins University
"A structural role for mRNAs in germ granules?"

1:40 pm
Wayne Hemphill, University of Colorado Boulder (15 minutes + 5 min Q&A)
"A General Mechanism for Tunable RNA-Mediated Regulation of Chromatin-Binding Proteins"

2:00 pm
Alicia Darnell, Duke University (15 minutes + 5 min Q&A)
"Metabolic control of protein synthesis by ribosome fall-off"

2:20 pm – 2:40 pm
Break

2:40 pm
Eric Wang, University of Florida
"RNA supply chain issues in repeat expansion diseases"

3:10 pm
Ru-pin Chi, NIEHS (15 minutes + 5 min Q&A)
"A death-promoting post-transcriptional mechanism in heat exposed cells"

3:30 pm
Taylor Niehoff, Duke University (15 minutes + 5 min Q&A)
"An unbiased screen identifies the PAXT connection as a modifier of rare codon-dependent expression"

3:50 pm
Lela Lackey, Clemson University (15 minutes + 5 min Q&A)
"Inflammation and alternative polyadenylation on alternative polyadenylation"

4:10 pm
Awards Presentation

4:30 pm – 6:00 pm
Reception, Hors d'oeuvres and bar service

6:00 pm
Meeting Adjourns

INVITED SPEAKER ABSTRACTS

Session I: RNA Regulation and Structure

KEYNOTE LECTURE 1
RNA promotes the formation of spatial compartments throughout the nucleus

Mitch Guttman, PhD
Professor, Division of Biology and Biological Engineering, Caltech;
Associate Director, UCLA-Caltech Medical Scientist Training Program (MD-PhD program);
Robertson Investigator, New York Stem Cell Foundation

The nucleus is a highly organized arrangement of RNA, DNA, and protein molecules that are compartmentalized within three-dimensional (3D) structures involved in shared functional and regulatory processes. Yet, what drives the molecular and spatial organization of these 3D compartments and what functional role(s) they play remains largely unknown. Here, we will discuss a widespread role for non-coding RNAs in seeding compartments throughout the nucleus and describe a unique mechanism for how ncRNAs drive concentration-dependent transitions from diffuse localization to compartmentalized structures. We will describe how these ncRNA-mediated compartments control major classes of nuclear functions, including RNA processing, heterochromatin assembly, and gene regulation. Finally, we will discuss several emerging mechanistic paradigms for how compartments ensure quantitative regulation in the nucleus.

High-throughput determination of RNA tertiary contact thermodynamics by quantitative DMS chemical mapping High-throughput determination of RNA tertiary contact thermodynamics by quantitative DMS chemical mapping

Joseph Yesselman, PhD
Assistant Professor, Department of Chemistry, University of Nebraska-Lincoln

Structured RNAs often contain long-range tertiary contacts that are critical to their function. Despite the importance of tertiary contacts, methods to measure their thermodynamics are low throughput or require specialized instruments. Here, we introduce a new quantitative chemical mapping method (qMaPseq) to measure Mg^{2+}-induced formation of tertiary contact thermodynamics in a high-throughput manner using standard biochemistry equipment. With qMaPseq, we measured the ΔG of 98 unique tetraloop/tetraloop receptor (TL/TLR) variants in a one-pot reaction. These results agree well with measurements from specialized instruments (R^2=0.64). Furthermore, the DMS reactivity of the TL directly correlates to the stability of the contact (R^2=0.68), the first direct evidence that a single DMS reactivity measurement reports on thermodynamics. Combined with structure prediction, DMS reactivity allowed the development of experimentally accurate 3D models of TLR mutants. These results demonstrate that qMaPseq is broadly accessible, high-throughput, and directly links DMS reactivity to thermodynamics.

Session II: RNA: Protein Interactions and RNA Modifications

Exploring the roles of RNA modifications in drug-resistant glioblastoma

Sigrid Nachtergaele, PhD
Assistant Professor of Molecular, Cellular & Developmental Biology, Yale University

RNA modifications represent a critical mechanism of gene expression regulation, controlling RNA processing, stability, and function. Recent years have yielded many insights into the molecular mechanism by which N6-methyladenosine (m6A) regulates mRNA decay, and the roles for m6A-mediated regulation of gene expression are increasingly being described in a wide variety of cancers. However despite this rapid progress, multiple studies focused on a single modification in a single cancer type still often yield conflicting conclusions as to whether the modification is functionally oncogenic or tumor suppressive. How RNA modifications may influence gene expression changes during the development of chemotherapeutic drug resistance also remains underexplored. Using temozolomide, the standard of care chemotherapeutic for glioblastoma, we have developed paired sensitive and drug-resistance cultured glioblastoma cell lines that we are using to dissect the functions of m6A and other modifications in this devastating disease. By combining our drug-resistant cell lines with novel RNA modification analysis tools, we have found that the MGMT transcript is not only massively upregulated in drug-resistant cells, but also stabilized in an RNA modification-dependent manner. This suggests that manipulating chemical modifications on MGMT and other key transcripts could re-sensitize glioblastoma cells to chemotherapy, and opens the door to understanding how chemical modifications may impact chemotherapeutic resistance in other cancers.

Gene silencing by cityRNAs

Kotaro Nakanishi, PhD
Professor, Department of Chemistry and Biochemistry, The Ohio State University

TinyRNAs (tyRNAs) are ≤17-nt guide RNAs associated with Argonaute proteins (AGOs). Certain 14-nt tyRNAs catalytically activate human Argonaute3 (AGO3) and are therefore called cleavage-inducing tyRNAs (cityRNAs). Notably, cityRNA-loaded Argonaute2 (AGO2) and AGO3 enhance their endonuclease activity when the immediate 5′ upstream region of the tyRNA-target site (UTy) contains sequences with low affinities for AGO. We propose a model in which cityRNA-loaded AGO2 and AGO3 efficiently cleave fully complementary tyRNA-target sites unless they directly recognize the UTy. Additionally, we developed systems to load endogenous AGOs with specific tyRNAs and demonstrated that, unlike microRNAs, cityRNA-mediated silencing heavily relies on target cleavage. Our study reveals that AGOs exploit cityRNAs for target recognition differently from microRNAs, thereby altering genes.

Session III: RNA Targeting and Therapeutics

Strategies to modulate the conformation and function of RNA with small molecules

Amanda E. Hargrove, PhD
Professor in Medicinal Chemistry, University of Toronto, Mississauga

Small molecules offer a unique opportunity to target structural and regulatory elements in therapeutically relevant RNAs, but understanding functional selectivity has been a recurrent challenge in small molecule:RNA recognition. RNAs offer less differentiating chemical functionality than proteins and sample multiple conformations that can each impact function. We have used organic synthesis, machine learning and a variety of biophysical and cell-based assays to reveal patterns in the chemical and structural properties of bioactive RNA ligands as well as RNA topological space privileged for differentiation. We have applied these principles to several disease-relevant systems. We have tuned diiminazene-based small molecules to functionally modulate different RNA tertiary structures in the oncogenic long noncoding RNA MALAT1, leading to either monofunctional degraders or tailored manipulation of RNA:protein interactions, respectively. We have also developed RNA-targeted antivirals for enterovirus (EV71) and SARS-CoV-2, revealing a novel allosteric mechanism of small molecule: RNA targeting.

Session IV: RNA Processing, Localization, and Translation

KEYNOTE LECTURE 2
A structural role for mRNAs in germ granules?

Geraldine Seydoux, PhD
Professor of Molecular Biology and Genetics, Johns Hopkins
Huntington Sheldon Professor in Medical Discovery, Johns Hopkins
Vice Dean for Basic Research, Johns Hopkins

Germ granules are RNA-rich condensates that segregate to the nascent germline in early embryos. We have identified over 400 transcripts that enrich in germ granules in *C. elegans* embryos. Most are low translation mRNAs that do not code for germ cell fate determinants and do not require granule localization for regulation. I will discuss the possibility that most mRNAs localize to germ granules, not for regulation, but to stabilize protein-rich condensates at the core of germ granules.

RNA supply chain issues in repeat expansion diseases

Eric Wang, PhD
Associate Professor, Molecular Genetics & Microbiology, University of Florida

SELECTED SPEAKER ABSTRACTS

Session I: RNA Regulation and Structure

Contributions of RNA Structure Near Splice Sites to the Splicing Regulatory Code

Anastacia Wienecke, **Claire Fleurisson**, Alain Laederach
UNC Chapel Hill

Mounting observations highlight the multi-faceted role of RNA structure in the regulation of RNA splicing. For instance, structure can scaffold splice signals, mask splice sites, and preclude the binding of regulatory proteins. My goal is to explore the general rules by which RNA structure contributes to splicing regulation. To this end, I focus specifically on RNA structure in spliceosome footprint regions, which from my analysis of recently-published cryo-EM data, are where the spliceosome assembles on the pre-mRNA. Structure must unpair within each footprint region for the spliceosome to assemble and cut the RNA. As such, I predict that less structure leads to greater footprint accessibility by the spliceosome, increasing splicing at that site. Our lab chemically probed the structures of RNA transcribed from three bacterial artificial chromosome (BAC) clones. The data is highly replicable and yields structure information for spliceosome footprints around 85 5' splice sites and 72 branchpoints. Twenty-four of these 157 footprints are associated with cryptic splice sites that we discovered via in-depth analysis of 18 RNA-sequencing datasets. Based on sequence data, dimensionality reduction techniques show a clear and expected separation between footprints around 5' splice sites and around branchpoints. Notably, footprints also form two distinct subclusters based on GC-content. Higher GC-content footprints are shown to have higher basepairing probabilities, reflecting their propensity for greater structure. Using dimensionality reduction techniques on structure-probing data, we investigate and compare structure between constitutive, alternative, and cryptic footprints at 5' splice sites and around branchpoints.

Structural and Biochemical Analyses Reveal that Spontaneous Base Flipping Helps Drive SARS-CoV-2 Protein Nsp15's Preferences in dsRNA Substrates

Zoe M. Wright, Kevin John Butay, Juno M. Krahn, Isha M. Wilson, Scott A. Gabel, Eugene F. DeRose, Israa S. Hissein, Jason G. Williams, Mario J. Borgnia, Meredith N. Frazier, Geoffrey A. Mueller, and Robin E. Stanley
NIEHS/NIH

Coronaviruses evade the host immune system with the help of the endoribonuclease Nsp15, which cleaves viral dsRNA (3' of uridine, "U") into pieces that are too short to be detected by the host. Nsp15's activity is known to have a direct effect on the severity of coronaviral infection, but Nsp15's physiological targets within viral RNA are not fully understood. While prior structural data shows that to cleave dsRNA, Nsp15's target U must be flipped out of the helix, it is not yet understood whether Nsp15 initiates flipping or captures

spontaneously flipped bases. We address this gap by designing fluorinated dsRNA substrates that allow us to directly relate a U's sequence context to both its tendency to spontaneously flip and its susceptibility to cleavage by Nsp15 via a combination of nuclease assays, 19F NMR spectroscopy, mass spectrometry, and single particle cryo-EM. Nsp15's cleavage efficiency was found to be directly related to that U's tendency to spontaneously flip – with one exception. U•G wobble pairs, which flip exclusively towards the major groove of dsRNA, are exceptionally resistant to cleavage by Nsp15. This suggests that Nsp15 primarily intercepts bases that have spontaneously flipped towards the minor groove, rather than actively initiating base flipping. Overall, our findings unify previous characterizations of Nsp15's cleavage preferences, and suggest that activity of Nsp15 during infection is partially driven by bulged or otherwise relatively accessible Us that appear at strategic positions in the viral RNA.

Session II: RNA: Protein Interactions and RNA Modifications

The role of DDX3X in the initiation of cardiac sex-differential protein expression

Kayla K. Mason, Kyle Park, Frank L. Conlon
UNC Chapel Hill

Sex disparities exist in both cardiac anatomy and physiology and prevalence in heart disease. These differences are often attributed to sex hormones. Through a series of genetic screens, we identified DEAD-box helicase 3 X-linked (Ddx3x) as a candidate for regulating sex chromosome-dependent differences in the mammalian heart. DDX3X has proposed functions in translation initiation, splicing, and nuclear export, having been shown to bind and resolve secondary structures in mRNA. Consistently, point mutations in DDX3X cause congenital heart disease in females predominantly. To date, no investigation into the cardiac function of DDX3X has been performed. We have generated a Ddx3x cardiomyocyte (CM) conditional allele in mice and have shown it is essential in CMs for embryonic viability in females but not males by E12.5. Additionally, at E10.5 proteins associated with cardiac development, cell cycle, and electrical signaling are downregulated in female hearts. Because DDX3X has multiple proposed functions in RNA biology, we have immunopurified the endogenous cardiac interactome and demonstrated that DDX3X interacts with ribosomal proteins in the female heart. To identify the DDX3X cardiac target mRNAs, we performed Enhanced Crosslinking and Immunoprecipitation (eCLIP) in parallel with ribosome profiling of female Ddx3x null hearts. Our results imply that DDX3X functions to regulate translation initiation of target mRNAs within the female mammalian heart. Based on this, I hypothesize that DDX3X binds mRNA targets to regulate cardiac sex-differential protein expression. These studies will provide insight into the basis both for cardiac sex differences and potential clinical relevance for DDX3X-associated CHD patients.

Single-cell m6A profiling in the mouse cortex reveals cell type-specific methylation profiles as well as changes during aging and stress

Matthew Tegowski, Anna K. Prater, Christopher L. Holley, Kate D. Meyer
Duke University

N6-methyladenosine (m6A) is an abundant mRNA modification that plays critical roles in neurodevelopment and brain function by regulating mRNA stability, localization, and translation. However, due to technical limitations, transcriptome-wide profiling of m6A sites within the diverse cell types of the brain has not been possible. To address this, we developed a mouse model to perform single-cell DART-seq (scDART-seq), which allows for global mapping of m6A sites at single-cell resolution in vivo. We used this model to investigate cell type-specific methylation within the mouse cortex and found that most methylation events are consistent across cell types, consistent with other studies done in bulk. However, we also found a striking lack of methylation in microglia, the resident immune cells of the brain, compared with other cell types, which has not been previously observed. Furthermore, it is also unknown how m6A changes during development or in response to specific stimuli. So, we used scDART-seq to identify cell type-specific changes in methylation during aging and in response to restraint stress. While the methylation of most mRNAs in most cell types is relatively unchanged, we also find many sites with age- and stress- dependent changes within neurons. Importantly, many neuronal mRNAs with altered m6A methylation are involved in aging-associated neurodegenerative disease, like App, which encodes the precursor to the β-amyloid peptide associated with Alzheimer's disease, or important regulators of neuronal function. Altogether, we present the first single-cell global m6A methylation profiles in vivo and identify important changes in methylation across cell types, aging, and stress.

Session III: RNA Targeting and Therapeutics

A multicomponent, out-of-equilibrium logic controls the assembly of nuclear paraspeckles

Wilton T. Snead, Krishna Shrinivas, Amy S. Gladfelter
Duke University

Biomolecular condensates mediate diverse functions in all cells. The mechanisms by which cells maintain coexisting condensates of distinct identity are poorly understood. We set out to uncover the rules that govern the assembly and identity of multi-layered nuclear paraspeckles, which mediate post-transcriptional gene expression. Paraspeckles comprise multiple RNA-binding proteins segregated into core and shell layers, the organization of which is mediated by the long noncoding RNA NEAT1. Previous work proposed that two sub-domains from the ends of NEAT1 localize to the shell, while three sub-domains from the middle localize to the core. How these sub-domains mediate layer formation is unclear. We synthesized RNA fragments corresponding to each sub-domain and assembled condensates using the core protein FUS. Unexpectedly, shell RNAs promoted FUS condensation more strongly compared to core RNAs, owing to a greater density of FUS binding motifs in shell RNAs. Given the lack of preference for core RNA, we wondered how paraspeckle

layers form. Importantly, paraspeckles assemble co-transcriptionally, meaning that proteins interact with different portions of NEAT1 during the minutes-long timescale of its transcription. To emulate this dynamic process, we performed RNA mixing experiments using RNAs labeled with different fluorophores. RNAs segregated into layers when shell RNA was introduced 10 minutes after core RNA, suggesting that core RNA promoted a gel-like state that prevented shell RNA mixing. These findings suggest that the NEAT1 "blueprint" for paraspeckles is dynamically decoded by proteins during transcription. We hypothesize that co-transcriptional assembly is a general mechanism for building coexisting nuclear condensates of distinct composition.

Chemical Visualization of Nascent RNA Folding with Co-Transcriptional TMO-MaP

Edgar Faison, Patrick Irving, Jeff Ehrhardt, Rhese Thompson, Kevin Weeks, Qi Zhang
UNC-Chapel Hill, Qi Zhang Lab

RNAs are functionally diverse biomolecules with complex conformational landscapes. During the process of transcription, RNAs explore their encoded conformational landscapes, where the timed and sequential folding events can diversify end-state conformations and functional consequences for a single transcript. Hence, elucidating the evolving conformational landscape along the transcriptional coordinates is critical to understanding the complete functional envelope of an RNA transcript. However, such a task remains technically and conceptually challenging. Here, we present Cotranscriptional-TMO Mutational Profiling, CT-MaP, for the facile and versatile time-resolved interrogation of nascent RNA secondary structures at single nucleotide resolution under active transcription. We demonstrated this technique on two transcriptional riboswitches, the Bacillus cereus fluoride riboswitch and the Bacillus subtilis FMN riboswitch, revealing uniquely links transcriptional dynamics to critical folding events for regulating ligand-dependent gene transcription. Our findings further suggest a possible resolution to a bona fide structural paradox with implications for developing RNA-targeted small molecule therapeutics.

A Complex issue: When 2 X's need to Xist

Diana R. Dou, Yanding Zhao, Julia A. Belk, Yang Zhao, Kerriann M. Casey, Derek C. Chen, Rui Li, Bingfei Yu, Suhas Srinivasan, Brian T. Abe, Katerina Kraft, Ceke Hellström, Ronald Sjöberg, Sarah Chang, Allan Feng, Daniel W. Goldman, Ami A. Shah, Michelle Petri, Lorinda S. Chung, David F. Fiorentino, Emma K. Lundberg, Anton Wutz, Paul J. Utz, Howard Y. Chang
Duke

Autoimmune diseases are the third most prevalent disease in America and disproportionately impact XX females 4x more than XY males. Irrespective of hormonal status in both humans and mice, the XX sex chromosome complement is strongly associated with susceptibility to autoimmunity, but the genetic risk

underlying autoimmune diseases from the second X chromosome in aggregate remain unresolved. The Xist long noncoding RNA (lncRNA) is expressed only in XX individuals and is required to randomly inactivate one of the two X chromosomes to achieve gene dosage compensation. Here, we show that the Xist ribonucleoprotein (RNP) complex, comprised of numerous autoantigenic components, is an important driver of sex-biased autoimmunity.

In this study, we utilized an inducible and non-silencing allele of Xist introduced into an autosome in mice to evaluate the impact of the XIST RNP in autoimmune predilection independent of sex chromosome or hormonal background. Inducible expression of transgenic Xist in male mice introduced Xist RNP complexes and sufficed to produce autoantibodies. Male SJL/J mice expressing transgenic Xist developed more severe multiorgan pathology in the pristane-induced model of lupus than wild-type males. Xist expression in males also reprogrammed T and B cell population and chromatin states to more resemble wild type females. Concurrently, antigen array studies showed autoimmune disease patients displayed significant seroactivity towards multiple components of the XIST RNP. Altogether, our data point to a significant role for Xist as a scaffold for ubiquitous RNP components and a driver for autoimmunity underlying the sex-biased female preponderance for developing autoimmune diseases.

Session IV: RNA Processing, Localization, and Translation

A General Mechanism for Tunable RNA-Mediated Regulation of Chromatin-Binding Proteins

Wayne O. Hemphill, Regan Fenske, Calvin K. Voong, Anne R. Gooding, Jiarui Song, Arthur J. Zaug, Karen J. Goodrich, James A. Goodrich, Thomas R. Cech
University of Colorado Boulder

Many DNA-binding proteins, like transcription factors and chromatin-modifying enzymes, have been shown to also bind RNA, leading to broadly relevant questions regarding the regulatory roles(s) of such RNA binding. The histone methyltransferase, Polycomb Repressive Complex 2 (PRC2), deposits the H3K27me3 epigenetic mark to negatively regulate expression at numerous target genes, and this activity has been implicated in embryonic development, cell differentiation, and various cancers. Numerous in vitro studies demonstrate that RNA can inhibit PRC2 histone methyltransferase activity through mutually antagonistic binding with nucleosomes, but how such RNA binding might regulate PRC2 activity in vivo remains an area of active investigation. We used biochemical, biophysical, and computational approaches to interrogate PRC2's RNA and DNA binding kinetics. Our findings suggest that PRC2 can 'direct transfer' between RNA and DNA without a free-enzyme intermediate, likely via a highly unstable ternary complex. Simulations based on these PRC2 kinetic data predict that direct transfer should allow chromatin-modifying activity to be both positively and negatively regulated by RNA binding, depending on the proportions of nucleosome-proximal (e.g., nascent) versus freely diffusing RNA. These predictions were affirmed by in vitro histone methyltransferase assays with covalently linked RNA-nucleosome arrays. Finally, additional experiments with numerous other

RNA-, DNA-, and small molecule-binding proteins demonstrate that direct transfer activity may be a common capability of nucleic acid binding proteins, suggesting that this mechanism could have general relevance to RNA-binding chromatin-associated proteins.

Metabolic control of protein synthesis by ribosome fall-off

Alicia Darnell[1,2], Christopher Chidley[3], Victoria Paradise[2], Kristian Davidsen[4], Sarah Lincoln[5], Danica Cui[5], Campbell Vander Heiden[2], Lucas Sullivan[4], Peter Sorger[3], Joseph Davis[5], Matthew Vander Heiden[2,5]
[1]Department of Pharmacology and Cancer Biology, Duke University
[2]Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology
[3]Harvard Medical School and Laboratory of Systems Pharmacology
[4]Fred Hutchinson Cancer Research Center
[5]Department of Biology, Massachusetts Institute of Technology

Amino acids are a key nutrient for cancer cells. If scarce, they can become limiting substrates for protein synthesis through charged tRNA depletion that leads to ribosome stalling and premature fall-off, highlighting a direct but poorly understood pathway through which cellular metabolism controls gene expression. With a reporter library to measure ribosome fall-off during amino acid limitation at any sense codon, we find that arginine limitation uniquely and consistently induces ribosome fall-off, specifically at the arginine codon AGA. This is driven by depletion of charged Arg-tRNATCT, and AGA usage in mRNAs correlates with reduced rates of protein synthesis across the proteome during arginine limitation. In FACS-based screens, we discovered that the ribosome quality control (RQC) pathway, known to orchestrate ribosome fall-off on defective or damaged mRNAs, also catalyzes ribosome fall-off during arginine limitation. Fall-off during arginine limitation does not require the canonical RQC sensor ubiquitin ligase ZNF598, suggesting that “metabolic” ribosome fall-off is downstream of a distinct RQC-initiating signal. Our work outlines a new mechanism for metabolism to control gene expression at the ribosome and reframes our understanding of the physiological targets of ribosome quality control.

A death-promoting post-transcriptional mechanism in heat exposed cells

Ru-pin Alicia Chi[1], Yin Li[1], Ankit Gupta[1], Brian Papas[2] and Marcos Morgan[1].
[1]Reproductive and Developmental Biology Laboratory, NIEHS/NIH, RTP, Durham, NC, USA.
[2]Integrative Bioinformatics, Epigenetics and Stem Cell Biology Laboratory, NIEHS/NIH, RTP, Durham, NC, USA.

When challenged with environmental stressors, cells need to respond rapidly to ensure the best outcome. This includes reprogramming of gene expression to minimize and repair the damage caused by the stressor, and in cases where repair is not possible, promote cell death. Post-transcriptional modification of mRNAs is

one way to rapidly alter gene expression without having to synthesize new transcripts or degrade existing ones. Here, we identified the cytoplasmic poly(A) polymerase enzyme TENT5A as a heat stress (HS) regulator in somatic and germ cells. Upon heat exposure, cells upregulated Tent5a expression indicating possible roles in HS response. Through global poly(A) profiling, we identified Atxn2 as a putative TENT5A target. We demonstrate that TENT5A likely elongated Atxn2 poly(A) tails which resulted in stabilization of transcript but reduced translation. ATXN2 itself is a stress responder by facilitating stress granule (SG) formation, and we subsequently show that TENT5A over-expressing cells downregulated SG after heat exposure, which correlated with reduced cell viability. This indicated that TENT5A functions to promote cell death under heat conditions. In mice, Tent5a is similarly induced after scrotal HS in the germ cells. Supporting the pro-death function of TENT5A we observed in somatic cells, we found that the germ cells in Tent5a knockout mice were protected from heat-induced apoptosis, ultimately resulting in improved reproductive outcomes compared to wildtype mice after scrotal heat exposure. Taken together, we show that TENT5A modulate cell response to heat by promoting apoptosis, potentially through post-transcriptionally modulating Atxn2.

An unbiased screen identifies the PAXT connection as a modifier of rare codon-dependent expression

Taylor Niehoff, Jackson Peterson, Siqi Li, Christopher Counter
Duke University School of Medicine

Rare codons, those occurring least often in the genome, typically dampen expression through affecting chromatin structure, mRNA stability, mRNA localization, and translation. However, there are a growing number of examples whereby this norm is broke, and instead rare codon-enriched genes are highly expressed. This suggests the intriguing possibility that unknown mechanism(s) exist to overcome the poor expression imposed by rare codons. Given this, we screened an sgRNA library for genes that when inactivated increased the level of fluorescent protein in a rare codon-dependent fashion. Ten genes were identified from this screen, six of which were either part of or associated with the PAXT connection. The PAXT connection is a multi-protein exosome complex that degrades polyadenylated RNA in the nucleus. I have verified that inactivating components of the PAXT connection increased the amount of reporter protein in a rare codon-dependent manner. However, rare and common codon-enriched reporter RNA were both increased and increasing the amount of the reporter reduced the effects of PAXT knockout on RNA and protein changes. This suggests that the rare and common codon-enriched reporter transcripts are targets of the PAXT connection, but the increase in rare mRNA manifests as a larger increase in protein expression. From analysis of RNA seq and proteomics from PAXT knockout cells, specifically rare codon-enriched endogenous genes are significantly increased at the RNA and protein levels with knockout– confirming that specifically rare PAXT targets are increased at the protein level. This work identifies the PAXT connection as a new mechanism of codon-dependent regulation.

Inflammation and alternative polyadenylation on alternative polyadenylation

Jiamutai, Abigail Hatfield, Austin Herbert, Debarati Majumdar, Vijay Shankar, **Lela Lackey**
Clemson University

Alternative polyadenylation (APA) results in different 3' isoforms of transcripts. When APA occurs in the 3' untranslated region (3'UTR) it can alter RNA localization, stability and translational efficiency. The SERPINA1 mRNA has two distinct 3' UTR isoforms, both of which code for the protease inhibitor α-1-antitrypsin (A1AT). A1AT is an acute phase protein that is predominately expressed in liver hepatocytes and upregulated during inflammation. We analyzed the dynamics of APA during cellular stress by treating the liver cell line HepG2 with the cytokine interleukin 6 (IL-6), ethanol or peroxide. SERPINA1 is transcriptionally up-regulated after IL-6 treatment and has altered polyadenylation, resulting in an increase in long 3'UTR isoforms. We find that the long 3'UTR strongly represses endogenous A1AT protein expression even with high levels of SERPINA1 mRNA. Changes in transcriptome-wide expression in HepG2 cells due to stress suggest changes in secretory protein processing. Our data show that IL-6-mediated inflammation influences polyA site choice for SERPINA1 precursor transcripts resulting in reduced A1AT protein expression.

POSTER ABSTRACTS

1

Decoding Repression: Recruitment of epigenetic silencers by RNA binding proteins and long non-coding RNAs

E.W. Abrash, M.M. Murvin, G.A. Goda, D. Dominguez, and J.M. Calabrese
UNC Chapel HIll

Gene regulation by epigenetic enzymes underlies proper development and homeostasis, and its dysregulation drives many forms of disease, including cancers. Therefore, understanding how epigenetic modifiers are recruited to chromatin is essential to untangling the mechanisms that underlie life. Epigenetic modifiers can be recruited by lncRNAs, the most potent of which, Xist, silences an entire X chromosome. However, it is unclear how lncRNAs encode the ability to recruit epigenetic modifiers. As the most powerfully repressive lncRNA known, Xist is an ideal model for decoding how lncRNAs recruit epigenetic modifiers and serves as a paradigm for other lncRNA-enzyme relationships. Xist requires an abundant RNA Binding Protein (RBP) known as hnRNPK to recruit the silencing enzyme complex Polycomb Repressive Complex 1 (PRC1). Yet, paradoxically, hnRNPK binds thousands of RNAs without contributing to transcriptional repression. Herein, we apply a large-scale in vitro binding assay to show that Xist's affinity for hnRNPK is substantially higher than that of other expressed RNAs and that its high-affinity hnRNPK sites are highly clustered within only ~10% of its sequence. We next use a series of mutants to test whether clustered hnRNPK-binding regions are sufficient to restore silencing function to a non-functional Xist fragment and discovered that two Xist hnRNPK-binding regions work together to recruit PRC1. Ultimately, defining the underlying RNA sequence features that enable RNA-mediated PRC1 recruitment will identify a paradigm that will guide studies of other lncRNAs, RBPs, and silencing enzymes, and may suggest new therapeutic strategies to control deleterious gene expression.

2

tts-1 lncRNA modulates C. elegans longevity pathways

Ben McCarthy, Eli Waite, Emily Mathew, Evan Lister-Shiumauchi, Shawn Ahmed
UNC Chapel Hill

Conserved TOR and insulin signaling pathways function in parallel to suppress aging in diverse species. The most highly expressed RNA 1) in stress-resistant dauer larvae, 2) in long-lived C. elegans daf-2/insulin/IGF-1 signaling mutants, and 3) in response to starvation, is a 711 nucleotide long non-coding RNA (lncRNA) called tts-1 (1). tts-1 was named telomere transcribed sequence-1 because it contains several telomere repeats. The trt-1 telomerase reverse transcriptase maintains telomere length in C. elegans, but the non-coding telomerase RNA subunit that contains the template for the (TTAGGC)n nematode telomere repeat sequence is not known. We asked if tts-1 might be the C. elegans telomerase RNA by deleting tts-1 using CRISPR/Cas9 genome editing, but tts-1 mutants did not display transgenerational sterility and telomere fusions

of telomerase mutants (2).We found that deletion of tts-1 does not affect aging in an otherwise wildtype background. However, tts-1 strongly represses aging in daf-2 mutants yet promotes longevity in rsks-1 ribosomal S6 kinase mutants are deficient for the Target Of Rapamycin (mTORC1) nutrient sensing pathway. We used RNA FISH to show that tts-1 is expressed in intestinal and neuronal nuclei and often localizes to prominent foci. Non-coding RNAs that occupy discrete locations within nuclei can regulate gene expression, so the tts-1 lncRNA may possess an epigenetic function. We are studying physical interactions of the tts-1 lncRNA in an effort to understand how it regulates aging.

1. Jones SJ, Riddle DL, Pouzyrev AT, Velculescu VE, Hillier L, Eddy SR, Stricklin SL, Baillie DL, Waterston R, Marra MA (2001), Changes in gene expression associated with developmental arrest and longevity in Caenorhabditis elegans. Genome Res.2. Meier B, Clejan I, Liu Y, Lowden M, Gartner A, Hodgkin J, Ahmed S (2006), trt-1 Is the Caenorhabditis elegans Catalytic Subunit of Telomerase. PLOS Genetics 2(2): e18.

3

Intronic Alu elements adopt diverse RNA structures

Scott R. Allen, Abigail C. Lehr, Alain Laederach, Kevin M. Weeks
UNC Chapel Hill

Alu elements, a class of transposable repeat elements in the human genome, derive from the 7SL RNA and have a widely reported secondary structure resembling that of the 7SL RNA Alu domain. There are ~1.4 million Alu elements inserted into the human genome, comprising ~13% of the transcriptome, many of which are encoded in intronic sequences and are transcribed as part of the pre-mRNA. Despite their abundance and numerous reported roles in both normal biology and disease states, very few examples of Alu elements exist for which there are chemical probing data or data-informed structural models. Here, we use DMS probing coupled with mutational profiling (DMS-MaP) to measure the secondary structure of several previously uncharacterized intronic Alu elements folded in the context of flanking intronic RNA sequence. Our findings suggest that intronic Alu elements do not systematically adopt the expected secondary structure model and often pair with flanking sequences. Each Alu element has, in essence, a unique RNA structure. These findings expand the known RNA structure space occupied by Alu elements and suggest that RNA structural differences underlie the diverse functional roles observed for Alu elements.

Defining how viral infection redirects targeting of the m6A methyltransferase complex to specific host transcripts

Caroline J. Aufgebauer, Theodore M. Nelson, Christopher E. Mason, Stacy M. Horner
Duke University

Modulation of chemical modifications to RNA facilitates reprogramming of gene expression in response to cellular stress. We previously found that Flaviviridae infection induces altered levels of the RNA modification

N6-methyladenosine (m6A) on specific host transcripts, leading to changes in RNA splicing and translation of genes important for infection. Notably, expression levels of the m6A-machinery are unchanged during infection, suggesting that Flaviviridae-driven changes in m6A are likely mediated by altered targeting of the m6A-machinery. However, it remains unknown how targeting of the m6A-machinery to specific host RNAs is altered in response to infection. We found that transcripts with increased m6A during infection also have increased binding by the m6A methyltransferase, METTL3, during infection. As previous studies have shown that RNA binding proteins can regulate METTL3 targeting, we used proximity proteomics to identify novel METTL3 protein interactions during infection. Our preliminary analysis revealed that many RNA cleavage and polyadenylation factors have increased interaction with METTL3 upon infection. Interestingly, several m6A-altered transcripts have shortened 3UTRs upon viral infection, thus we are currently studying the crosstalk between RNA targeting of cleavage/polyadenylation factors and METTL3.

5

Identifying roles of nuclear DVL1, a Wnt-pathway protein, in mRNA regulation and splicing

Rachel L. Babcock[1,3], Dalia Martinez-Marin[2], Geetha Boligala[2], Grace Stroman[2], Rakhshanda Layeequr Rahman[3,4], Jannette M. Dufour[1], and Kevin Pruitt[2].

[1]Department of Cell Biology and Biochemistry, Texas Tech University Health Sciences Center, Lubbock, TX, 79430, USA

[2]Department of Pharmacology, Lineberger Comprehensive Cancer Center, The University of North Carolina at Chapel Hill, Chapel Hill, NC, 27599, USA

[3]Breast Center of Excellence, Texas Tech University Health Sciences Center, Lubbock, TX, 79430, USA

[4]Cancer Institute, MetroHealth System, Cleveland, OH, 44109, USA

Dishevelled 1 (DVL1) protein is a scaffold protein most known for its cytoplasmic role to regulate Wnt signaling, a pathway integral for various developmental processes, including tumorigenesis. We recently reported DVL1 is enriched in triple negative breast cancer (TNBC), promotes tumor growth, and localizes to the nucleus. While DVL is critical for Wnt signaling, surprisingly little is known about its nuclear role. Herein, we performed co-immunoprecipitation (co-IP) of nuclear DVL1 from two TNBC cell lines (MDA-MB-231 and MDA-MB-468) and a control, non-cancerous mammary cell line (MCF10A), followed by affinity purification mass spectrometry (AP-MS) to identify novel protein-protein interactions of nuclear DVL1. Pathway analyses on the ~2600 identified DVL1 binding factors suggests DVL1 likely participates in the regulation of mRNA processing/translation, a new and undefined role for DVL1. In the MDA-MB-231 cell line, we found DVL1 bound with numerous splicing factors, including serine/arginine-rich splicing factor 1 (SRSF1). Follow-up co-IP assays validated SRSF1 interaction with nuclear DVL1 in two TNBC and a non-TNBC cell line. Excitingly, our preliminary differential exon usage analysis from our RNA-Sequencing assay indicates differential splicing events between TNBC cells over-expressing DVL1, compared to empty vector controls. Our future assays will determine mechanisms by which DVL1 directly affects SRSF1 function and how this contributes to mRNA splicing of key genes related to tumorigenesis. We anticipate completion of this study will uncover novel roles of nuclear DVL1, strengthening our broader understanding of DVL1 as a potential therapeutic target in TNBC to improve women's health.

6

Mechanical stretch induces changes in CLK1 and SR protein functions in myotubes

Gabrielle B Bais, Emma R Hinkle, R Eric Blue, Matthew Combs, Joan M Taylor, Jimena Giudice
UNC Chapel Hill

Cells use mechanotransduction to sense and respond to mechanical stimuli from their physical environment. The impact of mechanical forces on the intracellular environment is especially important in muscle cells because they are constantly generating and responding to mechanical forces. Alternative splicing is an RNA processing mechanism that allows single genes to produce multiple mRNA transcripts. Skeletal muscles have one of the highest levels of tissue-specific alternative splicing, due to specialized muscle cell functions. Numerous components of the contractile machinery of muscle cells (sarcomeres) undergo alternative splicing regulation impacting contractility. Our lab has previously shown that cyclic mechanical stretch of skeletal muscle cells induces global changes in transcription and alternative splicing. Several differentially expressed and spliced transcripts identified in our deep RNA-sequencing study were targets of the serine/arginine rich splicing factor 4 (SRSF4). CDC2-like kinase 1 (CLK1) regulates the subcellular localization of serine/arginine-rich (SR) proteins and their splicing regulatory functions via phosphorylation. Interestingly, Clk1 pre-mRNA is itself a splicing target of SR proteins. We therefore investigated the interrelationship between CLK1 kinase production and activity, SR protein phosphorylation, and mechanical stretch in muscle cells. Mechanical stretching of muscle cells led to an increased production of functional Clk1 transcripts and increased SRSF4 phosphorylation. Inhibition of CLK1 kinase activity (TG003) resulted in an increase in CLK1 mRNA and protein production and a decrease in SRSF4 phosphorylation. These findings suggest that CLK1 phosphorylation of SRSF4 regulates the production of functional CLK1 protein via a feedback loop that connects mechanical stretching, with CLK1 splicing regulation by phosphorylated SRSF4. The function of SRSF4 as a mechanically sensitive RNA-binding protein that is regulated by CLK1 expands the current understanding of the intricate interplay between mechanotransduction and alternative splicing.

7

Revealing RNA poly(A) tail dynamics during spermiogenesis and its function to support male fertility

Marine Baptissart, Ankit Gupta, Alexander C. Poirot, Brian Papas and Marcos Morgan
NIEHS

Spermiogenesis is the final process of spermatogenesis where round spermatids are elongating to shape the future sperm. Because spermatids are silenced transcriptionally, their differentiation relies exclusively on post-transcriptional regulations. In the 80s, changes in mRNA poly(A) tail length have been shown to control the translation of Transition proteins and Protamines transcripts during spermiogenesis. Since, our understanding of how poly(A) tail processing might shape the transcriptome of spermatids remains limited. Here, we characterize the whole transcriptome poly(A) tail dynamics during spermiogenesis and its impact on mRNA metabolism and fertility. We used FACS sorting to isolate round and elongated spermatids from

mice testis. Long read RNA sequencing was adapted to profile poly(A) tails in both cell types. As the germ cells enter spermiogenesis, we show a global increase in poly(A) tail length associated with transcript stabilization. During spermiogenesis progression, the poly(A) tail of transcripts is shortened to 60-nts leading to translational activation. In addition, unstable transcripts show high proportion of uridylated poly(A) tails, a 3'-end modification commonly associated with mRNA decay. TENT5C and TUT7 are terminal nucleotidyltransferases responsible for mRNA polyadenylation and 3'-end uridylation, respectively. They both are required for spermatogenesis, but their role in shaping the poly(A) tail of the spermatid's transcriptome remains limited. We demonstrate for the first time that the poly(A) polymerase activity of TENT5C is required for fertility and spermiogenesis: male mice expressing catalytically dormant TENT5c (TENT5C dCat) are sterile and produce headless spermatozoa - a phenotype associated with a handful of transcripts showing abnormally short poly(A) tails in elongated spermatids. With the analysis on TUT7-depleted spermatids, we revealed a global loss of uridylated transcripts in elongated spermatids; however, this process is not required for fertility as TUT7 null mice remain fertile. In summary, we comprehensively characterize the dynamic of poly(A) tail during spermiogenesis and its role to support mRNA metabolism and fertility. This study opens new perspectives with the potential to exploit poly(A) tailing to develop male contraceptive and therapeutics for infertility.

8

The E2-ubiquitin conjugase Rad6 mediates translation control under oxidative stress

Gessica C. Barros, Sezen Meydan, Chia-Yu Chen, Blanche Cizubu, Nicholas R Guydosh, Gustavo M. Silva
Duke University

Regulation of cells' response to stress at the translation level involves translation repression of non-stress responsive mRNAs to allow translation of stress responsive proteins. The Redox control of Translation by Ubiquitin (RTU) is a new pathway controlling translation elongation in yeast exposed to oxidative stress. RTU is characterized by an accumulation of K63-linked polyubiquitin chains that modify ribosomal proteins leading to ribosome arrest through action of the ubiquitin enzymes Rad6 (E2-conjugase) and Bre1 (E3-ligase). However, how these enzymes contribute to the translation pause is unknown. To search for ribosomal pause signatures controlled by Rad6 and Bre1, we performed Ribosome-sequencing (RiboSeq) analysis that revealed a preferential ribosome stall at isoleucine and proline motifs (XIP). These pauses are abolished upon deletion of RAD6 and dependent on Rad6's ubiquitination activity. Intriguingly, bre1 minimally reduces XIP pauses. Our analysis on the biological meaning of these pauses suggests a global control of translation instead a selective pathway targeting the antioxidant response. To test this model, we evaluate global translation and determined that Rad6's ubiquitination activity controls translation repression under stress. Bre1 also participates in the translation repression, but our data suggests the existence of compensatory mechanisms that compensate Bre1's absence. Future studies are needed to uncover the connection and temporal organization among all translation phenotypes controlled by the RTU enzymes. Extended studies will also clarify the mechanisms and functionality of the XIP pauses.

9

Defining how the m6a-methyltransferase complex targets hepatitis C viral RNA

Katherine M. Bland, Jordan V. Reaves, Stacy M. Horner
Duke University

Targeting of N6-methyladenosine (m6A) to cellular mRNA molecules is governed by proteins in the m6A-methyltransferase complex (m6A-MTC). The m6A-MTC can also add m6A to viral RNA molecules. Indeed, we have shown that methylation of the viral RNA of hepatitis C virus (HCV), a positive-strand RNA virus, requires several members of the m6A-MTC. Specifically, we have found that WTAP, the central scaffolding protein of the m6A-MTC, recruits the catalytic core proteins METTL3 and METTL14 to HCV RNA for methylation. The mechanisms by which these m6A-MTC proteins target HCV RNA for methylation is unknown. Additional proteins in the m6A-MTC are known to regulate its recruitment to cellular mRNA molecules. However, as the generation of HCV RNA molecules is distinct from cellular mRNA, it is unclear if and how these proteins regulate HCV RNA methylation. Here, we aim to identify the m6A-MTC proteins that regulate methylation of HCV RNA. This study will define the role of the known m6A-MTC protein VIRMA in the recruitment of the m6A-MTC to HCV RNA for m6A-modification. VIRMA has previously been described to have site-specific roles in m6A deposition to cellular RNAs. Additionally, VIRMA has been identified as a potential interactor with several HCV proteins. Thus, we hypothesize that VIRMA is required for proper m6A modification of viral RNA and subsequent regulation of HCV infection. Preliminary results suggest VIRMA is required for methylation of HCV RNA. Overall, this work will lead to new insights into how the m6A-MTC can be differentially recruited to viral RNA molecules during infection.

10

Alternative splicing regulation of clathrin in heart development and function

R. Eric Blue, Adam Black, Matheus Sadovsky, Gabrielle Gentile, Brian Cooley, Haifeng Yin, Jimena Giudice
UNC Chapel Hill

Alternative splicing mechanisms regulate membrane trafficking genes during the first four postnatal weeks of heart development. These splicing-trafficking events are highly tissue-specific and evolutionarily conserved. One trafficking gene regulated by alternative splicing during heart development is the clathrin heavy chain (Cltc). Alternative splicing of a microexon (exon 31, 21 nucleotides) in Cltc generates two isoforms: a long CLTC form that includes the microexon and a short CLTC form that skips it. This microexon is tissue specific for adult heart and skeletal muscles. The short form is predominantly expressed in embryonic stages and there is a gradual transition towards primary expression of the long form in adult hearts and skeletal muscles. Several human diseases cause adult hearts to revert to fetal splicing programs which fail to support the adult functions of the organ. However, little is known about the cellular mechanisms underlying these diseases and less is known about microexons' regulation in heart function. We found that Cltc splicing is altered in both human diseases (i.e. myotonic dystrophy type 1) and mouse models of cardiac pressure overload

(TAC) and myocardial infarction (left anterior descending coronary artery ligation, LAD). To investigate the mechanisms by which Cltc splicing regulates cardiac biology, we utilized CRISPR/Cas9 editing in mice to block endogenous Cltc splicing by deleting Cltc microexon 31 from the genome. Cltc-homozygous (Cltc-HO) hearts express only the short CLTC form across development, while Cltc-wildtype (Cltc-WT) hearts undergo normal Cltc splicing regulation during development. Interestingly, by challenging the heart with pressure overload (via transverse aortic constriction, TAC), we found Cltc-HO mice had a less severe heart failure response across 8 weeks following TAC. We hypothesize that Cltc exon-skipping can precondition adult hearts to better respond to the effects of pressure overload and other cardiac challenges. Moreover, although Cltc-WT and Cltc-HO hearts similarly express hypertrophic and fibrotic markers following the TAC challenge we found interesting differences in the expression of mRNAs and proteins involved in the regulation of iron metabolism and ferroptosis. Our ongoing efforts are directed to define the mechanisms by which Cltc splicing regulates heart physiology and failure and set the stage for potential therapies that redirect splicing to help treat severe heart challenges.

11

Decoding SENP3 and PELP1 Dynamics within the RNA Rixosome

Saisamhita R. Bommu, Jacob Gordon, Robin E. Stanley
National Institute of Environmental Health Sciences

Ribosome biogenesis is a critical cellular process essential for ribosome production. The RNA Rixosome, a nuclear protein complex involved in RNA processing and heterochromatin maintenance, facilitates this process by helping mature ribosomal RNA (rRNA). Proteins in the Rixosome are often modified by SUMO (Small Ubiquitin-like Modifier), influencing their function. This study examines the interaction between SUMO-specific protease 3 (SENP3) and PELP1, both stable components of the Rixosome complex. SENP3 regulates deSUMOylation, while PELP1, a key scaffold protein, is involved in ribosome maturation, cell cycle regulation, and hormone signaling. The interplay between SENP3 and PELP1 remains poorly understood. Through a combination of immunoprecipitation and western blot analysis, we confirmed the direct interaction between SENP3 and PELP1, identifying a specific binding region within PELP1 that mediates this interaction. AlphaFold Multimer computational modeling suggested that a highly conserved Val-Glu-Ile motif in PELP1 (776aa-778aa) is critical for supporting the interaction. This was further validated biochemically through site-directed mutagenesis of the conserved motif, which demonstrated that SENP3 was unable to bind when the motif was mutated. Additionally, we observed that the stability of the SENP3-PELP1 complex is enhanced in the presence of other Rixosome components, such as TEX10 and WDR18. Sequence alignment analysis indicates that this SENP3-binding motif is evolutionarily conserved in vertebrates, underscoring its potential biological significance in regulating Rixosome function in higher eukaryotes. These findings not only deepen our understanding of the regulatory mechanisms governing SUMO dynamics within the RNA Rixosome but also reveal a previously uncharacterized structural interaction between SENP3 and PELP1.

12

Cloneable nanotechnology for investigating molecular mechanisms of preribosome processing in vivo with electron cryotomography

Kanda Borgognoni, Zoe Wright, Robin Stanley
NIEHS/NIH

Mapping native protein structure at atomic resolution requires electron cryotomography (cryoET), and localizing proteins within cells using cryoET requires unique labeling strategies. Our approach utilizes a novel enzymatic tag denoted a cloneable selenium nanoparticle (cSeNP) to produce contrast in tomograms for easy identification of proteins in vivo. This tag is analogous to GFP in that cSeNP is expressed concatenated to a protein of interest, yet it is distinct in that inorganic cations are added to cell medium to produce the electron dense label. Production of cSeNPs requires optimization of bioorthogonal inorganic syntheses to produce NPs while mitigating toxicity. I optimized these reactions in bacteria and am currently translating the NP biosynthesis to yeast. We are targeting ribosome assembly since isolation of specific preribosome intermediates with current methods is nontrivial and because ribosomes are easily detected in cryoET. Ribosomes are essential for organism survival. Ribosome biogenesis involves codependent processing mechanisms of hundreds of ribosome assembly factors that aid in RNA quaternary structuring. The precision involved is critical to ribosome function, yet the mechanisms of in vivo assembly remain largely unexplored due to its complexity. As recent work has confirmed the link between ribosome malformation and human diseases, it is paramount that we uncover the molecular basis of ribosome assembly. We will use cSeNPs to localize distinct preribosomes in situ with cryoET, which will pave the way for unveiling precise steps involved in ribosome biogenesis.

13

Identification of chromatin associated RNAs with repressive-like regulatory potential

Samuel P. Boyson, Quinn Eberhard, Shuang Li, McKenzie M. Murvin, J. Mauro Calabrese
UNC Chapel Hill

Over seventy percent of the mammalian genome is transcribed into RNA, the vast majority of which is non-coding (ncRNA). While much of this ncRNA remains functionally uncharacterized, a growing number of long non-coding RNAs (lncRNAs) have emerged as important regulators of gene expression. One of the most extreme examples of lncRNA mediated gene regulation is the lncRNA Xist, which represses transcription across an entire X chromosome by recruiting various repressive enzyme complexes along the length of the X chromosome. In addition to Xist, other lncRNAs such as Airn and Kcnq1ot1 exhibit similar repressive function, with the ability to repress transcription across megabase regions of the genome. However, the exact features that confer function to these RNAs and the prevalence of these features in the rest of the transcriptome remains unclear. Using known repressive lncRNAs to define repressive characteristics, we employed both experimental and computational approaches in mouse trophoblast stem cells (TSCs) to broadly assay the transcriptome for regions harboring repressive-like characteristics such as associations with similar RNA

binding proteins and kmer content. We identify many regions within long chromatin associated transcripts containing repressive-like functional characteristics, which may be suggestive of more widespread functional activity for chromatin associated RNA across the transcriptome. Our data show that many ncRNAs harbor functional characteristics like those of known repressive lncRNAs and we predict that many of these chromatin associated transcripts may utilize these domains to repress transcription through similar mechanisms to Xist.

14

Predicting RNA-Protein Binding Energy Using the Novel IRIS Model Integrated with AlphaFold 3

Eduardo Cisneros de la Rosa, Yafan Zhang, Xingcheng Lin
NC State University

RNA-protein interactions are crucial for cellular function and regulation, making the accurate prediction of their binding energies is a key goal in molecular biology and bioinformatics. We introduce IRIS (Interpretable protein-RNA interactions Informed by Structure), a novel model that leverages both computational predictions and experimental data to achieve high-precision RNA-protein binding energy estimates. By integrating AlphaFold 3, IRIS assesses the structural impacts of mutations on RNA sequences, particularly enhancing prediction quality for structures with four or more mutations. For sequences with three or fewer mutations, the integration does not degrade prediction accuracy.To validate our computational predictions, we utilized data from an RNA array high-throughput assay, focusing on the 2bu1 RNA-protein complex. This assay generated a diverse collection of RNA sequences through variations of the hairpin, where different sequences were mutated using doped oligonucleotides. Barcoded RNAP initiation sequences were introduced to reduce sequencing errors and identify individual molecules within the population. These mutated sequences were then sequenced using Illumina technology to provide a comprehensive dataset for validation. Notably, structures with four or more mutations from the native sequence yielded unpredictable results until we incorporated AlphaFold 3 PDB inferred structures to calculate the new energy matrix. Our results highlight a significant improvement in the accuracy of RNA-protein binding energy calculations, particularly when combining computational models with machine-learning techniques to advance our understanding of RNA-protein interactions and their broader biological implications.

15

Multivalency of mRNA impacts the assembly of biomolecular condensates necessary for cell cycle control in multinucleate cells

Sierra J. Cole, Christine Roden, Benjamin Stormo, Amy S. Gladfelter
Duke University

Despite being a syncytial cell where multiple nuclei share the same cytoplasm, Ashbya gossypii nuclei behave independently and divide asynchronously. This independence is partly due to cyclin mRNA (CLN3) forming

biomolecular condensates with the protein Whi3, organizing RNA around individual nuclei. Interestingly, wild isolates of A. gossypii show more synchronous nuclear division at certain temperatures, possibly due to mutations in the number of Whi3 binding sites within the CLN3 transcript. To explore the role of multiple protein binding sites in RNA condensation, we generated CLN3 mRNA binding site mutants with identical valence but different binding site locations. Despite having the same number of binding sites, these mutants exhibited varying condensate properties depending on the mutated site. Additionally, RNA Bind-n-Seq data showed that oligos with the same valence could have different affinities. These results suggest the number of binding sites is not sufficient to determine condensate properties. Instead, it is likely that differences primary sequence surrounding binding sites or alterations to RNA secondary structure impact both RNA-protein interactions and potentially RNA-RNA interactions. We plan to confirm structural differences between mutants and observe their behavior within the cell to determine if changes in condensate properties affect the cell cycle. This model system reveals the complexity of RNA molecular grammar and the challenges in understanding how RNA valency impacts condensate formation.

16

Epigenetic CRISPR Editors encoded as mRNA for engineering primary human T cells

Rachel Conover, Charles Gersbach
Duke University

Cell engineering methods which include transgene integration have led to breakthroughs in cancer treatments such as chimeric antigen receptor (CAR) T cells. However, attempts to replicate these successes with non-integrating strategies such as expression of a CAR from mRNA have failed due to the transient nature of expression from the delivered genetic material. Here, we report development of CRISPR epigenetic editors for gene interference (CRISPRi) as mRNA constructs. These constructs, when delivered to primary human T cells, allow for knockdown for up to 10 days post-delivery. When combined with a library of sgRNAs targeting transcription factors delivered as lentivirus, these mRNA constructs can be used to screen for proteins which affect T cell phenotype after transient knockdown of endogenous expression.

17

Collective dosage homeostasis of condensation-prone proteins via their mRNAs

Neve Costello Heaven, Rupert Faraway, Jernej Ule
Francis Crick Institute

The concentration of proteins containing intrinsically disordered regions (IDRs) must be tightly controlled to maintain cellular homeostasis. However, mechanisms for collective control of these proteins, which tend to localise to membraneless condensates, are less understood compared to proteins at membrane-bound organelles. Here we report ‘interstasis’, a homeostatic mechanism that senses the concentration of co-condensing proteins and controls their gene expression through mRNA-mediated negative feedback.

Interstasis relies on multivalent mRNA regions that encode IDRs, which are reinforced by conserved codon biases and recognised by specific RNA-binding proteins. TRA2 proteins are strongest binders of multivalent purine-rich regions that encode charged IDRs, including arginine-enriched mixed charge domains (R-MCDs). Accumulation of R-MCD proteins increases the cohesion of nuclear speckles, a protein-RNA condensate, which recruits TRA2 proteins that selectively retain the purine-rich mRNAs in the speckles. This decreases further synthesis of charged proteins that are most highly prone to phase separation and are encoded by bidirectionally dosage-sensitive genes. Cdc2-like kinase (CLK) activity controls the localisation of TRA2 proteins to speckles, thereby modulating the setpoint of interstasis. Thus, interstasis is a collective feedback loop that senses the accumulation of condensation-prone speckle proteins, and then sequesters mRNAs that encode these proteins to promote their mutual homeostasis.

18

Alternative splicing of clathrin is essential for proper skeletal muscle development

Jessica L. Cote, R. Eric Blue, Matheus Sadovsky, Sumaya Addish, Olivia Perez, Adam J. Black, Jimena Giudice
UNC Chapel Hill

The clathrin triskelion, composed of three clathrin heavy chains (CLTC) and bound light chains, regulates clathrin-mediated endocytosis and cytoskeletal rearrangement. Alternative splicing generates two Cltc transcripts: a long mRNA that includes a 21-nucleotide micro-exon (exon 31) and a short mRNA that skips it. Inclusion of micro-exon 31 is almost exclusive to striated muscles. In healthy skeletal muscles, the short CLTC isoform is predominant in fetal stages, whereas the long form is the primary isoform in adulthood. Cltc splicing is mis-regulated in skeletal muscle diseases such that the short form is predominant throughout development and adulthood. It was not known, however, whether Cltc mis-splicing disrupts skeletal muscle development. To address this knowledge gap, we forced the skipping of micro-exon 31 in mice using CRISPR/Cas9 gene editing. We found that adult homozygous (HO) mice, which express only the short CLTC isoform throughout development, displayed increased skeletal muscle mass. Despite their large muscles, HO mice exhibited decreased muscle endurance and grip strength. The Cltc splicing transition occurs gradually during the first four postnatal weeks, which is precisely when two specialized muscle components, transverse tubules and neuromuscular junctions, undergo maturation. Transmission electron microscopy revealed that transverse tubules and neuromuscular junctions were strikingly abnormal in muscles of HO mice. In vitro experiments demonstrated that Cltc mis-splicing reduced the formation of flat clathrin plaques, structures that can scaffold proteins at membranes. Our data suggest that Cltc mis-splicing disrupts the anchoring of actin-associated proteins that is critical for membrane remodeling during the maturation of specialized muscle components.

19

Regulation of translation by cell-cell adhesion

Alec D'Alessandro, Kwabena Badu-Nkansah, Terry Lechler
Duke University

Desmosomes are protein complexes at the cell cortex of keratinocytes that ensure cell-to-cell adhesion and structural integrity in stress-exposed tissues like the epidermis and cardiac muscle. In addition, desmosomes also help organize microtubules, actin filaments, and participate in cell signaling. Here, we investigate a non-canonical function of desmosomes as an organizing center for mRNA localization and protein synthesis. Our aim is to understand more about how loss of cell adhesion in keratinocytes can affect other cellular processes that have implications for cortical organization, survival, and stress response.

20

Detecting the Functional Potential of Chromatin-Associated lncRNAs

Q. Eberhard, M. Murvin, M. Calabrese
University of North Carolina at Chapel Hill

Thousands of long non-coding RNAs (lncRNAs) are expressed in humans, however, the functions of very few have been successfully identified. One of these few is XIST, a 17 kb lncRNA that is responsible for repression of the additional X chromosome of female placental mammals and is vital for healthy development. Xist's function is largely attributed to chromatin proximity and ability to recruit RNA binding-proteins (RBPs) that deposit repressive histone markers in cis. Using XIST as a model, we can evaluate other RNAs for similar features and abilities to predict whether they have XIST-like function. One group of RNAs to consider are nascent (or newly transcribed) RNAs as they are tethered to chromatin by RNA polymerase during transcription. Here, I investigate whether these RNAs are likely to have XIST-like function based on sequence properties, expression, stability, degree of chromatin-association, and ability to bind combinations of RBPs similar to XIST. In K562 cells, I identify 4,677 nascent transcripts that are expressed above a TPM threshold of 0.125, are chromatin associated, and are more stable than the least stable functional cis-repressive lncRNA expressed in this cell line, KCNQ1OT1. Further, the RBP-binding profiles of all RNA transcripts were calculated via a combinatorial state HMM and 230 of the expressed, stable, and chromatin-associated nascent transcripts were in the preliminary top 5% of transcripts with RBP profiles similar to XIST, indicating possible XIST-like functionality. These possibly functional lncRNAs with be examined for function in the near future through wet-lab interrogation.

21

Loss of tumor endothelial QKI expression results in pronounced reductions in metastasis and remodeling of the tumor microenvironment

Lincy Edatt, Li Danyan, Jennifer Modliszewski,William David Green, Jared Michael Green, Anderson Carlton, Justin Milner , Andrew Dudley and Chad V. Pecot
UNC-Chapel Hill

The TME encompasses cross talk between different cells and stromal elements within the tumor. Over the years, only a few TME targets have resulted in approved therapeutic agents against highly metastatic lung cancers. Quaking (QKI) is a highly conserved RNA-binding protein that plays critical roles in vascular development during embryogenesis. The objective of our study is to elucidate the role QKI expression in the tumor endothelial cells (TECs) have on the TME during metastatic progression of lung cancer. In a conditionally deleted QKI knock-out model, LLC or JH18-LN2A tumor cell lines was injected either subcutaneously or orthotopically. The QKIKO group showed a 30-60% improved survival, 45-60% less tumor burden and 50% reduction in metastasis and angiogenesis compared to the WT control groups. Spectral flow analysis of tumors from both LLC and LN2A models revealed that WT control mice had higher infiltration of phagocytic immune cells, which showed varied levels of uptake of TEC derived extracellular vesicles(EVs). ScRNA sequencing on TECs and phagocytic immune cells populations that received TEC derived EVs revealed 10-40 fold upregulation of PYDC3, an inhibitor of inflammatory responses. In addition, we observed TEC derived PYDC3 mRNA could be packaged and delivered into immune cells through EVs. Further siRNA silencing, RIP assay, mutation studies and immunofluorescence revealed that QKI, an RNA binding protein could bind to PYDC3 resulting in its degradation and thus activation of inflammasome responses. These findings suggest that targeting tumor endothelial QKI could be novel therapeutic strategy for controlling tumor angiogenesis and metastasis in lung cancer.

22

Human tRNA Splicing Endonuclease Complex Cleaves Diverse RNA Substrates

Jessica Elder, Israa Hissein, Robin Stanley
NIEHS/NIH

tRNAs are essential adapter molecules that ensure the fidelity of protein translation. They carry amino acids to ribosome and position them for incorporation into growing peptide chains via base-pairing of their anti-codon with codons in the template mRNA. tRNAs are transcribed in an immature form (pre-tRNAs) that must be processed before they are functional for protein translation. About 5-10% of human pre-tRNAs encode an intron that disrupts the clover-leaf structure and must be spliced out for the tRNAs to be functional in protein translation. The tRNA Splicing Endonuclease complex (TSEN) is a heterotetrameric protein complex that splices introns from pre-tRNAs. This complex is essential, and mutations are associated with a group of neurodevelopmental/neurodegenerative diseases known as pontocerebellar hypoplasia (PCH). TSEN is composed of two catalytic subunits (TSEN2/TSEN34) and two structural subunits (TSEN15/TSEN54). TSEN2 cleaves introns at the 5' end, while TSEN34 cleaves introns at the 3' end. To date, no kinetic

investigation has been performed on the human TSEN complex. Here we present data for the human TSEN complex that supports the lack of a defined cleavage order but differing kinetics of the two cleavage sites. We also demonstrate that human TSEN can cleave introns in non-canonical positions in tRNA, as well as multiple introns within the same tRNA molecule. We are continuing our kinetic investigation of tRNA intron splicing to elucidate the molecular mechanisms of substrate recognition and cleavage.

23

Detection and validation of snoRNA-transcript interactions with snoCLASH

Brittany A. Elliott, Gene Yang, Alex Choi, Yinzhou Zhu, Christopher L. Holley
Duke University

Small nucleolar RNAs (snoRNAs) have traditionally been linked to ribosomal RNA (rRNA) modification, playing key roles in 2'-O-methylation (Nm) and pseudouridylationâ€"two epitranscriptomic modifications vital for ribosome functionality and protein synthesis. More recently, the functional roles of these RNAs have expanded beyond the processing of non-coding RNAs to include alternative splicing and regulation of transcript expression, indicating a broader biological significance. This expanded functionality suggests that snoRNAs may influence various cellular processes and contribute to the pathogenesis of diseases, including cancer, neurodegenerative disorders, and metabolic diseases. To explore the potential role of Box C/D snoRNAs in directing Nm to transcripts, which could significantly impact gene expression, we employed isolation of snoRNA interactions with rRNA elimination followed by cross-linking ligation and sequencing of hybrids (snoCLASH). We detected 1.8 million snoRNA-RNA chimeric hybrids using an established chimeric read analysis (ChiRA) pipeline. To further evaluate chimeras, high copy number targets and low complexity reads were discarded, resulting in 55,000 hybrids for analysis. These hybrids were analyzed against the POSTAR3 RNA binding protein (RBP) database to identify regions associated with RBP regulation. Among the most significant RBPs associated with snoRNA-target regions, 43% are involved with other RNA epitranscriptomic modifications, despite this class making up only 10% of the total POSTAR3 database. High scoring hybrids meeting this criteria were further validated for changes in transcript Nm and gene expression. These data stand to enrich our understanding of Box C/D snoRNA roles and open new therapeutic avenues for treating snoRNA-related diseases.

24

X-Chromosome linked miRNAs regulate sex differences in cardiac physiology

James Emerson, Frank Conlon
University of North Carolina- Chapel Hill

Males and females exhibit distinct anatomical and functional characteristics of the heart, which make them susceptible to specific disease states. For example, females have a higher prevalence of the dangerous ventricular arrhythmia, torsade de pointes, due to sex differences in cardiac electrophysiology. To understand the molecular basis of sex differences, we previously uncovered differentially expressed proteins in the

mouse heart and determined that post-transcriptional mechanisms control sex differences in protein expression. Therefore, we hypothesized that differential miRNA expression could account for proteomic sex differences. Using small RNA sequencing, we identified miRNAs with sex-differential expression in mouse hearts. Four conserved miRNAs were present in a single locus on the X-chromosome and are expressed at higher levels in females than males. A miRNA from this locus, miR-871, is responsible for decreased expression of the protein Sarcalumenin (SRL) in females. The human orthologue of miR-871, miR-888 co-evolved with the SRL 3' UTR and regulates human SRL. SRL is involved in calcium signaling, and we demonstrated it can account for differences in cardiac electrophysiology between males and females. Using animal models, we confirmed that miR-871 promotes sex differences as miR-871 over-expression mimics the effects of conditional cardiomyocyte-specific Srl null mice by prolonging ventricular repolarization. In contrast, inhibiting miR-871 with an antagomir in females shortened ventricular repolarization times. Using a drug-inducible arrhythmia model, we showed that miR-871 inhibition diminishes the female predisposition to developing torsade de pointes. These findings highlight the importance of sex-differential miRNAs in mediating sex differences in cardiac physiology and disease.

25

Structural analysis of the lncRNA SChLAP1

James P. Falese, Emily J. McFadden, Christopher A. Dâ€™Inzeo, Amanda E. Hargrove
NIEHS

The lncRNA Second Chromosome Locus Associated with Prostate 1 (SChLAP1) was identified in 2013 as robustly associated with metastatic prostate cancer and poor patient outcomes. Despite its established clinical significance, its mechanism of action has remained contested. We hypothesized that structural analysis of SChLAP1 would facilitate identification of functionally important regions. Using in vitro, in cellulo, and ex cellulo SHAPE-MaP, we developed the first experimentally informed structure model for SChLAP1. Using the ΔSHAPE algorithm, we identified putative protein binding regions (reduced reactivity in cells) alongside several regions that were significantly unfolded in cells (increased reactivity in cells). Further analysis of our probing data indicated both well-structured and conformationally heterogenous regions within SChLAP1. We found that one heterogenous region overlapped with a primate-specific retroviral insertion named THE1B, which was located at the 3'end of SChLAP1. Using native gel electrophoresis, we uncovered additional evidence for conformational heterogeneity in this insertion, suggesting that a retroviral insertion within SChLAP1 resulted in a complex RNA structural domain. We then performed biotin pulldowns followed by proteomics of several in-cell protected SChLAP1 substructures to identify candidate SChLAP1-binding proteins. We identified one known and several previously unestablished protein binding partners for SChLAP1, including for a strongly predicted substructure in the THE1B insertion. Further, each of the proteins we identified have previously been implicated in prostate cancer progression. Ongoing work will evaluate the roles for these SChLAP1:protein interactions in prostate cancer progression and also further characterize the conformational landscapes within the 3'end of SChLAP1.

26

Pervasive high-order structure in human pre-mRNAs

Claire Fleurisson, Kevin Weeks, Alain Laederach
UNC Chapel Hill

A first step in gene expression is synthesis of pre-mRNA. Pre-mRNA undergoes intricate processing events before maturing into functional mRNA, including alternative splicing which is observed in most multi-exon genes in humans. Despite its importance, our understanding of pre-mRNA structure remains elusive due to inherent biological constraints, notably its short half-life within cells. In this study, we examine pre-mRNA structure on a large scale. Bacterial artificial chromosome (BAC) clones, containing intact gene copies, were fragmented, amplified, and transcribed into RNA. The structures of these transcribed pre-mRNAs were assessed by chemical probing, including using probing strategies specifically sensitive to tertiary structure. Using a tailored bioinformatics pipeline, we sorted the data by strand and conducted comprehensive analyses to define the structural characteristics of pre-mRNA. Among intronic repetitive elements, especially Alu elements, complex three-dimensional structures are widespread. We also validated our approach by identifying known RNA structures in the E. Coli transcriptome. Our study reveals an unanticipated complex structural landscape of pre-mRNA and hints at multiple potential new RNA-mediated mechanisms of transcriptional regulation.

27

In Search of Iron-Sensitive RNA Binding Proteins

Kwame Forbes, Maria Aleman, Daniel Dominguez, Conner Breen, Grant Goda
UNC CHAPEL HILL

Iron, a vital micronutrient essential for respiration and DNA synthesis, is tightly balanced at the cellular level by iron regulatory proteins 1 and 2 which bind to iron regulatory elements in the untranslated regions of mRNAs involved in iron homeostasis. Surprisingly, this important system regulates merely 13 genes. Encouraged by recent studies having unveiled additional RBPs as iron-sensitive regulators of the transcriptome during iron deficiency (ID) stress, we hypothesize ID evokes a broader post-transcriptional program than previously understood. Suspecting that variations in iron demand across cell types and tissues results in distinct post-transcriptional responses to ID, we treated two prominent cell lines, K562 and HepG2, with an iron chelator to induce ID and performed deep mRNAseq for analysis of differential gene expression and changes in alternative splicing. Preliminary analyses show differences in enriched gene sets between the cell lines, suggesting differences in response to ID. Further, we have compared alternative splicing changes in response to iron chelation to a large RNAseq dataset of ~ 150 RNA binding protein knockdown experiments. Using this approach we have identified a set of candidate iron-sensitive RNA binding proteins. To explore how iron levels changes affect tissues, we utilized a mouse model of dietary iron restriction in c57Bl/6J mice. Upon validating for systemic ID, extracting the liver and performing preliminary RNAseq analysis showed a wide range of gene expression changes. We expect our ongoing analysis to unveil RBPs that show RNA regulation mediated by changes in iron levels.

28

Structural and functional characterization of nidoviral endoribonucleases

Erik J. Daquilanea, Jaci M. Fleming, Kedar Sharma, Robert Dutcher, Mario J. Borgnia, Robin Stanley, Meredith N. Frazier
Department of Chemistry and Biochemistry, College of Charleston

NendoU, a uridine specific endoribonuclease conserved across nidoviruses, processes viral RNA to evade detection by host defense systems. While the enzymatic function is conserved, the oligomeric state necessary for activity does not appear to be conserved. Structural and biochemical studies with arterivirus and delta-coronavirus NendoU have shown the active form to be a dimer, while studies with alpha- and beta-coronavirus NendoU proteins form hexamers. How oligomerization affects NendoU RNA recognition and processing remains poorly understood. We have used SEC-MALS and nuclease assays to confirm gamma-coronavirus NendoU forms hexamers to cleave uridine containing RNA. While gamma-coronaviruses are more closely related to delta-coronaviruses, our data suggests the gamma-coronavirus NendoU is more similar to alpha- and beta-coronavirus NendoU. We have also performed the first molecular biology studies on the torovirus NendoU protein. Toroviruses were originally classified in the coronavirus family, but were recently moved to their own family; thus they represent a more distant virus. We have demonstrated uridine specific nuclease activity and shown by SEC-MALS that torovirus NendoU forms hexamers. Preliminary cryo-em data showed the torovirus NendoU protein forms a different shape than hexameric coronavirus NendoU; however, significant orientation bias has prevented 3D reconstruction. Collectively, these findings advance our knowledge of the conservation of NendoU structure and function across diverse nidoviruses, which will provide new insight into targets for viral inhibition.

29

Alternative splicing of the Snap23 microexon is regulated by QKI and RBFOX2 in a tissue-specific manner and is mis-spliced in striated muscle diseases

Gabrielle M. Gentile, R. Eric Blue, Grant A. Goda, Bryan B. Guzman, Rachel A. Szymanski, Eunice Y. Lee, Nichlas M. Engels, Emma R. Hinkle, Hannah J. Wiedner, Aubriana N. Bishop, Hua Zhang, Xander H. T. Wehrens, Daniel Dominguez, Jimena Giudice
UNC Chapel Hill

The reprogramming of alternative splicing networks during development is a hallmark of tissue maturation and identity. Alternative splicing of microexons (small, genomic regions ~ 51 nucleotides) functionally regulate protein-protein interactions in the brain and are mis-spliced in neuronal diseases. However, little is known about the regulation and function of alternatively spliced microexons in striated muscle. Here, we investigated alternative splicing of the synaptosome-associated protein 23 (Snap23) microexon. We found that inclusion of this microexon is developmentally regulated and tissue-specific, as it occurs exclusively in adult heart and skeletal muscle. The alternative region is highly conserved in mammalian species and encodes an in-frame sequence of 11 amino acids. Furthermore, alternative splicing of this microexon is mis-regulated in mouse models of heart and skeletal muscle diseases. We identified the RNA-binding proteins (RBPs) quaking (QKI)

and RNA binding fox-1 homolog 2 (RBFOX2) as the primary splicing regulators of the Snap23 microexon. These RBPs additively bind downstream of the Snap23 microexon to promote its inclusion, and this regulation can be escaped when the weak splice donor is mutated to the consensus 5' splice site. Our results are the first report of how microexon alternative splicing is controlled during mammalian striated muscle development.

30

Navigating the Roles of Cis and Trans Environment on Species-Specific Gene Expression

Gilbert Giri, Daniel Dominguez
University of North Carolina

Immediate cis and trans environment of a gene play a major role in regulating gene expression. Closely related species can have similar DNA composition yet significantly different transcriptome driven by differential regulation of genetic elements. Here, we investigate how much of this differential regulation is driven by cis vs trans elements. Using a mouse-human hybrid cell line, we looked at the differential expression and splicing of human and mouse one-to-one orthologs. We found a significant correlation between the expression of mouse and human ortholog gene pairs, and their sequence and functional conservation. Additionally, we identified novel transcripts that were not annotated in either mouse or human systems which suggests an interspecies mRNA regulatory mechanism unique to the hybrid cells. Using RNA-crosslinking experiments, we will unravel whether mRNA motif conservation or amino acid sequence conservation is primarily driving these novel interactions between mRNA and proteins. Overall, our work disentangles the effect cis and trans environment have in regulating gene expression adding a new dimension to the evolutionary and functional development of a species.

31

Hyperactivation of an RNA Binding Protein by Cancer-Associated Mutation

Grant Goda, Gilbert Giri, Bryan Guzman, Md Siraj, Raquel Gaither, Rachel Cherney, Jose Martinez, Corey Davis, Michael Sullivan, Conner Breen, Luisa Escobar-Hoyos, Maria M. Aleman, Daniel Dominguez
UNC-CH

RNA binding proteins (RBPs) are a class of regulatory factors that directly interact with RNA to control its fate; therefore, they are master regulators of gene expression. Mutations in RBPs are associated with cancers and neurodegenerative diseases. In colorectal adenocarcinoma and Burkitt's lymphoma, RBP poly c binding protein 1 (PCBP1) is recurrently mutated. PCBP1 is essential for life and has been shown to control developmental processes by binding C-rich RNA sequences to regulate alternative splicing, mRNA stability, and translation. Recently the PCBP1 (L100Q) hotspot mutations were shown to increase PCBP1 RNA association with no known mechanism. Here, we detail the molecular mechanism by which the L100Q mutation elicits a structural change resulting in homo-oligomerized and hyperactive PCBP1 with increased RNA affinity.

We further show L100Q PCBP1 has altered cellular localization which shifts RNA regulatory control from nuclear splicing to cytoplasmic RNA regulation. Owing to this nuclear to cytoplasmic shift RNA regulation, RNA sequencing, ribosome profiling, and proteomics revealed changes in alternative splicing and translation in cells harboring PCBP1 mutations. We then probed L100Q PCBP1 bound genes and discovered enrichments of bound RNAs with roles in epithelial to mesenchymal transition and wnt-signaling pathways crucial for colon cancer progression. Accordingly, cells harboring L100Q PCBP1 displayed increased migration and invasion. Finally, we show PCBP1 mutations in Burkitt lymphoma and colorectal cancers are functionally distinct hinting at multiple ways for PCBP1 to drive cancer. This work sets a precedence for understanding RBP hyperactivation by mutation and the reinforces importance of RNA mis-regulation in cancer.

32

Uncovering the architecture and function of the human rixosome complex in ribosome synthesis

Jacob Gordon, Alan J. Warren, and Robin E. Stanley
NIEHS/NIH

Biogenesis of eukaryotic ribosomes is essential for life. This intricate pathway requires processing of pre-ribosomal RNA to remove four spacer regions (5' ETS, ITS1, ITS2, 3' ETS) and generate the mature ribosomal RNAs. Removal of the ITS2 spacer RNA is initiated by an endoribonuclease-kinase complex RNase PNK (LAS1L-NOL9). RNase PNK is further integrated within a conserved mega-complex called the Rixosome which aids in coordinating ITS2 processing. Additional proteins in the Rixosome (PELP1-WDR18-TEX10-SENP3-MDN1) play other distinct enzymatic roles in ribosome synthesis. Members of the Rixosome are linked to human diseases including cancer and neurological disorders, underscoring their importance. How does the Rixosome coordinate ITS2 processing and other steps of ribosome synthesis? We set out to characterize the architecture of the human Rixosome and reveal how RNase PNK is integrated to facilitate ITS2 processing. We successfully reconstituted the human Rixosome and determined that PELP1 modularly scaffolds each Rixosomal enzyme in distinct structural regions. A non-catalytic domain within the endoribonuclease of RNase PNK (LAS1L) was determined to tether the Rixosome scaffold to the catalytic domains of RNase PNK. We used cryo-EM to determine structures of the Rixosome's scaffolding core and RNA processing domains, RNase PNK. Our human RNase PNK structure reveals the conserved ribonuclease-kinase domains and mammalian specific features, such as an enlarged RNA binding cleft. Our data provide a composite understanding of the human Rixosome's architecture and how the various enzymatic domains are assembled. This begins to elucidate how the Rixosome might coordinate many steps in ribosome synthesis, including ITS2 spacer removal.

33

Insights into Coronavirus Replication Regulated by Poly(A) Tail Modifications

Ankit Gupta1, Yin Li[1], Marine Baptissart[1], Shih-Heng Chen[2], Brian N. Papas[3], Negin P. Martin[2], and Marcos Morgan[1]

[1]Reproductive and Developmental Biology Laboratory,
[2]Viral Vector Core Facility,
[3]Integrative Bioinformatics Core Facility, National Institute of Environmental Health Sciences, National Institutes of Health, Durham, NC NIH-NIEHS

Coronaviruses, including the mouse hepatitis virus (MHV), are positive-strand RNA viruses with 3' polyadenylated genomes and subgenomic transcripts. The lengths of their viral poly(A) tails change dynamically during infection through mechanisms that remain poorly understood. We developed a splint-ligation method to measure poly(A) tail lengths and terminal modifications, offering a novel approach to investigate these modifications. Upon infecting mouse fibroblast 17-CL1 cells with MHV, a member of the Betacoronavirus genus, we observed that viral RNA poly(A) tails are continuously shortened, and these short tails (~22 nucleotides) are selectively uridylated by the mammalian terminal uridylyl-transferases TUT4/7. Notably, TUT4/7 depletion enhances viral replication, leading to increased viral RNA loads. Building on these findings, we further examined the role of TUT4/7 using an ex vivo macrophage model and an in vivo mouse model. Direct RNA sequencing revealed that viral RNA poly(A) tails shorten more rapidly in macrophages than in fibroblasts, correlating with accelerated viral replication. In macrophages, only short poly(A) tails (~22 nucleotides) undergo uridylation, unlike the longer tails (~44 nucleotides) in fibroblasts. Collectively, our findings suggest a mechanism in which a fraction of MHV transcripts is uridylated by TUT4/7 and subsequently targeted by cellular decay machinery. Studying viral RNA processing in different cellular systems can provide a clearer understanding of how viral RNA is processed during infection. This knowledge could lead to the development of novel strategies to combat COVID-19 and other coronaviruses.

34

Protein Disorder and RNA Binding

B.B. Guzman, G.A. Goda, D.I. Dominguez
UNC-Chapel Hill

RNA binding proteins (RBPs) interact with and tightly regulate the fate of messenger RNAs. Despite their essential role in normal biology, how RBPs bind their RNA targets is not well understood. Generally, RBPs are known to bind short RNA motifs through well-folded RNA binding domains (RBDs). However, many RBPs also contain low complexity domains (LCDs), which are protein domains characterized by disorder and overrepresentation of one or two amino acids. The biological function of these disordered domains is poorly understood, and their ability to bind RNA with any specificity, or at all, is an open question. Furthermore, mutations in RBP disordered regions are associated with human disease. Here, we aim to understand the role of disordered domains in RBP function. We focus on HNRNPR, an RBP with three RBDs and an LCD enriched in arginine and glycine. HNRNPR regulates RNA localization in neurons and mutations to the disordered domain are associated with neurodevelopmental disorders. We define the specificity landscape of HNRNPR RBDs and show that the LCD also interacts with guanine-rich structured RNAs with remarkable specificity. We detail the molecular mechanism by which the disordered domain binds to specific RNA. We further probe how disease-associated mutations in HNRNPR's LCD impact RNA binding, localization, and dynamics in

cells. Finally, we detail the complex ways in which HNRNPR selects RNA targets which involves both RBDs and the disordered domain. This work highlights the complexity of protein-RNA interactions and motivates future work to consider LCDs as a functional RNA binding domain within RBPs.

35

The contribution of isoform diversity to cell fate determination in the mammalian retina

Sarah E. Hadyniak and Jeremy N. Kay
Duke University

During development, progenitor cells adopt a spectrum of cell fates via differential expression of cell-type-specific molecules. While many studies have focused on gene-level expression differences during embryogenesis, little is known about how mRNA isoforms contribute to cell fate. We hypothesize that isoform diversity is responsible for generating the necessary number of recognition molecules used by the ~100 neurons in the mouse retina to provide cell subtypes with unique molecular identity. All homotypic neurons in the retina tile their dendritic projections to prevent blind spots from occurring in visual processing. Here, we aim to discover the molecules responsible for this phenomenon in a dataset of isoform profiles from the mouse retina. We are establishing single cell, long-read sequencing in the mouse retina to describe isoform diversity within 15 bipolar cell subtypes. We plan to integrate our bipolar cell dataset with whole retina time course data to explore changes in isoform expression as a developmental trajectory from progenitor to specific bipolar cell type. Future directions of this study aim to identify alternatively spliced molecules involved in cell type-specific neuronal repulsion. Ultimately, these data will allow us to define the contribution of mRNA isoform diversity to cell fate.

36

Precursor RNA structural patterns at SF3B1 mutation sensitive cryptic 3' splice sites

Austin Herbert, Abigail Hatfield, Alexandra Randazza, Valeria Miyamoto, Katie Palmer, Lela Lackey
Clemson Center for Human Genetics

A hallmark of SF3B1 mutation is an increased use of upstream cryptic 3' splice sites (C3SS) in a broad number of genes, a finding that is recapitulated across multiple isogenic and patient cell types. Here, we investigate the precursor RNA sequence and structural properties of 73 C3SS utilized in SF3B1 mutants defined from bulk-RNA sequencing data across different patient and isogenic SF3B1 MT cell lines. Occurring at a mean distance of 19 base pairs upstream, SF3B1 sensitive C3SS show no significant difference in length compared to C3SS utilized in SF3B1 WT cells. Both SF3B1 sensitive and control C3SS contain a distinct polypyrimidine tract and a strong consensus AG splice site motif. A combination of experimental and in-silico methods revealed no significant difference in SHAPE reactivity or predicted base accessibility between SF3B1 sensitive C3SS AG motif, their canonical AG counterparts, or C3SS AG used in a SF3B1 wildtype background. Independent from cryptic or canonical context, the adenine of the 3' AG motif displays

significantly increased SHAPE reactivity compared to flanking nucleotides. We will further investigate the presence of structural variance at the branch point or splicing regulatory motifs that may drive splice site choice. Our observations yield insights into the structural context of cryptic and canonical 3' splice site motifs. Finally, we will use these findings to investigate a generalized mechanistic role of precursor RNA structures in contributing to regulation of cryptic 3' splicing.

37

Kinetic Investigation of the tRNA Splicing Endonuclease Complex (TSEN)

Israa Hissein, Jessica Elder, Robin Stanley
NIEHS/NIH

The removal of pre-tRNA introns is crucial for the functional maturation of tRNAs in all eukaryotic organisms. In humans, this process is mediated by the tRNA splicing endonuclease complex (TSEN), which comprises four subunits: TSEN2, TSEN15, TSEN34, and TSEN54. Mutations in TSEN have been linked to neurodegenerative diseases such as pontocerebral hypoplasia (PCH), and deletion of any of these core proteins results in embryonic lethality. We aim to investigate the kinetics of tRNA splicing by the TSEN complex by analyzing its activity on various pre-tRNA substrates. In addition, we aim to determine if the TSEN complex exhibits a splice site order preference given that two separate reactions occur involving cleavage at the 5' and 3' splice sites by the nucleases TSEN2 and TSEN34, respectively. The heterotetrameic TSEN complex was expressed in E. coli using a polycistronic vector and then purified by affinity and size exclusion chromatography. A series of endoribonuclease assays were carried out to test TSEN's activity across a range of pre-tRNA substrate concentrations measuring the rate of tRNA cleavage at each concentration. The samples were run on a TBE-Urea gel and the data was analyzed to elucidate TSEN's splicing efficiency and specificity. We observed that full-length pre-tRNA-Arg, exhibits tighter binding to TSEN compared to a minimal RNA substrate. Furthermore, within the TSEN reaction our results support that, the 3'end of the tRNA is cleaved more rapidly than the 5' end. These findings reveal that TSEN displays differential cleavage rates based on substrate structure and site order.

38

High Resolution Interrogation of Flavivirus Protein Interactions with Host Transcriptomes

Rose Homoelle, JohnCarlo Kristofich, Hannah Schmidt, Stacey Horner, Christopher Nicchitta
Duke University

Flaviviruses are single-stranded RNA arboviruses that have emerged as prevalent human pathogens over the past several decades, with half of the world's population currently at risk for the mosquito-borne dengue virus (DENV). Conserved between flaviviruses is an ~11 kb RNA genome that encodes a single polyprotein which is translated, translocated, and proteolytically processed on the endoplasmic reticulum (ER) membrane. Prior work in our lab has found that DENV hijacks translation machinery on the ER membrane to support viral propagation and suggests that understanding the flavivirus-ER interaction landscape may reveal novel

strategies for antiviral therapeutic development. LEAP-RBP is a recent technique developed in our lab for the selective and quantitative recovery of UV-crosslinked RNA-protein complexes. LEAP-RBP circumvents limitations in data interpretation of other RBP enrichment methods by providing high signal-to-noise (S/N) by virtue of its high specificity and yield. Using LEAP-RBP, we were able to quantify the RNA-binding activity of the flavivirus proteins encoded by zika virus (ZIKV) to inform future hypotheses regarding how viral proteins intersect with host biology via RNA interactions.

39

TurboID Identifies Key Interacting Proteins of the CCR4-Not Complex in Mouse Embryonic Stem Cells

Brad Lackford, James Ward, Xiukun Wang, and Guang Hu
NIEHS

The CCR4-Not complex has been identified as an essential regulator self renewal and pluripotency in mouse and human embryonic stem cells (ESC). The Cnot complex is the primary deadenylase in the cell, but has been shown to play a role in many post transcription regulatory processes. Teasing apart its molecular functions to understand how self renewal and pluripotency are maintained in ESCs has been confounded by the many roles it plays. To that end, we undertook biotin proximity labeling using TurboID to identify interacting proteins in order to isolate individual molecular functions of the CCR4-Not complex. We have identified 174 high confidence hits across 10 different subunits of the complex. We have validated several of these hits (Atxn2l, hnRNPL, Upf1, and PabpC1) as essential regulators of the pluripotent state via conditional KO cell lines. Of note, Atxn2l appears to regulate the translation efficiency of pluripotency factors. Deletion of Atxn2l quickly leads to loss of pluripotency factors Esrrb and Klf4 on a protein level, resulting in up regulation of ectoderm transcripts and differentiation.

40

Systematic study of RNA G-quadruplex folding stability

Yue Hu, Bryan Guzman*, Justin Martyr, Daniel Dominguez
UNC-CH

RNA G-quadruplexes (rG4s) are four-stranded RNA structures that form from G-rich sequences. In mammalian transcriptomes, rG4s are overrepresented and have been implicated in regulating various aspects of gene expression and RNA processing, from transcription to mRNA decay. However, uncertainty remains regarding their folding in vivo, with some studies suggesting that most rG4sremain unfolded.
To investigate rG4 stability, we designed a controlled pool of natural rG4 sequences that incorporates biologically relevant features such as varied folding strengths and topologies, evidence of RNA-binding protein interaction in vivo, prior identification in rG4 profiling assays, and presence in different transcript regions (e.g., UTRs versus CDS). With this pool, we optimized an RT-stop assay coupled to quantitative sequencing to detect rG4 folding, leveraging the fact that strong rG4s inhibit reverse transcription. This method enabled us to measure the folding strength of thousands of rG4s in parallel with a dynamic range

spanning several orders of magnitude. Additionally, we introduced folding modifiers such as potassium, lithium, and 7-deaza-GTP substitutions to assess their effects on rG4 stability. Our results demonstrate that factors such as the number of guanines, spacing length, and protein interactions significantly influence the strength of rG4. Furthermore, for proteins that are known to interact with rG4s in vivo and in vitro, we show that binding strength is proportional to rG4 folding strength, highlighting the interplay between RNA structure and protein binding. This assay offers a powerful method for understanding rG4 folding mechanisms and their role in gene regulation.

41

ZATT (ZNF451) is a TOP2 Regulated Ribonuclease

Ananda A Jaguva Vasudevan, Amanda R. Riccio, and R. Scott Williams
National Institute of Environmental Health Sciences

Zinc-finger Associated with TDP2 and TOP2 (ZATT) is a multifunctional protein participating in the repair of TOP2 mediated DNA damage, replication, and transcription. These nuclear activities have been linked to its intrinsic N-terminal E3 SUMO2/3 ligase activity. ZATT also contains an array of C2H2 class Zinc finger motifs (Znf), and a C-terminal PIN/NYN family nuclease domain of unknown function. We report that purified ZATT harbors intrinsic ribonuclease (RNase) activity. The broad Mg2+ and Mn2+ dependent RNase activity prefers RNA hairpin>RNA/DNA>dsRNA>ssRNA but fails to cleave ssDNA or dsDNA substrates in vitro. ZATT displays an RNASEH1-like activity, and efficiently cleaves embedded ribonucleotides in DNA harboring 4/more ribonucleotides. A minimal active nucleolytic core encompasses both the Znf repeat and the PIN domain, while the intact PIN core's activity was much less than the Znf-PIN. To define the basis for ZATT catalytic activity, we determined a 2.3 Ã… resolution crystal structure of ZATT-PIN that unveils an extensive dimeric interface. SEC-MALS demonstrates that ZATT exists as a stable dimer in solution, and mutating the interface abrogates RNase activity. TOP2 directly binds and inhibits ZATT ribonuclease activity. Structural analysis of the ZATT-TOP2 complex implicates a TOP2-mediated allosteric switch that promotes restructuring and occlusion of the ZATT active center. Consistent with a proposed role for ZATT in processing cellular RNA/DNA hybrids, we find that FL-ZATT co-immunoprecipitates with the RNA/DNA hybrid-specific anti-S9.6 antibody. Together, results from biochemical and structural analysis point to possible roles for ZATT in the nucleolytic regulation of complex nucleic acid structures like R-loops, which left unresolved are a critical threat to genome integrity.

42

Blunting of the estrogen response dynamics may underlie the relative insensitivity of mammary luminal epithelium to endocrine therapy in Luminal B breast cancer

Laura G Kammel, Xin Xu, Robert N Wine MS, Jackson Hoffman, Fleur Chapus, Maria Sifre, Carl Bortner, Jianying Li, Erin Quist, and Joseph Rodriguez
NIEHS

Luminal B breast cancers are characterized by increased proliferation, lower sensitivity to endocrine therapy, and worse clinical prognosis compared to the Luminal A subtype. While several potential drivers of this more aggressive phenotype have been investigated, no definitive mechanism underlying the relative endocrine insensitivity has been described. Tumors derived from the MMTV-PyMT (MP) mouse model resemble the expression profiles of Luminal B breast cancers. We used this model to investigate if the relative endocrine insensitivity of Luminal B tumors could reflect an alteration in transcriptional regulation and chromatin accessibility of estrogen receptor alpha (ER) targets in mammary epithelial cells (MECs). MP mice were crossed to an ER-reporter strain (WT) in order to isolate ER+ MECs. Single-cell RNAseq confirmed enrichment of ER+ MECs and differential analysis revealed a broader induction of ER targets in ER+ luminal MECs from WT compared to MP mice. Curiously, the expression of transcription factors directly involved in mediating the estrogen response, including Esr1, Gata3, and Myc, were altered in WT but not MP cells after estrogen treatment. To explore whether differential gene expression could be explained by differences in the chromatin landscape, we performed ATACseq using the same induction paradigm. We found that while many regions with an estrogen-induced gain in accessibility were shared between genotypes, overall, there was a larger gain in accessible sites that were unique to WT mice. Together, our results suggest that while hyperplastic ER+ MECs are responsive to estrogen, induction is less dynamic and less robust than in their normal counterparts.

43

Structural and functional characterization of torovirus EndoU

Jaci M. Fleming, Aasim Khan, Patrick O'Reilly, Ashley Hubbard, Kedar Sharma, Robert Dutcher, Mario Borgnia, Robin Stanley, Meredith N. Frazier
College of Charleston

Toroviruses (ToVs) are enveloped single stranded RNA viruses in the Nidovirales order, which also includes coronaviruses. ToVs cause disease in animal industries worldwide, leading to economic losses; additionally, there have been cases of ToVs causing mild disease in humans. Many non-structural proteins (nsps) involved in viral replication and transcription are conserved across nidoviruses; therefore, understanding the structure and function of these proteins will help identify possible therapeutic targets and inform rational drug design. One viral protein of interest, ToV nsp12, is an endoribonuclease with specificity for Us (EndoU). Nidoviral EndoU family members regulate viral RNA to evade host immune systems. Bioinformatic analyses were performed to identify the protein boundaries for bovine ToV nsp12; it was then successfully overexpressed and purified using affinity and size exclusion chromatography. Cryo-EM data sets have been collected to determine a high-resolution structure but show significant orientation bias. SEC-MALS analysis revealed nsp12 oligomerizes into a hexamer, similar to the coronavirus EndoU nsp15. Nuclease assays confirm nsp12 has EndoU activity; substrate preference studies are continuing. This research will provide insight into the evolution of EndoU through comparison of ToV nsp12 nuclease activity, specificity, and structure to other nidoviral EndoU proteins.

44

In vitro characterization of RDE-3, a poly(UG) polymerase required for RNAi maintenance in C. elegans

Bradley P. Klemm, Andrew P. Sikkema, Emily R. Schugardt, Traci M. Tanaka Hall
National Institute of Environmental Health Sciences

RNAi-deficient 3 (RDE-3) mediates transgenerational siRNA production and transposon silencing in C. elegans. RDE-3 is a poly(UG) polymerase that adds tails of alternating U and G nucleotides to targets of RNAi. The so-called pUG tails fold into a novel quadruplex structure which recruits RNA-dependent RNA polymerase to generate the second strand, thus generating additional small interfering RNAs targeting the original transcript. We have begun to evaluate the substrate specificity and kinetic parameters of RDE-3 in vitro. We find that RDE-3 NTP specificity (kcat/KM) is affected by the nucleotide at the 3' end of the RNA. For an RNA with a 3'-G nucleotide, RDE-3 is highly specific for UTP over GTP. For 3'-U RNA, RDE-3 conversely favors the addition of GTP over UTP. For RNAs with 3'-A or 3'-C, RDE-3 displayed minimal preference for either UTP or GTP. These results confirm that RDE-3 has the intrinsic NTP specificity required for non-templated polymerization of pUG tails. We have additionally interrogated the RNA specificity of RDE-3, finding that the intrinsic affinity is weak. We developed an assay to monitor processive pUG tail extension in vitro. Even after 24-hours, the pUG tails reached a maximum of 10-13nt, fewer than the minimum 23nt required for the pUG quadruplex structure. RDE-3 extension displays high fidelity when the UTP and GTP are relatively balanced. However, when either UTP or GTP are in excess, we observe increased UU and GG misincorporations.

45

Machine Learning Predictions of RNA-Protein Interactions to Infer lncRNA Function

Megan B Kratz and Keriayn N Smith
UNC

Researchers have extensively studied and manipulated thousands of messenger RNAs (mRNAs), proteins, and other vital molecules across many human tissues. Long noncoding RNAs (lncRNAs) hold potential regulatory importance, with their genes outnumbering protein-coding genes in the human genome. However, lncRNA functions remain largely unknown. One major challenge in understanding lncRNA function is the lack of broad sequence conservation, both across species and among lncRNAs with similar functions within a species. Because lncRNA functions are enacted through their interactions, we employed a machine learning (ML)-based approach that focuses on lncRNA interactions with RNA binding proteins (RBPs). This predictive method effectively infers conserved and functionally relevant interactions, even among evolutionarily distant species, thereby shedding light on sequence-function relationships for these enigmatic RNAs.

46

RNA Origami-Based Thrombin Inhibition: A Novel Anticoagulant System with Rapid, Reversible Control

Abhichart Krissanaprasit, Thomas LaBean, David Harrell
NC State

Anticoagulants are critical for preventing thrombosis during medical procedures, yet current therapies are limited by the absence of fast and specific reversal agents, increasing the risk of excessive bleeding. Heparin, though widely used, is associated with variable patient responses and adverse side effects, underscoring the need for safer, more controllable alternatives to improve patient safety. We present an innovative RNA origami-based anticoagulant with fast action reversal agent (HEX01/02 system). Multivalent RNA aptamers target thrombin on RNA origami (HEX01) showed more potent and efficient thrombin inhibition than single aptamer and heparin in vitro and in vivo. The multivalency of the RNA origami structure further enhances binding specificity and anticoagulant efficacy, representing a significant advancement in anticoagulant design. To meet the critical need for rapid reversibility, we developed HEX02, a single-molecule DNA antidote capable of reversing HEX01's anticoagulant effects within just 30 seconds in vitro and 5 minutes in vivo. In murine models, HEX01 demonstrated effective, dose-dependent anticoagulation, while HEX02 enabled swift and specific reversal. Biodistribution studies confirmed HEX01's primary accumulation in the liver, with no observed cytotoxicity, hemolysis, or immune activation. RNA origami provides a versatile platform for RNA-based therapeutics, offering the potential for safe, responsive therapies with broad clinical and surgical applications. Beyond anticoagulation, this platform could be applied to areas such as targeted drug delivery and gene therapy, demonstrating the expansive potential of RNA origami technology.

47

The Diverse RNA Structures of Alu Elements

Abigail C. Lehr, Scott R. Allen, Alain Laederach, Kevin M. Weeks
UNC Chapel Hill Department of Chemistry

Alu elements are a family of largely inactive retrotransposons that comprise roughly 13% of the human transcriptome, encoded primarily in intronic and untranslated regions. Alu elements assume diverse roles in both normal biology and in disease states. Despite the known diversity in Alu sequence and function, we lack data-driven, robust structural information beyond a single widely-cited canonical structure. Here we model the RNA secondary structures of previously uncharacterized Alu elements, using DMS probing coupled with mutational profiling (DMS-MaP) measured for both simple transcripts and in cells. Our findings suggest that Alu elements do not systematically conform to the expected RNA secondary structure; instead, each Alu element appears to adopt a unique structure. Moreover, Alu elements routinely interact with their native flanking sequences, in contrast to the assumption that the RNA structure of Alu elements is self-contained. Our structural data expand the known RNA structure space assumed by Alu elements and suggest that diversity in RNA structures underlies the diverse biological functions attributed to Alu elements.

48

hmSEEKR: a Hidden Markov model that identifies regional similarities based on base on kmer profile

Shuang Li, Daniel Sprague, Mauro Calabrese
University of North Carolina at Chapel Hill, Department of Pharmacology

Functional analysis of nucleotide species is hindered using traditional linear alignment tools, such as BLAST, for sequences that are poorly conserved, evolve rapidly, and hence rarely harbor long stretches of linear similarity, such as long noncoding RNAs (lncRNAs). Although lncRNAs have been shown to regulate fundamental cellular processes, the majority remain functionally uncharacterized because of the inability to detect recurrent relationships among sequences. Previously, we developed and updated SEquence Evaluation through k-mer Representation (SEEKR), which is a method of sequence comparison that utilizes sequence substrings called k-mers to quantify non-linear similarity between nucleic acid species. Here we describe the development of a new tool called hmSEEKR, which is a SEEKR based Hidden Markov model that identifies regional similarities base on kmer profile between any set of sequences regardless of linear similarities. hmSEEKR can locate regions among any given sequences (the whole transcriptome) that bears high similarities to a query sequence (Xist repeat A), therefore facilitating more accurate search for functional similar elements. hmSEEKR is implemented with easy application for users with little to no coding experience as well as professional bioinformaticians.

49

A novel B7H3-targeting macrocyclic peptide conjugate efficiently delivers siRNAs to cancer cells and tumor endothelium

Danyan Li, Lincy Edatt, Jillian Perry, Avery Huber, William David Green, Jarred Michael Green, Justin Milner, Albert Bowers and Chad V. Pecot
UNC Chapel Hill

Short-interfering RNAs (siRNAs) can treat a wide range of human diseases from viral outbreaks to cancer, however, targeted siRNA delivery remains a significant challenge. Recently, significant efforts have focused on developing novel delivery technologies for siRNAs, targeting tumors and other extrahepatic tissues. B7-H3 (CD276),a B7 family of immune checkpoint proteins, is highly expressed in cancer cells and tumor endothelial cells (TECs). We employed an mRNA display campaign against the hB7-H3 protein and discovered a macrocyclic binding peptide, JP5. Using surface plasmon resonance, we confirmed specific binding of JP5 to both human and murine B7-H3 with a KD of approximately 770nM and 600nM, respectively. Characterization of protease stability showed that greater than 80% of the peptide remains intact after incubation with 50% serum for 24h. Further, our team conjugated JP5 to chemically stabilized siRNAs labeled with a Cy5 fluorophore to observe biodistribution of siRNAs, including cancer cells and TECs. Mice were implanted with A431 WT or B7H3 KO cells, and upon xenograft tumor formation, mice were given PBS (control) or JP5-siRNA-Cy5 (n=5 mice/group). Gross biodistribution by IVIS imaging was found

in the tumors, skin, bowel, liver and kidney, with minimal accumulation in the heart and lungs. Flow cytometry and immunofluorescence revealed that JP5 conjugation directed siRNAs to ~80% of TECs and to 40-60% of cancer cells, with a significant reduction in B7H3 KO cells. In conclusion, the JP5 macrocycle delivers siRNAs to B7H3-expressing cancer cells and TECs in the tumor microenvironment, paving a path forward for sophisticated cancer therapeutics.

50

IDEA: Integrating Structure and Sequence to Predict Protein-Nucleic Acid Interactions

Yafan Zhang, Eduardo Cisneros, Irene Silvernail, Xingcheng Lin*
North Carolina State University

Sequence-specific nucleic acid recognition underpins essential processes in gene regulation, yet experiment-independent methods for simultaneous prediction of nucleic acid recognition sites and their binding affinity remain limited. Here, we introduce the Interpretable protein-DNA/RNA Energy Associative (IDEA) model, an interpretable residue-level biophysical model capable of predicting binding sites and affinities of nucleic-acid-binding proteins without relying on experimental binding data. By integrating the structures and sequences of known protein-nucleic acid complexes into an optimized energy model, IDEA enables direct interpretation of the physicochemical interactions among individual amino acids and nucleotides. Using transcription factors as examples, we demonstrate that this energy model accurately predicts genomic DNA recognition sites and their binding strengths. Further, by leveraging deep-learning structure predictors, we show IDEA can precisely capture protein-single-stranded nucleic acid specificities, including RNA and single-stranded DNA. Additionally, we incorporate IDEA into a coarse-grained simulation framework that quantitatively captures the absolute protein-nucleic acid binding free energies. Overall, IDEA provides an integrated computational platform that alleviates experimental costs and biases in the assessment of nucleic acid recognition and can be utilized for mechanistic studies of various genetic and epigenetic processes.

51

Desmosome-dependent enrichment of RISC at the cell cortex

Sophia Link, Alec D'Alessandro, Daniel Hlavaty, Kwabena Badu-Nkansah, Terry Lechler
Duke University

Post-transcriptional regulation allows cells to rapidly modulate gene expression patterns in response to stimuli. RNA-binding proteins can affect this regulation by sequestering mRNAs within non-membranous compartments of the cytoplasm. The cell cortex of keratinocytes is enriched for many such RNA-binding proteins, including ribosomes, translation initiation factors, and the RNA-induced silencing complex (RISC), as well as, select mRNAs. Upon perturbation of the desmosome, RISC moves away from the cortex and RISC-associated mRNAs become upregulated in the translatome. This suggests that the cortical RISC population sequesters mRNAs that can be released upon wounding to meet the immediate needs of the

damaged tissue. In order to test this hypothesis, it is necessary to identify the molecular mechanism of RISC cortical localization to allow for rational disruption. We have determined cytoplasmic constructs of the tail domain of the desmosomal protein, desmoplakin, are sufficient to recruit RISC while constructs with just the first of three plakin repeat domain, PrdA, from the tail are insufficient. Additionally, the RISC with phosphorylation on serine-387 of the catalytic component protein is present in the cytoplasmic pool; this post-translational modification favors sequestration over degradation of RISC-bound mRNAs. These data represent an important first step in determining the mechanism of RISC cortical localization by identifying the minimum component of desmoplakin sufficient for recruitment. Additionally, the enrichment of RISC phosphorylated at serine-387 on Ago2 at the cortex and translation of select RISC associated mRNAs upon wounding support the model of RISC sequestering mRNAs for release upon desmosome perturbation.

52

Highly Programmable RNA Scaffolds for Multivariate CRISPR/Cas-based Effector Recruitment

Barun Mahata, Alan Cabrera, Daniel R. Reed, Yannie Guo, Suchir Misra, Hailey Szadowski, Jing Li, Alex Ma, Selvalakshmi Selvaraj Anand, Guy Bedford, Sunghwan Kim, Amaury F. Bittar, Sandy Luong and Isaac B. Hilton
Rice University

Rational engineering of RNA has enabled diverse applications ranging from COVID vaccines to improved guide RNAs (gRNAs) for CRISPR/Cas systems in vivo and in vitro. For instance, chemical modifications to gRNAs are often required for genome editing in cell and gene therapy contexts, as well as CRISPR activation (CRISPRa) strategies. Further, gRNAs are frequently modified to incorporate stem loops that can, in turn, recruit RNA-binding proteins fused to an array of enzymatic and/or regulatory domains. While powerful, these current platforms offer limited stoichiometric and spatiotemporal control and often interfere with gRNA function and/or editing efficacy. Here, we developed a designer multi-stem-loop RNA intermediated effector recruitment platform, called DRIMER, that permits robust stochiometric and spatiotemporal control over effector recruitment to targeted human loci when used with CRISPR/Cas systems. We show that DRIMER can recruit up to four unique effector domains to human loci and that these effectors can be functionally distinct and recruited in user-defined combinations. Using this platform, we demonstrate that combinations of key transcription factors and epigenetic modifiers harbor powerful synergies at endogenous human loci. Further, we demonstrate that DRIMER-mediated recruitment can be precisely tuned by incorporating chemically controlled riboswitches or optogenetically regulated stem-loops. Overall, the DRIMER platform is a highly programmable CRISPR-based system that allows stoichiometric and combinatorial effector recruitment using designer, structured RNAs for the precise modulation of human gene expression and other genomic activities, which is an attractive new capability for a wide range of applications spanning complex epigenome editing as well as gene and cell therapies.

53

not3 is required for male germ cell development and spermatogonial stem cell maintenance

Safia Malki, Qing Chen, Guang Hu
NIEHS

The foundation of spermatogenesis and lifelong fertility is provided by spermatogonial stem cells (SSCs). SSCs divide asymmetrically to either replenish their numbers (self-renewal) or produce undifferentiated progenitors that proliferate before committing to differentiation. However, regulatory mechanisms governing SSC maintenance is poorly understood. Here, we show that the CCR4-NOT mRNA deadenylase complex subunit CNOT3 plays a critical role in maintaining spermatogonial populations in mice. Cnot3 is highly expressed in undifferentiated spermatogonia, and its deletion in spermatogonia resulted in germ cell loss and infertility. Single cell analyses revealed that Cnot3 deletion led to the de-repression of transcripts encoding factors involved in SSC differentiation, including those in the glutathione redox pathway that are critical for SSC maintenance. Together, our study reveals that CNOT3 – likely via the CCR4-NOT complex – actively degrades transcripts encoding differentiation factors to sustain the spermatogonial pool and ensures the progression of spermatogenesis, highlighting the importance of CCR4-NOT-mediated post-transcriptional gene regulation during male germ cell development.

54

m6A Regulates Breast Cancer Proliferation and Migration through Stage-dependent Changes in Epithelial to Mesenchymal Transition Gene Expression

Kyle D. Mansfield and Mohammed G. Dorgham1
Brody School of Medicine, East Carolina University

Recent studies have implicated the mRNA modification N6-methyladenosine (m6A) in breast cancer progression but often with conflicting data as to whether it is promoting or inhibiting the process. To gain a clearer picture of the role of m6A in breast cancer progression, we used the MCF10 cell culture model to mimic stages of progression from benign (MCF10A), to tumorigenic (MCF10-AT1), to metastatic (MCF10-Ca1h) breast cancer. m6A levels were reduced by targeting the main m6A methyltransferase, METTL3, by CRISPR-Cas9 deletion in exon 1. Decreased m6A had cell line specific effects suggesting unique roles for m6A within each stage. For example, the MCF10A cells showing an increased proliferative phenotype, while the MCF10-Ca1h cells showed no significant effects of the METTL3 decrease. The MCF10-AT1 line showed the most dramatic effect, with the decrease in m6A modification leading to increased migration, proliferation, and gene expression changes indicative of an Epithelial to Mesenchymal transition (EMT). Current work is ongoing to identify the m6A RNA binding protein(s) that mediate this effect. Furthermore, as the AT1's were derived by stable integration of HRASG12V into the MCF10A lines, these results suggest that the status of the RAS pathway may influence the impact of m6A manipulation. Thus, the results of this study may begin to address the controversy of m6A's role in cancer and suggest that m6A may have a dynamic role in cancer that depends on the stage of progression and/or specific pathways mutated in that cancer.

55

Unraveling the Impact of Splicing Factor Mutations: An Integrated Genomic Approach

Jose C. Martinez, Sally A. Hansucker, Paul Armistead, Daniel Dominguez
UNC

Mutations in splicing factors significantly impact treatment outcomes in AML and MDS by disrupting RNA splicing, promoting leukemogenesis, and contributing to chemotherapy resistance. However, the specific downstream effects on cellular function, as well as the commonalities and differences among these mutations, remain uncharacterized. To investigate the oncogenic impact of splicing factor mutations, we used CRISPR/Cas9 to engineer K562 erythroleukemia cell lines with heterozygous mutations in SF3B1, SRSF2, and U2AF1. These models closely replicate the splicing abnormalities seen in AML. We employed high-throughput PacBio long-read sequencing to uncover a wide range of previously unannotated transcript isoforms, providing insights into the unique transcriptional landscape induced by these mutations. Short-read RNA sequencing revealed differential exon skipping events, particularly impacting RNA translation regulation. In parallel, ATAC-Seq analysis assessed chromatin accessibility, supporting novel transcript discovery. To explore how these mutations affect protein synthesis, we adapted ribosome profiling using the RiboLace method, revealing significant disruptions in mitochondrial oxidative phosphorylation and transcription regulation pathways.Our comprehensive genomic approach uncovered a rich diversity of novel transcript isoforms exclusive to cells with splicing factor mutations. Prominent exon skipping events involved RNA translation regulation. Additionally, disruptions in mitochondrial respiration and transcriptional regulation highlighted the extensive effects of these mutations on cellular metabolism and gene expression. These findings provide critical insights into how splicing factor mutations drive oncogenesis in AML and MDS by altering RNA splicing and cellular function, identifying new therapeutic targets for intervention in these malignancies.

56

Identification of novel iron-responsive elements

Justin G. Martyr, Francisco F. Cavazos, Maria M. Aleman, Daniel Dominguez
UNC Chapel Hill School of Medicine Department of Pharmacology

Iron is a critical biomolecule with essential functions in electron and oxygen transport on the systemic and cellular levels. Thus, iron is regulated extensively through many mechanisms, including through iron regulatory proteins 1 and 2 (IRP1, IRP2) and their interactions with structured RNA iron-responsive elements (IREs). In depleted iron conditions, these proteins gain RNA-binding functionality, leading to well-studied bidirectional modulation of gene expression in IRE-containing mRNAs involved in iron metabolism. Despite a push to find IREs through computational approaches, there has been difficulty in linking computational prediction to biological impacts. In this work, we utilized a new IRE prediction tool (ASPIIRE) in combination with proteomic and transcriptomic analyses to identify novel IRE-containing (nIRE) and iron-responsive genes in K562s after treatment with an iron chelator. After the selection of 18 promising nIREs, we evaluated binding of IRP1 and IRP2 through a modular biotinylated handle pulldown approach, allowing facile

assessment of IRP1 and IRP2 binding from cellular lysate. This approach has led to the identification of nIREs which bind to IRP1 and one specific nIRE able to bind both IRP1 and IRP2. These results provide initial biological validation of these binding interactions and constitute a novel group of IRP-IRE targets with evident changes in gene expression upon iron deficiency. Current investigation turns toward understanding the implications of these IRP-IRE interactions, providing insight into biological pathways regulated by these RNA-binding proteins in cases of iron deficiency.

57

Uncovering human-specific post-transcriptional regulatory elements in neurogenesis

Nicole D. Moss, Debra L. Silver
Duke University

The cortex is responsible for many of our uniquely human cognitive abilities. During development of the cortex, nearly every stage is shaped by post-transcriptional gene regulation including RNA splicing, modifications, localization, and stability. These regulatory events may also contribute to phenotypic differences between species, particularly humans and their most recent ancestor, chimpanzees. Humans diverged from chimpanzees ~7mya and have accumulated several genomic differences that could help explain human-specific aspects of cortical development. Specifically, two sets of uniquely human genomic regions have been identified, the highly conserved regions (Human Accelerated Regions, HARs) and rapidly evolving regions (Human Ancestry, Quickly Evolved Regions, HAQERs). These have been predominantly studied for their function as transcriptional enhancers and repressors. We propose that there are additional roles for these human-specific sequences in post-transcriptional gene expression. Over 40% of these human-specific sequences are found within introns, many near intron-exon boundaries, where we predict, they may contribute to the regulation of splicing during human-specific neurodevelopment. To explore this novel function of HARs and HAQERs we are developing a high throughput splicing reporter system to screen these elements. Independently, we are testing candidate elements for their role in splicing in neural cells. These experiments will reveal new unappreciated functions for intronic HARs/HAQERs as cis-regulatory elements and putative RNA binding protein sites in human-specific post-transcriptional RNA regulation.

58

A Proximity-Inducing G-Quadruplex Assembly as a Plausible Model for Prebiotic Peptide Formation

Alex E Neary, Laurie Betts, Leonard Collins, Peter Dykeman-Bermingham, Abigail Knight, Charles W Carter, Jr., Qi Zhang
UNC Chapel Hill

How life originated on Earth remains one of the most elusive mysteries. The most vexing question posed by the origin of life may be the creation of coded peptide synthesis. Translating a gene amplifies its chemical and structural diversity enormously. However, it is hard to guess what created the ribosome, 20 aminoacyl-

tRNA cognate pairs, and GTP-dependent assembly factors that assemble proteins today. What selection process produced that machinery? We describe a G4 quadruplex ribosome-free translation system that requires only a single RNA sequence acylated with amino acids. The system is driven by the observation that G4 structures have been found in the genomes of all domains of life and represent some of the most stable naturally occurring sequences. In addition, it has been shown that RNA G4s are even more stable than their DNA counterparts, further supporting the RNA world hypothesis. We demonstrate that dimerized amino-acylated G4s accelerate the assembly of dipeptides by bringing the local amino acid concentration to approximately 7M. We describe the experimental data and analysis that demonstrate the feasibility of this system in synthesizing short peptide chains without the need for the complex machinery found in modern cells. Molecular adaptations that enhanced the speed or precision of such a system could be selected for that function. We hypothesize that the system resembles the primordial process that selected the succession of accessory components in the modern translation system.

59

Protein binding in an mRNA 5'-UTR hinders translation

Simon Felder, Irma M. Nelson, Breanne M. Hatfield, Kevin M. Weeks
UNC

Structures in the 5' untranslated regions (UTRs) of mRNAs can physically modulate translation efficiency by impeding the scanning ribosome or by sequestering the translational start site. Here we assessed the underexplored impact of stable protein binding in 5'- and 3'-UTRs on translation efficiency by targeting the MS2 coat protein to a reporter RNA via its hairpin recognition site. We evaluated translation from the reporter RNA when co-expressed with MS2 coat proteins of varying affinities for the RNA, and at different expression levels. Binding of high-affinity proteins in the 5'-UTR hindered translation, whereas no effect was observed when the coat protein was targeted to the 3'-UTR. The decrease in reporter RNA translation was larger at higher concentrations of the coat protein and for higher affinity variants, reaching a maximum translational inhibition of 50-70%. MS2 proteins engineered to bind two reporter mRNAs simultaneously had a stronger effect than proteins that bound a single mRNA. Protein binding in a 5'-UTR can thus physically impede translation, and the effect is governed by affinity, concentration, and sterics.

60

Engineering an Endoplasmic Reticulum (ER) Stress Responsive RNA Structural Switch for Gene Therapies

Roza Ogurlu, Alan Rosales, Sophia M. Fergione, Aravind Asokan
Duke University

Protein overexpression in the endoplasmic reticulum (ER) can trigger ER stress and activate the unfolded protein response (UPR) pathway, which may ultimately result in apoptosis. ER stress induction following administration of adeno-associated virus (AAV) delivered Factor VIII was linked to low secretion efficiency,

declined gene expression after the first year, and liver injuries in previous preclinical and clinical studies. Excessive expression of various endogenous and exogenous proteins, such as monoclonal antibodies, can provoke the same response regardless of the delivery modality, thereby reducing the therapy's efficacy. Therefore, there is an unmet need for a comprehensive engineering approach for controlling UPR while preserving high long-term protein expression in gene therapies. Inositol-requiring enzyme 1a (IRE1a), a kinase/endoribonuclease transmembrane protein, splices out a 26-nucleotide intron from X-box-binding protein 1 (XBP1) mRNA during UPR by recognizing a structural RNA element and shifts the protein reading frame. In order to leverage unconventional XBP1 mRNA splicing for therapeutic transcripts, we incorporated fragments of the XBP1 mRNA (XBP1F1-5) upstream of the therapeutic transcript. We verified mRNA splicing and protein expression regulation with reporters and therapeutic genes, such as Factor VIII and Leronlimab. We observed a significant decrease in ER stress marker levels when therapeutic protein expression was modulated by XBP1F. With additional testing and refinement, our IRE1a-mediated splicing based ER stress responsive structural RNA switch will provide a therapeutic protein and delivery agent independent mechanism for maintaining homeostasis in cells following ER stress causing protein overexpression, thus adding to the toolbox of safety switches for gene therapy.

61

A higher order PUF complex is central to regulation of C elegans germline stem cells

Chen Qiu[1], Sarah L. Crittenden[2], Brian Carrick[2], Lucas B. Dillard[1], Venkata P. Dandey[1], Elizabeth G. Viverette[1], Robert N. Wine[1], Zachary T. Campbell[1], Marvin Wickens[2], Mario J. Borgnia[1], Judith Kimble[2], Traci M. Tanaka Hall[1]
[1]National Institute of Environmental Health Sciences, National Institutes of Health, NC
[2]University of Wisconsin

PUF (for Pumilio and FBF) RNA-binding proteins are evolutionarily conserved stem cell regulators. In C. elegans, four PUF proteins together with two partner proteins govern germline stem cell renewal, with the gld-1 mRNA as a major target for repression. Here we focus on the FBF-2 and LST-1 partnership. LST-1 harbors two PUF-interacting motifs (PIMs), while FBF-2 possesses an autoregulatory PIM in its C-terminal tail (CT). We discovered that the intramolecular interaction between the FBF-2 CT and RNA-binding domain brings a negatively charged acidic cluster of amino acid residues near the 5´end of FBF-2-bound RNA, resulting in reduced RNA-binding affinity. Binding of either of the two LST-1 PIMs to FBF-2 relieves this autoinhibition by displacing the CT PIM, consequently enhancing FBF-2 RNA-binding affinity. Furthermore, we demonstrate that LST-1 can simultaneously interact with two FBF-2 proteins via its two PIMs. We identify a new FBF-binding element (FBE) called FBEa* adjacent to the established FBEa in the 3'-UTR of gld-1, and uncover functional synergism of these two FBEs. We suggest that full FBF-mediated repression of gld-1 in germline stem cells is achieved through formation of a higher-order complex, wherein two FBF-2 proteins linked by the dual PIMs of LST-1 bind to the dual FBEs in gld-1 RNA. Discovery of the LST-1–FBF-2 ternary complex, the gld-1 adjacent FBEs, and their in vivo significance predicts an expanded regulatory repertoire of different assemblies of PUF-partner complexes in nematode germline stem cells.

62

Transcriptome Wide Discovery of RNA tertiary structures

David Qiu, Jeffrey Ehrhardt, Kevin Weeks
UNC Chapel Hill

Complex higher-order RNA structures govern diverse biological processes. Many examples, spanning riboswitches to catalytic RNAs, are currently known. However, the number of broad classes of function-driving RNA tertiary structures are relatively few, and even fewer have been described for human RNAs. Current chemical probing strategies, integrated with structure modeling, now allow robust discovery and characterization of base pairing structure in RNA. In contrast, we lack strategies for high-confidence, de novo, transcriptome-scale discovery of RNAs with complex three-dimensional folds. Here, we introduce a high-throughput chemical probing technology, highly selective for identifying RNA tertiary structure, called T-site probing. The strategy compares mutational profiling (MaP) sequencing signals from two chemical probes, dimethyl sulfate and the trimethyloxonium cation, and identifies RNA tertiary structures as T-sites: nucleotides where preferential interaction of the trimethyloxonium cation with a local electronegative pocket created by the underlying tertiary structure fold induces a strongly enhanced trimethyloxonium signal over dimethyl sulfate. T-site probing in human cells identifies hundreds of novel tertiary structures across the human transcriptome. The functional significance of T-sites in mRNAs is readily confirmed via translational assays. T-site probing is poised to redefine and expand how we discover and understand the biological regulatory roles of pervasive RNA tertiary structures.

63

Structural Analysis of Neural Microexon Splicing

Alexandra Randazza, Lela Lackey
Clemson University

Microexons are a highly conserved, primarily neural-specific class of exons ranging from 3 to 27 nucleotides in length. They are enriched in surface protein interaction domains and under strong selection to maintain reading frames. Dysregulation of microexons is linked to multiple neurodevelopmental disorders. Of the ~500 alternatively spliced microexons detected post-mortem in the superior temporal gyrus, approximately 40% are mis-spliced in a portion Autism Spectrum Disorder (ASD) cases whereas only ~5% of longer exons are dysregulated. While average length exons utilize exonic splicing enhancers to promote exon inclusion, microexons lack the length for this type of regulation. I hypothesize that RNA structure contributes to the regulation of microexon splicing in neural cells. I investigate the role of RNA structure in microexon splicing by structure probing the splice sites of microexons to uncover local, secondary structures using short, average, and long exons as controls. Structures are compared using RNAsmc and PERFUMES. I will also investigate long-distance interactions in these RNAs through a reversible crosslinking technique called spatial 2'-hydroxyl acylation reversible crosslinking (SHARC). To identify the effect of specific structures on splicing,

I will conduct mutational studies with minigenes in vitro, in differentiating neural cells, and HeLa cells. These results open doors for the development of new interventions for neurodevelopmental disorders associated with microexon dysregulation.

64

Alternative splicing regulation of clathrin in heart development and function

R. Eric Blue, Adam Black, Matheus Sadovsky, Gabrielle Gentile, Brian Cooley, Haifeng Yin, Jimena Giudice
UNC Chapel Hill

Alternative splicing mechanisms regulate membrane trafficking genes during the first four postnatal weeks of heart development. These splicing-trafficking events are highly tissue-specific and evolutionarily conserved. One trafficking gene regulated by alternative splicing during heart development is the clathrin heavy chain (Cltc). Alternative splicing of a microexon (exon 31, 21 nucleotides) in Cltc generates two isoforms: a long CLTC form that includes the microexon and a short CLTC form that skips it. This microexon is tissue specific for adult heart and skeletal muscles. The short form is predominantly expressed in embryonic stages and there is a gradual transition towards primary expression of the long form in adult hearts and skeletal muscles. Several human diseases cause adult hearts to revert to fetal splicing programs which fail to support the adult functions of the organ. However, little is known about the cellular mechanisms underlying these diseases and less is known about microexons' regulation in heart function. We found that Cltc splicing is altered in both human diseases (i.e. myotonic dystrophy type 1) and mouse models of cardiac pressure overload (TAC) and myocardial infarction (left anterior descending coronary artery ligation, LAD). To investigate the mechanisms by which Cltc splicing regulates cardiac biology, we utilized CRISPR/Cas9 editing in mice to block endogenous Cltc splicing by deleting Cltc microexon 31 from the genome. Cltc-homozygous (Cltc-HO) hearts express only the short CLTC form across development, while Cltc-wildtype (Cltc-WT) hearts undergo normal Cltc splicing regulation during development. Interestingly, by challenging the heart with pressure overload (via transverse aortic constriction, TAC), we found Cltc-HO mice had a less severe heart failure response across 8 weeks following TAC. We hypothesize that Cltc exon-skipping can precondition adult hearts to better respond to the effects of pressure overload and other cardiac challenges. Moreover, although Cltc-WT and Cltc-HO hearts similarly express hypertrophic and fibrotic markers following the TAC challenge we found interesting differences in the expression of mRNAs and proteins involved in the regulation of iron metabolism and ferroptosis. Our ongoing efforts are directed to define the mechanisms by which Cltc splicing regulates heart physiology and failure and set the stage for potential therapies that redirect splicing to help treat severe heart challenges.

65

The regulation of RNA stability during cortical development

Lucas D. Serdar, Jacob R. Egol, Brad Lackford, Brian D. Bennett, Guang Hu, Debra L. Silver
Duke University

RNA abundance is controlled by rates of synthesis and degradation. Although mis-regulation of RNA turnover is linked to neurodevelopmental disorders, how it contributes to cortical development is largely unknown. Here, we discover the landscape of RNA stability regulation in the cerebral cortex and demonstrate that intact RNA decay machinery is essential for corticogenesis in vivo. We use SLAM-seq to measure RNA half-lives transcriptome-wide across multiple stages of cortical development. Leveraging these data, we discover cis-acting features associated with RNA stability and probe the relationship between RNA half-life and developmental expression changes. Notably, RNAs that are upregulated across development tend to be more stable, while downregulated RNAs are less stable. Using compound mouse genetics, we discover CNOT3, a core component of the CCR4-NOT deadenylase complex linked to neurodevelopmental disease, is essential for cortical development. Conditional knockout of Cnot3 in neural progenitors and their progeny in the developing mouse cortex leads to severe microcephaly due to altered cell fate and p53-dependent apoptosis. Finally, we define the molecular targets of CNOT3, revealing it controls expression of poorly expressed, non-optimal mRNAs in the cortex. Collectively, our findings demonstrate that fine-tuned control of RNA turnover is crucial for brain development.

66

Tissue targeting specificity of the RNA binding protein Zfp36L2

George S. Stephenson, Alain Laederach, Silvia B. Ramos
UNC Chapel Hill

ZFP36L2 (zinc finger protein 36, C3H type-like 2) an RNA-binding protein targeting transcripts rich in adenine-uridine elements (AREs), exhibits remarkable distinct transcript preferences, ‘specific activity’, across different tissues. In this study we obtained differential expression transcriptomic data on a Zfp36l2 knock-out mouse model to delve into the mechanisms governing this tissue-specific targeting. Transcriptomic analyzes in six tissues, bone marrow, ovary, spleen, liver, kidney, and lung, of the Zfp36l2-deficient mouse confirmed that there is high tissue preference in ZFP36L2 targets. We observed no common regulated gene among these six different tissues. No gene is up regulated in more than four common tissues. We do observe common trends, specifically an enrichment in protein coding genes in the up regulated genes, consistent with these RBP primarily targeting genes on its 3’ UTRs. Curiously, we observed a significant increase in the proportion of IG (immunoglobulin) genes beingup regulated. More specifically, IGV genes encoding the variable region of the immunoglobulin heavy chain were overrepresented in our up regulated genes. These variable regions of the antibody are the domain that determines its unique antibody specificity against antigens on the target pathogen. This finding raises the interesting possibility that ZFP36L2 interfere on the recombination process of the V(D)J gene, a key mechanism to produce antibody diversity and B cell

clonal selection. Our findings provide new insights into the intricate regulatory network orchestrated by ZFP36L2, opening avenues for exploring its potential roles in B cell physiological processes.

67

Novel recognition of m6A in viral RNA by an antiviral ISG

Matthew Thompson, Moonhee Park, Matthew Lanahan, Stacy M. Horner
Duke University

Recognition of pathogen-associated molecular patterns (PAMPs) by cellular pattern recognition receptors (PRRs) is a core tenet of activating an antiviral response. During RNA virus infection, viral RNA can be recognized by PRRs via specific features such as secondary structure and 5' end processing. In turn, PRR binding can be deterred by certain RNA features. For example, the PRR, RIG-I, recognizes dsRNA and misprocessed 5' caps but is deterred by the m6A RNA modification. m6A has been observed in the genomes of several RNA viruses. For example, m6A in a region of the Hepatitis C virus (HCV) genome inhibits RIG-I binding. However, HCV, is known to be recognized by PRRs other than RIG-I. Whether m6A contributes these other PRRs binding HCV is unknown. In this study, we determine if m6A alters binding of the IFIT protein family of PRRs to RNA. but the role of m6A in RNA recognition by other PRRs, such as IFIT1, is unknown. IFIT1 functions in complex with IFIT3 to sense modifications in the 5' end of RNA, with IFIT1 directly binding the RNA and IFIT3 stabilizing the interaction. As the IFIT1-IFIT3 complex restricts infection by several RNA viruses containing m6A, including hepatitis C virus (HCV), we sought to determine if m6A impacts RNA sensing and antiviral functions of IFIT1 and IFIT3.To test m6A-dependence on IFIT binding, we assayed binding of the IFIT1-IFIT3 protein complex to synthetic RNA probes +/- m6A. Surprisingly, we observed IFIT1-3 to modified probes with a specificity than not. As previous work shows the mechanism of IFIT1-3 RNA binding consists direct IFIT1 RNA binding and IFIT3 indirect interaction, we wondered if m6A is bound similarly. We tested IFIT1 and IFIT3 direct RNA binding using UV-crosslinking assays and show that in the case m6A, IFIT3 directly binds RNA and IFIT1 is an indirect co-factor. Further, we show specific domains that are independent of IFIT1 interaction, are required for IFIT3 RNA binding. As IFIT3 has not been described to bind RNA in any context, we sought to determine if IFIT3 m6A RNA binding is important for antiviral function. First, IFIT3 binding to HCV RNA was mapped using a HyperTRIBE approach in which A-I editing by an IFIT3-ADAR fusion was used to measure IFIT3 binding. We identified IFIT3 binding at nine regions in the HCV genome. To test is IFIT3 RNA binding was m6A-dependent, we show that treatment with an m6A inhibitor decreased both IFIT3 TRIBE editing as well as editing by an m6A-reader, YTH-ADAR. Likewise, mutation of m6A sites near TRIBE editing resulted in decreased IFIT3- and YTH-ADAR RNA binding. Toward attributing function to this binding, we show overexpression of IFIT3 decreases production of HCV virions. However, if RNA-binding is required is not yet known. In summary, we show the antiviral protein, IFIT3, to be a novel binder of m6A-modified viral RNA.

68

Molecular basis for RNA discrimination by human DNA ligase 1

Percy P. Tumbale, Thomas J. Jurkiw, Juno M. Krahn, Lars C. Pedersen, Jessica S. Williams, Tom A. Kunkel, Patrick J. O'Brien, and R. Scott Williams
NIEHS/NIH

Faithful maintenance and propagation of eukaryotic genomes is ensured by actions of DNA ligases to finalize DNA replication, repair, and recombination by catalyzing the joining of DNA nicks. The mammalian replicative DNA ligase 1 (LIG1) is highly specific for DNA-DNA junctions over DNA-RNA junctions, discriminating strongly against a single ribonucleotide at the 5'side of the nick (5'-rNMP) characteristic of Okazaki Fragment Maturation (OFM) or RNase H2 incision during Ribonucleotide Excision Repair (RER). This selectivity of LIG1 prevents futile and potentially mutagenic DNA-RNA cleavage involving RNase H1 or RNase H2 and re-ligation cycles. Despite the long-observed ability of LIG1 to abort ligation on DNA-RNA junctions, the molecular mechanism behind it is ill-defined. Here, we report structural and kinetic analysis of LIG1 DNA-RNA complexes showing that LIG1 employs an aromatic steric gate to stabilize the enzyme substrate complex and directly excludes 5'-rNMP containing polynucleotides. Mutation of the steric gate compromises LIG1's discrimination against an rNMP-containing substrate by >5000-fold, in addition to having a significant impact on ligation efficiency with an all-DNA substrate. Our results establish the role of the conserved steric gate in ribonucleotide discrimination by high-fidelity DNA ligases at each step of the ligation reaction, which has parallels to the ribonucleotide discrimination by high-fidelity DNA polymerases, providing key insights into the specialization of high-fidelity DNA ligases that are tuned for faithful replication of eukaryotic genome.

69

Ligand-binding pockets in RNA, and where to find them

Seth D. Veenbaas, Jordan T. Koehn, Patrick S. Irving, Nicole N. Lama and Kevin M. Weeks
University of North Carolina at Chapel Hill

RNAs are critical regulators of gene expression, and their functions are often mediated by complex secondary and tertiary structures. Structured regions in RNA can selectively interact with small molecules – via well-defined ligand binding pockets – to modulate the regulatory repertoire of an RNA. The broad potential to modulate biological function via RNA-ligand interactions remains unrealized, however, due to challenges in identifying RNA motifs with the ability to bind ligands with good physiochemical properties (often termed drug-like). Here, we devise fpocketR, a computational strategy that accurately detects pockets capable of binding drug-like ligands in RNA structures. Remarkably few, less than 60, of such pockets had ever been visualized previously. We experimentally confirmed the ligandability of novel pockets detected with fpocketR using a fragment-based approach introduced here, Frag-MaP, that detects ligand-binding sites in cells. Analysis of pockets detected by fpocketR and validated by Frag-MaP reveals dozens of newly identified sites able to bind drug-like ligands, supports a model for RNA secondary structural motifs able to bind quality ligands, and creates a broad framework for understanding the RNA ligand-ome.

70

RNA structural determinants of Alu-mediated backsplicing

Justin M. Waldern, Scott R. Allen, Patrick S. Irving, Colin D. Taylor, Catherine A. Giannetti, Kevin M. Weeks, and Alain Laederach
UNC Chapel Hill

Circular RNA (circRNA) are exons that form a covalently linked circle by spliceosome-mediated backsplicing. Circularizing exons often require flanking inverted sequence repeats. It is hypothesized that these complementary repeats form inter-intronic RNA base pairs that bring the exon ends into proximity to facilitate backsplicing. However, these long-range inter-intronic RNA interactions have only been hypothesized to form and have never been experimentally characterized. Here, we describe the long-range RNA structural interaction between the intronic inverted repeat Alu elements required for backsplicing of the HIPK3 exon 2 circRNA. Using SHAPE-MaP chemical probing combined with structure modeling, we have inferred that the Alu elements are likely base pairing across roughly 85% of their approximately 300 nucleotide length, despite being more than 2,000 nucleotides apart in linear space. Furthermore, we have used SHAPE-JuMP technology to directly detect and confirm the through-space interactions between the two Alu elements flanking the HIPK3 exon 2. We have validated the functional requirement of the Alu-Alu interaction for backsplicing by measuring circularization efficiency in cell culture with RT-qPCR of a native HIPK3 derived construct, as well as in circularization deficient mutant constructs, where each Alu element has been deleted. Taken together, our data reveals important structural determinants of back-splicing, including the need for imperfect pairing interactions. In addition our data strongly suggest that the 3â€™ backbone of the pairing region is highly flexible, suggesting important differences in base-pairing potential of the inverted repeats.

71

Novel Role of DEAD-box Helicase Prp28 in Regulating Spliceosomal Catalytic Activation

Zhongshi Wang, Jonathan Staley
The University of Chicago

Spliceosome is highly dynamic as it undergoes de novo assembly and disassembly for each splicing reaction. ATP-dependent RNA helicases participate in the transitions between intermediate conformations. The chemical reaction is catalyzed by U6 snRNA. U6 snRNA, however, is usually not in its catalytic active form. It base-pairs with U4 snRNA in the U4/U6.U5 tri-snRNP conformation and can only be recruited to the pre-spliceosome via the U4/U6.U5 tri-snRNP. After spliceosome assembly, U4 snRNA is removed from the spliceosome by the ATP-dependent RNA helicase Brr2 to form the catalytic active conformation. Since Brr2 is a constituent component of the tri-snRNP, its unwinding activity needs to be repressed so that an intact tri-snRNP could be assembled on to the pre-spliceosome. However, it is unclear to this date what serves as the gatekeeper for Brr2 repression. Prp28 is the ATP-dependent RNA helicase in the prior step which removes U1 snRNA from the 5' splice site after tri-snRNP recruitment to the pre-spliceosome. Prp28 is present in the human tri-snRNP structure and has long been thought to be absent from the budding yeast tri-snRNP.

Here I'm showing that Prp28 together with Sad1 associates with yeast tri-snRNP in S. cerevisiae as well as pathogenic yeast Cryptococcus neoformans and functions in preventing premature activation of Brr2. The inhibitory activity of Prp28 to Brr2 depends on Prp28's ATP helicase activity. The findings demonstrate a new role of Prp28 in regulating spliceosomal catalytic activation as well as clarifying the shared pathway of spliceosome assembly from yeast to humans.

72

APE2 Structural and Functional Characterization

Jessica L. Wojtaszek, Tejas Patel, Kristen Segars, Giancarlo Gonzalez, C. Denise Appel, Bret D. Wallace, Geoff A. Mueller, and R. Scott Williams
NIEHS/NIH

APE2, an important member of the endonuclease/exonuclease/phosphatase (EEP) class of genome protective enzymes, consists of an N-terminal EEP catalytic domain, a flexible region with an embedded PCNA-interacting motif, and a C-terminal Zf-GRF domain. APE2 catalyzes robust PCNA-stimulated 3'-5' exonuclease and 3' phosphodiesterase activities and weak AP endonuclease activity. The Zf-GRF domain of APE2 is required for 3'-5' end resection of DNA damage. These functions support the main role of APE2 in reversing blocked 3' DNA ends, problematic lesions that preclude DNA synthesis. Blocked 3' DNA ends include DNA strand breaks with damaged 3' termini in the form of 3' phosphate, 3' phosphoglycolate, or 2'-3' cyclic phosphate that result from the likes of ribonucleotide incorporation into DNA as well as oxidative DNA damage. Loss of APE2 is lethal in cells with mutated BRCA1 or BRCA2, making APE2 a prime target for homologous recombination-defective cancers. TOP1 processing of genomic ribonucleotides is the main source of 3'-blocking lesions relevant to APE2-BRCA1/2 synthetic lethality. Here we present integrated biochemical and structural characterization of APE2 that serves as the preliminary steps toward targeted inhibition of APE2 in BRCA1/2-deficient cancers.

73

Multiomic Analyses Uncover the Role of SNORD34 in Mitochondrial Metabolism

Pei-Chen Wu, Chris Holley
Duke University

SNORD34 is a member of the Rpl13a family of small nucleolar RNAs (snoRNAs), which also includes SNORD32A, SNORD33, and SNORD35A. These snoRNAs have been implicated in the regulation of metabolic processes and oxidative stress. Our study reveals that the loss of Rpl13a snoRNAs in snoRNA knockout (snoKO) mice, which lack all four Rpl13a snoRNAs, leads to a reduction in reactive oxygen species (ROS) levels in aortic smooth muscle cells (SMCs), and limits the development of atherosclerosis in ApoE knockout mice on a high-fat diet. Using the Seahorse platform, we discovered that SNORD34 plays a major role in regulating mitochondrial oxygen consumption rate (OCR) in 293T cells. Similarly, SMCs derived from Snord34 knockout mice exhibit a marked reduction in OCR, indicating that Snord34 has a critical role in

mitochondrial respiration and function. Given this, we aimed to uncover the molecular mechanisms linking SNORD34 to mitochondrial function. Although SNORD34 is canonically a C/D box snoRNA responsible for guiding 2'-O-methylation on ribosomal RNA (rRNA), our previous work showed that another Rpl13a snoRNA, SNORD32A, targets mRNA for 2'-O-methylation. This methylation influences both mRNA abundance and translation efficiency, increasing mRNA levels while suppressing protein expression. To identify SNORD34's targets, we performed RNA-seq and proteomics, filtering for targets with discordant RNA and protein expression. We identified several mitochondrial genes affected in SNORD34 knockout SMCs. Moreover, metabolic flux analysis of wild-type and snoKO SMCs revealed alterations in mitochondrial metabolism, including increased fatty acid oxidation and decreased glucose metabolism. These findings suggest that SNORD34 may play a role in regulating mitochondrial function by influencing mitochondrial protein expression and modulating mitochondrial metabolic pathways. This would provide new insights into SNORD34's role in mitochondrial metabolism regulation.

74

Interpretable Protein-DNA Interactions Captured by Structure-based Optimization

Yafan Zhang, Xingcheng Lin
North Carolina State University

Understanding sequence-specific DNA recognition is crucial for unraveling the mechanisms underlying gene regulation. While there are increasing numbers of sequence-based machine learning methods that accurately quantify sequence-affinity relationships or predict genome recognition sites for certain proteins, their success largely depends on the availability and quality of experimental high-throughput (HT) sequences. Moreover, the interpretation of the interaction patterns governing protein-DNA binding specificities remains challenging. Here, we introduce IDEA, an interpretable residue-level biophysical model that leverages structural and sequence information from a limited number of experimentally determined protein-DNA complexes to accurately predict protein-DNA binding affinities. With recent advancements in AI prediction models such as AlphaFold3 and RoseTTAFoldNA, IDEA has the potential to reliably characterize interactions for any protein-DNA pair, even in the absence of experimental structures, offering a significant promise for expanding our understanding of protein-DNA interactions. Additionally, the trained IDEA model can be directly integrated into our recently developed residue-resolution protein-DNA simulation protocol, enabling accurate predictions of absolute binding free energies for specific protein-DNA pairs. Our computational approach balances efficiency and predictive accuracy, completing chromatin-level simulations in days while maintaining high accuracy, which is particularly advantageous for studying the dynamics and mechanics of large protein-DNA complexes. Overall, IDEA offers an integrated computational platform that reduces experimental costs and biases in assessing DNA recognition and provides valuable insights into the mechanistic aspects of DNA-recognition processes.

75

Structural and functional characterization of the 3' splice site of cryptic exon 3 in the androgen receptor pre-mRNA

Marek D. Zorawski, James P. Falese, Amanda E. Hargrove
Duke University

Alternatively spliced mRNAs can contribute to aggressiveness and treatment resistance in castration-resistant prostate cancer (CRPC). One splice isoform of the androgen receptor (AR), called AR-V7, can participate in androgen receptor signaling even in the absence of androgens (a state known as androgen independence) and clinically correlates to treatment resistance and poor patient outcomes in CRPC. AR-V7 incorporates cryptic exon 3 (CE3) within the AR gene that prematurely terminates the full-length transcript. While several proteins have been described that may promote splicing of CE3, the impact of the pre-mRNA structures in this region is unknown. Using selective 2'-hydroxyl acylation analyzed by primer extension with mutational profiling (SHAPE-MaP), we developed the first experimentally informed pre-mRNA secondary structure model of the CE3 3' splice site from cellular RNA. In this structural model, we observe a novel stem loop structure, which we hypothesized could inhibit AR-V7 biogenesis. We used an AR-V7 minigene assay to evaluate structure-modulating mutations in AR-V7 biogenesis and observed concordant changes in AR-V7 abundance in response to stabilization of this repressive structure, leading us to term this structure the cryptic exon 3-splicing inhibitory stem loop (CE3-SISL). Protein binding site identification of the region through the delta-SHAPE pipeline suggests a repressive protein interaction that may be lost in androgen independence. Together, these data suggest the importance of RNA structure in AR-V7 biogenesis and indicate a potentially targetable interface for the development of AR-V7-depleting agents, which would represent the first RNA-targeted splicing modulators being developed for patients with AR-V7-driven CRPC.