

Report Release Webinar | March 21, 2024

Our New Report

Toward Sequencing and Mapping of RNA Modifications Committee

Sponsors:

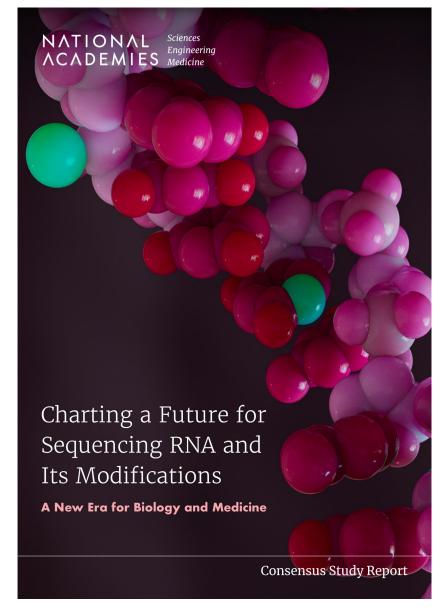
The Warren Alpert Foundation

National Human Genome Research Institute (NHGRI)

National Institute of Environmental Health Sciences (NIEHS)

Board on Life Sciences Board on Health Sciences Policy





Consensus Study Committee

Brenda L. Bass (Co-chair)

University of Utah School of Medicine

Taekjip Ha (Co-chair)

Harvard Medical School and Boston Children's Hospital

Nick M. Adams

Thermo Fisher Scientific

Juan D. Alfonzo

Brown University

Jeffrey Baker

National Institute for Innovation in Manufacturing and Biopharmaceuticals

Susan Baserga

Yale University and the Yale School of Medicine

NATIONAL Sciences Engineering Medicine

Lydia M. Contreras

The University of Texas at Austin

Markus Hafner

National Institute of Arthritis and Musculoskeletal and Skin Diseases

Sarath C. Janga

Indiana University Indianapolis

Patrick A. Limbach

University of Cincinnati

Julius B. Lucks

Northwestern University

Mary A. Majumder

Baylor College of Medicine

Nicole M. Martinez

Stanford University

Kate D. Meyer

Duke University School of Medicine

Keith R. Nykamp

Invitae Corporation

Tao Pan

The University of Chicago

National Academies Staff Team

Trisha Tucholski

Study Director

Steven Moss (until August 2023)

Former Study Director

Kathryn Asalone

Associate Program Officer

Kavita Berger

Board on Life Sciences Director

Consultants Michael Zierler

RedOx Scientific Editing

Jessica De Mouy

Research Associate

Victoria Hernandez

Mirzayan Fellow

Lyly Luhachack

Program Officer

Nam Vu

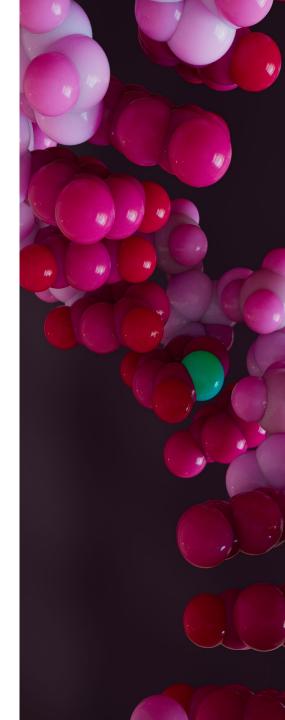
Senior Program Assistant

Special Thanks

Janet Iwasa and Rachel Torrez

University of Utah

For Report figures and Cover image



Report Reviewers

William Bonvillian

Massachusetts Institute of Technology

Cynthia Burrows

University of Utah

Vivian Cheung

University of Michigan

Robert Cook-Deegan

Arizona State University

Chuan He

The University of Chicago

Mark Helm

Johannes Gutenberg University Mainz

Tracy Johnson

University of California, Los Angeles

Todd Lowe

University of California, Santa Cruz

Intawat Nookaew

University of Arkansas for Medical Sciences

Samuel Schaffter

National Institute of Standards and

Technology

Elizabeth A. Snell

Oxford Nanopore Technologies

James Williamson

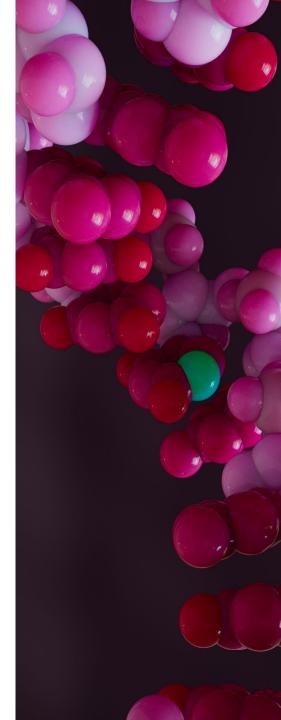
Scripps Research

Paul Ahlquist (Coordinator)
University of Wisconsin-Madison

Dan Blazer (Monitor)

Duke University School of Medicine

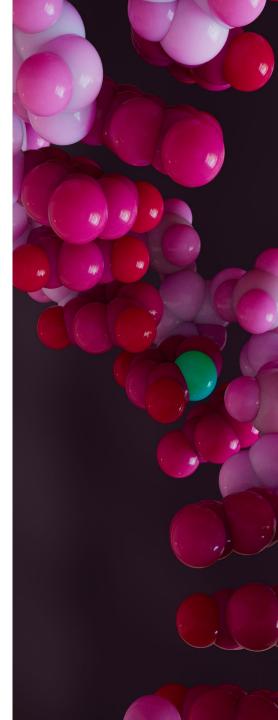




Overview of Statement of Task

The committee will develop a roadmap for achieving sequencing of any RNA, including its modifications. The committee will examine:

- Scientific needs
- Current methodologies and their limitations
- The state of current RNA databases
- Challenges related to using information for scientific, clinical, and public health analysis needs
- Computational and analytical needs
- Data ecosystems
- Policy, workforce, and infrastructure needs
- Potential new technologies



Committee's Approach

- NIEHS/NHGRI workshop May 2022
- NASEM Workshop March 2023
- Information gathering sessions
- Meetings of experts
- Ideation challenge June 2023
- Commissioned papers

Other National Academies reports:

- Mapping and Sequencing the Human Genome (1988)
- Transforming Glycoscience: A Roadmap for the Future (2012)

Capturing RNA Sequence and Transcript Diversity - From Technology Innovation to Clinical Application

May 24 - 25, 2022 • 11:00 a.m. - 4:00 p.m. EDT May 26, 2022 • 11:00 a.m. - 4:30 p.m. EDT

Virtual Event





Proceedings of a Workshop—in Brief

Toward Sequencing and Mapping of

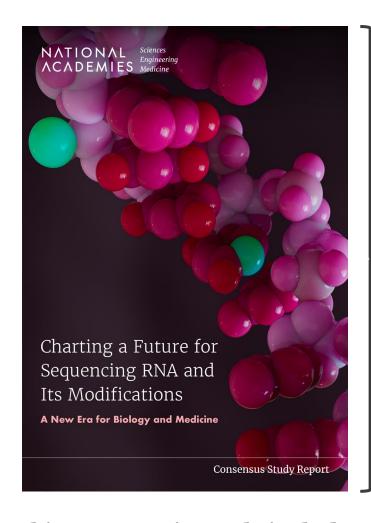
RNA Modifications

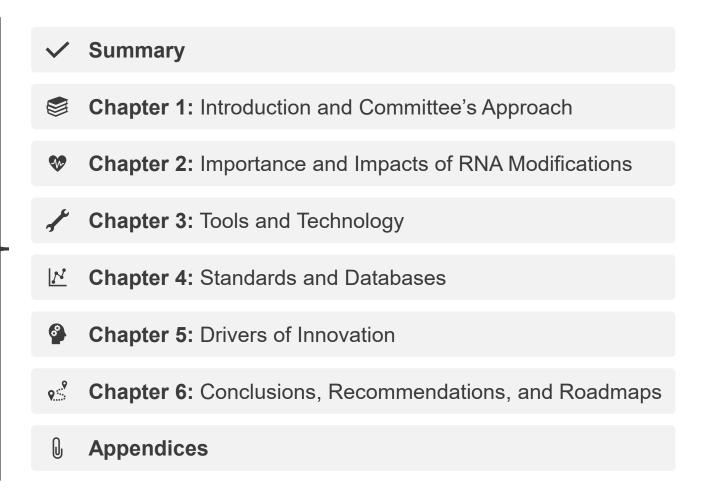
Proceedings of a Workshop—in Brief





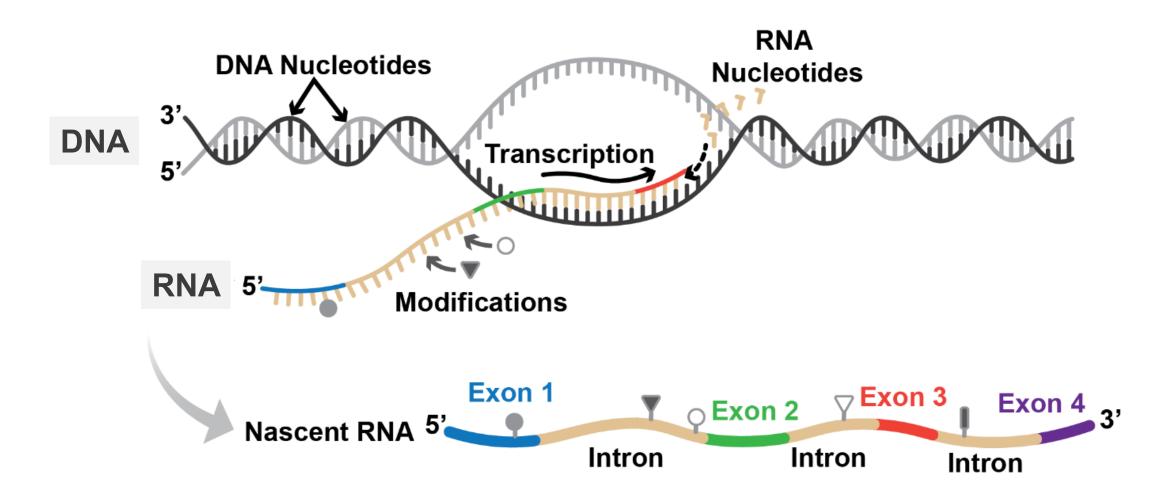
Organization of the Report



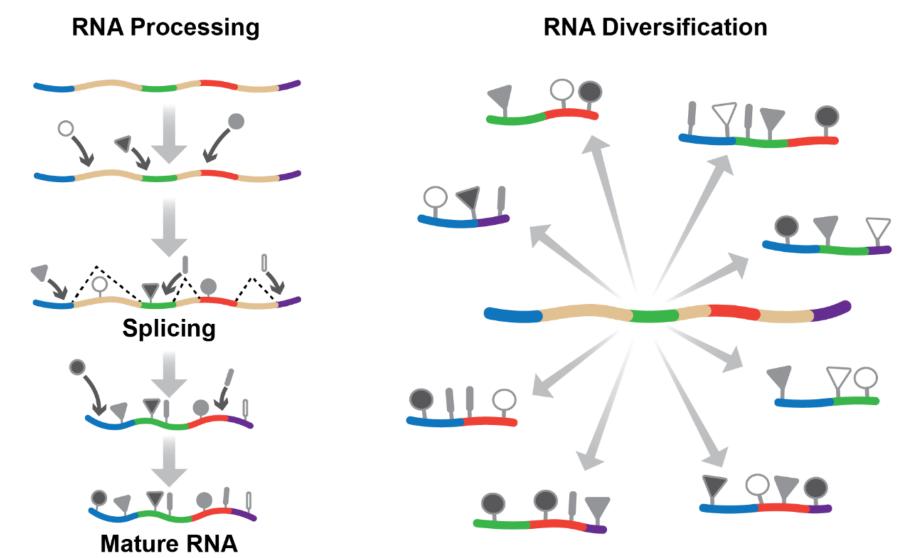


This presentation only includes *selected* highlights and recommendations from the full report.

From DNA "blueprint" to RNA



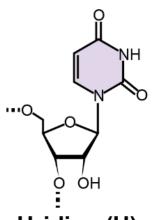
RNA and the complexity of life

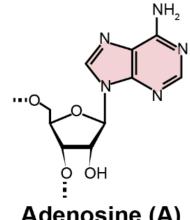


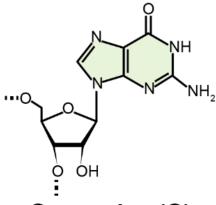
Many distinct RNA molecules derive from each gene!

What exactly are "RNA modifications"?

Nucleosides







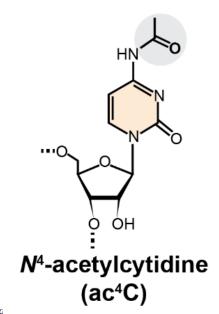
Cytidine (C)

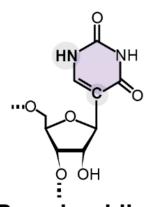
Uridine (U)

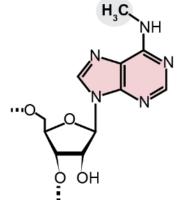
Adenosine (A)

Guanosine (G)

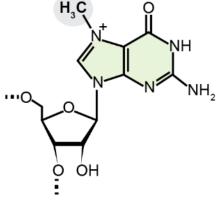
Modified Nucleosides (4 of ~170 known)







 (m^6A)



 N^6 -methyladenosine N^7 -methylguanosine (m^7G)

RNA modifications can:



Modified Nucleosides (4 of ~170 known) ÓН ÓΗ ÓН ÓН N^6 -methyladenosine N^7 -methylguanosine N⁴-acetylcytidine **Pseudouridine** (Ψ) (m^6A) (m^7G) (ac⁴C) NATIONAL ACADEMIES Medicine

Why are RNA modifications so important?

- Crucial in nearly every life process!
 - Regulation of metabolism
 - Circadian rhythm
 - Immune system regulation (e.g., differentiation between host and viral RNAs)
- Implicated in human diseases and disorders
 - Neurological disorders (e.g., intellectual disability, microcephaly)
 - Heart disease (e.g., hypertrophic cardiomyopathy)
 - Cancer (e.g., breast, skin, colorectal, bladder)
 - Diabetes
 - "tRNA modopathies"



Visit Chapter 2, Table 2-1 for a list of human diseases caused by aberrant tRNA modifications

RNA modifications in vaccines against COVID-19

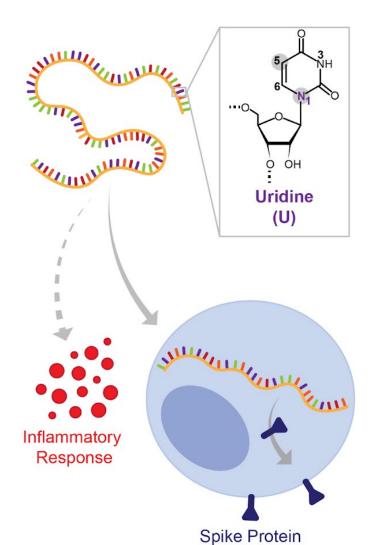
- Modifications permit the COVID-19 mRNA vaccines to succeed in protecting against infection and serious disease
- Foundational research awarded the 2023 Nobel Prize in Physiology or Medicine



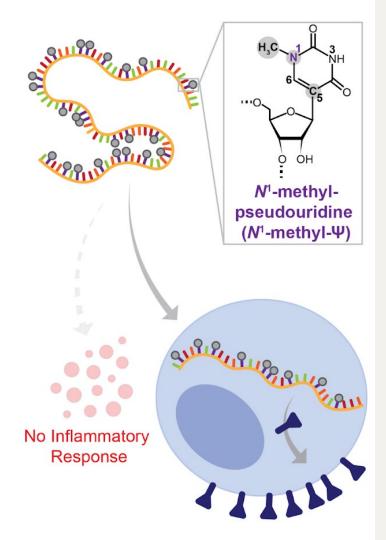
Dr. Katalin Karikó Dr. Drew Weissman



Unmodified RNA



Modified RNA



RNA vaccines and medicines

- mRNA vaccines against cancer (in clinical trials)
- mRNA vaccines for other pathogens (e.g., HIV, influenza, bacteria)
- Antivirals and antibiotics
- Therapeutics to treat rare diseases (e.g., spinal muscular atrophy)
- Enhancement to gene-editing, CRISPR, a promising tool for gene therapy



Visit Chapter 2, Table 2-2 for a list of RNA therapeutics approved by U.S. Food and Drug Administration (FDA) or European Medicines Agency (EMA)

RNA modifications beyond health and disease

- Agriculture
- Synthetic biology
- Nanotechnology
- Bioeconomy







Visit Chapter 2, Figure 2-4

Yu et al., Nat Biotech 2021

Conclusion 1: RNA modifications are a critical, but underexplored area of research

A more complete understanding of RNA modifications will be important for:

- Significantly advancing the fundamental knowledge of living systems
- Maintaining the health of humans, plants, animals, and the environment
- Preventing and treating disease
- Improving crop yields and resilience
- Stimulating the bioeconomy
- Addressing other issues of societal importance





















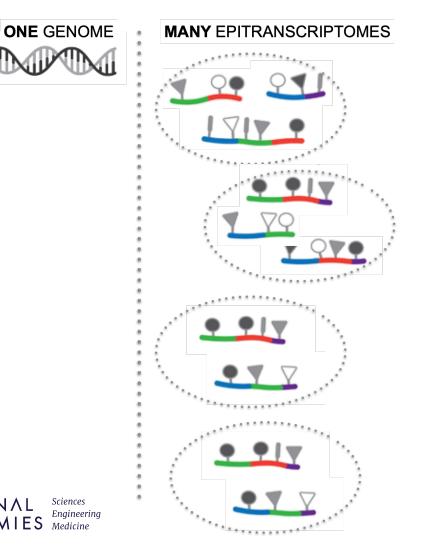


Conclusion 2: Enabling the determination of any epitranscriptome will be the most impactful goal

- "Epitranscriptome" = the sequences of a set of RNA molecules and their modifications
- Epitranscriptomes vary between cells and tissues, and with age, sex, and environment.

Conclusion 2: Enabling the determination of any epitranscriptome will be the most impactful goal

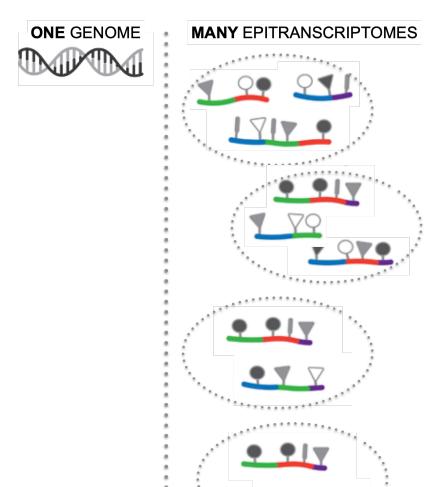
To define an individual requires sequencing:



- "Epitranscriptome" = the sequences of a set of RNA molecules and their modifications
- Epitranscriptomes vary between cells and tissues, and with age, sex, and environment.

Conclusion 2: Enabling the determination of any epitranscriptome will be the most impactful goal

To define an individual requires sequencing:



- "Epitranscriptome" = the sequences of a set of RNA molecules and their modifications
- Epitranscriptomes vary between cells and tissues, and with age, sex, and environment.
- There are many important epitranscriptomes to determine even for one individual!!
- Developing technology and infrastructure to enable the sequencing of *any* epitranscriptome will be the most impactful goal

Key efforts needed to unlock any epitranscriptome



A Path Forward

Phase I (5 years) Phase 2 (10 years) Phase 3 (15 years) Chapter 3, Figure 3-6 Standardize sample ▼ Enable end-to-end Establish high-throughput, A Roadmap for Advancing isolation, preparation, and sequencing and mapping of low-cost, highly sensitive handling for global analyses all RNA molecules and their clinical assays for multiple RNA modifications modifications Tools and Technology Refine and optimize existing experimental and Facilitate broad access to Apply tools and technologies computational approaches epitranscriptome sequencing to answer important by driving down cost and questions and provide Centralize information about improving usability insights into human health available technology, tools, and disease methods, and computational Accurately measure changes toolkits in modification stoichiometry Phase I (5 years) Phase 2 (10 years) Phase 3 (15 years) across multiple defined Explore novel experimental Chapter 4, Figure 4-3 cellular conditions for tRNAs and computational Establish modified Synthesize longer modified ✓ Readily available. A Roadmap for Developing Standards and Databases approaches Understand modification nucleoside standards for 25oligonucleotides and expand affordable Advancing cross-talk via modification 30 known modifications repertoire of available oligonucleotides of any Technology RNA core centers that measurements at the single custom-ordered sequence. modifications offer a range of Create 50-100 modified molecular level for tRNAs length, modification sequencing and oligonucleotides for use as Develop isotopically enriched stoichiometry, and ✓ Epitranscriptomes for computational services reference standards modified nucleosides for **Physical** structure and associated expertise cultured cell lines and quantification of modification Standards Commercially available in house tissues in single, wellstoichiometry standard kit defined states Maps for abundant RNA ✓ Commercial kits with types (e.g., tRNAs) for ✓ Complete expanded repertoire of Phase I (5 years) Phase 2 (10 years) Phase 3 (15 years) multicellular eukaryotic epitranscriptome for available modifications organisms in single, wellsimple multicellular and isotopically labeled defined states eukaryotic organisms Design multidisciplinary ▼ Include information about the A well-trained. oligonucleotide standards (e.g., flies, worms, zebra modules and training importance and impacts of impassioned, diverse U.S. Viral epitranscriptomes fish) in single, well-defined curricula for different RNA modifications in high workforce with (e.g., SARS-CoV-2, HIV, Establish universal Put in place a financial state educational levels school and college courses interdisciplinary expertise influenza) standards for RNA mechanism for ensuring the that is engaged in Develop experiential learning Equip members of the modification nomenclature long-term sustainability of advancing and applying modules for high school and existing workforce to and common ontologies for critical RNA data resources RNA biology and college students contribute to advancing and use in publications and epitranscriptomics across A sustainably funded. applying RNA science databases Create learning materials for the public and private stable, integrated, and through professional sectors, academia, and professional development Establish guidelines and centrally managed data development and training and retraining programs industry standards for raw data resource that is a long-Data and programs submissions lasting and always current Database Organize courses and Cultivating the source of curated Standards workshops that provide Require that raw and Future information about hands-on training in wet-lab processed datasets be made Workforce modified RNAs publicly available techniques Chapter 5, Figure 5-1 A Roadmap for Cultivating the Future Workforce A model for RNA Create dedicated modification data resource conferences and funding with guidelines for data mechanisms to convene professionals from different inclusion, sharing, and disciplines accessibility A suite of crossdisciplinary educational resources and training materials well-suited for a variety of levels

A Roadmap for Developing Standards (Figure 4-3)

	Phase I (5 years)	Phase 2 (10 years)	Phase 3 (15 years)
Physical Standards	 Establish modified nucleoside standards for 25–30 known modifications 	Synthesize longer modified oligonucleotides and expand repertoire of	 ✓ Readily available, affordable oligonucleotides of any
	 Create 50–100 modified oligonucleotides for use as reference standards ✓ Commercially available standard kit 	available modifications Develop isotopically enriched modified nucleosides for quantification of modification stoichiometry	custom-ordered sequence, length, modification stoichiometry, and structure
		✓ Commercial kits with expanded repertoire of available modifications and isotopically labeled oligonucleotide standards	

Conclusion 3: Large-scale, coordinated efforts in the life sciences have been vital in driving science and technology innovation

- E.g., Human Genome Project, Glycoscience Program, Human Microbiome Project, BRAIN Initiative
- Such efforts hold value in their ability to:
 - Align federal agencies
 - Support public—private partnerships
 - Organize consortia
 - Fund individual laboratories
 - Prioritize closing gaps in technology development, synthesis of standards, infrastructure buildout, workforce training, and public awareness



Conclusion 4: A large-scale epitranscriptomics effort is needed to accelerate technological innovation and progress in the field

- We envision it will impact multiple sectors (e.g., health, agriculture)
- Require expertise spanning multiple scientific disciplines (e.g., engineering, computer science, life science, social science)
- Require dedicated funding to key federal entities—such as NSF, NIH, NIST, DOD, and the DOE—to enhance their ability to work with academia, industry, philanthropic organizations, and international partners
- An endeavor of this scale and scope will entail a substantial investment of time and resources



Key Recommendations



Recommendation 1 (abbr.): An **established oversight body**, such as the Office of Science and Technology Policy or a similar entity with appropriate breadth and authority, **should catalyze and coordinate efforts** supporting a large-scale epitranscriptomics initiative



Recommendation 2 (abbr.): Federal funders of research such as NIH, NSF, DOD, and DOE should invest in and prioritize closing gaps in the existing tools, exploring new technologies, and centralizing resources for available tools and methods.



Recommendation 3 (abbr.): NIST should **develop, curate, and promote standards** to support the field of epitranscriptomics. Specifically, modified RNA reference materials should be developed with a focus on making them widely available and affordable.

Visit the **Summary** or **Chapter 6** for more details on Recommendations

Key Recommendations (cont.)



Recommendation 4 (abbr.): NCBI should establish and **promote standards** for databases, data deposition and exchange, and nomenclature for RNA modifications.



Recommendation 5 (abbr.): NIH should establish and maintain a sustainably funded and centrally **managed database** that maintains upto-date, curated information about RNAs and their modifications.



Recommendation 6 (abbr.): To develop a **strong workforce**, institutions and funding agencies, such as HHMI, NIGMS, and NSF, in partnership with education experts, scientific societies, and industry groups, should build upon existing educational materials and training opportunities for students and the public.

Visit the **Summary** or **Chapter 6** for more details on Recommendations

Potential outcomes and broader impacts



Novel sequencing tools, technologies, and computational methods



Improved data infrastructure



A well-trained, impassioned, diverse U.S. workforce





Discovery of new linkages between RNA modifications and genetic disorders



Improved disease control, prevention, and preparation for future pandemics



Development of better vaccines to fight viruses



New immunotherapies to treat cancer



New antibiotics and antivirals



Enhanced efficacy of RNAs used in gene therapy



New diagnostic protocols for viral detection, tumor testing, and detection of inherited diseases



New methods for improving crop yields and drought resilience, thus enhancing agricultural production and greater global food security



Bolstered synthetic biology capabilities



Advances in computer science





Thanks for listening!

We invite your questions



Q&A Panel

Brenda Bass

TJ Ha

Juan Alfonzo

Lydia Contreras

Patrick Limbach

Kate Meyer

Download the report

Critical challenges and roadblocks



Despite notable progress and economic value in this growing field, much about how RNA modifications affect the fate and function of RNA molecules in living systems is still unknown.



Existing technologies cannot currently discover all RNA modifications, let alone comprehensively sequence them on every RNA molecule.



Several types of standards are needed, specifically modified RNA reference materials, data standards around nomenclature and clear guidelines for data deposition and exchange, and robust and sustainable platforms for the curation and indexing of vast amounts of RNA data.



A sustainably funded and centrally managed database (or ensemble of databases) that maintains up-to-date, curated information about RNAs and their modifications is urgently needed.



To meet the future demands, the field needs a well-trained, impassioned, diverse U.S. workforce with interdisciplinary expertise and strong quantitative and computational skills

Highlights: Expanding Research



Conclusion 5: Discovery efforts and fundamental research in the field of epitranscriptomics will reinforce the importance and impact of RNA modifications and fuel technological advances that will improve scientists' ability to sequence them.

(cont.) New funding mechanisms, public and private, that encourage collaboration, spur innovation, and increase interest in RNA modifications will be critical.

Highlights: Advancing Technology



Conclusion 6: The current tools, technologies, and methodologies for end-to-end sequencing of RNA and all of its modifications are insufficient.

(cont.) The field of RNA biology will be driven forward by improving upon existing approaches and advancing new technologies that are robust and quantitative, and that preserve the information of full-length RNAs.



Conclusion 7: Improving the sensitivity of methodologies for cataloging and quantifying all RNA modifications in a sample, even without positional information, is an important enabling step that will inform the development of future RNA sequencing technologies and facilitate discovery of additional RNA modifications.

(cont.) Achieving this crucial intermediate goal will be spurred by an expanded repertoire of modified nucleosides for use as reference standards and more sensitive instrumentation.

Highlights: Advancing Technology (cont.)



Conclusion 8: Efforts directed toward enabling end-to-end sequencing of RNA and its modifications will accelerate innovation in the life sciences research enterprise but will also pave the way for developing new biotechnologies (e.g., biotherapeutics, vaccines, diagnostics, nanomaterials) and novel approaches that open new doors in life sciences research and other areas that are not yet apparent.



Recommendation 2 (abbr.): Federal funders of research such as NIH, NSF, DOD, and DOE should invest in and prioritize closing gaps in the existing tools, exploring new technologies, and centralizing resources for available tools and methods.

Highlights: Developing Standards



Conclusion 9: Several types of standards are needed:

- Technology-agnostic modified RNA reference materials
- Data standards around nomenclature and clear guidelines for data deposition and exchange
- Robust and sustainable platforms for the curation and indexing of vast amounts of RNA data



Recommendation 3 (abbr.): NIST should develop, curate, and promote standards to support the field of epitranscriptomics. Specifically, modified RNA reference materials should be developed with a focus on making them widely available and affordable.



Recommendation 4 (abbr.): NCBI should establish and promote standards for databases, data deposition and exchange, and nomenclature for RNA modifications.

Highlights: Centralizing Data Resources



Conclusion 10: The prevalence of "home-grown," small-group-supported RNA databases has been vital to advancing the field of RNA biology.

(cont.) Nonetheless, a major concern is the loss of resources (e.g., funding, staff) leading to a lack of maintenance of these laboratory-housed databases. Abandoning carefully curated databases may limit scientific growth and understanding, and waste time, effort, and resources.



Recommendation 5 (abbr.): NIH should establish and maintain a sustainably funded and centrally managed database that maintains up-to-date, curated information about RNAs and their modifications.

Highlights: Cultivating Innovation



Conclusion 11: Greater emphasis on RNA science in undergraduate courses is needed to build a better infrastructure for embracing future generations in the workforce.

(cont.) In addition to further education, the existing and future workforce needs interdisciplinary training with strong quantitative and computational skills.



Conclusion 12: Educational efforts in the RNA modifications field need to:

- Use methods that promote engagement
- Reflect the interdisciplinary nature of the science in education and related workforce development efforts,
- Invest in reaching and engaging students and trainees from diverse backgrounds, and
- Scale up proven strategies for retaining trainees in piloted programs.

Highlights: Cultivating Innovation (cont.)



Recommendation 6 (abbr.): To develop a strong workforce, institutions and funding agencies, such as HHMI, NIGMS, and NSF, in partnership with education experts, scientific societies, and industry groups, should build upon existing educational materials and training opportunities for students and the public.

(cont.) The materials and opportunities should be age appropriate and engaging for the interests of different groups, while covering the basic biological and chemical principles of RNA modifications, the tools available for their study, and their potential application in future medicines and useful biotechnologies.

A Roadmap for Advancing Technology

	Phase I (5 years)	Phase 2 (10 years)	Phase 3 (15 years)
Advancing Technology	Standardize sample isolation, preparation, and handling for global analyses	 Establish high-throughput, low-cost, highly sensitive clinical assays for multiple RNA modifications 	Enable end-to-end sequencing and mapping of all RNA molecules and their modifications
	 Refine and optimize existing experimental and computational approaches Centralize information about available 	 Facilitate broad access to epitranscriptome sequencing by driving down cost and improving usability 	 Apply tools and technologies to answer important questions and provide insights into human health and disease
	technology, tools, methods, and computational toolkits	Accurately measure changes in modification stoichiometry across	Leverage knowledge of RNA modifications for applications beyond
	Explore novel experimental and computational approaches	 multiple defined cellular conditions for tRNAs Understand modification cross-talk 	 health and medicine Continue to move epitranscriptomics toward single-cell and single-
	 ✓ RNA core centers that offer a range of sequencing and computational services and 	via modification measurements at the single molecular level for tRNAs	molecule applications ✓ Complete epitranscriptome of
	associated expertise in house	Epitranscriptomes for cultured cell lines and tissues in single, well-	complex multicellular organisms under multiple defined cellular
	 ✓ Maps for abundant RNA types (e.g., tRNAs) for multicellular eukaryotic organisms in single, well-defined states ✓ Viral epitranscriptomes (e.g., SARS-CoV-2, HIV, influenza) 	defined states	conditions
		✓ Complete epitranscriptome for simple multicellular eukaryotic organisms (e.g., flies, worms, zebra fish) in single, well-defined state	 ✓ Epitranscriptomes for human disease profiling
			41

A Roadmap for Developing Standards and Databases

	Phase I (5 years)	Phase 2 (10 years)	Phase 3 (15 years)
Physical Standards	 Establish modified nucleoside standards for 25–30 known modifications Create 50–100 modified oligonucleotides for use as reference standards ✓ Commercially available standard kit 	 Synthesize longer modified oligonucleotides and expand repertoire of available modifications Develop isotopically enriched modified nucleosides for quantification of modification stoichiometry Commercial kits with expanded repertoire of available modifications and isotopically labeled oligonucleotide standards 	✓ Readily available, affordable oligonucleotides of any custom-ordered sequence, length, modification stoichiometry, and structure

A Roadmap for Developing Standards and Databases (cont.)

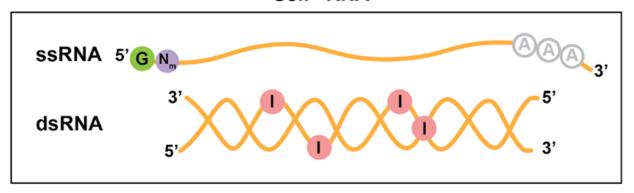
t in place a financial ✓ Seamless access to universally available and shareable information of ical RNA data resources defined set of cellular
ble, integrated, and an analysis of the solutions conditions condi
er es es ou

A Roadmap for Cultivating the Future Workforce

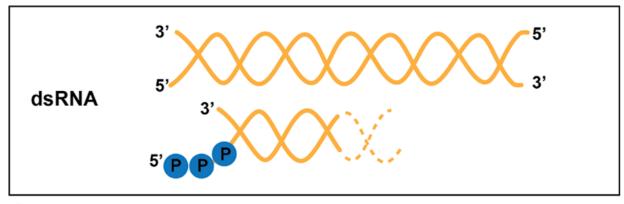
		<u> </u>				
	Phase I (5 years)		Phase 2 (10 years)		Phase 3 (15 years)	
Cultivating the Future Workforce	 Design multidisciplinary modules and training curricula for different educational levels Develop experiential learning modules for high school and college students Create learning materials for professional development and retraining programs Organize courses and workshops that provide hands-on training in wet-lab techniques Create dedicated conferences and funding mechanisms to convene professionals from different disciplines ✓ A suite of cross-disciplinary 	•	Include information about the importance and impacts of RNA modifications in high school and college courses Equip members of the existing workforce to contribute to advancing and applying RNA science through professional development and training programs	✓	A well-trained, impassioned, diverse U.S. workforce with interdisciplinary expertise that engaged in advancing and applying RNA biology and epitranscriptomics across the public and private sectors, academia, and industry	3
	educational resources and training materials well-suited for a variety of levels					

Chapter 2, Figure 2-1

"Self" RNA



"Non-self" RNA



- **G** 7-methyl-guanosine (m⁷G)
- N 2'-O-methyl
- Inosine
- A Poly(A) Tail
- Phosphate

Chapter 2, Figure 2-2

