Enabling Exploration of the Eukaryotic Epitranscriptome (E^4)

John Satterlee, NIDA

on behalf of the Common Fund Epitranscriptomics Work Group

DPCPSI Council of Councils, January 30, 2015

E4 Work Group Members

Co-chairs: Nora Volkow (NIDA), Dinah Singer (NCI)

Co-coordinators: John Satterlee (NIDA), Randy Knowlton (NCI)

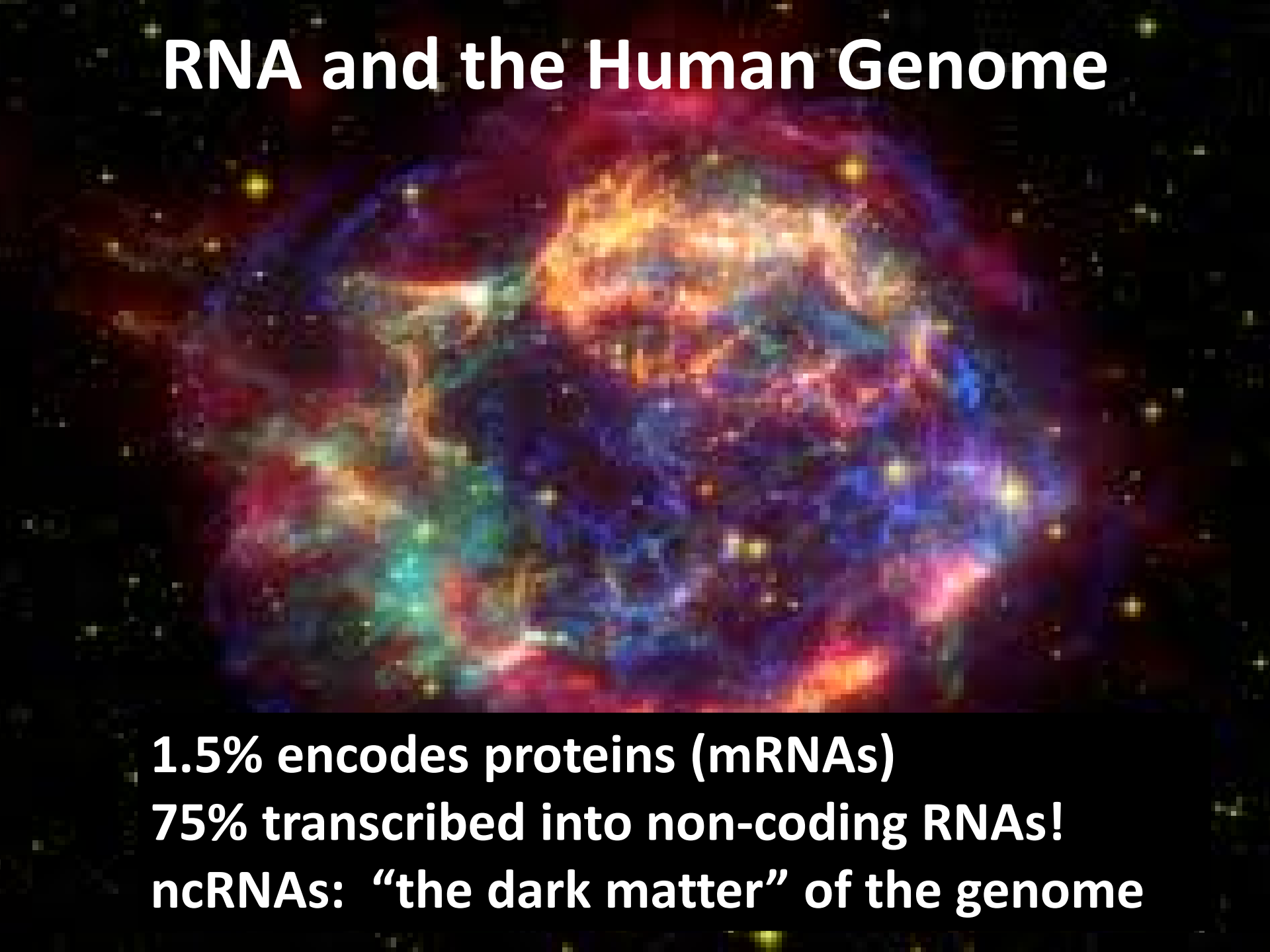
Richard Panniers	CSR
Lillian Kuo	NCATS
Dorit Zuk	NCATS
Carol Pontzer	NCCAM
Sean Hanlon	NCI
Lisa Neuhold	NEI
Michael Smith	NHGRI
PJ Brooks	NCATS
Conrad Mallia	NIAID
Elizabeth Stansell	NIAID
James Coulombe	NICHD
Stuart Moss	NICHD
Roger Little	NIDA
Dena Procaccini	NIDA
Silvio Gutkind	NIDCR
Isaac Rodriguez-Chavez	NIDCR
Terry Bishop	NIDDK
Lisa Chadwick	NIEHS
Peter Preusch	NIGMS
Darren Sledjeski	NIGMS
Anjene Addington	NIMH
Geetha Senthil	NIMH
David Owens	NINDS
Leslie Derr	OD
Taylor Gilliland	OD
Rebecca Lenzi	OD

MISSION: Identify key scientific issues in the area of Epitranscriptomics for development into a potential new Common Fund program.

Outline

- **Background**
- **RNA Modification Functions**
- **E4 Gaps and Opportunities**
- **E4 Program Components & Impact**

RNA and the Human Genome



1.5% encodes proteins (mRNAs)

75% transcribed into non-coding RNAs!

ncRNAs: “the dark matter” of the genome

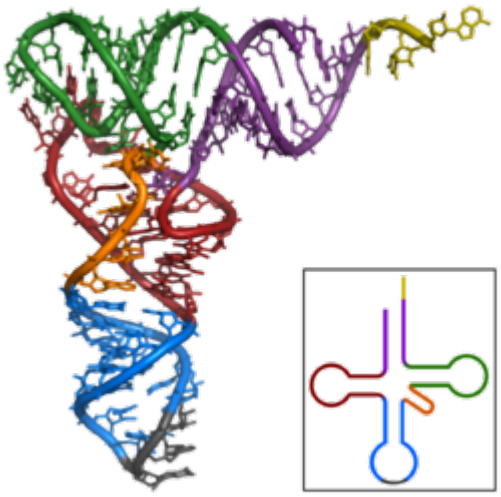
The RNA World

First self-replicating molecule ?

Many classes of RNA: **mRNA, tRNA, rRNA, microRNA, siRNA, piRNA, long non-coding RNA, snoRNA, etc.**

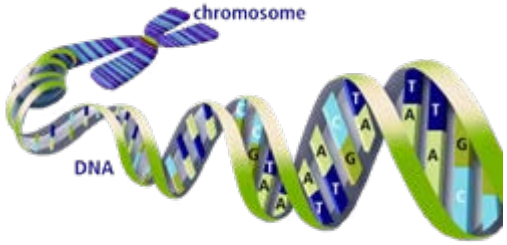
Many RNA functions:

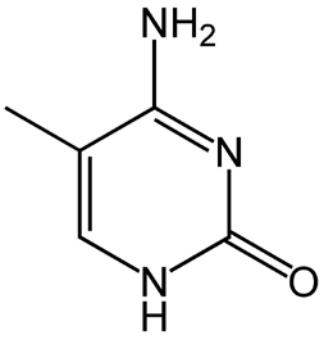
Protein translation and localized translation
Splicing
Structural scaffold
Chromatin recruitment
Ribozymes
Environmental sensing
Post-transcriptional regulation
Gene silencing
Defense against germline transposons
Intercellular signaling
RNA modification



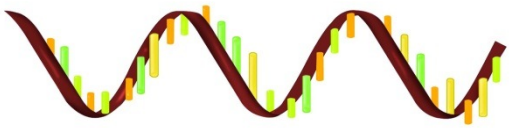
DNA and proteins undergo chemical modifications

DNA



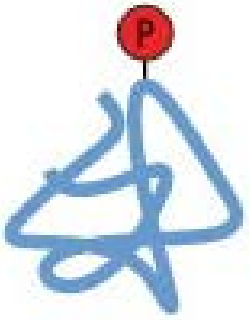


RNA



?

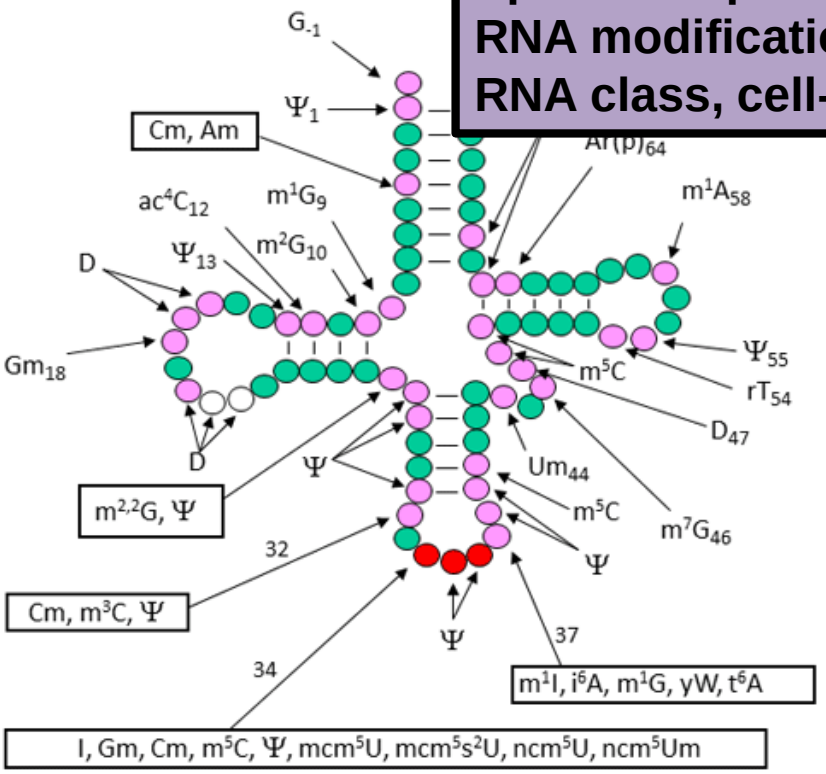
Proteins



Adapted from Samie Jaffrey

>110 RNA Modifications Are Known

Epitranscriptome: all of the RNA modifications for a given RNA class, cell-type, or organism.



Nucleosides
adenosine (A)
ntenyladenosine (i6A)
hyladenosine (m6A)
hyladenosine (m1A)
inosine (I)
ethylinosine (m1I)
guanosine (G)
1-methylguanosine (m1G)
7-methylguanosine (m7G)
N2-methylguanosine (m2G)
2'-O-methylguanosine (Gm)
uridine (U)
2'-O-methyluridine (Um)
5-methyluridine (m5U)
6-methyluridine (m6U)
5,6-dimethyluridine (m5m6U)
2-thiouridine (s2U)
4-thiouridine (s4U)
5-methoxyuridine (mo5U)
5-m
5
5
1-methyl-3-(3-amino-3-carboxypropyl)pseudouridine (m1acp3Y)
dihydrouridine (D)
cytidine (C)
5-methylcytidine (m5C)
2'-O-methylcytidine (Cm)
N4-acetylcytidine (ac4C)
2-thiocytidine (s2C)
5-formylcytidine (f5C)

66 known in eukaryotes

13 in eukaryotic mRNA

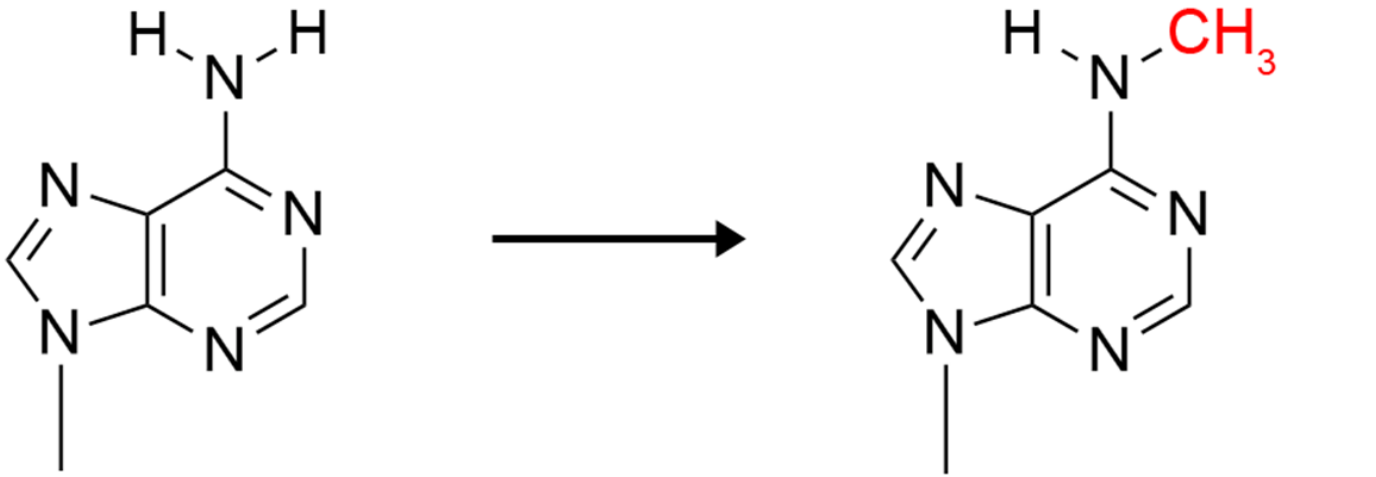
RNA Modification Database, 2013

Modified from Maria Basanta-Sanchez

Outline

- **Background**
- **RNA Modification Functions**
- **E4 Gaps and Opportunities**
- **E4 Program Components & Impact**

m6A: The 5th base in mRNA

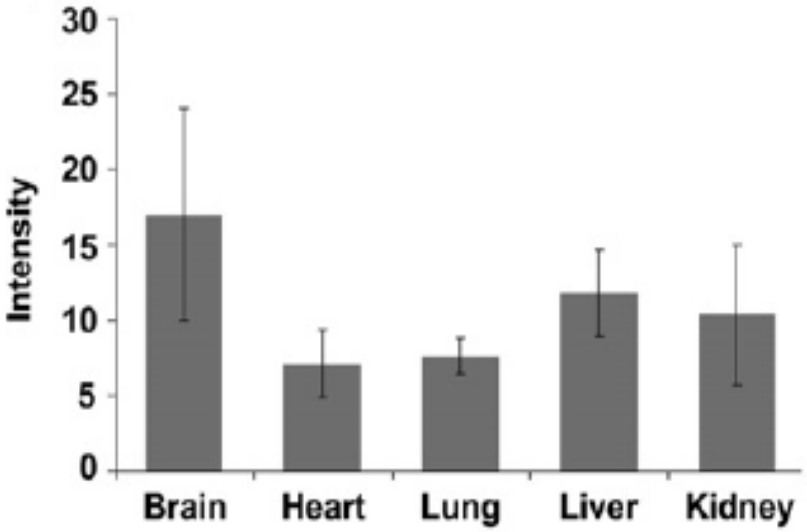


Adenosine **N^6-methyladenosine**

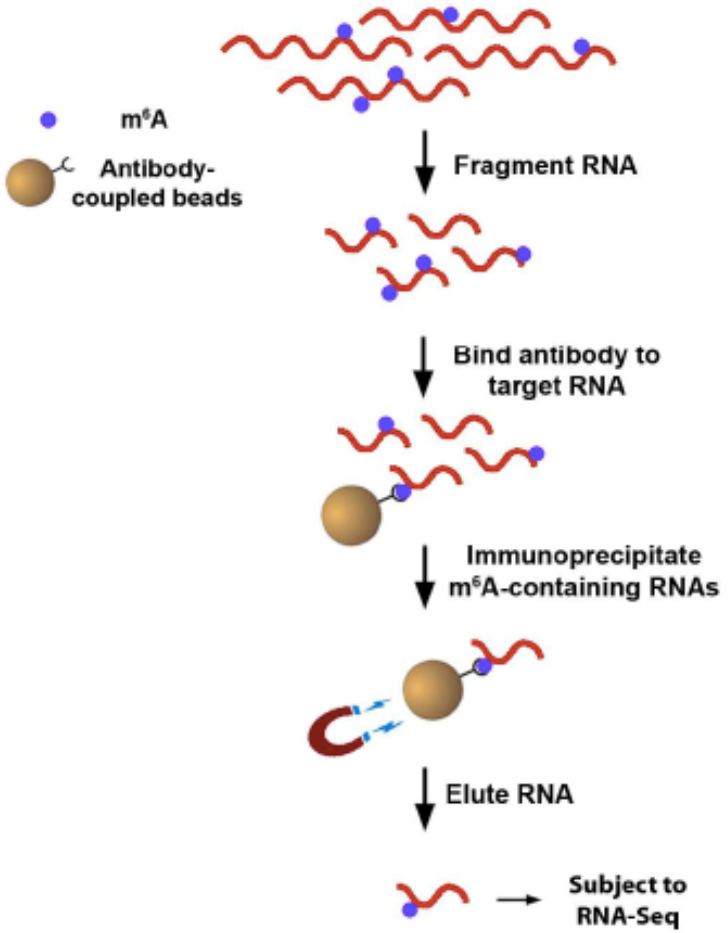
- **Discovered in 1975**
- **Base pairing is not affected, so not easy to detect m6A**
- **Found in tRNA, snRNAs, ribosomal RNA, one mRNA**
- **Interest in m6A languished**

Adapted from Samie Jaffrey

m6A: Antibodies and Transcriptome-wide Mapping



Methylated RNA IP-Seq (MeRIP-seq)

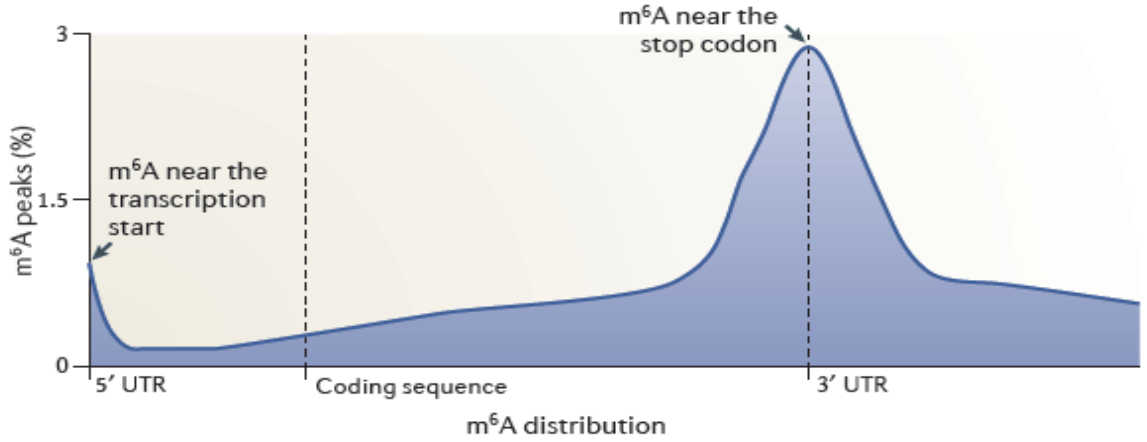


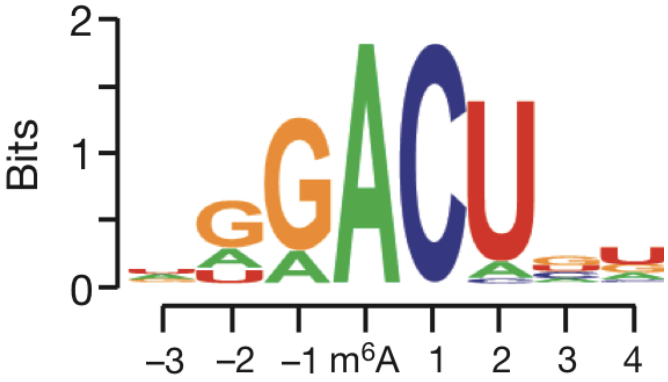
Meyer et al., Cell, 2012

Where is the m6A modification found?

>7,000 genes encode m6A methylated mRNAs

>400 m6A peaks mapped to non-coding RNAs

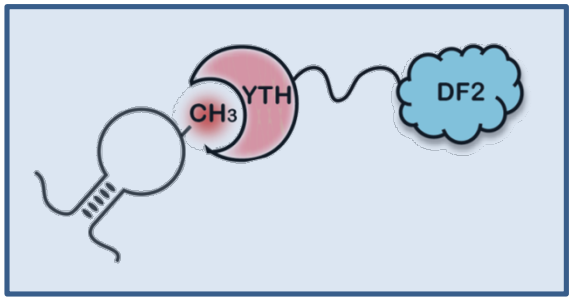




Nature **2012**, *485*, 201
Cell **2012**, *149*, 1635

m^{6}A: Readers, Writers, and Erasers (RWEs)

Writers: METTL3/14/WTAP
m6A-methyltransferase



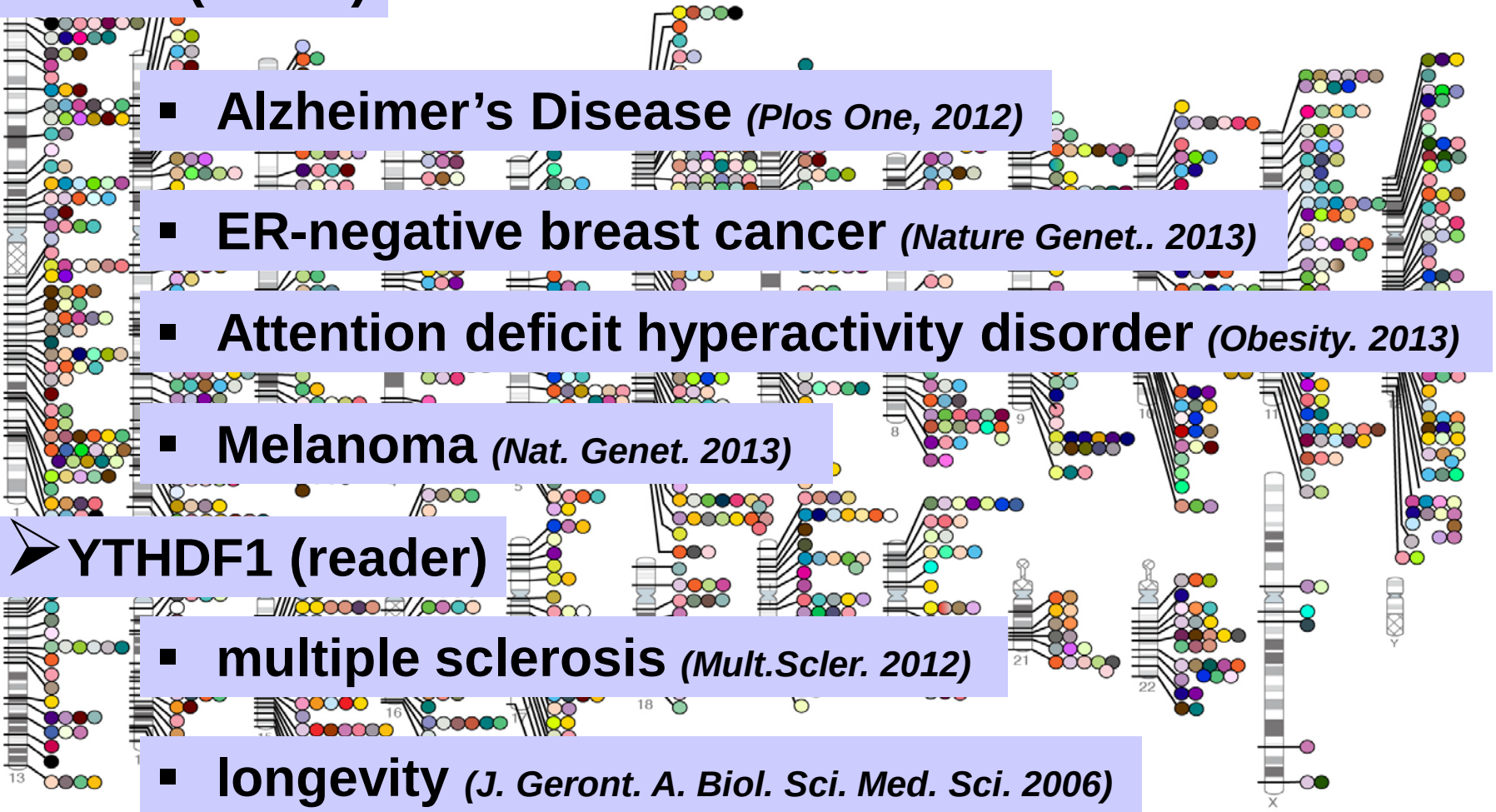
Readers: YTHDF1,2,3

Erasers: FTO and ALKBH5
m6A demethylases

Nature Chem. Biol. **2011**, *7*, 885
Mol. Cell. **2013**, 18.

m6A, RWEs and Human Disease

- **FTO (writer)**
  - **Alzheimer's Disease** *(Plos One, 2012)*
  - **ER-negative breast cancer** *(Nature Genet.. 2013)*
  - **Attention deficit hyperactivity disorder** *(Obesity. 2013)*
  - **Melanoma** *(Nat. Genet. 2013)*
- **YTHDF1 (reader)**
  - **multiple sclerosis** *(Mult.Scler. 2012)*
  - **longevity** *(J. Geront. A. Biol. Sci. Med. Sci. 2006)*



2011 2nd Quarter

m6A: A Diversity of Functions

nature

Sciencexpress Reports

m^{6}A mRNA methylation facilitates resolution of naïve pluripotency toward differentiation

Shay Geula,[1*] Sharon Moshitch-Moshkovitz,[2*] Dan Dominissini,[3*] Abed AlFatah Mansour,[1*] Nitzan Kol,[2] Mali Salmon-Divon,[2] Vera Hershkovitz,[2] Eyal Peer,[2] Nofar Mor,[1] Yair S. Manor,[1] Moshe Shay Ben-Haim,[2] Eran Eyal,[2] Sharon Yunger,[2] Yishay Pinto,[4] Diego Adhemar Jaitin,[5] Sergey Viukov,[1] Yoach Rais,[1] Vladislav Krupalnik,[1] Elad Chomsky,[1] Mirie Zerbib,[1] Itay Maza,[1] Yoav Rechavi,[1] Rada Massarwa,[1] Suhair Hanna,[1,6] Ido Amit,[5] Erez Y. Levanon,[4] Ninette Amariglio,[2,4] Noam Stern-Ginossar,[1] Noa Novershtern,[1†‡] Gideon Rechavi,[2†‡] Jacob H. Hanna[1†‡]

[1]The Department of Molecular Genetics, Weizmann Institute of Science, Rehovot, Israel. [2]Cancer Research Center, Chaim

[3]Chemica

[4]The Univ

Japan

*Correspondence: okamuran@pharm.kyoto-u.ac.jp

http://dx.doi.org/10.1016/j.cell.2013.10.026

rely on some of these factors (Fig. 1A and fig. S1). Regulators that specifically inhibited the stability of Oct4-GFP+ primed cells included the epigenetic repressors Dnmt1, Eed, and Suz12 polycomb components, Mbd3, and N^6-adenosine methyltransferase Mettl3, a component of the N^6-methyladenosine (m^{6}A) mRNA methylating complex (*3*) (Fig. 1A and fig. S1).

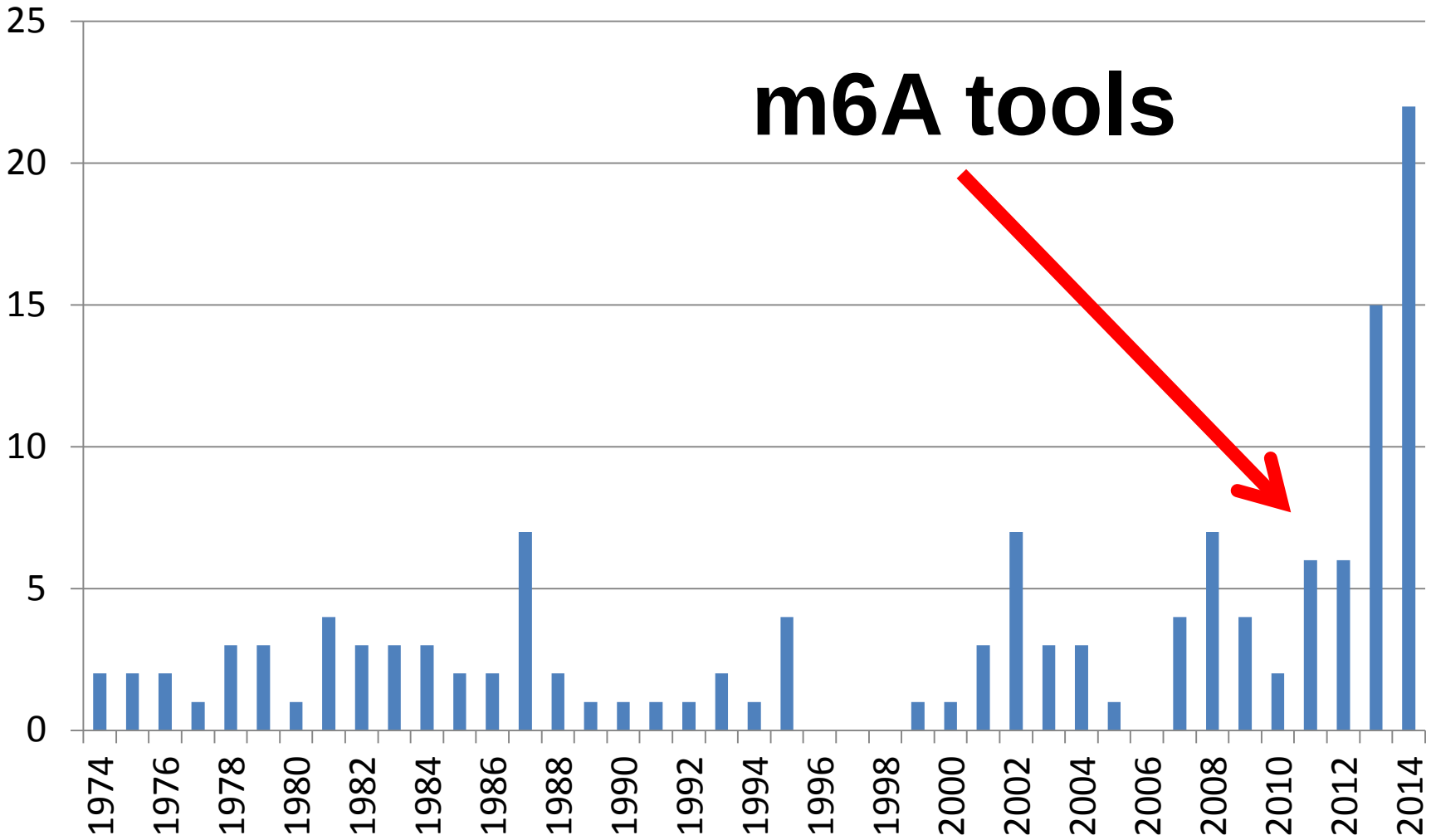
We subsequently focused on the role of m^{6}A in pluripotency transitions since the biological role of RNA modifications is only starting to be unveiled (*4*). m^{6}A is an RNA modification catalyzed by Mettl3 (methyl trans-

[1]Department of Chemistry, Institute for Biophysical Dynamics, The University of Chicago, 929 East 57th Street, Chicago, IL 60637, USA

[2]Genome Structure & Stability Group, BIG CAS-OSLO Genome Research Cooperation, Disease Genomics and Individualized Medicine Laboratory, Beijing Institute of Genomics, Chinese Academy of Sciences, No.7 Beitucheng West Road, Chaoyang District, Beijing 100029

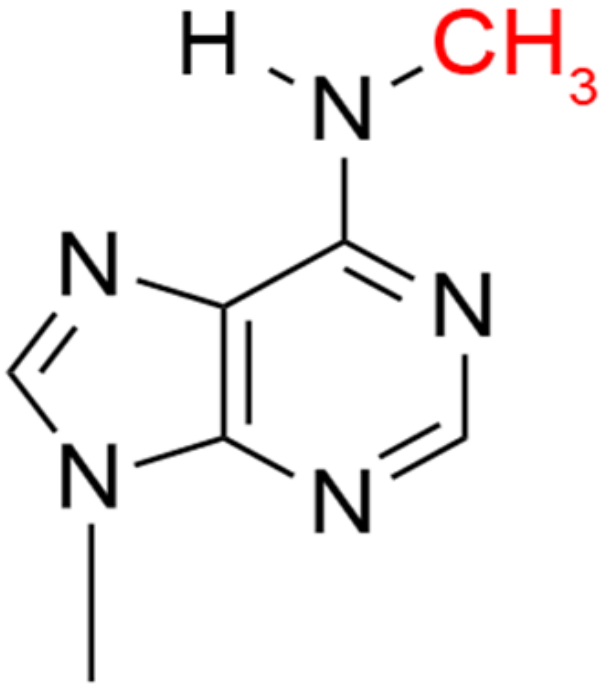
16, 2015

m6A Publication History



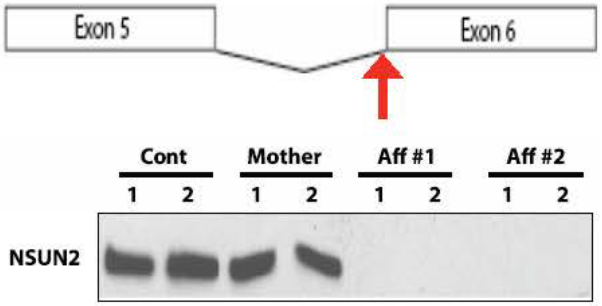
Pubmed search using 6-methyladenosine

Is m6A special?

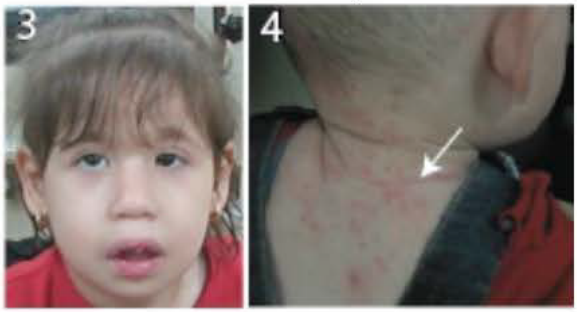


***N*^{6}-methyladenosine**

The 5mC Methyltransferase NSUN2 and Intellectual Disability



Dubowitz-like syndrome



Neurological	microcephaly, intellectual disability, behavioral deficit, abnormal gait
Development	growth retardation
Skin	Cutaneous abnormalities, sparse hair
Fertility	ovaries atrophy

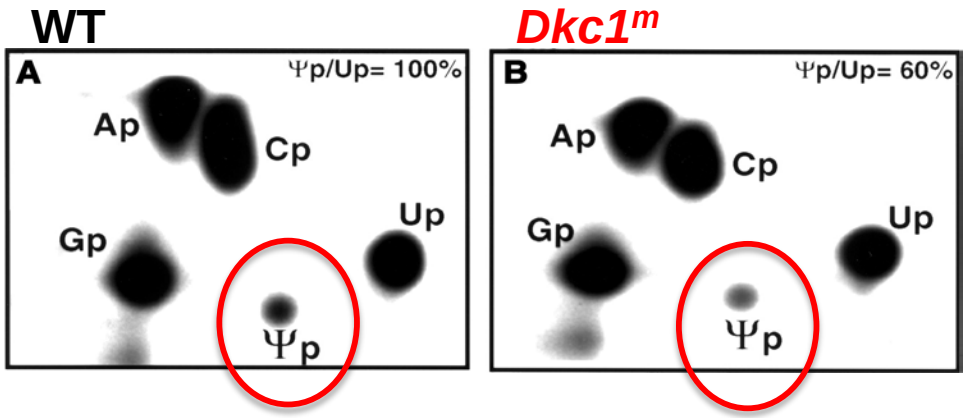
Sandra Blanco, JAHG 2012
JMG 2012, Clin. Genet. 2014

Pseudouridine (ψ)

DKC1 mutations reduce Ψ levels in rRNA

Dyskeratosis congenita:

- **often fatal**
- **premature aging**
- **bone marrow failure**
- **increased suscept. to cancer**



DKC1 and p27 genetically interact in pituitary tumorigenesis

Ψ present in mRNA (238 coding transcripts) and non-coding RNA

Regulated by environmental signals

(Science, 2003; Cancer Res. 2010; Nature 2014; Cell 2014)

Circular RNAs

ARTICLE

doi:10.1038/nature11928

Circular RNAs are a large class of animal RNAs with regulatory potency

Sebastian Memczak[1]*, Marvin Jens[1]*, Antigoni Elefsinioti[1]*, Francesca Torti[1]*, Janna Krueger[2], Agnieszka Rybak[1], Luisa Maier[1], Sebastian D. Mackowiak[1], Lea H. Gregersen[3], Mathias Munschauer[3], Alexander Loewer[4], Ulrike Ziebold[1], Markus Landthaler[3], Christine Kocks[1], Ferdinand le Noble[2] & Nikolaus Rajewsky[1]

Circular RNAs (circRNAs) in animals are an enigmatic class of RNA with unknown function. To explore circRNAs systematically, we sequenced and computationally analysed human, mouse and nematode RNA. We detected thousands of well-expressed, stable circRNAs, often showing tissue/developmental-stage-specific expression. Sequence analysis indicated important regulatory functions for circRNAs. We found that a human circRNA, antisense to the cerebellar degeneration-related protein 1 transcript (CDR1as), is densely bound by microRNA (miRNA) effector complexes and harbours 63 conserved binding sites for the ancient miRNA miR-7. Further analyses indicated that CDR1as functions to bind miR-7 in neuronal tissues. Human CDR1as expression in zebrafish impaired midbrain development, similar to knocking down miR-7, suggesting that CDR1as is a miRNA antagonist with a miRNA-binding capacity ten times higher than any other known transcript. Together, our data provide evidence that circRNAs form a large class of post-transcriptional regulators. Numerous circRNAs form by head-to-tail splicing of exons, suggesting previously unrecognized regulatory potential of coding sequences.

>2000 human circular RNAs predicted

Mature messenger RNAs are linear molecules with 5
that reflect start and stop of the RNA polymerase or

LETTER

doi:10.1038/nature11993

Natural RNA circles function as efficient microRNA sponges

Thomas B. Hansen[1], Trine I. Jensen[1], Bettina H. Clausen[2], Jesper B. Bramsen[1,3], Bente Finsen[2], Christian K. Damgaard[1] & Jørgen Kjems[1,3]

MicroRNAs (miRNAs) are important post-transcriptional regulators of gene expression that act by direct base pairing to target

miRNA-mediated endocleavage[12,13] (Supplementary Fig. 1b). The miR-671 target site exhibits near-perfect complementarity and very little

m6A
5mC
pseudouridine

Table 1 - List of Base Modifications Covered by Claims

Abbreviation	Chemical name
$m^1acp^3\Psi$	1-methyl-3-(3-amino-3-carboxypropyl) pseudouridine
m1A	1-methyladenosine
m^1G	1-methylguanosine
m1I	1-methylinosine
$m^1\Psi$	1-methylpseudouridine
m^1Am	1,2'-O-dimethyladenosine
m^1Gm	1,2-O-dimethylguanosine
m1Im	1,2'-O-dimethylinosine
m^2A	2-methyladenosine
m2,6A	2-methyl-N6-(cis-hydroxyisopentenyl) adenosine
ms^2hn^6A	2-methylthio-N6-hydroxynorvalyl carbamoyladenosine
m2,6A	2-methylthio-N6-isopentenyladenosine
ms^2m^6A	2-methylthio-N6-methyladenosine
m2,6A	2-methylthio-N6-threonyl carbamoyladenosine
s^2Um	2-thio-2'-O-methyluridine
s2C	2-thiocytidine
s^2U	2-thiouridine
Am	2'-O-methyladenosine
Cm	2'-O-methylcytidine
Gm	2'-O-methylguanosine
Im	2'-O-methylinosine
m5Um	2'-O-methyl-5-methyluridine
Um	2'-O-methyluridine
Ar(p)	2'-O-ribosyladenosine (phosphate)
Gr(p)	2'-O-ribosylguanosine (phosphate)
m1acp3U	3-(3-amino-3-carboxypropyl)uridine
m^3C	3-methylcytidine
m3C	3-methylcytidine
m^3U	3-methyluridine
m3Um	3,2'-O-dimethyluridine
imG-14	4-demethylwyosine
s4U	4-thiouridine
chm^5U	5-(carboxyhydroxymethyl)uridine
mchm5U	5-(carboxyhydroxymethyl)uridine methyl ester
inm^2s^2U	5-(isopentenylaminomethyl)-2-thiouridine
mnm5s2U	5-(methylaminomethyl)-2-thiouridine
inm^5U	5-(isopentenylaminomethyl)uridine
mnm5U	5-(methylaminomethyl)uridine
ncm^5Um	5-carbamoylmethyl-2'-O-methyluridine
ncm5U	5-carbamoylmethyluridine
$cmnm^5Um$	5-carboxymethylaminomethyl- 2'-O-methyluridine
cmnm5s2U	5-carboxymethylaminomethyl-2-thiouridine
$cmnm^5U$	5-carboxymethylaminomethyluridine
cm5U	5-carboxymethyluridine
f^5Cm	5-formyl-2'-O-methylcytidine
f5C	5-formylcytidine
hm^5C	5-hydroxymethylcytidine
h5U	5-hydroxyuridine
mcm^5s^2U	5-methoxycarbonylmethyl-2-thiouridine
mcm5s2U	5-methoxycarbonylmethyl-2-thiouridine
mcm^5U	5-methoxycarbonylmethyluridine
mo5U	5-methoxyuridine
m^5s^2U	5-methyl-2-thiouridine
mnm5s2U	5-methylaminomethyl-2-thiouridine
mnm^5s^2U	5-methylaminomethyl-2-thiouridine
mnm5U	5-methylaminomethyluridine
m^5C	5-methylcytidine
m5D	5-methyldihydrouridine
m^5U	5-methyluridine
m5Um	5,2'-O-dimethyluridine
τm^5U	5-taurinomethyluridine
m5s2U	5-taurinomethyl-2-thiouridine
m^5Um	5,2'-O-dimethyluridine
m5U	5-uridinoxyacetic acid methyl ester
preQ1	7-cyano 7-deaza guanosine
m7G	7-methylguanosine
G^+	archaeosine
D	dihydrouridine
oQ	epoxyqueuosine
galQ	galactosyl-queuosine
OHyW	hydroxywybutosine
I	inosine
imG2	isowyosine
k2C	lysidine
manQ	mannosyl-queuosine
m1A	methyladenosine
m^1G	N^1-methylguanosine
m1acp3Y	N^1-(3-amino-3-carboxypropyl) pseudouridine
m^2_2G	N^2,7-dimethylguanosine
m2,2G	N^2,N^2-dimethylguanosine
$m^2,^2G$	N^2,N^2-dimethylguanosine
m2,2,7G	N^2,N^2,7-trimethylguanosine
$m^{2,2,7}G$	N^2,N^2,7-trimethylguanosine
ac^4C	N^4-acetylcytidine
m4C	N^4-methylcytidine
m^4Cm	N^4,2'-O-dimethylcytidine
m4,4c	N^4,N^4-dimethylcytidine
io^6A	N^6-(cis-hydroxyisopentenyl)adenosine
g6A	N^6-glycinylcarbamoyladenosine
g^6A	N^6-glycinylcarbamoyladenosine
i^6A	N^6-isopentenyladenosine
m1g6A	N^6-methyl-N6-threonylcarbamoyladenosine
m^6A	N^6-methyladenosine
m6t6A	N^6-(cis-hydroxyisopentenyl)-adenosine
m^6Am	N^6,2'-O-dimethyladenosine
m6t6A	N^6-threonylcarbamoyladenosine
m^6,Am	N^6,N^6,2'-O-trimethyladenosine
o5U	uridine-5-oxyacetic acid
Ψ	pseudouridine
Q	queuosine
OHyW	undermodified hydroxywybutosine

110 known RNA mods

How Do RNA Modifications Impact RNA Classes and Functions?

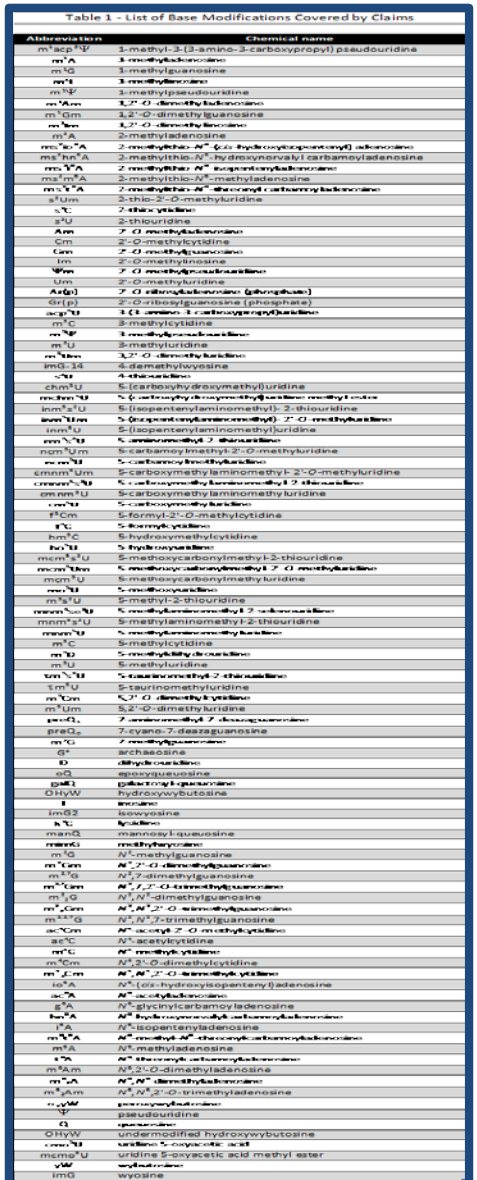


Table 1 - List of Base Modifications Covered by Claims

Abbreviation	Chemical name
m¹acp³Ψ	1-methyl-3-(3-amino-3-carboxypropyl) pseudouridine
m¹A	1-methyladenosine
m¹G	1-methylguanosine
m¹I	1-methylinosine
m¹Ψ	1-methylpseudouridine
m¹Am	1,2'-O-dimethyladenosine
m¹Gm	1,2'-O-dimethylguanosine
m¹Im	1,2'-O-dimethylinosine
m²A	2-methyladenosine
ms²i⁶A	2-methylthio-N⁶-(cis-hydroxyisopentenyl) adenosine
ms²t⁶A	2-methylthio-N⁶-hydroxynorvalylcarbamoyladenosine
ms²t⁶A	2-methylthio-N⁶-isopentenyladenosine
m₂²m²A	2-methylthio-N⁶-methyladenosine
ms²t⁶A	2-methylthio-N⁶-threonylcarbamoyladenosine
s²Um	2-thio-2'-O-methyluridine
s²C	2-thiocytidine
s²U	2-thiouridine
Am	2'-O-methyladenosine
Cm	2'-O-methylcytidine
Gm	2'-O-methylguanosine
Im	2'-O-methylinosine
Ψm	2'-O-methylpseudouridine
Um	2'-O-methyluridine
Ar(p)	2'-O-ribosyladenosine (phosphate)
Gr(p)	2'-O-ribosylguanosine (phosphate)
acp³U	3-(3-amino-3-carboxypropyl)uridine
m³C	3-methylcytidine
m³Ψ	3-methylpseudouridine
m³U	3-methyluridine
m⁴Cm	4,2'-O-dimethylcytidine
imG-14	4-demethylwyosine
s⁴U	4-thiouridine
chm⁵U	5-(carboxyhydroxymethyl)uridine
mchm⁵U	5-(carboxyhydroxymethyl)uridine methyl ester
inm⁵s²U	5-(isopentenylaminomethyl)-2-thiouridine
inm⁵Um	5-(isopentenylaminomethyl)-2'-O-methyluridine
inm⁵U	5-(isopentenylaminomethyl)uridine
nm⁵s²U	5-aminomethyl-2-thiouridine
ncm⁵Um	5-carbamoylmethyl-2'-O-methyluridine
ncm⁵U	5-carbamoylmethyluridine
cmnm⁵Um	5-carboxymethylaminomethyl-2'-O-methyluridine
cmnm⁵s²U	5-carboxymethylaminomethyl-2-thiouridine
cmnm⁵U	5-carboxymethylaminomethyluridine
cm⁵U	5-carboxymethyluridine
f⁵Cm	5-formyl-2'-O-methylcytidine
f⁵C	5-formylcytidine
hm⁵C	5-hydroxymethylcytidine
ho⁵U	5-hydroxyuridine
mcm⁵s²U	5-methoxycarbonylmethyl-2-thiouridine
mcm⁵Um	5-methoxycarbonylmethyl-2'-O-methyluridine
mcm⁵U	5-methoxycarbonylmethyluridine
mo⁵U	5-methoxyuridine
m⁵s²U	5-methyl-2-thiouridine
mnm⁵se²U	5-methylaminomethyl-2-selenouridine
mnm⁵s²U	5-methylaminomethyl-2-thiouridine
mnm⁵U	5-methylaminomethyluridine
m⁵C	5-methylcytidine
m⁵D	5-methyldihydrouridine
m⁵U	5-methyluridine
Um⁵U	5-taurinomethyl-2-thiouridine
τm⁵U	5-taurinomethyluridine
m⁵Cm	5,2'-O-dimethylcytidine
preQ₀	7-aminomethyl-7-deazaguanosine
preQ₁	7-cyano-7-deazaguanosine
m⁷G	7-methylguanosine
G⁺	archaeosine
D	dihydrouridine
oG	epoxyqueuosine
galQ	galactosyl-queuosine
OHyW	hydroxywybutosine
I	inosine
imG2	isowyosine
k²C	lysidine
manQ	mannosyl-queuosine
mimG	methylguanosine
m¹G	N¹-methylguanosine
m¹Gm	N¹,2'-O-dimethylguanosine
m²²G	N²,7-dimethylguanosine
m²²Gm	N²,7,2'-O-trimethylguanosine
m²₂G	N²,N²-dimethylguanosine
m²₂Gm	N²,N²,2'-O-trimethylguanosine
m²²²G	N²,N²,7-trimethylguanosine
ac⁴Cm	N⁴-acetyl-2'-O-methylcytidine
ac⁴C	N⁴-acetylcytidine
m⁴C	N⁴-methylcytidine
m⁴Cm	N⁴,2'-O-dimethylcytidine
m⁴₂Cm	N⁴,N⁴,2'-O-trimethylcytidine
io⁶A	N⁶-(cis-hydroxyisopentenyl)adenosine
ac⁶A	N⁶-acetyladenosine
g⁶A	N⁶-glycinylcarbamoyladenosine
hn⁶A	N⁶-hydroxynorvalylcarbamoyladenosine
i⁶A	N⁶-isopentenyladenosine
ms²t⁶A	N⁶-methyl-N⁶-threonylcarbamoyladenosine
m⁶A	N⁶-methyladenosine
t⁶A	N⁶-threonylcarbamoyladenosine
m⁶Am	N⁶,2'-O-dimethyladenosine
m⁶₂A	N⁶,N⁶-dimethyladenosine
m⁶₂Am	N⁶,N⁶,2'-O-trimethyladenosine
n⁵cmW	peroxywybutosine
Ψ	pseudouridine
Q	queuosine
OHyW	undermodified hydroxywybutosine
cmo⁵U	uridine 5-oxyacetic acid
mcmo⁵U	uridine 5-oxyacetic acid methyl ester
yW	wybutosine
imG	wyosine

RNA Classes:

mRNA, tRNA, rRNA, microRNA, siRNA, piRNA, long non-coding RNA, snoRNA, etc.

RNA Functions:

Protein translation and localized translation
Splicing
Structural scaffold
Chromatin recruitment
Ribozymes
Environmental sensing
Post-transcriptional regulation
Gene silencing
Defense against germline transposons
Intercellular signaling
RNA modification

Outline

- **Background**
- **RNA Modification Functions**
- **E4 Gaps and Opportunities**
- **E4 Program Components & Impact**

Scientists Consulted

Paul Agris, RNA Institute
Cheryl Arrowsmith, U. of Toronto
Thomas Begley, U of Albany
Howard Char
Jim Eberwine
Dan Fabris, SUNY Albany
Chuan He, U of Chicago
Samie Jaffrey, Cornell U
Stuart Legrice, NCI

30 RFI responses

Patrick Limbach, U of Cincinnati
Chris Mason, Cornell U
Thomas Misteli, NCI
J of Rochester
, Austr. Natl. U
Tamar Schlik, NYU
Gabriele Varani, U of Washington
Kevin Weeks, U of North Carolina
Jamie Williamson, Scripps
Crystal Zhao, Sanford Burnham

RNA structure-function
RNA modifying enzymes
Small molecules
Antibody development
Mass spectrometry
Chemistry

Obesity
Neurodevelopmental disorders
Cancer biology
Embryonic stem cells
Single cell imaging
Computational technology

What are the Scientific Needs in Epitranscriptomics?

A. Tools

- **Affinity reagents**
- **Small molecule modulators**
- **Computational tools**

B. Technologies

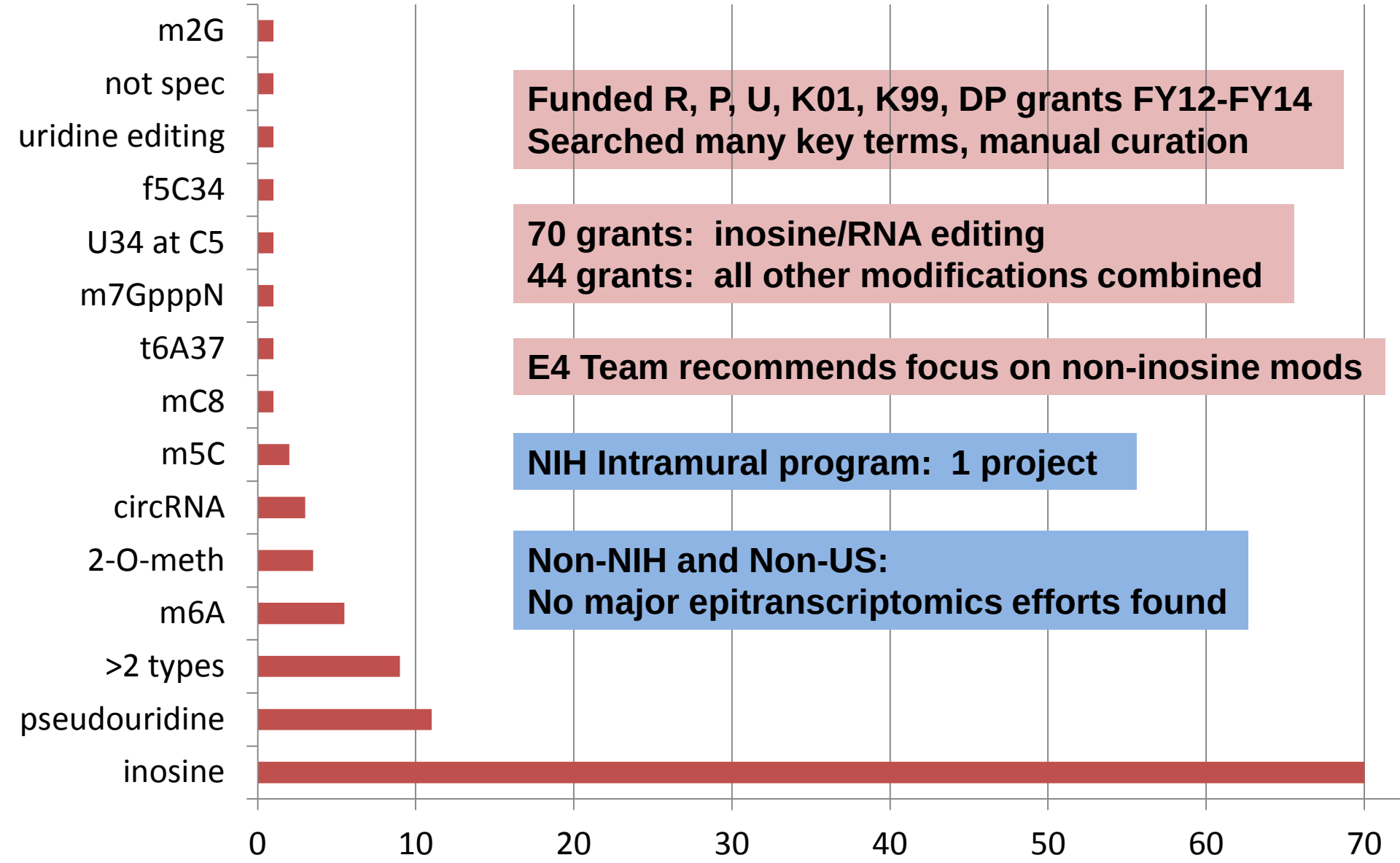
- **RNA mod low abundance detection, genomewide, single base resolution**
- **Detecting effects of RNA mods on RNA structure**
- **Imaging of RNA modifications**
- **Manipulation of RNA modifications**

C. Survey of the RNA modification landscape

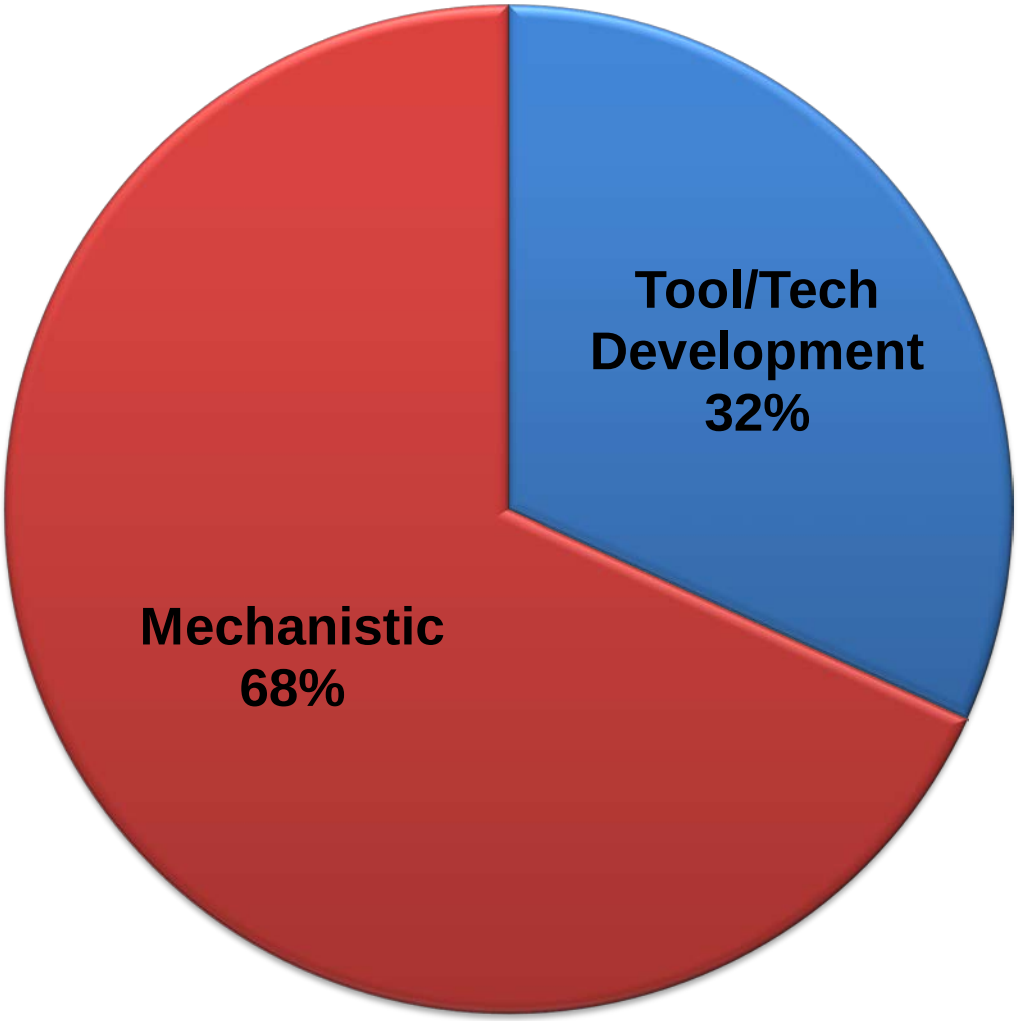
- **Discovery of novel RNA mods, RWEs**
- **Inventory of known RNA mods, RWEs**

D. Functions of RNA mods/RWEs in biological processes, health, and disease

NIH Portfolio in RNA Modifications



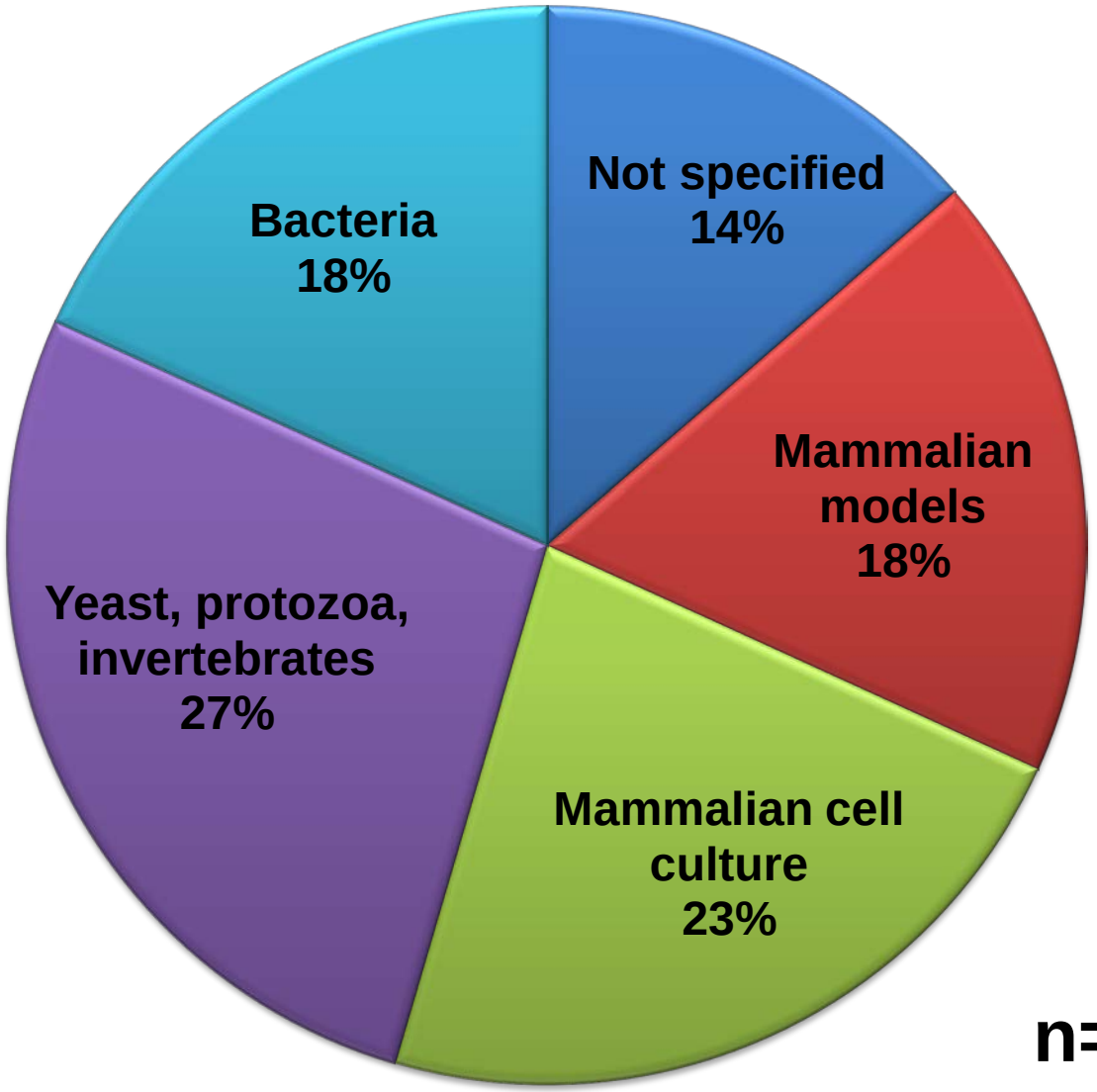
Limited NIH Support for RNA Mod Tool & Technology Development



Half are SBIR/STTR grants from 2014 NIDA RFA

n=44 grants

Comparatively Few Studies Using *in vivo* Mammalian Models



n=44 grants

Outline

- **Background**
- **RNA Modification Functions**
- **E4 Gaps and Opportunities**
- **E4 Program Components & Impact**

Key E4 Program Themes

A. Tools

- Affinity reagents
- Small molecule modulators
- Computational tools

B. Technologies

- RNA mod low abundance detection, genomewide, single base resolution
- Detecting effects of RNA mods on RNA structure
- Imaging of RNA modifications
- Manipulation of RNA modifications

CF NCI Proposal

C. Catalog of the RNA modification landscape

- Discovery of novel RNA mods, RWEs
- Inventory of known RNA mods, RWEs

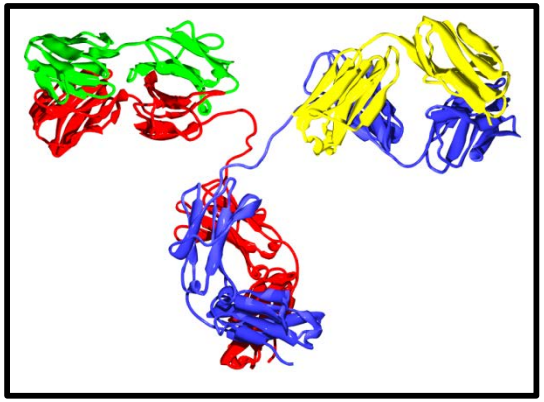
D. Functions of RNA mods/RWEs in biological processes, health, and disease

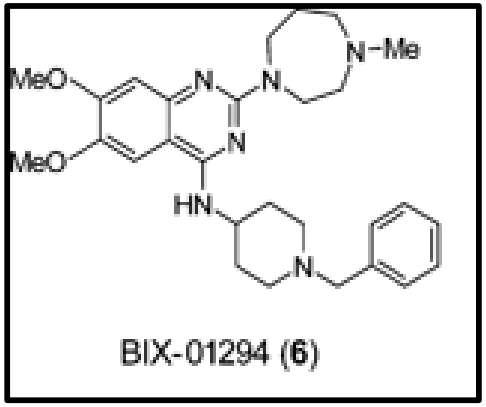
Epitranscriptomic Tools

GAP: Lack of essential tools to monitor and manipulate most RNA modifications

INITIATIVE: Develop enabling and transformative:

- **affinity reagents**
- **genetic tools and models**
- **small molecule modulators**



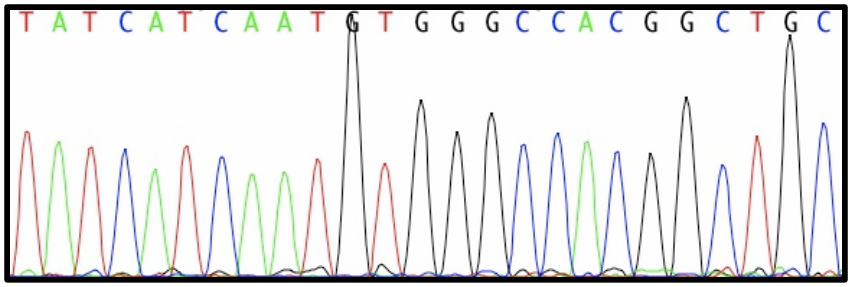


MECHANISM: CF-supported UH2 and IC-supported SBIR/STTR

DELIVERABLES: Defined # of tools for a diversity of RNA mods

Epitranscriptomic Technologies

GAP: **Lack critical technologies to monitor & manipulate most RNA mods**



INITIATIVE: **Develop enabling and transformative technologies:**

- **Detect low levels of RNA mods**
- **Transcriptome-wide RNA mod assays at single nucleotide resolution**
- **Tools to image and manipulate RNA mods or RWEs in vivo**
- **Computational strategies---predict effects on RNA structure/function**

MECHANISM: **CF-supported UH2/UH3 and IC-supported SBIR/STTR**

DELIVERABLES: **Defined # of technologies for a diversity of RNA mods**

Novel Epitranscriptomic Components

GAP: Only partial inventory of RNA mod reader, writer, and eraser (RWE) (proteins or ribozymes)

INITIATIVE: Support discovery or confirmation of currently unknown readers, writers, erasers, and modifications in any model system (UH2)

Follow up work on function and role in vertebrate systems (UH3)

MECHANISM: UH2/UH3

DELIVERABLES: Defined # new RWEs or RNA mods.

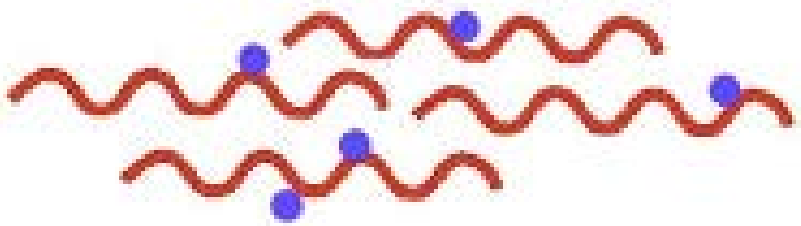
Epitranscriptome Catalog: Phase I and II

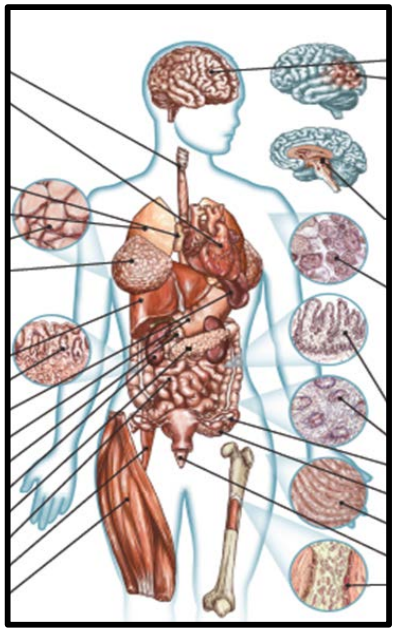
GAP: Very limited knowledge of the landscape (abundance and dynamics) of known RNA Mods in different tissues and RNA biotypes

INITIATIVE: <u>Phase I</u>. Exploit current and emerging technologies to inventory <u>known</u> RNA Mods and RWEs

<u>Phase II</u>. Expand to catalog key modified RNAs quantitatively at single nucleotide resolution (the Epitranscriptome). Focus on:

- RNA classes (e.g. mRNA, tRNA, microRNA, long non-coding RNA)
- Disease-relevant mammalian cell types and tissues
- Environmental exposures and RNA mod dynamics





MECHANISM: U01

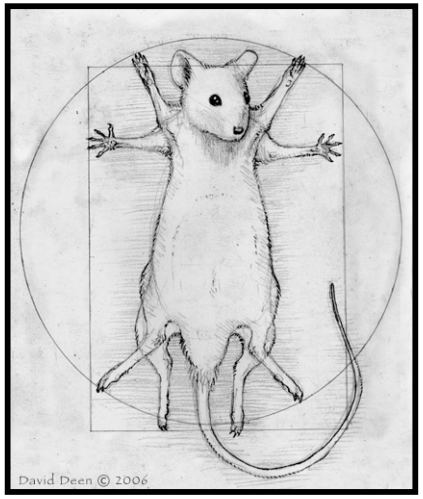
DELIVERABLES: The Epitranscriptome Catalog for a specified # of RNA mods at single nucleotide resolution for a defined #of tissues/conditions.

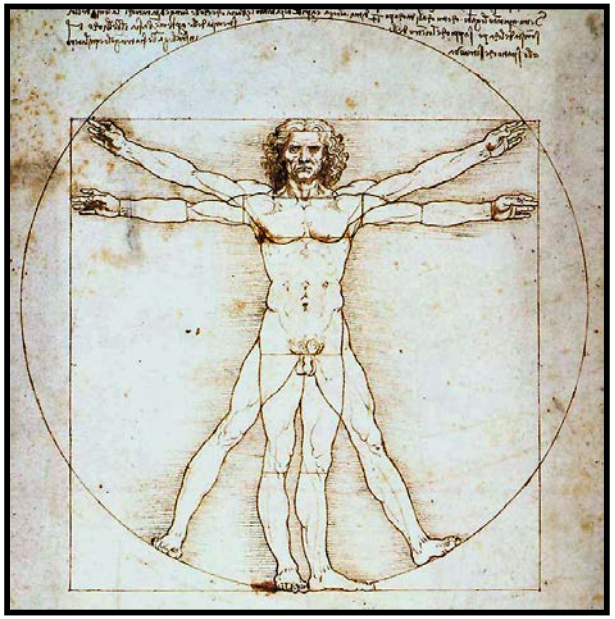
E4 Demonstration Projects

GAP: **We do not fully understand the potential roles of RNA mods in biological processes and disease.**

INITIATIVE: **Demonstration projects exploiting E4 Program deliverables to investigate the function of RNA modifications.**

Emphasize RNA mod quantitation and dynamics





MECHANISM: **U01 or R01 Competitive Revision**

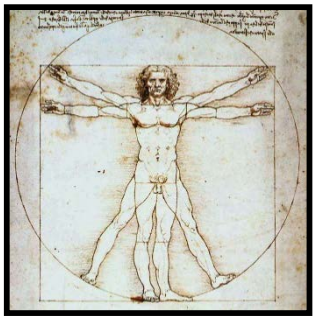
DELIVERABLES: **Characterize the roles of a specific # of RNA mods in a diverse selection of biological processes and diseases**

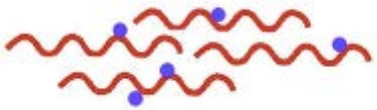
Epitranscriptomics Data Hub

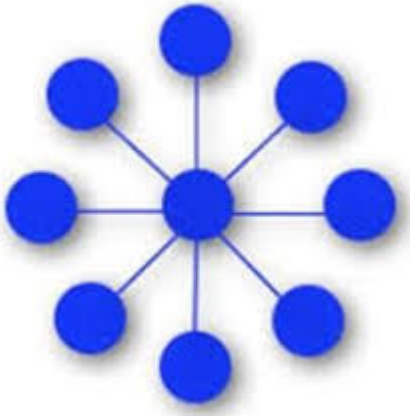
GAP: No centralized place to integrate RNA mod knowledge with disease and other datatypes

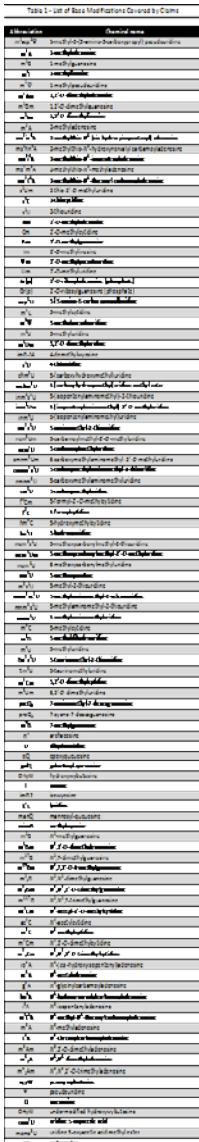
INITIATIVE: An Epitranscriptomics Data Hub for access to E4 deliverables:

- protocols, tools, technologies, publications
- links to existing RNA mod resources
- data coordination and access to Epitranscriptome catalog
- RNA mod analysis tools and guidance
- scientific outreach (meetings, workshops, etc)
- Link RNA mod phenotypes to human disease









MECHANISM: U54

DELIVERABLES: Website, defined # of protocols, data sets, outreach activities
Enable the scientific community to exploit E4 deliverables

Set aside/ YR

Phase I Epitranscriptome Catalog (U01); Phase II Epitranscriptome Catalog (U01)	$3M, $6M
Discovery: Novel RNA Mods, Writers, Erasers, and Readers (UH2/UH3); Genetic Models of Human Disease (U01)	$4M, $4M
Tech Dev: Improved Computational Tools for RNA Mods (UH2/UH3); RNA Mod Dynamics: Demonstration Projects in Health, and Disease (U01)	$3M, $4M
Tech Dev: Improved Detection/Sequencing of RNA Mods (UH2/UH3); Tech Dev: Improved imaging and Manipulation Tools for RNA Mods (UH2/UH3)	$5M, $4M
Data Coordination, Analysis, Outreach (U54)	$2M
Tools: Affinity Reagents (UH2); Tools: Small Molecule Modulators (UH2/UH3); Tech Dev: Exploiting RNA Mods to Develop. Therapeutics (UH2)	$3M, $3M, $3M
Develop Tools and Technologies for Epitranscriptomics (IC-supported SBIR/STTR)	$2M

CF $	18M	20M	20M	23M	23M	22M	22M	22M	16M	16M	$185M CF
IC $	2M	2M	2M	2M	2M	2M	2M	2M	2M	2M	$20M IC
Ttl $	20M	22M	22M	25M	25M	24M	24M	24M	18M	18M	$205M Total

Phase I | Phase II

2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	Set aside/ YR
Phase I Epitranscriptome Catalog (U01)				Phase II Epitranscriptome Catalog (U01)						$3M, $6M
										
Genetic Models of Human Disease (U01)										$4M
RNA Mod Dynamics: Demonstration Projects in Health, and Disease (U01)										$3M, $4M
Tech Dev: Improved imaging and Manipulation Tools for RNA Mods (UH2/UH3)										$5M, $4M
, Analysis, Outreach (U54)										$2M
Molecule UH2/UH3)					Tech Dev: Exploiting RNA Mods to Develop. Therapeutics (UH2)					$3M, $3M, $3M
nscriptomics (IC-supported SBIR/STTR)										$2M

CF $	18M	20M	20M	23M	23M	22M	22M	22M	16M	16M	$185M CF
IC $	2M	2M	2M	2M	2M	2M	2M	2M	2M	2M	$20M IC
Ttl $	20M	22M	22M	25M	25M	24M	24M	24M	18M	18M	$205M Total

Leveraging Other Projects and Resources

- ENCODE/Epigenomics/4DN? Integrate E4 data with their datasets
- BD2K: Align E4 data sets into future BD2K framework
- GWAS? Integration of RNA modification sites with SNPs and gene expression levels for particular tissues and diseases
- GTEx: Leverage GTEx post-mortem tissues samples for inventory if sufficient quantities and collection
- exRNA: Leverage body fluid samples from exRNA program to explore RNA mods in exRNAs
- SGC: Work together to develop small molecule modulators for Epitranscriptomic readers, writers, and erasers
- International: German funding agencies have expressed interest in this topic. Others may also be interested.
- KOMP: use to help develop mouse models of E4 disease

Beautify!

E4 Program Outcomes

New RNA Mod Tools & Technologies:

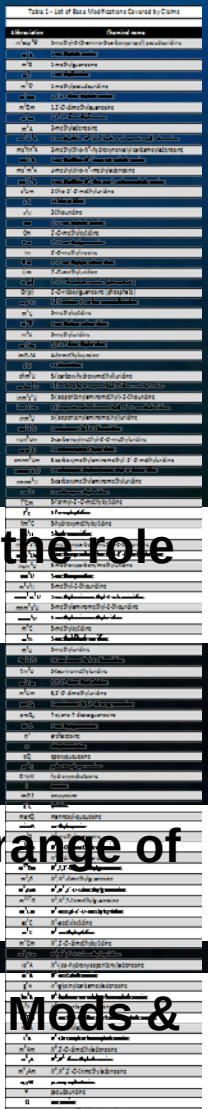
- **Catalyze the scientific community to investigate the role of RNA mods in biological & disease processes**

RNA Mod Discovery & Catalog:

- **See the RNA Mod “landscape” for the first time**
- **Enable the scientific community to generate hypotheses about the role of RNA mods in biological processes and diseases**

Overall:

- **Transform our understanding of the role of RNA Mods in wide range of critical biological & disease processes**
- **Provide a firm foundation to exploit new knowledge about RNA Mods & their function to prevent, diagnose, & treat disease**



Impact of finding a new modification?

update # hmC

266 hmC papers

Cell

Is an Epigenetic mark of Melanoma

Christine Guo Lian,[1,2,13] Yufei Xu,[1,13] Craig Ceol,[3,6] Feizhen Wu,[9] Allison Larson,[5] Karen Dresser,[7] Wenqi Xu,[9] Li Tan,[9] Yeguang Hu,[1] Qian Zhan,[2] Chung-wei Lee,[2] Di Hu,[1] Bill Q. Lian,[1,8] Sonja Kleffel,[5] Yijun Yang,[10] James Neiswender,[6]

Marian Mellén,[1,4] Pinar Ayata,[1,4] Scott Dewell,[2] Skirmantas Kriaucionis,[3,*] and Nathaniel Heintz[1,*]
[1]Laboratory of Molecular Biology, Howard Hughes Medical Institute

fC 5-formylcytosine

caC 5-carboxylcytosine

Joseph R. Ecker*, Patrice M. Milos*, Suneet Agarwal* & Anjana Rao*

mechanistic basis for transgenerational epigenetic inheritance.


QUESTIONS?