Title of proposed program: Enabling Exploration of the Eukaryotic Epitranscriptome (E^4)

What is the major obstacle/challenge/opportunity that the Common Fund should address?
Chemical modifications play a crucial role in the regulation of biological processes. For example, the function of a protein is often modulated by its stable tagging with different chemical groups, such as phosphates, sugars, or lipids, while specific chemical marks made along the chromatin (i.e., the DNA and/or its packaging proteins) can dial gene expression up or down. One area that lags behind most others in this context is the systematic characterization of all the chemical modifications that can befall RNA molecules (both coding and non-coding), sometimes referred to as the "epitranscriptome" (Saletore et al. Genome Biol. 13:175 (2012). A few covalent RNA modifications, such as 5'mRNA capping, alternative splicing, and polyadenylation, have been studied extensively. However, the functional roles of many other covalent RNA modifications remain unknown although they are likely to influence important parameters affecting RNA function, such as its stability, trafficking, localization, enzymatic/sensing/regulatory activity, and patterns of interactions with other molecules.

The RNA Modification Database indicates that there are at least 65 RNA modifications that occur in eukaryotic cells (Cantara et al., 2011, Nucl. Acids Res. 39:D195-201). Transfer and ribosomal RNA can be heavily modified and some of these modifications can also occur in messenger RNA. For example, recent studies have identified N6-methyladenosine sites in thousands of human mRNAs and suggest that this modification may play a role in regulation of alternative splicing and gene expression (Dominissini et al, 2012, Nature 485:201-6; Meyer et al., 2012, Cell 149:1635-1646). This mechanism appears to speed up the circadian clock (Fustin et al., 2013, Cell 155:793-806). Interestingly, the Fat Mass and Obesity (FTO) gene has been found to enzymatically demethylate N6-methyladenosine in RNA (Jia et al, 2011, Nat. Chem. Biol. 7:885-887), an activity that appears to affect dopamine function in the midbrain and striatum (Hess et al., 2013, Nat. Neuro. 16:1042-1048). Another RNA modification, 5-methylcytosine, has recently been found to occur on long non-coding RNAs involved in X inactivation (Xist) and homeotic gene regulation (HOTAIR) (Amort et al. 2013 RNA Biol. 10:1003-8). A final example is provided by oxidative processes, which may generate covalent RNA modifications that are associated with disease states or aging (Shan and Lin, 2006, Neurobiol. Aging 27:657–662).

The main obstacles hampering our efforts to better understand RNA modifications and their role in both healthy and diseased biological processes are fundamentally technical in nature. Presently, we lack user-friendly tools and technologies for investigating the epitranscriptome. The problem is compounded by the fact that the most effective and powerful assays we have today to study the transcriptome begin with a reverse transcriptase step that erases all the native marks that exist on the original RNA molecule.

What would the goals of the program be?

- Generate tools and technologies to monitor and manipulate eukaryotic RNA modifications
- Survey the diversity of known RNA modifications
- Discover previously unknown RNA modifications and modifying enzymes
- Generate a Mammalian Epitranscriptome Catalog
- Develop computational strategies to predict the presence of modifications on RNAs
- Explore the biogenesis and mechanistic functions of modified RNAs
- Develop small molecule modulators as probes and potential therapeutics

Why is a trans-NIH strategy needed to achieve these goals?

PLEASE LIMIT TO 2 PAGES

No single institute is moving forward in this emerging and important scientific area. A 2013 portfolio analysis of the five largest neuroscience ICs reveals only two grants investigating RNA modifications. NIDA has released a 2014 SBIR RFA to generate a few tools for this area, but the impact of this initiative will be limited. Existing evidence suggests that RNA modifications, which are typically lost during library construction, are almost certain to be extremely relevant to the understanding of many diseases.

What initiatives might form the strategic plan for this topic?

Phase 1 (years 1-5**)**

- Generate tools and technologies to enable the scientific community to monitor and manipulate eukaryotic RNA modifications. This would include antibodies/affinity reagents for probing novel RNA modifications, techniques for determining RNA modifications at single base resolution, strategies to look at the spectrum of RNA modifications in a single RNA sample, techniques for retaining RNA modifications for subsequent analysis and for monitoring and manipulating RNA modifications *in vivo*.
- Use sensitive approaches (e.g. mass spectrometry) to survey the diversity of known RNA modifications in a defined number of disease-relevant mouse and human tissues and in different RNA classes (e.g. mRNA, long non-coding RNAs, microRNAs).
- Discover novel RNA modifications in mRNA and regulatory RNAs as well as the protein and/or RNA enzymes that write, erase, or read these modifications.

Phase 2 (years 6-10)

- Generate a Mammalian Epitranscriptome Catalog of the mRNAs and regulatory RNAs from a defined number of disease-relevant mouse and human tissues that have a particular RNA modification. For example, affinity approaches could be used to pull down modified RNAs from a particular tissue or cell type which would be characterized to identify the modification sites at single base resolution. RNAs that lack a given modification could be characterized in parallel by RNA-seq to provide the complementary dataset for a given cell/tissue.
- Develop computational strategies to understand the relevant features of a given RNA substrate to understand and predict how RNA modifications are targeted to specific positions.
- Explore the biogenesis and mechanistic functions of modified RNAs, and the readers, writers, and erasers of these modifications in biological and disease processes.
- Develop small molecules as probes and early stage validation of new therapeutic targets.

If a Common Fund program on this topic achieved its objectives, what would be the impact?

- Provide researchers with the tools to explore the biological functions of Epitranscriptomic modifications in their system of choice.
- Produce a hypothesis-generating catalog for modified regulatory and mRNAs for a variety of disease-relevant mouse and human tissues.
- Identify the roles of RNA modifications in important biological and disease processes. Although for the most part the functions of epitranscriptomic modifications are not known, they are likely to play critical roles in the regulation of RNA structure/function, which would have broad and profound, largely untapped implications in health and disease.
- Epitranscriptomic writers, erasers, and readers bind to small molecules. Thus, like epigenomic proteins, they are likely to serve as a superfamily of novel druggable targets for therapeutic development.