The EpiPlex Assay for Epitranscriptomics

Get the Full Message: Uncover RNA Modifications in your RNA-Seq Data

INTRODUCTION

The epitranscriptome is the set of chemical modifications (mods) found naturally on RNA which play key roles in regulating nearly all aspects of RNA biology, including RNA–protein interactions, splicing, structure, transcript lifetime, trafficking, and translation. The distribution and number of mods, and their % abundance at each locus, changes between cell types, during development, and in disease. The ability to profile multiple RNA mods transcriptome-wide will enable basic and clinical research, as well as novel RNA therapeutics. For example, RNA mods have become highly relevant in therapeutic RNA editing. RNA editing introduces or removes modified RNA bases to alter biology and treat disease without the risk of long-term side effects associated with CRISPR-based DNA editing. Additionally, the effectiveness of RNA-targeting therapies can be influenced by the presence of RNA mods at or near the targeted region. Accordingly, pharma researchers

HIGHLIGHTS

Read m6A, I, and m5C in one reaction

EpiPlex adds multiplexed RNA mod detection to RNA-seq. Your data will enable you to correlate changes between multiple mod types at all loci.

Quantification of RNA modifications

Spike-in standards enable relative differences to be measured across samples.

Low sample input and sequencing costs

Analyzing all mods in one assay yields excellent data quality with only 20ng RNA input and 5–10M reads per mod type.

Full bioinformatics support

Our bioinformatics reporting makes it easy for you to get the most out of your data.

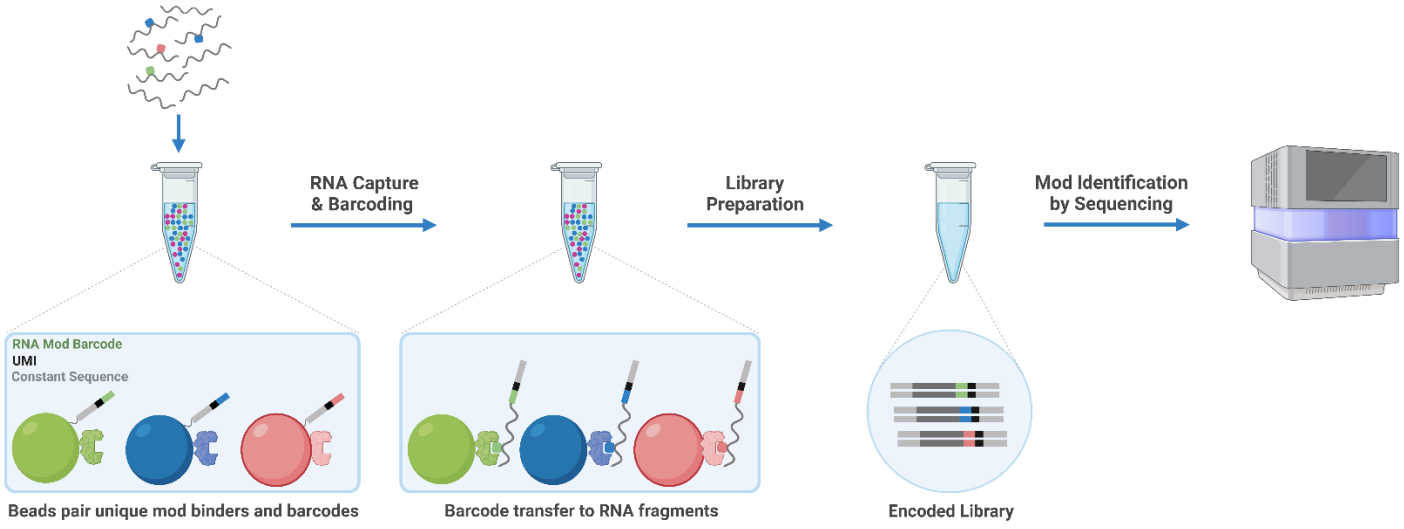

FIGURE 1: Schematic of the multiplexed EpiPlex barcoding workflow. Fragmented RNA is captured by a bead pool where each bead exhibits a distinct RNA modification barcode and protein binder. The barcode is enzymatically transferred to the captured RNA, followed by reverse transcription and library preparation. Sequencing infers the presence of an RNA modification based on the barcode. Depending on the modification, the exact location can be determined informatically.

Contact us
alidabio.com | info@alidabio.com

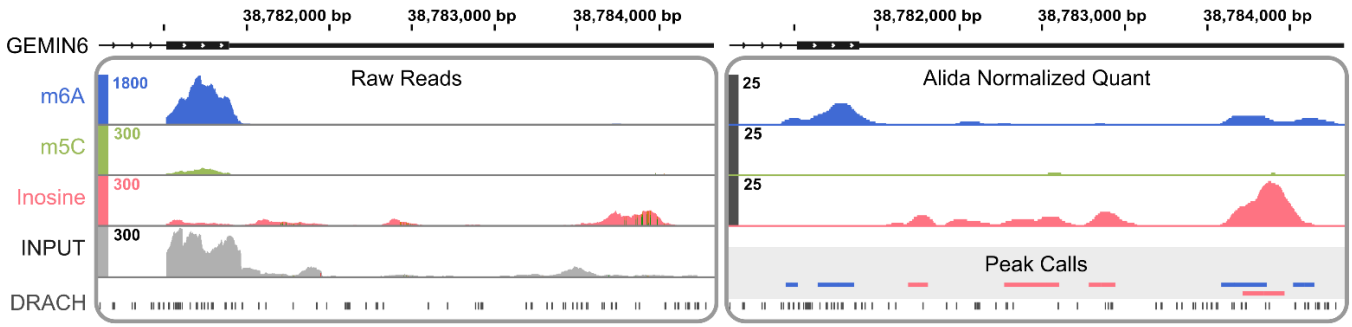

FIGURE 2: Representative NGS read pileups indicating the positions of m6A, m5C and inosine along a gene. Raw reads (left panel) are corrected by the input and normalized relative to spike-in controls. Peak calling is performed after data transformation and RNA modification positions are written to a bed file (right panel). In addition, the gene tracks display the A→G mutations (vertical lines, left inosine track) and the DRACH motif, which is recognized by the m6A writer METTL3.

will benefit from an analytical tool that measures on- and off-target effects transcriptome wide.

NEED FOR COMMERCIAL TOOLS

Existing methods for reading the epitranscriptome typically are limited to a single mod, some require high RNA input and/or employ long workflows and lack quantitative information. In addition, few methods are commercially available, forcing researchers to adopt academic protocols and tailoring them to their needs. The Alida EpiPlex Assay was designed to read multiple RNA mods concurrently with relative quantification and low sample input to make it fit not only with general research needs but also with medically significant applications such as tissue biopsies. The multiplexed assay design affords less variability across experiments and access to new information, such as co-localization of mods, and the opportunity to observe correlative changes in the epitranscriptome. We built a user-friendly customer pipeline to take advantage of the new data type generated by the EpiPlex Assay, making the epitranscriptome as accessible as conventional RNA sequencing.

Together, our novel assay chemistry and analysis tools empower the customer by adding a quantitative RNA mod landscape to their RNA sequencing data, making it an ideal discovery tool for understanding the transcriptome wide effects.

EPIPLEX ASSAY

The EpiPlex assay employs a proprietary barcoding technology to record the presence of mods on fragmented RNA (FIGURE 1). Molecular recognition of the mods using custom binding agents enriches modified fragments onto beads specific to each mod type. Our EpiPlex 1.0 RNA assay interrogates *N*6-methyladenosine (m6A), 5-methylcytidine (m5C), and inosine (I), which are significant mods of mRNA and ncRNA. Alida's intra-bead barcoding chemistry appends a mod-specific barcode during library preparation, recording the identity of each mod into the fragment. Unique molecular identifiers are included to improve the accuracy of deduplication.

DATA OUTPUT AND ANALYSIS

Data analysis with the EpiPlex analysis pipeline combines features of open-source Chip-seq software with our own proprietary algorithms for maximal robustness and accuracy. FIGURE 2 depicts representative EpiPlex data for a given gene viewed in a genome browser. The colored tracks show the enriched, modified fragments ("reads") obtained from binder-mediated barcoding after deduplication and

Contact us
alidabio.com | info@alidabio.com

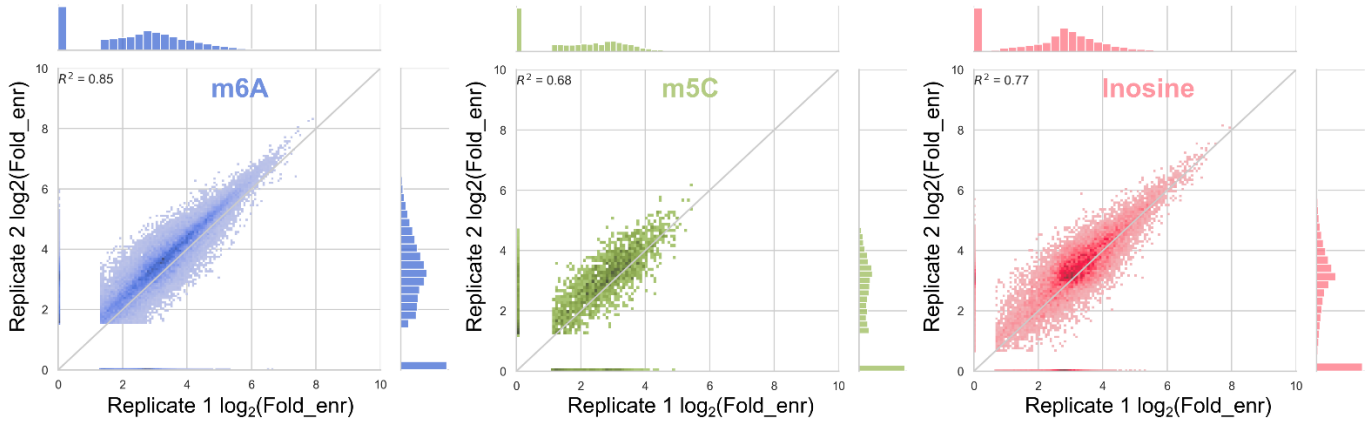

FIGURE 3: Differential analysis of RNA modifications. The correlation plots compare fold-enrichment values across technical replicates. Fold-enrichment measures the magnitude of the peak signal. High fold-enrichment values are indicative of a high copy number of modified transcripts relative to unmodified transcripts, and/or high modification density. Each data point represents a peak call, darker shading indicates overlapping data points. The R^2 values are a measure of reproducibility.

separation of the alignment files by barcode. Each track represents a different RNA mod. The mod sites ("peaks") are detected by identifying regions with significant enrichment of the sample reads over the unenriched input (gray) and over the RNA mod in the other barcode tracks. Our bioinformatics pipeline uses a machine-learning algorithm to exploit sequence motifs and additional data features such as A→G mutations for inosine to maximize the accuracy and precision of peak calls. In addition, the pipeline assigns a "fold-enrichment" value to each peak, which represents the peak height relative to a non-enriched input and normalized to spike-in standards. This relative quantification feature is non-standard for other IP-based methods (e.g. CHIP-seq or MeRIP) and greatly facilitates the observation of even subtle changes in mod levels across different samples and experiments.

DATA QUALITY

The EpiPlex Assay is designed for rigor and reproducibility. Technical replicates (A549 cells) show a tight correlation of fold-enrichment with RNA mod dependent R^2 values (FIGURE 3). In A549 cells, the EpiPlex assay detects 37,399 high confidence m6A sites, 13,561 inosine sites and 1,179 m5C sites. The user has the option to weight the results towards higher sensitivity by allowing a higher false positive rate for dynamic noise thresholding during analysis. Our pipeline routinely generates differential analysis plots of this style, facilitating the identification of global changes across samples.

Motif searching is one component of Alida Bio's peak calling validation (FIGURE 4). EpiPlex identifies m6A, m5C, and inosine site enrichment at A549 mRNA

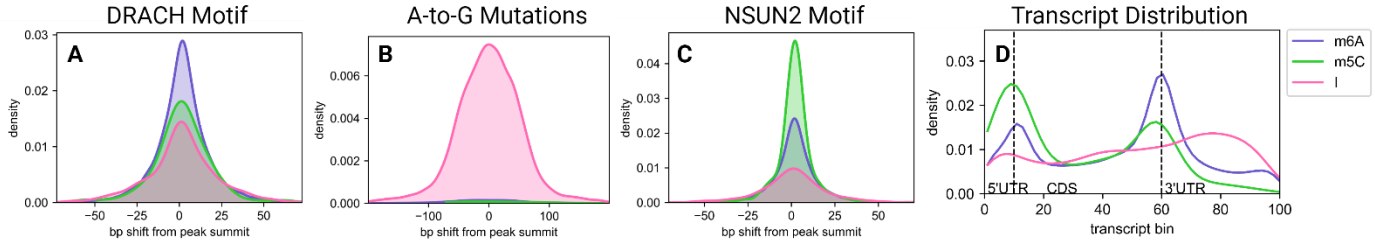

FIGURE 4: Global Distribution of RNA Mod Sites. Shown is the average distance of RNA modification peak calls from the nearest DRACH motif (A), A→G mutation (B), or the NSUN2 motif (C). The density plots are generated by normalizing all peak calls recorded for a given barcode to 1 and considering motifs found within the boundaries of the peaks. RNA modifications are known to locate to characteristic regions of the transcript, as depicted in the transcript distribution plot (D).

Contact us
alidabio.com | info@alidabio.com

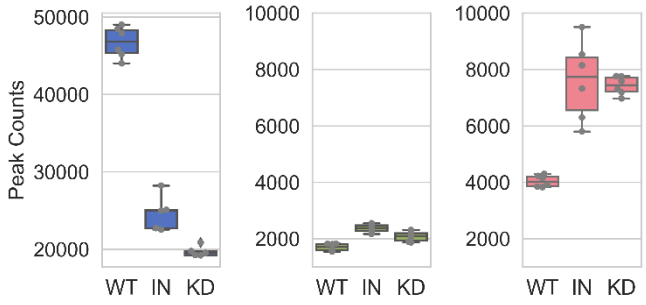

FIGURE 5: Case Study – Effect of METTL3 Inhibition on RNA Modification Frequency. HEK293T cells (WT) were treated with METTL3 inhibitor (IN) or depleted of METTL3 (KD). Depicted are the number of RNA modifications found in two biological and three technical replicates per condition. The dramatic drop in m6A modifications (blue) with attenuated METTL3 activity is associated with an increase in inosine modifications (red) and, to a much smaller degree in m5C (green), demonstrating the value of our multiplexed approach.

transcript positions (4D) consistent with published studies that used methods for reading each of these mods individually. m6A is enriched at the DRACH motif (4A), which is the preferred motif for adenine methylation by METTL3. Similarly, inosine peak calls are corroborated by A→G mismatches (4B) and m5C (4C) is enriched at the NSUN2 motif according to the sequence preference of its methyltransferase.

CASE STUDY: METHYLTRANSFERASE INHIBITION

m6A is the most common mRNA modification. It plays key roles in splicing, nuclear export, transcript lifetime, and translation, depending on sequence context. m6A is installed primarily by the methyltransferase complex METTL3 and METTL14. Excitingly, both proteins have been identified as cancer drug targets.

To validate the EpiPlex assay's ability to report on changes in the epitranscriptome, we measured responses to inhibition of METTL3 activity in HEK293T cells. Mock inhibition was performed using DMSO treatment as a control and knockdown of METTL3 activity was induced using (i) the small molecule inhibitor STM2457 (Storm Therapeutics) and (ii) METTL3 depletion. Three technical and two biological replicates were performed for each experiment.

Our EpiPlex assay reports a 3- to 5-fold reduction in the number of m6A sites called under inhibition (FIGURE 5). The spread between technical replicates is less than the spread between biological replicates in all cases. Chemical inhibition of METTL3 was less potent than METTL3 depletion under these conditions and there was more variability between biological replicates. While the number of m5C peak calls changed little under inhibition, we observed a 2-fold increase in the number of inosine sites. This observation is consistent with the known inhibitory role of m6A on A→I editing.

These correlative results demonstrate the benefit of performing epitranscriptome profiling in multiplex as offered by Alida Bio's EpiPlex assay technology.

ORDERING INFORMATION

Alida Bio is currently offering the EpiPlex Assay as an early access program ahead of the release of the reagent kit. For pricing information, and other inquiries, visit https://alidabio.com or email info@alidabio.com.

The tables below summarize basic service information and describe the data output the customer will receive. Alida Bio is happy to discuss custom projects beyond this standard assay offering.

Table 1: RNA requirements and turnaround time.

RNA types	**mRNA, ncRNA**
Species	**human, mouse**
RNA amount	**500ng total RNA**
Turnaround time	**2 – 4 weeks**

Contact us
alidabio.com | info@alidabio.com

Table 2: Analysis outputs.

SUMMARY REPORT	**HTML, CSV, PNG**
Peaks table with annotated gene features, #reads, fold-enrichment, and q-values	
Motif & mutation enrichment plots	
Transcript distribution plots	
RNA-seq TPM analysis	
Differential clustering of RNA modifications	
Fold-enrichment correlation plots	
PEAK LOCATIONS	**BED**
Transcript regions ("peaks") with RNA modifications for visualization in any genome viewer.	
COVERAGE TRACKS	**BIGWIG**
Read coverage corrected for the non-enriched input coverage and normalized to spike-in controls for visualization in any genome viewer. One file per RNA modification.	
GENOME ALIGNMENT	**BAM**
Aligned and deduplicated reads for visualization in any genome viewer. One file per RNA modification.	
RAW SEQUENCING DATA	**FASTQ**
Unprocessed sequencing reads, demultiplexed per sample for publication and data storage.	

Contact us
alidabio.com | info@alidabio.com