What are the data structures in eVERSE?

Artificial intelligence (AI) is becoming increasingly important to RNA therapeutic research, but developers need data to help train and use these machine learning models. At Eclipsebio, our eVERSE database provides developers with reproducible, AI-ready data from our assays. In this FAQ, we will review these assays and how they can help AI-powered drug development.

miR-eCLIP+

The miR-eCLIP+ assay directly identifies where and how microRNAs (miRNAs) and small interfering RNAs (siRNAs) bind to mRNA. Developers can gather insights into the miRNA activity of their therapeutic, which can help them understand how miRNAs impact stability and translation. With custom eVERSE datasets, they can also locate siRNA off-targets for improving models for target discovery and siRNA design.

With eVERSE we offer a number of ready-to-download, transcriptome-wide assessments of miRNA binding such as with our K562 dataset. We also offer custom dataset generation from specific cell lines or diseases. All datasets include an interactive summary of the data and identification of the specific sites bound by miRNAs. These data can be used to detect patterns in miRNA regulation across the transcriptome or to train models for designing effective siRNA therapies.

miR-eCLIP+ Protocol

1. Crosslink and immunoprecipitate the RNA-induced silencing complex (RISC), target mRNA, and the bound miRNA or siRNA.
2. Ligate the miRNAs and siRNAs and sequence the RNA.
3. Perform bioinformatics analysis to determine where and how miRNAs and siRNAs bind.

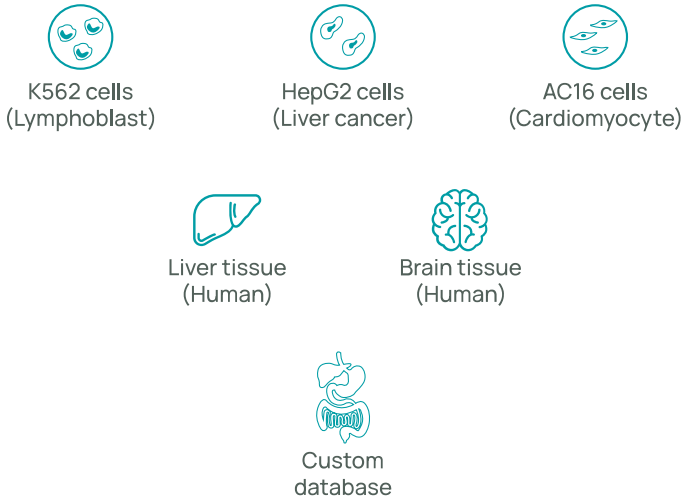
eSHAPE

The eSHAPE assay uses chemical probing to directly determine secondary structure in the RNA's environment, comparing how RNAs change their folding across different cell lines or disease conditions. eSHAPE can also be used to determine protein binding by comparing reactivity scores between two conditions.

In the eVERSE database, we offer datasets of RNA secondary structures from various cell lines using eSHAPE. The datasets consider the impact of cellular factors and different conditions on structure, providing a more accurate prediction than basic secondary structure prediction models. The datasets provide interactive reports on reactivity and inferred protein binding at a nucleotide level. eVERSE can also create these datasets with a custom cell line.

eSHAPE Protocol

1. Probe RNA in vitro, in LNPs, or in cells using the chemical NAI to form adducts on unpaired bases. These adducts appear as mutations after reverse transcription.
2. Prepare mutated and control libraries and sequence the RNA.
3. Perform bioinformatics analysis to determine the RNA secondary structure and nucleotide accessibility.



End-Seq

End-Seq identifies active UTRs and shows how different UTRs act under different disease conditions. This reveals druggable regions for therapeutic developers to target to treat certain diseases. In addition, the assay can discover which of the active UTRs have the cellular machinery required for optimal protein production.

In the eVERSE database, we offer ready-to-download datasets of 5′ and 3′ UTR mapping from various cell lines and human tissues created using End-Seq. We can also create this dataset with a custom cell line. The dataset includes an interactive report with information on the location and quality of UTRs, the expression rates of different UTRs under certain diseases, and a list of druggable regions for treating each disease.

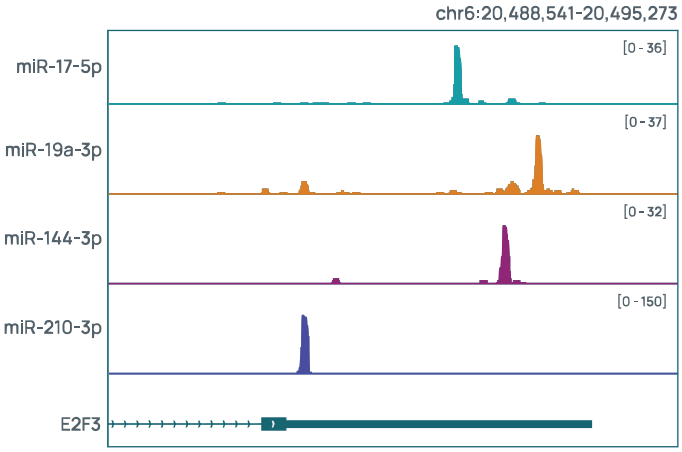
End-Seq Protocol

1. Extract RNA from the 5′ cap or the 3′ tail.
2. Prepare libraries and sequence the extracted RNA.
3. Map where active UTRs are located across the transcriptome using bioinformatics analysis.

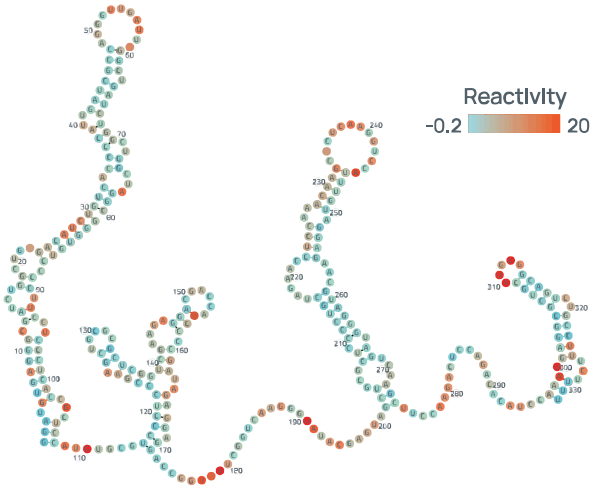
eVERSE datasets at Ecipsebio

At Eclipsebio, we use these assays to continuously improve datasets for the eVERSE platform. Whether you want to train AI models, validate your design, or start creating your new mRNA design, these as well as custom datasets can give you the information you need for RNA innovation.

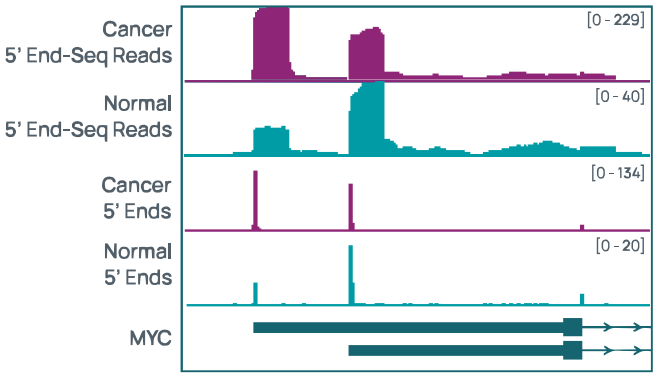
Ready to get your eVERSE dataset? Contact Eclipsebio to get started.



miR-eCLIP+: Direct detection of miRNA binding in the 3′ UTR of E2F3. Each row is a different miRNA. Peaks represent where each specific miRNA binds.



eSHAPE: eSHAPE-supported structure of a gene in the context of cellular factors such as RNA-binding proteins.



End-Seq: A 5′ UTR is more active in cancerous samples than in healthy samples, making it a potential target for selective regulation in cancer.