Comprehensive Annotation of the Human Coding Transcriptome: A Resource for Designing RNA-Targeting Therapeutics

The design of effective RNA-targeting therapeutics, such as oligonucleotides (ASOs, siRNAs) and small molecules, critically depends on identifying accessible and functional regions within target RNA strands

To address this need, we are developing an open-access database that provides comprehensive structural and functional annotations for the entire human coding transcriptome. We have comprehensively modeled the 2D structure of every precursor messenger RNA (pre-mRNA) to facilitate the identification of regions accessible to complementary sequences or small molecules. Moreover, this resource facilitates the annotation of pre-mRNA sequences with a range of complementary data that can be used to infer functional properties. These include RNA Binding Protein (RBP) maps, both experimentally probed (eCLIP) and *in silico* modeled (RBPMap), as well as miRNA binding sites, nucleotide modifications, splicing variants, along with information on sequence and structural conservation and covariation. We have also generated predictions for intronic and exonic splicing enhancers and silencer (ISE, ESE, ISS, ESS) sites to facilitate splicing regulatory RNA structures.

Our database supports targeting RNA through multiple modalities. It aids in identifying unstructured and functional sequences suitable for ASO or siRNA design, while also providing a collection of 2D and 3D folded motifs of target RNA sequences for small-molecule drug development. In summary, this database provides a significant resource for RNA-targeting therapeutic design.