



**Symposium: Beyond Sequence Alignment | 2017-10-16 | Abstracts**

**STRUCTURE-FUNCTION RELATIONSHIPS FOR INDIVIDUAL LONG NON-CODING RNAs**

Karissa Sanbonmatsu, Theoretical Biology & Biophysics Group, Los Alamos National Laboratory

Long non-coding RNAs have emerged as players in development in mammalian systems. Fundamental questions regarding the structure of lncRNAs have yet to be addressed, including, (1) are lncRNAs organized into modular sub-domains or linear chains of stem loops?, and (2) what do structure-function relations look like in the case of lncRNA systems?

We have developed an experimental strategy, called Shotgun Secondary Structure (3S) determination, to derive secondary structures of intact, individual lncRNAs. In this method, we chemically probe the entire lncRNA transcript. Next, we repeat the probing on several overlapping fragments on the transcript. If the probing profile for the fragment is similar to the profile for the same region of RNA in the context of the full transcript, this suggests the fragment folds into a modular sub-domain or modular secondary motif, eliminating a large number of other possible folds. We applied this method to the mouse Braveheart lncRNA and plant COOLAIR lncRNA, both of which possess clear phenotypes. Braveheart and COOLAIR each show highly structured sub-domains and share a unique expansive internal loop (r-turn motif).

In the case of Braveheart, CRISPR-Cas9 analysis demonstrates that the expansive loop motif is necessary for cardiomyocyte lineage commitment. This loop was found to interact with a zinc finger transcription factor (CNBP). In the case of COOLAIR, we used the 3S-generated secondary structure for *A. thaliana* to correctly predict the existence of COOLAIR in several other species. We then performed 3S on COOLAIR in *B. rapa* and demonstrated that the structure of COOLAIR is conserved across ~20 million years of evolution. Comparison of FLC expression levels in different species, for different SNPs, and for different splicing isoforms suggests that the length of the terminal helix of the central domain (H4) may contribute to the regulation of FLC levels.

**STAY  
CONNECTED:**

[bivt.edu](http://bivt.edu)

[facebook.com/biocomplexity](https://facebook.com/biocomplexity)

[twitter.com/biocomplexityvt](https://twitter.com/biocomplexityvt)

[youtube.com/c/biocomplexity](https://youtube.com/c/biocomplexity)

[google.com/+biocomplexity](https://google.com/+biocomplexity)