**SLEQ**

Statistical modeling of RNA structure profiling (SP) experiments enables parsimonious reconstruction of structure landscape.

**What Is It?**

Structure Landscape Explorer and Quantifier (SLEQ) is an algorithm for sparse

SP-guided reconstruction of RNA structure landscapes. It features a statistical model that selects a parsimonious set of structures that best explain SP data from a pre-determined candidate set (e.g.: Boltzmann ensemble) and estimates their relative abundances.

**Getting Started**

These instructions will help you run SLEQ on your local machine.

**Prerequisites:**

C++ studio (e.g.: Xcode) and R studio are required for SLEQ.

**Installation:**

All codes can run directly using C++ studio and R studio without any specific installation.

**Usage:**

In version v1.0, SLEQ supports the reproduction of the results in the paper “Statistical modeling of RNA experiments enables parsimonious reconstruction of structure landscape”. It can also be used for processing other datasets with proper parameter setting. Here I briefly introduce how to use SLEQ with a fluoride riboswitch dataset.

**1)** Open the folder “Fluoride\_riboswitch” and then open the file “SHAPE-SEQ.xcodeproj” with Xcode.

**2)** Set parameters as follows:

* Change the path of parameter “Folder” with your own path where ensemble files are deposited. The ensemble files exist in the folder “Fluoride\_riboswitch” and its format is FLUORIDE\_EQ/CO\_XXnt\_F/NF.txt “XX” here is the transcript length. “EQ” stands for equilibrium state and “CO” stands for cotranscription. “F/NF” stands for with/without fluoride.
* Change the parameter “RNAname” to the one you want to process. The format is FLUORIDE\_CO/EQ/\_XXnt\_NF/F.
* Change the parameter “Ed” = transcript length – 1.
* Change the parameter “RNALen” = transcript length.
* Choose the corresponding initialization for the parameter “AddFold”. 0 if the transcript length < 55. 3 if the 55 <= transcript length < 61. 6 if transcript length >=62. This parameter is the number of pseudo-knot structures spiked-in.
* Parameter “basicFoldNums” is the number of candidate structures generated by the stochastic sampling.
* Choose the corresponding initialization for parameters “SEQPlusReads” and “SEQMinusReads”.

**3)** Build and run. It will output three files:

“RNAname + SEQResponse.txt” is the pattern frequencies Y of linear model.

“RNAname + SEQDSMatrix.txt” is the design matrix X of the linear model.

“RNAname + SEQpara.txt” is the index of structures after prefiltering.

These output files are saved where ensemble files are deposited.

**4)** Close Xcode.

**5)** Open the file “NNLS\_FLUORIDE\_RIBOSWITCH.R” with R studio.

**6)** Set parameters as follows:

* Change the path of parameter “Folder” with your own path where above output files are deposited.

**7)** Run and plot parameters “idxt” and “rhot”. “idxt” is the index of selected structures. “rhot” is the corresponding relative abundances.

**Citation**

TBA

**Reporting bugs and requesting features**

To report a bug, open a ticket in the [issues tracker](https://github.com/AviranLab/patteRNA/issues). Features can be requested by opening a ticket in the [pull request](https://github.com/AviranLab/patteRNA/pulls).

**Contributors**

Hua Li – Initial implementation

Sharon Aviran – Supervisor