***SynNetModified Quick User Manual***

**Step 1:** Install MATLAB from <https://www.mathworks.com/products/matlab.html>

**Step 2:** Download SynNetModified.zip from <https://github.com/jwon0408/SynNetModified> and unzip it in *Documents > MATLAB*

**Step 3:** Navigate to *SynNetModified > SynNetModified-master > SynNet\_1.1 > 02\_SeekNet > Auxiliary\_Functions > fastAUC* and open ‘install.m’. MATLAB will launch and press “run” on the Editor ribbon. Make sure you get the following messages.

*MEX completed successfully.*

*Test succeeded*

**Step 4:** Start Microsoft Excel and navigate to *02\_SeekNet > Example\_Data* and open ‘LipchinaRawData.txt’ with tab as the text delimiter

This is an example of a single input dataset. Every RNA sequencing dataset with a different sequencing depth (i.e. sample cell number) must be provided in a separate text file. For every dataset, the following 4 kinds of data must be provided. See Fig. 1 for the specific format.

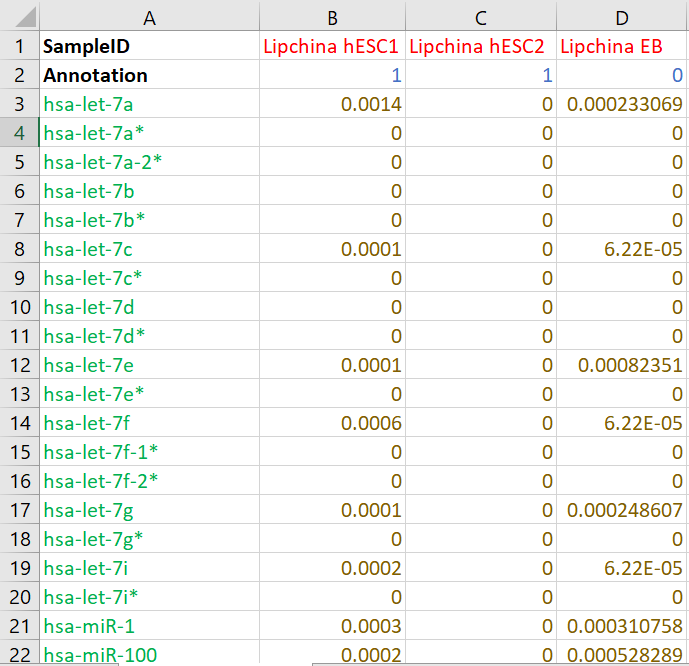
1. SampleID (row 1, starting from column B): unique identifiers (not just numbers) for individual sample cell populations
2. Annotation (row 2, starting from column B): 1 for target populations and 0 for other non-target populations
3. miRNA name (column A, starting from row 3): names of miRNAs identified
4. Expression level or count (columns from B, rows starting from 3): miRNA counts determined by the sequencing

Figure : Single Input Dataset Format. Shown in the figure is ‘LipchinaEarlyData.txt’

**Step 5:** Still using Excel, within the folder *Example\_Data*, open ‘miRNA\_Data\_files.txt’ (tab-delimited). In column A, provide the name of input data files (without the ‘.txt’ extension). In column B, for each input data file, provide the corresponding version number of miRBase. In other words, the version of miRBase ([www.mirbase.org](http://www.mirbase.org)) used by the source literature to align and identify the given miRNA sequences. IMPORTANT NOTE: if miRBase was not used or not specified by the authors, DO NOT leave column B blank for that datafile but type in 21 (i.e. assume the latest version).

**Step 6:** Navigate back up to *SynNet\_1.1 > 02\_SeekNet* and open ‘Seek\_Net.m’.

**Step 7:** In the Command Window on the bottom, type the following command to launch the program. The program will take minutes to run.

Seek\_Net(‘Example\_Data/Constraints.txt’);

**Step 8:** Navigate to *02\_SeekNet > F03\_Results* > *Constraints\_Results* and using Excel, open ‘C.txt’ (tab-delimited).

* Look at from cell A6 down to identify the Best Circuit Found. Note that ‘~’ stands for NOT.
* AUC (cell A5) and cMargin (cell B5) are key performance predictors. AUC value of 100% represents 100% classification accuracy. cMargin indicates the gap between the output score of target and non-target populations. Ideally, it should be greater than 1 (target ouput/non-target output).

**Step 9:** Within the folder *Constraints\_Results*, there are three other files that may be useful:

* ‘C-Inputs’ (.fig and .pdf): colormap of expression of the selected miRNA inputs and the circuit output and truth values
* ‘C-Inputs-Clegend’ (.fig and .pdf): colormap legend to supplement ‘C-Inputs’
* ‘C-OutputLvl\_C’ (.fig and .pdf): scatter plot of circuit output of each cell sample. The difference in the mean output values between target (positive) and non-target (negative) is indicated with a bar.

**Notes Regarding Files in *Constraint\_Results*:**

* If the Best Found Circuit is too complex to be implemented, the software also suggests a simpler or “pruned” circuit without significant performance reduction. All the output files are also generated for the pruned circuit with the filename extension ‘-Prune’.
* The remaining files are not so informative to non-experts of the software. Please, refer Esther for any inquiry regarding those files.