Tutorial for using *TFOFinder* Python program to design parallel triple helix forming oligonucleotides

A pdf copy of this tutorial was uploaded on GitHub.

1. Download TFOFinder either directly from GitHub as a zip file or using the git bash command: "$ git clone https://github.com/icatrina/TFOFinder.git"
2. Select the RNA target of interest. Example: *ovo-RE* (*D. melanogaster ovo* gene encodes a Zn finger transcription factor) and download the RNA sequence as a "fasta" file. For this tutorial the ovoRE.fasta file was used.
3. Fold the RNA target using one of the freely available software for the prediction of RNA secondary structure. We used a text-interface local copy of *RNAstructure*. Alternatively, a web server can be used, and depending on the server selected, the length of the target may be limited (*e.g.*, < 2,400 nucleotides for *mfold*, <10,000 nucleotides for *Sfold batch*; please check each web server for the exact length limit). Links for some of the web servers offering this service are given below:
    - *mfold* or *Quickfold*
    - *RNAstructure*
    - *RNAfold* (ViennaFold)
    - *Sfold*
4. After the folding of the RNA target is completed, download the output "ct" file (all structures or an individual structure of interest) that will serve as the input file for *TFOFinder*. For this tutorial, the "67_ovo_RE.ct" file will be used. The output files will be automatically saved in the folder containing the input file, and the name of the input file, without its extension, will be appended to the name of each output file. For comparison with the original files, a separate output folder was uploaded on GitHub, which contains the two output files obtained from running this tutorial. The two output files are:
    - "67_ovo_RE_sscount.csv" – the input file is converted to the ss-count file format as defined for the *mfold* application, except it does not contain the number of total structures on the very first line. The number of total structures (minimum free energy – MFE – and suboptimal – SO – structures) is printed during the run, see the text highlighted with yellow in Fig. 1.
      This file can also be used as input for *PinMol* to design molecular beacons, after the number of structures is inserted as the first line (*i.e.*, in this example this number would be 20).

```
 TFOFinder  Copyright (C) 2022  Irina E. Catrina
This program comes with ABSOLUTELY NO WARRANTY;
This is free software, and you are welcome to redistribute it
under certain conditions; for details please read the GNU_GPL.txt file

->->->->->->->->->->->->->->->->->->->->->->->->->->->->->->->->->->->->

WARNING: Previous files will be overwritten or appended!  Save them in a
different location than the current input file, or rename them.

->->->->->->->->->->->->->->->->->->->->->->->->->->->->->->->->->->->->
Enter the ct file path and name: /Users/ficatrina/Downloads/TFOFinder/Example/input/67_ovo_RE.ct
Enter the length of TFO probe; a number between 4 and 30: 10
Number of Structures = 20 ...Please wait...
```

Figure 1. Snapshot of text-interface with the number of structures contained in the input file highlighted in yellow.

- "67_ovo_RE_TFO_probes.txt" – the output file containing the information for the parallel TFO probes for the target of interest. The "Start Position" is the number of the first 5' nucleotide that it is part of the purine duplex identified in the RNA target. This sequence is made up of only purines or 100 % purines (G and A nucleotides - %GA). The sscount fraction (sscount) gives information about the likelihood that the 10-nucleotide region is double stranded in all structures (MFE and SO). A value of zero for the sscount fraction means all 10 nucleotides are predicted to be double stranded in all 20 structures included in the input file for *ovo-RE* mRNA.

```
Results for /Users/ficatrina/Downloads/TFOFinder/Example/input/67_ovo_RE.ct using 10 as parallel TFO probe length
Start Position,%GA,sscount,Paralell TFO Probe Sequence,Tm
331,100,0.18,CCCCCUUUUU,50
393,100,0.95,UUUUUUUCUU,21
5523,100,0.0,UUUUCUUUUU,21
5524,100,0.0,UUUCUUUUUU,21
5525,100,0.0,UUCUUUUUUU,21
```

Figure 2. Snapshot of the output file. The target, ovo-RE contains five possible regions amenable to forming a parallel triple helix.

5. Additional information, such as original files described in our manuscript, will be made available upon request; please contact Irina Catrina at iecatrina at gmail.com.