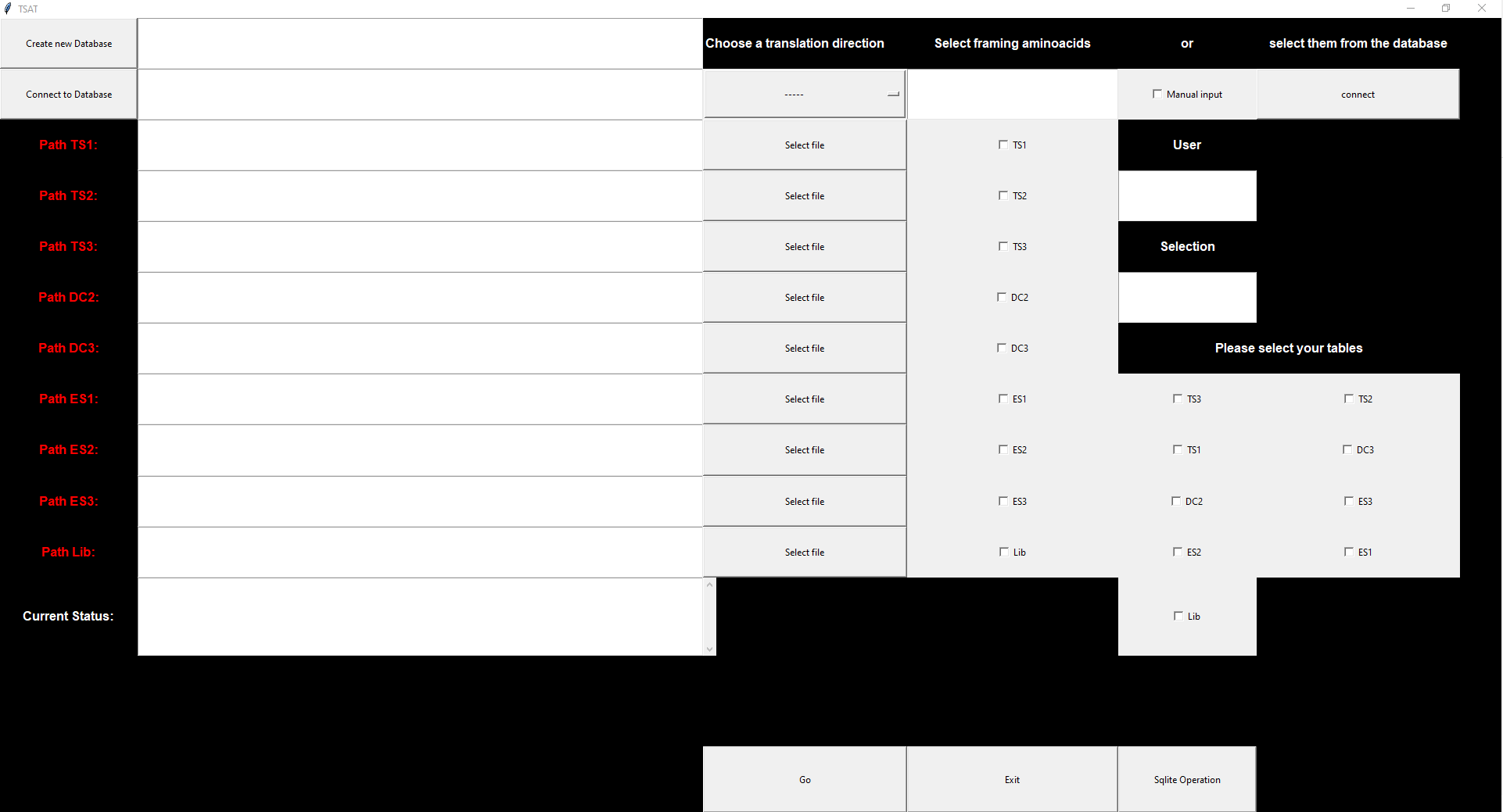
**How to use TSAT**

This guide is a step-by-step instruction for users of TSAT on how to use the software. The set of test data is provided on the following Link:

1. **Startup**

After successful startup the user is able to see the following Window.

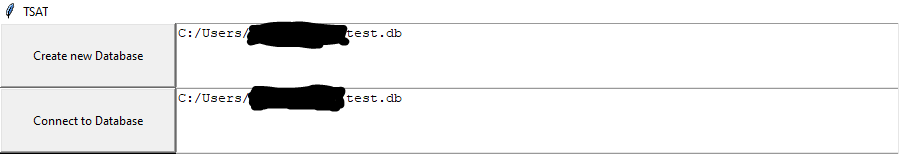


1. **Creation/connection to a database.**

TSAT needs a database where it can save the results of a run. A new database can be created by using the “Create new Database” Button on the top left side of TSAT. Alternatively, the user can connect to an already existing Database with the “Connect to Database” button.

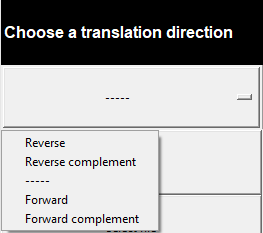


After creation of a new Database, it is important to connect to the new Database.



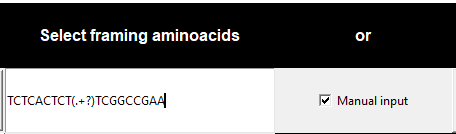
1. **Choose a translation direction**

Depending on your data, you can adjust how TSAT should translate the obtained DNA. This is done by choosing one of for options under the “Choose a translation direction” button.

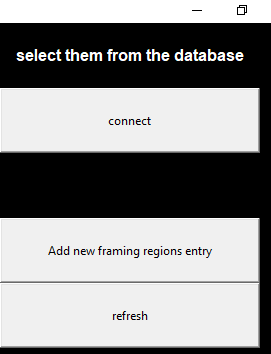


1. **Input the framing amino acids**

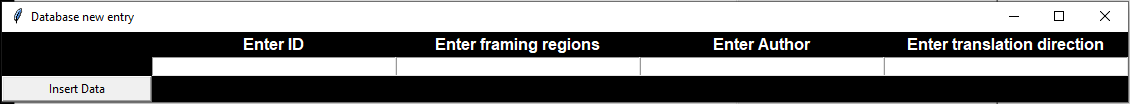
This step is of utmost importance as it provides TSAT with the information which random region is the region of interest. For this you have two possibilities. The first one is to manually input the required information into TSAT. This option can be useful if you have a dataset that you only need to process once. For this you have to check the box “Manual Input” and insert the framing regions into the text box under “Select framing amino acids”. Note that you have to give a defined amino acid composition before the randomised region and after it. The randomised region is displayed by (.+?)



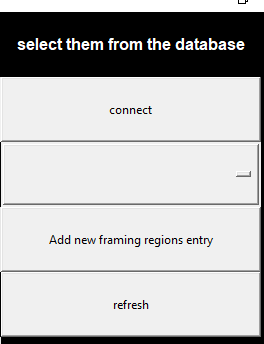
Alternatively, you can save your framing regions in a database and connect TSAT to this database. This is especially useful if you have many selections with the same randomised region, as it allows you to input the framing regions in a reduced amount of time with no spelling errors. In order to use this feature, you need to create a fresh database. This can be done clicking the “Create new Database” button in the top left corner of TSAT. After that you can connect to this database via the “Connect” button in the top right corner of TSAT. If you created a new database without any framing regions and connected to it two new buttons will appear below the “Connect” button.

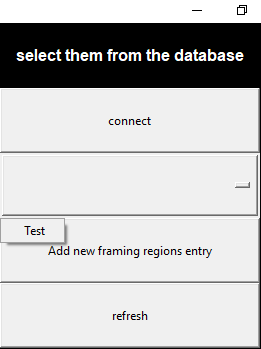


The next step is to add the wanted framing regions into the newly created database. This is done by clicking the button “Add new framing regions entry”. By clicking this button, a new window will open where you can insert the needed Information.



Here you need to give this framing region entry a name. This name will later be displayed in TSAT so that you can differentiate between different framing regions. Under framing regions, you have to enter the information that TSAT needs to find the desired randomized region. This is the same information you would have to enter in the manual mode. For this example the framing regions that will be added are TCTCACTCT(.+?)TCGGCCGAA with (.+?) standing for the unknown randomized region. Afterwards you can add an Author name which will only be saved in the database and not be displayed. Finally, you can enter a translation direction. At this moment the entry is also only saved in the database. It does not replace the drop-down selection within the main TSAT software. With the “Insert Data” button the information is transferred to the database and saved within. Be sure that you have inserted everything correctly as you can not change the data with TSAT. After this you can use the “refresh” button to update you connection with the Database. A new dropdown menu should appear in which you can select the entry you just created.





1. **Input user and selection**



Here you can insert a user and a name for the selection. Both will be saved in the database. This is optional and TSAT will leave these areas blank if no entry was given.

1. **Select files to be processed**

As a next step the files that TSAT should process have to be selected. TSAT takes fastq data with a coding of UTF-8. This can be achieved by pressing the “select file” button beside the textbox of “Path TS1” to Lib. The abbreviations stand TS = target selectin 1-3, DC = direct control 2-3, ES = empty selection 1-3 and Lib stands for the library. If the sorting and filtering functions of TSAT are not needed this differentiation can be ignored. However, in this case a separate software is needed to access the saved data in the database. An example for this could be DB browser for SQLite.



As a final preparatory step, the user needs to check the Checkbox to the right of the “Select file” button if a file was selected.



1. **Start TSAT**

TSAT is started via the “Go” button on the bottom of the software. The user may check the progress of processing in the “Current Status” box which will update in real time. An example of the displayed information would be the number of sequences that were found in a certain file. The textbox will also display the information when the processing has finished.

1. **Filtering and sorting**

If filtering and sorting of the processed data is intended, it is important that the target selection, control selection, empty selection and library are provided in the indicated paths. After TSAT finished the processing of the data the user may check the checkboxes of the provided files in the lower right of TSAT. Afterwards the user can start filtering by pressing the “SQLite Operation” button. At least one Target selection, one empty selection and a library file need to be provided. After filtering the user will be prompted to save a fasta file containing the results. This file can be opened with a text editor for manual inspection or used in further data analysis using different software.