

# Entropic Profiler User Manual

The “Entropic Profiler” tool is available through an easy to use web based interface. Next we introduce the main application’s input and output screens along with a description of all its features.

Figure 1 presents the initial main page of the application’s web interface seen from a common web browser. Figure 2 briefly describes each of the commands and options available to the user. The user can interact with the application through the web form in that page.

The screenshot shows the Entropic Profiler web interface in a Windows Internet Explorer browser window. The address bar displays the URL: <http://isabelle.math.ist.utl.pt/~448659/ep/index.php>. The page title is "Entropic Profiler".

**Information**

**Entropic profiles** of DNA sequences are local information plots of the relative over and under-expression of motifs per position. They are calculated based on Chaos Game Representation (CGR) using a recently proposed fractal kernel and Parzen's window density estimation method. They allow the visualization of motif densities for two different parameters: resolution  $L$  and smoothing parameter  $\Phi$ . This method detects biological significant regions of DNA, here exemplified for the genomes of *E.coli* and *H.influenzae*. An important simplification allows it's calculation using segment counts, explored in this application through suffix trees. More information is available in the help section.

**Reference:**  
Vinga, S. and Almeida, J.S. [Local Renyi entropic profiles of DNA sequences](#), BMC Bioinformatics 2007, 8:393 (Oct 16).

**Form Fields:**

- E-mail:**  **Job name:**
- Sequence:**
  - ☒ FASTA Text:
  - ☐ FASTA File:
  - ☐ Load Example:
  - ☐ Load Last Work:
- Study by:**
  - ☒ Position:
  - ☐ Motif:
- Resolution Length:**
- Smoothing Parameter:**
- [Options](#)
- 

Figure 1 – Entropic Profiler’s interface viewed from a web browser.

The diagram illustrates the main window of the Entropic Profiler, showing the flow of data and user interaction. The interface is divided into several sections:

- Input:** A section at the top left for entering data.
- User contact and task description:** A section at the top right for providing contact information and job details.
- Sequence:** A section for selecting the type of sequence input (FASTA Text, FASTA File, Load Example, Load Last Work).
- Study by:** A section for selecting the study type (Position, Motif).
- Parameters:** A section for setting parameters like Resolution Length and Smoothing Parameter.
- Buttons:** A section at the bottom containing buttons for 'Reset', 'Get Entropic Profiles', 'Select study type', 'Run application', 'Reset form', and 'Set parameters'.

Red arrows indicate the flow of data and user interaction. The flow starts from the 'Input' section, goes to the 'Sequence' section, then to the 'Study by' section, and finally to the 'Run application' button. The 'User contact and task description' section also feeds into the 'Run application' button. The 'Reset' button leads to the 'Reset form' button, and the 'Get Entropic Profiles' button leads to the 'Set parameters' button.

Figure 2 – Entropic Profiler’s main window.

### Loading sequences

To be able to get results in the Entropic Profiler, a DNA sequence is required. There are several ways to input a DNA sequence through the web interface:

- a sequence can be typed symbol by symbol or copied and pasted from another text source directly to the text area
- a file in the FASTA format containing the sequence can be selected and uploaded to the server
- some predefined sample sequences are already stored on the server and ready to be used
- if the user has previously worked on a sequence and it is still stored on the server, an option to load the last saved sequence and parameters will also be available to easily continue the previous work

These distinct ways of loading a sequence from the web interface are shown in Figure 3.

**(a) Type or copy/paste the text of the sequence**

Sequence

☒ FASTA Text:
 

TAAAAACACCTGACAAAAAAGAAATGGATCTATTGATCCTTTGAAAGATCAGGTTTG  
 TATTAGTAAATAAGATCTCTTTTATATATAAAGATCTTATTATTGTTATTATTAA  
 GATCTTTTTTGGTTGTGAGTAACCTTATTCGATCCTTATGGAACAGGTGTTATAG  
 GATCCAAAAATCTTGTGAAAAATGCGATCATTTTCTTTAAAAATCTTGTGTGAAAA  
 TGGAAAGTTATTAACAACCTTTATTTTCTCAACGTTACTAAGTGAGTAAAAACAG  
 TTTTATGACAGATTATTTATGAGTTATCCACAGATAAAAAATGAATTTTATTGAAA  
 TCCCTATAATTTAGTTTTTTAAACAATTAATTTTATTAAATGTACGTTTTATGATT  
 TGACTTTGTAGTTACTACAGGTTTTATGCTTCCTTCAAATCAAATAGTTAAGGAAA  
 TAAAAATCAAAGATTTTTTACAGTCTATTGTTACCGAAAGTGAACATAATGAAGAGG

☐ FASTA File:

☐ Load Example:

☐ Load Last Work:

**(b) Upload a sequence stored in a FASTA file**

Sequence

☐ FASTA Text:

☒ FASTA File:
 

C:\DNA\sequence.fasta
Browse...

☐ Load Example:

☐ Load Last Work:

**(c) Load a pre-defined example sequence**

Sequence

☐ FASTA Text:

☐ FASTA File:

☒ Load Example:
 

Escherichia coli K12
▼

☐ Load Last Work:

**(d) Load the last project the user worked on**

Sequence

☐ FASTA Text:

☐ FASTA File:

☐ Load Example:

☒ Load Last Work:
 

<b>Description</b>	Experimental promoter regions of Bacillus subtilis
<b>Submitted</b>	13-Dec-2007 10:12:43
<b>File</b>	<example3.seq>
<b>Size</b>	2.053 bytes
<b>Length</b>	2.000 basepairs

**Figure 3 – Different ways to load a sequence.**

### Study by position

After selecting a sequence, the user has the ability to choose to study the sequence by a particular position or by a motif of interest. After specifying the position to study and the parameters for the application and pressing the “Get Entropic Profiles” button, the results are presented in a new webpage. Figures 4 and 5 show the input and output screens for the position study as well as a description of all the parameters and output plots and values.

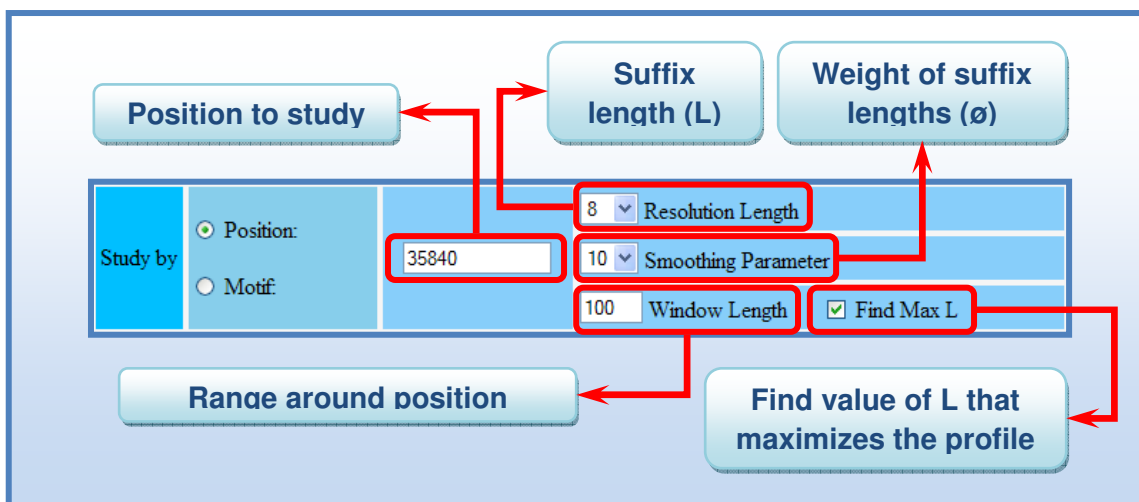


Figure 4 – Studying a sequence by position.

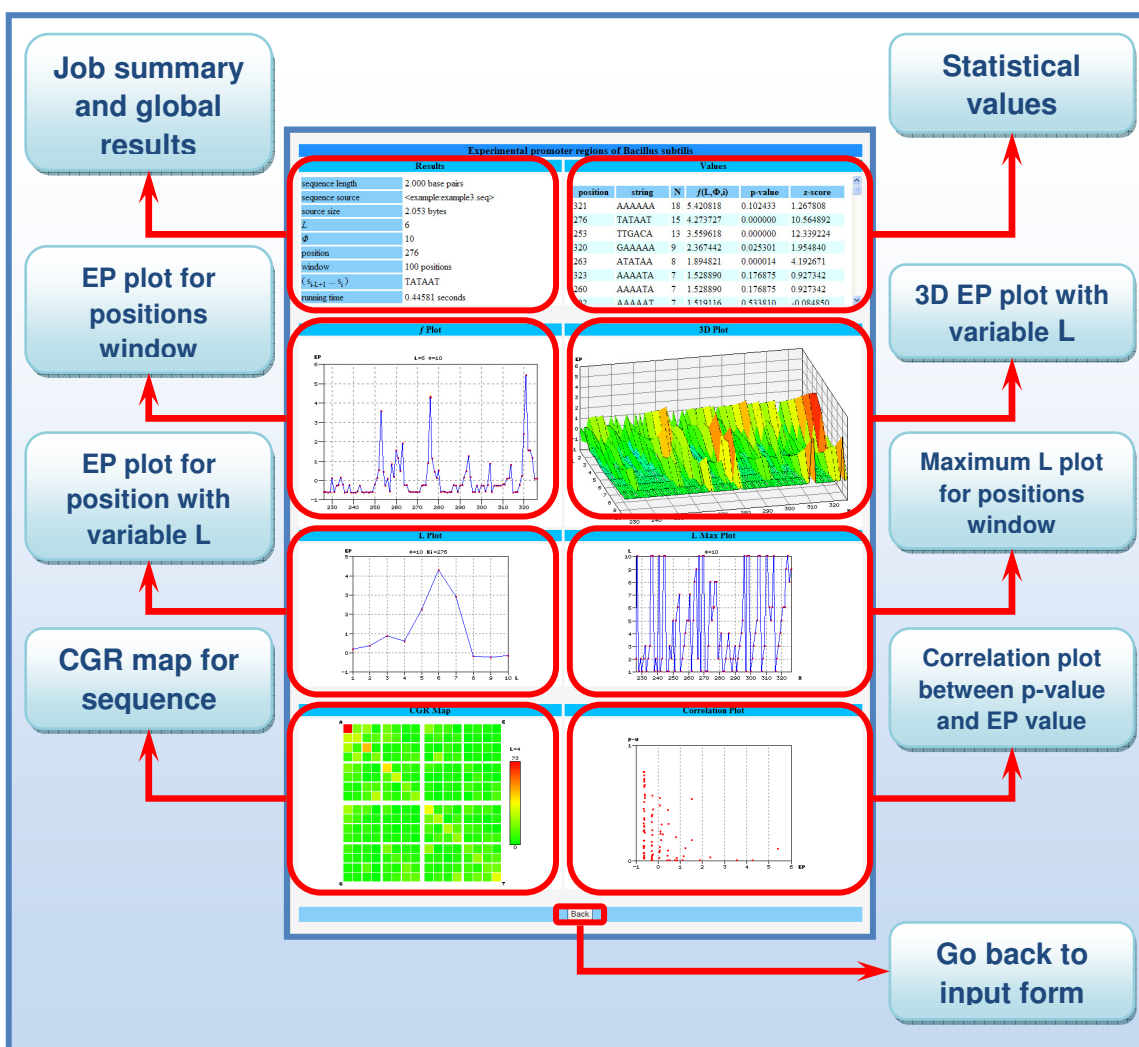


Figure 5 – The results of the study by position.

## Study by motif

If the user is interested on a specific motif inside the sequence, a search by motif can also be performed. The user has the ability to input the motif string by introducing the symbols one by one using the available buttons. A specific section of the sequence can also be searched for instead of the whole sequence.

Figures 6 and 7 describe the parameters and the results of the motif study.

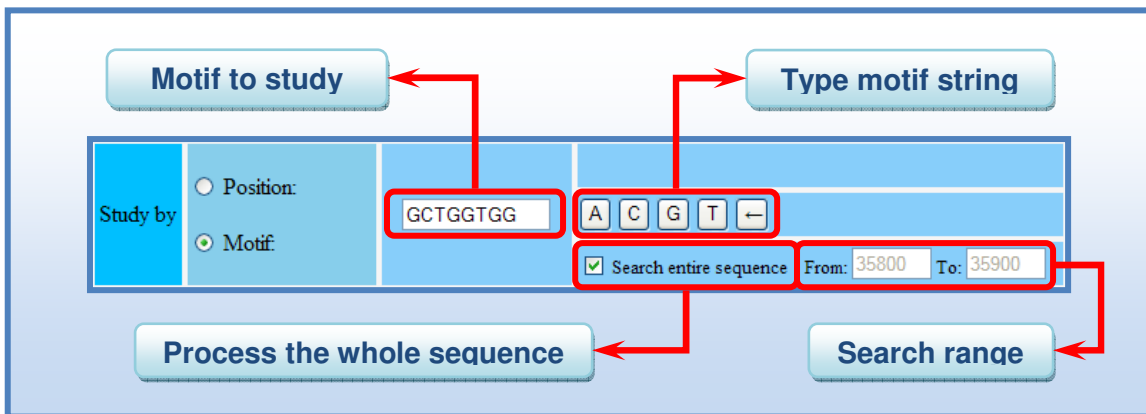


Figure 6 – Performing a study by motif.



Figure 7 – Results of the study by motif.