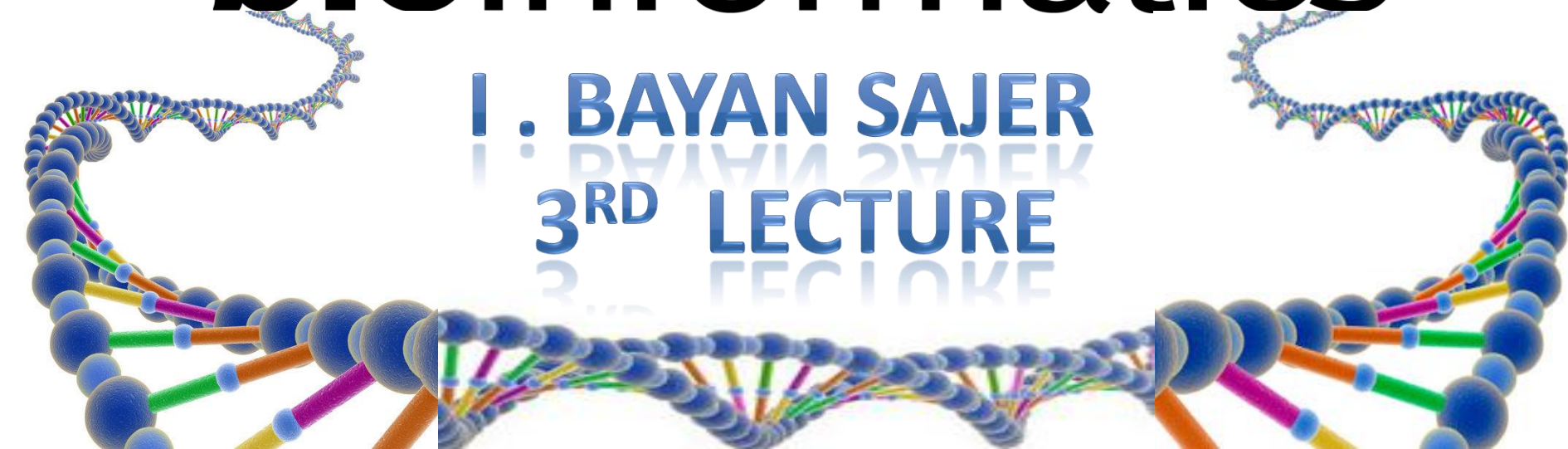


bioinformatics

I . BAYAN SAJER
3RD LECTURE



As you learned previously in the lectures :

There are three *parameters to consider for a dotplot*:

1) **The scoring scheme.**

2) **The cut-off score**

3) **The word size**

10110100101010110
0110100101101010
10110101101001010
0110100110001101010110100
1101001010010101111010110

bioinformatics

The Cut-off score.

**The higher the cut-off score the less dots will be plotted.
But, each dot is more likely to be significant.**

**The lower the cut-off score the more dots will be plotted.
But, dots are more likely to indicate a chance match
(noise).**

```
01110100101010110
0110100101101010
10110101101001010
011010011001101010110100
1101001010010101111010110
```

The Word size.

Is the size or the length of sequences being tested

For sequences with regions of small matching features.

**Small words pick small features
Individually**

**Larger words show matching
regions more clearly.**

***The lack of detail can be
an advantage***

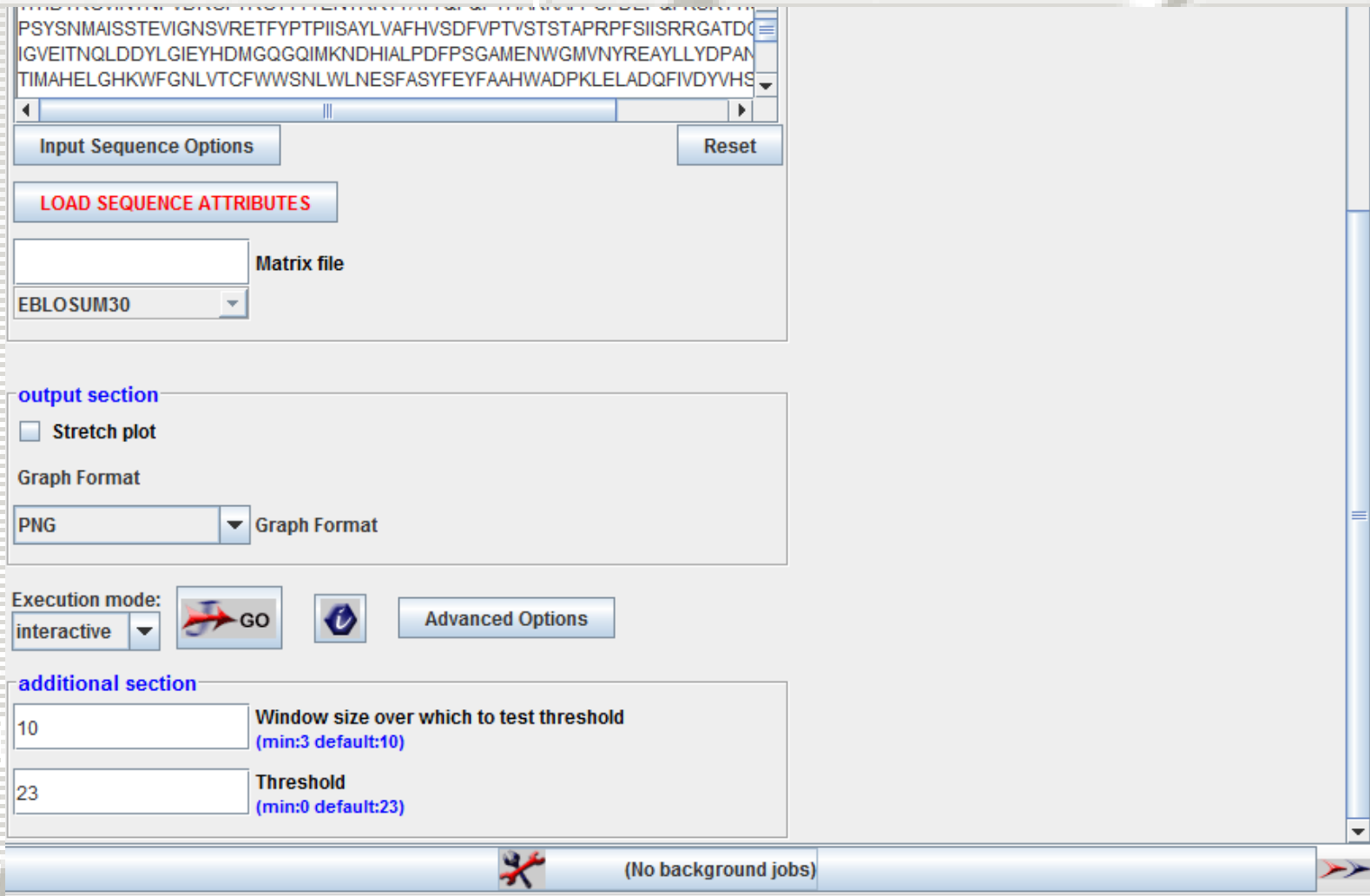
Changing the cut off score and word size will give different graphs in the mEMBOS the word size is called (window size) While the cut off score is called threshold .

The screenshot displays the mEMBOS web interface. At the top, there is a text area containing a protein sequence: PSYSNMAISSTEVIGNSVRETFTPIISAYLVAFHVSDFVPTVSTSTAPRPFSIISRRGATDCIGVEITNQLDDYLIEYHDMGQGQIMKNDHIALPDFSGAMENWGMVNYREAYLLYDPANTIMAHELGHKWFGLVTCFWWSNLWLNESFASYFEYFAAHWADPKLELADQFVDYVHS. Below the sequence is a progress bar. The interface is divided into several sections:

- Input Sequence Options:** Includes a "Reset" button and a "LOAD SEQUENCE ATTRIBUTES" button.
- Matrix file:** A dropdown menu currently set to "EBLOSUM30".
- output section:** Contains a "Stretch plot" checkbox (unchecked) and a "Graph Format" dropdown menu set to "PNG".
- Execution mode:** A dropdown menu set to "interactive", accompanied by a "GO" button with a red arrow icon and an "Advanced Options" button.
- additional section:** Contains two input fields: "Window size over which to test threshold" (set to 10, with a note "(min:3 default:10)") and "Threshold" (set to 23, with a note "(min:0 default:23)").

The bottom status bar shows a red wrench icon and the text "(No background jobs)".

If you want to change the window size and the threshold you should click on advanced option button then screw down



The screenshot shows a software interface for sequence analysis. At the top, a sequence viewer displays a DNA sequence: PSYSNMAISSTEIVGNSVRETFYPTPIISAYLVAFHVSDFVPTVSTSTAPRPFSIISRRGATDCIGVEITNQLDDYLIEYHDMGQGQIMKNDHIALPDFSGAMENWGMVNYREAYLLYDPANTIMAHELGHKWFGLVTCFWWSNLWLNESFASYFEYFAAHWADPKLELADQFVDYVHS. Below the sequence is a progress bar and a 'Reset' button. The 'Input Sequence Options' panel includes a 'LOAD SEQUENCE ATTRIBUTES' button, a 'Matrix file' dropdown set to 'EBLOSUM30', and an 'output section' with a 'Stretch plot' checkbox and a 'Graph Format' dropdown set to 'PNG'. The 'Execution mode' is set to 'interactive', with 'GO' and 'Advanced Options' buttons. The 'additional section' contains two input fields: 'Window size over which to test threshold' (set to 10, min:3 default:10) and 'Threshold' (set to 23, min:0 default:23). The status bar at the bottom shows '(No background jobs)'.

PSYSNMAISSTEIVGNSVRETFYPTPIISAYLVAFHVSDFVPTVSTSTAPRPFSIISRRGATDCIGVEITNQLDDYLIEYHDMGQGQIMKNDHIALPDFSGAMENWGMVNYREAYLLYDPANTIMAHELGHKWFGLVTCFWWSNLWLNESFASYFEYFAAHWADPKLELADQFVDYVHS

Input Sequence Options Reset

LOAD SEQUENCE ATTRIBUTES

Matrix file
EBLOSUM30

output section

☐ Stretch plot

Graph Format
PNG Graph Format

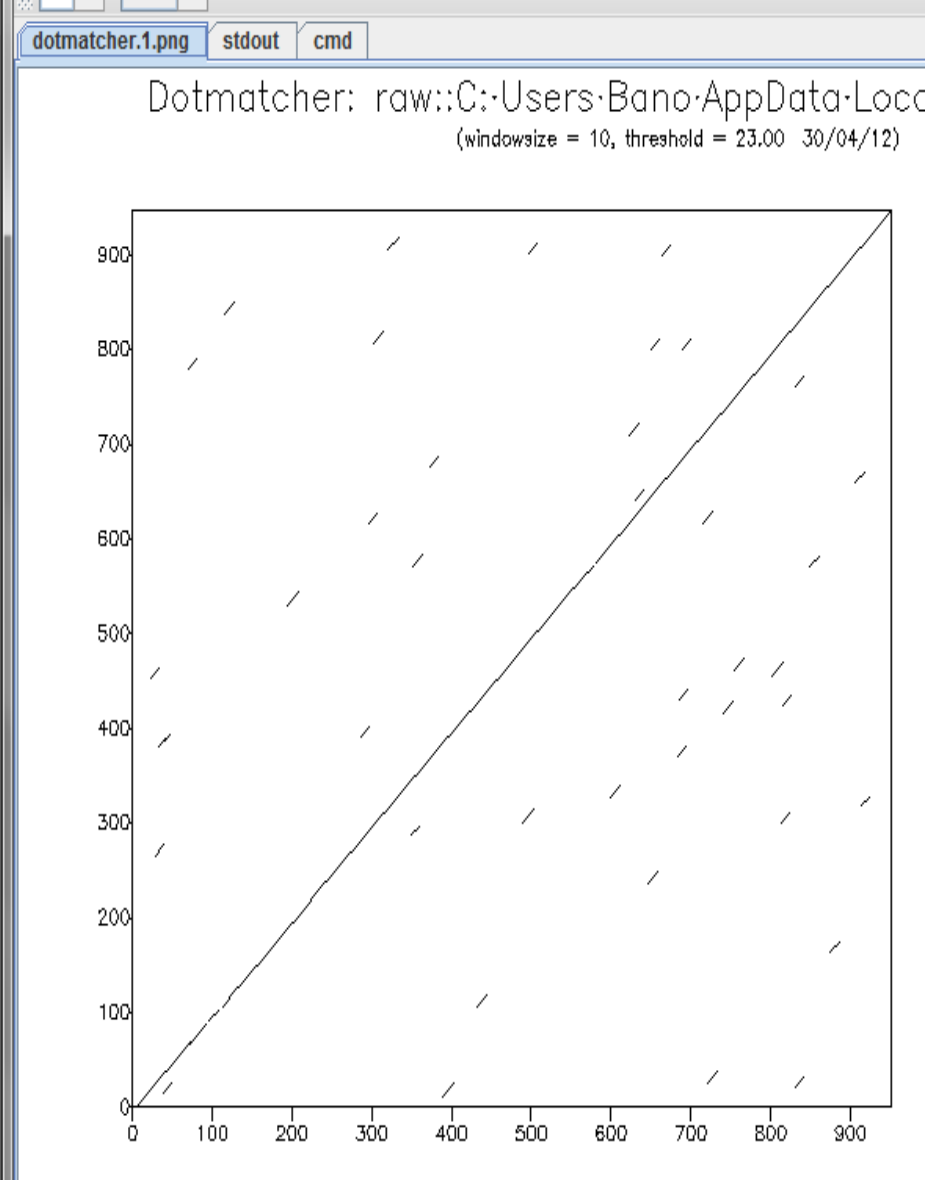
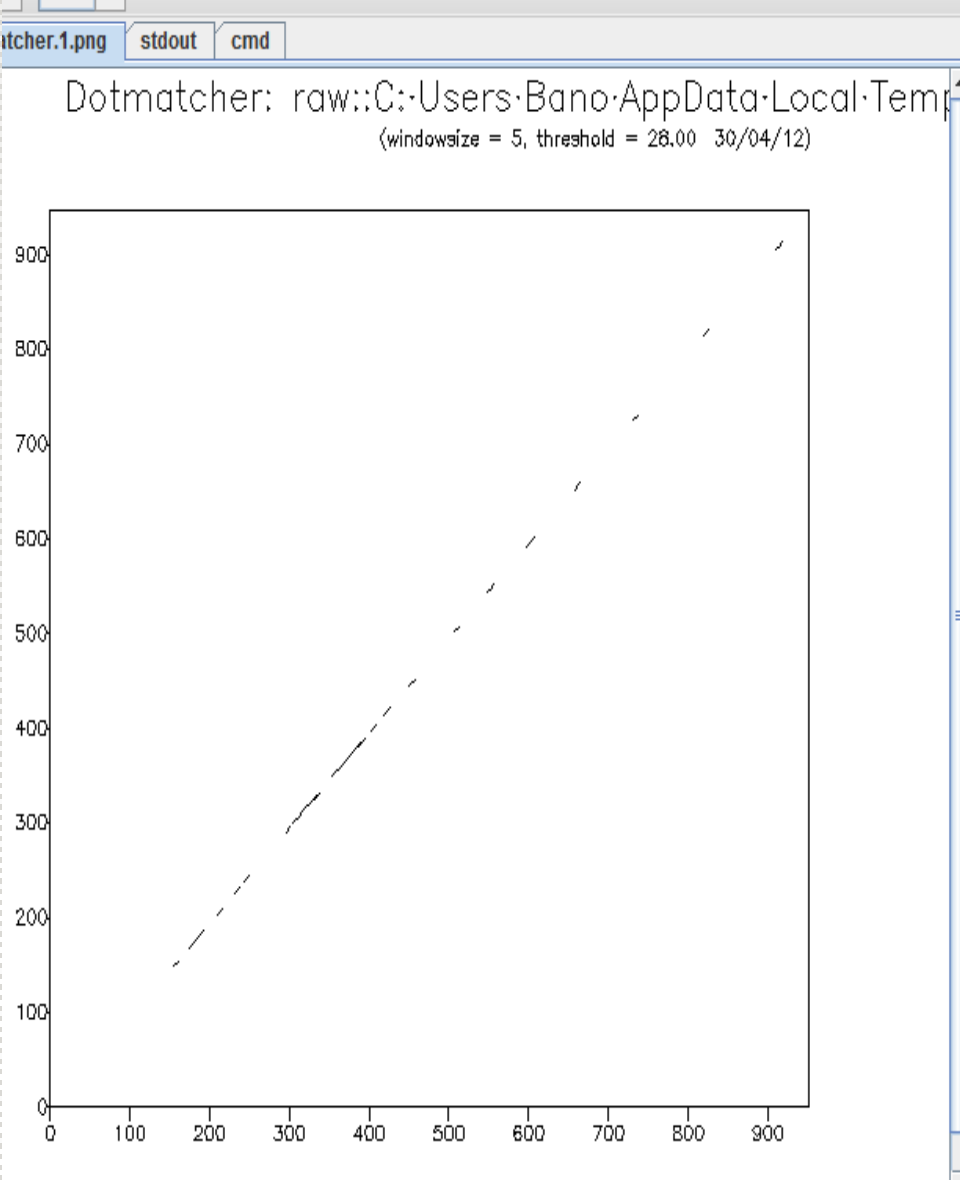
Execution mode:
interactive GO Advanced Options

additional section

10 Window size over which to test threshold
(min:3 default:10)

23 Threshold
(min:0 default:23)

(No background jobs)



Two graphs for the same alignment but with two different window size and threshold

You can see them in the top of the pages