



Introduction to **EMBOSS**

Shahid Manzoor

Adnan Niazi

SLU Global Bioinformatics Centre, Uppsala, Sweden

What is

The logo for the emboss software package. The word "emboss" is written in a bold, blue, sans-serif font with a 3D effect and a drop shadow. The letter 'e' is stylized with three red spheres of varying sizes positioned around its top and left sides, resembling a molecular structure or a protein binding site.

?

- A free Open Source software analysis package developed for molecular biology
- Programs share a common look and feel
- Incorporates many small and large programs
- Easy to run from the command line
- Retrieval of sequence data from the web
- Easy to call from other programs(e.g. perl)
- Easy to set up behind GUIs and Web interfaces

Scope of applications

- There are many EMBOSS programs (200+)
- Many sequence analysis & display programs
- Protein 3D structure prediction being developed
- Other programs e.g enzyme kinetics

What is wEMBOSS?

- A web interface to the EMBOSS package for sequence analysis
- It's developed by Martin Sarachu (Argentina) and Marc Colet (Belgium)



EMBOSS program: **wosname**

- Easy to forget a program name
- To find programs, use **wosname**
- **wosname** finds programs by looking for keywords in the description or the name of the program
- E.g a search for the keyword **protein** will list all the protein analysis associated tools

EMBOSS program: **seqret**

- Reads in a sequence and writes it out
- Reformat sequences
- Get sequences from databases

EMBOSS program: **showdb**

- Displays information of the currently available databases

Examples of other EMBOSS programs

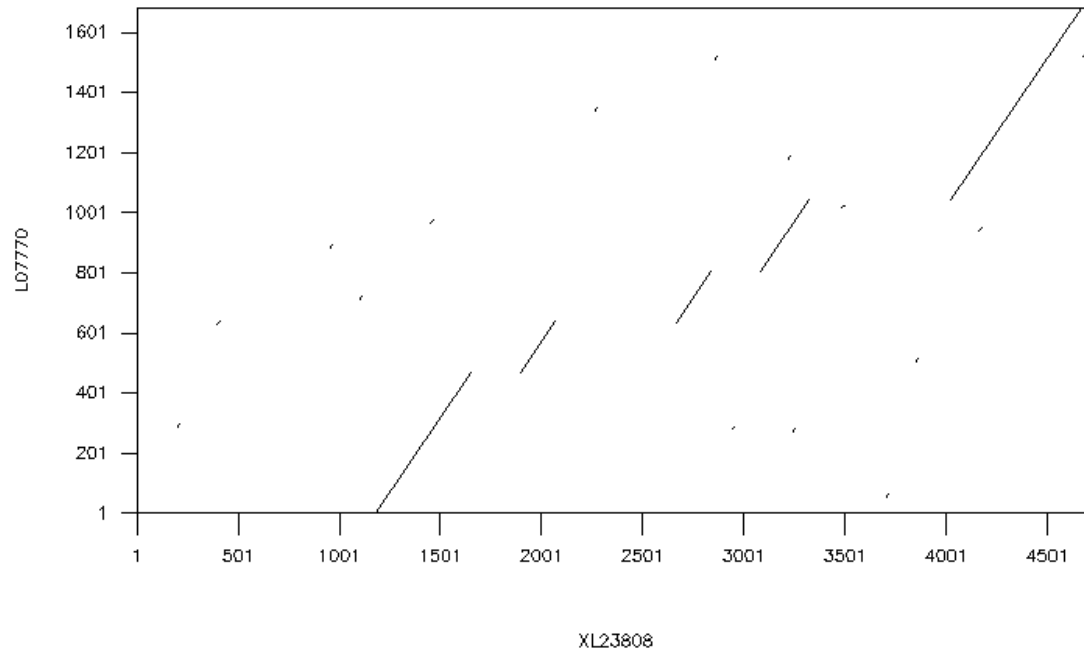
- Pairwise alignment -Dotup
- Global and local alignment – needle, water
- Protein translation – plotorf, getorf, transeq
- Protein fingerprint - pscan
- Primers - eprimer3
- Microsatellites -equicktandem

Pairwise alignment - Dotup

- One sequence is represented on each axis and significant matching regions are distributed along diagonals in the matrix.

Dottup: fasta::xl23808:XL23808[1:4734] vs fasta::xlrhodo...

Mon 21 Apr 2008 13:46:34



EMBOSS program: water

- Does a local alignment of two sequences

```
L07770           88  aggactttatagagctgctaccatgaacggaacagaaggtccaaatttt    137
      |||||.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|
BMU59922        10  AGGACTTTAAAGAGCCGCCAATATGAACGGAACAGAAGGCCCAAACTTTT    59

L07770          138  atgtccccatgtccaacaaaactgggggtggtacgaagccattcgattac    187
      |..|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|
BMU59922        60  ACATACCCCATGTCCAACAAGACTGGGGTGGTACGAAGCCATTCGAATAC    109

L07770          188  cctcagtattacttagcagagccatggcaatattcagcactg-gctgctt    236
      |||||.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|
BMU59922        110  CCTCAGTATTACCTGGCAGAGCCATGGCAATACTCCGTTTTGTGC-GCAT    158

L07770          237  acatgttcctgctcatcctgcttgggttaccaatcaacttcatgaccttg    286
      |||||.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|
BMU59922        159  ACATGTTCTCTGCTCATTCTTCTTGGGTTCCCAATCAACTTCATGACCTTG    208

L07770          287  tttgttaccatccagcacaagaaactcagaacaccccctaaactacatcct    336
      |..|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|
BMU59922        209  TACGTCACCATCCAGCACAAGAAGCTCCGGACACCCTTAAACTATATCCT    258

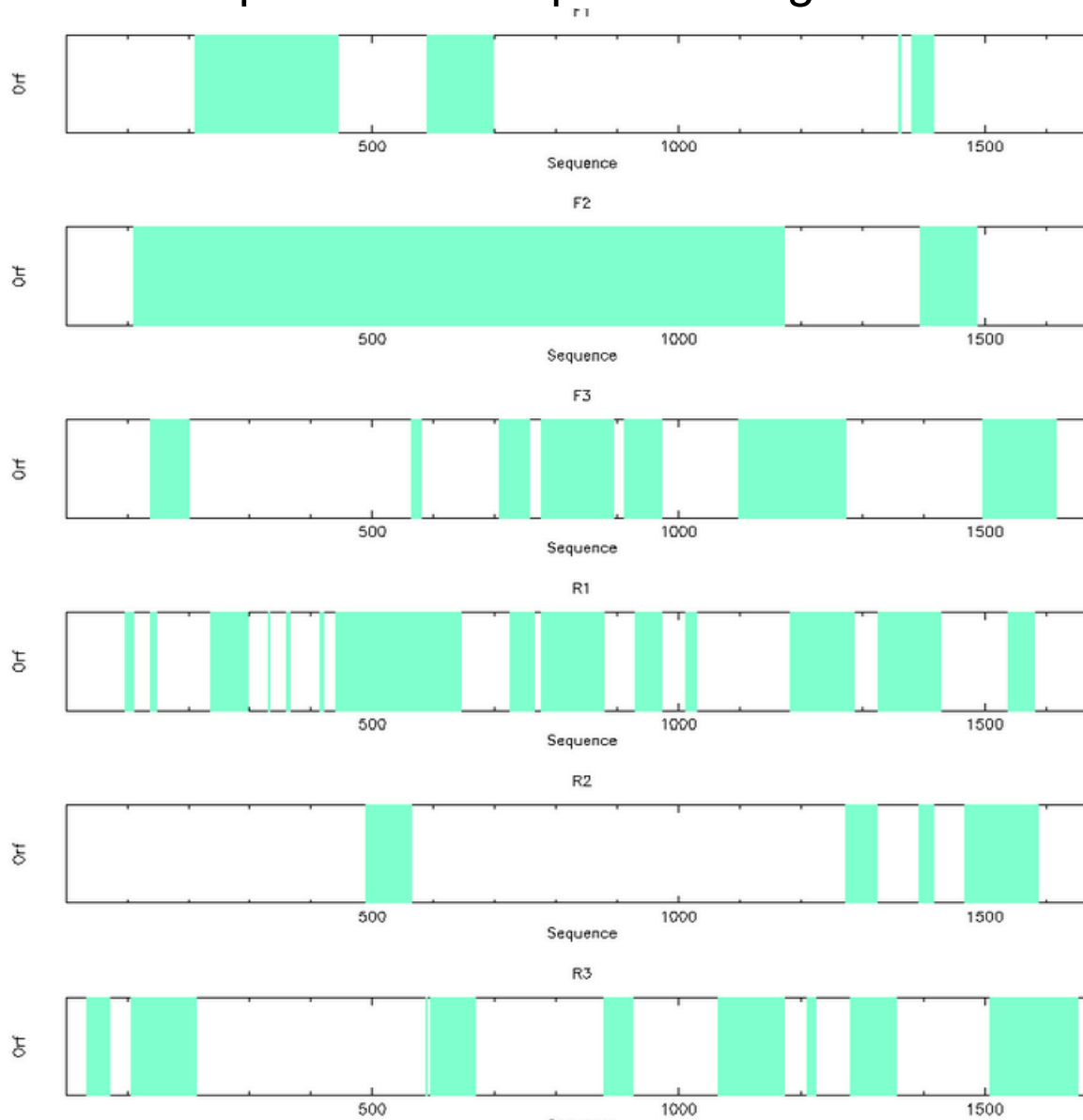
L07770          337  gctgaacctggtatttggccaatcacttcatggtcctgtgtgggttcacgg    386
      |||||..|||..|||..|||..|||..|||..|||..|||..|||..|
BMU59922        259  GCTGAATTTGGCCTTTGCCAACCACTTCATGGTCCTGTGTGGATTACCG    308

L07770          387  tgacaatgtacacctcaatgcacggctacttcatctttggccaaactggt    436
      |.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|
BMU59922        309  TCACAATGTACTCCTCAATGAACGGATACTTCGTCTTCGGACAAACCGGT    358

L07770          437  tgctacattgaaggcttcttggctacacttggtggtgaagtggccctctg    486
      |||||..|||..|||..|||..|||..|||..|||..|||..|||..|
BMU59922        359  TGCTATGTTGAAGGCTTCTTCGCTACCCTTGGTGGTGAAATCGCCCTTGG    408
```

EMBOSS program: plotorf

- Translates sequences to 6 open reading frames.



Start working on the tutorial

