

# Bioinformatics I

## Pairwise Sequencing For DNA & Protein Using Dynamic Programming

Dr Manaf A Guma

University Of Anbar- college of Applied sciences-Hit

Department of applied chemistry

1

## BLAST database for sequencing?

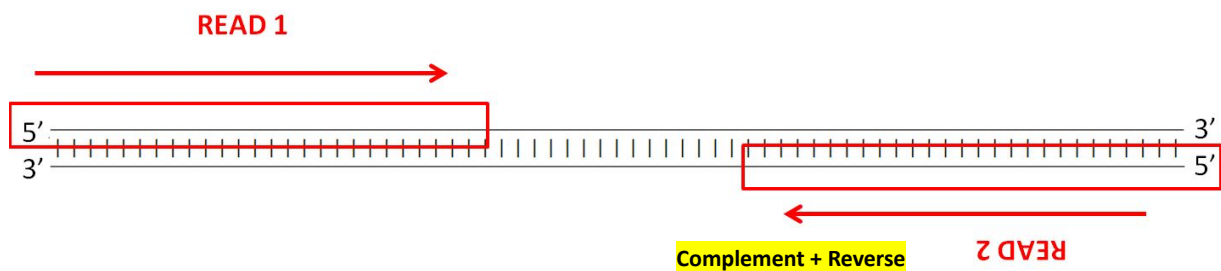
- **BLAST** Basic Local Alignment Search Tool
- A frequently used application of pairwise alignments is the search **for similar protein or nucleotide sequences in sequence databases.**
- With older dynamic alignment algorithms such as those designed by Smith and Watermann (1981) or Needleman and Wunsch (1970), this is too slow to perform even on current computers.
- **What BLAST look for?** It looks for the score of the alignment.
- **What does it indicate?** The BLAST Score indicates the quality of the best alignment between the query sequence and the found sequence (**hit**).
- **What does high/ low score represent?** The higher the score, the better the alignment.
- **When the score of the alignment is reduced?** Scores **are** reduced by mismatches and gaps in the best alignment

2

2

## How should we order the sample to check?

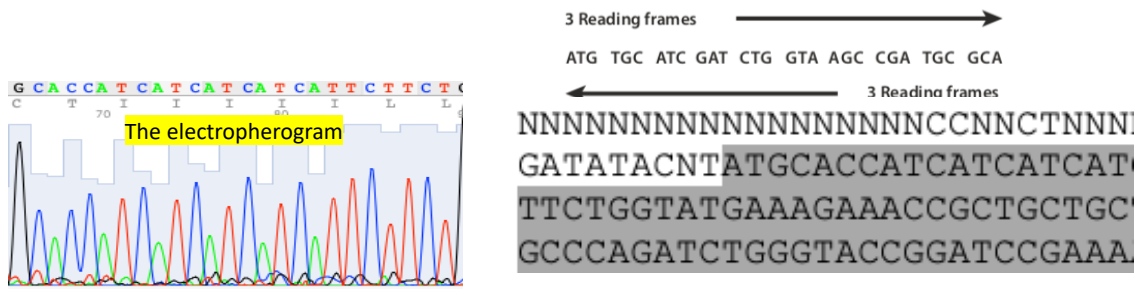
- Follow the rules of the companies, they ask for **specific concentrations (conc.)**.
- Usually, 0.375 ng/ 10  $\mu$ L for both (Forward **F** and Revers **R** )
- =. 75ng/ 20  $\mu$ L of DNA or RNA
- + 10ng/ 1  $\mu$ L of each **F** or **R**.



3

## Receiving the sample & starting check!

- Commonly, three files
- **name. seq.** (Forward **F** or Revers **R** ) =====Microsoft WORD
- **name. ab1** (Forward **F** or Revers **R** )===== needs a Software
- We need both (Forward **F** & Revers **R** ), if ... The seq is toooo long ! more than 400 bp.



4

## BLAST - general

- For a BLAST similarity search we need:
  1. input sequence (**query**)
  2. algorithm (implemented in the BLAST software)
  3. database (of protein or nucleotide sequences)
  4. Finally, we need to understand the output

5

## There are different types of BLAST

Basic BLAST

	Nucleotide BLAST	Protein BLAST	BLASTx	tBLASTn	tBLASTx
Query	DNA	protein	Translated DNA (6 frames)	protein	Translated DNA (6 frames)
Subject	DNA	Protein	Protein	Translated DNA in database (6 frames)	Translated DNA in database (6 frames)

6

## Translates DNA sequence into all six possible reading

- For example: if you translate the following seq
- AACCTGTATTTTCAGGGCGCCATG
- You will find:
- 6 possible proteins depends on the frameshift.

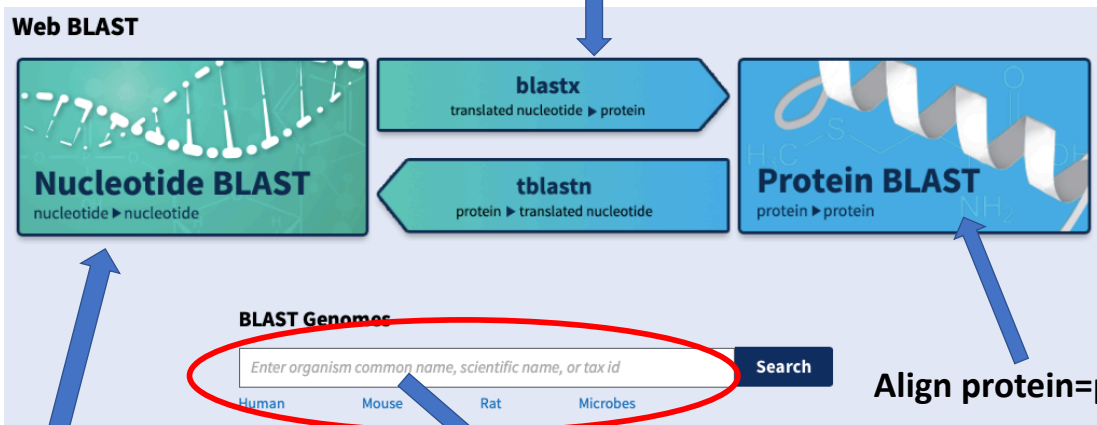
The diagram shows six possible protein translations from a DNA sequence in different reading frames:

- 5'3' Frame 1:** NLYFQGA**M**
- 5'3' Frame 2:** T**C**IFRAP
- 5'3' Frame 3:** P**V**FSGRH
- 3'5' Frame 1:** H**G**ALKIQV
- 3'5' Frame 2:** M**A**P-**K**YR
- 3'5' Frame 3:** W**R**PENTG

7

## The webs page

Align translated DNA=protein



Align DNA=DNA

Input the organism name

Align protein=protein

8

## Use BLAST database (alignment) to find the similarity?

- **The query:**

```
GAMDAIKKKMQMLKLDKENALDRAEQAEADKKAEDRSKQLEDELVSLQKKLKGTEDELDKYSEALKDAQEKLELAEKKATD
AADVASLNRRIQLVEEELDRAQERLATALQKLEEAKEKADESERGMKVIESRAQKDEEKMEIQEIQLKEAKHIAEDADRKY
EEVARKLVIIESDLERAEEERAELSEGKCAELEEEELKTVTNNLKSLEAQAEEKYSQKEDRYEEEEIKVLSDKLKEAETRAEFAER
SVTKLEKSIDDLEDELYAQKLYKAISEELDHALNDMTSI
```

- **The subject:**

- ASMDAIKKKMQLKLDKENALDRAEQAEADKKAEDRSKQLEDELVSLQKKLKGTEDELDKYSEALKDAQEKLELAEKKATD  
AADVASLNRRIQLVEEELDRAQERLATALQKLEEAKEKADESERGMKVIESRAQKDEEKMEIQEIQLKEAKHIAEDADRKY  
EEVARKLVIIESDLERAEEERAELSEGKCAELEEEELKTVTNNLKSLEAQAEEKYSQKEDRYEEEEIKVLSDKLKEAETRAEFAER  
SVTKLEKSIDDLEDELYAQKLYKAISEELDHALNDMTSI

9

## Translation the sequence to protein ( make it easy)

1. Check the alignment with your original copy using:

<https://blast.ncbi.nlm.nih.gov/Blast.cgi>

2. If you do not know the original copy of your protein, just type it

here: <https://www.uniprot.org>

10

Different styles of alignment

### Example: Protein Sequence check up

```

Query      1      MDAIKKKMQLKLDKENALDRAEQAEADKKAEDRSKQLEDELVSLQKKLKGTEDEL 60
Query_182929 47      .....E..... 106

Query      61      RISEALKDAQEKLEIAEKKATDAEADVASLNRRIQLVVEELDRAQERLATALQKLEEAEK 120
Query_182929 107      ..... 166

Query      121     AADESERGMKVIESRAQKDEEKMEIQEIQLKEAKHIAEDADRKYEEVARKLVIIESDLR 180
Query_182929 167      ..... 226
    
```

### Example: DNA Sequence check up

Sequence ID: Query\_35585 Length: 861 Number of Matches: 1

Range 1: 1 to 861 [Graphics](#) ▼ Next Match ▲ Previous

Score	Expect	Identities	Gaps	Strand
1591 bits(861)	0.0	861/861(100%)	0/861(0%)	Plus/Plus
Query 1	GCCAGCATGGACGCGATCAAGAAGAAGATGCAAATGCTGAAACTGGACAAAGAAAATGCG	60		
Sbjct 1	GCCAGCATGGACGCGATCAAGAAGAAGATGCAAATGCTGAAACTGGACAAAGAAAATGCG	60		
Query 61	CTGGACCGTGCCGAACAGGCGGAGGCGGACAAAGAAAGCGGCGGAGGATCGTAGCAAGCAG	120		
Sbjct 61	CTGGACCGTGCCGAACAGGCGGAGGCGGACAAAGAAAGCGGCGGAGGATCGTAGCAAGCAG	120		
Query 121	CTGGAAGACGAGCTGGTGAGCCTGCAAAAAGAAACTGAAGGGCACCGAAGACGAGCTGGAT	180		
Sbjct 121	CTGGAAGACGAGCTGGTGAGCCTGCAAAAAGAAACTGAAGGGCACCGAAGACGAGCTGGAT	180		

11

## Tutorial

- Check the pairwise similarity between two DNA sequences (given by you tutor) .
- What did you find?
- Discuss!

12