

# Förarbete bioinformatik

Inför ditt besök på Vetenskapens Hus behöver du förbereda dig genom att göra två uppgifter. En bioinformatik uppgift och en quiz på mentimeter.

## 1. Sekvensjämförelser: Identifiera konservativa sträckor

a. Markera all text i rutan nedanför:

```
>4
MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHS

>3
MFVFLVLLPLVSSQCVNLITRTQSYTNSFTRGVYYPDKVFRSSVLHSTQD

>2
MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHS

>1
MFVFLVLLPLVSSQCVNLRTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHS
```

b. Gå till länken <https://www.ebi.ac.uk/Tools/msa/clustalo/>

c. I rutan step 1 klistra in dina sekvenser enligt:

```
STEP 1 - Enter your input sequences
Enter or paste a set of
PROTEIN
sequences in any supported format:
MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHS
>3
MFVFLVLLPLVSSQCVNLITRTQSYTNSFTRGVYYPDKVFRSSVLHSTQD
>2
MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHS
>1
MFVFLVLLPLVSSQCVNLRTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHS
Or, upload a file: [Välj fil] Ingen fil har valts
```

d. Klicka på "submit":

```
STEP 3 - Submit your job
 Be notified by email (Tick this box if you wa
Submit
```

e. Efter ett tag syns följande resultat:

```
CLUSTAL O(1.2.4) multiple sequence alignment

 4   MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHS---   50
 2   MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHS---   50
 1   MFVFLVLLPLVSSQCVNLRTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHS---   50
 3   MFVFLVLLPLVSSQCVNLITRTQ---SYTNSFTRGVYYPDKVFRSSVLHSTQD   50
      *****      *      ;*****
```

Titta på nedersta raden, en \* betyder att aminosyran i den positionen är identiskt i alla fyra sekvenserna. Den är alltså konserverat bevarad.

Där det saknas en \* så har det skett förändringar:

- 1) I den första positionen som saknar en \* ser vi att två sekvenser har T medan de andra har R och I.
- 2) I den andra positionen ser vi att sekvens 3 har tre aminosyror deleterade medan de övriga tre har LPP som aminosyror.

De sekvenser som du har jämfört tillhör ett protein som heter Spike, som vi kommer att arbeta med under besöket på Vetenskapens Hus. Utifrån ditt resultat ovan, kan du nämna en sträcka på minst fyra aminosyror som är konserverat i proteinet? (Vi besvarar den frågan när du besöker oss)

## 2. Läsramar och kodoner

a. En DNA-sträng kan översättas på tre olika sätt:

1  
↓  
GGATCTGATCGAATGGATACGGTAGCC

1:a läsramen:  
GGATCTGATCGAATGGATACGGTAGACC  
G S D R M D T V D

2  
↓  
GGATCTGATCGAATGGATACGGTAGCC

2:a läsramen  
GGATCTGATCGAATGGATACGGTAGACC  
D L I E W I R STOP

3  
↓  
GGATCTGATCGAATGGATACGGTAGCC

3:e läsramen  
GGATCTGATCGAATGGATACGGTAGACC  
I STOP

Du kan välja första, andra eller tredje position för att hitta en läsram för ditt protein.

Om du matar in en okänd sekvens i ett translationsprogram (översätter DNA till protein) så kommer programmet att översätta till alla tre läsramarna eftersom det inte vet vilket protein som är det rätta.

b. Att beskriva sekvensskillnader korrekt

En sekvens jämförs alltid mot en referenssekvens. I vårt fall (SARS-CoV2) så är referenssekvensen den ursprungliga sekvensen, den sekvens som först isolerades i Wuhan-området.

Anta att vi jämför två sekvenser mot varandra som är 11 respektive 12 aminosyror långa.

Den nedre sekvensen är vår referenssekvens:

```

1 MFCVNLITRTQ- 11
a. MFCVNLRLTRTQL 12
   ***** **
  
```

I position 7 så har proteinet muterat från R till I. Då skriver vi det som R7I.

I position 10 så har en aminosyra försvunnit, deleterats. Då skriver vi det som L12del.

c. I quizlänken finns det några uppgifter med att förstå vilka punktmutationer som har skett på DNA-nivå. Nedanstående tabell kan då vara till nytta

	<b>T</b>	<b>C</b>	<b>A</b>	<b>G</b>
<b>T</b>	T T T Phe F T T C Phe F T T A Leu L T T G Leu L	T C T Ser S T C C Ser S T C A Ser S T C G Ser S	T A T Tyr Y T A C Tyr Y T A A Stop T A G Stop	T G T Cys C T G C Cys C T G A Stop T G G Trp W
<b>C</b>	C T T Leu L C T C Leu L C T A Leu L C T G Leu L	C C T Pro P C C C Pro P C C A Pro P C C G Pro P	C A T His H C A C His H C A A Gln Q C A G Gln Q	C G T Arg R C G C Arg R C G A Arg R C G G Arg R
<b>A</b>	A T T Ile I A T C Ile I A T A Ile I A T G Met M	A C T Thr T A C C Thr T A C A Thr T A C G Thr T	A A T Asn N A A C Asn N A A A Lys K A A G Lys K	A G T Ser S A G C Ser S A G A Arg R A G G Arg R
<b>G</b>	G T T Val V G T C Val V G T A Val V G T G Val V	G C T Ala A G C C Ala A G C A Ala A G C G Ala A	G A T Asp D G A C Asp D G A A Glu E G A G Glu E	G G T Gly G G G C Gly G G G A Gly G G G G Gly G

Gör ett test på <https://www.menti.com/p2xqw3bkij> och testa dina kunskaper.