



Schimpanse är en nära besläktad art till oss människor

Foto: Bildkollage med bilder från Pixabay

Människans närmaste släkt

– Multipel sekvensinpassning av mtDNA

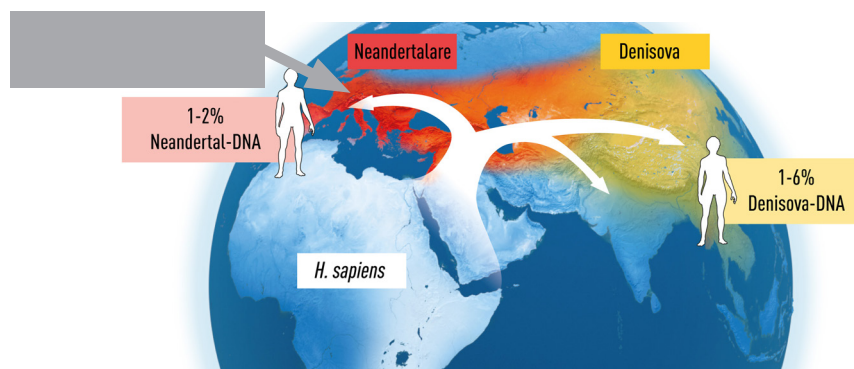
2022 fick Svante Pääbo Nobelpris för sin forskning om utdöda människors arvs massa. Metoderna som utvecklades för att studera gammalt DNA öppnade upp ett nytt vetenskapsfält, paleogenomik.

De första studierna av DNA från utdöda människor gjordes på mitokondrie-DNA (mtDNA). Människans referensgenom* för mtDNA består av 16 569 bp (baspar), medan kärnans DNA (förkortas nDNA) innehåller ca 6 000 000 000 bp, om man räknar in alla 46 kromosomer som finns i kärnan. Det är lättare att få fram mtDNA än nDNA ur gamla benbitar, eftersom det finns många mitokondrier i celler (men bara en cellkärna, se bild till höger). Så det var så man började forskningen.

I övningen ska du undersöka vad mtDNA säger om släktskapet mellan nutida människor och utdöda människoslag. Läs först ur pressmeddelandet till Nobelpriset:

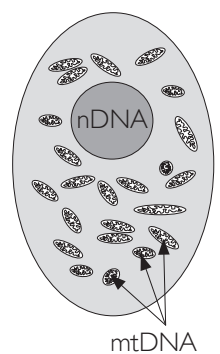
* Ett referensgenom tas oftast fram utifrån flera olika individer. Det ger ett slags "medelvärde" för vilken DNA-sekvens som är vanligast för en art.

"Pääbos upptäckter har gett oss en ny förståelse för vår evolutionära historia. Vid tiden för *Homo sapiens* migration ut ur Afrika levde åtminstone två nu utdöda homininer i Eurasien. Neandertalare levde huvudsakligen i västra Eurasien, medan denisova levde i de östra delarna av kontinenten. Under expansionen utanför Afrika korsades *Homo sapiens* med neandertalare och under vandringen vidare österut, även med denisova."



© The Nobel Committee for Physiology or Medicine. Ill. Mattias Karlén

Pääbos upptäckter har lärt oss hur världen var befolkad vid den tidpunkt då grupper av *Homo sapiens* lämnade Afrika och spred sig över världen. Neandertalare levde i väster och denisova i öster på den Eurasiatiska kontinenten. När *Homo sapiens* spred sig över kontinenten skedde möten som än idag kan spåras i vårt DNA.



CCeller kan innehålla många mitokondrier.

Syfte och övergripande uppgift

Du kommer att öva problemlösning med genetiska data. Du får tillgång till fem sekvenser av mtDNA. Genom att använda ett bioinformatiskt verktyg, *Multiple sequence alignment*, ska du ta reda på vilken av de fem mtDNA-sekvenserna som hör till vilken av dessa arter:

- Schimpans (*Pan troglodytes*)
- Neandertalare (*Homo sapiens neanderthalensis*) *
- Nutida människor (*Homo sapiens*)
- Heidelbergmänniskan (*Homo heidelbergensis*) *
- Denisovamänniska (*Homo sapiens ssp. Denisova*) *

* Utdöda människoarter som uppkallats efter platserna där man gjorde tidiga fynd av fossil: I Neanderdalen i Tyskland gjordes fynd redan på 1800-talet. Söder om staden Heidelberg i Tyskland hittades i början av 1900-talet en käke som visade sig ha tillhört en person som levde för ca 600 000 år sedan. I Denisova-grottan i Sibirien har man bland annat hittat ett fingerben från en ung kvinna.

Lite kort om bioinformatik och upplägg för övningen

Bioinformatik handlar om att samla, sortera och analysera biologiska data, som exempelvis DNA-sekvenser. Det finns många olika program som kan användas för att göra detta. Här använder ni ett verktyg som kallas *Multiple Sequence Alignment* via programmet ClustalW som finns tillgängligt online. Instruktioner för hur man gör finns på nästa sida.

Programmet jämför alla sekvenser man matar in parvis och gör en matchning (*alignment*) mellan alla sekvenser. Om två sekvenser liknar varandra mycket ger det ett högt så kallat *Score*-värde. På nästa sida visas exempel på hur ett resultat kan se ut, och vad man kan läsa ut av det.

Efter alignment kan man ta fram ett släktträd. I den här övningen väljer du en metod för dessa som kallas för maximum likelihood

Efter att du fått fram ett förslag på släktträd ska du klura ut vilken art som passar in på de olika grenarna i släktträdet. För att komma framåt med det har vi delat upp övningen i olika deluppgifter och det finns även ledtrådar till dem (sista sidan).

Uppgift 1

Följ metoden som visas på nästa sida för *Multiple Sequence Alignment* i programmet ClustalW. Använd informationen du får fram i resultatfönstret för att svara på frågorna:

- a. Vilken sekvens tillhör mest troligt en nutida människa? Hur vet du det?
- b. Vilken sekvens är mest lik nutida människan i det resultat du fått fram?
- c. Vilken sekvens borde tillhöra schimpansen? Motivera hur du tänker!

Multipel sekvensinpassning

Metod

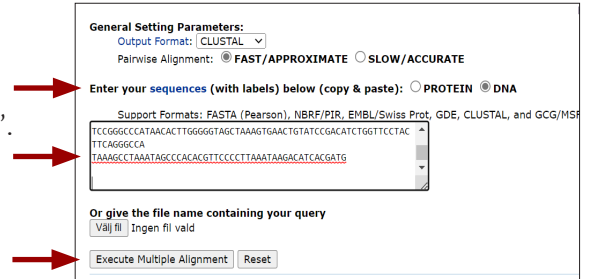
1. Hämta DNA-sekvenser i filen *mtDNAsekvenser*.
2. Markera allt och kopiera (tips på kort-kommandon till höger).
3. Öppna webbsidan ClustalW:
<https://www.genome.jp/tools-bin/clustalw>
4. På raden "Enter your sequences..." markera "DNA".
5. Klistra in kopierade sekvenser i den vita rutan.
6. Klicka på *Execute Multiple Alignment*. Vänta.
7. Resultatet är en lång textside.

Exemplet nedanför hjälper dig att lösa uppgifterna.

Kortkommandon

För PC/Windows, tryck CTRL + bokstav.
För Mac/macOS, tryck CMD + bokstav.

CTRL/CMD + A = Markera allt
CTRL/CMD + C = Kopiera
CTRL/CMD + X = Klipp Ut
CTRL/CMD + V = Klistra In
CTRL/CMD + F = Sök



```

CLUSTALW Result
[clustalw.aln][clustalw.dnd][readme]
Select tree menu v Exec

CLUSTAL 2.1 Multiple Sequence Alignments
Sequence type explicitly set to DNA
Sequence format is Pearson
Sequence 1: FJ490764.1 416 bp
Sequence 2: KJ157767.1 868 bp
Sequence 3: HM006842.1 1775 bp
Start of Pairwise alignments
Aligning...

Sequences (1:2) Aligned. Score: 67.5481
Sequences (1:3) Aligned. Score: 61.5385
Sequences (2:3) Aligned. Score: 68.6636
Guide tree file created: [clustalw.dnd]

There are 2 groups
Start of Multiple Alignment

Aligning...
Group 1: Sequences: 2      Score:13280
Group 2: Sequences: 3      Score:5840
Alignment Score 8117

CLUSTAL-Alignment file created [clustalw.aln]
    
```

Exempel på ett resultatfönster

Sekvensernas namn och längd (t ex 416 kvävebaser, förkortat bp för baspar).

1:2 betyder att sekvens 1 och 2 har jämförts och fått en poäng (*Score*). Ju högre värde desto mer lika varandra är sekvenserna.

Längre ned finns sekvenserna med 60 bokstäver (kvävebaser) per rad. Sekvenserna är matchade (*alignment*) och ligger under varandra.

```

KJ157767.1 TAAGAGGAAAAATCCGTTGACTTTAAAAATCGTGAGGGTTCAAGTCCCTCTATCCCCAAAA
HM006842.1 TAAAAGGAAAAATCCGTCGACTTTAAAAATCGTGAGGGTTCAAGTCCCTCTATCCCCAAAA
FJ490764.1 -----GGTTCAAGTCCCTCTATCCCCAAAA
*****

KJ157767.1 AGCCCATTTGACTCCTTAACATACTATTTATCTTATCTTTTTTTCGTTAGTAGTTC
HM006842.1 AGACTATTTAACTCCCAACTATT-----TATCCGACCCCTTTCC--TTAGCGGTTT
FJ490764.1 AGCCCATTTCAACTCCT----TATT-----ATCTTATCTTTTTTTCGTTAGTAGTTC
** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** *

KJ157767.1 -AAATTCGTTATCTTTCTTATTCACCCCTACTCTTTTACAAACGGATC-----CTAGAGA
HM006842.1 CAAATTCCTTATCTTTCTCATTCACTCTATTCTTTAGAAATGGATTTTTTTTCTAAGCG
FJ490764.1 AAAATTCGTTATCTTTCTTATTCACCCCTACTC-----AACGGATC-----CGAGAGC
*****
    
```

Eftersom sekvenserna är längre än 60 bp så följer flera "block", i det här fallet tre stycken.

↑ Stjärnor (*) markerar positioner som är lika för alla sekvenser. Ju fler * i en alignment, desto mer lika varandra är sekvenserna.

Uppgift 2

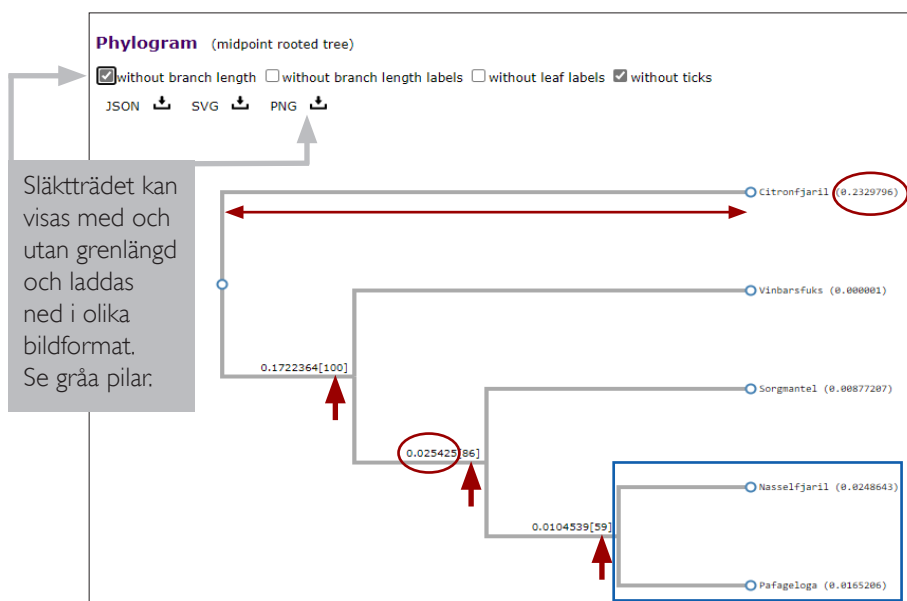
Använd den alignment du gjort för att ta fram ett släkträd. Använd släkträdet för att svara på uppgifterna här nedanför.

Metod

1. Överst på sidan finns "Select tree menu".
Välj RAxML bootstrap och klicka på "Exec".
2. En ny flik öppnas i webbläsaren. Vänta.
3. När resultatet är klart kan du scrolla ner för att se släkträdet under rubriken *Phylogram*. Ett exempel på ett släkträd visas här nedanför. Där ges en enkel förklaring till hur trädet ska läsas och också till vad siffrorna betyder.
 - a. Rita en skiss av trädet du får fram i det tomma utrymmet här nedanför. Skriv ut siffrorna 1–5 vid spetsen på grenarna.
 - b. Skriv in nutida människan vid rätt nummer för sekvensen (se uppgift 1 a).
 - c. Skriv in schimpans vid rätt gren i släkträdet (se uppgift 1 c).



Exempel på släkträd (gjort för fem vanliga vårffjärilar)



Grenlängd (se röda ringar och dubbelpil). Större siffra betyder fler mutationer sedan närmaste förgrening i släkträdet.

Datorprogrammet bygger olika träd från det data vi matat in. Bootstrapvärden inom [] vid de röda pilarna talar om i hur många av 100 träd som den grenen finns med. Ett slags "stabilitetsmått" för en viss gren/förgrening. Maxvärdet är [100]. Värdena ska helst vara över 70.

Den blå rutan visar två arter som är närmare släkt med varandra än vad de är med de andra arterna i trädet. De delar en närmaste gemensam förfader som de inte delar med någon annan i släkträdet.

Uppgift 3

Följande sekvens är känd från ett tidigare fynd av mtDNA hos en neandertalare:

CTAGCCCCAAACCCACTCCACCTTACTACCAAACAACCTTAGCCAAA

För att avslöja vilken av sekvenserna 1–5 som tillhör neandertalaren **gör så här**:

1. Kopiera sekvensen här ovanför.
2. Gå till resultatfönstret i ClustalW (den första fliken i webbläsaren).
3. Öppna webbläsarens sökfunktion. Du får upp en sökruta uppe till höger.
4. Klistra in sekvensen i sökrutan.
5. Se efter om sekvensen matchar någon av sekvenserna 1–5. Om du får en matchning visas den med en gulmarkering på webbsidan (det syns också i resultatfönstret 1/1 om du får en träff, annars 0/0).
 - a. Vilken sekvens tillhör neandertalaren?
 - b. Skriv in neandertalaren vid rätt gren i släkträdet från förra uppgiften.

Uppgift 4

Släkträdet har kvar ett par grenar med bara nummer. Du ska placera ut denisovamänniskan och heidelbergmänniskan. Som ledtråd får du veta att denisovamänniskan har visat sig ha det längsta mtDNA-genomet av alla homininer.

Skriv in denisova och heidelbergmänniskan vid rätt grenar i släkträdet.

Uppgift 5

Vilken är den moderna människans närmaste släkting enligt släkträdet?

- a. Vilken är den närmaste släktingen till heidelbergmänniskan enligt dina resultat?
- b. Ringa in och markera med ett "A" vilken förgrening i trädet som visar att schimpansen och alla människoarter har en och samma gemensamma förfader.
- c. Ringa in och markera med ett "B" den gemensamma förfadern för alla människoarter som ingår i uppgiften.

Uppgift 6

Släktrådet som du fått fram bygger på mtDNA på ca 16 500 bp jämfört med kärnans nDNA på ca $2 \times 3\,000\,000\,000$ bp (vi är diploida och har två kopior av varje kromosom). Mitokondrier följer med ägget, spermerna bidrar i regel bara med nDNA. Därför ärvs mtDNA bara via mamman. Svante Pääbo började med att kartlägga mtDNA men fortsatte sedan med nDNA.

I pressmeddelandet för Svante Pääbos Nobelpris avslutade man med en bild där ett släktråd kan anas som vita "vägar" som en del av illustrationen:



Pääbos banbrytande arbete har skapat en grund för att förklara vad som gör oss unikt mänskliga.

- I förra uppgiften markerade du den gemensamma förfadern för människoarterna i uppgiften med ett "B". Markera samma sak med ett "B" i illustrationen här ovanför.
- Hur kan man rita in en "väg" för schimpansen i illustrationen?
- Varför står vår egen art längre fram i bilden än denisova och neandertalare?
- Jämför hur närmaste släktingen till denisovamänniskan visas i illustrationen med hur det ser ut i ditt släktråd. Ge förslag på tänkbara förklaringar till det du ser när du jämför.

Ledtrådar

Uppgift 1

- I bakgrundstexten på sida 1 finns fakta om längden på mtDNA för nutida människor (och högst upp på resultatsidan i ClustalW finns information om sekvensernas längder). Den sekvens som stämmer överens motsvarar nutida människa.
- Jämför värden på *score* för parvisa jämförelser där sekvensen för nutida människor ingår (du måste alltså först lösa uppgift 1 a). Högst värde för de två sekvenser som liknar varandra mest.
- Du kan anta att schimpansens sekvens borde ge låga värden på *score* för alla parvisa jämförelser (jämför värden på *score* och se om det finns någon sekvens som ofta ger låga värden).

Uppgift 2

- Skissen behöver inte vara exakt med alla siffror. Det viktiga är att du får till ett liknande mönster med förgreningar och att du sätter siffrorna på rätt gren.
- Ta hjälp av ditt svar i uppgift 1 a.
- Ta hjälp av ditt svar i uppgift 1 c.

Uppgift 3

Använd gärna kortkommandon för att kopiera och klistra in (se tips på kortkommandon på sida 3).

När du får en träff så ser du numret på sekvensen längst till vänster på sidan. Det ska vara en träff på sekvens 1, 2, 3, 4 eller 5.

Uppgift 4

Gå tillbaka och titta i resultatfilen i ClustalW och leta efter den sekvens som har längst antal baser (bp). Den sekvensen motsvarar denisovamänniska.

Uppgift 5

Närmaste släktingar hittar man genom att följa en arts gren bakåt i tiden ned mot närmaste förgreningspunkt. Gå dit och följ sedan den andra grenen ut i spetsen och läs av vilken som är den närmaste släktingen.

Uppgift 6

Läs faktadelen på övre delen av sidan en gång till. Illustrationen från pressmeddelandet syftade till att summera den senaste forskningen som Svante Pääbo och kollegor har kommit fram till.