



Schimpansen är en nära besläktad art till oss människor

Foto: Bildkollage med bilder från Pixabay

Människans närmaste släkt

– Lärarhandledning

Med inspiration av Nobelpriset 2022 till Svante Pääbo har vi tagit fram en bioinformatikövning där man tar fram ett fylogenetiskt träd baserat på mitokondrie-DNA och utifrån ett antal ledtrådar tar reda på vilken art som hör hemma på vilken gren i släktträdet.

Syftet med övningen är att öva på problemlösning utifrån genetiska data och att träna på att tolka fylogenetiska träd.

För vem? Övningen har prövats ut av lärare och elever på både högstadiet och gymnasiet (kurserna biologi 1 och naturkunskap 1b). Lärares utvärdering pekar på att det kan vara stor spridning i hur lätt eller svårt elever har att klara övningen. Detta återspeglar sig även i lärarnas skattning i hur lång tid de uppfattat att övningen tar för elever. Beroende på elevgruppens förkunskaper kan de ledtrådar som ges sist i elevinstruktionen delas ut redan från början eller så behåller du dem som lärare om du vill utmana eleverna mer.

Hur lång tid tar det? Elevutvärderingarna visade på ganska stor spridning i hur lång tid det tog att göra själva datorövningen. De flesta har angett mellan 30-60 minuter för att göra alla uppgifter. Beroende på belastning i datortrafik ingår viss väntetid för att analysen av datat ska bli klart.

Vad behövs? Datorer med internetuppkoppling. Möjlighet att ladda ner en fil med mtDNA-sekvenser. Elevinstruktionen digitalt och/eller utskrivet.

På följande sidor ger vi dels förslag på hur övningen kan ramas in i biologiundervisningen, förslag på svar till uppgifterna samt några didaktiska råd baserade på forskning om användning av bioinformatik i skolan.

Mer information

Övningen "Människans närmaste släkt - Multipel sekvensinpassning av mtDNA", finns på Bioreurs webbplats under Resurser och Bioinformatik.

Övningens kontext

Innan ni kör igång med övningen är det viktigt att fundera över vilket sammanhang den får i undervisningen. Kanske vill ni fokusera på vad forskning kan vara och uppmärksamma ett Nobelpris? Kanske arbetar ni med evolution och vill knyta ihop det här med släktskap med genetik och DNA? Eller vill du utmana eleverna i att följa en instruktion för att navigera i en digital miljö? Beroende på sammanhang kan övningen användas för att nå olika lärandemål.

Oavsett vilka lärandemål du väljer att fokusera på så behöver övningen ges en inledning och då finns material att hämta via Nobelprislektionen:

<https://nobelprizemuseum.se/medicinpriset-2022/>

I bildspelet som kan hämtas där finns bilder från pressmeddelandet som används i elevinstruktionen samt bilder som visar skillnaden i storlek mellan DNA från mitokondrier (mtDNA) och från kärnan (nDNA). I ett tillhörande bildspelsmanus får du tips på vad som kan berättas kring de olika bilderna.



Något om bioinformatik och alignment

I elevinstruktionen finns en kort text om vad bioinformatik är. För att ge övningen en metodologisk kontext kan man gärna ta upp detta med att nya typer av data och metoder gör att vi kan undersöka saker i biologin på nya sätt. Bioinformatik används idag inom nästan all biologisk forskning.

För att förstå begreppet alignment kan det vara bra att med ett enkelt exempel förklara att det handlar om matchning, eller "inpassning". *Exempel:* två sekvenser liknar varandra, men är inte identiska. Den ena är lite längre än den andra. Men många baser stämmer bra överens. Hur kan man passa in dem mot varandra så att de hamnar så rätt som möjligt? Du får lägga in luckor med streck för att få till det:

ATTCGGATT GATTACGGTTT

En tänkbar lösning:

- A T T - C G G A T T

G A T T A C G G T T T

På detta sätt är de matchade, inpassade, alignade (uttal: alajnade). För att bedöma hur lika eller olika sekvenserna används en algoritm för att få fram Score-värden. Varje kvävebas som matchar perfekt får en viss poäng, och varje skillnad ger "straffpoäng" eller ett slags avdrag. Om två sekvenser skiljer sig åt mycket ger det ett lågt score-värde. Det finns många olika algoritmer och inpassningen kan styras så att den söker fram den matchning som ger maximala "poäng". Beroende på vilka algoritmer och inställningar man gör i programmen kan man få olika resultat. I den här övningen ändrar vi inga inställningar.

Oavsett vilka karaktärer vi använder för att studera evolution är det viktigt att tänka på att egenskaperna vi jämför är homologa, det vill säga att de har ett gemensamt ursprung. Ett antagande vi gör vid sekvensinpassningen är att kvävebaser från olika arter som matchas ovanför varandra är homologa och har ett gemensamt ursprung i någon "uranfader"-kvävebas. Vi måste leta efter enheter som vi tror har samma evolutionära ursprung oavsett vilken organisationsnivå (DNA/protein/egenskap) vi studerar organismerna på. En hjälp i detta är att jämföra med en avlägsen besläktad art.

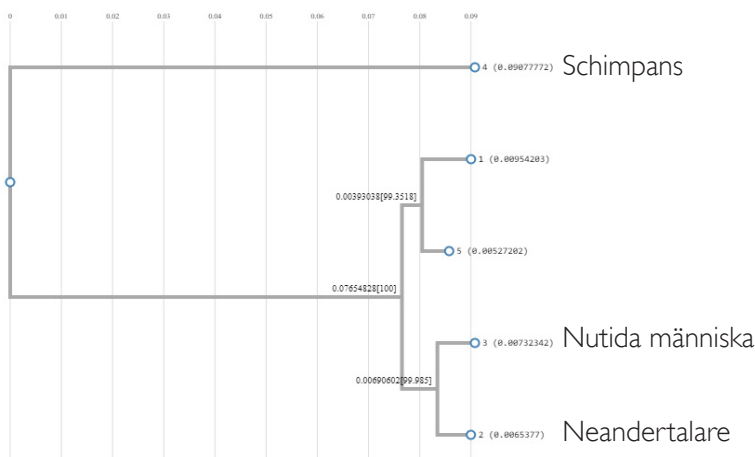
Svarsförslag och kommentarer till uppgifterna

Uppgift 1

- Den nutida människan passar in på sekvens 3 med motivering att den sekvensen är exakt lika lång som referensgenomet för människor som anges i den inledande faktatexten (16569 bp). OBS! Det finns en variation mellan människor i exakt hur långt mtDNA är, men vi har här valt att använda referensgenomet som riktmärke för att kunna använda som en ledtråd. Bland de övriga sekvenserna finns tre som har längder mycket nära detta värde (16567-16570 bp), medan en sekvens (nr 4) avviker mer (16554 bp). Det kan vara bra att diskutera styrkan i denna typ av "bevis". Enbart längden på mtDNA säger egentligen inte så mycket.
- Sekvensen som är mest lik den nutida människan är sekvens 2 vilket man ser genom att titta igenom Score-värdena i resultattabellen. Score 98.654 för den parvisa jämförelsen mellan sekvenserna är både det högsta värdet för alla jämförelser, och framförallt det högsta värdet för alla jämförelser för gjorts för just sekvens 3 med de övriga.
- Sekvens 4 borde tillhöra schimpansen med argumentet att den borde vara mest olika de andra som alla är olika typer av människor. Sekvens 4 har dels en avvikande längd på sitt mtDNA (se uppgift 1a) och dels ger alla jämförelser med sekvens 4 låga Score-värden.

Uppgift 2

- En skiss kan ritas för hand, men även laddas ned från resultatsidan. I elevinstruktionen finns en figur med förklarande text på sidan 4. Ett tillägg är att man kan summera värden som anges inom parentes efter numret på sekvenserna i spetsarna på grenarna (som motsvarar avståndet till närmaste förgreningspunkt) med de värden som ges på förgreningarna för att få fram olika distansmått.
- Nutida människan skrivs in på gren med siffran 3 på.
- Schimpans skrivs in på grenen för sekvens 4.

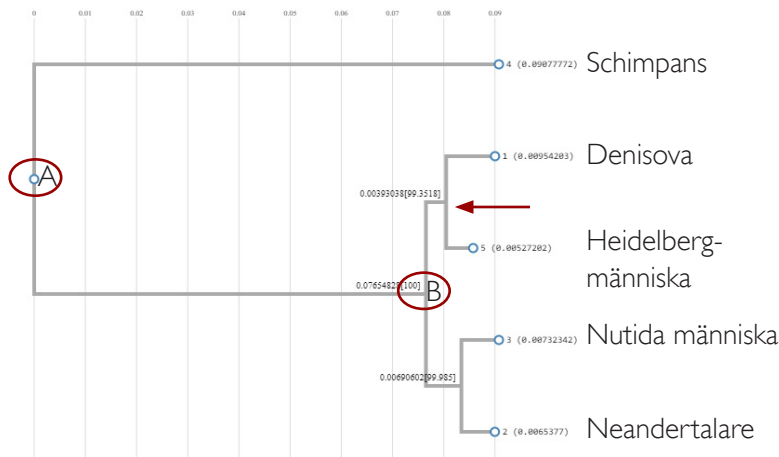


Uppgift 3

- Neandertalare motsvarar sekvens 2.
- Neandertalare skrivs in på grenen med siffran 2, det vill säga närmast släkt till nutida människan i detta släktträd.

Uppgift 4

Denisovamänniskan har det längsta mtDNA-genomet av alla homininer som ingår i just denna uppgift. Då tittar vi igen på sidan med resultat där den längsta mtDNA-sekvensen är på 16570 bp hos sekvens 1. OBS! Återigen är variationen i längd på mtDNA inget säkert data, men i uppgiften har vi använt detta värde för att skapa en ledtråd som fungerar för uppgiften. Genom utestlutningsmetoden finns nu bara en gren kvar som är för sekvens 5 vilket motsvarar Heidelbergmänniskan.

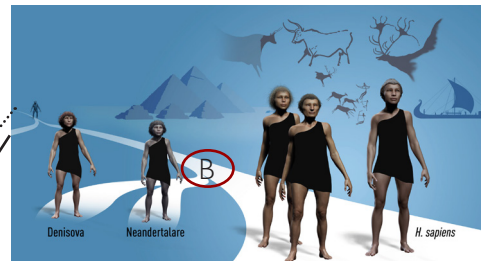


Uppgift 5

- Den närmaste släktingen till Heidelbergmänniskan är Denisovamänniskan enligt det släkttred man får fram utifrån mtDNA. Den röda pilen pekar på den närmaste gemensamma förfadern för dessa två människoarter.
- Se "A" i figuren ovan.
- Se "B" i figuren ovan.

Uppgift 6

- Se "B" i figuren här intill.
- Två olika förslag på släktgren för schimpans finns inritad i figuren intill. I illustrationen finns en liten figur långt bak som kan tolkas som en tidig människa - då borde man rita schimpansens gren som den prickade linjen (våra arter skildes åt längre tillbaka i tiden). Om den lilla figuren tolkas som en gemensam förfader för schimpanser och människor så kan den heldragna linjen fungera. Notera att linjerna ska nå lika långt fram i kanten som nu levande människor eftersom denna illustration visar en skillnad i grenlängd mellan nu levande och utdöda arter.
- Nutida människor står längre fram i bilden då vi lever idag till skillnad ifrån de arter som dött ut (som har grenar som avslutas en bit bakåt i bild).
- Illustrationen visar att neandertalare och denisova är närmast släkt med varandra och lika nära släkt med nutida människor då de delar samma gemensamma förfader (se markering "B"). I mtDNA-släkttredet var denisovas närmaste släkting heidelbergmänniskan, och nutida människor var närmast neandertalare. Släkttreds utseende beror på hur många arter som ingår. Utan sekvens 5 (Heidelbergmänniskan) blir ändå neandertalare närmaste släkting till nutida människor. Viktigt att tänka på är att illustrationen sammanfattar undersökningar av mer data då man idag kan analysera nDNA och inte bara mtDNA som vi gjort i övningen.



Några didaktiska råd

Forskning kring hur elever tar sig an och förstår bioinformatikövningar visar bland annat att de ofta klarar av själva knapptryckandet men att typen av instruktion belastar arbetsminnet olika mycket. Eleverna behöver ofta mest stöd i att knyta an görandet i datormiljön med själva biologin i den kontext man arbetar inom. Det finns flera didaktiska val som kan ge stöd för reflektion och koppling till det biologiska innehållet under genomförandet. Här ger vi några tips:

Arbeta själv eller i grupp? Ett förslag är att låta eleverna arbeta två och två vid en dator. De kan turas om att ha olika roller i själva genomförandet. En roll kan vara "Ordning och reda": läser instruktionen högt och leder på så sätt arbetet steg för steg. Den andra rollen kan vara "Nyfiken - varför?": klickar och genomför de olika stegen men har i uppgift att efter varje steg ställa frågan - varför gjorde vi det här? Ett annat alternativ är att gruppera eleverna om 3-4 per grupp där alla arbetar vid varsin dator men där en får rollen "Hänger du med?" för att se till att alla i gruppen tar sig framåt i övningen. Det är inte helt lätt att själv som lärare hantera frågor från många elever samtidigt i klassrummet. Det kan då underlätta att vissa problem löses inom elevgrupperna där de kan stötta varandra.

Utskriven eller digital instruktion? En fördel med att ha en utskreven instruktion är att man kan bocka av och skriva till anteckningar vartefter man tar sig framåt i övningen. En fördel med en digital instruktion är att länkarna i dokumentet är klickbara. I elevinstruktionen finns plats för att skriva korta svar på uppgifterna för hand. Nackdelen med enbart en utskreven instruktion är att man då behöver se till att man kommer till rätt webbplats för ClustalW (man behöver skriva in webbadressen eller söka fram den). Uppgift 3 bygger på att man ska kopiera en kortare sekvens för att matcha resultatet i ClustalW. Om man enbart har elevinstruktionen i pappersform behöver eleverna manuellt skriva in sekvensen i sökfönstret vilket kan ge skrivfel och tar viss onödig extra tid. Elevinstruktionen innehåller ett antal skärmdumpar med inringade områden för att uppmärksamma var man ska klistra in data och var man läser av olika resultat. Balansen mellan text och bild kan behöva anpassas beroende på elevers behov. Ett alternativ eller komplement till elevinstruktionerna kan vara att man som lärare visar på storbild först hur den digitala miljön ser ut som eleverna ska arbeta i. ClustalW är ett enkelt och användbart verktyg som går att använda för många andra uppgifter. I ett kortare videoklipp via [Bioresurs Youtubekanal](#) visar vi hur man arbetar i programmet ClustalW.

Lärostopp? Att stanna upp och gå igenom någon uppgift kan vara bra både för att få koll på hur det går för elevgrupperna men också ge tillfälle att gemensamt diskutera syftet med de olika steg som ingår i uppgiften. Visa delar av övningen via storbild. Här kan man även använda extra frågor som inte ingår i elevinstruktionen. Det kan hjälpa till att reda ut begrepp och fakta som kanske inte är självklara för alla, knyta an till begrepp som ni kanske har arbetat med tidigare för att underlätta det fortsatta arbetet. Några förslag på "checkfrågor":

Hur ser mtDNA-datat ut? Visa att sekvenserna i filen består av långa rader av kvävebaserna ATCG. Över 16500 bokstäver för varje sekvens. bp betyder baspar. Här kan man också ta upp en jämförelse: mtDNA-filen består av ca 1200 rader med vardera 70 bokstäver där filens storlek är 84 kb (kilobyte). Om vi skulle använt helgenomdata från kärnans DNA skulle vi behövt en fil med över 200 miljoner rader vilket motsvarar en fil på över 3 GB (gigabyte). Frågor som kan dyka upp här är varför det bara är en rad med bokstäver och inte två om man tänker på att vi lärt oss om DNAs

byggnad med basparning A-T och G-C. Svaret blir då att det räcker att analysera den ena strängen på grund av den komplementära basparningen. Det spelar ingen roll vilken sida vi jämför, men det förutsätter att vi jämför just samma sträng. Här kan man koppla till att DNA-strängarna har en riktning (3' och 5'-ändar skiljer sig åt kemiskt) vilket gör att vi kan veta vilken sträng vi jämför.

Hur ser en mutation ut? När man kört multiple sequence alignment och får fram resultatsidan så ser man hur de fem sekvenserna matchats, om man scrollar nedåt. Visa att alla positioner som har en stjärna (*) under är identiska mellan alla fem sekvenser. Leta efter positioner som saknar en stjärna för att visa exempel på mutationer. Här finns många exempel på där det skett ett utbyte av en kvävebas mot en annan. Det finns också exempel på insertioner och deletioner som visas med hjälp av streck (-). Om man tittar igenom datat ser man att särskilt sekvens 5 har avsnitt där det står flera "N" (istället för A, T, G eller C). Det är inte mutationer utan motsvarar positioner i sekvensen som visat sig vara svåra att få fram ett säkert resultat för när man sekvenserat materialet.

Vad finns det för koppling mellan mutationer och Score-värden? Ju fler skillnader som finns mellan två sekvenser desto lägre score-värde. Om vi tänker oss en ursprunglig sekvens så innebär fler mutationer över tid större skillnader mellan sekvenserna om vi följer dem framåt i tiden.

Hur ser kopplingen ut mellan mutationer och släktskap? Ett antagande som görs när man bygger släkträd på det sätt som görs i uppgiften är att ju fler skillnader som finns mellan två sekvenser desto längre tid har passerat sedan de tillhörde samma art. Om man vet hur många mutationer som sker per tidsenhet kan man räkna om antalet genetiska skillnader till tid - alltså få fram ett mått på hur länge sedan det är sedan två sekvenser var lika. Man brukar prata om en "molekylär klocka". Metoden är inte okontroversiell, för det har visat sig finnas en ibland stor variation i hur ofta mutationer uppkommer i olika delar av DNA-sekvenser och hos olika arter. Dessutom behöver man ta hänsyn till att organismer har olika generationstider om man vill räkna om till tid i termer av år.

Hur vet man vilken sekvens som är den "ursprungliga"? Det är svårt att veta. Det man ofta gör i den här typen av analyser är att man tar med en art som man vet är lite mer avlägset släkt, en så kallad "utgrupp". I den här uppgiften kan vi betrakta schimpanse som en utgrupp. I elevinstruktionen har vi kort nämnt att det finns olika sätt att bygga släkträd. I övningen används så kallade distansmått som grund för släktskapet som byggs upp. Det finns många olika metoder för att bygga släkträd som bygger på olika algoritmer och antaganden. Att enbart använda värden för hur stora skillnaderna är mellan olika arter ger oss inte tillräcklig information för att kunna uttala oss om hur evolutionen gått till. Men de är relativt snabba metoder som ändå ger intressanta mönster att diskutera.