

# Slægtskab, bioinformatik og biodiversitet

I 2010 rekonstruerede danske forskere som de første i verden et genom på grundlag af DNA fra en ca. 4.000 år gammel hårtot af et menneske fundet i et lag permafrost i Grønland. De kunne med stor sikkerhed fastslå at der var tale om en mand af en anden afstamning end de nuværende grønlandere, og at der derfor er sket indvandring til Grønland i flere omgange. Forskerne kunne også ud fra hårtottens DNA give en ret præcis beskrivelse af egenskaber hos manden, bl.a. at han var lille og tætbygget med tendens til skaldethed, havde brune øjne, mørk hud, skovl-formede fortænder og tørt ørevoks. Han havde blodtype A rhesus positiv og var disponeret for at udvikle bestemte sygdomme, se figur 182.

Sammenlignende DNA-analyser viste at manden er tættest beslægtet med nulevende befolkningsgrupper i det nordøstlige Sibirien, og det afslører at individer fra dette område sandsynligvis er migreret østpå til Grønland via Alaska og Canada for mellem 4.400 og 6.400 år siden.

For at nå frem til denne viden har forskerne fået bestemt baserækkefølgen af de bevarede dele af genomet, og derefter har de analyseret deres data ved brug af *bioinformatik* som er et nyere tværfagligt naturvidenskabeligt forskningsfelt. Bioinformatik er sammenligning af biologiske data med tilsvarende informationer i store databaser. De biologiske data kan fx være baserækkefølgen af hele genomer eller stumper af genomer fra forskellige organismer.

Det kan også være baserækkefølgen af RNA-molekyler eller større eller mindre dele af aminosyrerækkefølgen i et polypeptid. Bioinformatik er et uvurderligt redskab til at fastlægge slægtskab mellem individer og populationer som historien om hårtotten fra Grønland er et eksempel på. Det kan på samme måde bruges til at fastlægge slægtskab mellem forskellige arter. Sådanne *fylogenetiske analyser* har i flere tilfælde ført til justeringer af vores hidtidige opfattelse både af hvordan forskellige levende, men også uddøde organismer, er beslægtet. De seneste år har bioinformatik også vist sig anvendelig til at identificere DNA af ukendt oprindelse. Dermed kan man få større kendskab til jordklodens *biodiversitet*, det vil sige dens mangfoldighed af arter. I takt med at det hele tiden bliver både hurtigere og billigere at sekventere DNA, vokser informationerne i databaserne, og muligheden for at sammenligne, identificere og systematisere vores viden om organismer og livsformer – som enten har været eller er – stiger derved.

## Klassifikation og systematik

I dag tager man det som en selvfølge at DNA danner grundlaget for vores forståelse af livets udvikling og slægtskabet mellem levende organismer. Sådan har det ikke altid været. Først efter 1859 hvor Darwins evolutionsteori blev offentliggjort, begyndte man at anerkende at de levende organismer er beslægtede via en fælles stamform.

Rige – Planter



Række – Blomsterplanter



Klasse – Tokimbladede



Orden – Rosales



Familie – Rosaceae



Slægt – Rosa



Art



Figur 183. Linnés klassifikationssystem.

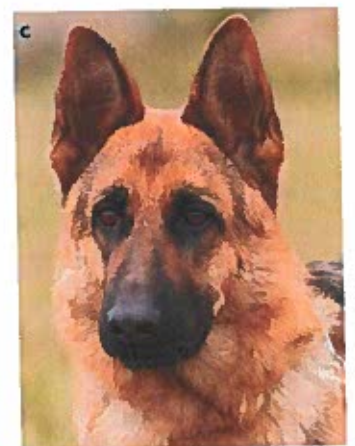
Man har dog siden oldtiden forsøgt at beskrive og klassificere levende organismer, og i 1700-tallet opfandt den svenske botaniker Carl von Linné et klassifikationssystem som man stadig anvender. Det er et hierarkisk system hvor Linné grupperede arter, der havde visse fællestræk i *slægter*. Slægter med fællestræk blev grupperet i *familier*, familier blev samlet i *ordener*, ordener i *klasser*, klasser i *rækker*, og rækker blev samlet i *riger*, se figur 183. Systemet anvendes som sagt stadig, men det har fået flere niveauer i hierarkiet efterhånden som man finder nye organismer, som kan være svære at placere inden for de niveauer som Linné definerede.

#### Arter har latinsk dobbeltnavn

Sammen med dette klassifikationssystem udviklede Linné også et såkaldt binomialt system hvor alle levende organismer får tildelt et latinsk dobbeltnavn svarende til slægt og artsbeskrivelse. Fx hedder en ulv på latin *Canis lupus*. Dens slægtsnavn er *Canis*, og dens artsbeskrivelse er *lupus*. Den er i samme slægt som prærieulven der på latin hedder *Canis latrans*. Hunden, *Canis lupus familiaris*, har vist sig at være en underart af ulve idet den godt kan parre sig med ulve og få fertilt afkom. Underarter giver man altså tre navne. *Canis*-slægten er beslægtet med slægten *Vulpes* der fx indeholder den almindelige røde ræv *Vulpes vulpes* og ørkenræven *Vulpes zerda*. Begge slægter (*Canis* og *Vulpes*) tilhører familien *Canidae* (hundefamilien) der igen tilhører ordenen *Carnivora* (rovdirene) osv., se figur 184. Ifølge Linné kunne de levende organismer inddeles i to riger: dyreriget og planteriget – en inddeling man har haft siden oldtiden. Denne inddeling har man nuanceret sidenhen. På Linnés tid skete klassifikationen af de



Figur 184.  
a. Ulv.  
b. Prærieulv.  
c. Hund.  
d. Rød ræv.  
e. Ørkenræv.



levende organismer på grundlag af systematiske observationer af morfologiske (form) og anatomiske (opbygning) træk hos organismene. Denne form for klassifikation bruger man stadig, men den er i dag suppleret med DNA-analyser, og det giver indimellem anledning til at om-døbe og flytte arter til nye placeringer i systemet.

#### Fra riger til domæner

Ifølge Darwins evolutionsteori fra 1859, se side 74, er alle arter udviklet fra en fælles stamform. Så i den sidste halvdel af 1800-tallet begyndte man at samle alle de dengang kendte arter i et fælles stamtræ – stadig på grundlag af morfologiske og anatomiske træk. I den periode blev bl.a. kendskabet til mikroorganismer større, og man inddelte derfor de levende organismer i tre riger: dyr, planter og protister. Protister var et rige der indeholdt de encellede organismer og de mest primitive af de flercellede organismer. Dette stamtræ har man siden bygget og justeret videre på, og der har været forskellige inddelinger, se figur 185.

En skelsættende forståelse af de levende organismers slægtskab kom da man i 1937 foreslog en inddeling i to imperier – pro-

karyoter og eukaryoter. Denne inddeling bruges stadig til at beskrive grundlæggende forskelle i cellers opbygning. Alle organismer uden cellekerne blev ved denne inddeling udskilt fra protister og fik i 1956 deres eget rige – monera.

I slutningen af 1960'erne blev svampe udskilt fra planteriget. Her består forskellen i at svampe ikke har kloroplaster (grøn-korn), og at opbygningen af deres cellevæg også er anderledes end planter.

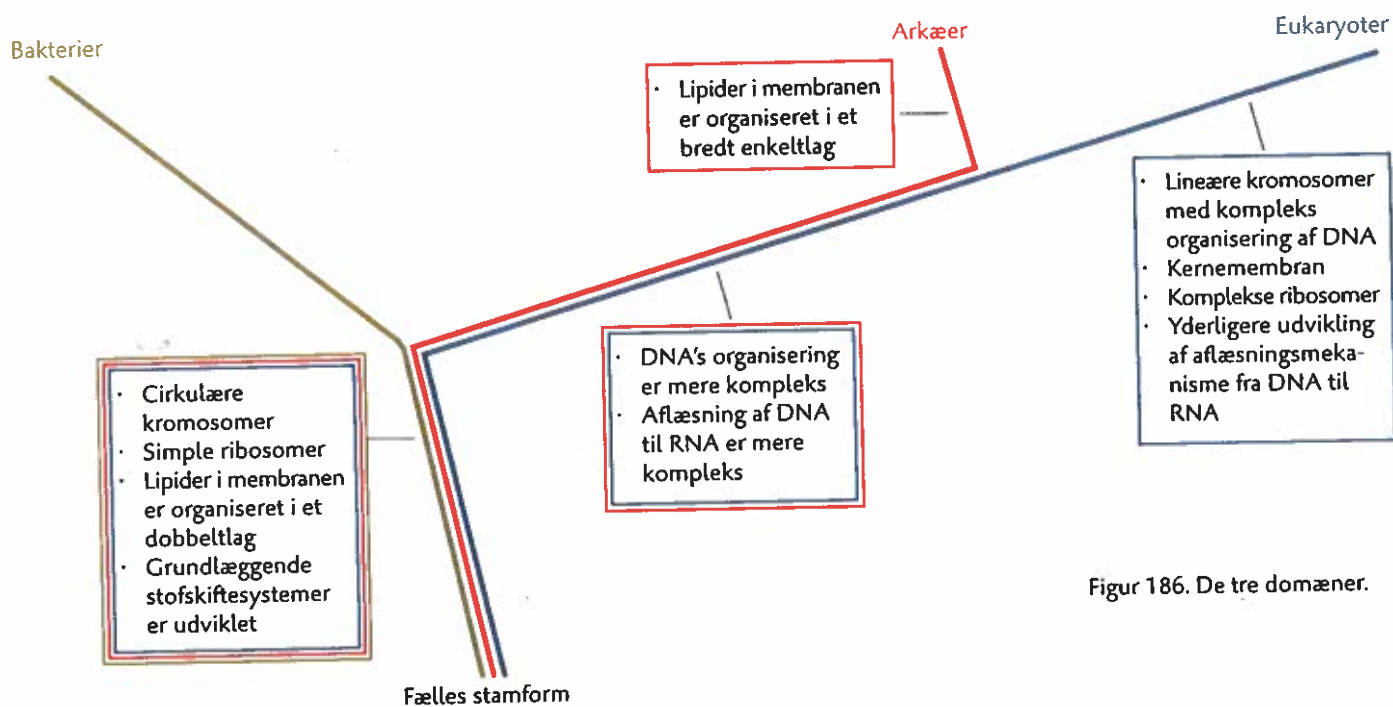
I 1977 blev monera (prokaryoterne) inddelt i to grupper, bakterier og arkæer. Forskellen på disse to grupper er den type lipider de har i cellemembranen og i opbygningen af deres cellevæg.

Efter 1990 er de fleste forskere derfor enige om at hovedinddele de levende organismer i tre domæner – bakterier, arkæer og eukaryoter, se figur 186.

Stamtræet viser at bakterier er de første organismer, der blev udskilt som en selvstændig udviklingslinje fra den fælles stamform. Det skyldes at arkæer også har en række fællestræk med organismer i det eukaryote domæne, som bakterier ikke har. Bl.a. er der visse ligheder i den måde DNA'et er organiseret omkring proteiner, og i den måde transskriptionen foregår på.

Figur 185. Videnskabelige inddelinger af de levende organismers slægtskab gennem tiden.

Linné 1735 2 riger	Haeckel 1866 3 riger	Chatton 1937 2 imperier	Copeland 1956 4 riger	Whittaker 1969 5 riger	Woese et al. 1977 6 riger	Woese et al. 1990 3 domæner
(Ikke behandlet)	Protister	Prokaryoter	Monera	Monera	Bakterier Arkæer	Bakterier Arkæer
Planter	Planter	Eukaryoter	Protister	Protister Svampe	Protister Svampe	Eukaryoter
Dyr	Dyr		Planter	Planter	Planter	
			Dyr	Dyr	Dyr	



Figur 186. De tre domæner.

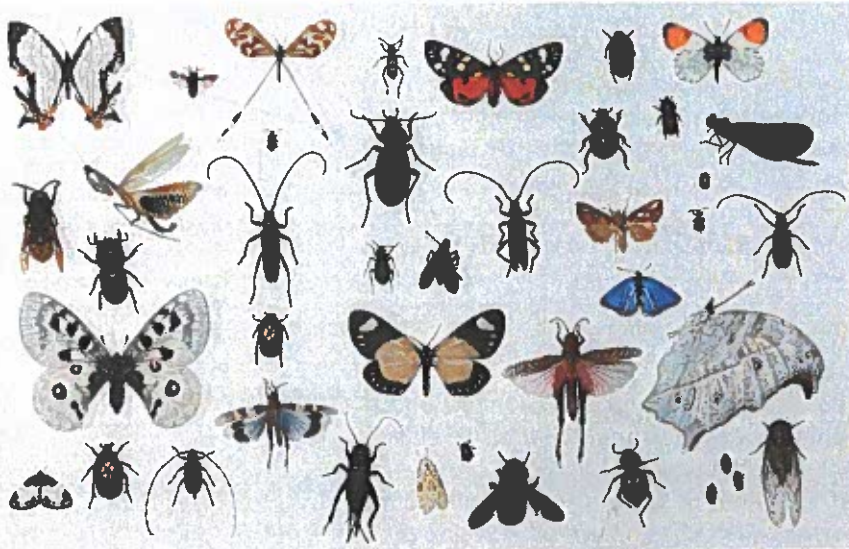
### Skal man tilbage til riger eller imperier?

Nogle forskere mener at udskillelsen af arkæer som et selvstændigt domæne er overdrevet sammenlignet med andre store og grundlæggende forskelle som man fx ser mellem eukaryote organismer. Undersøgelser af arkæers genomer viser typisk at de har ca. 43 % af deres genom fælles med bakterier og kun ca. 13 % af deres genom fælles med eukaryoter. Resten er fælles for alle levende organismer. Ser man på eukaryoterne er det en stor og meget blandet gruppe. Udviklingen af den eukaryote celle viser stor variation, bl.a. findes der forskellige typer af mitokondrier, og de mere primitive eukaryote organismer er lidt af en rodebutik hvad angår klassifikation. Hvis man sammenligner mere komplekse eukaryoters DNA med hinanden, fx en plantes DNA med et dyrs DNA, vil man også se store forskelle. Kritikerne af

domæne-inddelingen kan derfor ikke se hvorfor man vil opdele prokaryoterne i to selvstændige grupper, men ikke eukaryoterne. Diskussionen viser at klassifikation er et menneskeskabt system, som der helt sikkert fortsat vil være forskellige holdninger til.

### Hvor mange arter?

Der er klassificeret ca. 6.500 nulevende arter af bakterier, knap 300 arter af arkæer og ca. 1,3 millioner arter af eukaryoter. Kun en mindre del af disse arter er klassificeret ud fra kendskab til deres genom, men andelen vokser hver dag. Af de klassificerede eukaryote arter er ca. 1 million dyrearter, og ca. 80 % af disse er insekter, se figur 187 næste side. Man anslår at der findes over 10 millioner arter, så flertallet er endnu slet ikke klassificeret. Man anslår at 99,9 % af de arter der



Figur 187. Insekter er en stor og meget forskelligartet gruppe af organismer.

har levet, er uddøde, så når man laver et stamtræ i dag, repræsenterer det kun et øjebliksbillede set med evolutionsbriller.

### Konstruktion af et stamtræ

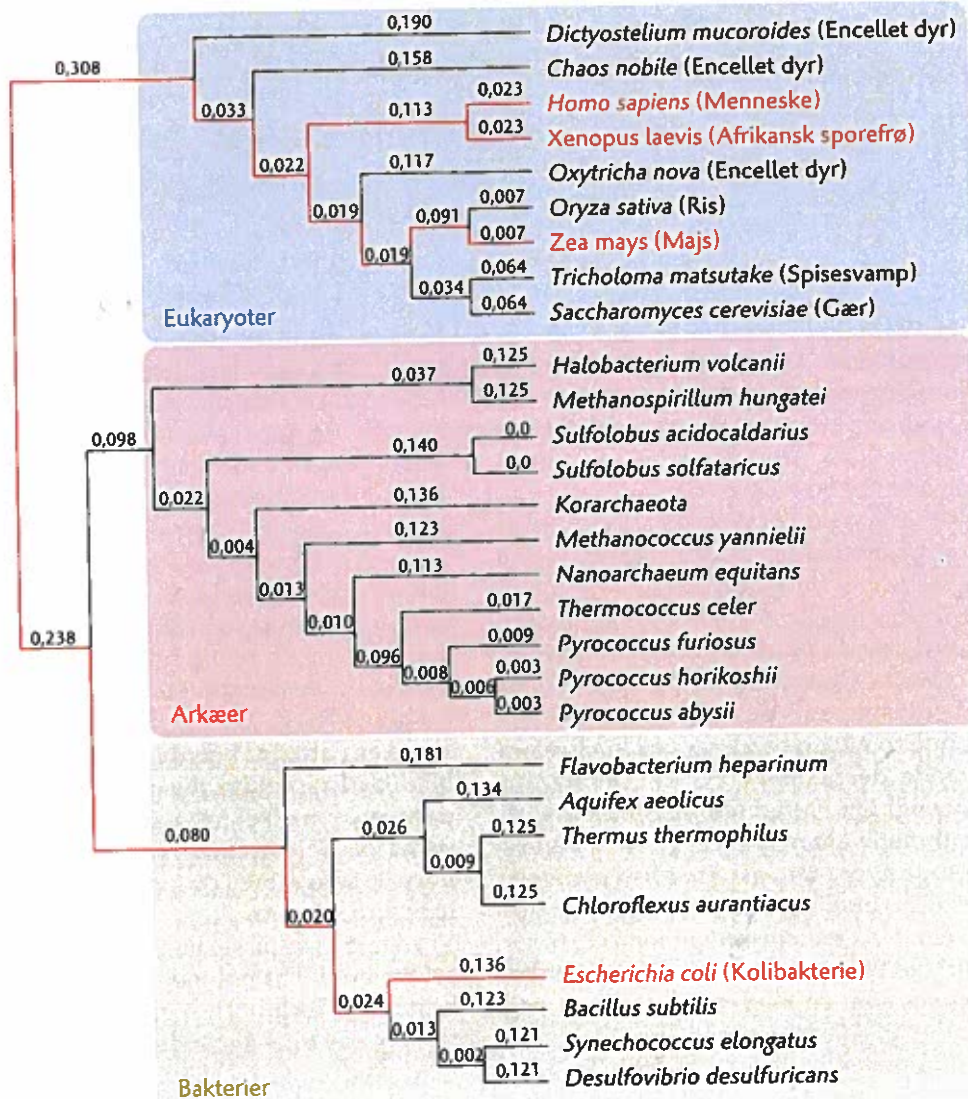
Alle organismer er beslægtede på et eller andet niveau i livets stamtræ, men man skal kende deres indbyrdes slægtskab for at forbinde dem korrekt i træet. Når man fx inden for den klassiske genetik bruger stamtræer til at vise nedarvning af en genetisk sygdom, er slægtskabsforholdene netop kendt. Men når de ikke er kendt, kan man i første omgang bestemme morfologiske og anatomiske forskelle og ligheder og dernæst bestemme baserækkefølgen af større eller mindre stykker RNA eller DNA – evt. hele genomer. Man kan også vælge at bestemme større eller mindre dele af aminosyrerækkefølgen af et polypeptid. Derefter kan man sammenligne resultaterne ved hjælp af bioinformatik, og derved bliver det muligt at konstruere et *fylogenetisk stamtræ*. At stamtræet er fylogenetisk betyder, at det er baseret på slægtskab.

### Genetiske forskelle

Det fylogenetiske stamtræ bygger på den forudsætning at der sker mutationer i individers arvemateriale. Hvis mutationerne sker i kønscellerne, kan de sendes videre til næste generation, og på den måde opstår genetisk variation inden for en art. Mutationerne kan forsvinde igen eller de kan opføres gennem tiden – begge dele på grund af selektion – og de kan derved udvikle en art i en bestemt retning. Arter kan forsvinde eller opsplittes i nye arter, og på den måde sker der over tid evolution.

Forskelle i baserækkefølgen mellem arvematerialet af repræsentanter fra forskellige arter kan derfor bruges som et mål for deres slægtskab og som en måleenhed når man skal konstruere stamtræer. Det er dog ikke helt ligegyldigt hvilke genetiske forskelle man bruger til at konstruere træet. Forskelle i ikke-kodende dele af genomet er ofte uden betydning i forhold til evolutionen, mens mutationer der er fremkommet i de kodende dele, kan have større konsekvenser og over tid føre til udvikling af nye arter. Mutationer i gener som er helt afgørende for individernes overlevelse, bliver oftest selekteret bort. Fx er baserækkefølgen i gener der koder for kromosomernes histon-proteiner, meget velbevaret. Det tolker man som at kromosomernes organisering er helt afgørende for cellernes mulighed for at lave korrekt replikation og celledeling.

Man må derfor vælge hvilke basesekvenser man sammenligner, ud fra hvor tæt beslægtet man forventer organismene er. Hvis man fx vil sammenligne populationer inden for den samme art, kan det være en fordel at se på ændringer i ikke-kodende dele af genomet, da der her kan opstå stor genetisk variation inden for relativt korte tidsrum. Hvis man derimod skal konstruere livets stamtræ, som repræsenterer



Figur 188. Stamtræ der viser slægtskab mellem levende organismer fra de tre domæner eukaryoter, arkæer og bakterier på grundlag af forskelle i rRNA-gener.

slægtskabet mellem arter i de tre domæner, altså fjernt beslægtede arter, må man tage udgangspunkt i meget velbevarede base-sekvenser som – når de varierer – sandsynligvis har haft afgørende betydning for evolutionen. Ændringerne er derved repræsentative for mange arter, og man kan derfor tegne de store linjer i stamtræet.

#### Stamtræer og genetiske afstande

Figur 188 viser et stamtræ konstrueret på grundlag af gener for rRNA der anses for at være meget velbevarede.

Som man kan se er stamtræet konstrueret således, at en gren i træet altid splittes i to. Forgreningspunktet repræsenterer en fælles forfar for de to arter eller grupper