

Livet skal sættes på strejkode

Et internationalt forskningsprojekt skal forsøge at få styr på alverdens levende organismer ved hjælp af små "DNA-strejkoder".

CTAGTGGTGTGACATGGCCATTTACACCGA

Af Jakob Damgaard, Gitte Petersen & Ole Seberg

■ Forstil dig, at du står midt i en tropisk regnskov. Omkring dig er et mylder af dyr og planter i alle størrelser. Du genkender kun ganske få af dem, men nysgerrig som du er, vil du naturligvis gerne vide, hvad resten er for nogle skabninger. Du trækker en lille transportabel maskine, en *barcode*, op af lommen, tænder den og etablerer satellitforbindelse til en stor central computer. Maskinen har en lille åbning ovenover en dataskærm. Du tager en plante, der har fanget din interesse, renser den for myrer, putter den ned i åbningen og venter i 10-15 minutter. Derpå lyser dataskærmen op, og en tekst viser sig, *Theobroma cacao* (kakao). Du klikker på navnet, og et nyt skærmbillede viser plantens placering i planteriget. Yderligere et par klik giver dig plantens totale udbredelse på jorden, dens anvendelse, dens kemiske indholdsstoffer osv.

Livet på strejkode

Nej, dette er ikke et manuskript til en hidtil ukendt episode af Star Trek. Det er derimod visionen for et seriøst internationalt forskningsprojekt, der med dannelsen af et uformelt konsortium, (CBOL), for nylig er blevet søsat. Helt grundlæggende er målet med CBOL at

sætte ikke-eksperter i stand til at artsbestemme dyr og planter på grundlag af korte, velkarakteriserede dele af disses DNA. Man vil simpelthen forsøge at karakterisere alle de hidtil kendte arter ved at sammenkoble hver enkelt art med en kort, unik DNA-sekvens, og tillige vil man sigte mod, at der fremstilles tilsvarende sekvenser for nybeskrevne arter. Samtlige af Jordens arter skal helt enkelt udstyres med en strejkode, der – i princippet – vil gøre det muligt, med stor præcision at henføre et hvilket som helst individ, eller en del deraf, til den korrekte art, og herved gøre al den information vi i tidernes løb har samlet om denne art tilgængelig.

Strejkode med fire farver

Sammenkædningen mellem strejkoder og arter stammer fra ligheden mellem de strejkoder, som vi bl.a. kender fra supermarkedet, og de farvede streger på det såkaldte kromatogram, som danner grundlaget for aflæsningen af en DNA-sekvens på en moderne sekvenator.

Mens de industrielle strejkoder adskiller sig ved antallet og tykkelsen af de enkelte streger, er "livets strejkoder" opbygget af fire farver, svarende til de fire baser (G, A, T, og C), der ind-



Foto: D. Janzen © University of Pennsylvania

Det er et af Consortium for the Barcode of Life's erklærede mål, at fremstille en transportabel barcode – en strejkodelæser – til feltbrug.

går i DNA-molekylet. De industrielle strejkoder er designet således, at de er unikke for det enkelte produkt, og kan aflæses elektronisk. Det er forventningen, at arternes strejkoder ligeledes vil kunne gøres unikke for hver enkelt art.

De industrielle strejkoder anvender 10 cifre på 10 positio-

ner (plus 2 cifre til at markere landet og 1 kontrolciffer, i alt 13 cifre) for at danne 10 milliarder koder, hvorimod DNA-strejkoder kun betjener sig af 4 forskellige farver. I teorien behøver man derfor kun 15 positioner til at danne 4^{15} (ca. 1 milliard) biologiske strejkoder – et tal der langt overstiger selv

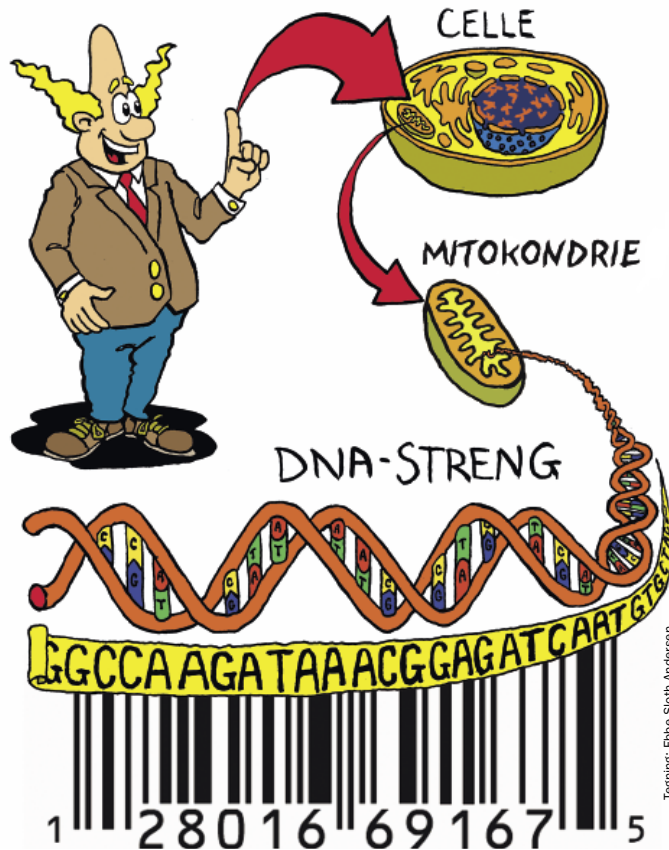
de mest fantasifulde forsøg på at anslå det totale antal nulevende arter på Jorden, der vurderes til at ligge i størrelsesordenen 10-15 millioner.

Udfordringen

De levende organismer grupperes gerne i en række forskellige riger. Dyre- og planteriget er kendte af enhver, men også svampe og protister (hovedsageligt encellede organismer) udgør ligesom forskellige grupper af bakterier deres egne riger. Fælles for alle organismer er, at de har en fælles oprindelse, og selvom der er stor forskel på f.eks. planter, bakterier og mennesker, så afspejles denne fælles begyndelse i arvematerialet. Sammensætningen af det genetiske materiale ændres med tiden ved mutationer, der f.eks. udskifter en base med en anden, eller indsætter eller fjerner en eller flere baser.

Nogle gensekvenser er næsten identiske på tværs af arter, f.eks. fordi de styrer helt basale processer i alle organismers liv. Man siger, at disse gensekvenser er konserverede. Sådanne sekvenser er ikke velegnede til at lave stregkoder ud fra, da selv mange tusinde base-positioner ikke indeholder tilstrækkelig variation til blot at beskrive en beskedne brøkdel af Jordens nulevende arter. Omvendt er andre gensekvenser meget variable (f.eks. fordi de ikke er koder for noget produkt), og derfor svære at arbejde med, idet de samme ændringer kan være opstået flere gange, og fordi stumper af korte sekvenser kan indsættes eller fjernes, uden at det betyder noget for gens funktion.

Når man skal lave DNA-stregkoder gælder det derfor om at finde en eller flere DNA-sekvenser, der lige præcis indeholder tilstrækkelig variation til at identificere arterne, men uden at disse sekvenser bliver unødigt lange. Sekvenser på omkring 5-700 baser vil være optimale, idet de dels indeholder en tilstrækkelig mængde information, og dels vil være overkommelige at fremstille.



Tegning: Ebbe Sloth Andersen

Når organismer fra dyreriget skal sættes på "DNA-stregkoder", er det i første omgang gener i cellernes såkaldte mitokondrier, forskerne har kig på. Det er dog håbet, at også DNA fra kromosomerne kan bruges til stregkoder.

Hvordan laves en stregkode?

I princippet er det ret let at fremstille stregkoderne. Man oprenser DNA fra den eller de organismer, man ønsker at fremstille en stregkode for, og formerer det ønskede gen

(eller sekvens) ved hjælp af *Polymerase Chain Reaction* (PCR), som er en metode, der kopierer DNA ud fra processer, der ikke er helt ulig dem som foregår i en levende celle, og endelig visualiseres sekvensen på en computerskærm via en automa-

tisk sekvenator. De således fremstillede stregkoder kan derefter sammenlignes med tidligere fremstillede stregkoder, og man kan afgøre, om den pågældende stregkode har været fremstillet før, og – ideelt set – hvilken art den tilhører. Det er indlysende,

Stregkoder og taxonomi

En vigtig forudsætning for, at systemet med genetiske stregkoder virker er, at der findes stregkoder for allerede kendte arter, som er identificeret af erfarne taxonomer. De fleste taxonomer er placeret på de naturhistoriske museer eller ansat ved genbanker osv., ligesom museerne ligger inde med de systematiske samlinger i form af herbarieark, nålede insekter, skeletter, skind m.m. Museerne vil derved få stigende betydning for at sikre, at der er en forbindelse mellem stregkoden og den organisme og det individ, der refereres til. Tillige vil det være nødvendigt, at DNA-prøverne opbevares på et sikkert sted, og igen vil det være naturligt at gøre det i museerne,

der har århundreders ekspertise på dette område.

Det har været foreslået, at man skulle fremstille stregkoderne ud fra de typeeksemplarer, som arterne er beskrevet ud fra, men da dette indebærer hel eller delvis ødelæggelse af dette individ, er det naturligvis ikke ønskværdigt endsige nødvendigt. Det vil naturligvis være ønskeligt, hvis beskrivelsen af nye arter blev ledsaget af en stregkode, og intet i de regelsæt, der styrer navngivning af dyr og planter, forhindrer dette. DNA-data kan – og bør måske i fremtiden – indgå i en ny beskrivelse på linie med andre former for information, f.eks. fra morfologi, vævstype, levested osv.

Selvom det i princippet er

relativt enkelt i dag at fremstille en brugbar DNA-sekvens er det ikke nogen triviel opgave at få en genetisk stregkode for alle arter. Mange beskrevne arter er kun kendt i ét (eller meget få) eksemplar(er) i samlingerne, og ofte er de i en tilstand eller så gamle, at DNA ikke, eller kun med meget stort besvær, kan ekstraheres fra materialet. Desuden er der mange grupper af organismer, hvor en taxonomisk ekspertise desværre slet ikke findes, hverken på lokalt eller globalt plan. Med det kolossale antal arter, der enten allerede er beskrevet eller venter på at blive det, er der derfor al mulig grund til at være stærkt bekymret for det alarmerende fald i taxonomisk ekspertise.

CBOL

Consortium for the Barcode of Life (CBOL) blev dannet på baggrund af et møde på National Museum of Natural History i Washington, D.C., og organisatorisk er projektet blevet støttet af Alfred P. Sloan Foundation i 2½ år. Konsortiet er meget bredt opbygget og består primært af naturhistoriske museer, botaniske haver, bioteknologisk industri og internationale organisationer (f.eks. frøbanker), der opbevarer og arbejder med biologiske materialer.

Herhjemme er det nydannede Statens Naturhistoriske Museum, som er en fusion af Geologisk, Botanisk og Zoologisk Museum samt Botanisk Have ved Københavns Universitet medlem af konsortiet. Også offentlige institutioner, der har behov for at identificere biologisk materiale som f.eks. ulovligt indførte dyre- og plantearter eller forarbejdet biologisk materiale, er tilknyttet.

at for overhovedet at få systemet til at virke meningsfyldt som generelt identifikationsværktøj, skal der være en større database af strekkodesekvenser til rådighed.

Som ovenfor beskrevet er det overordnede mål at fremstille et håndholdt apparat, en slags biologisk strekkodelæser (en *barcode*). Et apparat, der vil muliggøre, at man putter sin organisme ind i apparatet, venter et stykke tid og derpå får et bud på hvilken art det er, samt mulighed for at koble sig op til yderligere oplysninger om arten f.eks. gennem det verdensomspændende netværk af biologidatabaser Global Biodiversity Information Facility (GBIF).

Der kan forudses mange vanskeligheder med at forene og automatisere de mange processer i en sekvensanalyse, men det største problem i denne sammenhæng er næppe at få sekventoren gjort tilstrækkelig lille, men derimod at få standardiseret ekstraktionen af DNA såle-

des, at den rutinemæssigt virker på det helt kolossale spektrum af forskellige organismer.

Arter og strekkoder

Uden at åbne en diskussion om, hvordan arter defineres, så er det klart, at strekkoden ikke i sig selv definerer en art. Hvorvidt en eller flere strekkoder svarer til vores opfattelse af en given art, er noget man bliver nødt til at undersøge på traditionel vis, og spørgsmålet kan med lige så stor ret vendes om: Hvilke(n) strekkode(r) kan bruges til at definere en given art?

Strekkoderne erstatter altså ikke traditionel artsbestemmelse (taxonomi), men er et supplement hertil. Strekkoden er – hvis den svarer til en anerkendt art – blot en alternativ måde at bestemme arten på. Strekkoder vil altså åbne op for muligheden for hurtigt at kunne foretage en sikker artsbestemmelse af organismegrupper, som det normalt kræver eksperter at bestemme.

Strekkodens fordele

På grund af sikkerheden i bestemmelse af arterne har strekkoderne en række indlysende fordele for vores forståelse af biodiversiteten omkring os:

1. De gør det meget nemmere at sammenkoble forskellige livsstadier hos dyre- og plantearter. Det kan f.eks. være æg, larve, puppe og *imago* ("voksen") hos insekter med fuldstændig forvandling (f.eks. biller og sommerfugle).
2. De kan bidrage til at adskille arter, der er svære at skelne på ydre træk alene, f.eks. *twillingearter*, mikroorganismer, parasitter, samt fragmenter af arter.
3. De vil kunne bruges til at give et fingerpeg om, hvorvidt en given prøve kan repræsentere en ny art, fordi strekkoden er ukendt (eller alternativt, at variationen i strekkoden for en kendt art skal revideres).
4. Sekvensering af ét enkelt gen

for en meget stor mængde arter vil kunne yde bidrag til vores viden om arternes slægtskab og indbyrdes variation.

Uden at gå i detaljer med den ofte kontroversielle diskussion om, hvordan man fremstiller et stamtræ, så kan de sekvenser, der danner baggrund for strekkoderne, også bruges til at give et groft skøn over, hvor en given art er placeret i en slægtskabsmæssig sammenhæng.

Fra fødevarer til retsmedicin

Biologiske strekkoder er også interessante i en række praktiske sammenhænge. På grund af PCR-reaktionens helt utrolige følsomhed, kan teknikken bl.a. bruges til at identificere selv ganske små fragmenter af organisk materiale. Det vil f.eks. gøre det muligt at opspore uønskede komponenter i vores fødevarer. Det vil også gøre det muligt at identificere ulovligt indførte dyre- og plantearter, samt produkter fremstillet af disse arter. Endelig vil teknikken gøre det muligt hurtigt og effektivt at identificere parasitter hos os selv og vores husdyr. Et af de steder, hvor man formodentlig først vil se strekkodning benyttet, vil være til hurtig diagnosticering af potentielt epidemiske sygdomme som SARS, fugleinfluenza, AIDS, eller sygdomme benyttet som biologiske våben.

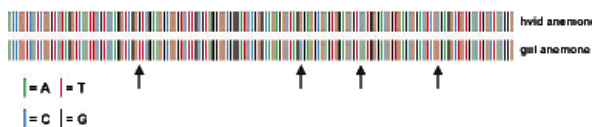
Retsmedicin er et af de forskningsfelter, der har været pioner i brugen af strekkoder. Det er velkendt, at man kan afsløre forbrydere ud fra efterladte biologiske spor på åstedet såsom blod, sæd og hår, og brugen af DNA-teknik har betydet, at en hel del alvorlige kriminelle sager er blevet genåbnet, ikke sjældent med en klar frifindelse af den mistænkte gerningsmand til følge.

Hvis man finder en død person, er noget af det første, man interesserer sig for, at finde ud af, hvem afdøde er, samt at fastlægge årsagen til og tidspunktet for dødens indtræden. Fastlæggelse af dødstidspunktet er forholdsvis enkelt, hvis der er tale om en nylig afdød, idet der

Mulige gener til strekkoder

Hos organismer med kønnet forering findes de fleste gener i to kopier i hver celle, nemlig en kopi fra hver af forældrene. En væsentlig undtagelse herfra er de gener, der indgår i de såkaldte organeller i cellen. Det drejer sig dels om mitokondrier, som har stor betydning for cellens stofskifte hos dyr, planter og svampe, og dels om grønkorn, der er ansvarlige for fotosyntesen hos grønne planter og alger. Disse organeller nedarves kun fra den ene af forældrene, typisk moderen, og den genetiske information nedarves derfor – når man ser bort fra eventuelle mutationer – uændret fra generation til generation. Hver eneste celle indeholder mange organeller, og hver eneste mitokondrie og grønkorn indeholder mange kopier af deres eget genom. Derfor er de et oplagt valg til strekkodebrug – sekvenser, der optræder i mange kopier, er simpelthen nemmest at arbejde med.

I dyreriget har interessen samlet sig omkring et DNA-område på ca. 650 basepar, der findes i mitokondriet, og som koder for et protein af betydning for cellens stofskifte. Sekvensen af dette gen er på en gang tilstrækkelig kon-



Biologiske strekkoder for to nært beslægtede anemone-arter. Pilerne peger på baseforskelle mellem strekkoderne.

serveret til, at man ved hjælp af PCR kan opformere det inden for et bredt spektrum af arter, men samtidig så tilpas variabelt, at det er bredt anvendeligt til fremstilling af strekkoder for såvel nært som fjernt beslægtede organismer.

Hos svampe og protister er variationen i mitokondrie-generne derimod for stor, mens den hos planter er for lille til, at man kan anvende dem som universelle strekkoder. Det er derfor nødvendigt at bruge forskellige typer af sekvenser hos de forskellige grupper af organismer.

Et problem med DNA fra mitokondrier og grønkorn er, at dette ikke nødvendigvis følger artens historie, idet det jo kun nedarves fra den ene forælder.

Der er derfor stor interesse for at finde gener i cellekernen, der har de samme fordele som gener fra organellerne, men som bedre

afspejler artsdannelse. Her samler interessen sig om de gener, der koder for ribosomerne, som deltager i cellens proteinsyntese, og som i modsætning til mange andre gener i kernen findes i et meget stort tal. Disse gener består dels af områder, som er fælles mellem ret forskellige organismegrupper, og dels af meget variable områder, der kan bruges til at adskille organismer på et lavere niveau, f.eks. arter.

Ud over de nævnte gener indeholder cellekernen enorme mængder af kodende og ikke-kodende sekvenser, der er potentielt anvendelige til strekkodning. Desværre er mange af sekvenserne yderst vanskelige at arbejde med, idet de kun findes i få kopier og ofte i forskellige versioner (genfamilier) i samme celle. Der foreligger således stadig et stort arbejde med at udvikle og teste nye genetiske strekkoder.



Foto: C. Clark ©

Biologiske samlinger, som her fuglesamlingen på The National Museum of Natural History i Washington vil, med dens 600.000 indsamlinger, få en forøget betydning, når de biologiske stegkoder skal fremstilles. Ikke kun fordi den indeholder så mange forskellige arter, men også som et opbevaringssted for DNA-prøver.

inden for det første døgn udvikles tydelige karakteristika, såsom afkøling, dødsstivhed osv. Hvis personen imidlertid har været død i længere tid, bliver bestemmelsen af dødstidspunktet mere usikker. Her kan det være nødvendigt at inddrage de larver af spyfluer, der ofte kan findes på liget, især hvis dette ligger ude i naturen. Ud fra forsøg med udlagte kadavere af større husdyr er det blevet demonstreret, at forskellige arter lægger æg på forskellige tidspunkter i forråd-

ningsprocessen. Efter at have sammenlignet larvernes størrelse med kendte standarder, samt udklækket og artsbestemt de voksne spyfluer, kan man regne baglæns således, at man får en nogenlunde nøjagtigt vurdering af, hvor længe personen har ligget død, selvom temperatur og vejrlig naturligvis giver en vis usikkerhed. Med kendskab til stregkoder for de relevante fluearter kan man imidlertid fremskynde processen betragteligt, idet man direkte kan arts-

bestemme larverne ved hjælp af deres DNA.

Som et kuriosum kan det nævnes, at lig kan blive flyttet omkring for at skjule sporene efter en forbrydelse, hvorfor man her eventuelt kan risikere kun at stå med nogle fluelarver som bevismateriale. Selv her er der dog håb om af opklare forbrydelsen, idet man ud fra larvernes maveindhold kan ekstrahere menneskeligt DNA, og derved har mulighed for at identificere afdøde. ■

Om forfatterne



*Jakob Damgaard er forskningsadjunkt
Tlf.: 35 32 21 51
E-mail: jdamgaard@bi.ku.dk*



*Gitte Petersen er lektor
Tlf.: 35 32 21 64
E-mail: gittep@bot.ku.dk*



*Ole Seberg er lektor
Tlf.: 35 32 21 53
E-mail: oles@bot.ku.dk*

*Alle er ved
Afdeling for Evolutionsbiologi
Biologisk Institut
Københavns Universitet*

Yderligere information:

Hebert, P.D.N., Cywinska, A., Ball, S.L. & deWard, J.R. 2003. Biological identifications through DNA barcodes. Proc. R. Soc. Lond. B 270: 313-321.

Stoeckle, M. 2003. Taxonomy, DNA, and the Bar Code of Life. - BioScience 53: 2-3.

*Interesserede kan også besøge
www.barcodinglife.com*