

DANSK DNA-DETEKTIV HJÆLPER ENGELSK POLITI I DRABSSAGER

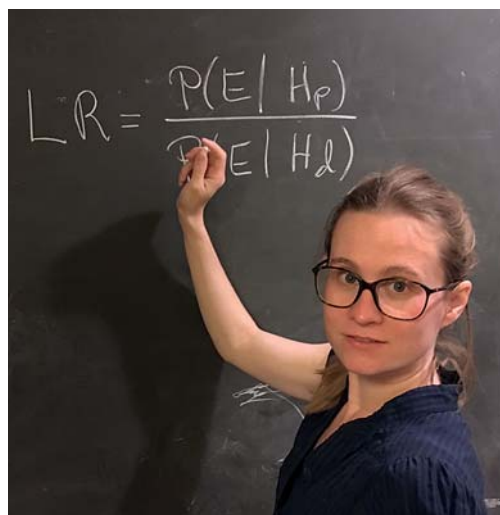
I en række drabssager har den danske forsker i statistik Therese Graversen fra Københavns Universitet hjulpet engelsk politi med at sandsynliggøre en gerningsmand. Hun bruger en unik metode, som allerede har skrevet engelsk retshistorie og kan sætte nye standarder for retsvæsnen i fremtiden.

I 2017 henvendte South Yorkshire Police sig til den danske forsker i statistik Therese Graversen fra Institut for Matematiske fag, Københavns Universitet med et problem. En mand var blevet banket ihjel i sit hjem i et meget voldeligt overfald. De engelske politifolk havde allerede anklaget en person, og de havde fundet et DNA-spor på vedkommende efter anholdelsen. Men fordi politiet skønnede, at DNA-sporet indeholdt DNA fra både offeret og mindst tre andre personer, bad politiet den danske forsker om hjælp til at lave en beregning, som kunne hjælpe med at kaste klarhed over det blandede spor.

»Der var fundet noget sammenblandet DNA på den anklagedes joggingbukser. Man havde en formodning om, at noget af det tilhørte den afdøde. Derfra skulle jeg så beregne vægten af, om det nu var den afdødes DNA på den anklagedes joggingbukser,« forklarer Therese Graversen.

Engelsk laboratorium måtte opgive

Når Therese Graversen får en opgave fra det engelske politi, er der altid fundet DNA fra fire eller



flere personer på for eksempel ofret eller en genstand fra gerningsstedet. For når mange personers DNA er blandet sammen, eksploderer antallet af kombinationsmuligheder i et puslespil bygget på opskriften på den menneskelige organisme. Selv for store private laboratorier, som det engelske retssystem normalt udliciterer opgaven til, volder udregningerne problemer. Det var hovedårsagen til, at den danske forsker i 2016 fik sin første sag.

»De sad med en DNA-blanding, hvor de kunne se, at der måtte være mindst fem personer. Men problemet var, at det computerprogram, de havde, ikke kunne regne på fem mennesker, det var for komplekst – der bliver simpelthen for mange

muligheder rent statistisk,« forklarer Therese Graversen, som efter den første sag i 2016 blev kontaktet af flere engelske politikredse, som havde hørt om hendes metode.

Sagen endte med at blive den første sag i England nogensinde, hvor den type beregninger var blevet brugt og accepteret i retten, og en af pionererne inden for DNA-analyse i England og

Europa, Peter Gill, bad efterfølgende om at få Therese Graversens rapport, så han fremover kunne bruge den som eksempel i engelske fora, hvor DNA-analyse diskuteres.

Ny og bedre metode

Hemmeligheden bag hendes beregningsmetode er slet ikke så hemmelig. Faktisk ligger den frit tilgængelig på nettet. Ved hjælp af et stykke software, som hun selv har kodet, analyserer hun den DNA, som politiet har sikret fra gerningsstedet. DNA-sekvenserne ligner små bjergtoppe i forskellige højder og indeholder oplysninger, som hendes software bruger til at lave udregningen.

»Før jeg udviklede min metode, var det her et område, hvor jeg nemt

Forfatter Michael Skov Jensen er kommunikationsmedarbejder ved Det Bio- og Naturvidenskabelige Fakultet, Københavns Universitet
msj@science.ku.dk

Kontakt til forskeren Therese Graversen:
graversen@math.ku.dk

Billedet: Therese Graversen ved tavlen.
Foto: Rasmus Lerchedahl Petersen.

Tekniske udfordringer ved analyse af DNA-spor

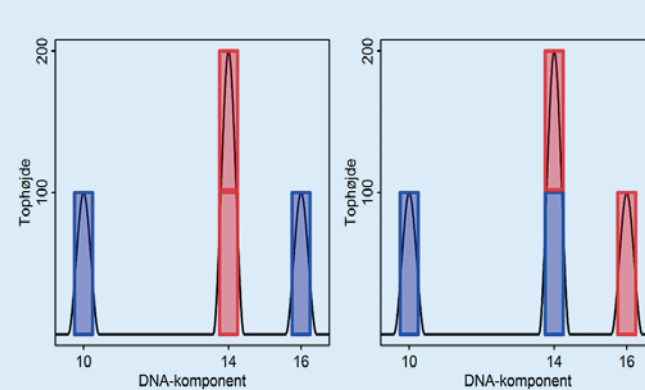
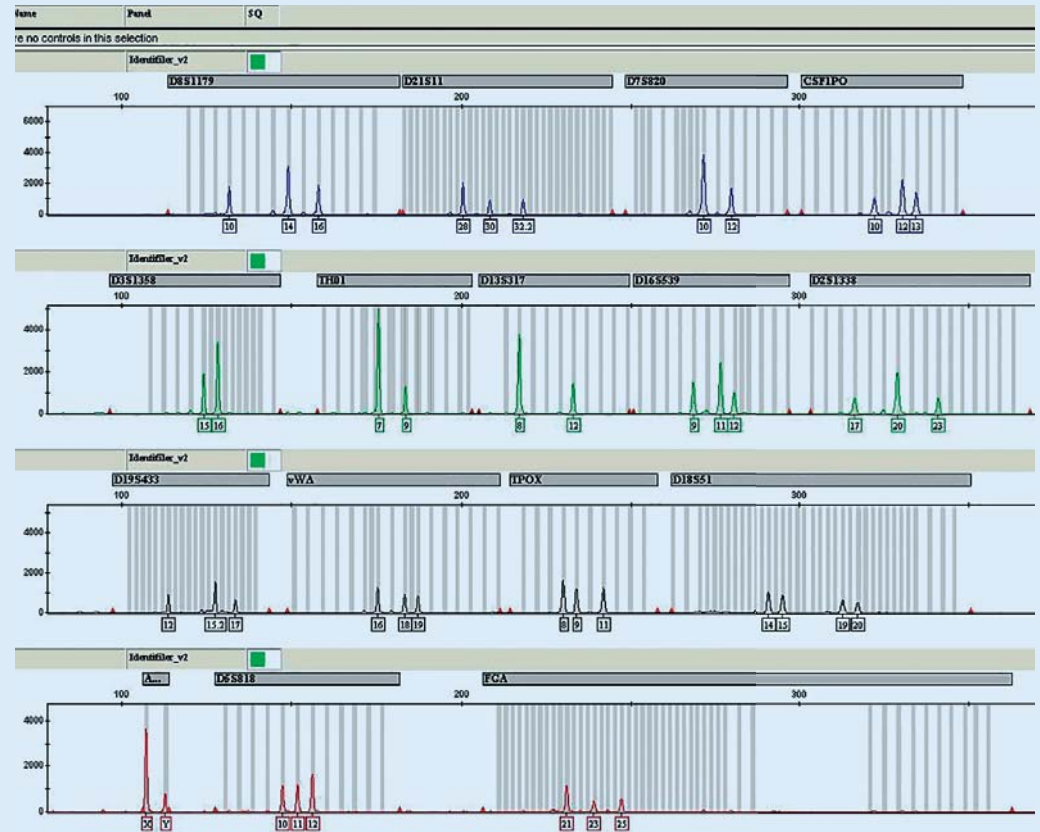
I DNA-profilering baseret på såkaldte STR-markører (STR står for Short Tandem Repeat) udnytter man, at nogle områder i vores DNA har en særlig opbygning, hvor sekvenser på typisk 4 basepar gentages et antal gange. Antallet af gentagelser i disse områder varierer fra person til person og er arvelige egenskaber. En markør angiver et område i hvert af personens to kromosomer i et af de 22 autosomale kromosompar, og målingen for markøren består i at registrere antallet af gentagelser i de to områder. De to områder for markøren kan have samme eller forskellige antal gentagelser. Selve DNA-profilen kommer således til at bestå af en eller to komponenter (alleller) for hver af de 16 markører, man normalt undersøger. Man inkluderer desuden typisk en 17. markør, som angiver personens køn.

Ved simple DNA-spor – dvs. DNA-spor af god kvalitet, som stammer fra en enkelt person – kan man aflæse personens DNA-profil direkte fra det såkaldte elektroferogram (EPG). Det er en graf bestående af små toppe, der indikerer hvilke DNA-komponenter personen har i sin profil; der vil være en eller to toppe for hver af de 17 markører, der udgør DNA-profilen.

Hvis DNA-sporet er en blanding af DNA fra flere personer, så er der naturligvis også flere forskellige DNA-komponenter i blandingen. Derfor vil man se hel flok toppe på EPG'et, og man kan ikke længere direkte aflæse den enkelte persons DNA-profil ud fra, hvilke komponenter der er i sporet. Hvis to personer eksempelvis har den samme (→)



Illustration af et område af vores DNA med såkaldte Short Tandem Repeats – i dette tilfælde gentages sekvensen GATA seks gange.



↑ Elektroferogram for et blandet DNA-spor. Kigger man på det øverste panel, er det inddelt i fire markører: D8S1179, D21S11, D7S820 og CSF1PO. De lodrette grå striber svarer omtrent til de mulige DNA-komponenter, en person kunne have for den markør. Der er registreret tre komponenter 10, 14, og 16 for markøren D8S1179, hvilket er tegn på at sporet indeholder mere end én persons DNA.

← To forskellige måder at fordele DNA-komponenterne for en markør på, hvis man antager, at det er en blanding af lige meget DNA fra hver af to personer.

kunne sidde og vente i dagevis, mens computeren regnede på for eksempel fire personers DNA,« fortæller Therese Graversen, der brugte ventetiden, hvor computeren gennemløb de astronomisk mange kombinationsmuligheder, som blandet DNA har, på at optimere koden

og det stykke software, hun bruger.

»Jeg begyndte at ændre hele udregningen mere fundamentalt og endte med en metode, der løber alle DNA-kombinationerne igennem på en smartere og mere selektiv måde,« siger Therese Graversen,

der med sin bærbare computer kan udregne mere komplekse sager på kortere tid, end de private laboratorier i England har mulighed for.

Kan skubbe til retspraksis

Forskerens metode er nyskabende, fordi den inddrager flere informa-

Arbejdsgang

1)



Et DNA-spor indsamles fra gerningsstedet. Sporet indeholder muligvis DNA fra flere personer. Foto: Shutterstock.

2)

Ved hjælp af PCR laves der milliarder af kopier af de to områder af DNA for hver markør. Kopierne får en af fire farver, alt efter hvilken af de 17 markører, de kommer fra. De enkelte komponenter optræder i omtrent samme forhold som i DNA-sporet. Det er normalt at kopieringen ikke er helt perfekt. Som regel bliver der dannet en lille mængde af nogle komponenter, som måske ikke allerede var i blandingen.

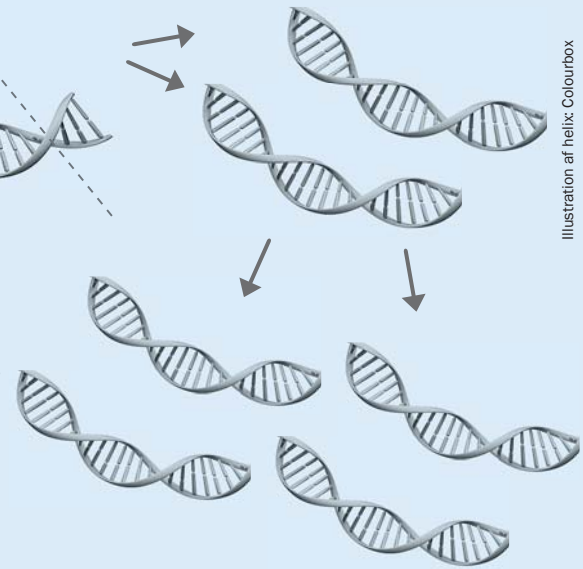
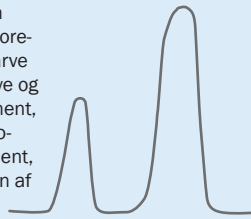


Illustration af helix: Colourbox

3)

De forskellige komponenter i blandingen registreres ved hjælp af kapillærelektroforese. Her sorteres komponenterne efter farve og længde. Ud fra kombinationen af farve og længde ved man præcis hvilken komponent, der er tale om. Der dannes et elektroferogram (EPG) med en top for hver komponent, hvor højden på toppen angiver mængden af komponenten.



5)

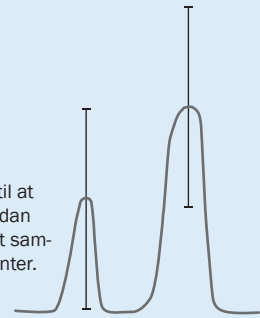
Ud fra sagens omstændigheder formuleres to scenarier, der skal holdes op mod hinanden. Disse kunne for eksempel være:

H_1 : DNA'et kommer fra anklagede og fire ukendte personer, der ikke er biologisk relaterede til den anklagede.

H_2 : DNA'et kommer fra fem ukendte personer, der ikke er biologisk relaterede til den anklagede.

4)

Statistisk modellering bruges til at beskrive variabiliteten af, hvordan EPG'et vil se ud for en bestemt sammensætning af DNA-komponenter.



6)

Der udfærdiges en vidneerklæring, som rapporterer resultatet af analysen i form af en likelihood-ratio:

$$LR = \frac{P(\text{EPG} | H_1)}{P(\text{EPG} | H_2)}$$

LR angiver hvor mange gange mere sandsynligt det ville være at elektroforesen gav netop det EPG, hvis H_1 var rigtig, end hvis H_2 var rigtig.

DNA-komponent i deres profiler, så vil man blot se én højere fælles top for den komponent, svarende til at de to personers toppe bliver "stabled" oven på hinanden.

Nøglen til de moderne metoder består i at udnytte den information, der ligger i toppenes højder. Højden af toppene angiver, hvor meget af de forskellige DNA-komponenter,

der er i sporet. Da alle toppe fra samme person cirka vil være lige høje (medmindre to er "stabled"), fortæller tophøjderne altså også noget om, hvilke DNA-komponenter der hører til hver enkelt person. Man kan forestille sig analysen af

EPG'et som et puslespil, hvor man for hver person må placere to små pinde af en bestemt højde indenfor toppene for en bestemt markør. Højden på personens pinde svarer til, hvor meget af personens DNA blandingen indeholder. (→)

Af Therese Graversen
graversen@math.ku.dk

Therese har en bevilling fra Danmarks Frie Forskningsfond | Teknologi og Produktion

tioner fra DNA-sekvenserne, som hidtil ikke er blevet udnyttet. Blandt andet bruger hun informationer om de såkaldte tophøjder. Højderne siger noget om, hvor stærkt signalet er og dermed også noget om, hvor meget DNA'et fylder i blandingen. Therese Graversens software er

også blevet anvendt i enkelte danske retssager, og hun håber på, at hendes metode vil fortsætte med at vinde indpas både i England og i Danmark i fremtiden.

»Det jeg har indset er, at små folk som mig, godt kan gå ind og skubbe

lidt til retspraksis. Min metode har gjort politiet både i Danmark og i England opmærksomme på, at man godt kan trække flere informationer ud af DNA'et end de plejer, og det genererer naturligt en efterspørgsel for de her mere sofistikerede metoder til at analysere DNA-spor,« siger hun. ■

DNA-spor er centrale i mange kriminalsager. Selv flere gamle sager bliver nu opklaret. Foto: Shutterstock.

Videre læsning

Principles of Forensic DNA for Officers of the Court. Webservice med små letforståelige træningsmoduler med alle de fundamentale principper. <https://projects.nfstc.org/otc>

Oversigtsartikel om DNA-metoder til retsbrug:

Steele, C.D. and Balding, D.J.: Statistical Evaluation of Forensic DNA Profile Evidence. Annual Review of Statistics and Its Application 2014 1:1, 361-384.

Standard håndbog:

Butler, J.M. (2005): Forensic DNA Typing: Biology, Technology, and Genetics of STR Markers (2nd Ed.). Elsevier Academic Press, New York.

Therese Graversens software, DNAmixtures: <http://dnamixtures.r-forge.r-project.org>

Therese Graversens ph.d.-afhandling fra Oxford 2014: Statistical and Computational Methodology for the Analysis of Forensic DNA Mixtures with Artefacts. <https://ora.ox.ac.uk/objects/ora:9362>



Opgaven går så ud på at udfylde alle toppene i EPG'et med pindene fra alle de forskellige personer.

Utallige kombinationer

I analysen af et DNA-spor bliver man nødt til at undersøge samtlige kombinationer af DNA-profiler og se, hvor godt hver enkelt passer med toppene i EPG'et. Ser vi på DNA-komponenterne for blot en enkelt markør, er der i størrelsesordenen 200 kombinationer for en person. Hvis man forestiller sig, at blandingen kommer

fra 2 personer, så er der 40.000 kombinationer at se på, fordi man for hver eneste af de 200 forskellige kombinationer, som den første person kunne have, også må se på de 200 mulige kombinationer, som den anden person kunne have. For en 5-personers blanding bliver det til 200^5 , altså 320 milliarder, kombinationer at løbe igennem. I praksis ved man desuden ikke altid, hvor mange personers DNA der er i blandingen, eller hvilket blandingsforhold det indgår i, og det gør naturligvis pusle-

spillet endnu mere komplekst.

En yderligere komplikation er, at hvis man genanalyserer det samme DNA-spor mange gange, vil man ikke få helt det samme EPG hver gang. Der er noget variabilitet i toppenes højde og derfor også i, hvilke toppe der ses på EPG'et. For eksempel bliver meget små toppe ikke altid vist på EPG'et. Man kan se det som måleusikkerhed i processen fra DNA-spor til EPG, og en statistisk model tager højde for den slags usikkerhed. ■

Bliv **STUDERENDE** **FOR EN DAG** på **SDU**

Tilmeld dig Anvendt matematik, Biokemi og molekylær biologi, Biologi, Biomedicin, Datalogi, Farmaci, Fysik, Kemi eller Matematik på sdu.dk/nat/studerendeforendag fra 1. okt. 2018