



UPPSALA  
UNIVERSITET

# Analys av ofullständiga DNA-profiler

Metoder, problem och lösningar

Minna Andersson

---

Independent Project in Biology

Självständigt arbete i biologi, 15 hp, vårterminen 2016

Institutionen för biologisk grundutbildning, Uppsala universitet

# Analys av ofullständiga DNA-profiler - Metoder, problem och lösningar

Minna Andersson

Självständigt arbete i biologi 2016

## Sammandrag

Alla människor DNA skiljer sig åt och unikt för varje individ, detta gör det möjligt att använda sig av DNA-analyser vid brottsutredningar. Många DNA-profiler från brott som begås är ofta ofullständiga på något sätt och detta kan leda till att fel person döms. Med en ofullständig profil avser jag vara DNA-prov som kan försvåra en analys till exempel DNA-prover som blivit nedbrutna eller prover som har fler än en bidragsgivare, en så kallad DNA-mix. Ett annat problem är att mängden DNA kan vara reducerad vilket gör att en analys kan vara svår att genomföra. Idag finns analyser för denna typ av prover, högkänslig DNA-analys klarar av att analysera mycket små mängder DNA och fungerar i stora drag på samma sätt som en vanlig DNA-analys. Från en högkänslig DNA-analys kan amplifieringen resultera i en del störningar som gör att tolkningen av analysen blir mer komplicerad. Det har gjorts flera studier där man försöker komma fram till nya metoder samt optimera de metoder som finns för att i möjligaste mån undvika dessa störningar. Resultatet från dessa studier är lovande. Man har bland annat visat att designen på primrarna som amplifierar det specifika området har betydelse. Rening innan amplifieringen som är nödvändig för att erhålla resultat har resulterat i både bättre och enklare tolkning av data som finns tillgänglig. Man har även vänt sig till andra typer av markörer än de vanliga korta tandemupprepningarna för att ge ett bättre resultat och möjliggöra användning av nästa generations sekvenseringsteknologi. Det pågår en intensiv forskning för att förbättra och utveckla högkänsliga DNA-analyser och stora framsteg görs, vilket förhoppningsvis inom en snar framtid skulle ge som resultat att färre personer blir oskyldigt dömda, om nästan ingen alls.

## Inledning

Brott begås dagligen och man jobbar alltid för att finna den skyldige. När ett brott begås lämnar nästan alltid vederbörande spår efter sig som innehåller DNA. Det kan vara allt från blod till sperma, hår och olika typer av vävnad. Av de spår som lämnas kan man med hjälp av DNA-profilering identifiera rätt person och döma den som är skyldig till brottet. Det händer dock att personer blir oskyldigt dömda och den skyldige till brottet går fri. Orsaken till att fel person döms kan vara på grund av misstag i DNA-analyser eller att analysen inte är tillräckligt noggrann.

En studie som gjordes år 2013 visade att av alla människor som blivit avrättade i USA var 4 % oskyldigt dömda (Gross *et al.* 2014). Forensiker som jobbar med att analysera DNA från exempelvis brottsplatser jobbar hela tiden för att förbättra dagens metoder och utveckla nya metoder. Det diskuteras idag om hur väl DNA-databaser är till ens fördel/nackdel då en del anser att det är integritetskränkande att samla på varje människas unika DNA. Detta är dock det enda sättet för dagens forensiker att kunna göra en DNA-analys och säkerhetsställa en identifiering.

## Syfte

Mot denna bakgrund och för att illustrera strategier för att minska antalet oskyldigt dömda, kommer detta arbete att fokusera på hur man kan identifiera den skyldige från en ofullständig DNA-profil. Med ofullständig DNA-profil avser jag DNA-prover med fler än en bidragsgivare (DNA-mix) och partiellt nedbrutet DNA som kan orsaka störningar i amlifieringen och på så sätt leda till misstag i tolkningen av en DNA-analys. Med ofullständig DNA-profil avses också problemet med att DNA kan förekomma i reducerad mängd vilket gör det svårt att överhuvudtaget utföra en analys. Texten kommer ta upp vilka störningar profilerna kan ge som kan leda till missbedömningar när den slutgiltiga DNA-profilen ska beskrivas och även summera vad man kan göra för att förbättra och utveckla nya metoder för att undvika dessa problem.

Det är inte bara störningar som kan orsaka en feltolkning utan även kontamination när prov tas från brottsplatsen samt kontamination under själva DNA-analysen men i denna uppsats tas det inte hänsyn till huruvida provet kan kontamineras utan fokus ligger på den molekylära analysprocessen.

## Bakgrund

### En individs korta tandemupprepningar i genomet

Majoriteten av alla celler som en människa består av är diploida och bär var och en på nästan identiskt DNA, förutom gameter och de röda blodcellerna som inte har någon cellkärna. Nästan hälften av en individs genomuppsättning består av upprepningar och en del av dessa är tandemupprepningar, dvs. korta sekvenser (ex. CAG) som sitter upprepade bredvid varandra (Duitama *et al.* 2014). De korta tandemupprepningarna (small tandem repeats, STR) eller som de även kallas; mikrosateliter, finns i varje cells icke-kodande DNA. Med icke-kodande DNA avses den region i genomet som inte skapar några proteiner. De är repetitiva regioner som är mellan 2-6 baspar långa och återfinns i princip över hela genomet. Varje STR-område som sitter på en allel skiljer sig mellan varje individ (figur 1) (El-Alfy & Abd El-Hafez 2012). På varje kromosom finns det två kopior av en allel vid ett givet lokus. En kopia av ett STR-område ärvs från varje förälder. Dessa två kopior av området behöver inte ha samma antal upprepningar. Om det finns samma antal repetitiva fragment inom STR-området på båda allelerna anses detta vara en homozygot men om antalet repetitiva fragment skiljer sig anses det vara en heterozygot (Reid *et al.* 2013).

Individ 1: CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG (12 upprepningar)

Individ 2: CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG (9 upprepningar)

Individ 3: CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG (11 upprepningar)

*Figur 1. Tre individer med olika antal upprepningar vid samma STR-område på kromatiden. Antalet upprepningar är unika för varje individ. Om antalet upprepningar inom STR-området på båda allelerna är av samma antal anses detta vara en homozygot, om antalet är olika klassas det som en heterozygot.*

STRs är väldigt polymorfa och unika för varje individ. På grund av dessa egenskaper används de ofta som markörer för kartlägningsstudier, sjukdomsdiagnostik och för identitetstestning inom forensiska analyser (Zhejiang University 2015). Införandet av STR-analysmetoden var en revolution inom revolutionen då man nu kunde lagra databaser med individens DNA för att

sedan använda sig av dessa data vid profileringen. Denna metod utvecklades och introducerades redan för 17 år sedan. Det var dock inte förrän år 2000 som metoden utvecklats till den grad att den fick bred spridning. Man menade att metoden skulle minska chansen för en felaktig match jämfört med andra metoder. Det finns databaser med data över individers STR-områden över hela världen, i USA som har den största databasen har den fått namnet CODIS (loci combined DNA index system) som jämför 13 specifika STR-områden i en analys (Gill *et al.* 2015).

### **Standard DNA-analys**

DNA-analyser använder sig normalt av amplifiering med Polymeras-kedjereaktionen (PCR). Denna metod beskrevs första gången 1985 av Kary Mullis och 1993 fick Mullis nobelpriset i kemi för detta revolutionerande verktyget inom molekylärbiologin (Butler 2005). PCR-metoden kan principiellt liknas vid kroppens egen celledelning. Denna reaktion som sker med hjälp av primers vilket är en sekvens av nukleinsyror som är komplementär till båda 3'-ändarna av sekvensen som ska amplifieras. En primer startar syntesen och gör att en specifik del av DNA-strängen kan kopieras och beroende på antal cykler som genomförs under en reaktion, kan mängden av det specifika DNA-fragmentet bli tillräckligt hög för att detekteras och karaktäriseras. Primrarna som amplifierar STR-områden är märkta med olika färgfluorescens-markörer, detta för att det ska kunna utöka antalet loci som kan analyseras i en enda PCR-reaktion och får att kunna detekteras (Butler 2005).

### **STR-analysen går till följande:**

#### *1. Kopiera de specifika STR-områdena*

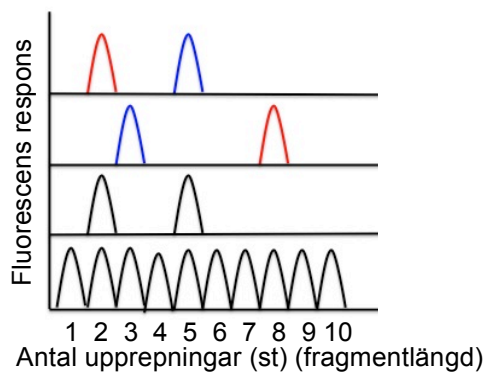
Med PCR amplifieras de STR-områden man vill analysera. Nästan alla STR-områden kan amplifieras, det viktiga är att det DNA som flankerar STR-området är intakt eftersom det är där de två primrarna binder. Även fast vissa STR-områden endast består av ett fåtal upprepningar, kan detta område, om DNA-området runt om är intakt ändå amplifieras (Butler 2005).

#### *2. Separering*

Efter att DNA-området har amplifierats med PCR separeras de amplifierade fragmenten med hjälp av kapillärelektrofores. Kortare fragment rör sig snabbare än längre fragment och detekteras därför först. De data man får fram från elektroforesen kan översättas till alleler i ett elektroferogram (EPG) (Taylor *et al.* 2016).

#### *3. Elektroferogram*

I EPGet får man olika toppar vars balans visar på hur bra en analys har genomförts. Man jämför därefter EPG:et med person av intresse (person of interest, POI) och som resultat får man antingen en match eller en icke match (figur 2). En match följs därefter oftast upp med andra analysmetoder för att stärka identifieringen (Kelly *et al.* 2014).



Figur 2. Förenkling över hur en match upptäcks. Två individer med olika antal upprepningar i två STR-områden. Områdena har amplifierats och fragmenten har detekterats med hjälp av fluorescens. Ett EPG har skapats som visar på att profilen från individ ett matchar med DNA från brottsplatsen. En laddarmall finns till för att kunna jämföra antalet upprepningar från de tre olika individerna. En match har upptäckts och därefter genomförs fler analyser för att stärka identifieringen.

## Högekänslig DNA-analys

### Hur går analysen till?

Om DNA-profilen från en brottsplats är ofullständig kan inte en standard STR-analys genomföras. Beroende på mängden DNA från en brottsplats kan forensiker behöva använda sig av andra metoder. Mängden DNA kan variera beroende på vad för sorts brott som begåtts, det kan vara en reducerad mängd epitelceller från en persons hud som fastnat på ett objekt, små blodstänk fläckar eller små ben/vävnads fragment som lämnats kvar på brottsplatsen (Turingan *et al.* 2016). Den metod som är mest vanlig när man analyserar en liten mängd DNA är högekänslig DNA-analys. Denna metod fungerar på liknande sätt som standard STR-analys-metoden. Det som skiljer dem åt är att man i högekänsliga DNA-analyser ofta ökar antalet cykler i PCR-amplifieringen, vilket resulterar i en ökad mängd DNA (McNevin *et al.* 2015).

### Tolkning av analysen

Tolkningen av en profil som skapats med högekänslig DNA-metod kan vara problematisk då det lätt uppstår problem med metoden om mängden DNA är så begränsad. En eller fler alleler hos en heterozygot individ kan missas i analysen, alleler kan hoppa in och det kan förekomma stamningar under PCR-reaktionen när DNA-området amplifieras (figur 3) (Gill *et al.* 2005).

### Variation av allelernas topphöjd

I ett EPG som fås när en högekänslig DNA-analys genomförs kan variation bland topparna ses även då replikaten är tagna från ett prov som blivit amplifierat på samma sätt. Denna metodvariation ökar då provet innehåller små mängder DNA. Fler cykler genomförs då mängden DNA är reducerad, detta för att öka DNA-mängden inför en analys. Då fler cykler genomförs blir amplifieringen mindre effektiv och därför ökar variationen. Vid ett visst lokus kanske en topp inte når över tröskelvärde (ger inte nog med fluorescens för att detekteras i EPG) för att räknas som en topp, medan samma allel i samma lokus men i ett annat replikat når över tröskelvärde och då räknas som en topp (Kelly *et al.* 2014). Variationen mellan replikatens topphöjd mäts genom att titta på heterozygotbalansen mellan de olika replikaten (Taylor *et al.* 2016). Om mängden DNA i utgångsprovet är mindre än 200 picogram finns en ökad risk att heterozygotbalansen rubbas och det genom att alleler hoppats över eller att fler alleler hoppar in

i en STR-profil (figur 3) (McNevin *et al.* 2015). Denna balans räknas ut genom att dividera det längre fragmentet med det kortare. Heterozygotbalansen visade sig vara starkt korrelerad med hur mycket DNA som finns tillgängligt. Även cyklernas antal hade påverkan på heterozygotbalansen, då mer negativt. Fler ensamma alleler upptäcktes. Alleler som egentligen skulle tillhöra ett heterozygot allel-par och blev då istället klassats som en så kallad falsk homozygot, en heterozygot vars ena allel hoppats över vid amplifieringen, vilket ger ett lägre värde för heterozygotbalansen (Hedell *et al.* 2015).

### **Antalet PCR-cyklers påverkan**

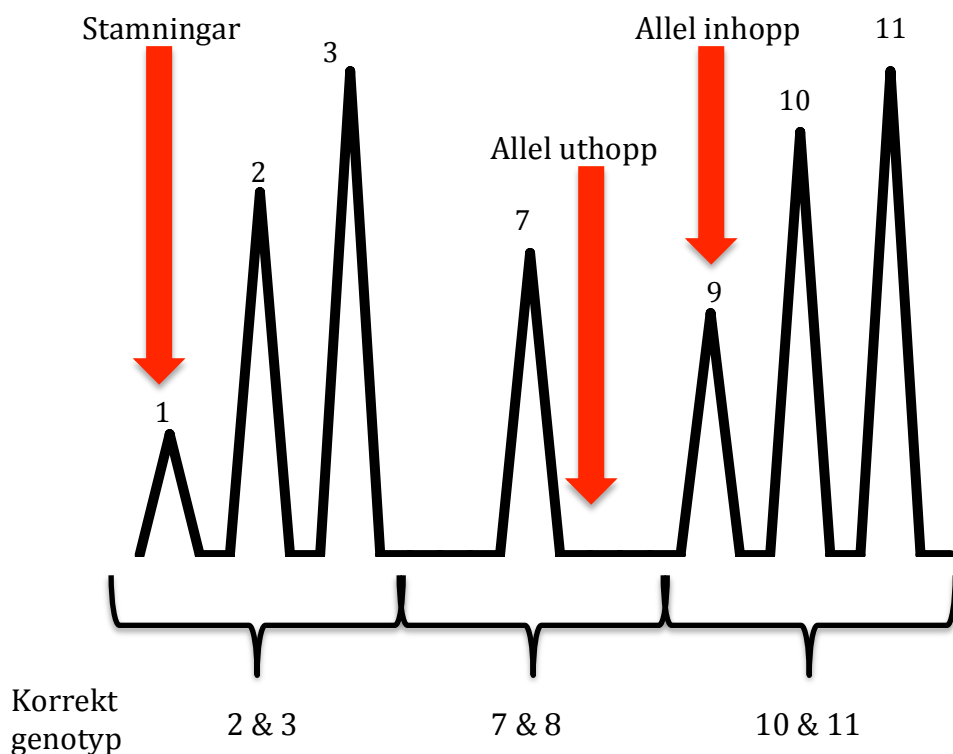
Vid en normal analys rekommenderas det att låta PCR-reaktionen genomföra 30-33 cykler men som nämnts tidigare ökar man ofta antalet cykler vid en högkänslig DNA-analys. Det visade sig att oavsett DNA-mängd gav resultatet att vid ett ökande av antalet cykler efter 34 stycken påverkades inte antalet detekterade alleler som hoppats över. En överhoppad allel är en allel som skulle ha räknats till en heterozygot men av någon anledning har en av allelerna missats. En heterozygot kan ha noll, en eller två alleler som hoppats över (Hedell *et al.* 2015).

### **Amplifieringsstörningar**

Platsen som primern ska binda till kan leda till störningar av topparnas balans i ett EPG. På bindningsplatsen på allelen för primern kan det finnas en mutation, vilket gör att amplifieringen blir mindre effektiv. Om mutationen sitter på 5' -ändan av platsen där primern ska binda till kan detta blocka kopieringen och allelen skulle då inte kopieras. Detta skulle då resultera i en falsk homozygot. Om mutationen är placerad på en annan plats inom primerns bindningsområde påverkar det temperaturen som gör att DNA-helixen separerar, vilket leder till negativ påverkan av DNA-områdets amplifiering (Clayton *et al.* 1998).

### **Stamningar**

Det inte alltid man får enbart två toppar vid ett lokus, vilket är maxantalet för antalet toppar för en individ. Om man får en STR-profil som visar fler än två toppar kan man dra slutsatsen att det är en mix av DNA eftersom att en individ bara har två kopior av varje allel vid ett visst lokus. Det behöver dock inte alltid vara en mix av DNA. Under PCR-processens bindningsfas när primrarna fäster till området av intresse för att kopiera DNA så kan, efter varje cykel i PCR-reaktionen, en, två eller fler primrar fästa till samma bindningsplats. Det som då syntetiseras i PCR-reaktionen kallas för stamningar. Man såg i en studie att en typ av stamningar uppkommer om den ny-syntetiserade DNA-strängen som produceras kan binda till sig själv och bilda en loop-liknande struktur. Detta gör att den ny-syntetiserade strängen blir några nukleotider kortare än strängen som var mall under transkriberingen. Fenomenet kan påverka STR-områden som då kan försvinna respektive läggas till. Stamningarnas sannolikhet att uppkomma beror på vilket lokus som kopieras (Bright *et al.* 2014). Topparna i EPG:et som stamningarna orsakar är sällan lika höga som de toppar som bildas då de "riktiga" STR-området avläses. Ibland kan det dock ställa till problem om huvudtoppen själv inte är speciellt hög (figur 3). Hedell *et al.* visade år 2015 att om antalet PCR-cyklar ökade hade det ingen effekt på närvaron av stamningar i EPG:et.



Figur 3. Förenklad bild på ett EPG som visar exempel på; Stamningar (1), som kan uppkomma då fler än en primer binder till samma område vid amplifieringen. Allel uthopp (8) /Allel inhopp (9), där en allel hoppat in/ut och ger på så vis en felaktig DNA-profil.

### N-band

Polymeraset som läser av DNA:t under amplifierandet kan lägga till en nukleotid när den nya strängen är syntetiserad. När sekvensen med den extra nukleotiden amplifieras skapas alltså en helt ny sträng som innehåller den nya "extra"-nukleotiden. Detta betyder att från en ursprungssträng kan de bildas två helt nya DNA-molekyler. Man får band på elektroforesgelen som är större än ursprung DNA-molekylen, detta kallas N-band. Dessa avläses därefter i ett EPG och visar då falska fragmentstorlekar som resultat (Clayton *et al.* 1998).

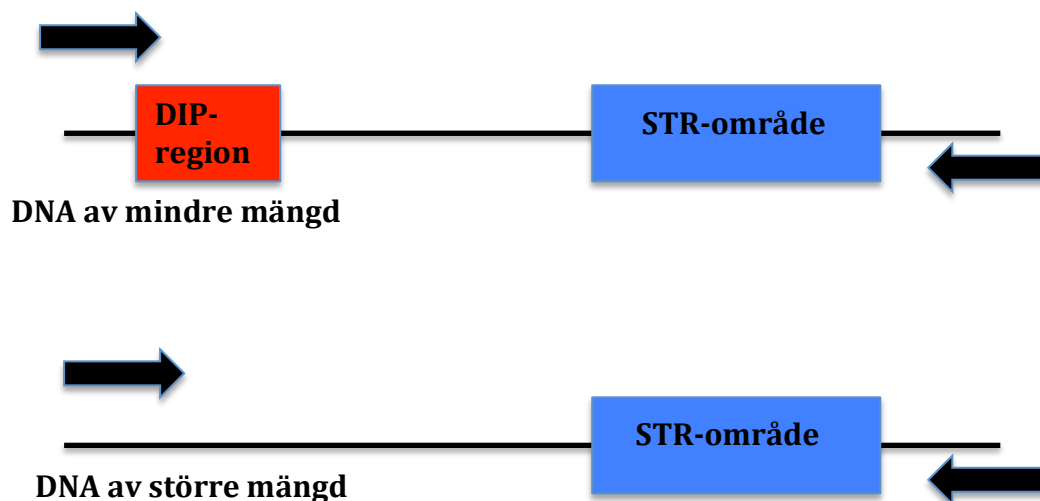
### Släktingars likheter

Människors DNA ser väldigt lika ut men skiljer sig en del mellan olika individer, dock mindre mellan släktingar. I en DNA-mix mellan två släktingar är det inte alltid man upptäcker att det är en blandning. Släktingarnas DNA-profiler kan sakna toppar från båda individerna och vad som egentligen skulle vara två olika alleler kan detekteras som samma. Det som gör det möjligt att detektera att det är två eller fler individers DNA trots bristen på extra toppar är genom att urskilja asymmetrin på topparna. Toppar som är asymmetriska på ett uppenbart heterozygot lokus kan visa på att DNA-provet är en mix trots frånvaro av de andra bidragsgivarnas toppar. Då mängden DNA är reducerad är det större chans för att missar mellan släktingars likheter i DNA:t kan uppstå och chansen minskar då DNA finns i större mängd (Clayton *et al.* 1998).

## Idéer till lösningar

### DIP-STRs markörer

Vid brottsfall där det inte alltid finns kroppsvätskor som blod och sperma på platsen och när förövarens DNA utgör mindre än 5 % av den totala mängden DNA som är tillgängligt i provet är det svårt att genomföra ett tillförlitligt DNA-test. I det flesta fall finns det DNA av förövaren med i DNA-provet som tas från brottsplatsen men detta kan i vissa fall förbli oupptäckt då mängden ursprungs DNA också är av liten mängd. För att reducera antalet fall då detta problem förekommer föreslogs en ny metod med nya markörer. Denna metod bygger på att para radering/insättnings polymorfism (deletion/insertion polymorfism, DIP) markörer med STRs. Markörerna fick namnet DIP-STR och är framförallt till störst användning då det finns en DNA-mix. En DIP-STR är två typer av markörer som parats ihop för att underlätta detektering av en mindre mängd DNA i en DNA-mix. DIP-STRs finns i hela genomet och är specifika för varje bidragsgivare i en mix. Med denna metod skulle man kunna analysera DNA-mixar med olika ursprung och ålder. I en mix har man ofta en större respektive mindre mängd av bidragsgivarnas DNA. Markörernas unika egenskap är att de kan bland en mix av DNA målsöka ett mindre specifikt område hos en liten mängd DNA, detta baserat på hur primrarna är designade. Man amplifierar hela DIP-STR regionen med en primer som överlappar den deleterade/införda regionen (DIP) på ena sidan och en primer som hybridiserar på andra sidan DIP-STR regionen. Båda primrarna är sekvensspecifika. Den ena primern binder antingen till alleler som har den specifika deletionen eller insättningen av alleler, den insättningen som bara det DNA av mindre mängd av den totala mängden DNA i provet har (figur 4). Den andra primern ser till att STR-regionen också blir amplifierad. Som resultat visade studien att dessa markörer var nästan 100 gånger mer känsliga vid detektering av DNA-blandningar än standard STR-markörerna som används vid dagens DNA-profilering. Beroende på polymorfismgraden varje DIP har ökar chansen för att identifiera alleler som är unika för den mindre DNA-mängden i provet och på så vis kunna hitta DNA som i ett annat fall kunnat förbli oupptäckt om man bara använt sig av STR-markörer (Oldoni *et al.* 2015).



Figur 4. Ett område av två DNA-regioner i samma DNA-mix. En region av bidragsgivaren som bidragit med mindre DNA och en region från bidragsgivaren som bidragit med en större mängd DNA. De svarta pilarna indikerar PCR-primrar som under amplifieringen läser av olika områden. DNA-regionen som kommer från bidragsgivaren som bidragit med mindre DNA till mixen har ett DIP-STRmarkör vilket gör den mer specifik vid bindningen av primern. Denna teknik gör det möjligt att från en mix med olika mängd av de två bidragsgivarnas DNA kunna hitta en specifik sekvens som inte skulle hittats då man bara använt sig av STRs. Koncept tagen från (Oldoni *et al.* 2015).



### **Primer design**

Ett av de svåraste problemen när mängden DNA på något sätt är ofullständigt, både om det är mixat, om det degraderats eller om det bara finns en liten mängd tillgänglig är att DNA måste bli kopierat på ett korrekt sätt.

Primrarna som används i PCR-reaktionen för amplifiering av en DNA-region har stor betydelse för huruvida DNA ska bli transkriberat eller ej. Butler et.al gjorde år 2003 ett försök till att minska primrars feltolkning genom att designa om amplikonerna och skapa "mini-STRs" som gör att primrarna får mindre att läsa av och blir således mer effektiva. Analyser av degraderat DNA visade sig ha stor nytta av denna skillnad då de inte behövde samma mängd DNA som när standard primrar används till vanliga STR-områden. Då inte samma mängd DNA behövdes kunde cyklerna i PCR-reaktionen minskas och färre störningar vid amplifieringen orsakades vilket gjorde att analysen förbättrades (Butler *et al.* 2003).

Det har gjorts fler studier på hur man kan förbättra primrarna och förhindra felaktig kopiering av målsekvenser. Genom att man jämfört vanliga PCR-primrar med primrar som är delvis gjorda av låsta nukleotider (locked nucleic acids, LNA). När en LNA induceras till antigen en DNA eller RNA sekvens ändrar det konformationen på DNA-helixen. Denna konformationsändring ger en ökad styrka i basparningen av primern och DNA-strängen då LNAs nu har starkare bindningsstyrka till den komplementära sekvensen på grund av konforamtionsändringen. Den ökade bindningsstyrkan ökar på så sätt känsligheten i PCR-reaktion vilket resulterar i bättre sekvensering och avläsningskvalité. Genom att jämföra de två olika primrarnas topphöjden (fluorens enheter) visade de på hur bra amplifieringen hade genomförts, en högre topphöjd gav mer fluorescens och på så vis en bättre amplifiering av primern.. Topphöjden minskade då templatmängden minskade, detta för alla loci och primrar. Av de loci som undersöktes så visade den mest framgångsrika LNA-primern en ökning med 30 % i topphöjd jämfört med den mest framgångsrika DNA-primern. Det man även kunde se var att den minst framgångsrika LNA-primern fortfarande hade en ökad topphöjd med 26 % jämfört med den minst framgångsrika DNA-primern. Svårigheten vid användandet av LNAs är designen av oligonukleotiden. Det finns inga riktiga guidelinjer för vilken primer-design som ger bäst resultat, men det finns studier som har visat att när man inkorporerar LNAs i primrar så har deras position och antal stor effekt. Om antalet LNAs som inkorporeras är för många kan amplifikationen försämrans. Beroende på vilken bas man inkorporerar ger de olika bindningsstyrka (C>T>G>>A). Det är dock viktigt att man vid oligonukleotidproduktionen inte har för stor bindningsstyrka mellan baserna eftersom det kan förstöra amplifikationen (Ballantyne *et al.* 2008).

### **Polymerasets effektivitet**

Kersbergen och Kloosterman (2003) visade i en studie de gjorde att om man efter 28 cykler ökar mängden taq polymeras och att reaktionen sedan får fortsätta 6 cykler förbättras amplifieringen. Denna strategi som skiljer sig åt från den vanliga högkänslighets analysen, gjorde att fler enstaka celler analyserades snabbare mellan de 28 första cyklarna än de 6 sista.

### **Nested PCR.**

Det finns andra sätt att gå tillväga än att öka antalet cykler i PCR-reaktionen vid en högkänslig DNA-analys. Nested PCR är en metod som använder två olika set av primrar i två separata PCR-reaktioner. I den första reaktionen amplifierar primrarna STR-regionen och dess flankerade region. Under den andra reaktionen är primrarna designade för att amplifiera en mindre produkt genom att använda den produkt som producerades i den första reaktionen

som mall. Denna metod kan analysera innehållet i en enda cell men nackdelen är att de lätt kan ske kontamination då de två PCR-reaktionerna sker i två olika rör (Gill *et al.* 2000).

### **Rening av DNA-provet före PCR-reaktionen**

Som nämnts ovan ökas antalet cykler vid en högkänslig DNA-analys men för att se om det fanns ett annat sätt att öka känsligheten vid detekteringen av alleler från ofullständiga DNA-prover gjordes en studie. Man testade huruvida olika reningsmetoder före PCR-reaktionen hade någon effekt på hur alleler i den kapillärelektroforetiska separationen detekterades med fluorescens. Genom att jämför fyra olika post-PCR rening metoders fluorescenssignalintensitet och grad av eluatets renhet. Qiagen MinElute silica column-metoden, vilket är en typ av reningsmetod för att extrahera DNA-prover som baseras på att DNA binder till en kiselbindare och tvättas därefter med alkohol och på så vis undviker man att själva provet tvättas bort. När denna reningsmetod användes ökade fluorescenssignalintensitet fyra gånger mer hos de renade än hos de orenade proven. Vid användandet av all PCR produkt som renats av Sephadex™ (reningsmetod med gelfiltrering) såg man att signalintensiteten ökade 3,5 gånger mer. Ytterligare en märkbar notering gjordes då man såg att vid användandet av Qiagen MinElute silica column-metoden kunde man förbättra känsligheten hos PCR-produkten och få en hel profil även om mängden DNA bara var 20 pg. Denna metod visade sig även kunna få signifikant data av endast 5 pg DNA (motsvara cirka en cells DNA) utan att öka antalet cykler under PCR-reaktionen. Genom att justera volymen av eluatet och mängden renad produkt i PCR-reaktionen kan känsligheten kontrolleras. Man kan bland annat använda denna reningsmetod för att förbättra allelers signaler som inte överstiger tröskelvärde i ett svagt prov (Smith & Ballantyne 2007).

### **Sannolikhetsförhållandet vid en DNA-mix**

Society of forensic genetics DNA comission rekommenderade för några år sedan att använda sig av sannolikhetsförhållandet (likelihood ratio, LR). En analys från en brottsplats kan ha två alternativa hypoteser som stöder endera den åtalades och den anklagandes hypotes. I en DNA mix kan den åtalade ha en hypotes om att både den skyldige och en okänd individs DNA finns i provet, medan den försvarades hypotes kanske menar på att det är två okända som bidrar till provet. LR jämför sannolikheten av bevisen under dessa alternativa hypoteser. Om LR blir större än 1 stödjer bevisen den åtalade och om de blir lägre än 1 stödjer bevisen den försvarades hypotes.

En av de goda effekterna denna metod har är att den gör det möjligt för forskare att samtidigt överväga och jämföra säkerheten i analyserna. Försvaren kan hävda att det till exempel finns tre personer som bidragit till DNA-provet istället för en, detta för att göra sig själv oskyldig och då kan en LR-analys antingen godkänna eller förneka denna hypotes (Meuwly *et al.* 2015).

### **Härstammingsmarkörer**

Ett annat sätt att analysera mixar av DNA är att använda sig av Y-kromosomens DNA som bara det manliga könet har. Detta sätt kan användas vid specifika sexualbrottsutredningar där mixen av DNA innehåller mer kvinnligt än manligt DNA och man analyserar då Y-kromosomens STR-områden (Y-STRs). Laplace metoden är en statistisk metod som kan separera två personers Y-STR mixar. I en studie som gjordes 2014 kunde man med hjälp av Laplace metoden separera Y-STR mixar om man kollade på 21 loci med upp till mellan 42-52 % och hela 92-99 % av fallen kunde separeras då man kollade på 10 loci (Andersen M.M *et al.* 2015). Mitokondriellt-DNA(mtDNA) är DNA som finns utanför cellkärnan och ärvs bara från mamman till barnen. Detta DNA är icke rekombinerat vilket betyder att det inte skiljer sig

något mellan olika generationer. Den icke rekombinerade delen av Y-kromosomen ärvs även den i exakt samma upplaga men denna ärvs från far till son. Dessa två kallas härstamningsmarkörer och analyseras på samma vis som de vanliga STR-markörerna. Dessa markörer används som verktyg vid sexualbrott och ifall DNA-provet undgått degradering (Buckleton *et al.* 2011).

### **Nästa generations sekvensering**

Det är många forskare som idag utreder huruvida användningen av nya sekvenseringsmetoder (next generation sequencing, NGS) kan tillämpas inom den forensiska vetenskapen. Detta som ytterligare en metod eller en alternativ metod till användandet av kapillärelektroforesen vilket idag är standardmetoden för att analysera STRs. Problemet med kapillärelektroforesmetoden kan vara att det är svårt att analysera många markörer då loci kan vara överlappade och antalet fluorescerande ämnen är begränsat (Kim *et al.* 2016). Man kan med NGS-metoden sekvensera mycket information med stor noggrannhet och man kan av NGS få information från flera individers markörer från bara en enstaka körning. Genom den snabba utvecklingen under de senaste åren har NGS ökat i popularitet. Av det stora antalet sekvenser som kan skapas med NGS-tekniken kan små variationer upptäckas (Lee *et al.* 2016). NGS har många fördelar, bland annat att det går relativt snabbt till en låg prissumma, man kan med NGS även göra en sekvensanalys på RNA, tillskillnad från när kapillärelektrofores-metoden används då det är DNA-sekvenser som analyseras. Dagens metod kan använda sig av verktyg för att sekvensera en specifik RNA sekvens. Med denna ”riktade” RNA-sekvenseringen kan man bland annat dra nytta vid identifieringen av kroppsvätskor då man nu kan ha betydligt fler biomarkörer med i analysen. År 2015 gjordes ett projekt där man ville utnyttja denna ”riktade” sekvenseringen och kunna återsekvensera mRNA-transkript för att identifiera kroppsvätska/vävnadens ursprung från en DNA-profil. Resultatet man fick i projektet var att med hjälp av multiplex RNA-NGS-analys kan man identifiera kroppsvätskor eller vävnadskälla av prover. Man använde sig av Illumina Miseq plattformen vid denna studie men det finns även andra NGS-plattformar (Hanson *et al.* 2015).

I ett experiment som gjordes av Kim *et al.* (2016) ville man studera om NGS-metoden och kapillärelektroforesmetoden gav samma resultat vid användandet inom kriminaltekniken. I en tub fanns 18 markörer som man analyserades både med NGS-metoden och med kapillärelektroforesmetoden. På så sätt ville man jämföra hur väl alleler detekterades med de två metoderna. Prover som kom från samma källa, DNA-mixar och DNA som blivit nedbrutet på konstgjord väg visade goda resultat. Ett multiplex PCR system kan producera ampliconer av denna typ av ofullständiga profiler för NGS-analys. Om NGS-metoden ska kunna användas inom kriminaltekniken som komplement till kapillärelektroforesmetoden behövs fler tester göras på enskilda prover samt test för att se hur provets känslighet.

### **Diskussion**

Det finns många problem som kan uppstå när en högkänslig DNA-analys genomförs. Både genom ny forskning och nya experiment utvecklar man såväl gamla som nya tekniker. Självklart uppstår det inte alltid problem och ibland kan de uppstå mer problem än dem jag valt att ta upp i denna uppsats. Det viktiga tycker jag är att man försöker komma på nya metoder och utvecklar de metoder som finns för att slippa problem med exempelvis stamningar, allel inhop och /uthopp av alleler, eftersom det kan få förödande konsekvenser som kan leda till att personer blir oskyldigt dömda. De 4 % som var oskyldigt dömda av alla som blev avrättade i USA 2013 är en siffra som jag tycker måste minska. Jag tycker att

det är ett viktigt med forskning inom detta område eftersom denna siffra inte får öka med anledning av att någon gör ett oavsiktligt fel på grund av feltolkningar.

Genom att man från olika infallsvinklar försöker hitta lösningar för att utveckla dagens metoder, är enligt mig en stor fördel. Man försöker lösa problem på olika sätt och denna fördel gör det möjligt att man snabbare kommer hitta den ultimata lösningen och de 4 % som avrättades trots att det var oskyldigt dömda kommer minska ofantligt mycket om nästan inte ned till 0 %. Från de studier som gjorts fram tills idag har man tittat på både, hur man kan med hjälp av andra markörer (DIP-STRs) än de vanliga STR-markörerna förbättra analysen av DNA-mixar. Hur man kan med unikt designade primrar förbättra amplifieringen och på så vis minska risken för att DNA ska bli fel läst. Även sätt att genom billig och relativt enkel rening av DNA-provet göra en förbättring av DNA-analysen. I denna uppsats redovisar jag endast ett axplock av alla studier som gjorts för att förbättra högkänsliga DNA-analyser och med tanke på kunskapen dagens forskare har är nog detta bara början på högkänslig DNA-analysers utveckling.

Självklart skiljer sig kostnaden för de olika metoderna, vilket är en viktig faktor att ha i baktanke när det kommer till vilken metod som skulle vara till bäst användning. Det beror dock inte bara på vad varje metod kostar utan också på vilken typ av ofullständig DNA-profil man får från brottsplatsen. Om man till exempel skulle hitta DNA från ett gammalt brott, säg en gammal vävnad, så skulle man mest troligt använda sig av metoden med "mini-STRs" då mängden intakt DNA inte behöver vara så stor vilken den i många fall inte är vi ett degraderat DNA-prov. Om man istället vid en våldtäkt skulle ta ett DNA-prov, där det skulle finnas en mindre mängd av en bidragsgivares DNA jämfört med den andre bidragsgivaren så skulle man inte använda sig av metoden med "mini-STRs". Vid denna typ av ofullständig profil skulle "mini-STRs" inte göra någon nytta utan här skulle man mest troligt försöka använda sig av DIP-STRs eller analysera Y-kromosomens DNA. Om man ser en varierande topphöjd i ett EPG och har svårt att tolka topparna kan man med en relativt enkel och billiga metod kunna rena provet, vilket kan vara en simpel lösning till att minska topphöjdsvariation. En sådan enkel teknik som en rening av DNA-provet kan i många fall räcka för att minska variationen och på så vis möjliggöra en förbättrad tolkning av en DNA-profil. Det sättet man kan förbättra amplifieringen på är genom att göra unik design på primrarna med låsta nukleotider. Det är inte bara primrarna som man kan göra unik design på utan också själv amplikonerna och på så sätt få en förbättrad amplifiering med mindre missar. I och med att alla nya studier som tagits fram för att lösa problemen skiljer sig ifrån varandra gör det svårt att med säkerhet säga vilken som skulle vara bäst lämpad för att vara den "ultimata" metoden. Enligt mig kanske det inte heller nödvändigtvis måste finnas en "ultimat" metod eftersom att brottsfall ofta skiljer sig ganska rejält ifrån varandra och att varje eget fall behöver sin unika metod för att en tolkning av en DNA-profil ska bli så säker som möjligt. Då dagens analyser inte använder sig av denna nya forskning är det svårt att kunna sätta metoderna i ett verklighetsperspektiv. Det finns idag bara studier som gjorts *in vitro* och där resultaten är varierande. NGS är den metod jag tror kan vara den som kommer göra mest framsteg, dels för att det är den metod som står för mest säkerhet vid transkriberingen men också för att den idag är ett användbart verktyg för många fler ändamål och inte bara inom kriminaltekniken.

Det finns många problem med analys av partiella DNA-prov, men så behöver det inte alltid vara. Inom en snar framtid finns det förhoppningsvis inte några problem alls då dagens forskning går framåt. Med tanke på de lösningar till problem som har belysts i denna uppsats och resultatet de studierna har fått så dröjer det nog inte länge innan det aldrig kommer

inträffa en oskyldigt dömd i dagens samhälle.

## Tack

Tack till min seminarie- & återkopplingsgrupp som kommit med tips längs skrivandet och tack till Stefan Bertilsson som guidat och gett mig vägledning under hela skrivprocessen.

## Källor:

- Andersen M.M, Eriksen P.S, Mogensen H.S & Morling N. 2015. Identifying the most likely contributors to a Y-STR mixture using the discrete Laplace method. *Forensic science international: Genetics* **15**:76-83.
- Ballantyne K.N, van Oorschot R.A.H & Mitchell R.J. 2008. Locked nucleic acids in PCR primers increase sensitivity and performance. *Genomics* **91**:301-305.
- Bright J, Stevenson K.E, Coble M.D, Hill C.R, Curran J.M & Buckleton J.S. 2014. Characterising the STR locus D6S1043 and examination of its effect on stutter rates. *Forensic science international: Genetics* **8**:20.
- Buckleton J.S, Krawczak M & Weir B.S. 2011. The interpretation of lineage markers in forensic DNA testing. *Forensic Science International: Genetics* **5**:78-83.
- Butler J.M 2005. *Forensic DNA typing: biology, technology, and genetics of STR markers*. 2:a uppl. Elsevier Academic Press, Amsterdam:Boston.
- Butler J.M, Shen Y, McCord B.R. 2003. The development of reduced size STR amplicons as tools for analysis of degraded DNA. *Journal of forensic sciences* **48**:1054-1064.
- Clayton T.M, Whitaker J.P, Sparkes R & Gill P. 1998. Analysis and interpretation of mixed forensic stains using DNA STR profiling. *Forensic Science International* **91**:55-70.
- Duitama J, Zablotzkaya A, Gemayel R, Jansen A, Belet S, Vermeesch J.R, Verstrepen K.J & Froyen G. 2014. Large-scale analysis of tandem repeat variability in the human genome. *Nucleic Acids Research* **42**:5728-5741.
- El-Alfy S.H & Abd El-Hafez A.F. 2012. Paternity testing and forensic DNA typing by multiplex STR analysis using ABI PRISM 310 Genetic Analyzer. *Journal of Genetic Engineering and Biotechnology* **10**:101-112.
- Gill P, Curran J & Elliot K. 2005. A graphical simulation model of the entire DNA process associated with the analysis of short tandem repeat loci. *Nucleic acids research* **33**:632-643.
- Gill P, Haned H, Bleka O, Hansson O, Dørum G & Egeland T. 2015. Genotyping and interpretation of STR-DNA: Low-template, mixtures and database matches-Twenty years of research and development *Forensic science international: Genetics* **18**:100.
- Gill P, Whitaker J, Flaxman C, Brown N & Buckleton J. 2000. An investigation of the rigor of interpretation rules for STRs derived from less than 100 pg of DNA. *Forensic Science International* **112**:17-40.
- Gross S.R, O'Brien B, Hu C & Kennedy E.H. 2014. Rate of false conviction of criminal defendants who are sentenced to death. *Proceedings of the National Academy of Sciences* **111**:7230-7235.
- Hanson E, Ingold S, Haas C & Ballantyne J. 2015. Targeted multiplexed next generation RNA sequencing assay for tissue source determination of forensic samples. *Forensic Science International: Genetics Supplement Series* **5**:e441-e443.
- Hedell R, Dufva C, Ansell R, Mostad P, Hedman J. 2015 Enhanced low-template DNA analysis conditions and investigation of allele dropout patterns. *Forensic science international* **14**:61-75.

- Kelly H, Bright J, Buckleton J.S & Curran J.M. 2014. A comparison of statistical models for the analysis of complex forensic DNA profiles. *Science & justice : journal of the Forensic Science Society* **54**:66-70.
- Kim E.H, Lee H.Y, Yang W.I, Yang I.S, Jung S & Shin K. 2016. Massively parallel sequencing of 17 commonly used forensic autosomal STRs and amelogenin with small amplicons. *Forensic Science International: Genetics* **22**:1-7.
- Kloosterman A.D & Kersbergeg P. 2003. Efficacy and limits of genotyping low copy number DNA samples by multiplex PCR of STR loci. *International Congress Series* **1239**:795-798.
- Lee H.Y, Lee Y, Lee E.Y, Oh S.Y, Jung S, Yang I.S, Yang W.I & Shin K. 2016. Massively parallel sequencing of the entire control region and targeted coding region SNPs of degraded mtDNA using a simplified library preparation method. *Forensic Science International:Genetics* **22**:37-43.
- McNevin D, Edson J, Robertson J, Austin J.J. 2015. Reduced reaction volumes and increased Taq DNA polymerase concentration improve STR profiling outcomes from a real-world low template DNA source: telogen hairs. *Springer Science+Business Media*, doi: 10.1007/s12024-015-9679-3.
- Meuwly D, Spreuwers L, Veldhuis R & Ali. 2015. Sampling variability in forensic likelihood-ratio computation: A simulation study. *Science & justice* **55**:499-508.
- Oldoni F, Castella, V & Hall D. 2015. A novel set of DIP-STR markers for improved analysis of challenging DNA mixtures. *Forensic science international:Genetics* **19**:156-164.
- Reid Y, Storts D, Riss T & Minor L. 2013. Authentication of Human Cell Lines by STR DNA Profiling Analysis. *Eli Lilly & Company and the National Center for Advancing Translational Sciences*.
- Smith P.J & Ballantyne J. 2007. Simplified Low-Copy-Number DNA Analysis by Post-PCR Purification. *Journal of Forensic Sciences* **52**:820-829.
- Study Results from Zhejiang University Provide New Insights into Genetic Research. Genetic analysis of STR markers on chromosome 21 in a Han population from southeast China. 2015. *China Weekly News* 266.
- Taylor D, Buckleton J & Bright J. 2016. Factors affecting peak height variability for short tandem repeat data. *Forensic science international:Genetics* **21**:126-133.
- Tekcan A, Tural S, Elbistan M, Kara N, Guven D & Kocak I. 2014. The combined QF-PCR and cytogenetic approach in prenatal diagnosis. *Molecular Biology Reports* **41**:7431-7436.
- Turingan R.S, Vasantgadkar S, Palombo L, Hogan C, Jiang H, Tan E & Selden R.F. 2016. Rapid DNA analysis for automated processing and interpretation of low DNA content samples. *Investigative genetics* **7**:2.

# **Analys av partiella DNA-profiler-metoder, problem och lösningar:** etisk bilaga

**Minna Andersson**

Självständigt arbete biologi 2016

## **Etik**

### **Forskningsetik**

I detta arbete har jag valt att använda mig av databaserna NCBI och Web of science för att leta efter originalartiklar. Jag började med att läsa review artiklar för att få en bredare förståelse för ämnet och sedan gick jag via deras källor till originalartiklarna. Jag valde att jämföra liknande studier med varandra allt eftersom jag läst fler artiklar. På det sättet har jag uteslutit fakta som jag ansett kunnat vara missledande och istället bara fokuserat på fakta som jag utifrån min förmåga ansett vara en relevant studie med tillförlitligt resultat.

### **Inskränkning på integriteten vs fördel för samhället**

I dagens läge när ett brott begås lämnar nästan alltid vederbörande DNA efter sig, det kan vara allt från blod, sperma, vävnader eller hår. Polisen har rätt att ta DNA-prover som lagras i ett DNA-register från personer som blivit dömd till grövre brott än böter. Detta register är självklart till en fördel och hjälp för samhället och polisen att kunna identifiera den skyldige till ett brott men är det etiskt rätt att lagra en människas DNA, inkräktar det på den brottsdömdes integritet? Informationen som lagras säger ingenting om hur man är som person, vad man har för egenskaper eftersom områdena som analyseras finns i den icke kodande regionen hos varje individ (El-Alfy & Abd El-Hafez 2012). Gränsen mellan vilken fördel det ger samhället och inskränkningar av en individs integritet är enligt mig hårfin.

### **DNA-registret och ansvaret**

Det finns självklart både för- & nackdelar med DNA register. Nackdelen med att det skapas sådana här register är inte bara att det kan anses vara etiskt kränkande för den enskilde individen. Om informationen om ens DNA hamnade i fel händer vad skulle det kunna leda till, vilka ändamål skulle det kunnat bli användbart för? Ansvaret som ligger hos forensiker vid en DNA-analys är väldigt stort. Det är viktigt att proven hanteras varsamt och med stor noggrannhet. Risken för kontamination och felanalyser minskar på så sätt samt att förtroendet hos de som lämnat DNA ökar för de som bär ansvaret, både med tanke på den etiska aspekten för vem som har tillgång till provet men också med tanke på hur de sköts och till vad det används för.

### **Användandet av PKU-registrets**

PKU-registret är en biobank med blodprover som samlats för användning vid forskning på Karolinska Universitetssjukhuset. Detta register kan polisen begära tillgång till för att identifiera en mördare. Vi mordet på Anna Lindh valde polisen att analysera en misstänkts blod från PKU-registret mot det som hittats på brottsplatsen och på så vis identifiera mördaren. Att använda sig av PKU-registret vid brottsutredningar kan vara till fördel då man kan avföra en misstänkt och givetvis en nackdel för den enskilde brottslingen. Att använda sig av en biobank som är ämnad för forskning för att identifiera en mördare, är enligt mig fel och inte etiskt korrekt. Varje nyfödd individ låter sitt DNA sparas för ett gott syfte och när detta sedan kan användas på ett annat sätt än vad den ursprungliga avsikten var, är enligt mig fel och inte etiskt korrekt.