



Lærervejledning

Sanger sekventering - fra mikroorganisme til DNA sekvens

Generelt om øvelsen.

Øvelsen Sanger sekventering - fra mikroorganisme til DNA sekvens er en øvelse, hvor bioinformatik er tænkt ind som en del af et tema om immunsystemet og infektionssygdomme til Biotek A elever. Elevforudsætningen er generel viden om DNA's opbygning og proteinsyntese, og det er en fordel hvis eleverne har arbejdet i laboratoriet med PCR og gelelektroforese før.

Øvelsen er delt op i 3 aktiviteter, som man kan undervise som et sammenhængende forløb.

Aktivitet 1 er et virtuel lab, hvor eleverne guides igennem processen fra DNA oprensning fra bakteriekulture fra 6 patienter, PCR opformering af 16S SSU rRNA, oprensning af PCR produktet, PCR cycle sequencing og DNA sekventering ved gelelektroforese. Det virtuelle lab er på engelsk og er et meget styret lab, dvs. at eleverne ikke rigtig kan lave fejl i det. Der står meget beskrivende tekst til hvad der sker, og det primære i denne øvelse er at eleverne sætter sig godt ind i stoffet for at kunne svare på de spørgsmål, der stilles i øvelsesvejledningen samt at de visuelt ser hvad der sker ved en Sanger sekventering.

Varigheden af aktivitet 1 er ca. 1 - 1 ½ time, afhængig af om eleverne svarer på spørgsmålene i timerne eller ej.

Aktiviteten 2 viser, hvordan man tjekker kvaliteten af en Sanger sekventering ved at se på kromatogramtoppene fra laser detektoren, samt hvordan man laver en Contig (=konsensussekvens) ud af de ca. 700 basepar lange DNA sekvenser som man får ved Sanger sekventering. Til begge dele skal de bruge bioinformatik programmet Ugene. I øvelsen ser eleverne et eksempel på en Sanger DNA sekvens, for at få en forståelse af, at der er nogle baser man ikke kan bestemme med sikkerhed. De skal de visuelt se hvad der sker med kvaliteten af kromatogramtoppene i begyndelsen af en DNA sekvens, så de forstår hvorfor at metoden kun giver ca. 700 bp lange reads af sekvenserne.

Anden del af aktiviteten er at sætte 4 sekvenser sammen til en ca. 1500 bp lang sekvens, hvilket programmet gør ved at lave en alignment. Eleverne vil her visuelt kunne se hvordan det sker og få en forståelse af at DNA sekvens bestemmelse ikke altid er helt så ligetil som når man aflæser en serie af baser i en DNA database.

Varigheden af denne aktivitet er ca. 30 min.

Aktivitet 3. I denne aktivitet skal eleverne matche de 6 bestemte 16 S SSU rRNA sekvenser fra det virtuelle laboratorium med NCBI databasen vha BLAST search søgemaskinen. Da det er real life søgninger kan der forekomme ventetid, og det er derfor svært at vide præcist hvor lang tid søgningerne kommer til at tage, da det afhænger af aktiviteten af søgninger på databasen. Men typisk tager en søgning ca. 2 minutter, dvs i alt skal eleverne bruge ca. 12 minutter til søgningerne.

Aktivitet 1: Svar

Oprensning af DNA.

1. Hvorfor er det nødvendigt at oprense DNA fra bakteriekolonierne? Fordi at DNA er inde i cellerne og derfor vil være utilgængeligt for opformering med mindre man får dem ud i en væskeanalyse med mindre de kommer ud af cellerne. Samtidig kan cellerester hæmme muligheden for at DNA kan opformeres ved PCR reaktionen.
2. Hvordan opnår man bakterie lysis i laboratoriet? (lysis=når bakterier sprænges). Ved at tilsætte proteolytiske enzymer, der nedbryder cellevæggen hos bakterierne.
3. Hvorfor skal prøven opvarmes til 100°C efter bakterie lysis? For at denaturere de proteolytiske enzymer, så de ikke senere ødelægger f.eks DNA polymerasen, der også består af aminosyrer koblet sammen med peptidbindinger.
4. Efter centrifugering af prøverne, hvor befinder DNA'et sig så i prøven, i pellet (bundfaldet) eller i supernatanten (væsken). Det befinder sig i supernatanten

PCR opformering.

5. Hvad består et master mix af? Af en termostabil DNA polymerase, buffer, ACTG DNA nucleotidtriphosphater, og de 2 primere, en forward og en reverse.
6. Forklar hvad en universel primer er og hvorfor det er smart at bruge universelle primere, når man ikke ved, hvilken bakterie man har dyrket. Universelle primere er primere der binder til (næsten) alle organismer i en større gruppe, f.eks bakterier, fordi de binder til nogle meget konservative områder af et gen, hvor mutationer derfor altid er skadelige for organismen. For at alle bakterier skal have genet, laver man universelle primere til gener, der indgår i helt vitale livsprocesser i alle organismer, f.eks gener der koder for proteiner eller RNA der indgår i opbygningen af ribosomerne.
7. Hvorfor er det vigtigt at have både en positiv og en negativ kontrol når man laver PCR opformering? Den negative kontrol bruges til at sikre, at PCR mikset ikke er forurenet med andet DNA, der vil give en falsk positiv. Når man bruger universelle primere, er det l. et at komme til at forurene master mixet med bakterier fra f.eks luften, og derved risikerer man at der kommer et falskt positivt svar. Den positive kontrol er for at sikre sig, at man har blandet sit mastermix korrekt og at det virker. Hvis man ikke får nogle resultater i sine prøver, men den positive kontrol har virket, så ved man at det er ens DNA oprensning, der er gået galt. Har man derimod ikke noget resultat i den positive kontrol, ved man at der er noget galt med ens mastermix, og derfor behøver man ikke at kassere DNA oprensningerne.
8. Forklar hvorfor det er vigtigt at genet er semikonservativt, for at man kan bruge det til artsidentifikation vha PCR opformering med universelle primere. Det er vigtigt at genet har meget konserverede områder, da det er der at de universelle primere skal binde sig til, se svar 6. Men det er også vigtigt at der er områder med meget variation, da man ellers ikke kan skelne mellem de forskellige arter. Er DNA'et for konserveret i hele områder vil der være flere bakterierarter der har det samme DNA, og så kan man ikke skelne imellem arterne.

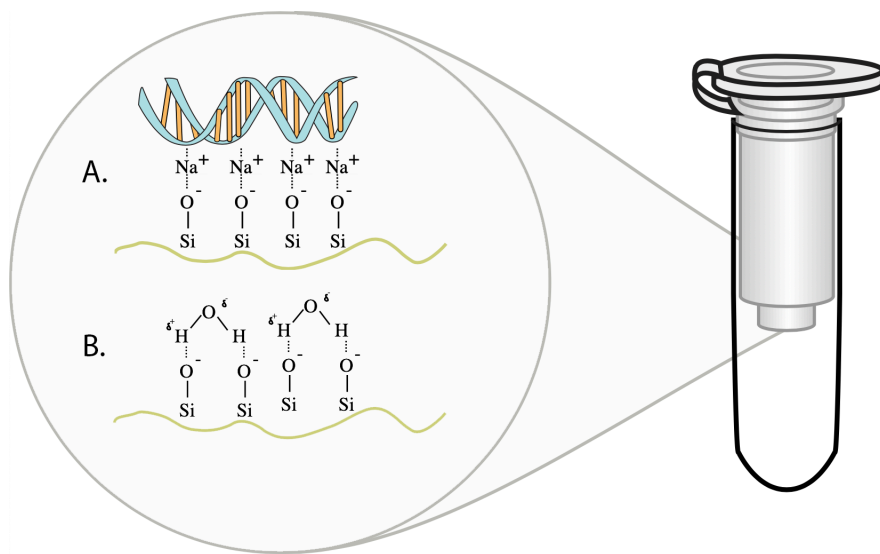


9. PCR maskinen kører 30 cykler med 3 forskellige temperaturer. Forklar hvad der sker med DNA'et ved hver temperatur og hvad man kalder de 3 steps. Der er tre forskellige steps. 1. Denaturering. Her varmes PCR prøven op til 95°C i 30 sek., hvilket bryder hydrogenbindingerne mellem DNA baserne, så DNA bliver enkeltstrengt. 2. Annealing. Her sænkes temperaturen til 60°C i 30 sek., så primerene binder sig til det enkeltstrengede DNA. Temperaturen er afhængig af hvilket primersæt man bruger, og er en afvejning mellem at hvis man sænker temperaturen for meget risikerer man at DNA igen bliver dobbeltstrengt og derfor ikke kan binde primerne. Men primere med et højt AT indhold danner færre hydrogenbindinger til det enkeltstrengede DNA og kræver derfor en lavere temperatur for at binde til det enkeltstrengede DNA end primere med højt GC indhold. 3. Extension. Her hæves temperaturen til 72°C, hvilket er optimumtemperaturen for den termostabile DNA polymerase. Hvor lang tid denne fase varer er afhængig af hvor langt et DNA stykke man skal opformere, jo længere stykke jo længere tid skal DNA polymerasen bruge. Ved 1500 bp bruger man typisk ca. 45 sek.

- Initial incubation step: 95°C for 10 minutes
- 30 cycles of the following sequence of steps:
 - Melt: 95°C 30 seconds
 - Anneal: 60°C 30 seconds
 - Extend: 72°C 45 seconds
- Final extension step: 72°C 10 minutes
- Final step: 4°C store at this temperature

Oprensning af PCR produkt.

1. Hvorfor er det smart først at lave en gelelektroforese af PCR produktet, den positive og den negative kontrol. For at kontrollere at PCR reaktionen har virket som den skal. Den negative kontrol viser at der ikke har været forurening i prøven, den positive at PCR reaktionen har virket som den skal og PCR reaktionen, at det fragment der er opformet har den korrekte længde, og at det derfor er det rigtige produkt man har fået.
2. Hvorfor er det nødvendigt at oprense PCR produktet? Fordi der vil være primerrester i PCR produktet, samt overskuds DNA nucleotidtriphosphater. I næste step, hvor man laver PCR cyclesequencing, er det meget vigtigt at der ikke er ekstra primere i prøven, da det vil medføre at der kommer fragmenter med uønsket længde pga at de laves ud fra en forkert primer. Uønskede DNA nucleotid triphosphater gør at forholdet mellem deoxy og didoexy nucleotider bliver forkert, hvilket gør at der er for få dideoxynucleotider, og at der derfor laves for få korte fragmenter.



10. Forklar ud fra figuren ovenfor, hvad der sker når bindingsbufferen tilsættes (A), og hvorfor det er vigtigt at bufferen både indeholder meget salt (NaCl) og har en høj pH, dvs lavt indhold af H^+ ioner. Silica pulveret indeholder negativt ladede oxygenatomer. Når der tilsættes en saltopløsning, vil frie Na^+ ioner tiltrækkes til de negativt ladede oxygenatomer. Det gør at silica materialet får en positiv overflade pga at det dækkes med Na^+ ioner. Denne positive overflade binder de negativt ladede fosfatgrupper i DNA backbone, og holder dem fast i søjlematerialet. Det er vigtigt at opløsningen er basisk, da det betyder at der er en meget lav koncentration af H^+ (H_3O^+) ioner i væsken, der ellers ville binde sig stærkere til de negative oxygen atomer end Na^+ ionerne.
11. Forklar igen ud fra figuren ovenfor, hvad der sker når elueringsbufferen tilsættes (B), og hvorfor det er vigtigt at vand er et polært molekyle. Når elueringsbufferen tilsættes vil vandets oxygen atomer, der jo har en delvist negativ ladning pga ladningsforskydning, sætte sig omkring Na^+ ionerne og derved bryde den ionbinding, der har holdt DNA'et bundet til silica gelen. Det delvist positive H-atomer vil sætte sig ved de negativt ladede fosfatgrupper og dække dem. Da der er meget mere vand end Na^+ ioner vil effekten blive at DNA frigives ud i vandet.
12. Overvej, hvorfor det er smart at vende søjlen på hovedet inden at elueringsbufferen tilsættes? Når DNA bindes i søjlematerialet vil det begynde med at binde til overfladen og når det er fyldt op med DNA vil der langsomt blive fyldt mere og mere op igennem søjlen. Så det højeste koncentration af DNA vil være når overfladen. Ved at vende søjlen på hovedet sikrer man sig at der er kortest muligt vej ud af søjlematerialet for mest muligt at DNA'et.

Cycle sequencing PCR

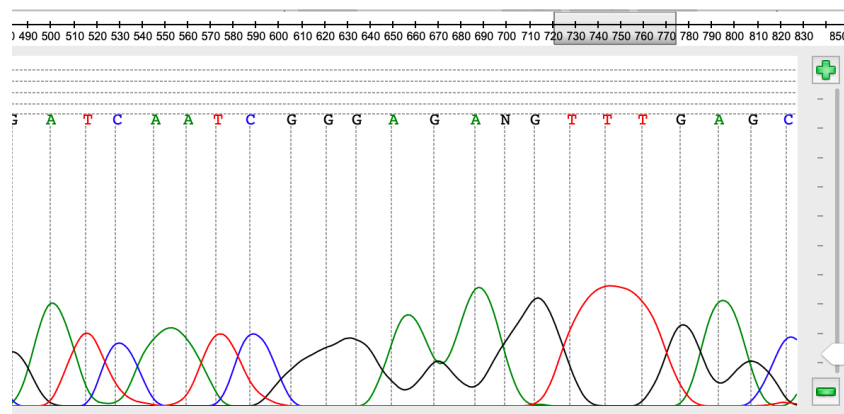
1. Hvad er formålet med at lave en cycle sequencing reaction? At lave DNA fragmenter af forskellig længde der hver især er mærket med en fluorescerende markør, der er specifik for den base, som fragmentet ender med.
2. Overvej hvorfor der kun tilsættes 1 primer per PCR cycle sequencing reaktion? **Primeren fungerer som begyndelsessted for DNA replikationen. Hvis der var tilsat to primere til reaktionen ville det betyde at der var to startsteder for DNA replikationen. Derfor ville der blive dannet to fragmenter af hver længde, der mange steder ville ende med forskellige baser, og det ville blive umuligt at aflæse hvilken base der er ved hvilken DNA position.**
13. Forklar, hvorfor at DNA polymerasen ikke kan sætte flere DNA nucleotider på efter at der er sat et DNA dideoxy nucleotid på den voksende DNA strengs kopi. **Fordi at DNA polymerasen katalyserer en kondensationsreaktion mellem -OH gruppen på carbonatom nummer 3 på det sidst tilkoblede DNA nucleotid og det næste DNA nucleotid triphosphats ene fosfatgruppe. På dideoxy nucleotider mangler -OH gruppen på carbon atom nummer 3, og denne reaktion kan derfor ikke foregå.**
14. Hvorfor skal man bruge flere cycle sequencing reaktioner med hver sin primer, for at lave en DNA sekventering af et DNA stykke på ca. 1500 basepars længde? **Fordi at hver DNA sekvens (read) man får ud af en Sanger sekventering kun er ca. 700 basepar lang. Der skal altså bruges flere DNA sekvenser for at dække hele den 1500 bp i den ønskede sekvens.**

DNA sekventering

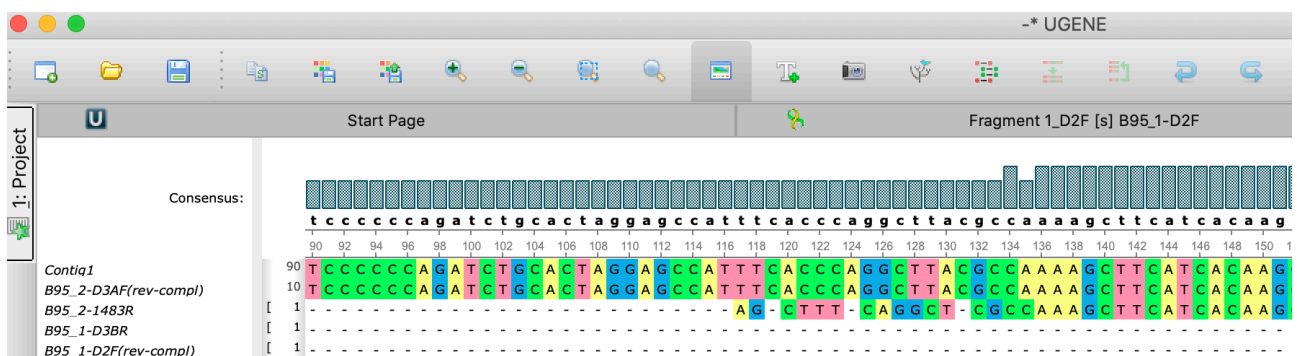
15. Hvorfor vandrer DNA, når der sættes strøm igennem en gel? **Fordi fosfat grupperne i DNA er negativt ladet, de vandrer derfor mod pluspolen.**
16. Hvorfor afskilles DNA fragmenterne af forskellige længde, når de vandrer igennem kapillær gelelektroforese gelen? **Fordi der er større modstand fra DNA fragmenter mod at blive trukket igennem gelen jo længere de er.**
17. Hvordan bestemmer Sekventerings maskinen, hvilken DNA base der afslutter et specifikt DNA fragment. **Ved at belyse DNA fragmenterne med en laser, der anslår den fluorescerende markør, og derved skaber et fluorescens signal. Farven på signalet er specifikt for de enkelte baser, da det er forskellige fluorescerende molekyler man har koblet på de specifikke baser i dideoxynucleotiderne.**
18. Hvorfor er det svært at få bestemt DNA sekvensen ved meget korte og meget lange DNA fragmenter? **Små fragmenter vandrer meget hurtigt og passerer derfor også laseren meget hurtigt. Laser detektoren kan derfor ikke nå at adskille alle de forskellige korte DNA fragmenters fluorescens, som kommer til at overlappe og blive utydelige. For de lange fragmenters vedkommende er det fordi at fragmenterne til slut bliver så store at de bevæger sig meget langsomt. DNA fragmenterne bliver meget lang tid om at passere laseren og signalerne bliver derfor trukket meget ud og det bliver meget svært at afgøre om det er en eller flere DNA fragmenter der er passeret.**

Aktivitet 2. Samling af lange DNA sekvenser.

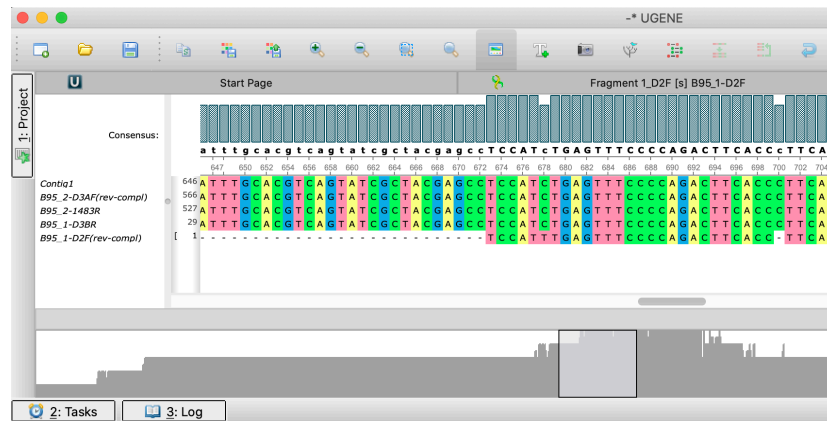
1. Hvilke farver har A, G, C og T basernes kromatogramtoppe? **A=grøn, G=sort, T=rød og C=blå**
2. Se på position 766 i DNA sekvensen. Baseparret er angivet med et N, der angiver at det kan være en hvilken som helst af de 4 DNA baser. Forklar ud fra kromatogrammet, hvorfor basen er angivet med et N? **Fordi det ikke kan aflæses om det er et A eller G, derfor angives det at der er en base med et N, men at man ikke ved hvad det er for en.**



3. Hvor godt matcher de to sekvenser hinanden i den første del af overlappet? **Ikke særlig godt, der er usikkerhed i sekvensen fragment 4 1483R er ikke god i begyndelsen, og derfor matcher den ikke perfekt op. Se man på dele af det bestemte kan man godt se at det ikke er helst skævt, men pga manglende aflæsning af baser eller ekstra aflæste baser kan man ikke matche det helt med fragment 3 D3AF fragmentet.**



4. Det område, man kan være mest sikker på er der hvor flest sekvenser overlapper. I hvilket sekvensområde er det? **Området, hvor der er flest sekvenser der overlapper er mellem 673 til ca. 850. Det kan ses grafisk på konsensus diagrammet for neden.**



Aktivitet 3. DNA database søgning med BLAST (basic local alignment search tool)

1. Skriv hvilke 6 bakterier I har identificeret og hvilken type prøve de kom fra.

Prøve 1: Fra lymfевæske. *Bartonella henselae*

Prøve 2: Afførings prøve. *Eschericia coli*

Prøve 3: Urin prøve. *Pseudomonas aeruginosa*

Prøve 4: Blod prøve. *Salmonella typhimurium* (ved BLAST search står den som en variant af *Salmonella enterica* var. *Typhimurium*, hvilket kan forvirre eleverne).

Prøve 5: Spyt prøve *Yersenia pestis*

Prøve 6: Afførings prøve. *Yersinia enterocolitica*

2. Skriv et kort referat af hvilke sygdomme og hvilke symptomer man har, hvis man er inficeret med bakterierne.

Prøve 1: *Bartonella henselae* - giver Cat Scratch Disease, på dansk kattekradsesyge. Patienter med et dårligt immunforsvar kan risikere blodforgiftning af infektion med bakterien.

Prøve 2. *Eschericia coli* - patogene stammer af denne tarmbakterie kan give diarré til patienter.

Prøve 3. *Pseudomonas aeruginosa* - en bakterie, der er kendt for at kunne vokse i meget forskellige miljøer. Svækkede patienter, f.eks patienter med cystisk fibrose, kan få alvorlige infektioner af denne bakterie i f.eks. luftvejene men også i urinvejene.

Prøve 4. *Salmonella typhimurium*. *Salmonella* bakterier er ofte patogene, og kan f.eks. give tyfus, en sygdom der tidligere medførte mange dødsfald. De mest almindelige salmonella forgiftninger er i tarmen, hvor bakteriesporer kan overleve turen forbi syrebadet i tarmen, og derved inficerer tynd- og tyktarm. Herfra kan de nogle gange inficere blodbanerne og lymfesystemet.

Prøve 5. *Yersenia pestis* - den bakterie der er skyld i sygdommen pest, der i middelalderen var skyld i massedød i Europa. Inficerer via loppebid, og spreder sig via blod og lymfekar, hvor den får lymfeknuder til at hæve voldsomt op. Kan også sætte sig i lungerne. Især sidstnævnte er meget dødelig, hvis det ikke bliver behandlet i tiden.

Prøve 6. *Yersinia enterocolitica* - en tarmbakterie, der naturligt findes i hos mange husdyr. Via kød og mejeriprodukter kan den spredes til mennesker, især fordi bakterien kan vokse ved 4°C.



Symtomer er diarré og voldsomme mavesmerter, hvor man først tror at det kunne være blindtarmsbetændelse. Kan i sjældne tilfælde give infektion i led.

3. Overvej, hvorfor det kan være et problem hvis man skal dyrke bakterierne før man kan kortlægge deres DNA. Fordi at der er mange bakterier der ikke kan dyrkes på et af de almindeligt brugte vækstmedier. Derfor er det ikke sikkert at man får bakterievækst selvom der er en bakterie i prøven.
4. Hvilke udfordringer tror I det kan det give at bruge universelle primere til at opformere DNA fra bakterieprøver? Universelle primere kan opformere DNA fra næsten alle bakterier, derfor vil forurening af mastermixet med f.eks luftbårne bakterier kunne risikere at give en forkert DNA sekvens og derfor i sidste ende en forkert diagnose. Det vil især være et problem fra f.eks afføringsprøver, da der er en meget stor bakterieflora i tarmen.