



En nytt multiplex PCR-protokoll för
identifiering och detektion av *Shigella* och
enteroinvasiv *E. coli* (EIEC) från livsmedel

Sofie Altgård, Sofia Berggren, Viktor Björklund, Sara
Lundsten, Thorsteinn Olafsson, Lovisa Pettersson

Beställare: Livsmedelsverket
Beställarrepresentant: Jan Andersson
Handledare: Karin Stensjö

SAMMANFATTNING

En av Livsmedelsverkets (SLV) uppgifter är kunna identifiera bakterien *Shigella* vid kontamination av livsmedel. Dagens detektion av *Shigella* är problematisk då enbart 10-100 bakterier ger sjukdomssymptom, vilket kräver hög precision och selektivitet av analysmetoderna. Projektet utformades i syfte att hjälpa SLV utveckla sina metoder för att detektera och odla *Shigella*.

Shigella är uppdelad i fyra serogrupper med flera serotyper i respektive grupp. Denna indelning är inte baserad på genetiskt ursprung. Projektet har lett fram till en ny detektionsmetod där man utnyttjar en multiplex PCR för att parallellt amplifiera ett antal gener som är unika för varje serogrupp av *Shigella*. Metoden gör det även möjligt att skilja *Shigella* från den mycket närbesläktade enteroinvasiva *E. coli* (EIEC). Möjligheterna att kunna odla *Shigella* selektivt på en platta har också granskats, men inget som direkt går att applicera har upptäckts. Dock diskuteras en del metaboliska reaktionsvägar som kan utnyttjas för odling.

Rapportens föreslagna metod löser i dagsläget detektion av serogrupper och i viss mån EIEC, men mer sekvensdata behövs för att säkerställa och utveckla resultatet ytterligare. Slutligen diskuteras även andra framtida lösningar på detektionsproblematiken.

Innehållsförteckning

INLEDNING	4
Bakgrund	4
Problemställning.....	5
Syfte och frågeställning.....	6
Avgränsningar	6
GENOMFÖRANDE OCH RESULTAT.....	7
Litteratursökning	7
Dataanalys	8
Genanalys	10
Primer- och probdesign	11
DISKUSSION.....	13
Sammanfattning av diskussion.....	13
En multiplex PCR rekommenderas	13
Vilka gener gav en bra detektion?.....	14
Så bra kommer den rekommenderade PCR-metoden kunna detektera <i>Shigella</i>	17
Så kan PCR-metoden utvecklas	17
I framtiden kan PCR-metoden ersättas med Mikroarray.....	17
Odlingsplattor och metaboliska reaktionsvägar	17
SLUTSATS OCH REKOMMENDATION.....	19
REFERENSER	21
BILAGA 1 – Genom undersökta i dataanalysen.....	24
BILAGA 2-7 – Sammanfattning av artiklar	25
BILAGA 8 - Lista på gener som undersöktes i dataanalysen.....	31
BILAGA 9 - Beskrivning av gener.....	32
BILAGA 10 - primer- och probdesign.....	34
BILAGA 11 - Primers och prober	35

INLEDNING

Bakgrund

Shigella är en gram-negativ bakterie som orsakar sjukdomen shigellos, en typ av bakteriell dysenteri, med utmärkande symptom så som blodig eller vattnig diarré. Bakterien kan hittas i mänsklig avföring och spridningen sker antingen genom direktkontakt med smittkällan, eller genom en kontaminerad källa (1). Trots att shigellos inte är så vanligt i Sverige finns det i dagsläget ändå ett behov av att snabbt kunna detektera *Shigella* då endast 10-100 bakterier behövs för att en person ska insjukna och få märkbara symptom (2). Eftersom förhållandevis få bakterier kan göra stor skada behöver man kunna arbeta med mycket små mängder, vilket försvårar detektion och isolering av *Shigella* från livsmedel och vatten (3).

Shigella delas numera in i fyra olika serogrupper (*S. dysenteriae*, *S. flexneri*, *S. sonnei* och *S. boydii*) efter att tidigare ha klassificerats som arter. Serogrupperna varierar i patogenicitet beroende på dess virulensplasmid, där *S. dysenteriae* och *S. flexneri* orsakar allvarligare symptom än de två andra (4). Indelningen i serogrupperna är gjord utifrån serologiska metoder, vilket betyder att det är ytstrukturen som avgör vilken serogrupp som bakterien tillhör. De olika serogrupperna har även utvecklats flertalet gånger från *Escherichia coli* (*E. coli*) genom historien. I deras genom är det därför svårt att hitta gemensamma mönster som kan användas för att stödja gruppindelningen. Detta gör det svårt att använda molekylärbiologiska metoder för att skilja mellan de olika grupperna.

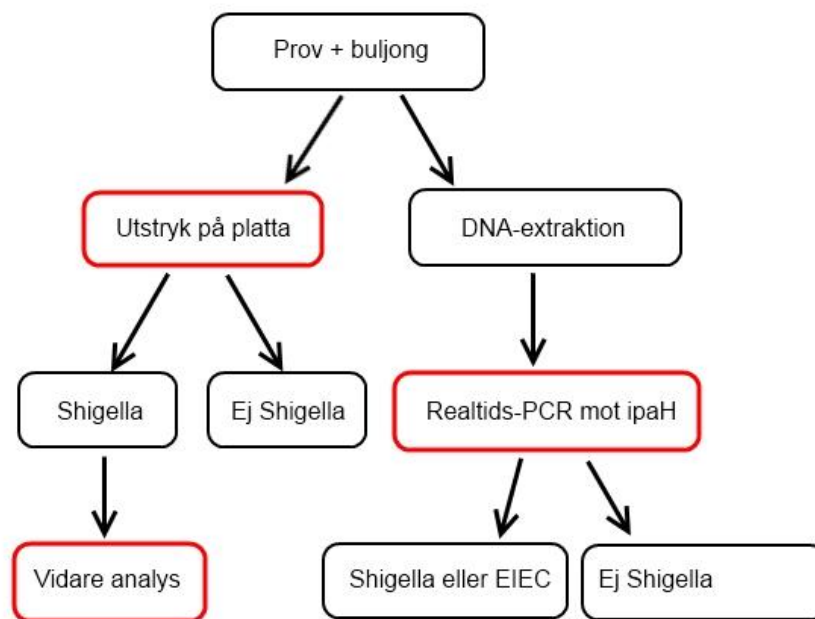
En annan sak som är problematisk vid detektion och tillbakaspårning är att serogrupperna är uppdelade i ett antal serotyper som även dessa skiljer sig åt (5). Det är även värt att notera att *Shigella* har ett dynamiskt genom vilket gör att det blir svårare att hitta gener som kan användas för detektion. Arten applicerar ofta så kallade "pathoadaptive mutations", vilket innebär att den inaktiverar gener för att bli en så effektiv patogen som möjligt. Dessa inaktiveringar kan ske på olika sätt inom gruppen. Ett vanligt fenomen är insättning av så kallade IS-element, en sekvens som sätts in mitt i en gen och stör dess öppna läsram (6).

Ytterligare ett problem som kan uppstå vid detektionen är att många unika gener sitter på bakteriens plasmid. Problemet med plasmider är att *Shigella* kan tappa dessa, exempelvis vid förvaring eller när de befinner sig i buljong (7).

Shigella är mycket närbesläktad med *E. coli*. De båda bakterierna är så pass lika genetiskt sett att *Shigella* skulle kunna klassas som en egen art inom släktet *Escherichia*. Det finns en specifik art, enteroinvasiv *E. coli* (EIEC), som är särskilt lik *Shigella*. En enteroinvasiv bakterie smittar genom att leva inuti en annan värdcell och utnyttja dess resurser (8). Även EIEC orsakar dysenteri och använder sig av en invasionsplasmid som är väldigt lik den hos *Shigella* (3). Liksom *Shigella* delas EIEC upp i bakterier med liknande egenskaper snarare än bakterier med samma ursprung (3). Det är dock viktigt att kunna skilja mellan EIEC och *Shigella* för att kunna spåra en kontaminerad källa, exempelvis vid ett bakterieutbrott.

Problemställning

Livsmedelsverket (SLV) har som uppgift vid bakterieutbrott att lokalisera smittkällan samt artbestämma den bakterie som orsakat utbrottet. I de fallen där *Shigella* tros vara orsaken till utbrottet tillkommer vissa svårigheter när det kommer till att odla och detektera denna bakterie. Detta beror dels på att bakterien förekommer i väldigt låga halter i livsmedel och att de har svårt att konkurrera med andra arter utanför deras naturliga miljö. Nedan visas en schematisk bild på den rutin som Livsmedelsverket använder idag för att detektera *Shigella*.



Figur 1 Denna figur beskriver Livsmedelsverkets nuvarande rutiner för att detektera *Shigella*. Varje ruta representerar ett steg i denna kedja. Det som är rödmarkerat betecknar möjliga förbättringsområden inom metoden (7).

Dagens metod har vissa brister. För att skilja de olika serogrupperna åt används idag immunologiska metoder (9), något som är väldigt tidskrävande och dessutom kräver en växande kultur. Eftersom att det vid bakterieutbrott är av stor vikt att snabbt identifiera smittkällan för att förhindra vidare spridning behöver denna metod utvecklas. Att gå över till molekylärbiologiska metoder kan vara mer tidseffektivt, men försvåras av *Shigellas* enorma diversitet och det faktum att de fyra grupperna är kategoriserade efter serologiska egenskaper (9).

De odlingsplattor som finns idag är inte tillräckligt selektiva för *Shigella*. I och med att ofta finns väldigt få *Shigella* i det livsmedel som orsakat smittan utkonkurreras den av andra bakterier som finns i större mängder. *Shigella*-bakterier blir helt enkelt överväxt av andra bakterier på odlingsplattan och det blir omöjligt att isolera en individuell koloni. I och med att nästan all vidare identifiering är beroende av en växande kultur försvåras arbetet kraftigt av detta problem. Skulle mer högkvalitativa odlingsmedier, som hämmar tillväxt av bakgrundsflora, tas fram finns det stora möjligheter att lösa dessa problem.

Syfte och frågeställning

Syftet med projektet var att underlätta Livsmedelsverkets detektion av *Shigella* från livsmedel, för att snabbt motverka vidare spridning av bakterien. På längre sikt kan projektet leda till en snabbare diagnostisering av patienter som blivit smittade av kontaminerat livsmedel och även en mer precis behandlingsmetod, då man vet vilken serogrupp som orsakat smittan. Även forskare utanför detta projekt som arbetar med *Shigella* och EIEC kan ha nytta av de kunskaper och metoder som projektet genererar.

Utifrån syftet utformades tre projektmål enligt följande prioritetsordning:

1. Ta fram en unik egenskap som skiljer *Shigella* från EIEC och föreslå en molekylärteknisk metod som kan appliceras på denna för att skilja de två bakteriestammarna åt.
2. Presentera ett antal egenskaper som på vilka en molekylärteknisk metod kan utformas för att särskilja serogrupperna hos *Shigella*.
3. Presentera en metod för att odla och isolera *Shigella*, en metod som innefattar ett förslag på en odlingsplatta som kan användas. Detta mål är något lägre prioriterat än de andra.

Avgränsningar

Eftersom projektet var rent teoretiskt utfördes inga laborationer för att bekräfta att de metoder som tagits fram fungerar i praktiken. En avgränsning blev därmed att det enbart lades fram resonemang och förslag. Mest tid lades på att uppnå delmål 1.

Ett krav från beställaren var att de framtagna metoderna skulle vara tidseffektiva och inom rimliga ekonomiska gränser. I praktiken betydde det en Polymerase Chain Reaction (PCR) på bakteriellt DNA. Projektet avgränsades till att undersöka framtagna multiplexa PCR-metoder, det vill säga en PCR där flera gener amplifieras samtidigt, samt vilka gener som kan tänkas användas i dessa reaktioner.

GENOMFÖRANDE OCH RESULTAT

Genomförandefasen innehöll tre större delar. Dels utfördes en omfattande litteratursökning, för att bland annat samla information om *Shigella*, hitta redan implementerade metoder för att skilja serogrupperna åt samt hitta nya odlingsplattor som SLV kan använda sig av. För att ytterligare undersöka de gener som hittades i litteraturen utfördes en dataanalys. Till sist sammanställdes och utvärderades den information som framkommit under respektive moment.

Litteratursökning

Metod

Informationssökningen var den period då projektgruppen läste och samlade på sig grundläggande information om ämnet. I arbetsuppgifterna innefattades att hitta och läsa vetenskapliga artiklar, sammanställa Livsmedelsverkets rutiner samt att undersöka vilka gener som ansågs vara av intresse. Inledningsvis delades informationssökningen upp utefter projektets tre mål.

Vid sammanställning av olika artiklar sågs ett tydligt mönster, nämligen att PCR var en relativt utstuderad och bra metod för att skilja mellan *Shigellas* serogrupper och i viss mån även EIEC. Vidare läsning av artiklar innefattade därför multiplex PCR samt vilka gener som kan vara av användning i en PCR för att på bästa sätt kunna uppnå projektets två högst prioriterade mål.

Under litteraturstudien undersöktes även metoder för odling av *Shigella*. Den mest lovande odlingsplattan som hittades under litteraturstudien (se resultat) utvärderades genom att undersöka hur ekonomiskt hållbart detta alternativ skulle vara. Även metaboliska reaktionsvägar, som en ny odlingsplatta skulle kunna baseras på, undersöktes.

Resultat av litteratursökning

Vid litteraturstudiens slut sammanfattades artiklar som ansågs relevanta för projektet. Nedan följer en kort beskrivning av dessa, för mer information se bilaga 2-7.

De artiklar som sammanfattas i tidigare nämnda bilagor tar upp genetiska skillnader mellan *Shigellas* serogrupper samt skillnader mellan *Shigella* och *E. coli*-arten (däribland EIEC). De gener som valdes ut för vidare dataanalys, och från vilken artikel dessa har tagits ifrån, kan ses i bilaga 8.

I bilaga 2, 3 och 6 sammanfattas de artiklar som behandlar olika metaboliska reaktionsvägar och odlingsplattor. En analys av dessa artiklar resulterade i två alternativ för selektion av *Shigella*. Alternativ ett är odlingsplattan Rainbow Agar, se bilaga 2. Enligt denna artikel ska plattan vara mer selektiv än odlingsmediumet XLD (10), som i dagsläget är en av de plattor som används av SLV. Plattans innehåll är ej publicerat och produceras enligt SLV endast i USA.

Alternativ två är den metaboliska reaktionsvägen för produktion av NAD, se bilaga 6. Enligt artikeln som sammanfattas i bilagan har *Shigella* stängt av vissa delar i denna metabolisk reaktionsväg vilket resulterar i att *Shigella* inte producerar NAD utan tillförsel av substrat, något som både *E. coli* och EIEC kan (6).

Dataanalys

Metod

För att få ökad förståelse om de genetiska likheter och skillnader mellan *Shigella*, EIEC och andra bakterier genomfördes en omfattande genomsökning på de gener som valdes ut under litteratursökningen, se bilaga 8. Genomen undersöktes med hjälp av NCBI:s BLAST, främst med verktyget tblastn.

Det första som utfördes var att hitta och sammanställa genom från olika serotyper av *Shigella*, EIEC samt patogena *E. coli*. För att se vilka genom som användes, se bilaga 1. Gensekvensen togs sedan fram med hjälp av NCBI:s graphic-verktyg eller GenBank (<https://www.ncbi.nlm.nih.gov>). En multipel sekvensinpassning utfördes med hjälp av BLAST mellan generna mot alla listade genom för att avgöra om genen fanns hos någon av dem, samt i vilken grad av likhet. Slutligen uträttades även en sekvensinpassning av genen mot hela NCBI:s databas för att se vilka eventuella andra organismer som hade liknande gener. Om den sista sökningen resulterade i nästan uteslutande *Shigella* så gjordes sökningen om, med ändringen att *Shigella* uteslöts så att andra arter kunde upptäckas.

Resultat av dataanalys

Dataanalysen gav en bild av hur de utvalda generna förekommer i olika arter. De gener som enligt dataanalysen var unika för en art sammanfattas i tabell 1 nedan. En lista på de sekvenser som användes under dataanalysen återfinns i bilaga 1.

Tabell 1 Sekvenslikhet mellan gener hos *Shigella* och EIEC. Procentsatserna anger hur stor längd av de jämförda sekvenserna som överensstämmer med den ursprungliga genen. De rödmarkerade rutorna är genen som BLAST:en utgått från.

“Art” / Gen	<i>cadA</i>	<i>shuA</i>	<i>iroN</i>	<i>rfbE</i>	<i>rfc</i>	<i>gspM</i>	<i>lacY</i>	<i>sitA</i>	<i>ipaH7</i>	<i>nadA</i>
<i>S. dysenteriae</i>	25%	100%	100%	-	-	95%	38%	100%	99%	32%
<i>S. sonnei</i>	100%	-	-	-	-	-	21%	100%	99%	100%
<i>S. flexneri</i>	25%	-	-	100%	100%	-	-	100%	100%	32%
<i>S. boydii</i>	25%	-	-	-	-	100%	-	100%	99%	32%
EIEC	25%	-	-	-	-	99%	Ej data	-	99%	32%
Övriga patogena <i>E. coli</i>	25%	EHEC 99%	-	-	-	ETEC 96%	Alla <i>E. coli</i> : 99-100%	APEC 99%	-	ETEC 38%
Övriga <i>E. coli</i>				Två st 99%	Tre st 99%	Flertal 97-99%				67%

(-) Sekvenslikheten var så pass låg att den anses vara irrelevant.

Genanalys

Metod

Efter dataanalysen utnyttjades tabell 1 för att undersöka de gener som var av störst intresse för projektet. Dessa gener utvärderades genom att läsa fler artiklar för att säkerställa att resultatet från dataanalysen kunde styrkas i litteraturen. Fokus låg på att kontrollera att generna som föreslagits var unika för en eller flera serogrupper, för EIEC eller *Shigella* i stort.

Resultat av genanalys

Genen *lacY* valdes för att den inte finns i *Shigella*-genus och kan därför användas för att särskilja EIEC från *Shigella* (11). Genen *shuA* valdes då den kan användas för att detektera den farligaste serotypen hos *Shigella*, *S. dysenteriae* serotyp 1 (12). Genen *gspM* valdes för att kunna detektera *S. boydii* och *S. dysenteriae*. Enligt dataanalys finns *gspM*-genen hos båda dessa serogrupper men har ett antal sekvensolikheter som kan utnyttjas för skilja dessa åt. Genen *rfc* valdes för att identifiera *S. flexneri* då denna sekvens är specifik nog för denna serogrupp (13). Eftersom *ipaH* genen är unik för EIEC och *Shigella*, och redan används för att särskilja dessa arter från andra bakterier (7), valdes det att även här använda denna. För mer information om varje gens funktion se bilaga 9.

Tabell 2 nedan sammanfattar de gener som användes som underlag vid primerdesign och vilken serogrupp eller art som identifieras.

Tabell 2 Detektion av arter med hjälp av unika gener för respektive genom.

Art / gen	<i>lacY</i>	<i>gspM1</i>	<i>gspM2</i>	<i>shuA</i>	<i>rfc</i>	<i>wzx</i>	<i>ipaH</i>
<i>S. dysenteriae</i>	-	+	-	+	-	-	+
<i>S. boydii</i>	-	+	+	-	-	-	+
<i>S. flexneri</i>	-	-	-	-	+	-	+
<i>S. sonnei</i>	-	-	-	-	-	+	+
EIEC	+	-	-	-	-	-	+

Gener som finns hos en art betecknas med (+), annars (-)

(*) *shuA* finns endast hos *S. Dysenteriae* typ 1

Primer- och probdesign

Metod

I slutet av projektet utfördes också en primer- och prob-design i programmet CLC Main Workbench 7.0.2. I och med att dessa sekvenser skulle användas i samma reaktion vid multiplex PCR fanns det många olika parametrar som behövde optimeras för att skapa fungerande primers. Bland annat behövde storleken på amplikonen ha olika storlek, för att kunna separeras på en agarosgel. Smälttemperaturen hos alla primers behövde också vara väldigt lika för att alla gener ska kunna amplifieras. För en mer specifik beskrivning av vilka parametrar som togs i beaktande under primer- och probdesignen, se bilaga 10.

Resultat av primer- och probdesign

Primers och prober har designats för att kunna detektera specifika gensekvenser hos *Shigella* genus, EIEC samt de olika serogrupperna av *Shigella* (tabell 3). För mer utförlig information om dessa sekvenser, se bilaga 11.

Tabell 3 Primers och prober designade i detta projekt. Tabellen visar de primers och prober som har designats i syftet att detektera de olika serogrupperna av *Shigella* samt *Shigella*-genus och EIEC. Tabellen innehåller amplikonstorlek, sekvens (5'-3') och GC-innehåll för respektive sekvens.

Detekterar art	Namn	Sekvens	Amplikonlängd
<i>Shigella</i> , EIEC	F_ <i>ipaH</i>	GCGTTTCTATGGCGTGTGTC	196 bp
	R_ <i>ipaH</i>	GTACTCATTCTCCAGCATCTC	
	Prob_ <i>ipaH</i>	CCTCCGCACTGCCGAAGCCA	
<i>S. boydii</i>	F_ <i>gspM1</i>	GGATAGTACAACAGGAGACAATG	322 bp
	R_ <i>gspM1</i>	ACATTCACCATCCCAGTCT	
	Prob_ <i>gspM1</i>	CCGCGCACTCGGTAGTCATCAAGC	
<i>S. boydii</i> , <i>S. dysenteriae</i>	F_ <i>gspM2</i>	TTAAGCCGTAGCGAACACT	122 bp
	R_ <i>gspM2</i>	TGCCAGATGAGGACGTAATAT	
	Prob_ <i>gspM2</i>	ACCTGGCGGGGCGTTCGC	
<i>E. coli</i> (EIEC)	F_ <i>lacY</i>	AGAAGCAGAAACAGACCAGAT	163 bp
	R_ <i>lacY</i>	ATTATTGGCTCATCGTTCGC	
	Prob_ <i>lacY</i>	CAGCCCACCAGCAGGAACGGTAC	
<i>S. flexneri</i>	F_ <i>rfc</i>	GGGACGAAAAAAGAACTCGC	281 bp
	R_ <i>rfc</i>	TTTATGGCTTCTTTGTTCGGC	
	Prob_ <i>rfc</i>	AAAAAGCCCCATCGGGACGAAAAAAGAAC	
<i>S. dysenteriae</i> typ 1	F_ <i>shuA</i>	TTGGCTTTGGCTGTTTCTG	79 bp
	R_ <i>shuA</i>	GCATTCCCCGTTGCCGTAA	
	Prob_ <i>shuA</i>	CCACCTTGCCAACGTTTGCTTTTGCT	
<i>S. sonnei</i>	F_ <i>wzx</i>	GCATCTATTCGTAGGGATATTG	-
	R_ <i>wzx</i>	CACACCAATACAATACGATTTAC	
	Prob_ <i>wzx</i>	CGCTCGCTTTGATTTCCACTACAC	

(-) Information saknas

(*) Har ej kunnat estimerats på samma vis som de andra primrarna

DISKUSSION

Sammanfattning av diskussion

Första målet - skilja *Shigella* från EIEC

Med ett rent prov skulle det vara möjligt att särskilja på EIEC med hjälp av den föreslagna PCR-metoden, där genen *lacY* amplifieras och ger positivt utslag för EIEC men negativt för *Shigella*. *lacY* tillsammans med *ipaH* kan bekräfta EIEC, men i dagsläget endast om provet med säkerhet är rent. Problem kan uppstå när en blandad bakterieflora analyseras i PCR, då primern som konstruerats till *lacY* fungerar hos ett stort antal *E. coli* och inte är specifik för EIEC. Metoden skulle ge ett falskt positivt utslag om provet till exempel innehåller en icke-patogen *E. coli* och en *Shigella*. I det EIEC-genom som användes vid dataanalysen hittades inte *lacY*-genen men den ska vara intakt hos EIEC (14). För att förbättra resultaten bör fler EIEC-genom sekvenseras för att sedan kunna designa en EIEC-specifik primer.

Andra målet - skilja mellan serogrupper

Med den PCR-metod som har föreslagits för att skilja på *Shigellas* serogrupper har målet som sattes uppnåtts. Primers och prober har designats på ett sätt som i teorin visar hög specificitet för de referensgenom som finns och kan på så sätt identifiera varje serogrupp. För genernas funktion och förekomst finns även vetenskapligt stöd, se bilaga 9. På grund av stora genomiska skillnader inom serogrupperna är det inte säkert att metoden kommer att fungera på alla serotyper, eftersom det inte finns sekvenserade genom för alla serotyper. Ett bra exempel på det är *S. flexneri* typ 6 vars genom inte finns sekvenserat och påstås vara en serotyp som skiljer sig från de andra inom serogruppen *flexneri* (15).

Tredje målet - att finna en mer selektiv odlingsplatta

I rekommendationen finns ej ett konkret förslag på en platta, det bästa förslaget hade redan testats och ansetts vara undermåligt av SLV. Litteraturstudien visade att det i forskningsvärlden inte läggs ner tillräckliga resurser på att hitta en bra och selektiv platta för *Shigella*. Det finns utrymme för sådant arbete i och med att den kartläggning av *Shigella*-genom som utförs ger underlag för studier av metaboliska vägar. Att studera metaboliska vägar kan vara ett sätt att lösa problemet, men det faktum att *Shigella* stänger av många metaboliska vägar gör att den har svårt att upptäckas bland annan bakgrundsflora.

En multiplex PCR rekommenderas

Utifrån en artikel från 2013 (13) erhöles en bra bild av hur ett multiplext PCR-protokoll kan utformas. Med utgångspunkt ur denna skapades en idé om hur man skulle kunna göra något liknande, men i en förbättrad version. Projektgruppen valde därför att en multiplex PCR ska rekommenderas till SLV.

För SLV är det redan en etablerad metod att använda PCR för detektion av *Shigella*. Vidare är multiplex PCR utifrån en ekonomisk aspekt hållbar och möter det krav som projektet har fått från beställaren. Kraven var att metoden skulle vara molekylärteknisk, helst någon form av PCR. Dessutom kan man med metoden spara in tid eftersom att flera tester kan göras i samma provrör. Under informationsökningen konstaterades också att dagens forskning är väldigt inriktad på att lösa problemet med hjälp av PCR-tekniker. Om problemet kan lösas helt med PCR, och om det fungerar

bra även vid rena eller blandade prover så kan man i framtiden minska användandet av odling på plattor. Detta kan i sin tur leda till att användandet av antibiotika för att selektera för *Shigella* kan minska, vilket kan medföra att även resistensen för antibiotika minskar.

Som tidigare nämnts har primer- och prob-designen gjorts för att passa en multiplex PCR. I kombination med multiplex PCR valdes också realtids-detektion med Taqman® -prober. Valet att även utgå från en realtids-PCR berodde på att detta är vad SLV använder sig av idag. Övergången från dagens metod till den rekommenderade metoden blir då så enkel som möjligt. Realtids-PCR tillför möjligheten att studera förloppet under körningen. Eftersom att alla primers har designats så att amplikonerna har olika storlek kommer också dessa kunna kontrolleras med gelelektrofores efter körningen.

PCR kan dessvärre även amplifiera gener från andra organismer. För att lösa detta problem har primerna och proberna i största möjliga utsträckning, designats för att endast binda till och amplifiera generna från de eftersökta arterna eller serogrupperna. De primrar och prober som designats är än så länge bara teoretiska, därför går det inte med säkerhet att säga att alla kommer att fungera perfekt. Många av de primers som detta arbete resulterat i har dock delar som är komplementära till ytterligare sekvenser i andra, icke önskvärda, arter av bakterier. Ingen utav dessa har dock en sekvenslikhet på 100 %. Det är därför viktigt att testa dessa i praktiken och utföra kontroller för att med säkerhet kunna säga att de är tillräckligt specifika.

Vad gäller tid och transportsträckor bör ett hållbart tankesätt appliceras för att minska kostnader och miljöpåverkan. Det finns en del svenska företag som designar och tillverkar primers och prober, detta kan utnyttjas för att spara in långa transporter. Om en multiplex PCR kan detektera alla serotyper direkt hos SLV behöver man heller ej skicka vidare *Shigella*-bakterier för serotypning.

Vilka gener gav en bra detektion?

De gener som undersöktes under dataanalysen valdes utifrån de artiklar som studerats under informationssökningen, se bilaga 2-7. Vid en närmare jämförelse av de resultat som erhållits från litteraturen, den bioinformatiska dataanalysen och designen av primers och prober, finns det både likheter och skillnader i vilka gener som föreslås lämpa sig för PCR:en. På grund av att *Shigella* kan tappa sin plasmid ansåg därför projektgruppen att de gener som skulle rekommenderas för en PCR skulle, i den mån det gick, sitta på *Shigellas* kromosom.

Nedan följer en diskussion runt de gener som behandlats under projektets gång. Som tidigare nämnts, kan mer grundläggande information om genernas funktion och placering i genomet hittas i bilaga 9.

Gener för att särskilja serogrupper

S. flexneri: rfbE och rfc

rfbE och *rfc* är båda gener som kan identifiera *S. flexneri*, enligt den dataanalys som genomfördes och två artiklar (15,16). Dessa gener hittades nästan bara i *S. flexneri* med undantag för två respektive tre icke-patogena *E. coli* (17). *rfc* rekommenderas istället för *rfbE* eftersom primern för *rfc*-genen passar bättre tillsammans med de andra primrarna. Den framtagna primern för *rfbE* var ej passande eftersom den hade lägre smälttemperatur gentemot de andra.

S. dysenteriae typ 1: shuA och iroN

För att kunna identifiera *S. dysenteriae* typ 1 hittades två passande gener, *shuA* och *iroN*, se tabell 1. Båda generna skulle i teorin fungera lika bra för att identifiera *S. dysenteriae* typ 1. *shuA*-genen återfinns hos EHEC (se tabell 1), men den är mycket mer välstuderad än *iroN* och därför rekommenderas användning av *shuA* framför *iroN*. SLV kan idag detektera *S. dysenteriae* med hjälp av en gen som heter *stx1* (7). Det är värt att notera att den primer som SLV använder för *stx1* ej är testad i den primerkombination som rekommenderas. Ska denna fortsätta användas krävs det att amplikonlängd och andra aktuella parametrar undersöks.

S. boydii och S. dysenteriae: gspM

För att kunna detektera *S. boydii* och *S. dysenteriae* utformades två primers för genen *gspM*. Denna gen går att hitta i båda serogrupperna och en del andra bakterier (se tabell 1), vilket inledningsvis utgjorde ett stort problem. Efter att ha sekvensinpassat dessa kunde dock vissa olikheter urskiljas i sekvenserna och därför designades två set av primers för *gspM*, se tabell 3 och 6. Ett primerset designades för att kunna amplifiera en viss region hos båda serogrupperna där sekvenserna överrensstämde. Det andra primersetet utformades så att endast *gspM* i *S. boydii* kunde amplifieras. Har man *S. boydii* i provet kommer det alltså resultera i två olika fragment av specifika storlekar. Om istället *S. dysenteriae* finns i provet kommer det resultera i ett fragment av en specifik storlek. Dessa primers ska även kunna detektera en *S. dysenteriae* som inte är av typ 1.

För detektion av *S. dysenteriae* så kan *gspM* dock vara problematisk då dataanalysen visade att den är en pseudogen och har en stor mängd mutationer som skiljer den från *gspM* i *S. boydii*. Detta tyder på att genen inte är under selektivt tryck och kan komma att undergå rekombination eller mutation.

S. sonnei: wzx

Projektgruppen valde att rekommendera genen *wzx* som idag redan används av SLV (7). Detta beror på att de två andra gener som undersökts, *cadA* och *nadA*, visade sig skapa en viss problematik, se diskussion om respektive gen nedan. Sekvenserna för de primers som SLV använder kunde dock inte hittas bland genomen som undersökts vilket gjorde att ingen amplikonstorlek kunde uppskattas. Eftersom detektionen av denna gen redan sker via realtids-PCR är troligtvis amplikonstorleken inte ett problem på samma sätt här. Däremot är det inte fastställt om amplikonstorleken hos *wzx* överlappar med någon av de andra amplikonen i den föreslagna metoden.

S. sonnei: cadA

Analyserna i BLAST indikerade att genen *cadA* skulle fungera bra för att särskilja *S. sonnei* vilket även bekräftades i litteratursökningen (se tabell 1). Vid försök att göra primer till denna gen visade det sig dessvärre att genen ej var lämplig att använda. *cadA* innehåller framförallt många IS-element och primern riskerar att passa hos både *S. flexneri* och *S. sonnei*, vilket alltså betyder att det ej med säkerhet skulle gå att skilja mellan de två serogrupperna.

S. sonnei: nadA

Genen *nadA* ska enligt Di Martino et al. (6) finnas i alla serogrupper av *Shigella*. På grund av ett IS-element ska dock amplikonstorleken hos *S. sonnei* skilja sig markant från resterande serogrupper. Problemet som då uppstod var att detta amplikon uppskattades vara så pass stort att det inte skulle fungera i en realtids-PCR. För att kunna detektera detta amplikon med gelelektrofores krävs en stor stege där de resterande generna inte skulle kunna skiljas åt. Därför har inga primers eller prober designats för denna gen, men den skulle kunna användas för att skilja *S. sonnei* från övriga *Shigella*.

Gener för att särskilja *Shigella* och EIEC

Shigella* och EIEC: *ipaH

Den dataanalys som utfördes visade att *ipaH*-genen är unik för *Shigella* och EIEC. Detta var förväntat, eftersom genen används flitigt för att identifiera de två grupperna i de artiklar som studerats. Resultat var en bra referens när de resterande analyserna bedömdes och diskuterades, eftersom det visade hur ett nästintill idealt BLAST-resultat såg ut.

EIEC: *LacY*

Det är bevisat att de flesta *Shigella* ej kan fermentera laktos och forskning tyder på att de har en muterad *lacY*-gen (11,12), alternativt saknar den helt. EIEC och många andra *E. coli* kan däremot fermentera laktos och detta tyder på sekvensskillnader mellan de som kan och inte kan fermentera laktos (14). Därför valdes *lacY* som markör för att skilja på *Shigella* och EIEC.

Eftersom *lacY* ej hittades i det EIEC-genom som fanns tillgängligt vid dataanalysen har primrar och prober designats utifrån ett annat genom, *E. coli* K12. Detta betyder att primern ej är specifik för enbart EIEC och medför, till skillnad från resterande primers, problem vid detektion i blandad bakterieflora. Det finns dessutom ingen garanti att primern binder till *lacY* i EIEC. För att få metoden så specifik som möjligt skulle *lacY*-genen hos fler stammar av EIEC behöva sekvenseras.

Shigella: SitA

Vad gäller genen *sitA* tydde dataanalysen på ett mycket lovande resultat för att skilja *Shigella* från EIEC, då genen inte fanns hos den senare, se tabell 1. Det fanns dock en osäkerhet i denna slutsats, då endast ett EIEC-genom användes i den bioinformatiska dataanalysen. En artikel från 2003 (18) tyder på att *sitA* kan finnas hos EIEC. På grund av detta tas *sitA*-genen ej med i rekommendationen.

Så bra kommer den rekommenderade PCR-metoden kunna detektera *Shigella*

Den föreslagna multiplexa PCR-metoden ska i teorin kunna skilja mellan *Shigellas* olika serogrupper i både rena och orena prover. Alla nydesignade primers, förutom *lacY* ska vara så specifika att en blandad bakterieflora, i teorin, inte ger något problem. Som nämnt ovan är primern för *lacY* designad utifrån ett *E. coli*-genom och är därför inte specifik för enbart EIEC, något som hade underlättat detektion i blandad flora.

Eftersom EIEC inte är uppdelad utifrån genetiskt ursprung är inte en stam av EIEC representativ för alla typer av EIEC. Då det endast finns ett sekvenserat genom för EIEC försvåras uppgiften ytterligare. Det skulle behövas fler sekvenseringar av deras diversa genom för att med säkerhet kunna särskilja EIEC med hjälp av molekylärtekniska metoder.

Så kan PCR-metoden utvecklas

Bristen på sekvenserade genom i kombination med en begränsad mängd information inom ämnet var den största begränsande faktorn för projektet. Eftersom uppdelningen i serogrupper inte är gjord utifrån fylogenetiskt ursprung krävs det sekvenser av samtliga serotyper inom varje grupp för att säkerställa att metoden fungerar på alla olika typer av *Shigella*. Det behövs dessutom mer forskning om hur grupperna är relaterade till varandra, det kanske rent av finns genetiska mönster som stärker den serologiska indelningen. Det finns ett tydligt syfte för att bedriva mer forskning i området, eftersom shigellos i många länder är ett utbrett problem.

I framtiden kan PCR-metoden ersättas med Mikroarray

Det finns redan nu mer utvecklade, men samtidigt dyrare metoder för att göra en mer specifik detektion av *Shigella*. En artikel från 2009 (19) visar att mikroarray går att tillämpa för att identifiera enskilda serotyper. För mer information om detektion med microarrayer, se bilaga 5. Som det ser ut idag är det inte ett lämpligt alternativ för SLV. Denna forskning visar dock att det finns sekvenser som är unika hos varje serotyp som på sikt kan medföra bättre och snabbare metoder för identifikation.

Odlingsplattor och metaboliska reaktionsvägar

Projektet hade som mål att finna en ny odlingsplatta som är mer selektiv för *Shigella* än nuvarande alternativ. Livsmedelsverkets tester av denna platta visade att den ej var tillräckligt selektiv för *Shigella* som Livsmedelsverket önskar. Dessutom innebär det faktum att plattan måste importeras från USA att kostnaden för plattan blir alldeles för hög.

Projektgruppen vill kunna presentera en hållbar lösning som kan användas under en längre tid. Rainbow agar kan optimeras genom att tillsätta antibiotika, men det finner vi inte som en acceptabel lösning. Att öka antibiotika-användande ökar risken för resistens hos bakterier och det är något som bör undvikas. Därför rekommenderas Livsmedelsverket att inte använda Rainbow Agar

Eftersom *Shigella* har som karaktärsdrag att stänga av många metaboliska processer för att bli mer patogen är detta ett övergripande problem när man letar efter processer som är unika för *Shigella*. En metabolisk väg av intresse är den tidigare beskrivna produktionen av NAD. Denna metaboliska process är dock något som inte är intakt hos *Shigella*, vilket gör att det orsakar negativ selektion.

Shigella kommer att dö medan övriga bakterier som innehar den metaboliska reaktionsvägen överlever.

Lösningen på detta odlingsproblem skulle istället kunna vara att rikta in sig på den anrikning som sker innan *Shigella* börjar odlas på plattor. Skulle den kunna vara mer selektiv och döda fler andra bakterier skulle kanske samma plattor användas för att få ett bättre resultat. Detta är dock inget som projektgruppen hittat lösningar på.

SLUTSATS OCH REKOMMENDATION

För att kunna detektera *Shigella*, dess olika serogrupper samt EIEC föreslås nedanstående PCR-protokoll. Reaktionen kan förslagsvis köras i två steg som tabellerna indikerar, där det i första steget fastslås om provet innehåller *Shigella* eller EIEC. Genom att inkludera *shuA* är det även möjligt att snabbt fastslå om källan till smittan är den farligaste serotypen, *S. dysenteriae* typ 1. I nästa steg identifieras vilken serogrupp som provet innehåller.

Tabell 4 : Denna tabell beskriver PCR-protokoll 1. Amplikonstorlek anges i antal baspar (bp) efter respektive primer. Varje x representerar en positiv detektion. Tillhörande primer- och probsekvenser hittas i bilaga 9.

Primer / Gen	<i>ipaH</i> (196 bp)	<i>shuA</i> (80 bp)	<i>lacY</i> (163 bp)
<i>S. dysenteriae</i> typ 1	X	X	
<i>S. dysenteriae</i> typ 2-6	X		
<i>S. boydii</i>	X		
<i>S. sonnei</i>	X		
<i>S. flexneri</i>	X		
EIEC	X		X

Tabell 5 : Denna tabell beskriver PCR-protokoll 2. Amplikonstorlek anges i antal baspar (bp) efter respektive primer. Varje x representerar en positiv detektion. Om inget amplikon (efter körning) och inget ljus (under körning) detekteras och tidigare PCR-protokoll har visat på närvaro av *ipaH*-genen, antas provet innehålla EIEC. Tillhörande primer- och probsekvenser återfinns i tabell 3.

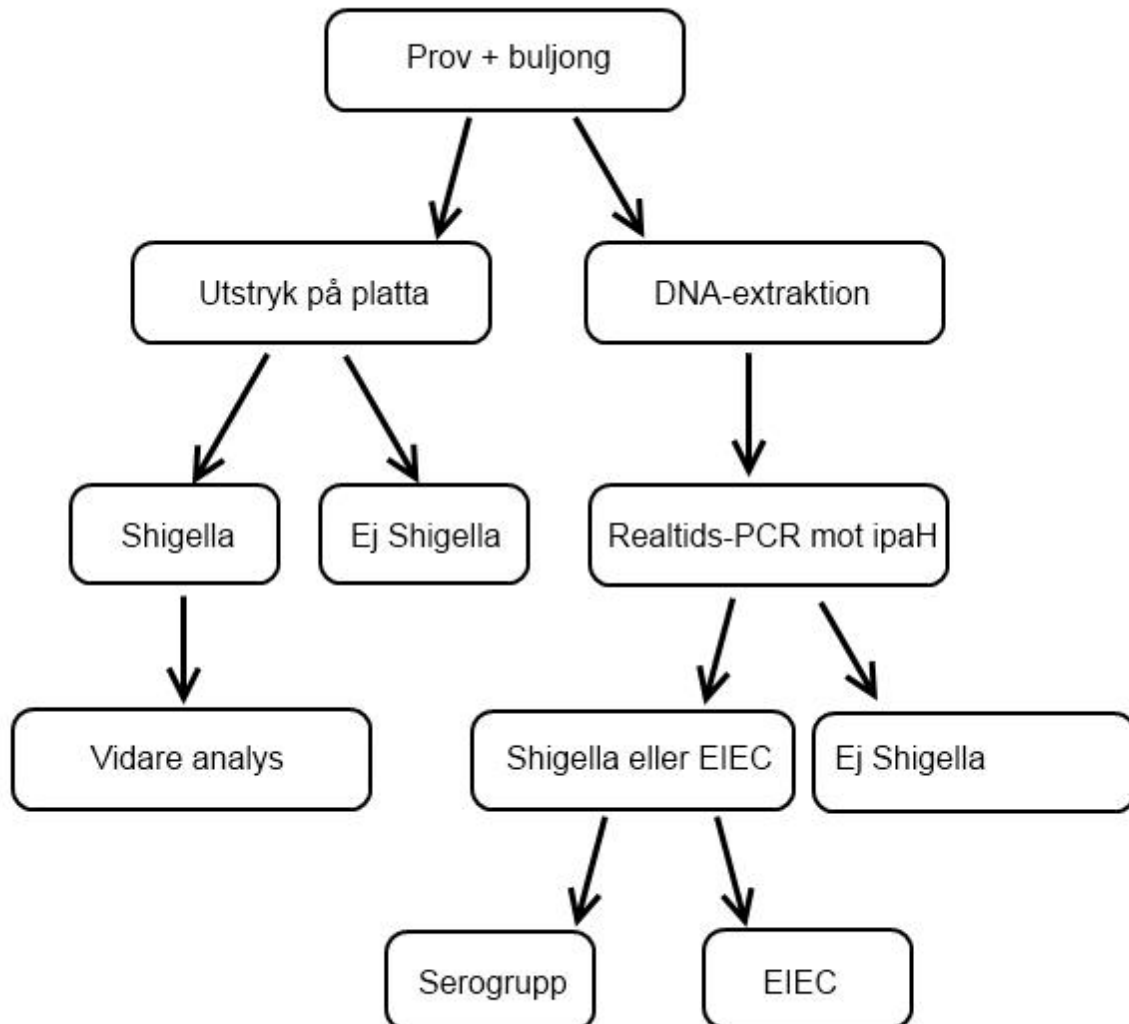
Primer / Gen	<i>gspM 1</i> (322 bp)	<i>gspM 2</i> (122 bp)	<i>wzx</i> (* bp)	<i>rfc</i> (281 bp)
<i>S. boydii</i>	X	X		
<i>S. dysenteriae</i>		X		
<i>S. sonnei</i>			X	
<i>S. flexneri</i>				X
EIEC				

(*) Uppgift saknas

Vid identifiering av *S. sonnei* har *wzx*-genen föreslagits men detta är ingen optimal lösning. PCR på denna gen, som redan används av SLV idag, kan appliceras med viss modifikation för att passa det nya protokollet. En annan möjlig lösning är att använda *nadA* som målgen, eftersom forskning visat att *S. sonnei* har ett IS-element insatt i genen vilket gör att den är större än hos andra serogrupper (6).

Primers är framtagna med syftet att vara specifika för just den serogrupp eller art som eftersöks och ska i teorin därför fungera i blandade prover. För att verifiera detta behövs testförsök. En komplett PCR behöver dessutom positiva, negativa samt interna kontroller. I dagsläget utför SLV positiva kontroller med stammar av *S. sonnei* och *S. boydii* från smittskyddsinstitutet, något som kan appliceras som positiv kontroll på den nya PCR som föreslås. Fler stammar från andra serogrupper bör också användas för bästa möjliga säkerhet.

De framtagna metoderna kan implementeras i ett uppdaterat flödesschema, baserat på Livsmedelsverkets nuvarande system:



Figur 2: Denna figur beskriver en föreslagen rutin för att detektera *Shigella*. Denna är en uppdaterad version av Livsmedelsverkets nuvarande rutiner (7). Med hjälp av den föreslagna metoden ska det nu gå att skilja *Shigella* och EIEC åt.

Vad gäller en ny odlingsplatta har projektet ej lett fram till någon förbättrad lösning. På grund av *Shigellas* förmåga att stänga av många metaboliska vägar har arten begränsad möjlighet att leva utanför sin naturliga miljö. En ideal platta skulle behöva imitera den näringsrika cytoplasman hos en tarmcell, men samtidigt vara så selektiv att andra bakteriearter inte klarar av att leva där. I dagsläget finns det ingen platta som kan möta dessa krav.

REFERENSER

1. Bhattacharya D, Bhattacharya H, Thamizhmani R, Sayi DS, Reesu R, Anwesh M, et al. Shigellosis in Bay of Bengal Islands, India: clinical and seasonal patterns, surveillance of antibiotic susceptibility patterns, and molecular characterization of multidrug-resistant *Shigella* strains isolated during a 6-year period from 2006 to 2011. *Eur J Clin Microbiol Infect Dis Off Publ Eur Soc Clin Microbiol*. 2014 Feb;33(2):157–70.
2. Hayford AE, Mammel MK, Lacher DW, Brown EW. Single nucleotide polymorphism (SNP)-based differentiation of *Shigella* isolates by pyrosequencing. *Infect Genet Evol J Mol Epidemiol Evol Genet Infect Dis*. 2011 Oct;11(7):1761–8.
3. Binet R, Deer DM, Uhlfelder SJ. Rapid detection of *Shigella* and enteroinvasive *Escherichia coli* in produce enrichments by a conventional multiplex PCR assay. *Food Microbiol*. 2014 Jun;40:48–54.
4. Ram PK, Crump JA, Gupta SK, Miller MA, Mintz ED. Part II. Analysis of data gaps pertaining to *Shigella* infections in low and medium human development index countries, 1984-2005. *Epidemiol Infect*. 2008 May;136(5):577–603.
5. Echeverria P, Sethabutr O, Pitarangsi C. Microbiology and diagnosis of infections with *Shigella* and enteroinvasive *Escherichia coli*. *Rev Infect Dis*. 1991 Apr;13 Suppl 4:S220–S225.
6. Di Martino ML, Fioravanti R, Barbabella G, Prosseda G, Colonna B, Casalino M. Molecular evolution of the nicotinic acid requirement within the *Shigella*/EIEC pathotype. *Int J Med Microbiol IJMM*. 2013 Dec;303(8):651–61.
7. Livsmedelsverk. Livsmedelsverkets rutiner för detektion av *Shigella*. Livsmedelsverk;
8. Van den Beld MJC, Reubsæet FAG. Differentiation between *Shigella*, enteroinvasive *Escherichia coli* (EIEC) and noninvasive *Escherichia coli*. *Eur J Clin Microbiol Infect Dis Off Publ Eur Soc Clin Microbiol*. 2012 Jun;31(6):899–904.
9. Civilingenjörsprogrammet i molekylär bioteknik. Projektbeställning för Självständigt arbete i molekylär bioteknik. 2014.
10. Zhang G, Lampel KA. Comparison of chromogenic Biolog Rainbow agar *Shigella*/Aeromonas with xylose lysine desoxycholate agar for isolation and detection of *Shigella* spp. from foods. *J Food Prot*. 2010 Aug;73(8):1458–65.
11. Ito H, Kido N, Arakawa Y, Ohta M, Sugiyama T, Kato N. Possible mechanisms underlying the slow lactose fermentation phenotype in *Shigella* spp. *Appl Environ Microbiol*. 1991 Oct;57(10):2912–7.
12. Yang F, Yang J, Zhang X, Chen L, Jiang Y, Yan Y, et al. Genome dynamics and diversity of *Shigella* species, the etiologic agents of bacillary dysentery. *Nucleic Acids Res*. 2005;33(19):6445–58.
13. Ojha SC, Yean Yean C, Ismail A, Singh K-KB. A pentaplex PCR assay for the detection and differentiation of *Shigella* species. *BioMed Res Int*. 2013;2013:412370.
14. Pavlovic M, Luze A, Konrad R, Berger A, Sing A, Busch U, et al. Development of a duplex real-time PCR for differentiation between *E. coli* and *Shigella* spp. *J Appl Microbiol*. 2011 May;110(5):1245–51.

15. Sun Q, Lan R, Wang Y, Zhao A, Zhang S, Wang J, et al. Development of a multiplex PCR assay targeting O-antigen modification genes for molecular serotyping of *Shigella flexneri*. J Clin Microbiol. 2011 Nov;49(11):3766–70.
16. *rfc* [Internet]. NCBI. [cited 2014 May 15]. Available from: <http://www.ncbi.nlm.nih.gov/ezproxy.its.uu.se/gene/1078521>
17. Morona R, Mavris M, Fallarino A, Manning PA. Characterization of the *rfc* region of *Shigella flexneri*. J Bacteriol. 1994 Feb;176(3):733–47.
18. Runyen-Janecky LJ, Reeves SA, Gonzales EG, Payne SM. Contribution of the *Shigella flexneri* Sit, Iuc, and Feo *iroN* acquisition systems to *iroN* acquisition in vitro and in cultured cells. Infect Immun. 2003 Apr;71(4):1919–28.
19. Li Y, Cao B, Liu B, Liu D, Gao Q, Peng X, et al. Molecular detection of all 34 distinct O-antigen forms of *Shigella*. J Med Microbiol. 2009 Jan;58(Pt 1):69–81.
20. Pupo GM, Lan R, Reeves PR. Multiple independent origins of *Shigella* clones of Escherichia coli and convergent evolution of many of their characteristics. Proc Natl Acad Sci. 2000 Sep 12;97(19):10567–72.
21. Lan R, Alles MC, Donohoe K, Martinez MB, Reeves PR. Molecular evolutionary relationships of enteroinvasive Escherichia coli and *Shigella* spp. Infect Immun. 2004 Sep;72(9):5080–8.
22. Nicotinate and nicotinamide metabolism - *Shigella flexneri* [Internet]. KEGG. [cited 2014 May 15]. Available from: http://www.kegg.jp/kegg-bin/highlight_pathway?scale=1.0&map=sfx00760&keyword=nad
23. Farfán MJ, Garay TA, Prado CA, Filliol I, Ulloa MT, Toro CS. A new multiplex PCR for differential identification of *Shigella flexneri* and *Shigella sonnei* and detection of *Shigella* virulence determinants. Epidemiol Infect. 2010 Apr;138(4):525–33.
24. Ashida H, Toyotome T, Nagai T, Sasakawa C. *Shigella* chromosomal *IpaH* proteins are secreted via the type III secretion system and act as effectors. Mol Microbiol. 2007 Feb;63(3):680–93.
25. Serdiuk T, Madej MG, Sugihara J, Kawamura S, Mari SA, Kaback HR, et al. Substrate-induced changes in the structural properties of *LacY*. Proc Natl Acad Sci U S A. 2014 Apr 22;111(16):E1571–1580.
26. Casalino M, Latella MC, Prosseda G, Ceccarini P, Grimont F, Colonna B. Molecular evolution of the lysine decarboxylase-defective phenotype in *Shigella sonnei*. Int J Med Microbiol IJMM. 2005 Mar;294(8):503–12.
27. Casalino M, Prosseda G, Barbagallo M, Iacobino A, Ceccarini P, Latella MC, et al. Interference of the CadC regulator in the arginine-dependent acid resistance system of *Shigella* and enteroinvasive E. coli. Int J Med Microbiol IJMM. 2010 Jun;300(5):289–95.
28. Mills M, Payne SM. Identification of *shuA*, the gene encoding the heme receptor of *Shigella dysenteriae*, and analysis of invasion and intracellular multiplication of a *shuA* mutant. Infect Immun. 1997 Dec;65(12):5358–63.
29. Lallemand M, Login FH, Guschinskaya N, Pineau C, Effantin G, Robert X, et al. Dynamic Interplay between the Periplasmic and Transmembrane Domains of GspL and GspM in the Type II Secretion System. Cascales E, editor. PLoS ONE. 2013 Nov 1;8(11):e79562.

30. *rfbE* [Internet]. NCBI. [cited 2014 May 15]. Available from: <http://www.ncbi.nlm.nih.gov.ezproxy.its.uu.se/gene/1078524>
31. Top Ten Pitfalls in Quantitative Real time PCR Primer Probe Design and Use [Internet]. Life Technologies. [cited 2014 May 14]. Available from: <http://www.lifetechnologies.com/se/en/home/references/ambion-tech-support/rtpcr-analysis/general-articles/top-ten-pitfalls-in-quantitative-real-time-pcr-primer.html>
32. TaqMan® Probes | TaqMan® probe design tips | TaqMan® design tutorial [Internet]. Premier Biosoft. [cited 2014 May 14]. Available from: http://www.premierbiosoft.com/tech_notes/TaqMan.html
33. GC Content Percentage [Internet]. nfstc. [cited 2014 May 14]. Available from: http://www.nfstc.org/pdi/Subject04/pdi_s04_m01_02_f.htm

BILAGA 1 – GENOM UNDERSÖKTA I DATAANALYSEN

Följande genom, eller delar av genom, användes under sökningarna i BLAST. GI-numrena listas samt vilken bakterie det representerar.

***Shigella* - Kromosom**

- gi|73854091 - *S. sonnei*
- gi|81239530 - *S. dysenteriae* typ 1
- gi|81244029 - *S. boydii* typ 4
- gi|344915202 - *S. flexneri* typ 2a
- gi|110804074 - *S. flexneri* typ 5
- gi|281599365 - *S. flexneri* typ X
- gi|187427012 - *S. boydii* typ 18
- gi|374269796 - *S. sonnei*

***Shigella* - Virulensplasmid**

- gi|73858315 - *S. sonnei*
- gi|81243805 - *S. dysenteriae* typ 1
- gi|81248166 - *S. boydii* typ 4
- gi|31983523 - *S. flexneri* typ 2a
- gi|384545966 - *S. flexneri* typ X
- gi|187426769 - *S. boydii* typ 18
- gi|374269797 - *S. sonnei*
- gi|58044990 - *S. flexneri* typ 5

EIEC

- gi|188501074 - virulensplasmid 53638
- gi|188491675 - kromosom 53638
- gi|188574176 - annan plasmid (75) 53638

Andra patogena *E. coli*

- gi|387610477 - ETEC H10407
- gi|215485161 - EPEC O127:H6 str. E2348/69
- gi|209395693 - EHEC O157:H7 str. EC4115
- gi|117622295 - APEC O1
- gi|595593103 - STEC, ST540

BILAGA 2 – SAMMANFATTNING AV COMPARISON OF CHROMOGENIC BIOLOG RAINBOW AGAR SHIGELLA/AEROMONAS WITH XYLOSE LYSINE DESOXYCHOLATE AGAR FOR ISOLATION AND DETECTION OF SHIGELLA SPP. FROM FOODS (10)

Artikeln presenterar en alternativ odlingsplatta som anses kunna isolera *Shigella* från andra bakterier. Plattan undersöks i artikeln genom att, utifrån tre olika experiment, jämföra dess kvalitet med redan etablerade odlingsplattor för denna tid (2010). Experimenten är utformade för att undersöka vilken av de testade plattorna som är mest selektiv för *Shigella*. En bra platta, enligt artikeln, ska kunna sålla bort onödig bakgrundsflora men samtidigt inte inhibera tillväxten av den efterfrågade bakterien.

I artikelns första experiment undersöktes de olika mediernas förmåga att inhibera kliniska isolat av *Shigella*. Totalt 18 olika medier testades varav 13 utav dessa var modifierad Rainbow Agar. Resultatet skiljde sig något mellan de fyra serogrupperna men visade ändå på att modifierad Rainbow Agar var den platta som gav bäst tillväxt av *Shigella*.

I artikelns andra experiment undersöktes de olika mediernas förmåga att detektera *Shigella* från kontaminerade livsmedel. Här skilde sig resultatet mer mellan de olika serogrupperna. Resultatet skilde sig även mycket åt beroende på vilket livsmedel som testades. För majoriteten tydde dock resultatet på att en modifierad Rainbow Agar var den minst inhiberande plattan.

I artikelns tredje experiment undersöktes de olika mediernas förmåga att inhibera bakgrundsflora. Även här tydde resultatet på att en modifierad Rainbow Agar var bättre än de andra testade plattorna. Det påpekas dock i artikeln att det var den specifika färg som *Shigella* fick på Rainbow Agarn som gjorde det lättare att urskilja denna från bakgrundsfloran.

Sammanfattningsvis tyder resultatet som presenteras i artikeln på att en Rainbow Agar med tillsatsen av 6.25 mg/L Streptomycin är det bäst isolerande odlingsmediet.

BILAGA 3 - SAMMANFATTNING AV GENOME DYNAMICS AND DIVERSITY OF SHIGELLA SPECIES, THE ETIOLOGIC AGENTS OF BACILLARY DYSENTERY (17)

För att få en bättre förståelse för de distinkta epidemiologiska och patologiska egenskaper som uppvisas hos *Shigellas* olika serogrupper utfördes en komplett sekvensering av genomen hos en serotyp inom varje serogrupp av *Shigella*. Genomen som sekvenserades var *S. dysenteriae* typ 1, *S. boydii* typ 4 och *S. sonnei*. Dessa jämfördes sedan med det tidigare sekvenserade genomet av *S. flexneri* typ 2a.

Tillsammans med en multilocus sekvensering (MLS) från Pupo et al. (20) och analys av de sekvenserade genomen så läggs starka bevis fram för att *Shigella* som art tillhör *E. coli*.

Shigella har samma replikationsorigin och -terminus som *E. coli* K12 stam MG1655 och har därför troligtvis samma replikationsmekanism som *E. coli*. En jämförelse av alla sekvenserade *Shigella*- och *E. coli*-genom som finns på GenBank visade att de delade ~3Mb gemensamt. Författarna undersökte även resultat från Lan et al. (21) där en MLS och en microarray användes för att visa att även EIEC har sitt ursprung från olika *E. coli*, och att de är mer besläktade med *Shigella* än andra patogena *E. coli*.

Shigellas genom har en stor diversitet vilket visas av distributionen av virulensgener. Många nya förmodade virulensgener identifierades som skulle kunna användas som mål för detektion eller behandlingsmetoder. *Shigellas* mångfald härstammar från horisontell genöverföring samt konvergent evolution som har drivits av den stora mängden IS-element som *Shigella* använder för att störa flertalet genfunktioner.

Författarna hävdar att genomens diversitet även kan ge stöd till hypotesen att *Shigella* har sitt ursprung från olika *E. coli*. Eftersom endast fyra genom har sekvenserats och *S. dysenteriae*- samt *S. sonnei*-genomen ligger utanför de tre stora fylogenetiska grupperna av *Shigella* så är detta dock inte en fullständig reflektion på mångfalden hos *Shigella*.

Då Shigellos är en av de största sjukdomarna som orsakar diarré i Kina så har sekvensering av *Shigella* fått stort stöd av regeringen. Författarna till denna artikel har varit med och byggt en hemsida vars mål är att ge det vetenskapliga samhället de verktyg som behövs för att utföra komparativa genomiska undersökningar av *Shigella*. Denna databas heter shiBASE (<http://www.mgc.ac.cn/ShiBASE>) och kan utnyttjas för framtida forskning om *Shigella*.

Genomen lades upp i GenBank och även på hemsidan shiBASE.

BILAGA 4 - SAMMANFATTNING AV POSSIBLE MECHANISMS UNDERLYING THE SLOW LACTOSE FERMENTATION PHENOTYPE IN SHIGELLA SPP. (11)

Förmågan att fermentera laktos är en egenskap som ofta används för att skilja *Shigella* från andra *E. coli*, eftersom i princip ingen serotyp av *Shigella* kan fermentera laktos. Hiteo et al. (11) utförde en studie för att genetiskt kartlägga lac-operonet hos ett antal serotyper av *Shigella* och jämföra dessa med *E. coli*.

Ett antal experiment utfördes för att undersöka operonet samt translaterat protein från lac-operonets tre gener; *lacA*, *lacY* och *lacZ*. Bland annat mättes aktivitet av β -galaktosidas (*lacZ*) som hydrolyserar laktos och laktospermeas (*lacY*) som sköter transport av laktos över membranet. För att undersöka generna utfördes Southern hybridization samt kloning av de motsvarande generna in i *E. coli* med fungerande laktosfermentation, där eventuella förändringar noterades.

Resultaten tyder på att de testade stammarna (*S. dysenteriae* typ 1, *S. flexneri* typ 1b och 3, *S. boydii* typ 2 och 4 samt tre stammar av *S. sonnei*) saknar identiska *lacY*-sekvenser med *E. coli*. Ingen hybridisering skedde mellan *E. coli* och någon *Shigella* förutom hos *S. sonnei*, som gav en något försvagad signal. Även *lacZ* gav liknande resultat, där endast gener från *S. sonnei* och *S. dysenteriae* typ 1 hybridiserade med *lacZ* från *E. coli*. Det är dessa serotyper som kan uppvisa svag laktosfermentation och orsaken till detta föreslås vara en intakt *lacZ*-gen men en skadad *lacY*-gen. I och med detta påverkas inte förmågan att hydrolysera laktos, men cellen kan inte transportera laktos över membranet.

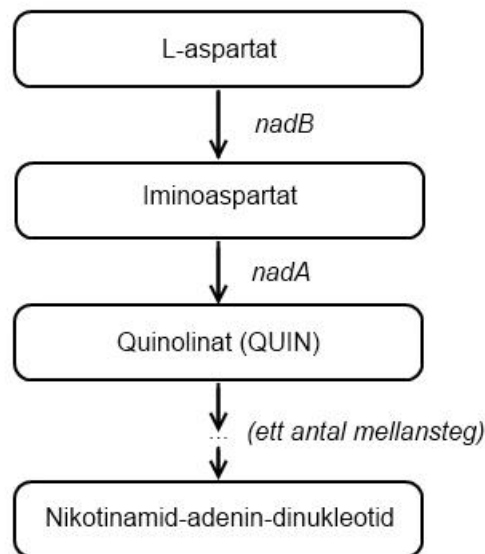
BILAGA 5 - SAMMANFATTNING AV MOLECULAR DETECTION OF ALL 34 DISTINCT O-ANTIGEN FORMS OF SHIGELLA (19)

I artikeln beskrivs en ny DNA mikroarray som designades för att snabbt kunna detektera de olika serotyperna hos *Shigella* från andra arter som är vanliga i livsmedel eller avföring. Mikroarrayen detekterar alla 34 distinkta O-antigenformer som finns hos *Shigella* och kan identifiera *S. boydii* typ 1-18, *S. dysenteriae* typ 1-13, *S. flexneri* typ 1-5 och 6 samt *S. sonnei*. *S. flexneri* typ 1-5 kan dock endast identifieras som en grupp då de delar till stor del samma egenskaper inom sina O-antigen. För att ta reda på den specifika serotypen krävs då fortfarande traditionell serotypning.

En mikroarray har många fördelar över nuvarande detektionsmetoder så som tidsaspekten då det kan göras under en dag. Vissa O-antigener är dock identiska eller nära besläktade hos *Shigella* och *E. coli*. För att kunna differentiera mellan dem så designades istället primers som kunde särskilja dem från varandra med olika amplikonstorlekar. När en av dessa O-antigener visar ett positivt resultat så krävs även en PCR-undersökning för att uppnå en säker identifiering .

BILAGA 6 - SAMMANFATTNING AV MOLECULAR EVOLUTION OF THE NICOTINIC ACID REQUIREMENTS WITHIN THE SHIGELLA/EIEC PATHOTYPE (6)

Nikotinamid-adenin-dinukleotid (NAD) är en kofaktor inblandad i ett antal kataboliska och anaboliska mekanismer. Hos *Shigella* är det observerat att NAD förhindrar intercellulär spridning och därmed minskar organismens patogenicitet. Bilden nedan visar hur bildandet av nikotinamid-adenin-dinukleotid går till och vilka andra ämnen som är inblandade i processen.



Figur 3 Illustration över bildandet av nikotinamid-adenin-dinukleotid (NAD) och i vilka mellansteg enzymen *nadA* och *nadB* är involverade (6,22).

Di Martino et al. genomförde 2013 försök för att undersöka huruvida olika stammar av *Shigella* och EIEC kunde växa i medium med och utan tillsats av nikotinsyra (en byggsten i NAD) eller QUIN, ett intermediat i bildandet av NAD. Det medium som användes var M9 minimal medium plate med 0,2% glukos samt med eller utan 10 µ/ml nikotinsyra eller 1 mM QUIN. Olika sorters antibiotika användes även vid behov. Resultat visade att stammarna av *Shigella* var strikt beroende av nikotinsyra eller QUIN, något som tyder på att gener som är viktiga för NAD-syntes har inaktiverats.

Vidare analys av de gensekvenser visade att inaktivering skett på olika sätt i de olika stammarna, med punktmutationer och/eller insättning av IS-element på skilda ställen i *nadA*- och *nadB*-generna. I *nadA*-genen hade alla serogrupper förutom *S. sonnei* punktmutationer som störde genens funktion. *S. sonnei* hade istället ett IS-element insatt i genen.

Resultat av inaktiveringen är att *Shigella* behöver tillskott av antingen nikotinamid-adenin-dinukleotid eller något intermediat som inte påverkas av de avstängda generna för att kunna växa i minimalt medium. De stammar av EIEC som undersöktes påvisade ej strikt beroende av varken niktotinsyra eller QUIN.

BILAGA 7 - SAMMANFATTNING AV A PENTAPLEX PCR ASSAY FOR THE DETECTION AND DIFFERENTIATION OF SHIGELLA SPECIES (12)

Artikeln presenterar en ny pentaplex PCR, där pentaplex syftar på att amplifieringen utförs med fem set av primers samtidigt och på så vis kan detektera förekomsten av fem gener samtidigt. PCR-metoden används för att särskilja *Shigellas* fyra serogrupper, och är enligt författarna till viss del framgångsrik. För detektion av serogrupperna används tre serogruppspecifika gener, en gen som är artspecifik samt en intern kontroll. Av dessa fem gener satt två stycken på *Shigellas* kromosom och resterande tre på dess invasionsplasmid. En av generna som användes är *rfc*, vilken i sin helhet är unik för *S. flexneri*.

Artikeln presenterade även lämpliga primers för alla gener författarna ville upptäcka i reaktionen. I diskussionen argumenterades det för att PCR:en även skulle kunna åtskilja *Shigella* och EIEC, vilket inte har gjorts tidigare. Dock kunde detta inte testas och då inte heller verifieras i studien.

BILAGA 8 - LISTA PÅ GENER SOM UNDERSÖKTES I DATAANALYSEN

Tabell 6 Lista på de gener som undersöktes under dataanalysen, tillsammans med information om vilken art de potentiellt kan detektera samt var information tagits från.

Gen	Kan användas på	Referens
<i>rfc</i>	<i>S. flexneri</i>	(13)
<i>rfpB</i>	<i>S. dysenteriae</i>	(12)
<i>she PAI</i>	<i>S. sonnei, S. flexneri</i>	(23)
<i>invC</i>	<i>Shigella genus</i>	(12)
<i>ipaH</i>	<i>Shigella genus, EIEC</i>	(24)
<i>wbgZ</i>	<i>S. sonnei</i>	(12)
<i>nadA</i>	<i>S. sonnei</i>	(6)
<i>uidA</i>	<i>Shigella genus, EIEC</i>	(14)
<i>wzx/wzy</i>	<i>S. sonnei</i>	(7)
<i>lacY</i>	<i>Shigella genus, EIEC</i>	(13)
<i>stxAB</i>	<i>S. dysenteriae</i>	(12)
<i>set1A, set1B</i>	<i>S. flexneri</i>	(12)
<i>iroN, iroBCDE</i>	<i>S. dysenteriae</i>	(12)
<i>shuA, shuS, shuTWXYUV</i>	<i>S. dysenteriae</i>	(12)
<i>gspC, gspM, gspL</i>	<i>S. dysenteriae, S. boydii</i>	(12)
<i>pic</i>	<i>S. flexneri, S. sonnei</i>	(12)

BILAGA 9 - BESKRIVNING AV GENER

Nedan följer en beskrivning av de gener som tas upp under resultatdelen i rapporten. För en tydlig överblick av sekvenslikheter mellan serogrupperna och olika arter, se tabell 1.

sitA

Genen *sitA* är involverad i järntransporten och är en del av *sit*-loket, vilket består av fyra gener som skapar ett specificerat transportsystem. Genen tros vara av vikt vid tillväxten i den intercellulära miljön. *sitA* kan hittas hos alla *Shigellas* serogrupper och flertalet EIEC-stammar, men genen finns inte hos andra patogena eller icke-patogena *E. Coli* (18). *sitA* är placerad på bakteriens kromosom (12).

lacY

Genen *lacY* kodar för proteinet laktospermeas som har en viktig roll i laktosfermentation. Proteinets transporterar socker över membranet mot protongradienten (25). *lacY*-genen existerar inte i *Shigella* men den finns i *E. coli*, däribland EIEC. *Shigella* är mestadels ej laktosfermenterande. Vissa serotyper av *S. sonnei* och *S. dysenteriae* kan fermentera laktos, men då väldigt långsamt. *LacY* är placerad på kromosomen (12).

CadA

CadA bildar tillsammans med bland annat *cadA* och *cadB* ett lokus som reglerar dekarboxylering av lysin. Resultatet blir en liten polyamid som heter kadaverin (26). I både *Shigella* och EIEC är *cad*-loket avstängt och *cadA* klassas därför som en pseudogen (26,27). Genen existerar i sin helhet, enligt analyserna i BLAST, enbart i *S. sonnei*.

shuA

Genen vid namn *shuA* kodar för ett ytprotein hos *Shigella* som har en receptorfunktion gentemot hem-grupper. *ShuA* är lokaliserat på kromosomen och är enligt de genomförda BLAST-sökningarna, (28) samt (12) unikt för serogruppen *S. dysenteriae*, primärt serotyp 1. Genen förekommer dock i vissa patogena *E. coli*-arter, som exempelvis EPEC, och en handfull andra gramnegativa bakterier.

iroN

Liksom *shuA* är *iroN* central i järntransport hos *S. dysenteriae*. *iroN* bildar tillsammans med andra *iro*-gener ett lokus som kodar för proteiner som huvudsakligen agerar receptorer för järnbaserade molekyler. Enligt (12) existerar *iroN* inom *Shigella*-familjen enbart i *S. dysenteriae* typ 1, vilket även bekräftas av BLAST-sökningarna.

GspM

Genen *gspM* är involverad i typ II sekretionssystemet (T2SS) tillsammans med *gspC* och *gspL* (29) T2SS är ett system för att transportera veckade proteiner genom yttermembranet hos gramnegativa bakterier, så som *Shigella*. *GspM* finns fullt funktionell i *S. boydii* men även som pseudogen i några serotyper av *S. dysenteriae*. När *gspM* från *S. boydii* BLASTades för att hitta liknande sekvenser visade det sig att många olika typer av *E. coli* innehåller gener med nästan identisk sekvens.

RfbE/wzx

rfbE är en gen som kodar för ett polysackarid-biosyntes-protein (30). Genen kallas även *wzx* och hittas i alla serotyper av *S. flexneri* förutom serogrupp 6 (15). Dataanalys visade att det finns stora sekvenslikheter mellan *S. flexneri* och två patogena *E. coli*. Genen *wzx* återfinns även hos *S. sonnei*, men har en annorlunda sekvens. SLV använder idag *wzx* för att identifiera *S. sonnei*. Hos *S. sonnei* sitter genen på plasmiden.

rfc

Genen *rfc* kodar för ett O-antigenpolymeras och återfinns i *S. flexneri* (16), på dess kromosom (13). BLASTning av genen resulterade i olika serotyper av *S. flexneri*, och även i detta fall tre icke-patogena *E. coli* (17). Genen sitter på kromosomen (13).

ipaH

ipaH är genen som kodar för invasionsplasmid-antigen H, vilket är involverat i att hämma inflammation hos den organism som *Shigella* infekterar. Genen finns i ett antal kopior i *Shigella*'s genom, både på kromosomen och invasionsplasmiden (24). Den används idag av Livsmedelsverket för att identifiera *Shigella* och EIEC, eftersom genen är unik för de två grupperna (enligt dataanalys utförd i detta projekt).

nadA

Genen *nadA* är inblandad i nikotinamid-adenin-dinukleotid (NAD)-syntes, där den katalyserar kondensation av iminoaspartat och dihydroxiacetonfosfat. *Shigella* har ej en fungerande NAD-syntes, något som hos vissa serotyper beror på en ej fungerande *nadA*-gen (6).

BILAGA 10 - PRIMER- OCH PROBDESIGN

Vid primer- och probdesign krävdes det att alla dessa sekvenser ska fungera i närvaro av varandra under en PCR. Primers och prober har utformats från de genom som samlades in under dataanalysen, se bilaga 1. Vid utvecklingen av specifika primers och prober togs ett antal parametrar och rekommendationer i övervägande. Nedan följer en lista över dessa.

- Amplikonstorlek skiljer sig
- Smälttemperatur (eng. melting temperature, T_m)
- GC-innehåll på 40-60%
- Eventuell sekundärstruktur (primers som binder till sig själv - hindrar korrekt hybridisering)
- Primers ska inte ha mer än två G eller C bland de fem sista nukleotiderna i 3'-änden
- Inga G i probens 5'-ände
- Fler C än G i probens sekvens
- Huruvida sekvensen var komplementär till andra regioner inom samma genom
- Huruvida sekvensen var komplementär till andra genom (specifikt andra serogrupper av *Shigella* och arter av *E. coli*)
- Huruvida primers och prober är komplementära till varandra ("primerdimers")

Förutom att de olika amplikonen skulle vara av olika storlek för att kunna separeras i en agarosgel (31) var det också viktigt att se till att de var anpassade efter TaqMan®-polymeraset. Primers designades därför så att amplikonstorlekarna låg mellan 80 och 322 baspar (31). Smälttemperaturen rekommenderades också att ligga mellan 58-60 °C för primers och 10 °C högre för en TaqMan®-prob (31,32). Vilket GC-innehåll en sekvens rekommenderades att ha varierade lite mellan olika källor. De primers som gjordes valdes därför så att de hade ett GC-innehåll på 40-60 % (32,33). Alla prober designades för genernas forward-sträng utom en ("Prob_gspM2"). Detta för att möta den rekommendation som fanns om att det skulle vara fler C än G i probsekvensen (32).

Primers och prober för generna *ipaH* och *wzx* togs direkt från SLVs protokoll (7). Primersekvenserna och probsekvensen för *wzx* kunde dock inte hittas i de genom som användes (se bilaga 1). Detta gjorde att amplikonstorleken ej kunde beräknas. Eftersom att dessa sekvenser ändå var tillhandahållna kunde en smälttemperatur uppskattas.

BILAGA 11 - PRIMERS OCH PROBER

Tabell 7 Denna tabell innehåller de parametrar som ansågs viktigast under designmomentet av de prober och primers som rekommenderas för den föreslagna metoden för respektive gen. Detta inkluderar, förutom vilken serogrupp av *Shigella* som dessa ämnar att detektera, varje primer och probsekvens, smälttemperatur (T_m) och GC-innehåll. För att underlätta för direkt detektion har också varje gens amplikonstorlek bifogats.

Detekterar art	Namn	Sekvens	Amplikon-längd (bp)	T_m (C°)	GC-innehåll (%)	Källa
<i>Shigella</i> , EIEC	F_ipaH	GCGTTTCTAIGGGGIGTC	196	58,29	56	Livsmedelsverkets protokoll
	R_ipaH	GTACTCAATTCAGCAATCTC		56,74	48	Livsmedelsverkets protokoll
	Prob_ipaH	CCTCGGCACATGCCGAAGCCA		68,36	70	Livsmedelsverkets protokoll
<i>S. boydii</i>	F_gspM1	GGATAGTACAACAGGAGACAATG	322	57,1	43	Egen design
	R_gspM1	ACATTCACCAATCCCAGTCT		57,29	47	Egen design
<i>S. boydii</i> , <i>S. dysenteriae</i>	Prob_gspM1	CCGGCACCTCGGTAGTCAATCAAGC	122	68,93	62	Egen design
	F_gspM2	TTAAGCCGTAGCGAACACT		58,32	47	Egen design
	R_gspM2	TGCCAGATGAGGACGTAATAT		57,16	43	Egen design
	Prob_gspM2	ACCTGGCGGGGCGTTCGC		67	55	Egen design
<i>E. coli</i> (EIEC)	F_lacY	AGAAGCAGAAAACAGACCAGAT	163	58,75	43	Egen design
	R_lacY	ATTATTGGCTCATCGTTCGC		58,44	45	Egen design
	Prob_lacY	CAGCCCACCAGCAGGAACGGTAC		68,23	65	Egen design
<i>S. flexneri</i>	F_rfc	GGGACGAAAAAAGAACTCGC	281	59,1	48	Egen design
	R_rfc	TTTAATGGCTTCTTTGTCCGC		58,15	45	Baserad på primer SflexDF1 [11]
	Prob_rfc	AAAAAGCCCCATCGGGACGAAAAAAGAAC		67,37	45	Egen design
<i>S. dysenteriae</i> typ 1	F_shuA	TTGGCTTTGGCTGTTCTG	79	58,08	47	Egen design
	R_shuA	GCAATCCCGTTGCCGGTAA		59,97	58	Egen design
	Prob_shuA	CCACCTTGCCAAACGTTTGTCTTTGCT		67,45	50	Egen design
<i>S. sonnei</i>	F_wzx	GCACTAATTCGTAGGATATTG	-	*56,32	43	Livsmedelsverkets protokoll
	R_wzx	CACACCAATACAATACGATTTAC		*53,82	32	Livsmedelsverkets protokoll
	Prob_wzx	CGCTCGCTTTGATTTCCACTACAC		*63,47	50	Livsmedelsverkets protokoll