UPTEC X 13 016
Examensarbete 30 hp
Mars 2014

UPPSALA
UNIVERSITET

Validation of PCR assays for detection of Shiga toxin-producing *E. coli* O104:H4 and O121 in food

Johanna Tawe

UPPSALA UNIVERSITET

Molecular Biotechnology Programme

Uppsala University School of Engineering

UPTEC X 13 016	**Date of issue 2014-03**

Author

Johanna Tawe

Title (English)

Validation of PCR assays for detection of STEC O104:H4 and O121 in food

Title (Swedish)

Abstract

Shigatoxin-producing *Escherichia coli* (STEC) can cause infections in humans which can be serious and sometimes fatal. There is a great need for methods that are able to detect different serogroups of STEC. In this project, conventional and real-time PCR assays for detection of STEC O104:H4 and O121, as recommended by the European Union Reference Laboratory (EU-RL) for STEC, were validated. The specificity, limit of detection, repeatability, efficiency and robustness were determined for three real-time PCR assays. The validation showed that the real-time PCR reactions were specific and sensitive although some additional tests are required.

Keywords
Pathogenic *E. coli*, STEC, VTEC, EHEC, PCR, validation

Supervisors

Catarina Nilsson
National Food Agency

Scientific reviewer

Margareta Krabbe
Uppsala university

Project name	Sponsors
Language **English**	Security
ISSN 1401-2138	Classification
Supplementary bibliographical information	Pages **35**

Biology Education Centre Biomedical Center Husargatan 3 Uppsala
Box 592 S-75124 Uppsala Tel +46 (0)18 4710000 Fax +46 (0)18 471 4687

Validation of PCR assays for detection of Shiga toxin-producing *E. coli* O104:H4 and O121 in food

Johanna Tawe

Populärvetenskaplig sammanfattning

Escherichia coli är den vanligaste bakterien i den normala tarmfloran hos människor. *E. coli* är vanligtvis ofarlig för människan men det finns patogena stammar som kan orsaka stor skada. Infektion av verocytotoxin-producerande *E.coli* (VTEC) kan framkalla både oblodiga och blodiga diarréer men även allvarligare tillstånd som njursvikt. VTEC smittar ofta människor via livsmedel som kontaminerats av gödsel från nötkreatur som är reservoarer av VTEC. Sommaren 2011 inträffade ett större utbrott av VTEC O104:H4 i Europa som drabbade över 3000 människor med totalt 15 inblandade länder. Smittan kunde slutligen spåras till groddar från bockhornsklöverfrön importerade från Egypten.

VTEC delas in i olika serogrupper baserat på vilket O-antigen de har. Det är av stor vikt att kunna detektera flera serogrupper av VTEC. Metoderna som Livsmedelsverket använder för att detektera serogrupper av VTEC är baserade på konventionell PCR samt realtids PCR. PCR-tekniken innebär att ett DNA-fragment unikt för den specifika serogruppen amplifieras och analyseras. Vid en realtids PCR detekteras DNA-fragmenten med hjälp av fluorescerande molekyler.

I denna studie har PCR reaktionerna för att detektera VTEC O104:H4 och O121 validerats. I valideringen bestämdes PCR reaktionernas specificitet, detektionsgräns, precision, effektivitet samt robusthet. Valideringen visade att PCR reaktionerna var både specifika och känsliga men ytterligare tester behöver utföras för att säkerställa resultaten.

Examensarbete 30 hp
Civilingenjörsprogrammet Molekylär bioteknik

Uppsala universitet, mars 2014

Index

Abbreviations

A/E lesion	Attaching and effacing lesion
BHI	Brain heart infusion
BLAST	Basic Local Alignment Search Tool
C_q	Quantification cycle
CV	Coefficient of variation
eae	*Escherichia coli* attaching and effacing gene
EAEC, EAggEC	Enteroaggregative *E. coli*
EHEC	Enterohemorrhagic *E. coli*
EU-RL	European Union Reference Laboratory
HUS	Haemolytic uremic syndrome
IMS	Immunomagnetic separation
LOD	Limit of detection
LEE	Locus of enterocyte effacement
NRL	National Reference Laboratory
NTC	Negative template control
PCR	Polymerase chain reaction
RFU	Relative fluorescence units
SMI	Smittskyddsinstitutet (Swedish Institute for Communicable Disease Control)
SSI	Statens Serum Institut, Denmark
STEC	Shiga toxin-producing *E. coli*
Stx	Shiga toxin
stx1, stx2	Shiga toxin genes
VTEC	Verocytotoxin-producing *E. coli*

1. Introduction

An outbreak of a Shiga toxin-producing *Escherichia coli* (STEC) infection in Germany from May to June in 2011 was the cause of 50 deaths in Europe. It is to date the largest outbreak of STEC in Europe, with 15 afflicted countries and a total of 3,168 infected patients. The outbreak was caused by a STEC serotype called O104:H4, and fenugreek sprouts were identified as the most likely origin (1). O104:H4 had not before been associated with severe infections (2). The outbreak was in fact not caused by a STEC strain but by another type of pathogenic *E. coli* known as enteroaggregative *E. coli* (EAEC or EAggEC). The EAEC strain had acquired the shiga toxin-encoding genes most probably by horizontal gene transfer (3). As a consequence of this outbreak the European Commission has passed a regulation (No 209/2013) that applies from July 1st 2013 concerning microbiological criteria for sprouts. The regulation states that all food business operators producing sprouts have to analyze their seeds for six serogroups of STEC which are recognized to be those causing the most of the haemolytic uremic syndrome (HUS) cases occurring in the EU, among them O104:H4. In order for the sprouts to be released on the market there cannot be any detection of STEC and none of the six serogroups.

The most common serotype associated with STEC outbreaks worldwide is O157:H7 but more than 200 serotypes of STEC are known (4, 5). Other serogroups frequently associated with human infections are O26, O103, O111, O145, O91, O113, O128, O45 and O121 (6). STEC O121, for example, is quite common in Sweden where O121 contributes to 7% of infections caused by non-O157 STEC (7). The ability to detect a variety of STEC serogroups is of concern for the public health and therefore the Swedish National Food Agency is in need of methods that are able to detect several different serotypes of STEC.

The method for identifying STEC in food is an ISO technical specification (ISO/TS 13136:2012) which, after an enrichment step, involves detection of the typical virulence genes of STEC (*stx1*, *stx2* and *eae*) with PCR. If both or one of the *stx* genes are detected isolation is attempted and the sample is evaluated further with serogroup specific PCR for O157, O111, O26, O145 and O103. A serogroups-specific isolation using e.g. immunomagnetic separation (IMS), and sometimes also selective agars, can be performed if one of the serogroups is detected in the PCR. If the *stx* genes are detected but none of the serogroups, a direct plating can be performed (8). Alternatively, immunoblotting can be used in order to isolate STEC but this is not included in the ISO technical specification.

The purpose of this project was to set up serogroup specific PCR assays for detection of STEC O104:H4 and O121. The PCR assays were TaqMan-based real time PCR assays for O104:H4 and a conventional PCR assay for O121 (2, 9). These assays also had to be validated to determine the precision, robustness, efficiency, sensitivity and specificity of the assays. The PCR assays were recommended by the European Union Reference Laboratory (EU-RL) for STEC. The EU-RLs are designated by the European Commission, and in addition there are National Reference Laboratories (NRLs) for the EU Members States for each respective field. The Microbiology Division of the Swedish National Food Agency is NRL for pathogenic *E. coli*, *Campylocacter*, *Listeria*, *Staphylococcus* and bacterial and viral contamination of bivalve mollusks.

1.1 Shiga toxin-producing *E. coli*

E. coli is a rod-shaped gram-negative bacterium that has been studied in laboratories for over 50 years. The high growth-rate and the fact that it is relatively easy to culture make *E. coli* an optimum modeling organism and it is used in many experiments in genetic and molecular biology. *E. coli* is also utilized to produce recombinant drugs (10, 11).

E. coli is the dominating microbe in the normal gut flora in humans and is usually non-pathogenic (11). However, there strains exist that have acquired the ability to produce virulence factors which makes them pathogenic and able to cause a variety of diseases. These virulence factors are transferred between strains on mobile genetic elements and can then become part of the genome. Pathogenic *E. coli* is divided into categories based on their virulence factors and pathogenicity; enteroaggregative *E. coli (*EAEC), enteropathogenic *E. coli* (EPEC), enterohaemorrhagic *E. coli* (EHEC), enterotoxigenic *E. coli* (ETEC), enteroinvasive *E. coli* (EIEC) and diffusely adherent *E. coli* (DAEC) among others (5).

EHEC is also called STEC or VTEC (verocytotoxin-producing *E. coli*) because of its ability to produce shiga toxins also known as verocytotoxins (4). The term EHEC refers to the clinical picture of many STEC infections characterized by hemorrhagic colitis and HUS (12). HUS is a life-threatening disease characterized by haemolytic anaemia (destruction of red blood-cells), thrombocytopenia (decrease of platelets in blood) and potentially acute renal failure (5). STEC infections also have been associated with cases of mild to bloody diarrhoea (6). In Sweden about 300 cases of STEC infection is reported every year. About five percent of the patients develop HUS, primarily children under the age of five (13). Cases of asymptomatic carriers of STEC have also been reported (14).

STEC is categorized into serotypes based on its O and H antigen. The O (Ohne) antigen is constituted by the polysaccharide in the cell wall lipopolysaccharide and the H (Hauch) antigen is determined by the flagella protein. The combination of the O and H antigen gives rise to a large quantity of serotypes of STEC but most of the over 200 known serotypes are not associated with human disease (5, 12).

1.2 Virulence factors

STEC can possess several virulence factors which are usually encoded by plasmids and pathogenicity islands that make it pathogenic to humans. Some of the virulence factors are well known but more information about genes involved in pathogenicity is needed. Presumably a combination of virulence genes is required for pathogenicity (4).

The shiga toxin (Stx) is the key virulence factor of STEC. Shiga toxins are a family of cytotoxins, divided into Stx1 and Stx2, with related structure and biological activity. The Stx genes, *stx1* and *stx2*, are encoded on temperate bacteriophages present in the bacterial chromosome (12). *Stx2* is more frequently associated with severe human infections than *stx1* (4). When Stx is produced in the colon it can be delivered to the kidney via the bloodstream. In the kidney, Stx causes damage on the renal endothelial cells, clogs up the microvasculature and causes renal inflammation which can develop into HUS. Stx also induces apoptosis of epithelial cells in the intestine and mediates local colon damage (5).

An important pathogenic property of certain serotypes of STEC is known as “attaching and effacing” (A/E) lesion. A/E lesion is a mechanism that occurs when STEC attaches to intestinal epithelial cells

and cause cytoskeletal changes and induces a histopathologic lesion (tissue damage). The microvilli in the epithelial cells are effaced and the bacterium becomes intimately attached to the cell membrane. The genes involved in A/E lesion are present in the bacterial genome as a large pathogenicity island called the Locus of Enterocyte Effacement (LEE) (4).

The LEE encodes several proteins involved in the A/E lesion, one of them being the adhesion protein Intimin. Intimin mediates the close attachment of STEC to the epithelial cells and is encoded by genes called *eae*. The presence of the LEE-island with the *eae* genes in STEC is often associated with disease upon infection in humans. The majority of STEC strains that cause bloody diarrhoea and HUS are *eae*-positive and possesses the LEE-island (4, 12).

Another important factor that contributes to the pathogenicity of STEC is its resistance to acidic environments. In fact acid resistance is a general characteristic of *E. coli*. The acidic environment in the gastrointestinal tract is an important host defence mechanism. Because of the acidic resistance STEC can survive in the gastrointestinal tract and in acidic foods such as apple juice and salami (12).

1.3 Sources of infection

Outbreaks of STEC infections have been traced to a number of different food sources. Ruminants, especially cattle, are the main reservoir for STEC and humans are infected by direct or indirect contact with the animals or their feces or by consuming their edible products. A large number of outbreaks have been caused by undercooked ground beef and other beef products contaminated with STEC (14, 15). The infectious dose of STEC is very low, for O157:H7 less than 50 organisms are required to cause infection (12).

Other food products that have been associated with infection are ready-to-eat sausages (such as salami), unpasteurized milk and apple juice, cheese, drinking water and vegetables such as lettuce, sprouts, cantaloupes and carrots. Foods that carry STEC have most probably been contaminated by contact with ruminant feces (12, 14).

2. Background

2.1 PCR

The polymerase chain reaction (PCR) was developed in the eighties by Kary Mullis (16) and has since been the most important tool in nucleic acid amplification and analysis. PCR is extremely sensitive and enables amplification of very small amounts of DNA fragments in vitro (17). Primers, nucleotides and a thermostable form of DNA polymerase are required in order to perform a PCR reaction. *Taq* polymerase from *Thermus aquatiqus* (a bacterium that lives at high temperatures) is most commonly used. Primers are short oligonucleotides that are complementary to the 3' ends of the DNA to be amplified (18). The primers are the foundation of the DNA strands to be synthesised. The PCR mixture also contains a buffer solution, that provides an appropriate chemical environment for the reaction, and divalent cations (usually Mg^{2+}) which facilitates nucleotide addition to the growing DNA strand (19).

The PCR procedure follows four general steps (20). First the DNA is denatured and separated into two strands by heating the mixture to around 94ºC. The temperature is then reduced to about 50-60ºC to enable annealing of the primers to the complementary 3' ends of the target DNA. The temperature is then increased to approximately 72ºC which is the optimum temperature for *Taq* polymerase. In this step DNA is synthesized by addition of complementary nucleotides to the 3' end of the primers. This results in a duplication of the target DNA. The annealing and synthesis steps are then repeated in several cycles and the amount of target DNA accumulates exponentially. After 32 cycles approximately 1 billion copies of target DNA are generated (17).

When conventional PCR is used as a detection method the amplified DNA target needs to be analysed further in order to ensure that the correct product has been produced. This is often done by gel electrophoresis which will reveal a single band with the expected amplicon size if the PCR was a success (20).

2.1.1 Real-time PCR

Real-time PCR is an improvement of conventional PCR that enables continuous monitoring of the PCR products after each cycle in real time (21). The amount of product that is formed during the PCR process can be observed by using fluorescent-tagged probes in the reaction mixture. The strength of the fluorescence signal is proportional to the quantity of product formed in the PCR. Initially the signal is undetectable but after a certain amount of cycles, as the amount of product increases, the signal becomes detectable and can be distinguished from the background. This is known as the threshold fluorescence signal level and the number of cycles where the fluorescence reaches the threshold is known as the quantification cycle, C_q (16, 22). Upon reaching the threshold level the signal will increase exponentially as the amount of product theoretically doubles after each cycle (Figure 1). After a while the signal levels off and reaches the plateau phase. This is the result of the consumption of components essential to the PCR reaction, i.e. primers, probe or dNTPs (16).

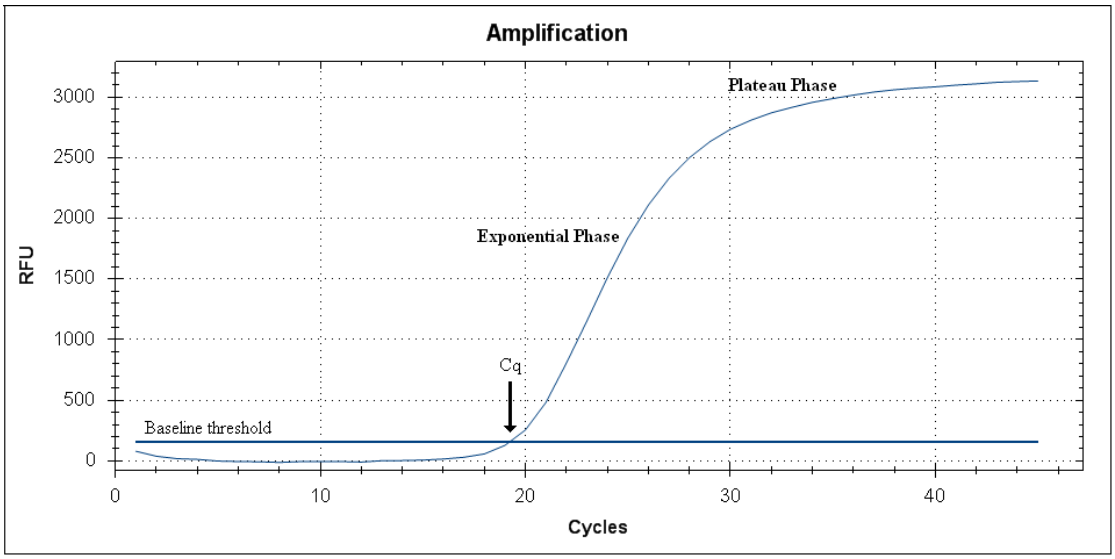


Figure 1. Real-time PCR response curve. The number of PCR cycles is plotted against the fluorescence. The Cq value is the number of cycles where the fluorescence reaches the threshold. The fluorescence then increases exponentially until it reaches the plateau phase and levels off.

The fluorescence signal can be obtained by using either sequence specific probes or non-specific labels (16). In this study TaqMan probes and SYBRGreen were used. TaqMan probes are sequence specific oligonucleotides with a fluorophore at the 5' end and a quencher dye at the 3' end. The quencher dye quenches the fluorescence of the fluorophore when the fluorophore and quencher are in proximity. Taqman probe binds to its complementary sequence on the target DNA during the annealing step in the PCR. When the DNA then is synthesised the exonuclease activity of *Taq* polymerase cleaves the probe and separates the quencher from the fluorophore. Since the fluorophore no longer is in proximity of the quencher a fluorescence signal will be detected (Figure 3) (21).

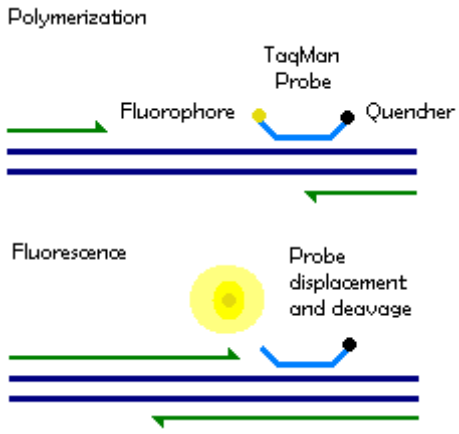


Figure 3. Mechanism of the TaqMan probe. When the quencher and fluorophore are in proximity no signal is detected. Polymerization of the nascent DNA cleaves the probe and separates the fluorophore from its quencher and a fluorescence signal is detected.

SYBR Green is a non-specific cyanine dye that binds to double-stranded DNA and emits a fluorescent signal (Figure 4). Since SYBR Green is non-specific it can be used to detect non-specific undesirable products, such as primer dimers, in PCR assays with melt curve analysis (21). A primer dimer consists of two primers that have hybridized to each other because of patches of complementary bases in their sequences.

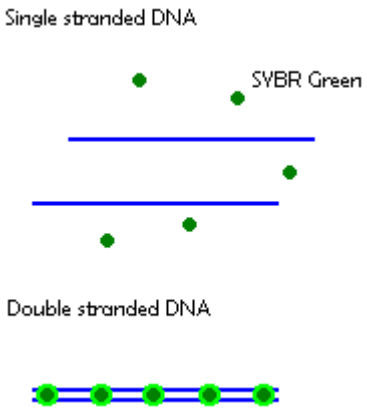


Figure 4. Mechanism of SYBR Green. Single stranded DNA doesn´t bind to SYBR Green and no signal is emitted. When the DNA becomes double stranded upon polymerization SYBR Green will bind and a fluorescence signal is detected.

2.2 Validation of real-time PCR assays

In order to secure accurate results the methods for detecting STEC with real-time PCR have to be validated. Every step in this method should be validated but in this study the focus is on the real-time PCR assays for detecting O104, H4 and O121.

The PCR assays in this study are qualitative, meaning that they only are used to determine if the template DNA is present or not. A quantitative PCR assay can be used to calculate the amount of template DNA that is present. The parameters measured were precision, specificity, limit of detection, efficiency and robustness. Other parameters that can be included in a validation study are accuracy, reproducibility, and limit of quantification (LOQ) among others.

2.2.1 Specificity

Specificity is the methods capacity to detect only the intended target and not any non-specific target that may be present in the sample (22). The specificity is divided into two parts, inclusivity and exclusivity. The inclusivity test verifies that the assay detects the intended target, while the exclusivity test confirms that the assay does not cross-react with any other organisms. The exclusivity test should include closely related organisms, organisms present in the normal flora of infected patients and in food that may be infected, and organisms that cause similar diseases as the target organism (23). The specificity of a PCR assay depends primarily on how unique the sequences of the probes and primers are.

2.2.2 Limit of Detection

The limit of detection (LOD), also known as analytical sensitivity, is the smallest amount of DNA that the assay is able to consistently detect (24). When detecting VTEC in foodstuff a low value of LOD is

preferable since a very small amount of STEC has the potential to cause infection in humans. Further the bacteria can be difficult to find in the food sample and therefore the assay needs to be able to detect low concentrations of DNA. The unit of LOD is given in gene copies per PCR reaction. The lowest LOD that is possible to obtain theoretically is three gene copies per reaction (assuming a Poisson distribution, single-copy detection and a 95% possibility of including a minimum of 1 copy in the PCR) (22).

2.2.3 Precision

The precision describes the variation within analytical results obtained from repeated tests performed under specified conditions. Precision can be described with different parameters but this study focuses on the *repeatability* and *reproducibility*. Repeatability measures the lowest degree of variation and is evaluated by performing the same tests on a number of replicates under identical conditions. Reproducibility, being at the opposite side of the spectra, measures the highest degree of variation and is determined by executing measurements under changed conditions (different laboratories, operators, equipment etc.). The intermediate precision can also be evaluated. This is done by varying some of the condition parameters and keeping other parameters constant (24). In this study only repeatability was measured due to limitations in available laboratories and equipment.

The precision of a PCR assay is estimated by calculating the standard deviation of the C_q-values obtained from several replicates. A poor precision can be caused by pipetting errors and stochastic variation. The precision usually increases with decreasing gene copy concentration (22).

2.2.4 Efficiency

The efficiency of a PCR assay is the fraction of target DNA that is successfully copied in each cycle. Ideally the amount of target DNA is doubled after each PCR cycle and the efficiency is 100%. A high efficiency implies that the method is robust and precise. The efficiency depends on the primer and probe design and can be reduced if there are undesirable products in the reaction mixture such as primer dimers (22).

The efficiency of a PCR assay can be calculated from a standard curve obtained from serial dilutions of known concentrations. This is however only an estimate of the PCR efficiency and a real test sample, such as food, may contain inhibitory substances that decrease the PCR efficiency (16).

2.2.5 Robustness

The robustness of a PCR assay is its capacity to resist small changes in the execution of the method. A robust method can therefore tolerate small changes and still give consistent results. By varying different parameters in the assay factors that have a negative effect on the assay may be identified. The varying parameters can be equipment, pH, temperature, concentration of reagents, etc.

3. Materials and methods

3.1 Control strains

The control strains used for the optimization and validation of the PCR assays (Table 1) were retrieved from Statens Serum Institut (SSI) in Denmark. EAggEC was the strain used for both the O104 and H4 assay. Each strain was inoculated onto nutrient agar and the plate was then incubated overnight at 37ºC.

Table 1. Control strains used in the optimization and validation of the PCR assays.

Strain	Reference	Origin
EAggEC O104:H4	D4116, C679-12	SSI
O121	39w	SSI

3.2 DNA extraction

DNA from the strains used in this study was prepared by using two different DNA extraction kits, NucleoSpin® Tissue (Machery-Nagel) and DNeasy® Blood & Tissue Kit (QIAGEN®), and by using the InstaGene™ matrix (Bio-Rad). NucleoSpin® Tissue (Machery-Nagel) was used for the control strains and DNeasy® Blood & Tissue Kit (QIAGEN®) for the control strains and an additional O121 strain used in the exclusivity study. The two extraction kits basically follow the same procedure but with different materials, i.e. buffers and filter columns.

An isolated colony of each strain was taken from the nutrient agar plates and inoculated in BHI broth overnight at 37ºC. A 1ml sample from each enrichment was used for the extraction procedure. Regardless of which extraction kit that was used the samples were first centrifuged in order to obtain a pellet containing the bacterial cells. The pellet was resuspended in a buffer and Proteinase K was added. The samples were then incubated at 56ºC for 1-3 hours to obtain complete lysis of the cells. To achieve RNA-free DNA RNase A was added to each sample. Following addition of another buffer and ethanol (96-100%) the samples were applied to a spin column that binds the DNA upon centrifugation. The samples were then washed repeated times by adding different buffers to the columns followed by centrifugation and disposal of flow-through. Finally, the DNA was eluted by addition of an elution buffer to the spin columns and a subsequent centrifugation (25, 26).

Non-VTEC bacterial DNA (to be used in the exclusivity study) and three strains of VTEC O121 DNA (used in the inclusivity study of the O121 assay) were extracted with the InstaGene™ matrix (Bio-Rad). A colony of each bacterium was resuspended in water and centrifuged to obtain a pellet to which the InstaGene matrix was added. The samples were incubated at 56ºC for approximately 20 minutes and then at 95ºC for 8 minutes. The samples were then centrifuged and the resulting supernatant was used for the PCR reactions (27).

3.3 Determination of DNA concentration and sample purity using NanoDrop

The DNA concentration and purity were evaluated by using NanoDrop® ND-1000 Spectrophotometer (Saveen & Werner). The concentration was given in ng/µl and the purity was evaluated from the 260:280 ratio (for DNA a 260:280 ratio of ~1.8 is considered acceptable). The concentration, *c*, was then converted into gene copies per µl by using equation 1.

$$\frac{6.022 \cdot 10^{23} mol^{-1} \cdot c}{650 g \cdot mol^{-1} bp^{-1} \cdot 5.4 \cdot 10^{6} bp} \quad (1)$$

Where $6.022 \cdot 10^{23}$ mol^{-1} is the Avogadro constant, 650 $g \cdot mol^{-1} bp^{-1}$ is the average molar mass of one base pair and $5.4 \cdot 10^{6}$ bp is the genome size of STEC O157.

3.4 PCR

3.4.1 Conventional PCR

EU-RL recommended a conventional PCR method to detect O121. Each PCR mixture consisted of 30 µl with 20 µl of reaction mixture and 10 µl of the DNA sample. The sample contained serial dilutions of target DNA ranging from 10^6 to 0.78 gene copies in two and three replicates. Each PCR contained three negative template controls (NTCs) which consisted of MilliQ water.

The components in the standard reaction mixture were 1x PCR Gold Buffer (Applied Biosystems®), 3mM $MgCl_2$ (Applied Biosystems®), 300µM dNTP (VWR International), 100nM of each primer (Eurofins Operon) and 2.5 units of AmpliTaq Gold® DNA polymerase (Applied Biosystems®). MilliQ was added to give a total of 30 µl of PCR mixture. The primers used for the conventional PCR are listed in Table 2. The standard reaction parameters were 95ºC for 15 min, 10 cycles of 95ºC for 30 s, 68-59ºC (decrease 1ºC/cycle) for 20 s, 72ºC for 52 s, followed by 35 cycles of 95ºC for 30 s, 59ºC for 20 s, 72ºC for 52 s and a final extension at 72ºC for 1 min. The PCR reactions were performed on GeneAmp® PCR System 2700 (Applied Biosystems). The generated PCR products were analyzed with 2% agarose gel electrophoresis followed by staining with GelRed Nucleic Acid Stain (Biotum) and visualisation with Molecular Imager® Gel Doc™ XR+ System (Bio-Rad). The amplicon size of the target gene of O121 is 651 bp.

Table 2. Primer and probe sequences for O121 used in the conventional PCR assay.

Serotype	Target gene	Forward primer (5'-3')	Reverse primer (5'-3')	Reference
O121	*wzx*	GTAGCGAAAGGTTAGACTGG	ATGGGAAAGCTGATACTGC	(28)

3.4.1.1 Optimization of the conventional PCR

The PCR reactions were also performed on Mastercycler® gradient (eppendorf) and the standard concentration of primers and Taq polymerase were changed to 300 nM and 1.25 U, respectively. The recommended PCR reaction parameters were also changed into a simplified program specified by 95ºC for 5 min, followed by 45 cycles of 95ºC for 30 s, 60ºC for 30 s, 72ºC for 1 min and a final extension at 72ºC for 10 min.

3.4.2 Real-time PCR

The real-time PCR assays for O104 and H4 were recommended by EU-RL. Each PCR mixture contained 25 µl with 20 µl of reaction mixture and 5 µl of the sample. Each PCR contained NTCs. The primers and probes used for the real-time PCR assays are listed in Table 3.

The O104 assay was tested with both 1x Taqman® Universal PCR Master Mix (Applied Biosystems®) and 1x PerfeCTa qPCR Toughmix (Quanta). The test consisted of 7 dilutions of the target DNA ranging from 10^6 to 10^0 in duplicates. The other reaction components were 500 nM of

each primer (Eurofins Operon) and 200 nM probe (Eurofins Operon). The PCR reaction parameters for both the O104 and H4 assays were 95ºC for 10 min followed by 45 cycles of 95ºC for 15 s and 60ºC for 1 min.

The reaction components for the H4 assay were 1x Taqman® Universal PCR Master Mix (Applied Biosystems®), 500 nM of each primer (Eurofins) and 200 nM probe (Eurofins).

For the O121 assay 1x PerfeCTa qPCR Toughmix (Quanta), 400 nM primers (Eurofins) and 100 nM probe (Eurofins) were used. PCR reaction parameters for O121 were 94ºC for 10 min followed by 45 cycles of 94ºC for 30 s, 52ºC for 30 s and 72ºC for 30 s.

MilliQ was added to each assay to give a total of 25 µl of PCR mixture. The real-time PCR was performed on CFX96™ Real-Time PCR Detection System (Bio-Rad) and the results were evaluated in CFX Manager ™ (Bio-Rad). The baseline threshold was set to 205, 52 and 50 RFU (relative fluorescence units) for the O104, H4 and O121 assay, respectively.

Table 3. Primer and probe sequences used in the real-time PCR assays.

Serotype	Target gene	Forward primer (5'-3')	Reverse primer (5'-3')	Probe	Reference
O104	*wzx*	TGTCGCGCAAAGAATTTCAAC	AAAATCCTTTAAACTATACGCCC	FAM-TTGGTTTTTTTGTATTAGCAATAAGTGGTGTC-BHQ1	(29)
H4	*fliC*	GCTGGGGGTAAACAAGTCAA	CCAGTGCTTTTAACGGATCG	HEX-TCTTACACTGACACCGCGTC-BHQ1	(29)
O121	*wzy*	TGGATGGCATTCCTCAGTAT	AGCAAGCCAAAACACTCAAC	FAM-TTAACACGGGCGTGGTTGGA-BHQ1	(30)

3.5 Validation of the real-time PCR assays

The real-time PCR assays were validated by determining the precision, specificity, LOD, efficiency and robustness of the assays. The validation was performed with the reaction components and reaction parameters declared above with 1x PerfeCTa qPCR Toughmix (Quanta) used for the O104 assay.

3.5.1 Specificity

The specificity of the primers and probes used in this study (Table 2 and 3) was analysed by using Standard Nucleotide BLAST (Basic Local Alignment Search Tool) administered by the NCBI National Center for Biotechnology Information (NCBI). BLAST is an algorithm that compares sequence information with sequence databases and identifies sequences with high similarity. The primers and probes should not show high sequence similarity with organisms other than the intended target.

The inclusivity of each PCR assay was studied with DNA from different strains of O104, O121 and H4. For the O104, H4 and O121 assays three, two and eleven strains were tested, respectively (Table 6). For the O104 and H4 assays approximately 5 ng of DNA per reaction was used and for O121 approximately 10 ng of DNA per reaction was used in duplicates.

The exclusivity of each assay was evaluated with a total of 45 strains of 30 different species (Table 7). The species included bacteria common in the normal microflora of humans and animals, pathogenic

bacteria (some of them common in foodborne illnesses), pathogenic bacteria known to cause infections with similar symptoms as STEC, plant pathogenic fungi, non-STEC *E. coli* and serogroups of STEC other than O104:H4 and O121. The PCR was performed for each assay with the bacterial species in duplicates with approximately 10 ng bacteria per reaction. Each assay also included two negative controls and two positive controls containing 100 gene copies per reaction of the template DNA.

3.5.2 Limit of Detection

The evaluation of LOD included six PCR runs with 12 dilutions of template DNA, ranging from 10^6 to 0.78 gene copies per reaction, in triplicates. The LOD for each assay was the lowest concentration consistently detected in all six PCR runs. 95 % of the replicates, which corresponds to all of the total 18 replicates, at a certain concentration had to give a positive result in order to be defined as detectable.

3.5.3 Repeatability

The repeatability of the PCR assays was evaluated with target DNA in 10 dilutions ranging from 10^6 gene copies per reaction for O104:H4 and from 10^5 gene copies per reaction for O121 down to the LOD of each assay. The dilutions were tested in six replicates in one PCR run. The standard deviation of the C_q-values of each concentration was then calculated by using equation 2.

$$STDAV = \sqrt{\frac{\sum (C_q - \overline{C_q})^2}{n}} \tag{2}$$

Where $\overline{C_q}$ is the mean C_q-value and n is the number of observations. The value obtained was used to calculate the coefficient of variation, CV, with equation 3.

$$CV = \frac{STDAV}{\overline{C_q}} \tag{3}$$

3.5.4 Efficiency

The efficiency of the assays was determined in one PCR run with template DNA in seven serial dilutions, ranging from 10^6 to 25 gene copies per reaction, in triplicates. The PCR run generated a standard curve with the C_q-value plotted against the logarithm of the concentration. The efficiency of the assay was then given by equation 4.

$$E(100\%) = \left(10^{\frac{-1}{k}} - 1 \right) \times 100 \tag{4}$$

Where k is the slope of the standard curve obtained by linear regression. The coefficient of determination, R^2, of the curve had to be >0.99. The value of the efficiency should lie within 90-100%, corresponding to a slope between -3.6 and -3.3.

3.5.5 Robustness

Evaluation of the robustness was executed by varying the primer, probe and buffer concentrations in the reaction mixture and the annealing temperature in the PCR program. For O104 and H4 five

annealing temperatures were tested within ±4ºC of the standard annealing temperature. The annealing temperatures were 64.3ºC, 62.4ºC, 60.0ºC, 58.0ºC and 56.0ºC. For O121 three annealing temperatures were tested (i.e. 54.4ºC, 52.0ºC and 50.1ºC) within ±2ºC of the standard annealing temperature. The primer, probe and buffer concentrations tested were ±20% of the standard reaction mixture concentrations (Table 4). The standard concentrations of the reaction mixture were also included in the test. Each of the conditions was tested on 10 replicates of 50 or 100 gene copies of the target DNA. For every combination of reaction mixture concentration and annealing temperature the standard deviation and coefficient of variation of the 10 replicates was calculated by using formula (2) and (3), respectively.

Table 4. Reaction mixture concentrations used in the robustness study.

	O104, H4			O121		
	Standard reaction mixture	+20% reaction mixture	-20% reaction mixture	Standard reaction mixture	+20% reaction mixture	-20% reaction mixture
Buffer	1x	1.2x	0.8x	1x	1.2x	0.8x
Primer	500 nM	600 nM	400 nM	400 nM	480 nM	320 nM
Probe	200 nM	240 nM	160 nM	100 nM	120 nM	80 nM

3.6 Melting curve analysis using SYBR Green

The melting curve analysis was performed with duplicates of 7 serial dilutions of template DNA ranging from 10^6 to 10^0 gene copies per reaction. The master mix components were 1x Power SYBR Green PCR Master mix together with 500 nM primers for the O104 and H4 assays and 400 nM primers for the O121 assay. The PCR reaction parameters were the same as in the validation followed by a temperature increase from 65°C to 95°C with an increment of 0.5°C for 5 s.

4. Results

4.1 Determination of DNA concentration and sample purity using NanoDrop

The results of the NanoDrop measurements are listed in Table 5. The concentration in gene copies per µl was obtained from equation 1 and was used to calculate the serial dilutions. The samples were pure according to the A260/A280 ratios (~1.8) when DNeasy® Blood & Tissue Kit (QIAGEN®) was used. The DNA samples extracted with NucleoSpin® Tissue (Machery-Nagel) gave A260/A280 ratios over and below 1.8. A low A260/A280 ratio indicates the presence of proteins in the sample whereas a high ratio isn´t an issue in general (31). The low ratio for one of the samples was considered acceptable for the purpose of this study.

Table 5. Results from the determination of DNA concentration and purity with NanoDrop ND-1000.

Strain	Extraction method	Concentration (ng/µl)	Mean value (ng/µl)	A260/A280	Concentration (gene copies/µl)
EAggEC D4116	DNeasy® Blood &Tissue Kit (QIAGEN®)	23.34 23.08 22.83	23.08	1.8 1.83 1.84	$3.93 \cdot 10^6$
EAggEC D4116	NucleoSpin® Tissue (Machery-Nagel)	16.64 18.58[a] 16.74	16.69	1.6 1.61 1.67	$2.84 \cdot 10^6$
E. coli O121 39w	DNeasy® Blood &Tissue Kit (QIAGEN®)	15.93 15.28 15.14	15.45	1.8 1.78 1.85	$2.63 \cdot 10^6$
E. coli O121 39w	NucleoSpin® Tissue (Machery-Nagel)	12.43 11.89 12.27	12.20	2.02 2.14 2.25	$2.08 \cdot 10^6$

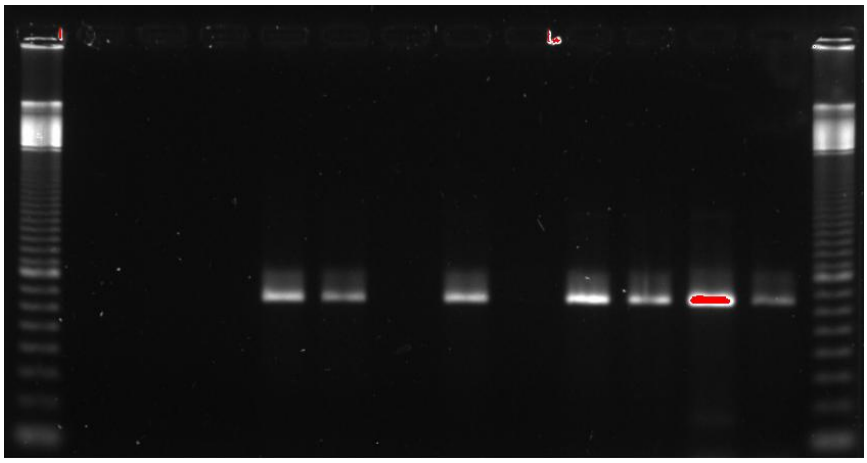
[a] The value was not used in the calculations.

4.2 PCR

4.2.1 Conventional PCR

The results of the conventional PCR assay for O121 recommended by EU-RL with the standard reaction mixture and standard reaction parameters are visualized in Figure 1. Although the assay was able to detect as few as 10^0 gene copies the detection was exceedingly irregular and for some experiments there was no detection at 10^6 gene copies.

a)



b)

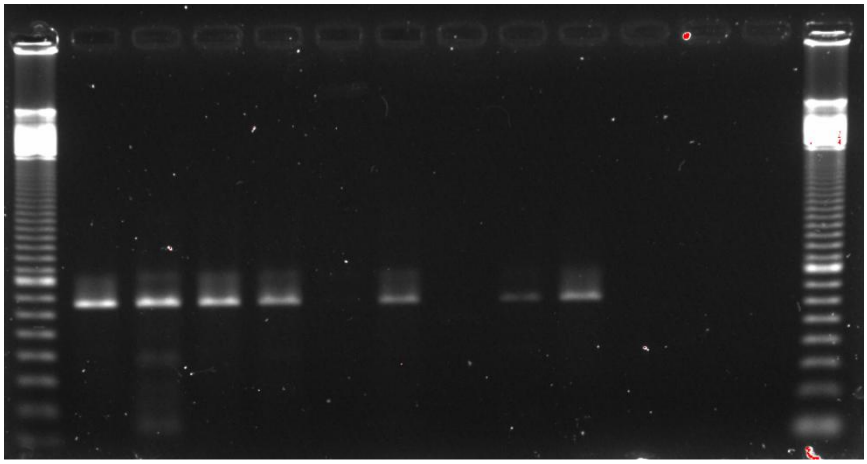
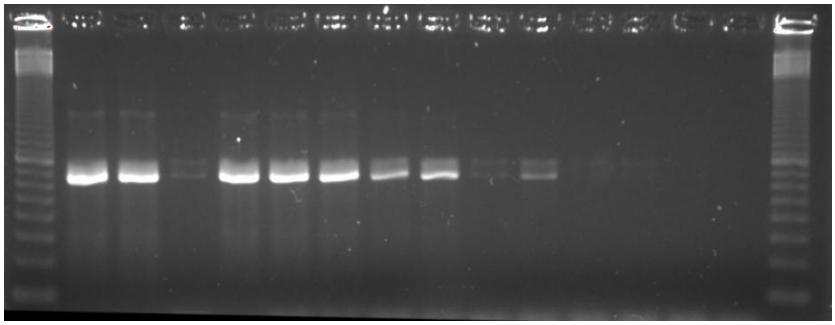


Figure 1. Image of electrophoresis gel loaded with PCR products generated by using the standard reaction parameters with the standard reaction mixture on GeneAmp® PCR System 2700 (Applied Biosystems). The samples (in three replicates) from left to right are in a) 10^6, 10^5, 10^4 and 10^3 gene copies per reaction and in b)10^2, 10^1 and 10^0 gene copies per reaction followed by three NTC:s. The first and last lanes are DNA size markers.

4.2.1.1 Optimization of the conventional PCR

The optimization of the conventional PCR assay for O121 also proved to be irregular and very unstable. Even though there were a few experiments that generated acceptable results with 100% detection down to 10^1 gene copies the results were inconsistent when the experiment was repeated. No improvement could be detected, neither when changing the reaction mixture, reaction parameters nor the PCR instrument. The results of O121 assay with 300 nM primers and 1.25 U *Taq* polymerase and the simplified program performed on Mastercycler® gradient (Eppendorf) is presented in Figure 2. From these results it was concluded that the conventional PCR assay for O121 was unreliable and could not be validated further.

a)



b)

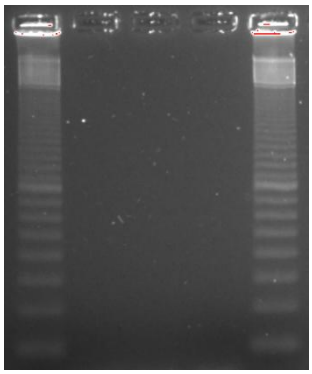


Figure 2. Image of electrophoresis gel loaded with PCR products generated by using the simplified reaction parameters with the optimized reaction mixture on Mastercycler® gradient (eppendorf). The samples (in two replicates) from left to right are in a) 10^6, 10^5, 10^4, 10^3, 10^2, 10^1 and 10^0 gene copies per reaction and in b) three NTC:s. The first and last lanes are DNA size markers.

4.2.2 Real-time PCR

The O104 real-time PCR assay became more sensitive when PerfeCTa qPCR Toughmix was used instead of Taqman® Universal PCR Master Mix (Figure 3). The mean C_q values of the duplicates of each concentration was decreased and the lowest concentration (10^0 gene copies) was detected with PerfeCTa qPCR Toughmix but not with Taqman® Universal PCR Master Mix. This resulted in the use of PerfeCTa qPCR Toughmix in the real-time PCR assay for O104.

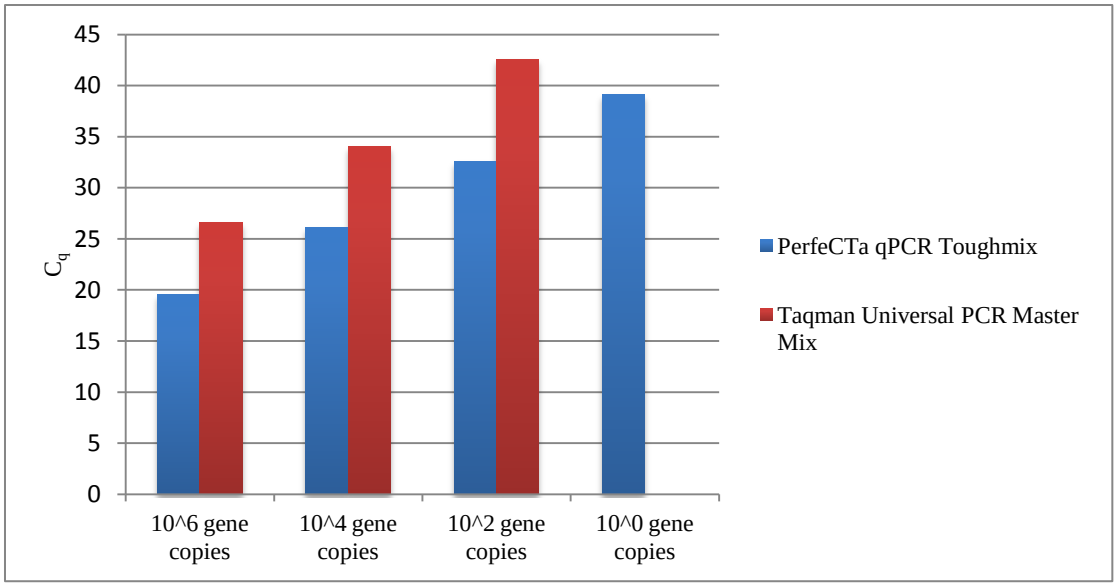


Figure 3. O104 real-time PCR assay tested with PerfeCTa qPCR Toughmix and Taqman® Universal PCR Master Mix. The x-axis represents the number of gene copies used in each reaction and the y-axis the mean C_q value of two replicates of each concentration. The lowest concentration was not detected when Taqman® Universal PCR Master Mix was used.

4.3 Validation of the real-time PCR assays

4.3.1 Specificity

The BLAST search showed that the primers and probes did not have a high sequence identity with other organisms. The primers were 100% sequence identical to their respective target.

Table 6 and 7 show the results of the inclusivity and exclusivity tests, respectively. All O104, H4 and O121 strains were detected correctly in the inclusivity part. There was no detection of the organisms and strains used in the exclusivity test which was as expected after performing the BLAST search.

Table 6. *E.coli* strains used for inclusivity real-time PCR tests and results for O104, H4 and O121.

Organism	Reference	Origin	O104 detection[a]	H4 detection[a]	O121 detection[a]
EAggEC	D4116	SSI	+	+	
E. coli O104:H4	2011 outbreak in Germany	EU-RL	+	+	
E. coli O104	H519	SSI	+		
E. coli O121	39w	SSI			+
E. coli O121	B08	EU-RL			+
E. coli O121	E202/10	SMI			+
E. coli O121	E203/10	SMI			+
E. coli O121	E238/09	SMI			+
E. coli O121	H17/97	SMI			+

[a] +, detection.

Table 7. Strains used for exclusivity real-time PCR tests and results.

Organism	Reference	O104 detection[a]	H4 detection[a]	O121 detection[a]
Bacillus anthracis	4429	-	-	-
B. anthracis	7702	-	-	-
B. cereus	B. cereus	-	-	-
Campylobacter coli	SLV-271	-	-	-
C. jejuni	SLV-542	-	-	-
Enterobacter cloceae	SLV-011	-	-	-
Enterococcus durans	SLV-078	-	-	-
Escherichia coli	B266	-	-	-
E. coli	S262	-	-	-
E. coli	SLV-O82	-	-	-
E. coli	U226	-	-	-
E. coli	XL-1 blue	-	-	-
E. coli EIEC	121	-	-	-
E. coli O113:H21	98NK2	-	-	-
E. coli O157	EDL933	-	-	-
E. coli O157	SLV-479	-	-	-
E. coli O157:H-	493/89	-	-	-
E. coli O26:H11	H2954/06	-	-	-
Francisella tularensis	T8	-	-	-
Fusarium culmorum	F.c	-	-	-
Fusarium graminearum	F.g	-	-	-
Klebsiella pneumonaie	SLV-186	-	-	-
Listeria ivanovii	SLV-348	-	-	-
L. monocytogones	SLV-444	-	-	-
L. monocytogones	SLV-513	-	-	-
Proteus mirabilis	SLV-374	-	-	-
Psuedomonas aeruginosa	SLV-395	-	-	-
P. aeruginosa	SLV-453	-	-	-
Salmonella dublin	SLV-242	-	-	-
S. typhimurium	SLV-248	-	-	-
Shigella boydii	33/08	-	-	-
S. dysenterieae	15/08	-	-	-
S. flexneri	100/08	-	-	-
S. sonnei	99/08	-	-	-
Staphylococcus aureus	SLV-438	-	-	-
S. xylosus	SLV-283	-	-	-
Vibrio cholarae	CCUG 4070	-	-	-
V. parahaemolyticus	CCUG 4224	-	-	-
V. vulnificus	CCUG 16397	-	-	-
Yersinia enterocolitica	SLV-408	-	-	-
Y. pestis	KIM	-	-	-
Y. pseudotuberculosis	TAVA81	-	-	-

[a] -, no detection.

4.3.2 Limit of Detection

The result of the LOD determination is presented with serogroup O121 as an example in Table 8. The results for the other serogroups are listed in appendix 1. The limit of detection was 3.12 gene copies or the O104 assay and 6.25 gene copies for both the H4 and O121 assays. The acceptable LOD was settled to be 10 gene copies and below and all assays had acceptable values.

Table 8. LOD of serogroup O121. The LOD was determined to be the lowest concentration where all replicates were detected.

Gene copies	Detection[a]
$5 \cdot 10^5$	+ (18/18)
$5 \cdot 10^4$	+ (18/18)
$5 \cdot 10^4$	+ (18/18)
$5 \cdot 10^2$	+ (18/18)
50	+ (18/18)
25	+ (18/18)
12.5	+ (18/18)
6.25	+ (18/18)
3.12	- (16/18)
1.56	- (15/18)
0.78	- (8/18)
0.39	- (5/18)

[a] +, detection in all replicates; -, all replicates not detected. The parenthesis specifies the number of detected replicates of the total 18 replicates.

4.3.3 Precision

The repeatability study for O104 and O121 show that the standard variation and the coefficient of variation both increase as the number of gene copies in the reaction decreases as expected (Table 9 and 10). For H4 the correlation between the standard deviation and the concentration is not as apparent (Table 9).

Table 9. Repeatability of detection for the O104 and H4 real-time PCR assays.

	O104		H4	
Gene copies per reaction	C_q value[a]	CV (%)	C_q value[a]	CV (%)
10^6	19.76 ± 0.079	0.4	18.32 ± 0.173	0.9
10^4	26.38 ± 0.046	0.2	25.76 ± 0.508	2.0
10^2	32.95 ± 0.300	0.9	32.95 ± 0.222	0.7
50	33.90 ± 0.258	0.8	35.33 ± 0.736	2.1
25	35.16 ± 0.424	1.2	37.27 ± 0.678	1.8
12.5	36.04 ± 0.517	1.4	38.71 ± 1.067	2.8
6.25	36.78 ± 0.587	1.6	37.90 ± 0.728	1.9
3.12	38.35 ± 0.690	1.8	ND	

[a] Mean C_q value ± standard deviation of the results of one PCR run with six replicates. ND, not detected.

Table 10. Repeatability of detection for the O121 real-time PCR assay.

Gene copies per reaction	C_q value[a]	CV (%)
$5 \cdot 10^5$	17.23 ± 0.067	0.4
$5 \cdot 10^4$	23.90 ± 0.093	0.4
50	30.90 ± 0.287	0.9
25	31.84 ± 0.240	0.8
12.5	32.95 ± 0.423	1.3
6.25	34.71 ± 0.739	2.1

[a] Mean C_q value ± standard deviation of the results of one PCR run with six replicates.

4.3.4 Efficiency

The efficiency of the PCRs for all three serotypes was >90% which is acceptable given predetermined guidelines (Table 11). The R^2 values also were within the acceptable guideline (R^2>0.99).

Table 11. Summary of the results from the efficiency study.

Serotype	Efficiency (%)	R^2 [a]	Slope[a]
O104	99	0.9991	-3.34
H4	93	0.9989	-3.51
O121	98	1	-3.37

[a] Obtained from the linear regression curve of the common logarithm of the gene copy number plotted against the C_q value.

Figure 3 shows the curve obtained from linear regression for serotype O104. The results of the linear regression of the other serotypes are presented in appendix 2.

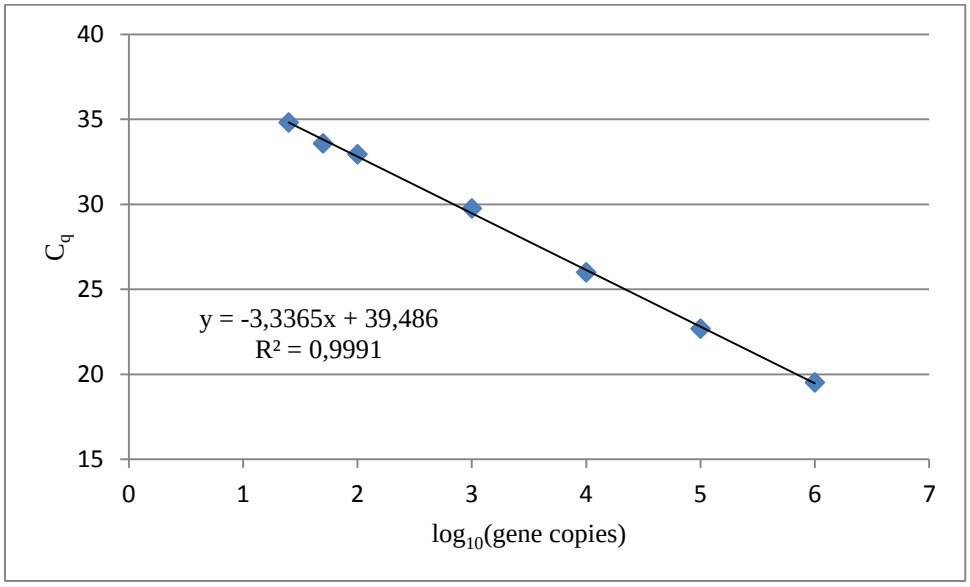


Figure 3. Efficiency curve for O104. The common logarithm of the gene copy number is plotted against the mean C_t value of the three replicates for each concentration.

3.4.5 Robustness

In the robustness study the assay for O104 produced a small increase in the C_q values for annealing temperatures higher than the standard and the variation was little for every condition except for when the concentrations were increased and the temperature was 64.3°C (Table 12). The H4 assay in general produced increased C_q values and variation when the annealing temperatures deviated from the standard (Table 13). The O121 assay was not affected by changes in either annealing temperature or reaction mixture concentration (Table 14). The O121 assay was however only tested on three annealing temperatures.

Table 12. Results of the robustness study for O104.

	Standard reaction mixture		+20% reaction mixture		-20% reaction mixture	
Temperature (°C)	C_q value[a]	CV (%)	C_q value[a]	CV (%)	C_q value[a]	CV (%)
56	32.77±0.199	0.6	32,70±0.237	0.7	32.69±0.166	0.5
58	32.69±0.217	0.7	32,85±0.181	0.5	32.63±0.154	0.5
60 (standard)	32.95±0.341	1.0	32.66±0.312	1.0	32.78±0.294	0.9
62.4	33.43±0.217	0.7	33,55±0.632	1.9	33.18±0.181	0.5
64.3	34,03±0.349	1.0	35.00±2.356	6.7	33.58±0.191	0.6

[a] Mean C_q value ± standard deviation of 10 replicates with 100 gene copies per reaction.

Table 13. Results of the robustness study for H4.

	Standard reaction mixture		+20% reaction mixture		-20% reaction mixture	
Temperature (°C)	C_q value[a]	CV (%)	C_q value[a]	CV (%)	C_q value[a]	CV (%)
56	34.97±0.747	2.1	36.52±1.092	3.0	34.05±0.890	2.6
58	33.38±0.391	1.2	34.27±0.308	0.9	32.94±0.262	0.8
60 (standard)	32.87±0.275	0.8	33.29±0.285	0.9	32.64±0.330	1.0
62.4	32.73±0.603	1.8	32.68±0.213	0.7	34.11±0.898	2.6
64.3	33.55±0.430	1.3	33.15±0.385	1.2	36.59±1.525	4.2

[a] Mean C_q value ± standard deviation of 10 replicates with 100 gene copies per reaction.

Table 14. Results of the robustness study for O121.

	Standard reaction mixture		+20% reaction mixture		-20% reaction mixture	
Temperature (°C)	C_q value[a]	CV (%)	C_q value[a]	CV (%)	C_q value[a]	CV (%)
50.1	31.09±0.312	1.0	31.41±0.333	1.1	31.73±0.163	0.5
52.0 (standard)	31.17±0.240	0.8	31.24±0.257	0.8	31.64±0.344	1.1
54.4	31.21±0.281	0.9	31.09±0.246	0.8	31.56±0.191	0.6

[a] Mean C_q value ± standard deviation of 10 replicates with 50 gene copies per reaction.

4.5 Melting curve analysis using SYBR Green

The melting curves for O104 and H4 are presented in figure 4 and 5, respectively. The melting curve for O121 is presented in appendix 3. Unspecific products were detected for both O104 and H4. For O104 there were two small peaks for the NTC samples around 70ºC which indicates that the PCR

produced primer dimers. For H4 there were apparent peaks for the two NTCs around 71ºC. There were also smaller peaks around the same temperature for the DNA template samples. This is an indication that there were unspecified products generated in the H4 PCR assay as well.

No unspecific product was detected for the O121 assay in any of the NTCs. One replicate of 10^0 gene copies per reaction did however generate a peak at a lower temperature than the other replicates with template DNA.

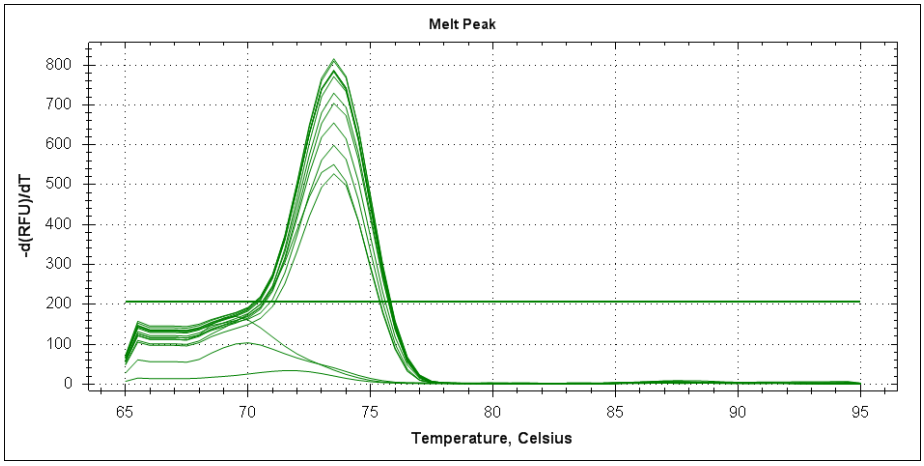


Figure 4. Melting curve of serogroup O104. The x axis represents the temperature and the y axis represents the negative derivative of the fluorescence divided by the derivative of the temperature. The peaks represent positive controls with the template DNA in concentrations ranging from 10^6 to 10^0 gene copies per reaction and two NTCs. The three smaller peaks represent the two NTCs and 10^0 gene copies, and the higher peaks represent the template DNA.

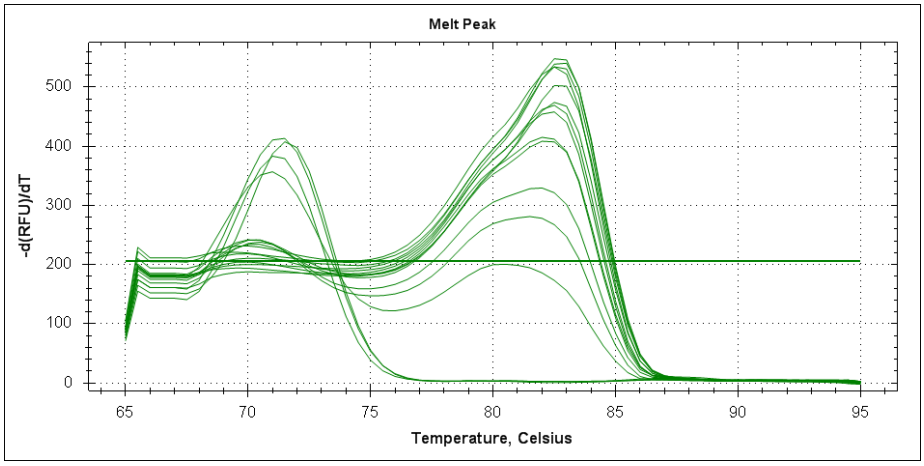


Figure 5. Melting curve of serogroup H4. The temperature is plotted against the negative derivative of the fluorescence divided by the derivative of the temperature. The peaks represent positive controls with the template DNA in concentrations ranging from 10^6 to 10^0 gene copies per reaction and two NTCs. The higher peaks around 71ºC represent the two NTCs and two replicates of 10^0 gene copies. The other peaks represent the other concentrations of the template DNA.

5. Discussion

The conventional PCR for O121 proved to be unreliable and was removed from the study and replaced by a real-time PCR. Real-time PCR is more sensitive and less time-consuming than conventional PCR and therefore real-time PCR is preferred to conventional PCR. The reason to why the conventional PCR did not perform as expected is unclear. It could be due to unspecific primers or that the optimal reaction parameters or reaction mixture concentrations were not found. The reason could also be malfunctions in the PCR performing equipment although using two different machines makes this less likely.

The validation of the real-time PCR assays included determination of the precision, specificity, LOD, efficiency and robustness. Since the validated assays are qualitative this means that certain parameters are more important than others in the study. A qualitative assay needs to be very sensitive and able to detect the target DNA in low concentrations. The precision however is of less concern in this study but very important when the assays are quantitative. The strains used in the inclusivity tests for O104, H4 and O121 were all detected in the assays. This implies that the assays are able to detect the intended target. None of the assays detected any of the strains used in the exclusivity study which means that there was no cross-reaction of the PCR assays with these strains. In addition the primers and probes proved to be specific for the target DNA when searching the sequences in BLAST. This indicates that the assays are specific and detects only the intended target DNA.

There are recommendations that state that the inclusivity study should include at least 50 different strains of the same serotype and that the exclusivity study should be performed on a minimum of 30 strains (32). Consequently the inclusivity of the assays has to be tested further to ensure that they detect the target DNA. The specificity is especially important when performing the PCR assays on real food samples since DNA from a variety of other organism will be present in the samples. The specificity therefore also needs to be evaluated using food matrices that consist of products that are often associated with STEC contamination, such as ground beef and raw milk. It is essential that the assays do not detect any of the DNA present in the matrices since this would give a false positive result.

An acceptable value of the LOD was determined to be 10 gene copies per reaction and below and each of the assays satisfied this demand. The LOD of serogroup O104 was 3.12 gene copies per reaction which is close to the theoretically lower limit of LOD (3 gene copies per reaction). The LOD is an important parameter in qualitative assays and is a measurement of how sensitive the assays are. STEC is often present in very small amounts in food samples and the fact that all assays gave acceptable results is satisfactory. However, the method is meant to be used on food samples containing DNA from several different organisms and it is therefore important to determine the LOD when using food matrices with added target DNA to simulate a real contaminated food sample.

In the repeatability study the PCR assays had CV values below 3% with H4 obtaining the overall highest CV values. As mentioned above the precision of a qualitative study is not a key parameter but the repeatability of a qualitative assay should give an estimate of the precision near the LOD (24). The precision generally decreases with decreasing concentration and for O104 and O121 the coefficient of variation peaks at the LOD. The results of the repeatability obtained for all serogroups are considered acceptable.

The efficiency of all three assays was acceptable given the predetermined requirement for the efficiency (within the range of 90-100%). The functions, obtained by linear regression, that were used to calculate the efficiency all had R^2 values > 0.99. This means that the calculated efficiencies are

reliable. The PCR assay for H4 was the lowest (93%) but still within the acceptable range. The efficiency always varies between repeated tests and this is due to small variations in the execution of the test, e.g. pipetting.

In the robustness study the assay for O104 generates a higher C_q value when the annealing temperature is increased. The assay for H4 generates a higher C_q value both when annealing is increased and decreased. An increased C_q value means that a greater number of cycles in the PCR is required in order for the fluorescent signal to be detected. This could thus lead to an increased LOD and lower sensitivity. In general the variation increases when the C_q value is increased. The instrument used for the real-time PCR reactions, CFX96™ Real-Time PCR Detection System (Bio-Rad), guarantees a temperature variation no greater than ±0.4ºC. Hence the scenario that the annealing temperature is changed with ±2ºC or more, as the annealing temperatures used in the robustness study, is highly unlikely. In addition the instrument also has a safety mechanism which decreases the temperature to 4ºC if there is more than a ±3ºC change. At 4ºC the PCR reaction is stopped. The O121 assay was not affected by changes in the temperature and none of the assays generated any remarkable change in the C_q values when the reaction mixture concentrations were changed. There was on the other hand only three temperatures tested for the O121 assay but, as mentioned, it is unlikely that the temperature will change that much.

In the melt curve analysis both O104 and H4 strongly indicates that unspecific products are formed in the PCR reactions. For H4 this was evident both in the samples with template DNA and in the NTCs, but for O104 the unspecific products were only detectable in the NTCs. The presence of unspecific products suggests that the primer and probe design is inadequate. If the primers are not specific enough they could potentially produce amplicons that are not the intended one. The unspecific products can also be primer dimers produced because the primers are complementary to each other. Unspecific products formation could have a negative impact on the efficiency of the PCR assays and could be the reason to why the assay for H4 shows a lower efficiency value. However, since the efficiency was within acceptable values this was not considered to be of any concern. The slightly decreased repeatability of the H4 assay could also possibly be explained by the formation of unspecific products.

6. Conclusions

The aim of this study was to validate methods for detecting STEC O104:H4 and O121 in food. The validation of the real-time PCR assays indicates that the methods are sensitive and specific. There is however additional tests that have to be performed in order to guarantee that the assays generate accurate results. Namely, the inclusivity have to be tested on additional strains of the target DNA, the exclusivity have to be evaluated on real food samples and the sensitivity of the method have to be determined by testing the assays on food matrices.

The ability to detect STEC O104:H4 in food is currently important for food business operators producing sprouts because of the European Commission regulation (No 209/2013), and this study is an indication that the real-time PCR assays recommended by the EU-RL generates accurate results. As O121 is more common in Sweden than in other EU countries, the need for a method for detection of the serotype is higher in Sweden. The real-time PCR assay for O121, validated in this study, shows great promise for being used to detect STEC O121.

7. References

1. Statens Serum Institut. 2011. VTEC O104 outbreak in Germany. Available: http://www.ssi.dk/English/News/EPI-NEWS/2011/No%2027-33%20-%202011.aspx, downloaded 2013-05-19.

2. EU Reference Laboratory, Istituto Superiore di Sanità. 2013. Detection and identification of Verocytotoxin-producing *Escherichia coli* (VTEC) O104:H4 in food by Real Time PCR. Available: http://www.iss.it/binary/vtec/cont/EU_RL_VTEC_Method_04_Rev_0.pdf , downloaded 2013-05-23.

3. Denamur E. 2011. The 2011 Shiga toxin-producing *Escherichia coli* O104:H4 German outbreak: a lesson in genomic plasticity. Clinical Microbiology and Infection **17**: 1124-1125.

4. Caprioli A, Morabito S, Brugère H, Oswald E. 2005. Enterohaemorrhagic *Escherichia coli*: emerging issues on virulence and modes of transmission. Veterinary Research **36**: 289-311.

5. Kaper JB, Nataro JP, Mobley HLT. 2004. Pathogenic *Escherichia coli*. Nature reviews. Microbiology **2**: 123-140.

6. Lin A, Sultan O, Lau HK, Wong E, Hartman G, Lauzon CR. 2011. O serogroup specific real time PCR assays for the detection and identification of nine relevant non-O157 STECs. Food Microbiology 28: 478-483.

7. Smittskyddsinstitutet. 2013. Statistik för Enterohemorragisk *E. coli* infektion (EHEC). Available: http://www.smittskyddsinstitutet.se/statistik/enterohemorragisk-e-coli-infektion-ehec/?t=com#statistics-nav, downloaded 2013-05-29.

8. ISO/TS 13136:2012. Microbiology of food and animal feed -- Real-time polymerase chain reaction (PCR)-based method for the detection of food-borne pathogens -- Horizontal method for the detection of Shiga toxin-producing Escherichia coli (STEC) and the determination of O157, O111, O26, O103 and O145 serogroups.

9. EU Reference Laboratory, Istituto Superiore di Sanità. 2013. Identification of the VTEC serogroups mainly associated with human infections by conventional PCR amplification of O-associated genes. Available: http://www.iss.it/binary/vtec/cont/EU_RL_VTEC_Method_003_Rev_0.pdf, downloaded 2013-05-29.

10. Carlson K, Linder C. 2008. Introduktion till mikrobiologi. Upplaga 1:1. Pozcal, Poland.

11. Garcia MI, Bouguénec CL. 1996. Role of adhesion in pathogenicity of human uropathogenic and diarrhoeogenic *Escherichia coli*. Bulletin de l'Institut Pasteur **94**: 201-236.

12. Gyles CL. 2007. Shiga toxin-producing *Escherichia coli*: An overview. Journal of Animal Science **85**: E45-E62.

13. Smittskyddsinstitutet. 2012. Sjukdomsinformation om Enterohemorragisk *E. coli*-infektion (EHEC). Available: http://www.smittskyddsinstitutet.se/sjukdomar/enterohemorragisk-e-coli-infektion/, downloaded 2013-05-22.

14. Livsmedelsverket, Statens Jordbruksverk, Statens veterinärmedicinska anstalt, Smittskyddsinstitutet, Socialstyrelsen och Naturvårdsverket. 2007. Verotoxinbildande E. coli – VTEC-bakteriers smittvägar, förekomst samt risker för folkhälsan.

15. Hussein HS. 2007. Prevalence and pathogenicity of Shiga toxin-producing *Escherichia coli* in beef cattle and their products. Journal of Animal Science **85**: E63-E72.

16. Kubista M, Andrade JM, Bengtsson M, Forootan A, Jonák J, Lind K, Sindelka R, Sjöback R, Sjögreen B, Strömbom L, Ståhlberg A, Zoric N. 2006. The real-time polymerase chain reaction. Molecular Aspects of Medicine **27**: 95-125.

17. Mathews CK, van Holde KE, Ahern KG. 1999. Biochemistry. 3rd edition. Pearson Education, Upper Saddle River, NJ.

18. Lodish HF, Berk A, Kaiser CA, Krieger M, Scott MP, Bretscher A, Ploegh H. 2008. Molecular cell biology. 6th edition. W.H. Freeman, New York.

19. Pavlov AR, Pavlova NV, Kozyavkin SA, Slesarev AI. 2004. Recent developments in the optimization of the thermostable DNA polymerases for efficient applications. Trends in Biotechnology **22**: 253–260.

20. Brown TA. 2002. Genomes. 2nd edition. Wiley-Liss, Oxford.

21. Houghton SG, Cockerill FR. 2006. Real-time PCR: Overview and applications. Surgery **139**: 1-5.

22. Bustin SA, Benes V, Garson JA, Hellemans J, Hugget J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT. 2009. The MIQE Guidelines: Minimum Information of Quantitative Real-Time PCR Experiments. Clinical Chemistry **55:4**: 611-622.

23. Sloan LM. 2007. Real-Time PCR in Clinical Microbiology: Verification, Validation, and Contamination Control. Clinical Microbiology Newsletter **29**: 87-95.

24. Burd EM. 2010. Validation of Laboratory-Developed Molecular Assays for Infectious Diseases. Clinical Microbiology Reviews **23**: 550-576.

25. Macherey-Nagel. 2014. Genomic DNA from tissue.User manual.Nucleopin tissue. Available: http://www.mn-net.com/Portals/8/attachments/Redakteure_Bio/Protocols/Genomic%20DNA/UM_gDNATissue.pdf , downloaded 2014-01-18.

26. Qiagen. 2014. DNeasy Blood and Tissue Handbook. Available: http://mvz.berkeley.edu/egl/resources/product%20inserts/DNeasy_Blood_&_Tissue_Handbook.pdf , downloaded 2014-01-18.

27. Bio-rad. 2014. InstaGene matrix. Available: http://www.bio-rad.com/webroot/web/pdf/lsr/literature/Bulletin_9289.pdf , downloaded 2014-01-18.

28. Monday SR, Beisaw A, Feng PC. 2007. Identification of Shiga toxigenic *Escherichia coli* seropathotypes A and B by multiplex PCR. Molecular and Cellular Probes **21**: 308-311.

29. Bugarel M, Beutin L, Martin A, Gill A, Fach P. 2010. Micro-array for the identification of Shiga toxin-producing *Escherichia coli* (STEC) seropathotypes associated with Hemorrhagic Colitis and Hemolytic Uremic Syndrome in humans. International Journal of Food Microbiology **142**: 318-329.

30. Lin A, Sultan O, Lau HK, Wong E, Hartman G, Lauzon CR. 2011. O serogroup specific real time PCR assays for the detection and identification of nine relevant non-O157 STECs. Food Microbiology **28**: 478-483.

31. Thermo Scientific. 2013. T042 – Technical bulletin. NanoDrop Spectrophotometers. Assessment of Nucleic Acid Purity. Available: http://www.nanodrop.com/Library/T042-NanoDrop-Spectrophotometers-Nucleic-Acid-Purity-Ratios.pdf , downloaded 2013-05-31.

32. NordVal. 2009. Protocol for the validation of alternative microbiological methods. Available: http://www.nmkl.org/NordVal/NordValprotocolmarch2009.pdf, downloaded 2013-05-31.

8. Appendix

Appendix 1: LOD of the PCR assays for O104 and H4

Table 15. LOD of O104 and H4. The LOD was determined to be the lowest concentration that was detected for every replicate.

O104		H4	
Gene copies	Detection	Gene copies	Detection
10^6	+ (18/18)	10^6	+ (18/18)
10^5	+ (18/18)	10^5	+ (18/18)
10^4	+ (18/18)	10^4	+ (18/18)
10^3	+ (18/18)	10^3	+ (18/18)
10^2	+ (18/18)	10^2	+ (18/18)
50	+ (18/18)	50	+ (18/18)
25	+ (18/18)	25	+ (18/18)
12.5	+ (18/18)	12.5	+ (18/18)
6.25	+ (18/18)	6.25	+ (18/18)
3.12	+ (18/18)	3.12	- (17/18)
1.56	- (11/18)	1.56	- (17/18)
0.78	- (11/18)	0.78	- (8/18)

[a] +, detection in all replicates; -, all replicates not detected. The parenthesis specifies the number of detected replicates of the total 18 replicates.

Appendix 2 : Efficiency curve for the H4 and O121 assays

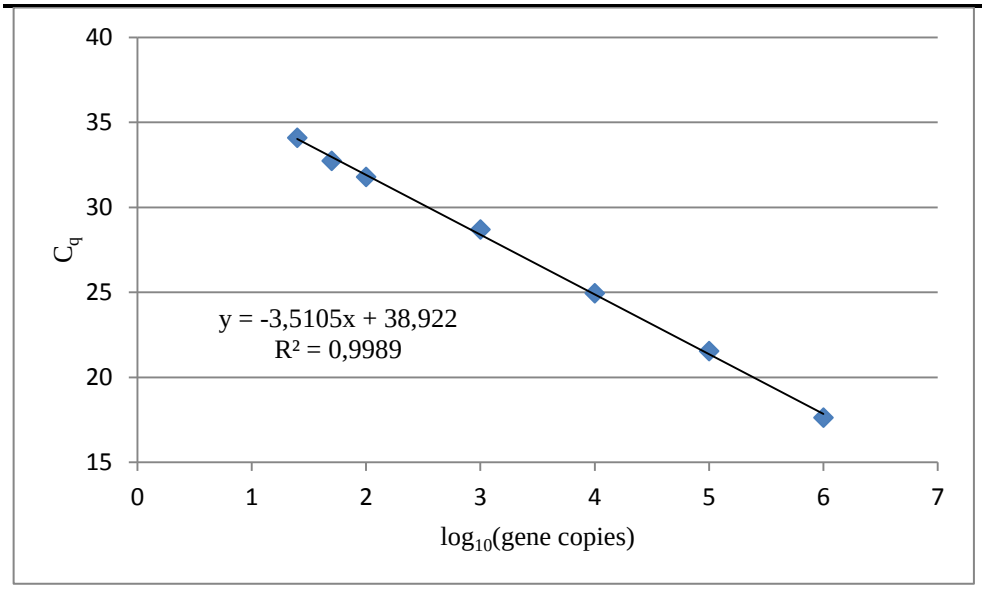


Figure 6. Efficiency curve for H4. The common logarithm of the gene copy number is plotted against the mean C_t value of the three replicates for each concentration.

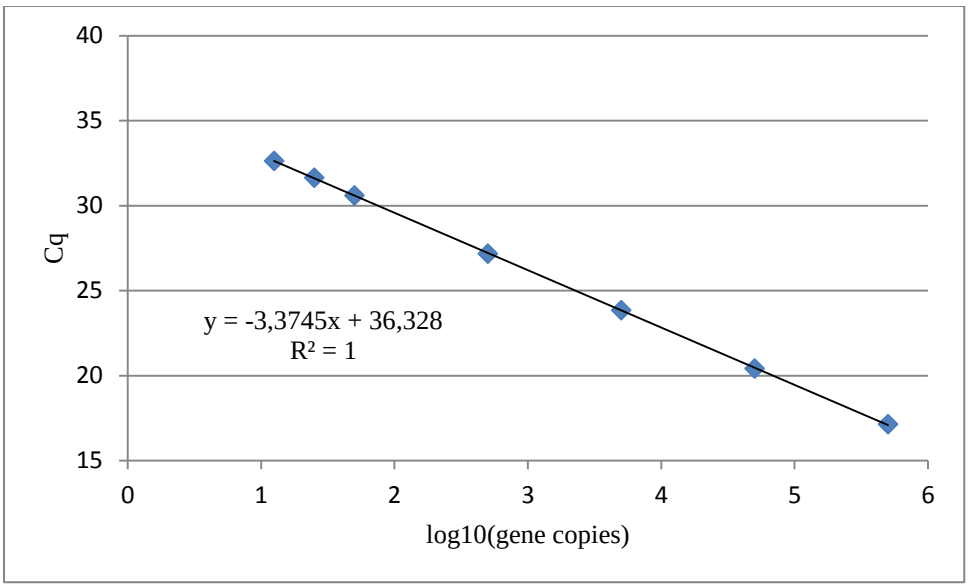


Figure 7. Efficiency curve for O121. The common logarithm of the gene copy number is plotted against the mean C_t value of the three replicates for each concentration.

Appendix 3: Melting curve for the O121 assay

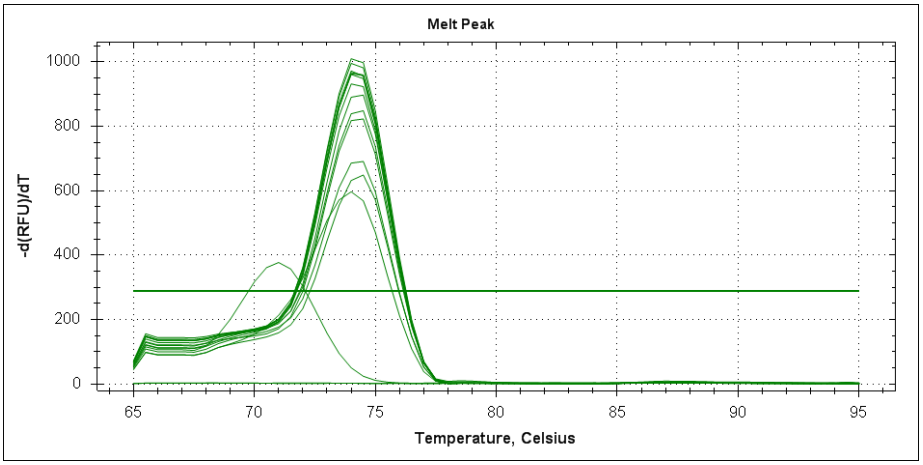


Figure 8. Melting curve of serogroup O121. The temperature is plotted against the negative derivative of the fluorescence divided by the derivative of the temperature. The peaks represent positive controls with the template DNA in concentrations ranging from 10^6 to 10^0 gene copies per reaction and two NTCs. The small peak represents the template DNA in 10^0 gene copies.