Detection of enterohaemorrhagic *Escherichia coli* in food by droplet digital PCR to detect simultaneous virulence factors in a single genome

Li He[1,2#], David Simpson[1*#], Michael G. Gänzle[1,3]

[1]University of Alberta, Department of Agricultural, Food and Nutritional Science

[2]Sichuan Agricultural University, College of Food Science, Ya'an, Sichuan, P.R. China.

[3] Hubei University of Technology, College of Bioengineering and Food Science, Wuhan, Hubei, P.R. China.

# These authors contributed equally

*Corresponding author: djsimpso@ualberta.ca

410 Agriculture/Forestry Centre, University of Alberta, Edmonton, Alberta, Canada T6G 2P5

Abstract

Shiga toxin producing *E. coli* are a problem for food producers. STEC's require a combination of virulence factors to cause disease, so ideally detection techniques should detect the presence of multiple virulence factors in a single cell directly from food. Droplet Digital PCR (ddPCR) is commonly used to quantify the number of copies of a gene in a sample, moreover it is able to link two genes to the same piece of DNA. Here *stx* and an O-antigen specific gene are detected simultaneously with taqman probes confirming that the cells are intact as well as distinguishing between strains based on their genotype. Using ddPCR *E. coli* O157:H7 and O104:H4 are quantified from apple juice, milk and spinach washings without an enrichment step, the detection limit of ddPCR in apple juice was 2 cfu/ml. Also, ddPCR was used to detect pathogenic bacterial cells in the presence of background strains which shared one or none of the target genes, including avirulent strains. Whole cell ddPCR is compared to several DNA extraction techniques demonstrating that whole cell ddPCR is more reliable for linking genes within an organism. Whole cell ddPCR is a promising technique for the rapid and specific detection of foodborne pathogens.

Keywords: Droplet digital PCR, EHEC, Pathogen detection, Shiga toxin,

1 Introduction

Shiga toxin producing *Escherichia coli* (STEC) are a major concern for food safety. STEC cause more than 300,000 cases of foodborne illness annually in the United States (Scallan et al., 2011) and are estimated to cost $965 million annually (Scharff, 2012).

Detection methods that simultaneously detect multiple virulence factors would be preferable for detection of STECs since the ability of STECs to cause disease in humans is dependent on the presence of multiple virulence factors; moreover, these virulence factors often have variants or alternatives that serve the same function. For instance, *E. coli* produce two principal variants of Shiga toxin, *stx1* and *stx2*, with *stx2* being the more toxic of the two (Jackson et al., 1987). STECs are commonly found in ruminant guts (Gyles, 2007) where the prophage acquired Shiga toxin (Croxen et al., 2013) poisons protozoan predators (Lainhart et al., 2009). The abundance of STEC in ruminant guts contributes strongly to the pervasiveness as a foodborne contaminant (Etcheverría and Padola, 2013). The pathogenesis of STEC also depends on additional virulence factors, including intimin, the translocated intimin receptor and a type 3 secretion system encoded on the Locus of Enterocyte Effacement (LEE) (Etcheverría and Padola, 2013). Intimin is responsible for adhesion to the epithelial cells but strains lacking the LEE may use other adhesion factors such as an autoagglutinating adhesion Saa (Paton et al., 2001). The genes involved in virulence can vary between strains and STECs often possess at least a few of additional genes that contribute to their virulence. Strains of *E. coli* that harbor the Shiga toxin in combination with a large virulence plasmid that allows invasion of epithelial cells are classified as *Shigella* spp. and cause shigellosis (Belotserkovsky and Sansonetti, 2018). Strains of *E. coli* that harbor the Shiga toxin encoding *stx* in combination with LEE cause the hemolytic uremic syndrome (HUS) and are termed enterohaemorrhagic *E. coli* (EHEC).

In the U.S., *E. coli* O157:H7 accounted for 36% of foodborne EHEC infections and approximately 65% of *E. coli* O157:H7 outbreaks in the United States were transmitted through food (Heiman et al., 2015). STECs are not only transmitted through beef and dairy products but also through wildlife contaminated produce and wheat flour or through contaminated water (Erickson and Doyle, 2007). Other *E. coli* serotypes that contribute to foodborne disease in the U.S. include the "Big Six" Shiga toxin serotypes, O26, O45, O103, O111, O121, and O145 (Brooks et al., 2005). Because a majority of foodborne EHEC infections are caused only by few serotypes, the U.S. regulates STEC on the basis of their serotypes rather than the presence of (a combination of) virulence factors (USDA Food Safety and Inspection Service, 2012). A positive PCR for any of the seven serotypes that are regulated necessitates follow-up by culture-dependent methodology to determine whether the strains also carry one or several genes coding for the Shiga toxin.

Strains of *E. coli* that harbor the Shiga toxin in combination with aggregative adhesion fimbria (AAF) also cause the HUS and have been termed enteroaggregative hemorrhagic *E. coli* (EAHEC) (Croxen and Finlay, 2009; Frank et al., 2011). *E. coli* O104:H4 caused one case of EAHEC infection in South Korea in 2005 (Bae et al., 2006); EAHEC became notorious when a strain with the same serotype caused an outbreak linked to fenugreek sprouts (Buchholz et al., 2011), which resulted in 3842 illnesses including 53 deaths in 2011 (Beutin and Martin, 2012). Th outbreak strain of *E. coli* O104:H4 has three plasmids coding for virulence factors. The β-lactamases TEM-1 and CTX-M-15 are encoded by the pEC_Bactec plasmid, while different versions of the AAF are present on the 55989p plasmid and the P042 plasmid.

Detection of pathogenic *Escherichia coli* in food is challenging since pathogenic and non-pathogenic strains of this species are differentiated only by the presence or absence of a specific combination of virulence factors. *E. coli* has been used as an indicator for fecal contamination,

however this status has recently been called into question with the discovery of free living environmental *E. coli* (Berthe et al., 2013; Ratajczak et al., 2010; Walk et al., 2007). Moreover, the distribution of virulence factors onto both plasmids and the chromosome creates difficulties in linking these virulence factors to the same strain without isolating the strain as pure culture.

Droplet digital polymerase chain reaction (ddPCR) is a technique which creates an emulsion of PCR reaction mixture in oil resulting in approximately 20 000 droplets before performing PCR and measuring the positive and negative droplets. Since ddPCR detects the presence or absence of a gene in 20 000 individual droplets it is likely that ddPCR is less susceptible to contaminants that effect PCR efficiency where reducing PCR efficiency would reduce the amount of bacteria quantified using qPCR. It has been primarily used for absolute quantification of gene copy numbers (Hindson et al., 2011), however, ddPCR has also been used to detect contaminants in food (Morisset et al., 2013) as well as STEC in bovine feces (Verhaegen et al., 2016) and viral DNA in clinical serum (Hayden et al., 2013; Strain et al., 2013). ddPCR is less sensitive to inhibition from matrices than comparable techniques such as qPCR and provides greater reproducibility (Hindson et al., 2013). Because ddPCR has its PCR reaction occurring inside a droplet if a droplet contained a single cell any target genes amplified must have been present in the cell. In this way it is possible to establish whether two genes are present in the genome of the same cell. As outlined above, the ability to determine whether two virulence factors are present on the same bacterial genome is critical for determination of the virotype of *E. coli*. By including bacterial cells instead of DNA in the reaction mixture, ddPCR has been used for the detection of multiple virulence factors in cells of *E. coli* (McMahon et al., 2017). Because of its improved robustness over qPCR ddPCR may be an excellent

This study aimed to provide proof of concept of the ability of ddPCR to simultaneously detect two genes in a single cell of *E. coli*, and to use intact cells ddPCR for detection of STEC in different food matrices. The suitability of the method was assessed by using EHEC *E. coli* O157:H7 and EAHEC *E. coli* O104:H4 as model organisms which share common genes such as *stx2* allow the specificity of the technique to be investigated.

2 Materials and methods

2.1 Growth and Preparation of *E. coli* strains

Three *E. coli* strains were selected on the basis of their virulence genes and serotypes: *E. coli* O104:H411-3088 (*stx2* and wzy_{O104}); *E. coli* O157:H7 LCDC 7236 (*stx1*, *stx2*, and wzy_{O157}) and *E. coli* DH5α, a reference strains without *stx1*, *stx2*, wzy_{O157} or wzy_{O104}. Cells of the strains were prepared by transferring growth from a single colony from Luria-Bertani (LB) agar into 5mL of LB broth and growing overnight at 37°C. The overnight culture was centrifuged at 5000×g for 10min and the cell pellet was resuspended in 1mL of LB broth. The cell counts of the strains were approximately 5×10^9cfu/mL, which was determined by duplicate plating of serial dilutions on LB agar. The strains used to generate mixed-cultures in both LB broth and in food were grown separately, diluted then mixed at the specified ratio.

2.2 Primers and Probes

Primers and probes were designed by the Primer3plus online tool and produced by IDT (Coralville, Iowa) **(Table 1**). The primers and probes were designed to amplify targets for pathogen detection, *stx*1, *stx*2, wzy_{O104} and wzy_{O157}, or DNA sequences that are 1kb, 10kb, 50kb, 100kb, 1000kb or 2000kb from *stx2* in the O104:H4 genome (**Figure 1)**. The gene coding for the O-antigen assembly

polymerase, *wzyE*, is present in all strains of *E. coli* and was used as positive control. Primers and probes were rehydrated to stock concentrations of 100μM with 1X Tris-EDTA pH8.0 (TE).

2.3 Extraction of Genomic DNA and Distance Testing

Cells from 1mL of overnight culture were used to extract the genomic DNA (gDNA). DNA was extracted manually with the phenol-chloroform method and with two different commercial kits, Qiagen Blood & Tissue DNA Kit (Qiagen Inc – Toronto, Canada) and Promega Genomic DNA Purification Kit (Promega Madison WI), as described in the manufacturer's instructions for Gram-negative bacteria.

For DNA extraction with the phenol-chloroform method, cells were lysed by bead beating using a BioSpec Mini-BeadBeater-8 and 0.1 mm Zirconia/Silica Beads for 3 minutes (BioSpec Products, Bartlesville, OK). DNA was then extracted twice with an equal volume of phenol/chloroform/isoamylalcohol (25:24:1), followed by an equal volume of chloroform, and precipitated in isopropanol.

All gDNA preparations were quantified by NanoDrop™ One (Thermofisher Scientific Inc.) and diluted to 10μg/L.

To determine the fragmentation of DNA with different isolation methods, multiplex ddPCR targeted genes with increasing distance from *stx2* on the genome of *E. coli* O104:H4 (**Figure 1**). Each PCR reaction was conducted with template DNA prepared with different DNA isolation methods, and with intact cells. The percent linkage between *stx2* and other genes with different distances in different templates are shown in **Figure 2**.

2.4 Optimization of the Multiplexed Intact cell (*E. coli* O104:H4) ddPCR Method

Gradient ddPCR was performed to determine optimal temperature for the primers and probes. Optimal distinction between the fluorescence amplitude intensity of positive and negative droplets was observed at annealing temperatures between 58°C and 60°C for the *stx2* and wzy_{O104} assays. An annealing temperature of 58°C was selected for subsequent experiments.

2.5 Detection of *E. coli* O104:H4 and O157:H7 in the presence of background *E. coli* strains

To determine the sensitivity of ddPCR for detection of *E. coli* O104:H4 in the presence of other strains of *E. coli*, *E. coli* O104:H4 was mixed with *E. coli* DH5α, *E. coli* O157:H7 or both at different ratios (**Table 2**). Mixed cultures were inactivated at 60 °C for 15 minutes as a precaution for safely generating droplets with intact pathogenic bacterial cells. One µL of inactivated culture was used as template in 25µL of ddPCR reaction mixture.

To evaluate the performance of the multiplexed intact cells ddPCR assay for distinguishing *E. coli* O157:H7 in the presence of multiple background strains, overnight cultures of *E. coli* DH5α and *E. coli* O104:H4 were mixed with *E. coli* O157:H7 at different ratios (Table 2). The combined cell count of *E. coli* strains was approximately 10^7 cfu/mL in all reactions. The targets *stx1*, *stx2*, wzy_{O157} and *wzyE* were used to detect *E. coli* strains with the following combinations of primers and probes: *stx1* and wzy_{O157}; *stx2* and wzy_{O157}; *stx1* and *wzyE*; and *stx2* and *wzyE*. *E. coli* O157:H7 harbors all four target genes; *stx2* is additionally present in *E. coli* O104:H4 and *wzyE* is present in all strains of *E. coli* (Figure 5).

2.6 Detection of *E. coli* O104:H4 in food matrix with different proportions of background strains

In order to detect *E. coli* O104:H4 in a food matrix, cells of *E. coli* O104:H4 and *E. coli* O157:H7 were mixed at a log ratio of 5:5 (10^5 cfu/ml of *E. coli* O104:H4 to 10^5 cfu/ml of *E. coli* O157:H7) or 5:6 and added 1ml of the mixture directly into 50mL of either apple juice or milk, or directly

onto 50g of fresh spinach leave. After the cells were added to the spinach the they were gently massaged into the spinach manually For apple juice and milk, 1μL of the inoculated samples was used directly as the ddPCR template, whereas the inoculated spinach was washed with 50mL LB broth by inverting the tube 50 times, 1μL of which was used as the ddPCR template. Uninoculated apple juice, milk and spinach wash were used for the respective negative controls.

To determine the detection limit of *E. coli* O104:H4 in the presence of a high microbial background, approximately 10^2 cfu of *E. coli* O104:H4 and 10^4 cfu of *E. coli* O157:H7 were inoculated into 50mL apple juice. Aliquots of 50mL of the inoculated apple juice were centrifuged at 5000×g for 10min. The pellet was resuspended in 1mL LB broth, and centrifuged at 5000×g for 10 minutes. Finally, the pellet was resuspended in 100μL LB broth and inactivated at 60 °C for 15 minutes prior to transferring 1μL to a 25μL reaction system. This experiment was performed with 22 independent replicates.

2.7 Droplet Digital PCR (ddPCR)

Droplet digital PCR was performed according to McMahon et al. 2017 with modifications (McMahon et al., 2017). In brief, the 25μl PCR reaction consisted of 1X supermix for probes (Bio-Rad), 1.8μM each of *stx1* and *stx2* primers; other primers were used at 2.25μM. Probes for *stx1* and *stx2* were used at 0.5μM and all other probes were used at 0.625μM. The PCR cycle was as follows: 95°C for 5min initial denaturation; 40 cycles of 95°C for 20s and between 52°C and 62°C for 60s; followed by 94°C for 10min and then held at 4°C. In the reactions with restriction enzyme 2.5μl of BamHI FastDigest (ThermoFisher Scientific Inc.) was used and an 80°C enzyme inactivation step was added to the beginning of the PCR cycle. The droplets were read using the QX200TM ddPCR system (Bio-Rad Laboratories, Inc.) A cutoff of no fewer than 10000 droplets was used for all

183 samples (Pinheiro et al., 2012).The association between the *stx* and *wzy* or other targets was
184 assessed by the "% linkage" value, which was calculated as follows:

185 $$\%\ linkage = \ 100\ (\frac{\#\ of\ droplets\ positive\ for\ target\ A\ and\ B}{\#\ of\ droplets\ positive\ for\ target\ A})$$

186 where A and B represent the two genes targeted by the assay.

187 All PCR reactions with *E. coli* inoculated in food employed un-inoculated food samples (apple
188 juice, milk, or spinach) as negative controls.

189 **3 Results**

190 3.1 Comparison of the Percent Linkage of *stx2* and other Genes in *E. coli* O104:H4

191 Extracted DNA and whole cells were used to determine if linkages of distant genes was possible
192 using extracted DNA. Amplification of all targets was observed with both whole cell and gDNA
193 extracted templates (Figure 2 and data not shown). When the target genes were 1 kbp apart, co-
194 amplification of both targets was observed when whole cells or DNA obtained with all of the four
195 DNA isolation methods was used as template (**Figure 2**). When the target genes were 10 or 50 kbp
196 apart, co-amplification of both targets was observed when whole cells or DNA obtained with the
197 Genomic DNA isolation kit was used as template while the % linkage obtained with other template
198 DNA was low **(Fig. 2a)**. When the target genes were more than 100 kbp apart, co-amplification was
199 observed only when whole cells were used as template **(Figure 2).** To assess the possibility that the
200 tertiary structure of the chromosomal DNA was preventing access to the target genes, the percent
201 linkage of the digested DNA through the integration of the restriction enzyme BamHI in the ddPCR
202 mixture was compared to the heat-treated cells (**Figure 2a**). The addition of the restriction enzyme
203 digestion did not result in significant increase in the percent linkage between the *stx* and other

targets. These results demonstrate that DNA isolation shear DNA to fragments that are sized between 10-50 kbp; intact cell ddPCR but not PCR with isolated gDNA is suitable to detect whether two genes are present on the same chromosome.

3.2 Evaluation of the ddPCR method in the presence of background strains

To determine whether the presence of closely related strains of *E. coli* interferes with ddPCR amplification of the target genes, or the % linkage of the two target genes, *E. coli* O104:H4 was mixed with other *E. coli* strains. Other *E. coli* strains were added to exceed the cell counts of the target strain up to 10,000 fold, i.e. exceeding the ratio of target to background that is expected in food such as meat or produce where the total cell count of *E. coli* is less than 10^4 cfu/g. In addition, a target to background ratio of 1:20,000 is the lowest ratio that can be detected in digital PCR systems that measure 20,000 droplets. When *E. coli* O104:H4 was mixed with the non-pathogenic *E. coli* DH5α **(Figure 3),** *E. coli* O104:H4 was detectable even when the ratio of target to background was 1:10,000. The abundance of the background strain had no effect on the detected copy number of either wzy_{O104} or *stx2*. Moreover, the % linkage was essentially unaffected by the background strain.

E. coli O104:H4 was also mixed with *E. coli* O157:H7 (**Figure 4**) in order to determine the accuracy of ddPCR when a background strain possesses one of the two genes being detected, in this case *stx2*. Once again the background strain had no apparent effect on the detected gene copy numbers. The gene copy number determined for wzy_{O104} matched the cell count of *E. coli* O104:H4 while the gene copy number determined for *stx2* matched the cell counts of (*E. coli* O104:H4 + O157:H7). The % linkage was calculated on the basis of wzy_{O104} varied when the two strains were used in a different ratio but values remained above 10%. The % linkage on the basis of *stx2* was not calculated as this target is also present in *E. coli* O157:H7 which does not carry wzy_{O104}.

3.3 Detection of *E. coli* O157:H7 in LB broth with *E. coli* O104:H4 and *E. coli* DH5α

E. coli O157:H7 was mixed with *E. coli* DH5α and *E. coli* O104:H4 The gene copies determined for each of the targets matched the cell counts of the respective target strains. The percent linkage of *stx1* and wzy_{O157} was clearly dependent on the overall cell count and the presence of other strains (**Figure 5)**. The % linkage was highest when low cell counts of *E. coli* O157:H7 were used as template and remained above 10% in all reactions **(Figure 5)**.

3.4 Detection of *E. coli* O104:H4 in food

E. coli O104:H4 was mixed with *E. coli* O157:H7in milk, apple juice and on spinach in order to determine the sensitivity of ddPCR in a food matrix. No positive droplets were detected in ddPCR reactions that used uninoculated milk, apple juice, or spinach as template, excluding interference from the food matrix. Surprisingly, the % linkages were higher in apple juice and milk (**Figure 6**) than in LB (**Figure 4**), however, the % linkage for the spinach wash samples was much lower. The gene copies of the respective targets corresponded to the number of cells containing the respective genes in the reaction. The presence of background strains did not affect the detection of *E. coli* O104:H4.

To determine the detection limit of the ddPCR away in liquid foods, 10^2 cfu/ml *E. coli* O104:H4 and 10^4 cfu/ml *E. coli* O157:H7 were added to 50ml of apple juice giving a concentration of 2 cfu/ml and 200 cfu/ml respectively and centrifuged and resuspended in 100µl. The concentration of cells from 50ml apple juice to 100µl LB resulted in a cell count of approximately 1 cfu/µl of *E. coli* O104:H4. One µl of this suspension was used in a ddPCR reaction using the *stx2* and wzy_{O104} primers and probes. Twenty one of the twenty two repeats generated droplets that were positive for

both *stx2* and wzy_{O104} (**Table 3**); one sample was positive only for wzy_{O104}. The assay thus detects *E. coli* O104:H4 in liquid food if the strain is present at a cell count of 2 cells/mL or higher.

4 Discussion

Digital PCR is an effective technique for the detection of bacterial DNA, both because of its precision and its specificity, especially in the presence of food matrices (Hindson et al., 2013; Morisset et al., 2013). In addition, digital PCR differs from other PCR methods in that each template molecule is amplified in a separate reaction chamber, which is generated by water droplets emulsified in oil or a microfluidics device (Hindson et al., 2011; Ottesen et al., 2006). This allows distinction between two target genes that are present on the same DNA strand, or in the same bacterial cell (McMahon et al., 2017). A previous study provided proof of concept that the ddPCR allows detection of EHEC in food enrichment cultures (McMahon et al., 2017); this study expands prior knowledge by detecting *E. coli* using serotype, detecting EAHEC in addition to EHEC and by performing direct detection of both EHEC and EAHEC in foods.

The virulence of most pathogenic strains of *E. coli* depends on the simultaneous presence of virulence factors; in particular, enterotoxigenic *E. coli* and EHEC cause severe disease only when the same genome encodes for adhesion factors and toxins (Croxen et al., 2013). In the United States and several other jurisdiction, however, pathogenic *E. coli* are regulated on the basis of their serotype; regulatory compliance therefore necessitates the detection of the serotype (USDA Food Safety and Inspection Service, 2012). The serotype of *E. coli* strains is readily determined by sequencing or PCR amplification of the O-antigen specific *wzy* genes (Pintara et al., 2018). This study employed two combinations of toxin genes and *wzy* genes to detect EHEC in foods, however, the assay is readily modified to any other combination of genes, provided that both genes are located on the same strand of template DNA, or in the same cell. The comparison of ddPCR with

intact cells as the template to reactions performed with isolated gDNA demonstrated that the distance between two genes on the chromosome influences the result when gDNA is used as the template but not when intact cells are employed in the reaction. ddPCR detects DNA fragmentation when two genes with a known distance from each other are amplified (Han et al., 2019). The fragment size of gDNA that we determined using ddPCR matched the read length achieved with long read sequencing (Figure 2 and Jain et al., 2018) suggesting that DNA extraction may be the limitation for that technique as well. Intact cell ddPCR but not PCR with gDNA thus enables detection of the presence of two genes in the same cell. Conversely, assays that aim to determine the copy number of genes on bacterial chromosomes should employ highly sheared gDNA (Bhat et al., 2009), e.g. DNA prepared by bead beating, which fragments DNA to template molecules sized between 1 and 10 kbp (Figure 2).

Several virulence factors of *E. coli* including the AAF are encoded on plasmids (Croxen et al., 2013). In addition, the temperate Shiga-toxin phage in STEC spontaneously or upon induction by stress reverts to the lytic cycle (Livny and Friedman, 2004; Łoś et al., 2009). The production of phage particles that contain *stx* genes may interfere with the % linkage as determined by ddPCR. *E. coli* O104:H4 11-3088 harbors an Stx2 prophage; *E. coli* O157:H7 LCDC 7236 harbors one Stx 1 prophage and one Stx2 prophage (Mercer et al., 2015). The induction of Shiga-Toxin prophages by stress differs in these two strains (Zhang et al., 2018). In the present study, the % linkage obtained with *stx* genes was comparable to the % linkage obtained with *wzy* genes which are not present in phage genomes. This result demonstrates that phage particles, or the presence of multiple copies of a target in the same cell do not interfere with intact cell ddPCR.

The % linkage determined in this study with EHEC and EAHEC ranged from 10% to 80%. Values for the % linkage that are below the theoretical value of 100% likely reflect the presence of DNA

from permeabilized cells in overnight cultures, or cell lysis during sample preparation. The % linkage obtained with single strains was higher when compared to the % linkage in mixed cultures, indicating that pipetting and centrifugation steps prior to droplet generation compromise cell integrity (compare Figure 2 to Figures 3-6). Sample preparation from spinach involved most pipetting, mixing, and centrifugation steps, which may account for these samples having the lowest values for the % linkage (Figure 6). The heating step that was included to eliminate the necessity to conduct ddPCR with intact cells in a biosafety level 1 laboratory (this study) did not decrease the % linkage when compared to a previous study with EHEC (McMahon et al., 2017). Conversely, the decrease of the % linkage during sample preparation may suggests that intact cell ddPCR only detectes intact cells that were viable at the time of sample preparation, which is not readily achieved with other PCR methods (Laidlaw et al., 2019; Sidari and Caridi, 2011).

McMahon *et al.* used ddPCR to detect bacteria in enrichment broths (McMahon et al., 2017) but this study is the first time that ddPCR has been used to detect bacteria directly from food. Concentration of bacterial cells by centrifugation increased the sensitivity of intact cell ddPCR to 2 cfu/mL; however, regulations require detection of *E. coli* O157 at a level of negative in 25mL to negative in 325g of sample (Canada Food Inspection Agency, 2019; USDA Food Safety and Inspection Service, 2019). Current methods for detection of EHEC all employ an enrichment step, which serves not only to increase the number of viable cells but also to ensure that only viable cells are detected (Canada Food Inspection Agency, 2019; Rohde et al., 2017; USDA Food Safety and Inspection Service, 2019; Wolffs et al., 2005)**.** The requirement for enrichment, however, also greatly extends the analysis time. While PCR assays including ddPCR are completed in less than 90 min, a single enrichment or a combination of selective and non-selective enrichments add 12-48 h to the analysis time (Law et al., 2015; Rohde et al., 2017). Reducing the requirement for enrichment

will thus greatly accelerate current methods even if the enrichment step cannot be completely eliminated. The ability of the ddPCR method as employed in the current study to detect less than 1000 cells per mL may allow for reduced enrichment time, and thus reduce the overall time required for detection. Rapid analysis is an important criterion for detection of EHEC in perishable foods including ground beef and fresh produce that are shipped only after confirmation that they are not contaminated by EHEC. Ultimately, intact cell ddPCR may allow a more accurate detection of pathogenic *E. coli* in comparison to current methods. Currently, EHEC contamination of foods is regulated on the basis of the O-antigen in the United States, and only the presence of *E. coli* is tested in the EU (Commission of the European Communities, 2005), however the virulence of individual strains is determined by a combination of virulence factors (González-Escalona and Kase, 2019). *E. coli* cause HUS only if they possess *stx1, stx2* or both in combination with virulence factors that mediate adhesion to the intestinal mucosa, most commonly the LEE (Schmidt, 2010) or, in the case of EAHEC, AAF (Brzuszkiewicz et al., 2011). While *E. coli* O157:H7 and the so-called "Big Six" non-O157 STECs account for most North American foodborne outbreaks linked to EHEC, not all strains carrying these O-antigens are EHEC as they lack *stx* or adhesion factors (Delannoy et al., 2013; Gamage et al., 2003), and not all EHEC belong to one of the 7 serotypes that are regulated in the U.S (Mathusa et al., 2010). Intact cell ddPCR enables the simultaneous detection of two virulence genes in a single cell, something that normally requires isolation of a strain into a pure culture, and thus more accurately predicts the potential of strain to cause severe disease in humans. Currently, ddPCR equipment is limited to simultaneous detection of only two targets, and current equipment does not allow multiplexing of PCR reactions on the basis of melt curves. These limitations, however, can currently be addressed by using

multiple PCR reactions per samples, and may be eliminated by improved design of ddPCR equipment.

5. Conclusions

This study demonstrated that intact cell ddPCR accurately detects EHEC in foods; in particular, intact cell ddPCR allows for the unprecedented determination of whether two genes (coding for virulence factors) are present in the same bacterial cell. Because the ability of EHEC to cause disease in humans is dependent on the simultaneous presence of multiple virulence factors in the genome of a single cell, ddPCR for simultaneous detection of virulence factors, or of *stx* and the serotype will greatly facilitate the accurate detection of pathogenic *E. coli* in foods.

Acknowledgements

We are grateful to Dr. Urmila Basu for her support and advice with using the ddPCR equipment. Alberta Agriculture and Forestry and China Scholarship Council are acknowledged for financial support; Michael Gänzle acknowledges the Canada Research Chairs Program for funding.

References

Bae, W.K., Lee, Y.K., Cho, M.S., Ma, S.K., Kim, S.W., Kim, N.H., Choi, K.C., 2006. A case of hemolytic uremic syndrome caused by *Escherichia coli* O104:H4. Yonsei Med. J. 47, 437–439. https://doi.org/10.3349/ymj.2006.47.3.437

Belotserkovsky, I., Sansonetti, P.J., 2018. *Shigella* and Enteroinvasive *Escherichia coli*, in: Frankel, G., Ron, E.Z. (Eds.), *Escherichia Coli*, a Versatile Pathogen. Springer International Publishing, Cham, pp. 1–26. https://doi.org/10.1007/82_2018_104

Berthe, T., Ratajczak, M., Clermont, O., Denamur, E., Petit, F., 2013. Evidence for coexistence of distinct *Escherichia coli* populations in various aquatic environments and their survival in estuary water. Appl. Environ. Microbiol. 79, 4684 LP – 4693. https://doi.org/10.1128/AEM.00698-13

Beutin, L., Martin, A., 2012. Outbreak of Shiga toxin–producing *Escherichia coli* (STEC) O104:H4 infection in Germany causes a paradigm shift with regard to human pathogenicity of STEC strains. J. Food Prot. 75, 408–418. https://doi.org/10.4315/0362-028X.JFP-11-452

Bhat, S., Herrmann, J., Armishaw, P., Corbisier, P., R.Emslie, K., 2009. Single molecule detection in

366 nanofluidic digital array enables accurate measurement of DNA copy number. Anal. Bioanal. Chem.
367 394, 457–467. https://doi.org/DOI 10.1007/s00216-009-2729-5

368 Brooks, J.T., Sowers, E.G., Wells, J.G., Greene, K.D., Griffin, P.M., Hoekstra, R.M., Strockbine, N.A., 2005.
369 Non-O157 Shiga toxin–producing *Escherichia coli* infections in the United States, 1983–2002. J. Infect.
370 Dis. 192, 1422–1429. https://doi.org/10.1086/466536

371 Brzuszkiewicz, E., Thürmer, A., Schuldes, J., Leimbach, A., Liesegang, H., Meyer, F.-D., Boelter, J., Petersen,
372 H., Gottschalk, G., Daniel, R., 2011. Genome sequence analyses of two isolates from the recent
373 *Escherichia coli* outbreak in Germany reveal the emergence of a new pathotype: Entero-Aggregative-
374 Haemorrhagic *Escherichia coli* (EAHEC). Arch. Microbiol. 193, 883–891.
375 https://doi.org/10.1007/s00203-011-0725-6

376 Buchholz, U., Bernard, H., Werber, D., Böhmer, M.M., Remschmidt, C., Wilking, H., Deleré, Y., an der
377 Heiden, M., Adlhoch, C., Dreesman, J., Ehlers, J., Ethelberg, S., Faber, M., Frank, C., Fricke, G., Greiner,
378 M., Höhle, M., Ivarsson, S., Jark, U., Kirchner, M., Koch, J., Krause, G., Luber, P., Rosner, B., Stark, K.,
379 Kühne, M., 2011. German outbreak of *Escherichia coli* O104:H4 associated with sprouts. N. Engl. J.
380 Med. 365, 1763–1770. https://doi.org/10.1056/NEJMoa1106482

381 Canada Food Inspection Agency, 2019. Preventive controls for *E. coli* O157/NM in raw beef products
382 [WWW Document]. URL http://www.inspection.gc.ca/food/requirements-and-guidance/preventive-
383 controls-food-businesses/meat/raw-beef-products/eng/1541538060346/1541538137261#fn1

384 Commission of the European Communities, 2005. Working document: Commission regulation on
385 microbiological criteria for foodstuffs. SANCO/4198/2001 Rev. 19 1–40.

386 Croxen, M.A., Finlay, B.B., 2009. Molecular mechanisms of *Escherichia coli* pathogenicity. Nat. Rev.
387 Microbiol. 8, 26.

388 Croxen, M.A., Law, R.J., Scholz, R., Keeney, K.M., Wlodarska, M., Finlay, B.B., 2013. Recent advances in
389 understanding enteric pathogenic *Escherichia coli*. Clin. Microbiol. Rev. 26, 822–880.

390 Delannoy, S., Beutin, L., Fach, P., 2013. Discrimination of Enterohemorrhagic *Escherichia coli* (EHEC) from
391 non-EHEC strains based on detection of various combinations of type III effector genes. J. Clin.
392 Microbiol. 51, 3257–3262. https://doi.org/10.1128/JCM.01471-13

393 Erickson, M.C., Doyle, M.P., 2007. Food as a vehicle for transmission of Shiga toxin–producing *Escherichia*
394 *coli*. J. Food Prot. 70, 2426–2449. https://doi.org/10.4315/0362-028X-70.10.2426

395 Etcheverría, A.I., Padola, N.L., 2013. Shiga toxin-producing *Escherichia coli*. Virulence 4, 366–372.
396 https://doi.org/10.4161/viru.24642

397 Frank, C., Werber, D., Cramer, J.P., Askar, M., Faber, M., an der Heiden, M., Bernard, H., Fruth, A., Prager,
398 R., Spode, A., Wadl, M., Zoufaly, A., Jordan, S., Kemper, M.J., Follin, P., Müller, L., King, L.A., Rosner, B.,
399 Buchholz, U., Stark, K., Krause, G., 2011. Epidemic profile of Shiga Toxin producing *Escherichia coli*
400 O104:H4 outbreak in Germany. N. Engl. J. Med. 365, 1771–1780.
401 https://doi.org/10.1056/NEJMoa1106483

402 Gamage, S.D., Strasser, J.E., Chalk, C.L., Weiss, A.A., 2003. Nonpathogenic *Escherichia coli* can contribute to
403 the production of Shiga toxin. Infect. Immun. 71, 3107–3115. https://doi.org/10.1128/IAI.71.6.3107-
404 3115.2003

405 González-Escalona, N., Kase, J.A., 2019. Virulence gene profiles and phylogeny of Shiga toxin-positive

406 *Escherichia coli* strains isolated from FDA regulated foods during 2010-2017. PLoS One 14, e0214620.
407 https://doi.org/10.1371/journal.pone.0214620

408 Gyles, C.L., 2007. Shiga toxin-producing *Escherichia coli*: An overview. J. Anim. Sci. 85, E45--E62.
409 https://doi.org/10.2527/jas.2006-508

410 Han, J., Lee, J.Y., Bae, Y.-K., 2019. Application of digital PCR for assessing DNA fragmentation in cytotoxicity
411 response. Biochim. Biophys. Acta - Gen. Subj. 1863, 1235–1242.
412 https://doi.org/10.1016/j.bbagen.2019.05.001

413 Hayden, R.T., Gu, Z., Ingersoll, J., Abdul-Ali, D., Shi, L., Pounds, S., Caliendo, A.M., 2013. Comparison of
414 droplet digital PCR to real-time PCR for quantitative detection of cytomegalovirus. J. Clin. Microbiol.
415 51, 540–546. https://doi.org/10.1128/JCM.02620-12

416 Heiman, K.E., Mody, R.K., Johnson, S.D., Griffin, P.M., Gould, L.H., 2015. *Escherichia coli* O157 outbreaks in
417 the United States, 2003–2012. Emerg. Infect. Dis. 21, 1293.

418 Hindson, B.J., Ness, K.D., Masquelier, D.A., Belgrader, P., Heredia, N.J., Makarewicz, A.J., Bright, I.J., Lucero,
419 M.Y., Hiddessen, A.L., Legler, T.C., Kitano, T.K., Hodel, M.R., Petersen, J.F., Wyatt, P.W., Steenblock,
420 E.R., Shah, P.H., Bousse, L.J., Troup, C.B., Mellen, J.C., Wittmann, D.K., Erndt, N.G., Cauley, T.H.,
421 Koehler, R.T., So, A.P., Dube, S., Rose, K.A., Montesclaros, L., Wang, S., Stumbo, D.P., Hodges, S.P.,
422 Romine, S., Milanovich, F.P., White, H.E., Regan, J.F., Karlin-Neumann, G.A., Hindson, C.M., Saxonov,
423 S., Colston, B.W., 2011. High-throughput droplet digital PCR system for absolute quantitation of DNA
424 copy number. Anal. Chem. 83, 8604–8610. https://doi.org/10.1021/ac202028g

425 Hindson, C.M., Chevillet, J.R., Briggs, H.A., Gallichotte, E.N., Ruf, I.K., Hindson, B.J., Vessella, R.L., Tewari,
426 M., 2013. Absolute quantification by droplet digital PCR versus analog real-time PCR. Nat. Methods
427 10, 1003–1008. https://doi.org/10.1038/nmeth.2633

428 Jackson, M.P., Neill, R.J., O'Brien, A.D., Holmes, R.K., Newland, J.W., 1987. Nucleotide sequence analysis
429 and comparison of the structural genes for Shiga-like toxin I and Shiga-like toxin II encoded by
430 bacteriophages from *Escherichia coli* 933. FEMS Microbiol. Lett. 44, 109–114.
431 https://doi.org/10.1111/j.1574-6968.1987.tb02252.x

432 Jain, M., Koren, S., Miga, K.H., Quick, J., Rand, A.C., Sasani, T.A., Tyson, J.R., Beggs, A.D., Dilthey, A.T.,
433 Fiddes, I.T., Malla, S., Marriott, H., Nieto, T., O'Grady, J., Olsen, H.E., Pedersen, B.S., Rhie, A.,
434 Richardson, H., Quinlan, A.R., Snutch, T.P., Tee, L., Paten, B., Phillippy, A.M., Simpson, J.T., Loman,
435 N.J., Loose, M., 2018. Nanopore sequencing and assembly of a human genome with ultra-long reads.
436 Nat. Biotechnol. 36, 338–345. https://doi.org/10.1038/nbt.4060

437 Laidlaw, A.M., Gänzle, M.G., Yang, X., 2019. Comparative assessment of qPCR enumeration methods that
438 discriminate between live and dead *Escherichia coli* O157:H7 on beef. Food Microbiol. 79, 41–47.
439 https://doi.org/10.1016/j.fm.2018.11.002

440 Lainhart, W., Stolfa, G., Koudelka, G.B., 2009. Shiga toxin as a bacterial defense against a eukaryotic
441 predator, *Tetrahymena thermophila*. J. Bacteriol. 191, 5116–5122. https://doi.org/10.1128/JB.00508-
442 09

443 Law, J.W.-F., Ab Mutalib, N.-S., Chan, K.-G., Lee, L.-H., 2015. Rapid methods for the detection of foodborne
444 bacterial pathogens: principles, applications, advantages and limitations. Front. Microbiol. 5, 770.
445 https://doi.org/10.3389/fmicb.2014.00770

446 Livny, J., Friedman, D.I., 2004. Characterizing spontaneous induction of Stx encoding phages using a

selectable reporter system. Mol. Microbiol. 51, 1691–1704. https://doi.org/10.1111/j.1365-2958.2003.03934.x

Łoś, J.M., Łoś, M., Węgrzyn, G., Węgrzyn, A., 2009. Differential efficiency of induction of various lambdoid prophages responsible for production of Shiga toxins in response to different induction agents. Microb. Pathog. 47, 289–298. https://doi.org/http://dx.doi.org/10.1016/j.micpath.2009.09.006

Mathusa, E.C., Chen, Y., Enache, E., Hontz, L., 2010. Non-O157 Shiga toxin–producing *Escherichia coli* in foods. J. Food Prot. 73, 1721–1736. https://doi.org/10.4315/0362-028X-73.9.1721

McMahon, T.C., Blais, B.W., Wong, A., Carrillo, C.D., 2017. Multiplexed single intact cell droplet digital PCR (MuSIC ddPCR) method for specific detection of enterohemorrhagic *E. coli* (EHEC) in food enrichment cultures. Front. Microbiol. 8, 332. https://doi.org/10.3389/fmicb.2017.00332

Mercer, R.G., Zheng, J., Garcia-Hernandez, R., Ruan, L., Gänzle, M.G., McMullen, L.M., 2015. Genetic determinants of heat resistance in *Escherichia coli*. Front. Microbiol. 6, 932. https://doi.org/10.3389/fmicb.2015.00932

Morisset, D., Štebih, D., Milavec, M., Gruden, K., Žel, J., 2013. Quantitative analysis of food and feed samples with droplet digital PCR. PLoS One 8, 1–9. https://doi.org/10.1371/journal.pone.0062583

Ottesen, E.A., Hong, J.W., Quake, S.R., Leadbetter, J.R., 2006. Microfluidic digital PCR enables multigene analysis of individual environmental bacteria. Science (80-.). 314, 1464–1467. https://doi.org/10.1126/science.1131370

Paton, A.W., Srimanote, P., Woodrow, M.C., Paton, J.C., 2001. Characterization of Saa, a Novel Autoagglutinating Adhesin Produced by Locus of Enterocyte Effacement-Negative Shiga-ToxigenicEscherichia coli Strains That Are Virulent for Humans. Infect. Immun. 69, 6999–7009. https://doi.org/10.1128/IAI.69.11.6999-7009.2001

Pinheiro, L.B., Coleman, V.A., Hindson, C.M., Herrmann, J., Hindson, B.J., Bhat, S., Emslie, K.R., 2012. Evaluation of a droplet digital polymerase chain reaction format for DNA copy number quantification. Anal. Chem. 84, 1003–1011. https://doi.org/10.1021/ac202578x

Pintara, A.P., Guglielmino, C.J.D., Rathnayake, I.U., Huygens, F., Jennison, A. V., 2018. Molecular prediction of the O157:H-negative phenotype prevalent in Australian Shiga toxin-producing *Escherichia coli* cases improves concordance of *in silico* serotyping with phenotypic motility. J. Clin. Microbiol. 56. https://doi.org/10.1128/JCM.01906-17

Ratajczak, M., Laroche, E., Berthe, T., Clermont, O., Pawlak, B., Denamur, E., Petit, F., 2010. Influence of hydrological conditions on the *Escherichia coli* population structure in the water of a creek on a rural watershed. BMC Microbiol. 10, 222. https://doi.org/10.1186/1471-2180-10-222

Rohde, A., Hammerl, J.A., Boone, I., Jansen, W., Fohler, S., Klein, G., Dieckmann, R., Dahouk, S. Al, 2017. Overview of validated alternative methods for the detection of foodborne bacterial pathogens. Trends Food Sci. Technol. 62, 113–118. https://doi.org/https://doi.org/10.1016/j.tifs.2017.02.006

Scallan, E., Hoekstra, R.M., Angulo, F.J., Tauxe, R. V., Widdowson, M.-A., Roy, S.L., Jones, J.L., Griffin, P.M., Giffin, P., 2011. Foodborne illness acquired in the United States—major pathogens. Emerg. Infect. Dis. 17, 7–15. https://doi.org/10.3201/eid1701.P11101

Scharff, R.L., 2012. Economic burden from health losses due to foodborne illness in the United States. J. Food Prot. 75, 123–131. https://doi.org/10.4315/0362-028X.JFP-11-058

Schmidt, M.A., 2010. LEEways: tales of EPEC, ATEC and EHEC. Cell. Microbiol. 12, 1544–1552. https://doi.org/10.1111/j.1462-5822.2010.01518.x

Sidari, R., Caridi, A., 2011. Methods for detecting enterohaemorrhagic *Escherichia coli* in food. Food Rev. Int. 27, 134–153. https://doi.org/10.1080/87559129.2010.535232

Strain, M.C., Lada, S.M., Luong, T., Rought, S.E., Gianella, S., Terry, V.H., Spina, C.A., Woelk, C.H., Richman, D.D., 2013. Highly precise measurement of HIV DNA by droplet digital PCR. PLoS One 8, 1–8. https://doi.org/10.1371/journal.pone.0055943

USDA Food Safety and Inspection Service, 2019. Detection, Isolation and Identification of Top Seven Shiga Toxin-Producing *Escherichia coli* (STEC) from Meat Products and Carcass and Environmental Sponges [WWW Document]. URL https://www.fsis.usda.gov/wps/wcm/connect/7ffc02b5-3d33-4a79-b50c-81f208893204/MLG-5B.pdf?MOD=AJPERES (accessed 7.10.19).

USDA Food Safety and Inspection Service, 2012. Shiga toxin-producing Escherichia coli in certain raw beef products. 9 CFR, parts 416, 417, and 430. Fed. Regist. 77, 31975–31981.

Verhaegen, B., De Reu, K., De Zutter, L., Verstraete, K., Heyndrickx, M., Van Coillie, E., 2016. Comparison of droplet digital PCR and qPCR for the quantification of Shiga toxin-producing *Escherichia coli* in bovine feces. Toxins (Basel). 8. https://doi.org/10.3390/toxins8050157

Walk, S.T., Alm, E.W., Calhoun, L.M., Mladonicky, J.M., Whittam, T.S., 2007. Genetic diversity and population structure of *Escherichia coli* isolated from freshwater beaches. Environ. Microbiol. 9, 2274–2288. https://doi.org/10.1111/j.1462-2920.2007.01341.x

Wolffs, P., Norling, B., Rådström, P., 2005. Risk assessment of false-positive quantitative real-time PCR results in food, due to detection of DNA originating from dead cells. J. Microbiol. Methods 60, 315–323. https://doi.org/https://doi.org/10.1016/j.mimet.2004.10.003

Zhang, L., Simpson, D., McMullen, L., Gänzle, M., 2018. Comparative genomics and characterization of the late promoter pR' from Shiga Toxin prophages in *Escherichia coli*. Viruses 10, 595.

Figure legends

Figure 1. Location of target genes on the circular genome of *E. coli* O104:H4. A schematic presentation of the location of the primer and probe combination on the *E. coli* O104:H4 2011C-3493 genome (NC_018658.1). The *wzyE* primer/probe conbination was used for the 2000kbp marker.

Figure 2. Linkage between *stx2* and other genes with different distances. **Panel A:** % linkage between *stx2* and other genes with increasing distances using whole cells or gDNA extracted by different methods as template. Data are shown as means ± standard deviation of three independent replicates. **Panel B:** Representative 2D plots generated by the QuantaSoft analysis software for O104:H4 cells or DNA isolated with either the promega genomic DNA purification kit or phenol-chloroform, targeting genes either 1 kbp or 2000 kbp apart.

Figure 3. Percent linkage (**Panels A and C)** and copy number of *stx2* and wzy_{O104} **(Panel B and D)** determined with *E. coli* O104:H4 and with *E. coli* DH5α as background at different proportions. Experiments were conducted to maintain a constant cell count of *E. coli* DH5α **(Panels A and B)** or a constant cell count of *E. coli* O104 (**Panels C and D).** Data are shown as means ± standard deviation of three independent replicates.

Figure 4. Percent linkage (**Panels A and C)** and copy number of *stx2* and wzy_{O104} **(Panel B and D)** determined with *E. coli* O104:H4 and with *E. coli*O157H7 as background at different proportions. Data are shown as means ± standard deviation of three independent replicates.

Figure 5. Percent linkage **(Panel A)** and gene copy number (**Panel B)** at different proportions of strains *E. coli* O157:H7, *E. coli* O104:H4 and *E. coli* DH5α. The % linkage shown in **Panel A** was calculated as follows: ▨ (*stx1*+wzy_{O157})/*stx1*, ▭ (*stx1*+wzy_{O157})/wzy_{O157}, ▬ (*stx2*+wzy_{O157})/*stx2*,

(*stx2*+*wzy*$_{O157}$)/*wzy*$_{O157}$, (*stx2*+*wzyE*)/*stx2*, (*stx2*+*wzyE*)/*wzyE*. Bars in **Panel B** show the copy number of the following target genes using the primer probe pair in brackets: - *stx1*(*stx1*+*wzy*$_{O157}$), - *wzy*$_{O157}$(*stx1*+*wzy*$_{O157}$), - *stx2*(*stx2*+*wzy*$_{O157}$), - *wzy*$_{O157}$(*stx2*+*wzy*$_{O157}$), - *stx2*(*stx2*+*wzyE*), - *wzyE*(*stx2*+*wzyE*). Data are shown as means ± standard deviation of three independent replicates.

Figure 6. Detection of *E. coli* O104:H4 in food with *E. coli* O157:H7 as background at different proportions. **Panel A**, % linkage between *stx2* and *wzy*$_{O104}$; **Panel B**: copy number of *stx2;* **Panel C,** copy number of *wzy* $_{O104}$. Log ratio of *E. coli* O104:H4 and *E. coli* O157:H7 5:5, Log ratio of *E.coli* O104:H4 and *E. coli* O157:H7 5:6. Data are shown as means ± standard deviation of three independent replicates. The horizontal line indicates the detection limit of 10^3 gene copies / mL. Gene copy numbers in uninoculated apple juice, milk, and spinach were all below the detection limit.

Table 1: Primers and probes used in this study

Oligos	Sequence(5' → 3')	Amplicon Size (bp)
Stx2-F	ACTCTGACACCATCCTCT	118
Stx2-R	CACTGTCTGAAACTGCTC	
O104wzy-F	GCAATAGCCAATTTGCACAT	160
O104wzy-R	CCCGGGGCAATTATCATTAA	
Stx1-F	TTACCCCCTCAACTGCTAAT	101
Stx1-R	GTTTCCTTCTATGTGTCCGG	
O157wzy-F	CTAACAGTTCTGCTCCATAC	111
O157wzy-R	AGGGAATAAAGCATCAAGAC	
1kbp-F	GTCTCCCCCAACACGTAATA	101
1kbp-R	GACGATGACGCGATGATTAT	
10kbp-F	TCTTATCAATGGTGGGCTGT	140
10kbp-R	AGCCTCCTTGATTGGTGAT	
50kbp-F	CGTGCTTCATTCACGATTTG	159
50kbp-R	GCTGGTCTTCCAGACTTTAC	
100kbp-F	GTCCACGGCTCATCATTAAT	124
100kbp-R	CGCAAAAGACGGTCGTTA	
1000kbp-F	GAAGTGATAAATACCGAACGC	115
1000kbp-R	TCTCCCCATCAGGAAACG	
wzyE-F	GTGAGCCTATAAGCGTAG	92
wzyE-R	GAGTATGGTGCTGAACTC	
Stx2	56-FAM/ACATTGCTG/ZEN/ATTCGCCCCCAGTT/3IABkFQ	
Stx1	56-FAM/TGCGCATCA/ZEN/GAATTGCCCCCAGAGTGG/3IABkFQ	
O104wzy	5HEX/AGGAGGGTT/ZEN/CGGCACTGGACTTGGGTT/3IABkFQ	
O157wzy	5HEX/AACAAGTAT/ZEN/TTGGAGACATGGGAGC/3IABkFQ	
1kbp	5HEX/CCTGTTAAG/ZEN/GCGAGCCATACACCCGCA/3IABkFQ	
10kbp	5HEX/TGTACAGTC/ZEN/CACCGCCAGCGAGCTT/3IABkFQ	
50kbp	5HEX/TGCCAACGA/ZEN/ACACATGCCCTGCGCA/3IABkFQ	
100kbp	5HEX/ATGCACCAT/ZEN/GACGGACAGCCTGGCT/3IABkFQ	
1000kbp	5HEX/TGGCCGCTG/ZEN/TTTGAGATTGCACCGG/3IABkFQ	
wzyE	5HEX/CGCCAGTCC/ZEN/GGAGTGGTTATTCA/3IABkFQ	

Table 2: Ratio of *E. coli* strains used in ddPCR experiments

Strain	Log_{10} cfu/ml																					
E. coli O104:H4	7	6	5	4	5	5	5	5	5	7	7	6	5	4	5	6	5	5	5	7	3	7
E. coli DH5 α	7	7	7	7	-	6	7	8	9	-	-	-	-	-	-	-	-	-	-	7	7	3
E. coli O157:H7	-	-	-	-	-	-	-	-	-	6	7	7	7	7	-	5	5	6	7	3	7	7

Table 3: Number of positive droplets in detection of *E. coli* O104:H4 in apple juice with *E. coli* O157:H7 as background 1:100 ratio (Ch1-*stx2*, Ch2-wzy_{O104}) n = 22

	Average # of droplets	**Std. dev**	**# of negative reactions**
stx2 and wzy_{O104} positive	2.18	1.07	1 of 22
stx2 positive / wzy_{O104}negative	44.50	19.18	0 of 22
stx2 negative / wzy_{O104}positive	1.00	1.35	10 of 22

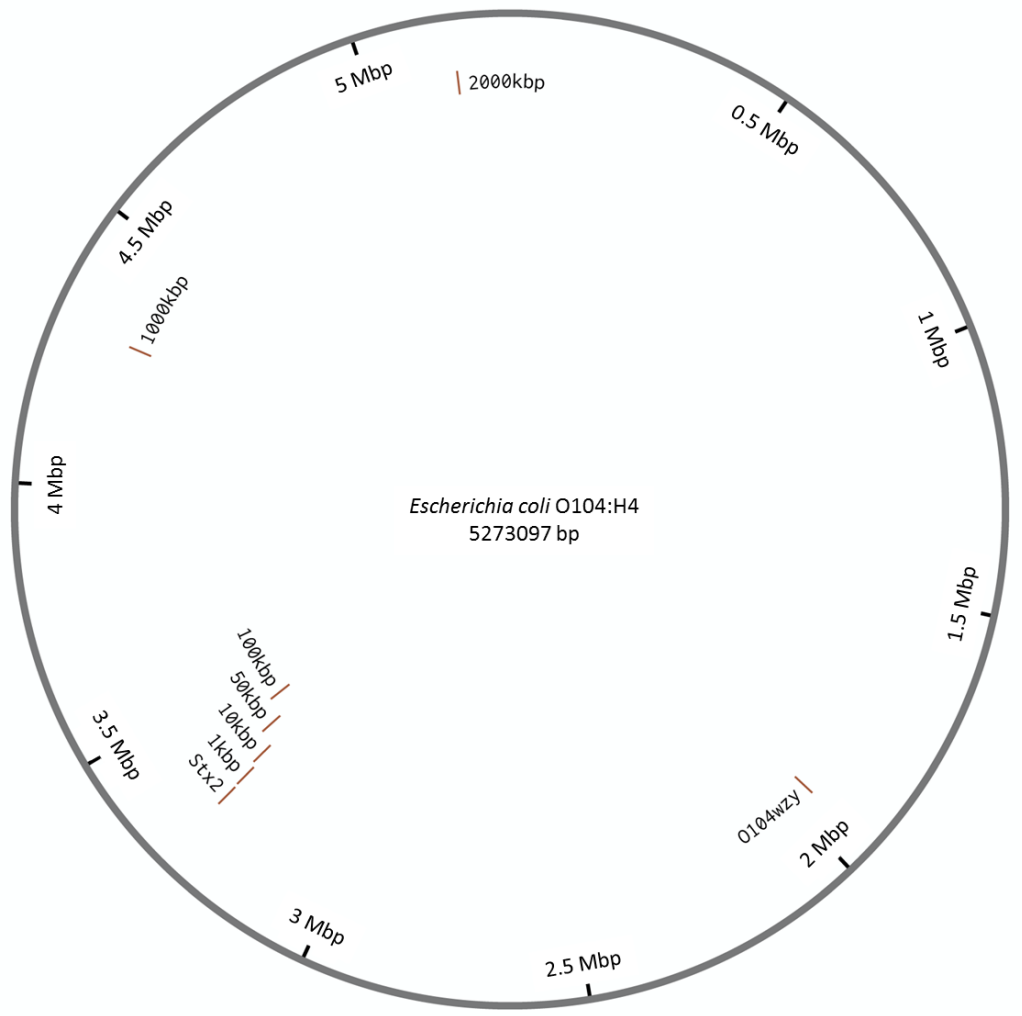


Figure 1: Location of target genes on the circular genome of *E. coli* O104:H4. A schematic presentation of the location of the primer and probe combination on the *E. coli* O104:H4 2011C-3493 genome (NC_018658.1). The *wzyE* primer/probe conbination was used for the 2000kbp marker.

Figure 2

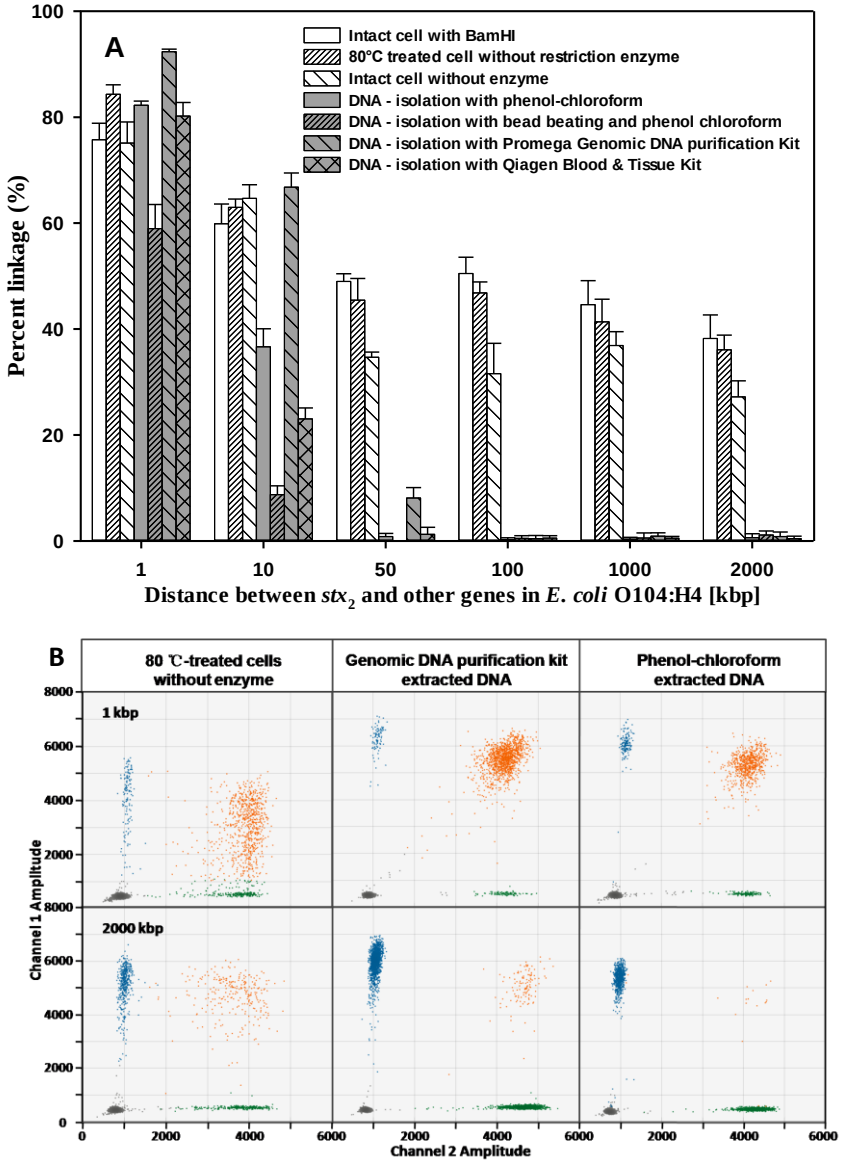


Figure2: Linkage between *stx2* and other genes with different distances. **Panel A:** % linkage between *stx2* and other genes with increasing distances using whole cells or gDNA extracted by different methods as template. Data are shown as means ± standard deviation of three independent replicates. **Panel B:** Representative 2D plots generated by the QuantaSoft analysis software for O104:H4 cells or DNA isolated with either the promega genomic DNA purification kit or phenol-chloroform, targeting genes either 1 kbp or 2000 kbp apart.

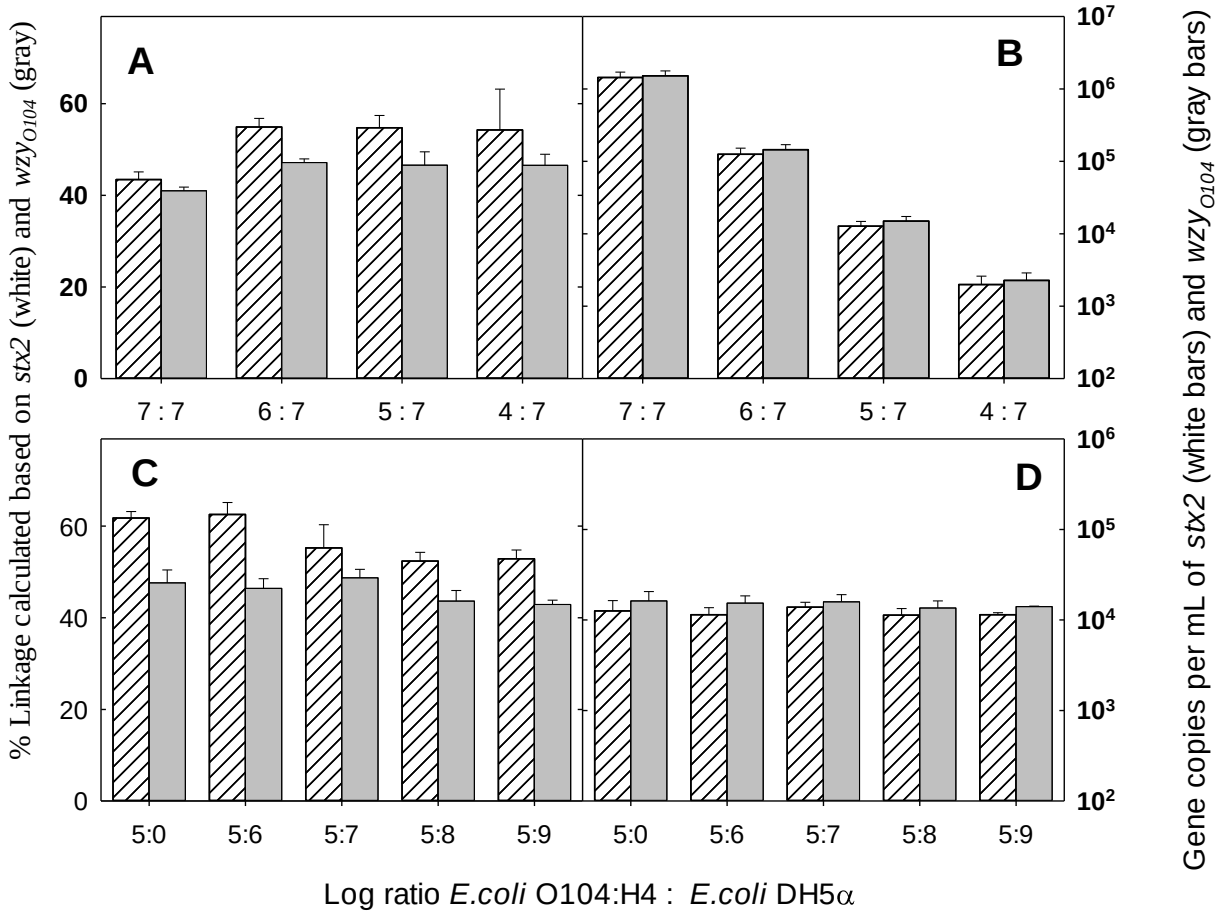


Figure 3: Percent linkage (**Panels A and C)** and copy number of *stx2* and wzy_{O104} **(Panel B and D)** determined with *E. coli* O104:H4 and with *E. coli* DH5α as background at different proportions. Experiments were conducted to maintain a constant cell count of *E. coli* DH5α **(Panels A and B)** or a constant cell count of *E. coli* O104 (**Panels C and D).** Data are shown as means ± standard deviation of three independent replicates.

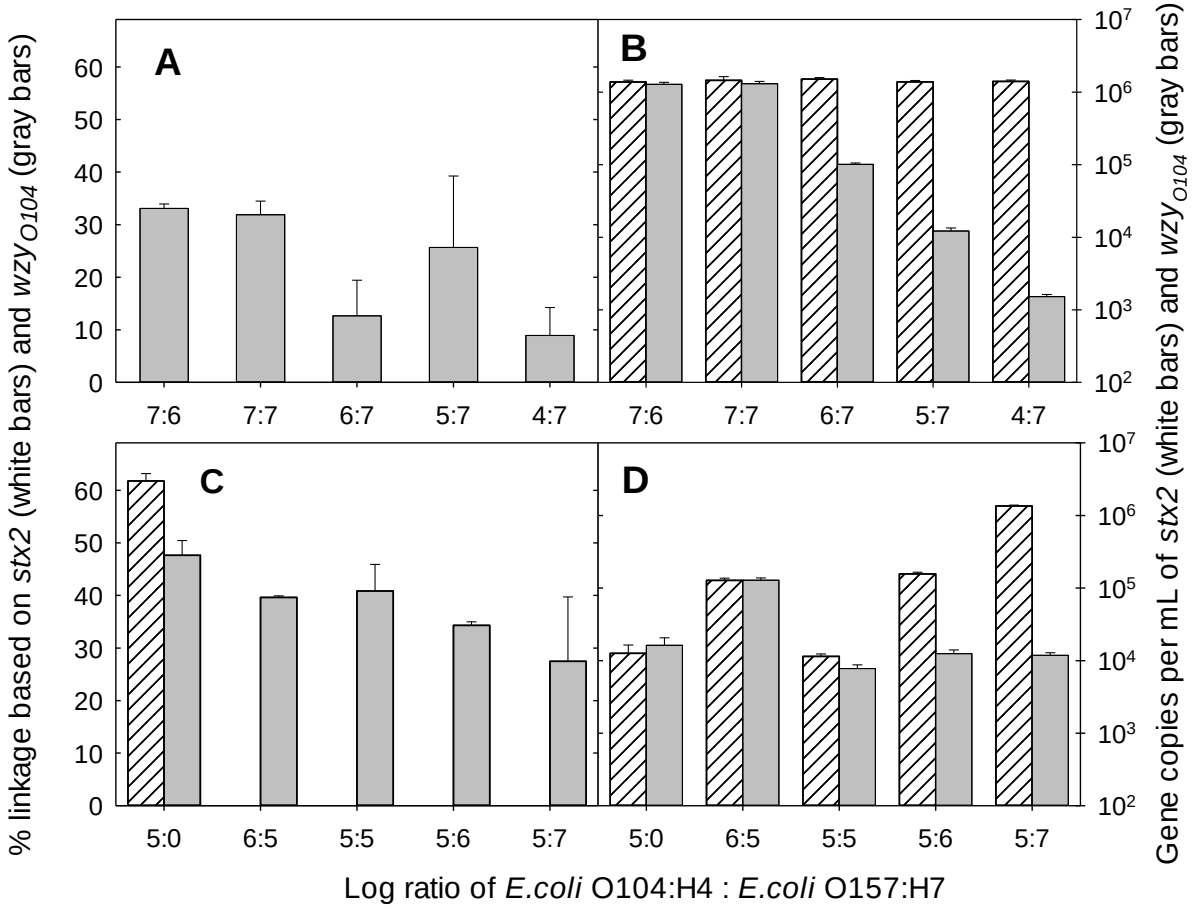
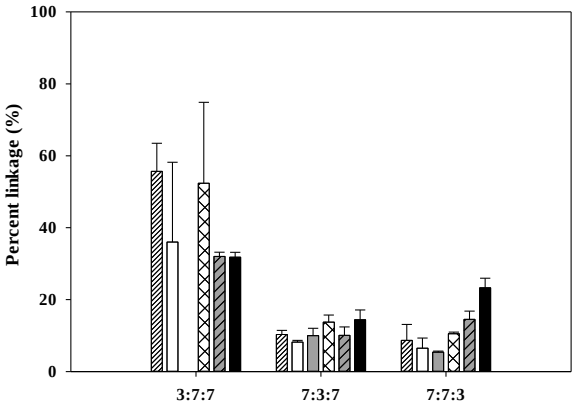
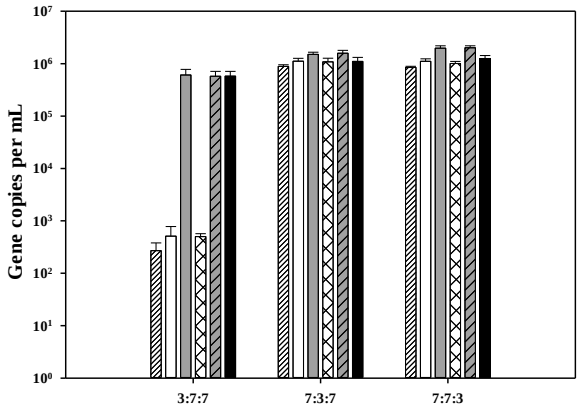


Figure 4. Percent linkage (**Panels A and C)** and copy number of *stx2* and wzy_{O104} **(Panel B and D)** determined with *E. coli* O104:H4 and with *E. coli* O157:H7 as background at different proportions. Data are shown as means ± standard deviation of three independent replicates.



(a) **Log ratio of *E. coli* O157:H7 to *E. coli* O104:H4 and *E. coli* DH5α**



(b) **Log ratio of *E. coli* O157:H7 to *E. coli* O104:H4 and *E. coli* DH5α**

Figure 5: Percent linkage **(Panel A)** and gene copy number (**Panel B)** at different proportions of strains *E. coli* O157:H7, *E. coli* O104:H4 and *E. coli* DH5α. The % linkage shown in **Panel A** was calculated as follows: ▨ (*stx1*+wzy_{O157})/*stx1*, ▭ (*stx1*+wzy_{O157})/wzy_{O157}, ▬ (*stx2*+wzy_{O157})/*stx2*, ⊠ (*stx2*+wzy_{O157})/wzy_{O157}, ▨ (*stx2*+*wzyE*)/*stx2*, ■ (*stx2*+*wzyE*)/*wzyE*. Bars in **Panel B** show the copy number of the following target genes using the primer probe pair in brackets: ▨ - *stx1*(*stx1*+wzy_{O157}), ▭ - wzy_{O157}(*stx1*+wzy_{O157}), ▬ - *stx2*(*stx2*+wzy_{O157}), ⊠ - wzy_{O157}(*stx2*+wzy_{O157}), ▨ - *stx2*(*stx2*+*wzyE*), ■ - *wzyE*(*stx2*+*wzyE*). Data are shown as means ± standard deviation of three independent replicates.

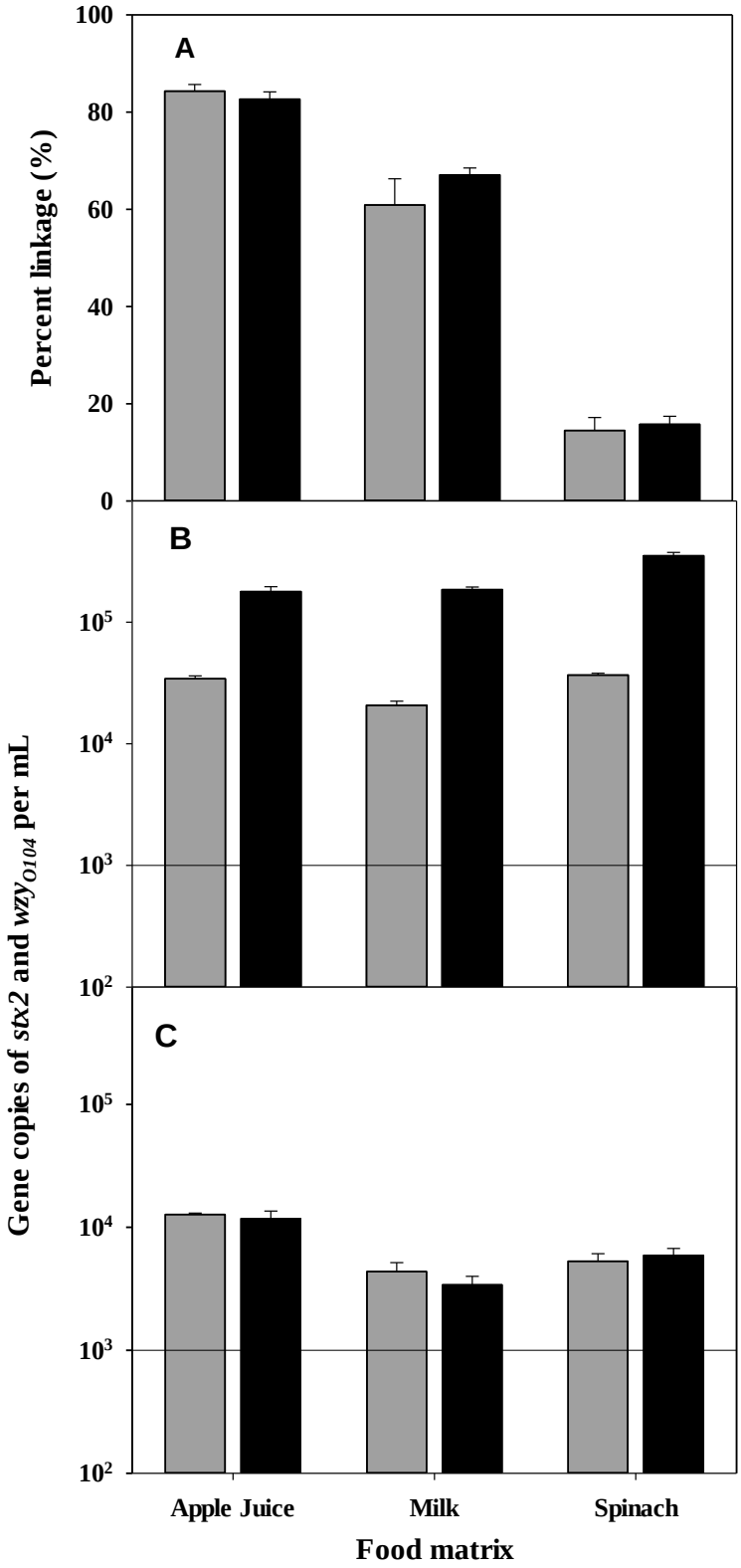

A
B
C
Percent linkage (%)
Gene copies of stx2 and wzy$_{O104}$ per mL
100
80
60
40
20
0
10^5
10^4
10^3
10^2
Apple Juice
Milk
Spinach
Food matrix

Figure 6: Detection of *E. coli* O104:H4 in food with *E. coli* O157:H7 as background at different proportions. **Panel A**, % linkage between *stx2* and *wzy*$_{O104}$; **Panel B**: copy number of *stx2;* **Panel C,** copy number of *wzy* $_{O104}$. ▭ Log ratio of *E. coli* O104:H4 and *E. coli* O157:H7 5:5, ▬ Log ratio of *E. coli* O104:H4 and *E. coli* O157:H7 5:6. Data are shown as means ± standard deviation of three independent replicates. The horizontal line indicates the detection limit of 10^3 gene copies/mL. Gene copy numbers in uninoculated apple juice, milk, and spinach were all below the detection limit.