

Table of Contents

Declaration	i
Abstract.....	ii
Dedication	v
Acknowledgements	vi
Table of Contents	viii
List of Tables	x
List of Figures.....	xi
1. Introduction and problem statement	1
2. Persistence of foodborne Diarrheogenic <i>Escherichia coli</i> in the agricultural and food production environment: implications for food safety and public health.....	5
2.1 <i>Escherichia coli</i>.....	5
2.2 Categorization of pathogenic <i>E. coli</i>.....	6
2.3 Isolation and identification of <i>E. coli</i>.....	7
2.3.1 Culture dependent methods	7
2.3.2 Culture independent methods.....	8
2.4 Diarrheogenic <i>Escherichia coli</i> a foodborne pathogen	12
2.5 The agricultural and food production environment as a hotbed for emerging Diarrheogenic <i>Escherichia coli</i>	12
2.6 Linking Diarrheogenic <i>Escherichia coli</i> persistence in the agricultural and food production environment to inherent bacterial attributes.....	13
2.6.1 Diarrheogenic <i>Escherichia coli</i> attachment to biotic and abiotic surfaces.....	14
2.6.2 Lateral gene transfer as a tool utilized by DEC for adaptation to the agricultural and food production environment	21
2.6.3 <i>E. coli</i> response to the changing environmental conditions with emphasis on Diarrheogenic <i>Escherichia coli</i>	25
2.6.3.1 <i>Escherichia coli</i> response to environmental stresses with emphasis on Diarrheogenic <i>Escherichia coli</i> in the open environment outside the human host.....	25
2.6.4 <i>Escherichia coli</i> bacterial secretion systems and their role in environmental persistence.....	26

viii

2.7 Conclusion and future perspectives.....	27
3. Hypothesis and Objectives	30
3.1 Hypotheses.....	30
3.2 Objectives.....	32
4. Research Chapter.....	33
4.1 Enteroaggregative <i>Escherichia coli</i> is the predominant diarrheagenic <i>E. coli</i> pathotype in irrigation water and food sources in South Africa.....	34
4.1.1 Introduction.....	36
4.1.2 Materials and Methods.....	37
4.1.3. Results	42
4.1.4. Discussion.....	52
4.2 <i>Escherichia coli</i> isolated from food sources and irrigation water: A potential risk for causing intestinal dysfunction?	55
4.2.1 Introduction.....	57
4.2.2. Materials and methods	58
4.2.3. Results	65
4.2.4. Discussion.....	76
4.2.5. Conclusion	78
4.3 Detection of Extended Beta Lactamase Cefotaxime Resistance and Virulence Genes in <i>Escherichia coli</i> by Duplex Quantitative Real Time PCR and Melt Curve Analysis.....	79
4.3.1 Introduction.....	80
4.3.2. Materials and methods	82
4.3.3. Results	90
4.3.4 Discussion.....	99
4.3.5. Conclusion	100
5. General Discussion.....	101
5.1 Critique of methodology.....	101
5.2 Discussion.....	103
6. Conclusions and Recommendations	106
6.1 Conclusions.....	106
6.2 Recommendations	106
7. References.....	108

List of Tables

Table 4.1.1 Primer sequences and the expected amplicon sizes for the multiplex polymerase chain reaction employed in the detection of Diarrheagenic <i>Escherichia coli</i>	40
Table 4.1.2 Percentage of <i>Escherichia coli</i> strains isolated from food sources and irrigation water in South Africa positive for virulence gene determinants (n=205) and cell adherence patterns (n=46) associated with diarrheagenic <i>E. coli</i>	45
Table 4.1.3 Nucleotide sequences of virulence gene determinants associated with enteroaggregative <i>Escherichia coli</i> (EAEC) in <i>Escherichia coli</i> strains isolated from food sources and irrigation water in South Africa.	46
Table 4.2.1 Strains, sources of isolation, pathotypes and assays used for characterizing <i>Escherichia coli</i> isolated from food sources and irrigation water in South Africa	64
Table 4.3.1: Source and characteristics of <i>E. coli</i> isolates from food sources and irrigation water in South Africa used in developing and validating a set of two duplex qPCR assays for detection of antibiotic resistance and virulence genes	83
Table 4.3.2: Genes and primer sequences used for development of a pair of duplex qPCR assays for detection of antibiotic resistance and virulence in <i>E. coli</i> isolated from food sources and irrigation water in South Africa	85
Table 4.3.3: Optimized reagents and consumables used in a pair of duplex qPCR assays for detection antibiotic and of virulence genes in <i>E. coli</i> isolated from food sources and irrigation water in South Africa.	88
Table 4.3.4: Thermocycling conditions for a pair of duplex qPCR assays based on different commercially available qPCR super mixes for detection of virulence (shigatoxin 1, shigatoxin 2, intimin) and Cefotaxime resistant (CTX-M) extended beta lactamase (ESBL) genes in <i>E. coli</i> isolated from food sources and irrigation water in South Africa	89
Table 4.3.5: Template volumes and melting temperatures of amplicons in monoplex and duplex qPCR assays for detection of antibiotic resistant and virulence genes in <i>E. coli</i> isolated from food sources and irrigation in South Africa using melt curve analysis	92

List of Figures

Figure 2.1: Different pathotypes of <i>E. coli</i> and their characterizing genetic determinants: <i>stx</i> , verocytotoxin/shiga toxin; <i>eae</i> , intimin; <i>bfp</i> , bundle-forming pili; <i>ipa</i> , invasion plasmid antigens; <i>tia</i> , toxigenic invasion loci A; <i>astA</i> , plasmid with enteroaggregative heat-stable toxin; <i>shET</i> , <i>Shigella</i> enterotoxin; CFA, colonization factor antigens; <i>ST</i> , heat-stable enterotoxin; <i>LT</i> , heat-labile enterotoxin. Source: (Ceuppens et al., 2014).....	7
Figure 2.2: Interplay of interdependent bacterial processes that may facilitate the persistence of Diarrheagenic <i>Escherichia coli</i> (DEC) in the open environment. The processes can be bi-directional depending on the prevailing environmental conditions influenced by anthropogenic, climatic and ecological factors	16
Figure 2.3: The phylogeny of <i>E. coli</i> influences the acquisition of virulence genes and subsequently the diarrheagenic pathotype (DEC) which may inadvertently lead to emergence of new pathotypes in varying environmental niches such as the agricultural and food production setting. “The phylogenetic neighbourhood of geographically remote <i>E. coli</i> supports the notion of a rapid worldwide spread of an evolutionary common ancestor (maybe with the advent of mammals and selection in specific habitats.” (Chaudhuri and Henderson, 2012). Phylogroups A (mostly commensals) and B1 (non-O157 enterohaemorrhagic (EHEC) strains) form the newest clades. Phylogroups D1 (uropathogenic (UPEC) and enteroaggregative (EAEC) strains) and E (separate clade of O157:H7 EHEC and 055:H7 enteropathogenic (EPEC) strains). Phylogroups D2 and B2 (EPEC strain E2348/69 and extraintestinal (ExPEC) strains) form the oldest clades.....	17
Fig 4.1.1. Evolutionary relationships of <i>Escherichia coli</i> strains in GenBank showing $\geq 80\%$ gene nucleotide sequence similarity to strain K2 isolated from producer distributor bulk milk in South African milk based on <i>aaiC</i> , the enteroaggregative <i>E. coli</i> (EAEC) virulence gene determinant. The evolutionary history was inferred using the Neighbour-Joining Method. Codes of strains used for comparison represent accession numbers from GenBank. G1, G2, G3 and G4 represent defined clusters of strains showing differences in genetic location of <i>aaiC</i> (plasmid or chromosome) in each strain as well as source of strain isolation (foodborne or human faeces).....	47
Fig. 4.1.2. Evolutionary relationships of <i>Escherichia coli</i> in GenBank showing $\geq 90\%$ gene nucleotide sequence similarity to <i>E. coli</i> strains isolated from producer distributor bulk milk (PDBM) and irrigation water in South Africa based on <i>aatA</i> , the enteroaggregative <i>E. coli</i> (EAEC) virulence gene determinant. The evolutionary history was inferred using the Neighbour-Joining Method. Strain sources: PDBM-L7, 57; Irrigation water-MPUW51, MPUW84. Codes of strains used for comparison represent accession numbers in GenBank.....	48
Fig. 4.1.3a. Aggregative adherence (AA) pattern ‘stacked-brick’ observed in <i>Escherichia coli</i> strains isolated from food sources and irrigation water in South Africa. Strains were grown on HeLa cells and	

showed strong to moderate AA characteristic of enteroaggregative *Escherichia coli* (EAEC). Strain code: Standard EAEC strain 042; B-H8; C-M28; D-NW(V)7(3); E-NW(V)10(1); F-NW(W)9(3). Source of isolation: A-Clinical strain; B and C-Producer distributor bulk milk; D and E-Irrigated lettuce; F-Irrigation water. Images were taken to 100X with a Zeiss microscope. Resolution=20 µm.49

Fig. 4.1.3b. Weak aggregative adherence (AA) observed in *Escherichia coli* strains isolated from food sources and irrigation water in South Africa. Strains were grown on HeLa cells and showed weak adherence in comparison to strains in Fig. 4a. Strain codes: G- LeK1; H-M12; I-K5; J-K16; K-M6; L-N23. Strain sources: G-Irrigated lettuce; H, I, J, K and L-Producer distributor bulk milk. Images were taken to 100X with a Zeiss microscope. Resolution= 20 µm..... 50

Fig.4.1.3c. Invasiveness observed in *Escherichia coli* strains isolated from food sources and irrigation water in South Africa. Invasiveness is characteristic of enteroinvasive *Escherichia coli* (EIEC) and therefore strains were identified as presumptive EIEC since no other subsequent confirmatory test was done. Strains were grown on HeLa cells. Strain codes: i and ii-LeK; iii and iv- 37. Strain sources: LeK-Irrigated lettuce; 37-Producer distributor bulk milk. Images were taken at 60X with a Zeiss microscope. Resolution=20 µm. 51

Figure. 4.2.1. Biofilm formation in Luria Bertani broth by diarrheagenic *E. coli* (DEC) and non-DEC isolated from food sources and irrigation water. DEC strains: 1 to 6 (PDBM); 31 to 33 (irrigated lettuce); 36 to 38 (DEC standard strains). Non-DEC: 7 to 20 (PDBM); 21 to 30 (irrigation water); 34 to 35 (irrigated lettuce). a, b: Biofilm formation significantly different ($p \leq 0.05$). DEC standard strains: 36-enteroaggregative *E. coli* strain 042; 37-enteropathogenic *E. coli* strain 2348; 38-diffusely adherent *E. coli* strain F1875. Strain codes: 1-M6, 2-N23, 3-K16, 4-K5, 5-M28, 6-M12, 7-M37, 8-N5, 9-M24, 10-M11, 11-58or85, 12-M4, 13-N26, 14-L5, 15-K3, 16-F8, 17-54, 18-N6, 19-M13, 20-513, 21-MPU(W)8(2), 22-MPU(W)1(1), 23-MPU(W)8(3)again, 24-MPU(W)9(3), 25-NW(W)9(1), 26-CR12, 27-MPU(W)6(2), 28-NW(W)6(1), 29-NW(W)3, 30-CR4, 31-LeK, 32-NW(V)10(1)again, 33-NW(V)7(3), 34-NW(V)10(1), 35-Vp. The horizontal line shows biofilm production for the low biofilm producing strain N26..... 66

Figure 4.2.2. Biofilm formation in Dulbecco's Modified Eagle Medium-F12+0.5% Mannose by diarrheagenic *E. coli* (DEC) and non-DEC previously isolated from food sources and irrigation water in South Africa. DEC strains: 1 to 6 (PDBM); 32 to 34 (irrigated lettuce); 37 to 38 (DEC standard strains). Non-DEC: 7 to 20 (PDBM); 21 to 31 (irrigation water); 35 to 36 (irrigated lettuce). a, b: Biofilm formation significantly different at $p \leq 0.05$. DEC standard strains: 37-enteroaggregative strain 042; 38-diffusely adherent *E. coli* strain F1875. Strain codes: 1-M6, 2-N23, 3-K16, 4-K5, 5-M28, 6-M12, 7-M37, 8-N5, 9-M24, 10-M11, 11-58or85, 12-M4, 13-N26, 14-L5, 15-K3, 16-F8, 17-54, 18-N6, 19-M13, 20-513, 21-MPU(W)8(2), 22-MPU(W)1(1), 23-MPU(W)8(3) again, 24-MPU(W)9(3), 25-NW(W)9(1), 26-CR12, 27-MPU(W)6(2), 28-NW(W)6(1), 29-NW(W)3, 30-CR4, 31-NW(W)9(3),

32-LeK, 33-NW(V)10(1)again, 34-NW(V)7(3), 35-NW(V)10(1), 36-Vp. PDBM-Producer Distributor Bulk Milk. The horizontal line shows biofilm production for the low biofilm producing strain N26.. 67

Figure 4.2.3. Extracellular polymeric substance (EPS) formation (amorphous extracellular mass attached to bacterial wall) by diarrheagenic (DEC) and non-DEC previously isolated from food sources and irrigation water in South Africa visualized under Transmission Electron Microscopy. Arrows point to EPS. Strain pathotype and source of isolation: 1- DEC isolated from PDBM; 2- Non-DEC isolated from PDBM; 3-non-DEC isolated from irrigation water; 4 and 5-non-DEC isolated from irrigated lettuce. Magnification X15000. Scale 0.5 to 1.0µm. Strain codes:1-K2, 2-N25, 3-NW(W)9(1), 4-Vp, 5-NW(V)3(1). PDBM-Producer Distributor Bulk Milk. 70

Figure 4.2.4 Cytotoxic activity (disruption of the HeLa cell cytoskeleton) in diarrheagenic *E. coli* (DEC) and Non-DEC previously isolated from food sources and irrigation water in South Africa. Images recorded at 60X. Scale=20µm. Strain pathotype and source of isolation: 1-Non-DEC isolated from coleslaw; 2-DEC isolated from PDBM; 3, 4 and 5-Non-DEC isolated from PDBM; 6-Non-DEC isolated from irrigated lettuce. Strains 7 and 8 (no cytotoxic activity), 7-Non-DEC isolated from PDBM; 8-DEC isolated from PDBM. Test strain codes: 1-Nana 10, 2-K16, 3-M24, 4-N25, 5-K3, 6-NW(V)10(1), 7-N5, 8-N23. PDBM-Producer Distributor Bulk Milk. Only strains exhibiting cytotoxicity are reported. 71

Figure 4.2.5. Percentage reduction in initial Transepithelial Electrical Resistance (TER) observed in polarized T-84 epithelial cells infected with diarrheagenic (DEC) and non-DEC previously isolated from food sources and irrigation water in South Africa. Strain pathotype and source of isolation: DEC strains: 1 to 4 (PDBM); 12 to 13 (irrigated lettuce); 19 (DEC standard strain). Non-DEC: 5 to 9 (PDBM); 10 to 11 (Irrigation water); 14 to 17 (Irrigated lettuce); 18 (Coleslaw); 20 (Standard commensal *E. coli* strain aafa). a, b: TER significantly different at $p \leq 0.05$. 19-DEC standard strain enteroaggregative *E. coli* strain 042; 20- Commensal standard *E. coli* strain aafa. Strain codes: 1-K2, 2-N23, 3-K5, 4-K16, 5-K3, 6-54, 7-513, 8-M37, 9-L5, 10-NW(W)9(1), 11-MPU(W)4, 12-LeK, 13-NW(V)10(1)again, 14-Vq, 15-NW(V)6(2), 16-V5, 17-V9, 18-Nana 10. PDBM-Producer Distributor Bulk Milk. Only strains exhibiting reduction in TER are reported..... 72

Figure 4.2.6. Interleukin 8 (IL-8) induction in polarized T-84 epithelial cells resulting from infection with diarrheagenic *E. coli* (DEC) and non-DEC previously isolated from food sources and irrigation water in South Africa. Strain pathotype and source of isolation: DEC strains: 1 to 4 (PDBM); 12 to 13 (irrigated lettuce); 19 (DEC standard strain). Non-DEC: 5 to 9 (PDBM); 10 to 11 (irrigation water); 14 to 17 (irrigated lettuce); 18 (Coleslaw); 20 (standard commensal *E. coli* strain). a, b, c, d induction of IL-8 significantly different ($p \leq 0.05$). 19-DEC standard control strain EAEC strain 042; 20- Standard commensal *E. coli* strain aafa. Strain codes: 1-K2, 2-N23, 3-K5, 4-K16, 5-K3, 6-54, 7-513, 8-M37, 9-L5, 10-NW(W)9(1), 11-MPU(W)4, 12-Lek, 13-NW(V)10(1) again, 14-Vq, 15-NW(V)6(2),

16-V5, 17-V9, 18-Nana 10. PDBM-Producer Distributor Bulk Milk. Only strains exhibiting induction of IL-8 are reported. 73

Figure 4.2.7. Principal Component Analysis variables correlation plot showing axes Factors 1 and 2 (F1 and F2) accounting for approximately 60% of variation observed after analysis of 4 factors in diarrheagenic *E. coli* (DEC) and non-DEC tested in-vitro potentially associated with human gut intestinal dysfunction. Analyses included biofilm formation in Luria Bertani and Dulbecco's Modified Eagle Medium-F12+0.5% Mannose media, Cytotoxic activity on HeLa cells, disruption of epithelial cell integrity (TER reduction) and induction of Interleukin-8 on polarized T-84 cells. The strains were previously isolated from food sources and irrigation water. The strains were compared to DEC standard strain enteroaggregative *E. coli* strain 042. TER-Trans epithelial electrical resistance 74

Figure 4.2.8. Principal Component Analysis biplot showing axes Factors 1 and 2 (F1 and F2) accounting for approximately 60% of variation observed in 11 strains of diarrheagenic (DEC) and non-DEC isolated from food sources and irrigation water. The strains were analysed in-vitro for 4 factors potentially associated with human gut intestinal dysfunction. Analyses included; Biofilm formation in Luria Bertani and Dulbecco's Modified Eagle -F12+0.5% Mannose media, cytotoxic activity, disruption of epithelial cell integrity (TER reduction) and induction of Interleukin-8 on polarized T-84 cells. The strains were previously isolated from food sources and irrigation water. DEC standard strain EAEC strain 042 was used for comparison. Strains and source of isolation: 1 to 3 (DEC isolated from PDBM); 4 to 7 (non-DEC isolated from PDBM); 8 to 9 (non-DEC isolated from irrigation water); 10 (DEC isolated from irrigated lettuce); 11 (non-DEC isolated from irrigated lettuce); 12 (EAEC standard strain 042). Strain codes: 1-N23, 2-K16, 3-K5, 4-M37, 5-L5, 6-K3, 7-513, 8-NW(W)9(1), 9-NW(W)6(1), 10-LeK, 11-NW(V)10(1). 75

Figure 4.3.1: Melting peak curves showing antibiotic resistant (*bla_{CTX-M}*) and virulence (*eae*, *stx1*, *stx2*) genes in *E. coli* isolated from food sources and irrigation water determined using monoplex qPCR assays with a commercial qPCR mix containing SYBR Green I dye. The horizontal line depicts the threshold for detection. *stx1*-shigatoxin 1; *stx2*-shigatoxin 2; *eae*- intimin; *bla_{CTX-M}*- general class of Cefotaxime resistant extended beta-lactamase (ESBL) producing *E. coli*. QPCR mix was SsoAdvanced™ Universal Inhibitor-Tolerant SYBR® Green Supermix. The peaks were determined by plotting the negative derivative of change in fluorescence increase over time (-d(RFU)\dt) with increase in temperature. 93

Figure 4.3.2: Melting peak curves showing antibiotic resistant (*bla_{CTX-M}*) and virulence (*eae*, *stx1*, *stx2*) genes in *E. coli* isolated from food sources and irrigation water in South Africa determined using a pair of duplex qPCR assays from a commercial qPCR mix containing SYBR Green I dye. The horizontal line depicts the threshold for detection. *stx1*-shigatoxin 1; *stx2*-shigatoxin 2; *eae*- intimin; *bla_{CTX-M}*- general class of Cefotaxime resistant extended beta-lactamase (ESBL) producing *E. coli*. QPCR mix was SsoAdvanced™ Universal Inhibitor-Tolerant SYBR® Green Supermix. The peaks

were determined by plotting the negative derivative of change in fluorescence over time $(-d(RFU)/dt)$ with increase in temperature. 94

Figure 4.3.3: Melting peak curves showing targeted antibiotic resistant (*bla_{CTX-M}*) and virulence (*eae*, *stx1*, *stx2*) genes in *E. coli* isolated from food sources and irrigation water determined with monoplex qPCR assays using a commercial qPCR mix containing EvaGreen dye. The horizontal line depicts the threshold for detection. *stx1*-shigatoxin 1; *stx2*-shigatoxin 2; *eae*- intimin; *bla_{CTX-M}*- general class of Cefotaxime resistant extended beta-lactamase (ESBL) producing *E. coli*. The QPCR mix was KAPA HRM Fast PCR mix. The peaks were determined by plotting the negative derivative of change in fluorescence over time $(-d(RFU)/dt)$ with increase in temperature. 95

Figure 4.3.4: Melting peak curves showing antibiotic resistant (*bla_{CTX-M}*) and virulence (*eae*, *stx1*, *stx2*) genes in *E. coli* isolated from food sources and irrigation water in South Africa determined using a pair of duplex qPCR assays with a commercial qPCR mix containing EvaGreen dye. The horizontal line depicts the threshold for detection. *stx1*-shigatoxin 1; *stx2*-shigatoxin 2; *eae*- intimin; *bla_{CTX-M}*- general class of Cefotaxime resistant extended beta-lactamase (ESBL) producing *E. coli*. The QPCR mix was KAPA HRM Fast PCR mix. The peaks were determined by plotting the negative derivative of change in fluorescence over time $(-d(RFU)/dt)$ with increase in temperature..... 96

Figure 4.3.5: QPCR duplex reactions showing amplification (fluorescence increase v/s quantification cycle (C_q)) of DNA dilution series (10^0 to 10^{-8}) using two commercially available qPCR mixes based on EvaGreen and SYBR Green I dyes. EvaGreen based qPCR mix- KAPA HRM Fast PCR mix. SYBR Green I dye based qPCR mix- SsoAdvanced™ Universal Inhibitor-Tolerant SYBR® Green Supermix. Data from these graphs was used to draw determine the limit of detection for each assay by drawing standard curves. 97

Figure 4.3.6: Standard curve derived from plotting the change in quantification cycle (C_q) with the quantity of serially diluted (10^0 to 10^8) template DNA for a pair of 2 duplex assays each based on commercial qPCR mixes containing SYBR Green I or EvaGreen dyes used in detection of antibiotic resistance and virulence genes in *E. coli* isolated from food sources and irrigation water in South Africa. 98

1. Introduction and problem statement

Foodborne illnesses present a great risk to global public health and well-being by causing morbidity and mortality, lost revenue to economies and rising health costs to the affected people (Anelich, 2014; Kirk et al., 2015; Newell et al., 2010). Outbreaks associated with food have the capacity to hamper lucrative international and regional trade as well as run affected companies into the ground (Commission of The European Communities, 2011). Diarrheagenic *Escherichia coli* (DEC) forms one of the common pathogens associated with foodborne illnesses causing diarrhea especially among vulnerable groups such as infants and immune compromised individuals e.g. HIV patients or the elderly as well as western travelers to developing countries causing traveler's diarrhea. The common DEC pathotypes associated with foodborne illnesses include enteropathogenic *E. coli* (EPEC), enterotoxigenic *E. coli* (ETEC), shigatoxin producing *E. coli* (STEC), enteroaggregative *E. coli* (EAEC) and enteroinvasive *E. coli* (EIEC) that have been reported to be transmitted through the oral fecal route (Croxen et al., 2013; Kaper et al., 2004). This scenario suggests a point within the food chain where sanitation was not observed thus comprising health and safety of the consumer. Historically these DEC pathotypes were associated with poor sanitation linked to contaminated water especially in developed countries but increasing infection rates are presently mostly observed in developing countries especially in Sub-Saharan Africa and South East Asia (Kotloff et al., 2013). This shift in prevalence rates within developing nations has been linked to poor hygiene levels with the situation being exacerbated by rising population levels without matched improvement in sewage and housing infrastructure (Troeger et al., 2017).

Similarly, *E. coli* strains (Non-DEC) not manifesting the normal genotypic and phenotypic characteristics associated with DEC may lead to negative health outcomes affecting the gastrointestinal tract (GIT) referred to as intestinal dysfunction (ID). This condition is associated with a change in the "normal" relative abundance of bacterial taxa in the gastrointestinal tract (GIT) and their associated metabolic functions within the human intestinal (Shin et al., 2015). The conditions ultimately lead to a breakdown in homeostasis and as well as the normal synergistic relationship between the human body and its microbial constituents (Shin et al., 2015). Additionally, ID is linked to reduced mucosal immunity and a subsequent increased risk of intestinal diseases causing reduced dietary nutrient intake and utilization in young infants.

This results into growth faltering and neurocognitive disorders later in life (Faith et al., 2015). For example, malnourished infants from Bangladesh and Malawi with Kwashiorkor showed a high rate of inflammation that was associated with a bloom of *Enterobacteriaceae* as well as a high DEC burden upon faecal analysis (Kau et al., 2015; Subramanian et al., 2014). These observations are particularly disturbing especially within developing countries where multiple factors such as inadequate dietary practices and poor environmental sanitation may increase the risk of contamination with *Enterobacteriaceae* and a subsequent bloom within the gut.

The rise in global trade has re-exposed developed countries to the risk of contamination that has been linked to outbreaks with common foodborne pathogens such as *E. coli* O157:H7 (Mangen et al., 2010) and O104: H4 have caused a serious continental outbreak in 2011 (Beutin and Martin, 2012). Of particular concern among such outbreak pathogens is their ability to resist treatment with antibiotics. Antibiotic resistance associated with emerging infectious pathogens poses a great risk to humanity in the 21st century (The Review on Antimicrobial Resistance, 2014). For example, the rise of extended beta lactamase enzyme producing bacteria (ESBL) within the last 20 years has increased this risk (Cantón et al., 2012; Shi et al., 2015). Similarly, within the last decade, high prevalence of Cefotaxime resistant (*CTX-M*) extended beta lactamase producing (ESBL) bacteria has increased the risk of spreading pathogens resistant to 4th generation cephalosporins that are usually the last line of treatment for acute bacterial infections (Shi et al., 2015). The world-wide dissemination of the highly virulent and antibiotic resistant *E. coli* ST131 strain exemplifies this emerging threat posed to food safety and public health.

These examples demonstrate how the risk of illnesses is not confined to a single geographical area but across the globe thereby spreading the risk beyond national and geographical borders. Therefore, understanding the sources of this contamination as well as the predominant pathotypes within the different foods and environmental sources is crucial to containing the spread of outbreaks. Additionally, some DEC pathotypes such as EAEC may not cause full blown outbreaks or illness but rather lead to recurrent infections and asymptomatic diarrhea (Steiner et al., 2006). Such a scenario leads to nutrient malabsorption in children leading to stunting and neurocognitive impairment (Acosta et al., 2016). Such effects present long-term health effects that are not reversible and can prevent a whole generation of children from achieving a full productive and healthy life resulting in which in a loss to affected nations.

In developing countries such as South Africa, continuous national monitoring of foodborne DEC is absent. Although, clinical diarrheal cases associated with DEC have been reported by the National Institute of Communicable Diseases (NICD) since 2012, these metrics fail to provide a link with potential sources of contamination associated with the high diarrheal disease burden noticed within the country especially among infants and immune compromised adults. Studies have shown that food (Caine et al., 2014; Njage and Buys, 2017; Ntuli et al., 2016) and environmental sources including irrigation water (Aijuka et al., 2015; Gemmell and Schmidt, 2012; Ijabadeniyi et al., 2011; Olaniran et al., 2009) harbor *E. coli* containing genes associated with DEC suggesting that contaminated food sources may be a major route of infection.

Therefore, based on this backdrop, this study sought to characterize previously isolated *E. coli* strains from various food sources (producer distributor bulk milk, irrigated lettuce, street vendor coleslaw) and irrigation water collected over 10 years from different provinces. This information would help identify sources susceptible to DEC contamination as well as potential points of cross-contamination. Additionally, such information would be vital for determining mitigation measures.

This study additionally sought to develop a rapid and accurate molecular based assay for the cost-effective detection of antibiotic resistant DEC isolated from food sources and irrigation water in South Africa. Such an assay would be vital in monitoring the prevalence and emergence of pathogens threatening food safety and public health within the country

**Chapter 2 is written in form of a mini-review article prepared for submission to the
Journal: *Food Research International***

2. Persistence of foodborne Diarrheagenic *Escherichia coli* in the agricultural and food production environment: implications for food safety and public health

Matthew Aijuka and Elna M. Buys*

Department of Consumer and Food Sciences, University of Pretoria, Corner Lynwood and Roper Roads, Hatfield 0028, Pretoria, Republic of South Africa.

* Corresponding author: elna.buys@up.ac.za

Abstract:

Diarrheagenic *Escherichia coli* (DEC) is a leading cause of foodborne illness associated with intestinal disease. While known over the years that contamination of food sources occurs via the oral faecal-route, the mechanisms underlying its persistence within the open environments including the food chain remains virtually known. Therefore, in this mini-review we will shed light on bacterial processes such as initial attachment, biofilm formation, horizontal gene transfer and response to environmental stresses. These factors may enable persistence of DEC as well as the emergence of potentially more virulent strains within the agricultural and food production environment. Mechanistic studies in clinical microbiology and immunology have elucidated infection pathways in the human and other animal bodies leading to diagnostic and treatment solutions. Therefore, understanding DEC behaviour in the agricultural and food production environment is crucial for ensuring food safety and public health by reducing the burden of foodborne illnesses.

Key words: Diarrheagenic *Escherichia coli*, Lateral gene transfer, environmental response, biofilm formation, foodborne illness, environmental response

2.1 *Escherichia coli*

Escherichia coli is a facultative Gram-negative facultative anaerobe of the phylum Proteobacteria and Family *Enterobacteriaceae* which forms a large part of the normal microflora of the human gut (Shin et al., 2015). The high numbers in the gut suggest a commensal or symbiotic relationship with other microbial constituents of the normal human microbiota. While many strains from this species present no viable health risk, certain strains do have the ability to acquire virulence factors enabling them to attack specific parts of the

human body and other animals causing diseases (Leimbach et al., 2013). Alternatively, *E. coli* has also been referred to as a laboratory ‘work-horse’, a term used to refer to its ease of culture at different growth conditions facilitating its convenient easy use for elucidating bacterial processes in laboratories around the world (Blount, 2015). For example, over the last century the commensal strain *E. coli* K-12 has been used to unravel many fundamental processes underlying biological phenomena, many of which have gone on to be utilized in various industries from biotechnology to health and other applications (Blount, 2015).

2.2 Categorization of pathogenic *E. coli*

Pathogenic strains of *E. coli* include those that cause disease in the human gastrointestinal tract (DEC) and extra-intestinal *E. coli* (ExPEC) which cause infections in the urinary tract (uropathogenic *E. coli*, UPEC), septicemia associated *E. coli* (SEPEC) and meningitis-associated *E. coli* (MNEC). Infections caused in other parts of the body include those causing blood infections (bacteremia). Pathogenic *E. coli* has also been shown to cause diarrheal disease in domestic farm animals such piglets and poultry (Avian Pathogenic *E. coli*, APEC) depicting its wide distribution amongst humans and animals and thereby potential circulation within the human agricultural and food production environment. For example, the global dissemination of the ExPEC strain *E. coli* sequence type (ST131) has been associated with chronic infections and antibiotic resistance affecting many people (Nicolas-Chanoine et al., 2014).

DEC have acquired virulence genes through horizontal genetic transfer (HGT) and to a lesser extent random mutation enabling the organisms to attach to the human gut surfaces leading to diseases among which diarrhea is a symptom (Croxen et al., 2013; Kaper et al., 2004). The virulence factors enable attachment to the gut mucosal lining leading to persistent infections and inflammation. DEC are categorized into pathotypes with each description based on which virulence factors it possesses as well as the mechanism of pathogenesis (Figure 2.1).

Common DEC pathotypes commonly associated with human disease include; enteropathogenic *E. coli* (EPEC), enterotoxigenic *E. coli* (ETEC), shigatoxin producing *E. coli* (STEC), enteroaggregative *E. coli* (EAEC), diffusely adherent (DAEC), enteroinvasive *E. coli* (EIEC) and adherently invasive *E. coli* (AIEC) (Croxen et al., 2013). The common virulence factors associated with many of these pathotypes include attachment proteins such as intimin (*eae*) and bundle forming pili (*bfp*) for EPEC, cytotoxin production as noticed with STEC

(shigatoxins 1, *stx1* and 2, *stx2* as well as the locus of enterocyte effacement, LEE) as well as heat labile (*LT*) and heat stable (*ST*) toxins for ETEC and epithelial cell adherence patterns for EAEC (stacked-brick pattern) and DAEC (random attachment along the whole cell surface) (Croxen et al., 2013). It has been suggested that driven environmental and genetic factors have led to the rise of multiple clones of these pathogens leading to the great heterogeneity observed among groups such as EAEC, ETEC and AIEC (Leimbach et al., 2013).

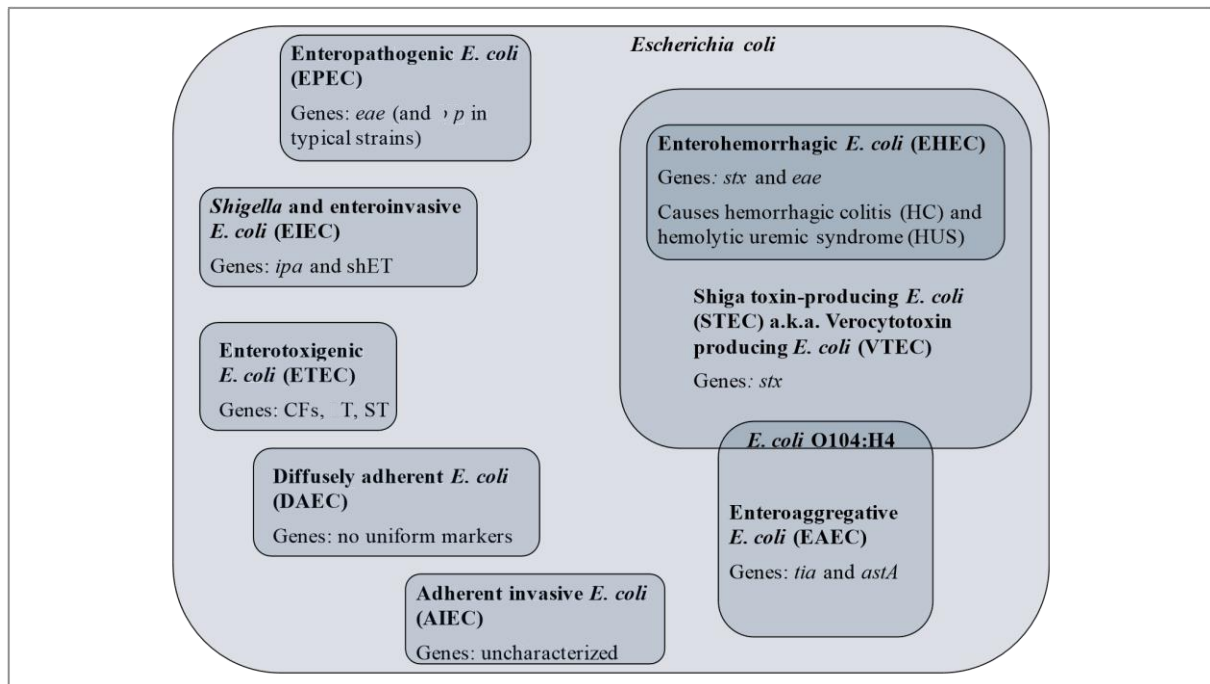


Figure 2.1: Different pathotypes of *E. coli* and their characterizing genetic determinants: *stx*, verocytotoxin/shiga toxin; *eae*, intimin; *bfp*, bundle-forming pili; *ipa*, invasion plasmid antigens; *tia*, toxigenic invasion loci A; *astA*, plasmid with enteroaggregative heat-stable toxin; *shET*, *Shigella* enterotoxin; *CFA*, colonization factor antigens; *ST*, heat-stable enterotoxin; *LT*, heat-labile enterotoxin. Source: (Ceuppens et al., 2014). STEC additionally carry the locus of enterocyte effacement (LEE).

2.3 Isolation and identification of *E. coli*

2.3.1 Culture dependent methods

Historically, before the advent of non-culture methods based on identification of nucleic acid directly within a sample, all bacterial species such as *E. coli* had to be grown on and isolated from commercial media containing 5-bromo-4-chloro-3-indolyl-Beta-D-glucoronic acidic (BCIG) (Prats et al., 2008). BCIG is a chromogenic substrate used for detection of Beta-

glucuronase activity in *E. coli* which is depicted by blue colonies (Prats et al., 2008). Commercially available validated culture based methods such as the MacConkey Agar, Chromogenic Chromocult (CC) and rapid *E. coli* (REC) media are a few examples, which allow simultaneous detection of total coliforms and *E. coli*, using fluorogenic substrates that provide distinct colony colours upon hydrolysis by *E. coli* (Jacob et al., 2014; Prats et al., 2008).

The isolation of strains using these methods is still vital because it enables downstream characterization of individual *E. coli* isolates and comparison of different strains. Identification of *E. coli* based on agar however is usually presumptive and therefore requires further confirmation using molecular techniques such as the polymerase chain reaction (PCR), biochemical methods such as phenotype microarray (PH) (Franz et al., 2011) or matrix assisted ionization coupled with time of flight spectroscopy (Jørgensen et al., 2017).

2.3.2 Culture independent methods

2.3.2.1 Serotyping

All pathogenic bacteria have been historically grouped according to possession of O, H and K-antigens thereby placing them into medically relevant serotypes based on reaction to rat antiserum (Orskov and Orskov, 1992). The O antigens describe the constituents of the of the bacterial cell wall which also play a role in virulence while the H-antigen describes possession of flagellar (Orskov and Orskov, 1992). Initially it was shown that medically relevant pathogens could be easily described using this scheme, but later studies found that it was cumbersome and multiple combinations could exist for the ever increasing number of strains (Nataro and Kaper, 1998). Incidentally, serotype characterization is still vital today for identifying common pathogenic foodborne pathogens implicated in major outbreaks such as *E. coli* O157:H7 and O104:H4 (Croxen et al., 2013). The use of molecular techniques has provided a less cumbersome and a more accurate way to determine serotypes compared to use of antibodies that are prone to false positives (Gao et al., 2016). Ultimately, serotyping identification is important just when you are sure that the strain is responsible for a certain outbreak, but its identification, with the exception of O157:H7 which seems to be distributed worldwide, should rely on the presence of virulence genes, since there are O157 non H7 strains that indeed do not harbor *stx* genes.

2.3.2.2 Molecular sequence typing

Once pathogenic *E. coli* strains have been accurately identified using the methods mentioned above it is crucial to infer genetic relatedness among them to decipher potential sources of contamination or emergence of new strains. Commonly used molecular typing methods include Multi locus sequence typing (MLST), Pulsed-field gel electrophoresis (PFGE), Bacterial repetitive extragenic palindromic sequence based PCR (Rep-PCR) and more recently with a great reduction in costs Whole genome sequencing (WGS) which is taken as the current gold standard for species and strain identification as well as for differentiation (Pérez-Losada et al., 2013).

Multi locus sequence typing

MLST: Involves the characterization of isolates using the sequences of internal fragments of usually seven house-keeping genes (i.e., constitutive genes required for the maintenance of basic cellular functions) (Pérez-Losada et al., 2013). Gene regions of approximately 450–500 bp are sequenced and those found unique within a species are assigned an allele number (Belén et al., 2009). Subsequently, each isolate is characterized by the alleles at each of the seven loci, constituting its allelic profile or sequence type (ST) (Pérez-Losada et al., 2013). MLST has been useful in clinical epidemiology because it is based on analysis of nucleotide sequences, has an easy methodology adaptable to different laboratories and the generated data sets can easily be shared over the internet (Belén et al., 2009).

Pulsed-field gel electrophoresis

PFGE: Was initially developed to enable bacterial strain differentiation based on large DNA molecules (Maule, 1998). The method relies on a direct current (DC) electric field that periodically changes direction and/or intensity relative to the agarose gel (Maule, 1998). The time interval, during which the field is in any one direction is called the pulse time, and its duration is the single most important factor in determining the molecular-size range over which separation is possible (Maule, 1998). The high discrimination ability of the method made it the gold standard for tracking food and environmental pathogens worldwide (Pichel et al., 2012). However, the work load involved in carrying out the analysis that takes up to 3 days has been suggested as a major drawback with the analysis for use in routine epidemiological settings. Uysal and Durak, 2012 used PFGE to track foodborne and antibiotic resistant *E. coli* strain within the food chain. The quick adoption of whole genome sequencing (WGS) enabled by

the reduced costs and availability of resource repositories to analyze bioinformatic data worldwide has led to reduced use of PFGE for tracking foodborne and environmental pathogens (Allard et al., 2016).

Repetitive extragenic palindromic sequence based PCR (Rep-PCR)

Short intergenic repeated sequences that are randomly distributed, but highly conserved within bacteria genomes such as repetitive extragenic palindromic elements (REPs) and enterobacterial repetitive intergenic consensus (ERIC) sequences can be used to assess the genetic diversity of strains (De Bruijn, 1992). Primers can be designed to target these conserved sequences and comparison can be carried out using PCR (Tobes and Pareja, 2006). The quickness of the method and similarity to commonly used PCR makes it a preferable option to methods such as PFGE (Moser et al., 2010). Rep-PCR has been used to compare the diversity of *E. coli* strains from different water and environmental sources (Chandran and Mazumder, 2014).

Whole genome sequencing

WGS: WGS provides the opportunity to unravel the full extent of genetic diversity within a given bacterial species and strain through providing the contents of their nucleotide sequences (Whitman, 2015). This ability is crucial for improving identification, elucidating functional properties of taxonomic groups as well as resolving many of the ambiguities in phylogeny noted with some of the methods mentioned above (Whitman, 2015). For example, genome sequencing can help in classifying an organism, comparing genomes sizes in cases where multiple genomes of the same species are available through calculating the pan- and core genomes (Whitman, 2015). For example, the comparison of more than 2000 *E. coli* genomes found an *E. coli* core genome of about 3100 gene families and a total of about 89,000 different gene families (Land et al., 2015). WGS has provided thorough identification, phylogeny, and almost perfect source tracking of *E. coli* strains from various sources (Kaas et al., 2012; Touchon et al., 2009).

2.3.2.3 PCR based methods

Involve amplification of any nucleic acid sequence present in a complex sample through a cyclic process to generate a large number of identical copies that can readily be analyzed present a viable option (Kubista et al., 2006). For example, in real time PCR the amount of product formed is monitored during the reaction by monitoring the fluorescence of dyes or

probes introduced into the reaction that is proportional to the amount of product formed, and the number of amplification cycles required to obtain a particular amount of DNA molecules is registered. Similarly, digital PCR (dPCR) is a quantitative PCR method that provides a sensitive and reproducible way of measuring the amount of DNA or RNA. (Biassoni and Raso, 2014). In dPCR, the initial sample is distributed into many wells before the amplification step, resulting in 1 or 0 targets being present in each well (Biassoni and Raso, 2014).

After amplification, the number of positive versus negative reactions is determined and the absolute quantification of the target is calculated using a Poisson distribution to amplify targeted amplicons (Biassoni and Raso, 2014). These methods provide a quick and accurate measure of pathogens within given samples.

2.3.2.4 Metagenomics

Since nucleic acids were first extracted directly from the environment and sequenced, metagenomics has grown into one of the most data-rich and pervasive techniques for understanding the taxonomic and functional diversity of microbial communities (Temperton and Giovannoni, 2012). A notable development in the second decade of bacterial genome sequencing was the generation of metagenomic data, which covers all DNA present in a given sample (Land et al., 2015). The outputs could be a mixture of bacterial 16S rRNA genes, of which few to numerous copies are sequenced, or random DNA fragments corresponding to microbial DNA, or retro-transcripts of RNA molecules of which a large number of short sequences are determined (Ceuppens et al., 2014). The sequences generated can be analyzed using bioinformatics to allow categorization of the bacterial content of the original sample and the power of these methods enables sequencing of the entire DNA content of an ecosystem such as a food system or product to be determined (Temperton and Giovannoni, 2012). However, one increasing problematic drawback with this method is the rapid technology changes enabling faster identification of bacterial genomes but with shorter reads making it hard to ably construct community associations (Land et al., 2015). Additionally, the lack of harmonization in data processing methods makes comparison of different studies more difficult (Temperton and Giovannoni, 2012). These issues aside, culture independent methods provide a lot of promise for understanding the bacterial communities within different habitats.

2.4 Diarrheagenic *Escherichia coli* a foodborne pathogen

Diarrheagenic *E. coli* (DEC) is a leading cause of foodborne disease around the world affecting all age groups and demographics with symptoms including but not limited to diarrhoea, gastroenteritis, inflammation and nutrient malabsorption (Croxen et al., 2013; Leimbach et al., 2013). At highest risk of illness and death are new-borns, infants, the elderly as well as individuals with compromised immunity such as those living with Human Immuno-Deficiency Virus (HIV). While specific numbers relating to different geographical areas differ, higher rates of infection and disease have been observed in developing and least developed nations of the world (Acosta et al., 2016; Kotloff et al., 2013). However, recurrent outbreaks are also frequently reported in the developed world (Kaur et al., 2010; Rasko et al., 2011). Interestingly, the causes of illness in both these areas differ with poor sanitation being blamed for endemic DEC prevalence in less developed nations (Ahmed et al., 2013; Baker et al., 2016). In contrast, within developed countries, cross-contamination of food products along the heavily internationalized logistics food chain involving heavy mechanization at large food processing and distribution centres is a significant risk factor attributed to outbreaks.

The main pathotypes which are defined by how they illicit disease and having significant association with foodborne illness include; enteropathogenic *E. coli* (EPEC), enterotoxigenic *E. coli* (ETEC), shigatoxin producing *E. coli* (STEC), enteroaggregative *E. coli* (EAEC), enteroinvasive *E. coli* (EIEC), diffusely adherent *E. coli* (DAEC) and adherent and invasive *E. coli* (AIEC). However, EPEC, ETEC, EHEC and EAEC are noted to constitute the highest burden to human disease. It has been suggested as mainly transferred through the oral-faecal route hence making them foodborne pathogens (Kirk et al., 2017). Proper sanitation within domestic and food processing environments is an effective prevention strategy.

2.5 The agricultural and food production environment as a hotbed for emerging Diarrheagenic *Escherichia coli*

The emergence of highly virulent strains causing large-scale outbreaks is a big concern. STEC O157:H7 as well as other shigatoxin (STEC) producing *E. coli* strains (non-O157 STEC) still cause many outbreaks in developed countries (Havelaar et al., 2010; Kirk et al., 2017; Tauxe et al., 2010). This is in spite of developed countries having better tracking systems for foodborne pathogens and illness than developing nations (Nhampossa et al., 2015; Oundo et al., 2008). In 2011, a highly virulent EAEC strain O104:H4 code named the ‘German outbreak

strain' producing Shiga toxins was clinically described as enteroaggregative haemorrhagic *E. coli* (EAHEC) (Beutin and Martin, 2012; Bielaszewska et al., 2011; Rasko et al., 2011). This strain caused the deadliest foodborne outbreak associated with *E. coli* in Europe. This outbreak affected many sectors of European society such as health, trade and politics (Commission of The European Communities, 2011). This outbreak leads to speculation that genetic mechanisms in the DEC group may have favoured the emergence of highly virulent foodborne within the agricultural and food production ecosystem.

2.6 Linking Diarrheagenic *Escherichia coli* persistence in the agricultural and food production environment to inherent bacterial attributes

A great body of work has elucidated the pathogenicity of most DEC pathotypes within the human body (Blount, 2015; Croxen et al., 2013; Leimbach et al., 2013; Nataro and Kaper, 1998). However, food and environmental microbiologists lack a comparable mechanistic understanding of phenotypic or genotypic factors enabling environmental prevalence (Berg et al., 2014; Dini-Andreote et al., 2012; Martínez, 2013; van Elsas et al., 2011) and subsequent survival of DEC.

Fortunately, new research has begun to shed more light on how these pathogens may survive within hostile environments including water, soil and plants (Chandran and Mazumder, 2015; Franz et al., 2011; Macarisin et al., 2012; Nagy et al., 2015; Saldaña et al., 2011). Reperant et al., 2012 suggest that gut microbes including DEC have co-evolved with humans over the past several thousand years specifically targeting and responding accordingly to lifestyle changes within their human hosts which enables seamless transfer into the host. Additionally, some of the DEC virulence genes may provide an evolutionary advantage helping to sustain life of the strains in the open environment (Martínez, 2013).

As environmental changes driven by climatic and anthropogenic activities escalate, understanding how these processes affect the prevalence and persistence of foodborne pathogens such as DEC is crucial. This is because DEC are still a menace to food safety, food security and human health. In this mini-review, we therefore aim to shed light on some of the work that has sought to understand the relationship of DEC in the open environment that provides a direct link with persistence along the human food chain (Figure 2.2). Most of the reported work is understandably focused on outbreak pathogens such as *E. coli* O157:H7 and *E. coli* O104:H4 which may obscure the risks presented by other strains because of the high

heterogeneity within many DEC pathotypes. However, because the former strains have previously been implicated in large-scale outbreaks, they can be used as standard models for studying DEC behaviour as seen in the clinical setting with the use of prototypes. The validation of these outbreak strains with a larger set of strains will shed more information regarding their persistence within the food and agricultural open environment and thereby risk posed to food safety and public health

Additionally, understanding the genotypic and phenotypic characteristics enabling DEC environmental persistence shall provide insight into how new pathotypes emerge (Figure 2.3). In our review, we shall focus on three broad categories that may work in tandem to enable DEC environmental persistence. They include; attachment to biotic and abiotic surfaces, evolution and adaptation as driven by lateral gene transfer and response to stressful conditions in the environment.

2.6.1 Diarrheagenic *Escherichia coli* attachment to biotic and abiotic surfaces

To survive within the external environment, DEC needs to establish itself onto a biotic host such as a plant or an environmental surface. Bacterial attachment subsequently leading to colonization and persistence is a complex interaction between planktonic bacteria, surface associated bacteria and the surface (Carter et al., 2016). The attachment process is dynamic and governed by bacterial cellular appendages (Wong et al., 2017). Additionally, it is sequential and bi-directional and depends on the prevailing environmental conditions (Beloin et al., 2008). The control of attachment occurs through different regulatory systems such as the *Rcs* two-way component regulatory pathway that aids remodelling of the bacterial surface. On the other hand, the *Enz/OmpR* two-component pathway helps sense changes in osmolarity (Beloin et al., 2008). For example the *ompR234* protein promotes biofilm formation by binding the *csgD* promoter region and stimulating its transcription (Prigent-combaret et al., 2001). Additionally, the *csgD* gene encodes the transcription regulator *csgD* which in turn activates the transcription of the *csgBA* operon encoding curli which are extracellular structures involved in bacterial attachment (Prigent-combaret et al., 2001).

E. coli attachment to surfaces is either reversible or irreversible (Beloin et al., 2008).

Reversible attachment involves physicochemical and electrostatic interactions between the bacterial envelop and environmental substrates while, irreversible attachment uses adhesive organelles of the fimbrial family such as; type I fimbriae, curli and conjugative pili (Beloin et al., 2008).

Additionally, irreversible attachment leads to progressive inter-bacterial attachment and build-up of the resulting biomass forming a biofilm, which describes a group of single or multispecies bacteria living together as a multifunctional unit of existence for the benefit of all organisms (Wood, 2008). Cellular appendages and bacterial secretion systems (SS) such as the Type 3SS (Saldaña et al., 2009), Type 6SS (Bernard et al., 2010; Boyer et al., 2009) and Type 7SS (Hufnagel et al., 2015) are associated with attachment and colonization of abiotic surfaces through biofilm formation in a complex interdependent way. This process follows different regulatory pathways and may differ from strain to strain (Niba et al., 2007).

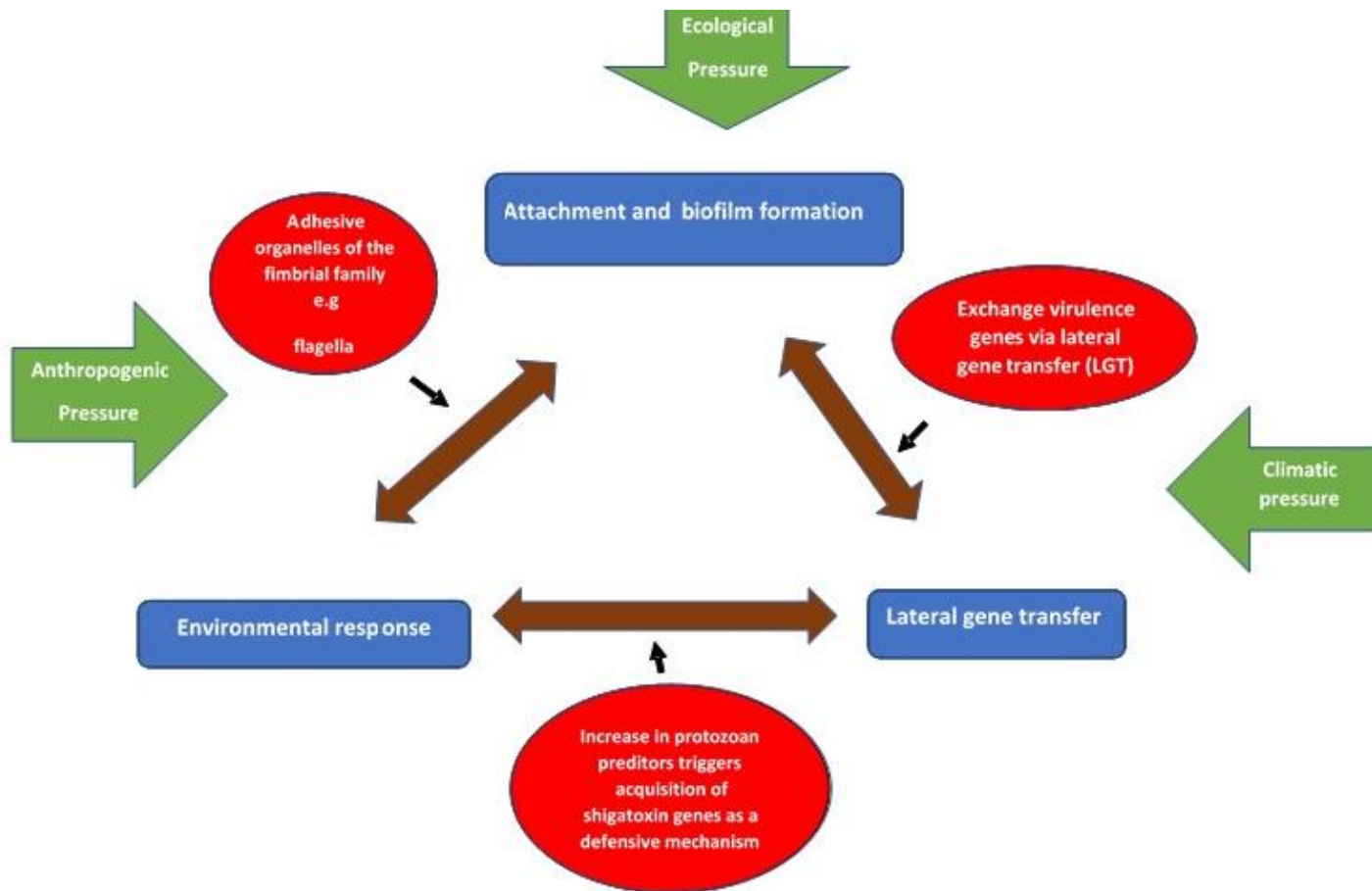


Figure 2.2: Interplay of interdependent bacterial processes that may facilitate the persistence of Diarrheagenic *Escherichia coli* (DEC) in the open environment. The processes can be bi-directional depending on the prevailing environmental conditions influenced by anthropogenic, climatic and ecological factors

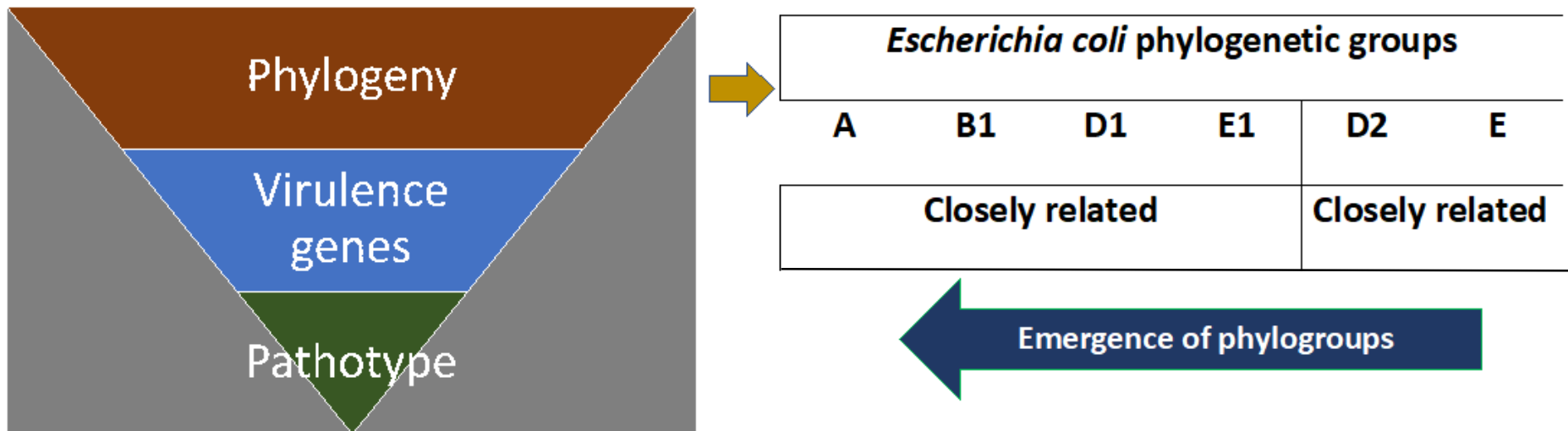


Figure 2.3: The phylogeny of *E. coli* influences the acquisition of virulence genes and subsequently the diarrheagenic pathotype (DEC) which may inadvertently lead to emergence of new pathotypes in varying environmental niches such as the agricultural and food production setting. “The phylogenetic neighbourhood of geographically remote *E. coli* supports the notion of a rapid worldwide spread of an evolutionary common ancestor (maybe with the advent of mammals and selection in specific habitats.” (Chaudhuri and Henderson, 2012). Phylogroups A (mostly commensals) and B1 (non-O157 shigatoxin producing (STEC) strains) form the newest clades. Phylogroups D1 (uropathogenic (UPEC) and enteroaggregative (EAEC) strains) and E (separate clade of O157:H7 STEC and O55:H7 enteropathogenic (EPEC) strains). Phylogroups D2 and B2 (EPEC strain E2348/69 and extraintestinal (ExPEC) strains) form the oldest clades.

2.6.1.1 Initial attachment as an adaptive strategy to initiate colonization of environmental surfaces and habitats by Diarrheagenic *Escherichia coli*

For DEC to persist within the open environment, the persistence usually begins with planktonic bacteria multiplying to form a whole community leading to colonization. The factors enabling this persistence may include cell surface appendages such as curli, type I fimbria, type IV pili and flagella (Beloin et al., 2008). Additionally, bacterial secretion systems aid in this process by facilitating the functioning of bacterial appendages through the type 2SS (Beloin et al., 2008), type 3SS (Saldaña et al., 2009), type 6SS (Silverman et al., 2012) and type 7SS (Hufnagel et al., 2015). Additionally, the initial attachment can be facilitated and reinforced by extracellular polysaccharide substances such as cellulose (Borgersen et al., 2018; Macarisin et al., 2012).

Different *E. coli* appendages were shown to play a crucial role in its initial attachment and stability to abiotic surfaces such as silica (Wong et al., 2017). Total internal reflection fluorescence (TIRF) microscopy was used to investigate the attachment and stability of *E. coli* deletion mutants of curli ($\Delta csgA$), flagella ($\Delta flhA\Delta c$) and type I fimbriae on silica surfaces (Wong et al., 2017). It was found that the absence of curli and flagella diminished ability to attach strongly in comparison to type I fimbriae deficient strains that were comparable to the wild-type strains thereby indicating that removal of fimbriae did not affect initial attachment to abiotic surfaces (Wong et al., 2017).

Due to many foodborne outbreaks associated with fresh green produce, studies have sought to elucidate the mechanisms through which DEC attaches and persists on these ready-to-eat minimally processed food products. For example, Macarisin et al., (2012) used two *E. coli* O157:H7 outbreak strains, to investigate the role of pili and cellulose expression in adherence to spinach leaves and noticed that curli strains had a stronger association with leaf surfaces than curli mutants. Additionally, attachment increased with incubation time (0, 24, 48 h). Cellulose mutants had strong attachment nonetheless and curli expressing strains had an extracellular matrix (Macarisin et al., 2012). Similarly, (Saldaña et al., 2009) had previously reported that curli and cellulose were co-expressed in the presence of the transcriptional activator *csgD* in attaching and effacing *E. coli* strains such as EHEC and EPEC. Additionally, they noticed that *Fis* (factor for inversion stimulation) was a negative transcriptional activator for curli expression (Saldaña et al., 2009).

The persistence of DEC strains in the open environment has been suggested to follow the same mechanism or at least share similarities with observations in the human host (Berg et al., 2014; Dini-Andreote et al., 2012). For instance, the locus of enterocyte effacement (LEE) pathogenicity island found in EPEC and STEC facilitated attachment of *E. coli* O157:H7 strains to baby spinach leaves, entry into the stomata, intracellular and vascular tissue (xylem and phloem) (Saldaña et al., 2011). This process occurred through the type III secretion system and effector genes that coordinated production of curli, the *E. coli* common pilus, haemorrhagic coli type 4 pilus and flagella (Saldaña et al., 2011). These bacterial surface structures have also been shown to attach equally strongly to abiotic surfaces such as glass and stainless steel which are common food contact surfaces. For example, filament deficient O157:H7 mutants were shown to bind less strongly to glass (Nagy et al., 2015) and steel (Carter et al., 2016) surfaces compared to the wild strains. However, curli based attachment was shown to vary largely depending on the strain, growth medium and abiotic surface (Carter et al., 2016).

Like STEC and EPEC attachment and persistence mechanisms within the open environment, EAEC strains have been shown to use virulence factors in human infections. The attachment of EAEC to human epithelial cells is facilitated by fimbrial and afimbrial (outer membrane proteins) adhesins that help in formation of the aggregative adherence (AA) phenotype in both typical and atypical EAEC strains (Estrada-Garcia and Navarro-Garcia, 2012). Aggregative adherence fimbriae (*AAF I to III*) are associated with attachment and the AA phenotype in typical EAEC while strains lacking these structures have been shown to use a type IV pilus (atypical strain C1096) encoded by genes located in an *IncI1* plasmid (Dudley et al., 2006a) while the *Hda* adhesin is responsible for attachment in strains lacking the *AAF* (Boisen et al., 2008).

To investigate the adhesive mechanism of the German outbreak *E. coli* O104:H4 strain on fresh produce and abiotic surfaces, Nagy et al., (2016) noticed that deletion of the *aggA* gene which codes for the *AAF/I* fimbriae subunit (*AggA*) greatly reduced colonization. This observation suggests that it plays a role in the adherence and subsequent colonization of spinach leaves. In contrast to the attachment mechanism of typical EAEC foodborne pathogen such as *E. coli* 104:H4 possessing the transcriptional activator *AggR*, that regulates virulence mechanisms involved in attachment (Estrada-Garcia and Navarro-Garcia et al. 2012), atypical EAEC (lacking *AggR*) were noted to mediate AA and attachment to abiotic surfaces such as glass through a type IV pilus carried on the Incompatibility class 1 (*IncI1*) plasmid of the EAEC strain C1096 (Dudley et al., 2006a).

This observation shows the various differing mechanisms that facilitate the persistence of DEC foodborne pathogens in the open environment thereby complicating potential intervention and prevention measures.

2.6.1.2 Biofilm formation as an adaptation utilized by Diarrheogenic *Escherichia coli* to enhance environmental persistence in the open environment

Biofilms describe sessile communities of single or multiple species that are characterized by long-term bacteria to bacteria and subsequently bacteria to surface adherence that favour attachment to surfaces in comparison to a planktonic state (Niba et al., 2007). Biofilm formation occurs through the Type 6 secretion system in and depending on the availability of favourable conditions occurs in a contact-dependent mechanism of effector delivery (Silverman et al. 2012). The ability of DEC to form biofilms has been shown to support persistence within the food production environment (Giaouris et al., 2015) and domestic setting (Ahmed et al., 2013) supporting its ability to cause human infections through cross-contamination. For example, the formation of biofilms by ETEC on drinking water contact surfaces is linked to the warm and humid months in Bangladesh, a period when high rates of diarrheal disease (epidemic levels) are noted among poor households (Ahmed et al., 2013). Similarly, biofilm formation by potentially pathogenic *E. coli* was noticed to occur under conditions relevant to the food production chain (Nesse et al., 2014). In this work, intimin positive (*eae+*) strains had higher biofilm formation capacity at lower temperatures (12° and 20°C) than those without intimin (*eae-*) (Nesse et al., 2014). Intimin negative (*eae-*) strains instead had higher biofilm capacity at a high temperature (37°C) suggesting that the intimin attachment protein played a role in environmental attachment albeit at lower temperatures (Nesse et al., 2014). The observations from this work suggest that DEC may use different mechanisms to promote biofilm formation and thereby persistence based on the prevailing conditions. Secondly, the same attachment structures such as intimin that facilitate persistence of EPEC and STEC in the human body may also be applicable in the open external environment. These scenarios further emphasize DEC versatility and the risk posed to food contamination and subsequent illness. Aside from temperature, environmental factors such as nutrient medium and the attachment surface affect the ability of DEC to form biofilms and consequently further persistence within a food production environment. On using two *E. coli* O157:H7 outbreak strains possessing curli fimbriae to assess biofilm formation in differential media and two different abiotic surfaces, (Carter et al., 2016) noticed higher biofilm formation

on stainless steel (Luria broth without salt) and polypropylene (spinach lysates). The change in physiological state of the strains when grown in spinach lysates preventing curli production coupled with the presence of other uncharacterized adherence factors was suggested as the cause of reduced biofilm production observed on stainless steel surfaces in the respective samples (Carter et al., 2016).

2.6.2 Lateral gene transfer as a tool utilized by DEC for adaptation to the agricultural and food production environment

The diversity of DEC is primarily associated with the acquisition of mobile elements such as plasmids, bacteriophages, genomic islands and chromosomal DNA through mechanisms such as conjugation, transduction and natural transformation (Blount, 2015; Leimbach et al., 2013). The foreign genetic material can be acquired through exchange with closely and distantly related strains within a specified host (human gut) or within the open environment by the acquisition of naked free flowing DNA packaged in a plasmid or bacteriophage.

Through the acquisition of foreign DNA, its survival of cellular defence mechanisms, establishment via Lateral gene transfer (LGT) into the main chromosome or as an extrachromosomal element and clonal multiplication in the new host, the newly acquired genetic information may become abundant in the host cell population. Additionally, via mutations, expression of the obtained DNA may be further regulated and tuned with the encoded proteins for better integration into the bacterial cellular networks (Skippington and Ragan, 2011). The core bacterial chromosomal genes that initiate cellular processes such as replication, transcription and translation are tightly regulated. Consequently, acquiring foreign DNA is crucial in helping adjust to new environmental pressures (Leimbach et al., 2013).

2.6.2.1 Mobile genetic elements assisting in the adaptation to a hostile open environment

It has suggested that only genes necessary for critical bacterial processes will be absent in the open environment (Touchon et al., 2009). This is because apart from the genes required for vital metabolic and nutritional bacterial processes necessary for survival, the rest are accessory genes that can be acquired from or dispelled into the environment (Soborg et al., 2013). Additionally, based on the prevailing environmental conditions, these accessory genes facilitate adaptation to a given environmental niche (Soborg et al., 2013; Touchon et al., 2009). This suggestion has been corroborated by (Sidhu et al. 2013) and Soborg et al. (2013) who found many bacterial virulence genes associated with DEC including toxins, adhesins, secretion systems and regulators of virulence factors within the soil and freshwater

environments. The acquisition of these genes correlated to several factors such as genome size, genome G-C composition, carbon utilization and oxygen tolerance (Jain et al., 2003; Skippington and Ragan, 2011a) with the process favouring more closely related strains.

The family *Enterobacteriaceae*, in which DEC resorts, forms part of a large group of bacterial species that have commensal, symbiotic or pathogenic relationships within the open environment (Van Overbeek et al., 2014). Such as scenario can potentially aid persistence and fitness of DEC through the exchange of mobile genetic elements (Fletcher et al., 2013; Van Overbeek et al., 2014). Foodborne DEC have larger genomes than commensals suggesting the ability to tolerate variable conditions facilitated by a more extensive repertoire of adaptive genes (Van Overbeek et al., 2014). This evolutionary adaptation has been suggested as a means by which DEC increase environmental persistence to heightening the chance of entering the human body through the faecal-oral route (Xicohtencatl-Cortes et al., 2009) because these pathogens have evolved to closely match changes in human diets (Reperant et al., 2012).

These findings suggest that DEC faces pressures within the open environment that select for survival traits, many of which are virulence genes (Soborg et al., 2013).

2.6.2.2 Lateral gene transfer and its role in the emergence of Diarrheogenic *Escherichia coli* pathotypes and subsequent environmental persistence

DEC pathotypes have evolved independently through multiple acquisitions of virulence factors by Lateral gene transfer (LGT) resulting in diverse and dynamic genomic structures that may represent overlapping ecological niches as they have a different distribution in humans, domesticated animals and wild animals (Tenailon et al., 2010). However, *E. coli* classification based on DEC pathotypes mostly provides a clinical/public health categorization of a small group of strains within the more abundant species with a usually vague description of environmental distribution. Therefore, this categorization unintentionally neglects crucial factors such as phylogeny that may influence the ecology and distribution of this group of pathogens within the open environment outside the human host (Robins-Browne et al., 2016).

2.6.2.3 DEC phylogeny and the emergence of new pathotypes

DEC pathotypes fall within all phylogroups of *E. coli* phylogeny including phylogroups A, B1, B2, D1, D2 and E (Chaudhuri and Henderson, 2012). A close relationship exists among phylogroups A, B1, D1 and E of which (A and B1) (Figure 2.3) form the most recently emerged phylogroups of strains (Chaudhuri and Henderson, 2012).

On the other hand, phylogroups B2 and D2 form the backbone of the ancient strains from which the other groups have emerged making them the earliest members on the ancestry tree (Leimbach et al., 2013). Group A is mostly composed of commensals, B1 has non-O157 EHEC, D1 has EAEC and EHEC O157 while ETEC does not fall in any phylogroup but rather occurs across the whole spectrum. This broad spectrum of characterization can complicate the ability to study the ecology of DEC strains (Leimbach et al., 2013). The close relationship of phylogroups A, B1, D1 and E1 which consist of common foodborne pathogens, as well as commensals, may facilitate the exchange of virulence factors within the open environment and hence enhance the emergence of new strains (commensal and pathogen) or more lethal pathogens (pathogen to pathogen). This could result from the high gene flux in *E. coli* which causes closely related strains to share a significant amount of accessory genes (Touchon et al., 2009). Additionally, at broader taxonomic scales lateral gene transfer is more frequent within than between taxonomic groups because such strains usually share a similar life style (Skippington and Ragan, 2012). For example, intra-group edges linking phylogenetic groups A, B1 and B2 with D are more frequent as most are commensal compared to extra-intestinal pathogens (B2 and D) which raises the possibility of preferential transfer (Skippington and Ragan, 2012). Additionally, phylogeny has been suggested as more important than ecology in influencing genetic exchange because genetic material exchanged among closely related individuals can be integrated by homologous recombination and has greater compatibility with the native host (Jain et al., 1999). Furthermore, the shared evolutionary history of closely related individuals biases the uptake of genetic material including phage host infection biases, DNA uptake specificity and quorum sensing (Thomas and Nielsen, 2005). The bacterial life style, environmental niche and phylogeny do not substantially hinder lateral gene transfer although transfer is biased towards closely related species (Skippington and Ragan, 2012).

Consequently, phylogeny may help food safety and public health practitioners to better understand the ecology as well as metabolic and phenotypic characteristics associated with the different DEC pathotypes within the open environment. This information can be used to infer DEC ecological distribution over geographic and temporal scales which is vital in coming up with appropriate interventions for preventing the contamination and subsequent proliferation of these pathogens on food and within environmental sources, thereby reducing the risk of foodborne illness. Additionally, by using phylogeny which describes the genetic ancestry of DEC through the broader spectrum of the whole *E. coli* species, we can potentially assess the

risk of emerging foodborne pathogens through the acquisition of virulence factors by Lateral gene transfer.

2.6.2.4 Increased environmental persistence and virulence of DEC through the acquisition of mobile genetic elements

By acquiring new virulence genes, DEC pathotypes may be able to survive longer in the open environment through increased virulence which usually correlates with resistance to increasing environmental pressure (Soborg et al., 2013). Additionally, new pathotypes may emerge that could increase the risk of causing human illness through the faecal-oral route. Phylogenetically related DEC such as non-O157 STEC (B1), O157 STEC (E) and EAEC (B1) may easily exchange genes leading to the emergence of new pathotypes (Chaudhuri and Henderson, 2012). This scenario is not unprecedented and has been reported in recent times such as during the 2011 German *E. coli* outbreak associated with *E. coli* O104:H4 strain that had characteristics of both EAEC and STEC (Rasko et al., 2011). This emerging foodborne pathogen was suggested to have acquired a plasmid containing the *AAF/I* locus but having lost the *AAF/II* locus while gaining a plasmid with the gene encoding *CTX-M-15* extended-spectrum beta-lactamase (ESBL) producing antibiotics (Van Overbeek et al., 2014). Additionally, this pathotype was noticed to have recently acquired the *Stx2* encoding phage leading to its notably increased virulence compared to prototypical EAEC strains (Rasko et al., 2011) further suggesting that emergence of new pathogenic subtypes is ongoing and presents a viable risk to food safety and public health.

Similarly, *E. coli* O157:H7 is suggested to have diverged from an O55:H7 EPEC precursor by the acquisition of virulence factors such as phage-encoded Shiga toxin (Leimbach et al., 2013). The ability to acquire this toxin has been suggested as a defensive mechanism against predation by protozoa in the open environment (Martínez, 2013) which indirectly ended up as a negative consequence to human health by causing haemorrhagic colitis (HC) and haemorrhagic uremic syndrome (HUS). Additionally, STEC strains such as O157:H7 strain EDL 933 use the locus of enterocyte effacement (LEE), a genomic island whose actions are facilitated through the T3SS to facilitate the persistence on green leafy vegetables as well as internalization into the stomata and plant vascular tissues (Xicohtencatl-Cortes et al., 2009). The LEE is a virulence factor in EPEC and EHEC intestinal infections that is associated with the attaching and effacing phenotype (Kaper et al., 2004).

In the food industry, heat resistance within bacterial pathogens is a primary concern that can complicate the provision of safe food since heat treatment is a conventional processing technique that helps kill pathogens such as DEC. However, foodborne *E. coli* strains have shown high heat resistance (Mercer et al., 2015). In this study, it was reported that strains were found to possess a D_{60} -value of greater than 10 minutes with subsequent genomic analysis of the strains revealing possession of a pathogenic island, the locus of heat resistance (LHT), with high homology to those found in *Cronobacter sakazakii* and *Klebsiella pneumoniae* (Mercer et al., 2015). The transfer of the LHT into commensal strains led to the acquisition of the heat resistant phenotype (Mercer et al., 2015) further exemplifying how these elements can lead to the emergence of highly resilient strains.

2.6.3 *Escherichia coli* response to the changing environmental conditions with emphasis on Diarrheagenic *Escherichia coli*

To survive the harsh conditions of the open environment, DEC have evolved a complex interplay of bacterial response systems that facilitate adjustment to the prevailing conditions. Principal among these interconnected response systems includes the alternate sigma factor, *RpoS* that is a global response mechanism to stressful conditions (Battesti et al., 2011). Additionally, bacterial secretion systems including the type 3 secretion systems, type 4 secretion systems and type 6 secretion systems form apparatuses that aid in translocation of proteins and DNA through and across cellular membranes enabling response to conditions such as changes in temperature, osmotic pressure, oxygen levels and moisture (Green and Meccas, 2015).

2.6.3.1 *Escherichia coli* response to environmental stresses with emphasis on Diarrheagenic *Escherichia coli* in the open environment outside the human host

E. coli when encountering stressful conditions such as temperature, pH, osmotic pressure and nutrient starvation can trigger global stress response mechanisms dependant on the alternate sigma factor *RpoS* (Battesti et al., 2011). The *RpoS* through interaction with the core RNA polymerase (RNAP) and in the process controlling approximately 500 genes allows bacteria to withstand stressful conditions and treatments (Ng and Bassler, 2009). For example, in the common foodborne pathogen *E. coli* O157:H7 this factor enables increased tolerance to low pH conditions (Chauret, 2011), a hurdle commonly used to improve the shelf life of food.

Additionally, the *RpoS* gene has been noted to control the expression of the heat shock (*dnak*) and cold shock (*CspA*, *yfiA*) proteins which have been shown to have high expression in *E. coli* O157: H7 during stress conditions (Chauret, 2011).

In a bid to study the proteome of *E. coli* O157:H7 and *E. coli* O104:H4 using tandem mass spectrometry under minimal nutrient conditions, Islam et al., (2016), noticed in both pathogens high prevalence of the virulence factors commonly associated with human disease. For example, proteins from O157:H7 include LEE proteins such as intimin and *Tir* involved in iron scavenging (Islam et al., 2016). Similarly, proteins from O104:H4 included *AAFI*, serine protease autotransporters, beta-lactamases and Shiga toxin 2 subunit B (Islam et al., 2016). These results suggest that to survive outside the human intestines, DEC must be able to respond to stresses of poor nutrient environments probably by facilitating expression of virulence genes (Islam et al., 2016). Increases in expression of virulence and antibiotic resistance genes maybe a response mechanism to stresses such as limited nutrient conditions as a means of facilitating adaption to environmental conditions outside the human host (Islam et al., 2016).

The ability to metabolize different environmental substrates may also provide an evolutionary advantage for pathogens such as *E. coli* O157:H7 to survive in the open environment (Franz et al., 2011). For example, the ability to grow on propionic, alpha-ketobutyric and alpha-hydroxybutyric acid correlated with an increased survival time of *E. coli* O157:H7 in manure-amended soils (Franz et al., 2011).

Additionally, the *E. coli* pentabolome which describes all its metabolic reactions is composed of mostly (57%) core reactions that are common to all strains (Leimbach et al., 2013). The majority of the pentabolome reactions are anabolic (molecule building) while most catabolic (molecule breakdown) reactions make up the dispensable metabolome (Leimbach et al., 2013). Therefore, *E. coli* and notably DEC have evolved mechanisms that can acquire or dispense genes to respond to the prevailing environmental conditions facilitating growth and colonization of a given environment.

2.6.4 *Escherichia coli* bacterial secretion systems and their role in environmental persistence

The type 3 secretion system spans three cellular membranes (inner, outer and eukaryotic host cell membrane) enabling bacteria to deliver effector proteins into host cells allowing bacterial

survival and colonization (Deng et al., 2017). In STEC O157:H7 strain EDL933, the Type 3 secretion system gene cluster was noted to enable persistence on leafy green vegetables (Xicohtencatl-Cortes et al., 2009). Similarly, the Type 4 secretion system clusters are large protein complexes traversing the cell envelopes of many bacteria and contain a channel through which protein-protein or protein DNA complexes can be translocated enabling transfer of virulence genes and hence facilitation of environmental adaptation (Skippington and Ragan, 2011b). For example, most type 4SS clusters in Gram-negative bacteria encode a small protein that resembles the *TraA* pilin (encoded by the *E. coli* F plasmid) which might help establish contact between donor and target cells (Skippington and Ragan, 2011b).

Furthermore, the Type 6 secretion system also negotiates interactions with eukaryotic and prokaryotic competitors in Gram-negative bacteria by encoding cytoplasmic, periplasmic and membrane proteins to form a trans envelope apparatus (Silverman et al., 2012). The type 6 secretion system is tightly regulated together with other virulence gene determinants such as quorum sensing and flagella synthesis to help the bacteria respond to changes in the different environments such as water, soil or specific host tissues (Leung et al., 2011).

This coordinated response occurs through the activation of bacterial secretion systems and flagella synthesis at the proper time (Leung et al., 2011). For example, in EAEC, expression of the *sci-2* gene cluster (a T6SS gene cluster), is positively regulated by *AggR*, and *Arac*-like transcriptional factors also inducing expression of plasmid carried genes mediating aggregative adherence (Dudley et al., 2006b).

2.7 Conclusion and future perspectives

In this mini-review, we have sought to explore some of the inherent bacterial properties responsible for the persistence of DEC in the open environment. Our specific focus on outbreak pathogens suggests they employ the same virulence genes at least in some part to navigate the harsh environmental conditions. This scenario implicates virulence genes in the facilitation of DEC persistence within the human body and open environment.

Elucidating these mechanisms of persistence will enable the use of effective food safety interventions that shall prevent contamination and reduce the persistence of DEC within the food chain thereby reducing the risk of foodborne illness. For instance, understanding how DEC attaches to different food contact surfaces coupled with the identification of processing

conditions favouring expression of attachment extracellular proteins (biofilm formation) shall enable the design of resilient food contact surfaces and optimized of processing conditions.

While a decent body of literature is emerging regarding the circulation of DEC within the open environment, there is still a lot of information lacking regarding the distribution of these pathogens in different geographical and temporal regions. Most large-scale studies have focused on characterization of DEC from clinical specimen although these pathogens spread via the oral-faecal route. Therefore, more studies tracking the prevalence of these pathogens within food and environmental sources is warranted since these pathotypes have emerged independently worldwide within different ecological niches (Chaudhuri and Henderson, 2012; Leimbach et al., 2013).

The rise of microbiome research (human, soil, water, built environment) should help shed more light on the ecology of DEC within the open environment. The total number of microbiota and genes within a given environment (microbiota) has been shown to influence not only the overall ecological health but also the health of humans and animals within it (Alivisatos et al., 2015; Mariadassou et al., 2015; Martiny et al., 2015; Ramirez et al., 2018). Additionally, the role of higher taxonomic groups in predicting the overall population of lower groups (species) has been suggested as a means of tracking prevalence of pathogenic and non-pathogenic microbial groups over geographical and temporal sites leading to what ecologists would describe as biogeography (Philippot et al., 2010).

Some questions that may require answers in the future include the following: What type of microbiome favours the emergence and persistence of DEC within an agricultural food processing and environment? What environmental pressures either anthropogenic through factors such as antibiotic resistance or climatic through warmer temperatures drive the emergence of new DEC pathotypes in previously unaffected foods and ecological habitats? How can we track such heterogeneous pathogens using quick diagnostics especially in the changing landscape of microbiology that is encouraging targeting of microbial communities rather than single strains?

Answers to many of these questions will be obtained through the collaborative effort of agricultural, clinical, environmental, food and veterinary practitioners seeking to understand the underlying mechanisms enabling proliferation of such pathogens (Fletcher et al., 2013). For instance, global initiatives such as the ‘one health’ approach that seeks to bring together

environmental and health professional and make their research collaborative (Atlas, 2012), is a step in the right direction.

Bestprof.com

3. Hypothesis and Objectives

3.1 Hypotheses

The study stated the following hypotheses;

- i.) *E. coli* strains isolated from food sources and irrigation water from South Africa will possess virulence genes associated with the different Diarrheagenic *E. coli* (DEC) pathotypes. South Africa has a high diarrheal disease burden mainly associated with vulnerable groups such as infants and immune compromised adults (Gray and Vawda, 2016). Additionally the country has been associated with the deteriorating bacteriological quality of surface water sources which predisposes the food chain to contamination with foodborne pathogens (Aijuka et al., 2015; Caine et al., 2014; Ijabadeniyi et al., 2011). DEC has been associated with many foodborne pathogens many of which are transferred through the fecal-oral route (Ceuppens et al., 2014; Croxen et al., 2013). Additionally, DEC has been noted to use virulence factors associated with common human disease such as aggregative adherence fimbriae I (AAF/I) (Nagy et al., 2016) and the type 3 secretion system (Saldaña et al., 2011) to facilitate adherence to food sources enabling persistence and subsequent transfer to the human host. Furthermore, DEC have been suggested to live a bi-phasic life style that enables extended persistence within the open environment outside the human host (Dini-Andreote et al., 2012)
- ii.) Non-DEC associated with food sources and irrigation water will be associated with phenotypes associated with intestinal dysfunction, a condition that describes destabilization of the normal microflora associated with a healthy gut. Intestinal dysfunction has been associated with a bloom in bacteria from the phylum Proteobacteria, family *Enterobacteriaceae* and species *E. coli* which increases the risk to inflammation and gastrointestinal disease (Blanton et al., 2016; Shin et al., 2015). Inflammation generates nitrate which is utilized by *Enterobacteriaceae* anaerobically thereby outcompeting the obligately anaerobic groups, Bacteroides and Firmicutes that are associated with a healthy gut (Shin et al., 2015). South Africa has a high diarrheal disease burden especially among vulnerable groups such as infants and the immune-compromised that may be linked to higher rate of inflammation. Additionally, a high prevalence *E. coli* has been isolated from food sources (Njage and Buys, 2017; Ntuli et al., 2016) and irrigation water sources

(Aijuka et al., 2015; Ijabadeniyi et al., 2011) suggesting a link between the high diarrheal disease burden among the country's vulnerable population with the equally high prevalence of inflammation causing *E. coli* strains in the food chain

- iii.) Real time quantitative PCR (qPCR) coupled with melt curve analysis provide the capacity for simultaneously detecting Cefotaxime-resistant (CTX-M) extended beta-lactamase enzyme producing genes and virulence factors (shigatoxin 1, shigatoxin 2 and intimin) in *E. coli* previously isolated from food sources and irrigation water in South Africa. QPCR technology merges the polymerase chain reaction chemistry with the use of fluorescent reporter molecules in order to monitor the production of amplification products during each cycle of the PCR reaction (Navarro et al., 2015). Intercalating dyes such as SYBR Green bind to the minor groove of double stranded DNA (dsDNA) causing fluorescence that increases and can be measured in the extension phase of each cycle of qPCR (Eischeid, 2011). DNA melting curve analysis using dsDNA-specific dyes produces complex and reproducible melting profiles, resulting in the detection of multiple melting peaks from a single amplicon and allowing the discrimination of different species (Rasmussen et al., 2007). Identification of the amplicon of interest can be achieved by examining the first derivative of the melting curve and identifying the characteristic "melt peak" (T_m), which is the temperature at which the rate of fluorescence change (DNA denaturation) is highest and is observed in the raw data as a sudden decrease in fluorescence (Rasmussen et al., 2007).

3.2 Objectives

The objectives of the study were to;

- i.) Characterize *E. coli* previously isolated from food sources and irrigation water in South Africa based on the major DEC pathotypes commonly associated with foodborne disease including; enteropathogenic *E. coli* (EPEC), enterotoxigenic *E. coli* (ETEC), enteroaggregative *E. coli* (EAEC) and diffusely adherent *E. coli* (DAEC) with the aim of determining which pathotypes are associated with food sources (producer distributor bulk milk, irrigated lettuce, street vendor coleslaw) and irrigation water and potentially by association the high diarrheal disease burden noticed with the country
- ii.) Determine whether non-DEC strains isolated from food sources and irrigation water in South Africa are associated with phenotypes such as biofilm formation, cytotoxicity, disruption of epithelial cell barriers and induction of pro-inflammatory cytokines *in-vitro* that may be linked to intestinal dysfunction in a bid to understand the risk faced by consumers from non-pathogenic but potentially harmful *E. coli* transmitted through the food chain
- iii.) Develop a multiplex real time qPCR assay for simultaneous detection of cefotaxime resistant (*CTX-M*) extended beta-lactamase producing (ESBL) and virulence (shigatoxin 1, shigatoxin 2 and intimin) genes in *E. coli* isolated from food sources and irrigation water in South Africa with the aim of providing a rapid, accurate and efficient method of characterizing and monitoring emerging foodborne and antibiotic resistant pathogens that pose an immediate but unquantified risk to food safety and public health in South Africa

4. Research Chapter

This section is composed of 3 research chapters written in the format of the Journal to which they were/are intended to be submitted:

- 4.1) Enteroaggregative *Escherichia coli* is the predominant diarrheagenic *E. coli* pathotype among irrigation water and food sources in South Africa. This paper was accepted and published in the International Journal of Food Microbiology: Aijuka, M., Santiago, A. E., Girón, J. A., Nataro, J. P., & Buys, E. M. (2018). International Journal of Food Microbiology Enteroaggregative *Escherichia coli* is the predominant diarrheagenic *E. coli* pathotype among irrigation water and food sources in South Africa. *International Journal of Food Microbiology*, 278 (April), 44–51. <http://doi.org/10.1016/j.ijfoodmicro.2018.04.018>
- 4.2) *Escherichia coli* isolated from food sources and irrigation water: A potential risk for causing intestinal dysfunction? This paper was prepared for submission to the Journal: *Food Research International*
- 4.3) Detection of Extended Beta Lactamase Cefotaxime Resistance and Virulence Genes in *Escherichia coli* by Duplex Quantitative Real Time PCR and Melt Curve Analysis. This paper was prepared for submission to the Journal: *Applied and Environmental Microbiology*

4.1 Enteroaggregative *Escherichia coli* is the predominant diarrheagenic *E. coli* pathotype in irrigation water and food sources in South Africa

Abstract

Diarrheagenic *E. coli* (DEC) has been implicated in foodborne outbreaks worldwide and has been associated with childhood stunting in the absence of diarrhoea. Infection is extraordinarily common, but the routes of transmission have not been determined. Therefore, determining the most prevalent pathotypes in food and environmental sources may help provide better guidance to various stakeholders in ensuring food safety and public health and advancing understanding of the epidemiology of enteric disease. We characterized 205 *E. coli* strains previously isolated from producer distributor bulk milk (PDBM)(118), irrigation water (48), irrigated lettuce (29) and street vendor coleslaw (10) in South Africa. Enteropathogenic *E. coli* (EPEC), enterotoxigenic *E. coli* (ETEC), enteroaggregative *E. coli* (EAEC) and diffusely adherent *E. coli* (DAEC) were sought. We used PCR and partial gene sequencing for all 205 strains while 46 out of 205 that showed poor resolution with PCR in comparison to control strains thereby yielding inconclusive results were subsequently characterized using cell adherence (HeLa cells).

PCR and partial gene sequencing of *aatA* and/or *aaiC* genes confirmed EAEC (2%, 5 out of 205) as the only pathotype. Phylogenetic analysis of sequenced EAEC strains with *E. coli* strains in GenBank showing $\geq 80\%$ nucleotide sequence similarity based on possession of *aaiC* and *aatA* generated distinct clusters of strains separated predominantly based on their source of isolation (food source or human stool) suggesting a potential role of virulence genes in source tracking. EAEC 24%, 11 out of 46 strains (PDBM=15%, irrigation water=7%, irrigated lettuce=2%) was similarly the predominant pathotype followed by strains showing invasiveness to HeLa cells, 4%, 2 out of 46 (PDBM=2%, irrigated lettuce=2%), among stains characterized using cell adherence.

Therefore, EAEC may be the leading cause of DEC associated food and water-borne enteric infection in South Africa. Additionally, solely using molecular based methods targeting virulence gene determinants may underestimate prevalence, especially among heterogeneous pathogens such as EAEC.

Key words: Enteroaggregative *E. coli*, Evolutionary relationship, Virulence genes, Cell adherence assay, Multiplex PCR, Foodborne pathogens

Abbreviations: DEC-diarrheagenic *E. coli*; EAEC-enteroaggregative *E. coli*; EIEC-enteroinvasive *E. coli*; ETEC-enterotoxigenic *E. coli*; EPEC-enteropathogenic *E. coli*; DAEC-diffusely adherent *E. coli*; AIEC-adherently invasive *E. coli*; DMEM-Dulbecco's Modified Eagle Medium; PBS-Phosphate Saline Buffer; MEGA- Molecular Evolutionary Genetics Analysis; NICD- National Institute of Communicable Diseases; TSB-Tryptone Soy Broth; UPEC-Uropathogenic *E. coli*

4.1.1 Introduction

Diarrheagenic *Escherichia coli* (DEC) has long been associated with foodborne illness and outbreaks worldwide, thereby posing a risk to global food safety and public health. *E. coli* pathotypes commonly associated with illness amongst varying age groups and geographical locations include enteropathogenic (EPEC), enterotoxigenic (ETEC), enterohaemorrhagic (EHEC), enteroaggregative (EAEC), diffusely adherent (DAEC) and enteroinvasive (EIEC) *E. coli* (Croxen et al., 2013; Kaper et al., 2004; Nataro and Kaper, 1998). However, recent studies have implicated less characterized strains in disease outbreaks such as EAEC producing shigatoxins (*E. coli* O104:H4 in 2011) that caused a large scale foodborne outbreak throughout Europe in 2011 (Rasko et al., 2011). Additionally, adherent and invasive *E. coli* (AIEC) has been linked to patients with Crohn's disease (Nash et al., 2010). Outbreaks such as that occurring in Germany in 2011 suggest risk of highly virulent pathotypes emerging, possibly driven by factors such as climate change (Carlton et al., 2016). In addition to diarrhoea, EAEC has been associated with growth faltering in children in the absence of diarrhoea (Acosta et al., 2016) and molecular studies have suggested that up to 80% of children may harbour EAEC at a point in time (Platts-Mills et al., 2015).

Variation in epidemiology of diarrhoeal disease and associated pathotypes has been linked to geographical, temporal and climatic conditions complicating food safety and public health prevention and control initiatives especially in resource limited countries (Carlton et al., 2016). South Africa like many developing and Sub-Saharan African countries experiences a high incidence of diarrhoea (Tau et al., 2012) in infants and immune compromised adults, such as HIV-positive patients (Samie et al., 2007). Such illness would disproportionately affect the low income population, which makes up a large part of the urban and rural population (World Bank, 2014) due to inadequate waste disposal and sanitation facilities (Baker et al., 2016). Within South Africa, The National Institute of Communicable Diseases (NICD) initiated a national surveillance program to monitor diarrheagenic pathotypes (Tau et al., 2012).

This initiative coupled with additional site specific studies from clinical specimens (Bisi-johnson et al., 2011; Samie et al., 2007) and food and environmental sources (Adefisoye and Okoh, 2016; Castro-Rosas et al., 2012; Farrokh et al., 2013; Gemmell and Schmidt, 2012; Newell et al., 2010) have helped shed light on the magnitude of the diarrheal disease burden associated with DEC as well as its prevalence within the environment.

Gaps remain in understanding the prevalence of DEC pathotype(s) in food and environmental sources over varying geographical, temporal and sample sources. Such information would help link observed infections/outbreaks with more specific food and environmental sources providing more informed guidance to food safety and public health interventions (Newell et al., 2010). The present study followed from two major previously concluded studies. The first study was initiated by The South African Department of Agriculture and Water Research Commission to help characterize the bacterial quality of South African irrigation water (Aijuka et al., 2015). A high prevalence of faecal indicators within milk sold by unregulated retailers alarmed the South African Dairy Standard Agency which subsequently initiated the second study to characterize its microbiological quality (Ntuli et al., 2017, 2016).

We sought to determine the most prevalent DEC pathotypes associated with *E. coli* previously isolated from food sources and irrigation water collected in South Africa over varying geographical and temporal spans. Additionally, we sought to determine food sources associated with highest prevalence of DEC.

4.1.2 Materials and Methods

4.1.2.1. *Escherichia coli* isolate source

A total of 205 *E. coli* strains previously isolated from irrigation water, irrigated lettuce, producer distributor bulk milk (PDBM) and street vendor coleslaw were used in this study. This bacterial collection included 48 isolates from irrigation water (Aijuka et al., 2015; Aijuka, 2014), 29 from irrigated lettuce (Aijuka, 2014), a total of 118 from milk (pasteurized and unpasteurized PDBM) sold by small holder sellers (producer-distributor) (Ntuli et al., 2016) and 10 from street-vendor coleslaw purchased in Pretoria City, South Africa. All isolates collected in these studies were stored at -80°C at the Department of Food Science, University of Pretoria, South Africa in Tryptone Soy Broth (TSB)(Biolab Diagnostics (Pty) Ltd, Midrand, South Africa) containing 30% glycerol (Sigma-Aldrich, St. Louis, MO, USA).

4.1.2.2. Isolate resuscitation and transportation

The isolates were regrown in TSB, incubated at 37°C overnight and transferred onto freshly prepared TSB agar slants in McCartney bottles. Slants were incubated at 37°C overnight and couriered to The Child Health Research Center, Department of Pediatrics, University of Virginia School of Medicine, Charlottesville, Virginia USA where all subsequent analyses were done. All isolates were resuscitated in Luria Broth Base, Miller's modification (LB;

AmericanBio Inc, Natick, MA, USA) and subsequently grown on LB plates with respective overnight incubation at 37°C prior to any analysis.

4.1.2.3. Haemolysin production in *E. coli* isolates

All 205 isolates were grown on 5% blood agar (Hardy Diagnostics, Santa Maria, CA, USA), incubated at 37°C and checked for beta or alpha haemolysis as an indicator of the presence of hemolysins (Greene et al., 2015).

4.1.2.4. Molecular characterization of EPEC, ETEC, EAEC and DAEC

ETEC, EPEC and EAEC pathotypes were identified using a multiplex polymerase chain reaction (PCR) (Panchalingam et al., 2012). PCR targets (Table 4.1.1) included ETEC heat-labile (*LT*) and heat-stable (*STh*) enterotoxin genes, the EPEC intimin (*eae* gene) outer membrane protein adhesin and *bfpA*, the gene encoding the bundle forming pili (BFP); the EAEC plasmid-encoded gene *aatA*; and the EAEC chromosomally encoded *aaiC* locus. The gene targets are known virulence determinants of their respective pathogens (Nataro and Kaper, 1998). Strains positive for *eae* but not BFP were designated as atypical EPEC. Strains positive for either ETEC enterotoxins were considered ETEC and strains positive for either EAEC factor were considered EAEC. A monoplex PCR reaction targeting an accessory gene (*daaC*) of a major fimbrial sub-unit, F1845, was used for identifying DAEC (Campos et al., 1999).

Template DNA was prepared by mixing a loop-full of an overnight culture with 500µL of BP 2819 Water (Fisher Scientific, Waltham, MA, USA) in a 2 mL eppendorf tube (Eppendorf AG, Hamburg, Germany) and heating it to boiling for 20 minutes. The mixture was rapidly cooled on ice and centrifuged at 5000 rpm for 2 minutes. The supernatant was collected and stored at -20°C and used as template DNA in all subsequent tests.

For the PCR reaction, 3 µL of template DNA was added to 10 µL of Quick-Load® *Taq* 2X Master Mix containing 20 mM Tris-HCl, 100 mM KCl, 3.0 mM MgCl₂, 0.2mM deoxynucleotide triphosphates (dNTPs) and 50 µg/mL Hot Start *Taq* DNA polymerase (New England Biolabs). Each primer pair (0.4 µL of 20 pmol/µL) was added together with 3µL of RNase-free water to a final volume of 20 µL. PCR was performed under the following conditions: preheating at 96°C for 4 min, denaturation at 95°C for 20 secs, annealing at 57°C for 20 secs, elongation at 72°C for 1 min. PCR was performed for 35 cycles with final extension at 72°C for 7 min in an Eppendorf Mastercycler Gradient thermal cycler (Eppendorf AG).

The amplification products were separated through a 2% agarose gel and visualized by ultraviolet light trans-illumination after ethidium bromide staining. The 1-kb plus A 100-bp DNA ladder (New England BioLabs, Ipswich, MA, USA) was used as a molecular size marker in gel. Control strains employed in every PCR reaction were ETEC H10407, EAEC 042 and for EPEC strains CVD 28 (*eae* positive) and HB101 (pMAR7) (*bfpA*-positive).

For DAEC, PCR reagents and consumables were as described in the multiplex assay above. Thermo-cycling conditions included the following: Preheating at 95°C for 3 min, denaturation at 95°C for 30 seconds, annealing at 60°C for 30 seconds and elongation at 68°C for 2 minutes. PCR was performed for 35 cycles with a final extension at 68°C for 10 minutes. DAEC strain C1845 was used as a positive control.

4.1.2.5 Partial gene sequencing of PCR positive DEC isolates

PCR amplification products were separated in a 2% agarose gel, visualized by ultraviolet light trans-illumination after ethidium bromide staining, excised, and purified using a QIAquick® PCR Purification Kit (Qiagen Inc, Germantown, MD, USA) per the manufacturer's instructions. The purified PCR product was mixed with a single primer of the respective gene target (1 µL), 10 µL of DNA template and 4 µL of RNase free water to make a total of 15 µL in a PCR grade tube. The samples were delivered for final analytical confirmation to the DNA Sequencing Service (GeneScript USA Inc.) at The University of Virginia.

Table 4.1.1 Primer sequences and the expected amplicon sizes for the multiplex polymerase chain reaction employed in the detection of Diarrheagenic *Escherichia coli*.

Pathogen	Primer	Target Gene	Primer Sequence (5'-3')	Amplicon (bp)
	LT-F	<i>Let</i>	CACACGGAGCTCCTCAGTC	508
	LT-R		CCCCCAGCCTAGCTTAGTTT	
	ST-F	<i>Est</i>	GCTAAACCAGTAG/AGGTCTTCAAAA	147
	ST-R		CCCGGTACAG/AGCAGGATTACAACA	
EPEC	BFPA-F	<i>bfpA</i>	GGAAGTCAAATTCATGGGGG	367
	BFPA-R		GGAATCAGACGCAGACTGGT	
	EAE-F	<i>Eae</i>	CCCGAATTCGGCACAAGCATAAGC	881
	EAE-R		CCCGGATCCGTCTCGCCAGTATTCG	
EAEC	CVD432F	<i>aatA</i>	CTGGCGAAAGACTGTATCAT	630
	CVD432R		CAATGTATAGAAATCCGCTGTT	
	AAIC F	<i>aaiC</i>	ATTGTCCTCAGGCATTTAC	215
	AAIC R		ACGACACCCCTGATAAACA	
DAEC	DAAC F	<i>daaC</i>	ATTACGTCATCCGGGAAGCACACA	146
	DAAC R		TTGTCTGCGGCTTTATGAGCAAGC	

Abbreviations: EAEC, enteroaggregative *Escherichia coli*; EPEC, enteropathogenic *E. coli*; ETEC, enterotoxigenic *E. coli*; EAEC, EPEC and ETEC were identified using a multiplex PCR and primer pairs according to (Panchalingam et al., 2012) and DAEC using a monoplex PCR and primer pairs according to (Campos et al., 1999)

4.1.2.5. Cell adherence assay

As HEp-2 or HeLa cell adherence remains the gold standard for discrimination of DEC pathotypes showing localized (EPEC), aggregative (EAEC) and diffuse (DAEC) adherence (Croxen et al., 2013; Kaper et al., 2004; Nataro and Kaper, 1998), 46 out of the total 205 strains were characterized. The Hep-2 cell adherence assay as described by (Nataro et al., 1987) for differentiating patterns of DEC was used with some modifications. HeLa cells instead of Hep-2 cells were used to determine the adherence patterns of 46 out of the 205 *E. coli* strains that showed poor resolution with PCR using the *daaC* gene (DAEC). However, all previously characterized EAEC strains based on PCR and partial gene sequencing of virulence gene determinants were excluded from this analysis. Selection of these strains was based on poor resolution of presumptive DAEC identified using PCR. HeLa cells at 80% confluence were aseptically transferred into 24-well plates (Fisher Scientific) containing 12 mm cover slips (Fisher Scientific) in each well and washed with Phosphate Buffer Saline (PBS) (Fisher Scientific) and 1 mL of Dulbecco's Modified Eagle's Medium (DMEM) (Sigma-Aldrich) +0.5% D-mannose (Sigma-Aldrich). The bacterial isolates were grown in 3 mL of LB broth in 13 mL plastic tubes while shaking at 37°C. Ten millilitres of the overnight culture was added to each well using standard strains EPEC E2348/69, EAEC 042 and DAEC 1845 as controls of localized, aggregative and diffuse adherence, respectively. Incubation was at 37°C in a CO₂ incubator for 3 and 6 hours (on replenishing with fresh medium after 3 h) for each isolate. The cells were washed gently (3X) with Peptone Buffered Saline (PBS) (Sigma-Aldrich) and 500 µL of 2% formalin were added to fix the samples for 20 min at room temperature. The samples were rinsed (3X) with distilled water (dH₂O) and stained with 500 µL of a Giemsa solution (Sigma-Aldrich) for 20 minutes. The samples were rinsed (3X) with dH₂O until the colour disappeared. The coverslips were removed from the 24-well plates, airdried and mounted with a tiny drop of Cytoseal (Fisher Scientific) mounting glue onto a glass slide. The samples were observed under a Zeiss light microscope and images were recorded at 60 X.

4.1.2.6. Statistical analysis

Means for prevalence were calculated for each identified DEC pathotype based on the method of analysis (molecular or cell adherence) for all strains isolated from food sources and irrigation water. We used the Basic Local Alignment Search Tool (BLAST) within National Center for Biotechnology Information (NCBI) to search for nucleotide sequences of *E. coli* strains deposited in GenBank, with $\geq 80\%$ similarity to *aaiC* or *aatA* sequences found in EAEC strains

in this study. Thereafter, the evolutionary relationship of each strain was inferred using the Neighbour-Joining method and computed with Molecular Evolutionary Genetics Analysis for Bigger Data Sets version 7.0 (MEGA 7) (Kumar et al., 2016).

4.1.3. Results

4.1.3.1. Haemolysin production

Of the 205 isolates tested on blood agar, only 1 strain from irrigation water (CR4) showed beta-haemolysis and none showed alpha-haemolysis. As haemolysis is correlated with the presence of uropathogenic *E. coli*, we concluded that these strains were not common in our samples.

4.1.3.2. PCR and partial virulence gene sequencing

Among the different DEC virulence gene determinants sought in our collection of 205 *E. coli* strains, only those specific for EAEC and DAEC were found by PCR. The remaining 196 isolates were negative for virulence gene determinants of the DEC pathotypes screened for in this study using PCR. Based on PCR (3.9%, 8 out of 205) and (0.5%, 1 out of 205) were EAEC and DAEC respectively (Table 4.1.2). EAEC was predominant among strains from PDBM (5.1%, 6 out of 118) and to a lesser extent irrigation water (4.2%, 2 out of 48) of which a single isolate (K2) carried the *aaiC* gene and 5 isolates (M1, 57, L7, N25 and 79) carried the *aatA* gene. The single strain (CR4) of DAEC was from irrigation water. Partial gene sequencing of individual virulence gene determinants associated with each pathotype among PCR positive strains (excluding strain 79) confirmed only EAEC.

Gel electrophoresis bands showing strain 79 after PCR consistently showed a faint signal compared to the other sequenced strains (data not shown) and hence was left out of the sequencing analysis. However, numbers of EAEC were lower compared to those observed based solely on characterization with PCR (Table 4.1.2) suggesting an initial over-estimation of EAEC prevalence. Based on partial gene sequencing, EAEC was found in 2.4%, 5 out of 205 strains. EAEC strains confirmed with partial gene sequencing included K2 positive for *aaiC* (Table 3). On the other hand, strains 57, L7, MPU(W)5(1) and MPU(W)8(4) were positive for *aatA* (Table 4.1.3). EAEC was confirmed in PDBM (2.5%, 3 out of 118) and irrigation water (4.2%, 2 out of 48).

4.1.3.3. Evolutionary relationship of EAEC strains isolated from food sources and irrigation in South Africa with genetically related *E. coli* strains in GenBank based on partial gene sequencing of *aaiC* and *aatA* genes

We used the evolutionary relationship of EAEC strains isolated from this study to infer relatedness to *E. coli* strains previously isolated from different sources and deposited within GenBank as a way of determining potential routes for DEC contamination. The nucleotide sequence of the *aaiC* gene in strain K2 from PDBM showed 80% identity to 15 *E. coli* strains in GenBank. The dendrogram constructed to infer the evolutionary relationship of strain K2 with these 15 strains generated 4 distinct clusters; G1, G2, G3 and G4. The strains predominantly clustered based on location of *aaiC* within the bacterium (chromosome or plasmid) well as on the source of strain isolation (food sources or human stool) (Fig. 4.1.1). G1 predominantly consisted of strains of *E. coli* serovar O104:H4 among which were outbreak strains 227-11 and 2011C-3493 isolated from patients in Germany and The United States respectively. The closest strain related to K2 was also isolated from a food source, *E. coli* strain 06-0048 (Accession number: CP012498.1) isolated from alfalfa sprouts in California USA in 2006. Strains isolated from humans (G1 and G4) were isolated from bloody and non-bloody stool in patients from Denmark, France, Georgia, Poland and the USA suggesting a link of the *aaiC* gene with diarrheal causing *E. coli*.

The nucleotide sequence of the *aatA* gene in strain 57 from milk showed between 90 to 93% identity to 22 *E. coli* strains in GenBank (Table 4.1.3). However, nucleotide sequences from strains L7 isolated from milk as well as MPUW51 and MPU84 both isolated from irrigation water did not have comparisons in GenBank. Nonetheless, we used the 22 strains showing similarity to the single strain 57 for drawing evolutionary relationships with all 4 strains. The dendrogram inferring evolutionary relationships generated 3 clusters G5, G6 and G7 (Fig. 4.1.2).

G5 comprised 3 of the 4 strains (PDBM=2; irrigation water=1) used for the analysis. In spite of clustering together, strains in G5 showed low evolutionary relatedness (Fig. 4.1.2). G6 contained 1 isolate from irrigation water (MPUW51) that showed closest relatedness to *E. coli* strains positive for the *Dr* family of adhesins. Additionally, most isolates in G5 and G7 were isolated from humans. G7 consisted of strains predominantly within the clonal group, sequence type ST131. Strains from G6 and G7 were predominantly isolated from clinical sources. Similarly, as previously noted with strains clustered based on the *aaiC* gene, strains

from this study predominantly clustered based on source of isolation (environmental or clinical). Strains isolated from humans (G6 and G7) were predominantly associated with extraintestinal infections such as urinary tract infections in patients from China, France, Poland, Spain and the UK suggesting a link of *aatA* with extraintestinal pathogenic *E. coli*.

4.1.3.4 Adherence tests

The strains predominantly exhibited the characteristic ‘stacked-brick’ pattern or aggregative adherence (AA) of EAEC and to a less extent invasiveness typical of EIEC on HeLa cells (Table 4.1.2). No strain showed the diffuse adherence phenotype typical of DAEC suggesting false positives with PCR that led to the poor resolution previously reported. Based on the adherence and invasive phenotypes, EAEC and EIEC were found in (24%, 11 out of 46) and (4%, 2 out of 46) of strains respectively. EAEC strains exhibited strong to moderate and weak AA capacity (Fig. 3a, 3b and 3c). Prevalence of EAEC in PDBM, irrigation water and irrigated lettuce was (15%, 7 out of 46), (7%, 3 out of 46) and (2%, 1 out of 46) respectively (Table 2). Prevalence of EIEC was 2%, 1 out of 46 in both PDBM and irrigated lettuce.

Table 4.1.2 *Escherichia coli* strains isolated from food sources and irrigation water in South Africa positive for virulence gene determinants (n=205) and cell adherence patterns (n=46) associated with diarrheagenic *E. coli*. Strains tested for cell adherence had shown inconclusive results using PCR when compared with control strains.

Sample	Method of characterization							
	Polymerase chain reaction			Partial gene sequencing		HeLa cell adherence pattern		
Total* number of isolates	<i>aaiC</i> (EAEC)	<i>aatA</i> (EAEC)	<i>daaC</i> (DAEC)	<i>aaiC</i> (EAEC)	<i>aatA</i> (EAEC)	Aggregative adherence (EAEC)	Cell invasiveness (EIEC)	
PDBM	118	1(1)	4(5)	-	1(1)	2(2)	15(7)	2(1)
Irrigation	48	-	4(2)	2(1)	-	4(2)	7(3)	-
Irrigated	29	-	-	-	-	-	2(1)	2(1)
Coleslaw	10	-	-	-	-	-	-	-
Total	205	1	7	1	1	4	24(11)	2(2)

(*) not a percentage; Percentages in table reported to the first significant whole number. Number of isolates in parentheses; - not detected/no adherence pattern; *aaiC*- Chromosomal virulence gene determinant in EAEC; *aatA*- Plasmid virulence gene determinant in EAEC; EAEC-enteroaggregative *E. coli*; *daaC*- a major fimbrial sub-unit in standard DAEC strain C1845; DAEC-diffusely adherent *E. coli*; EAEC-enteroaggregative *Escherichia coli*; EIEC-enteroinvasive *E. coli*. All strains tested for cell adherence had previously shown poor resolution using PCR compared to control strains.

Table 4.1.3 Nucleotide sequences of virulence gene determinants associated with enteroaggregative *Escherichia coli* (EAEC) in *Escherichia coli* strains isolated from food sources and irrigation water in South Africa.

Isolate code	Isolate	Sequenced gene product (<i>aaiC</i> or <i>aatA</i>)
K2	PDBM	CTTGGACGGCCTGTTTTTGATTTGGTTATATTAAGTTAAAGGTCACTAAGGAAATACAACCTTATTA GAGCATTTTTTTAAAAGAGGTGAGAAAGATATTGTAATTGAAATTGTTTCGCAAGGAATCTACTAAAT CAGGTAGTGCATATTCATCCTTTAAGGTTGTTTATCAGGGGTGTCGTAA
57	PDBM	TAAACCGCCTGGGTCAGCGCATCGTGGGTGGGGAACACCGGGCTGAACACGGGGCCACCGGACTGG CAGACCGTGGAATAAGGCATCGGGGGTGAACGTTGTCTGCGGCTTTATGAGCAAGCAGCCCTT
L7	PDBM	TAGCTGTAGCGCATCGTGAGGGCGA ACCCGGCTGACCCGGGCCACCGGAC TGCGTACGTGTAAAACAGACAGCGT GCGCT GAGCGTTGTCTGCGCTTTATGAGCA GCTCGCTCTTTTAGAGGG
MPU(W)5(1)	Irrigation water	ATTACTCTTTTCCATTTGCATCCCCT TGGCTAAACCAGCTTACTCTATCACC GGACTGGCCGATAGTTTATAATCAA TCC GTCTGGTTTTAACAGAGTTATGAGA ATCAGTTCCATTTTATTTGGCCTTGG TTTCTCTTGATTTCTGGTGAAA
MPU(W)8(4)	Irrigation water	ACTGGCGTCCCGCCGCATCGTGTTTG GGCGAACACCGGGCACAGCCGGCGC CACCGGATGGGGACCGGGAAATTAC GAAT GGGCGTGAGCGTTGGCGGCGGCTTT GGAGAGGCAGCCCCTTTTGAATAAAA G

PDBM-Producer Distributor Bulk Milk; K2-positive for *aaiC*; 57, L7, MPU(W)5(1) and MPU(W)8(4)-positive for *aatA*

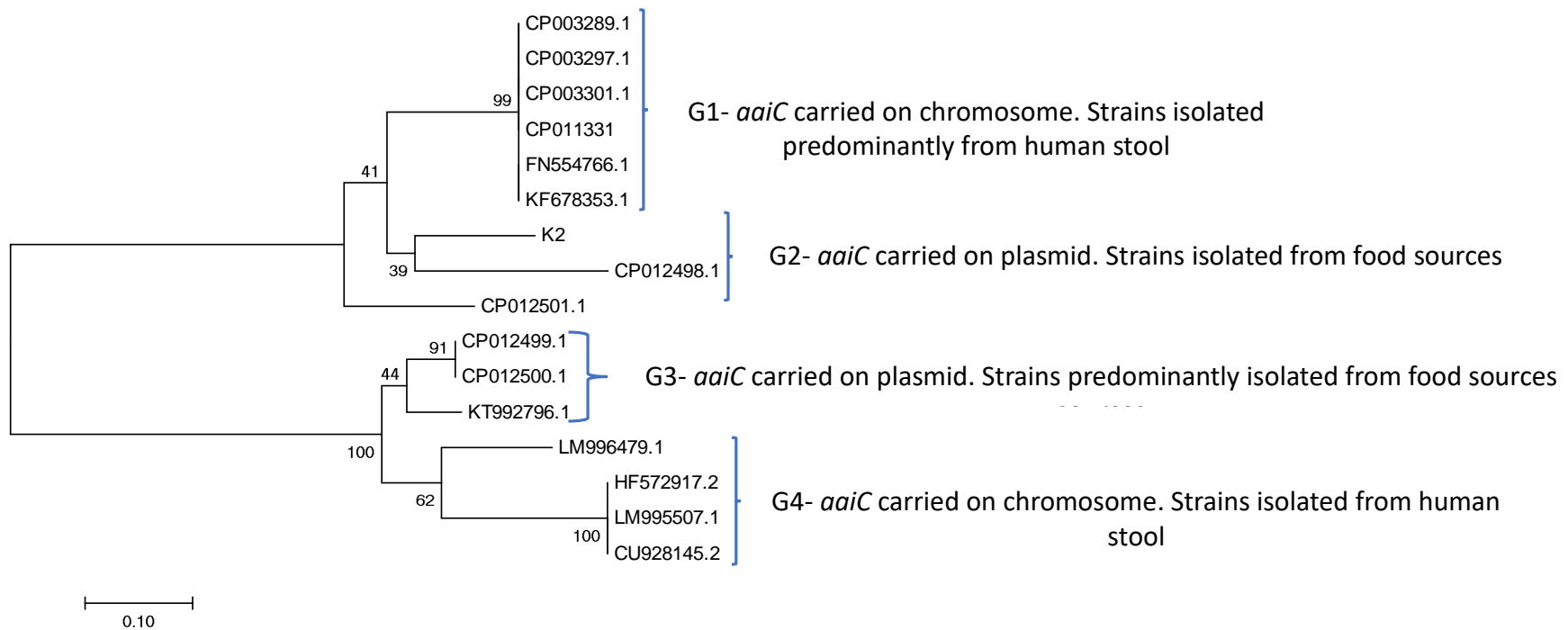


Figure 4.1.1. Evolutionary relationships of *Escherichia coli* strains in GenBank showing $\geq 80\%$ gene nucleotide sequence similarity to strain K2 isolated from producer distributor bulk milk in South African milk based on *aaiC*, the enteroaggregative *E. coli* (EAEC) virulence gene determinant. The evolutionary history was inferred using the Neighbour-Joining Method. Codes of strains used for comparison represent accession numbers from GenBank. G1, G2, G3 and G4 represent defined clusters of strains showing differences in genetic location of *aaiC* (plasmid or chromosome) in each strain as well as source of strain isolation (foodborne or human faeces).

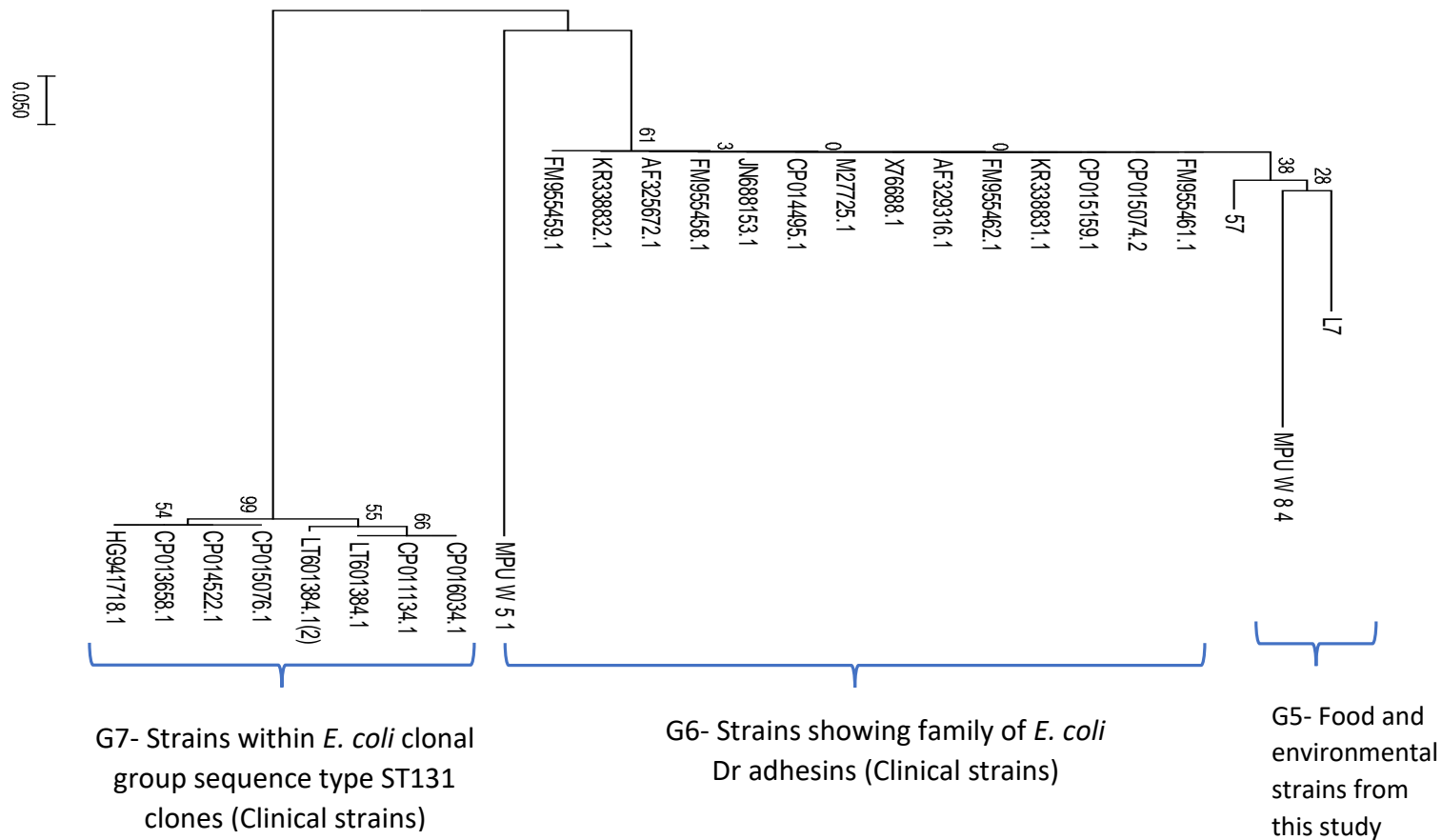


Figure 4.1.2. Evolutionary relationships of *Escherichia coli* in GenBank showing $\geq 90\%$ gene nucleotide sequence similarity to *E. coli* strains isolated from producer distributor bulk milk (PDBM) and irrigation water in South Africa based on *aatA*, the enteroaggregative *E. coli* (EAEC) virulence gene determinant. The evolutionary history was inferred using the Neighbour-Joining Method. Strain sources: PDBM-L7, 57; Irrigation water-MPUW51, MPUW84. Codes of strains used for comparison represent accession numbers in GenBank

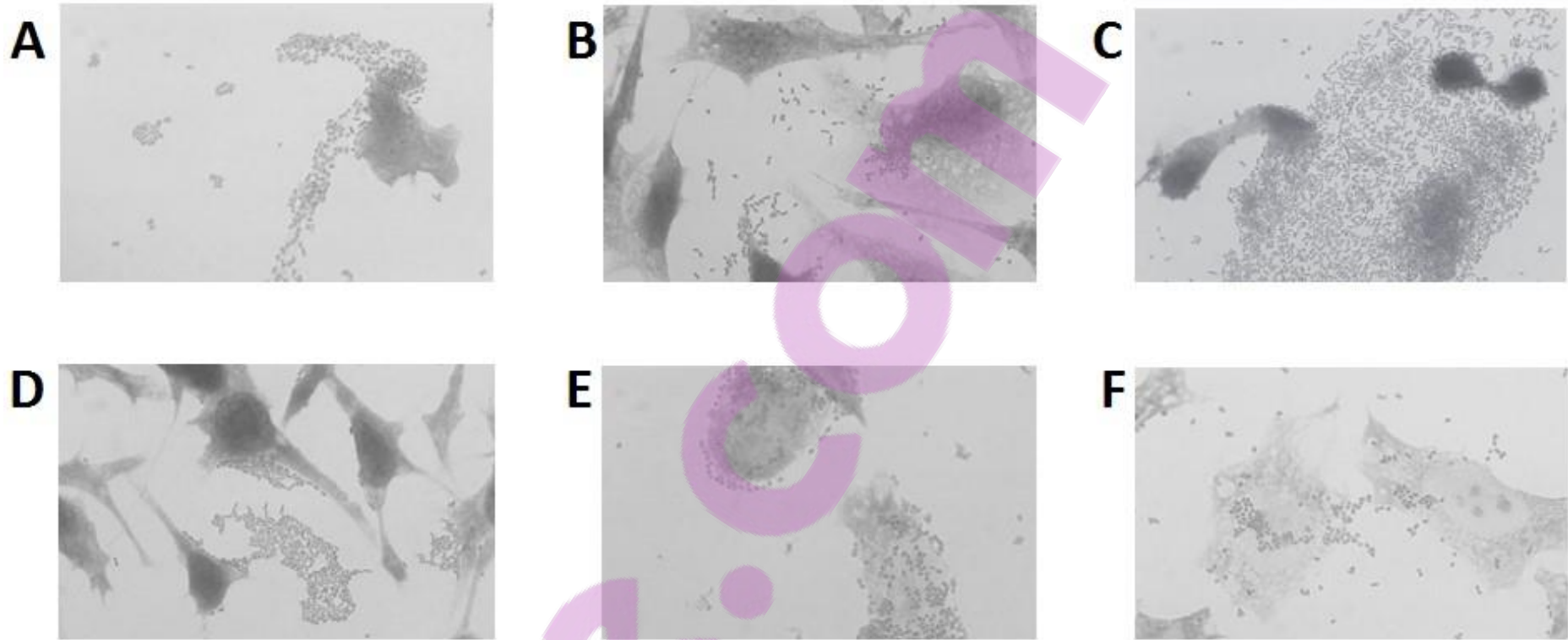


Figure 4.1.3a. Aggregative adherence (AA) pattern 'stacked-brick' observed in *Escherichia coli* strains isolated from food sources and irrigation water in South Africa. Strains were grown on HeLa cells and showed strong to moderate AA characteristic of enteroaggregative *Escherichia coli* (EAEC). Strain code: Standard EAEC strain 042; B-H8; C-M28; D-NW(V)7(3); E-NW(V)10(1); F-NW(W)9(3). Source of isolation: A-Clinical strain; B and C-Producer distributor bulk milk; D and E-Irrigated lettuce; F-Irrigation water. Images were taken to 100X with a Zeiss microscope. Resolution=20 μm .

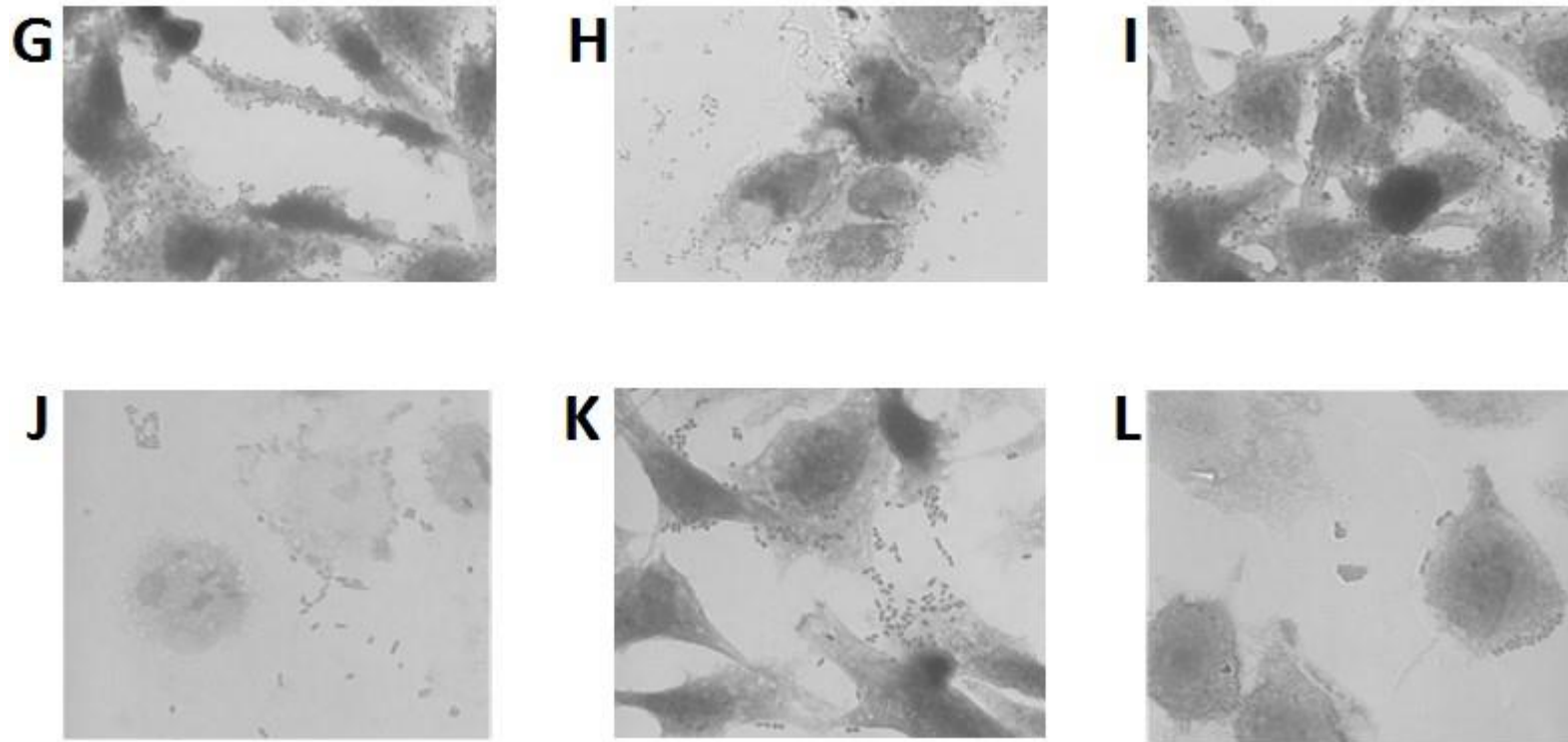


Figure 4.1.3b. Weak aggregative adherence (AA) observed in *Escherichia coli* strains isolated from food sources and irrigation water in South Africa. Strains were grown on HeLa cells and showed weak adherence in comparison to strains in Fig. 4a. Strain codes: G- LeK1; H-M12; I-K5; J-K16; K-M6; L-N23. Strain sources: G-Irrigated lettuce; H, I, J, K and L-Producer distributor bulk milk. Images were taken to 100X with a Zeiss microscope. Resolution= 20 μm .

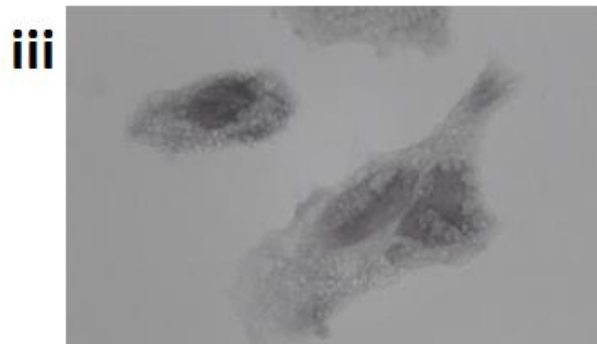
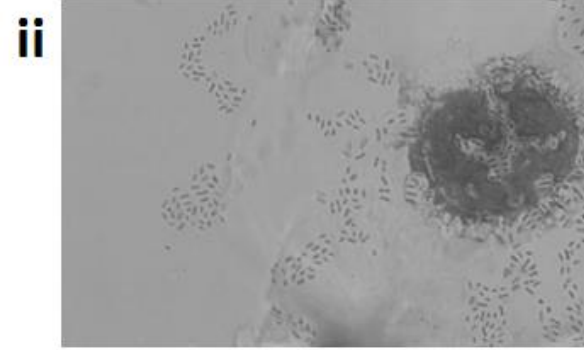
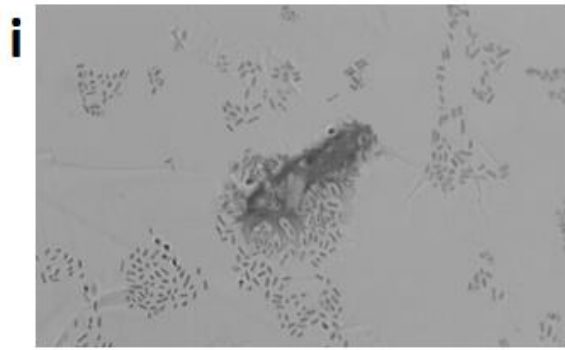


Figure 4.1.3c. Invasiveness observed in *Escherichia coli* strains isolated from food sources and irrigation water in South Africa. Invasiveness is characteristic of enteroinvasive *Escherichia coli* (EIEC) and therefore strains were identified as presumptive EIEC since no other subsequent confirmatory test was done. Strains were grown on HeLa cells. Strain codes: i and ii-LeK; iii and iv- 37. Strain sources: LeK-Irrigated lettuce; 37-Producer distributor bulk milk. Images were taken at 60X with a Zeiss microscope. Resolution=20 μm .

4.1.4. Discussion

Determining the most prevalent DEC pathotypes associated with diarrheal disease within a given locale can help provide better guidance to food safety and public health interventions. Predominance of EAEC in food sources and irrigation water in South Africa suggests it is the most prevalent DEC pathotype in these sources. These sources subsequently provide routes for EAEC infection within the general population. Therefore, EAEC may be the leading cause of food and water-borne diarrheal disease caused by pathogenic *E. coli* in South Africa. EAEC is an emerging food-borne pathogen largely affecting infants, immune compromised adults and travellers to developing countries (Croxen et al., 2013; Estrada-Garcia and Navarro-Garcia, 2012). Additionally, EAEC has been associated with inflammation and malnutrition among infants in developing countries (Acosta et al., 2016) thereby heightening the risk of a similar scenario playing out locally (South Africa) especially within the low income urban and rural population who may lack adequate food and water safety systems. Previous studies in South Africa have also noted high prevalence of EAEC in food and environmental (Caine et al., 2014) as well as clinical sources (Adefisoye and Okoh, 2016; Bisi-johnson et al., 2011; Samie et al., 2007; Tau et al., 2012).

Aaic forms part of a group of genes localized to a 117kb pathogenicity island inserted at *pheU* in typical EAEC having homologues in other Gram-negative bacteria and proposed to constitute a Type VI secretion system (T6S) (Dudley et al., 2006b). Bacteria have evolved different regulatory components for the expression of these genes which may be acquired by horizontal transfer for specific adaptation to varying hosts and niches such as marine life, plants or animals (Boyer et al., 2009) an attribute that was noticed among strains from our work (Fig 4.1.1). The regulatory mechanism governing type VI gene expression varies widely from species to species and even from strain to strain within a given species given its wide distribution among different bacterial groups (Miyata et al., 2013).

Additionally, the aggregative adherence trait in EAEC strains is a bacterial outcome of co-evolution with human hosts and has been acquired by different *E. coli* lineages some of which may share similar but non-identical mobile elements and so no single strain can be considered representative of the EAEC group (Okeke et al., 2010). The *aatA* gene in EAEC forms part of a plasmid encoded locus also under the control of *aggR* coding for an ABC transporter complex that channels the virulence factor dispersin out of the bacterial cell (Nishi et al. 2003).

On the other hand the Dr- family of adhesins consists of clones such as *Dr*, *Afa-I*, *Afa-III* and *F1845* commonly associated with Uropathogenic *E. coli* (UPEC) but also found in DEC. Incidentally *agg* and *aaf* adhesins of standard EAEC strains 17-2 and 042 respectively are similar to *afaD-III* an afimbrial adhesin in UPEC and DEC strains (Garcia et al., 2000, 1996). While the specific role of the *aatA* in EAEC and Dr family of adhesins in pathogenic *E. coli* may differ, both are involved in colonization of the host environments and thereby persistence.

Predominance of EAEC within liquid environments (milk and irrigation water) compared to solids (irrigated lettuce and coleslaw) suggests provision of more favourable conditions for persistence, thereby presenting higher risk as potential routes for foodborne infection. Additionally, EAEC form biofilms on abiotic surfaces (Estrada-Garcia and Navarro-Garcia, 2012), an adaptation that may facilitate persistence within a secondary environment providing protection against external stresses as well as capturing nutrients. This adaptation poses a risk factor for contamination of food contact surfaces.

The low prevalence of potentially invasive *E. coli* (presumptive EIEC) and absence of other DEC pathotypes suggests that food sources and irrigation water in South Africa may not be major reservoirs/carriers of these pathotypes. Additionally, these absent pathotypes may not serve as causes of DEC associated illness in South Africa. Therefore, compared to other DEC pathotypes, EAEC strains may be more suitably adapted for longer persistence within the open environment as has previously been noted with outbreak strains of enteroaggregative haemorrhagic *E. coli* (EAHEC) serotype O104:H4 in Europe. Prevalence of DEC within this study may have been underestimated based on characterization of strains using molecular tools (PCR and sequencing). This is because subsequent characterization of poorly resolved strains (PCR) with cell adherence assays identified more DEC pathotypes (EAEC and invasive *E. coli* (presumptive EIEC)). This observation highlights shortcomings associated with molecular diagnostic tools targeting virulence gene determinants during monitoring and surveillance (due to the impracticality of tissue culture assays) studies especially when targeting pathogens with great heterogeneity such as the EAEC group. Therefore, studies investigating the prevalence of DEC within food and environmental sources maybe underestimating the levels and subsequently the risk posed to food safety and public health.

However, on a brighter note, partial gene sequencing of virulence gene determinants followed by inference of evolutionary relationships seems to provide an adequate and cost-effective means of source tracking foodborne, environmental and clinical DEC. We show that using a

single sequenced virulence gene coupled with its comparison to strains of the same species within GenBank having the same gene with high nucleotide sequence similarity ($\geq 80\%$) provides a quick and accurate tracker for contamination sources. This approach unlike sophisticated and expensive techniques such as Pulse Field Gel Electrophoresis (PFGE) and Whole Genome Sequencing (WGS) may be more applicable within resource limited areas.

Lastly, the large diarrheal disease burden noticed in South Africa especially among infants (Chola et al., 2010) and immune compromised adults such as HIV-positive individuals (Moshabela et al., 2012) may be associated with EAEC being transmitted through contaminated food and water. Therefore, routine screening for this pathogen in food and environmental sources especially among low income communities may help monitor and control risks of potential outbreaks and long-term health and nutritional effects emanating from recurrent infections.

4.2 *Escherichia coli* isolated from food sources and irrigation water: A potential risk for causing intestinal dysfunction?

Abstract

We have previously shown that diarrheagenic *Escherichia coli* (DEC) and non-DEC are prevalent in food sources and irrigation water in South Africa. Recent data suggest that an increased relative abundance of faecal *Enterobacteriaceae* is associated with poorer health outcomes among children in developing countries. Thus, exposure to non-DEC from environmental sources may incur adverse effects, although the mechanisms underlying these effects remain obscure. To further elucidate this phenomenon, we assayed non-DEC strains from environmental sources in South Africa for phenotypes that may be associated with intestinal dysfunction (ID). DEC strains were also used. The strains had previously been isolated from Producer Distributor Bulk Milk (PDBM), irrigated lettuce, street vendor coleslaw and irrigation water.

In-vitro assays identified; biofilm formation (n=38), extracellular polymeric substance (EPS) formation (n=38), cytotoxic activity (n=10), disruption of tight junctions and induction of Interleukin 8 (IL-8) on polarized T-84 cells (n=20). The number of strains tested for each assay differed, depending on prior molecular and phenotypic characterization that signalled potential pathogenicity. Subsequently, all strains having data points for all analyses were used to compute Principal Component Analysis (PCA) plot curves to infer associations amongst test strains and a standard DEC pathogenic strain (042).

Biofilm formation on glass cover slips after strains were grown in nutrient rich media (LB and DMEM-F12+0.5% D-Mannose) at 37°C varied based on pathotype (DEC and non-DEC) and source of isolation (food, irrigation water, clinical) suggesting that pathotype and source isolation influence persistence within a defined environmental niche. Additionally, DEC isolated from irrigated lettuce had a significantly higher ($p \leq 0.05$) propensity for biofilm formation in both media compared to all strains including DEC standard controls. This suggested the propensity for irrigated lettuce as a potential source of persistent pathogenic strains. Furthermore, all strains were able to form EPS suggesting the ability to form mature biofilms under conditions relevant for food processing (20 to 25°C). Of the (60%, 6 out of 10) strains that showed cytotoxic activity, most (83%, 5 out of 6 strains) were non-DEC isolated from food sources many of which are consumed with minimal processing.

Mean percentage reduction in initial TEER (a measure of intestinal disruption), did not significantly differ ($p=0.05$) in all test strains from that observed in the standard DEC. Additionally, IL-8 induction from strains isolated from PDBM (139pg/mL), irrigation water (231.93pg/mL) and irrigated lettuce (152.98pg/mL) was significantly higher ($p\leq 0.05$) than in the commensal strain aafa. PCA categorized strains based on sources of isolation showed potential for use in source tracking especially when comparing many strains from various environmental sources. We show that non-DEC strains along the food chain possess characteristics that may lead to ID. Further investigations using a larger collection of strains may provide a clearer link to these reported observations that could be associated with the high diarrheal disease burden within the country especially among infants.

Key words: Intestinal Dysfunction; Diarrheagenic *E. coli*, Foodborne pathogens, epithelial cell integrity, Interleukin-8

4.2.1 Introduction

A healthy gut has been associated with increased prevalence of bacteria from the phyla Bacteroides and Firmicutes and a decrease in those from the group Proteobacteria which includes the family *Enterobacteriaceae*. The family is home to a variety of intestinal foodborne pathogens such as Diarrheagenic *Escherichia coli* (DEC)(Guinanane and Cotter, 2013). Intestinal dysfunction (ID) is associated with a change in the “normal” relative abundance of bacterial taxa in the gastrointestinal tract (GIT) and their associated metabolic functions within the human intestinal (Shin et al., 2015). The conditions ultimately lead to a breakdown in homeostasis and as well as the normal synergistic relationship between the human body and its microbial constituents (Shin et al., 2015). ID is linked to reduced mucosal immunity and a subsequent increased risk of intestinal diseases causing reduced dietary nutrient intake and utilization in young infants. This results into growth faltering and neurocognitive disorders later in life (Faith et al., 2015). For example, malnourished infants from Bangladesh and Malawi with Kwashiorkor showed a high rate of inflammation that was associated with a bloom of *Enterobacteriaceae* as well as a high DEC burden upon faecal analysis (Kau et al., 2015; Subramanian et al., 2014). These observations are particularly disturbing especially within developing countries where multiple factors such as inadequate dietary practices and poor environmental sanitation may increase the risk of contamination with *Enterobacteriaceae* and a subsequent bloom within the gut. Therefore, the long term need for a healthy intestinal gut community is paramount to healthy living right from the early stages of life into adulthood. After the first 24 months, the microbial composition of the gut microflora stabilises and the relative abundance of the various component bacterial species are determined by the microbiology of the prevailing human and environmental surroundings (Faith et al., 2015).

In South Africa, the large diarrheal disease burden associated with infants and immune compromised adults such as HIV patients has been linked to a high entero-pathogen prevalence in the GIT (Moshabela et al., 2012; Tau et al., 2012), a condition that may be exacerbated by ID. Additionally, recent studies have shown that DEC is a commonly isolated pathogen in food sources and irrigation water in South Africa (Aijuka et al., 2018, 2015, Ntuli et al., 2017, 2016) suggesting a quantifiable food safety risk within the population (Ntuli et al., 2018). However, while extremely valuable, these studies have failed to pick up strains that do not exhibit the characteristics (genotypic and phenotypic) commonly associated with DEC because of the heterogeneous nature of these pathotypes.

This is because environmental, food safety and health practitioners predominantly characterize DEC strains based on a shared combination of virulence genes and immunological responses common in clinical outbreak strains such as *E. coli* O157:H7. Such characterization may potentially neglect a slew of other strains that pose a health risk but are not genetically like common outbreak DEC. These atypical strains nevertheless have the ability to cause mostly ‘non-fatal’ but equally health detrimental recurrent infections and inflammation especially among immune compromised groups.

For this reason, we aimed to ascertain whether *E. coli* strains previously isolated from food sources and irrigation water in South Africa are associated with factors linked to ID. We hypothesized that non-DEC isolated from food sources and irrigation may not be pathogens but cause intestinal microbial imbalance especially among vulnerable groups such as infants and immune-compromised adults. This hypothesis would be based on the ability to form biofilms, elicit cytotoxins, disrupt epithelial cell barriers and induce Interleukin-8 *in vitro*. Since ID is associated with a bloom of *Enterobacteriaceae* and inflammation within the GIT, these factors may potentially be used as markers for accessing the risk to this condition associated with non-DEC and subsequently infection by DEC. Our data may provide more definitive information regarding the risk of ID by non-DEC transmitted via food sources and irrigation water. To the best of our knowledge, our work provides the first attempt to link non-DEC isolated from food sources and irrigation water to factors associated with ID. This information may help food safety and health officials look beyond outbreak pathogens when investigating the aetiology of foodborne illness associated with pathogenic *E. coli*.

4.2.2. Materials and methods

4.2.2.1 Sources of *E. coli* strains included in this study

We used previously characterized DEC and non-DEC strains isolated from food sources and irrigation water in South Africa (Aijuka et al., 2018) (Table 4.2.1). Briefly, the strains had been characterized for presence of virulence genes associated with common DEC pathogens according to Panchalingam et al. (2012) as well as proof of attachment to epithelial cells (Nataro et al., 1987). The strains that did not possess any of the sought-after virulence genes or adherence patterns on epithelial cells were labelled as non-DEC while those positive for any of the characteristics were described as DEC. The strains had been isolated from producer distributor bulk milk (PDBM), irrigated lettuce, irrigation water and street vendor coleslaw purchased in Pretoria, South Africa.

4.4.2.2 Biofilm formation

Biofilm formation was carried out according to Mohamed et al. (2007). The conditions were selected to mimic those within a human host environment of approximately 37°C and high nutrient rich media. A total of 38 strains including 12 DEC strains (PDBM=6; irrigated lettuce=3, standard DEC strains=3) and 26 non-DEC strains (PDBM=14, irrigated lettuce=2; irrigation water=10) were grown in LB broth (AmericanBio, Inc) and DMEM-F12 (Thermo Fisher Scientific) + 0.5% D-Mannose. D-Mannose has been shown to influence Type I pili mediated biofilm formation in *E. coli* (Pratt and Kolter, 1998). The standard strains EAEC strain 042, EPEC E2348/69 and DAEC E2348/69 strain F1845 were used as positive pathogenic controls since they have ability to cause intestinal infection suggesting ability to persist and colonize human epithelial gut surfaces. The foodborne (PDBM) N26, previously shown to have low biofilm formation capacity (optical density, $OD_{620nm}=0.03$) under the conditions used in this study (data not shown) was used as the negative control for non-biofilm producers. However, in DMEM-F12+0.5% Mannose one standard DEC strain (EPEC E2348/69) was excluded while an extra strain from irrigation water was added keeping the total number of strains at 38. Strains were grown with shaking (at 250 rpm) overnight (12 h) at 37°C.

Glass cover slips were placed in 24 well plates and 1mL of LB and DMEM-F12+0.5% D-Mannose added. Bacterial suspensions (30 μ L) of each strain were added and incubated for 6h. Washing was done 3X with PBS(Sigma-Aldrich) after which 500 μ L of 2% formalin was added for 20 min followed by washing and crystal violet staining (Sigma-Aldrich). The dry biofilms from each strain previously stained with crystal violet had 1 mL methanol added and left to stand at room temperature for 10 min. The supernatant was transferred into 96 well plates and OD at 620nm (OD_{620}) read using a microplate reader (Biochrome® Anthos MultiRead 400). Each assay was performed in quadruplicate and repeated on three different occasions.

4.2.2.3 Extracellular polymeric substance formation

We screened all strains tested for biofilm formation above for their ability to form EPS by growing them on DMEM-HG agar (Sigma-Aldrich) at 25°C for 72 h. EPS formation was confirmed by presence of mucoid colonies (Flemming and Wingender, 2010). Only 5 strains (DEC=1, non-DEC=4; PDBM=2, irrigated lettuce=1, irrigation water=1) showing the most mucoid colonies were selected and processed for imaging using negative staining and TEM to observe the structure of EPS.

4.2.2.4 Visualization of EPS using TEM

For each test strain, a loopful of 24 h culture was immersed in an Epon resin mixture made from mixtures A and B. Mixture A consisted of 5mL of EMbed 812 (Electron Microscopy Sciences) and 8mL of Dodecenylsuccinic anhydride (DDSA) (Sigma-Aldrich). Mixture B was composed of 8mL of Embed 812 and 7mL of Nadic Methyl Anhydride (Ted Pella Inc). The samples were immediately delivered to The Advanced Microscopy Unit, University of Virginia Medical School for further processing and observation.

4.2.2.5 Cytotoxic activity

Production of toxins is a common feature among pathogenic *E. coli* such as EAEC, ETEC and shigatoxin producing *E. coli* (STEC) that cause diarrhoea. Therefore, we examined these strains for in-vitro production of cytotoxins as an indicator of potential pathogenicity by disruption of epithelial cell integrity. A total of 10 strains were used for this assay. The selection of these strains was based on their ability to lyse HeLa cells during DEC pathotype characterization with the cell adherence assay (Aijuka et al., 2018). During this assay (cell adherence assay) some strains had shown the capacity to lyse HeLa cells after 3 h of incubation at 37°C under CO₂ conditions. Therefore, it was these strains that were selected for this assay. Three DEC (PDBM= 2; irrigated lettuce=1) and 7 Non-DEC (PDBM=4; coleslaw=1; irrigation water=2) comprised the strains. HeLa cells at 80% confluence were aseptically transferred into 24-well plates (Fisher Scientific) containing 12 mm cover slips (Fisher Scientific) in each well and washed with PBS and 1 mL of DMEM-F12+ 0.5% D-Mannose was added. Strains were grown overnight in 3 mL of LB broth in 13 mL plastic tubes while shaking at 37°C.

The bacterial cultures were centrifuged (5000 rpm) to separate the bacteria from the growth medium, the supernatants filter sterilized and 200µL of this mixture added to the HeLa cell monolayers and incubated for 6 h at 37°C in a CO₂ incubator.

The HeLa cells were washed gently 3X with PBS, fixed with 500 µL of 2% formalin for 20 min at room temperature. The samples were rinsed 3X with distilled water (dH₂O) and stained with 500 µL of a Giemsa staining solution (Fisher Scientific) for 20 min. The samples were rinsed (3X) with dH₂O until the colour disappeared. The coverslips were removed from the 24-well plates, air dried and mounted with a tiny drop of Cytoseal (Fisher Scientific) mounting glue onto a glass slide. The samples were observed under a Zeiss light microscope and images were recorded at 60X.

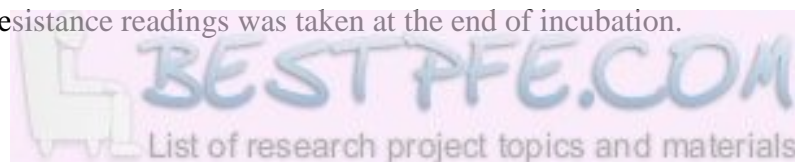
4.2.2.6 Disruption of epithelial cell integrity and Cytokine specific protein secretion

We investigated the ability of these strains to reduce TER and induce IL-8 on polarized T-84 cells in a bid to examine the potential risk posed to human gut inflammation, infection and subsequent disease. In this assay 7 DEC (PDBM=4, irrigated lettuce=2, DEC standard =1) and 13 Non-DEC (PDBM=5, irrigated lettuce=4, irrigation water=2, commensal standard strain=1) strains were used. The assay was carried out according to (Strauman et al., 2010) with a few modifications.

Human colonic T84 cells (American Type Culture Collection CCL-248) were routinely maintained in DMEM-F-12 media (Sigma-Aldrich) containing 10% FBS (Sigma-Aldrich), 50U/mL penicillin, and 50 mg/mL of streptomycin. Polarized T84 cells were generated as follows: T84 cells passages 8 to 20 were seeded at a density of 3×10^5 cells/mL onto collagen coated 12-mm polycarbonate Transwell permeable support cell culture inserts (0.4 μ m pore size, Corning) and grown for 10 to 14 days, during which time the cells were fed daily. Monolayer resistance was determined using an EVOM ohmmeter with the Endohm 12 and STX2 electrodes (World Precision Instruments Inc). Monolayers were considered polarized when resistance was equal to or greater than $1000\Omega/\text{cm}^2$ and not more than $2000\Omega/\text{cm}^2$. Background resistance for collagen-coated cell free membranes was subtracted from initial resistance values to obtain resistance values to be used in statistical analyses.

One hour before infection, polarized T84 cells were washed three times with PBS to remove FBS and antibiotics. Fresh DMEM-F-12 containing 1% D-mannose was then added to the apical and basolateral compartments and incubated at 37°C in 5% CO₂ for 60 min. Overnight cultures of bacteria were standardized in DMEM-HG to OD₆₀₀ of 0.30 ± 0.02 , which is approximately 2×10^8 to 4×10^8 CFU/mL. D-Mannose was added to each culture to a concentration of 1%.

For each sample, 100 μ L of culture were added to the apical side of three separate wells. DMEM-HG+1% D-Mannose was added to uninfected wells and the well with no cells. The infected cells were incubated for 3 h at 37°C in 5% CO₂, at the end of which time the bacteria were aspirated from the upper wells and the cells were washed three times with PBS. Fresh DMEM-F12+1% D-mannose was added to the apical chamber, 100 μ g/mL gentamicin was added to the apical and basolateral chambers, resistance was measured, and the cells incubated for a further 21 h. Resistance readings was taken at the end of incubation.



For detection of IL-8 secretion into culture supernatants a Human IL-8 Elisa kit (Invitrogen) was used per the manufacturer's instructions. Briefly 50µL of standards (prepared via serial dilutions using 160 µL) and samples were added to anti-human IL-8 pre-coated 96 well strip plates and incubated at room temperature (25°C) for 1h. Plates were washed 3X and 50 µL of Biotinylated Antibody Reagent was added to each well followed by incubation at room temperature for 1h. Plates were washed 3X and 100 µL of Streptavidin-HRP solution added to each well. Plates were covered with adhesive plate covers and incubated (25°C) for 30 min followed by washing (3X) and addition 100 µL of TMB (3,3',5,5'-tetramethylbenzidine) reagent to each well. Plates were placed in a dark room for 30 min. The reaction was stopped by adding 100 µL of stop solution to each well. The (OD₄₅₀-OD₅₅₀) was measured using a microplate reader (Biochrome® Anthos MultiRead 400).

4.2.2.7 Statistical analyses

ANOVA main and interaction effects looking at multiple independent variables was used. Independent variables analysed included pathotype (DEC and Non-DEC) and source of *E. coli* strain isolation (PDBM, irrigation water, irrigated lettuce, coleslaw, DEC standard strain, commensal standard *E. coli* strain). The dependent variables for each analysis included; biofilm formation (in LB and/or DMEM-F12+0.5% D-Mannose), disruption of epithelial cell integrity (TER reduction) and lastly induction of IL-8 on polarized T-84 cells. IBM SPSS software version 20 (IBM Corporation) was used for all the analyses.

4.2.2.7.1 Principal component analysis

We used Principal Component Analysis (PCA) to provide insights into associations amongst the different dependent variables (analyses). Additionally, PCA was used to cluster strains based on the associations amongst the 4 variables assayed in the previous sections above. The collected data (biofilm formation, cytotoxic activity, disruption of epithelial cell barriers and IL-8 induction in polarized T-84 cells) from 11 strains (DEC=4, Non-DEC=7; PDBM=7, irrigation water=2, irrigated lettuce=2) that had complete and reproducible information for all analyses was used. Standard EAEC strain 042 was used as a comparison pathogenic control.

The PCA biplots were aimed at determining which of the tested variables had most influence on *E. coli* strain differentiation with regards to pathotype (DEC or non-DEC) and source of strain isolation (PDBM, irrigation water, irrigated lettuce) when compared to DEC standard strain, EAEC strain 042. PCA biplots were constructed using XLSTAT 2017 (Microsoft Corporation). It was vital to investigate any potential relationship among the strains because

they were isolated over a wide range of ecological, geographical and temporal scales (Aijuka et al., 2015; Ntuli et al., 2016) which is crucial in accessing source tracking and understanding habitats of foodborne pathogens.

Table 4.2.1 Strains, sources of isolation, pathotypes and assays used for characterizing *Escherichia coli* isolated from food sources and irrigation water in South Africa

Strain code1	Source	<i>E. coli</i> strain type	Assay			
			Biofilm	Human IL-8	Cytotoxicity	TER
N26	PDBM	Non-DEC	+	-	-	-
M37	PDBM	Non-DEC	+	+	+	+
NW(V)10(1)	Irrigated lettuce	Non-DEC	+	+	+	+
MPU(W)8(2)	Irrigation water	Non-DEC	+	-	-	-
MPU(W)1(1)	Irrigation water	Non-DEC	+	-	-	-
N5	PDBM	Non-DEC	+	-	+	-
M24	PDBM	Non-DEC	+	-	+	-
M11	PDBM	Non-DEC	+	-	-	-
58 or 85	PDBM	Non-DEC	+	-	-	-
MPU(W)8(3)again	Irrigation water	Non-DEC	+	-	-	-
M4	PDBM	Non-DEC	+	-	-	-
MPU(W)9(3)	water	Non-DEC	+	-	-	-
NW(W)9(1) ²	water	Non-DEC	+	+	-	+
L5	PDBM	Non-DEC	+	+	-	+
CR12	Irrigation water	Non-DEC	+	-	+	-
MPU(W)6(2)	Irrigation water	Non-DEC	+	-	-	-
NW(W)6(1)	Irrigation water	Non-DEC	+	-	-	-
K3	PDBM	Non-DEC	+	+	+	+
F8	PDBM	Non-DEC	+	-	-	-
54	PDBM	Non-DEC	+	+	-	+
N6	PDBM	Non-DEC	+	-	-	-
NW(W)3	Irrigation water	Non-DEC	+	-	-	-
Vp ²	Irrigated lettuce	Non-DEC	+	-	-	-
CR4	Irrigation water	Non-DEC	+	-	+	-
M13	PDBM	Non-DEC	+	-	-	-
513	PDBM	Non-DEC	+	-	-	-
Nana 10	Coleslaw	Non-DEC	-	+	+	+
Vq ¹	Irrigated lettuce	Non-DEC	-	+	-	+
V5	Irrigated lettuce	Non-DEC	-	+	-	+
V9	Irrigated lettuce	Non-DEC	-	+	-	+
Vd	Irrigated lettuce	Non-DEC	-	-	-	+
NW(V)6(2)	Irrigated lettuce	Non-DEC	-	+	-	+
N25 ²	PDBM	Non-DEC	-	-	+	-
MPU(W)4	Irrigation water	Non-DEC	-	+	-	+
M6	PDBM	DEC	+	-	-	-
N23	PDBM	DEC	+	+	+	+
K2 ²	PDBM	EAEC (<i>aaiC</i>)	-	+	-	+
NW(V)7(3)	Irrigated lettuce	DEC	+	-	-	-
NW(V)10(1)again	Irrigated lettuce	DEC	+	+	-	+
LeK	Irrigated lettuce	DEC	+	+	-	+
M12	PDBM	DEC	+	-	-	-
M28	PDBM	DEC	+	-	-	-
K5	PDBM	DEC	+	+	-	+
K16	PDBM	DEC	+	+	+	+

¹; In subsequent analyses shown in the results section, all strains were analysed based solely on *E. coli* strain type/pathotype (DEC or Non-DEC) and source of isolation (food or irrigation water) rather than on individual strains. Only strains showing a positive output for each assay are were reported in the results sections.

(+) assay carried out

(-) Assay not carried out

All DEC were EAEC apart from strain coded (LeK) that showed invasiveness on HeLa cells and presumptively identified as enteroinvasive *E. coli*

²; Tested for colonic acid formation with Transmission Electron Microscopy

TER-Trans epithelial electrical resistance. Strains showing TER reduction were subsequently tested for interleukin-8 (IL-8) induction.

4.2.3. Results

4.2.3.1 Biofilm formation in LB broth

Biofilm formation varied based on *E. coli* strain type (DEC or non-DEC) and source of strain isolation (PDBM, irrigation water, irrigated lettuce, DEC human pathogenic standard strain) i.e. There was a significant interaction ($p \leq 0.05$) between these factors aiding or hampering biofilm formation. Higher ($p \leq 0.05$) mean biofilm formation was noted in standard DEC and DEC isolated from irrigated lettuce with mean OD of 0.098 and 0.092 respectively compared to all other strains (Figure 4.2.1). Conversely, lower ($p \leq 0.05$) mean biofilm formation was noted in non-DEC isolated from PDBM (0.060), irrigation water (0.051), irrigated lettuce (0.051) and DEC isolated from PDBM (0.050). However, the high variability noted among the strains and sources suggests using a larger set of strains in order to draw meaningful biological conclusions.

4.2.3.2 Biofilm formation in DMEM-F12+0.5% Mannose

Similarly, as noted above in LB broth, biofilm formation of the same strains in this media also varied based on the strain type (DEC and non-DEC) and isolation source (PDBM, irrigation water, irrigated lettuce and human pathogen) i.e. There was a significant interaction ($p \leq 0.05$) between these factors aiding or hampering biofilm formation (Figure 4.2.2). The highest mean biofilm formation was noted in DEC isolated from irrigated lettuce with mean OD_{620nm} of 0.123. However, in contrast to growth in LB broth, significantly ($p \leq 0.05$) lower mean biofilm formation was noted with all other *E. coli* strain types regardless of isolation source (Figure 4.2.2). The mean biofilm formation for DEC isolated from PDBM, control human pathogenic *E. coli* as well as non-DEC isolated from PDBM, irrigation water and irrigated lettuce ranged from 0.052 to 0.075. Similarly, as previously noted with LB broth, the high variability in biofilm formation among the strains suggests using a larger set of strains in order to draw meaningful biological conclusions.

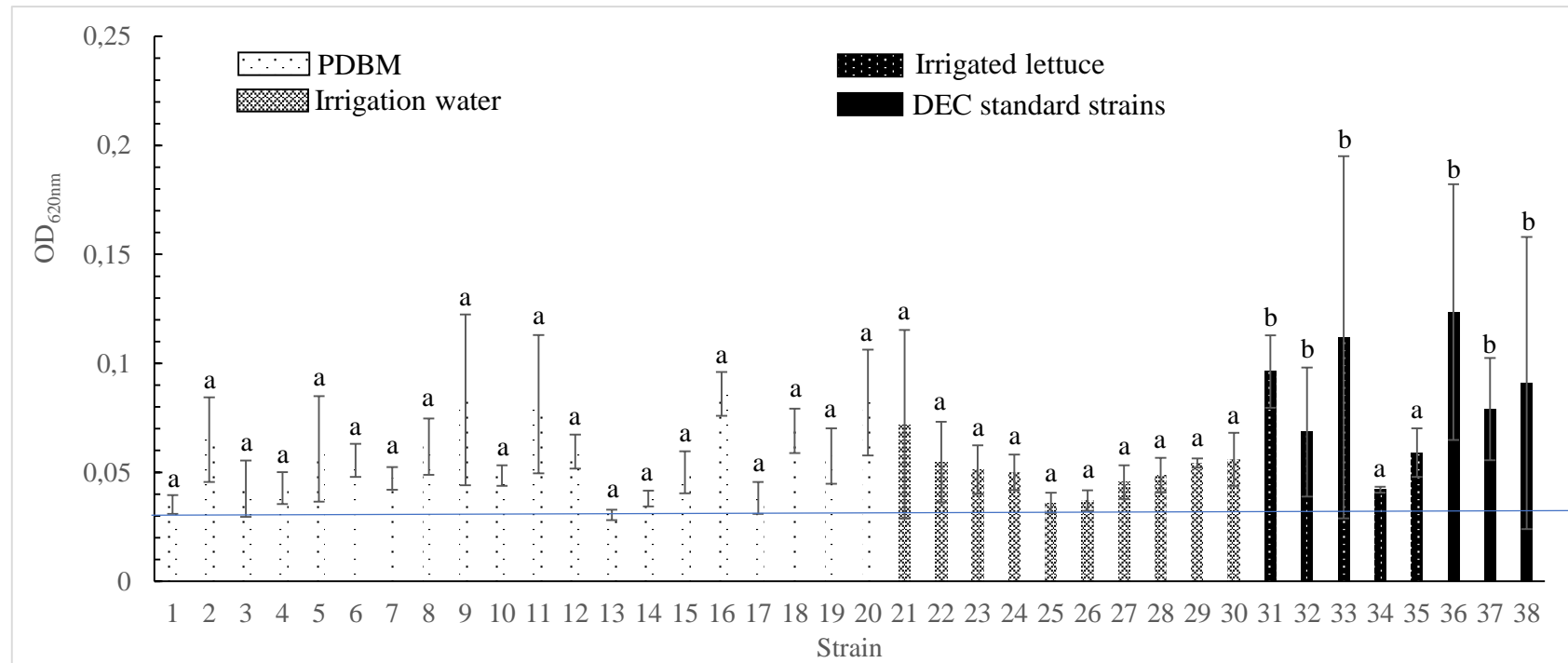


Figure. 4.2.1. Biofilm formation in Luria Bertani broth by diarrheagenic *E. coli* (DEC) and non-DEC isolated from food sources and irrigation water. DEC strains: 1 to 6 (PDBM); 31 to 33 (irrigated lettuce); 36 to 38 (DEC standard strains). Non-DEC: 7 to 20 (PDBM); 21 to 30 (irrigation water); 34 to 35 (irrigated lettuce). a, b: Biofilm formation significantly different ($p \leq 0.05$). DEC standard strains: 36-enteroaggregative *E. coli* strain 042; 37-enteropathogenic *E. coli* strain 2348; 38-diffusely adherent *E. coli* strain F1875. Strain codes: 1-M6, 2-N23, 3-K16, 4-K5, 5-M28, 6-M12, 7-M37, 8-N5, 9-M24, 10-M11, 11-58or85, 12-M4, 13-N26, 14-L5, 15-K3, 16-F8, 17-54, 18-N6, 19-M13, 20-513, 21-MPU(W)8(2), 22-MPU(W)1(1), 23-MPU(W)8(3)again, 24-MPU(W)9(3), 25-NW(W)9(1), 26-CR12, 27-MPU(W)6(2), 28-NW(W)6(1), 29-NW(W)3, 30-CR4, 31-LeK, 32-NW(V)10(1)again, 33-NW(V)7(3), 34-NW(V)10(1), 35-Vp. The horizontal line shows biofilm production for the low biofilm producing strain N26.

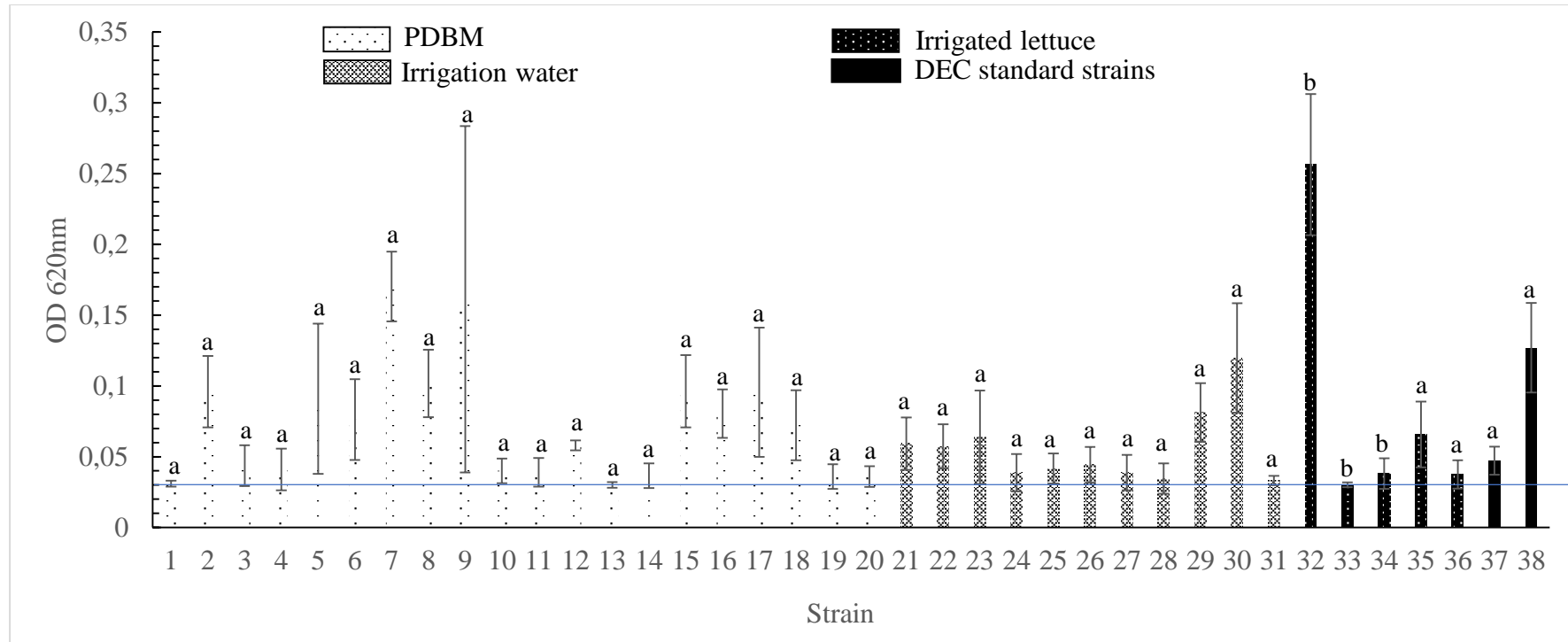


Figure 4.2.2. Biofilm formation in Dulbecco's Modified Eagle Medium-F12+0.5% Mannose by diarrheagenic *E. coli* (DEC) and non-DEC previously isolated from food sources and irrigation water in South Africa. DEC strains: 1 to 6 (PDBM); 32 to 34 (irrigated lettuce); 37 to 38 (DEC standard strains). Non-DEC: 7 to 20 (PDBM); 21 to 31 (irrigation water); 35 to 36 (irrigated lettuce). a, b: Biofilm formation significantly different at $p \leq 0.05$. DEC standard strains: 37-enteroaggregative strain 042; 38-diffusely adherent *E. coli* strain F1875. Strain codes: 1-M6, 2-N23, 3-K16, 4-K5, 5-M28, 6-M12, 7-M37, 8-N5, 9-M24, 10-M11, 11-58or85, 12-M4, 13-N26, 14-L5, 15-K3, 16-F8, 17-54, 18-N6, 19-M13, 20-513, 21-MPU(W)8(2), 22-MPU(W)1(1), 23-MPU(W)8(3)again, 24-MPU(W)9(3), 25-NW(W)9(1), 26-CR12, 27-MPU(W)6(2), 28-NW(W)6(1), 29-NW(W)3, 30-CR4, 31-NW(W)9(3), 32-LeK, 33-NW(V)10(1)again, 34-NW(V)7(3), 35-NW(V)10(1), 36-Vp. PDBM-Producer Distributor Bulk Milk. The horizontal line shows biofilm production for the low biofilm producing strain N26.

4.2.3.3 EPS formation

All 36 test strains showed mucoid colonies suggestive of EPS formation. The 5 representative strains (showing the highest visual production of mucoid colonies) observed under TEM were characterized by an amorphous extracellular mass attached to individual bacteria providing a qualitative assessment of EPS production (Figure 4.2.3).

4.2.3.4 Cytotoxic activity

Of the 10 strains tested for cytotoxic activity, 6 (60%) strains were positive with 5 (50%) of them non-DEC and 1 (10%) DEC (Figure 4.2.4). Therefore, more non-DEC (that would have previously been characterized as potentially non-pathogenic) than DEC produced toxins suggesting a previously unquantified health risk to consumers. Additionally, of the 6 strains that were positive for cytotoxic activity all were isolated from food sources which usually undergo minimal-to-no processing prior to human consumption. Four (67%, 4 out of 6 strains) were isolated from PDBM and 1 (17%, 1 out of 6) strain each from coleslaw and irrigated lettuce. These observations show the direct risk posed by non-DEC to the health of consumers through the food chain on the basis of cytotoxic activity.

4.2.3.5 Epithelial cell barrier disruption and IL-8 induction in polarized T-84 epithelial cells

Mean percentage reduction in initial Transepithelial electrical resistance, TER (53.5 to 73.8%) from strains isolated from PDBM, irrigation water, irrigated lettuce and coleslaw was comparable to the DEC standard strain (EAEC strain 042) (Figure 4.2.5). For all strains assayed, IL-8 induction varied with isolation source (Figure 4.2.6). Although, the highest IL-8 induction was noted in DEC standard strain *E. coli* 042 (349.07pg/mL), significantly higher ($p \leq 0.05$) mean levels were observed in strains isolated from irrigation water (231.93pg/mL), irrigated lettuce (152.98pg/mL) and PDBM (139.21pg/mL) compared to commensal strain aafa (99.27pg/mL). Comparison studies with a larger set of strains are suggested in order to draw more meaningful conclusions based on the high variation noted using a single pathogenic and commensal strain.

4.2.3.5 Association of DEC and non-DEC with factors associated with intestinal dysfunction *in-vitro*

Approximately 60% of the variation in all strains was explained by factors 1 and 2 (F1 and F2) (Figure 4.2.7). The PCA correlation curve showed positive correlations between the following variables; biofilm formation in LB and induction of IL-8, induction of IL-8 and cytotoxic activity, biofilm formation in LB and DMEM-F12+0.5% Mannose.

Strains clustered into 3 distinct groups (1, 2 and 3) with two strains (1 and 10) grouping separately from the rest (Figure 4.2.8). The strains clustered together irrespective of the source of strain isolation although food and environmental strains clustered far from the clinical standard DEC strain *E. coli* 042. All strains in group 1 showed low biofilm formation in LB broth and DMEM-F12+0.5% Mannose, low IL-8 induction as well as percentage mean reduction in initial TER on polarized T-84 cells. On the other hand, strains in group 2 showed high biofilm formation in DMEM-F12+0.5% Mannose, low mean percentage reduction in initial TER and no cytotoxic activity. Furthermore, strains in group 3 induced IL-8 in polarized T84 cells and had cytotoxic activity. Lastly strain 1 showed biofilm formation in LB broth and DMEM-F12+0.5% Mannose while strain 10 showed highest biofilm formation of all strains in DMEM-F12+0.5% Mannose. Overall, pathogenic predictors (biofilm formation in LB, induction of IL-8 and cytotoxic activity) which showed positive correlation with EAEC strain 042, showed poor, inverse or no relationship with most DEC and Non-DEC strains (Groups 1 and 2) isolated from food sources and irrigation water with exception of those in Group 3 (strains 6 and 11).

We do acknowledge that the number of strains used for this assay was minimal and would thereby affect the validity of any deductions made. However, the study was reductionist in approach and meant to carry out analyses only of strains that had previously tested positive for a phenotypic or genotypic characteristic potentially linked to pathogenicity *in-vitro* e.g possession of virulence genes and cytotoxic activity. This meant that fewer strains would subsequently be assayed downstream but the characterization (isolation sources, pathotypes) would be maintained and provide an acceptable representation of the health risk faced by the consumer.

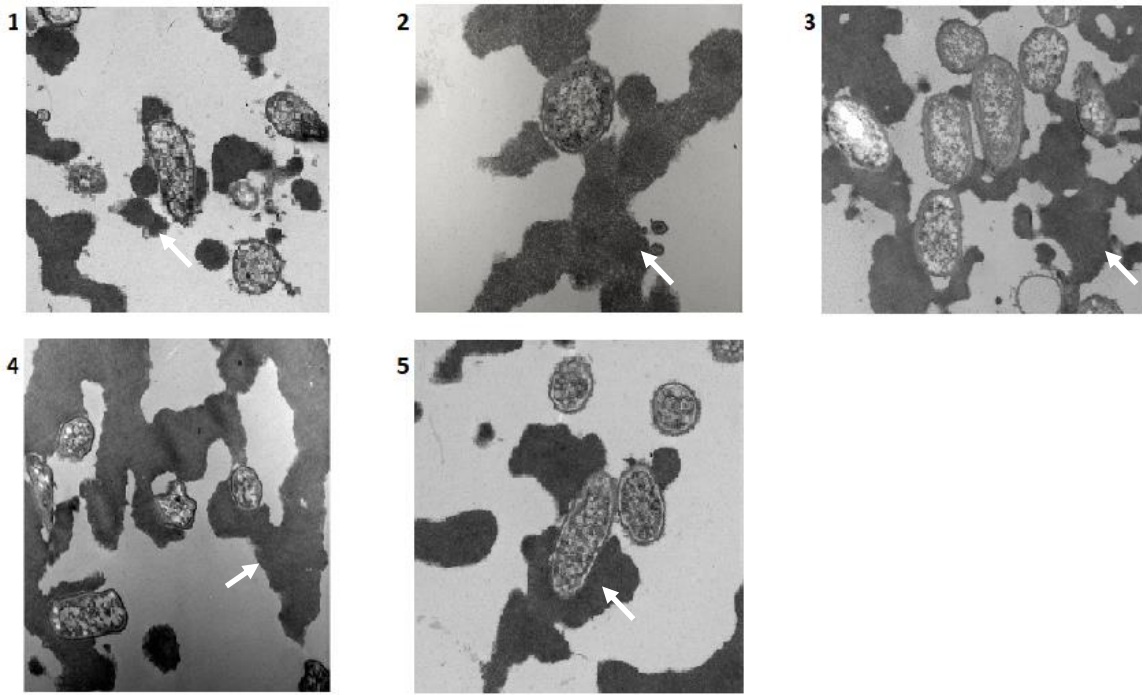


Figure 4.2.3. Extracellular polymeric substance (EPS) formation (amorphous extracellular mass attached to bacterial wall) by diarrheagenic (DEC) and non-DEC previously isolated from food sources and irrigation water in South Africa visualized under Transmission Electron Microscopy. Arrows point to EPS. Strain pathotype and source of isolation: 1- DEC isolated from PDBM; 2- Non-DEC isolated from PDBM; 3-non-DEC isolated from irrigation water; 4 and 5-non-DEC isolated from irrigated lettuce. Magnification X15000. Scale 0.5 to 1.0 μ m. Strain codes:1-K2, 2-N25, 3-NW(W)9(1), 4-Vp, 5-NW(V)3(1). PDBM-Producer Distributor Bulk Milk.

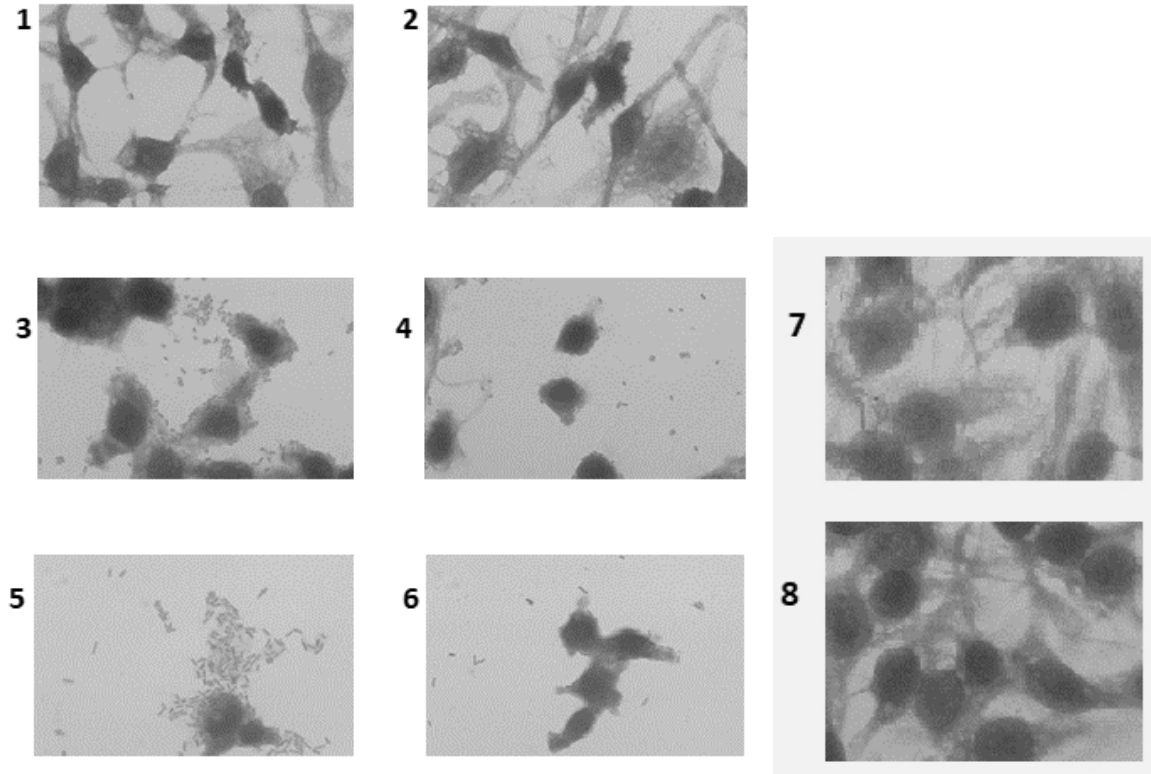


Figure 4.2.4 Cytotoxic activity (disruption of the HeLa cell cytoskeleton) in diarrheagenic *E. coli* (DEC) and Non-DEC previously isolated from food sources and irrigation water in South Africa. Images recorded at 60X. Scale=20 μ m. Strain pathotype and source of isolation: 1-Non-DEC isolated from coleslaw; 2-DEC isolated from PDBM; 3, 4 and 5-Non-DEC isolated from PDBM; 6-Non-DEC isolated from irrigated lettuce. Strains 7 and 8 (no cytotoxic activity), 7-Non-DEC isolated from PDBM; 8-DEC isolated from PDBM. Test strain codes: 1-Nana 10, 2-K16, 3-M24, 4-N25, 5-K3, 6-NW(V)10(1), 7-N5, 8-N23. PDBM-Producer Distributor Bulk Milk. Only strains exhibiting cytotoxicity are reported.

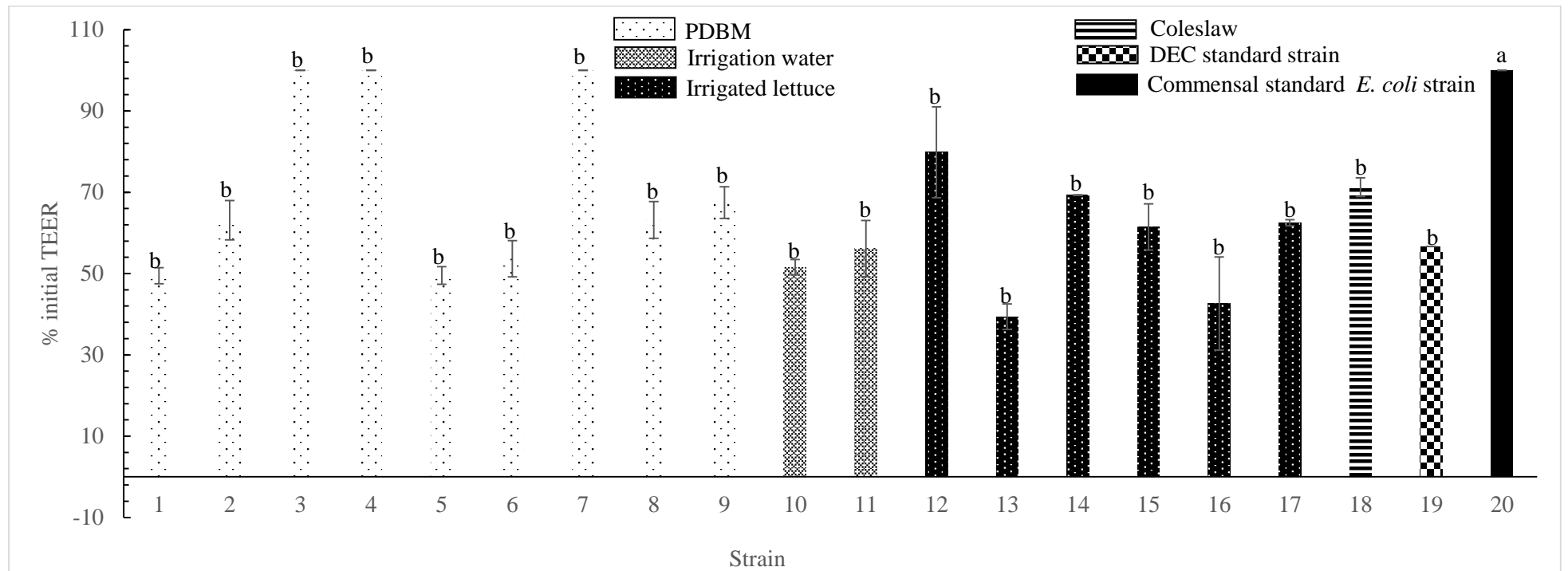


Figure 4.2.5. Percentage reduction in initial Transepithelial Electrical Resistance (TER) observed in polarized T-84 epithelial cells infected with diarrheagenic (DEC) and non-DEC previously isolated from food sources and irrigation water in South Africa. Strain pathotype and source of isolation: DEC strains: 1 to 4 (PDBM); 12 to 13 (irrigated lettuce); 19 (DEC standard strain). Non-DEC: 5 to 9 (PDBM); 10 to 11 (Irrigation water); 14 to 17 (Irrigated lettuce); 18 (Coleslaw); 20 (Standard commensal *E. coli* strain aafa). a, b: TER significantly different at $p \leq 0.05$. 19-DEC standard strain enteroaggregative *E. coli* strain 042; 20- Commensal standard *E. coli* strain aafa. Strain codes: 1-K2, 2-N23, 3-K5, 4-K16, 5-K3, 6-54, 7-513, 8-M37, 9-L5, 10-NW(W)9(1), 11-MPU(W)4, 12-LeK, 13-NW(V)10(1)again, 14-Vq, 15-NW(V)6(2), 16-V5, 17-V9, 18-Nana 10. PDBM-Producer Distributor Bulk Milk. Only strains exhibiting reduction in TER are reported.

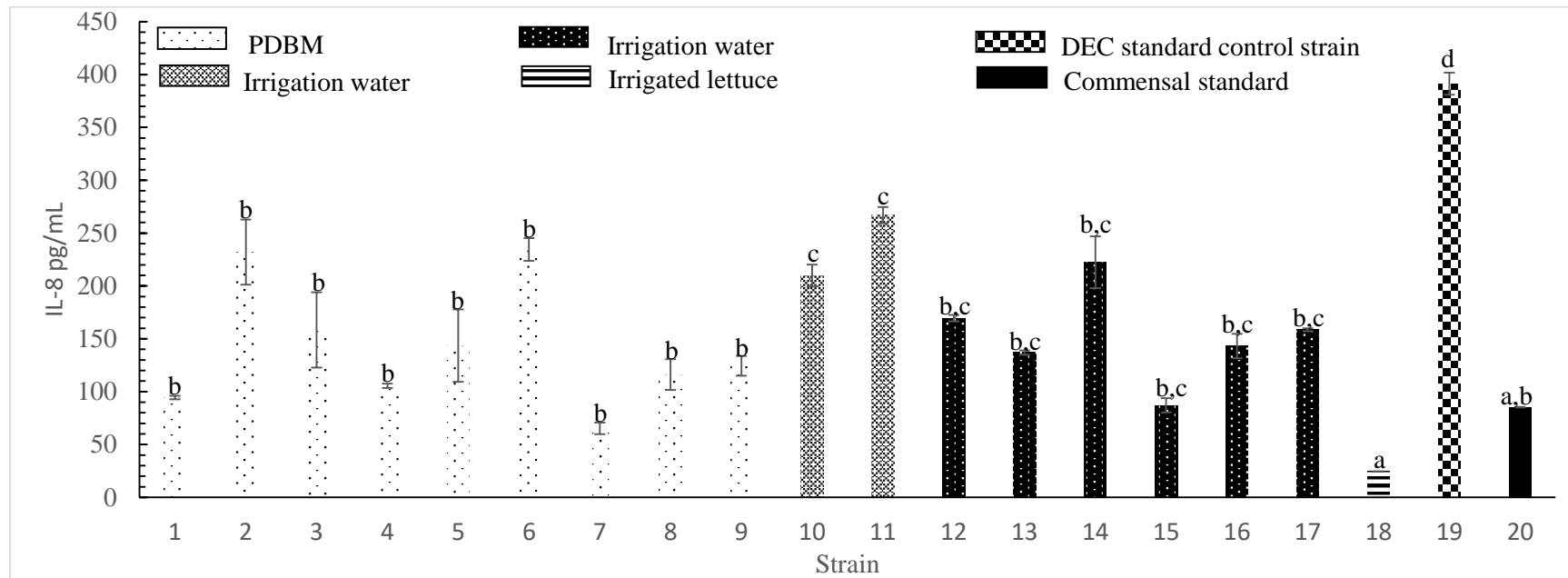


Figure 4.2.6. Interleukin 8 (IL-8) induction in polarized T-84 epithelial cells resulting from infection with diarrheagenic *E. coli* (DEC) and non-DEC previously isolated from food sources and irrigation water in South Africa. Strain pathotype and source of isolation: DEC strains: 1 to 4 (PDBM); 12 to 13 (irrigated lettuce); 19 (DEC standard strain). Non-DEC: 5 to 9 (PDBM); 10 to 11 (irrigation water); 14 to 17 (irrigated lettuce); 18 (Coleslaw); 20 (standard commensal *E. coli* strain). a, b, c, d induction of IL-8 significantly different ($p \leq 0.05$). 19-DEC standard control strain EAEC strain 042; 20-Standard commensal *E. coli* strain aafa. Strain codes: 1-K2, 2-N23, 3-K5, 4-K16, 5-K3, 6-54, 7-513, 8-M37, 9-L5, 10-NW(W)9(1), 11-MPU(W)4, 12-Lek, 13-NW(V)10(1)again, 14-Vq, 15-NW(V)6(2), 16-V5, 17-V9, 18-Nana 10. PDBM-Producer Distributor Bulk Milk. Only strains exhibiting induction of IL-8 are reported.

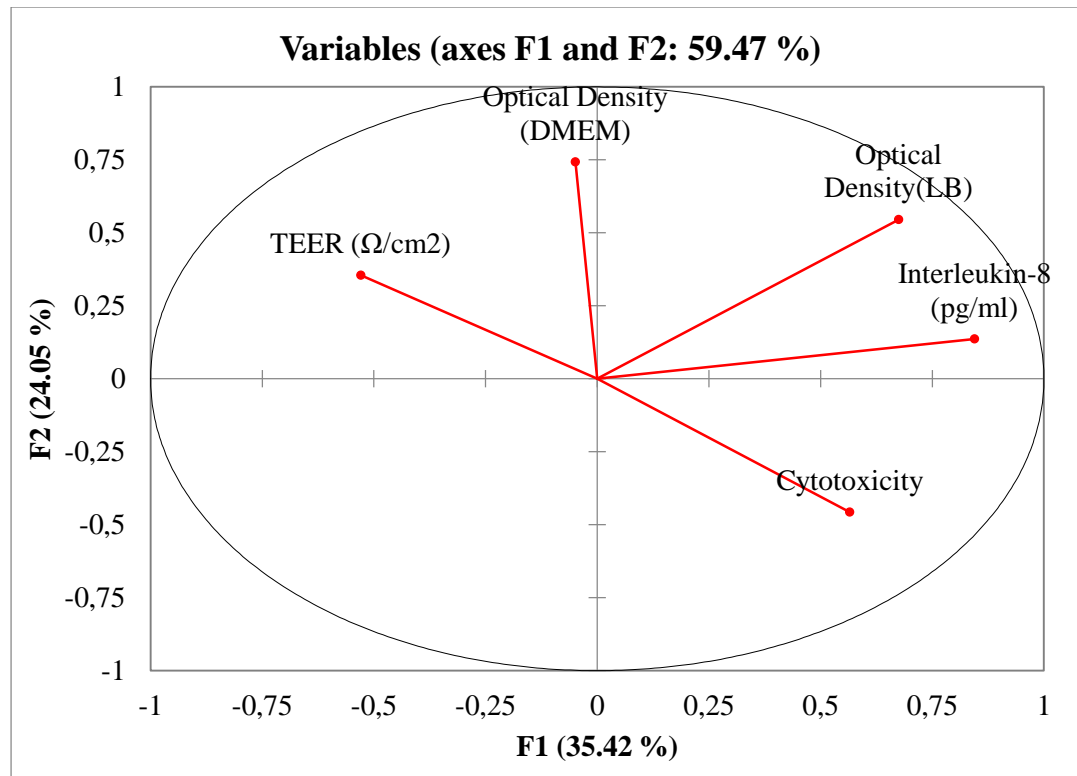


Figure 4.2.7. Principal Component Analysis variables correlation plot showing axes Factors 1 and 2 (F1 and F2) accounting for approximately 60% of variation observed after analysis of 4 factors in diarrheagenic *E. coli* (DEC) and non-DEC tested in-vitro potentially associated with human gut intestinal dysfunction. Analyses included biofilm formation in Luria Bertani and Dulbecco's Modified Eagle Medium-F12+0.5% Mannose media, Cytotoxic activity on HeLa cells, disruption of epithelial cell integrity (TER reduction) and induction of Interleukin-8 on polarized T-84 cells. The strains were previously isolated from food sources and irrigation water. The strains were compared to DEC standard strain enteroaggregative *E. coli* strain 042. TER-Trans epithelial electrical resistance

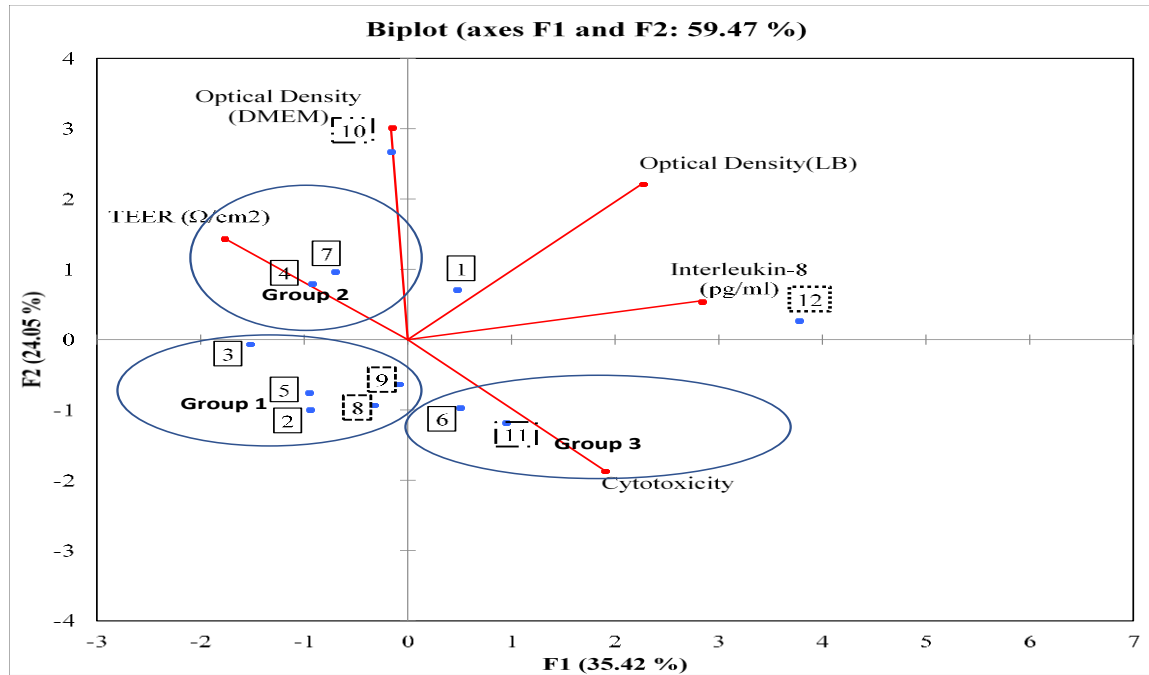


Figure 4.2.8. Principal Component Analysis biplot showing axes Factors 1 and 2 (F1 and F2) accounting for approximately 60% of variation observed in 11 strains of diarrheagenic (DEC) and non-DEC isolated from food sources and irrigation water. The strains were analysed in-vitro for 4 factors potentially associated with human gut intestinal dysfunction. Analyses included; Biofilm formation in Luria Bertani and Dulbecco's Modified Eagle -F12+0.5% Mannose media, cytotoxic activity, disruption of epithelial cell integrity (TER reduction) and induction of Interleukin-8 on polarized T-84 cells. The strains were previously isolated from food sources and irrigation water. DEC standard strain EAEC strain 042 was used for comparison. Strains and source of isolation: 1 to 3 (DEC isolated from PDBM); 4 to 7 (non-DEC isolated from PDBM); 8 to 9 (non-DEC isolated from irrigation water); 10 (DEC isolated from irrigated lettuce); 11 (non-DEC isolated from irrigated lettuce); 12 (EAEC standard strain 042). Strain codes: 1-N23, 2-K16, 3-K5, 4-M37, 5-L5, 6-K3, 7-513, 8-NW(W)9(1), 9-NW(W)6(1), 10-LeK, 11-NW(V)10(1).

4.2.4. Discussion

The variation of biofilm formation with pathotype (DEC or non-DEC) and isolation source (food or irrigation water) suggests that these two factors may affect the ability of a strain to persist within a given environment such as a food source, open environment or human body. However, large variations in biofilm formation of strains in both growth media probably due to the few isolates used prevented the drawing of definitive conclusions. Therefore, the characterization of more strains from different food and environmental sources based on biofilm formation may provide useful information on which food and environmental sources frequently harbour strains that can proliferate under conditions like those in the human body. Such information is vital for controlling the risk of foodborne illnesses associated with potentially pathogenic *E. coli*. Additionally, food sources such as irrigated lettuce may be more susceptible to carrying DEC associated with human disease since they have similar biofilm formation capacity as standard DEC strains in LB broth, which is an indicator of persistence. The persistence of DEC on irrigated lettuce may provide a direct route for entry into the human gut through consumption while the ability to form biofilms in high nutrient conditions at high temperatures (37°C) enables its persistence and subsequent infection of the human gut. Since all DEC strains in this study were EAEC, biofilm formation may have been facilitated by aggregative adherence fimbriae I (AAF/I) in typical EAEC (carrying the global virulence master regulator *aggR*) or the Inc11 plasmid having a Type IV pilus in atypical EAEC (lacking *aggR*). Foodborne, DEC have been suggested to use the same virulence factors to enable persistence within the non-human environments (Nagy et al., 2016). Similarly, biofilm formation is a hall mark of EAEC gastrointestinal pathogenesis (Estrada-Garcia and Navarro-Garcia, 2012) which may also be a strategy used for persistence within the open environment. The high biofilm formation capacity by DEC isolated from irrigated lettuce in D-Mannose containing media (DMEM-F12) suggests its ability to withstand varying environmental growth conditions compared to all other strains in this study. This is because Mannose can contribute to biofilm inhibition as a result of with binding of the Type I pili mannose specific adhesin (*fimH*) which is involved in non-specific binding to abiotic surfaces (Pratt and Kolter, 1998).

The ability to form EPS by all strains tested suggests that they can all form mature biofilms which can persist and spread under environmental conditions. The formation of EPS in *E. coli* is characterized by formation of among other compounds, of colanic acid that helps in reinforcement of the 3-dimensional structure of the biofilm (Beloin et al., 2008). For example EAEC persistence on sprouts was directly influenced by the ability to form colanic acid which

enabled long term persistence within this secondary environment compared to isogenic mutants lacking the colanic acid biosynthesis gene (Borgersen et al., 2018).

The ability of gastrointestinal pathogens such as *E. coli* transmitted through the faecal-oral route to persist at different stages within the food chain has been suggested as an evolutionary adaptation for gaining entry into their principal niche, the human host (Reperant et al., 2012). This is particularly of concern in small unregulated food processing establishments in rural and low income settings similar to areas where milk from this study was purchased (Ntuli et al., 2016) that are particularly vulnerable to poor sanitation and hygiene practices.

The observation that most strains (83%, 5 out of 6 strains) with cytotoxic activity were non-DEC and isolated from food sources such as PDBM (pasteurized and non-pasteurized), irrigated lettuce and coleslaw presents a direct food safety risk especially to vulnerable groups such as infants and immune compromised adults. This is especially concerning because these foods are sometimes consumed without any further processing to eliminate foodborne pathogens. This observation provides a vivid example of how standard characterization of foodborne *E. coli* may fail to determine strains that can potentially cause disease.

Cytotoxic activity exhibited by pathogenic *E. coli* is linked to the possession of virulence genes such as Serine Protease Autotransporters produced by *Enterobacteriaceae* (SPATES), that are known to degrade host intracellular and extracellular substrates causing a variety of adverse effects (Ruiz-Perez and Nataro, 2014). Additionally, in EAEC, the most prevalent DEC pathotype in this study, *pet* (plasmid encoded toxins) and *ShET1* (Shigella enterotoxins) have been associated with cytotoxic activity. Further characterization of toxins from non-DEC should help in assessment of the risk to human health by comparison with known cytotoxins produced by standard DEC strains.

The capacity to reduce Transepithelial electrical resistance, TER as well as induce the inflammatory cytokine IL-8 in polarized epithelial cells by strains previously isolated from a food source and irrigation water suggests their risk of causing infection and inflammation respectively to potential consumers. Additionally, this scenario shows that even strains previously characterized as potentially non-pathogenic (non-DEC) based on standard protocols may have the ability to cause disease or at least ID. The persistent loss of epithelial integrity in epithelial cells by enteric pathogens is due to disruption of tight junction proteins (Ochieng et al., 2014; Strauman et al., 2010) requiring AAF/I or AAF/II in EAEC and a cascade of cellular

disruption events in intracellular pathogens such as EIEC (Croxen et al., 2013). These effects result in the loss of host ions and proteins escalating diarrheal disease (Strauman et al., 2010).

On the other hand the innate immune response is caused by recognition of external bacterial structures such as AAF or flagellin resulting from bacterial attachment that lead to formation of lipid rafts and recognition of external structures providing a specified pattern recognition receptor (PRR) inducing IL-8 (Edwards et al., 2011). The further characterization of adhesins associated with non-DEC from this study will shed light on the mechanisms by which they disrupt epithelial cell barriers and induce IL-8.

PCA analysis did not provide definitive relationships among the different strains based on the observed pathogenic predictors probably because of the few isolates used (having all reproducible data points for all assays). This was in turn affected by the reductionist approach of the study that only subsequently tested strains positive for the pathogenic parameter of interest. Future studies using a larger number of strains including those previously implicated in foodborne outbreaks should help provide a clearer link than one shown in this study regarding the potential risk to food safety and public health.

However, PCA was able to cluster strains based on their isolation source (food, environmental strains and clinical). This suggests that the method is potentially useful in comparing the risk of illness posed by strains from different sources along the food chain as well as identifying points along the food chain that are most prone to contamination with potentially pathogenic *E. coli* enabling microbial source tracking.

4.2.5. Conclusion

This preliminary study using a limited number but ecologically and temporarily diverse group of foodborne and environmental *E. coli* from South Africa suggests that they are associated with factors such as reduction of TER, Induction of IL-8 and cytotoxic activity. Therefore this may potentially lead to intestinal dysfunction especially among vulnerable groups such as infants and immune compromised adults. Additionally, this observation is particularly of concern because some of these foods such as PDBM (milk) may be used for weaning infants thus exacerbating the risk of disrupting development of the normal healthy gut microbiota.

4.3 Detection of Extended Beta Lactamase Cefotaxime Resistance and Virulence Genes in *Escherichia coli* by Duplex Quantitative Real Time PCR and Melt Curve Analysis

Abstract

The rise of emerging highly virulent and antibiotic resistant pathogens presents a global public health risk. Therefore, routine monitoring of their prevalence within the clinical, environmental and food production setting is vital. Real time qPCR coupled with melting curve analysis (MCA) presents an avenue for rapid and accurate characterization of such pathogens. We evaluated commercial qPCR mixes based on SYBR Green I and EvaGreen dyes for development of an assay for simultaneous detection of cefotaxime resistant (CTX-M) extended beta lactamase (ESBL) producing and virulent (shigatoxin 1, shigatoxin 2 and intimin) genes in previously isolated and characterized *E. coli* (n=12) from food sources and irrigation water and irrigated vegetable products in South Africa. Both SYBR Green 1 and EvaGreen dye-based mixes in each duplex assay were able to simultaneously detect two amplicons (shigatoxin 1 and bla_{CTX-M}) and (shigatoxin 2 and intimin) within a single closed tube reaction based on their separation in melting temperature (T_m). However, a higher mean T_m separation between targeted amplicons and smoother melting curves in monoplex and duplex reactions were observed with the EvaGreen based qPCR mix suggesting better performance when targeting multiple amplicons. Therefore, through simple step-wise optimization of DNA template, PCR conditions, primer pair volume and the scanning rate during melting curve analysis, we adopted a conventional PCR assay for detection of large amplicons (375 to 1580 bp) into one based on qPCR. Additionally, the qPCR assay could simultaneously detect antibiotic resistance and virulence genes within a single closed tube reaction. This robust technique can facilitate the development of tailor made assays for rapid and accurate detection as well as the characterization of emerging foodborne and environmental pathogens that pose a risk to public health in different regions of the world.

Key words: *Escherichia coli*, Duplex real time-qPCR, melt curve analysis, Cefotaxime resistance, virulence, extended beta lactamase

4.3.1 Introduction

Foodborne outbreaks are linked to bacterial pathogens such as *Escherichia coli* that exhibit increased virulence and antibiotic resistance. *E. coli* strains isolated from food and environmental sources have exhibited virulence and antibiotic resistance (Aijuka et al., 2015; Heiman et al., 2015; Koga et al., 2015; Ntuli et al., 2018, 2017) presenting a risk to food safety and public health. Additionally, the spread of pathogenic *E. coli* across large geographical areas within developed, developing and least developed countries presents an emerging threat to global food safety, public health and international trade. Pathogenic *E. coli* associated with foodborne illnesses is clinically categorized into a range of diarrheagenic *E. coli* (DEC) pathotypes that cause intestinal illnesses. DEC pathotypes commonly associated with food and water include; enterohaemorrhagic *E. coli* (EHEC), enterotoxigenic *E. coli* (ETEC), enteroaggregative *E. coli* (EAEC) and enteroinvasive *E. coli* (EIEC). These groups can be identified by molecular characterization of virulence genes in isolated *E. coli* strains such as toxins (shigatoxin, *stx1* and *stx2*) as well as attachment proteins (intimin, *eae*), that are commonly used to identify EHEC.

Antibiotic resistance associated with emerging infectious pathogens together with climate change and terrorism are amongst the greatest risks posed to humanity in the 21st century (The Review on Antimicrobial Resistance, 2014). This risk has equally been weighed in by the governments of The United States and European Union (Karp et al., 2017; Roca et al., 2015). For example, the rise of extended beta lactamase enzyme producing bacteria (ESBL) within the last 20 years has increased this risk (Cantón et al., 2012; Shi et al., 2015). Similarly, within the last decade, high prevalence of Cefotaxime resistant (CTX-M) extended beta lactamase producing (ESBL) bacteria has increased the risk of spreading pathogens resistant to 4th generation cephalosporins that are usually the last line of treatment for acute bacterial infections (Shi et al., 2015). The world-wide dissemination of the highly virulent and antibiotic resistant *E. coli* ST131 strain exemplifies this emerging threat posed to food safety and public health.

Developing countries in the BRICS (Brazil, Russia, India, China and South Africa) economic block and specifically South Africa, Brazil and India present an interesting scenario, because they have large sections of the food chain having both a modern and highly controlled system whilst the informal sector lacks similar checks and balances. Recent studies have shown that informally sold food products and environmental sources in these countries are contaminated

with antibiotic resistant pathogenic *E. coli* bacterial (Aijuka et al., 2018, 2015; Hoffmann et al., 2014; Ntuli et al., 2017, 2016; Rebello and Regua-Mangia, 2014) These studies suggest that both the informally sold food products as well as environmental sources near the food chain are reservoirs of potentially pathogenic and antibiotic resistant *E. coli*.

While foodborne illnesses present a public health burden in developing countries such as South Africa (Gray and Vawda, 2016), accurate documentation of these risk factors is unavailable hampering effective monitoring of the most relevant pathogens and subsequently the ability to maintain food safety and public health. Therefore, the use of molecular based diagnostic tools such as quantitative real time polymerase chain reaction (qPCR) for detection of bacterial pathogens presents an avenue for quick diagnosis as well as well-informed treatment options that may reduce morbidity and mortality. Additionally, qPCR is the most powerful tool for quantitative nucleic acids analysis (Kubista et al., 2006) that enables the simultaneous amplification and detection of specific DNA sequences with the amount of amplified product formed and monitored by fluorescence emitted by an intercalating dye or hydrolysis probe (Navarro et al., 2015). Furthermore, qPCR can be used to detect multiple virulence genes in food and environmental pathogens (Ahberg et al., 2015; Chassagnea et al., 2009; Mendes et al., 2007; Singh and Mustapha, 2014). This is because most pathogens have varying pathogenic traits making the method ideal for wide use in detection and epidemiological studies (Singh and Mustapha, 2014). Additionally, qPCR using intercalating dyes such as SYBR Green and (EvaGreen) coupled with melting curve analysis (MCA), present a fast, accurate and efficient means of detecting pathogenic characteristics associated with foodborne bacteria (Guion *et al.*, 2008).

Therefore, the objective of this study was to develop a qPCR assay using commercially available qPCR mixes based on SYBR Green 1 and EvaGreen fluorescence binding coupled with melt curve analysis. The assay was aimed at enabling the simultaneous detecting of virulence (*stx1*-Shigatoxin 1; *stx2*-shigatoxin 2; *eae*-intimin) and Cefotaxime resistant CTX-M-CTX-M) extended beta lactamase (ESBL) producing genes in *E. coli* previously isolated from food sources and irrigation water.

4.3.2. Materials and methods

4.3.2.1 Source, virulence and antibiotic resistance characterization of *E. coli* isolates

E. coli isolates used for the assay development were previously isolated from irrigation water (Aijuka et al. 2015) and irrigated lettuce (Aijuka 2014) in South Africa.

Additionally, the isolates (Table 4.3.1) had previously been characterized for possession of CTX-M ESBL (Njage and Buys, 2014) and virulence (shigatoxin; *stx1* and *stx2* and intimin; *eae*) genes (Aijuka *et al.* 2015). All isolates were stored in Tryptone Soy broth (TSB) (Biolab, Wadeville Johannesburg, South Africa) containing 30% glycerol (Saint Louis, Missouri, USA) at -80°C in the Department of Consumer and Food Sciences, University of Pretoria. Prior to use in the assay development process as well as during assay validation for sensitivity and specificity, all isolates were regrown in TSB at 37°C for 24 h. The isolates were subsequently streaked on TSB and incubated at 37°C for 24 h.

Table 4.3.1: Source and characteristics of *E. coli* isolates from food sources and irrigation water in South Africa used in developing and validating a set of two duplex qPCR assays for detection of antibiotic resistance and virulence genes

Source of isolation	Isolate code	<i>bla</i> _{CTX-M}	Virulence genes (<i>stx 1</i> , <i>stx 2</i> and <i>eae</i>)			DNA concentration (ng/μL)
			<i>stx 1</i>	<i>stx 2</i>	<i>eae</i>	
Irrigation water	MPU(W)8(3)	+	+	+	-	20.1
Irrigation water	MPU(W)9(1)	+	+	+	+	50.2
Irrigation water	NW(W)6(1)	+	+	+	-	55.0
Irrigation water	NW(W)9(1)	-	+	+	+	94.8
Irrigation water	NW(W)9(2)	-	+	+	+	95.0
Irrigation water	NW(W)5(1)2	-	-	-	+	74.1
Irrigation water	NW(W)5(3)	-	+	+	+	98.1
Irrigation water	MPU(W)5(7)	-	+	+	-	24.0
Irrigation water	MPU(W)5(1)	-	-	-	-	33.2
Irrigated lettuce	NW(V)5(3)	-	-	-	-	36.0
Irrigated lettuce	NW(V)7(3)	+	+	+	-	96.0
Irrigated lettuce	NW(V)3	+	+	+	+	99.0

*bla*_{CTX-M}- Cefotaxime resistant (CTX-M) extended beta-lactamase (ESBL) producing *E. coli*; *stx 1*-shigatoxin 1; *stx 2*-Shigatoxin 2; *eae*-Intimin;

+: positive for gene

-: negative for gene

4.3.2.2 Extraction and determination of genomic DNA purity

E. coli isolates were grown in 2mL of TSB and incubated at 37 °C for 24 h. The cultures were harvested by centrifugation (10 000 rpm for 10 min). For each isolate, resuspension of the bacterial pellet was done using 200µL of Triton™ X-100 buffer (Sigma Aldrich, St. Louis MO, USA) according to (Xue et al., 2009). The mixture was incubated at 95 °C for 30 min with agitation. The mixture was subsequently cooled to 4 °C for 10 min and centrifuged at 13000 rpm for 10 min. The purity (260/280) of DNA was measured using a Thermo Scientific Nanodrop 2000 spectrophotometer (Thermo Fisher Scientific, Wilmington DE, USA) with all isolates having a range of 1.8 to 2.0. The stock solutions for DNA were diluted (1:10) with double distilled water to give working concentrations for all isolates ranging from 20 to 100 ng/µL. The final volume of template DNA in each 20 µL closed tube reaction was adjusted for the different duplex reactions to achieve the required specificity. (Table 4.3.2).

4.3.2.3 Design of primers

The primer pairs for detection of antibiotic resistance (CTX-M ESBL) and virulence (*stx1*, *stx2* and *eae*) are shown in Table 4.3.2. The sequence homology and specificity of each primer for detection of the relevant genes was tested in silico and synthesised by Whitehead Scientific (Pty) Ltd/Integrated DNA Technologies (IDT) (Brakenfell, Cape Town, South Africa). All primer pairs were subsequently tested for specificity in amplifying the respective genes experimentally in monoplex, duplex and multiplex reactions by varying their volume (forward: reverse) in relation to one another within each single tube duplex reaction. The final concentrations of primer pairs used for the reactions are shown in Table 4.3.2.

Table 4.3.2: Genes and primer sequences used for development of a pair of duplex qPCR assays for detection of antibiotic resistance and virulence in *E. coli* isolated from food sources and irrigation water in South Africa

Gene	Primers	Primer sequence (5'-3')	Target gene product size (bp)	GC content (%)	Reference
Shigatoxin 1	<i>stx1</i> -F	TGT CGC ATA GTG GAA CCT CA	655	50	
	<i>Stx1</i> -R	TGC GCA CTG AGA AGA AGA GA		50	
Shigatoxin 2	<i>stx2</i> -F	CCA TGA CAA CGG ACA GCA GTT	477	52.4	
	<i>stx2</i> -R	TGT CGC CAG TTA TCT GAC ATT C		45.5	
Intimin	<i>eae</i> -F	CAT TAT GGA ACG GCA GAG GT	375	45.5	(Bai et al., 2010)
	<i>eae</i> -R	ACG GAT ATC GAA GCC ATT TG		45	
¹ CTX-M type-ESBLs	<i>bla_{CTX}</i> -F	ATG TGC AGY ACC AGT AAR GTK ATG GC	1354	48.1	(Hasman et al., 2005; Hendriksen et al., 2009)
	<i>bla_{CTX}</i> -R	TGG GTR AAR TAR GTS ACC AGA AYS AGC GG	1580	51.7	

1- Targets the whole group of this class of antibiotics that is categorized into 5 groups based on their amino acid sequence identity

bla_{CTX-M}- general class of Cefotaxime resistant (CTX-M) extended beta-lactamase (ESBL) producing *E. coli*

4.3.2.4 Quantitative real time PCR (qPCR)

Four commercially available qPCR mixes were evaluated for use in developing the assay. Two of the mixes contained SYBR Green 1 dye. The other two contained EvaGreen dye. These dyes have different modes of binding to double stranded DNA and thereby adoption in qPCR applications with melting curve analysis (MCA) (Jansson et al., 2017). SYBR Green 1 dye-based mixes were KAPA SYBR®Fast qPCR Kit Master mix (2X) Universal (KAPA SYBR Fast DNA polymerase, ROX reference dye, Fluorescein dye (10nM), SYBR Green I, MgCl₂) (KAPA Biosystems, Cape Town, South Africa) and SsoAdvanced™ Universal Inhibitor-Tolerant SYBR® Green Supermix (Antibody-mediated hot start Sso7d fusion polymerase, dNTPs, MgCl₂, SYBR Green 1, enhancers, stabilizers and blend of reference dyes (ROX and Fluorescein)) (Bio-Rad Laboratories Inc. Hercules CA, USA). EvaGreen based mixes include KAPA HRM Fast PCR Kit (EvaGreen, Taq DNA polymerase) (KAPA Biosystems) and Precision melt Supermix (Hot start iTaq™ DNA polymerase, dNTPs, MgCl₂, EvaGreen dye, enhancers and stabilizers) (Bio-Rad Laboratories Inc). Final quantities and concentrations of optimized reagents and consumables in each 20µL closed tube reaction are shown in Table 4.3.3. The harmonized protocols for each qPCR super mix used are shown in Table 4.3.3. All qPCR reactions were carried out in 0.2mL PCR tubes with domed caps high profile clear #twi0201 (Bio-Rad Laboratories Inc). The optimised PCR reactions and subsequent MCA were carried out using the CFX96 Touch™ Real Time PCR detection system (Bio-Rad Laboratories Inc) and fluorescent analysis in the FAM/SYBR Green 1 Channel.

4.3.2.5 The limit of detection (LOD) and specificity of the assay

To determine the limit of detection (LOD), serial dilution (10⁻¹ to 10⁻⁸) of DNA for all test isolates were made and used to run a qPCR reaction with the commercial mixes used in the final duplex assays. A plot of the quantification cycle (C_q) versus log starting quantity of DNA was generated after the qPCR step. The LOD for each duplex assay was expressed in terms of detection per log DNA concentration. The plot curve (C_q v/s Log starting DNA) and standard curve were calculated using the CFX Manager™ software Industrial Diagnostics Version 2.0 (Bio-Rad Laboratories Inc). Determination of the analytical specificity of the assay was done with using the 12 isolates (Table 4.3.1).

4.3.2.6 Data analysis

All qPCR data including means and standard deviations for the duplex reactions including melting temperatures (T_m) and C_q values were determined using the CFX Manager™ software Industrial Diagnostics Version 2.0 (Bio-Rad Laboratories Inc). Additionally, all melting curves were determined using the same software. A student T-test was carried out using Microsoft Excel 2016 (Microsoft Corporation, Redmond, WA USA) to compare T_m 's of targeted amplicons (genes) in monoplex and duplex reactions at 95% confidence level.

Table 4.3.3: Optimized reagents and consumables used in a pair of duplex qPCR assays for detection antibiotic and of virulence genes in *E. coli* isolated from food sources and irrigation water in South Africa.

qPCR mix	Gene	DNA (μ l) ¹	Primer		MgCl ₂ (mM)
			Concentration (nM)		
			Forward	Reverse	
KAPA HRM Fast PCR Kit	<i>stx1</i>	0.5	500	250	1.5
	<i>stx2</i>	1.0	250	100	1.0
	<i>eae</i>	0.5	100	250	1.0
	<i>bla_{CTX-M}</i>	2.0	250	250	2.0
Precision melt supermix	<i>stx1</i>	0.5	500	250	-
	<i>stx2</i>	0.5	100	100	-
	<i>eae</i>	0.5	100	250	-
	<i>bla_{CTX-M}</i>	2.0	500	100	-
KAPA SYBR®Fast qPCR Kit	<i>stx1</i>	2.0	500	250	-
	<i>stx2</i>	2.0	250	100	-
	<i>eae</i>	2.0	250	250	-
Master mix (2X) Universal	<i>bla_{CTX-M}</i>	2.0	250	250	-
SsoAdvanced™ Universal	<i>stx1</i>	1.0	500	250	-
Inhibitor-Tolerant SYBR® Green	<i>stx2</i>	0.5	250	100	-
Supermix	<i>eae</i>	1.0	100	250	-
	<i>bla_{CTX-M}</i>	1.5	250	250	-

qPCR-Quantitative real time polymerase chain reaction

¹; quantity of DNA in 20ul reaction tube. Concentration of DNA ranged from 50 to 500 μ g/ml

-; MgCl₂ was contained in the qPCR reaction mix

stx1- shigatoxin 1; *stx2*- shigatoxin 2; *eae*- intimin; *bla_{CTX-M}*- general class of Cefotaxime resistant (CTX-M) extended beta-lactamase (ESBL) producing *E. coli*

Table 4.3.4: Thermocycling conditions for a pair of duplex qPCR assays based on different commercially available qPCR super mixes for detection of virulence (shigatoxin 1, shigatoxin 2, intimin) and Cefotaxime resistant (CTX-M) extended beta lactamase (ESBL) genes in *E. coli* isolated from food sources and irrigation water in South Africa

qPCR mix	Cycling step	Temperature (°C)	Duration	Cycles
KAPA HRM Fast PCR Kit	Enzyme activation	95	3 minutes	hold
	Denaturation	95	5 seconds	
	Annealing/extension	55 to 57.5 (Gradient)	30 seconds	45
Precision melt supermix	Denaturation	95	1 minute	hold
	High Resolution Melt	65 to 90°C (in 0.01/s increments)	1 second/step	N/A
	Initial DNA denaturation	95	2 minutes	hold
	Denaturation	95	10 seconds	
KAPA SYBR®Fast	Annealing/extension (+ plate read)	55 to 58 (Gradient)	30 seconds	
	Extension + plate read (optional)	72	30 seconds	45
	Heteroduplex formation	95	30 seconds	hold
SsoAdvanced™	High Resolution Melt + plate read	60	1 minute	hold
	Enzyme activation	65 to 90°C (in 0.01/s increments)	1 second/step	hold
	Denaturation	95	3 minutes	hold
Universal Inhibitor-	Denaturation	95	3 seconds	
	Annealing/extension	55 to 58 (Gradient)	30 seconds	40
	Melt curve analysis (+plate read)	60 to 90°C (in 0.01/s)	1 second/step	hold
Universal Inhibitor-	Polymerase activation and DNA denaturation	98°C	3 minutes	hold
	Denaturation	98°C	15 seconds	hold
	Annealing/extension and plate read	55 to 58 (Gradient)	60 seconds	40
	Melt curve analysis (+plate read)	65 to 95°C (in 0.01/s increments)	1 second/step	hold

4.3.3. Results

4.3.3.1 Optimization of qPCR conditions for detection of antibiotic resistance and virulence genes in *E. coli* isolates

We optimized the reagents, consumables and qPCR thermocycling and subsequent melt curve conditions (Tables 4.3.3 and 4.3.4). The optimization of these variables included varying the volume of DNA, primer pairs, MgCl₂ (only for Kapa HRM fast PCR mix for which it came separate) for each 20 μ L closed tube reaction. The thermocycling conditions were adjusted (using a gradient temperature range) to get the best annealing temperatures for all targeted amplicons for each duplex. To target all amplicons at their T_m during MCA, the slowest ramp rate (with 0.01^oC/s increments) was used.

4.3.3.2 Evaluation of a SYBR Green I and EvaGreen based qPCR commercial mix for development of a duplex pair of qPCR assays for detection of antibiotic resistance and virulence genes in *E. coli*

Two qPCR mixes, A (SsoAdvancedTM Universal Inhibitor-Tolerant SYBR® Green Supermix containing SYBR Green I dye and B (KAPA HRM Fast PCR mix) provided the most reproducible results (T_m and C_q values) for detection of the target genes. Mixes A and B were only able to provide ample T_m separation for each target gene within a duplex reaction (Table 4.3.5). The selection of the genes for duplexing was based on size (bp) of the amplified product with each pair of targeted genes consisting of a larger and smaller amplicon. The two genes detected in each reaction duplex reaction had ample mean T_m separation based on MCA. The first duplex reaction from both qPCR mixes targeted *stx1* and *bla_{CTX-M}* and while the second, *stx2* and *eae*.

4.3.3.3 Melt curve analysis using a SYBR Green I dye based qPCR mix

When qPCR mix A was used, the mean T_m separation between genes *stx 1* (80.5^oC) and *bla_{CTX-M}* (85.85^oC) was 5.35^oC in monoplex reactions. Similarly, a sufficiently large T_m separation of 4.4 ^oC was noticed between genes *stx 2* (86.8^oC) and *eae* (82.4^oC). Therefore, based on this sufficiently distinguishing T_m separation, the gene pairs were used to design the duplex assay. The duplex single tube reaction yielded melting peaks with T_m's similar (not statistically different p \leq 0.05) to monoplex reactions (Figures 4.3.1 and 4.3.2).

The volume of DNA used in each reaction for monoplex and duplex reactions varied depending on the targeted gene and reaction (monoplex or duplex) (Table 4.3.5).

4.3.3.4 Melt curve analysis using an EvaGreen dye-based qPCR mix

On the other hand, when qPCR mix B was used for duplexing, a mean T_m difference of 7.85°C was observed between genes *Stx 1* (79.85°C) and *bla_{CTX-M}* (87.7°C). Additionally, a similarly large mean T_m difference of 7.5°C was noticed between genes *stx 2* (88.15°C) and *eaec* (80.65°C). Using mixture B containing the EvaGreen dye (7.68°C) as opposed to mixture A having SYBR Green dye 1 (4.88°C) provided a larger mean T_m difference between amplicons suggesting a larger range for amplicon detection. This larger mean T_m separation observed between amplified genes when using mixture B allowed simultaneous detection of the target genes with similar T_m as in the monoplex reactions (Figures 4.3.3 and 4.3.4). The melting curve peaks generated using the EvaGreen dye-based qPCR mix for both monoplex and duplex reactions were smoother and generated more fluorescence than those with the SYBR Green I dye mix. Additionally, the volume of DNA used for each reaction in monoplex and duplex when using mixture B did not vary too much (Table 4.3.5) compared with when mixture A was used.

4.3.3.5 The limit of detection and analytical specificity of duplex assays based on qPCR mixes containing SYBR Green 1 and Eva Green 1 dyes

The limit of detection for template DNA in both assays was 0.02ng with a linear dynamic range of 0.02 ng to 100 ng of template. The C_q values for amplification of template DNA over this range are shown in Figure 5 while the resulting standard curve R^2 and slope of the standard curve is shown in Figure 6. Both assays were able to detect all the relevant genes from each of the test strains previously used for development (Table 4.3.1).

Table 4.3.5: Template volumes and melting temperatures of amplicons in monoplex and duplex qPCR assays for detection of antibiotic resistant and virulence genes in *E. coli* isolated from food sources and irrigation in South Africa using melt curve analysis

Quantitative real time PCR super mix				
Monoplex/ Duplex reaction	SsoAdvanced™ Universal Inhibitor-Tolerant SYBR® Green Supermix		KAPA HRM Fast PCR Kit	
Targeted gene (s)	DNA quantity (µL)	Melting Temperature (°C)	DNA quantity (µL)	Melting Temperature (°C)
<i>eae</i>	1.5	82.4±0.07	1.5	80.65±0.07
<i>stx1</i>	0.9	80.5±0.07	1.5	79.85±0.07
<i>bla_{CTX-M}</i>	0.5	85.85±0.07	1.5	87.7±0.21
<i>stx2</i>	0.5	86.8±0.02	1.5	88.15±0.07
<i>stx1</i> and <i>bla_{CTX-M}</i>	1	80.4±0.07 and 84.40±0.01	1	79.35±0.07 and 87.50±0.07
<i>eae</i> and <i>stx2</i>	0.5	82.8±0.07 and 85.70±0.07	0.5	81.10±0.07 and 87.80±0.07

Targeted genes: *stx1*-shigatoxin 1; *stx2*-shigatoxin 2; *eae*- intimin; *bla_{CTX-M}*- general class of Cefotaxime resistant extended beta-lactamase (ESBL) producing *E. coli*

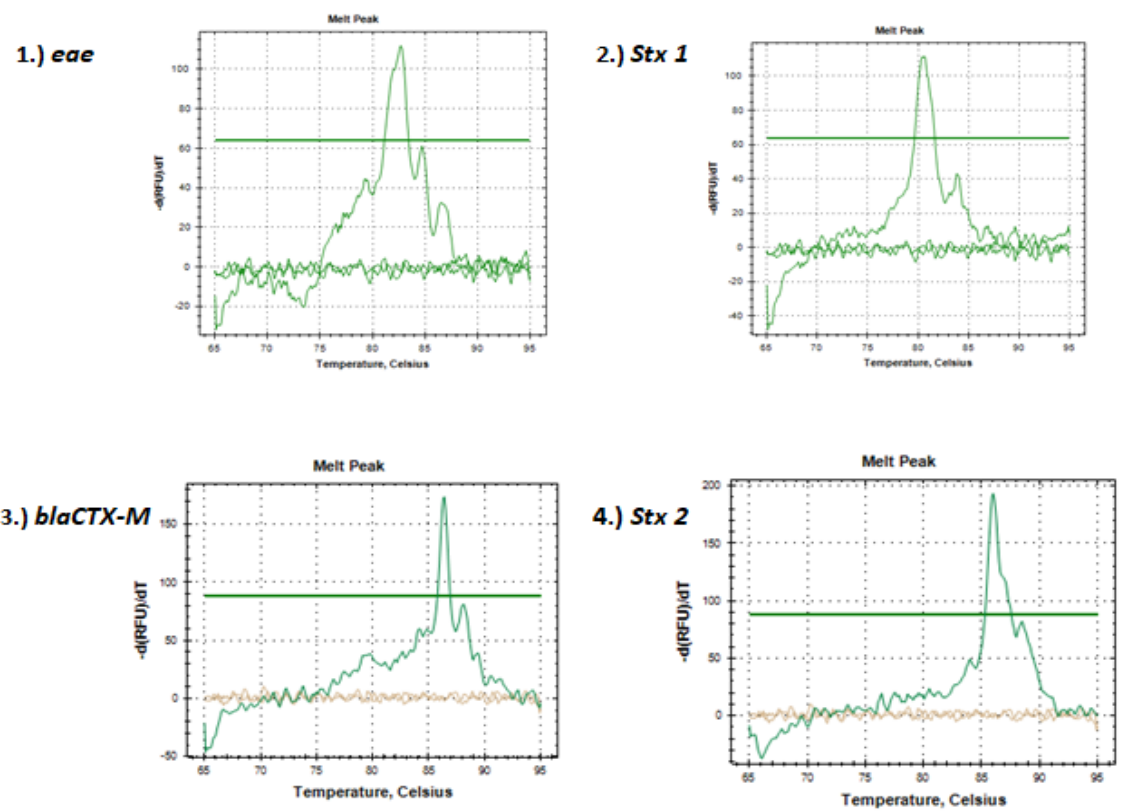
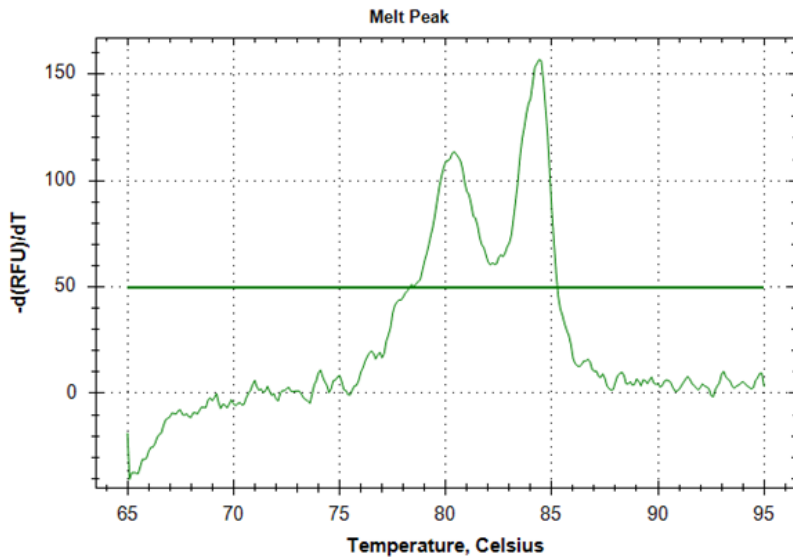
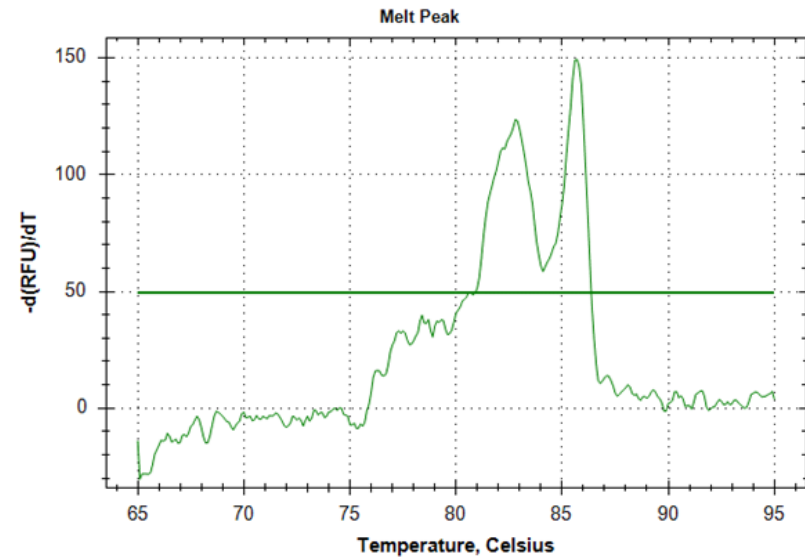


Figure 4.3.1: Melting peak curves showing antibiotic resistant (*bla_{CTX-M}*) and virulence (*eae*, *stx1*, *stx2*) genes in *E. coli* isolated from food sources and irrigation water determined using monoplex qPCR assays with a commercial qPCR mix containing SYBR Green I dye. The horizontal line depicts the threshold for detection. *stx1*-shigatoxin 1; *stx2*-shigatoxin 2; *eae*- intimin; *bla_{CTX-M}*- general class of Cefotaxime resistant extended beta-lactamase (ESBL) producing *E. coli*. QPCR mix was SsoAdvanced™ Universal Inhibitor-Tolerant SYBR® Green Supermix. The peaks were determined by plotting the negative derivative of change in fluorescence increase over time (-d(RFU)\dt) with increase in temperature.



1.) *Stx 1* and *blaCTX-M*



2.) *Stx 2* and *eae*

Figure 4.3.2: Melting peak curves showing antibiotic resistant (*bla_{CTX-M}*) and virulence (*eae*, *stx1*, *stx2*) genes in *E. coli* isolated from food sources and irrigation water in South Africa determined using a pair of duplex qPCR assays from a commercial qPCR mix containing SYBR Green I dye. The horizontal line depicts the threshold for detection. *stx1*-shigatoxin 1; *stx2*-shigatoxin 2; *eae*- intimin; *bla_{CTX-M}*- general class of Cefotaxime resistant extended beta-lactamase (ESBL) producing *E. coli*. QPCR mix was SsoAdvanced™ Universal Inhibitor-Tolerant SYBR® Green Supermix. The peaks were determined by plotting the negative derivative of change in fluorescence over time ($-d(RFU)/dt$) with increase in temperature.

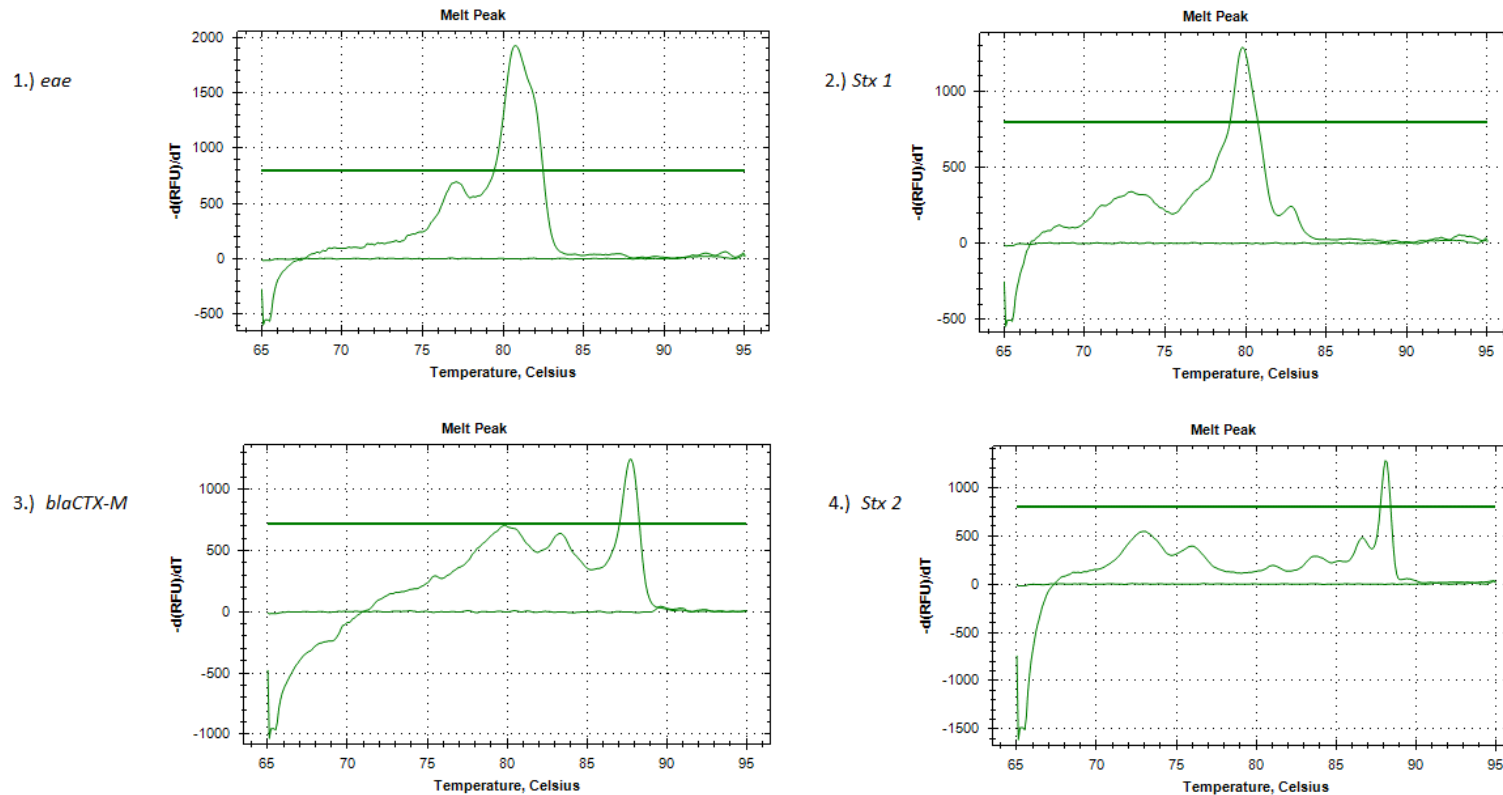
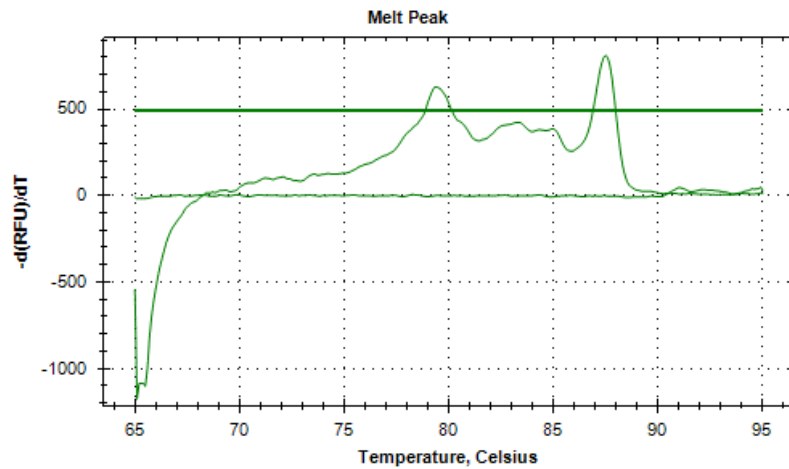
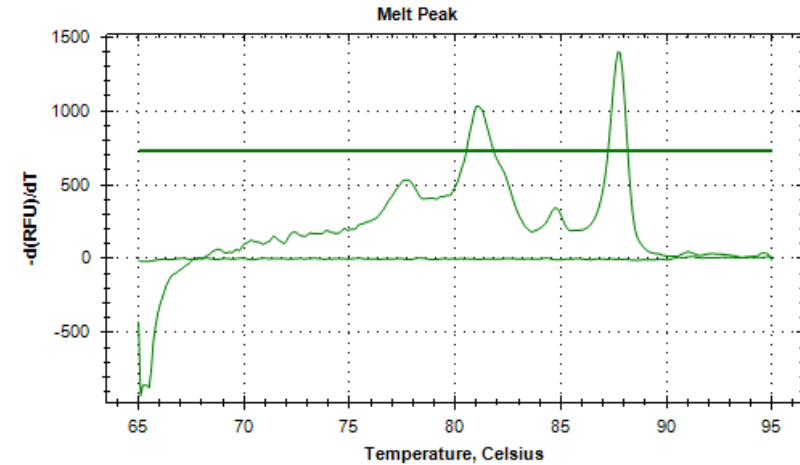


Figure 4.3.3: Melting peak curves showing targeted antibiotic resistant (*bla_{CTX-M}*) and virulence (*eae*, *stx1*, *stx2*) genes in *E. coli* isolated from food sources and irrigation water determined with multiplex qPCR assays using a commercial qPCR mix containing EvaGreen dye. The horizontal line depicts the threshold for detection. *stx1*-shigatoxin 1; *stx2*-shigatoxin 2; *eae*- intimin; *bla_{CTX-M}*- general class of Cefotaxime resistant extended beta-lactamase (ESBL) producing *E. coli*. The QPCR mix was KAPA HRM Fast PCR mix. The peaks were determined by plotting the negative derivative of change in fluorescence over time ($-d(RFU)/dt$) with increase in temperature.



1.) *Stx 1* and *blaCTX-M*



2.) *Stx 2* and *eae*

Figure 4.3.4: Melting peak curves showing antibiotic resistant (*blaCTX-M*) and virulence (*eae*, *stx1*, *stx2*) genes in *E. coli* isolated from food sources and irrigation water in South Africa determined using a pair of duplex qPCR assays with a commercial qPCR mix containing EvaGreen dye. The horizontal line depicts the threshold for detection. *stx1*-shigatoxin 1; *stx2*-shigatoxin 2; *eae*- intimin; *blaCTX-M*- general class of Cefotaxime resistant extended beta-lactamase (ESBL) producing *E. coli*. The QPCR mix was KAPA HRM Fast PCR mix. The peaks were determined by plotting the negative derivative of change in fluorescence over time ($-d(RFU)/dt$) with increase in temperature.

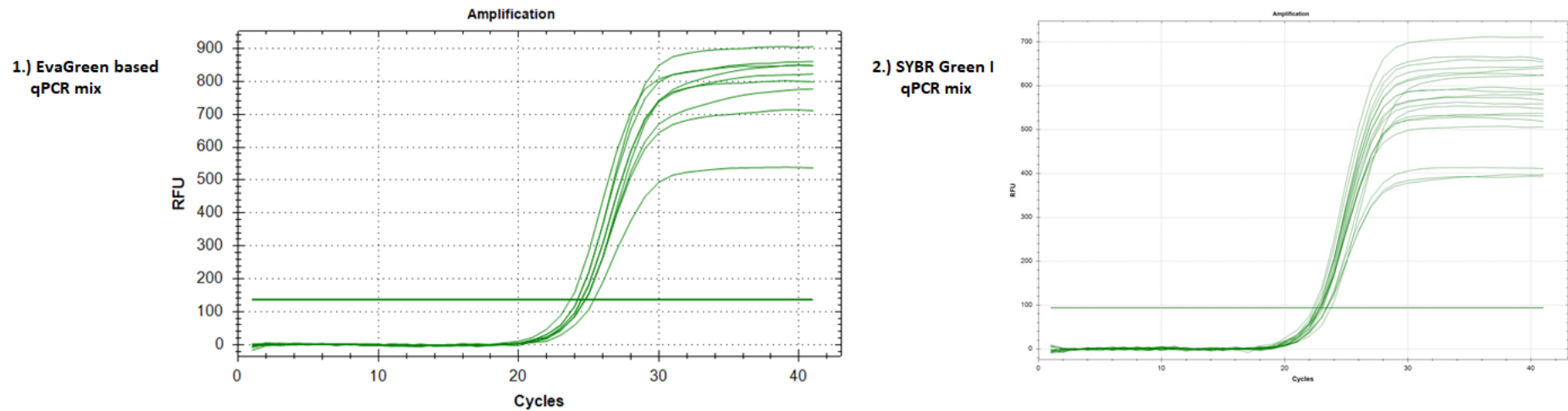


Figure 4.3.5: QPCR duplex reactions showing amplification (fluorescence increase v/s quantification cycle (C_q)) of DNA dilution series (10^0 to 10^{-8}) using two commercially available qPCR mixes based on EvaGreen and SYBR Green I dyes. EvaGreen based qPCR mix- KAPA HRM Fast PCR mix. SYBR Green I dye based qPCR mix- SsoAdvancedTM Universal Inhibitor-Tolerant SYBR[®] Green Supermix. Data from these graphs was used to draw determine the limit of detection for each assay by drawing standard curves.

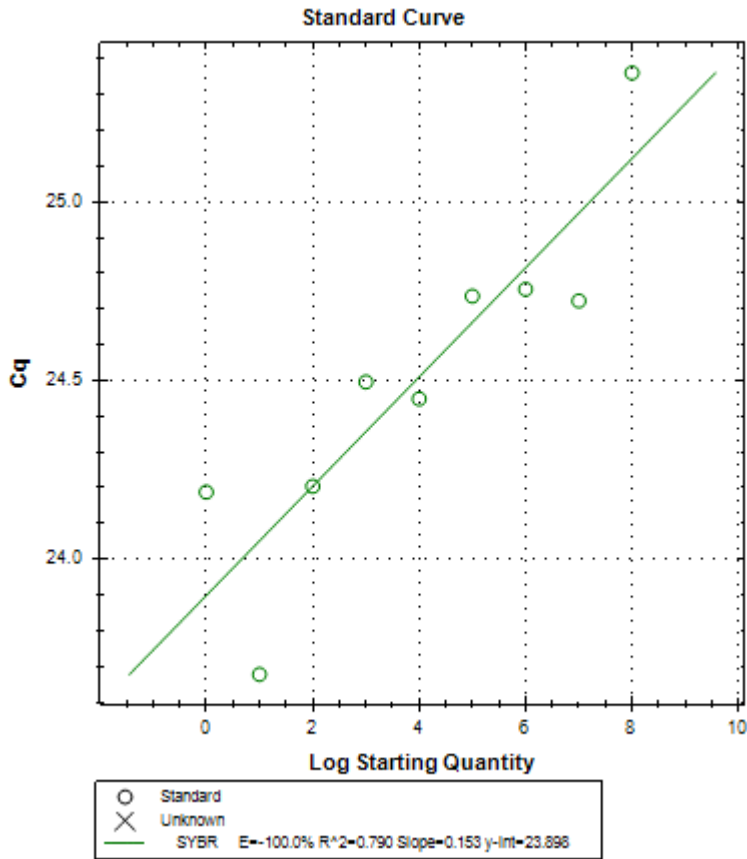


Figure 4.3.6: Standard curve derived from plotting the change in quantification cycle (C_q) with the quantity of serially diluted (10^0 to 10^8) template DNA for a pair of 2 duplex assays each based on commercial qPCR mixes containing SYBR Green I or EvaGreen dyes used in detection of antibiotic resistance and virulence genes in *E. coli* isolated from food sources and irrigation water in South Africa.

4.3.4 Discussion

The differences observed in specificity of the two assays from each commercial qPCR mix in terms of DNA template, smoothness of the melt curves and mean difference in T_m separation of the targeted amplicons may be related to the DNA binding mechanism of each dye. SYBR Green is an asymmetric cyanine compound that binds to the minor groove of DNA leading to increase in formation of non-specific products thereby limiting its applicability in real time PCR over a broad range of temperatures (Gudnason et al., 2007). Additionally, it has a high affinity for G-C rich amplicons regardless of size that might interfere with the PCR reaction process (Gudnason et al., 2007). On the other hand, the symmetrical cyanine structure of EvaGreen may aid in formation of fewer non-specific products since the binding of double stranded (dsDNA) occurs evenly over the whole molecule. This suggested scenario has been observed in saturating dyes of the SYTO family that showed a broader range of melting temperatures based on their saturation ability once bound to dsDNA providing suitable conditions for qPCR multiplexing compared to SYBR Green based dyes (Eischeid, 2011; Monis et al., 2005).

The robustness of qPCR coupled with MCA in diagnostic testing makes the above drawbacks easy to circumvent. In this study, we optimised the template concentration, ramp rate of the T_m acquisition step, primer pair and $MgCl_2$ volumes within each duplex closed tube reaction. Several studies have shown the robustness of qPCR in detection of amplicons using different strategies. For instance, a low scanning rate for intercalating dyes helps to control PCR inhibition since every system (thermocycler) exhibits a delay between temperature at the sensor and the sample, therefore such a rate will precisely determine T_m 's (Ahberg et al., 2015). Additionally, when using commercial qPCR mixes, the dye concentration is constant, therefore varying primer pair concentration can provide optimization of the assay (Monis et al., 2005). Furthermore, the assays were able to simultaneously detect 6 Metallo-beta-lactamase encoding genes with amplicons ranging from 72 to 1449 bp in different *Enterobacteriaceae* species through optimization of primer pair concentration, sequence (designed to have 1°C T_m separation) and volume (total reaction volume 48 μ L). Similarly, Ahberg et al. (2015) used MCA to design a qPCR assay that quantified Avian flu at each melting scan step rather than at the end of the reaction.

These studies depict how the robustness of qPCR can enable tailor-made solutions each investigator or laboratory. Commercial qPCR mixes provide standardization of many of the proposed assay reagents and consumables, therefore an investigator only needs to vary a few conditions to develop assays that are both specific and valid for routine and research diagnostic laboratories where savings on time and resources are critical.

4.3.5. Conclusion

We have developed a pair of duplex qPCR assays based on MCA using commercially available mixes based on two intercalating dyes, SYBR Green I and EvaGreen for detection of antibiotic resistance and virulence genes in *Escherichia coli* isolated from food sources and irrigation water. The simple optimization of a few reagents and conditions for the assay highlights its robustness for diagnostic application in a food safety and clinical health. This is a crucial step in providing fast, accurate and economical emerging pathogen monitoring programs.

5. General Discussion

5.1 Critique of methodology

Like all scientific studies, this study also encountered some shortcoming in its findings. These shortcomings are explored below and correspond to the three research sections of the study.

In the first research section aimed at characterizing *E. coli* isolated from food sources and irrigation water in South Africa based on the DEC pathotype, the study may have been influenced by the method for analysis. For example, molecular identification using virulence genes yielded few positive targets while the subsequent use of cell adherence assays with epithelial cells (after many strains had shown poor results with PCR) provided better scrutiny and identification. This raises the question of how many previous studies that used virulence genes underestimated the prevalence of these pathogens within the food and environmental setting? In this study initial characterization was done with PCR and only after repeated poor resolution of some strains was cell adherence used to realize the practical significance of the specific pathotype.

The difficulty of getting more representative characterization arises from the heterogeneity inherent in many of these pathotypes such as EAEC, ETEC and DAEC which hinders more representative identification (Croxen et al., 2013; Leimbach et al., 2013). On the other hand, the use of cell adherence assays is cumbersome, costly and requires highly trained personnel making its adoption in routine diagnostic laboratories problematic. This scenario is especially an issue in developing countries like South Africa that are resource constrained. However, the popularity of WGS as an analytical method coupled with the reduced capacity for analysis presents a great leap forward for studying the genetic variation among strains of the same species (Ronholm et al., 2016). This holds a lot of promise in identifying genetic markers for differentiating closely related clades of strains such as DEC pathotypes since early methods were based on partial genetic information and thereby less accurate compared to WGS.

In the second research section, the study investigated whether non-DEC strains isolated from food and irrigation water sources in South Africa were linked with phenotypes associated with intestinal dysfunction. Intestinal dysfunction (ID) is associated with a change in the “normal” relative abundance of bacterial taxa in the gastrointestinal tract (GIT) and their associated metabolic functions within the human intestinal (Shin et al., 2015). The conditions ultimately lead to a breakdown in homeostasis and as well as the normal synergistic relationship between

the human body and its microbial constituents (Shin et al., 2015). ID is linked to reduced mucosal immunity and a subsequent increased risk of intestinal diseases causing reduced dietary nutrient intake and utilization in young infants. This results into growth faltering and neurocognitive disorders later in life (Faith et al., 2015). Additionally, ID maybe a precursor to diarrhoea caused by DEC based on symptoms previously observed in malnourished children and individuals with irritable bowel syndrome (IBS) (Faith et al., 2015). The mechanism by which the symptoms resulting from ID are caused are still being elucidated and the clearest indicator is usually abundance of different taxa such as *Enterobacteriaceae*. Recent studies had suggested that a bloom of *E. coli* within the human GIT predisposes individuals to intestinal disease (Shin et al., 2015) and thereby foodborne *E. coli* provided a potential source of infection. Therefore, as a preliminary study using non-DEC we sought to use phenotypes (*in-vitro*) associated with DEC as a proxy for measuring their ability of non-DEC to withstand and bloom within the GIT potentially leading to ID through distorting the healthy microbial population within the GIT. The resulting data showed that some non-DEC strains may have the capacity to produce cytotoxins, induce the pro-inflammatory cytokine IL-8 and disrupt epithelial cell barriers. These conditions may ultimately lead to destabilization of the normal gut community subsequently predisposing the host to intestinal disease such as DEC infection.

However, the small number of isolates used in the study made it hard to draw more accurate conclusions related to the risk to food safety and public health. Therefore, the use of a larger number of strains might have helped provide a more direct link between non-DEC from food and irrigation water in South Africa and the risk of intestinal dysfunction. Additionally, to effectively link a phenotypic phenomenon associated with a given immunological response, it is crucial to understand the bacterial surface properties of a given species. All strains from this study were sourced from food and environmental sources and thereby had not been characterized for bacterial surface properties such as adhesins and outer membrane proteins, factors that are known to influence how bacteria interact with epithelial cell (and the associated immunological response) and abiotic surrounding.

In the last research chapter, the study developed a duplex qPCR assay for simultaneous detection of cefotaxime-resistant (CTX-M) beta lactamase enzyme producing extended beta lactamase (ESBL) and virulence genes in *E. coli* previously isolated from food sources and irrigation water which might been limited by the following factors.

QPCR using melting curve analysis can be limited by the size of amplicons. This may be ascribed to competition for the dye binding to double stranded DNA releasing fluorescence, an indication of amplification and detection of targeted amplicon. This makes it a limiting factor for continued progress of a given reaction (Monis et al., 2005). In the study we used amplicons of sizes ranging from 300 to 1500bp. Although such limitations can be overcome as we mentioned in chapter 3, redesign of primers to fit smaller sizes (50 to 200bp) might have provided a faster optimization process for the assay. It could also have provided improved detection of all 4 genes (*bla_{CTX-M}*, *stx1*, *stx2* and *eae*) within a single closed tube reaction as opposed to the duplex reaction using the reported two closed tube reactions.

Additionally, the use of a larger panel of strains for determining the specificity of the developed assay for validation would have increased its reliability. This is because molecular detection assays though accurate can also be prone to false-positives if poorly designed thereby reducing the specificity and wide scale adoption of the method as a routine monitoring and diagnostic tool.

5.2 Discussion

This study was initiated based on the backdrop of reports about the increased prevalence of diarrhea among infants and immune compromised adults in South Africa. Based on past studies and observations that most diarrheal disease is associated with the fecal-oral route, this study sought to characterize previously isolated *E. coli* strains from different food sources and irrigation water collected over 10 years from varying geographical regions of South Africa.

Observations from such a study were envisaged to provide more reliable information regarding the DEC pathotypes commonly associated with food sources and irrigation water and providing a direct link to diarrheal illness. To the best of the author's knowledge, this is the first original contribution to carry out such a link using such diverse (ecological, temporal and geographic) isolates in South Africa.

This study reports that EAEC was the most prevalent DEC pathotype from all sources and that they were presumably associated with the observed diarrheal illness. EAEC has been associated with sporadic and outbreak infections across the world in both developing and developed countries and may, according to our results be a leading cause of diarrheal infection in South Africa. Additionally, it has previously been suggested that EAEC has emerged separately in different parts of the world leading to multiple lineages that are genotypically diverse (Chattaway et al., 2014).

This observation can be corroborated by our findings although all of the strains in our studies were derived from food and environmental sources unlike in the study by (Chattaway et al., 2014) that used clinical strains. Future comparative studies should help elucidate differences among the strains based on sources of isolation.

The inherent heterogeneity of the pathotype facilitated by possession of a diverse suite of virulence genes complicates its identification but might also aid in the ability to survive outside the human host within the open environment. Interestingly the genetic heterogeneity in EAEC provided a means for separating strains based on their sources of isolation (food, water, clinical) suggesting an application for source tracking. Virulence genes carried over by horizontal transfer are reflective of genetic history which can also be driven by environmental niches (Dini-Andreote et al., 2012; Touchon et al., 2017). Further exploration of this potential application may assist in validating the usefulness of this resource when compared to presently used source tracking methods such as MLST and Rep-PCR that enable detection of potential sources of EAEC contamination along the food chain.

This study shows that the pathotype has got the capacity to survive within the food and agricultural environment presenting a viable risk to food safety and by association public health especially among vulnerable groups such as infants and immune compromised adults. Limited information is available regarding the prevalence of EAEC within the food and agricultural environment around the world, although the outbreak associated with bean sprouts in 2011 in Germany caused by the EAEC Shigatoxin producing strain O104:H4 (Rasko et al., 2011) prompted studies into environmental persistence. These studies have shown the AAF/I (Nagy et al., 2016) are used for attachment and biofilm formation within this strain as well as its ability to form colanic acid that was shown to help persistence of the organisms in sprouts (Borgersen et al., 2018). Just like the past outbreaks associated with the common foodborne pathogen *E. coli* O157:H7 led to investigations into how it survives within the agricultural and food production setting, similar studies with O104:H4 should elucidate more about EAEC related strains. Although, as previously mentioned, its (EAEC) inherent genotypic heterogeneity might not be as straight forward as with O157:H7. For instance, EAEC strains isolated from animals were found to not carry the same virulence factors common to human pathogens although such a study might have only reflected conditions within a given geographical setting (Uber et al., 2006). Therefore, further characterization of these strains should shed more light on the mechanisms used by pathogenic foodborne and environmental EAEC to navigate a hostile environment such as the agricultural and food production setting. Additionally, comparison of

these strains with outbreak pathogens and clinical strains previously implicated in disease should be able to provide more insight relating to potential sources of contamination.

This study employed a novel approach of comparing the capacity of previously isolated non-DEC (that would usually be regarded as safe) to illicit phenotypes associated with intestinal dysfunction. By doing this, the study shows that many potentially pathogenic *E. coli* strains can end up in food sources causing disease even when routine microbiological testing clears them from being a health risk. This is because most molecular tests only target genes commonly shared with the outbreak pathogens and neglect those which are distantly related but also pathogenic. The heterogeneity of DEC pathotypes could be attributed to this (Croxen et al., 2013; Leimbach et al., 2013). Therefore, the high diarrheal disease burden may in part be also attributed to non-DEC strains that fall out of the detection range of currently validated methods of identification. However, further elucidation of this initial finding will require use of more strains providing more data points to clearly illustrate associations between strains.

Lastly, this study showed that the advent of a rapid and accurate molecular identification method such as qPCR is a great tool for developing tailor-made assays for use in characterizing bacterial pathogens within a given laboratory. Using commercially available qPCR reaction mixes containing intercalating dyes, this study was able to develop a rapid and accurate duplex qPCR assay based on melt curve analysis for the simultaneous detection of antibiotic resistance (*CTX-M* ESBL) and virulence (*stx1*, *stx2* and *eae*) genes in *E. coli* strains isolated from food sources and irrigation water in South Africa. Optimization of the assay involved adjustment of a few parameters on the thermocycler as well as reagents within the reactions, an activity that does not require highly specialized training in comparison to many similarly effective and rapid diagnostic assays. Such a scenario depicts how these molecular tools can be adapted for uses based on a researcher's need for providing cost effective alternatives that would have previously been unavailable. Such assays are increasingly crucial in the era of increasing antibiotic resistance among bacterial pathogens.

6. Conclusions and Recommendations

6.1 Conclusions

EAEC is the predominant DEC pathotype associated with food sources and irrigation water in South Africa and maybe associated with the high diarrheal disease pattern associated especially among vulnerable groups such as infants and immune compromised adults. Additionally, the phylogenetic characterization of EAEC strains based on the presence of virulence gene markers is a potential source tracking tool.

Non-DEC isolated from food sources and irrigation water in South Africa has the potential to cause phenotypes associated with intestinal dysfunction and thereby might in part be associated with the diarrheal disease burden noticed in the country. However, further elucidation of these associations using a large set of strains as well as by characterizing the strains should provide a clearer link of the risk to food safety and public health of these organisms.

QPCR coupled with melting curve analysis presents a rapid, efficient, accurate and cost-effective method for characterizing DEC and non-DEC strains isolated from food sources and irrigation water. Such an analytical tool would be highly beneficial for use by food safety and health practitioners carrying out routine diagnostics for academic research. Such a tool can prove vital especially in the era of emerging foodborne and environmental pathogens that threaten the food chain and general health well being

6.2 Recommendations

The following recommendations are suggested to help supplement this work. These studies will add to the body of knowledge linking foodborne and environmental DEC to food safety and health.

The use of whole genome sequencing is now considered the gold standard for identifying bacteria. Using this method for characterizing heterogeneous pathogens such as DEC will provide more insight into relationships at different genome levels ranging from virulence factors and metabolic capacity to phylogeny. Additionally, this method provides the advantage of comparing recently sequenced strains to others in data bases around the world, providing ideal comparison based on geography and source of isolation, two factors that greatly influence the emergence of pathogens such as DEC.

The use of metagenomics for analyzing samples from food and environmental sources is another relevant analytical tool. Just as suggested with WGS, the reducing cost of both these applications provides enormous advantages compared to single-strain-isolation-and-comparison. With metagenomics, a researcher can identify whole communities of species and their associated higher taxonomic affiliations. Such information is crucial in understanding which communities of bacteria are associated with higher prevalence of DEC pathotypes providing more useful information regarding mitigation measures. Additionally, this method provides the ability to study the functional flow of genes associated with these communities. However, as this method gains wider popularity in routine monitoring of pathogenic bacterial populations, it presents some limitations such as its limited sensitivity. This is usually because of the lower quality of genetic material leading to generation shorter and less accurate reads of the genomes making up the sampled bacterial community. Such a drawback is crucial as it can hinder the ability to ably detect contaminated food and environmental sources during routine analysis by food safety and public health officials. Pathogens such as *E. coli* O157:H7 with a low infective dose are particularly of concern here since they have previously been associated with large scale outbreaks.

The use of more strains from food and environmental sources for analyses aimed at linking non-DEC to intestinal dysfunction is crucial because it will provide more data points for analyzing the associated phenotypes. Additionally, by having more data points, different associations among the variables will be clearly seen, something that this work was not able to adequately depict.

The further characterization of all the DEC and non-DEC strains used in this study especially relating to bacterial surface properties. This characterization especially involving adhesins and transports systems will provide greater insight into how they contribute to their ability to survive in the environment. Additionally, it will provide a better understanding of bacterial attachment properties and how such strains compare (genotypically and phenotypically) with characterized foodborne outbreak pathogens and clinical isolates previously implicated in gastrointestinal disease.

This information can prove vital since it can assist in unravelling the mechanism underlying persistence of these strains in the environment providing better assessment of the risk posed to food safety and public health.

7. References

- Acosta, G.J., Vigo, N.I., Durand, D., Riveros, M., Arango, S., Zambruni, M., Ochoa, T.J., 2016. Diarrheagenic *Escherichia coli*: Prevalence and Pathotype Distribution in Children from Peruvian Rural Communities. *Am. J. Trop. Med. Hyg.* 95, 574–579. <https://doi.org/10.4269/ajtmh.16-0220>
- Adefisoye, M.A., Okoh, A.I., 2016. Identification and antimicrobial resistance prevalence of pathogenic *Escherichia coli* strains from treated wastewater effluents in Eastern Cape, South Africa. *Microbiologyopen* 5, 143–151. <https://doi.org/10.1002/mbo3.319>
- Ahberg, C.D., Manz, A., Neuzil, P., 2015. Single fluorescence channel-based multiplex detection of avian influenza virus by quantitative PCR with intercalating dye. *Sci. Rep.* 5, 1–8. <https://doi.org/10.1038/srep11479>
- Ahmed, D., Islam, M.S., Begum, Y.A., Janzon, A., Qadri, F., Sjöling, Å., 2013. Presence of enterotoxigenic *Escherichia coli* in biofilms formed in water containers in poor households coincides with epidemic seasons in Dhaka. *J. Appl. Microbiol.* 114, 1223–1229. <https://doi.org/10.1111/jam.12109>
- Aijuka, M., Charimba, G., Hugo, C.J., Buys, E.M., 2015. Characterization of bacterial pathogens in rural and urban irrigation water. *J. Water Health* 13, 103–117. <https://doi.org/10.2166/wh.2014.228>
- Aijuka, M., Santiago, A.E., Girón, J.A., Nataro, J.P., Buys, E.M., 2018. Enteroaggregative *Escherichia coli* is the predominant diarrheagenic E . coli pathotype among irrigation water and food sources in South Africa. *Int. J. Food Microbiol.* 278, 44–51. <https://doi.org/10.1016/j.ijfoodmicro.2018.04.018>
- Aijuka, M.E.O., 2014. Bacteriological quality of South African irrigation water and its role as a source of contamination on irrigated lettuce. University of Pretoria, Pretoria.
- Alivisatos, A.P., Blaser, M.J., Brodie, E.L., Chun, M., Dangl, J.L., Donohue, T.J., Dorrenstein, P.C., Gilbert, J.A., Green, J.L., J.K., J., Kinight, R., Maxon, M.E., McGall-Ngai, M.J., Miller, J.F., Pollard, K.S., Ruby, E.G., Taha, S., Consortium, U.M.I., 2015. A unified initiative to harness Earth’s microbiomes. *Science* (80-.). 350, 507–508. <https://doi.org/10.1126/science.aac8480>

- Allard, M.W., Strain, E., Melka, D., Bunning, K., Musser, S.M., Brown, E.W., Timme, R., 2016. Practical Value of Food Pathogen Traceability through Building a Whole-Genome Sequencing Network and Database. *J. Clin. Microbiol.* 54, 1975–83. <https://doi.org/10.1128/JCM.00081-16>
- Anelich, L., 2014. Foodborne Diseases: Prevalence of Foodborne Diseases in Africa, in: *Encyclopedia of Food Safety*. Elsevier, pp. 262–275. <https://doi.org/10.1016/B978-0-12-378612-8.00072-X>
- Atlas, R.M., 2012. One Health : Its Origins and Future. *Curr. Top. Microbiol. Immunol.* 365, 1–13. <https://doi.org/10.1007/82>
- Bai, J., Shi, X., Nagaraja, T.G., 2010. A multiplex PCR procedure for the detection of six major virulence genes in *Escherichia coli* O157:H7. *J. Microbiol. Methods* 82, 85–89. <https://doi.org/10.1016/j.mimet.2010.05.003>
- Baker, K.K., O'Reilly, C.E., Levine, M.M., Kotloff, K.L., Nataro, J.P., Ayers, T.L., Farag, T.H., Nasrin, D., Blackwelder, W.C., Wu, Y., Alonso, P.L., Breiman, R.F., Omoro, R., Faruque, A.S.G., Das, S.K., Ahmed, S., Saha, D., Sow, S.O., Sur, D., Zaidi, A.K.M., Quadri, F., Mintz, E.D., 2016. Sanitation and Hygiene-Specific Risk Factors for Moderate-to-Severe Diarrhea in Young Children in the Global Enteric Multicenter Study, 2007-2011: Case-Control Study. *PLoS Med.* 13, e1002010. <https://doi.org/10.1371/journal.pmed.1002010>
- Battesti, A., Majdalani, N., Gottesman, S., 2011. The RpoS-Mediated General Stress Response in *Escherichia coli*. *Annu. Rev. Microbiol.* 65, 189–213. <https://doi.org/10.1146/annurev-micro-090110-102946>
- Belén, A., Pavón, I., Maiden, M.C.J., 2009. Molecular Epidemiology of Microorganisms 551, 1–11. <https://doi.org/10.1007/978-1-60327-999-4>
- Beloin, C., Roux, A., Ghigo, J., 2008. *Escherichia coli* biofilms. *Curr. Top. Microbiol. Immunol.* 322, 249–289. https://doi.org/https://doi.org/10.1007/978-3-540-75418-3_12
- Berg, G., Erlacher, A., Smalla, K., Krause, R., 2014. Vegetable microbiomes: Is there a connection among opportunistic infections, human health and our “gut feeling”? *Microb. Biotechnol.* 7, 487–495. <https://doi.org/10.1111/1751-7915.12159>

- Bernard, C.S., Brunet, Y.R., Gueguen, E., Cascales, E., 2010. Nooks and crannies in type VI secretion regulation. *J. Bacteriol.* 192, 3850–60. <https://doi.org/10.1128/JB.00370-10>
- 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>
- Biassoni, R., Raso, A., 2014. *Quantitative Real-Time PCR Methods and Protocols*. Humana Press, New York. <https://doi.org/10.1007/978-1-4939-0733-5>
- Bielaszewska, M., Mellmann, A., Zhang, W., Köck, R., Fruth, A., Bauwens, A., Peters, G., Karch, H., 2011. Characterisation of the *Escherichia coli* strain associated with an outbreak of haemolytic uraemic syndrome in Germany, 2011: A microbiological study. *Lancet Infect. Dis.* 11, 671–676. [https://doi.org/10.1016/S1473-3099\(11\)70165-7](https://doi.org/10.1016/S1473-3099(11)70165-7)
- Bisi-johnson, M.A., Obi, C.L., Vasaikar, S.D., Baba, K.A., Hattori, T., 2011. Molecular basis of virulence in clinical isolates of *Escherichia coli* and *Salmonella* species from a tertiary hospital in the Eastern Cape, South Africa. *Gut Pathog.* 3, 1–8. <https://doi.org/10.1186/1757-4749-3-9>
- Blanton, L. V., Barratt, M.J., Charbonneau, M.R., Ahmed, T., Gordon, J.I., 2016. Childhood undernutrition, the gut microbiota, and microbiota-directed therapeutics. *Science* (80-.). 352, 1–6. <https://doi.org/10.1126/science.aad9359>
- Blount, Z.D., 2015. The unexhausted potential of *E. coli*. *Elife* 4, 1–12. <https://doi.org/10.7554/eLife.05826>
- Boisen, N., Struve, C., Scheutz, F., Krogfelt, K.A., Nataro, J.P., 2008. New adhesin of enteroaggregative *Escherichia coli* related to the Afa/Dr/AAF family. *Infect. Immun.* 76, 3281–92. <https://doi.org/10.1128/IAI.01646-07>
- Borgersen, Q., Bolick, D.T., Kolling, G.L., Aijuka, M., Ruiz-Perez, F., Guerrant, R.L., Nataro, J.P., Santiago, A.E., 2018. Abundant production of exopolysaccharide by EAEC strains enhances the formation of bacterial biofilms in contaminated sprouts. *Gut Microbes.* <https://doi.org/10.1080/19490976.2018.1429877>

- Boyer, F., Fichant, G., Berthod, J., Vandenbrouck, Y., Attree, I., 2009. Dissecting the bacterial type VI secretion system by a genome wide in silico analysis: what can be learned from available microbial genomic resources? *BMC Genomics* 10, 1–14. <https://doi.org/10.1186/1471-2164-10-104>
- Caine, L.-A., Nwodo, U., Okoh, A., Ndip, R., Green, E., 2014. Occurrence of Virulence Genes Associated with Diarrheogenic *Escherichia coli* Isolated from Raw Cow's Milk from Two Commercial Dairy Farms in the Eastern Cape Province, South Africa. *Int. J. Environ. Res. Public Health* 11, 11950–11963. <https://doi.org/10.3390/ijerph111111950>
- Campos, L.C., Vieira, M.A., Trabulsi, L.R., da Silva, L.A., Monteiro-Neto, V., Gomes, T.A., 1999. Diffusely adhering *Escherichia coli* (DAEC) strains of fecal origin rarely express F1845 adhesin. *Microbiol. Immunol.* 43, 167–170.
- Cantón, R., González-Alba, J.M., Galán, J.C., 2012. CTX-M enzymes: Origin and diffusion. *Front. Microbiol.* 3, 1–19. <https://doi.org/10.3389/fmicb.2012.00110>
- Carlton, E.J., Woster, A.P., DeWitt, P., Goldstein, R.S., Levy, K., 2016. A systematic review and meta-analysis of ambient temperature and diarrhoeal diseases. *Int. J. Epidemiol.* 45, 117–130. <https://doi.org/10.1093/ije/dyv296>
- Carter, M.Q., Louie, J.W., Feng, D., Zhong, W., Brandl, M.T., 2016. Curli fimbriae are conditionally required in *Escherichia coli* O157: H7 for initial attachment and biofilm formation. *Food Microbiol.* 57, 81–89. <https://doi.org/10.1016/j.fm.2016.01.006>
- Castro-Rosas, J., Cerna-Cortés, J.F., Méndez-Reyes, E., Lopez-Hernandez, D., Gómez-Aldapa, C.A., Estrada-Garcia, T., 2012. Presence of faecal coliforms, *Escherichia coli* and diarrheogenic *E. coli* pathotypes in ready-to-eat salads, from an area where crops are irrigated with untreated sewage water. *Int. J. Food Microbiol.* 156, 176–180. <https://doi.org/10.1016/j.ijfoodmicro.2012.03.025>
- Ceuppens, S., Li, D., Uyttendaele, M., Renault, P., Ross, P., Ranst, M. Van, Cocolin, L., Donaghy, J., 2014. Molecular methods in food safety microbiology: Interpretation and implications of nucleic acid detection. *Compr. Rev. Food Sci. Food Saf.* 13, 551–577. <https://doi.org/10.1111/1541-4337.12072>

- Chandran, A., Mazumder, A., 2015. Pathogenic potential, genetic diversity, and population structure of *Escherichia coli* strains isolated from a forest-dominated watershed (Comox Lake) in British Columbia, Canada. *Appl. Environ. Microbiol.* 81, 1788–1798.
<https://doi.org/10.1128/AEM.03738-14>
- Chandran, A., Mazumder, A., 2014. Occurrence of diarrheagenic virulence genes and genetic diversity in *Escherichia coli* isolates from fecal material of various avian hosts in British Columbia, Canada. *Appl. Environ. Microbiol.* 80, 1933–1940.
<https://doi.org/10.1128/AEM.03949-13>
- Chassagnea, L., Pradelb, N., Frédéric Robina, F., Valérie Livrelli, V., Richard Bonnet, R., Delmas, J., 2009. Detection of stx1, stx2, and eae genes of enterohemorrhagic *Escherichia coli* using SYBR Green in a real-time polymerase chain reaction. *Diagn. Microbiol. Infect. Dis.* 64, 98–101. <https://doi.org/10.1016/j.diagmicrobio.2009.01.031>
- Chattaway, M.A., Jenkins, C., Rajendram, D., Cravioto, A., Talukder, K.A., Dallman, T., Underwood, A., Platt, S., Okeke, I.N., Wain, J., 2014. Enteroaggregative *Escherichia coli* have evolved independently as distinct complexes within the *E. coli* population with varying ability to cause disease. *PLoS One* 9, 1–17.
<https://doi.org/10.1371/journal.pone.0112967>
- Chaudhuri, R.R., Henderson, I.R., 2012. The evolution of the *Escherichia coli* phylogeny. *Infect. Genet. Evol.* 12, 214–26. <https://doi.org/10.1016/j.meegid.2012.01.005>
- Chauret, C., 2011. Survival and control of *Escherichia coli* O157:H7 in foods, beverages, soil and water. *Virulence* 2, 593–601. <https://doi.org/10.4161/viru.2.6.18423>
- Chola, L., Michalow, J., Tugendhaft, A., Hofman, K., 2010. Reducing diarrhoea deaths in South Africa: costs and effects of scaling up essential interventions to prevent and treat diarrhoea in under-five children. *BMC Public Health* 15, 1–10.
<https://doi.org/10.1186/s12889-015-1689-2>
- Commission of The European Communities, 2011. Commission Staff Working Document: Lessons learned from the 2011 outbreak of Shiga toxin-producing *Escherichia coli* (STEC) O104:H4 in sprouted seeds. Brussels, Belgium.

- Croxen, M.A., Law, R.J., Scholz, R., Keeney, K.M., Wlodarska, M., Finlay, B.B., 2013. Recent advances in understanding enteric pathogenic *Escherichia coli*. Clin. Microbiol. Rev. 26, 822–880. <https://doi.org/10.1128/CMR.00022-13>
- De Bruijn, F.J., 1992. Use of Repetitive (REP and ERIC) sequences and the PCR to fingerprint the Genomes of Rhizobium meliloti isolates and other Soil Bacteria. Appl. Environ. Microbiol. 58, 2180–2187.
- Deng, W., Marshall, N.C., Rowland, J.L., McCoy, J.M., Worrall, L.J., Santos, A.S., Strynadka, N.C.J., Finlay, B.B., 2017. Assembly, structure, function and regulation of type III secretion systems. Nat. Rev. Microbiol. 15, 323–337. <https://doi.org/10.1038/nrmicro.2017.20>
- Dini-Andreote, F., Andreote, F.D., Araújo, W.L., Trevors, J.T., van Elsas, J.D., 2012. Bacterial Genomes: Habitat Specificity and Uncharted Organisms. Microb. Ecol. 64, 1–7. <https://doi.org/10.1007/s00248-012-0017-y>
- Dudley, E.G., Abe, C., Ghigo, J.-M., Latour-Lambert, P., Hormazabal, J.C., Nataro, J.P., 2006a. An IncII Plasmid Contributes to the Adherence of the Atypical Enterotoxigenic *Escherichia coli* Strain C1096 to Cultured Cells and Abiotic Surfaces. Infect. Immun. 74, 2102–2114. <https://doi.org/10.1128/IAI.74.4.2102-2114.2006>
- Dudley, E.G., Thomson, N.R., Parkhill, J., Morin, N.P., Nataro, J.P., 2006b. Proteomic and microarray characterization of the AggR regulon identifies a pheU pathogenicity island in enterotoxigenic *Escherichia coli*. Mol. Microbiol. 61, 1267–1282. <https://doi.org/10.1111/j.1365-2958.2006.05281.x>
- Edwards, L.A., Bajaj-Elliott, M., Klein, N.J., Murch, S.H., Phillips, A.D., 2011. Bacterial-epithelial contact is a key determinant of host innate immune responses to enteropathogenic and enterotoxigenic *Escherichia coli*. PLoS One 6, 1–15. <https://doi.org/10.1371/journal.pone.0027030>
- Eischeid, A.C., 2011. SYTO dyes and EvaGreen outperform SYBR Green in real-time PCR. BMC Res. Notes 4, 2–6. <https://doi.org/10.1186/1756-0500-4-263>

- Estrada-Garcia, T., Navarro-Garcia, F., 2012. Enteroaggregative *Escherichia coli* pathotype: A genetically heterogeneous emerging foodborne enteropathogen. FEMS Immunol. Med. Microbiol. 66, 281–298. <https://doi.org/10.1111/j.1574-695X.2012.01008.x>
- Faith, J.J., Colombel, J.-F., Gordon, J.I., 2015. Identifying strains that contribute to complex diseases through the study of microbial inheritance. Proc. Natl. Acad. Sci. 112, 633–640. <https://doi.org/10.1073/pnas.1418781112>
- Farrokh, C., Jordan, K., Auvray, F., Glass, K., Oppegaard, H., Raynaud, S., Thevenot, D., Condron, R., De Reu, K., Govaris, A., Heggum, K., Heyndrickx, M., Hummerjohann, J., Lindsay, D., Miszczucha, S., Moussiegt, S., Verstraete, K., Cerf, O., 2013. Review of Shiga-toxin-producing *Escherichia coli* (STEC) and their significance in dairy production. Int. J. Food Microbiol. 162, 190–212. <https://doi.org/10.1016/j.ijfoodmicro.2012.08.008>
- Flemming, H.C., Wingender, J., 2010. The biofilm matrix. Nat. Rev. Microbiol. 8, 623–633. <https://doi.org/10.1038/nrmicro2415>
- Fletcher, J., Leach, J.E., Eversole, K., Tauxe, R., 2013. Human pathogens on plants: designing a multidisciplinary strategy for research. Phytopathology 103, 306–15. <https://doi.org/10.1094/PHYTO-09-12-0236-IA>
- Franz, E., van Hoek, A.H.A.M., Bouw, E., Aarts, H.J.M., 2011. Variability of *Escherichia coli* O157 strain survival in manure-amended soil in relation to strain origin, virulence profile, and carbon nutrition profile. Appl. Environ. Microbiol. 77, 8088–8096. <https://doi.org/10.1128/AEM.00745-11>
- Gao, F., Isabel, A., Switt, M., Fratamico, P.M., Pm, F., Debroy, C., Liu, Y., Debroy, C., Liu, Y., Needleman, D.S., Baranzoni, G.M., Feng, P., 2016. Advances in Molecular Serotyping and Subtyping of *Escherichia coli*. Front. Microbiol 7, 1–8. <https://doi.org/10.3389/fmicb.2016.00644>
- Garcia, M.-I., Gounon, P., Courcoux, P., Labigne, A., Le Bouguéneq, C., 1996. The afimbrial adhesive sheath encoded by the afa-3 gene cluster of pathogenic *Escherichia coli* is composed of two adhesins. Mol. Microbiol. 19, 683–693. <https://doi.org/10.1046/j.1365-2958.1996.394935.x>

- Garcia, M.-I., Jouve, M., Nataro, J.P., Gounon, P., Le Bouguéneq, C., 2000. Characterization of the AfaD-like family of invasins encoded by pathogenic *Escherichia coli* associated with intestinal and extra-intestinal infections. *FEBS Lett.* 479, 111–117.
[https://doi.org/10.1016/S0014-5793\(00\)01898-6](https://doi.org/10.1016/S0014-5793(00)01898-6)
- Gemmell, M.E., Schmidt, S., 2012. Microbiological assessment of river water used for the irrigation of fresh produce in a sub-urban community in Sobantu, South Africa. *Food Res. Int.* 47, 300–305. <https://doi.org/10.1016/j.foodres.2011.07.016>
- Giaouris, E., Heir, E., Desvaux, M., Hébraud, M., Mørretrø, T., Langsrud, S., Doulgeraki, A., Nychas, G., Kačániová, M., Czaczyk, K., Ölmez, H., Simões, M., 2015. Intra and interspecies interactions within biofilms of important foodborne bacterial pathogens. *Front. Microbiol.* 6, 1–26. <https://doi.org/10.3389/fmicb.2015.00841>
- Gray, A., Vawda, Y., 2016. South African Health Review, in: Padarath A, King J, Mackie E, C.J. (Ed.), *South African Health Review 2016*. Health Systems Trust, Durban, South Africa, p. 369. <https://doi.org/10.1093/heapol/czq021>
- Green, E.R., Meccas, J., 2015. Bacterial Secretion Systems – An overview. *Am. Soc. Microbiol.* 4, 1–32. <https://doi.org/10.1128/microbiolspec>.
- Greene, N.P., Crow, A., Hughes, C., Koronakis, V., 2015. Structure of a bacterial toxin-activating acyltransferase. *Proc. Natl. Acad. Sci. U. S. A.* 112, 3058–3066.
<https://doi.org/10.1073/pnas.1503832112>
- Gudnason, H., Dufva, M., Bang, D.D., Wolff, A., 2007. Comparison of multiple DNA dyes for real-time PCR: Effects of dye concentration and sequence composition on DNA amplification and melting temperature. *Nucleic Acids Res.* 35, 1–8.
<https://doi.org/10.1093/nar/gkm671>
- Guinanane, C.M., Cotter, P.D., 2013. Role of the gut microbiota in health and chronic gastrointestinal disease: understanding a hidden metabolic organ. *Ther Adv Gastroenterol* 6, 295–308. <https://doi.org/10.1177/>
- Hasman, H., Mevius, D., Veldman, K., Olesen, I., Aarestrup, F.M., 2005. β -Lactamases among extended-spectrum β -lactamase (ESBL)-resistant *Salmonella* from poultry, poultry products and human patients in The Netherlands. *J. Antimicrob. Chemother.* 56, 115–121. <https://doi.org/10.1093/jac/dki190>

- Havelaar, A.H., Brul, S., Jong, A. De, Jonge, R. De, Zwietering, M.H., Benno, H., 2010. Future challenges to microbial food safety. *Int. J. Food Microbiol.* 139, S79–S94. <https://doi.org/10.1016/j.ijfoodmicro.2009.10.015>
- Heiman, K.E., Mody, R.K., Johnson, S.D., Griffin, P.M., Hannah Gould, L., 2015. *Escherichia coli* O157 Outbreaks in the United States, 2003–2012. *Emerg. Infect. Dis.* 21, 1293–1301. <https://doi.org/10.3201/eid2108.141364>
- Hendriksen, R.S., Mikoleit, M., Kornschöber, C., Rickert, R.L., Duyne, S. Van, Kjelsø, C., Hasman, H., Cormican, M., Mevius, D., Threlfall, J., Angulo, F.J., Aarestrup, F.M., 2009. Emergence of multidrug-resistant *Salmonella* concord infections in Europe and the United States in children adopted from Ethiopia, 2003–2007. *Pediatr. Infect. Dis. J.* 28, 814–818. <https://doi.org/10.1097/INF.0b013e3181a3aeac>
- Hoffmann, S.A., Pieretti, G.G., Fiorini, A., Patussi, E.V., Cardoso, R.F., Mikcha, J.M.G., 2014. Shiga-Toxin Genes and Genetic Diversity of *Escherichia coli* Isolated from Pasteurized Cow Milk in Brazil. *J. Food Sci.* 79, 1175–1180. <https://doi.org/10.1111/1750-3841.12477>
- Hufnagel, D.A., Depas, W.H., Chapman, M.R., 2015. The Biology of the *Escherichia coli* Extracellular Matrix. *Microbiol Spectr* 3, 1–24. <https://doi.org/10.1128/microbiolspec.MB-0014-2014>
- Ijabadeniyi, O.A., Debusho, L.K., Vanderlinde, M., Buys, E.M., 2011. Irrigation water as a potential preharvest source of bacterial contamination of vegetables. *J. Food Saf.* 31, 452–461. <https://doi.org/10.1111/j.1745-4565.2011.00321.x>
- Islam, N., Nagy, A., Garrett, W.M., Shelton, D., Cooper, B., Nou, X., 2016. Different Cellular Origins and Functions of Extracellular Proteins from *Escherichia coli* O157:H7 and O104:H4 as Determined by Comparative Proteomic Analysis. *Appl. Environ. Microbiol.* 82, 4371–8. <https://doi.org/10.1128/AEM.00977-16>
- Jacob, M.E., Bai, J., Renter, D.G., Rodgers, A.T., Shi, X., Nagaraja, T.G., 2014. Comparing Real-Time and Conventional PCR to Culture-Based Methods for Detecting and Quantifying *Escherichia coli* O157 in Cattle Feces. *J. Food Prot.* 77, 314–319. <https://doi.org/10.4315/0362-028X>

- Jain, R., Rivera, M.C., Lake, J.A., 1999. Horizontal gene transfer among genomes: The complexity hypothesis. *Proc. Natl. Acad. Sci.* 96, 3801–3806.
<https://doi.org/10.1073/pnas.96.7.3801>
- Jain, R., Rivera, M.C., Moore, J.E., Lake, J.A., 2003. Horizontal gene transfer accelerates genome innovation and evolution. *Mol. Biol. Evol.* 20, 1598–1602.
<https://doi.org/10.1093/molbev/msg154>
- Jansson, L., Koliiana, M., Sidstedt, M., Hedman, J., 2017. Blending DNA binding dyes to improve detection in real-time PCR. *Biotechnol. Reports* 14, 34–37.
<https://doi.org/10.1016/j.btre.2017.02.002>
- Jørgensen, S.B., Søråas, A. V, Arnesen, L.S., Leegaard, T.M., Sundsfjord, A., Jennum, P.A., 2017. A comparison of extended spectrum β -lactamase producing *Escherichia coli* from clinical, recreational water and wastewater samples associated in time and location. *PLOS ONE* | 12, 1–15. <https://doi.org/10.1371/journal.pone.0186576>
- Kaas, R.S., Friis, C., Ussery, D.W., Aarestrup, F.M., 2012. Estimating variation within the genes and inferring the phylogeny of 186 sequenced diverse *Escherichia coli* genomes. *BMC Genomics* 13, 1–13. <https://doi.org/10.1186/1471-2164-13-577>
- Kaper, J.B., Nataro, J.P., Mobley, L.T., 2004. Pathogenic *Escherichia coli*. *Nat. Rev. Microbiol.* 2, 123–140. <https://doi.org/10.1038/nrmicro818>
- Karp, B.E., Tate, H., Plumblee, J.R., Dessai, U., Whichard, J.M., Thacker, E.L., Robertson Hale, K., Wilson, W., Friedman, C.R., Griffin, P.M., McDermott, P.F., 2017. National Antimicrobial Resistance Monitoring System: Two Decades of Advancing Public Health Through Integrated Surveillance of Antimicrobial Resistance. *Foodborne Pathog. Dis.* 14, 545–557. <https://doi.org/10.1089/fpd.2017.2283>
- Kau, A.L., Planer, J.D., Liu, J., Rao, S., Yatsunenko, T., Trehan, I., Manary, M.J., Liu, T.-C., Stappenbeck, T.S., Maleta, K.M., Ashorn, P., Dewey, K.G., Houpt, E.R., Hsieh, C.-S., Gordon, J.I., 2015. Functional characterization of IgA-targeted bacterial taxa from malnourished Malawian children that produce diet-dependent enteropathy. *Sci Transl Med* 25, 276. <https://doi.org/10.1126/scitranslmed.aaa4877>

- Kaur, P., Chakraborti, A., Asea, A., 2010. Enteroaggregative *Escherichia coli* : An Emerging Enteric Food Borne Pathogen. *Interdiscip. Perspect. Infect. Dis.* 2010, 1–10.
<https://doi.org/10.1155/2010/254159>
- Kirk, M.D., Angulo, F.J., Havelaar, A.H., Black, R.E., 2017. Diarrhoeal disease in children due to contaminated food. *Bull. World Health Organ.* 95, 233–234.
<https://doi.org/10.2471/BLT.16.173229>
- Kirk, M.D., Pires, S.M., Black, R.E., Caipo, M., Crump, J.A., Devleeschauwer, B., Döpfer, D., Fazil, A., Fischer-Walker, C.L., Hald, T., Hall, A.J., Keddy, K.H., Lake, R.J., Lanata, C.F., Torgerson, P.R., Havelaar, A.H., Angulo, F.J., 2015. World Health Organization Estimates of the Global and Regional Disease Burden of 22 Foodborne Bacterial, Protozoal, and Viral Diseases, 2010: A Data Synthesis. *PLoS Med.* 12, 1–21.
<https://doi.org/10.1371/journal.pmed.1001921>
- Koga, V.L., Rodrigues, G.R., Scandorieiro, S., Vespero, E.C., Oba, A., de Brito, B.G., de Brito, K.C.T., Nakazato, G., Kobayashi, R.K.T., 2015. Evaluation of the Antibiotic Resistance and Virulence of *Escherichia coli* Strains Isolated from Chicken Carcasses in 2007 and 2013 from Paraná, Brazil. *Foodborne Pathog. Dis.* 12, 479–485.
<https://doi.org/10.1089/fpd.2014.1888>
- Kotloff, K.L., Nataro, J.P., Blackwelder, W.C., Nasrin, D., Farag, T.H., Panchalingam, S., Wu, Y., Sow, S.O., Sur, D., Breiman, R.F., Faruque, A.S.G., Zaidi, A.K.M., Saha, D., Alonso, P.L., Tamboura, B., Sanogo, D., Onwuchekwa, U., Manna, B., Ramamurthy, T., Kanungo, S., Ochieng, J.B., Omere, R., Oundo, J.O., Hossain, A., Das, S.K., Ahmed, S., Qureshi, S., Quadri, F., Adegbola, R.A., Antonio, M., Hossain, M.J., Akinsola, A., Mandomando, I., Nhampossa, T., Acácio, S., Biswas, K., O'Reilly, C.E., Mintz, E.D., Berkeley, L.Y., Muhsen, K., Sommerfelt, H., Robins-Browne, R.M., Levine, M.M., 2013. Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): A prospective, case-control study. *Lancet* 382, 209–222. [https://doi.org/10.1016/S0140-6736\(13\)60844-2](https://doi.org/10.1016/S0140-6736(13)60844-2)
- Kubista, M., Andrade, J.M., 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. *Mol. Aspects Med.* 27, 95–125.
<https://doi.org/10.1016/j.mam.2005.12.007>

- Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874.
<https://doi.org/10.1093/molbev/msw054>
- Land, M., Hauser, L., Jun, S.-R., Nookaew, I., Leuze, M.R., Ahn, T.-H., Karpinets, T., Lund, O., Kora, G., Wassenaar, T., Poudel, S., Ussery, D.W., Ussery, D.W., Hauser, L., Leuze, M.R., Ahn, T., Kora, G., Lund, O., Wassenaar, T., Poudel, S., 2015. Insights from 20 years of bacterial genome sequencing. *Funct Integr Genomics* 15, 141–161.
<https://doi.org/10.1007/s10142-015-0433-4>
- Leimbach, A., Hacker, J., Dobrindt, U., 2013. *E. coli* as an All-Rounder: The Thin Line Between Commensalism and Pathogenicity. *Curr. Top. Microbiol. Immunol.* 358, 3–32.
https://doi.org/10.1007/82_2012_303
- Leung, K.Y., Siame, B.A., Snowball, H., Mok, Y.K., 2011. Type VI secretion regulation: Crosstalk and intracellular communication. *Curr. Opin. Microbiol.* 14, 9–15.
<https://doi.org/10.1016/j.mib.2010.09.017>
- Macarisin, D., Patel, J., Bauchan, G., Giron, J.A., Sharma, V.K., 2012. Role of Curli and Cellulose Expression in Adherence of *Escherichia coli* O157:H7 to Spinach Leaves. *Foodborne Pathog. Dis.* 9, 160–167. <https://doi.org/10.1089/fpd.2011.1020>
- Mangen, M.-J.J., Batz, M.B., Annemarie, K., Hald, T., Morris, J.G., Taylor, M., Havelaar, A.H., 2010. Integrated Approaches for the Public Health Prioritization of Foodborne and Zoonotic Pathogens. *Risk Anal.* 30, 782–797. <https://doi.org/10.1111/j.1539-6924.2009.01291.x>
- Mariadassou, M., Pichon, S., Ebert, D., 2015. Microbial ecosystems are dominated by specialist taxa. *Ecol. Lett.* 18, 974–982. <https://doi.org/10.1111/ele.12478>
- Martínez, J.L., 2013. Bacterial pathogens: From natural ecosystems to human hosts. *Environ. Microbiol.* 15, 325–333. <https://doi.org/10.1111/j.1462-2920.2012.02837.x>
- Martiny, J.B.H., Jones, S.E., Lennon, J.T., Martiny, A.C., 2015. Microbiomes in light of traits: A phylogenetic perspective. *Science* (80-.). 350, 1–8.
<https://doi.org/10.1126/science.aac9323>
- Maule, J., 1998. Pulsed-field gel electrophoresis. *Mol. Biotechnol.* 9, 107–26.
<https://doi.org/10.1007/BF02760813>

- Mendes, R.E., Kiyota, K.A., Monteiro, J., Castanheira, M., Andrade, S.S., Gales, A.C., Pignatari, A.C.C., Tufik, S., 2007. Rapid detection and identification of metallo-Beta-lactamase-encoding genes by multiplex real-time PCR assay and melt curve analysis. *J. Clin. Microbiol.* 45, 544–547. <https://doi.org/10.1128/JCM.01728-06>
- 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, 1–13. <https://doi.org/10.3389/fmicb.2015.00932>
- Miyata, S.T., Bachmann, V., Pukatzki, S., 2013. Type VI secretion system regulation as a consequence of evolutionary pressure. *J. Med. Microbiol.* 62, 663–676. <https://doi.org/10.1099/jmm.0.053983-0>
- Mohamed, J.A., Huang, D.B., Jiang, Z.-D., DuPont, H.L., Nataro, J.P., Belkind-Gerson, J., Okhuysen, P.C., 2007. Association of putative enteroaggregative *Escherichia coli* virulence genes and biofilm production in isolates from travelers to developing countries. *J. Clin. Microbiol.* 45, 121–6. <https://doi.org/10.1128/JCM.01128-06>
- Monis, P.T., Giglio, S., Saint, C.P., 2005. Comparison of SYTO9 and SYBR Green I for real-time polymerase chain reaction and investigation of the effect of dye concentration on amplification and DNA melting curve analysis. *Anal. Biochem.* 340, 24–34. <https://doi.org/10.1016/j.ab.2005.01.046>
- Moser, S.A., Mitchell, S.C., Ruby, J.D., Momeni, S., Osgood, R.C., Whiddon, J., Childers, N.K., 2010. Repetitive extragenic palindromic PCR for study of *Streptococcus mutans* diversity and transmission in human populations. *J. Clin. Microbiol.* 48, 599–602. <https://doi.org/10.1128/JCM.01828-09>
- Moshabela, M., MacPherson, P., Ezard, N., Frean, E., Mashimbye, L., Elliott, J.H., Oldenburg, B., 2012. Clinical and social determinants of diarrhoeal disease in a rural HIV/AIDS clinic, South Africa: a case-control study. *Int. J. STD AIDS* 23, 346–350. <https://doi.org/10.1258/ijsa.2011.011285>
- Nagy, A., Mowery, J., Bauchan, G.R., Wang, L., Nichols-Russell, L., Nou, X., 2015. Role of Extracellular Structures of *Escherichia coli* O157:H7 in Initial Attachment to Biotic and Abiotic Surfaces. *Appl. Environ. Microbiol.* 81, 4720–4727. <https://doi.org/10.1128/AEM.00215-15>

- Nagy, A., Xu, Y., Bauchan, G.R., Shelton, D.R., Nou, X., 2016. Aggregative adherence fimbriae I (AAF/I) mediate colonization of fresh produce and abiotic surface by Shiga toxinogenic enteroaggregative *Escherichia coli* O104:H4. *Int. J. Food Microbiol.* 229, 44–51. <https://doi.org/10.1016/j.ijfoodmicro.2016.04.007>
- Nash, J.H., Villegas, A., Kropinski, A.M., Aguilar-Valenzuela, R., Konczy, P., Mascarenhas, M., Ziebell, K., Torres, A.G., Karmali, M. a, Coombes, B.K., 2010. Genome sequence of adherent-invasive *Escherichia coli* and comparative genomic analysis with other *E. coli* pathotypes. *BMC Genomics* 11, 1–15. <https://doi.org/10.1186/1471-2164-11-667>
- Nataro, J., Kaper, J., 1998. Diarrheagenic *Escherichia coli* Strains. *Clin. Microbiology Rev.* 11, 142–201.
- Nataro, J.P., Kaper, J.B., Robins-Browne, R., Prado, V., Vial, P., Levine, M.M., 1987. Patterns of adherence of diarrheagenic *Escherichia coli* to HEp-2 cells. *Pediatr. Infect. Dis. J.* 6, 829–831. <https://doi.org/10.1097/00006454-198709000-00008>
- Navarro, E., Serrano-Heras, G., Castaño, M.J., Solera, J., 2015. Real-time PCR detection chemistry. *Clin. Chim. Acta* 439, 231–250. <https://doi.org/10.1016/j.cca.2014.10.017>
- Nesse, L.L., Sekse, C., Berg, K., Johannesen, K.C.S., Solheim, H., Vestby, L.K., Urdahl, A.M., 2014. Potentially Pathogenic *Escherichia coli* Can Form a Biofilm under Conditions Relevant to the Food Production Chain. *Appl. Environ. Microbiol.* 80, 2042–2049. <https://doi.org/10.1128/AEM.0333113>
- Newell, D.G., Koopmans, M., Verhoef, L., Duizer, E., Aidara-Kane, A., Sprong, H., Opsteegh, M., Langelaar, M., Threlfall, J., Scheutz, F., der Giessen, J. van, Kruse, H., 2010. Food-borne diseases - The challenges of 20years ago still persist while new ones continue to emerge. *Int. J. Food Microbiol.* 139, 3–15. <https://doi.org/10.1016/j.ijfoodmicro.2010.01.021>
- Ng, W.-L., Bassler, B., 2009. Bacterial Quorum-Sensing Network Architectures. *Annu. Rev. Genet* 43, 197–222. <https://doi.org/10.1146/annurev-genet-102108-134304>.

- Nhampossa, T., Mandomando, I., Acacio, S., Quintó, L., Vubil, D., Ruiz, J., Nhalungo, D., Saco, C., Nhabanga, A., Nhacolo, A., Aide, P., Machevo, S., Sigauque, B., Nhama, A., Kotloff, K., Farag, T., Nasrin, D., Bassat, Q., Macete, E., Levine, M.M., Alonso, P., 2015. Diarrheal disease in rural Mozambique: Burden, risk factors and etiology of diarrheal disease among children aged 0-59 months seeking care at health facilities. *PLoS One* 10, 1–19. <https://doi.org/10.1371/journal.pone.0119824>
- Niba, E.T.E., Naka, Y., Nagase, M., Mori, H., Kitakawa, M., 2007. A genome-wide approach to identify the genes involved in biofilm formation in *E. coli*. *DNA Res.* 14, 237–46. <https://doi.org/10.1093/dnares/dsm024>
- Nicolas-Chanoine, M.-H., Bertrand, X., Madec, J.-Y., 2014. *Escherichia coli* ST131, an intriguing clonal group. *Clin. Microbiol. Rev.* 27, 543–74. <https://doi.org/10.1128/CMR.00125-13>
- Njage, P.M.K., Buys, E.M., 2017. Quantitative assessment of human exposure to extended spectrum and *AmpC* β -lactamases bearing *E. coli* in lettuce attributable to irrigation water and subsequent horizontal gene transfer. *Int. J. Food Microbiol.* 240, 141–151. <https://doi.org/10.1016/j.ijfoodmicro.2016.10.011>
- Ntuli, V., Njage, P.M.K., Buys, E.M., 2017. Extended-spectrum β -lactamase, shigatoxin and haemolysis capacity of O157 and non-O157 *E. coli* serotypes from producer-distributor bulk milk. *Int. Dairy J.* 66, 126–134. <https://doi.org/10.1016/j.idairyj.2016.11.008>
- Ntuli, V., Njage, P.M.K., Buys, E.M., 2016. Characterization of *Escherichia coli* and other *Enterobacteriaceae* in producer-distributor bulk milk. *J. Dairy Sci.* 99, 9534–9549. <https://doi.org/10.3168/jds.2016-11403>
- Ntuli, V., Njage, P.M.K., Buys, E.M., Bonilauri, P., Serraino, A., 2018. Quantitative risk assessment of hemolytic uremic syndrome associated with consumption of bulk milk sold directly from producer to consumer in South Africa. *J. Food Prot.* 81, 472–481.
- Ochieng, J.B., Boisen, N., Lindsay, B., Santiago, A., Ouma, C., Ombok, M., Fields, B., Stine, O.C., Nataro, J.P., 2014. *Serratia marcescens* is injurious to intestinal epithelial cells. *Gut Microbes* 5, 729–736. <https://doi.org/10.4161/19490976.2014.972223>

- Okeke, I.N., Wallace-Gadsden, F., Simons, H.R., Matthews, N., Labar, A.S., Hwang, J., Wain, J., 2010. Multi-locus sequence typing of enteroaggregative *Escherichia coli* isolates from nigerian children uncovers multiple lineages. PLoS One 5, 1–12. <https://doi.org/10.1371/journal.pone.0014093>
- Olaniran, A.O., Naicker, K., Pillay, B., 2009. Antibiotic resistance profiles of *Escherichia coli* isolates from river sources in Durban, South Africa. World J. Microbiol. Biotechnol. 25, 1743–1749. <https://doi.org/10.1007/s11274-009-0071-x>
- Orskov, F., Orskov, I., 1992. *Escherichia coli* serotyping and disease in man and animals. Can. J. Microbiol. 38, 699–704. <https://doi.org/10.1139/m92-115>
- Oundo, J.O., Kariuki, S.M., Boga, H.I., Muli, F.W., Iijima, Y., 2008. High incidence of enteroaggregative *Escherichia coli* among food handlers in three areas of Kenya: A possible transmission route of travelers' diarrhea. J. Travel Med. 15, 31–38. <https://doi.org/10.1111/j.1708-8305.2007.00174.x>
- Panchalingam, S., Antonio, M., Hossain, A., Mandomando, I., Ochieng, B., Oundo, J., Ramamurthy, T., Tamboura, B., Zaidi, A.K.M., Petri, W., Houpt, E., Murray, P., Prado, V., Vidal, R., Steele, D., Strockbine, N., Sansonetti, P., Glass, R.I., Robins-Browne, R.M., Tauschek, M., Svennerholm, A.M., Kotloff, K., Levine, M.M., Nataro, J.P., 2012. Diagnostic microbiologic methods in the GEMS-1 case/control study. Clin. Infect. Dis. 55, 294–302. <https://doi.org/10.1093/cid/cis754>
- Pérez-Losada, M., Cabezas, P., Castro-Nallar, E., Crandall, K.A., 2013. Pathogen typing in the genomics era: MLST and the future of molecular epidemiology. Infect. Genet. Evol. 16, 38–53. <https://doi.org/10.1016/j.meegid.2013.01.009>
- Philippot, L., Andersson, S.G.E., Battin, T.J., Prosser, J.I., Schimel, J.P., Whitman, W.B., Hallin, S., 2010. The ecological coherence of high bacterial taxonomic ranks. Nat. Rev. Microbiol. 8, 523–529. <https://doi.org/10.1038/nrmicro2367>
- Pichel, M., Brengi, S.P., Cooper, K.L.F., Ribot, E.M., Al-Busaidy, S., Araya, P., Fernández, J., Vaz, T.I., Kam, K.M., Morcos, M., Nielsen, E.M., Nadon, C., Pimentel, G., Pérez-Gutiérrez, E., Gerner-Smidt, P., Binsztein, for the Shigella flexner, N., 2012. Standardization and International Multicenter Validation of a PulseNet Pulsed-Field Gel Electrophoresis Protocol for Subtyping *Shigella flexneri* Isolates. Foodborne Pathog. Dis. 9, 418–424. <https://doi.org/10.1089/fpd.2011.1067>

Platts-Mills, J.A., Babji, S., Bodhidatta, L., Gratz, J., Haque, R., Havt, A., McCormick, B.J.J., McGrath, M., Olortegui, M.P., Samie, A., Shakoor, S., Mondal, D., Lima, I.F.N., Hariraju, D., Rayamajhi, B.B., Qureshi, S., Kabir, F., Yori, P.P., Mufamadi, B., Amour, C., Carreon, J.D., Richard, S.A., Lang, D., Bessong, P., Mduma, E., Ahmed, T., Lima, A.A.A.M., Mason, C.J., Zaidi, A.K.M., Bhutta, Z.A., Kosek, M., Guerrant, R.L., Gottlieb, M., Miller, M., Kang, G., Houpt, E.R., Chavez, C.B., Trigoso, D.R., Flores, J.T., Vasquez, A.O., Pinedo, S.R., Acosta, A.M., Ahmed, I., Alam, D., Ali, A., Rasheed, M., Soofi, S., Turab, A., Yousafzai, A.K., Bose, A., Jennifer, M.S., John, S., Kaki, S., Koshy, B., Muliylil, J., Raghava, M.V., Ramachandran, A., Rose, A., Sharma, S.L., Thomas, R.J., Pan, W., Ambikapathi, R., Charu, V., Dabo, L., Doan, V., Graham, J., Hoest, C., Knobler, S., Mohale, A., Nayyar, G., Psaki, S., Rasmussen, Z., Seidman, J.C., Wang, V., Blank, R., Tountas, K.H., Swema, B.M., Yarrot, L., Nshama, R., Ahmed, A.M.S., Tofail, F., Hossain, I., Islam, M., Mahfuz, M., Chandyo, R.K., Shrestha, P.S., Shrestha, R., Ulak, M., Black, R., Caulfield, L., Checkley, W., Chen, P., Lee, G., Murray-Kolb, L.E., Schaefer, B., Pendergast, L., Abreu, C., Costa, H., Moura, A. Di, Filho, J.Q., Leite, Ivaro, Lima, N., Maciel, B., Moraes, M., Mota, F., Ori, R., Quetz, J., Soares, A., Patil, C.L., Mahopo, C., Mapula, A., Nesamvuni, C., Nyathi, E., Barrett, L., Petri, W.A., Scharf, R., Shrestha, B., Shrestha, S.K., Strand, T., Svensen, E., 2015. Pathogen-specific burdens of community diarrhoea in developing countries: A multisite birth cohort study (MAL-ED). *Lancet Glob. Heal.* 3, 564–575. [https://doi.org/10.1016/S2214-109X\(15\)00151-5](https://doi.org/10.1016/S2214-109X(15)00151-5)

Prats, J., Garcia-Armisen, T., Larrea, J., Servais, P., 2008. Comparison of culture-based methods to enumerate *Escherichia coli* in tropical and temperate freshwaters. *Lett. Appl. Microbiol.* 46, 243–248. <https://doi.org/10.1111/j.1472-765X.2007.02292.x>

Pratt, L.A., Kolter, R., 1998. Genetic analysis of *Escherichia coli* biofilm formation: roles of flagella, motility, chemotaxis and type I pili. *Mol. Microbiol.* 30, 285–293. <https://doi.org/10.1046/j.1365-2958.1998.01061.x>

Prigent-combaret, C., Vidal, O., Lejeune, P., Landini, P., Dorel, C., 2001. Complex Regulatory Network Controls Initial Adhesion and Bio lm Formation in. *Society* 183, 7213–7223. <https://doi.org/10.1128/JB.183.24.7213>

- Ramirez, K.S., Knight, C.G., De Hollander, M., Brearley, F.Q., Constantinides, B., Cotton, A., Creer, S., Crowther, T.W., Davison, J., Delgado-Baquerizo, M., Dorrepaal, E., Elliott, D.R., Fox, G., Griffiths, R.I., Hale, C., Hartman, K., Houlden, A., Jones, D.L., Krab, E.J., Maestre, F.T., McGuire, K.L., Monteux, S., Orr, C.H., Van Der Putten, W.H., Roberts, I.S., Robinson, D.A., Rocca, J.D., Rowntree, J., Schlaeppli, K., Shepherd, M., Singh, B.K., Straathof, A.L., Bhatnagar, J.M., Thion, C., Van Der Heijden, M.G.A., De Vries, F.T., 2018. Detecting macroecological patterns in bacterial communities across independent studies of global soils. *Nat. Microbiol.* 3, 189–196.
<https://doi.org/10.1038/s41564-017-0062-x>
- Rasko, D.A., Webster, D.R., Sahl, J.W., Bashir, A., Boisen, N., Scheutz, F., Paxinos, E.E., Sebra, R., Chin, C.-S., Iliopoulos, D., Klammer, A., Peluso, P., Lee, L., Kislyuk, A.O., Bullard, J., Kasarskis, A., Wang, S., Eid, J., Rank, D., Redman, J.C., Steyert, S.R., Frimodt-Møller, J., Struve, C., Petersen, A.M., Krogfelt, K.A., Nataro, J.P., Schadt, E.E., Waldor, M.K., 2011. Origins of the *E. coli* Strain Causing an Outbreak of Hemolytic–Uremic Syndrome in Germany. *N. Engl. J. Med.* 365, 709–717.
<https://doi.org/10.1056/NEJMoa1106920>
- Rasmussen, J.P., Saint, C.P., Monis, P.T., 2007. Use of DNA melting simulation software for in silico diagnostic assay design: Targeting regions with complex melting curves and confirmation by real-time PCR using intercalating dyes. *BMC Bioinformatics* 8, 1–11.
<https://doi.org/10.1186/1471-2105-8-107>
- Rebello, R.C. de L., Regua-Mangia, A.H., 2014. Potential enterovirulence and antimicrobial resistance in *Escherichia coli* isolates from aquatic environments in Rio de Janeiro, Brazil. *Sci. Total Environ.* 490, 19–27. <https://doi.org/10.1016/j.scitotenv.2014.04.040>
- Reperant, L.A., Cornaglia, G., Osterhaus, A.D.M.E., 2012. The Importance of Understanding the Human – Animal Interface From Early Hominins to Global Citizens. *Curr. Top. Microbiol. Immunol.* 365, 49–81. <https://doi.org/10.1007/82>
- Robins-Browne, R.M., Holt, K.E., Ingle, D.J., Hocking, D.M., Yang, J., Tauschek, M., 2016. Are *Escherichia coli* Pathotypes Still Relevant in the Era of Whole-Genome Sequencing? *Front. Cell. Infect. Microbiol.* 6, 1–9.
<https://doi.org/10.3389/fcimb.2016.00141>

- Roca, I., Akova, M., Baquero, F., Carlet, J., Cavaleri, M., Coenen, S., Cohen, J., Findlay, D., Gyssens, I., Heur, O.E., Kahlmeter, G., Kruse, H., Laxminarayan, R., Liébana, E., López-Cerero, L., MacGowan, A., Martins, M., Rodríguez-Baño, J., Rolain, J.M., Segovia, C., Sigauque, B., Taconelli, E., Wellington, E., Vila, J., 2015. The global threat of antimicrobial resistance: Science for intervention. *New Microbes New Infect.* 6, 22–29. <https://doi.org/10.1016/j.nmni.2015.02.007>
- Ronholm, J., Nasheri, N., Petronella, N., Pagotto, F., 2016. Navigating microbiological food safety in the era of whole-genome sequencing. *Clin. Microbiol. Rev.* 29, 837–857. <https://doi.org/10.1128/CMR.00056-16>
- Ruiz-Perez, F., Nataro, J.P., 2014. Bacterial serine proteases secreted by the autotransporter pathway: classification, specificity, and role in virulence. *Cell. Mol. Life Sci.* 71, 745–770. <https://doi.org/10.1007/s00018-013-1355-8>
- Saldaña, Z., Sánchez, E., Xicohtencatl-Cortes, J., Puente, J.L., Girón, J.A., 2011. Surface Structures Involved in Plant Stomata and Leaf Colonization by Shiga-Toxigenic *Escherichia coli* O157:H7. *Front. Microbiol.* 2, 1–9. <https://doi.org/10.3389/fmicb.2011.00119>
- Saldaña, Z., Xicohtencatl-Cortes, J., Avelino, F., Phillips, A.D., Kaper, J.B., Puente, J.L., Girón, J.A., 2009. Synergistic role of curli and cellulose in cell adherence and biofilm formation of attaching and effacing *Escherichia coli* and identification of Fis as a negative regulator of curli. *Env. Microbiol* 11, 992–1006. <https://doi.org/10.1111/j.1462-2920.2008.01824.x>
- Samie, A., Obi, C.L., Dillingham, R., Pinkerton, R.C., Guerrant, R.L., 2007. Enteroaggregative *Escherichia coli* in Venda, South Africa: Distribution of virulence-related genes by multiplex polymerase chain reaction in stool samples of human immunodeficiency virus (HIV)-positive and HIV-negative individuals and primary school child. *Am. J. Trop. Med. Hyg.* 77, 142–150.
- Shi, H., Sun, F., Chen, J., Ou, Q., Feng, W., Yong, X., Xia, P., 2015. Epidemiology of CTX-M-type extended-spectrum beta-lactamase (ESBL)-producing nosocomial *-Escherichia coli* infection in China. *Ann. Clin. Microbiol. Antimicrob.* 14, 1–5. <https://doi.org/10.1186/s12941-015-0063-7>

- Shin, N.-R., Whon, T.W., Bae, J.-W., 2015. Proteobacteria: microbial signature of dysbiosis in gut microbiota. *Trends Biotechnol.* 33, 496–503.
<https://doi.org/10.1016/j.tibtech.2015.06.011>
- Sidhu, J.P.S., Ahmed, W., Hodgers, L., Toze, S., 2013. Occurrence of virulence genes associated with diarrheagenic pathotypes in *Escherichia coli* isolates from surface water. *Appl. Environ. Microbiol.* 79, 328–335. <https://doi.org/10.1128/AEM.02888-12>
- Silverman, J.M., Brunet, Y.R., Cascales, E., Mougous, J.D., 2012. Structure and Regulation of the Type VI Secretion System 66, 453–472. <https://doi.org/10.1146/annurev-micro-121809-151619>
- Singh, P., Mustapha, A., 2014. Development of a real-time PCR melt curve assay for simultaneous detection of virulent and antibiotic resistant *Salmonella*. *Food Microbiol.* 44, 6–14. <https://doi.org/10.1016/j.fm.2014.04.014>
- Skippington, E., Ragan, M.A., 2012. Phylogeny rather than ecology or lifestyle biases the construction of *Escherichia coli-Shigella* genetic exchange communities. *Open Biol.* 2, 1–14. <https://doi.org/10.1098/rsob.120112>
- Skippington, E., Ragan, M.A., 2011. Lateral genetic transfer and the construction of genetic exchange communities. *FEMS Microbiol. Rev.* 35, 707–735.
<https://doi.org/10.1111/j.1574-6976.2010.00261.x>
- Soborg, D.A., Hendriksen, N.B., Kilian, M., Kroer, N., 2013. Widespread occurrence of bacterial human virulence determinants in soil and freshwater environments. *Appl. Environ. Microbiol.* 79, 5488–5497. <https://doi.org/10.1128/AEM.01633-13>
- Steiner, T.S., Samie, A., Guerrant, R.L., 2006. Infectious diarrhea: new pathogens and new challenges in developed and developing areas. *Clin. Infect. Dis.* 43, 408–410.
<https://doi.org/10.1086/505874>
- Strauman, M.C., Harper, J.M., Harrington, S.M., Boll, E.J., Nataro, J.P., 2010. Enteroaggregative *Escherichia coli* Disrupts Epithelial Cell Tight Junctions. *Infect. Immun.* 78, 4958–4964. <https://doi.org/10.1128/IAI.00580-10>

- Subramanian, S., Huq, S., Yatsunenkov, T., Haque, R., Mahfuz, M., Alam, M.A., Benezra, A., Destefano, J., Meier, M.F., Muegge, B.D., Barratt, M.J., Vanarendonk, L.G., Zhang, Q., Province, M.A., Jr, W.A.P., Ahmed, T., Gordon, J.I., 2014. Persistent gut microbiota immaturity in malnourished Bangladeshi children. *Nature* 510, 417–421. <https://doi.org/10.1038/nature13421>
- Tau, N.P., Meidany, P., Smith, A.M., Sooka, A., Keddy, K.H., 2012. *Escherichia coli* O104 associated with human diarrhea, South Africa, 2004–2011. *Emerg. Infect. Dis.* 18, 1314–1317. <https://doi.org/10.3201/eid1808.111616>
- Tauxe, R. V, Doyle, M.P., Kuchenmüller, T., Schlundt, J., Stein, C.E., 2010. Evolving public health approaches to the global challenge of foodborne infections. *Int. J. Food Microbiol.* 139, S16–S28. <https://doi.org/10.1016/j.ijfoodmicro.2009.10.014>
- Temperton, B., Giovannoni, S.J., 2012. Metagenomics: Microbial diversity through a scratched lens. *Curr. Opin. Microbiol.* 15, 605–612. <https://doi.org/10.1016/j.mib.2012.07.001>
- Tenaillon, O., Skurnik, D., Picard, B., Denamur, E., 2010. The population genetics of commensal *Escherichia coli*. *Nat. Rev. Microbiol.* 8, 207–217. <https://doi.org/10.1038/nrmicro2298>
- The Review on Antimicrobial Resistance, 2014. Antimicrobial Resistance: Tackling a Crisis for the Health and Wealth of Nations. London, United Kingdom. <https://doi.org/10.1038/510015a>
- Thomas, C.M., Nielsen, K.M., 2005. Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nat. Rev. Microbiol.* 3, 711–721. <https://doi.org/10.1038/nrmicro1234>
- Tobes, R., Pareja, E., 2006. Bacterial repetitive extragenic palindromic sequences are DNA targets for Insertion Sequence elements. *BMC Genomics* 7, 1–12. <https://doi.org/10.1186/1471-2164-7-62>

- Touchon, M., Hoede, C., Tenailon, O., Barbe, V., Baeriswyl, S., Bidet, P., Bingen, E., Bonacorsi, S., Bouchier, C., Bouvet, O., Calteau, A., Chiapello, H., Clermont, O., Cruveiller, S., Danchin, A., Diard, M., Dossat, C., El Karoui, M., Frapy, E., Garry, L., Ghigo, J.M., Gilles, A.M., Johnson, J., Le Bouguéneq, C., Lescat, M., Mangenot, S., Martinez-Jéhanne, V., Matic, I., Nassif, X., Oztas, S., Petit, M.A., Pichon, C., Rouy, Z., Ruf, C. Saint, Schneider, D., Turret, J., Vacherie, B., Vallenet, D., Médigue, C., Rocha, E.P.C., Denamur, E., 2009. Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. *PLoS Genet.* 5, 1–25. <https://doi.org/10.1371/journal.pgen.1000344>
- Touchon, M., Moura de Sousa, J.A., Rocha, E.P., 2017. Embracing the enemy: The diversification of microbial gene repertoires by phage-mediated horizontal gene transfer. *Curr. Opin. Microbiol.* 38, 66–73. <https://doi.org/10.1016/j.mib.2017.04.010>
- Troeger, C., Forouzanfar, M., Rao, P.C., Khalil, I., Brown, A., Reiner, R.C., Fullman, N., Thompson, R.L., Abajobir, A., Ahmed, M., Alemayohu, M.A., Alvis-Guzman, N., Amare, A.T., Antonio, C.A., Asayesh, H., Avokpaho, E., Awasthi, A., Bacha, U., Barac, A., Betsue, B.D., Beyene, A.S., Boneya, D.J., Malta, D.C., Dandona, L., Dandona, R., Dubey, M., Eshrati, B., Fitchett, J.R.A., Gebrehiwot, T.T., Hailu, G.B., Horino, M., Hotez, P.J., Jibat, T., Jonas, J.B., Kasaeian, A., Kisseon, N., Kotloff, K., Koyanagi, A., Kumar, G.A., Rai, R.K., Lal, A., El Razek, H.M.A., Mengistie, M.A., Moe, C., Patton, G., Platts-Mills, J.A., Qorbani, M., Ram, U., Roba, H.S., Sanabria, J., Sartorius, B., Sawhney, M., Shigematsu, M., Sreeramareddy, C., Swaminathan, S., Tedla, B.A., Jagiellonian, R.T.-M., Ukwaja, K., Werdecker, A., Widdowson, M.-A., Yonemoto, N., El Sayed Zaki, M., Lim, S.S., Naghavi, M., Vos, T., Hay, S.I., Murray, C.J.L., Mokdad, A.H., 2017. Estimates of global, regional, and national morbidity, mortality, and aetiologies of diarrhoeal diseases: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet Infect. Dis.* 17, 909–948. [https://doi.org/10.1016/S1473-3099\(17\)30276-1](https://doi.org/10.1016/S1473-3099(17)30276-1)
- Uber, A.P., Trabulsi, L.R., Irino, K., Beutin, L., Ghilardi, A.C.R., Gomes, T.A.T., Liberatore, A.M.A., De Castro, A.F.P., Waldir, P.E., 2006. Enteroaggregative *Escherichia coli* from humans and animals differ in major phenotypical traits and virulence genes. *FEMS Microbiol Lett* 256, 251–257. <https://doi.org/10.1111/j.1574-6968.2006.00124.x>

- Uysal, A., Durak, Y., 2012. Pulsed-field gel electrophoresis typing, antibiotic resistance, and plasmid profiles of *Escherichia coli* strains isolated from foods. *Can. J. Microbiol.* 58, 1278–1287. <https://doi.org/10.1139/w2012-108>
- van Elsas, J.D., Semenov, A. V, Costa, R., Trevors, J.T., 2011. Survival of *Escherichia coli* in the environment: fundamental and public health aspects. *ISME J.* 5, 173–183. <https://doi.org/10.1038/ismej.2010.80>
- Van Overbeek, L.S., Van Doorn, J., Wichers, J.H., Van Amerongen, A., Van Roermund, H.J.W., Willemsen, P.T.J., 2014. The arable ecosystem as battleground for emergence of new human pathogens. *Front. Microbiol.* 5, 1–17. <https://doi.org/10.3389/fmicb.2014.00104>
- Whitman, W.B., 2015. Genome sequences as the type material for taxonomic descriptions of prokaryotes. *Syst. Appl. Microbiol.* 38, 217–222. <https://doi.org/10.1016/j.syapm.2015.02.003>
- Wong, K.K.W., Olsson, A.L.J., Asadishad, B., Van der Bruggen, B., Tufenkji, N., 2017. Role of Cell Appendages in Initial Attachment and Stability of *E. coli* on Silica Monitored by Nondestructive TIRF Microscopy. *Langmuir* 33, 4066–4075. <https://doi.org/10.1021/acs.langmuir.7b00314>
- Wood, T.K., 2008. Minireview Insights on *Escherichia coli* biofilm formation and inhibition from whole-transcriptome profiling. *Environ. Microbiol.* 11, 1–15. <https://doi.org/10.1111/j.1462-2920.2008.01768.x>
- World Bank, 2014. Economics of South African townships : special focus on Diepsloot. World Bank, Washington, D.C, United States of America. <https://doi.org/10.1596/978-1-4648-0301-7>
- Xicohtencatl-Cortes, J., Chacón, E.S., Saldaña, Z., Freer, E., Girón, J.A., 2009. Interaction of *Escherichia coli* O157:H7 with Leafy Green Produce. *J. Food Prot.* 72, 1531–1537. <https://doi.org/10.4315/0362-028X-72.7.1531>
- Xue, X., Teare, M.D., Holen, I., Zhu, Y.M., Woll, P.J., 2009. Optimizing the yield and utility of circulating cell-free DNA from plasma and serum. *Clin. Chim. Acta* 404, 100–104. <https://doi.org/10.1016/j.cca.2009.02.018>