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

Matthew Aijuka[a]; Araceli E. Santiago[b]; Jorge A. Girón[b]; James P. Nataro[b] and Elna M. Buys[a]*

[a]Department of Food Science, University of Pretoria, Republic of South Africa

[b]Child Health Center, Department of Pediatrics, School of Medicine, University of Virginia and University of Virginia Children's Hospital, Charlottesville, Virginia, USA

*Corresponding author at Department of Food Science, University of Pretoria, Lynnwood Rd, Pretoria, 0002, South Africa

Email addresses: aijumatt@yahoo.com (M. Aijuka), aes8j@hscmail.mcc.virginia.edu (A. E Santiago), jagiron@yahoo.com (J. A Girón), jpn2r@hscmail.mcc.virginia.edu (J. P Nataro), elna.buys@up.ac.za (E. M Buys)

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 *in-vitro*. Subsequently, all strains having data points for all analyses were used to compute Principal Component Analysis (PCA) plot curves to infer potential 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

List of Abbreviations: Intestinal Dysfunction-ID; DEC-Diarrheagenic *Escherichia coli*; EAEC-enteroaggregative *E. coli*; Producer Distributor Bulk Milk (PDBM); TER-Trans epithelial electrical resistance; IL-8- Interleukin-8; LB- Luria-Bertani; DMEM-F12: Dulbecco's Modified Eagle's Medium and Ham's F12 Nutrient Mixture (1:1); DMEM-HG-DMEM high glucose; EAEC-enteroaggregative *E. coli*; EPEC-enteropathogenic *E. coli*; DAEC-diffusely adherent *E. coli*; PBS-Peptone buffer saline; TEM-Transmission Electron Microscopy; DDSA-Dodecynlsuccinic anhydride; NMA-Methyl Nadic Anhydride; FBS-Fetal Bovine Serum; ANOVA-Analysis of Variance; OD-Optical Density; AAF-Aggregative Adherence Fimbriae; TMB-Tetramethylbenzidine; EPS-Extracellular Polymeric Substance

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 *Enterobacteriacea*e 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 adult hood. 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 may 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 by persisting within the human gastrointestinal tract 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. Although still explorative and far from definitive, our study may help provide another informative avenue for food safety and health officials to look beyond outbreak pathogens when investigating the aetiology of foodborne illness associated with pathogenic *E. coli*.

Our study is a follow up from our recent work (Aijuka et al. 2018) that has shown that enteroaggregative *E. coli* (EAEC) is the predominant Diarrheagenic *E. coli* (DEC) pathotype isolated from food sources and irrigation water in South Africa. However, despite the high prevalence of EAEC compared to other DEC pathotypes, most of the *E. coli* strains screened (>90%) did not fall into the major DEC categories (non-DEC) commonly associated with food or waterborne illness in spite of being collected over multiple years and across the different provinces of South Africa. Therefore, we sought to use simple phenotypic pathogenic indicators for determining whether non-DEC strains potentially posed a risk to human health since standard diagnostic measures could not clearly provide this distinction. Therefore, our present study rather than draw definitive conclusions about the risk posed by all individual non-DEC to human health would serve as a launch pad for more in-depth characterization.

2. Materials and methods

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 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 and were subsequently denoted as non-DEC while those positive for any of the characteristics associated with DEC were described as such. The number of strains tested for each assay differed, depending on prior molecular and phenotypic characterization that signalled potential pathogenicity. In other words, only strains from each assay that showed a positive result from a previous assay was used in a

Table 1 Strains, sources of isolation, pathotypes and assays used for characterizing *Escherichia coli* isolated from food sources and irrigation water in South Africa

Strain code[1]	**Source**	***E. coli* 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 (*aaiC*)	-	+	-	+
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-Transepithelial electrical resistance. Strains showing TER reduction were subsequently tested for interleukin-8 (IL-8) induction.

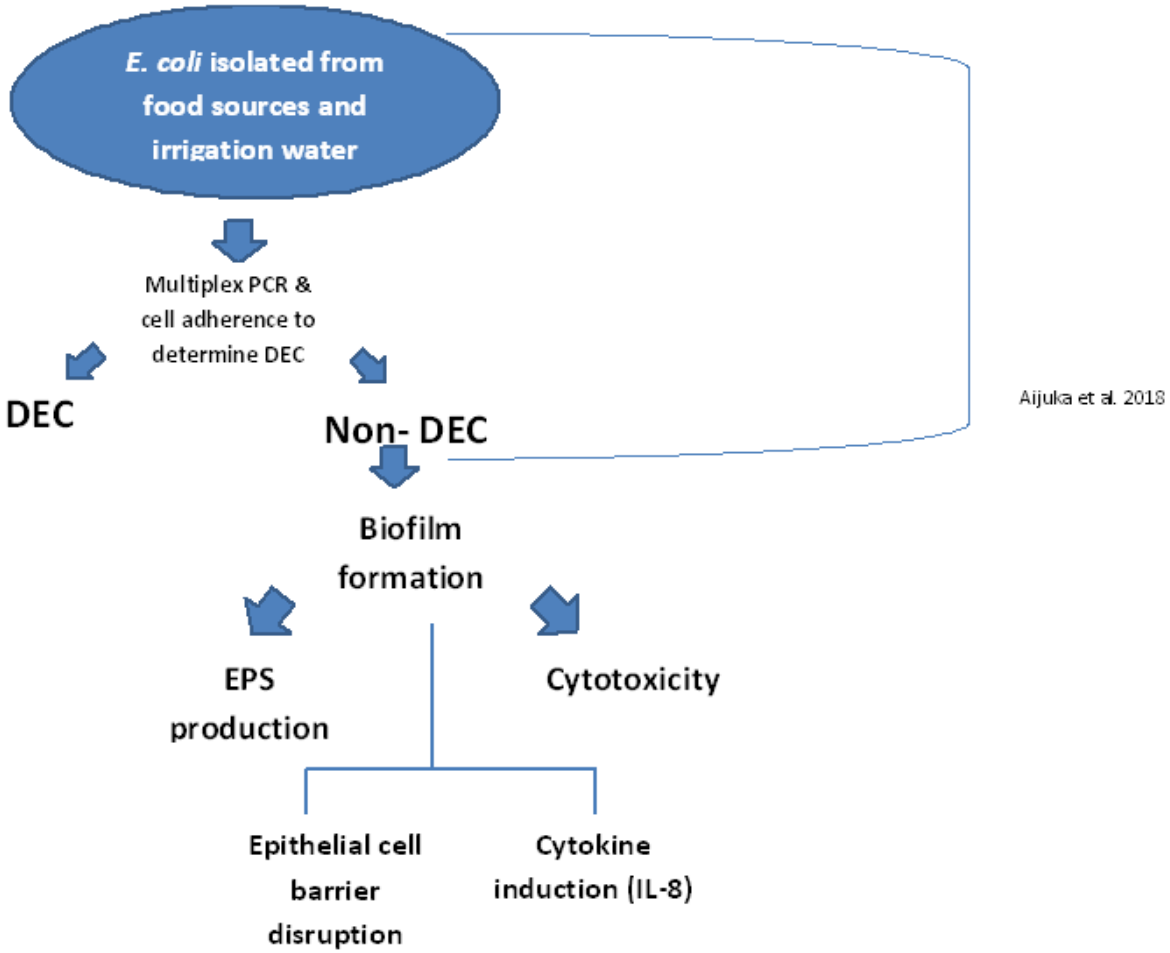

Figure 1. The strain selection strategy used for characterizing non-DEC previously isolated from food sources and irrigation water in South Africa for phenotypes potentially associated with intestinal dysfunction *in-vitro*. DEC, Diarrheagenic *Escherichia coli*; EPS, Extracellular polymeric substance; IL-8, interleukin 8

downstream test in order to remain with strains that had data points for multiple assays and thereby a better correlation with the risk for pathogenicity. The following systematic order of strain exclusion was followed for screening. Strains showing qualitative extracellular polymeric substance formation were subsequently tested for biofilm formation, cytotoxicity, reduction in transepithelial electrical resistance and induction of IL-8 in that descending order (Figure 1). Admittedly, this method of strain characterization limited the comparison and subsequently how much we can infer from our data. However, it provided a first step in phenotypically characterizing foodborne strains that would have previously been designated as 'safe' (non-DEC) based on molecular and tissue culture characterization (Aijuka et al., 2018). The strains had been isolated from producer distributor bulk milk (PDBM), irrigated lettuce, irrigation water and street vendor coleslaw purchased in Pretoria, South Africa.

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 (30µL) of each strain were added and incubated for 6h. Washing was done 3X with PBS(Sigma-Aldrich) after which 500 µL 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.

2.2.1 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. 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.

2.2.2.1 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.

2.3 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. 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 CO_2 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 CO_2 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 (dH_20) and stained with 500 µL of a Giemsa staining solution (Fisher Scientific) for 20 min. The samples were rinsed (3X) with dH_20 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.

2.4 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 $3x10^5$ cells/mL onto collagen coated 12-mm polycarbonate Transwell permeable support cell culture inserts (0.4μm 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 $1000\Omega/cm^2$ and not more than $2000\Omega/cm^2$. 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% CO_2 for 60 min. Overnight cultures of bacteria were standardized in DMEM-HG to OD_{600} of 0.30±0.02, which is

approximately $2x10^8$ to $4x10^8$ CFU/mL. D-Mannose was added to each culture to a concentration of 1%.

For each sample, 100 µL 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% CO_2, 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 µg/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 (OD_{450}-OD_{550}) was measured using a microplate reader (Biochrome® Anthos MultiRead 400).

2.5 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.

2.5.1 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.

3. Results

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 2.). 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 PBDM (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.

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 3.). 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 3.). 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.

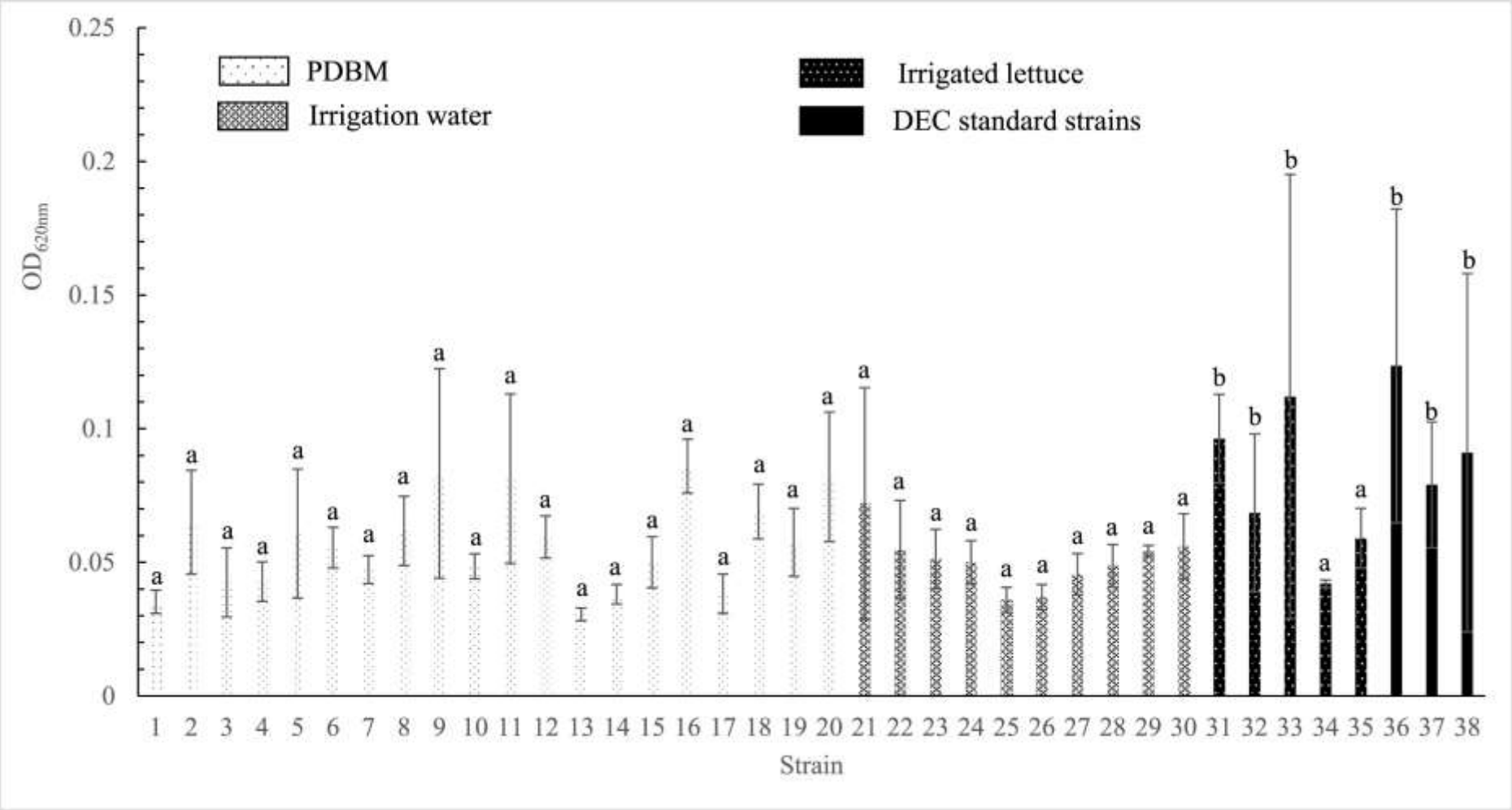

Figure 2. Biofilm formation in Luria Bertani broth by diarrheagenic E. coli (DEC) and non-DEC isolated from food sources and irrigation water. DEC strains: 1 to 6 (PDBM); 31 to 33 (irrigated lettuce); 36 to 38 (DEC standard strains). Non-DEC: 7 to 20 (PDBM); 21 to 30 (irrigation water); 34 to 35 (irrigated 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. Strain codes in the figure have been changed to Arabic numerals to save on space for comparison. Otherwise, alpha-numerical strain codes in the legend represent the original codes for the strains used throughout the study as previously shown in Table 1.

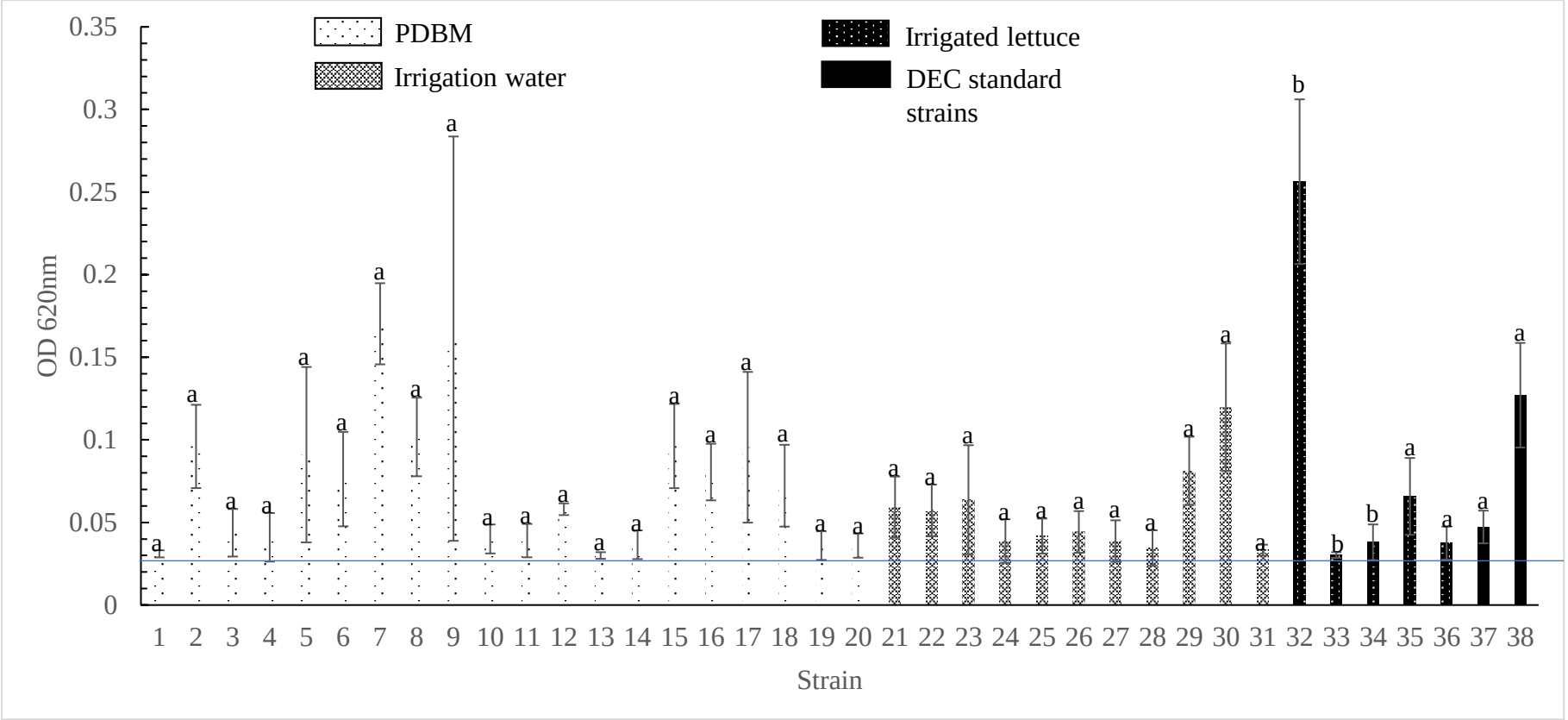

Figure 3. Biofilm formation in Dulbecco's Modified Eagle Medium-F12+0.5% Mannose by diarrheagenic *E. coli* (DEC) and non-DEC previously isolated from food sources and irrigation water in South Africa. DEC strains: 1 to 6 (PDBM); 32 to 34 (irrigated lettuce); 37 to 38 (DEC standard strains). Non-DEC: 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 *E. coli* 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. Strain codes in the figure have been changed to Arabic numerals to save on space for comparison. Otherwise, alpha-numerical strain codes in the legend represent the original codes for the strains used throughout the study as previously shown in Table 1.

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.

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.).

3.3.1 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 5.). 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 based on cytotoxic activity

3.3 .2 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 6.). For all strains assayed, IL-8 induction varied with isolation source (Figure 7.).

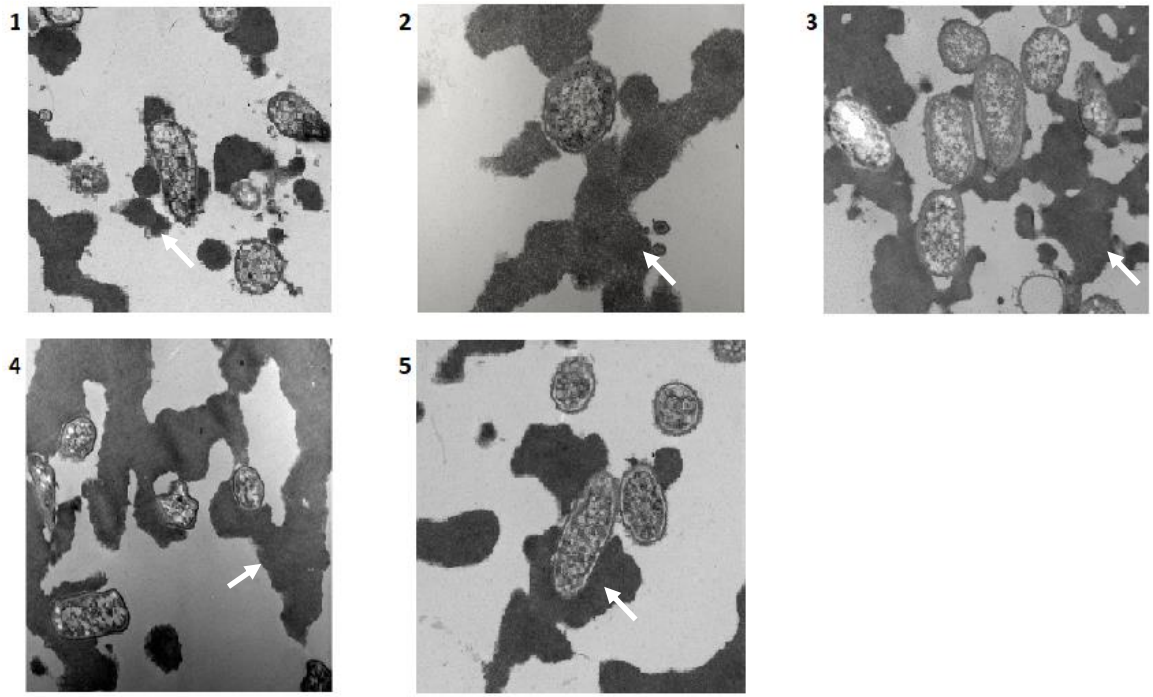

Figure 4. Extracellular polymeric substance (EPS) formation (amorphous extracellular mass attached to bacterial wall) by diarrheagenic (DEC) and non-DEC previously isolated from food sources and irrigation water in South Africa visualized under Transmission Electron Microscopy. Arrows point to EPS. Strain 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. Strain codes in the figure have been changed to Arabic numerals to save on space for comparison. Otherwise, alpha-numerical strain codes in the legend represent the original codes for the strains used throughout the study as previously shown in Table 1.

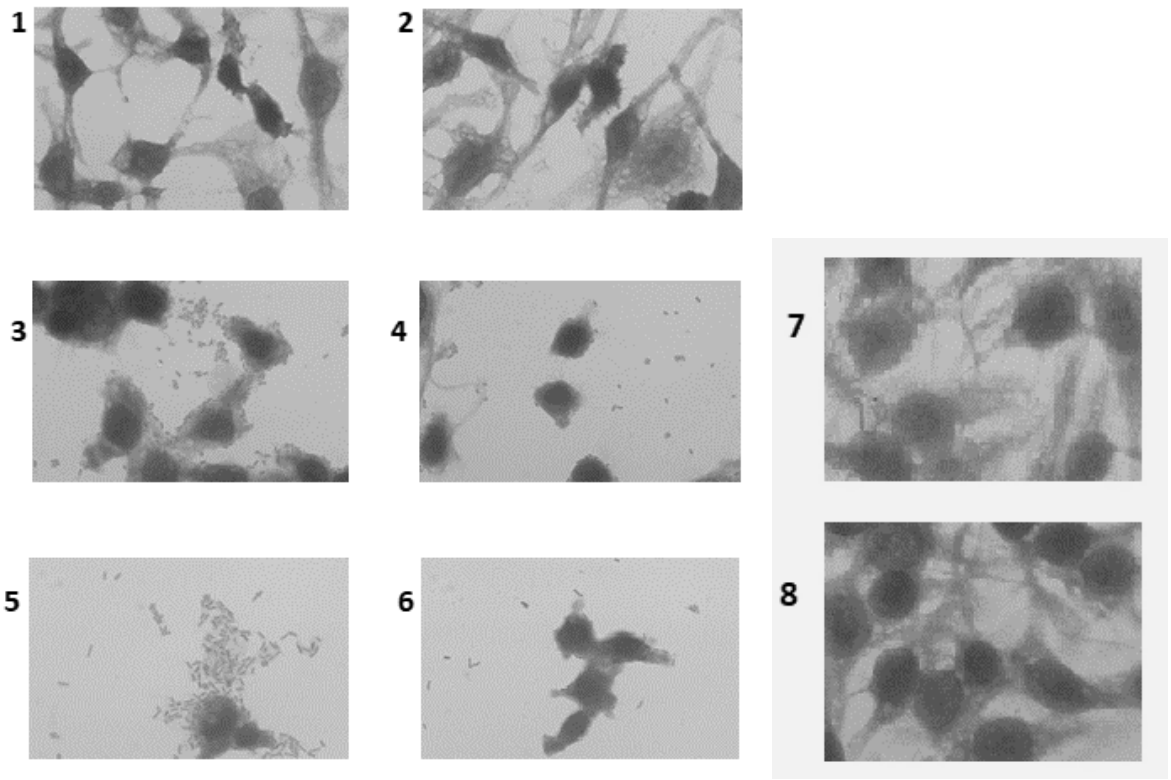

Figure 5. Cytotoxic activity (disruption of the HeLa cell cytoskeleton) in diarrheagenic *E. coli* (DEC) and Non-DEC previously isolated from food sources and irrigation water in South Africa. Images recorded at 60X. Scale=20μm. Strain pathotype and source of isolation: 1-Non-DEC isolated from coleslaw; 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. Strain codes in the figure have been changed to Arabic numerals to save on space for comparison. Otherwise, alpha-numerical strain codes in the legend represent the original codes for the strains used throughout the study as previously shown in Table 1.

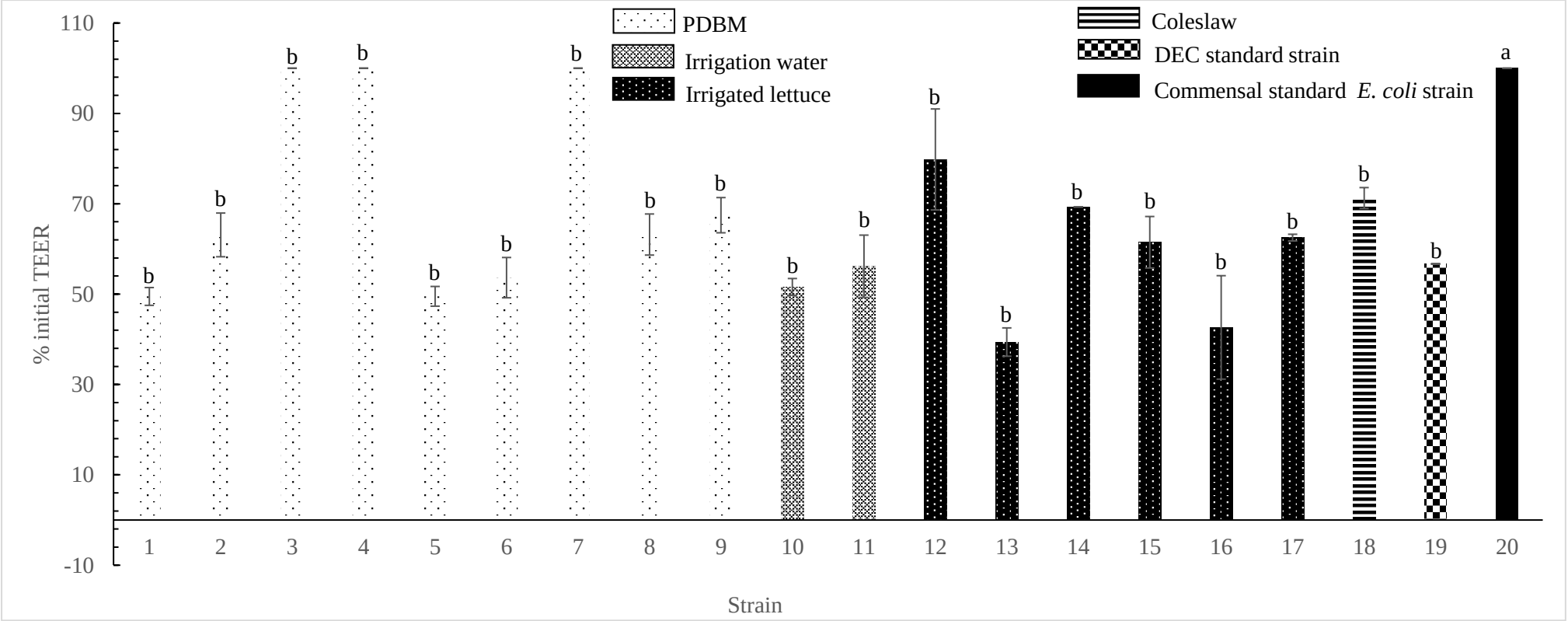

Figure 6. Percentage reduction in initial Transepithelial Electrical Resistance (TER) observed in polarized T-84 epithelial cells infected with diarrheagenic (DEC) and non-DEC previously isolated from food sources and irrigation water in South Africa. Strain pathotype and source of isolation: DEC strains: 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. Strain codes in the figure have been changed to Arabic numerals to save on space for comparison. Otherwise, alpha-numerical codes in the legend represent the original codes for the strains used throughout the study as previously shown in Table 1.

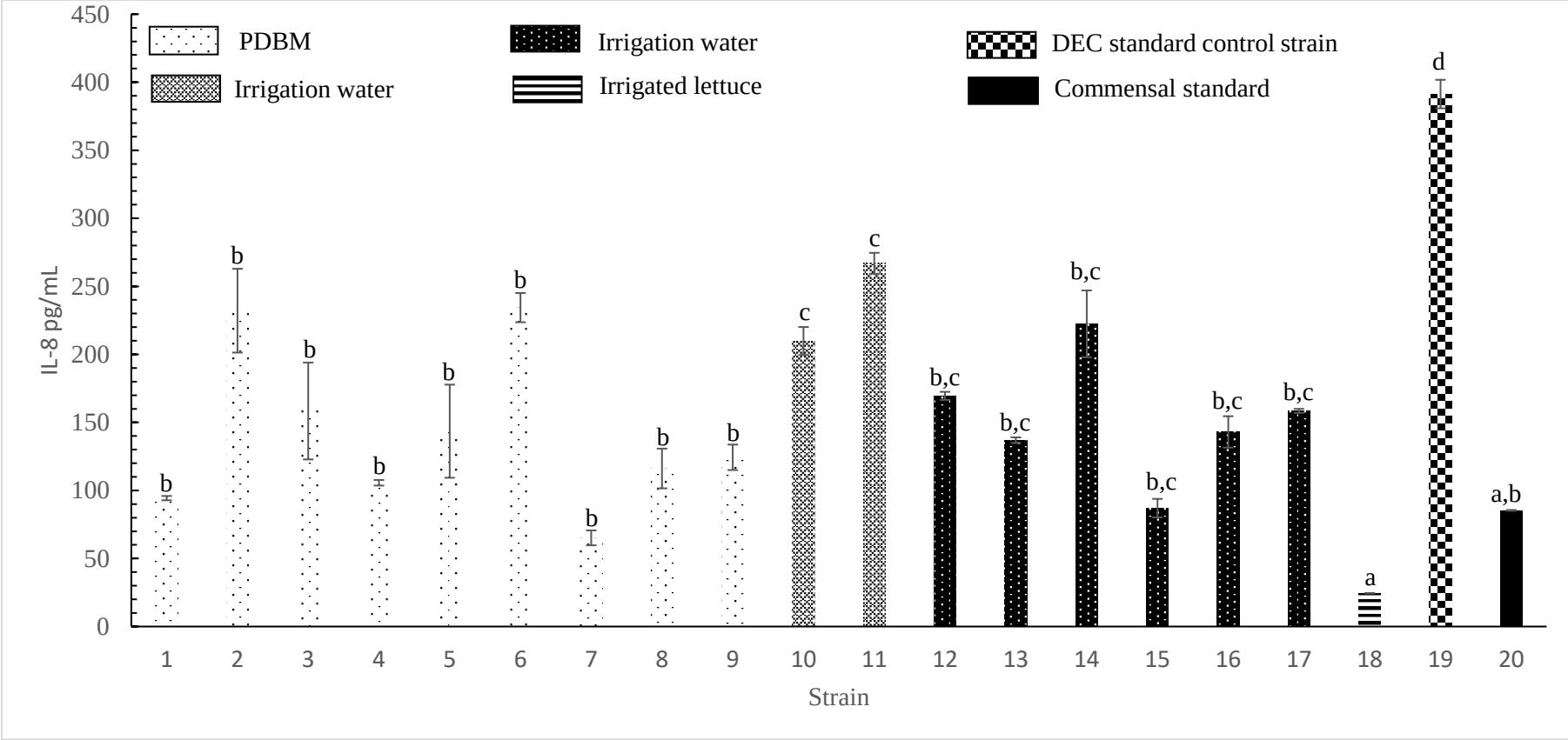

Figure 7 . Interleukin 8 (IL-8) induction in polarized T-84 epithelial cells resulting from infection with diarrheagenic *E. coli* (DEC) and non-DEC previously isolated from food sources and irrigation water in South Africa. Strain pathotype and source of isolation: DEC strains: 1 to 4 (PDBM); 12 to 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. Strain codes in the figure have been changed to Arabic numerals to save on space for comparison. Otherwise, alpha-numerical strain codes in the legend represent the original codes for the strains used throughout the study as previously shown in Table 1.

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.

3.3.3 Association of DEC and non-DEC with factors potentially 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 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 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

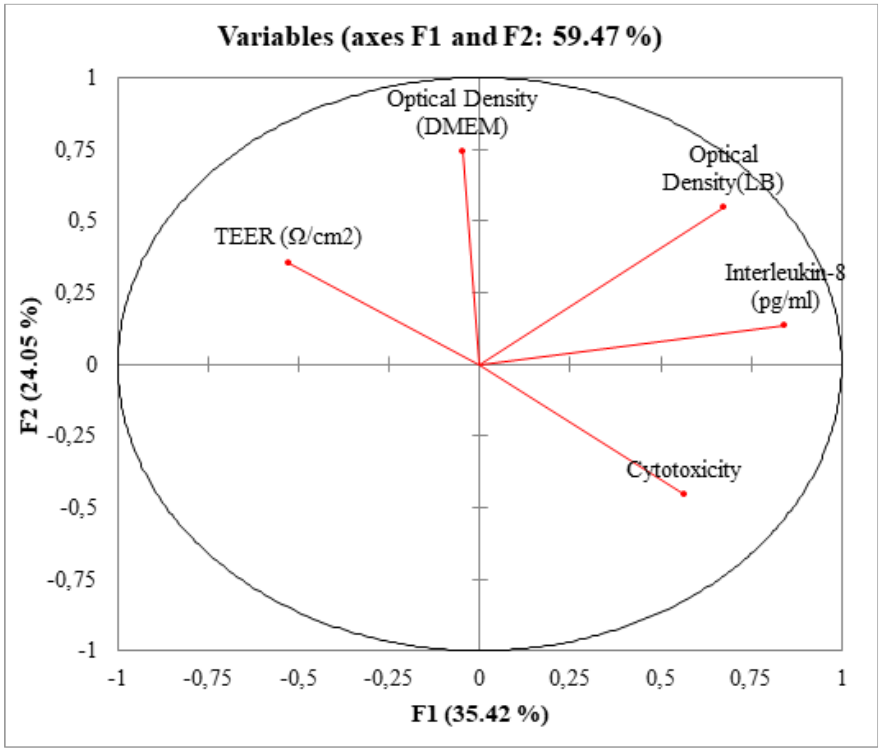

Figure 8. Principal Component Analysis variables correlation plot showing axes Factors 1 and 2 (F1and F2) accounting for approximately 60% of variation observed after analysis of 4 factors in diarrheagenic *E. coli* (DEC) and non-DEC tested in-vitro potentially associated with human gut intestinal dysfunction. 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-Transepithelial electrical resistance.

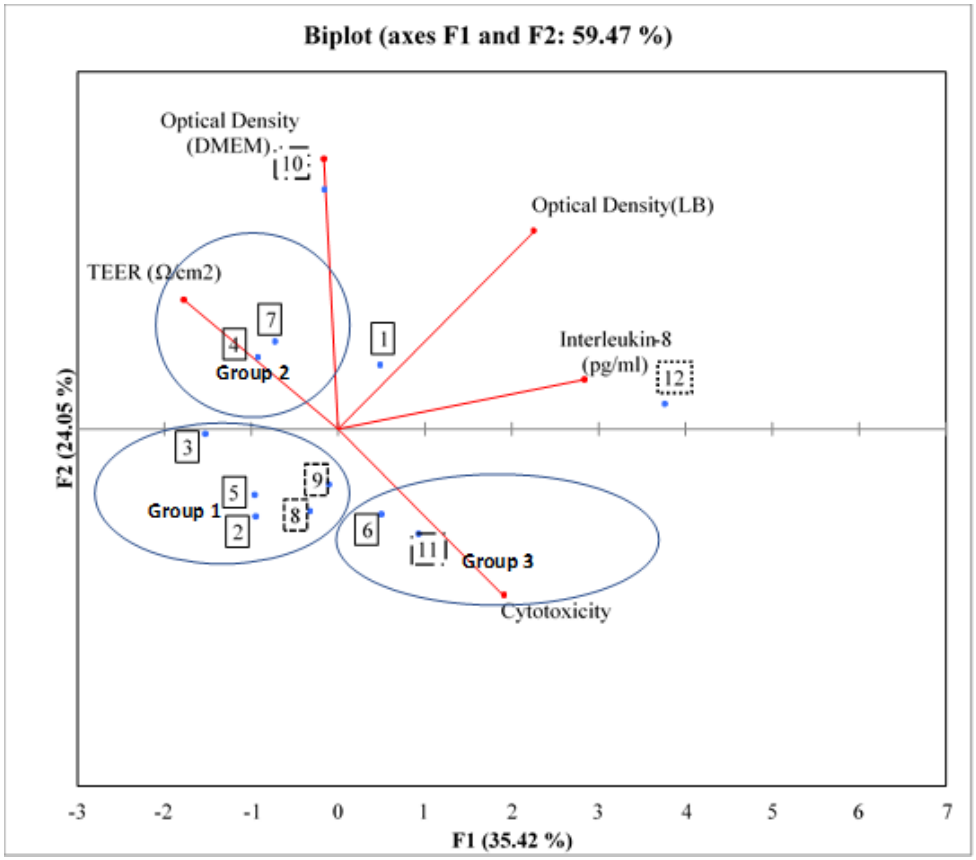

Figure 9. Principal Component Analysis biplot showing axes Factors 1 and 2 (F1 and F2) accounting for approximately 60% of variation observed in 11 strains of diarrheagenic (DEC) and non-DEC isolated from food sources and irrigation water. The strains were analysed in-vitro for 4 factors potentially 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). Strain codes in the figure have been changed to Arabic numerals *to* save on space for comparison. Otherwise, alpha-numerical strain codes in the legend represent the original codes for the strains used throughout the study as previously shown in Table 1.

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 elimination based and meant to carry out analyses only of strains that had previously tested positive for a phenotypic or genotypic characteristic potentially liked to pathogenicity *in-vitro* for example 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.

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 Incl1 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 elimination-based approach of the study that only subsequently tested strains positive for the potential pathogenic parameters of interest in this study. 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.

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, these effects 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. Our aim is to further characterize these strains based on multiple pathogenic parameters to yield more conclusive data regarding the link between non-DEC and ID *in-vitro*.

Conflict of interest

The authors declare no conflict of interest

Acknowledgements

The Department of Research and Innovation, University of Pretoria for a post-graduate travel bursary to Matthew Aijuka to travel to James P. Nataro's laboratory at The University of Virginia. Work in the Nataro lab was supported by US National Institutes of Health grant AI-33096 to JPN.

References

Aijuka, M., Charimba, G., Hugo, C. J., & Buys, E. M. (2015). Characterization of bacterial pathogens in rural and urban irrigation water. *Journal of Water and Health*, *13*(1), 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). 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. https://doi.org/10.1016/j.ijfoodmicro.2018.04.018

Beloin, C., Roux, A., & Ghigo, J. (2008). Escherichia coli biofilms. *Current Topics in Microbiology and Immunology*, *322*, 249–289. https://doi.org/https://doi.org/10.1007/978-3-540-75418-3_12

Borgersen, Q., Bolick, D. T., Kolling, G. L., Aijuka, M., Ruiz-Perez, F., Guerrant, R. L., … 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

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. *Clinical Microbiology Reviews*, *26*(4), 822–880. https://doi.org/10.1128/CMR.00022-13

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 enteroaggregative Escherichia coli. *PLoS ONE*, *6*(10), 1–15. https://doi.org/10.1371/journal.pone.0027030

Estrada-Garcia, T., & Navarro-Garcia, F. (2012). Enteroaggregative Escherichia coli pathotype: A genetically heterogeneous emerging foodborne enteropathogen. *FEMS*

Immunology and Medical Microbiology, *66*(3), 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. *Proceedings of the National Academy of Sciences*, *112*(3), 633–640. https://doi.org/10.1073/pnas.1418781112

Flemming, H. C., & Wingender, J. (2010). The biofilm matrix. *Nature Reviews Microbiology*, *8*(9), 623–633. https://doi.org/10.1038/nrmicro2415

Guinanane, C. M., & Cotter, P. D. (2013). Role of the gut microbiota in health and chronic gastrointestinal disease: understanding a hidden metabolic organ Caitriona. *Ther Adv Gastroenterol*, *6*(4), 295–308. https://doi.org/10.1177/

Kau, A. L., Planer, J. D., Liu, J., Rao, S., Yatsunenko, T., Trehan, I., … Gordon, J. I. (2015). Functional characterization of IgA-targeted bacterial taxa from malnourished Malawian children that produce diet-dependent enteropathy. *Sci Transl Med*, *25*(7), 276. https://doi.org/10.1126/scitranslmed.aaa4877

Kim, M., Ashida, H., Ogawa, M., Yoshikawa, Y., Mimuro, H., & Sasakawa, C. (2010). Bacterial interactions with the host epithelium. *Cell Host & Microbe*, *8*(1), 20–35. https://doi.org/10.1016/j.chom.2010.06.006

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. *Journal of Clinical Microbiology*, *45*(1), 121–126. https://doi.org/10.1128/JCM.01128-06

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. *International Journal of STD & AIDS*, *23*(5), 346–350. https://doi.org/10.1258/ijsa.2011.011285

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 toxigenic enteroaggregative Escherichia coli O104:H4. *International Journal of Food Microbiology*, *229*, 44–51. https://doi.org/10.1016/j.ijfoodmicro.2016.04.007

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. *Pediatric Infectious Disease Journal*, *6*(9), 829–831. https://doi.org/10.1097/00006454-198709000-00008

Ntuli, V., Njage, P. M. K., & Buys, E. M. (2016). Characterization of Escherichia coli and other Enterobacteriaceae in producer-distributor bulk milk. *Journal of Dairy Science*, *99*(12), 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. *Journal of Food Protection*, *81*(3), 472–481. Retrieved from http://jfoodprotection.org/doi/pdf/10.4315/0362-028X.JFP-17-199

Ochieng, J. B., Boisen, N., Lindsay, B., Santiago, A., Ouma, C., Ombok, M., … Nataro, J. P. (2014). Serratia marcescens is injurious to intestinal epithelial cells. *Gut Microbes*, *5*(6), 729–736. https://doi.org/10.4161/19490976.2014.972223

Panchalingam, S., Antonio, M., Hossain, A., Mandomando, I., Ochieng, B., Oundo, J., …

Nataro, J. P. (2012). Diagnostic microbiologic methods in the GEMS-1 case/control study. *Clinical Infectious Diseases*, *55*(SUPPL. 4), 294–302. https://doi.org/10.1093/cid/cis754

Petrova, O. E., & Sauer, K. (2012). Sticky situations: Key components that control bacterial surface attachment. *Journal of Bacteriology*, *194*(10), 2413–2425. https://doi.org/10.1128/JB.00003-12

Pratt, L. A., & Kolter, R. (1998). Genetic analysis of Escherichia coli biofilm formation: roles of flagella, motility, chemotaxis and type I pili. *Molecular Microbiology*, *30*(2), 285–293. https://doi.org/10.1046/j.1365-2958.1998.01061.x

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. *Current Topics in Microbiology and Immunology*, *365*, 49–81. https://doi.org/10.1007/82

Ruiz-Perez, F., & Nataro, J. P. (2014). Bacterial serine proteases secreted by the autotransporter pathway: classification, specificity, and role in virulence. *Cellular and Molecular Life Sciences : CMLS*, *71*(5), 745–770. https://doi.org/10.1007/s00018-013-1355-8

Shin, N.-R., Whon, T. W., & Bae, J.-W. (2015). Proteobacteria: microbial signature of dysbiosis in gut microbiota. *Trends in Biotechnology*, *33*, 496–503. https://doi.org/10.1016/j.tibtech.2015.06.011

Strauman, M. C., Harper, J. M., Harrington, S. M., Boll, E. J., & Nataro, J. P. (2010). Enteroaggregative Escherichia coli Disrupts Epithelial Cell Tight Junctions. *Infection and Immunity*, *78*(11), 4958–4964. https://doi.org/10.1128/IAI.00580-10

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. *Emerging Infectious Diseases*, *18*(8), 1314–1317. https://doi.org/10.3201/eid1808.111616

Wood, T. K. (2008). Minireview Insights on Escherichia coli biofilm formation and inhibition from whole-transcriptome profiling. *Environmental Microbiology*, *11*, 1–15. https://doi.org/10.1111/j.1462-2920.2008.01768.x

Highlights

- Non-DEC formed biofilms, cytotoxins, disrupted epithelial cells and induced interleukin-8
- Epithelial cell disruption was comparable to a standard DEC strain EAEC 042
- PCA clustered strains based on isolation source showing potential use in source tracking
- Further studies with more strains required to provide clearer link of Non-DEC in causing ID

Graphical abstract