Prevalence, characterisation and potential origin of *Escherichia coli* found in surface and ground waters utilized for irrigation of fresh produce

by
Nika Anna Schoeman

Thesis is presented in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE IN FOOD SCIENCE

In the Department of Food Science, Faculty of AgriSciences
University of Stellenbosch

Supervisor:

Professor T.J. Britz

Co-supervisors:

Dr G.O. Sigge and Dr C. Lamprecht

March 2013

Stellenbosch University http://scholar.sun.ac.za

DECLARATION

By submitting this thesis electronically, I declare that the entirety of the work contained herein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously submitted it, in its entirety or in part, for obtaining any qualification.

____________________________	____________________________
Nika Schoeman	Date

Copyright © 2013 Stellenbosch University

All rights reserved

Stellenbosch University http://scholar.sun.ac.za

ABSTRACT

Over the past two decades, there has been an increase in the use of water sources for irrigation, as well as an increase in *Escherichia coli* outbreaks linked to fresh produce. The full extent and type of *E. coli* contamination present in natural water sources is unknown and the contamination sources have also not been confirmed. The aim of this study was to enumerate and characterise *E. coli* from both irrigation water and potential contamination source sites.

Total coliform and *E. coli* counts found in contamination source sites were as high as log 7.114 and log 6.912 MPN.100 mL^{-1}, respectively. Total coliform and *E. coli* counts for irrigation sites were lower, with maximum counts of log 5.788 and log 5.768 MPN.100 mL^{-1}, respectively. It was found that more than one third (5/14 = 35.71%) of the irrigation sites had *E. coli* counts exceeding the guidelines (<1 000 counts.100 mL^{-1}) for 'safe' irrigation water for fresh produce (<1 000 counts.100 mL^{-1}) as set by the Department of Water Affairs (DWA) and World Health Organisation (WHO), making the water unsuitable for the irrigation of fresh produce.

Phylogenetic subgroups (A_0, A_1, B1, $B2_2$, $B2_3$, D_1 and D_2) and the MALDI Biotyper system (PCA dendrogram) were used to create a fingerprint of each *E. coli* isolated from the environment. These were then used to link *E. coli* strains from irrigation water to their most probable contamination origin. *Escherichia coli* population structure was found in this study, to be better suited for linking *E. coli* strains from irrigation water to their most likely source, than just applying the phylogenetic grouping. The MALDI Biotyper data in combination with the phylogenetic subgroup assignment was then used to group similar strains and link *E. coli* from irrigation water to their contamination sources by comparing population structures. Strains isolated from surface and groundwater showed similar distribution patterns, but groundwater strains showed a population structure more indicative of porcine and bovine origin, while surface water showed population characteristics which could not be used to make conclusive links between the irrigation water and suspected contamination sources.

When investigating the population structures of individual sample sites, it was found that phylogenetic subgroups A_0, A_1 and B1 frequently made up the bulk of the *E. coli* population. It was also found that linking individual irrigation sites to contamination sources was successful, as irrigation site Berg-2 was found to have a similar population structure to contamination source site Plank-1 which represents human pollution from an informal settlement. This led to the

Stellenbosch University http://scholar.sun.ac.za

conclusion that Berg-2 was being contaminated by human pollution, most probably from an informal settlement. Upon further investigation it was found that Berg-2 is downstream of an informal settlement, proving that *E. coli* population structure is a successful means of microbial source tracking (MST).

Virulence factors of the 153 *E. coli* isolated during the study were identified and the potential risk associated with using the investigated irrigation water for irrigation of fresh produce, was determined. Two enteropathogenic *E. coli* (EPEC) strains were isolated from the irrigation water, one from the Plankenburg River water, and the other from a borehole in the Drakenstein area. The latter indicates that borehole water is not as safe as was once thought, and that there are bacterial contaminants finding their way into groundwater. The occurrence of an EPEC strain in river water shows that neither ground nor surface water is guaranteed to be safe, and that treatment of water being used for the irrigation of fresh produce should be implemented.

Stellenbosch University http://scholar.sun.ac.za

UITTREKSEL

Oor die afgelope twee dekades was daar nie net 'n toename in die gebruik van alternatiewe waterbronne vir besproeiing nie, maar ook 'n toename in uitbrake van *Escherichia coli* uitbrake wat aan vars produkte gekoppel kan word. Die tipe *E. coli*-besmetting wat in natuurlike waterbronne teenwoordig is, is onbekend en die besmettingsbron is ook nog nie bevestig nie. Daarom was die doel van hierdie studie om die voorkomssyfer van *E. coli* van beide besproeiingswater en potensiële kontaminasiebronne te bepaal, asook om die *E. coli* te karakteriseer.

Totale kolivorme en *E. coli*-tellings wat in kontaminasiebronne gevind is, het 'n maksimum van log 7,114 en log 6,912 MPN.100 mL^{-1} onderskeidelik bereik, terwyl die totale kolivorme en *E. coli*-tellings vir besproeiingswater laer was, met 'n maksimum van log 5,788 en 5,768 MPN.100 mL^{-1}, onderskeidelik. Dit is bevind dat meer as 1/3 (5/14 = 35,71%) van die besproeiingswaterbronne meer *E. coli* bevat as wat toegelaat word in die riglyne vir "veilige" besproeiingswater vir vars produkte (<1 000 fekale koliforme.100 mL^{-1}) wat deur die Departement Waterwese (DWA) en die Organisasie vir Wêreldgesondheid (WHO) aanbeveel word.

Filogenetiese subgroepe (A_0, A_1, B1, $B2_2$, $B2_3$, D_1 en D_2) en die 'MALDI Biotyper'-stelsel (PKA dendrogram) is gebruik om unieke profiele vir elke geïsoleerde *E. coli* te skep. Dié profiele is daarna gebruik om *E. coli*-stamme van besproeiingswater te koppel aan die mees waarskynlike oorsprong van kontaminasie. Daar is in hierdie studie bevind dat die *E. coli*-populasiestruktuur beter geskik was vir die koppeling van *E. coli*-stamme van besproeiingswater na hul mees waarskynlikste bron, as net die toepassing van die filogenetiese groepering. Dit was toe gevind dat *E. coli* wat uit oppervlak- en grondwater geïsoleer is, soortgelyke verspreidingspatrone het, maar grondwaterstamme se bevolkingstruktuur is meer aanduidend van fekale besmetting deur varke en beeste, terwyl oppervlakwater se bevolkingseienskappe nie duidelik genoeg was om 'n gevolgtrekking oor moontlike bronne van besmetting te maak nie.

Toe die populasiestruktuur van die individuele bemonsteringspunte ondersoek is, is daar bevind dat die filogenetiese subgroepe A_0, A_1 en B1 dikwels die meeste tot die *E. coli* bevolking bydra. Daar is ook bevind dat die koppeling van isolate in individuele besproeiingswaterbronne met hul mees waarskynlike bronne van kontaminasie suksesvol was. Besproeiingswater van Berg-2 het 'n soortgelyke populasiestruktuur as Plank-1 wat beskou is as 'n kontaminasiebron. Dit het gelei tot die gevolgtrekking dat Berg-2 heel waarskynlik deur menslike besoedeling beïnvloed

Stellenbosch University http://scholar.sun.ac.za

word, soos Plank-1, en dat daar moontlik ook 'n informele nedersetting by Berg-2 betrokke is. Na verdere ondersoek is gevind dat Berg-2 inderdaad ook stroomaf van 'n ander informele nedersetting geleë is, wat bewys dat die *E. coli*-populasiestruktuur 'n suksesvolle manier is om *E. coli* kontaminasie te verbind met besproeiingswater.

Patogeniese faktore, wat in *E. coli* voorkom en maagkieme veroorsaak, is vooraf getoets in elkeen van die 153 *E. coli*-isolate wat tydens die studie geïdentifiseer is. Twee 'enteropathogenic' *E. coli* (EPEC)-stamme is uit die besproeiingswater geïsoleer: een uit die Plankenburgrivier (Plank-3), en die ander uit 'n boorgat in die Drakenstein-gebied (Boorgat A1). Hierdie inligting dui aan dat boorgatwater nie so veilig is as wat voorheen vermoed is nie, en dat bakteriese kontaminasie wel vookom wat nie alleen die grondwater besmet nie, maar ook daarin oorleef.

Die voorkoms van die EPEC-stamme in hierdie studie is 'n aanduiding dat beide grond- en opppervlakwater ewe gevaarlik kan wees, en dat daar dus geen waarborg vir die veiligheid van die water is nie. Die behandeling van grond- en oppervlakwater, wat vir die besproeiing van vars produkte gebruik word, moet daarom ernstig oorweeg word om moontlike uitbrake van *E. coli* op vars produkte te verhoed.

Stellenbosch University http://scholar.sun.ac.za

ACKNOWLEDGEMENTS

I would like to extend my sincerest gratitude and appreciation to the follow people and institutions for all their help, time, hard work and funding that went into this study:

Prof Trevor Britz for the guidance, motivation and endless hours of work as my supervisor

Dr Gunnar Sigge and Dr Corné Lamprecht for the advice and invaluable input as my co-supervisors

The National Research Foundation (NRF-Incentive Funding and Scarce Skills Bursary) for their financial contribution towards this study

SAAFoST (Food-Bev SETA bursary) for their financial contribution towards this study

This study was part of an ongoing solicited research project (K5/1773) (A quantitative investigation into the link between irrigation water quality and food safety), funded and managed by the Water Research Commission and co-funded with the Department of Agriculture;

Carel Muller, Wendy Lilje, John Morris, Lario Moolman, Layne Lategan, Nadine Pretorius, Stephan du Toit, Ian Grimbacher, Mynhardt Theron, Deidré February and any other organisations, companies and individuals who gave inputs through personal communication, wonderful suggestions and advice, additional funding, and allowing sampling on their premises or by commenting on documents and workshop results.

My lab partners Marlize Jordaan, Amanda Brand and Anneri Carinus for the love, friendship and willingness to help, the time spent together made every moment worthwhile

Students and staff of the Department of Food Science for their assistance, friendship and motivation

Finally to my family and friends who always stood by me and helped me though the difficult times, your love, encouragement and support were unwavering throughout my studies.

Stellenbosch University http://scholar.sun.ac.za

TABLE OF CONTENTS

Declaration		iii
Abstract		iii
Uittreksel		v
1s		iiii
Chapter 1:	Introduction	1
Chapter 2:	Literature Review	7
Chapter 3:	Enumeration and characterisation of *Escherichia coli* from irrigation water and potential contamination sites	42
Chapter 4:	Evaluation of phylogenetic groupings and abundant ribosomal proteins as MST markers for assessment of the origin of faecal contamination of irrigation water	85
Chapter 5:	General discussion and conclusions	131

Language and style used in this thesis are in accordance with the requirements of the *International Journal of Food Science and Technology*. This thesis represents a compilation of manuscripts where each chapter is an individual entity and some repetition between chapters has, therefore, been unavoidable.

Stellenbosch University http://scholar.sun.ac.za

CHAPTER 1
INTRODUCTION

Worldwide, especially in recent years, the increasing levels of faecal coliform bacteria found in natural water systems being used for both direct and indirect human use has become a major health risk. Various practices cause the faecal contamination in the rivers, but overall runoff from cattle farms as well as that from overloaded municipal sewage plants and general pollution as a result of inadequate sewage facilities in rural settlements all play major roles in causing contamination.

Escherichia coli, along with other pathogenic microorganisms such as *Listeria* and *Salmonella* spp. are major contaminants in fresh water systems (Müller *et al.,* 2001; Melloul *et al.,* 2002; Islam *et al.,* 2004; Ackerman, 2010; Lötter, 2010; Linscott, 2011). Although many studies have been undertaken to determine total coliforms present in irrigation water used in South Africa (Steynberg *et al.,* 1995; Ackerman, 2010; Lötter, 2010), not many studies have looked specifically at the *E. coli* loads in the water. *Escherichia coli* loads in these waters are of great importance, as *E. coli* can lead to a variety of gastro-intestinal diseases with acute and chronic symptoms (Percival *et al.,* 2004; Bhunia, 2008). *Escherichia coli* are also of particular importance when it comes to irrigation water of fresh produce. This is because it has been shown in a number of studies, that contaminated irrigation water can lead to the occurrence and survival of *E. coli* and other pathogenic microorganisms on fresh produce (Brackett, 1999; Müller *et al.,* 2001; Beuchat, 2002; Okafo *et al*., 2003; Islam *et al.,* 2004; Ibenyassine *et al.,* 2006; Avery *et al*., 2008; Franz *et al*., 2008; Fremaux *et al*., 2008; Ackerman, 2010).

Escherichia coli are a commensal in the gut of warm blooded animals and cattle in particular, and as a result outbreaks have historically been associated with the consumption of ground beef (Willey *et al*., 2008; Linscott, 2011). Here the contamination of the food product is due to insufficient aseptic protocols being followed during slaughtering and evisceration of the animal (Avery *et al*., 2008). More recently however, leafy vegetables such as lettuce, cabbage and spinach, as well as other fresh produce have been responsible for numerous large scale outbreaks throughout the world (Islam *et al.,* 2004; Franz *et al*., 2008; Fahs *et al*., 2009; IUFoST, 2011; Linscott, 2011). This has led to fresh produce being seen as one of the common sources of *E. coli* contamination causing outbreaks.

Stellenbosch University http://scholar.sun.ac.za

One of the largest outbreaks to date occurred in Germany in 2011 (Frank *et al*., 2011; IUFoST, 2011). In this outbreak, sprouts were linked to 3 800 confirmed infections caused by *E. coli* O104:H4. Of these 3 800 individuals, 44 deaths were confirmed (Frank *et al*., 2011). The link between the *E. coli* and the sprouts was later found to be the origin of the seeds in Egypt (Anon., 2011). The fenugreek seeds used to cultivate the sprouts were found to be contaminated with the same *E. coli* O104:H4 strain. It was also later speculated that the source of the contamination was from the irrigation water (Anon., 2011). Some other outbreaks linked to fresh produce include Montana in 1995, where there was a confirmed outbreak of *E. coli* O157:H7 caused by the consumption of lettuce (Ackers *et al.,* 1998). Numerous outbreaks linked to lettuce and spinach have also occurred due to contaminated irrigation water used in the Salina's Valley in California (O'Brien, 2007; Fahs *et al*., 2009), and a large outbreak in Japan has also been linked to the consumption of beet sprouts (Islam *et al.,* 2004).

These outbreaks have led to several studies being undertaken in order to determine the contamination source of the *E. coli* found in fresh produce. Although microbial contamination can arise at any step of the production process: harvesting, transport or packaging, a study done by Islam *et al*. (2004) showed that the step in the production of fresh produce, such as lettuce, which carries the highest risk for contamination is the soil preparation step. Soil preparation includes the application of fertiliser, water and any other additives. It was also found by independent studies that *E. coli* can be transferred to fresh produce during irrigation of crops with contaminated water (Müller *et al.,* 2001; Islam *et al.,* 2004; Ibenyassine *et al.,* 2006; Avery *et al*., 2008; Franz *et al*., 2008; Fremaux *et al*., 2008; Ackerman, 2010). Therefore if there is contamination present in the irrigation water, the risk of contamination of fresh produce is greatly increased (Islam *et al.,* 2004).

Escherichia coli O157:H7 has also been shown to survive for several months within the soil of vegetable fields which have been exposed to contaminated irrigation water or bovine manure enriched fertilisers (Islam *et al.,* 2004). This just shows that farm practices are extremely important as one application of contaminated water or fertiliser could cause several months of potentially harmful produce. Contaminated irrigation water can therefore lead to the spread of pathogenic microorganisms such as pathogenic *E. coli* strains. This can lead to huge economic losses, as well as a decrease in trust by the consumers; outbreaks occurring around the world have shown time and again that once consumer trust is lost, it is difficult to gain back.

Another problem surrounding infections caused by Enterohaemoragic *E.* coli (EHEC) is that it is often not reported and therefore no clinical detection and confirmation of the causative agent is done. Only in the case of a large outbreak are sufficient laboratory tests conducted and a data

Stellenbosch University http://scholar.sun.ac.za

set compiled. This means that South Africa holds no real data correlating to amounts of *E. coli* infecting people on an annual basis. Only in some rare cases, and large outbreaks do the data get captured and data sets compiled.

When looking at food safety, water quality is therefore a very important factor to consider, which in turn also places emphasis on the quality of the water being used for the irrigation of fresh produce. In South Africa, which has minimal rainfall, farmers rely largely on natural water sources to provide them with irrigation water (Woodford *et al*., 2005). Rivers, dams, springs and boreholes are frequently utilised by farmers, but often the microbial quality of the water is unknown. A number of studies (Bezuidenhout *et al*., 2002; Ackermann, 2010; Lötter, 2010) have been conducted in South Africa to determine the microbial quality of river water, but very little work has been done on any source of groundwater such as boreholes and springs. Results from surface water show that in most cases the river water contains faecal coliform loads higher than 1 000 cfu.100 mL^{-1}, which is the guideline for faecal coliform loads in water being used for irrigation of fresh produce (WHO, 1989; DWAF*, 1996). By using water with a coliform load of >1 000 cfu.100 mL^{-1} to irrigate fresh produce, the risk of being infected by pathogenic *E. coli* is vastly increased. (*DWAF = Departnment of Water Affairs and Forestry, post 2006 this national department was re-named to DWA = Department of Water Affairs.)

The high microbial loads of surface water, as well as sewage treatment facilities and rain water seeping through landfills, questions whether it is possible that the contamination also leaches through the ground and reaches underground aquifers. Groundwater is an overexploited source in South Africa (Hughes, 2004), but very little information is known about the microbial quality of the groundwater used. Various studies have been done on groundwater, but many focus primarily on the hardness, salinity and mineral content of the water (Adams *et al*., 2001; Onstott *et al*., 2006). Where studies have been done concerning the water quality, *E. coli* was detected in numerous instances (Jain *et al*., 2009; Bezuidenhout *et al*., 2011; Haramoto *et al*., 2011). This leaves a void in known information, which is important to note.

The main problem present is that there is a gap in the information known about groundwater in South Africa, as well as the diversity and type of *E. coli* present. This is an important problem to solve as groundwater is used across the country as a source of irrigation water for all kinds of crops, including fresh produce (Hughes, 2004). Fresh produce is of particular importance as it is eaten raw or minimally processed, meaning that pathogens will not be eliminated by heat treatment. Due to the huge variety of *E. coli* strains which exist (Groisman,

Stellenbosch University http://scholar.sun.ac.za

2001; Johnson & Russo, 2002; Johnson & Russo, 2003; Percival *et al.,* 2004), the type and potential risk of the *E. coli* strains present in irrigation water are also of great importance.

To fill this gap that has been left by previous studies, *E. coli* will be isolated from a variety of sources which include both surface and groundwater being specifically used for irrigation. Various known sources of *E. coli* contamination, such as sewage treatment plants, informal settlements and animal farms, will also be tested to determine which types of *E. coli* can be associated with which source. This will also help to facilitate microbial source tracking (MST) later on. There will also be a focus on groundwater as very little information is available, especially in South Africa on groundwater quality. By completing this study, it is hoped that the prevalence as well as the characteristics of the *E. coli* present in both irrigation and potential contamination sources can be determined. In conjunction with this main objective, we hope to be able to use the information collected to see how different *E. coli* strains cluster together and potentially use MST to link *E. coli* in irrigation water to its contamination source. This will help with a risk assessment of contaminated irrigation water used on fresh produce.

REFERENCES

Ackermann, A. (2010). *Assessment of microbial loads of the Plankenburg and Berg Rivers and the survival of Escherichia coli on raw vegetables under laboratory conditions.* MSc in Food Science. University of Stellenbosch. South Africa.

Ackers, M.L., Mahon, B.E., Leahy, E., Goode, B., Damrow, T., Hayes, P.S., Bibb, W.F., Rice, D.H., Barrett, T.J., Hutwangner, L., Griffin, P.M. & Slutsker, L. (1998). An outbreak of *Escherichia coli* 0157:H7 infections associated with leaf lettuce consumption. *Journal of Infectious Diseases*, **177**, 1588-1593.

Adams, S., Titus, R., Pietersen, K., Tredoux, G. & Harris, C. (2001). Hydrochemical characteristics of aquifers near Sutherland in the Western Karoo, South Africa. *Journal of Hydrology*, **241**, 91-103.

Anonymous (2011). EHEC O104:H4 outbreak event in Germany clarified: sprouts of fenugreek seeds imported from Egypt as underlying cause. *Bundesinstitut fur Risikobewertung.* http://www.bfr.bund.de/en/press_information/2011/21/ehec_o104_h4_outbreak_event_in_germany_clarified__sprouts_of_fenugreek_seeds_imported_from_egypt_as_underlying_cause-83273.html, Accessed 08/05/2012.

Avery, L.M., Williams, A.P., Killham, K. & Jones, D.L. (2008). Survival of *Escherichia coli* O157:H7 in waters from lakes, rivers, puddles and animal-drinking troughs. *Science of the Total Environment*, **389**, 378-385.

Beuchat, L.R. (2002). Ecological factors influencing survival and growth of human pathogens on raw fruits and vegetables. *Microbes and Infection*, **4**, 413-423.

Bezuidenhout, C.C., Mthembu, N. & Lin, J. (2002). Microbiological evaluation of the Mhlathuze River, KwaZulu-Natal (RSA). *Water SA*, **28**(3), 281-286.

Bezuidenhout, C.C. & the North-West University Team (2011). A scoping study on the environmental water (groundwater and surface water) quality and management in the

Stellenbosch University http://scholar.sun.ac.za

North-West Province, South Africa. Water Research Commission (WRC) Report No. KV 278/11. WRC, Pretoria.

Bhunia, A.K. (2008). *Escherichia coli.* In: *Foodborne Microbial Pathogens: Mechanisms and Pathogenesis* (edited by Heldman, D.R.). Pp. 183-200. Washington: Springer.

Brackett, R.E. (1999). Incidence, contributing factors, and control of bacterial pathogens in produce. *Postharvest Biology and Technology*, **15**, 305-311.

DWAF (Department of Water Affairs and Forestry). (1996). Agricultural Use: Irrigation. In: *South African Water Quality Guidelines. Volume 4.* (2nd Ed.). (Edited by Holmes, S. CSIR Environmental Services). Pretoria: Department of Water Affairs and Forestry.

Fahs, F., Mittelhammer, R.C. & McCluskey, J.J. (2009). *E. coli* outbreaks affect demand for salad vegetables. *Choices*, **24**(2), 26-29.

Frank, C., Werber, D., Cramer, J.P., Askar, M., van der Heiden, M., Barnett, H., Fruth, A., Prager, R., Spode, A., Wadj, M., Zoufaly, A., Jordan, S., Stark, K. & Krause, G. (2011). Epidemic profile of shiga-toxin-producing *Escherichia coli* O104:H4 outbreak in Germany. *New England Journal of Medicine*, **365**(19), 1771-1780.

Fremaux, B., Boa, T., Chaykowski, A., Kasichayanula, S., Gritzfeld, J., Braul, L. & Yost, C. (2008). Assessment of the microbial quality of irrigation water in a prairie watershed. *Journal of Applied Microbiology*, **106**, 442-454.

Franz, E., Semenov, A.V. & van Bruggen, A.H.C. (2008). Modelling the contamination of lettuce with *Escherichia coli* O157:H7 from manure-amended soil and the effect of intervention strategies. *Journal of Applied Microbiology*, **105**, 1569-1584.

Groisman, E.A. (2001). *Principles of Bacterial Pathogenesis*. Pp. 387-456. New York: Academic Press.

Haramoto, E., Yamada, K. & Nishida, K. (2011). Prevalence of protozoa, viruses, coliphages and indicator bacteria in groundwater and river water in the Katmandu Valley, Nepal. *Transactional of the Royal Society of Tropical Medicine and Hygiene*, **105**, 711-716.

Hughes, D.A. (2004). Incorporating groundwater recharge and discharge functions into an existing monthly rainfall-runoff model. *Hydrological Sciences Journal*, **49**(2), 297-311.

Ibenyassine, K., AitMhand, R., Karamoko, Y., Cohen, N. & Ennaji, M.M. (2006). Use of repetitive DNA sequences to determine the persistence of enteropathogenic *Escherichia coli* in vegetables and in soil grown in fields treated with contaminated irrigation water. *Letters in Applied Microbiology*, **43**, 528-533.

IUFoST (The International Union of Food Science and Technology) (2011). Shiga toxin producing *Escherichia coli*: Germany 2011 *Escherichia coli* O104:H4 outbreak linked to sprouted seeds. *IUFoST Scientific Information Bulletin (SIB)*.

Islam, M., Doyle, M.P., Phatak, S.C., Millner, P. & Jiang, X. (2004). Persistence of Enterohemoragic *Escherichia coli* O157:H7 in soil and on leaf lettuce and parsley grown in fields treated with contaminated manure composts or irrigation water. *Journal of Food Protection*, **67**(7), 1365-1370.

Jain, C.K., Bandyopadhyay, A. & Bhadra, A. (2009). Assessment of ground water quality for drinking purpose, District Nainital, Uttarakhand, India. *Environment Monitoring and Assessment*, **166**, 663-676.

Johnson, J.R. & Russo, T.A. (2002). Extraintestinal pathogenic *Escherichia coli*: "The other bad *E coli*". *Journal of Laboratory and Clinical Medicine*, **139**(3), 155-162.

Johnson, J.R. & Russo, T.A. (2003). Review: Medical and economic impact of extraintestinal infections due to *Escherichia coli*: focus on an increasingly important endemic problem. *Microbes and Infection*, **5**, 449-456.

Linscott, A.J. (2011). Food-borne illnesses. *Clinical Microbiology Newsletter*, **33**, 41-45.

Stellenbosch University http://scholar.sun.ac.za

Lötter, M. (2010). *Assessment of microbial loads present in two Western Cape Rivers used for irrigation of vegetables.* MSc in Food Science, University of Stellenbosch. South Africa.

Melloul, A., Amahmid, O., Hassani, L. & Bouhoum, K. (2002), Health effect of human wastes use in agriculture in El Azzouzia (the wastewater spreading area of Marrakesh city, Morocco). *International Journal of Environmental Health Research*, **12**, 17-23.

Müller, E.E., Ehlers, M.M. & Grabow, W.O.K. (2001). The occurrence of *E. coli* O157:H7 in South African water sources intended for direct and indirect human consumption. *Water Research*, **35**(13), 3085-3088.

Onstott, T.C., Lin, L.H., Davidson, M., Mislowack, B., Borcsik, M., Hall, J., Slater, G., Ward, J., Sherwood Lollar, B., Lippmann-Pipke, J., Boice, E., Pratt, L.M., Pfiffner, S., Moser, F. D., Gihring, T., Kieft, T.L., Phelps, T.J., Van Heerden, E., Litthaur, I.D., Deflaun, M., Rothmel, R., Wanger, G. & Southam, G. (2006). The origin and age of biogeochemical trends in deep fracture water of the Witwatersrand Basin, South Africa. *Geomicrobiology Journal*, **23**, 369-414.

Okafo, C.N., Umoh, V.J. & Galadima, M. (2003). Occurrence of pathogens on vegetables harvested from soils irrigated with contaminated streams. *The Science of the Total Environment*, **311**, 49-56.

O'Brien, C. (2007). Killer Spinach. *Earth Island Journal*, 25-28. Available online at: http://www.earthisland.org/journal/index.php/eij/article/killer_spinach/, Accessed 17/02/2011.

Percival, S., Chalmers, R., Embrey, M., Hunter, P., Sellwood, J. & Wyn-Jones, P. (2004). *Microbiology of Waterborne Diseases: Microbiological Aspects and Risks.* New York: Academic Press.

Steynberg, M.C., Venter, S.N., de Wet, C.M.E., du Plessis, G., Holhs, D., Rodda, N. & Kfir, R. (1995). Management of microbial water quality: New perspectives for developing areas. *Water Science And Technology*, **32**, 183-191.

WHO (World Health Organization) (1989). Health guidelines for the use of wastewater in agriculture and aquaculture. Geneva, World Health Organization. Technical Report Series No 776. World Health Organization. Switzerland, Geneva.

Willey, J.M., Sherwood, L.M. & Woolverton, C.J. (Eds.). (2008). Microbiology of food. In: *Prescott, Harley & Klien's Microbiology*, 7th Ed. Pp. 1023-1048. New York: McGraw Hill.

Woodford, A., Rosewaene, P. & Girman, J. (2005). How much groundwater does South Africa have? DWAF (Department of Water Affairs & Forestry) Pretoria, South Africa.

Stellenbosch University http://scholar.sun.ac.za

CHAPTER 2

LITERATURE REVIEW

A. INTRODUCTION

In 2006, the Water Research Commission (Backeberg, G.R., 2006, Water Research Commission (WRC), Pretoria, South Africa. Personal communication) initiated a research project aimed at determining the microbial loads of South African rivers that are used for the irrigation of fresh produce. In addition to measuring the microbial loads, another aim was to define and characterise the consortium of microorganisms present in the rivers. The study was initiated as a result of the high levels of microbial contamination in river water and the potential carry-over of microbes from irrigation water to fresh produce being irrigated with the contaminated water (Beuchat & Ryu, 1997; Solomon *et al*., 2002; Okafo *et al*., 2003; Islam *et al*., 2004).

Another factor influencing the initiation of this project is the recorded increase in the consumption of fresh fruit and vegetables over the last 20 years (Matthews, 2006). This increase has forced farmers to escalate their total output of fresh produce, putting more strain on production of a large yield. As a result of the change in consumer needs, and the emphasis on the consumption of at least five fresh fruit and vegetables per day (Guenther *et al.,* 2006), farmers are also in the position where they have an increased water requirement (which is already a limiting commodity). In arid areas, such as southern Africa, where water has become a problem, farmers are forced to use water from natural sources such as rivers, dams, boreholes and springs to irrigate crops. Water quality from these sources is however not always tested, and maintenance of water quality of an acceptable standard can therefore not be guaranteed. Furthermore there is a gap in available information on the microbial quality of groundwater, as very few studies have been done on a local or global scale (Bezuidenhout *et al*., 2011; Haramoto *et al*., 2011).

Guidelines regarding irrigation water quality have been compiled by the World Health Organisation (WHO), but these guidelines cannot be enforced (Ackermann, 2010; Lötter, 2010). The WHO states that irrigation water being used for fresh produce must have a faecal coliform count of no more than 1 000 cfu.100 mL^{-1} (WHO, 1989; DWAF, 1996), as high coliform loads can facilitate carryover of pathogens from water to produce. Farmers are often not even aware of these guidelines or the possible risks posed by contaminated water being used for irrigation of fresh produce. Consumers are also frequently left in the dark regarding the fresh produce they eat and the potential risks associated with said fresh produce. Only when there are large scale

Stellenbosch University http://scholar.sun.ac.za

outbreaks, such as that of an EHEC (enterohaemoragic *Escherichia coli*) strain causing haemolytic uraemic syndrome (HUS) in Germany in 2011 (Anon., 2011a and b), do consumers and retailers become aware of the potential risks.

Escherichia coli were first mentioned as an emerging causative agent of foodborne diseases in the USA in 1997 (Beuchat & Ryu, 1997; Tauxe, 1997). It was also shown that of the disease outbreaks on a global scale, linked to fresh produce between 1990 and 1995, approximately 29% were linked to *E. coli* O157:H7 (Tauxe, 1997). In the years following, more emphasis was placed on eating more healthily and consuming more fresh fruit and vegetables (Guenther *et al.,* 2006). By actively making consumers more aware of the importance of eating fresh fruit and vegetables, the consumption of these products subsequently increased (Harris *et al*., 2003). Matthews (2006) reported that the increase in consumption of vegetables was 29% from 1980 - 2000, which was even before the importance of fruits and vegetables was stressed to the public. Therefore, it can also be said that even if the contamination level stayed the same, more incidences would be recorded due to the increase in consumption (Harris *et al*., 2003).

It is however hard to say what percentage of disease outbreaks linked to fresh produce worldwide are caused by *E. coli* as many countries' reporting systems are deficient. The USA, Germany and the UK are three of the few countries in the world which have a reporting system of foodborne diseases that has been successfully implemented. The USA also tends to determine the exact disease causing organism, whereas in other countries such as South Africa people either recover at home or just receive general medication for stomach ailments. The information available on *E. coli* outbreaks linked to fresh produce, for third world countries, is therefore limited, and where available it is more often than not based on incidences occurring in the USA and usually focused only on *E. coli* O157:H7. Between 1998 and 2002, *E. coli* O157:H7 was found to be responsible for 6.2% of the disease outbreaks linked to fresh produce in the USA (Lynch *et al*., 2006). This emphasises that there is definitely a problem regarding food safety and irrigation water, which needs to be documented more carefully and that more information is desperately needed.

B. MICROBIAL QUALITY OF IRRIGATION WATER SOURCES

Escherichia coli is an organism which has over the years been observed in an assortment of environmental niches. As studies on natural water sources intensified and detection methods

Stellenbosch University http://scholar.sun.ac.za

became more reliable and easier to conduct, more and more *E. coli* strains were found in places previously thought to host none (Jain *et al*., 2009; Bezuidenhout *et al*., 2011; Haramoto *et al*., 2011). Due to the increase in *E. coli* originating outbreaks from consumption of fresh produce, and the link between fresh produce and disease causing *E. coli* being narrowed down to contaminated irrigation water, much more work has recently gone into studying the microbial quality of irrigation water (Müller *et al*., 2001; Bezuidenhout *et al*., 2002; Rai & Tripathi, 2007; Jain *et al*., 2009; Paulse, 2009; Ackermann, 2010; Lötter, 2010; Bezuidenhout *et al*., 2011; Haramoto *et al*., 2011).

In South Africa, both surface and groundwater are used for irrigation of fresh produce as municipal water is expensive and not always readily available. Rivers and underground water reservoirs, on the other hand, are frequently more accessible and much cheaper to use. Adversely, natural irrigation water sources are also more likely to be polluted with a vast array of microbial contaminants as natural water sources are not monitored (Müller *et al*., 2001; Bezuidenhout *et al*., 2002; Rai & Tripathi, 2007; Ackermann, 2010; Lötter, 2010). Contaminants found in natural water sources can come from a variety of sources, some of which will be explored further in this study.

i. Surface water

Over the previous four years, various data sets on the faecal pollution in the Western Cape have been obtained, each with a slightly different, but still related, focal point (Ackermann, 2010; Lötter, 2010; Kikine, 2011). It was found in the first research projects, which focused on the microbial loads of the Plankenburg, Mosselbank and Berg Rivers in the Stellenbosch and surrounding areas, that the microbial loads were all extremely high. It was found that the faecal coliform counts across the three study sites ranged from 0 to 1 700 000 MPN.100 mL^{-1} (Ackermann, 2010; Lötter, 2010). It was also found that the water being drawn for irrigation from these rivers not only had a high microbial load, but that a vast array of microorganisms was also present. The consortium of microorganisms in the river water consisted of both pathogenic as well as non-pathogenic microorganisms (Ackermann, 2010; Lötter, 2010), but the exact organisation, in terms of *E. coli* groupings present in the whole population is unknown.

Other studies done throughout South Africa showed similar results; a study done in 1998/1999 (Bezuidenhout *et al*., 2002) along the Mhlathuze River in Kwa-Zulu Natal, found that total as well as faecal coliform counts exceeded the guidelines set by DWAF and WHO. It was found in this particular study that the faecal coliform counts ranged from 280 – 3 620 cfu.100 mL^{-1},

Stellenbosch University http://scholar.sun.ac.za

while the total coliform counts ranged between 820 and 12 520 cfu.100 mL^{-1} (Bezuidenhout *et al*., 2002). It was also speculated that the contamination of the river was mainly due to industrial practices and eutrophication. A similar study done in Bloemfontein by Griesel & Jagals (2002) on the water coming from the Renoster Spruit sub catchment also drew similar conclusions. The water in this catchment was tested for *E. coli*, amongst other microorganisms, and *E. coli* counts were also found to exceed the recommended guideline for irrigation water of 1 000 cfu.100 mL^{-1}. The contamination in the catchment area was thought to originate from eutrophication, agricultural and industrial waste (Griesel & Jagals, 2002).

Although industrial effluent plays a role in introducing microbial contaminants as well as increasing the carbon, nitrogen and phosphorous levels in the water, thus boosting growth and survival, in areas lacking industrial activity, contamination must come from an alternative source. In these cases it has been concluded that contamination comes mainly from informal settlements which lack adequate sewerage systems, insufficient wastewater treatment, agricultural runoff and leachate originating from land-fills (Zamxaka *et al*., 2004).

It can therefore be concluded, by looking at previous studies, that natural water sources are contaminated, as well as that a portion of the microbial population present, in natural water sources, is *E. coli*. Types of *E. coli* present, as well as the percentage of pathogenic *E.* coli present in the rivers have not yet been looked at. It is therefore not known if there is a variation of strains present in natural water sources or not, and if there is, its composition is unknown (Ackermann, 2010; Lötter, 2010; Kikine, 2011).

ii. Ground water

Other than rivers, groundwater such as boreholes and springs, also contributes to irrigation water in this country. A study by Hughes *et al*. in 2004 showed that 64% of South Africa's extracted groundwater is used solely for agricultural irrigation. In the Western Cape the use of groundwater as an irrigation source is especially widespread due to the lack of water in the dry, hot summer months. The existence of fractured rock, quartzitic and primary aquifers found along the coast also makes groundwater considerably more accessible in these areas (Woodford *et al*., 2005).

Over the years, groundwater has always been seen as a pristine water source as a result of water travelling through several natural sand and stone 'filters' while the water percolates down to the groundwater catchment area (Bezuidenhout *et al*., 2011). As a result of this mind-set associated with groundwater, there is a substantial gap in the qualitative information known about water originating from boreholes and springs, not only in South Africa but around the world

Stellenbosch University http://scholar.sun.ac.za

(Bezuidenhout *et al*., 2011; Haramoto *et al*., 2011). The known information regarding groundwater is also more to do with salinity, hardness and mineral content opposed to microbial water quality, which further emphasises the lack of information.

By not knowing in what state the water being used for irrigation is, farmers may be irrigating their fresh produce and other crops with water that is of a substandard quality which may also contain potentially pathogenic organisms. These factors may then lead to health related problems for consumers of the fresh produce. It is however important to take into consideration that, although the water found in underground catchments has been 'filtered' by the layers of rock and soil, it may still have carried along some human pathogens in the leachate which is then also present in the groundwater (Jain *et al*., 2009; Bezuidenhout *et al.*, 2011; Haramoto *et al*., 2011). The type of sand through which the water seeps, also seems to play a role in the affectivity of this natural filter system (Smith *et al*., 1985).

As with surface water, inappropriate or leaking sewerage systems and runoff containing animal faecal matter are the main contributors to contamination. When looking at ground water, highly contaminated surface waters can also siphon through the ground and eventually reach the underground catchment, while still containing harmful human pathogens (Jain *et al*., 2009; Bezuidenhout *et al.*, 2011; Haramoto *et al*., 2011). As little is known regarding groundwater, it is extremely important that more studies be conducted to determine the state of groundwater in our country.

Through studies that have been done recently on groundwater quality, it was found that bacteria (coliforms and *E. coli* in particular), viruses, and protozoa all find their way into groundwater albeit some in very low concentrations (Jain *et al*., 2009; Bezuidenhout *et al*., 2011; Haramoto *et al*., 2011). It was also found by two independent studies (Jain *et al*., 2009; Haramoto *et al*., 2011) that tube wells and wells exceeding 40 m in depth were less likely to contain bacterial contamination, when compared to hand dug wells. This means that it is expected that wells which exceed a depth of 40 m are less likely to contain microbial contamination.

In the Eastern Cape, a study was done on water quality of domestic water used in rural areas. One of the sites included in the study was a windmill-powered water tank which extracts underground water (Zamxaka *et al*., 2004). Even though both the total and faecal coliform counts found in this water were relatively low, there were still detectable levels of both total and faecal coliforms, showing that microbial contaminants do find their way into groundwater (Zamxaka *et al*., 2004).

Stellenbosch University http://scholar.sun.ac.za

An array of other factors also contributes to the likelihood of contamination reaching underground water reservoirs. Soil type is possibly the main contributing factor, as it determines the efficiency of the 'filter' through which the water needs to pass to reach the reservoirs (Smith *et al*., 1985).

Contamination of Natural Water Sources

Escherichia coli outbreaks are historically associated with the consumption of ground beef (Willey *et al*., 2008; Linscott, 2011). This is because *E. coli* is a commensal organism in the gut of both cattle and humans, together with other warm-blooded animals. Human pathogens are also often present in cattle and diseased individuals, and are then expelled in their faecal matter (Avery *et al*., 2008). Now that *E. coli* is also seen as a regular contaminant of fresh produce and leafy vegetables, such as lettuce, cabbage, spinach and herbs (Beuchat & Ryu, 1997; Harris *et al*., 2003; CDC, 2006), more studies have been initiated to address some of the problems prevailing throughout present data (Van Blommestein, 2012).

Worldwide, the increasing levels of faecal coliforms present in river systems used for both direct and indirect human use have become a major health risk, especially in the last 10 years (Müller *et al.,* 2001; Bezuidenhout *et al*., 2002; Melloul *et al.,* 2002; Linscott, 2011). Various practices lead to faecal contamination of rivers, but overall runoff from cattle farms (Avery *et al*., 2008; Fremaux *et al*., 2008) as well as overflow from overloaded municipal sewage treatment plants and general eutrophication as a result of inadequate sewerage systems in rural settlements (Islam *et al.,* 2004; De Villiers, 2007) play a notable role in causing contamination.

Escherichia coli, along with other pathogenic microorganisms such as *Listeria* and *Salmonella* spp., are major contaminants in natural fresh water systems (Müller *et al.,* 2001; Melloul *et al.,* 2002; Islam *et al.,* 2004; Ackerman, 2010; Lötter, 2010; Linscott, 2011). Although many studies have been undertaken to look at total coliforms present in rivers across South Africa (Steynberg *et al.,* 1995; Ackerman, 2010; Lötter, 2010), not many studies have looked in depth at the *E. coli* strains present in the water.

Microbial contamination can come from a variety of sources and find its way into surface, as well as groundwater sources which are then used for irrigation, as water settling in dams, rivers, boreholes and springs all have to originate from somewhere. Rain water is the most obvious water source to end up in catchments, and although rain water cannot contain microbes, it may still play a role in the distribution of bacteria in various water sources. This is because rain water

Stellenbosch University http://scholar.sun.ac.za

causes sediment to be re-suspended within a water body as well as acting as a carrier for runoff from various agricultural practices (Avery *et al.*, 2008; Fremaux *et al.*, 2008; Gerba & Choi, 2009).

Return flow is also a key contributor, and all water that remains after irrigation (especially flood irrigation) and then finds its way back to the irrigation canal falls in this category (Gerba & Choi, 2009). This means that all pesticides, herbicides and microbes present in fertiliser applied to crops are washed back into the irrigation water and applied to the produce via irrigation (Gerba & Choi, 2009). Rural settlements and insufficient or completely lacking sewerage systems play a big role in microbial contamination of water sources, particularly *E. coli* contamination (Islam *et al.,* 2004; De Villiers, 2007). In the Western Cape, as well as South Africa as a whole, there are vast areas covered by informal settlements where people are lacking adequate sewerage systems and as a result, human waste often ends up in the nearest river or dam (Kfir *et al*., 1995; De Villiers, 2007).

Effluents from waste water treatment plants and industrial practices are also sources of microbial contamination. Additionally, industrial effluent also contributes to the increased chemical oxygen demand (COD) of the natural waterways (Geldreich, 1978; De Villiers, 2007). This increase of carbon, nitrogen and phosphorous can then be expected to facilitate the growth of the introduced microbes, especially *E. coli*. When looking specifically at *E. coli* sources; agricultural practices, along with human sewage, are probably the greatest contributing factors. This is because runoff coming from cattle, pig and chicken farms contain faecal matter which is known to contain high numbers of *E. coli* (Avery *et al*., 2008; Fremaux *et al*., 2008). Cattle farms, in particular, are important *E. coli* reservoirs because as mentioned, human pathogenic *E. coli* was first associated with cattle (Avery *et al*., 2008).

When it comes to groundwater, the source of the contamination is assumed to be the same as that of surface water, except the way in which the microbial contamination reaches the catchment is varied. In the case of groundwater, surface water and runoff first filters through the various layers of soil and rock until it reaches the groundwater reservoir (Bezuidenhout *et al*., 2011). This means that there is some sort of natural filtration system present, which has led to the misconception that groundwater is unspoiled (Bezuidenhout *et al*., 2011). Although the water filters through these soil and rock layers, it has been shown that when the contamination levels in the water are too high to be filtered out, some residual contamination may still find its way into groundwater reservoirs (Smith *et al*., 1985; Jain *et al*., 2009; Bezuidenhout *et al*., 2011; Haramoto *et al*., 2011).

Stellenbosch University http://scholar.sun.ac.za

It has also been found that the way in which a well has been sunk, as well as the depth of the well both play a part in the extent of the microbial contamination present. Jain *et al*. found in 2009 that hand-pump and tube wells had no bacterial contamination present, whereas some (21%) hand dug wells and spring water samples contained more than ten coliforms per 100 mL. Hand-dug wells tend not to be as deep as wells sunken using industrial machinery, while tube wells on the other hand characteristically extend further than 40 m into the ground (Jain *et al*., 2009). This could explain the possible differences in contamination, although due to the lack of follow-up studies, it is difficult to validate this data. Springs also often have a reservoir which is close to the surface and as a result that reservoir may become more easily contaminated by outside factors. It can therefore be said that deeper wells contain less bacterial contamination as a result of the water having undergone more extensive filtration by the soil and rocks. Shallow wells and springs tend to be more contaminated due to less filtration being undergone (Jain *et al*., 2009).

Growth and Survival of *Escherichia coli*

It has been shown (Van Blommestein, 2012) that one of the main contributing factors controlling the growth and survival of *E. coli* in natural water sources is in fact not temperature, but rather the amounts of carbon, nitrogen, phosphates and other growth factors in the water. In-depth studies have been done on the Plankenburg River in the Stellenbosch area, and these studies have shown that the chemical oxygen demand (COD) in the river was positively correlated to the bacterial load present (Van Blommestein, 2012). It was also shown that the initial COD in the water affects both the growth, and decline rate, and as an effect of this the survival of the organism is also affected.

Other studies have been conducted with the aim of looking at possible causes for the increased COD, phosphate and nitrogenous wastes present in the natural water sources in the Western Cape. In 2007 it was found that the leading cause of increased phosphate and nitrogenous wastes in the Berg River in the Western Cape is due to the overloading of municipal sewage treatment plants, agricultural lands, informal settlements which have no real sewerage systems and the runoff associated with these facilities (De Villiers, 2007).

Once *E. coli* finds its way into water systems which then get used for irrigation, a host of other problems also arise. One of these is the persistence of *E. coli* and its capacity to survive under harsh conditions in the environment (Tal & Schuldiner, 2009). *Escherichia coli* O157:H7, for example, has been shown to survive for seven months within the soil of vegetable fields which

Stellenbosch University http://scholar.sun.ac.za

have been exposed to contaminated irrigation water or bovine manure (Islam *et al.,* 2004). This means that after only a single application of contaminated manure or irrigation water it is possible that the bacteria can be introduced into the soil and survive for extended periods. If conditions in the soil are ideal for bacterial growth the *E. coli* can multiply, exponentially increase in numbers and pose a threat to the consumers of fresh produce (Islam *et al.,* 2004; Islam *et al.,* 2005). This just reiterates that monitoring of farm practices is extremely important as one application of contaminated water or manure could cause several months of problems as the *E. coli* attaches and survives on the plant material as well as in the soil (Islam *et al*., 2004).

Escherichia coli are also known to be robust with a high survival rate. When dealing with fresh produce the bacteria can survive both in the soil as well as, in some cases, the actual plant tissue. It has been found that when dealing with leafy crops, such as lettuce and spinach, the leafy portion of the plant that is eaten by consumers can become contaminated with *E. coli* O157:H7 without constant or even prolonged exposure directly to the contamination source (Islam *et al.,* 2004). This is possibly due to internal transportation of the *E. coli* through the root system, primary contact with contaminated irrigation water or via external means such as rain or wind dispersal of the bacteria.

Bacteria may also be introduced into the plant tissue via mechanical damage. This, along with previously mentioned factors, all lead to lasting survival and multiplication of the bacteria in the environment and plant tissue (Islam *et al.,* 2004). In the same study in 2004, it was concluded that it is possible for *E. coli* O157:H7 to infect lettuce leaves via contaminated soil or water. It was also shown that there is a definite carryover of the pathogen, even though an exact percentage of carryover is not known (Islam *et al*., 2004).

When dealing with contaminated irrigation water and the carryover of pathogens from the water to the fresh produce, various factors influence the attachment and internalisation, therefore also having an impact on the survival of the bacteria (Gerba & Choi, 2009). In particular; type of irrigation, surface properties of the fresh produce, where (in relation to the soil) the edible portion is located as well as climate, all play important roles when it comes to bacterial carryover of pathogens (Gerba & Choi, 2009). Looking at irrigation water application in particular, flood and overhead irrigation methods seem to pose a greater risk than other irrigation methods to facilitate the adherence of bacteria to leafy vegetables (Solomon *et al*., 2002; Okafo *et al*., 2003; Islam *et al.,* 2004; Song *et al*., 2006; Rai & Tripathi, 2007). Drip and flood irrigation however, seem to be more successful at prolonging survival, as well as facilitating internalisation of *E. coli* via the plant roots (Solomon *et al*., 2002; Mootian *et al*., 2008). Therefore it is also important for farmers to also

Stellenbosch University http://scholar.sun.ac.za

consider type of irrigation when wanting to minimise risk of contamination and survival of microorganisms.

Risks Associated with *E. coli* Contamination in Irrigation Water Sources

When looking at contaminated water which is being used for irrigation of fresh produce, a number of associated risks are involved. These risks are not always associated solely with the consumer of the fresh produce, retailer or supplier but it is rather associated with everyone involved in the production process. Each one is also affected in its own way; some risks have to do with food safety or health of the consumer, while other risks include loss of usable yield or consumer trust. Risk, in this instance, is therefore defined as any potential problem associated with contaminated irrigation water being used on fresh produce which can have detrimental effects to anyone involved in the production process or consumption of fresh produce.

Although microbial contamination can arise at any step of the fresh produce production process: harvesting; transport or packaging, a study done by Islam and associates (2004) showed that the step in the production of fresh produce, such as lettuce, which carries the highest risk for contamination is the soil preparation step. Soil preparation includes the application of fertiliser as well as water. Therefore if contamination is present in either the water or fertiliser source, contamination risk of the fresh produce is drastically increased (Islam *et al.,* 2004). This is because bacteria in general, and *E. coli* in particular, are extremely robust and can survive for extended periods of time in soil, even if they are not soil borne pathogens, while still maintaining pathogenicity. Some plasmids carrying virulence factors can however be lost over time due to environmental stress (Snyder & Champness, 2007; Willey *et al.,* 2008). This means that over a period of time, in non-ideal environmental conditions, some pathogenic *E. coli* may lose a portion of their virulence factors.

The number of *E. coli* present in the water source also plays a direct role in the risk involved when using the water. This is because of the reasoning that when a higher level of contamination is present in the irrigation water, the better the chances of the *E. coli* attaching to the plants and posing a risk to consumers. A study conducted in Canada, showed an increase in *E. coli* contamination levels during the rainy season (Fremaux *et al*., 2008). Together with the increase in *E. coli* numbers during the rainy season, a rise in the genetic variation was also noted when using randomly amplified polymorphic DNA (RAPD) (Fremaux *et al*., 2008). This shows that not only is there an increased risk due to shear number of viable pathogens, but there is also an

Stellenbosch University http://scholar.sun.ac.za

increased risk as *E. coli* variation increases. With the increase in runoff from the bacterial, nitrogen and phosphate sources, not only will there be an increase in bacterial counts and species variation, but it will also be expected that there will be an increase in growth rate of the bacteria present (Van Blommestein, 2012).

When isolates were typed during the rainy season by RAPD, *E. coli* showed a definite increase in variation of strains to that of the population present in the dry season (Fremaux *et al*., 2008). This means that there is also a greater chance of gene acquisition occurring due to the increased number of genes present as well as the selection pressure acting on the population (Snyder & Champness, 2007). This increase in adaption rate can cause the acquisition of more plasmids which can, in turn, increase pathogenicity by various means. The pathogenicity can be increased by acquisition of antibiotics resistant genes as well as numerous other virulence factors (Snyder & Champness, 2007). By increasing the adaption (gene acquisition) rate of the bacteria, the potential risk associated with the contaminated water sources will consequently increase.

In conjunction with the increased risk due to gene acquisition, the lack of information and the naivety of farmers further increases the potential risks present. This is because although there are a number of ways in which water can be treated to minimise the *E. coli* contamination before irrigation, many farmers are either not aware of the risk associated with the microbial contamination present or they are completely unaware of the contamination as a whole.

Treatment of irrigation water is also expensive, especially when hundreds of thousands of litres are used at a time, and it is also not standard practice. No regulations exist which promote farmers to treat their irrigation water before application, and it is then subsequently sidestepped. There are no regulations set for irrigation water in South Africa, but the WHO and the DWA has a global recommended guideline for irrigation water which states that no more than 1 000 faecal coliforms should be present per 100 mL (WHO, 1989; DWAF, 1996). Responsibility therefore lies with the retailers to whom the farmers supply stock to, to ensure that the farms they get their product from are enforcing the guidelines and that the irrigation water used is checked regularly to maintain an acceptable water quality. Farmers are often not prepared to adopt additional cost when they are not necessary or forced to do it. Many farmers are also unaware of the substandard quality of water that they use to irrigate their produce, amplifying the associated risks.

Stellenbosch University http://scholar.sun.ac.za

C. *ESCHERICHIA COLI* LINKED DISEASE OUTBREAKS ASSOCIATED WITH FRESH PRODUCE

Historically, pathogenic *Escherichia coli* have primarily been seen as a foodborne pathogen associated with improperly cooked ground beef or other products of bovine origin. This is due to the fact that *E. coli* makes up a large portion of the gut commensals of warm-blooded animals, such as cattle (Willey *et al*., 2008; Linscott, 2011). Contamination of the food product (beef in this case) is caused due to insufficient aseptic protocols being followed during slaughtering and evisceration of the animal (Avery *et al*., 2008). In recent years however, *E. coli* outbreaks have been caused less and less by consumption of ground beef and more by consumption of fresh leafy vegetables and salad products including lettuce, cabbage, spinach and vegetable sprouts (Islam *et al.,* 2004; Franz *et al*., 2008; Fahs *et al*., 2009; Linscott, 2011). The origin of foodborne *E. coli* outbreaks shifted so much so, that it reached a point where one third of all *E. coli* outbreaks were linked to fresh produce. This shift of outbreak origins has led to the inclusion of these leafy vegetables and salad products in the list of food sources associated with pathogenic *E. coli* (Beuchat & Ryu, 1997; Harris *et al*., 2003; CDC, 2006; SGM, 2007; ASM, 2008; FAO/WHO, 2008).

It can therefore be said that throughout the preceding 25 – 30 years, *E. coli* has gone from being a foodborne pathogen virtually never associated with fresh produce to a waterborne pathogen which contaminates fresh produce (SGM, 2007; ASM, 2008; FAO/WHO, 2008). Consumer mentality has shifted as well, making people more wary of fresh produce. Pathogenic microorganisms associated with fresh produce have also been linked to the microbial contamination in natural water sources being used for irrigation of these products (Solomon *et al*., 2002; Matthews, 2006). Lettuce has often been pin-pointed as the source of the contamination in large scale outbreaks (Islam *et al*., 2004; Gil & Selma, 2006; Mena, 2006; ASM, 2008), but outbreaks have also been linked to other fresh produce sources such as radishes (Mena, 2006), parsley (Islam *et al*., 2004), alfalfa sprouts (Gil & Selma, 2006; Mena, 2006) and cilantro (Johnston *et al*., 2006).

When looking at *E. coli* outbreaks across the globe (Doyle, 1991; Ackers *et al*., 1998; Michino *et al.*, 1999), Southern Africa does not feature very often. One of the few large recorded outbreaks occurring in Southern Africa was in 1992 in Swaziland (Effler *et al*., 2001). In general, *E. coli* outbreaks have historically been spread more across the Northern hemisphere and less around the Southern hemisphere, for example, in Japan (1996) there was a large scale outbreak of

Stellenbosch University http://scholar.sun.ac.za

E. coli O157:H7 which resulted in 9 000 people falling ill and 11 deaths (Johnston *et al*., 2006; Matthews, 2006). This outbreak was linked to the consumption of contaminated radish sprouts which were thought to have been exposed to the contaminant before germination (Johnston *et al*., 2006; Matthews, 2006). Sprouts are also grown under environmental conditions which are ideal for bacteria to thrive. The combination of both high temperatures (±37°C) and moisture in the environment create the optimal growing conditions for bacteria such as *E. coli* (Holt *et al.,* 1994; Willey *et al*., 2008).

Other outbreaks of *E. coli* O157:H7 in the USA led to damage being incurred by the economy, consumer trust as well as the health of consumers. In Montana (1995) lettuce was said to be the source of an *E. coli* O157:H7 outbreak (Ackers *et al.,* 1998). Numerous studies have also linked the Salinas Valley, California, to numerous outbreaks of *E. coli* O157:H7 originating from fresh produce (O'Brien *et al.*, 1992; Fahs *et al*., 2009). In the Salinas Valley, the source of the contamination, on numerous occasions, has been hypothesised to be the irrigation water being used on the lettuce and spinach (O'Brien *et al.*, 1992; Fahs *et al*., 2009). In September 2006 an *E. coli* O157:H7 outbreak was linked to both lettuce and spinach grown in this area. In this outbreak there were 204 confirmed illnesses, and according to the Centre for Disease Control (CDC) three people died (CDC, 2006).

In 2011 however, the largest and most documented *E. coli* outbreak (linked to fresh produce) occurred in Northern Germany. This outbreak was not as a result of the O157:H7 strain, which has dominated literature and all major outbreaks so far, but instead the causative agent was a little known strain O104:H4 (Anon., 2011b). In this outbreak 3 792 people became ill, 862 of those with HUS, and 42 deaths had been reported by 22 June 2011, just over two months after the first reported case (Anon., 2011a). The actual source of the bacteria was not clearly confirmed, but first cucumbers from Spain were blamed, causing a drastic drop in sales, loss of finances and an immediate disinterest in any vegetables and salad products across Europe. Once cucumbers were found to not be the source, an organic farm in Bienenbüttel, Lower Saxony, Germany was investigated as it was thought that their vegetable sprouts were to blame (Anon., 2011b). The farm was shut down, but it was neither confirmed nor denied that they were responsible for the outbreak (Anon., 2011b). It was later found that the origin of the *E. coli* was actually the fenugreek seeds used to cultivate the sprouts, which were imported from Egypt (Anon., 2011c). It was speculated that the seeds were harvested from plants that were irrigated with contaminated irrigation water. The seeds were therefore contaminated before reaching Germany and germinated to cultivate sprouts (Anon., 2011c). When attempting to pin-point the exact

Stellenbosch University http://scholar.sun.ac.za

contamination source of a large scale outbreak, various hurdles need to be overcome. Many of these hurdles have also been amplified by international trade, as seen in the outbreak in Germany, as it complicates the production process.

D. ESCHERICHIA COLI

Brief History

Escherichia coli was first identified and described by Theodor Escherich in 1885 (Escherich, 1988) during his work on the intestinal flora of infants. The bacterium was first described as *Bacterium coli commune*, and only post 1919 was the genus *Escherichia* and the type species *E. coli* used (Percival *et al.,* 2004). *Escherichia coli* is an extremely prevalent organism in literature and is often used as an indicator organism for faecal contamination to measure water quality. This way of testing water quality is now also being investigated as *E. coli* does not have all the qualities needed for an organism to be classified as an 'indicator' (Solo-Gabriele *et al*., 2000; Ishii & Sadowsky, 2008).

Escherichia coli are Gram-negative, non-spore forming bacteria and are most often motile by means of peritrichous flagella (Holt *et al.,* 1994; Percival *et al.,* 2004). Their cells are regarded as rod-shaped, but can often range from coccoidal to long and fibrous in appearance, which is dependent on environmental conditions. Several of the species found in the genus *Escherichia* have an optimal growth temperature of 37°C; they are all facultative anaerobes that catabolise D-glucose as well as other carbohydrates with acid and gas production (Holt *et al.,* 1994).

In the 1920's further studies on the organism showed that *E. coli* was an extremely heterogeneous species, and that further classification would be necessary. In the 1940's, 20 years later, the classification scheme which led to the division of the species, *Escherichia coli*, was developed. This classification scheme successfully allocated 70 new serogroups to the original *E. coli* species (Percival *et al.,* 2004). The division was mainly based on two antigens, namely the O (somatic) antigen, and H (flagella) antigen (Kauffmann, 1947).

As a result, numerous *E. coli* strains, producing a combination of polysaccharides and extracellular polymers have been identified and categorised (Percival *et al.,* 2004). Extremely mucoid strains can sometimes synthesise what is referred to as K and M antigens. The K antigens are extracellular polymers; while the M antigens are acid polysaccharides comprised mainly of

Stellenbosch University http://scholar.sun.ac.za

colonic acid (Holt *et al.,* 1994; Percival *et al.,* 2004). Different strains also have different kinds of fimbriae which determine adhesion to, and invasion of, host as well as biofilm initiation (Percival *et al.,* 2004).

Later it was decided that the use of both the K- and M-antigens would be adopted for more accurate classification of the species into what is known as 'serotypes' (Percival *et al*., 2004). Along with the O antigen, H (flagella) antigens and K (capsular) antigens are now also used for grouping of strains. Therefore as a whole, *E.* coli can be divided into various pathotypes/virotypes, determined by virulence factors, as well as serogroups and serotypes. Serogroup divisions are determined by deviations in the 'O' antigens, while serotypes are determined by those in the 'H' and 'K' antigens (Bhunia, 2008).

In 1973 the first new species *Escherichia blattae*, since the type species *E. coli*, was described by Burgess and colleagues during their work on cockroaches (Burgess *et al*., 1973). Since 1973, only four other *Escherichia* species have been identified and described in addition to a metabolically inactive *E. coli* species (Holt *et al.,* 1994; Percival *et al.,* 2004). These additional species are *E. albertii* (Huys *et al*., 2003), *E. fergusonii* (Farmer *et al*., 1985a), *E. hermannii* (Brenner *et al*., 1982a) and *E. vulneris* (Brenner *et al*., 1982b), see Tables 1 and 2.

Table 1. Initial discovery of source and species belonging to the *Escherichia* genus

Organism	Origin	Reference
E. blattae	Cockroaches	Burgess *et al*., 1973
E. albertii	Birds	Huys *et al*., 2003
E. fergusonii	Human intestine	Farmer *et al*., 1985a
E. hermannii	Human wounds	Brenner *et al*., 1982a
E. vulneris	Human wounds	Brenner *et al*., 1982b
E. coli	Human intestine (babies)	Escherich, 1988

Stellenbosch University http://scholar.sun.ac.za

Table 2. Selected biochemical properties of all species listed in the genus *Escherichia* represented by percentage positive tests within the sample group

Characteristic	*E. coli*[a]	*E. coli* (inactive)[a]	*E. blattae*[a]	*E. fergusonii*[a]	*E. hermanii*[a]	*E. vulneris*[a]	*E. albertii*[b]
Indole production	98	80	0	98	99	0	0
Voges-Proskauer	0	0	0	0	0	0	0
Citrate (Simmons)	1	1	50	17	1	0	0
Hydrogen sulphide	1	1	0	0	0	0	0
Urea hydrolysis	1	1	0	0	0	0	0
Lysine decarboxylase	90	40	100	95	6	85	100
Arginine dihydrolase	17	3	0	5	0	30	0
Ornithine decarboxylase	65	20	100	100	100	0	100
Motility	95	5	0	93	99	100	0
Gelatin hydrolysis (22°C)	0	0	0	0	0	0	0
Nitrate oxidised to nitrite	100	98	100	100	100	100	100
Oxidase, Kovacs	0	0	0	0	0	0	0
ONPG test	95	45	0	83	98	100	100
D-Glucose, acid production	100	100	100	100	100	100	100
D-Glucose, gas production	95	5	100	95	97	97	100
L-Arabinose	99	85	100	98	100	100	100
myo-Inositol	1	1	0	0	0	0	0
D-Mannitol	98	93	0	98	100	100	100
Melibiose	75	40	0	0	0	100	0
L-Rhamnose	80	65	100	92	97	93	0
D-Sorbitol	94	75	0	0	0	1	0
Sucrose	50	15	0	0	45	8	0

[a] Information adapted from Farmer *et al*., 1985b

[b] Information adapted from Abbott *et al*., 2003

Stellenbosch University http://scholar.sun.ac.za

Clinical Features and Associated Risks

Background

Since 1921 *E. coli* (then *Bacterium coli*) has been associated with diarrhoea, especially linked to food sources, as well as infections of both the urinary and intestinal tracts (Percival *et al*., 2004). As more studies were conducted on the organism, it was also concluded that not all *E. coli* strains are pathogenic and that the non-pathogenic (or commensal) strains contribute a large portion of the naturally occurring microbes in the lower gut of humans and other, warm-blooded animals (Forsythe, 2000; Bhunia, 2008; Ishii & Sadowsky, 2008). Only in some cases do they acquire pathogenicity which can cause acute, chronic and general infections (Puebte & Finlay, 2001; Percival *et al*., 2004).

Escherichia coli is seen as the leading cause of urinary tract sepsis and infections (Svanborg-Edén *et al*., 1976), though it has also been linked to neonatal meningitis, general sepsis, acute enteritis and 'traveller's diarrhoea' (Puebte & Finlay, 2001; Percival *et al*., 2004). Various pathotypes of *E. coli* cause different infections and have different infectious doses. The most common groupings, and also the groups which are most often researched, include enteropathogenic *E. coli* (EPEC), enterotoxigenic *E. coli* (ETEC), enteroinvasive *E. coli* (EIEC), diffusely adhering *E. coli* (DAEC), enteroaggregative *E. coli* (EAEC), enterohaemoragic *E. coli* (EHEC) falling under the major group of Vero cytotoxigenic *E. coli* (VTEC) and most recently extraintestinal *E. coli* (EXPEC) (Puebte & Finlay, 2001; Johnson & Russo, 2002; Johnson & Russo, 2003; Percival *et al.,* 2004). VTEC is the only group of pathogenic *E.* coli which has a truly low infectious dose, as less than 100 bacterial cells need to be ingested to cause disease (Karmali, 1989; Bolton *et al*., 1996; Percival *et al*., 2004). The other enteropathogenic *E. coli* groups have much higher infectious doses, with some groups having an infectious dose as high as 10^5–10^{10} bacterial cells (Karmali, 1989; Bolton *et al*., 1996; Percival *et al*., 2004). Sensitivity of individuals to infection is however determined by sex, age, immunity and pH of the host's gastro-intestinal tract (GIT) (Percival *et al.,* 2004), therefore the necessary infectious dose may vary between individuals.

Gene Acquisition and Pathogenicity

As previously discussed, *E. coli* is a naturally occurring constituent of the microbes naturally found in the guts of both humans and most warm-blooded animals (Forsythe, 2000; Bhunia, 2008). Naturally occurring strains, mainly commensal strains, are non-pathogenic and therefore cause no

Stellenbosch University http://scholar.sun.ac.za

harm to the host; they are in fact essential to their hosts' digestive system (Selander *et al*., 1987). On the contrary, the pathogenic strains, such as the enterohaemoragic and enteropathogenic *E. coli* strains can lead to serious illness and sometimes even death, even though they are close relatives to the non-pathogenic denizens (Snyder & Champness, 2007). To say that they are close relatives is also not an accurate description, as pathogenic strains are in actual fact just non-pathogenic strains which have, over time, acquired virulence genes which have made them pathogenic (Hacker *et al*., 1997). Therefore the only differences between these non-pathogenic and pathogenic variants are the presence of one, or numerous, genes acquired by the bacteria through mutation, transformation, transposition (horizontal gene transfer), conjugation or transduction. These genes may then be 'stored' in a small section of DNA. This DNA is most often a, usually closed circle, plasmid (Snyder & Champness, 2007), but the genes may also become integrated into the chromosomal DNA over time.

Plasmids are extra-chromosomal DNA which gives the hosting bacteria an advantage over and above other bacteria in the surrounding environment. Plasmids do not carry genes which are needed for day to day crucial functions such as growth or multiplication, but instead plasmids carry advantageous genes which make the survival of the bacterium more successful (Snyder & Champness, 2007). Such genes include genes for antibiotic resistance, toxin production and various other virulence factors. In *E. coli* in particular, plasmids are the distinguishing factors between strains and also determine how they infect their host cells (Snyder & Champness, 2007).

Escherichia coli O157:H7 for example, harbours a plasmid called pO157. This large plasmid codes for a number of virulence factors used by the bacterium, one of these being a toxin which affects GTPase. GTPase regulates actin structures in eukaryotic cells, and when disrobed can cause the actin fibres (microvilli) on the intestinal epithelial cells to lose structure, collapse or rearrange (Lathem *et al*., 2002). This toxin also cleaves human C1 esterase inhibitor, which then leads to increased inflammation and extensive tissue damage. These are all symptoms characteristic of the disease caused by *E. coli* O157:H7 (Lathem *et al*., 2002). Therefore, if *E. coli* O157:H7 were to lose this plasmid, it would also lose these associated virulence factors.

It can consequently be seen that plasmids play a fundamental role in the differential virulence and survival of *E. coli*. Plasmids code for antibiotic resistance and often pick up genes from neighbouring bacteria, even if they are not from the same genus or species, making the selection pressure present in the environment very important. This is because if selection pressure is increased in the immediate environment, bacteria of all genera need to compete for commodities. This causes the adaption rate (rate of gene acquisition) to increase and therefore

Stellenbosch University http://scholar.sun.ac.za

the occurrence of antibiotic resistance as well as new virulence factors may also be increased (Snyder & Champness, 2007). This means that if there are more microorganisms present, the available resources will be exploited much sooner, increasing the selection pressure in the environment. In 2008 Fremaux *et al*. conducted a study in Canada which showed that during the rainy season microbial counts as well as variation increased. Fremaux *et al*. (2008) therefore concluded that it is expected that selection pressure increases during the rainy season, as population numbers and competition increase. It can therefore be deduced that the occurrence of antibiotic resistance and new virulence factors will also increase during the rainy season in this studied environment.

Classification of Pathogenic *E. coli*

Background

The pathogenic *E. coli* group is divided into various pathotypes/virotypes. These divisions are determined by the presence, or absence, of certain virulence factors (Bolton *et al*., 1996; Puebte & Finlay, 2001; Percival *et al*., 2004; Bhunia, 2008). All *E. coli* strains within a pathotype grouping have similar virulence genes and pathogenicity factors, therefore leading to similar symptoms being caused by members of a particular pathotype. As mentioned, virulence factors leading to discrimination between pathotypes are therefore usually present in plasmids (Snyder & Champness, 2007). By dividing *E. coli* strains into pathotypes, we can see which groups have the ability to cause disease, as well as how severe the said disease may be. Pathotyping also allows for potential risk assessment as it helps put a measure on the associated risk when a certain pathotype is present in a certain concentration.

Enteropathogenic E. coli (EPEC)

Enteropathogenic *E. coli* (EPEC) was the first pathotype to be described, and is mainly found in children and infants (Levine & Edelman, 1984; Bhunia, 2008). Symptoms of EPEC include watery diarrhoea as well as vomiting, accompanied by a low fever and fluid loss of varying degrees, causing severe dehydration and sometimes even death in infants under the age of six months (Puebte & Finlay, 2001). EPEC cells closely adhere to the epithelial cells in the gut, form attachment/affacing lesions but do not produce toxins (Bhunia, 2008). Once ingested, the

Stellenbosch University http://scholar.sun.ac.za

bacterial cells attach to the enterocytes, located on the intestinal wall of the small bowel and cause infection (Levine & Edelman, 1984; Puebte & Finlay, 2001). The type three secretion system (TTSS) is the method then used to introduce translocated intimin receptor (TIR) as well as a consortium of effector molecules directly into the host cell (Bhunia, 2008).

An "attaching and effacing" (EAF) lesion is then formed, which can easily be identified by the formation of a structural pedestal formed by extensive rearrangement of actin molecules (Levine & Edelman, 1984; Bhunia, 2008). The rearrangement of cellular actin also causes a breakdown in the cytoskeleton of the host cell, and as a result, microvilli are deformed and disappear (Lathem *et al.,* 2002). Membrane permeability is also affected by the effector protein EspF, which targets tight junctions. This change in permeability means that the epithelial cells are unable to absorb nutrients efficiently, and therefore cause the extensive fluid loss and mineral imbalance causing disease (Percival *et al.,* 2004; Bhunia, 2008).

Enterohaemoragic E. coli (EHEC) or Vero Cytotoxigenic E. coli (VTEC)

Due to the extremely low infectious dose (<100 bacterial cells) of VTEC it is seen as the most virulent and dangerous group of *E. coli* serotypes. EHEC, which falls into this group, is a producer of shiga-like toxins which are cytotoxic to Vero cells (Karmali, 1989; Percival *et al*., 2004). These toxins, especially Stx1 and Stx2, are extremely virulent and are the causative agents, in EHEC, of haemorrhagic colitis and haemolytic uraemic syndrome (HUS) (Karmali, 1989). The best example of VTEC and EHEC, which dominates literature, is *E. coli* O157:H7 (O'Brien & Holmes, 1987; Müller *et al.,* 2001; Islam *et al.,* 2004; Percival *et al.,* 2004; Bhunia, 2008; Franz *et al*., 2008; Willey *et al*., 2008; Linscott, 2011). This particular strain is often associated with foodborne diseases, and symptoms caused as a result of the shiga-like toxins include abdominal pain, diarrhoea and sometimes vomiting, both of which can become bloody (Karmali, 1989; Percival *et al*., 2004).

EHEC and *E. coli* O157:H7 in particular, are often associated with foodborne diseases and have traditionally been isolated in insufficiently cooked ground beef, raw milk, uncooked sausage and fermented hard salami. These sources have often been found to be the source in major outbreaks of foodborne diseases (O'Brien and Holmes, 1987; Müller *et al.,* 2001; Islam *et al.,* 2004; Percival *et al.,* 2004; Bhunia, 2008; Franz *et al*., 2008; Willey *et al*., 2008; Linscott, 2011). In more recent years however, fresh produce such as lettuce and spinach has been identified as the source of large EHEC outbreaks, causing deaths throughout the United States of America (O'Brien *et al.*, 1992; Bhunia, 2008; Fahs *et al*., 2009).

Stellenbosch University http://scholar.sun.ac.za

Many EHEC strains are also acid resistant and if exposed to a low pH, such as fermented hard salami or apple cider, the bacterial cells are able to survive even lower pH environments such as the stomach and small intestine (Bhunia, 2008). As with EPEC, EHEC cells also intimately bind to epithelial cells in the gut and form attaching/effacing lesions, but EHEC also produces toxins (Bhunia, 2008). Once in the small intestine, bacterial cells attach to the epithelial cells and cause attaching and effacing lesions, similar to EPEC (Percival *et al*., 2004; Bhunia, 2008). After fimbriae-mediated attachment to the intestinal epithelial cells, the bacteria use the TTSS to transmit a signal to the host cell which initiates actin polymerisation, rearrangement of the host cytoskeleton and effacement of the microvilli. As the microvilli deform, the actin molecules rearrange to form the pedestal, such as with EPEC (Puebte & Finlay, 2001; Percival *et al.,* 2004; Bhunia, 2008). The enteropathogenic *E. coli* adherence factor (EAF) then causes the onset of diarrhoea by sloughing enterocytes, causing inflammation and disrupting tight junctions which changes cell membrane permeability and as a result disrupts mineral absorption (Puebte & Finlay, 2001; Bhunia, 2008).

The Shiga toxins produced by EHEC have been shown to be very similar to those produced by the genus *Shigella*. Stx1 was found to be identical to the toxins produced by *Shigella dysenteriae,* and as a result also cause the same symptoms. Stx2 is not quite as similar, but does show homology ranging between 55-57% when compared to the Shiga toxins of *S. dysenteriae* (O'Brien *et al.,* 1992; Puebte & Finlay, 2001; Percival *et al*., 2004). Both Stx1 and 2 are comprised of six smaller parts, one subunit A and six subunit B's, which each have similar functions in both toxins. Subunit A is the biologically active subunit which holds the toxicity, while subunit B's main function is linked to the attachment of the toxin to the host cell surface, therefore aiding deliverance of the toxin (Percival *et al*., 2004; Puebte & Finlay, 2001). The variation in Stx2 is mainly observed in the B subunits, and as a result the toxins are both antigenically different as well as use different methods of attachment and can therefore invade different cells sometimes (DeGrandis *et al.,* 1989; Melton-Celsa & O'Brien*,* 1998; Puebte & Finlay, 2001).

Enteroaggregative E. coli (EAEC)

Enteroaggregative *E. coli* (EAEC) are named as such due to their distinguishing characteristic to aggregate and adhere to Hep-2 cells (Puebte & Finlay, 2001; Percival *et al.,* 2004; Bhunia, 2008; Okhuysen & Du Pont, 2010). The EAEC cells attach to epithelial cells, form aggregates but do not invade/enter cells at all. EAEC cells do however produce toxins which determine virulence (Bhunia, 2008). The methods used by the bacteria are not fully understood, and the exact toxin

Stellenbosch University http://scholar.sun.ac.za

responsible for symptoms is not yet known (Bhunia, 2008). Symptoms caused by EAEC include watery diarrhoea, which is often mucoid in nature and can last up to 14 days, as well as a low fever. Vomiting is not a common symptom of EAEC infections and diarrhoea can become bloody, although it is not common (Puebte & Finlay, 2001).

EAEC infects the host cells in three stages, the first two of which involves attachment to the mucosal membrane in the intestinal tract as well as the initiation of mucous production and secretion. Both these actions then lead to the formation of a thick, mucoidal biofilm (Percival *et al.,* 2004). The bacterial cells then tend to clump together (aggregate) in a 'stacked brick' formation (Percival *et al.,* 2004; Bhunia, 2008; Okhuysen & Du Pont, 2010). Adhesion between cells, as well as attachment to host cells, is fimbrial mediated by adhesins coded for on a large plasmid (Puebte & Finlay, 2001). EAEC cells also have the ability to increase mucus production and secretion of the host goblet cells. This leads to a thick mucoid layer surrounding the host, as well as bacterial cells. This makes for better adherence of bacterial cells and aggregate formation within a mucoid biofilm (Puebte & Finlay, 2001; Bhunia, 2008). Some EAEC strains are also able to cause tissue damage which results in villus atrophy as well as various other cytotoxic effects.

Enteroinvasive E. coli (EIEC)

Enteroinvasive *E. coli* (EIEC) is often mistaken for an infection caused by *Shigella flexneri*. This is because symptoms caused by EIEC are very similar to those caused by *S. flexneri* as a result of genetics, biochemistry and pathogenicity traits shared by EIEC and species falling under the *Shigella* genus (Puebte & Finlay, 2001; Bhunia, 2008). EIEC cells invade epithelial cells in the gut by adhesion, followed by internalisation and subsequent lateral cell-to-cell movement and intracellular multiplication. EIEC does however not produce any toxins, and the most prominent symptom caused by this pathotype is dysentery, which it shares with *Shigella* spp. (Puebte & Finlay, 2001; Percival *et al.,* 2004; Bhunia, 2008).

EIEC strains are also mostly non-motile as well as lysine decarboxylase and lactose negative. EIEC invades the colonial epithelial cells by interaction with the mucous secreted by goblet cells and subsequent invasion of the epithelial cells. Once internalised, bacterial cells multiply intracellularly, move to adjacent host cells and cause lysis of the host cell vacuole (Puebte & Finlay, 2001). Macrophages are also invaded by the bacterial cells, and in this case host cells are killed by induction of apoptosis (Menard *et al*., 1996; Puebte & Finlay, 2001). Along with dysentery-like symptoms, watery diarrhoea, with mucous, blood and pus in faeces can also be

Stellenbosch University http://scholar.sun.ac.za

observed. Most cases exhibit fever, although due to the lack of shiga, or shiga-like, toxin production, EIEC does not cause haemolytic uremic syndrome (HUS) (Puebte & Finlay, 2001).

Diffusely Adhering E. coli (DAEC)

Similar to EAEC, diffusely adhering *E. coli* (DAEC) also adheres to the Hep-2 cells. Conversely, DAEC does not invade the epithelial cells or produce toxic compounds (Puebte & Finlay, 2001; Percival *et al.,* 2004; Bhunia, 2008). DAEC is usually associated with persistent watery diarrhoea in young children, especially between the ages of two and five years, and infants (Levine *et al*., 1993). Disease is however most prominent in older children and severity of symptoms also increase with the age of the host (Puebte & Finlay, 2001; Bhunia, 2008). Although it has been associated with diarrhoea, some speculation still exists regarding exact virulence factors used by the pathogen in order to cause disease. This being said, studies have shown that attachment of the bacterial cells to host cell is mainly fimbrial-mediated, but also dependent on afimbrial adhesins (Afa) arising from the Afa/Dr family of adhesins (Puebte & Finlay, 2001; Bhunia, 2008).

Enterotoxigenic E. coli (ETEC)

Enterotoxigenic *E. coli* (ETEC), which is also often referred to as the causative agent of 'travellers' diarrhoea', is one of the leading causes of diarrhoea from a foodborne source (Puebte & Finlay, 2001; Percival *et al.,* 2004; Bhunia, 2008). The first ETEC infection was reported in 1956 in Calcutta, India, and was noted as exhibiting cholera-like symptoms (Bhunia, 2008). Symptoms include watery diarrhoea of varying degrees, severe fluid-loss, and in some cases abdominal cramps, vomiting and fever are evident (Puebte & Finlay, 2001; Percival *et al*., 2004; Bhunia, 2008). ETEC is also one of the few *E. coli* pathotypes which often causes disease in adults as well as children, although severity is worse in children under the age of two, and is especially prevalent in the tropics (Bhunia, 2008).

ETEC, once ingested, adheres to the epithelial cells in the gut but does not invade the epithelial cells. Toxins are produced by the bacterial cells and secreted into the host cell where they cause numerous metabolic disruptions. These metabolic disruptions inhibit sodium and chloride ion absorption into the host cell (Puebte & Finlay, 2001; Bhunia, 2008). The general lack of fever, vomiting and non-mucoidal diarrhoea in case studies is also consistent with acute, toxin-mediated, infections and not systemic infections (Puebte & Finlay, 2001). Local inhabitants in the

Stellenbosch University http://scholar.sun.ac.za

tropics are typically immune to the most prevalent ETEC serogroups as a result of continual exposure to the pathogen and subsequent antibody build up. Infections are customarily treated with antibiotics but treatment is however not always effective due to the increasing antibiotic resistance among ETEC serogroups (Puebte & Finlay, 2001) as a result of plasmid acquisition. As with cholera; patients are then treated with rehydration combined with salt and mineral replacement because with proper hydration the illness is self-limiting and will eventually cease (Puebte & Finlay, 2001; Bhunia, 2008).

ETEC has well documented virulence factors, and produces various toxins which fall under either heat stable (ST) toxins, or heat-labile toxins (LT). Approximately a third of all ETEC serogroups produce only LT, a third produce only ST and the remaining proportion produce both LT and ST (Puebte & Finlay, 2001; Bhunia, 2008). The heat labile toxins closely resemble those produced by *Vibrio cholerae* and also cause similar symptoms; hence the mistaken diagnosis of ETEC infections in some cases (Bhunia, 2008). Two heat-labile toxins, LT-I and LT-II, are produced by ETEC. Where LT-I is expressed by ETEC infecting both humans and animals while LT-II is predominantly found in animal infections (Bhunia, 2008).

The heat-stable toxins (ST) are small, 2 kDa, peptide toxins which are extremely heat stable. These toxins have been known to be stable at 100°C for a time period of up to 30 minutes (Bhunia, 2008). Two main types of heat stable toxins also exist, namely STh isolated from human hosts only, and STp which has been isolated from pigs but are also found in ETEC strains infecting humans (Puebte & Finlay, 2001; Bhunia, 2008). STh's main function is to bind guanylate cyclase C, which results in the activation of cyclic guanyl monophosphate (cGMP). Activation of the cGMP in the host cell signals a change in GMP which causes disruptions in the cell ion pumps. Cell ion pumps then pump out large amounts of chloride ions into the extracellular matrix, and chloride and sodium ion uptake back into the host cell is inhibited (Puebte & Finlay, 2001; Bhunia, 2008). LT-I works in a similar way as STh, with the only difference being that LT-I affects cyclic adenosine monophosphate (cAMP), and not cGMP (Puebte & Finlay, 2001; Bhunia, 2008).

Extra-intestinal Pathogenic E. coli (ExPEC)

Most *E. coli* strains, and all those already discussed, are associated with or cause infections related to the gut and intestinal epithelial cells. Those that do cause infection show symptoms such as diarrhoea, vomiting and fever. There are however a few strains which are not associated with the intestinal tract, and cause infections elsewhere in the human host (Puebte & Finlay, 2001; Du

Stellenbosch University http://scholar.sun.ac.za

Plessis *et al.*, 2011). These pathotypes are known as extra-intestinal pathogenic *E. coli* (ExPEC) (Johnson & Russo, 2002).

Although ExPEC is not the main pathotype reported about in literature and is not responsible for major food-borne disease outbreaks, it is still extremely dangerous and causes numerous deaths and serious illnesses worldwide (Phillips *et al.,* 1988). ExPEC infects other regions of the body other than the intestinal cells and is composed of a whole range of serotypes and strains which all infect the host in a slightly different way and have different target cells (Johnson & Russo, 2002). Aside from being the most common cause of urinary tract infections (UTI's) ExPEC also cause a host of other dangerous infections such as neonatal meningitis, neonatal sepsis, nosocomial pneumonia and wound infections (Johnson & Russo, 2002). These infections can occur at varying degrees of severity, but in the cases of neonatal sepsis and meningitis, the infections can often be fatal (Louvois, 1994).

Uropathogenic E. coli (UPEC)

Uropathogenic *E. coli* (UPEC) is one of these ExPEC strains, and infects the urinary tract causing urinary tract infections (UTI's), urethritis, cystitis, and pyelonephritis (Puebte & Finlay, 2001). The majority (80%) of all community-acquired urinary tract infections, as well as approximately 30% of UTI's picked up in hospitals are caused by UPEC (Puebte & Finlay, 2001; Du Plessis *et al.*, 2011). Nosocomial infections are often as a result of catheters being left in for an extended period of time, which eradicates the 'flushing' action of urine through the urinary tract. This 'flushing' action of urine is often the factor which minimises UPEC infections as it stops the bacterial cells from attaching to the urinary tract wall and causing infection (Puebte & Finlay, 2001).

UPEC can be introduced into the urinary tract in a number of ways as UPEC is sometimes present in the colon, and can then move, along with water and other fluids, to the kidneys and bladder and attach to cells in the urinary tract. Infection can also occur at the distal tip of the urethra, as the rest of the urinary tract is typically sterile, and infection can then result from ascending migration of the bacteria due to colonisation (Puebte & Finlay, 2001). Further colonisation of the infection can cause infection in the bladder as well as kidneys if untreated. In extreme cases, the bacteria can enter into the blood stream of the host and cause a systemic infection (Puebte & Finlay, 2001; Du Plessis *et al.*, 2011).

Infections as a result of UPEC are usually characterised by painful urination as well as difficulties to urinate. Symptoms do however vary depending on the infected region as well as the

Stellenbosch University http://scholar.sun.ac.za

severity of the infection. If the infection reaches the kidneys; lower back pain, fever and even vomiting can result (Puebte & Finlay, 2001; Du Plessis *et al.*, 2011). Many UPEC strains, especially those causing nosocomial infections, often carry some antibiotic resistant genes, making treatment problematic. Even when taking this into account, antibiotics are still the most widely used, and successful treatment available for UPEC infections (Puebte & Finlay, 2001, Manges *et al.*, 2001; Du Plessis *et al.*, 2011). In infections resulting from the use of a catheter, removal of the catheter reinstates the 'flushing' action of urine in the urinary tract which can dislodge the bacteria and help clear up the infection if it is not as severe yet.

The virulence factors used by UPEC strains are numerous as they need to use strong virulence factors to attach to the urinary tract wall and cause infection, as urine flow is the biggest obstacle faced by the bacteria. UPEC genes code for a number of adhesins to facilitate attachment to the uroepithelial cells. These adhesins are both afimbrial and fimbrial in nature but not much is known about the exact adhesins used and how they function as many of the adhesins' functions overlap (Kuehn *et al.*, 1992; Puebte & Finlay, 2001). Two pili have been identified; type 1 pili as well as P pili (Kuehn *et al.*, 1992). Along with these pili, UPEC strains also often produce haemolysin and CNF-1 toxin to cause disease (Kuehn *et al.*, 1992; Puebte & Finlay, 2001). Non-toxigenic, but still pathogenic variants can also produce a number of different capsular proteins (usually K1 and/or K5), or use lipopolysaccharides to induce cytokine production in the host cells (Puebte & Finlay, 2001).

E. *ESCHERICHIA COLI* PHYLOGENY

Phylogenetic analysis of *E. coli* can also be used in order to better understand the grouping patterns of *E. coli* isolates coming from various sources. It has been shown that *E. coli* strains fall under one of four main phylogenetic groups, namely A, B1, B2 and D (Herzer *et al*., 1990; Denamur *et al*., 1999; Carlos *et al*., 2010). These groups can be further divided into several subgroups, namely A_0, A_1, B1, $B2_2$, $B2_3$, D_1 and D_2 (Clermont *et al*., 2000). Numerous studies have been done on the phylogeny of *E. coli* over the years, and as a result many deductions can be made regarding the origin and potential virulence of an *E. coli* strain by determining just its phylogenetic grouping (Herzer *et al*., 1990; Denamur *et al*., 1999; Clermont *et al*., 2000; Kotlowski *et al*., 2007; Carlos *et al*., 2010).

Stellenbosch University http://scholar.sun.ac.za

The four main phylogenetic groups have been used for years (Herzer *et al.*, 1990) to categorise *E.* coli strains into groups, and then linking these groups to various possible origins and potential virulence factors. For example, according to Carlos *et al*. 2010, most environmental strains fall under the B1 phylogenetic group, while extraintestinal strains are commonly found in phylo-groups B2 and D (Denamur *et al.*, 1999; Johnson & Stell, 2000). *Escherichia coli* strains which are classified as intestinal pathogenic strains are spread across groups A, B1 and D (Pupo *et al.*, 1997), whereas most commensal strains have been found to fall under phylo-group A (Bingen *et al.*, 1998).

In 2000, Clermont *et al*. developed and started using a PCR based technique which allowed for the further division of *E. coli* strains into the seven subgroups. A combination of three genetic markers, *chuA, yjaA* and DNA fragment TspE4.C2, are tested for and markers visualised. By being able to determine which combination of markers are present (in the genetic material of the bacterial strain), each strain can be divided into its respective phylogenetic subgroup (Clermont *et al.*, 2000). These subgroups can then be used to make deductions about each strain, such as a possible origin as well as potential virulence factors on a more reliable basis than when using only the four main phylo-groups.

A study in 2010 by Carlos *et al*. looked particularly at *E. coli* strains originating from different hosts (humans, cows, pigs, chickens, sheep and goats). It was found that *E. coli* strains in the $B2_3$ subgroup, meaning it contains all thee target markers, were exclusively strains isolated from human sewage and were therefore of a human origin. Samples isolated from chicken faecal matter all fell in subgroup A_0, showing that they did not contain any one of the three markers (Carlos *et al.*, 2010). It was also observed that diet definitely influences what kinds of *E. coli* strains were present, with the main divisions being between species that consumed meat (omnivores and carnivores) and those that do not (herbivores). Strains originating from humans and pigs were noted to have a similarity of 88.3%, and cows, goats and sheep were found to have a similarity of 96% (Carlos *et al.*, 2010). This also clearly demonstrates the role of diet on the *E. coli* strains present. Strains originating from chickens showed to be significantly different to all the mammal-originating strains (Carlos *et al.*, 2010).

Phylogenetic groups are of great value when determining the origin of *E. coli* contamination. This is because phylogenetic groups, as said before, allude to the source of the *E. coli* being studied. When linking contamination in an irrigation source to a contamination source origin, this will become extremely beneficial and will hopefully play a part in microbial source tracking.

Stellenbosch University http://scholar.sun.ac.za

F. CONCLUSION

Naturally occurring bodies of water may easily become contaminated via a number of sources which often carry a variety of microbial contamination as well as nitrogen, phosphate and carbon (De Villiers, 2007). This combination of increased nitrogen, phosphates and carbon, as well as the inclusion of microbial contamination in an ever changing water body is a problem, especially in South Africa where water is a dwindling commodity (SAICE, 2006). As a result of the lack of water, farmers turn to natural water sources for irrigation purposes. The excessive use of natural water, lack of decent, and operational, sewerage systems in parts of South Africa as well as agriculture and industry's influence on water quality leads to unsafe water being used for irrigation (Ackermann, 2010; Lötter, 2010).

What magnifies this problem is the fact that recent studies have shown that when water containing microbial contamination is used for irrigation of fresh produce, such as lettuce, tomatoes and herbs, that a portion of the microbial contamination can actually attach to, and in some cases be internalised by the produce (Solomon *et al*., 2000; Islam *et al*., 2004; Mootian *et al*., 2008). In the case of *E. coli,* contamination is particularly important to be looked at in greater detail. This is because, although not all *E.* coli strains are pathogenic, some of those that are, such as those falling under the EHEC classification, have a very low infectious dose (<100 bacterial cells) (Karmali, 1989; Percival *et al*., 2004). This makes them especially dangerous as they can easily cause infection which can lead to fatal diarrhoea and Haemorrhagic Uremic syndrome (HUS) (Bolton *et al*., 1996; Percival *et al*., 2004).

It has also been noted that there is a notable gap in information available when looking at groundwater as opposed to surface water (Bezuidenhout *et al*., 2011; Haramoto *et al*., 2011). This is possibly due to the misconception that groundwater originates from an underground water reservoir and can therefore not be contaminated by outside factors (Bezuidenhout *et al*., 2011). This is however untrue, as the few studies which have looked at the microbial quality of groundwater has found it to be contaminated with coliforms, and *E. coli* amongst others (Jain *et al*., 2009; Bezuidenhout *et al*., 2011; Haramoto *et al*., 2011). It is therefore extremely important to further research the quality of groundwater that is being used for irrigation of fresh produce in order to make farmers, as well as consumers, aware of the dangers and potential risk associated with fresh produce being irrigated with water of a substandard quality.

Stellenbosch University http://scholar.sun.ac.za

Furthermore the type of *E. coli* present in the water is also of great value in order to assess potential risk associated with the water, as not all *E. coli* strains are pathogenic. The majority of the

E. coli population are harmless to humans and form part of the human gut denizens (Avery *et al*., 2008), but pathogenic variants are also often found. All pathogenic *E. coli* are not the same, as some strains are more dangerous than others and their infectious dose varies from strain-to-strain (Puebte & Finlay, 2001; Percival *et al.,* 2004; Bhunia, 2008). Although various studies have been done on microbial contamination in river water, there have not been any studies looking in depth at *E. coli*, particularly not the exact types of *E. coli* present.

This leaves a substantial gap in the information known regarding this subject, and it is thus essential to compare *E. coli* from a variety of water sources before risk models can be setup. In addition to irrigation water sources, potential contamination sources also need to be looked at and *E. coli* found there characterised. By completing this, it is expected that vital gaps in information will be filled, and potential origins of contamination can be speculated using microbial source tracking.

G. References

Abbott, S., O'Connor, J., Robin, T., Zimmer, B.L. & Janda, J.M. (2003). Biochemical properties of a newly described *Escherichia* species, *Escherichia albertii*. *Journal of Clinical Microbiology*, **41**(10), 4852-4854.

Ackermann, A. (2010). *Assessment of microbial loads of the Plankenburg and Berg Rivers and the survival of Escherichia coli on raw vegetables under laboratory conditions*. MSc in Food Science. University of Stellenbosch, Stellenbosch, South Africa.

Ackers, M.L., Mahon, B.E., Leahy, E., Goode, B., Damrow, T., Hayes, P.S., Bibb, W.F., Rice, D.H., Barrett, T.J., Hutwagner, L., Griffin, P.M. & Slutsker, L. (1998). An outbreak of *Escherichia coli* 0157:H7 infections associated with leaf lettuce consumption. *The Journal of Infectious Diseases*, **177**(6), 1588-1593.

ASM (American Society for Microbiology) (2008). Foodborne outbreaks from leafy greens on the rise. *Science Daily*, 17 March 2008. Accessed 22 May 2012 from: http://www.sciencedaily.com/releases/2008/03far/080317164356.htm

Anonymous (2011a). "Outbreak of Shiga toxin-producing E. coli in Germany (22 June 2011, 11:00)". *ECDC*. 22 June 2011. Accessed 9 July 2011 from: http://ecdc.europa.eu/en/activities/sciadvice/Lists/ECDC%20Reviews/ECDC_DispForm.aspx?List=512ff74f-77d4-4ad8-b6d6-bf0f23083f30&ID=1120&RootFolder=%2Fen%2Factivities%2Fsciadvice%2FLists%2FECDC%20Reviews

Stellenbosch University http://scholar.sun.ac.za

Anonymous (2011b). "German-grown food named likely culprit in deadly outbreak". CNN.com. 2010-06-05. Accessed 9 July 2011 from: http://edition.cnn.com/2011/WORLD/europe/06/05/europe.e.coli/index.html?hpt=hp_t2

Anonymous (2011c). EHEC O104:H4 outbreak event in Germany clarified: sprouts of fenugreek seeds imported from Egypt as underlying cause. *Bundesinstitut fur Risikobewertung.* http://www.bfr.bund.de/en/press_information/2011/21/ehec_o104_h4_outbreak_event_in_germany_clarified__sprouts_of_fenugreek_seeds_imported_from_egypt_as_underlying_cause-83273.html, Accessed 08 May 2012.

Avery, L.M., Williams, A.P., Killham, K. & Jones, D.L. (2008). Survival of *Escherichia coli* O157:H7 in waters from lakes, rivers, puddles and animal-drinking troughs. *Science of the Total Environment*, **389**, 378-385.

Beuchat, L.R. & Ryu, J.H. (1997). Produce handling and processing practices. *Emerging Infectious Diseases*, **3**(4), 459-465.

Bezuidenhout, C.C., Mthembu, N. & Lin, J. (2002). Microbiological evaluation of the Mhlathuze River, KwaZulu-Natal (RSA). *Water SA*, **28**(3), 281-286.

Bezuidenhout, C.C & the North-West University Team (2011). A scoping study on the environmental water (groundwater and surface water) quality and management in the North-West Province, South Africa. Water Research Commission (WRC) Report No. KV 278/11. WRC Printers, Pretoria.

Bhunia, A.K. (2008). *Escherichia coli.* In: *Foodborne Microbial Pathogens: Mechanisms and Pathogenesis*. (edited by Helderman, D.R.). Pp. 183-200. New York: Springer.

Bingen, E., Picard, B., Brahimi, N., Mathy, S., Desjardins, P., Elion, J. & Denamur, E. (1998) Phylogenetic analysis of *Escherichia coli* strains causing neonatal meningitis suggests horizontal gene transfer from a predominant pool of highly virulent B2 group strain. *Journal of Infectious Diseases,* **177**, 642-650.

Bolton, F.J., Crozier, L. & Williamson, J.K. (1996). Isolation of *Escherichia coli* 01 57 from raw meat products. *Letters in Applied Microbiology*, **23**, 317-321.

Brenner, D.J., Davis, B.R., Steigerwalt, A.G., Riddle, C.F., McWhorter, A.C., Allen, S.D., Farmer, J.J., Saitoh, Y. & Fanning, G.R. (1982a). Atypical biogroups of *Escherichia coli* found in clinical specimens and description of *Escherichia hermannii* sp. nov. *Journal of Clinical Microbiology*, **15**, 703-713.

Brenner, D.J., McWhorter, A.C., Leete Knudson, J.K. & Steigerwalt, A.G. (1982b). *Escherichia vulneris:* a new species of *Enterobacteriaceae* associated with human wounds. *Journal of Clinical Microbiology*, **15**, 1133-1140.

Burgess, N.R.H., McDermott, S.N. & Whiting, J. (1973). Laboratory transmission of *Enterobacteriaceae* by the oriental cockroach, *Blatta orientalis*. *Journal of Hygiene*, **71,** 9-14.

Carlos, C., Pires, M.M., Stoppe, N.C., Hachich, E.M., Sato, M.I.Z., Gomes, T.A.T., Amaral, L.A. & Ottoboni, L.M.M. (2010). *Escherichia coli* phylogenetic group determination and its application in the identification of the major animal source of fecal contamination. *BMC Microbiology*, **10**, 161-171.

CDC (Centers for Disease Control and Prevention) (2006). Ongoing multistate outbreak of *Escherichia coli* serotype O157:H7 infections associated with consumption of fresh spinach – United States. *Morbidity and Mortality Weekly Report*, **55**, 1045-1046.

Clermont, O., Bonacorsi, S. & Bingen, E. (2000). Rapid and simple determination of the *Escherichia coli* phylogenetic group. *Applied and Environmental Microbiology*, **66**(10), 4555-4558.

Denamur, E., Elion, J., Duriez, P., Brahimi, N., Bingen, E., Picard, J.P., Garcia, J.S. & Gouriou, S. (1999). The link between phylogeny and virulence in *Escherichia coli e*xtraintestinal infection. *Infection and Immunity*, **67**(2), 546-553.

Stellenbosch University http://scholar.sun.ac.za

DeGrandis, S., Law, H., Brunton, J., Gyles, C. & Lingwood, C.A. (1989). Globotetraosylceramide is recognized by the pig edema disease toxin. *Journal of Biological Chemistry,* **264**, 12520-12525.

De Villiers, (2007). The deteriorating nutrient status of the Berg River, South Africa. *Water SA*, **33**(5), 659-664.

Doyle, M. P. (1991). *Escherichia coli* 0157:H7 and its significance in foods. *International Journal of Food Microbiology,* **12**, 289-302.

Du Plessis, C., Warkentien, T. & Bavaro, M. (2011). Uropathogenic *Escherichia coli*. *The Female Patient,* **36**, 18-23.

DWAF (Department of Water Affairs and Forestry). (1996). Agricultural Use: Irrigation. In: *South African Water Quality Guidelines. Volume 4.* (2nd Ed.). (Edited by Holmes, S. CSIR Environmental Services). Pretoria: Department of Water Affairs and Forestry.

Effler, P., Isaäcson, M., Arntzen, L., Heenan, R., Canter, P., Barrett, T., Lee, L., Mambo, C., Levine, W., Zaidi, A. & Griffin, P.M. (2001). Factors contributing to the emergence of *Escherichia coli* O157 in Africa. *Emerging Infectious Diseases*, **7**(5), 812-819.

Escherich, T. (1988). The intestinal bacteria of the neonate and breast-fed infant: 1884 (reprint). *Review of Infectious Diseases*, **10,** 1220–1225.

Fahs, F., Mittelhammer, R.C. & McCluskey, J.J. (2009). *Escherichia coli* outbreaks affect demand for salad vegetables. *Choices*, **24**(2), 26-29.

FAO/WHO (Food and Agriculture Organization of the United Nations/ World Health Organization) (2008). Microbial hazards in fresh fruit and vegetables. *Microbial Risk Assessment Series: pre-publication version*. Accessed 22 May 2012 from: http://www.who.int/foodsafety/publications/micro/MRA_FruitVeges.pdf

Farmer, J.J., Fanning, G.R., Davis, B.R., O'Hara, C.M., Riddle, C., Hickman-Brenner, F.W., Asbury, M.A., Lowery, V.A. & Brenner, D.J. (1985a). *Escherichia fergusonii* and *Enterobacter taylorae*, two new species of *Enterobacteriaceae* isolated from clinical specimens. *Journal of Clinical Microbiology*, **21**, 77-81.

Farmer, J.J., Davis, B.R., Hickman-Brenner, F.W., McWhorter, A., Huntley-Carter, G.P., Asbury, M.A., Riddle, C., Watehen-Grady, H.G., Elias, C. & Fanning, G.R. (1985b). Biochemical identification of new species and biogroups of *Enterobacteriaceae* isolated from clinical specimens. *Journal of Clinical Microbiology*, **21**(1), 46-76.

Forsythe, S.J. (Ed.). (2000). Food poisoning microorganisms. In: *The Microbiology of Safe Food*. Pp. 143 – 145. Oxford: Blackwell Science Ltd.

Franz, E., Semenov, A.V. & van Bruggen, A.H.C. (2008). Modelling the contamination of lettuce with *Escherichia coli* O157:H7 from manure-amended soil and the effect of intervention strategies. *Journal of Applied Microbiology*, **105**, 1569-1584.

Fremaux, B., Boa, T., Chaykowski, A., Kasichayanula, S., Gritzfeld, J., Braul, L. & Yost, C. (2008). Assessment of the microbial quality of irrigation water in a prairie watershed. *Journal of Applied Microbiology*, **106**, 442-454.

Geldreich, E. (1978). Bacterial populations and indicator concepts in feces, stormwater and solid wastes. In: *Indicators of Viruses in Water and Food*. (edited by Berg, G.). Pp. 51-65. Michigan: Ann Arbor Science Publishers Inc.

Gerba, C.P. & Choi, C.Y. (2009). Water Quality. In: *The Produce Contamination Problem*. (edited by Sapers, G.M., Solomon, E.B. & Matthews, K.R.). Pp. 105-118. London: Elsevier Inc.

Gil, M.I. & Selma, M.V. (2006). Overview of hazards in fresh-cut produce production: control and management of food safety hazards. In: *Microbial Hazard Identification in Fresh Fruits and Vegetables* (edited by James, J.). Pp. 95-109. New Jersey: John Wiley & Sons.

Stellenbosch University http://scholar.sun.ac.za

Griesel M. & Jagals, P. (2002) Faecal indicator organisms in the Renoster Spruit system of the Modder-Riet River catchment and implications for human users of the water. *Water SA*, **28**(2), 227-234.

Guenther, P.M., Dodd, K.W., Reedy, J. & Krebs-Smith, S.M. (2006). Most Americans eat much less than recommended amounts of fruits and vegetables. *Journal of the American Dietetic Association*, **106**(9), 1371-1379.

Hacker, J., Blum-Oehler, G., Muhldorfer, I., & Tschape, H. (1997). Pathogenicity islands of virulent bacteria: Structure, function and impact on microbial evolution. *Molecular Microbiology*, **23**, 1089-1097.

Haramoto, E., Yamada, K. & Nishida, K. (2011). Prevalence of protozoa, viruses, coliphages and indicator bacteria in groundwater and river water in the Katmandu Valley, Nepal. *Transactional of the Royal Society of Tropical Medicine and Hygiene*, **105**, 711-716.

Harris, L.J., Farber, J.N., Beuchat, L.R., Parish, M.E., Suslow, T.V., Garrett, E.H. & Busta F.F. (2003). Outbreaks associated with fresh produce: incidence, growth, and survival of pathogens in fresh and fresh-cut produce. *Comprehensive Reviews in Food Science and Food Safety*, **2** (supplement), 78-141.

Herzer, P.J., Inouye, S., Inouye, M. & Whittam, T.S. (1990). Phylogenetic distribution of branched RNA-linked multicopy single-stranded DNA among natural isolates of *Escherichia coli*. *Journal of Bacteriology*, **172**(11), 6175-6181.

Holt, J.G., Krieg, N.R., Sneath, P.H.A., Staley, J.T. & Williams, S.T. (Eds.) (1994). *Bergey's Manual of Determinative Bacteriology* (9th Ed.). Pp. 179-178. Baltimore: Williams & Wilkins.

Hughes, D.A. (2004). Incorporating groundwater recharge and discharge functions into an existing monthly rainfall-runoff model. *Hydrological Sciences Journal*, **49**(2), 297-311.

Huys, G., Cnockaert, M., Janda, J.M. & Swings, J. (2003). *Escherichia albertii* sp. nov., a diarrhoeagenic species isolated from stool specimens of Bangladeshi children. *International Journal of Systematic and Evolutionary Microbiology*, **53**, 807–810.

Ishii, S. & Sadowsky, M.J. (2008). *Escherichia coli* in the environment: Implications for water quality and human health. *Microbes and Environments*, **23**(2), 101-108.

Islam, M., Doyle, M.P., Phatak, S.C., Millner, P. & Jiang, X. (2004). Persistence of enterohemorrhagic *Escherichia coli* O157:H7 in soil and on leaf lettuce and parsley grown in fields treated with contaminated manure composts or irrigation water. *Journal of Food Protection*, **67**(7), 1365-1370.

Islam, M., Doyle, M.P., Phatak, S.C., Millner, P. & Jiang, X. (2005). Survival of *Escherichia coli* O157:H7 in soil and on carrots and onions grown in fields treated with contaminated manure composts or irrigation water. *Journal of Food Protection*, **22**, 63-70.

Jain, C.K., Bandyopadhyay, A. & Bhadra, A. (2009). Assessment of ground water quality for drinking purpose, District Nainital, Uttarakhand, India. *Environment Monitoring and Assessment*, **166**, 663-676.

Johnson, J.R. & Russo, T.A. (2002). Extraintestinal pathogenic *Escherichia coli*: "The other bad *E. coli*". *Journal of Laboratory and Clinical Medicine*, **139**, 155-162.

Johnson, J.R. & Russo, T.A. (2003). Review: Medical and economic impact of extraintestinal infections due to *Escherichia coli*: focus on an increasingly important endemic problem. *Microbes and Infection*, **5**, 449-456.

Johnson, J.R. & Stell, A.L. (2000). Extended virulence genotypes of *Escherichia coli* strains from patients with urosepsis in relation to phylogeny and host compromise. *Journal of Infectious Diseases*, **181**, 261-272.

Johnston, L.M., Moe, C.L., Moll, D. & Jaykus, L. (2006). The epidemiology of produce-associated outbreaks of foodborne disease. In: *Microbial Hazard Identification in Fresh Fruits and Vegetables* (edited by James, J.). Pp. 38-52. New Jersey: John Wiley & Sons.

Stellenbosch University http://scholar.sun.ac.za

Karmali, M. A. (1989). Infection by verocytotoxin-producing *Escherichia coli*. *Clinical Microbiology Review*, **2**(1), 15-38.

Kauffmann, F. (1947). Review: the serology of the coli group. *Journal of Immunology*, **57**, 71–100.

Kfir, R., Hilner, C., Du Preez, M. & Bateman, B. (1995). Studies on the prevalence of giardia cysts and cryptosporidium oocysts in South-African water. *Water Science and Technology*, **31**, 435-438.

Kikine, T.N.F. (2011). *Profiling of Potential Pathogens From Plankenburg River Water Used for the Irrigation of Fresh Produce*. MSc in Food Science Thesis. Stellenbosch University, Stellenbosch, South Africa.

Kotlowski, R., Bernstein, C.N., Sepehri, S. & Krause, D.O. (2007). High prevalence of *Escherichia coli* belonging to the B2+D phylogenetic group in inflammatory bowel disease. *Gut*, **56**, 669-675.

Kuehn, M.J., Heuser, J., Normark, S. & Hultgren, S.J. (1992). P pili in uropathogenic *E. coli* are composite fibres with distinct fibrillar adhesive tips. *Letters to Nature*, **356**, 252-255.

Lathem, W.W., Grys, T.E., Witowski, S.E., Torres, A.G., Kaper, J.B., Tarr, P.I & Welch, R.A. (2002). StcE, a metalloprotease secreted by *Escherichia coli* O157:H7, specifically cleaves C1 esterase inhibitor. *Molecular Microbiology*, **45**, 277-288.

Levine, M.M. & Edelman, R. (1984). Enteropathogenic *Escherichia coli* of classic serotypes associated with infant diarrhoea: epidemiology and pathogenesis. *Epidemiologic Review*, **6**, 31-51.

Levine, M.M., Ferreccio, C., Prado, V., Cayazzo, M., Abrego, P., Martinez, J., Maggi, L., Baldini, M.M., Martin, W., Maneval, D. (1993). Epidemiologic studies of *Escherichia coli* diarrheal infections in a low socioeconomic level peri-urban community in Santiago, Chile. *American Journal of Epidemiology*, **138**, 849-869.

Linscott, A.J. (2011). Food-borne illnesses. *Clinical Microbiology Newsletter*, **33**, 41-45.

Lötter, M. (2010). *Assessment of microbial loads present in two Western Cape Rivers used for irrigation of vegetables*. MSc in Food Science. University of Stellenbosch, Stellenbosch, South Africa.

Louvois, J. (1994). Acute bacterial meningitis in the new born. *Journal of Antimicrobial Chemotherapy,* **34**(A), 61-73.

Lynch, M., Painter, J., Woodruff, R. & Braden, C. (2006). Surveillance for foodborne-disease outbreaks: United States 1998-2002. *Morbidity and Mortality Weekly Report*, Centres for Disease Control and Prevention. Accessed 06 March 2012 from: http://www.cdc.gov/mmwr/preview/mmwrhtml/ss5510a1.htm?s_cid=ss5510a1_e.

Manges, A.R., Johnson, J.R., Foxman, B. & O'Bryan, T.T. (2001). Widespread distribution of urinary tract infections caused by a multidrug-resistant *Escherichia coli* clonal group. *The New England Journal of Medicine*, **345**(14), 1007-1014.

Matthews, K.R. (Ed.). (2006). Microorganisms associated with fruits and vegetables. In: *Microbiology of Fresh Produce*. Pp. 1-19. Washington DC: ASM Press.

Melloul, A., Amahmid, O., Hassani, L. & Bouhoum, K. (2002). Health effect of human wastes use in agriculture in El Azzouzia (the wastewater spreading area of Marrakesh city, Morocco). *International Journal of Environmental Health Research*, **12**, 17–23.

Melton-Celsa, A. R., and O'Brien, A. D. (1998). Structure, biology, and relative toxicity of shiga toxin family members for cells and animals. In: *Escherichia coli 0157:H7 and Other Shiga Toxin-Producing E. coli Strains* (Edited by Kaper, J. B. & O'Brien, A.D). Pp. 121-128. Washington DC: American Society for Microbiology.

Mena, K.D. (2006). Produce quality and foodborne disease: assessing water's role. In: *Microbial Hazard Identification in Fresh Fruits and Vegetables* (Edited by James, J.). Pp. 95-109. New Jersey: John Wiley & Sons.

Stellenbosch University http://scholar.sun.ac.za

Menard, R., Dehio, C. & Sansonetti, P. J. (1996). Bacterial entry into epithelial cells: The paradigm of *Shigella. Trends in Microbiology*, **4**, 220-226.

Michino, H., Araki, K., Minami, S., Takaya, S., Sakai, N. & Miyazaki, M. (1999). Massive outbreak of *Escherichia coli* 0157:H7 infection in school hildren in Sakai City, Japan, associated with consumption of white radish sprouts. *American Journal of Epidemiology*, **150**(8), 787-796.

Mootian, G., Wu, W., Pang, H., and Matthews, K. R. (2008). Transfer prevalence of *Escherichia coli* O157:H7 from soil, water, and manure contaminated with low numbers of the pathogen to lettuce plants of varying age. In: *Program and Abstract Book IAFP Annual Meeting*. p. 67. Oklahoma: Columbus.

Müller, E.E., Ehlers, M.M. & Grabow, W.O.K. (2001). The occurrence of *E. coli* O157:H7 in South African water sources intended for direct and indirect human consumption. *Water Research*, **35**(13), 3085-3088.

O'Brien A. D. & Holmes R. K. (1987). Shiga and Shiga-like toxins. *Microbiological Reviews*, **51**, 206–220.

O'Brien, A. D., Tesh, V. L., Donohue-Rolfe, A., Jackson, M. P., Olsnes, S., Sandvig, K., Lindberg, A. A. & Keusch, G. T. (1992). Shiga toxin: Biochemistry, genetics, mode of action, and role in pathogenesis. *Current Topics in Microbiology and Immunology,* **180**, 65-94.

Okafo, C.N., Umoh, V.J. & Galadima, M. (2003). Occurrence of pathogens on vegetables harvested from soils irrigated with contaminated streams. *The Science of the Total Environment*, **311**, 49-56.

Okhuysen, P.C. & Du Pont, H.L. (2010). Entero-aggregative *Escherichia coli* (EAEC): A cause of acute and persistent diarrhoea of worldwide importance. *Journal of Infectious Diseases,* **202**(4), 503–505.

Paulse, A.N., Jackson, V.A. & Khan, W. (2009). Comparison of microbial contamination at various sites along the Plakenburg and Diep Rivers, Western Cape, South Africa. *Water* SA, **35**(4), 469-478.

Percival, S., Chalmers, R., Embrey, M., Hunter, P., Sellwood, J. & Wyn-Jones, P. (Eds.). (2004). *Escherichia coli.* In: *Microbiology of Waterborne Diseases: Microbiological Aspects and Risks*. New York: Academic Press.

Phillips, I., Eykyn, S., King, A., Grandsden, W.R., Rowe, B. & Frost, J.A. (1988). Epidemic multiresistant *Escherichia coli* infection. *West Lambeth Health District*, **1**, 1038-1041.

Puebte, J.L. & Finlay, B.B. (2001). Pathogenic *E. coli*. In: *Principles of Bacterial Pathogenesis*. Pp. 387-456. New York: Academic Press.

Pupo, G.M., Karaolis, D.K.R., Lan, R. & Reeves, P.R. (1997) Evolutionary relationships among pathogenic and non-pathogenic *Escherichia coli* strains inferred from multilocus enzyme electrophoresis and *mdh* sequence studies. *Infection and Immunity*, **65**, 2685-2692.

Rai, P.K. & Tripathi, B.D. (2007). Microbial contamination in vegetables due to irrigation with partially treated municipal wastewater in a tropical city. *International Journal of Environmental Health Research*, **17**(5), 389-395.

SAICE (South African Institute for Civil Engineering). (2006). *The SAICE infrastructure report card for South Africa: 2006*. Pp. 1-16. Midrand: SAICE House.

Selander, R. K., Caugant, D. A. & Whittam, T. S. (1987). Genetic structure and variation in natural populations of *Escherichia coli*. In: *Escherichia coli and Salmonella typhimurium: cellular and molecular biology*, (Edited by: Neidhardt, F. C., Ingraham, J.L., Low, K.B., Magasanik, B., Schaechter, M. & Umbarger, H.E.). Pp. 1625–1648. Washington DC: ASM Press.

Snyder, L. & Champness, W. (Eds.). (2007). A brief history of bacterial genetics. In: *Molecular Genetics of Bacteria* (3rd Ed.), Pp. 8-10. Washington DC: ASM Press.

Stellenbosch University http://scholar.sun.ac.za

SGM (Society for General Microbiology) (2007). Lettuce, leafy greens and *E. coli*. *Science Daily*, 2 September 2007. Accessed 22 May 2012 from: http://www.sciencedaily.com/releases/2007/09/070902193834.htm

Smith, M. S., Thomas, G. W., White, R. E. & Ritonga, D. (1985). Transport of *Escherichia coli* through intact and disturbed soil columns. *Journal of Environmental Quality*, **14**, 87-91.

Solo-Gabriele, H.M., Wolfert, M.A., Desmarais, T.R. & Palmer, C.J. (2000). Sources of *Escherichia coli* in a coastal subtropical environment, *Applied and Environmental Microbiology*. **66**(1), 230-237.

Solomon, E.B., Yaron, S. & Matthews, K.R. (2002). Transmission of *Escherichia coli* O157:H7 from contaminated manure and irrigation water to lettuce plant tissue and its subsequent internalization. *Applied and Environmental Microbiology,* **68**(1), 397-400.

Song, I., Stine, S.W., Choi, C.Y. & Gerba, C.P. (2006). Comparison of crop contamination by microorganisms during subsurface drip and furrow irrigation. *Journal of Environmental Engineering,* **132**, 1243–1248.

Steynberg, M.C., Venter, S.N., de Wet, C.M.E., du Plessis, G., Holhs, D., Rodda, N. & Kfir, R. (1995), Management of microbial water quality: New perspectives for developing areas. *Water Science Technology*, **32**, 183-191.

Svanborg-Edén, C., Jodal, U., Hanson, L.A., Lindberg, U. & Sohl Åkerlund, A. (1976). Variable adherence to normal human urinary tract epithelial cells of *Escherichia coli* strains associated with various forms of urinary tract infections. *The Lancet*, **2**, 490-492.

Tal, N. & Schuldiner, S. (2009). A coordinated network of transporters with overlapping specificities provides a robust survival strategy. *Proceedings of the National Academy of Science of the United States of America*, **106**(22), 9051-9056.

Tauxe, R.V. (1997). Emerging foodborne diseases: an evolving public health challenge. *Emerging Infectious Diseases*, **3**(4), 425-434.

Van Blommestein, A. (2012). *Impact of selected environmental factors on E. coli growth in river water and an investigation of carry-over to fresh produce during irrigation*. MSc in Food Science. University of Stellenbosch, Stellenbosch, South Africa.

WHO (World Health Organization) (1989). Health guidelines for the use of wastewater in agriculture and aquaculture. Geneva, World Health Organization. *Technical Report Series No 776. World Health Organization*, Switzerland, Geneva.

Willey, J.M., Sherwood, L.M. & Woolverton, C.J. (Eds.). (2008). Microbiology of food. In: *Prescott, Harley & Klien's Microbiology*, 7th Ed. Pp. 1023-1048. New York: McGraw Hill.

Woodford, A., Rosewaene, P. & Girman, J. (2005). How much groundwater does South Africa have? DWAF (Department of Water Affairs & Forestry) Pretoria, South Africa. Government Printers.

Zamxaka, M., Pironcheva, G. & Muyima, N.Y.O. (2004). Microbiological and physico-chemical assessment of the quality of domestic water sources in selected rural communities of the Eastern Cape Province, South Africa. *Water SA*, **30**(3), 333-340.

Stellenbosch University http://scholar.sun.ac.za

CHAPTER 3

ENUMERATION AND CHARACTERISATION OF *ESCHERICHIA COLI* FROM IRRIGATION WATER AND POTENTIAL CONTAMINATION SITES

SUMMARY

In the Western Cape, farmers frequently draw water from natural water sources as an alternative means to irrigate crops, fresh produce and fruit. In this study nineteen sites, comprising both contamination sources and irrigation sites, were sampled with the aim of enumerating and characterising *Escherichia coli* from irrigation water and to determine the diversity within the *E. coli* population present in natural water sources.

Total coliform and *E. coli* counts found in contamination source sites were recorded as high as log 7.114 MPN.100 mL^{-1} and log 6.912 MPN.100 mL^{-1}, respectively. Total coliform and *E. coli* counts for irrigation sites were much lower (by approximately one log) than those of contamination sites, with maximum counts of log 5.768 MPN.100 mL^{-1} and log 5.788 MPN.100 mL^{-1}, respectively. It was found that more than one third (5/14) of the irrigation sites had *E. coli* counts exceeding the guidelines for safe irrigation water for fresh produce (of 1 000 counts.100 mL^{-1}), making the water unsuitable for the irrigation of fresh produce (World Health Organisation (WHO, 1989; DWAF, 1996). It was also concluded that groundwater is a better option than surface water when using natural water sources for irrigation of fresh produce.

The Jaccard statistical method was used to create dendrograms which clustered similar *E. coli* strains on the basis of their biochemical profiles. In total 38 clusters were formed, with one to 30 isolates in each. Evidence was also found which confirms that *E. coli* is capable of gene acquisition resulting in acclimatisation and probably an increased survival rate in environmental niches with a fluctuating carbon source. API 20E and matrix assisted laser desorption/ionisation time-of-flight mass spectrometry (MALDI TOF-MS) was used to identify isolated strains, and their identification abilities compared. It was found that MALDI BIOTYPER was a more objective method to use for identification. It was also found that MALDI BIOTYPER was able to identify *E. coli*, and other bacterial strains, when API 20E was not able to. Ultimately it was concluded that the variation of *E. coli* strains present as well as the prevalence of *E. coli* in irrigation water is a matter of concern that needs to be investigated in more detail.

Stellenbosch University http://scholar.sun.ac.za

INTRODUCTION

With the increase in consumption of fresh fruit and vegetables, farmers have been required to increase their annual output of fresh produce (USDA, 1999; Guenther *et al.,* 2006). Due to the low annual rainfall in South Africa, in conjunction with the winter rainfall of the Western Cape, the hot summer months are extremely dry (SAICE, 2006). This means that farmers need to find alternative water sources to irrigate their crops, which usually leads to the use of surface and/or groundwater where available.

Although water quality guidelines have been published by the World Health Organisation (WHO) as well as the Department of Water Affairs and Forestry (DWA), these guidelines are most often not acknowledged and as a result irrigation water used for fresh produce often has a faecal coliform count higher than the recommended *E. coli* guideline of 1 000 counts.100 mL^{-1} (WHO, 1989; DWAF, 1996). Additionally the natural water sources are often polluted with a wide range of microbes, some of which are pathogenic (Ackermann, 2010; Lötter, 2010).

Escherichia coli is one of the most important microorganisms to look for as some strains are pathogenic and *E. coli* is also a gut denizen of humans, cattle and other warm-blooded animals. This means that *E. coli* is a good indicator of faecal contamination as well as an organism which is of great concern to the health of consumers. This has therefore led to a lot more effort going into studying the quality of alternative water sources (Müller *et al.*, 2001; Bezuidenhout *et al.*, 2002; Rai & Tripathi, 2007; Jain *et al.*, 2009; Ackermann, 2010; Lötter, 2010; Bezuidenhout *et al.*, 2011; Haramoto *et al.*, 2011) used for irrigation.

Over the last 25 years there has also been a striking increase in the role of *E. coli* when it comes to food safety (USDA, 1999; Guenther *et al.,* 2006). This is because *E. coli* was mainly associated with the consumption of insufficiently prepared beef (Willey *et al.*, 2008; Linscott, 2011; Charimba *et al.*, 2012). Over the last 15 years, *E. coli* has been found more and more as a pathogen also associated with the consumption of fresh produce, eaten raw or minimally processed (Beuchat & Ryu, 1997; Tauxe, 1997; Gemmell & Schmidt, 2012). After numerous large-scale outbreaks of pathogenic *E. coli* associated with fresh produce (Islam *et al.,* 2004; Franz *et al.*, 2008; Fahs *et al.*, 2009; Linscott, 2011), studies have shown that the source of the *E. coli* found on the fresh produce is the irrigation water (Beuchat & Ryu, 1997; Beuchat, 2002; Solomon *et al.*, 2002; Okafo *et al.*, 2003; Islam *et al.*, 2004; Van Blommestein, 2012). In other words, the irrigation water used to irrigate the fresh produce is of a substandard quality and could possibly contain pathogenic *E. coli*. The *E. coli* can then be carried over to the fresh produce, where it can adhere

Stellenbosch University http://scholar.sun.ac.za

and survive on the fresh produce (Beuchat & Ryu, 1997; Beuchat, 2002; Solomon *et al*., 2002; Okafo *et al*., 2003; Islam *et al*., 2004; Van Blommestein, 2012).

Further problems arise when farmers start using groundwater as an alternative water source, as there is a lack of information concerning the microbial quality of groundwater. Some studies have been done on groundwater, but unfortunately almost all of these studies have looked mainly at the chemical attributes, and not the microbial quality (Adams *et al*., 2001). This is partially due to the misconception that groundwater is clean by default (Bezuidenhout *et al*., 2011). A few studies, both locally and in other countries, have shown recently that groundwater is, however, not as clean as previously thought (Bezuidenhout *et al*., 2011; Haramoto *et al*., 2011). Coliform bacteria and *E. coli* were found in groundwater and therefore pose the same risks as surface water when it comes to irrigation of fresh produce.

Microbial contamination found in natural water ways can come from a number of sources, but the most widely accepted source of *E. coli* contamination is faecal matter (Avery *et al*., 2008). Faecal matter from humans and cattle are of the greatest importance, as these most often carry pathogens (Willey *et al*., 2008; Linscott, 2011), but faecal matter from other warm-blooded animals is also significant. Aside from faecal contamination; human pollution, industry, agriculture and insufficient sewage works also need to be investigated when determining the origin of *E. coli* contamination (Islam *et al.,* 2004; De Villiers, 2007).

It is therefore essential to look at irrigation water as well as potential contamination sources when isolating and characterising *E. coli*. Additionally this could possibly lead to a better understanding of the composition of the *E. coli* population found in natural water sources. The aim of this study was therefore to determine the prevalence of environmental *Escherichia coli* strains in natural water sources as well as contamination sources. This was achieved by firstly characterising *E.* coli strains using biochemical tests and matrix assisted laser desorption/ionisation time of flight mass spectrometry biotyper (MALDI biotyper), after which strains were compared and dissimilarities between strains determined.

Materials and Methods

Site Selection

A total of 19 sites were selected (Tables 1 and 2) and each sampled twice over a period of five months. Of the 19 selected sites, 12 sites were classified strictly as irrigation sources, five strictly as contamination sources and two sites overlapped categories and were classified both as

Stellenbosch University http://scholar.sun.ac.za

irrigation and contamination sources. This meant that in total there were 14 irrigation sites (Table 1) and seven contamination source sites (Table 2) investigated. The sites were chosen mostly in the Stellenbosch and surrounding areas, while there was one site adjacent to the Cape Flats area near Muizenburg, one further North West near Lutzville, one in the Drakenstein area and four sites in Wellington/Paarl area. For the purpose of this study, groundwater was defined as borehole and spring water and surface water as rivers and dams.

Irrigation source sites

The irrigation source sites were selected from a range of different types of water sources including; six boreholes, five rivers, one dam, one spring and one cellar effluent sample which has had sewage added to the effluent (Table 1). Sample sites were chosen on the grounds that the water being sampled must be extracted at the same point as that being used for irrigation of fresh produce.

Table 1. Irrigation sites, their geographical locations and water application

Water source	Geographic Location	Used to irrigate
Borehole A1	Drakenstein	Herbs
Borehole A2	Drakenstein	Herbs
Borehole D	Wellington	Fresh produce, grapes and used in winemaking
Borehole N1	Strandfontein (adj. to Cape Flats)	Herbs and salad products
Borehole N2	Strandfontein (adj. to Cape Flats)	Herbs and salad products
Borehole P	Raithby	Fresh produce as well as home use
Spring	Wellington	Fresh produce as well as home use
Plankenburg River (Plank-3)	Stellenbosch	Fresh produce on various farms
Veldwagters River	Stellenbosch	Fresh produce on various farms
Olifants River	Lutzville	Fresh produce and crops also used in houses
Mosselbank River	Kraaifontein	Fresh produce
Berg River (Berg-2)	Paarl/Franschhoek	Fresh produce and fruit
Winery effluent (Effluent L1)	Stellenbosch (outlying)	Fresh produce consumed by farm workers
Dam	Wellington	Fresh produce, grapes, used in winemaking and also used in house

Stellenbosch University http://scholar.sun.ac.za

Contamination source sites

The second set of sampling sites was chosen to represent potential contamination sources from where the *E. coli* may have originated from. These sites were therefore chosen expecting that they would have a high microbial load, especially in terms of total coliform bacteria and *E. coli* content. The contaminated water also had to be able to reach another water source which could contribute to a river or dam which is used for irrigation (Table 2). Agricultural activity was represented by both a dairy and a piggery, in both cases the samples were taken from the water being used to wash faecal matter out of the pig and cow stalls. The wash water (high in faecal contamination) is then directed to a dam slightly further along which could over-flow and reach a nearby marsh which signals the start of the Plankenburg River. In both cases the subsequent dams were also sampled. Industrial activity was represented by partially treated wine cellar effluent. This water was expected to have a high microbial count as untreated human sewage is added to the effluent and then the water is treated by means of a constructed wetland. This is done with the hopes of decreasing the concentration of unwanted elements (carbon, nitrogen and phosphate). This industrial water source is also used to irrigate fresh produce consumed by the workers on the wine farm.

Table 2. Contamination source sites, their geographical locations and the main contributor of contamination

Water source	Geographic Location	Contamination source
Piggery effluent (2 sites)	Stellenbosch	Porcine
Plankenburg River (Plank-0)	Stellenbosch	Environmental (control for Plankenburg River 1)
Plankenburg River (Plank-1)	Stellenbosch	Human pollution
Veldwagters River	Stellenbosch	Sewage
Winery effluent (Effluent L1)	Stellenbosch (outlying)	Industrial effluent
Cow Farm effluent	Stellenbosch	Bovine

The human factor (any human originating contamination such as sewage, litter or wash water) was represented by two sample sites; one being the effluent from a local sewage treatment facility, and the second represented human pollution as a whole as water was sampled from the Plankenburg River before and after flowing through Kayamandi, a large informal

Stellenbosch University http://scholar.sun.ac.za

settlement in Stellenbosch. The effluent from the sewage treatment facility flows directly into the Veldwagters River, which then converges with the Eerste River. Water from the Veldwagters and Eerste Rivers are then used by nearby farms for irrigation purposes, and as a result the Veldwagters River sample was classified as both a contamination source (Table 2) and irrigation site (Table 1). The winery effluent was also classified under both contamination sources and irrigation water samples.

Sampling Frequency

Samples were collected over a period of five months, from January to May 2012. During this time all sample sites were sampled twice, and it was decided that if the two sample results drastically contradicted one another, that a third sample would be taken. Samples were taken in such a way that follow-up samples from a single site were approximately two months apart. This was to ensure that sufficient time could pass between sampling opportunities so that it was confirmed that the contamination in the water system was constant and not just a once-off contamination. Seasonal fluctuations were expected, but if contamination was only recorded once, it may be that the contamination source was not a constant factor.

Sample Collection

When collecting water samples, the safety of the sampler was also taken into account and stringent safety measures were applied to ensure this. The sampling method changed between different water sources as the nature of the sample site determined how the samples could be taken. The standard methods set out in SANS 5667-6 (SANS, 2006a) and 5667-11 (SANS, 2006b) for rivers and streams and groundwater respectively were followed as closely as possible. For the river water, sample bottles were submerged prior to opening and water was collected approximately 30 cm under the surface of the water with the mouth of the bottle facing upstream. The bottle was then closed before removing the closed bottle from the water.

When sampling from the boreholes, the borehole pump was first left to run for at least a minute to ensure that the sample came directly from the borehole pump, and had not been stored in a reserve tank or dam. The bottle was then opened and placed in the stream of water from the borehole pump until filled. In the case of the spring water, the spring water in question was collected in a dug-out hole with a manhole cover. Here the water was collected in the same way as in the rivers; the bottle was submerged and positioned approximately halfway between the

Stellenbosch University http://scholar.sun.ac.za

surface of the water and the bottom of the collection well. After being submerged, the bottle was opened and allowed to fill up with water and closed while still submerged. Dam water, including piggery overflow and winery effluent, was collected in a similar way, with the exception that water was collected as far into the dam as was safe and possible. Water samples were also collected with great care so as to insure minimal disturbance to the sediment in the dam.

Autoclaved bottles (Schott) were used to collect the 1 L water samples. Bottles were transported to and from the sample sites on ice, and stored at 4°C until analysed. Samples were analysed within six hours of collection to minimise any microbial growth which could result in false enumeration data.

Sample Analysis

Total Coliforms and Escherichia coli counts

Water analysis was done according to the standard methods described by SANS 9308 (SANS, 2012). The QuantiTray system was used to enumerate total coliforms and *E. coli*, using the Colilert 18 kit (IDEXX, SOUTH AFRICA). The dilutions of the samples used with the QuantiTray system varied from source types and potential contamination load. QuantiTrays were incubated at 35°C for 18 h and subsequently examined, and total coliforms and *E. coli* levels determined using a conversion table.

Isolation of E. coli isolates

Wells showing fluorescence on the QuantiTrays were marked, and the area of large wells of each tray was divided into quarters. A maximum of two fluorescent wells were chosen at random from each quarter, and 1 mL of the contents of each chosen well was removed, aseptically, and placed in a sterile McCartney bottle. A maximum sample size of 8 mL was therefore generated for further analysis. A loop-full of this isolated sample was placed in a McCartney bottle with 9 mL sterile saline (0.83% m/v), and vortexed (Centrotec). This was used as the 'concentrated' solution, and from this, a 10^{-2} dilution was prepared using sterile saline (0.83%). This 10^{-2} dilution was then used to prepare spread plates on Eosin Methylene-blue Lactose Sucrose Agar (L-EMB) (Oxoid CM0069, South Africa), and then incubated for 24 h at 35°C.

After incubation, colonies showing a metallic green sheen and denoting typical *E. coli* growth (Merck, 2007), were regarded as a presumptive *E. coli* strain. A minimum of five colonies showing typical *E. coli* growth on the L-EMB agar were isolated using the Harrison Disk Method

Stellenbosch University http://scholar.sun.ac.za

(Harrigan & McCance, 1976). These colonies were then streaked onto Brilliance™ *E. coli* coliform selective medium (Oxoid CM1046, South Africa), and incubated for 24 h at 35°C. Streaking out of each isolate on Brilliance™ *E. coli* coliform selective agar (Oxoid CM1046, England), was repeated, if necessary, until pure cultures were obtained. On Brilliance™ *E. coli* coliform selective agar (Oxoid CM1046, England), *E. coli* typically forms round, smooth, convex colonies which are deep purple in colour. Any atypical (morphologically different) cultures resulting from the purification process were also isolated and carried forward.

Characterisation and confirmation of E. coli identification

Each isolate was then streaked out on Nutrient Agar (NA) (Biolab, South Africa), and the API 20E system (BioMérieux, South Africa) was used in conjunction with Gram stains and motility tests (Murray, 1981) to create a unique 'profile number' for each isolate. This profile number was then entered into the APIweb™ (BioMérieux, South Africa) database and the isolates were identified. After identification and characterisation, isolates were stored at -80°C in the presence of 40% (v/v) glycerol (Fluka Analytical, Germany).

Reference Strains

Reference strains of *E. coli* (Table 3) were added to the dataset of isolates to act as comparative controls when identifying isolates using API 20E (BioMérieux, South Africa), as the properties of the reference strains are well documented. When using the MALDI Biotyper, the reference samples were used to determine whether the instrument was accurate when identifying isolates.

Table 3. List of reference strains [ATCC (American Type Culture Collection)]

Strain	ATCC code
Escherichia coli 58	ATCC 11775
Escherichia coli 157	ATCC 4350
ATCC 35218	ATCC 35218
ATCC 25922	ATCC 25922

Numerical Analysis of Isolate Profiles

The isolate profiles created by API 20E (BioMérieux, France) were converted into a series of ones and zeros denoting positive and negative attributes respectively. Jaccard (SJ) and Sokal &

Stellenbosch University http://scholar.sun.ac.za

Michener (SM) coefficients were used to create dissimilarity matrices. The unsorted matrices were then sorted by means of average linkage cluster analysis (Lockhart & Liston, 1970), and dendrograms were created. Dendrograms resulting from SJ analyses were used to determine variation within the *E. coli* species. Dendrograms were also used to illustrate the similarities and dissimilarities between *E. coli* strains, which allowed for a comparison of isolate behaviour, biochemical attributes and phenotype.

MALDI Biotyper Analysis

Sample preparation

Identification of all isolates were confirmed using the MALDI Biotyper software (Bruker, Germany) to do MALDI BIOTYPER (matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry) based fingerprinting analysis. Single colonies were collected from NA (Biolab, South Africa) after an incubation period of 24 h at 35°C. Colonies were picked up using sterile plastic 1 µL loops (Looplast, South Africa) and placed in 1.5 mL tubes (Eppendorf, Germany) containing a mixture (1:3 ratio) of GC grade water (Fluka Analytical, Germany) and absolute ethanol (Sigma Aldrich, Germany), to make a final volume of 1 200 µL. Samples were then vortexed (Centrotec) and stored at -18°C until needed (Duvenage (nee Collignon), S., 2012, Applications specialist – mass spectrometry, Bruker South Africa, Stellenbosch, South Africa. Personal communication).

Ribosomal protein extraction

Protein extracts were prepared from the microbial samples using the standard formic acid/acetonitrile extraction method developed by Bruker (Duvenage (nee Collignon), S., 2012, Applications specialist – mass spectrometry, Bruker South Africa, Stellenbosch, South Africa. Personal communication). Each protein extract (1 µL) was then spotted directly onto a prepared polished steel MTP 384 target plate with transponder technology (Bruker Daltonik GmbH, Germany) and allowed to air dry. Once dry, the sample extract spots were overlaid with one µL HCCA matrix solution (α-cyano-4-hydroxycinnamic acid) from Bruker, Germany. HCCA matrix was prepared by dissolving contents according to manufacturer's instructions (Bruker, Germany).

The instrument was calibrated each time before biotyping, using Bruker's bacterial test standard (BTS). BTS is an extract made of ribosomal proteins from *E. coli* DH5 alpha which is then re-suspended in 50 µL organic solvent, and stored as 5 µL aliquots at -18°C. BTS was spotted in the

same way as the ribosomal protein extract of the samples, and overlaid with HCCA matrix. The calibration acts as a positive control calibration when identifying unknown bacterial strains.

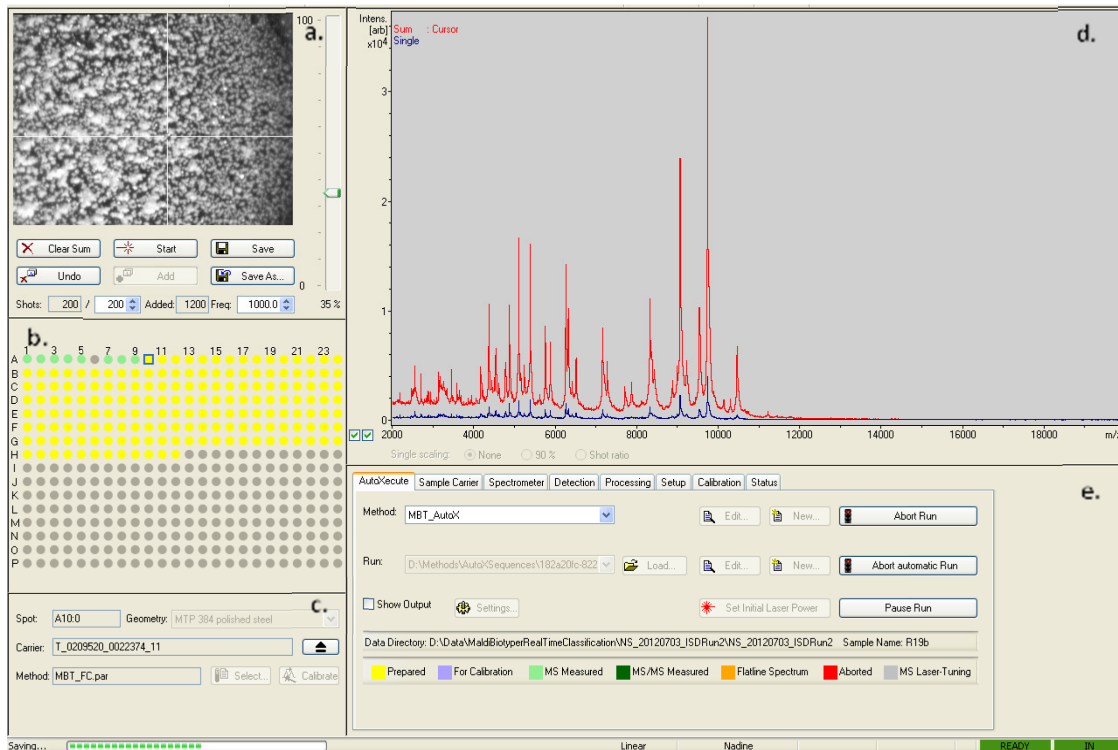

Figure 1. Illustration of MALDI-TOF flex control software while identifying an isolate

a. Live view of sample on target plate while being shot with the laser
b. Visual representation of progress bar, this allows for an accurate determination of which sample spot the instrument is busy scanning
c. Details pertaining to the sample position being scanned as well as the target plate and laser method being used
d. Live view of the spectra created while sample is scanned
e. Details pertaining to the analysis method being used for identification

The target plate was then mounted on a target frame #74115 (Bruker Daltonik GmbH, Germany), and loaded into the Bruker UltrafleXtreme MALDI-TOF MS. The instrument was first calibrated by selecting the BTS spot and running it on 'auto calibrate' in the FlexControl software (ver 3.3.108.0). Once the calibration was saved the sample spots were analysed using the MBT_FC.par method (for laser) in conjunction with the pre-programmed sample analysis autoExecute function MBT_autoX.axe to identify the samples. The spectra collected by the FlexControl software was then compared to the spectra of 3 771 bacterial strains in the BDAL database (Bruker, Germany) using the MALDI Biotyper 3.0 software (Bruker, Germany) to give an

Stellenbosch University http://scholar.sun.ac.za

identification. The BDAL database also contains 11 *E. coli* strains' spectra which are used for comparison when identifying unknown strains.

The functioning of the FlexControl software used in conjunction with the Bruker UltrafleXtreme can be seen in Fig. 1. In Fig. 1a, a magnified view of the spot being analysed can be seen, also observe the crystallisation pattern as a result of the HCCA matrix overlay. This (Fig. 1a) is a live magnified view, and the laser shots can be seen as well as manipulated in this window. Fig. 1b is a visual representation of a progress bar, and provides an accurate determination of which position on the target plate is being scanned. Details pertaining to the sample position being scanned (Fig. 1c) can be seen, and methods relating to the laser (Fig. 1c) as well as the analysis method (Fig. 1e) can be set here. The spectrum of each sample spot can be seen while it is being created (Fig. 1d), and resultant spectra are seen after each sample spot has been scanned.

RESULTS AND DISCUSSION

Prevalence of Escherichia coli in natural water sources

As *E. coli* only makes up a portion of total coliforms it was observed, as expected, that coliform counts surpassed *E. coli* counts in all collected samples (Fig. 2). It was observed that a high coliform count is not always indicative of the presence of *E. coli*. This was seen in certain samples (Borehole A2, N2 and P1) where high coliform counts were observed (as high as log 3.3742 MPN.100 mL^{-1}) but no *E. coli* was detected (Fig. 2). It was also found that Borehole D1, which had the lowest total coliform count (log 2.6818 MPN.100 mL^{-1}), showed higher *E. coli* levels (log 1.8129 MPN.100 mL^{-1}) than 37% (7/19) of the sample sites. It should also be noted that there may be some underestimation of *E. coli*, as some *E. coli* strains are MUG negative and a MUG positive reaction is what Colilert 18 uses as a discriminator.

Although there was no correlation between coliform and *E. coli* counts, it is still important to note the variation of coliform and *E. coli* counts from site-to-site, as well as the number of sites (10/19 = 52.6%) showing *E. coli* counts above the recommended *E. coli* guidelines of 1 000 counts.100 mL^{-1} (WHO, 1989; DWAF, 1996). The DWAF guidelines also associate risk with various levels of *E. coli* present in irrigation water utilised for fresh produce. These associated risks say that water containing no *E. coli* is classified as having 'no risk' associated when using the water for irrigation of fresh produce. Water with *E. coli* counts ranging between one and 999 *E. coli* counts.100 mL^{-1} is classified as 'low risk' and water with an *E. coli* level ranging between 1 000 and 3 999 *E. coli* counts.100 mL^{-1} is 'high risk' (DWAF, 1996). It must be mentioned that although in

Stellenbosch University http://scholar.sun.ac.za

Figs 2 and 3 Borehole A1, D1, Spring C1 and Dam D3 do not show any detected *E. coli*, *E. coli* was present, albeit in very low concentrations. These concentrations were not shows in Figs 2 and 3 when using a log scale.

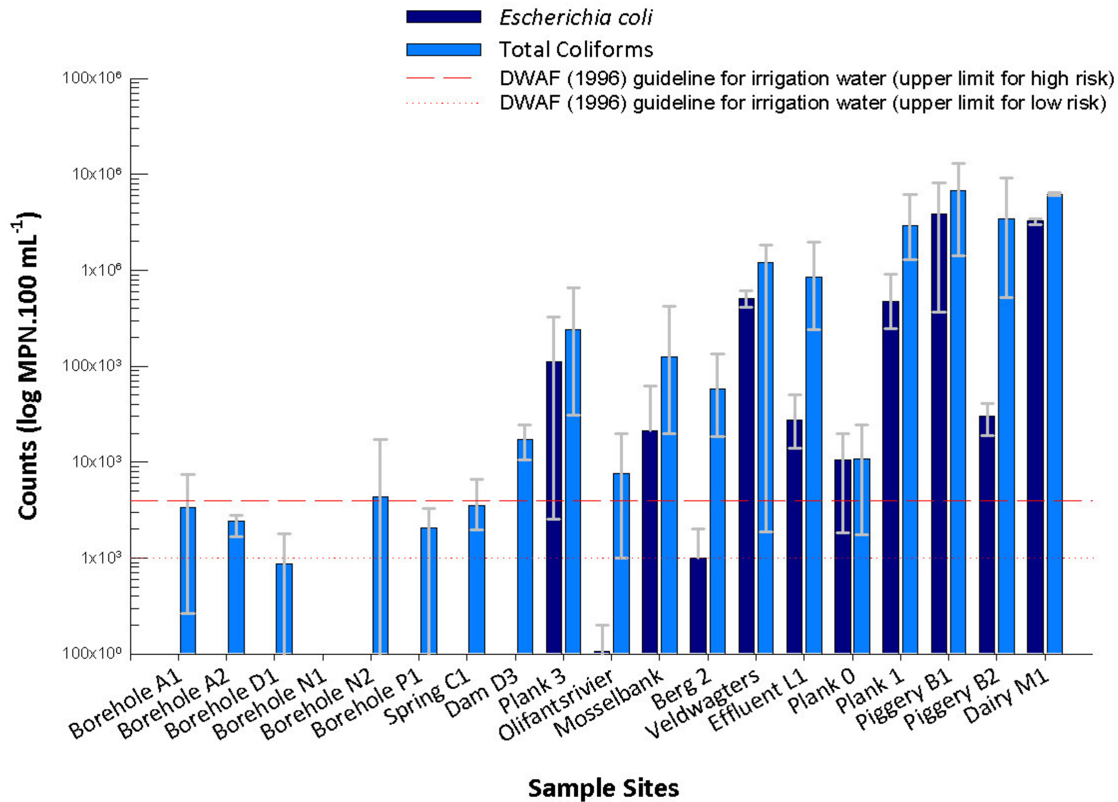

Figure 2. Comparison of mean *E. coli* and total coliform mean counts observed in all sample sites, including minimum and maximum error bars. (Borehole A1, D1, spring C1 and Dam D3 did have *E. coli* counts ranging between log 0.00 and log 2.00) (Upper limit high risk = 4 000 *E. coli* counts.100 mL^{-1}, Upper limit low risk = 1 000 *E. coli* counts.100 mL^{-1})

Irrigation sites

In all irrigation sites sampled, both total coliforms and *E. coli* counts (Fig. 3) were substantially lower than those of the contamination source sites. In this case, total coliforms and *E. coli* counts across all irrigation sites ranged from not detected to log 5.7679 MPN.100 mL^{-1} and not detected to log 5.7875 MPN.100 mL^{-1}, respectively. These ranges were similar to those of surface water being used for irrigation (Fig. 3). When looking at groundwater (boreholes A1, A2, D1, N1, N2, P1 and spring C1) used for irrigation on the other hand, the total coliforms and *E. coli* counts ranged between not detected to log 4.3837 MPN.100 mL^{-1} and not detected to log 2.00 MPN.100 mL^{-1}, respectively (Fig. 3).

Stellenbosch University http://scholar.sun.ac.za

Even though the groundwater *E. coli* counts do not exceed the recommended guidelines for irrigation water for fresh produce stated by DWAF (1996), it is still important to note that there is *E. coli* present in the groundwater, albeit at low concentrations. It is also important to note that some pathogenic *E. coli* strains have a very low infectious dose (Karmali, 1989; Percival *et al*., 2004) and can cause disease, even when only a few bacterial cells are ingested. The data represented by Fig. 2 illustrates the difference between total coliforms and *E. coli* counts when it comes to irrigation water, as well as the lower *E. coli* and total coliforms levels present in groundwaters (boreholes A1, A2, D1, N1, N2, P1 and spring C1), when comparing them to that of surface water (Dam 3, Olifants River, Plank 3, Veldwagters, Berg 2 and Mosselbank Rivers). It can therefore be concluded that groundwater is a safer option when considering the *E. coli* and total coliform counts found in this study.

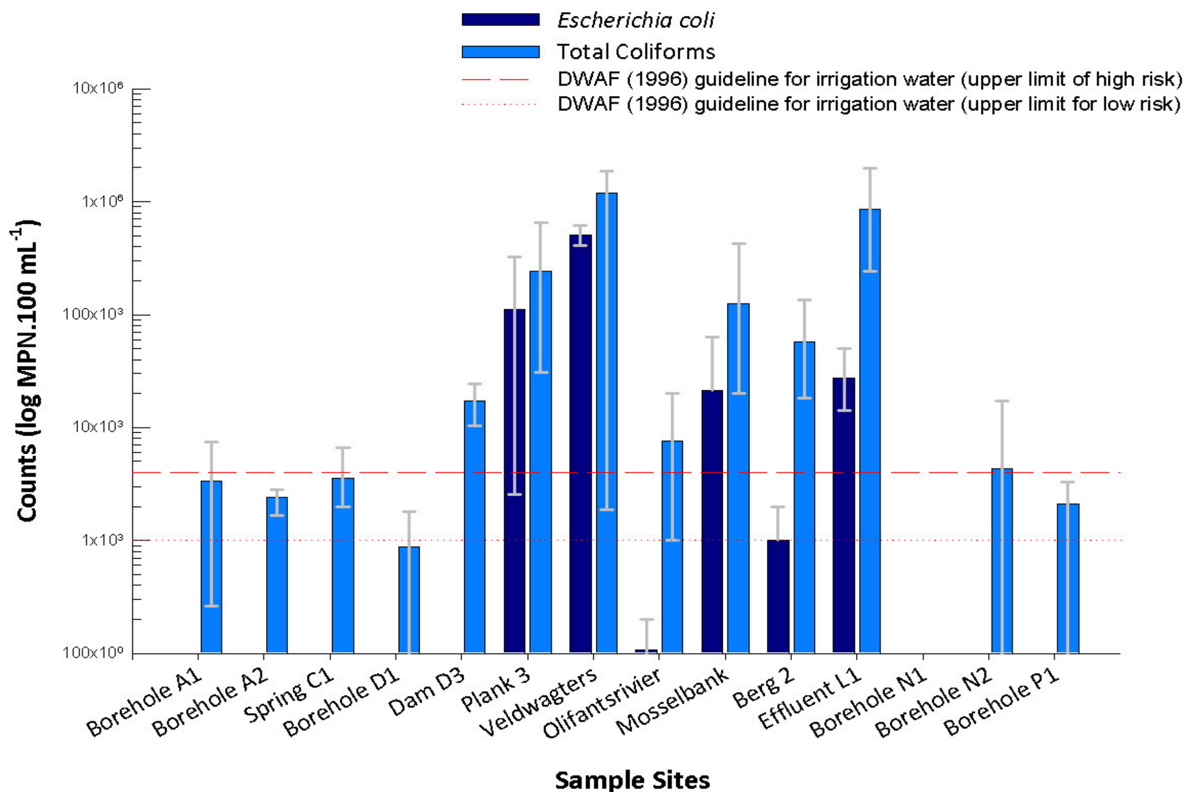

Figure 3. Comparison of *E. coli* and total coliform mean counts observed only at irrigation sites, including minimum and maximum error bars. (Borehole A1, D1, spring C1 and Dam D3 did have *E. coli* counts ranging between log 0.00 and log 2.00)(Upper limit high risk = 4 000 *E. coli* counts.100 mL^{-1}, Upper limit low risk = 1 000 *E. coli* counts.100 mL^{-1})

When comparing the high and low risk guidelines (DWAF, 1996) to the *E. coli* counts in irrigation water, it can be seen that Boreholes A2, N1, N2 and P1 are the only sites which have no

Stellenbosch University http://scholar.sun.ac.za

associated risk as there was no *E. coli* detected. Olifants River, Borehole D1 and A1, Spring C1 and Dam D3 all have mean *E. coli* counts ranging between log 0 and log 3.00 counts.100 mL^{-1}, which means that this water has a low risk associated when being used to irrigate fresh produce, according to the guidelines set by DWAF (1996). Only one irrigation site (Berg 2) had a mean *E. coli* level of greater than log 3, but less than log 3.6021 counts.100 mL^{-1}, showing that it has a high risk (DWAF, 1996) associated when used for irrigation of fresh produce. On the top end of the scale, and exceeding the upper limit for 'high risk' water as stated in the guideline published by DWAF (DWAF, 1996) are Plank-3, Veldwagters and Mosselbank rivers and the winery effluent. This concurs with the finding that groundwater is less contaminated than surface water, and there is less risk associated when using groundwater to irrigate fresh produce than when using surface water.

It is important to note that the spring (C1) also contains *E. coli* contamination, albeit very low (mean *E. coli* count = log 0.5720 *E. coli* counts.100 mL^{-1}). This is important to note, as it is thought that spring water does not contain any contamination of faecal origin. This may also lead to the discovery of definitive 'environmental strains' of *E. coli* which have not been investigated in much detail. The idea of purely environmental strains is also not easy to test for as water sources are so intertwined and cross contamination occurs on a regular basis. Spring water on the other hand cannot be affected by other river water and the contamination which it carries; springs originate at higher altitudes, and flow down towards a collection point. This means that spring water has little or no contact with potential contamination sources, and leachate from upstream sewage treatment facilities or landfills are also unlikely to play a role.

Contamination Sources

In this study contamination sources were considered as *E. coli* reservoirs, and as such it was expected that *E. coli* counts recorded in these water samples would all exceed the recommended guidelines for irrigation water used for fresh produce (WHO, 1989; DWAF, 1996). Once the results confirmed this (Fig. 4), the ranges of recorded *E. coli* and total coliform counts were looked at in more detail. The total coliform and *E. coli* counts ranged between log 5.5126 and log 7.1138 MPN.100 mL^{-1} and log 4.1503 and log 6.9119 MPN.100 mL^{-1}, respectively. When mean values were plotted (Fig. 4), it was observed that all total coliform and *E. coli* counts exceeded the recommended guidelines set by DWAF and WHO (WHO, 1989; DWAF, 1996). This therefore confirmed that runoff from cattle and pig farms, industrial effluent (with added sewage) such as

Stellenbosch University http://scholar.sun.ac.za

that coming from wine farms, and effluent coming from a sewage treatment plant could carry high loads of coliform bacteria as well as *E. coli.*

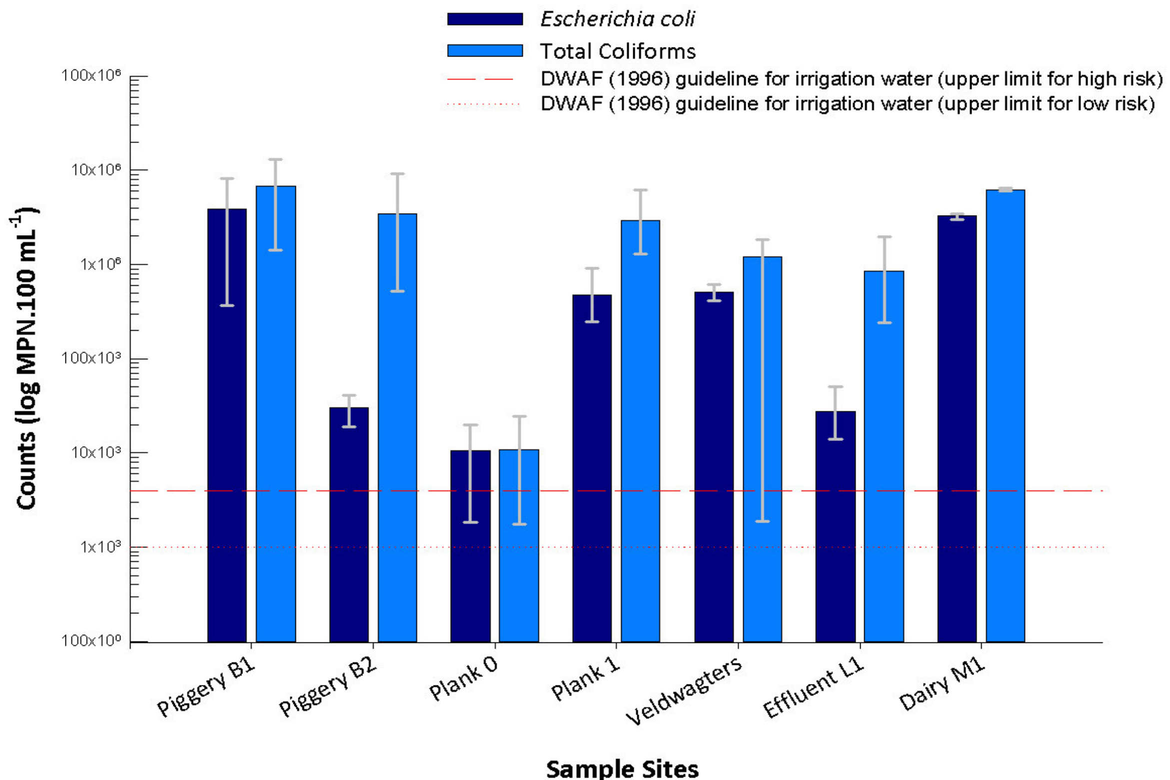

Figure 4. Comparison of *E. coli* and total coliform mean counts observed only at contamination source sites, including minimum and maximum error bars. (Upper limit high risk = 4 000 *E. coli* counts.100 mL^{-1}, Upper limit low risk = 1 000 *E. coli* counts.100 mL^{-1})

This is of foremost importance as, although water from the contamination sources (except Veldwagters River and the winery effluent) are not directly used for irrigation, the water from these contamination sources do all eventually merge with, or cross paths with a river or dam which is used for irrigation. It can also be deduced that the higher the microbial counts present in water, the higher the chances of the water still containing microbial contamination when it reaches groundwater catchment areas. Along with this, sampling of contamination sources and characterisation of the *E. coli* present will possibly allow for microbial source tracking (MST) in the future.

By looking at the Plankenburg River site before and after an informal settlement, it can also be concluded that human pollution plays a significant role in introducing microbial contaminants into water sources. The Plankenburg River was sampled both before (Plank 0) and after (Plank 1)

Stellenbosch University http://scholar.sun.ac.za

the informal settlement, Kayamandi, and it was found that there was a substantial increase in both total coliform (log 3.7758 MPN.100 mL^{-1} to log 6.3400 MPN.100 mL^{-1}), and *E. coli* (log 3.0858 MPN.100 mL^{-1} to log 5.5891 MPN.100 mL^{-1}) mean levels. Where water was collected (Plank 0) five km before the informal settlement it has not yet passed through any informal settlements or industrial areas at this point. While the test site (Plank 1), which is immediately after the informal settlement showed an increase in both total coliforms and *E. coli* counts (Fig. 3). This gives a possible indication of the effect of an informal settlement on water quality, and it can be deduced that this increase can be classified as 'the human factor'.

The winery waste water (effluent L1) first undergoes a step where human sewage is added to the waste water before treatment in the constructed wetland. This is done with the hopes of the microorganisms present being able to breakdown the unwanted compounds (carbon, nitrogen and phosphate) in the water before it is re-introduced into a river system. This water is however also used to water a subsistence vegetable garden, and the vegetables are consumed by the workers on the farm. The mean level of *E. coli* in this water (log 4.4403 MPN.100 mL^{-1}) far exceeds *E. coli* guidelines (WHO, 1989; DWAF, 1996) for irrigation water and is in fact classified as 'high risk' water. This also increases the chance of carry-over of bacteria to produce occurring if the water is used for irrigation (Van Blommestein, 2012) or carry-over to other water sources. If *E. coli* strains found in this water are human pathogens, using this water would be very dangerous and there is a higher risk of foodborne disease outbreaks.

Isolate Identification

Both API 20E and the MALDI Biotyper system were used to identify the isolated, presumptive *E. coli* strains. It was however found that *Klebsiella*, *Citrobacter* and *Enterobacter* species, in addition to *E. coli*, were also identified amongst the isolated strains originating from both contamination sources and irrigation sites (Appendix A and B). The occurrence of *Klebsiella* and *Enterobacter* strains, from both contamination sources and irrigation sites, can be explained by the inclusion of L-EMB agar as a growth medium to isolate *E. coli*. The reasoning behind this is that although typical growth of *E. coli* on L-EMB agar is a colony showing a metallic green sheen, it is known that both *Klebsiella* and *Enterobacter cloacae* may also show this characteristic (Merck, 2007). *Citrobacter*, on the other hand, has a biochemical profile almost identical to *E. coli* (Farmer *et al*., 1985) and can therefore overcome most of the isolation hurdles in the same ways that *E. coli* can. Additionally API 20E, which identifies a strain according to its biochemical profile, is unable to

Stellenbosch University http://scholar.sun.ac.za

detect a clear difference between *E. coli* and *Citrobacter freundii*. Other studies done on natural water sources in the Stellenbosch and surrounding areas have also found high levels of *Citrobacter* in the water (Lötter, 2010; Van Blommestein, 2012). This could mean that *Citrobacter* spp. have been isolated due to their prevalence or because they possibly out-competed some of the *E. coli* strains. However, as *Citrobacter* spp. was not particularly tested for, no concrete conclusions can be made at this time.

Analytical Profile Index (API) 20E

The API 20E system uses 27 biochemical tests to characterise the microorganism and the APIweb™ program is then used to compare results and give an identification based on the biochemical profiles of the isolates. This method of testing relies a great deal on the individual conducting the testing as well as the number of comparative isolates in the database used for identification. The results from the API test strips are also not always a clear positive or negative result as some colour variation may occur in the API test wells which are not always easily distinguishable as positive or negative reactions. This also means that only presence/absence is evaluated and no degree of presence is measured. The API database that is used to compare identification profiles can also be a limiting factor as environmental and clinical *E. coli* isolates can be biochemically different, while still being part of the same species (Lan & Reeves, 2000). The database used by API can only distinguish between two *E. coli* 'types', namely biochemical strains one and two (Van Den Munckhof, J., 2012, Industrial sales consultant – coastal region, BioMérieux South Africa (Pty) Ltd. Stellenbosch, South Africa. Personal communication). It was however observed that a variety of biochemical profiles were generated and that not only two main *E. coli* types were present in natural water systems (Fig. 6). This means that the isolates gathered from the irrigation sites and contamination sources were less likely to be identified accurately by the API 20E system, as more than two *E. coli* types exist.

As a means of showing the likelihood of the identification given by API being correct, APIweb gives a 'percentage certainty' along with each isolate identification (Appendix A). This percentage given with each isolate identification shows the probability of the identification being correct. Along with this percentage, an additional identification rating is also given which can either be; excellent identification, very good identification, good identification, acceptable identification, acceptable identification to genus level, unacceptable profile, doubtful profile and

Stellenbosch University http://scholar.sun.ac.za

low discrimination. All these ratings, together with the percentage certainty can be used to determine accuracy of the identification (Appendix A).

When looking particularly at the percentage certainty and identification ratings of the isolates, it was noted that a large portion (51/154 = 33.11%) of the *E. coli* identifications made using the API system showed a very high percentage certainty (>99%), while still being classified as having 'doubtful profiles'. This means that all positive and negative attributes separately fall within the model of a typical *E. coli* biochemical profile; however the combination of positive and negative attributes is uncommon. In other words, the identification profiles generated in these cases are unusual. This could be as a result of *E. coli* acquiring genes from other microorganisms also found in the environment. These acquired genes could then lead to a change in the biochemical profile of the *E. coli* in question, resulting in a 99.9% certainty in conjunction with a 'doubtful profile' rating.

As a result of these unusual combinations of attributes, a review of which attributes were queried by the API system was conducted. This was done to determine whether a single advantageous characteristic could be singled out which may be advantageous to the organism under certain environmental conditions, and therefore representing a possible acclimatisation. In this study it was found that the test result which was most often queried by the APIweb software was the positive result of the Voges-Proskauer (VP) test. Of the 145 isolates identified by API 20E, 48% (70 of 145) of the isolates showed a positive result for VP. When comparing this percentage of positive results to those given by API 20E, it was found that in the API database, 0% of their samples showed a positive result for VP.

When investigated further, it was found that a strain which is VP positive is able to produce the metabolite acetoin (3-hydroxy-2-butanone) when the Embden-Meyerhof (EM) pathway is used to degrade glucose and other fermentable carbon sources (Levine, 1916; Lopez *et al*., 1975; Huang *et al*., 1999). Formation of acetoin is often used as a microbial classification marker, and on a physiological level helps the microorganisms avoid acidification, as well as playing a role in carbon storage and constant regulation of the NAD/NADH ratio within the cell (Xiao & Xu, 2007).

Acetoin may also help in carbon storage, as acetoin can be re-utilised by microorganisms to a carbon source used during stationary growth phase (Grundy *et al*., 1993). If glucose is depleted in the surrounding environment, acetoin will therefore be converted and utilised. This means that the ability to produce and reutilise acetoin may help *E. coli* in environmental niches which are forever changing and not always rich in a usable carbon source (Johansen *et al*., 1975; Mayer *et al*., 1995). Acetoin biosynthesis is also seen as an energy-saving pathway, which would also be

Stellenbosch University http://scholar.sun.ac.za

advantageous to microorganisms living in a nutrient variable environment (Johansen *et al*., 1975; Mayer *et al*., 1995). It can therefore be speculated that *E. coli* testing positive for VP may be better equipped to survive in environments with a fluctuating carbon source and concentration, such as natural water systems. Even though it is known that *E. coli* is a very genetically fluid group of bacteria (LeClerc *et al*., 1996; Matic *et al*., 1997) which easily acquires genes from other bacteria, this data shows that there is directionality in their gene acquisition. This means that *E. coli* can acquire and preserve genes which give it a survival advantage above other bacteria in their immediate environment.

MALDI Biotyper Analysis

MALDI Biotyper uses the presence and concentration of the five most abundant ribosomal proteins in the sample to create a spectrum which is used for identification purposes (Duvenage (nee Collignon), S., 2012, Applications specialist – mass spectrometry, Bruker South Africa, Stellenbosch, South Africa. Personal communication). Due to the identification being conducted mostly by an automated system, there is less chance of the human factor becoming a problem in reading and interpreting results. This also means that the measurements from one sample to another are much more controlled and reliable, and consequently the results observed are also more consistent. Due to the way in which MALDI Biotyper identifies strains using presence/absence as well as concentration, the data can also be used more reliably for statistical purposes when analysed (Duvenage (nee Collignon), S., 2012, Applications specialist – mass spectrometry, Bruker South Africa, Stellenbosch, South Africa. Personal communication).

In a similar way to API 20E's 'percentage certainty', MALDI Biotyper identifications are ranked by the Biotyper 3.0 software (Bruker) by using a scoring system between 0 and 3.00. This scoring system specifies that any identification which has a score of 2.300-3.000 shows highly probable species identification. Below this, 2.000-2.299 shows secure genus identification, probable species identification; 1.700-1.999 denotes probable genus identification and any score below 1.700 cannot be seen as a reliable identification. This scoring system therefore allows for a degree of certainty when identifying isolates (Appendix B).

Comparison of API 20E and MALDI Biotyper as an identification system

In total, 154 bacterial strains were isolated as discussed in the materials and methods. After identification of isolates using the API 20E system, and confirmation of identification by MALDI

Stellenbosch University http://scholar.sun.ac.za

Biotyper, it was found that 148 of the isolated strains were positively identified as *Escherichia coli*. When using only the API 20E system to identify the isolates, a much larger percentage (22/154 = 14.29%) of the isolates could either not be positively identified or were identified as bacterial cultures other than *E. coli*. When compared with the results obtained using MALDI Biotyper, where only 5.84% (9/154) of the isolates were confirmed not to be *E. coli*. Three of the non-*E. coli* strains were identified as such by both API 20E and MALDI Biotyper.

It can therefore be concluded that using the MALDI Biotyper system is much better suited than using biochemical tests, which allow for a certain degree of subjectivity. The data shown in Table 3 shows all the isolates (19/154) which were either unidentifiable, or identified as a bacterium other than *E. coli* by the API system, and compared to the identification given by APIweb to that of Biotyper 3.0. It can be seen that all isolates which could not be identified by API could be identified using MALDI Biotyper and Biotyper 3.0. It is also important to note that all the identifications conducted by Biotyper 3.0 showed a certainty score above 2.299, denoting a positive identification at species level.

Stellenbosch University http://scholar.sun.ac.za

Table 3. Identification and score variations between non-*E. coli* strains obtained using the API 20E system and the MALDI Biotyper

Sample	API 20E		MALDI-TOF-MS	
	Identification	% certainty	Identification	Score*
A11.1	*Klebsiella pneumoniae*	99.5	*Escherichia coli*	2.354
A11.2	*Serratia marcescens*	87.1	*Escherichia coli*	2.465
A11.3	*Escherichia fergusonii*	-	*Escherichia coli*	2.339
A22.2	*Serratia odorifera*	99.1	*Escherichia coli*	2.323
B12.1	No positive identification	-	*Paenibacillus amylolyticus*	2.550
C11.1	No positive identification	-	*Citrobacter freundii*	2.386
C11.2	No positive identification	-	*Citrobacter freundii*	2.408
C11.3	*Citrobacter koseri/farmeri*	91.2	*Citrobacter freundii*	2.397
C11.4	*Citrobacter koseri/farmeri*	91.2	*Citrobacter freundii*	2.356
E12.4	*Citrobacter braakii/ Enterobacter intermedius*	59.3/ 34.2	*Escherichia coli*	2.299
G11.2	*Pantoea spp.*	-	*Enterobacter cloacae*	2.337
M21.5	*Serratia odorifera*	82.1	*Escherichia coli*	2.351
L21.1	No positive identification	-	*Escherichia coli*	2.326
L21.2	No positive identification	-	*Escherichia coli*	2.481
L21.3	No positive identification	-	*Escherichia coli*	2.479
L21.4	No positive identification	-	*Escherichia coli*	2.287
L21.5	No positive identification	-	*Escherichia coli*	2.276
H23.5	No positive identification	-	*Escherichia coli*	2.310
F21.2	*Pantoea/Serratia*	-	*Klebsiella oxytoca*	2.325

*2.300-3.000 shows highly probable species identification
2.000-2.299 shows secure genus identification, probable species identification
- no percentage

Stellenbosch University http://scholar.sun.ac.za

Variation within isolates of the *Escherichia coli* species

When assessing the dissimilarity dendrogram created using the Jaccard statistical method (SJ) of all isolates (Fig. 5), the variation within the dataset is evident. There are 58 small clusters in total, each containing one to 30 isolates. This shows that 58 biochemical profiles were generated by API 20E, all of which vary by at least one biochemical test result. By removing isolates which were identified as non-*E. coli* strains (Fig. 5), the dendrogram is simplified and shows less variation (Fig. 6).

Across the entire dataset of *E. coli* isolates found, and positively identified, there was a large intrinsic variation between isolates, especially when looking at their biochemical test results. When looking at the dissimilarity matrix illustrated in Fig. 6, this becomes quite apparent. It can be seen that there are 38 different clusters containing anything from one single isolate to a group of 30 isolates. The organisms contained in each of these groups all have an identical biochemical profile, resulting in a 0% variance within a cluster. By statistically cutting off the groups (at the level illustrated by the dotted line in Fig. 6), the 38 clusters are reduced to just four (ExelStat). Each of these clusters had a maximum variation amounting to 45.43%, while the variation between clusters is at least 54.57% (Ntushelo, N., 2012, Biometry unit, ARC-Infruitec/Nietvoorbij, Stellenbosch, South Africa. Personal communication). This 45.43% can be seen as intrinsic error in this case.

By statistically dividing the clusters into larger groups, namely clusters A to E (Fig. 6), the clusters are easier to work with and grouping of isolates from the same sources can be seen more conclusively. In Table 4 the distribution of the clusters can be seen more clearly, and the first thing that is noted is cluster E which contains only one isolate. This isolate (D24.2) was the only isolate which was identified by API as *E. coli* biochemical type 2. As previously mentioned, APIweb can only distinguish between two main *E. coli* types, namely biochemical strains 1 and 2. All other isolates were identified as biochemical type 1, while D24.2 was identified as type 2. When looking at the identification of this isolate with MALDI Biotyper however, it was identified as *E. coli* and nothing suggested that on a protein level that it was so vastly different to any other strain isolated. The only real difference that could be found when looking at the API identification profile was that isolate D24.2 was unable to ferment rhamnose, sucrose and melibiose. There were other isolates which exhibited some of these properties as well, but none that exhibited all. This is an isolate which originates from a dam in Wellington, and possibly needs to be investigated with molecular methods to determine whether any other differences can be found.

Stellenbosch University http://scholar.sun.ac.za

When considering that all isolates shown in Fig. 6 were positively identified as *E. coli* strains, it is important to note just how much variation still occurs within the *E. coli* species. This reinforces the findings of *E. coli* being a very genetically diverse species which could easily acquire genes from other *E. coli* strains, and even other microbial species (LeClerc *et al.*, 1996; Matic *et al.*, 1997). The dendrogram creates clusters by comparing isolate profiles from API, and at zero dissimilarity level, all isolates in that cluster have the same identification profile.

Subsequently, it is important to state that *E. coli* isolates acquire genes which would be most advantageous to their survival in a particular environmental niche, as even a slight advantage over other microorganisms in the same environment is imperative to survival (Levin & Bull, 1994). Although no conclusive link could be made between particular biochemical attributes and a particular environmental niche, more research comprising a larger isolate dataset as well as more varied sample sites could lead to more conclusive findings. Repeated sampling of one site at an hourly interval might also be advantageous to the study, as this would show variation of the *E. coli* population on a daily basis.

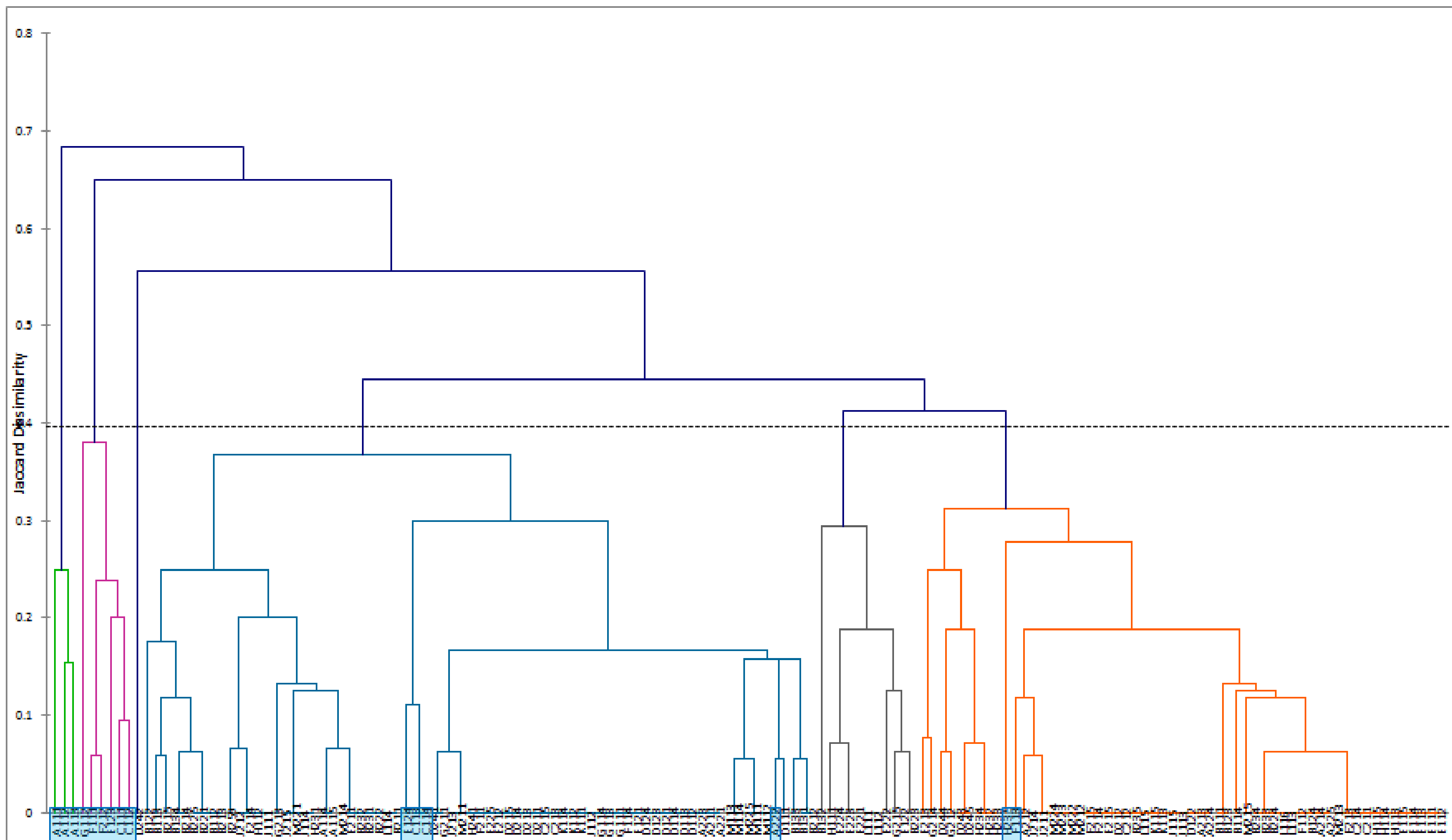

Figure 5. Dissimilarity dendrogram, based on the SJ dissimilarity coefficient, of all isolated cultures, with non-*E. coli*

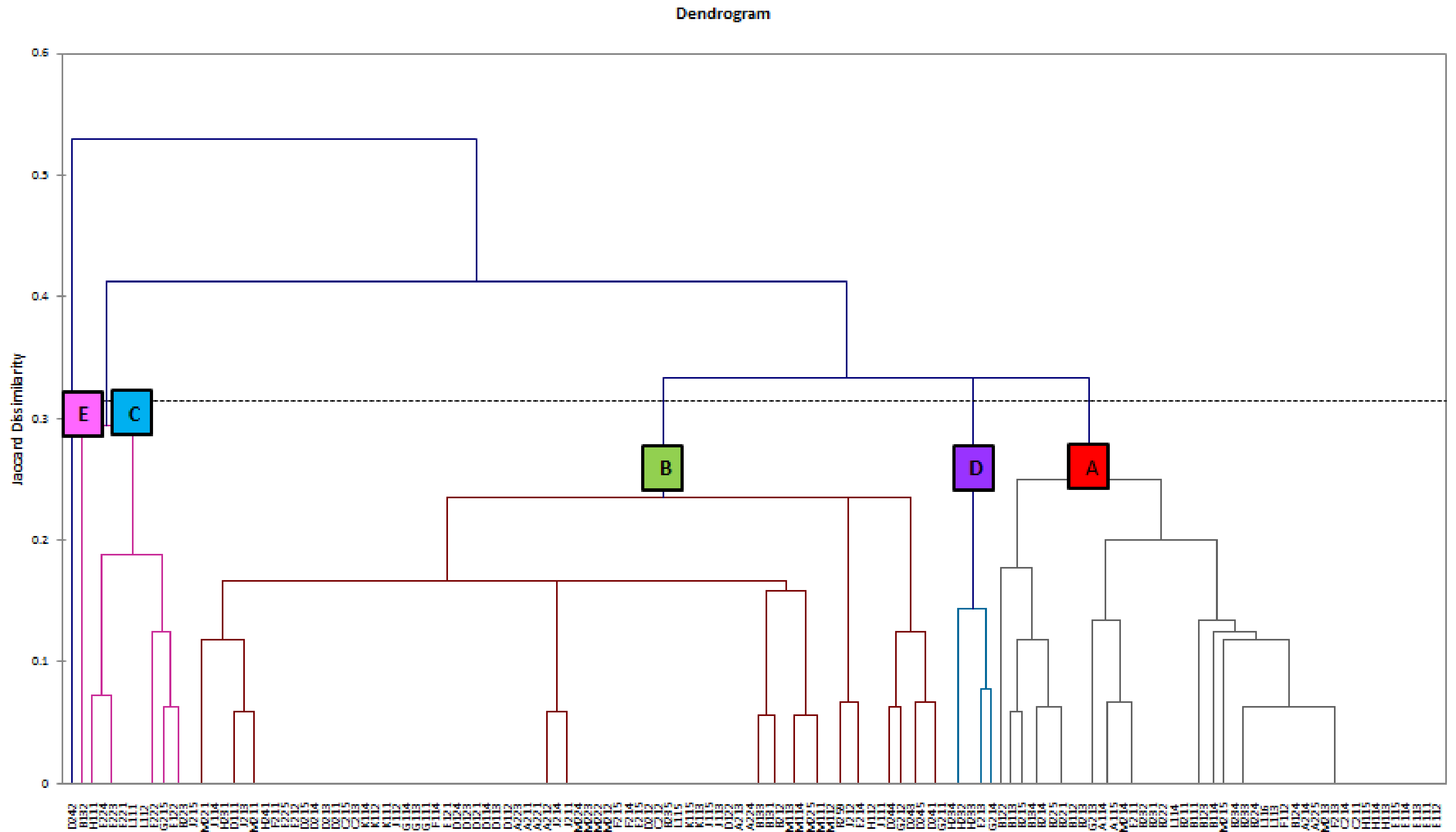

Figure 6. Dissimilarity dendrogram, based on the SJ dissimilarity coefficient, of all isolates confirmed as *E. coli* cultures

Stellenbosch University http://scholar.sun.ac.za

In other words, similarity/dissimilarity matrices use positive and negative attributes to determine the variation between clusters and subsequently showing how closely they are related to one another. By using these methods to determine variation and affiliation between *E. coli* strains, the amount of variation (on a biochemical level) is determined within the species. This information was also used to track potential origins of particular strains, as contamination source strains and irrigation water strains which fall in the same cluster possibly share an origin.

When looking at the data in Table 4, and the isolates present in each major cluster, it is important to note that isolates with the same starting letter come from the same source (Appendix A). The highlighted cells (Table 4) are isolates which have been isolated from contamination sources, while the cells which are surrounded by a border show strains which have been isolated from sample sites which are classified as both contamination sources and irrigation sites. It can be seen that there is a relatively even distribution of strains originating from contamination sources and irrigation sites present in each cluster.

As contamination source irrigation site strains are grouped together, this may facilitate the linking of *E. coli* strains in irrigation water to their most probable source. Cluster B (the largest cluster) contains strains from surface water, groundwater, pigs, cows and human sewerage. This means that cluster B is a more diverse and possibly a more generalised cluster of *E. coli* strains. It can also be possible that the strains in this cluster are environmental strains, as most (90%) of the strains isolated from Plank 0, before the informal settlement and industrial section of Stellenbosch, fall within this cluster. The isolates from the spring (C1) also fell within this cluster, and a large portion (23 of 27) of the borehole water isolates as well. Cluster B also contains majority (11 of 14) of the strains isolated from the dairy. If the conclusion regarding environmental strains can be validated and confirmed, it may also be said that cow faeces play a minor role in contamination of water systems. This is because 78.57% of the strains which are contributed by the bovine faeces are similar to environmental strains which are in the water already. On the other hand, it can also be speculated that the reasoning behind the bovine-originating strains being clustered with possible environmental strains is due to nomadic farmers and small populations living, and raising livestock on the river beds throughout history. This could mean that bovine-originating strains have been introduced into the natural water systems many years ago and have therefore found their way into almost all types of water systems due to the merging and cross contamination between water ways.

Stellenbosch University http://scholar.sun.ac.za

Table 4. *Escherichia coli* strain distribution as determined by SJ dendrogram from API 20E generated data

Cluster A	Cluster B	Cluster C	Cluster D	Cluster E	Cluster B ...continued
A114	A211	A212	D241	D242	F114
A115	A213	A214	D243		F211
B112	A221	A225	D244		F213
B122	A223	B111	D245		F214
B211	A224	B114	E213		F215
B213	B113	B123	G211		G111
B221	B131	B124	G212		G113
B222	B133	B132	G214		G114
B225	B134	B223			G215
B231	B212	B224			H112
B232	B214	B233			H113
E211	B215	B234			H114
G213	B235	E122			H115
H231	C211	E221			H241
H232	C212	E222			J111
H233	C213	E223			J112
H234	C214	E224			J113
J114	C215	F112			J115
J215	D111	H111			J211
L114	D112	L111			J212
M214	D113	L112			J213
M221	D114	L113			J214
	D121	L116			K111
	D122	M215			K112
	D123				K113
	D124				K114
	D211				K115
	D212				L115
	D213				M111
	D214				M112
	D215				M113
	E111				M114
	E112				M211
	E113				M212
	E114				M213
	E115				M222
	E121				M223
	E212				M224
	E214				M225
	E215				ATCC 25922
	E225				

Shaded blocks show contamination sites and bordered blocks show sites which overlap contamination and irrigation sites. The rest were isolated from strict irrigation sites.

Stellenbosch University http://scholar.sun.ac.za

Cluster D contains strains from Veldwagters River, the dam in Wellington and a single strain from Plank-0. As Veldwagters River includes water effluent of a sewage treatment facility, it could be speculated that the *E. coli* strains present are human in origin. This however, is only a theory and needs to be investigated further. This means that there is human sewerage finding its way into the dam in question. This could be very dangerous as faecal contamination may contain pathogenic *E. coli* strains as well as other pathogenic microorganisms. Clusters A and C contain contaminants from a variety of sources (piggery, dairy, sewerage, human pollution and industrial waste). Along with the strains from contamination sources, there are also isolates present which have been collected from both surface and groundwater. It can therefore be speculated that contaminants from a variety of sources play a role in contamination of both surface and groundwater. This is very important to note, as groundwater is often used without prior testing, due to the misconception that groundwater is clean. It has however been proven in this study that this is not always the case and that water coming from any naturally occurring water source should be approached with caution.

Conclusions and Recommendations

When investigating the prevalence of *E. coli* at the sample sites, more than half (10/19) of the sample sites had *E. coli* levels exceeding the WHO and DWAF guidelines for water being utilised for the irrigation of fresh produce to be consumed raw or minimally processed (WHO, 1989; DWAF, 1996). When looking just at the irrigation sites, 71% (5/7) of the surface waters sampled were deemed as 'unsafe' for irrigation purposes, while all groundwater tested contained less than log 3 MPN.100 mL^{-1} *E. coli*.

This shows that groundwater is a safer option to use as an alternative irrigation source. Some groundwater samples however were still classified as 'low risk' waters when being used for irrigation of fresh produce (DWAF, 1996). Even though *E. coli* was only detected at low concentrations in the samples, it must be remembered that some pathogenic *E. coli* have very low infectious doses and could therefore still cause disease, even when only a few bacterial cells are present. Total coliform bacteria was seen to be as high has log 3.6383 MPN.100 mL^{-1} in groundwater which could be detrimental to the health and safety of fresh produce consumers, as coliform bacteria is seen as an indicator for other potential pathogens (CAC, 1978). It is also important to note that only seven groundwater sites were investigated in this study, and as a result this data will need to be verified using a broader outlook focussed on groundwater.

Stellenbosch University http://scholar.sun.ac.za

In terms of prevalence of *E. coli* and total coliform bacteria in water used for irrigation, it can be concluded from this study that there is definitely a potential problem for farmers using natural water sources to irrigate fresh produce. It is however uncertain as to the extent of the potential problem, as well as where the responsibility lies with regards to water screening and treatment.

Comparing the API 20E system to the use of MALDI Biotyper, it can be concluded that MALDI Biotyper is a better suited and more consistent method to be used for isolate identification. The reason for this is due to the measurement of a concentration as well as a presence/absence of the five most abundant proteins. The increased use of an automated process as well as computer-driven analysis also guarantees objectivity and removes a great deal of human error which is prominent in API 20E. Although direct comparison of API 20E and MALDI Biotyper is difficult due to the varied focus and principles that the identification methods are based on, it is still valid to say that because they are both used for identification purposes, their abilities can be paralleled. This is clearly seen in the examples given where a number of isolates could not be identified by API 20E, but were identifiable by MALDI Biotyper, to a species level. Thus confirming the value of MALDI Biotyper as a bacterial identification method.

Variation amongst the bacterial strains collected, within the *E. coli* species alone, can be seen as a confirmation of the genetic fluidity of the bacterium as well as a potential factor of concern. This large variation alludes to the ability of the bacteria to acclimatise to environmental niches, making them particularly dangerous if they are pathogenic. Acclimatisation can include genes which account for pathogenicity, antibiotic resistance, toxin production or any other gene which allows the bacterium to survive better in a particular environment. It can therefore be concluded that the variation of *E. coli* strains present in the environment is a matter of concern and should be investigated further.

In order to confirm the results found in this study, as well as to facilitate more in depth data analysis opportunities, a larger study with more sample sites and more sample replicates would be beneficial. Samples coming from easily definable environmental niches could also potentially help in characterising strains and linking particular biochemical tests to environmental niches where they could provide an advantage to the bacterium.

Stellenbosch University http://scholar.sun.ac.za

REFERENCES

Ackermann, A. (2010). *Assessment of Microbial Loads of the Plankenburg and Berg Rivers and the Survival of Escherichia coli on Raw Vegetables under Laboratory Conditions*. MSc in Food Science Thesis. Stellenbosch University, Stellenbosch, South Africa.

Adams, S., Titus, R., Pietersen, K., Tredoux, G. & Harris, C. (2001). Hydrochemical characteristics of aquifers near Sutherland in the Western Karoo, South Africa. *Journal of Hydrology*, **241**, 91–103.

Avery, L.M., Williams, A.P., Killham, K. & Jones, D.L. (2008). Survival of *Escherichia coli* O157:H7 in waters from lakes, rivers, puddles and animal-drinking troughs. *Science of the Total Environment*, **389**, 378-385.

Beuchat, L.R. & Ryu, J.H. (1997). Produce handling and processing practices. *Emerging Infectious Diseases*, **3**(4), 459-465.

Beuchat, L. R. (2002). Ecological factors influencing survival and growth of human pathogens on raw fruits and vegetables. *Microbes and Infection*, **4**, 413–423.

Bezuidenhout, C.C., Mthembu, N. & Lin, J. (2002). Microbiological evaluation of the Mhlathuze River, KwaZulu-Natal (RSA). *Water SA*, **28**(3), 281-286.

Bezuidenhout, C.C & The North-West University Team (2011). A scoping study on the environmental water (groundwater and surface water) quality and management in the North-West Province, South Africa. Water Research Commission (WRC) Report No. KV 278/11.

CAC (California Administrative Code). (1978). Wastewater reclamation criteria. *California Admistrative Code*, title 22, division 4. California Department of Health Services, Berkeley.

Charimba, G., Hugo, C. & Hugo, A. (2012). The incidence of diarrhoeagenic *Escherichia coli* in minced beef and boerewors. *Food Research International*, **47**, 353-358.

De Villiers, S. (2007). The deteriorating nutrient status of the Berg River, South Africa. *Water SA*, **33**(5), 659-664.

DWAF (Department of Water Affairs and Forestry). (1996). Agricultural Use: Irrigation. In: *South African Water Quality Guidelines. Volume 4.* (2nd Ed.). (Edited by Holmes, S. CSIR Environmental Services). Pretoria: Department of Water Affairs and Forestry.

Fahs, F., Mittelhammer, R.C. & McCluskey, J.J. (2009). *Escherichia coli* outbreaks affect demand for salad vegetables. *Choices*, **24**(2), 26-29.

Farmer, J. J., Davis, B. R., Hickman-Brenner, F. W., McWhorter, A., Huntley-Carter, G. P., Asbury, M. A., Riddle, C., Wathen-Grady, H. G., Elias, C., Fanning, G. R., Steigerwalt, A. G., O'Hara, C. M., Morris, G. K., Smith, P. B. & Brenner, D.J. (1985). Biochemical identification of new species and biogroups of *Enterobacteriaceae* isolated from clinical specimens. *Journal of Clinical Microbiology,* **21**, 46-76.

Franz, E., Semenov, A.V. & Van Bruggen, A.H.C. (2008). Modelling the contamination of lettuce with *Escherichia coli* O157:H7 from manure-amended soil and the effect of intervention strategies. *Journal of Applied Microbiology*, **105**, 1569-1584.

Gemmell, M.E. & Schmidt, S. (2012). Microbial assessment of river water used for the irrigation of fresh produce in a sub-urban community in Sobantu, South Africa. *Food Research International*, **47**, 300-305.

Guenther, P.M., Dodd, K.W., Reedy, J. & Krebs-Smith, S.M. (2006). Most Americans eat much less than recommended amounts of fruits and vegetables. *Journal of the American Dietetic Association*, **106**(9), 1371-1379.

Grundy, F.J., Waters, D.A., Takova, T.Y. & Henkin, T.M. (1993). Identification of genes involved in utilization of acetate and acetoin in *Bacillus subtilis*. *Molecular Microbiology*, **10**, 259–271.

Stellenbosch University http://scholar.sun.ac.za

Haramoto, E., Yamada, K. & Nishida, K. (2011). Prevalence of protozoa, viruses, coliphages and indicator bacteria in groundwater and river water in the Katmandu Valley, Nepal. *Transactional of the Royal Society of Tropical Medicine and Hygiene*, **105**, 711-716.

Harrigan, W.F. & McCance, M.E. (1976). Methods for the selection and examination of microbial colonies. In: *Laboratory Methods in Food and Dairy Microbiology*, (Edited by Harrigan, W.F., & McCance, M.E.). Pp. 47-49. London: Academic Press.

Huang, M., Oppermann-Sanio, F.B. & Steinbüchel, A. (1999). Biochemical and molecular characterization of the *Bacillus subtilis* acetoin catabolic pathway. *Journal of Bacteriology*, **181**, 3837–3841.

Islam, M., Doyle, M.P., Phatak, S.C., Millner, P. & Jiang, X. (2004). Persistence of enterohemoragic *Escherichia coli* O157:H7 in soil and on leaf lettuce and parsley grown in fields treated with contaminated manure composts or irrigation water. *Journal of Food Protection*, **67**(7), 1365-1370.

Jain, C.K., Bandyopadhyay, A. & Bhadra, A. (2009). Assessment of ground water quality for drinking purpose, District Nainital, Uttarakhand, India. *Environment Monitoring and Assessment*, **166**, 663-676.

Johansen, L., Bryn, K. & Stormer, F.C. (1975). Physiological and biochemical role of the butanediol pathway in *Aerobacter* (*Enterobacter*) *aerogenes*. *Journal of Bacteriology*, **123**, 1124–1130.

Karmali, M.A. (1989). Infection by verocytotoxin-producing *Escherichia coli*. *Clinical Microbiology Review*, **2**(1), 15-38.

Lan, R. & Reeves, P.R. (2000). Intraspecies variation in bacterial genomes: the need for a species genome concept. *Trends in Microbiology*, **8**, 396-401.

LeClerc, J.E., Li, B., Payne, W.L. & Cebula, T.A. (1996). High mutation frequencies among *Escherichia coli* and *Salmonella* pathogens. *Science*, **274**(5290), 1208-1211.

Levin, B.R. & Bull, J.J. (1994). Short-sighted evolution and the virulence of pathogenic microorganisms. *Trends in Microbiology,* **2**, 76–81.

Linscott, A.J. (2011). Food-borne illnesses. C*linical Microbiology Newsletter*, **33**, 41-45.

Levine, M. (1916). On the significance of the Voges-Proskauer reaction. *Journal of Bacteriology*, **1**, 153–164.

Lockhart, W.R. & Liston, J. (1970). *Methods for Numerical Taxonomy*. American Society for Microbiology, Washington DC, USA.

Lopez, J.M., Thoms, B. & Rehbein, H. (1975). Acetoin degradation in *Bacillus subtilis* by direct oxidative cleavage. *European Journal of Biochemistry,* **57**, 425–430.

Lopez-Suacedo, C., Cerna, J.F., Villegas-Sepulveda, N., Thompson, R., Velazquez, F.R., Torres, J., Tarr, P.I. & Estrada-Garcia, T. (2003). Single multiplex polymerase chain reaction to detect diverse loci associated with diarrheagenic *Escherichia coli*. *Emerging Infectious Diseases*, **9**(1), 127-131.

Lötter, M. (2010). *Assessment of Microbial Loads Present in Two Western Cape Rivers Used for Irrigation of Vegetables***.** MSc in Food Science Thesis. Stellenbosch University, Stellenbosch, South Africa.

Matic, I., Radman, M., Taddei, F., Picard, B., Doit, C., Bingen, E., Denamur, E., Elison, J., LeClerc, J.E. & Cebula, T.A. (1997). Highly variable mutation rates in commensal and pathogenic *Escherichia coli*. *Science*, **277**(5333), 1833-1834.

Mayer, D., Schlensog, V. & Bock, A. (1995). Identification of the transcriptional activator controlling the butanediol fermentation pathway in *Klebsiella terrigena*. *Journal of Bacteriology*, **177**, 5261–5269.

Merck (2007). EMB Agar (Eosin methylene-blue lactose sucrose agar). In: *Merck Microbiology Manual*, Pp. 276-277. Darmstadt, Germany.

Stellenbosch University http://scholar.sun.ac.za

Moses, A.E., Garbati, M.A., Egwu, G.O. & Ameh, J.A. (2006). Detection of *E. coli* O157 and O26 serogroup in human immunodeficiency virus – infected patients with clinical manifestations of diarrhea in Maiduguri, Nigeria. *Research Journal of Medicine and Medical Science*, **1**(4), 140-145.

Müller, E.E., Ehlers, M.M. & Grabow, W.O.K. (2001). The occurrence of *E. coli* O157:H7 in South African water sources intended for direct and indirect human consumption. *Water Research*, **35**(13), 3085-3088.

Murray, R.G.E. (Editor) (1981). 3.3.4 Characterization: Motility. In: *Manual of Methods for General Bacteriology* (Editor in chief: Gerhardt, P.). Pp. 25-26. Washington DC: American Society for Microbiology.

Okafo, C.N., Umoh, V.J. & Galadima, M. (2003). Occurrence of pathogens on vegetables harvested from soils irrigated with contaminated streams. *The Science of the Total Environment*, **311**, 49-56.

Omar, K.B. & Barnard, T.G. (2010). The occurrence of pathogenic *Escherichia coli* in South African wastewater treatment plants as detected by multiplex PCR. *Water SA*, **36**(2), 172-176.

Pass, M.A., Odedra, R. & Batt, R.M. (2000). Multiplex PCR for identification of *Escherichia coli* virulence genes. *Journal of Clinical Microbiology*, **38**, 2001–2004.

Percival, S., Chalmers, R., Embrey, M., Hunter, P., Sellwood, J. & Wyn-Jones, P. (Editors) (2004). *Escherichia coli.* In: *Microbiology of Waterborne Diseases: Microbiological Aspects and Risks*. New York: Academic Press.

Rai, P.K. & Tripathi, B.D. (2007). Microbial contamination in vegetables due to irrigation with partially treated municipal wastewater in a tropical city. *International Journal of Environmental Health Research*, **17**(5), 389-395.

Solomon, E.B., Yaron, S. & Matthews, K.R. (2002). Transmission of *Escherichia coli* O157:H7 from contaminated manure and irrigation water to lettuce plant tissue and its subsequent internalization. *Applied and Environmental Microbiology,* **68**(1), 397-400.

SAICE (South African Institute for Civil Engineering). (2006). *The SAICE Infrastructure Report Card for South Africa: 2006*. Pp. 1-16. Midrand: SAICE House.

SANS ISO 5667-6 (2006a). Water quality – Sampling – Part 6: Guidance on sampling of rivers and streams. Published by Standards South Africa, Pretoria.

SANS ISO 5667-11 (2006b). Water quality – Sampling – Part 11: Guidance on sampling of groundwaters. Published by Standards South Africa, Pretoria.

SANS ISO 9308 (2012). Microbial analysis of water – General test methods. Total coliforms and *Escherichia coli* in water: Defined substrate technology (Colilert) method. Published by Standards South Africa, Pretoria.

Tauxe, R.V. (1997). Emerging foodborne diseases: An evolving public health challenge. *Emerging Infectious Diseases*, **3**(4), 425-434.

USDA (U.S. Dept. of Agriculture, Economic Research Service). (1999). Fruit and Tree Nut Situation and Outlook Report, Pp. 258-287.

Van Blommestein, A. (2012). *Impact of Selected Environmental Factors on E. coli Growth in River Water and an Investigation of Carry-over to Fresh Produce during Irrigation*. MSc in Food Science Thesis. Stellenbosch University, Stellenbosch, South Africa.

WHO (World Health Organization). (1989). Health guidelines for the use of wastewater in agriculture and aquaculture. Geneva, World Health Organization. *Technical Report Series No 776. World Health Organization*, Switzerland, Geneva.

Willey, J.M., Sherwood, L.M. & Woolverton, C.J. (Eds.). (2008). Microbiology of food. In: *Prescott, Harley & Klien's Microbiology*, 7th Ed. Pp. 1023-1048. New York: McGraw Hill.

Xiao, Z. & Xu, P. (2007). Acetoin metabolism in bacteria. *Critical Reviews in Microbiology*, **33**, 127-140.

Stellenbosch University http://scholar.sun.ac.za

APPENDICES

Appendix A. Isolate identification according to API 20E, including percentage certainty and identification rating

Isolate Name	Source	Identification	Percentage certainty (%)	Identification rating
A11.1	Borehole A1	*Klebsiella pneumoniae*	99.5	Good identification
A11.2	Borehole A1	*Serratia marcescens*	87.1	Acceptable identification to genus
A11.3	Borehole A1	*Escherichia fergusonii*	-	Unacceptable profile
A11.4	Borehole A1	*Escherichia coli*	99.9	Doubtful profile
A11.5	Borehole A1	*Escherichia coli*	99.9	Doubtful profile
A21.1	Borehole A1	*Escherichia coli*	99.5	Very good identification
A21.2	Borehole A1	*Escherichia coli*	99.7	Doubtful profile
A21.3	Borehole A1	*Escherichia coli*	99.2	Doubtful profile
A21.4	Borehole A1	*Escherichia coli*	99.6	Doubtful profile
A22.1	Borehole A1	*Escherichia coli*	99.5	Very good identification
A22.2	Borehole A1	*Serratia odorifera*	99.1	Acceptable identification
A22.3	Borehole A1	*Escherichia coli*	99.5	Very good identification
A22.4	Borehole A1	*Escherichia coli*	99.2	Doubtful profile
A22.5	Borehole A1	*Escherichia coli*	99.6	Doubtful profile
B11.1	Piggery Ferm. Dam	*Escherichia coli*	99.6	Doubtful profile
B11.2	Piggery Ferm. Dam	*Escherichia coli*	99.9	Good identification
B11.3	Piggery Ferm. Dam	*Escherichia coli*	-	Unacceptable profile
B11.4	Piggery Ferm. Dam	*Escherichia coli*	95.1	Doubtful profile
B12.1	Piggery Overflow	?	-	Unacceptable profile
B12.2	Piggery Overflow	*Escherichia coli*	96.3	Good identification
B12.3	Piggery Overflow	*Escherichia coli*	98.8	Doubtful profile
B12.4	Piggery Overflow	*Escherichia coli*	99.6	Doubtful profile

Stellenbosch University http://scholar.sun.ac.za

B13.1	Piggery Big Dam	*Escherichia coli*	99.2	Very good identification
B13.2	Piggery Big Dam	*Escherichia coli*	-	Unacceptable profile
B13.3	Piggery Big Dam	*Escherichia coli*	-	Unacceptable profile
B13.4	Piggery Big Dam	*Escherichia coli*	73.5	Low discrimination
B21.1	Piggery Ferm. Dam	*Escherichia coli*	99.9	Excellent identification
B21.2	Piggery Ferm. Dam	*Escherichia coli*	99.2	Very good identification
B21.3	Piggery Ferm. Dam	*Escherichia coli*	97.5	Good identification
B21.4	Piggery Ferm. Dam	*Escherichia coli*	73.5	Low discrimination
B21.5	Piggery Ferm. Dam	*Escherichia coli*	73.5	Doubtful profile
B22.1	Piggery Overflow	*Escherichia coli*	97.5	Good identification
B22.2	Piggery Overflow	*Escherichia coli*	99.9	Excellent identification
B22.3	Piggery Overflow	*Escherichia coli*	99.9	Doubtful profile
B22.4	Piggery Overflow	*Escherichia coli*	99.6	Doubtful profile
B22.5	Piggery Overflow	*Escherichia coli*	97.5	Good identification
B23.1	Piggery Big Dam	*Escherichia coli*	99.9	Excellent identification
B23.2	Piggery Big Dam	*Escherichia coli*	99.9	Excellent identification
B23.3	Piggery Big Dam	*Escherichia coli*	99.6	Doubtful profile
B23.4	Piggery Big Dam	*Escherichia coli*	99.6	Doubtful profile
B23.5	Piggery Big Dam	*Escherichia coli*	99.2	Doubtful profile
C11.1	Spring C1	?	-	Doubtful profile
C11.2	Spring C1	?	-	Unacceptable profile
C11.3	Spring C1	*Citrobacter spp.*	91.2	Doubtful profile
C11.4	Spring C1	*Citrobacter spp.*	91.2	Doubtful profile
C21.1	Spring C1	*Escherichia coli*	99.9	Doubtful profile
C21.2	Spring C1	*Escherichia coli*	99.2	Doubtful profile
C21.3	Spring C1	*Escherichia coli*	99.5	Very good identification

Stellenbosch University http://scholar.sun.ac.za

C21.4	Spring C1	*Escherichia coli*	99.9	Doubtful profile
C21.5	Spring C1	*Escherichia coli*	99.5	Doubtful profile
D11.1	Borehole D1	*Escherichia coli*	87.3	Doubtful profile
D11.2	Borehole D1	*Escherichia coli*	99.5	Very good identification
D11.3	Borehole D1	*Escherichia coli*	99.5	Very good identification
D11.4	Borehole D1	*Escherichia coli*	99.5	Very good identification
D12.1	Dam D1	*Escherichia coli*	99.5	Very good identification
D12.2	Dam D1	*Escherichia coli*	99.2	Doubtful profile
D12.3	Dam D1	*Escherichia coli*	99.5	Very good identification
D12.4	DamD1	*Escherichia coli*	99.5	Very good identification
D21.1	Borehole D1	*Escherichia coli*	99.5	Very good identification
D21.2	Borehole D1	?	-	Low discrimination
D21.3	Borehole D1	*Escherichia coli*	99.5	Very good identification
D21.4	Borehole D1	*Escherichia coli*	99.5	Very good identification
D21.5	Borehole D1	*Escherichia coli*	99.5	Very good identification
D24.1	DamD1	*Escherichia coli*	99.7	Very good identification
D24.2	DamD1	*Escherichia coli 2*	99.7	Very good identification
D24.3	DamD1	*Escherichia coli*	99.3	Very good identification
D24.4	DamD1	*Escherichia coli*	99.3	Doubtful profile
D24.5	DamD1	*Escherichia coli*	99.3	Very good identification
E11.1	Plank 0	*Escherichia coli*	99.9	Doubtful profile
E11.2	Plank 0	*Escherichia coli*	99.9	Doubtful profile
E11.3	Plank 0	*Escherichia coli*	99.9	Doubtful profile
E11.4	Plank 0	*Escherichia coli*	99.9	Doubtful profile
E11.5	Plank 0	*Escherichia coli*	99.9	Doubtful profile
E12.1	Plank 1	*Escherichia coli*	99.5	Very good identification

E12.2	Plank 1	*Escherichia coli*	99.9	Doubtful profile
E12.3	Plank 1	*Klebsiella pneumoniae*	98.4	Good identification
E12.4	Plank 1	*Citrobacter braakii*	59.3	Doubtful profile
E21.1	Plank 0	*Escherichia coli*	99.9	Excellent identification
E21.2	Plank 0	*Escherichia coli*	99.5	Very good identification
E21.3	Plank 0	*Escherichia coli*	98.8	Good identification
E21.4	Plank 0	*Escherichia coli*	99.9	Excellent identification
E21.5	Plank 0	*Escherichia coli*	99.2	Doubtful profile
E22.1	Plank 1	*Escherichia coli*	99.9	Excellent identification
E22.2	Plank 1	*Escherichia coli*	99.8	Doubtful profile
E22.3	Plank 1	*Escherichia coli*	99.9	Excellent identification
E22.4	Plank 1	*Escherichia coli*	99.9	Excellent identification
E22.5	Plank 1	*Escherichia coli*	99.5	Very good identification
F11.1	Plank 3	*Citrobacter freundii*	97.5	Doubtful profile
F11.2	Plank 3	*Escherichia coli*	99.6	Doubtful profile
F11.3	Plank 3	*Erwinia spp?*	-	Low discrimination
F11.4	Plank 3	*Escherichia coli*	99.5	Very good identification
F21.1	Plank 3	*Escherichia coli*	99.5	Very good identification
F21.2	Plank 3	*Pantoea/Serratia*	-	Low discrimination
F21.3	Plank 3	*Escherichia coli*	99.9	Doubtful profile
F21.4	Plank 3	*Escherichia coli*	99.2	Doubtful profile
F21.5	Plank 3	*Escherichia coli*	99.2	Doubtful profile
G11.1	Veldwagters River	*Escherichia coli*	99.5	Very good identification
G11.2	Veldwagters River	*?*	-	Unacceptable profile
G21.1	Veldwagters River	*Escherichia coli*	99.7	Very good identification
G21.2	Veldwagters River	*Escherichia coli*	99.6	Doubtful profile

Stellenbosch University http://scholar.sun.ac.za

G21.3	Veldwagters River	*Escherichia coli*	95.8	Good identification
G21.4	Veldwagters River	*Escherichia coli*	82.7	Very good identification to genus
G21.5	Veldwagters River	*Escherichia coli*	99.2	Doubtful profile
G22.1	Veldwagters River	?	-	Low discrimination
H11.1	Olifants River	*Escherichia coli*	99.9	Excellent identification
H11.2	Olifants River	*Escherichia coli*	99.9	Excellent identification
H11.3	Olifants River	*Escherichia coli*	99.9	Doubtful profile
H11.4	Olifants River	*Escherichia coli*	99.9	Doubtful profile
H11.5	Olifants River	*Escherichia coli*	99.9	Doubtful profile
H23.1	Olifants River	*Escherichia coli*	99.9	Very good identification
H23.2	Olifants River	*Escherichia coli*	99.6	Very good identification
H23.3	Olifants River	*Escherichia coli*	99.6	Very good identification
H23.4	Olifants River	*Escherichia coli*	99.6	Very good identification
H23.5	Olifants River	?	-	Unacceptable profile
H24.1	Olifants River	*Escherichia coli*	995	Very good identification
J11.1	Mosselbank River	*Escherichia coli*	99.9	Excellent identification
J11.2	Mosselbank River	*Escherichia coli*	99.5	Very good identification
J11.3	Mosselbank River	*Escherichia coli*	99.2	Doubtful profile
J11.4	Mosselbank River	*Escherichia coli*	99.9	Excellent identification
J11.5	Mosselbank River	*Escherichia coli*	99.2	Doubtful profile
J21.1	Mosselbank River	*Escherichia coli*	99.5	Very good identification
J21.2	Mosselbank River	*Escherichia coli*	99.5	Very good identification
J21.3	Mosselbank River	*Escherichia coli*	99.2	Doubtful profile
J21.4	Mosselbank River	*Escherichia coli*	99.9	Excellent identification
J21.5	Mosselbank River	*Escherichia coli*	99.2	Doubtful profile
K11.1	Berg 2	*Escherichia coli*	99.5	Very good identification

Stellenbosch University http://scholar.sun.ac.za

K11.3	Berg 2	*Escherichia coli*	99.2	Doubtful profile
K11.4	Berg 2	*Escherichia coli*	99.5	Very good identification
K11.5	Berg 2	*Escherichia coli*	99.2	Doubtful profile
L11.1	Winery effluent	*Escherichia coli*	99.9	Excellent identification
L11.2	Winery effluent	*Escherichia coli*	99.9	Excellent identification
L11.3	Winery effluent	*Escherichia coli*	99.6	Doubtful profile
L11.4	Winery effluent	*Escherichia coli*	99.9	Excellent identification
L11.5	Winery effluent	*Escherichia coli*	99.2	Doubtful profile
L11.6	Winery effluent	*Escherichia coli*	99.6	Doubtful profile
L21.1	Winery effluent	?	-	Low discrimination
L21.2	Winery effluent	?	-	Low discrimination
L21.3	Winery effluent	?	-	Low discrimination
L21.4	Winery effluent	?	-	Low discrimination
L21.5	Winery effluent	?	-	Low discrimination
M11.1	Dairy	*Escherichia coli*	99.8	Very good identification
M11.2	Dairy	*Escherichia coli*	99.8	Very good identification
M11.3	Dairy	*Escherichia coli*	-	Unacceptable profile
M11.4	Dairy	*Escherichia coli*	-	Unacceptable profile
M21.1	Dairy	*Escherichia coli*	99.5	Very good identification
M21.2	Dairy	*Escherichia coli*	99.2	Doubtful profile
M21.3	Dairy	*Escherichia coli*	99.9	Doubtful profile
M21.4	Dairy	*Escherichia coli*	99.9	Excellent identification
M21.4	Dairy	*Escherichia coli*	82.1	Acceptable identification
M22.1	Dairy	*Escherichia coli*	99.9	Excellent identification
M22.2	Dairy	*Escherichia coli*	99.2	Doubtful profile
M22.3	Dairy	*Escherichia coli*	99.2	Doubtful profile

Stellenbosch University http://scholar.sun.ac.za

M22.4	Dairy	*Escherichia coli*	99.2	Doubtful profile
M22.5	Dairy	*Escherichia coli*	99.8	Very good identification

? = Unknown isolate

- = no percentage given

Appendix B. Isolate identification according to MALDI-TOF, including identification score

Isolate Name	Source	Identification	Identification Score
A11.1	Borehole A1	*Klebsiella pneumoniae*	2.354
A11.2	Borehole A1	*Serratia marcescens*	2.465
A11.3	Borehole A1	*Escherichia coli**	2.339
A11.4	Borehole A1	*Escherichia coli**	2.410
A11.5	Borehole A1	*Escherichia coli**	2.531
A21.1	Borehole A1	*Escherichia coli**	2.518
A21.2	Borehole A1	*Escherichia coli**	2.458
A21.3	Borehole A1	*Escherichia coli**	2.491
A21.4	Borehole A1	*Escherichia coli**	2.450
A22.1	Borehole A1	*Escherichia coli**	2.437
A22.2	Borehole A1	*Serratia odorifera*	2.323
A22.3	Borehole A1	*Escherichia coli**	2.448
A22.4	Borehole A1	*Escherichia coli**	2.464
A22.5	Borehole A1	*Escherichia coli**	2.445
B11.1	Piggery Ferm. Dam	*Escherichia coli**	2.447
B11.2	Piggery Ferm. Dam	*Escherichia coli***	2.515
B11.3	Piggery Ferm. Dam	*Escherichia coli**	2.533
B11.4	Piggery Ferm. Dam	*Escherichia coli**	2.506
B12.1	Piggery Overflow	*Escherichia coli**	2.485
B12.2	Piggery Overflow	*Paenibacillus amylolyticus*	2.555
B12.3	Piggery Overflow	*Escherichia coli**	2.460
B12.4	Piggery Overflow	*Escherichia coli**	2.437
B13.1	Piggery Big Dam	*Escherichia coli**	2.394

Stellenbosch University http://scholar.sun.ac.za

B13.2	Piggery Big Dam	*Escherichia coli**	2.314
B13.3	Piggery Big Dam	*Escherichia coli**	2.378
B13.4	Piggery Big Dam	*Escherichia coli**	2.406
B21.1	Piggery Ferm. Dam	*Escherichia coli**	2.462
B21.2	Piggery Ferm. Dam	*Escherichia coli**	2.412
B21.3	Piggery Ferm. Dam	*Escherichia coli**	2.428
B21.4	Piggery Ferm. Dam	*Escherichia coli**	2.391
B21.5	Piggery Ferm. Dam	*Escherichia coli**	2.416
B22.1	Piggery Overflow	*Escherichia coli**	2.461
B22.2	Piggery Overflow	*Escherichia coli**	2.315
B22.3	Piggery Overflow	*Escherichia coli**	2.403
B22.4	Piggery Overflow	*Escherichia coli**	2.451
B22.5	Piggery Overflow	*Escherichia coli**	2.397
B23.1	Piggery Big Dam	*Escherichia coli**	2.321
B23.2	Piggery Big Dam	*Escherichia coli**	2.330
B23.3	Piggery Big Dam	*Escherichia coli**	2.349
B23.4	Piggery Big Dam	*Escherichia coli**	2.420
B23.5	Piggery Big Dam	*Escherichia coli**	2.434
C11.1	Spring C1	*Citrobacter freundii*	2.386
C11.2	Spring C1	*Citrobacter freundii*	2.408
C11.3	Spring C1	*Citrobacter freundii*	2.397
C11.4	Spring C1	*Citrobacter freundii*	2.356
C21.1	Spring C1	*Escherichia coli**	2.384
C21.2	Spring C1	*Escherichia coli**	2.383
C21.3	Spring C1	*Escherichia coli**	2.439
C21.4	Spring C1	*Escherichia coli**	2.349
C21.5	Spring C1	*Escherichia coli**	2.419
D11.1	Borehole D1	*Escherichia coli**	2.335
D11.2	Borehole D1	*Escherichia coli**	2.418
D11.3	Borehole D1	*Escherichia coli**	2.428
D11.4	Borehole D1	*Escherichia coli**	2.430

Stellenbosch University http://scholar.sun.ac.za

D12.1	Dam D1	*Escherichia coli**	2.448
D12.2	Dam D1	*Escherichia coli**	2.330
D12.3	Dam D1	*Escherichia coli**	2.343
D12.4	Dam D1	*Escherichia coli**	2.465
D21.1	Borehole D1	*Escherichia coli**	2.368
D21.2	Borehole D1	*Escherichia coli**	2.507
D21.3	Borehole D1	*Escherichia coli**	2.512
D21.4	Borehole D1	*Escherichia coli**	2.398
D21.5	Borehole D1	*Escherichia coli**	2.445
D24.1	Dam D1	*Escherichia coli**	2.486
D24.2	Dam D1	*Escherichia coli**	2.567
D24.3	Dam D1	*Escherichia coli**	2.549
D24.4	Dam D1	*Escherichia coli**	2.461
D24.5	Dam D1	*Escherichia coli**	2.300
E11.1	Plank 0	*Escherichia coli**	2.567
E11.2	Plank 0	*Escherichia coli**	2.471
E11.3	Plank 0	*Escherichia coli**	2.542
E11.4	Plank 0	*Escherichia coli**	2.535
E11.5	Plank 0	*Escherichia coli**	2.483
E12.1	Plank 1	*Escherichia coli**	2.379
E12.2	Plank 1	*Escherichia coli**	2.549
E12.3	Plank 1	*Klebsiella pneumoniae*	2.348
E12.4	Plank 1	*Escherichisa coli*	2.212
E21.1	Plank 0	*Escherichia coli**	2.477
E21.2	Plank 0	*Escherichia coli**	2.524
E21.3	Plank 0	*Escherichia coli**	2.442
E21.4	Plank 0	*Escherichia coli**	2.507
E21.5	Plank 0	*Escherichia coli**	2.452
E22.1	Plank 1	*Escherichia coli**	2.351
E22.2	Plank 1	*Escherichia coli**	2.143
E22.3	Plank 1	*Escherichia coli**	2.393

Stellenbosch University http://scholar.sun.ac.za

E22.4	Plank 1	*Escherichia coli**	2.438
E22.5	Plank 1	*Escherichia coli**	2.510
F11.1	Plank 3	*Citrobacter freundii*	2.548
F11.2	Plank 3	*Escherichia coli**	2.366
F11.3	Plank 3	*Escherichia coli**	2.579
F11.4	Plank 3	*Escherichia coli**	2.514
F21.1	Plank 3	*Escherichia coli**	2.501
F21.2	Plank 3	*Klebsiella oxytoca*	2.325
F21.3	Plank 3	*Escherichia coli**	2.300
F21.4	Plank 3	*Escherichia coli**	2.231
F21.5	Plank 3	*Escherichia coli**	2.475
G11.1	Veldwagters River	*Escherichia coli**	2.495
G11.2	Veldwagters River	*Enterobacter spp.*	2.337
G21.1	Veldwagters River	*Escherichia coli**	2.461
G21.2	Veldwagters River	*Escherichia coli**	2.382
G21.3	Veldwagters River	*Escherichia coli**	2.344
G21.4	Veldwagters River	*Escherichia coli**	2.312
G21.5	Veldwagters River	*Escherichia coli**	2.263
G22.1	Veldwagters River	*Escherichia coli**	2.373
H11.1	Olifants River	*Escherichia coli**	2.456
H11.2	Olifants River	*Escherichia coli**	2.440
H11.3	Olifants River	*Escherichia coli**	2.489
H11.4	Olifants River	*Escherichia coli**	2.399
H11.5	Olifants River	*Escherichia coli**	2.506
H23.1	Olifants River	*Escherichia coli**	2.381
H23.2	Olifants River	*Escherichia coli**	2.363
H23.3	Olifants River	*Escherichia coli**	2.349
H23.4	Olifants River	*Escherichia coli**	2.415
H23.5	Olifants River	*Escherichia coli**	2.310
H24.1	Olifants River	*Escherichia coli**	2.478
J11.1	Mosselbank River	*Escherichia coli**	2.250

Stellenbosch University http://scholar.sun.ac.za

J11.2	Mosselbank River	*Escherichia coli**	2.550
J11.3	Mosselbank River	*Escherichia coli**	2.299
J11.4	Mosselbank River	*Escherichia coli**	2.311
J11.5	Mosselbank River	*Escherichia coli**	2.470
J21.1	Mosselbank River	*Escherichia coli**	2.527
J21.2	Mosselbank River	*Escherichia coli**	2.351
J21.3	Mosselbank River	*Escherichia coli**	2.434
J21.4	Mosselbank River	*Escherichia coli**	2.300
J21.5	Mosselbank River	*Escherichia coli**	2.391
K11.1	Berg 2	*Escherichia coli**	2.443
K11.3	Berg 2	*Escherichia coli**	2.396
K11.4	Berg 2	*Escherichia coli**	2.436
K11.5	Berg 2	*Escherichia coli**	2.480
L11.1	Winery effluent	*Escherichia coli**	2.347
L11.2	Winery effluent	*Escherichia coli**	2.463
L11.3	Winery effluent	*Escherichia coli**	2.388
L11.4	Winery effluent	*Escherichia coli**	2.433
L11.5	Winery effluent	*Escherichia coli**	2.362
L11.6	Winery effluent	*Escherichia coli**	2.277
L21.1	Winery effluent	*Escherichia coli**	2.326
L21.2	Winery effluent	*Escherichia coli**	2.481
L21.3	Winery effluent	*Escherichia coli**	2.479
L21.4	Winery effluent	*Escherichia coli**	2.287
L21.5	Winery effluent	*Escherichia coli**	2.276
M11.1	Dairy	*Escherichia coli**	2.497
M11.2	Dairy	*Escherichia coli**	2.438
M11.3	Dairy	*Escherichia coli**	2.386
M11.4	Dairy	*Escherichia coli**	2.477
M21.1	Dairy	*Escherichia coli**	2.343
M21.2	Dairy	*Escherichia coli**	2.209
M21.3	Dairy	*Escherichia coli**	2.422

Stellenbosch University http://scholar.sun.ac.za

M21.4	Dairy	*Escherichia coli**	2.416
M21.5	Dairy	*Serratia odorifera*	2.351
M22.1	Dairy	*Escherichia coli**	2.383
M22.2	Dairy	*Escherichia coli**	2.513
M22.3	Dairy	*Escherichia coli**	2.498
M22.4	Dairy	*Escherichia coli**	2.433
M22.5	Dairy	*Escherichia coli**	2.450

*Closely related to *Shigella* and not definitely distinguishable at the moment

Stellenbosch University http://scholar.sun.ac.za

CHAPTER 4

EVALUATION OF PHYLOGENETIC GROUPINGS AND ABUNDANT RIBOSOMAL PROTEINS AS MST MARKERS FOR ASSESSMENT OF THE ORIGIN OF FAECAL CONTAMINATION OF IRRIGATION WATER

SUMMARY

Natural water sources are often used for irrigation of fresh produce, especially in the Western Cape as a result of the winter rainfall. This is a problem, as natural water sources are laden with coliforms and other potentially harmful bacteria. *Escherichia coli* are considered a potential risk in these water sources, and the types of *E. coli* as well as microbial source tracking (MST) were further investigated in this study.

In total, 153 *E. coli* isolated from 19 samples sites were characterised using molecular methods to determine phylogeny and pathogenicity. The majority of the strains (143) represented irrigation and contamination water in the Western Cape, while 10 strains were selected from an "environmental" site to act as a control. Phylogeny, together with MALDI-TOF MS data, was used to link *E. coli* from irrigation water to their most probable contamination source. Isolated *E. coli* strains were grouped into one of the seven phylogenetic subgroups and the population structures of the *E. coli* present at each sample site was compared. Phylogenetic group B1 has been reported to contain strains which survive in the environment with ease. In this study 36.57% of isolates were grouped in this phylogenetic group. The data from this study showed that *E. coli* isolates originating from humans, cows and pigs all have different population structures based on phylogenetic groups. *Escherichia coli* originating from pigs were of the main phylogenetic group A as the majority (80%) while strains originating from a bovine source, showed phylogenetic subgroup B1 to be the most prevalent (71.43%). *Escherichia coli* strains with human origins were most commonly (50%) grouped in phylogenetic subgroup A_1.

Strains from irrigation water showed similar phylogenetic distribution patterns and phylogenetic subgroup B1 was seen as the most common group for both surface and groundwater isolates (59.26% and 44.44% respectively). Strains from groundwater sites did however have a population structure more indicative of porcine origin, while surface water showed population characteristics more in line with contamination arising from a dairy farm. The small "environmental" group (10 strains) exhibited a completely different population structure, with the prominent grouping as $B2_3$ (50%).

Stellenbosch University http://scholar.sun.ac.za

Two enteropathogenic *E. coli* (EPEC) strains were isolated, one from Plankenburg river site 3, and one from Borehole A in the Drakenstein area. EPEC was previously isolated from the Plankenburg River (Van Blommestein, 2012) which suggests a consistent EPEC contamination source. The isolation of a pathogenic strain from borehole water was unexpected as borehole water is traditionally seen as 'clean' (Bezuidenhout *et al*., 2011). The borehole was also found to have much lower *E. coli* counts than the other water sources investigated.

It was concluded that MST is not easily conducted with *E. coli* when applying it to natural water sources. This is due to the constant cross contamination occurring between various water sources, especially during the rainy season. Some links were observed between irrigation water and contamination sources, but a more in depth study would help strengthen the conclusions made here.

INTRODUCTION

It is known that the natural water sources used for irrigation of fresh produce contain microbial contaminants and a portion of the microbial pollution, especially in rivers, consists of *E. coli* (Ackermann, 2010; Lötter, 2010; Kikine, 2011). This however does not mean very much without analysing the composition of the *E. coli* population present, due to the diversity of *E. coli* strains, all with a variety of associated risk factors (Groisman, 2001; Johnson & Russo, 2002; Johnson & Russo, 2003; Percival *et al.,* 2004). *Escherichia coli* are also known for their high mutation rate and ability to acquire genes from other bacteria, whether from the same genus or not (LeClerc *et al*., 1996; Matic *et al*., 1997; Ochman *et al*., 2000; Reid *et al.,* 2000). If a variety of *E. coli* strains are present, it also means that some strains may be pathogenic while others are harmless. Of the pathogenic strains, some are also more infectious than others, with the EHEC group having the lowest infectious dose (<100 bacterial cells), making it particularly important to determine pathogenicity of the existing *E. coli* in water being used for irrigation of fresh produce (Karmali, 1989; Bolton *et al*., 1996; Percival *et al*., 2004). Therefore, without further investigation of the *E. coli* population structure in the water, potential risk associated with the contaminated water cannot be determined.

Phylogeny of *E. coli* strains has been shown to be successful in helping group *E. coli* strains according to their origin. Four major, and seven subgroupings, of *E. coli* phylogeny exist (Herzer *et al*., 1990; Denamur *et al*., 1999; Carlos *et al*., 2010), all of which are used to better understand *E. coli* origin, as well as potential virulence. It has been reported that extraintestinal strains are

Stellenbosch University http://scholar.sun.ac.za

most commonly found in groups B2 and D, while environmental strains most often occur in phylo-group B1 (Denamur *et al*., 1999; Johnson & Stell, 2000; Carlos *et al*., 2010). Groups A, B1 and D contain intestinal pathogens (Pupo *et al*., 1997) and commensal strains are most commonly classified in phylo-group A (Bingen *et al*., 1998). When using all seven sub-groups, a better, and more reliable, division is however possible. In 2010, Carlos *et al*. used all seven groups to successfully link *E. coli* strains to their origins (humans, cows, pigs, chickens, sheep and goats). The use of phylogenetic grouping of *E. coli* is therefore especially useful when conducting MST, and will be used in this study as a basis for tracking the microbial source.

The aims of this study were therefore to use both molecular methods and the MALDI Biotyper system to create a fingerprint of *E. coli* isolated from the environment, and to use this fingerprint to link *E. coli* strains from irrigation water to its most probable contamination source origin. Pathogenicity of *E. coli* strains and potential risk associated with using the investigated irrigation water for irrigation of fresh produce will be determined. For the purpose of clarity, throughout this study associated risk will be defined as a measure of pathogenic *E. coli* in the river which may potentially cause disease if ingested by a susceptible host in adequate numbers.

MATERIALS AND METHODS

Site Selection

A total of 19 sites were selected (Tables 1 and 2) and each sampled twice over a period of five months, from January to May 1012. Samples were taken in such a way that follow-up samples from a single site were approximately two months apart. This was to ensure that sufficient time could pass between sampling opportunities so that a representative sample could be attained. Of the 19 selected sites, 12 sites were classified strictly as irrigation sources, five strictly as contamination sources and two sites overlapped categories and were classified both as irrigation and contamination sources. This meant that in total there were 14 irrigation sites (Table 1) and seven contamination source sites (Table 2) investigated. The sites were chosen mostly in the Stellenbosch and surrounding areas, while there was one site adjacent to the Cape Flats area near Muizenburg, one further North West near Lutzville, one in the Drakenstein area and four sites in Wellington/Paarl area.

Stellenbosch University http://scholar.sun.ac.za

Irrigation source sites

The irrigation source sites were selected from a range of different types of water sources including; six boreholes, five rivers, one dam, one spring and one cellar effluent sample which has had sewage added to the effluent (Table 1). Sample sites were chosen on the grounds that the water being sampled must be extracted at the same point as that being used for irrigation of fresh produce.

Table 1. Irrigation sites, their geographical locations and water application

Water source	Geographic Location	Used to irrigate
Borehole A1	Drakenstein	Herbs
Borehole A2	Drakenstein	Herbs
Borehole D	Wellington	Fresh produce, grapes and used in winemaking
Borehole N1	Strandfontein (adj. to Cape Flats)	Herbs and salad products
Borehole N2	Strandfontein (adj. to Cape Flats)	Herbs and salad products
Borehole P	Raithby	Fresh produce as well as home use
Spring	Wellington	Fresh produce as well as home use
Plankenburg River (Plank-3)	Stellenbosch	Fresh produce on various farms
Veldwagters River	Stellenbosch	Fresh produce on various farms
Olifants River	Lutzville	Fresh produce and crops also used in houses
Mosselbank River	Kraaifontein	Fresh produce
Berg River (Berg-2)	Paarl/Franschhoek	Fresh produce and fruit
Winery effluent (Effluent L1)	Stellenbosch (outlying)	Fresh produce consumed by farm workers
Dam	Wellington	Fresh produce, grapes, used in winemaking and also used in house

Contamination source sites

The second set of sampling sites was chosen to represent potential contamination sources from where the *E. coli* may have originated from. These sites were therefore chosen expecting that they would have a high microbial load, especially in terms of total coliform bacteria and *E. coli* content. The contaminated water also had to be able to reach another water source which could contribute to a river or dam which is used for irrigation (Table 2). Agricultural activity was represented by both a dairy and a piggery, in both cases the samples were taken from the water being used to wash faecal matter out of the pig and cow stalls. The wash water (high in faecal contamination) is then directed to a dam slightly further along which could over-flow and reach a

Stellenbosch University http://scholar.sun.ac.za

nearby marsh which signals the start of the Plankenburg River. In both cases the subsequent dams were also sampled. Industrial activity was represented by partially treated wine cellar effluent. This water was expected to have a high microbial count as untreated human sewage is added to the effluent and then the water is treated by means of a constructed wetland. This is done with the hopes of decreasing the concentration of unwanted elements (carbon, nitrogen and phosphate). This industrial water source is also used to irrigate fresh produce consumed by the workers on the wine farm.

The human factor (any human originating contamination such as sewage, litter or wash water) was represented by two sample sites; one being the effluent from a local sewage treatment facility, and the second represented human pollution as a whole as water was sampled from the Plankenburg River before and after flowing through Kayamandi, a large informal settlement in Stellenbosch. The effluent from the sewage treatment facility flows directly into the Veldwagters River, which then converges with the Eerste River. Water from the Veldwagters and Eerste Rivers are then used by nearby farms for irrigation purposes, and as a result the Veldwagters River sample was classified as both a contamination source (Table 2) and irrigation site (Table 1). The winery effluent was also classified under both contamination sources and irrigation water samples.

Table 2. Contamination source sites, their geographical locations and the main contributor of contamination

Water source	Geographic Location	Contamination source
Piggery effluent (2 sites)	Stellenbosch	Porcine
Plankenburg River (Plank-0)	Stellenbosch	Environmental (control for Plankenburg River 1)
Plankenburg River (Plank-1)	Stellenbosch	Human pollution
Veldwagters River	Stellenbosch	Sewage
Winery effluent (Effluent L1)	Stellenbosch (outlying)	Industrial effluent
Cow Farm effluent	Stellenbosch	Bovine

Stellenbosch University http://scholar.sun.ac.za

Sample Collection

Water samples were collected in sterile bottles (Schott) using the standard methods set out in SANS 5667-6 (SANS, 2006a) and 5667-11 (SANS, 2006b) for rivers and streams and groundwater respectively. Stringent safety measures were applied to ensure the safety of the sampler, and water samples were transported back to the laboratory at 4°C and examined within six hours of collection to minimise bacterial growth before analysis.

Sample Analysis

Isolation of Escherichia coli

Total coliforms and *E. coli* counts were determined using the standard methods laid out by SANS 9308 (SANS, 2012) describing the defined substrate technology (Colilert) method. Fluorescent wells (positive for *E. coli* and total coliforms) were then removed from the QuantiTray trays and used further to isolate *E. coli* strains. Eosin Methylene-blue Lactose Sucrose Agar (L-EMB) (Oxoid CM0069, South Africa) and Brilliance™ *E. coli* coliform selective agar (Oxoid CM1046, England) were used to isolate presumptive colonies, and the API 20E system (BioMérieux, South Africa) was used to obtain an initial identification. A ribosomal protein extract was made from each isolate, and matrix assisted laser desorption/ionisation time-of-flight mass spectrometry (MALDI-TOF MS) (Bruker, Germany) was used to confirm the identification made by the API 20E system (BioMérieux, South Africa).

Reference strains

Table 3. List of reference strains (ATCC = American Type Culture Collection)

Strain	ATCC code
Escherichia coli 58	ATCC 11775
Escherichia coli 157	ATCC 4350
ATCC 35218	ATCC 35218
ATCC 25922	ATCC 25922

Reference strains of *E. coli* (Table 3) were added to the dataset of isolates to act as controls. Reference strains were used to assist in isolate identification and characterisation as a comparative control sample, as the properties of the reference strains are well documented.

Stellenbosch University http://scholar.sun.ac.za

When using the MALDI-TOF MS, the reference samples were used to determine whether the instrument was accurate when identifying isolates.

Numerical analysis of isolate profiles

The spectra obtained in the baseline study (chapter 3 of this thesis) to confirm identification of all isolates using MALDI-TOF MS (matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry) fingerprinting, were further analysed in this chapter. Spectra obtained were viewed using the Biotyper 3.0 real time classification system (Bruker, Germany) and principle component analysis (PCA) was then performed on the data. The unsorted data was then organised in a hierarchical fashion and a PCA dendrogram was created.

One spectra of each environmental isolate and duplicate spectra from each of the reference cultures were included in the analysis. The addition of duplicates for the reference cultures was included to determine to what extend spectrum variation occurs during MALDI-TOF analysis.

PCR Methods

DNA template preparation

Isolates were cultivated on tryptone soy agar (TSA) (Oxoid) for 24 h at 35°C. Following this, a loop of each culture was boiled in a 1.5 mL micro centrifuge tube with 100 µL nuclease free water for 13 min. The tubes were cooled on ice and centrifuged (Vacutec) for 15 min at 14 000 x g, and the supernatant was transferred to a sterile tube, and stored at -18°C until needed (Altahi & Hassen, 2009)

Phylo-group polymerase chain reaction

Triplex PCR (t-PCR) was conducted on *E. coli* strains to determine the phylogenetic groups, based on the method of Claremont *et al*. (2000). Each 12.5 µL reaction volume consisted of 1 x KAPA 2G fast multiplex PCR master mix (KAPA Biosystems, South Africa), 0.2 µM of each primer (Table 4) and 0.25 µL template DNA.

The positive control which was included in all analyses was *E. coli* strain (ATCC 25922), which contains genetic markers, *chuA* and *yjaA,* as well as the DNA fragment TspE4.C2. Similarly, a negative control was included in each PCR run, where RNase free water was used to replace the

Stellenbosch University http://scholar.sun.ac.za

DNA template. All tubes were then placed in the G-storm thermal cycler (Vacutec, South Africa) and reaction conditions summarised in Table 5, were applied.

Table 4. List of the primer sequences and amplicon sizes used for t-PCR to determine the phylogenetic groups of the *E. coli* isolates (Clermont *et al.*, 2000)

Primer*	Primer sequence (5' - 3')	Size (bp)
yjaA.1 (F)	TGAAGTGTCAGGAGACGCTG	211
yjaA.2 (R)	ATGGAGAATGCGTTCCTCAAC	
chuA.1 (F)	GACGAACCAACGGTCAGGAT	279
chuA.2 (R)	TGCCGCCAGTACCAAAGACA	
TSPE4.C2.1 (F)	GAGTAATGTCGGGGCATTCA	152
TSPE4.C2.2 (R)	CGCGCCAACAAAGTATTACG	

F - Forward primer; R - Reverse primer

Table 5. Summary of t-PCR reaction conditions

	Step number	Action	Temperature (°C)	Time (mm:ss)
	1	Initial denaturation	95	03:00
Together form one cycle (x30)	2	Denaturation	95	00:30
	3	Primer annealing	60	00:30
	4	Elongation	72	00:30
	5	Final elongation	72	05:00
	6	cooling	4	10:00

Gel electrophoresis, using a 2% agarose (SeeKem, Switzerland) gel was used for the analysis of PCR products. The gel contained 1 μg.mL^{-1} ethidium bromide (Sigma, Germany) to visualise banding patterns using a UV light. Gel electrophoresis was executed for 20 min at 210 V thereafter banding patterns were observed and it was determined which combination of the three

Stellenbosch University http://scholar.sun.ac.za

possible DNA fragments were amplified in each strain. This allowed each strain to be assigned to one of four main phylogenetic groups (A, B1, B2 or D), and further grouped into seven phylogenetic subgroups, as given in Table 6 (A_0, A_1, B1, $B2_2$, $B2_3$, D_1 or D_2).

Table 6. Distribution of genetic markers *chuA* and *yjaA* and DNA fragment TspE4.C2 denoting each of the seven phylogenetic subgroups (Carlos *et al*., 2010)

Phylogenetic group	*chuA*	*yjaA*	TspE4.C2
A_0	-	-	-
A_1	-	+	-
B1	-	-	+
$B2_2$	+	+	-
$B2_3$	+	+	+
D_1	+	-	-
D_2	+	-	+

Pathotype polymerase chain reaction

A multiplex PCR (m-PCR) method was used to determine the pathotype of the *E. coli* isolates. This in turn will help assess potential risk associated with contaminated water being used for irrigation of fresh produce. Methods described by Omar & Barnard (2010) were modified to optimise results. Each 12.5 µL reaction volume consisted of 1 x KAPA 2G fast multiplex PCR master mix (KAPA Biosystems, South Africa), 0.2 µM of each primer (Table 7) and 0.25 µL template DNA.

As positive control, a standard culture mix (SCM) was prepared, consisting of equal amounts of DNA from each pathotype investigated (ETEC, EAEC, EHEC, EPEC and EIEC). The positive control reaction tube was made up using 1.25 µL SCM to replace the DNA template. A negative control was prepared by replacing the DNA template with RNase-free water. Positive and negative controls were included with each m-PCR run. All tubes were then placed in the G-storm thermal cycler (Vacutec, South Africa), and reaction conditions outlined in Table 8 were applied.

Stellenbosch University http://scholar.sun.ac.za

Table 7. List of primer sequences and amplicon sizes used for the multiplex PCR to detect *Escherichia coli* pathotypes

Pathotype	Primer*	Primer sequence (5' - 3')	Size (bp)	Reference
Commensal (*mdh*)	Mdh01	GGTATGGATCGTTCCGACCT	300	Tarr *et al.*, 2002
	Mdh02	GGCAGAATGGTAACACCAGAGT	300	
EPEC/EHEC (*eaeA*)	L-eaeA(F)	GACCCGGCACAAGCATAAGC	384	Lopez-Suacedo *et al.*, 2003
	L-eaeA(R)	CCACCTGCAGCAACAAGAGG	384	
EHEC (*stx1*, *stx2*)	Stx1(F)	ACACTGGATGATCTCAGTGG	614	Moses *et al.*, 2006
	Stx1(R)	CTGAATCCCCCTCCATTATG	614	
	Stx2(F)	CCATGACAACGGACAGCAGTT	779	Moses *et al.*, 2006
	Stx2(R)	CCTGTCAACTGAGCACTTTG	779	
EIEC (*ial*)	L-ial(F)	GGTATGATGATGATGAGTCCA	650	Lopez-Suacedo *et al.*, 2003
	Ial(R)	GGAGGCCAACAATTATTTCC	650	
EAEC (*eagg*)	Eagg(F)	AGACTCTGGCGAAAG ACT GTATC	194	Pass *et al.*, 2000
	Eagg(R)	ATGGCTGTCTGTAATAGATGAGAAC	194	
ETEC (*LT*, *ST*)	LT(F)	GGCGACAGATTATACCGTGC	450	Lopez-Suacedo *et al.*, 2003
	LT(R)	CGGTCTCTATATTCCCTGTT	450	
	ST(F)	TTTCCCCTCTTTTAGTCAGTCAACT	160	Omar & Barnard, 2010
	ST(R)	GGCAGGATTACAACAAAGTTCACA	160	

F - Forward primer; R - Reverse primer

Table 8. Summary of pathotype m-PCR reaction conditions

	Step number	Action	Temperature (°C)	Time (mm:ss)
	1	Initial denaturation	95	15:00
Together form one cycle (x35)	2	Denaturation	94	00:45
	3	Primer annealing	55	00:45
	4	Elongation	68	02:00
	5	Final elongation	72	05:00
	6	cooling	4	00:30

Gel electrophoresis, using a 1.5% agarose (SeeKem, Switzerland) gel was used to analyse PCR products. The gel contained 1 $\mu g.mL^{-1}$ ethidium bromide (Sigma, Germany) to visualise banding patterns using a UV light, and a 100 bp DNA ladder (Promega, South Africa) was added to an open well when loading the gel to use as a reference. Gel electrophoresis was executed for 90 min at 120 V, and banding patterns were observed and analysed. It was determined whether each

Stellenbosch University http://scholar.sun.ac.za

isolate was a commensal or pathogenic *E. coli* isolate according to which genes had been amplified. *Escherichia coli* strains were grouped into one of the six pathotypes represented by the strains in the SCM or classified as a commensal strain (Table 9).

Table 9. Distribution of genetic markers *mdh, ial, eaeA, stx1, stx2, lt, st* and *eagg* across the seven pathogenic *E. coli* groups

Combination of genes present	Type of *E. coli*	Reference
mdh	Commensal *E. coli*	Tarr *et al*., 2002
mdh, ial	Enteroinvasive *E. coli* (EIEC)	Lopez-Suacedo *et al*., 2003
mdh, eaeA	Enteropathogenic *E. coli* (EPEC)	Lopez-Suacedo *et al*., 2003
mdh, eaeA, stx1	Enterohaemoragic *E. coli* (EHEC)	Lopez-Suacedo *et al*., 2003; Moses *et al*., 2006
mdh, eaeA, stx2	Enterohaemoragic *E. coli* (EHEC)	Lopez-Suacedo *et al*., 2003; Moses *et al*., 2006
mdh, stx1	Shigatoxigenic *E. coli* (STEC)	Moses *et al*., 2006
mdh, stx2	Shigatoxigenic *E. coli* (STEC)	Moses *et al*., 2006
mdh, lt	Enterotoxigenic *E. coli* (ETEC)	Lopez-Suacedo *et al*., 2003
mdh, st	Enterotoxigenic *E. coli* (ETEC)	Omar & Barnard, 2010
mdh, eagg	Enteroaggregative *E. coli* (EAEC)	Pass *et al*., 2000

Results and Discussion

Escherichia coli isolates from natural water sources

In total 148 *E. coli* strains were collected, 138 from irrigation sites and contamination sources in Stellenbosch and surrounding areas, and 10 from the "environmental" control site Plankenburg River site 0 (Plank-0). Of the 138 isolates, 61 (44.20%) were from contamination sites, including 19 strains from sites classified as both irrigation and contamination sites. Isolates from irrigation water were made up of 96 (69.57%) strains; not including the 19 isolates which came from sites classified as both irrigation and contamination sites. The 10 isolates from Plank-0 were not considered to represent either an irrigation site or contamination source, but were later seen as representing environmental strains.

All *E. coli* strains were investigated using the MALDI Biotyper system as well as pathotype and phylogenetic PCR. A character fingerprint was created for each isolate, and used to determine

Stellenbosch University http://scholar.sun.ac.za

whether it is possible to track *E. coli* strains from irrigation water to their most probable contamination source.

Phylogenetic Groupings of *Escherichia coli*

Two genetic markers (*chuA* and *yjaA*) as well as a DNA fragment (TspE4.C2) were used to determine phylogenetic subgroups (Carlos *et al.*, 2010). Fig. 1 shows an example of the PCR amplified genetic markers and DNA fragment after separation on a 1.25% agarose gel. It can be seen that the banding patterns of lanes 2-7 each represent a different phylogenetic subgroup (Fig. 1). Lane 1 is a 100 bp DNA ladder (Promega, South Africa) which is used as a size reference. Using the positive control lane 9 (Fig. 1) and the DNA ladder as reference markers, it can be determined which DNA fragments are amplified in each isolate. The combination of amplified fragments led to grouping of each isolate to a specific phylogenetic subgroup (A_0, A_1, B1, $B2_3$, D_1 or D_2) as shown in Fig. 2.

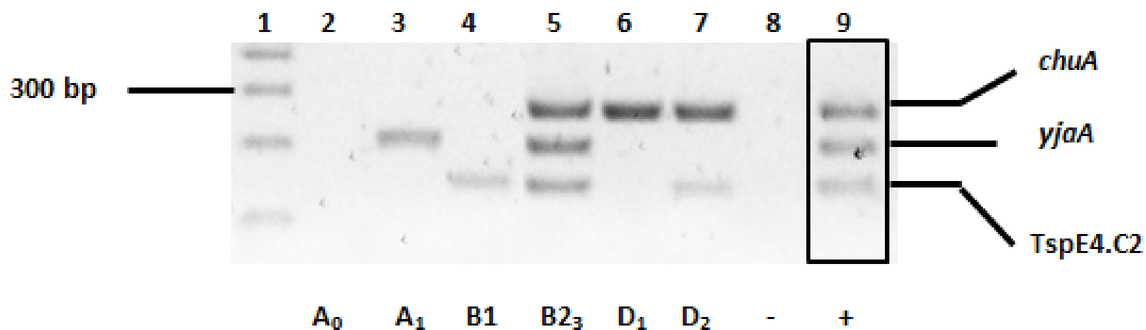

Figure 1. Visual representation of amplified genetic markers and DNA fragments

Porcine

The 25 *Escherichia coli* isolates from the piggery clustered mainly under subgroups A_0 (40%) and A_1 (40%), which means that 80% of all strains isolated from the piggery were classified in main phylogenetic group A (Table 10). In previous studies, group A has been reported to contain both intestinal pathogenic and commensal strains which are most commonly isolated from carnivores, omnivores and non-human mammals (Baldy-Chudzik *et al.*, 2008). Group A is however also seen as a 'generalist' group as strains clustered in group A can be isolated from any vertebrate group (Gordon & Cowling, 2003). It was reported that strains from humans may also regularly occur in phylogenetic group A.

Stellenbosch University http://scholar.sun.ac.za

The other 20% of the strains isolated from the piggery fall under B1, $B2_2$ and D1 (Table 10). B1 contains only three strains, which amounts to 12% of strains isolated from the piggery. Phylogenetic group B1 is also seen as a 'generalist' group (Gordon & Cowling, 2003) as studies have found it to contain *E. coli* strains isolated from a number of hosts (Carlos *et al*., 2010). Strains in phylogenetic group B1 have also been found to survive well in the environment (Walk *et al*., 2007). As viable *E. coli* was isolated from natural water sources in this study, it was expected that a large proportion of the strains would be assigned to phylogenetic group B1. This was the case, as seen in Fig. 2, which therefore concurs with the findings made by Walk *et al*. (2007).

The single strain which was classified in subgroup $B2_2$ is important as it is the only isolated *E. coli* strain which shows the characteristics of this phylo-group (Fig. 2), containing only the genetic markers (*chuA* and *yjaA*). As only one of the 25 (4%) isolates from the piggery was grouped in phylogenetic subgroup $B2_2$, it can possibly indicate that only a small proportion of the *E. coli* population from pigs falls in the subgroup. Since only 8 of 241 isolates (3.32%) were found to belong to phylogenetic subgroup $B2_2$ were also found in a previous study by Carlos *et al*. (2010), it may be that there are very few naturally occurring sources of phylo-group $B2_2$ strains or that standard laboratory methods are not suitable to isolate these strains efficiently.

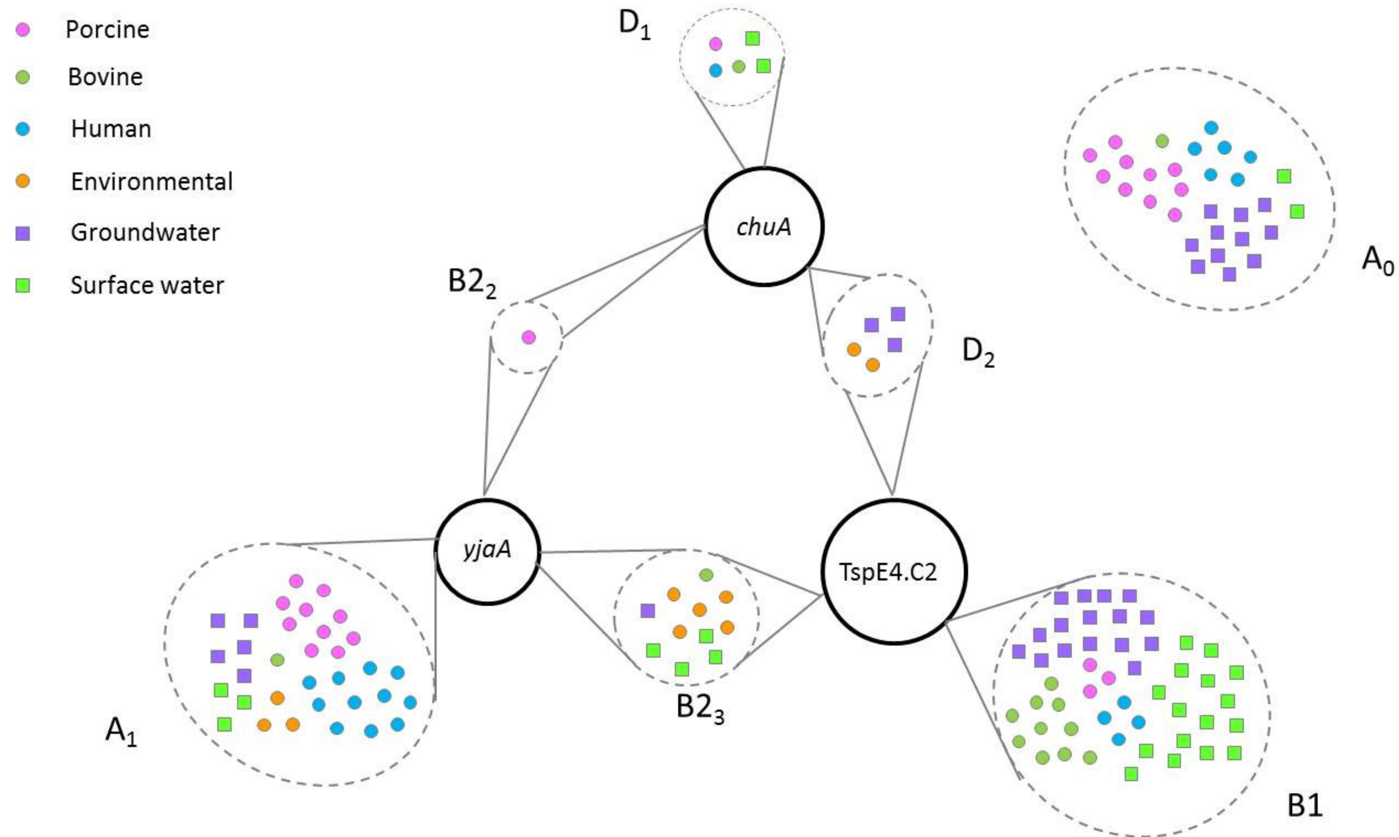

Figure 2. Graphic representation of the occurrence of genetic markers in *E. coli* strains isolated from different hosts (Circles with a solid outline represent the genetic markers (*chuA* and *yjaA*) and the DNA fragment (TspE4.C2). Strains from different contamination sources are represented by small coloured circles, while strains from irrigation sites are represented by coloured squares. Lines leading from the genetic markers to subgroups (outlined in dotted lines) show that that marker was present in all the strains present in that subgroup. Subgroups have been labelled accordingly) (Strains from contamination source sites which are also used for irrigation were only included as contamination source strains and not as irrigation site strains).

Table 10. Distribution of the *E. coli* phylogenetic subgroups among the contamination sources and irrigation sites analysed

Phylogenetic subgroup	Pig	Cow	Human	Environmental	Groundwater	Surface water
A_0	10	1	6	-	11	2
A_1	10	1	11	3	5	3
B1	3	10	4	-	16	16
$B2_2$	1	-	-	-	-	-
$B2_3$	-	1	-	5	1	4
D_1	1	1	1	-	-	2
D_2	-	-	-	2	3	-
Total	**25**	**14**	**22**	**10**	**36**	**27***

*Strains from contamination source sites which are also used for irrigation were only included as contamination source strains and not as irrigation site strains.

Stellenbosch University http://scholar.sun.ac.za

Bovine

Escherichia coli isolated from the dairy (14), were mainly (10/14 = 71.43%) assigned to phylogenetic group B1 (Fig. 2). This confirms that strains isolated from bovine origin are able to persist in the environment with ease (Walk *et al*., 2007) and that they share characteristics with other *E. coli* strains, as B1 is seen as a 'generalist' group (Gordon & Cowling, 2003). Because group B1 can persist in the environment it can possibly be concluded, from the data found in this study that cow manure has been a consistent contaminant of river water, and that is why strains belonging to phylogenetic group B1 are so often isolated from the environment. The rest of the strains from the dairy were grouped in A_0 (7.14%), A_1 (7.14%), $B2_3$ (7.14%) and D_1 (7.14%) (Table 10) with one strain assigned to each group. This may mean a variety of things, but it could be concluded that there is a possibility that extraintestinal *E. coli*, intestinal pathogenic *E. coli* and commensal *E. coli* may be present, as phylogenetic subgroups A_0, A_1, $B2_3$ and D_1 are all present, albeit in low numbers. This was expected, as cows are known to carry commensal and human pathogenic *E. coli* (Avery *et al*., 2008; Willey *et al*., 2008; Linscott, 2011).

Human

The human host was represented by the 22 *E. coli* strains collectively isolated from the Veldwagters River and winery effluent. Veldwagters River strains are classified as being of human origin because the sampling site is downstream from the Stellenbosch sewage treatment facility, and treated effluent from the waste treatment facility drains into the river system (RMS, 2012). The winery effluent water sample is taken from a wine farm where the cellar effluent has sewage (from the farm) added to the effluent before the water is passed through a built wetland. Although neither of these samples were pure untreated sewage, strains from human origin could still contaminate water systems in this way, if sewage treatment was insufficient.

The main group (11/22 = 50%) represented by human hosts was reported to be phylogenetic subgroup A_1 with distribution also across subgroups A_0 (27.27%) and B1 (18.18%) as seen in Table 10. This distribution does not correspond with results conveyed by Carlos *et al*. (2010), as they reported that isolates of human origin were spread among all phylogenetic subgroups. In this study however, no human strains were found to belong to either phylogenetic subgroup $B2_2$, $B2_3$ or D_2 (Table 10). In the study by Carlos *et al*. (2010) however, raw sewage was used as the sample representing strains with human origin. As the raw sewage had not been

Stellenbosch University http://scholar.sun.ac.za

treated, as it was in this study, the difference in population structure may be explained by survival ability of certain strains in the environment.

It is therefore possible that strains which are part of phylogenetic subgroups $B2_2$, $B2_3$ and D_2 are unable to survive the fermentation step that the winery effluent undergoes, once the human sewage is added. Similarly, the treatment that the human sewage undergoes at the sewage treatment facility could mean that the same phylogenetic subgroups are not able to persist. An alternative explanation could be that phylogenetic subgroups $B2_2$, $B2_3$ and D_2 do not survive well outside their host environment, the human gut in this case, and that the strains do not survive once the environment is no longer raw sewage. This could mean that it is the treatment or fermentation steps that cause the loss of these subgroups from raw sewage, or that once the treated or fermented water drains back into a river or dam that those strains cannot survive in the natural water environment.

Surface water

The phylogenetic distribution of the 27 *E. coli* strains form surface water (Plank-3, Veldwagters River, Olifants River, Mosselbank River, Berg-2 and Dam D), showed a majority (59.26%) grouping in phylogenetic subgroup B1 (Table 10). When considering that Walk *et al*. found in 2007 that B1 strains survive in the environment with ease, it is possible to assume that the majority of *E. coli* strains found in the environment will fall within this group. Some of the surface water strains were also distributed in phylogenetic subgroups A_0 (6.67%), A_1 (11.11%), $B2_3$ (14.81%) and D_1 (7.40%) as seen in Table 10. This array of phylogenetic groups found in smaller percentages, can be explained by river water continuously flowing while outside factors constantly introduce new contaminants from a variety of sources. Surface water may also cross contaminate other rivers and dams, particularly in the rainy season when rivers and dams overflow (RMS, 2012).

Groundwater

Groundwater (Boreholes A1, A2, D, N1, N2 and P and spring C) showed similar *E. coli* population structures to those observed in surface water, with the majority (16/36 = 44.44%) of the isolates being assigned to group B1 (Table 10). The main difference between surface and groundwater isolates was that phylogenetic subgroup A_0 was a main group in groundwater, containing 11 isolates (Table 10). Furthermore, smaller proportions of the 36 isolates from groundwater were divided amongst phylogenetic subgroups A_1 (13.89%), $B2_3$ (2.78%) and D_2 (8.33%).

Stellenbosch University http://scholar.sun.ac.za

It was found by a study in 2008 (Gordon *et al*., 2008) that *E. coli* strains that were grouped in phylogenetic subgroup A_0, were not correctly classified when confirming genotype assignment using multi-locus sequence typing (MLST). When using MLST, it was observed that a proportion of the strains assigned to phylogenetic group A, should have been assigned to Group B1. If the same was true for the data in this study, it would mean that more isolates would have been assigned to group B1, being the group which is able to flourish in the environment. This would also have helped confirm the premise that most of the *E. coli* strains found in groundwater may in fact be purely environmental strains which naturally occur in the water systems.

Environmental

In this study, 'environmental strains' have been defined as *E. coli* strains which are found in rivers before any potential contamination sources sites have been able to affect the water quality. These strains therefore are those which are considered to be naturally present in water systems, hence they do not come from a particular contamination source should be disregarded when determining contamination. The isolates classed as 'environmental' were all from the Plank-0 site and were seen as environmental strains because, at this point in the Plankenburg River, the river has not yet passed through any informal settlements, industrial areas or demarcated farm lands. This means that the *E. coli* strains present in these water samples were most probably environmental in origin. When investigating the population structure of the 10 *E. coli* strains from this sample site (Table 10), it was found that most (50%) of the strains were characterised as phylogenetic subgroup $B2_3$. When taking the work of Walk *et al*. (2008) into consideration, it was thought that environmental strains would fall under phylogenetic group B1, due to this group's ability to survive outside their host's gut.

The conflicting results in this study may be explained by the type of environment studied, as the geographic location of this study differed from that of Walk *et al*. (2008). Another possible explanation is that although strains in phylogenetic group B1 are able to survive easily outside of their host, this does not mean that they are environmental in origin. They could merely be strains from faecal contamination which survive better in the environment than other strains. It may also be that strains in subgroup $B2_3$, which were dominant in the environmental population, are environmental strains. In other words, strains in phylogenetic group B1 may originally come from another source, but still survive well in the environment away from their host.

Stellenbosch University http://scholar.sun.ac.za

No isolates were characterised as members of groups A_0, B1, $B2_2$ and D (Table 10). This corresponds to the findings of Gordon & Cowling (2003) who concluded that main phylogenetic group B2 is a 'specialist' group, and that sources which have a high prevalence of strains representing group B2 are less likely to be associated with strains which are part of other phylogenetic groups.

In this study, it was concluded that it is not possible to isolate a group of *E. coli* strains from the environment and say that they are purely environmental isolates that have originated in the environment. This is because there is so much cross contamination as well as historical practices which could have led to the introduction of *E. coli* into the river systems and that they have now adapted and become part of the normal population structure. In other words, it may be speculated that no natural water source is free of some kind of contamination and that no *E. coli* strains are naturally present in natural water ways. All *E. coli* strains present in natural water systems have originated from another contamination source, such as faecal contamination, agricultural or industrial practises.

Escherichia coli population structures across all sample sites (types)

Table 10 shows the distribution of the *Escherichia coli* isolates from all sample sites across the seven phylogenetic subgroups. From this information, it can be concluded that pigs and humans have a similar distribution of phylogenetic subgroups within their respective sample groups, resulting in a similar population structure across both contamination sources.

One explanation for this occurrence is the similar dietary requirements of pigs and humans, as well as both being omnivorous and monogastric vertebrates (Apajalahti, 2005). As *E. coli*'s natural habitat is the intestines of warm blooded animals, it was expected that dietary requirements of the host would play a role in the *E. coli* population recovered from a particular source. In a study conducted by Carlos *et al*. (2010) it was found that dietary requirements and type of digestive tracks do play a role in determining the *E. coli* population structure.

Isolates from the dairy had a very different distribution when compared to that of humans and pigs. This is visualised in Fig. 2, where the majority (10/14) of the bovine isolates are found in phylogenetic subgroup B1, which also correlates well with results reported by Carlos *et al.* (2010). The other contamination source sites did not exhibit such a large group of B1 isolates, but when further investigating the irrigation water sampled it was found that both ground and surface waters contained mostly B1 strains (16/36 and 16/27 respectively). This may be useful when

Stellenbosch University http://scholar.sun.ac.za

undertaking source tracking of *E. coli* strains, as this information allows for linking a population attribute from irrigation to contamination water.

Microbial source tracking

Background

Microbial source tracking (MST) is often used to trace where faecal contamination originates (Meays *et al*., 2004). As there is no standard MST detection method for MST (Meays *et al*., 2004) various molecular and biochemical methods have been used to track the natural host of faecal bacteria, which in the case of this study was specifically *E. coli*. When using molecular methods, only selected microorganisms are tested for and these are chosen as they give good host discrimination. An ideal microorganism to use for MST should have similar properties to that of an indicator organism; such that it should have similar survival characteristics to the pathogen it is indicative of, be non-pathogenic and easily and rapidly detectable and enumerated (Scott *et al*., 2002). In addition to the properties mentioned, microorganisms used for MST ought also to have some way of discriminating between hosts (Farber, 1996). As these are merely guidelines, organisms used for MST do not always conform to all these characteristics, but when screening for a new microorganism to use for MST these guidelines may be helpful (Meays *et al*., 2004). In literature, it has been shown that bacteroidetes, enterococcus or faecal viruses specific to the presumptive host can be successfully used during MST (Meays *et al.,* 2004).

In this study, the MALDI Biotyper system as well as phylogenetic grouping was used to create a set of predefined attributes for each isolate. These attributes allowed for the division of *E. coli* strains using Biotyper 3.0 (Bruker, Germany) into PCA dendrograms based on the MALDI Biotyper data as well as phylogenetic groups. The attributes were also used to see whether *E. coli* isolated from irrigation water can be traced back to its original contamination source. *Escherichia coli* does not fit all the criteria for an ideal MST organism as it may be pathogenic, but as these criteria are just guidelines, and because *E. coli* is often used as an indicator organism (Ashbolt *et al*., 2001; LeClerc *et al*., 2001) its ability to be applied to MST was investigated further in this study.

In this study, the initial attribute sets were compared, but it was found that it was not possible to link an individual strain in irrigation water to a probable contamination source. From this it was therefore concluded that *E. coli,* when investigating individual strains, gives little information which can be used to conclude geographical structure or host discrimination. Other work where phylogenetic groupings were studied (Gordon, 2001; Carlos *et al*., 2010) recorded

Stellenbosch University http://scholar.sun.ac.za

similar conclusions that only a small percentage of genetic diversity observed in the *E. coli* genome between strains can be accounted for by geographical location and their preferred host.

Following these findings, Carlos *et al*. (2010) concluded that instead of expecting a single phylogenetic group to represent a particular source, the population structure of each contamination source needs to be taken into account and compared. The distribution of particular genes can therefore be used to create a population structure which is comparable with the structure of an *E. coli* population from another source. Thus in this study, the clustering resulting from the MALDI Biotyper data was then combined with the phylogenetic groupings and the population structure of each cluster and source type was compared.

MST on the basis of phylogenetic groupings

Tables 11 and 12 show the distribution (just presence/absence) of the isolates from irrigation and contamination sample sites respectively, across the seven phylogenetic subgroups. It was noted that the main phylogenetic groups B2 and D usually occur together, with the exception of isolates from the wine effluent as well as Plank-3. In all other examples, if isolates from either group D (D_1 or D_2) or B2 ($B2_2$ or $B2_3$) are present, the other is also present (Tables 11 and 12). These groups have been noted in a previous study (Johnson *et al.*, 2001) to contain more pathogenic strains than the main phylogenetic groups A (A_0 or A_1) and B1. The irrigation sites which contain strains from groups B2 and D are Mosselbank River and Borehole D. This pattern does however not bear a resemblance to any of the distribution patterns of the *E. coli* strains isolated from contamination source sites (Table 12). As a result, no concrete conclusions could be made regarding the contamination sources which contribute to the *E. coli* present in these three sites (Plank-0, Mosselbank River and Borehole D). It was assumed that strains from surface and groundwater will come from a variety of sources due to the ease with which surface water gets influenced by outside factors, and therefore a direct match to one of the contamination sites was unlikely.

It can however be seen that the distribution of isolates from the irrigation water (both ground and surface waters) was comparable to both the distributions of the isolates from the piggery and human source sites, as well as those isolated from the dairy (Fig. 3). It can therefore be speculated that due to the presence of phylogenetic subgroups A_0 (11/36 = 30.56%) and B1 (16/36 = 44.44%) in the isolates from groundwater samples, the contamination sources playing a role were most likely a combination of the dairy and piggery (Fig. 3). This is because the high prevalence of phylogenetic subgroup B1 in the groundwater could have come from the high

Stellenbosch University http://scholar.sun.ac.za

prevalence of phylogenetic group B1 in cows, while the high prevalence of phylogenetic subgroup A_0 could be linked to the high prevalence of A_0 in the isolates from the piggery. The high prevalence (11/36 = 30.56%) of phylogenetic subgroup A_0, and a lower frequency of A_1 (5/36 = 13.89%) in the groundwater population (Table 10), could also mean that the A_0 contamination did not come from pigs (Tables 10 and 12), but rather from an alternative/unknown source, as the piggery population also showed high prevalence of phylogenetic subgroup A_1.

The *E. coli* population structure observed in surface water (Fig. 3) shows prevalence for only one phylogenetic subgroup namely B1. This means that the most likely faecal contamination source is of bovine origin, but as there are many population attributes which are not shared by surface water strains and strains isolated from the dairy, there may be an alternative source which is contributing to the surface water contamination. It was therefore concluded that more than just the investigated contamination sources are likely to be contributing factors, and that further investigations need to be done in order to confirm this statement.

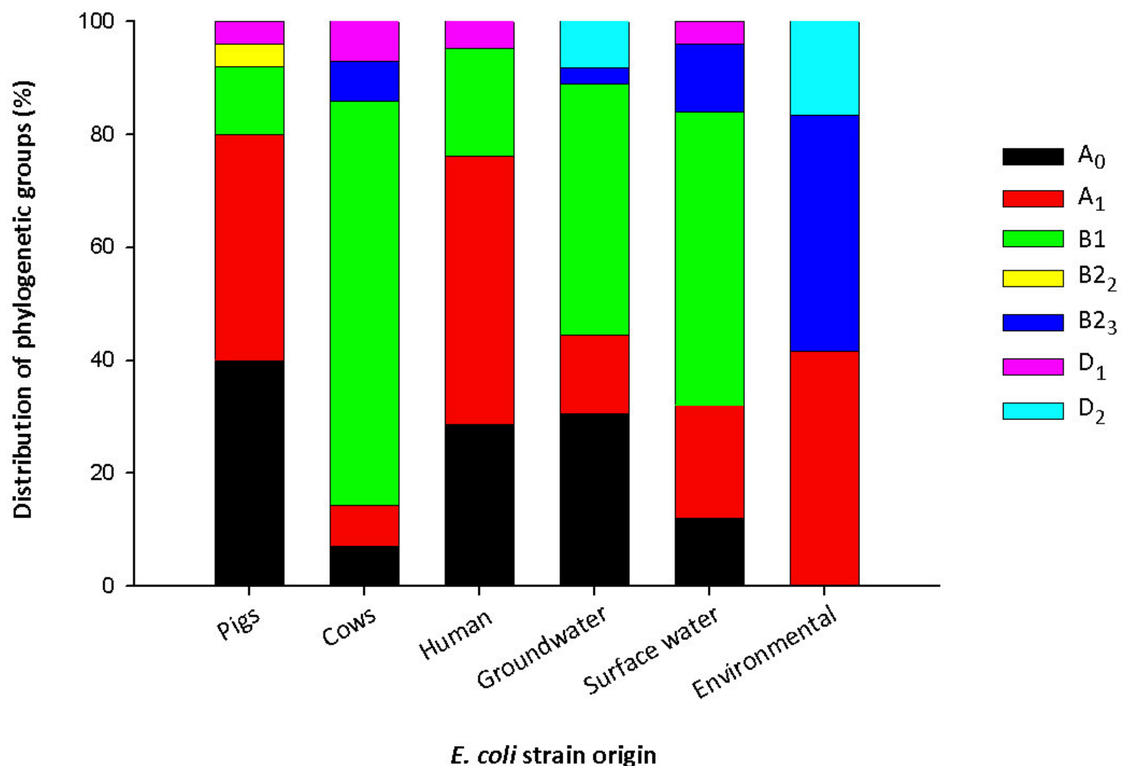

Figure 3. Distribution of the *Escherichia coli* phylogenetic subgroups within each group of isolates from various origins

Table 11. Detailed distribution (presence/absence) of *Escherichia coli* phylogenetic subgroups observed in the irrigation sites analysed

phylogenetic subgroup	Surface water							Groundwater		
	Winery Effluent	Veldwagters River	Plank-3	Mosselbank River	Olifants River	Berg-2	Dam	Spring C	Borehole A	Borehole D
A_0	✓	✓	✓	-	✓	-	✓	-	✓	✓
A_1	✓	✓	✓	-	✓	✓	-	-	✓	-
B1	✓	-	✓	✓	✓	✓	-	✓	✓	✓
$B2_2$	-	-	-	-	-	-	-	-	-	-
$B2_3$	-	-	✓	✓	-	-	-	-	-	✓
D_1	✓	-	-	✓	-	-	-	-	-	-
D_2	-	-	-	-	-	-	-	-	-	✓

Table 12. Distribution of *Escherichia coli* phylogenetic subgroups observed in the contamination sources sites analysed

phylogenetic subgroup	Plank-0	Plank-1	Veldwagters River	Winery Effluent	Dairy	Piggery
A_0	-	-	✓	✓	✓	✓
A_1	✓	✓	✓	✓	✓	✓
B1	-	✓	-	✓	✓	✓
$B2_2$	-	-	-	-	-	✓
$B2_3$	✓	-	-	-	✓	-
D_1	-	-	-	✓	✓	✓
D_2	✓	-	-	-	-	-

Stellenbosch University http://scholar.sun.ac.za

Three of the five contamination source sites (winery effluent, dairy and piggery) (Table 12) contained strains from phylogenetic subgroup D1, while it was absent in all but one (Mosselbank River) of the irrigation water (Table 11). As *E. coli* grouped in phylogenetic main groups D and B2 which are more likely to contain virulence factors than strains of groups A or B1 (Johnson *et al*., 2001), it can be concluded that even though contamination source sites were more likely to contain pathogenic *E. coli*, it can be seen that members of phylogenetic subgroup D1 were not readily carried over. This may be explained when considering that *E. coli* strains in main phylogenetic group D are seen as 'specialist strains' meaning that they are specific to a particular environmental niche (Walk *et al*., 2007). This can then lead to the conclusion that isolates in this phylogenetic group are less likely to be able to survive when removed from their preferred host environment. The implications of this conclusion are that isolates which are more likely to be pathogenic, such as phylogenetic group D (Johnson *et al*., 2001), did not survive well outside their natural host environment as they were not often isolated from the environment (Table 10).

When studying the population structure of the *E. coli* isolated from each sample site (Figs. 4 and 5), some linking can be made between irrigation sites and contamination sources. The population structure observed at Borehole A resembles that of the *E. coli* isolated from the winery effluent (Figs. 4 and 5). This leads to the conclusion that the groundwater, at Borehole A, might have been contaminated by either human sewage, industrial waste or a combination of the two. This is because of the factors which influence the contamination composition of effluent collected from the winery. The winery effluent has an additional phylogenetic group within the population structure (D_1), but as the rest of the structure is very similar to that of Borehole A these two sites could potentially be linked. As there is no physical evidence of a link, it was assumed that the similarities in the *E. coli* population structures from both sites (Borehole A and the winery effluent) was due to similar contamination sources having an influence on the *E. coli* present at the sites.

In addition to linking the population structures of *E. coli* from Borehole A and the winery effluent, the winery effluent can further be linked to the *E. coli* populations from the piggery as well as the Olifants River. This population structure, comprising mainly of phylogenetic subgroups A_0, A_1 and B1 also seems to be the most common basis for an *E. coli* population structure (Figs. 4 and 5). As phylogenetic group B1 is known to survive well in the environment (Walk *et al.*, 2007), this may explain the prevalence of this group in the environment. Subgroups A0 and A1 are both considered generalist groups (Gordon & Cowling, 2003) which have been linked to a variety of contamination sources. This may also strengthen the premise that environmental water contamination comes from a number of sources.

Stellenbosch University http://scholar.sun.ac.za

The *E. coli* population structures observed at Berg-2 and Plank-1 closely resemble one another (Figs. 4 and 5), which makes the most likely source for contamination of the Berg River human pollution. Plank-1 is situated downstream from an informal settlement which is likely to have insufficient infrastructure and sanitary facilities. This means that the water from this river may be used for washing, sanitation and recreational purposes, which could all play a role in introducing *E. coli* into the water system. Some people in this informal settlement also have dogs or other animals as pets which have access to the river water and may therefore also contribute to contamination. Runoff from the areas surrounding the river during the rainy season would also increase the pollution and potentially contribute to the *E. coli* population. It was therefore concluded that a similar informal settlement could be a main factor contributing to the contamination load at the Berg-2 sample site. After further investigation it was found that there are informal settlements upstream from the Berg-2 sample site. This therefore confirmed that source tracking by means of population structure of *E. coli* strains could be an accurate and a viable way of conducting MST.

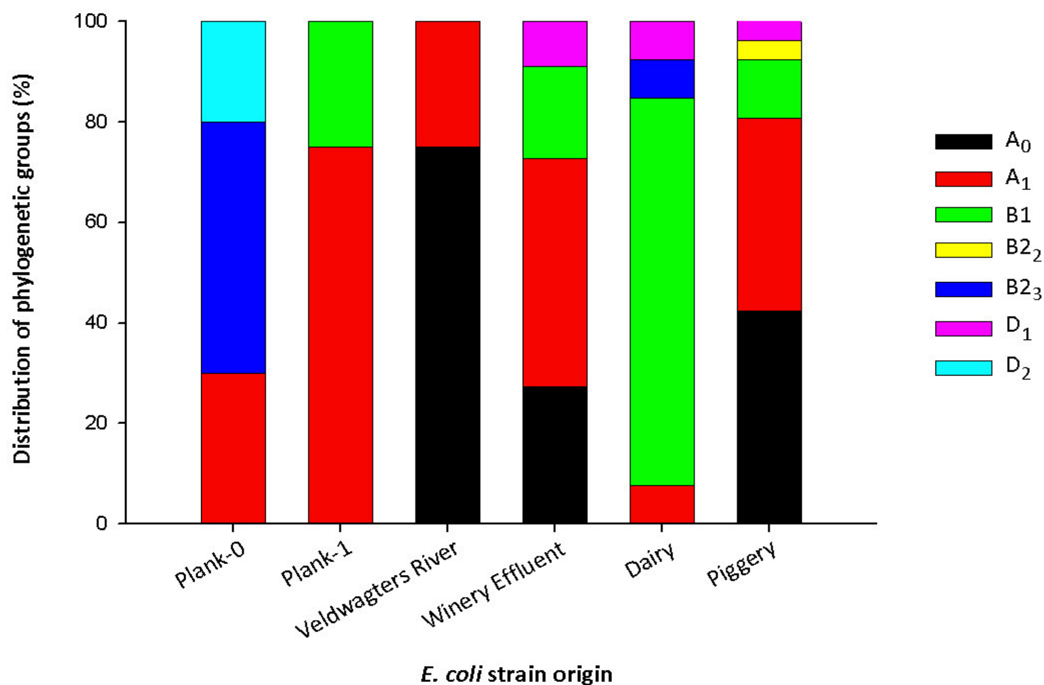

Figure 4. Population structure of *E. coli* isolated from the contamination source sites

Stellenbosch University http://scholar.sun.ac.za

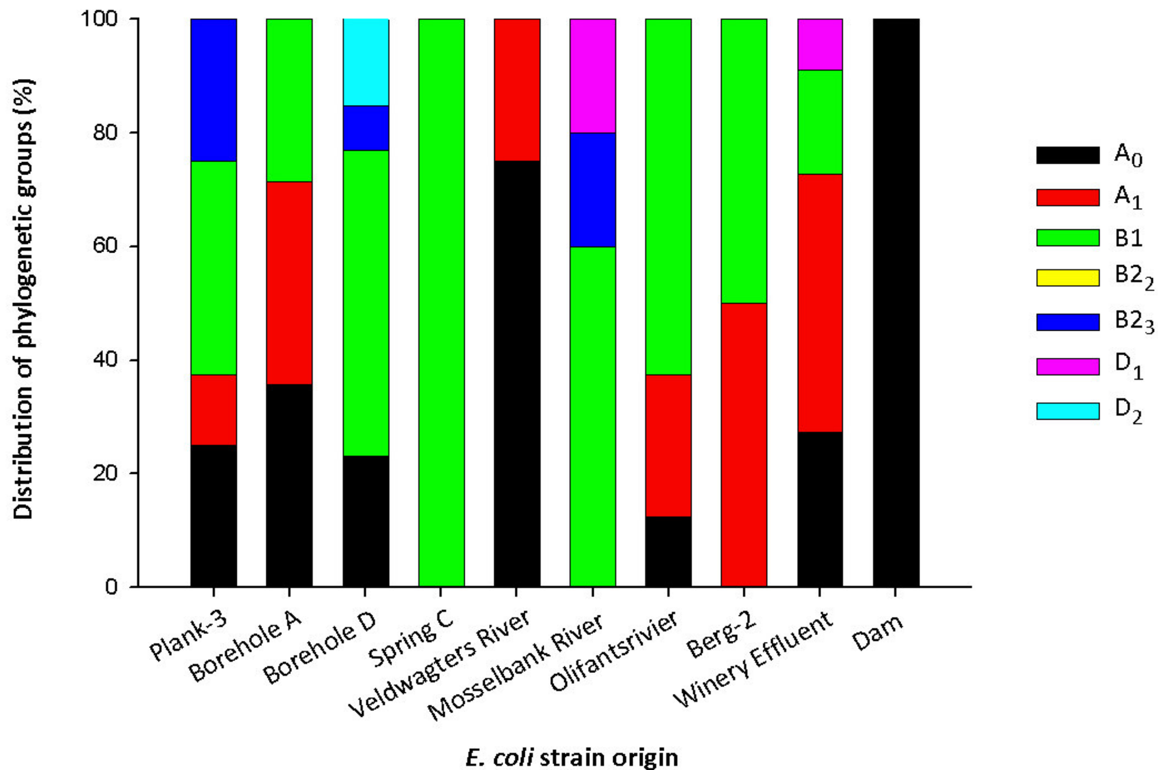

Figure 5. Population structure of *E. coli* isolated from the irrigation sites

The one strain found in this study, which was unique in its phylogeny was strain B13.3 as it was the only strain with the characteristics of phylogenetic subgroup $B2_2$ (Appendix A). B13.3 was isolated from a secondary dam at the piggery. The water sampled at this site first undergoes a separation step (separates solids from liquids), fermentation step and then reaches a larger dam (sample site) which, when it overflows, reaches a marsh which signals the start of the Plankenburg River (Morris, J., 2012, Farm manager, Department Agricultural Sciences, Stellenbosch University, Stellenbosch, South Africa. Personal communication). It is therefore possible that this strain (B13.3) is in fact not of porcine origin at all, and could have come from a bird or another animal in contact with the dam.

Another explanation for the single strain in phylogenetic subgroup $B2_2$ could be that strains in phylogenetic subgroup $B2_2$ are present in small numbers in the environment. This is supported by the study done by Carlos *et al*. (2010) who found that subgroup $B2_2$ was represented by only 3.32% (8 of 241) of their total isolates. In this study $B2_2$ was represented by <1% (1 of 134) of the total *E. coli* strains, which could be an indication of the phylogenetic subgroup's low presence in the environment. According to Gordon (2001) and Gordon & Cowling (2003), if the phylogenetic

Stellenbosch University http://scholar.sun.ac.za

group is not abundant in the environment, standard microbial methods would also not necessarily allow for isolation of that particular group of strains.

MST using the MALDI Biotyper system

In the previous chapter of this thesis, the MALDI Biotyper (Bruker, Germany) and API 20E (BioMérieux, France) systems' identification abilities were compared. It was concluded that the MALDI Biotyper was better suited for identifying environmental strains with consistent accuracy, due to the increased objectivity of the system. The API 20E data was therefore not used for MST. A PCA dendrogram (Fig. 6) (Bruker, Germany) created from the resultant spectra from the MALDI Biotyper system was used instead.

In the PCA dendrogram (Fig. 6), based on hierarchal sorting of the MALDI spectra, showing the clustering of the *E. coli* strains it can be seen that the sorting of the spectra resulted in four main clusters. Cluster 1 (red) is the largest cluster containing 57 *E. coli* strains, when not counting reference strains and duplicates, Cluster 3 (blue) is the smallest cluster; containing only eight strains, not inclusive of reference strains (Table 13). In Table 13, the shaded cells are *E. coli* strains which originated from contamination sites. The outlined cells are *E. coli* strains which originate from sample sites which are seen as both contamination sources as well as irrigation sites (Veldwagters River and the winery effluent), and the remaining cells are *E. coli* strains which are from irrigation sites.

The four clusters from the PCA dendrogram were then combined with the phylogenetic grouping data and the phylogenetic distribution within each cluster was determined (Table 14). As previously found, linking of individual strains from irrigation water to contamination source is not possible and as a result phylogenetic population structures within each cluster was determined for all the contamination sources as well as the irrigation sites (Figs. 7, 8 and 9).

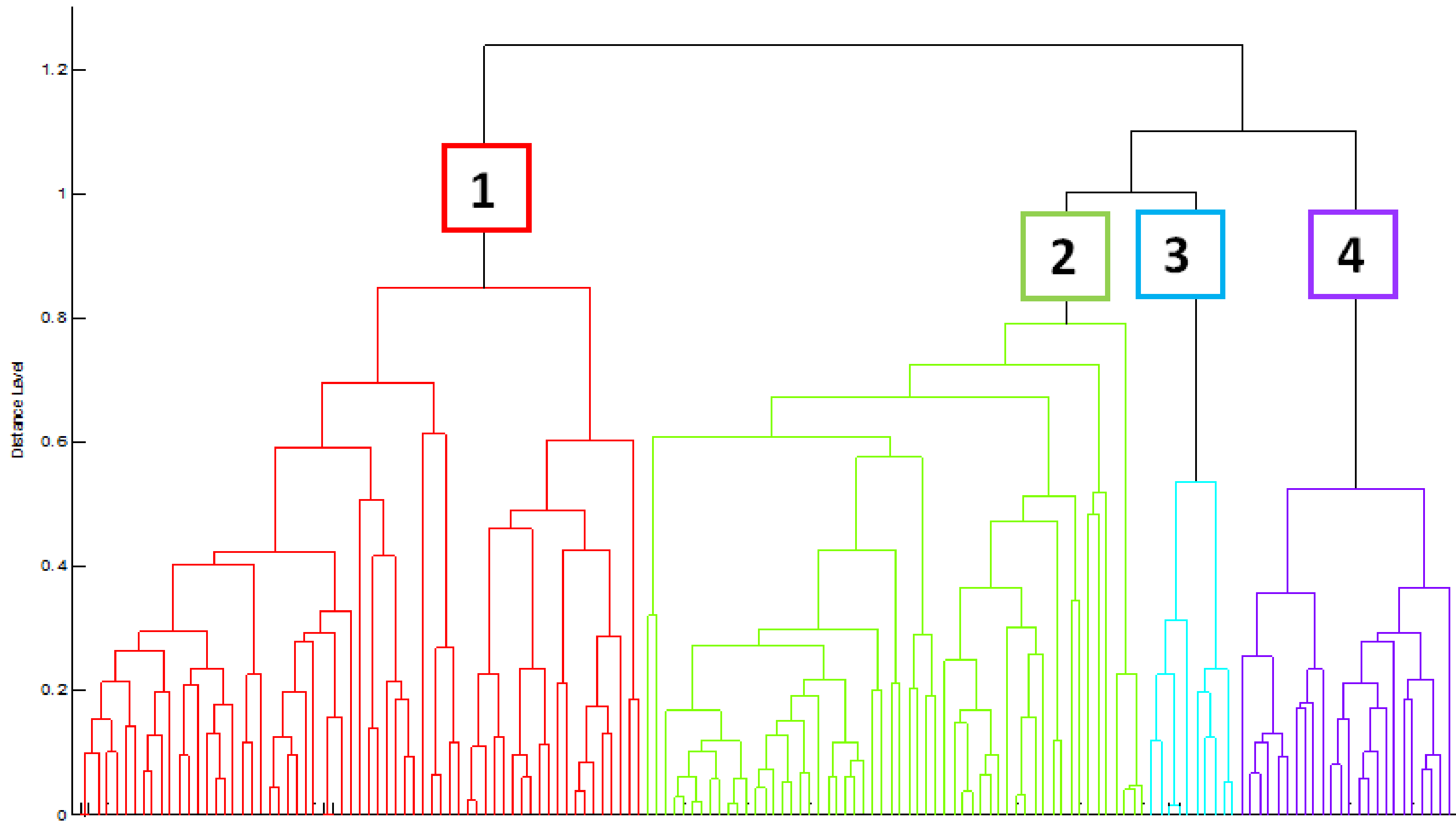

Figure 6. PCA dendrogram showing clustering of *E. coli* strains isolated from contamination sources and irrigation sites

Stellenbosch University http://scholar.sun.ac.za

Table 13. *Escherichia coli* strain distribution as determined by MALDI-TOF MS data

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 1 ...continued	Cluster 2 ...continued
A11.4	A11.1	ATCC 35218	A11.3	E11.2	G21.2
A11.5	A11.2	ATCC 35218	B13.4	E11.3	G21.3
A21.2	A21.1	E11.1	C21.2	E11.4	G21.4
A21.3	A22.1	E11.5	C21.3	E12.1	G22.1
A21.4	A22.2	E21.1	C21.5	E12.2	H11.4
A22.3	A22.5	E21.3	D12.2	E21.4	J11.1
A22.4	B13.1	F21.5	D21.1	E21.5	J11.3
ATCC 25922	B13.3	J11.2	D21.3	E22.3	J11.4
ATCC 25922	B21.2	J21.1	E12.4	E22.4	J21.2
B11.1	B22.2	L21.4	E22.2	E22.5	J21.3
B11.2	B22.3		G21.5	EC 157	J21.4
B11.3	B22.4		H23.1	EC 157	J21.5
B11.4	B23.1		H23.2	EC 58	L11.1
B12.2	B23.2		H23.3	EC 58	L11.3
B12.3	B23.3		H23.4	F11.3	L11.5
B12.4	B23.4		H23.5	F11.4	L21.1
B13.2	B23.5		H24.1	G11.1	L21.5
B21.1	D11.1		K11.1	G21.1	M11.1
B21.3	D11.2		K11.3	H11.1	M11.2
B21.4	D11.3		K11.4	H11.2	M11.3
B21.5	D11.4		L11.6	H11.3	M11.4
B22.1	D12.3		M21.1	H11.5	M21.4
B22.5	D12.4		M21.2	J11.5	M21.5
C21.1	D21.4		M21.3	K11.5	M22.1
C21.4	D21.5			L11.2	M22.4
D12.1	D24.5			L11.4	
D21.2	E22.1			L21.2	
D24.1	F11.2			L21.3	
D24.2	F21.1			M22.2	
D24.3	F21.3			M22.3	
D24.4	F21.4			M22.5	

Shaded cells show contamination sites and bordered cells show sites which overlap contamination and irrigation sites. The remaining cells are strains from irrigation sites.

Table 14. Distribution of the *Escherichia coli* phylogenetic subgroups within each cluster from the PCA dendrogram

Phylogenetic subgroup	Cluster 1	Cluster 2	Cluster 3	Cluster 4
A_0	16	10	0	4
A_1	16	14	0	6
B1	18	19	2	10
$B2_2$	0	1	0	0
$B2_3$	4	4	3	0
D_1	3	1	1	0
D_2	0	3	2	0
Total	57	52	8	20

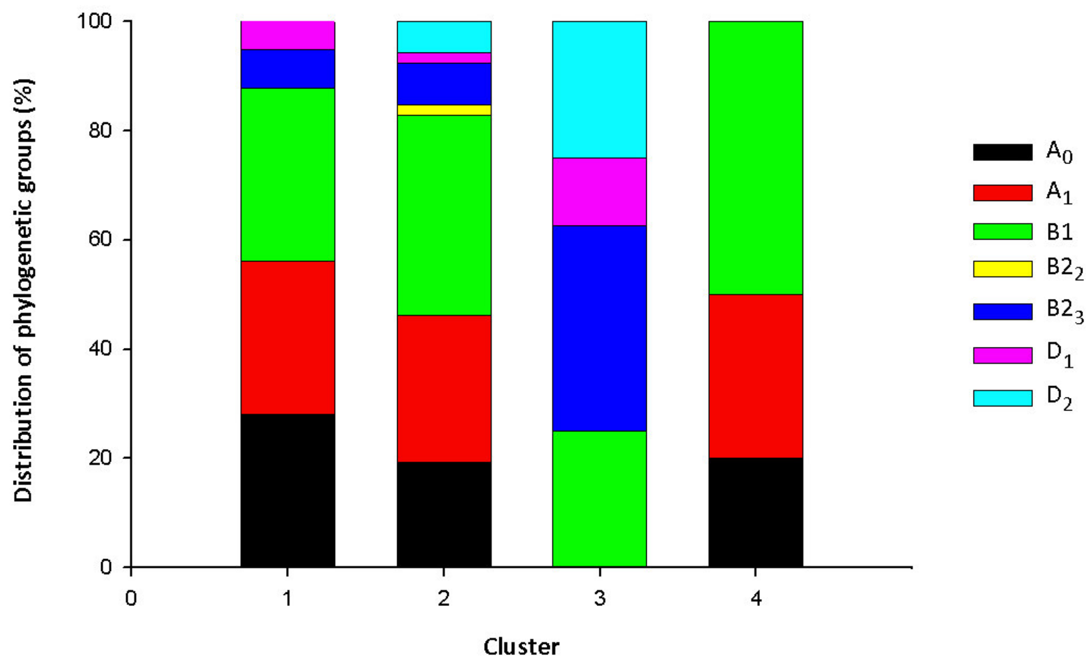

Figure 7. Phylogenetic population structure of *E. coli* isolated from both irrigation sites and contamination sources, grouped according to PCA dendrogram clusters

When looking at Clusters 1 and 2 (Fig. 7), it can be seen that distributions of phylogenetic subgroups A_0, A_1, B1, $B2_3$ and D_1 within both clusters are similar. Not much more than this could be seen from the combined groupings. However, when the dendrogram clusters are further divided into strains from contamination (Fig. 8) and irrigation (Fig. 9) sites it becomes easier to link irrigation water to contamination sources. Cluster 1 contamination and irrigation populations are both similar in phylogenetic population structure (Figs. 8 and 9). The population structure of this cluster also closely resembles both the population structures of *E. coli* said to be bovine or human in origin (Fig. 3). When looking at the isolates found in cluster 1, it can be seen that there are

Stellenbosch University http://scholar.sun.ac.za

11 groundwater and 23 surface water strains (Table 13). This could therefore indicate that human pollution and any agricultural practices involving cattle may have a notable influence on both surface and ground water. In this case there was however no link to a particular geographical area or type of irrigation sample site, but merely that a combination of *E. coli* isolated from different ground and surface waters have similar attributes to *E. coli* populations isolated from humans and cows.

The population structures of cluster 4 of both the contamination source and irrigation site populations strongly resembled one another (Figs. 8 and 9). The combination of phylogenetic subgroups A_0, A_1, and B1 seen in cluster 4 is similar to the population structure of *E. coli* isolated from Borehole A and the Olifants River (Fig. 5), as well as the contamination believed to be human in origin (Fig. 3). For this it was concluded that the *E. coli* population in Cluster 4 is most likely from a human contamination source, even though very few contamination source strains were grouped in this cluster (Table 13).

Irrigation sample Clusters 2 and 3 do not resemble any of the contamination source sites' population structure (Fig. 8), but Cluster 3's population structure does look similar to that of strains from the Mosselbank River (Fig. 5). It was therefore concluded that the *E. coli* populations from irrigation water in Clusters 3 and 4 probably have a contamination source polluting the water which was not investigated in this study. This could mean that any number of other animals could be responsible for the contamination of these water sources. When looking at the strains which are grouped together in cluster 3, it can be seen that Cluster 3 contains only one strain which comes from a contamination source (winery effluent) (Table 13) and as a result, it can be said that isolates in this cluster were most probably environmental strains, or strains which had come from faecal contamination other than those which were investigated. The strains in this cluster were mainly isolated from Plank-0, which had been selected as an 'environmental' source as it has not yet passed through industrial areas, informal settlements or farming land.

As pigs, humans and cattle were the only contamination reservoir hosts investigated, it also made the linking of *E. coli* populations from irrigation water to contamination source more challenging. To address this problem, a larger study needs to be done with a more even distribution of samples coming from a more diverse set of sample sites, as well as additional contamination sources being investigated. Contamination sources representing *E. coli* populations from water birds, chickens, ruminants and horses would also be beneficial for more accurate MST.

Stellenbosch University http://scholar.sun.ac.za

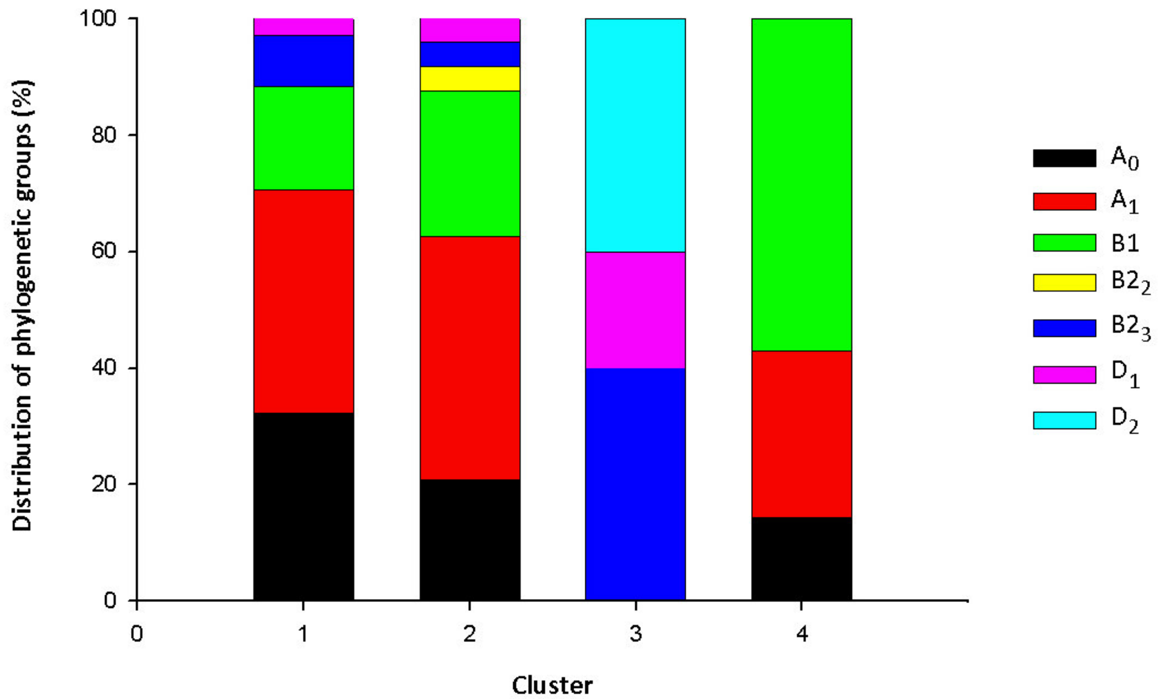

Figure 8. Phylogenetic population structure of *E. coli* isolated from contamination sources, grouped according to PCA dendrogram clusters

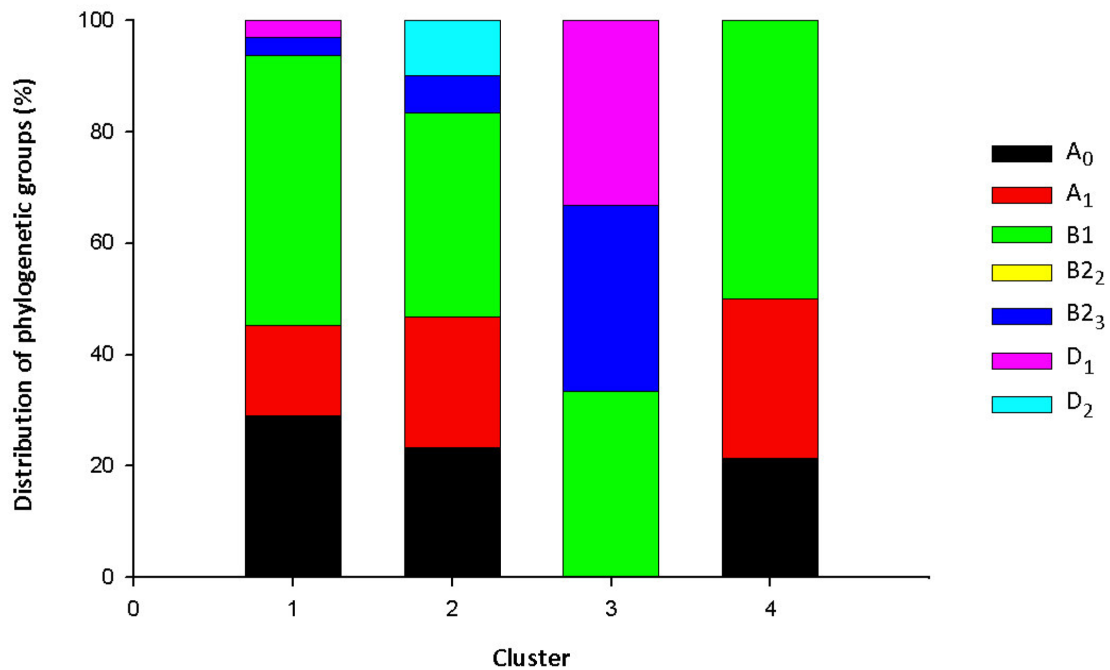

Figure 9. Phylogenetic population structure of *E. coli* isolated from irrigation water, grouped according to PCA dendrogram clusters

Stellenbosch University http://scholar.sun.ac.za

Potential risk assessment and pathogen determination

It has been shown that bacterial contamination present in irrigation water may be carried over to fresh produce and affect the health of consumers (Beuchat & Ryu, 1997; Solomon *et al*., 2002; Okafo *et al*., 2003; Islam *et al*., 2004; Van Blommestein, 2012). Thus the potential risk associated with using irrigation water from natural water sources was determined by identifying pathogenic strains among the 153 *E. coli* isolates. This was done as a means to determine the associated risk with using contaminated irrigation water.

Multiplex PCR was used to identify the presence of intestinal pathogenic *E. coli* (InPEC) strains, and it was found that of the 153 *E. coli* strains, only two showed amplification of virulence genes (Appendix A). Both these strains were confirmed to be enteropathogenic *E. coli* (EPEC) strains based on the presence of the *eaeA* gene(Lanes 2 and 3 in Fig. 10), which is responsible for the formation of attaching and effacing lesions typical in infection (Jerse *et al*., 1990; Bhunia, 2008). These strains (A11.3 and F11.3) were both isolated from different irrigation sites used for irrigation of fresh produce.

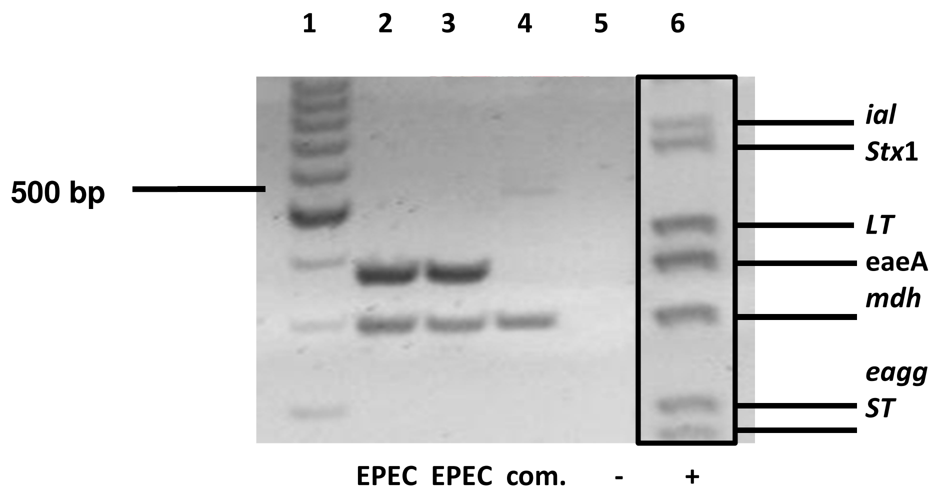

Figure 10. Visual representation of amplified pathogenic genes when viewed under UV light (Com. = commensal strain)

Strain A11.3 was isolated from groundwater (Borehole A) which was unexpected because in Chapter 3 of this thesis, borehole water was concluded to be the safest alternative to use for irrigation as a result of the lower levels of *E. coli* and total coliforms present. Many farmers choose to use borehole water, without prior treatment, as a result of the theory that borehole water is ‘clean’ (Bezuidenhout *et al*., 2011). This incidence of contaminated borehole water is of concern, as the source of the contamination cannot be identified. No obvious contamination

Stellenbosch University http://scholar.sun.ac.za

sources were in direct contact with the water of Borehole A, which indicates that the contamination must come from elsewhere.

The other EPEC strain (F11.3) was isolated from the irrigation site, Plank-3. The water at this site is extracted from the river and redirected into an irrigation furrow, from where it is used directly for irrigation of fresh produce. This water therefore poses the same risk to consumers as using the contaminated water from Borehole A. In a previous study done on the Plankenburg River, including the Plank-3 site (Van Blommestein, 2012), two EPEC strains were isolated on separate occasions, showing that there are indicators that a consistent source of EPEC contamination is finding its way into the Plankenburg River. It was therefore speculated that this contamination came from the informal settlement which is situated next to the river, upstream from where the EPEC strains were isolated. No EPEC strains were however isolated from the Plank-1 site which is just downstream from the informal settlement, and where it was expected that contamination from the settlement would be prevalent. The lack of EPEC strains isolated from other contamination sites, and Plank-1 in particular may be due to seasonal fluctuations, random isolation when using the Harrison Disk Method (Harrigan & McCance, 1976) or possibly a misrepresentation of isolates which occur in low numbers in the environment (Gordon, 2001; Gordon & Cowling, 2003).

It is not possible to assign a risk factor to the use of this water as there is no defined success rate of *E. coli* carry-over from irrigation water to fresh produce. It can however be said that if one EPEC was isolated from a river, it can be assumed that more EPEC strains may be present in the water as only a one litre sample was tested. To confirm these results, numerous samples need to be taken from the same sample site and a more selective medium must be used to isolate a larger number of *E. coli* strains from a single sample site. This might give a better estimation of the level of EPEC contamination present in the Plankenburg River and borehole A1. Borehole A is especially a concern as the source of the EPEC isolate was unidentified.

When comparing the phylogeny of the two EPEC strains, it was found that strain A11.3 was classified as a member of the phylogenetic group B1, while F11.3 was categorised as subgroup $B2_3$. Intestinal pathogenic *E. coli* strains are most commonly found in groups A, B1 and D according to Pupo *et al*. (1997) but *E. coli* strains in phylo-groups B2 and D, according to Johnson *et al*. (2001), are more likely to contain extraintestinal virulence genes, than those in phylo-groups A and B1. This also means that these two strains do not share either of the genetic markers (*chuA* and *yjaA*) or the DNA fragment (TspE4.C2), as A11.3 contains only the DNA fragment while F11.3 contains only the genetic markers. This goes to show just how different *E. coli* can be; even when two

Stellenbosch University http://scholar.sun.ac.za

strains will ultimately cause the same disease, they can be from different phylogenetic groups and have different biochemical profiles. When comparing these two isolates using MALDI-TOF MS, they do not have a similar spectra resulting from their five most abundant ribosomal proteins, as they are grouped in two different PCA clusters (Table 13 and Fig. 6).

In a previous study, where the Plankenburg River was also sampled, and *E. coli* isolated (Van Blommestein, 2012), it was reported that three EPEC strains were isolated from the Plankenburg River, all of which belonged to phylogenetic subgroup $B2_3$. This could mean that either there is a continual source which introduces EPEC strains into the Plankenburg River, or that the EPEC strains in phylogenetic subgroup $B2_3$ have become well adapted to the water environment and are able to survive and multiply, even though subgroup $B2_3$ is not known to survive as well in the environment as phylogenetic group B1. The EPEC strain from the Borehole A (A11.3) was found to belong to phylogenetic group B1, which means that it may survive well in the environment. This could therefore result in horizontal gene transfers which facilitate the carry-over of the *eaeA* virulence gene from one *E. coli* strain to another.

It is also important to note that when *E. coli* is isolated from the environment, a representative sample is not always collected. It was reported by various studies (Gordon, 2001; Gordon & Cowling, 2003) that sampling of bacterial isolates from the environment is often biased and that only strains which are present in abundance are isolated. In a study by Gordon & Cowling (2003) it was concluded that only strains which are present in a capacity exceeding 1% of the bacterial population present would stand a chance of being isolated. This being said, the enumeration method used in this study (Colilert 18) may detect *E. coli* even if only one cell is present in the 100 mL water sample being tested, assuming that the single cell does not get out-competed during culturing on the defined substrate. Isolating a particular strain of *E. coli* which is present in low numbers is however much more problematic. As a result only the abundant strains are isolated using standard methods.

Relating the results above to this study, it could be that even though high levels of *E. coli* contamination may be present in the samples, it does not mean that a representative sample of the *E. coli* strains would have been isolated. It may therefore be implicit that strains which have been isolated in low numbers within this dataset are misrepresented and their abundance could be underestimated. As a result, the population structure of the *E. coli* isolated from the environment is a misrepresentation of the true *E. coli* population structure in the river. This was however not a real problem in the source tracking portion of this study, as all samples were analysed using the same microbial standard methods, and isolates were collected by one

Stellenbosch University http://scholar.sun.ac.za

individual, maintaining accuracy throughout the study. This was also confirmed by the use of reference strains when identifying isolates using the MALDI Biotyper system. Prevalence of less common strains was seen as a problem when the pathogenicity, and hence associated risk, was determined. This is because, according to Gordon (2001) and Gordon & Cowling (2003), the low prevalence (<2%) abundance of *E. coli* strains in the irrigation water is a misrepresentation of the true prevalence.

As a result, the potential misrepresentation of less abundant strains could mean that although only two pathogenic strains were found amongst the isolated *E. coli*, the potential risk may be greater than initially assumed. Although no solid proof is available for this statement, it was assumed that due to the lower abundance of pathogenic strains, in comparison to commensal strains, meant that pathogenic strains and thus their prevalence had been underestimated.

CONCLUSIONS

In total 153 *E. coli* strains were collected, 143 from irrigation sites and contamination sources in Stellenbosch and surrounding areas. An additional 10 strains were isolated from the Plankenburg site (Plank-0) which was taken as representing an uncontaminated environmental site.

It was found that linking a single *E. coli* strain from irrigation water to another strain from a contamination source based on phylogenetic grouping was unsuccessful, as the variation of different *E. coli* strains was too vast. It was however found in this study that linking an irrigation site with one or more contamination sources was possible when taking into consideration the total phylogenetic group variation present amongst the *E. coli* population isolated at each site. It was therefore concluded that the *E. coli* population structure, as determined by phylogenetic groupings, was better suited for source tracking. When taking the contamination sources into consideration, each site was found to have a distinct population structure. The isolates from the piggery, therefore porcine in origin, were largely represented by phylogenetic main group A (80%). Isolates from the dairy on the other hand were mainly (71.43%) assigned to phylogenetic group B1, and *E. coli* with direct human origins showed prevalence (50%) of phylogenetic subgroup A_1.

The irrigation water was divided into ground and surface water isolates, and their population structures were also found to differ, with prevalence of phylogenetic subgroup B1 (44.44% and 59.26%, respectively). This was expected, as isolates grouped in phylo-group B1 have been reported to be able to survive well in the environment (Walk *et al*., 2007). In addition to phylo-group B1, groundwater isolates also exhibited an additional major grouping of phylogenetic

Stellenbosch University http://scholar.sun.ac.za

subgroup A_0. This combination within the population structure gives groundwater sources a unique population structure.

It was concluded that the *E. coli* population structures in human and pig faecal contamination sources are similar and share characteristics while the population structure of *E. coli* isolates from the dairy source was different. Data from this study led to the conclusion and confirmation of previous studies' results (Gordon & Cowling, 2003; Carlos *et al*., 2010) that dietary requirements and gut properties play an important role in determining the population structure of the *E. coli* isolates present.

From the population structure data in this study it was furthermore concluded that the contamination sources most likely to play a role in contamination of groundwater are those with faecal matter from pigs and cattle. With surface water however, the most likely faecal contamination came from bovine sources, but as no other resemblances in the *E. coli* population structures from any one particular contamination site could be linked to that of surface water, it was concluded that more than just the investigated contamination sources play a role.

When dividing the *E. coli* isolated from contamination sources and irrigation sites into smaller 'populations' representing each individual sample site, further conclusions could be made to link particular contamination sources with irrigation sites. It was found that Borehole A and the winery effluent had similar population structures and it is possible that winery effluent could have been a contamination source of Borehole A. Borehole A does, however, not have any contact with the winery effluent sampled in this study, but it may be an indication that a contamination source, similar to the winery effluent influenced the population structure of Borehole A.

Data from this study showed that isolates from the piggery, the Olifants River and the winery effluent all had similar population structures. It was therefore concluded that the combination of human and winery contamination of the winery effluent could be a major contamination contributor of irrigation water.

Berg-2 and Plank-1 sites also showed similar population structures. Plank-1 is known to become contaminated after passing through an informal settlement, thus it was concluded that the most probable source of contamination of the Berg River is also human pollution from an informal settlement. Upon further investigation it was found that there is an informal settlement upstream from the Berg-2 sampling site. This showed that by using unique population structures, irrigation sites can be linked to 'probable' contamination sources.

The potential risk associated with using irrigation water from natural water sources was determined by identifying pathogenic strains among the 153 *E. coli* isolates. This was done as a

Stellenbosch University http://scholar.sun.ac.za

means to determine the associated risk with using contaminated irrigation water. It was found that a small portion (2/153 = 1.49%) of the *E. coli* from both irrigation water and contamination source sites contained virulence genes consistent with EPEC. One of these strains was isolated from surface water (Plank-3), while the other strain was isolated from groundwater (Borehole A). This contradicts the statement that groundwater is a safer option for irrigation water because of the lower *E. coli* and total coliform counts observed in this study. It was also clear that the *E. coli* and total coliform counts are not an indication of associated risk and cannot be used as an indicator of potential harm to consumers.

When comparing the phylogeny of the two EPEC strains, it was found that strain A11.3 was classified as a member of the phylogenetic group B1, while F11.3 was categorised as subgroup $B2_3$. In a previous study, where the Plankenburg River was also sampled and *E. coli* isolated (Van Blommestein, 2012), it was reported that three EPEC strains were isolated from the Plankenburg River, all of which belonged to phylogenetic subgroup $B2_3$. The implications of this are that EPEC strains which are in phylogenetic group B1 are able to survive well in the environment (Walk *et al*., 2007), which means that they survive, multiply and could therefore result in horizontal gene transfers which facilitates the carry-over of the *eaeA* virulence gene from one *E. coli* strain to another. Phylogenetic subgroup $B2_3$ on the other hand is not known to survive in the environment as well as group B1 (Walk *et al*., 2007), and therefore the results in this and previous studies shows that there is most likely a continual contamination source which contributes to the contamination in the Plankenburg River. Another explanation could be that these EPEC strains, in phylogenetic subgroup $B2_3$, have adapted in such a way that they are able to survive and multiply in a fluvial environment. This could also mean that the associated risk needs to be re-evaluated, as EPEC strains in phylogenetic group B1 may multiply quickly and survive for long periods of time and therefore increase the potential risk associated with using the water to irrigate fresh produce.

Data from this and other studies (Gordon, 2001; Gordon & Cowling, 2003) also showed that standard *E. coli* isolation methods must be taken into account when applying source tracking techniques, as they can be heavily biased. When sampling water which contains high coliform loads, strains which are present in the water in low numbers (<1% of total bacterial consortium) will not be isolated (Gordon, 2001 & Gordon & Cowling, 2003). In other words, strains which are reported at a low prevalence, such as the pathogenic EPEC strains in this study, could be more abundant in the population than perceived from the dataset. It can therefore be concluded even if no pathogenic strains are isolated from a water sample; the water sample may still not be as safe as anticipated. This is because the number of pathogenic *E. coli* in water systems is often

Stellenbosch University http://scholar.sun.ac.za

underestimated as a result of biased isolation methods. Natural water sources, both ground and surface waters which show coliform contamination, should be used with great caution.

REFERENCES

Ackermann, A. (2010). *Assessment of Microbial Loads of the Plankenburg and Berg Rivers and the Survival of Escherichia coli on Raw Vegetables under Laboratory Conditions*. MSc in Food Science Thesis. Stellenbosch University, Stellenbosch, South Africa.

Altahi, A.D. & Hassan, S.A. (2009). Bacterial quality of raw milk investigated by *Escherichia coli* and isolates analysis for specific virulence-gene markers. *Food Control*, **20,** 913-917.

Apajalahti, J. (2005). Comparative gut microflora, metabolic challenges, and potential opportunities. *The Journal of Applied Poultry Research*, **14**, 444-453.

Ashbolt, N.J., Grabow, W.O.K. & Snozzi, M. (2001). Indicators of microbial water quality. In: *Water Quality: Guidelines, Standards and Health*, Pp. 289-316. London: IWA Publishing.

Avery, L.M., Williams, A.P., Killham, K. & Jones, D.L. (2008). Survival of *Escherichia coli* O157:H7 in waters from lakes, rivers, puddles and animal-drinking troughs. *Science of the Total Environment*, **389**, 378-385.

Baldy-Chudzik, K., Mackiewics, P. & Stosik, M. (2008). Phylogenetic background, virulence gene profiles, and genomic diversity in commensal *Escherichia coli* isolated from ten mammal species living in one zoo. *Veterinary Microbiology*, **131**, 173-184.

Beuchat, L.R. & Ryu, J.H. (1997). Produce handling and processing practices. *Emerging Infectious Diseases*, **3**(4), 459-465.

Bezuidenhout, C.C. & The North-West University Team (2011). A scoping study on the environmental water (groundwater and surface water) quality and management in the North-West Province, South Africa. Water Research Commission (WRC) Report No. KV 278/11.

Bhunia, A.K. (2008). *Escherichia coli.* In: *Foodborne Microbial Pathogens: Mechanisms and Pathogenesis*. (edited by Helderman, D.R.). Pp. 183-200. New York: Springer.

Bingen, E., Picard, B., Brahimi, N., Mathy, S., Desjardins, P., Elion, J. & Denamur, E. (1998). Phylogenetic analysis of *Escherichia coli* strains causing neonatal meningitis suggests horizontal gene transfer from a predominant pool of highly virulent B2 group strain. *Journal of Infectious Diseases,* **177**, 642-650.

Bolton, F.J., Crozier, L. & Williamson, J.K. (1996). Isolation of *Escherichia coli* 0157 from raw meat products. *Letters in Applied Microbiology*, **23**, 317-321.

Carlos, C., Pires, M.M., Stoppe, N.C., Hachich, E.M., Sato, M.I.Z., Gomes, T.A.T., Amaral, L.A. & Ottoboni, L.M.M. (2010). *Escherichia coli* phylogenetic group determination and its application in the identification of the major animal source of fecal contamination. *BMC Microbiology*, **10**, 161-171.

Clermont, O., Bonacorsi, S. & Bingen, E. (2000). Rapid and Simple Determination of the *Escherichia coli* Phylogenetic Group. *Applied and Envivormental Microbiology*, **66**(10), 4555-4558.

Denamur, E., Denamur, E., Duriez, P., Brahimi, N., Bingen, E., Picard, J.P., Garcia, J.S. & Gouriou, S. (1999). The link between phylogeny and virulence in *Escherichia coli* extraintestinal infection. *Infection and Immunity*, **67**(2), 546-553.

Farber, J.M. (1996). An introduction to the hows and whys of molecular typing. *Journal of Food Protection*, **59**(10), 1091–1101.

Stellenbosch University http://scholar.sun.ac.za

Gordon, D.M. (2001). Mini review: Geographical structure and host specificity in bacteria and the implications for tracing the source of coliform contamination. *Microbiology*, **147**, 1079-1085.

Gordon, D.M., Clermont, O., Tolley, H. & Denamur, E. (2008). Assigning *Escherichia coli* strains to phylogenetic groups: multi-locus sequence typing versus the PCR triplex method. *Environmental Microbiology*, **10**(10), 2484-2496.

Gordon, D.M. & Cowling, A. (2003). The distribution and genetic structure of *Escherichia coli* in Australian vertebrates: host and geographic effects. *Microbiology*, **149**, 3575-3586.

Groisman, E.A. (2001). Principles of Bacterial Pathogenesis. Pp. 387-456. Academic Press. New York.

Harrigan, W.F. & McCance, M.E. (1976). Methods for the selection and examination of microbial colonies. In: *Laboratory Methods in Food and Dairy Microbiology*, (Edited by Harrigan, W.F. & McCance, M.E.), Pp. 47-49. London: Academic Press

Herzer, P.J., Inouye, S., Inouye, M. & Whittam, T.S. (1990). Phylogenetic distribution of branched RNA-linked multicopy single-stranded DNA among natural isolates of *Escherichia coli*. *Journal of Bacteriology*, **172**(11), 6175-6181.

Islam, M., Doyle, M.P., Phatak, S.C., Millner, P. & Jiang, X. (2004). Persistence of enterohemorrhagic *Escherichia coli* O157:H7 in soil and on leaf lettuce and parsley grown in fields treated with contaminated manure composts or irrigation water. *Journal of Food Protection*, **67**(7), 1365-1370.

Jerse, A.E., Yu, J., Tall, B.D. & Kaper, J.B. (1990). A genetic locus of enteropathogenic *Escherichia coli* necessary for the production of attaching and effacing lesions on tissue culture cells. *Proceedings of the National Academy of Science of the United States of America*, **87**, 7839-7843.

Johnson, J.R., Delavari, P., Kuskowski, M. & Stell, A.L. (2001). Phylogenetic distribution of extraintestinal virulence-associated traits in *Escherichia coli*. *Journal of Infectious Diseases*, **183**, 78-88.

Johnson, J.R. & Russo, T.A. (2002). Extraintestinal pathogenic *Escherichia coli*: "The other bad *E coli*". *Journal of Laboratory and Clinical Medicine*, **139**,155-162.

Johnson, J.R. & Russo, T.A. (2003). Review: Medical and economic impact of extraintestinal infections due to *Escherichia coli*: focus on an increasingly important endemic problem. *Microbes and Infection*, **5**, 449-456.

Johnson, J.R. & Stell, A.L. (2000). Extended virulence genotypes of *Escherichia coli* strains from patients with urosepsis in relation to phylogeny and host compromise. *Journal of Infectious Diseases*, **181**, 261-272.

Karmali, M.A. (1989). Infection by verocytotoxin-producing *Escherichia coli*. *Clinical Microbiology Review*, **2**(1), 15-38.

Kikine, T.N.F. (2011). *Profiling of Potential Pathogens From Plankenburg River Water Used for the Irrigation of Fresh Produce*. MSc in Food Science Thesis. Stellenbosch University, Stellenbosch, South Africa.

LeClerc, J.E., Li, B., Payne, W.L. & Cebula, T.A. (1996). High mutation frequencies among *Escherichia coli* and *Salmonella* pathogens. *Science*, **274**(5290), 1208-1211.

LeClerc, H., Mossel, D.A.A., Edberg, S.C. & Stuijk, C.B. (2001). Advances in the bacteriology of the coliform group: their suitability as markers of microbial water safety. *Annual Review of Microbiology*, **55**, 201-234.

Le Gall, T., Clermont, O., Gouriou, S., Picard, B., Nassif, X., Denamur, E. & Tenaillon, O. (2007). Extraintestinal virulence is a coincidental by-product of commensalism in B2 phylogenetic group *Escherichia coli* strains. *Molecular Biology and Evolution*, **24**, 2373-2384.

Linscott, A.J. (2011). Food-borne illnesses. *Clinical Microbiology Newsletter*, **33**, 41-45.

Stellenbosch University http://scholar.sun.ac.za

Lopez-Suacedo, C., Cerna, J.F., Villegas-Sepulveda, N., Thompson, R., Velazquez, F.R., Torres, J., Tarr, P.I. & Estrada-Garcia, T. (2003). Single multiplex polymerase chain reaction to detect diverse loci associated with diarrheagenic *Escherichia coli*. *Emerging Infectious Diseases*, **9**(1), 127-131.

Lötter, M. (2010). *Assessment of Microbial Loads Present in Two Western Cape Rivers Used for Irrigation of Vegetables***.** MSc in Food Science Thesis. Stellenbosch University, Stellenbosch, South Africa.

Matic, I., Radman, M., Taddei, F., Picard, B., Doit, C., Bingen, E., Denamur, E., Elison, J., LeClerc, J.E. & Cebula, T.A. (1997). Highly variable mutation rates in commensal and pathogenic *Escherichia coli*. *Science*, **277**, 1833-1834.

Meays, C.L., Broersma, K., Nordin, R. & Mazumder, A. (2004). Source tracking fecal bacteria in water: a critical review of current methods. *Journal of Environmental Management*, **73**, 71–79.

Moses, A.E., Garbati, M.A., Egwu, G.O. & Ameh, J.A. (2006). Detection of *E. coli* O157 and O26 serogroup in human immunodeficiency virus – infected patients with clinical manifestations of diarrhea in Maiduguri, Nigeria. *Research Journal of Medicine and Medical Science*, **1**(4), 140-145.

Ochman, H. J., Lawrence, G. & Groisman, E.A. (2000). Lateral gene transfer and the nature of bacterial innovation. *Nature*, **405**, 299–304.

Okafo, C.N., Umoh, V.J. & Galadima, M. (2003). Occurrence of pathogens on vegetables harvested from soils irrigated with contaminated streams. *The Science of the Total Environment*, **311**, 49-56.

Omar, K.B. & Barnard, T.G. (2010). The occurrence of pathogenic *Escherichia coli* in South African wastewater treatment plants as detected by multiplex PCR. *Water SA*, **36**(2), 172-176.

Pass, M.A., Odedra, R. & Batt, R.M. (2000). Multiplex PCR for identification of *Escherichia coli* virulence genes. *Journal of Clinical Microbiology*, **38**, 2001-2004.

Percival, S., Chalmers, R., Embrey, M., Hunter, P., Sellwood, J. & Wyn-Jones, P. (2004). *Escherichia coli*. In: *Microbiology of Waterborne Diseases: Microbiological Aspects and Risks*. Pp. 71-90. New York: Academic Press.

Picard, B., Garcia, J.S., Gouriou, S., Duriez, P., Brahimi, N., Bingen, E., Elion, J. & Denamur, E. (1999). The link between phylogeny and virulence in *Escherichia coli* intestinal infection. *Infection and Immunity*, **67**, 546-553.

Pupo, G.M., Karaolis, D.K.R., Lan, R. & Reeves, P.R. (1997). Evolutionary relationships among pathogenic and non-pathogenic *Escherichia coli* strains inferred from multilocus enzyme electrophoresis and *mdh* sequence studies. *Infection and Immunity*, **65**, 2685-2692.

Reid, S.D., Herbelin, C.J., Bumbaugh, A.C. Selander, A.K. & Whittam, T.S. (2000). Parallel evolution of virulence in pathogenic *Escherichia coli*. *Nature*, **406**, 64–67.

RMS (Resource Management Services). (2012). Stellenbosch Municipality: Proposed establishment of waste management facilities at the Stellenbosch landfill site. Report no. RMS/SM/SLFS/WA/02/12.

SANS 5667-6 (2006a). Water quality – Sampling – Part 6: Guidance on sampling of rivers and streams. Published by Standards South Africa, Pretoria.

SANS 5667-11 (2006b). Water quality – Sampling – Part 11: Guidance on sampling of groundwaters. Published by Standards South Africa, Pretoria.

SANS 9308 (2012). Microbial analysis of water – General test methods. Total coliforms and *Escherichia coli* in water: Defined substrate technology (Colilert) method. Published by Standards South Africa, Pretoria.

Stellenbosch University http://scholar.sun.ac.za

Scott, T.M., Rose, J.B., Jenkins, T.M., Farrah, S.R. & Lukasik, J. (2002). Microbial source tracking: current methodology and future directions. *Applied and Environmental Microbiology*, **68**(12), 5796-5803.

Solomon, E.B., Yaron, S. & Matthews, K.R. (2002). Transmission of *Escherichia coli* O157:H7 from contaminated manure and irrigation water to lettuce plant tissue and its subsequent internalization. *Applied and Environmental Microbiology,* **68**(1), 397-400.

Tarr, C.L., Large, T.M., Moeller, C.L. Lacher, D.W., Tarr, P.I., Acheson, D.W. & Whittan, T.S. (2002). Molecular characterization of a serotype O121:H19 clone, a distinct shiga toxin-producing clone of a pathogenic *Escherichia coli*. *Infection and Immunity*, **70**(12), 6853-6859.

Van Blommestein, A. (2012). *Impact of Selected Environmental Factors on E. coli Growth in River Water and an Investigation of Carry-over to Fresh Produce during Irrigation*. MSc in Food Science Thesis. Stellenbosch University, Stellenbosch, South Africa.

Walk, S.T., Alm, E.W., Calhoun, L.M., Mladonicky, J.M. & Whittam, T.S. (2007). Genetic diversity and population structure of *Escherichia coli* isolated from freshwater beaches. *Environmental Microbiology*, **9**, 2274-2288.

Willey, J.M., Sherwood, L.M. & Woolverton, C.J. (Eds.). (2008). Microbiology of food. In: *Prescott, Harley & Klien's Microbiology*, 7th Ed. Pp. 1023-1048. New York: McGraw Hill.

Stellenbosch University http://scholar.sun.ac.za

Appendix

Appendix A PCR Results of all *E. coli* strains investigated using molecular methods

Isolate Name	Source	PCR Results	
		Pathotype	Phylogenetic Group
A11.1	Borehole A1	-	A_0
A11.2	Borehole A1	-	A_1
A11.3	Borehole A1	EPEC	B1
A11.4	Borehole A1	-	A_1
A11.5	Borehole A1	-	A_1
A21.1	Borehole A1	-	A_0
A21.2	Borehole A1	-	B1
A21.3	Borehole A1	-	A_0
A21.4	Borehole A1	-	A_1
A22.1	Borehole A	-	A_0
A22.2	Borehole A	-	B1
A22.3	Borehole A	-	B1
A22.4	Borehole A	-	A_0
A22.5	Borehole A	-	A_1
B1.1	Piggery Ferm. Dam	-	A_1
B1.2	Piggery Ferm. Dam	-	A_1
B1.3	Piggery Ferm. Dam	-	A_1
B1.4	Piggery Ferm. Dam	-	A_0
B12.1	Piggery Overflow	-	A_0
B12.3	Piggery Overflow	-	A_1
B12.4	Piggery Overflow	-	A_1
B13.1	Piggery Big Dam	-	A_0
B13.2	Piggery Big Dam	-	D_1
B13.3	Piggery Big Dam	-	$B2_2$
B13.4	Piggery Big Dam	-	A_0
B21.1	Piggery Ferm. Dam	-	A_0
B21.2	Piggery Ferm. Dam	-	A_0
B21.3	Piggery Ferm. Dam	-	A_0
B21.4	Piggery Ferm. Dam	-	A_0
B21.5	Piggery Ferm. Dam	-	A_0
B22.1	Piggery Overflow	-	B1

Stellenbosch University http://scholar.sun.ac.za

B22.2	Piggery Overflow	-	B1
B22.3	Piggery Overflow	-	B1
B22.4	Piggery Overflow	-	A_1
B22.5	Piggery Overflow	-	A_0
B23.1	Piggery Big Dam	-	A_1
B23.2	Piggery Big Dam	-	A_1
B23.3	Piggery Big Dam	-	A_1
B23.4	Piggery Big Dam	-	A_4
C21.1	Spring C1	-	B1
C21.2	Spring C1	-	B1
C21.3	Spring C1	-	B1
C21.4	Spring C1	-	B1
C21.5	Spring C1	-	B1
D11.1	Borehole D1	-	D_2
D11.2	Borehole D1	-	D_2
D11.3	Borehole D1	-	D_2
D11.4	Borehole D1	-	B1
D12.1	Dam D1	-	B1
D12.2	Dam D1	-	B1
D12.3	Dam D1	-	B1
D12.4	Dam D1	-	B1
D21.1	Borehole D1	-	A_0
D21.2	Borehole D1	-	B1
D21.3	Borehole D1	-	A_0
D21.4	Borehole D1	-	$B2_3$
D21.5	Borehole D1	-	B1
D24.2	Dam D1	-	A_0
D24.3	Dam D1	-	A_0
D24.4	Dam D1	-	A_0
E11.1	Plank 0	-	$B2_3$
E11.2	Plank 0	-	$B2_3$
E11.3	Plank 0	-	$B2_3$
E11.4	Plank 0	-	$B2_3$
E11.5	Plank 0	-	$B2_3$
E12.1	Plank 1	-	B1
E12.2	Plank 1	-	A_1

Stellenbosch University http://scholar.sun.ac.za

E12.4	Plank 1	-	A_1
E21.1	Plank 0	-	D_2
E21.2	Plank 0	-	A_1
E21.3	Plank 0	-	D_2
E21.4	Plank 0	-	A_1
E21.5	Plank 0	-	A_1
E22.1	Plank 1	-	A_1
E22.2	Plank 1	-	A_1
E22.3	Plank 1	-	A_1
E22.4	Plank 1	-	A_1
E22.5	Plank 1	-	B1
F11.2	Plank 3	-	A_1
F11.3	Plank 3	EPEC	$B2_3$
F11.4	Plank 3	-	B1
F21.1	Plank 3	-	B1
F21.3	Plank 3	-	B1
F21.4	Plank 3	-	A_0
F21.5	Plank 3	-	$B2_3$
G11.1	Veldwagters River	-	A_0
G21.1	Veldwagters River	-	A_0
G21.3	Veldwagters River	-	A_1
G21.4	Veldwagters River	-	A_0
H11.1	Olifantsrivier	-	B1
H11.2	Olifantsrivier	-	B1
H11.3	Olifantsrivier	-	B1
H11.4	Olifantsrivier	-	B1
H11.5	Olifantsrivier	-	B1
H23.4	Olifantsrivier	-	A_1
H23.5	Olifantsrivier	-	A_1
H24.1	Olifantsrivier	-	A_0
J11.1	Mosselbank River	-	B1
J11.2	Mosselbank River	-	B1
J11.3	Mosselbank River	-	$B2_3$
J11.4	Mosselbank River	-	B1
J11.5	Mosselbank River	-	D_1
K11.1	Berg 2	-	B1

Stellenbosch University http://scholar.sun.ac.za

K11.3	Berg 2	-	A_1
K11.4	Berg 2	-	A_1
K11.5	Berg 2	-	B1
L11.1	Winery Effluent	-	A_1
L11.2	Winery Effluent	-	A_1
L11.3	Winery Effluent	-	A_0
L11.4	Winery Effluent	-	A_1
L11.5	Winery Effluent	-	A_1
L11.6	Winery Effluent	-	B1
L21.1	Winery Effluent	-	A_1
L21.2	Winery Effluent	-	A_0
L21.3	Winery Effluent	-	A_0
L21.4	Winery Effluent	-	D_1
L21.5	Winery Effluent	-	B1
M11.1	Dairy	-	B1
M11.2	Dairy	-	B1
M11.3	Dairy	-	B1
M11.5	Dairy	-	A_0
M21.1	Dairy	-	B1
M21.2	Dairy	-	B1
M21.3	Dairy	-	B1
M21.4	Dairy	-	A_1
M21.5	Dairy	-	$B2_3$
M22.1	Dairy	-	D_1
M22.2	Dairy	-	B1
M22.3	Dairy	-	B1
M22.4	Dairy	-	B1
M22.5	Dairy	-	B1

Stellenbosch University http://scholar.sun.ac.za

CHAPTER 5

GENERAL DISCUSSION AND CONCLUSIONS

Naturally occurring surface and groundwater is often used by farmers as an alternative source for irrigation water. Microbial quality of such water is not readily known, as no regulations have been set for irrigation water in this country and only guidelines exist (WHO, 1989; DWAF, 1996). As a result, water used for the irrigation of fresh produce frequently surpasses the upper limit for low risk water application of 1 000 faecal coliform counts.100 mL^{-1} (DWAF, 1996). Microbial quality of groundwater is also not tested as a standard practice, as groundwater is believed to be safer than surface water when applied for irrigation purposes (Adams *et al*., 2001; Bezuidenhout *et al*., 2011; Haramoto *et al*., 2011).

When investigating the prevalence of *Escherichia coli* at the sample sites, more than half (10/19) of the sample sites had *E. coli* levels exceeding the WHO and DWAF guidelines for water being utilised for the irrigation of fresh produce to be consumed raw or minimally processed (WHO, 1989; DWAF, 1996). When considering only the irrigation sites, 71% (5/7) of the surface waters sampled were deemed as 'unsafe' for irrigation purposes, while all groundwater tested contained less than log 3 MPN.100 mL^{-1} *E. coli* (*E. coli* levels ranged from not detected to log 2.00 MPN.100ml^{-1}). It was therefore concluded that due to the low prevalence of *E. coli*, that groundwater is a safer option to use as an alternative water source.

The *E. coli* prevalence in irrigation water is however not the only important factor to consider, as the type of *E. coli* contamination present and by extension, their characteristics, also play a role in determining potential risk of using contaminated water for irrigation purposes. In this study, risk was defined as a measure of pathogenic *E. coli* strains present in the water. Characterising *E. coli* is also important for microbial source tracking which may lead to determining which sources play a role in contamination of certain waters.

In this study two methods were used to characterise and group *E. coli* strains isolated from irrigation and contamination source sites. A matrix assisted laser desorption-ionisation time-of-flight mass spectrometer (MALDI-TOF MS) was used for its biotyping capabilities, based on the identification of the five most abundant proteins present. API 20E was the second method employed where 27 biochemical tests were used to identify each strain using a unique biochemical profile and APIweb™. From this data, it was concluded that the MALDI Biotyper was a more accurate and less subjective method to use, but the biotyper is expensive and requires much more preparation and training before being able to use the instrument efficiently. The API 20E

Stellenbosch University http://scholar.sun.ac.za

system is however more accessible, cheaper and can be used even when only characterising a few strains in a small laboratory. On the other hand, it is a very subjective set of results to read and relies heavily on the individual doing the analysis. MALDI Biotyping also allows for more accurate identification, as shown by the data in this study where some strains could not be identified at all using API 20E, but could be identified when using the MALDI Biotyper. It was therefore concluded that MALDI Biotyping is better suited for characterisation and identification of *E. coli* strains. As a result of the MALDI Biotyper being able to identify strains which were unidentifiable when using API 20E, the MALDI Biotyper data was used to cluster strains in a PCA dendrogram, which was used in conjunction with molecular methods for microbial source tracking.

When comparing surface and groundwater *E. coli* counts as found in this study, it was noted that groundwater counts were lower than those of surface water. This is however still a matter of concern as some pathogenic strains have low infectious doses (Karmali, 1989; Percival *et al*., 2004) and can cause disease, even when only a few bacterial cells are ingested. It must also be taken into account that strains which are less prevalent in the environment are unlikely to be isolated using standard methods (Gordon, 2001; Gordon & Cowling, 2003). The implications of this is that strains which are reported at a low prevalence, such as the EPEC isolated in this study, could be more abundant in the natural population than the data indicates. In addition to this it may be expected that even if water is tested for pathogenic *E. coli*, and declared to be suited for irrigation of fresh produce, that it might not be true. This is because the pathogenic *E. coli* may not always be isolated and could still cause disease to consumers of the fresh produce irrigated with the contaminated water.

Enteropathogenic *E. coli* (EPEC) strains were isolated in this study from both groundwater and surface water irrigation sites which imply that groundwater, similarly to surface water, is also influenced by outside factors and that contaminants can find their way into groundwater. The faecal coliform counts at this irrigation site were also found to be consistently under the recommended guideline limit (WHO, 1989; DWAF, 1996). This aspect is of importance as it proves that *E. coli* load is not an indication of pathogenic *E. coli* present, and therefore also not of potential risk. When taking into account the biased sampling, it may be concluded that the actual prevalence of EPEC in the natural water systems was not accurately represented, and that the prevalence of EPEC in the water systems may be higher than expected. This also concludes that untreated groundwater is not as safe an option for irrigation of fresh produce as originally thought. *Escherichia coli* prevalence was also found not to give a clear indication of associated risk, as pathogens can be present even when *E. coli* prevalence is lower than the recommended

Stellenbosch University http://scholar.sun.ac.za

guidelines (WHO, 1989; DWAF, 1996), as was the case with Borehole A1 where an EPEC strain was isolated. Although EPEC is not known to have an extremely low infectious dose (<100 cells), ingestion of small amounts may still cause disease in infants and immune deficient individuals (Karmali, 1989; Bolton *et al*., 1996; Percival *et al*., 2004). Therefore, to make an accurate assessment of associated risk, further work needs to be done and different sample methods need to be used to overcome the biased sampling.

The source of this particular pathogen is another problem, as boreholes are influenced minimally by outside factors which could contribute to contamination. As an EPEC strain was also isolated from surface water (Plank 3) in this study, as well as previously (Van Blommestein, 2012), it was concluded that EPEC contamination is not just a once-off occurrence, but that it is probably a continual contamination source which plays a role. This also emphasises the importance of microbial source tracking to be used to curb pathogenic contamination from reaching natural water systems.

When comparing the phylogeny of the two EPEC strains, it was found that strain A11.3 was classified as a member of the phylogenetic group B1, while F11.3 was categorised as subgroup $B2_3$. In a previous study, where the Plankenburg River was also sampled, and *E. coli* isolated (Van Blommestein, 2012), it was reported that three EPEC strains were isolated from the Plankenburg River, all of which belonged to phylogenetic subgroup $B2_3$. The implications of this are that EPEC strains which are in phylogenetic group B1 are able to survive well in the environment (Walk *et al*., 2007), which means that they survive, multiply and could therefore result in horizontal gene transfers which facilitates the carry-over of the *eaeA* virulence gene from one *E. coli* strain to another. Phylogenetic subgroup $B2_3$ on the other hand is not known to survive in the environment as well as group B1 (Walk *et al*., 2007), and therefore the results in this and previous studies shows that there is most likely a continual contamination source which contributes to the contamination in the Plankenburg River. Another explanation could be that these EPEC strains, in phylogenetic subgroup $B2_3$, have adapted in such a way that they are able to survive and multiply in a fluvial environment.

Microbial source tracking of *E. coli* from irrigation water to the contamination source is important because if the most probable source for pathogenic contamination can be identified, the contamination can potentially be prevented. It could also go a long way in reducing pathogenic *E. coli* finding its way onto fresh produce and ultimately increasing food safety. In this study it was found that linking a particular strain to a probable source based solely on that strain's phylogeny was not possible, especially in river systems where bacteria in the irrigation water

Stellenbosch University http://scholar.sun.ac.za

probably come from a number of contamination sources. To solve this problem of source tracking, phylogenetic grouping of *E. coli* isolated from various sample sites was then used to create a population structure for each site. These population structure patterns from irrigation sites were compared to those of contamination sources, and it was concluded that population structure based on phylogenetic grouping can successfully be employed. It was also found that population structure comparisons are well suited for source tracking.

The *E. coli* population structure observed in groundwater showed a close resemblance to the population structure of *E. coli* from both dairy and piggery sources. Isolates from surface water were the most difficult to link back to a possible source, as the structural patterns from surface water populations appeared to be linked to almost all the contamination sources investigated. It was concluded that, due to the number of outside factors which influence surface water, narrowing the source of contamination down to a single type of source was not possible.

When using the population structure of individual sample sites instead of sample site 'types', a better and more accurate linking could be facilitated. For example, data from this study showed that isolates from the piggery, the Olifants River and the winery effluent all had similar population structures. It was therefore concluded that the combination of human and winery contamination of the winery effluent could be a major contamination contributor to irrigation water.

In another example, the Berg-2 and Plank-1 sites were found to have similar population structures. Plank-1 is situated downstream from an informal settlement with insufficient infrastructure and sanitary facilities. This means that the water from this river may be used for washing, sanitation and recreational purposes, which could all play a role in introducing *E. coli* into the water system. Plank-1, as was found in this study, is therefore seen as a contamination source representing human pollution, and as Plank-1 and Berg-2 have such similar population structures, it was concluded that the most probable source of contamination of the Berg River is also human pollution from an informal settlement. Upon further investigation it was found that there is an informal settlement upstream from the Berg-2 sampling site. This showed that by using unique population structures, irrigation sites can be linked to contamination sources.

Phylogenetic group prevalence was also observed when looking at the population structures based on phylogenetic grouping. As a result, it was concluded that strains from natural water systems (ground and surface waters) showed an abundance of phylogenetic group B1. *Escherichia coli* strains assigned to phylogenetic group B1 were reported in a study by Walk *et al.* (2007) to be able to survive in the environment with ease. The population structure of *E. coli* from

Stellenbosch University http://scholar.sun.ac.za

dairy samples also showed a definitive numerical dominance of phylo-group B1. This led to the conclusion that cattle, and their faecal contamination, may have been a continual contamination source in surface water over the years. It is possible that cow manure had contaminated the river systems and that the *E. coli* present in the contamination then acclimatised to aid in their survival. This acclimatisation caused the strains to change in such a way that they no longer fully resemble the original bovine strains. This caused these strains to become prevalent in natural water and to potentially be seen as 'environmental strains' which are natural inhabitants of river water.

As *E. coli* is readily able to acquire genes, it was speculated in this study that those strains which survive in the environment may have acquired genes (LeClerc *et al*., 1996; Matic *et al*., 1997) to help aid in their survival. This speculation was based on the fact that 48% of the *E. coli* strains in this study were found to be VP (Voges-Proskauer) positive, an attribute which is not normally associated with *E. coli*. A positive VP result means that an organism can produce acetoin (3-hydroxy-2-butanone). In identification systems the formation of acetoin is often used as a microbial identification marker, and on a physiological level helps the organisms avoid acidification, as well as playing a role in carbon storage and constant regulation of the NAD/NADH ratio within the cell (Xiao & Xu, 2007). This means that the ability to produce and reutilise acetoin may help *E. coli* in environmental niches which are forever changing and not always rich in a usable carbon source (Johansen *et al*., 1975; Mayer *et al*., 1995). It was concluded in this study that the positive VP result for many strains showed an acclimatisation which aided in the survival of the *E. coli* in the more hostile environments.

Some recommendations for future studies to facilitate the identification of contamination sources would be to include a larger range of contamination sites. This would aid in the source tracking as population structures of *E. coli* strains from horses, ruminants, waterbirds and chickens are yet to be determined. If more contamination source sites can be characterised and the *E. coli* population structure determined, irrigation sites which show population structures which do not correlate with those of isolates from pigs, cows or humans may be linked to a different source. Biased isolation techniques are another important aspect that needs examination. This could be overcome by the addition of specific enrichment steps which allow for the rapid growth of *E. coli* as well as isolation of more strains from each water sample. By including these steps, the *E. coli* count in the water sample will be increased by the enrichment step, and some of the less abundant strains could be isolated if more strains were isolated.

As there were only two pathogenic strains found when testing for intestinal pathogenic *E. coli* (InPEC), extra-intestinal pathogenic *E. coli* (ExPEC) should also be tested for to enable a better

Stellenbosch University http://scholar.sun.ac.za

determination of associated risk with regards to pathogenic strain prevalence. Following the isolation of the EPEC strain from borehole water in this study, research which explores more groundwater sources, the *E. coli* strains present as well as the underground catchment area would be of value to examine. This would give a much broader range of information pertaining to groundwater, which may facilitate better source tracking and highlight possible ways in which the contamination could have reached the underground water reservoir.

REFERENCES

Adams, S., Titus, R., Pietersen, K., Tredoux, G. & Harris, C. (2001). Hydrochemical characteristics of aquifers near Sutherland in the Western Karoo, South Africa. *Journal of Hydrology*, **241**, 91–103.

Bezuidenhout, C.C. & The North-West University Team (2011). A scoping study on the environmental water (groundwater and surface water) quality and management in the North-West Province, South Africa. Water Research Commission (WRC) Report No. KV 278/11.

Bolton, F.J., Crozier, L. & Williamson, J.K. (1996). Isolation of *Escherichia coli* 01 57 from raw meat products. *Letters in Applied Microbiology*, **23**, 317-321.

DWAF (Department of Water Affairs and Forestry). (1996). Agricultural Use: Irrigation. In: *South African Water Quality Guidelines. Volume 4.* (2nd Ed.). (Edited by Holmes, S. CSIR Environmental Services). Pretoria: Department of Water Affairs and Forestry.

Gordon, D.M. (2001). Mini review: Geographical structure and host specificity in bacteria and the implications for tracing the source of coliform contamination. *Microbiology*, **147**, 1079-1085.

Gordon, D.M. & Cowling, A. (2003). The distribution and genetic structure of *Escherichia coli* in Australian vertebrates: host and geographic effects. *Microbiology*, **149**, 3575-3586.

Haramoto, E., Yamada, K. & Nishida, K. (2011). Prevalence of protozoa, viruses, coliphages and indicator bacteria in groundwater and river water in the Katmandu Valley, Nepal. *Transactional of the Royal Society of Tropical Medicine and Hygiene*, **105**, 711-716.

Johansen, L., Bryn, K. & Stormer, F.C. (1975). Physiological and biochemical role of the butanediol pathway in *Aerobacter* (*Enterobacter*) *aerogenes*. *Journal of Bacteriology*, **123**, 1124–1130.

Karmali, M.A. (1989). Infection by verocytotoxin-producing *Escherichia coli*. *Clinical Microbiology Review*, **2**(1), 15-38.

LeClerc, J.E., Li, B., Payne, W.L. & Cebula, T.A. (1996). High mutation frequencies among *Escherichia coli* and *Salmonella* pathogens. *Science*, **274**(5290), 1208-1211.

Matic, I., Radman, M., Taddei, F., Picard, B., Doit, C., Bingen, E., Denamur, E., Elison, J., LeClerc, J.E. & Cebula, T.A. (1997). Highly variable mutation rates in commensal and pathogenic *Escherichia coli*. *Science*, **277**(5333), 1833-1834.

Mayer, D., Schlensog, V. & Bock, A. (1995). Identification of the transcriptional activator controlling the butanediol fermentation pathway in *Klebsiella terrigena*. *Journal of Bacteriology*, **177**, 5261–5269.

Percival, S., Chalmers, R., Embrey, M., Hunter, P., Sellwood, J. & Wyn-Jones, P. (Editors) (2004). *Escherichia coli.* In: *Microbiology of Waterborne Diseases: Microbiological Aspects and Risks*. New York: Academic Press.

Stellenbosch University http://scholar.sun.ac.za

Van Blommestein, A. (2012). *Impact of Selected Environmental Factors on E. coli Growth in River Water and an Investigation of Carry-over to Fresh Produce during Irrigation*. MSc in Food Science Thesis. Stellenbosch University, Stellenbosch, South Africa.

Walk, S.T., Alm, E.W., Calhoun, L.M., Mladonicky, J.M. & Whittam, T.S. (2007). Genetic diversity and population structure of *Escherichia coli* isolated from freshwater beaches. *Environmental Microbiology*, **9**, 2274-2288.

WHO (World Health Organization). (1989). Health guidelines for the use of wastewater in agriculture and aquaculture. Geneva, World Health Organization. *Technical Report Series No 776. World Health Organization*, Switzerland, Geneva.

Xiao, Z. & Xu, P. (2007). Acetoin metabolism in bacteria. *Critical Reviews in Microbiology*, **33**, 127-140.