final report

Project code: V.MFS.0333

Prepared by: Dr Robert Barlow
CSIRO Agriculture & Food

Date published: 20-06-2017

PUBLISHED BY
Meat and Livestock Australia Limited
Locked Bag 1961
NORTH SYDNEY NSW 2059

Molecular characterisation of bacteria for continued market access

Meat & Livestock Australia acknowledges the matching funds provided by the Australian Government and contributions from the Australian Meat Processor Corporation to support the research and development detailed in this publication.

Executive summary

Microbiological requirements for red meat export exist in a number of markets, with the USA and EU having the most stringent current or potential requirements for pathogens. Australia implements a variety of safety and quality testing programs in order to maintain access to global markets. However, these markets may update their requirements at any time and Australian red meat exporters must subsequently demonstrate that they meet the updated requirements through equivalence agreements. The EU and USA have held discussions in relation to expanding the list of adulterants for red meat products entering their markets with proposals aimed at banning all raw meat containing any Shiga toxin-producing *Escherichia coli* (STEC) in the EU and the addition of specific antimicrobial resistant (AMR) bacteria to the adulterant list in the USA. In response to the changes that have already occurred and in anticipation of changes that may occur, Meat and Livestock Australia (MLA) and CSIRO have previously conducted a number of projects investigating the prevalence of STEC and AMR in Australian beef cattle populations. Whilst previous scientific outcomes reinforce Australia's position as a supplier of safe beef products, a set of knowledge gaps were identified that required immediate attention in order to equip the Australian red meat export industry to meet current and emerging market access requirements. The knowledge gaps addressed in this project include: phenotypic and genotypic characterisation of *Enterococcus* from Australian beef cattle, understanding the relatedness of non-O157 (Big6) STEC from human and cattle sources, and understanding the performance of current STEC test systems and compare with other commercially available STEC test systems.

AMR in Australian beef cattle

Antimicrobial agents are used in cattle production systems for the prevention and control of bacterial associated diseases. Bacteria that are resistant to antimicrobials are of increased concern to public health officials throughout the world as they may compromise the ability of treatment regimens to address disease and infection in humans. This study attempted to further understand the basis of AMR in *Enterococcus faecium* and *Enterococcus faecalis* arising from project G.MFS.0285. The results of the AMR investigation confirmed that high levels of resistance to antimicrobials that are not critically or highly important to human medicine with resistance to flavomycin (80.2%) and lincomycin (85.4 – 94.2%) are routinely observed. Conversely, resistance to antibiotics considered critically or highly important to human medicine such as tigecycline, daptomycin, vancomycin and linezolid were not present in this study. The latter conclusion was formed only after the initial phenotypic test results (see G.MFS.0285) were shown to contain major errors with respect to daptomycin and tigecycline. These data corroborate previous studies of AMR

in *E. coli* and *Salmonella* that demonstrate there is minimal evidence that Australian cattle production practices are responsible for disproportionate contributions to AMR development and in general, resistance to antimicrobials of critical and high importance in human medicine was low.

Subtyping non-O157 (Big6) STEC

In 2012 the export STEC testing requirements for USA changed to include six additional serogroups (O26, O45, O103, O111, O121 and O145), collectively known as non-O157 STEC or Big6. Previous studies (G.MFS.0286) have determined the prevalence of these organisms in beef cattle at slaughter, however, little is known about the capacity of these organisms to cause human clinical disease. Investigations on O157 STEC have demonstrated that some strains are more likely to cause disease than others. Determining the relatedness of human and cattle non-O157 STEC may enable the identification of a similar non-random distribution of disease causing strains in Australian cattle. A total of 106 predominantly Australian human and cattle sourced non-O157 STEC (i.e. O26, O45, O103, O111, O121 and O145) were selected from 170 isolates on the basis of pulsed field gel electrophoresis and *stx* subtype profiles and subsequently characterised using next generation sequencing based analysis for multi-locus variable number tandem repeats (MLVA), comparative genome fingerprinting, pangenome content, Shiga toxin bacteriophage insertion sites (SBI) and single nucleotide polymorphisms (SNPs). Additionally, the production of Shiga toxin by each isolate was assessed. Cattle carriage and the frequency of human clinical disease in Australia is dominated by O26 and O111 with the remaining non-O157 serotypes of regulatory importance to export beef producers present at very low prevalence, if at all. The results of this study demonstrated that there were notable differences observed between isolates belonging to different serotypes, however no major differences were observed between cattle and human isolates of the same serotype. Indeed, there was substantial correlation between isolates of cattle and human origin. However, while cattle are a reservoir for non-O157 isolates associated with human disease, the low incidence of human disease due to such isolates may be correlated with the low prevalence of these same serogroups in cattle.

Comparison of STEC detection systems

Detection of STEC remains a significant challenge for the Australian red meat industry. Australian exporters currently use two main STEC test systems to determine if a manufacturing beef enrichment is potentially positive (PP) for STEC. Enrichments that are PP are subsequently culture confirmed at a Department of Agriculture and Water Resources approved laboratory, however, the

conversion of PP to confirmed positive is low for non-O157. A total of nine STEC test systems comprising five traditional (BAX, RapidFinder, Qiagen, Biotecon and FSIS), three advanced (GDS, Roka and PALL), and one confirmatory (NeoSeek) test system were assessed using 100 potential positive broths generated as part of industry-based STEC testing. Of the 100 samples assessed, traditional test systems detected between 64-85 PP's compared to the advanced systems which detected between 39 and 56 PP's. A total of twelve out of the 100 samples had confirmed for non-O157 STEC (all O26) and all screening test systems identified at least 10. NeoSeek was the only confirmatory test system assessed in this study and it determined that 16 samples were positive for Big6 STEC. Whilst the NeoSeek approach eliminated PP's it did result in a higher number of confirmed positives than the current culture confirmation process.

Metagenomic analysis determined that manufacturing beef enrichment broths are often dominated by non-*E. coli* organisms such as *Clostridium perfringens*, with *E. coli* often comprising less than 10% of the enrichment. Furthermore, the target serogroups are seldom the dominant serogroups in each broth and can be present at ratios lower than 1 in 1000 *E. coli*. Attempts were made to identify novel genetic markers, away from *stx*, *eae* and O-serogroups, that may be suitable for incorporation into novel screening test systems that are best suited for Australian conditions. The sequencing depth (i.e the amount of genomic DNA analysed with respect to the overall concentration of genomic DNA) achieved in this study was insufficient to identify additional markers and may need to be revisited in coming years as sequencing technology improves further. The results of this study indicate that the test systems being currently used in Australia's STEC testing program are comparable to other available test systems that belong to the same test kit category (e.g. traditional screening or novel screening). Although this study did not evaluate the cost, capital investment and labour intensity of each test system, it is clear that systems that utilise additional or alternative genetic markers may substantially reduce the number of PP samples requiring confirmation. The use of such systems could reduce testing costs but more importantly would substantially reduce the costs of holding product prior to export. In addition, the use of PCR-based confirmation as opposed to culture confirmation is being assessed internationally for a range of foodborne pathogens. The results of this study provide encouragement for the industry to continue to explore the implementation of novel STEC confirmation systems such as NeoSeek STEC as a means of reducing the costs and timeframes associated with the STEC testing program.

When taken together with recent MLA/CSIRO research on STEC and AMR, the scientific outcomes of this project equip the Australian red meat export industry with data and knowledge to meet current and identified emerging market access requirements. The conclusions and key messages that arise from this project primarily stem from the application of a range of molecular and next generation sequencing tools. Tools such as whole genome sequencing and metagenomics analysis are now common place globally and are replacing frequently used typing and investigative methods. Advancements in these areas provides both a challenge and an opportunity for the Australian red meat industry to continue to reinforce its global position as a preferred supplier of premium red meat products with assured safety and quality.

Contents

Executive summary 2

AMR in Australian beef cattle 2

Subtyping non-O157 (Big6) STEC 3

Comparison of STEC detection systems 3

Background 11

Project objectives 12

Success in achieving milestone 12

Recommendations 13

Milestone 2 – AMR characterization – Draft publication on the genetic and phenotypic investigation of AMR reported in 2013 survey prepared and submitted to MLA 14

Abstract 14

Project objectives 16

Success in achieving milestone 16

Introduction 17

Materials and Methods 18

Sample collection 18

Isolation of *Enterococcus* 18

Phenotypic detection of antimicrobial resistance 18

Genotypic detection of antimicrobial resistance 19

Results 19

Prevalence and identity 19

Antimicrobial susceptibility testing 20

Re-categorisation of daptomycin, tigecycline and quinupristin/dalfopristin results 22

Antibiogram profiles 23

Discussion 23

References 27

Appendix A: Genetic characterisation of AMR in *E. coli* from Australian beef cattle populations 38

Background 38

Materials and Methods 38

Samples 38

Phenotypic AMR testing 38

Genotypic AMR testing 41

Results and Discussion 41
Conclusion 41
References 42
Overall progress of the project 43
Recommendations 43
Milestone 3 – Key collaborator meetings – Attend and present pSTEC research findings at IAFP 2015 and VTEC 2015. 44
Abstract 44
Milestone description 46
Project objectives 46
Success in meeting the milestone 46
Discussion 46
International Association for Food Protection (IAFP 2015) 46
Presentation of MLA/CSIRO co-funded research 46
Collaborator meetings at IAFP 2015 47
Summary of conference 48
Symposium on Shiga Toxin (Verocytotoxin) Producing *Escherichia coli* Infections (VTEC 2015) 51
Presentation of MLA/CSIRO co-funded research 51
Collaborator meetings at VTEC 2015 52
Summary of conference 52
Conclusions/recommendations 60
Milestone 4 – Non-O157 subtyping – Draft publication prepared and submitted to MLA – Comparative analysis of up to 100 Australian non-O157 EHEC from cattle and clinical sources 61
Abstract 61
Project objectives 63
Success in achieving milestone 63
Introduction 64
Materials and Methods 67
Strain selection 67
Pulsed field gel electrophoresis 67
Shiga toxin gene subtyping 68
Shiga toxin production 68
Sequencing 68
Multi-locus variable number tandem repeat analysis (MLVA) 69

Comparative genome fingerprinting (CGF) 69
Shiga toxin bacteriophage insertion site (SBI) typing 70
Pangenome 70
Single nucleotide polymorphisms (SNPs) 73
Results 73
Strain selection 73
Shiga toxin production 78
Sequencing statistics 79
MLVA 79
Comparative genome fingerprinting (CGF) 79
Shiga toxin bacteriophage insertion site (SBI) typing 82
Pangenome 85
Single nucleotide polymorphisms (SNPs) 88
Discussion 88
Acknowledgements 92
References 93
Appendix 1 97
Milestone 5 – Comparison of pSTEC test systems and metagenomics profiling of manufacturing beef enrichment broths 98
Executive summary 98
Project objectives 101
Success in achieving milestone 101
Introduction 103
Materials and Methods 105
Sample collection 105
Initial STEC screening 105
Additional STEC testing 105
Quantitative PCR 107
Bioinformatic methods 107
Results and Discussion 111
Methods comparison 111
Sample summary 111
Detection of potential positives 111
Detection of confirmed samples 111

Concentration of *eae* and O-antigen genes....113
Metagenomic assessment....116
Taxonomic abundance....116
E. coli virulence markers and O antigen genes....118
Mapping of *eae*, stx_1 and stx_2....122
Conclusions....124
Acknowledgements....126
References....127
Appendix 1....129
Appendix2....136
Appendix 3....140
Industry workshop – Shiga toxin-producing *Escherichia coli* in manufacturing beef: Where have we been? Where should we be going?....141
The complexity of STEC testing – Kate McMillan....141
Background....141
Australian perspective....141
STEC testing is now more complex....142
Low confirmation rates....142
Samples that don't culture confirm....142
Conclusions....143
Comparison of STEC detection systems – Robert Barlow....143
Background....143
What does a good test look like?....144
Study design....144
Results....145
Conclusions....146
The future of STEC testing – Glen Mellor....147
Current concept....147
Not all STEC are equal....147
Molecular risk assessment?....148
Future testing systems....148
Conclusions....148
Future typing methods – here now – P. Scott Chandry....148
Background....148

Applications of NGS technology 149
Industry adoption 149
Next steps 150

Background

Microbiological requirements for beef export exist in a number of markets, with the USA and EU having the most stringent current or potential requirements for pathogens. International concerns are being raised about antimicrobial resistance of microorganisms (that may be found on meat) as a potential technical barrier to trade. In 1994 the USA declared Shiga toxin-producing *E. coli* (STEC) O157:H7, as an adulterant of raw, non-intact beef products. In 2012 the list of adulterants was extended to include an additional six STEC serogroups (O26, O45, O103, O111, O121 and O145). Collectively the additional six adulterants are known as non-O157 STEC (Big6) and when STEC O157 is included the group of seven STEC serogroups are referred to as Top7 STEC. The EU has proposed banning all raw meat containing any STEC and the USA has proposed the addition of multidrug resistant *Salmonella* to the list of adulterants of raw, non-intact beef products. The World Health Organisation has identified antimicrobial resistant infections as a major public health concern, and often cites overuse of antibiotics in animal agriculture as a factor in antimicrobial resistance in humans. The potential for declaration of strains resistant to several antibiotics as adulterants in the USA may be the first example of a round of technical barriers to trade that may require extensive surveys or lot-by-lot testing and rejections/recalls.

In response to the changes that have already occurred and in anticipation of changes that may occur, Meat and Livestock Australia (MLA) and CSIRO have conducted a number of projects investigating the prevalence of STEC and *Salmonella* in Australian beef cattle populations. Additionally, the antimicrobial status of *Salmonella* and indicator organisms *E. coli* and *Enterococcus* from cattle has been assessed. Along with understanding the prevalence of STEC or antimicrobial resistance (AMR) in Australian cattle populations, effort has been put into understanding the effectiveness of accepted and proposed methods for the detection and confirmation of STEC. While the scientific outcomes of the recent projects continue to reinforce Australia as a supplier of safe beef products, a number of further scientific questions were identified when considering the likely impact on the red meat industry and export markets. Further investigation of these outstanding questions is intended to equip Australia's red meat export industry to most efficiently meet current and emerging market access requirements.

Project objectives

1. Determine the relatedness of 'Big 6' Shiga toxin-producing *Escherichia coli* strains isolated from Australian cattle with those isolated from Australian human clinical cases
2. Genetically characterise pathogenic Shiga toxin-producing *E. coli* strains from Australian beef and cattle samples and assess improvements in specificity of emerging commercial confirmation test systems
3. Engage with leading international researchers, share information with these researchers and others and contribute to international consensus on defining the risk of *E. coli* found in cattle and beef to human health
4. Phenotypically and genetically characterise the antibiotic resistance status of enterococci isolated from beef cattle at slaughter

Success in achieving milestone

The project objectives of V.MFS.0333 have been successfully completed. The project extended the results and conclusions from previous MLA/CSIRO co-funded research by confirming:

- That there is minimal evidence that cattle production practices are responsible for disproportionate contributions to AMR development and in general, resistance to antimicrobials of critical and high importance in human medicine was low.
- That there are notable differences between STEC isolates belonging to different serogroups, however no major differences were observed between cattle and human isolates of the same serogroup and therefore cattle are confirmed as a reservoir of non-O157 STEC isolates associated with human clinical disease.
- STEC test systems currently being used in Australia are comparable to other available test systems in terms of their ability to identify samples that will ultimately culture confirm. The use of additional or alternative markers in STEC test systems may substantially reduce the number of potential positives samples that require culture confirmation, thereby reducing testing and holding costs.

The outcomes of this project have been communicated through a range of forums including international conferences, peer-reviewed publications, and industry workshops. This document is the compilation of a series of reports, publications and workshop documents produced as a result of this project's successful completion. The document includes:

- Report – AMR characterisation – Draft publication on the genetic and phenotypic investigation of AMR reported in the 2013 survey. This report was subsequently published: Barlow, R.S., McMillan, K.E., Duffy, L.L., Fegan, N., Jordan, D., and Mellor, G.E. (2017) Antimicrobial resistance status of *Enterococcus* from Australian cattle populations at slaughter. ***PLoS One*** 12(5):e0177728. http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0177728
- Report – Key collaborator meetings – Attend and present STEC research findings at IAFP 2015 and VTEC 2015.
- Report – Non-O157 subtyping – Comparative analysis of up to 100 Australian non-O157 EHEC from cattle and clinical sources.
- Report – Comparison of STEC test systems and metagenomics profiling of manufacturing beef enrichment broths.
- Industry workshop – Shiga toxin-producing *Escherichia coli* in manufacturing beef: Where have we been? Where should we be going?
    - The complexity of STEC testing – Kate McMillan
    - Comparison of STEC detection systems – Robert Barlow
    - The future of STEC testing – Glen Mellor
    - Future typing methods – here now – P. Scott Chandry

Recommendations

Technological and computational advances in the sequencing of DNA has transformed food safety and quality research. The conclusions and key messages that arise from this project stem from the application of a range of next generation sequencing tools. Tools such as whole genome sequencing and metagenomics analysis are now common place globally and are replacing standard typing and investigative methods. There is opportunity for the Australian red meat industry to identify gaps in existing knowledge and to address them using these new applications. These could include:

- Determination of the contributors, flow and diversity of commensal and pathogenic bacteria through production and processing of red meat animals and products,
- Defining the molecular profile of Australian STEC for use in a PCR based STEC confirmation test system,
- Broad scale assessment of the AMR status of red meat production systems i.e culture independent evaluation.

Milestone 2 – AMR characterization – Draft publication on the genetic and phenotypic investigation of AMR reported in 2013 survey prepared and submitted to MLA

Abstract

Antimicrobial agents are used in cattle production systems for the prevention and control of bacterial associated diseases. A consequence of their use is the potential development of antimicrobial resistance (AMR). *Enterococcus faecium* and *Enterococcus faecalis* that are resistant to antimicrobials are of increased concern to public health officials throughout the world as they may compromise the ability of various treatment regimes to control disease and infection in human medical settings. Australia is the world's third largest exporter of beef; however it does not have an ongoing surveillance system for AMR in cattle or foods derived from these animals. This study examined 910 beef cattle, 290 dairy cattle and 300 veal calf faecal samples collected at slaughter for the presence of enterococci and determined the phenotypic and genotypic AMR status of 800 enterococci. *Enterococcus* were isolated from 805 (88.5%) beef cattle faeces, 244 (84.1%) dairy cattle faeces and 247 (82.3%) veal calf faeces. The results of AMR testing identified high levels of resistance to antimicrobials of limited human clinical significance with resistance to flavomycin (80.2%) and lincomycin (85.4 – 94.2%) routinely observed. Major errors were observed with the Sensititre test system when used to evaluate tigecycline, daptomycin and quinupristin/dalfopristin resistance. Additional phenotypic and genotypic AMR testing determined that resistance to tigecycline, daptomycin, vancomycin and linezolid was not present in this study. Resistance to quinupristin/dalfopristin was observed in two (1.7%) of *E. faecium* although the genetic mechanism behind this resistance is yet to be elucidated. There is minimal evidence that cattle production practices are responsible for disproportionate contributions to AMR development and in general resistance to antimicrobials of critical and high importance in human medicine was low regardless of the isolate source. The low level of antimicrobial resistance in *Enterococcus* from Australian cattle is likely to result from comprehensive controls around the use of antimicrobials in food-production animals in Australia. Nevertheless, continued monitoring of the effects of all antimicrobial use is required to support Australia's reputation as a supplier of safe and healthy food.

Contents

Abstract....14
Project objectives....16
Success in achieving milestone....16
Introduction....17
Materials and Methods....18
Sample collection....18
Isolation of *Enterococcus*....18
Phenotypic detection of antimicrobial resistance....18
Genotypic detection of antimicrobial resistance....19
Results....19
Prevalence and identity....19
Antimicrobial susceptibility testing....20
Re-categorisation of daptomycin, tigecycline and quinupristin/dalfopristin results....22
Antibiogram profiles....23
Discussion....23
References....27
Appendix A: Genetic characterisation of AMR in *E. coli* from Australian beef cattle populations....38
Background....38
Materials and Methods....38
Samples....38
Phenotypic AMR testing....38
Genotypic AMR testing....41
Results and Discussion....41
Conclusion....41
References....42
Overall progress of the project....43
Recommendations....43

Project objectives

- Determine the relatedness of 'big 6' Shiga toxin-producing *Escherichia coli* strains isolated from Australian cattle with those isolated from Australian human clinical cases
- Genetically characterise pathogenic Shiga toxin-producing *Escherichia coli* strains from Australian beef and cattle samples and assess improvements in specificity of emerging commercial confirmation test systems
- Engage with leading international researchers, share information with these researchers and others and contribute to international consensus on defining the risk of *E. coli* found in cattle and beef to human health
- Phenotypically and genetically characterise the antibiotic resistance status of enterococci and *E. coli* isolated from beef cattle at slaughter

Success in achieving milestone

Additional phenotypic and genotypic investigations into the antimicrobial resistances observed in *E. coli* and *Enterococcus* from cattle at slaughter have been completed. The resistances in question were identified during project G.MFS.0285 and represented resistances to antimicrobials of high and critical importance to human medicine or that were inconsistent with the results observed with related antimicrobials. A total of 354 observations from previous AMR testing using the Sensititre test system were re-evaluated using additional phenotypic and genotypic methods. Major errors were found to have occurred with the testing of *Enterococcus* with respect to quinupristin/dalfopristin, daptomycin and tigecycline resistance and re-categorisation of these results from resistant to susceptible was required. This study confirms that resistance to antimicrobials of high or critical importance such as vancomycin, linezolid, daptomycin and tigecycline was not identified in Enterococcal isolates from Australian cattle at slaughter. Follow up investigations of *E. coli* strains exhibiting amoxicillin/clavulanic acid and ceftiofur resistance determined that resistance to amoxicillin/clavulanic acid was reproducible across phenotypic test systems although a genetic basis for resistance has not been identified. Reproducibility of the ceftiofur resistance was variable across phenotypic test system and it is likely that the reduced susceptibility observed is not due to the presence of genes that typically confer ceftiofur resistance. The *E. coli* strains in question will be analyses further using whole genome sequencing and any additional findings will be communicated to MLA (see Appendix A).

Introduction

Australia is one of the world's most efficient producers of cattle and third largest exporter of beef, exporting 67% of its total beef and veal production in 2012-13 (Meat & Livestock Australia 2013). Antimicrobial agents are used in cattle production systems for the prevention and control of bacterial associated diseases. A consequence of the use of antimicrobials is the potential for antimicrobial resistance (AMR) to develop in bacteria, including zoonotic pathogens which can be transferred to the human population via the food chain or by direct exposure to animals (Collignon and Angulo 2006; Heuer, Hammerum et al. 2006). Novel resistance phenotypes continue to emerge in zoonotic foodborne pathogens and commensal bacteria isolated from food production animals (Walsh and Fanning 2008; Szmolka and Nagy 2013). In particular, *Enterococcus faecalis* and *Enterococcus faecium* have become of increasing performance over recent decades because of life-threatening hospital-acquired infections. Consequently, understanding, assessing and mitigating the risks of non-human use of antimicrobials on human health outcomes remains a high priority. The World Health Organisation (WHO) has developed and maintains criteria and ranks antimicrobials based on their importance to human medicine (World Health Organisation 2011). These lists will help regulators and stakeholders identify appropriate antimicrobials for use in food animal production systems (Collignon, Powers et al. 2009).

In comparison to Australia, a number of countries do have established AMR surveillance programs in place. Whilst the main focus of these programs revolves around AMR in bacteria from humans there is considerable and increasing demand to enhance their impact by assessing AMR in bacteria from animals during production and from foods at the retail level. Multi-focus surveillance programs enable trends in AMR development to be further evaluated with respect to production practices, animal type and clinical use and are particularly useful in addressing concerns from regulators about the overall impact of antimicrobial use. Countries that lack sophisticated multi-focus surveillance programs instead rely on relatively short-term intensive surveys to evaluate the prevalence and AMR status of bacteria from an animal type, production practice or as a result of clinical use. The aim of this study was to determine the prevalence and AMR status of *Enterococcus* isolates from Australian cattle populations.

Materials and Methods

Sample collection

Australian beef cattle destined for export can be classified into three animal groups: beef cattle, dairy cattle, and veal calves. A total of 31 abattoirs representing >85% of total beef exports agreed to participate in the survey. The number of cattle to be sampled at each abattoir was stratified based on animal group and slaughter volumes. Sample collection targets of 900, 300 and 300 were established for beef cattle, dairy cattle and veal calves, respectively. Samples were collected across two sampling windows with sample numbers collected from each participating abattoir ranging from 8-80 (mean 24) per sampling window. Systematic random sampling was used to collect the samples across a consecutive two day period in each of the sampling windows. A sampling day consisted of eight hours of production with each abattoir expected to sample evenly across the day. Abattoirs were expected to collect up to a maximum of 40 samples per sampling day therefore all samples were expected to be collected a minimum of 12 minutes apart. Each sampling window occurred over an eight week period with the first window occurring in February and March, 2013 and the second sampling window occurring in August and September, 2013. Faecal samples were collected post-evisceration by cutting the intestine 15-30 cm from the rectal end and squeezing at least 40 g of material into a sterile jar. Samples were kept chilled and returned to the laboratory by overnight courier for processing.

Isolation of *Enterococcus*

The presence of *Enterococcus* was determined by enriching 1 g of faeces in 10 ml of BBL Enterococcosel Broth (BD, Maryland, USA) for 18-24 h at 35 ± 2°C. Enriched broths were then plated onto BBL Enterococcosel Agar (BD) and incubated for 18-24 h at 35 ± 2°C. Translucent colonies with brownish-black to black zones were then streaked onto SBA and incubated for 18-24 h at 35 ± 2°C. Isolates were confirmed as *Enterococcus* spp. by PCR (Ke, Picard et al. 1999). A species specific PCR was then used to identify *E. faecalis* and *E. faecium* strains (Dutkamalen, Evers et al. 1995; Dutkamalen, Evers et al. 1995). Further speciation was not performed and the remaining isolates were labelled *Enterococcus spp*.

Phenotypic detection of antimicrobial resistance

The AMR phenotype of isolates was initially determined using the broth microdilution method and the Sensititre apparatus. Custom susceptibility panels for *Enterococcus* (AUSVP2; TREK Diagnostic Systems, UK) were used to test all isolates. The dilution ranges and breakpoints for each

antimicrobial are shown in Table1. Interpretation of the MIC values was based on CLSI interpretive criteria when available; otherwise EUCAST and NARMS values were used. The breakpoint listed for florfenicol is the susceptible breakpoint. Isolates that exceeded the MIC value of the susceptible breakpoint were reported as non-susceptible. *Enterococcus faecalis* ATCC 29212 was used as the control strains.

Isolates that demonstrated resistance to antimicrobials (daptomycin, tigecycline and quinupristin/dalfopristin) of human clinical significance in the Sensititre testing process were further evaluated using MIC evaluator strips (daptomycin and tigecycline; Thermofisher Scientific, UK) and/or disc susceptibility testing (quinupristin/dalfopristin). Susceptibility testing was conducted as per the manufacturer's recommendations with each isolate suspended in cation adjusted Mueller-Hinton broth at 0.5 MacFarland standard. Each isolate was subsequently spread plated onto Mueller-Hinton agar and overlaid with the appropriate MIC evaluator strip or antibiotic susceptibility disc. The MIC or zone of clearance was measured after 24 hours incubation at 37°C.

Genotypic detection of antimicrobial resistance

Isolates that demonstrated resistance to daptomycin, tigecycline or quinupristin/dalfopristin were tested for the presence of AMR genes or SNPs that have previously been shown to be associated with resistance to the aforementioned antimicrobials {Arias, 2011 #416;Cattoir, 2015 #411;Diaz, 2014 #417;Jung, 2010 #413;Soltani, 2000 #412}. The primers, cycling conditions and expected product sizes are shown in Table 2. Detection of SNPs in *liaR*, *liaS* and *rpsJ* was conducted by Sanger sequencing of PCR products (AGRF, Brisbane) and subsequent analysis in Vector NTi (Life Technologies, Australia).

Results

Prevalence and identity

In total, 1500 faecal samples comprising 910 beef cattle faeces, 290 dairy cattle faeces and 300 veal calf faeces were tested for the presence of *Enterococcus*. *Enterococcus* were isolated from 805 (88.5%) beef cattle faeces, 244 (84.1%) dairy cattle faeces and 247 (82.3%) veal calf faeces. Species specific PCR determined that 6.4% of all isolates were *E. faecalis* and 8.0% were *E. faecium*. Veal samples (14.3%) were significantly ($p < 0.05$) more likely to contain *E. faecalis* than dairy (3.1%) or

beef (4.8%) samples. No significant differences in prevalence were observed between the three animal groups for *E. faecium*.

Antimicrobial susceptibility testing

Sensititre evaluation

A total of 800 *Enterococcus* isolates comprising 96 *E. faecalis*, 120 *E. faecium*, and 584 *Enterococcus spp.* were submitted for AMR analysis using the Sensititre test system. The distribution of MICs for each antimicrobial and species group is shown in Table 3. Breakpoints are not available for unspeciated *Enterococcus* isolates and therefore resistance data is only shown for *E. faecium* and *E. faecalis*. Streptogramin MIC values for *E. faecalis* are not presented as this species is intrinsically resistant. Similarly, flavomycin MIC values for *E. faecium* are not shown as they are inherently nonsusceptible. Irrespective of animal group and species, resistance to flavomycin (77.3 – 88.9%) and lincomycin (77.8 – 100.0%) was common. Resistance to quinupristin / dalfoprisitin was observed in 38 – 48% of all *E. faecium* isolates but was not correlated with similar resistances to virginiamycin. There was a strong association between daptomcyin resistant *E. faecalis* and veal calves, however this was not considered to be statistically significant ($p < 0.05$). Resistance to tetracycline (2.3 – 13.0%) and erythromycin (0.0 – 13.6%) were observed in all three animal groups except for erythromycin resistance in *E. faecium* from veal calves. Furthermore, tigecycline resistance was only observed in *E. faecium* and *E. faecalis* from grass-fed animals, and whilst tetracycline resistance in *E. faecalis* was more common in grain-fed isolates, the opposite relationship existed in *E. faecium* with tetracycline resistance only detected in isolates from grass-fed animals.

Genotypic investigation of AMR

Initial evaluation of AMR in the *Enterococcus* isolates identified resistance to antimicrobials of human clinical significance. In particular, resistance to quinupristin/dalfopristin, daptomycin and tigecycline was noted. Critically the quinupristin/dalfopristin resistance was observed in the absence of resistance to the other streptogramin antimicrobial tested, virginiamycin. Similarly, resistance to daptomycin and tigecycline was higher than anticipated and worthy of follow up investigation. All isolates including those not identified as *E. faecalis* or *E. faecium*, exhibiting MICs greater than the clinical breakpoints for quinupristin/dalfopristin, daptomycin or tigecycline were tested by PCR for a range of genetic markers known to be associated with resistance to these antimicrobials. In total, 314 quinupristin/dalfopristin resistant isolates, 42 daptomycin resistant isolates and 22 tigecycline resistant isolates were tested further.

Isolates exhibiting quinupristin/dalfopristin resistance in the Sensititre system were tested for the presence of the streptogramin acetyltransferases *satA*, *satG*, *vatG* and the ABC transporter *vgaD*. All 314 isolates tested negative for *satA*, *satG*, *vatG* and *vgaD*. Strains lacking these genes were further tested using the oligo set strepto-M and strepto-N which are designed to identify conserved motifs of potentially novel streptogramin resistance genes. Five (1.6%) of 314 isolates gave a PCR product of expected size using the strepto-M and strepto-N oligo set. Four of the five isolates were *E. faecium* from adult beef cattle slaughtered at separate abattoirs. Two of the cattle had been grass-fed and two were grain-fed. The remaining positive was identified an unspeciated *Enterococcus* isolated from a grain-fed adult beef animal.

For tigecycline resistant isolates, fragments of *rpsJ* were amplified, sequenced and analysed for a SNP that encodes a predicted amino acid change of Asp60 to Tyr. Fragments of *rpsJ* were amplified from all three *E. faecium* isolates and from 14 (82.4%) of 17 *Enterococcus* spp. isolates. Fragments of *rpsJ* were not amplified in either of the two *E. faecalis* isolates. Analysis of the 17 *rpsJ* fragments determined that none of the isolates harboured the SNP that has been shown to be associated with reduced susceptibility to tigecycline. Daptomycin resistant isolates were tested for the presence of SNPs in *liaR* and *liaS*. PCR products for *liaR* or *liaS* were generated from eight (88.9%) of nine *E. faecalis* isolates and all three *E. faecium* isolates but was found in only two (6.7%) of 30 *Enterococcus* spp. isolates. *E. faecalis* strains were most likely to harbour *liaR* on its own whereas *E. faecium* were more likely to harbour *liaS*. One *E. faecalis* and one *E. faecium* isolate were shown to contain both *liaR* and *liaS*. Sequencing of the PCR fragments determined that none of the isolates possessed the Thr120 to Ala SNP in *liaR* or the Trp73 to Cys SNP in *liaS*.

Additional phenotypic AMR testing

In the absence of corroborating genotypic data for the observed Sensititre test results all isolates that had demonstrated resistance to quinupristin/dalfopristin, daptomycin or tigecycline were subjected to an additional round of phenotypic AMR. Quinupristin/dalfopristin resistant isolates were tested in duplicate using disc susceptibility testing. Of the 50 *E. faecium* isolates tested two isolates demonstrated resistance to quinupristin/dalfopristin with a further two isolates showing reduced susceptibility. The remaining 46 isolates would be considered susceptible to quinupristin/dalfopristin. Although breakpoints don't exist for non *E. faecium* and *E. faecalis* isolates,

application of the same criteria to the *Enterococcus* spp. isolates would have identified two resistant isolates, 49 isolates with reduced susceptibility and 213 isolates that would be considered susceptible. Of the four *E. faecium* isolates demonstrating resistance or reduced susceptibility to quinupristin/dalfopristin three had previously tested positive using the strepto-M/strepto-N oligo set suggesting that these isolates may harbour an as yet unidentified genetic basis for the observed results.

Testing of daptomycin and tigecycline was completed in duplicate on each of the isolates that had previously demonstrated resistance to these antimicrobials in the Sensititre system. MICs for tigecycline were all below the clinical breakpoint and therefore all isolates should be considered susceptible to tigecycline. The three *E. faecium* and two *E. faecalis* isolates all had an MIC of 0.12 µg/mL and are consistent with wild-type strains. None of the *Enterococcus* spp. isolates had MICs greater than the clinical breakpoint used for *E. faecium* and *E. faecalis*. Similarly, the three *E. faecium* and nine *E. faecalis* isolates previously identified as resistant to daptomycin all had MICs below the clinical breakpoint on re-testing. The *E. faecium* isolates all had MICs of 2 µg/mL whereas the *E. faecalis* isolates ranged from 0.25 to 2 µg/mL. One *Enterococcus* spp. isolate had an elevated MIC of 8 µg/mL, however all remaining isolates were below the clinical breakpoint concentration.

Re-categorisation of daptomycin, tigecycline and quinupristin/dalfopristin results

The inability to reproduce the findings of the primary phenotypic antimicrobial testing conducted using the Sensititre test system with custom AMR plates and the absence of the identification of AMR-linked genetic markers suggests that the original phenotypic assessment for resistance to daptomycin, tigecycline and quinupristin/dalfopristin is comprised of major errors. The major error rates for *E. faecalis* isolates against daptomycin and tigecycline were 9.4% and 2.1%, respectively. Whilst the major errors rates for *E. faecium* isolates against daptomycin and tigecycline were low at 2.5% for both, the major error rate associated with quinupristin/dalfopristin testing was 40%. The acceptance of major errors in the assessment of the three antimicrobials results in the re-categorisation of the results for the *E. faecalis and E. faecium* isolates in questions. A full summary of the revised results are shown in Table 4 and Figure 1.

Antibiogram profiles

Resistance to three or more classes of antimicrobial (MDR) was observed in 18 (8.3%) of all *E. faecium* and *E. faecalis* isolates. Table 5 shows the antibiograms for each *E. faecium* and *E. faecalis*. MDR was observed in 6 (5.0%) *E. faecium* and 12 (12.5%) *E. faecalis* isolates. Resistance to four or more antimicrobial classes was less commonly observed with only four (3.3%) *E. faecium* isolates falling into this category. Antibiogram profiles of *E. faecium* were dominated by resistance to lincomycin (78.3%). FLV-LIN was the most common antibiogram associated with *E. faecalis* isolates with 53 (55.2%) of 96 isolates harbouring this combination. The main MDR profiles for *E. faecium* and *E. faecalis* were ERY-LIN-TET (3.3%) and ERY-FLV-LIN (5.2%), respectively.

Discussion

Bacteria that are resistant to antimicrobials are of increased concern to public health officials throughout the world as they may compromise the ability of various treatment regimes to address disease and infection in human medical settings. Knowledge and understanding of the types of AMR present in food production animals is key to determining the ongoing risk that AMR bacteria pose to human health. Australia currently does not have a nationally coordinated program for the ongoing surveillance and analysis of AMR bacteria in animals, bacteria in food derived from animals, or bacteria from humans. Consequently it relies heavily on routine testing of human and animal clinical isolates as well as infrequent surveys of isolates from animals or from food of animal origin to understand AMR development and trends.

Enterococci are ubiquitous bacteria that demonstrate intrinsic resistance to a number of first-line antimicrobial agents and have also demonstrated capacity to rapidly acquire resistance to antimicrobials of high clinical importance including quinolones, macrolides, tetracyclines, streptogramins and glycopeptides (Hammerum 2012). They are also frequently associated with mobile genetic elements harbouring AMR genes and have the potential for resistance to virtually all antimicrobials of importance to human medicine (Ramos, Igrejas et al. 2012). The importance of Enterococci as the third most commonly isolated nosocomial pathogen (Hidron, Edwards et al. 2008) and the clear relationship between exposure to parental antimicrobials and the development of resistance (Hollenbeck and Rice 2012) warrants their ongoing inclusion in any human, animal or food AMR surveillance program. *Enterococcus faecalis* and *E. faecium* were recovered from 6.4% and 8.0% of samples in this survey and although they are the two Enterococcal species most associated with human infections, monitoring of environmental enterococci is useful as it may provide insights

to trends of MIC's which may be of concern to the more clinically relevant species. From a human clinical perspective, resistance in *E. faecalis* and *E. faecium* to ampicillin, vancomycin, linezolid, daptomycin, quinupristin / dalfopristin (*E. faecium* only) and tigecycline are the key issues. Resistance to other older antimicrobials such as lincomycin, flavomycin, tetracycline and erythromycin are seldom considered as either resistance is common or the antimicrobials are seldom used in human medicine (Hollenbeck and Rice 2012). The findings of this study reinforce this segregation of concern with increased levels of resistance to lincomycin, flavomycin (*E. faecalis* only), tetracycline and erythromycin observed in *E. faecium* and *E. faecalis* isolates from all animal groups.

Conversely, resistance to antimicrobials of critical and high importance to human medicine is of much greater concern to the ongoing treatment strategies for Enterococcal infections. Resistance to ampicillin, linezolid and vancomycin was not observed in this study in any *E. faecium* or *E. faecalis* isolates. This is significant as ampicillin remains the preferred therapy for uncomplicated enterococcal infections. Similarly, the absence of vancomycin resistant enterococcus assists in maintaining optimal treatment options. Linezolid and quniupristin / dalfopristin are suggested therapies for vancomycin resistant enterococcus infections. Although resistance to linezolid was not observed, 41.7% of *E. faecium* isolates were shown to be resistant to quinupristin / dalfopristin when tested using the Sensititre test system. Surprisingly, resistance to quinupristin / dalfopristin did not correlate well with the resistance observed to another streptogramin antimicrobial, virginiamycin. A variety of genes have been identified that give rise to streptogramin resistance in Enterococci (Soltani, Beighton et al. 2000; Jung, Shin et al. 2010). Examination of the *E. faecium* isolates for these genes did not, in general, identify their presence and additional phenotypic testing did not support the original phenotypic findings of the Sensititre testing suggesting that the quinupristin/dalfopristin results were major errors of the Sensititre testing system. Two *E. faecium* isolates did have reproducible quinupristin/dalfopristin resistance despite known streptogramin resistance genes being absent. Both strains appear to harbour conserved core regions of streptogramin resistance genes and require whole genome sequence analysis to identify potentially yet to be described streptogramin resistance genes.

Resistance to daptomycin and tigecycline initially observed with the Sensititre system could not be replicated using gradient diffusion techniques. Publications detailing genes conferring resistance to

these antimicrobials in *Enterococcus* isolates are extremely limited, however whole genome analysis of strains demonstrating reduced susceptibility have identified a number of single nucleotide polymorphisms (SNPs) present in those isolates when compared with wild-type populations (Aarestrup 1999; Arias, Panesso et al. 2011; Diaz, Tran et al. 2014; Cattoir, Isnard et al. 2015). Investigation of the SNPs in the *liaFSR* regulon and *rpsJ* determined that the isolates in this study are identical to wild-type isolates. When combined with the agar dilution results the original Sensititre test results are believed to be major errors and the overall data set has been modified to reflect these findings. As a consequence this study reports that resistance to the critical or high importance antimicrobials linezolid, daptomycin, tigecycline and vancomycin was not present in Enterococcal isolates from Australian cattle populations regardless of source.

The generation of discordant AMR results after testing with multiple phenotypic test systems is concerning though not confined to this study alone. Several studies have detailed discrepancies in essential agreement and categorical agreement between test systems when single antimicrobial / bacteria combinations are considered. The United States Food and Drug Administration will approve the marketing of AMR tests system provided that very major errors (false-negatives) and major errors (false-positives) do not exceed 1.5% and 3% respectively and essential MIC agreement within one doubling MIC dilution of >90% occurs between the test system and the reference CLSI method (Jorgensen and Ferraro 2009). This study has identified major errors with quinupristin/dalfopristin, daptomycin and tigecycline, however only the combination of quinupristin/dalfopristin with *E. faecium* and daptomycin with *E. faecalis* strains exceed the allowable 3% major error rate. It is not possible to suggest an alternative explanation for the quinupristin errors other than to accept there is likely to have been undetermined issues that arose during the custom plate manufacturing process. On the other hand, the *in vitro* evaluation of daptomycin and tigecycline resistance has been shown to be highly dependent on the culture conditions used may provide an explanation for the higher than acceptable major error rate observed in this study (Butaye, Devriese et al. 1998; Rathe, Kristensen et al. 2010; Kelesidis, Humphries et al. 2011).

Overall, the results corroborate previous Australian based animal and retail food surveys that have shown a low level of AMR, relatively small proportions of MDR and most importantly the maintenance of susceptibility to most antimicrobials of critical and high importance to human health (Barton, Pratt et al. 2003; DAFF 2007; Barlow and Gobius 2008). Importantly, it would appear that

the production practices at work in Australian cattle populations are not generating pools of resistance that are likely to result in the inability to treat human infections caused by enterococci. Nevertheless, it is necessary to maintain strict guidelines and controls around the use of antimicrobials in food-production animals in Australia and monitoring the effects of all antimicrobial use is required to support Australia's reputation as a supplier of safe and healthy food.

References

Aarestrup, F. M. (1999). "Association between the consumption of antimicrobial agents in animal husbandry and the occurrence of resistant bacteria among food animals." Int J Antimicrob Agents **12**(4): 279-285.

Arias, C. A., D. Panesso, et al. (2011). "Genetic Basis for In Vivo Daptomycin Resistance in Enterococci." New England Journal of Medicine **365**(10): 892-900.

Barlow, R. S. and K. S. Gobius (2008). Pilot survey for antimicrobial resistant (AMR) bacteria in Australian food. CSIRO. Cannon Hill, Queensland.

Barton, M. D., R. Pratt, et al. (2003). "Antibiotic resistance in animals." Commun Dis Intell **27 Suppl**: S121-126.

Butaye, P., L. A. Devriese, et al. (1998). "Effects of different test conditions on MICs of food animal growth-promoting antibacterial agents for enterococci." Journal of Clinical Microbiology **36**(7): 1907-1911.

Cattoir, V., C. Isnard, et al. (2015). "Genomic Analysis of Reduced Susceptibility to Tigecycline in Enterococcus faecium." Antimicrobial Agents and Chemotherapy **59**(1): 239-244.

Collignon, P. and F. J. Angulo (2006). "Fluoroquinolone-resistant Escherichia coli: food for thought." The Journal of infectious diseases **194**(1): 8-10.

Collignon, P., J. H. Powers, et al. (2009). "World Health Organization Ranking of Antimicrobials According to Their Importance in Human Medicine: A Critical Step for Developing Risk Management Strategies for the Use of Antimicrobials in Food Production Animals." Clinical Infectious Diseases **49**(1): 132-141.

DAFF (2007). "Pilot surveillance program for antimicrobial resistance in bacteria of animal origin." Australian Government Department of Agriculture, Fisheries and Forestry.

Diaz, L., T. T. Tran, et al. (2014). "Whole-Genome Analyses of Enterococcus faecium Isolates with Diverse Daptomycin MICs." Antimicrobial Agents and Chemotherapy **58**(8): 4527-4534.

Dutkamalen, S., S. Evers, et al. (1995). "Detection of Glycopeptide Resistance Genotypes and Identification to the Species Level of Clinically Relevant Enterococci by Pcr." Journal of Clinical Microbiology **33**(1): 24-27.

Dutkamalen, S., S. Evers, et al. (1995). "Detection of Glycopeptide Resistance Genotypes and Identification to the Species Level of Clinically Relevant Enterococci by Pcr (Vol 33, Pg 25, 1995)." Journal of Clinical Microbiology **33**(5): 1434-1434.

Hammerum, A. M. (2012). "Enterococci of animal origin and their significance for public health." Clinical Microbiology and Infection **18**(7): 619-625.

Heuer, O. E., A. M. Hammerum, et al. (2006). "Human health hazard from antimicrobial-resistant enterococci in animals and food." Clinical infectious diseases : an official publication of the Infectious Diseases Society of America **43**(7): 911-916.

Hidron, A. I., J. R. Edwards, et al. (2008). "Antimicrobial-Resistant Pathogens Associated With Healthcare-Associated Infections: Annual Summary of Data Reported to the National Healthcare Safety Network at the Centers for Disease Control and Prevention, 2006-2007." Infection Control and Hospital Epidemiology **29**(11): 996-1011.

Hollenbeck, B. L. and L. B. Rice (2012). "Intrinsic and acquired resistance mechanisms in enterococcus." Virulence **3**(5): 421-433.

Jorgensen, J. H. and M. J. Ferraro (2009). "Antimicrobial Susceptibility Testing: A Review of General Principles and Contemporary Practices." Clinical Infectious Diseases **49**(11): 1749-1755.

Jung, Y. H., E. S. Shin, et al. (2010). "Characterization of Two Newly Identified Genes, vgaD and vatG, Conferring Resistance to Streptogramin A in Enterococcus faecium." Antimicrobial Agents and Chemotherapy **54**(11): 4744-4749.

Ke, D. B., F. J. Picard, et al. (1999). "Development of a PCR assay for rapid detection of *Enterococci*." Journal of Clinical Microbiology **37**(11): 3497-3503.

Kelesidis, T., R. Humphries, et al. (2011). "Daptomycin Nonsusceptible Enterococci: An Emerging Challenge for Clinicians." Clinical Infectious Diseases **52**(2): 228-234.

Meat & Livestock Australia (2013). Fast Facts 2013; Australia's beef industry.

Ramos, S., G. Igrejas, et al. (2012). "Antibiotic resistance and mechanisms implicated in fecal enterococci recovered from pigs, cattle and sheep in a Portuguese slaughterhouse." Annals of Microbiology **62**(4): 1485-1494.

Rathe, M., L. Kristensen, et al. (2010). "Vancomycin-resistant Enterococcus spp.: validation of susceptibility testing and in vitro activity of vancomycin, linezolid, tigecycline and daptomycin." Apmis **118**(1): 66-73.

Soltani, M., D. Beighton, et al. (2000). "Mechanisms of resistance to quinupristin-dalfopristin among isolates of Enterococcus faecium from animals, raw meat, and hospital patients in western Europe." Antimicrobial Agents and Chemotherapy **44**(2): 433-436.

Szmolka, A. and B. Nagy (2013). "Multidrug resistant commensal Escherichia coli in animals and its impact for public health." Frontiers in microbiology **4**: 258.

Walsh, C. and S. Fanning (2008). "Antimicrobial resistance in foodborne pathogens--a cause for concern?" Current drug targets **9**(9): 808-815.

World Health Organisation (2011). Critically important antimicrobials for human medicine - 3rd revision, World Health Organization.

Table 1. Dilution ranges and breakpoints for antimicrobial susceptibility testing

Antimicrobial	*Enterococcus*	
	Range	**Breakpoint**
Ampicillin	0.5-16	≥16
Chloramphenicol	2-32	≥32
Daptomycin	0.125-4	≥8
Erythromycin	0.25-8	≥8
Flavomycin	1-32	≥32
Gentamicin	32-1024	≥512
Kanamycin	128-1024	≥1024
Lincomycin	1-32	≥8
Linezolid	0.5-8	≥8
Penicillin	0.5-16	≥16
Quinupristin / dalfopristin	0.25-8	≥4
Streptomycin	256-1024	≥1024
Teicoplanin	0.125-4	2
Tetracycline	2-16	≥16
Tigecycline	0.016-0.5	≥0.5
Vancomycin	0.25-32	≥32
Virginiamycin	1-32	>8

Table 2. Primers, cycling conditions and expected product sizes of *Enterococcus* AMR gene PCRs

Resistance to:	**Oligo (5' – 3')**	**Cycling conditions**	**Products size (bp)**	**Reference**
Quinupristin/ Dalfopristin	satA-1: GCTCAATAGGACCAGGTGTA satA-2: TCCAGCTAACATGTATGGCG	1 min 94°C, 1 min 55°C, 1 min 72°C x 35	273	(Soltani, Beighton et al. 2000)
	satG-1: ACTATACCTGACGCAAATGC satG-2: GGTTCAAATCTTGGTCCG	25s 94°C, 40s 55°C, 50s 72°C x 30	513	
	strep-M:ATHATGAAYGGIGCIAAYCAYMGIATG strep-N:ICCDATCCAIACRTCRTTICC	2 min 40°C, 90s 72°C, 30s 95°C x 35	144 or 147	
	vatG-1:GTGGGAAAAGCATACACCT vatG-2:TTGCAGGATTACCACCAAC	30s 94°C, 30s 55°C, 30s 72°C x 30	200	(Jung, Shin et al. 2010)
	vgaD-1:CAACTGGAGCGAGCTGTTA vgaD-2:GACAGCCGGATAATCTTTTG	30s 94°C, 30s 55°C, 30s 72°C x 30	201	
Daptomycin	liaR-F:GGTCCGATCATCCACATCTA liaR-R:CCGTTTAGGCGTTTCATCAT	30s 94°C, 30s 60°C, 30s 72°C x 30	553	This study
	liaS-F:AAAGTCATTGGTGGGGAGAA liaS-R:GACTGGGAAGCGTTGATGAT	30s 94°C, 30s 60°C, 30s 72°C x 30	526	

Tigecycline	rpsJ-F:AGAGGTTGCGACACGCCCGG rpsJ-R:TCTACAACAGTTACTGGAAT	30s 94°C, 30s 60°C, 30s 72°C x 30	525	(Cattoir, Isnard et al. 2015)

Table 3. Distribution of MICs and occurrence of resistance among *Enterococcus* isolates from beef cattle, dairy cattle and veal calf faecal samples using the Sensititre test system

Class	Antimicrobial	Species	N =	% Resistant	95% CI	Antimicrobial concentration (µg/ml)																
						0.015	0.03	0.06	0.12	0.25	0.5	1	2	4	8	16	32	64	128	256	512	1024	>1024
Aminoglycoside	Gentamicin	*Enterococcus faecalis*	96	0.0	0.00 - 3.77												100.0						
		Enterococcus faecium	120	0.0	0.00 - 3.03												99.2	0.8					
		Enterococcus spp	584	NA	NA												99.7	0.3					
	Kanamycin	*Enterococcus faecalis*	96	1.0	0.03 - 5.67														92.7	5.2	1.0	1.0	
		Enterococcus faecium	120	0.8	0.02 - 4.56														99.2			0.8	
		Enterococcus spp	584	NA	NA														95.0	3.8	0.3	0.3	0.5
	Streptomycin	*Enterococcus faecalis*	96	1.0	0.03 - 5.67															99.0			1.0
		Enterococcus faecium	120	0.0	0.00 - 3.03															100.0			
		Enterococcus spp	584	NA	NA															99.1			0.9
Glycopeptides	Teicoplanin	*Enterococcus faecalis*	96	0.0	0.00 - 3.77				30.2	46.9	17.7	5.2											
		Enterococcus faecium	120	0.0	0.00 - 3.03				42.5	40.0	14.2	3.3											
		Enterococcus spp	584	NA	NA				33.2	41.1	19.7	6.0											
	Vancomycin	*Enterococcus faecalis*	96	0.0	0.00 - 3.77					1.0	40.6	34.4	11.5	10.4	2.1								
		Enterococcus faecium	120	0.0	0.00 - 3.03						46.7	22.5	15.0	15.0	0.8								
		Enterococcus spp	584	NA	NA					0.5	41.6	28.3	13.4	12.3	3.9								
Glycylcycline	Tigecycline	*Enterococcus faecalis*	96	2.1	0.25 - 7.32	2.1	45.8	41.7	3.1	5.2	2.1												
		Enterococcus faecium	120	2.5	0.52 - 7.13		43.3	44.2	5.8	4.2	2.5												
		Enterococcus spp	584	NA	NA	2.1	45.0	37.7	8.0	4.3	2.7	0.2											
Lincosamide	Lincomycin	*Enterococcus faecalis*	96	85.4	76.74 - 91.79							10.4	1.0	3.1	8.3	25.0	34.4	17.7					
		Enterococcus faecium	120	94.2	88.35 - 97.62							4.2	0.8	0.8	2.5	27.5	50.0	14.2					
		Enterococcus spp	584	NA	NA							12.3	1.9	2.1	6.2	30.7	37.5	9.4					
Lipopeptide	Daptomycin	*Enterococcus faecalis*	96	9.4	4.38 - 17.05								75.0	15.6	9.4								
		Enterococcus faecium	120	2.5	0.52 - 7.13							5.8	45.0	46.7	2.5								
		Enterococcus spp	584	NA	NA				0.2		3.3	24.8	40.6	24.5	6.7								
Macrolide	Erythromycin	*Enterococcus faecalis*	96	10.4	5.11 - 18.32					33.3	15.6	14.6	20.8	5.2	1.0	9.4							
		Enterococcus faecium	120	8.3	4.07 - 14.79					53.3	13.3	12.5	9.2	3.3	4.2	4.2							
		Enterococcus spp	584	NA	NA					45.9	13.2	11.0	17.6	6.3	1.2	4.8							

Class	Antimicrobial	Species	N =	% Resistant	95% CI	Antimicrobial concentration (µg/ml)																
						0.015	0.03	0.06	0.12	0.25	0.5	1	2	4	8	16	32	64	128	256	512	1024	>1024
Oxazolidinones	Linezolid	*Enterococcus faecalis*	96	0.0	0.00 - 3.77							6.3	87.5	6.3									
		Enterococcus faecium	120	0.0	0.00 - 3.03							5.8	87.5	6.7									
		Enterococcus spp	584	NA	NA						0.3	5.0	88.2	6.5									
Penicillins	Ampicillin	*Enterococcus faecalis*	96	0.0	0.00 - 3.77						44.8	52.1	3.1										
		Enterococcus faecium	120	0.0	0.00 - 3.03						30.0	60.0	10.0										
		Enterococcus spp	584	NA	NA						45.9	29.5	22.3	2.4									
	Penicillin	*Enterococcus faecalis*	96	0.0	0.00 - 3.77						13.5	28.1	26.0	31.3	1.0								
		Enterococcus faecium	120	0.0	0.00 - 3.03						12.5	34.2	30.8	20.8	1.7								
		Enterococcus spp	584	NA	NA						19.3	27.1	22.8	21.2	9.2	0.2	0.2						
Phenicol	Chloramphenicol	*Enterococcus faecalis*	96	0.0	0.00 - 3.77									17.7	82.3								
		Enterococcus faecium	120	0.0	0.00 - 3.03								1.7	51.7	40.8	5.8							
		Enterococcus spp	584	NA	NA								0.3	38.4	59.4	1.9							
Phosphoglycolipid	Flavomycin#	*Enterococcus faecalis*	96	80.2	70.83 - 87.64							15.6	1.0		2.1	1.0	2.1	78.1					
		Enterococcus faecium	NA	NA	NA																		
		Enterococcus spp	584	NA	NA							11.8	1.0	0.7	1.0	0.7	2.1	82.7					
Streptogramins	Quinupristin / dalfopristin*	*Enterococcus faecalis*	NA	NA	NA																		
		Enterococcus faecium	120	41.7	32.74 - 51.02						5.0	1.7	51.7	30.0	6.7	5.0							
		Enterococcus spp	584	NA	NA						11.8	6.2	36.8	35.1	7.2	2.9							
	Virginiamycin*	*Enterococcus faecalis*	NA	NA	NA																		
		Enterococcus faecium	120	0.0	0.00 - 3.03							85.0	5.8	9.2									
		Enterococcus spp	584	NA	NA							86.6	6.8	6.0	0.2	0.3							
Tetracycline	Tetracycline	*Enterococcus faecalis*	96	7.3	2.98 - 14.45								86.5	2.1	4.2	3.1	4.2						
		Enterococcus faecium	120	11.7	6.53 - 18.80								78.3	8.3	1.7	1.7	10.0						
		Enterococcus spp	584	NA	NA								83.2	3.9	2.1	1.9	8.9						

[*] *Enterococcus faecalis* isolates are intrinsically resistant to streptogramins; [#] *Enterococcus faecium* isolates are inherently nonsuceptible to flavomycin.

Solid vertical lines indicate breakpoints for resistance. The white fields indicate the dilution range tested for each antimicrobial. Values in the shaded area indicate MIC values greater than the highest concentration tested.

Table 4. Distribution of MICs and occurrence of resistance among *Enterococcus* isolates from beef cattle, dairy cattle and veal calf faecal samples following additional phenotypic and genotypic assessment.

Class	Antimicrobial	Species	N =	% Resistant	95% CI	Antimicrobial concentration (µg/ml)																
						0.015	0.03	0.06	0.12	0.25	0.5	1	2	4	8	16	32	64	128	256	512	1024	>1024
Aminoglycoside	Gentamicin	*Enterococcus faecalis*	96	0.0	0.00 - 3.77												100.0						
		Enterococcus faecium	120	0.0	0.00 - 3.03												99.2	0.8					
		Enterococcus spp	584	NA	NA												99.7	0.3					
	Kanamycin	*Enterococcus faecalis*	96	1.0	0.03 - 5.67														92.7	5.2	1.0	1.0	
		Enterococcus faecium	120	0.8	0.02 - 4.56														99.2			0.8	
		Enterococcus spp	584	NA	NA														95.0	3.8	0.3	0.3	0.5
	Streptomycin	*Enterococcus faecalis*	96	1.0	0.03 - 5.67															99.0			1.0
		Enterococcus faecium	120	0.0	0.00 - 3.03															100.0			
		Enterococcus spp	584	NA	NA															99.1			0.9
Glycopeptides	Teicoplanin	*Enterococcus faecalis*	96	0.0	0.00 - 3.77				30.2	46.9	17.7	5.2											
		Enterococcus faecium	120	0.0	0.00 - 3.03				42.5	40.0	14.2	3.3											
		Enterococcus spp	584	NA	NA				33.2	41.1	19.7	6.0											
	Vancomycin	*Enterococcus faecalis*	96	0.0	0.00 - 3.77					1.0	40.6	34.4	11.5	10.4	2.1								
		Enterococcus faecium	120	0.0	0.00 - 3.03						46.7	22.5	15.0	15.0	0.8								
		Enterococcus spp	584	NA	NA					0.5	41.6	28.3	13.4	12.3	3.9								
Glycylcycline	Tigecycline	*Enterococcus faecalis*	96	0.0	0.00 - 3.77	2.1	45.8	41.7	5.2	5.2													
		Enterococcus faecium	120	0.0	0.00 - 3.03		43.3	44.2	8.3	4.2													
		Enterococcus spp	584	NA	NA	2.1	45.0	38.0	10.3	4.6													
Lincosamide	Lincomycin	*Enterococcus faecalis*	96	85.4	76.74 - 91.79							10.4	1.0	3.1	8.3	25.0	34.4	17.7					
		Enterococcus faecium	120	94.2	88.35 - 97.62							4.2	0.8	0.8	2.5	27.5	50.0	14.2					
		Enterococcus spp	584	NA	NA							12.3	1.9	2.1	6.2	30.7	37.5	9.4					
Lipopeptide	Daptomycin	*Enterococcus faecalis*	96	0.0	0.00 - 3.77					2.1	4.2	2.1	76.0	15.6									
		Enterococcus faecium	120	0.0	0.00 - 3.03							5.8	47.5	46.7									
		Enterococcus spp	584	NA	NA				0.2	0.7	5.3	25.5	40.9	27.2	0.2								
Macrolide	Erythromycin	*Enterococcus faecalis*	96	10.4	5.11 - 18.32					33.3	15.6	14.6	20.8	5.2	1.0	9.4							
		Enterococcus faecium	120	8.3	4.07 - 14.79					53.3	13.3	12.5	9.2	3.3	4.2	4.2							
		Enterococcus spp	584	NA	NA					45.9	13.2	11.0	17.6	6.3	1.2	4.8							

Class	Antimicrobial	Species	N =	% Resistant	95% CI	Antimicrobial concentration (µg/ml)																
						0.015	0.03	0.06	0.12	0.25	0.5	1	2	4	8	16	32	64	128	256	512	1024	>1024
Oxazolidinones	Linezolid	*Enterococcus faecalis*	96	0.0	0.00 - 3.77							6.3	87.5	6.3									
		Enterococcus faecium	120	0.0	0.00 - 3.03							5.8	87.5	6.7									
		Enterococcus spp	584	NA	NA						0.3	5.0	88.2	6.5									
Penicillins	Ampicillin	*Enterococcus faecalis*	96	0.0	0.00 - 3.77						44.8	52.1	3.1										
		Enterococcus faecium	120	0.0	0.00 - 3.03						30.0	60.0	10.0										
		Enterococcus spp	584	NA	NA						45.9	29.5	22.3	2.4									
	Penicillin	*Enterococcus faecalis*	96	0.0	0.00 - 3.77						13.5	28.1	26.0	31.3	1.0								
		Enterococcus faecium	120	0.0	0.00 - 3.03						12.5	34.2	30.8	20.8	1.7								
		Enterococcus spp	584	NA	NA						19.3	27.1	22.8	21.2	9.2	0.2	0.2						
Phenicol	Chloramphenicol	*Enterococcus faecalis*	96	0.0	0.00 - 3.77									17.7	82.3								
		Enterococcus faecium	120	0.0	0.00 - 3.03								1.7	51.7	40.8	5.8							
		Enterococcus spp	584	NA	NA								0.3	38.4	59.4	1.9							
Phosphoglycolipid	Flavomycin#	*Enterococcus faecalis*	96	80.2	70.83 - 87.64							15.6	1.0		2.1	1.0	2.1	78.1					
		Enterococcus faecium	NA	NA	NA																		
		Enterococcus spp	584	NA	NA							11.8	1.0	0.7	1.0	0.7	2.1	82.7					
Streptogramins	Quinupristin / dalfopristin*	*Enterococcus faecalis*	NA	NA	NA																		
		Enterococcus faecium	120	1.7	0.20 - 5.89						5.0	40.0	53.3	1.7									
		Enterococcus spp	584	NA	NA						11.8	42.6	45.2	0.2	0.2								
	Virginiamycin*	*Enterococcus faecalis*	NA	NA	NA																		
		Enterococcus faecium	120	0.0	0.00 - 3.03							85.0	5.8	9.2									
		Enterococcus spp	584	NA	NA							86.6	6.8	6.0	0.2	0.3							
Tetracycline	Tetracycline	*Enterococcus faecalis*	96	7.3	2.98 - 14.45								86.5	2.1	4.2	3.1	4.2						
		Enterococcus faecium	120	11.7	6.53 - 18.80								78.3	8.3	1.7	1.7	10.0						
		Enterococcus spp	584	NA	NA								83.2	3.9	2.1	1.9	8.9						

[*] *Enterococcus faecalis* isolates are intrinsically resistant to streptogramins; [#] *Enterococcus faecium* isolates are inherently nonsuceptible to flavomycin.

Solid vertical lines indicate breakpoints for resistance. The white fields indicate the dilution range tested for each antimicrobial. Values in the shaded area indicate MIC values greater than the highest concentration tested.

Yellow highlighted fields indicate differences between original and revised AMR results.

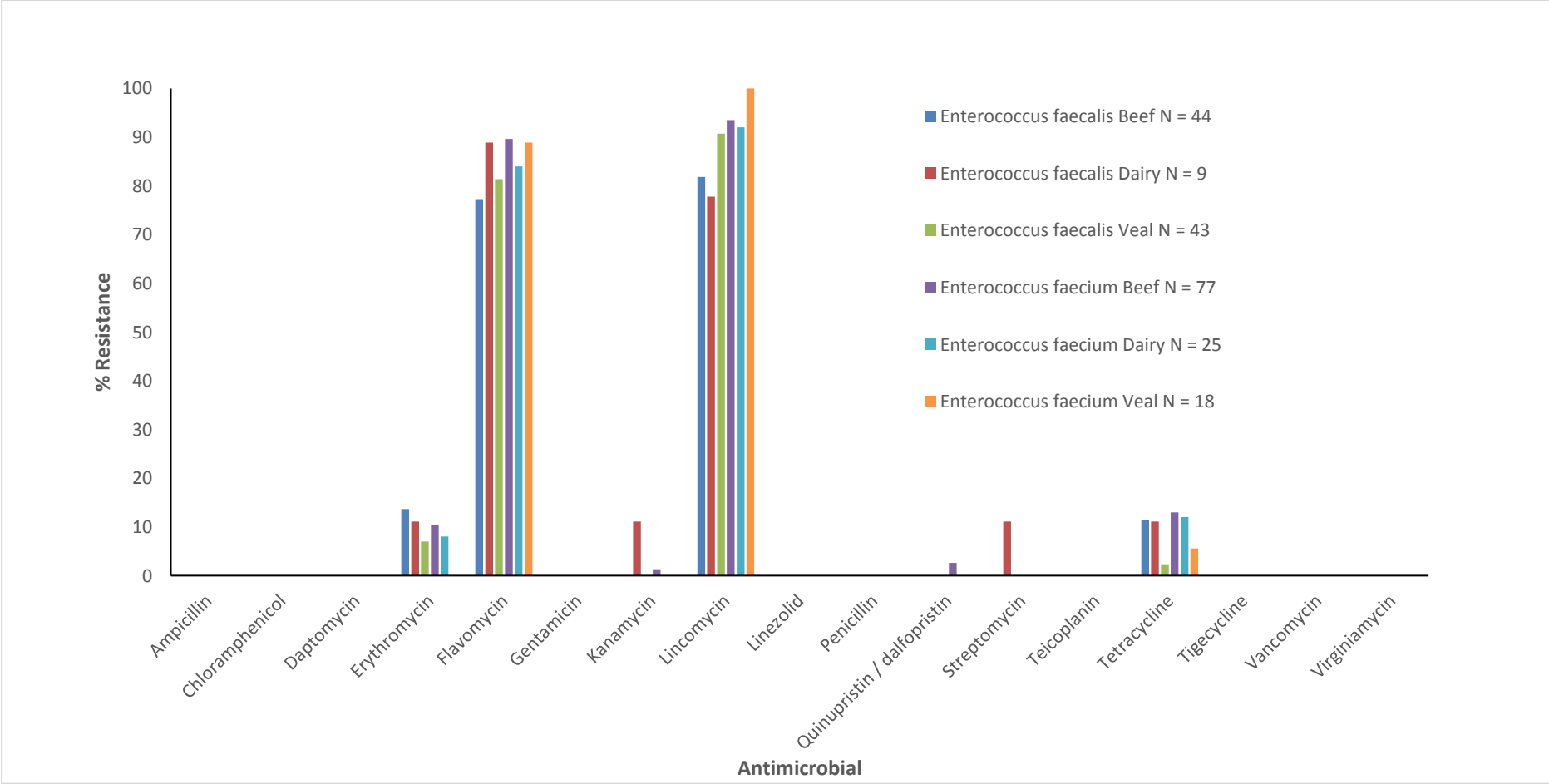


Figure 1. Prevalence of AMR in *Enterococcus* isolates from beef cattle, dairy cattle and veal calf faecal samples

Table 5. Antibiograms of *Enterococcus faecium* and *faecalis* isolates from beef cattle, dairy cattle and veal calf faecal samples

Antibiograms	*E. faecium* (N=120)	*E. faecalis* (N=96)
ALL SENSITIVE	6	1
FLV		13
LIN	94	17
TET	1	
ERY LIN	4	
FLV LIN		53
LIN SYN	1	
LIN TET	8	
ERY FLV LIN		5
ERY LIN TET	4	1
ERY LIN SYN	1	
ERY KAN LIN	1	
FLV LIN TET		2
ERY FLV LIN TET		3
ERY FLV KAN LIN STR TET		1

* FLV – flavomycin, LIN – lincomycin, TET – tetracycline, ERY – erythromycin, SYN – quinupristin/dalfopristin, KAN – kanamycin, STR - streptomycin

Appendix A: Genetic characterisation of AMR in *E. coli* from Australian beef cattle populations

Background

A 2013 survey of the phenotypic AMR status of 800 *E. coli* isolates collected from beef cattle, dairy cattle and veal calves concluded that the Australian beef industry has continued to minimise the development of AMR with resistance to antimicrobials of high or critical importance to human medicine remaining low. Nevertheless resistance to ceftiofur and amoxicillin/clavulanic acid was observed in a maximum of two *E. coli* isolates. Detailed characterisation of the resistances observed is required in order to assist in understanding how the resistance may have developed and what potential there may be for the resistance determinants to transfer from animals to humans via the food chain.

Materials and Methods

Samples

E. coli isolates PSEC240 and PSEC1780 are unrelated strains that were originally isolated from veal calves. Evaluation of their AMR status in the 2013 survey determined that EC240 was resistant to amoxicillin/clavulanic acid, ampicillin and cefazolin whereas EC1780 was resistant was resistant to amoxicillin/clavulanic acid, ampicillin, cefazolin, ceftiofur and tetracycline. Both isolates had been stored at -80°C prior to recovery at 37°C for 18-24h on 5% sheep blood agar (Biomerieux, Australia).

Phenotypic AMR testing

Isolates were evaluated for their reported resistances using both broth microdilution and disk diffusion assays. Broth microdilution testing was conducted using the Sensititre test system and the custom made AUSVN2 test plate. Antimicrobial susceptibility disks for ceftiofur, ampicillin and amoxicillin/clavulanic acid were used in conjunction with Mueller-Hinton agar (Oxoid, Australia) and incubated at 37°C for 24h.

.

Table A1: PCR protocols used in this study

Target gene	Oligo	Sequence (5' to 3')	Annealing temp (°C)	Reference
bla$_{CTX-M}$	CTX-M/F	TTTGCGATGTGCAGTACCAGTAA	60	(Edelstein, Pimkin et al. 2003)
	CTX-M/R	CGATATCGTTGGTGGTGCCATA		
bla$_{CMY}$	CMY-F	CCGGACACCTTTTTGCTTTT	60	(Sidjabat, Paterson et al. 2009)
	CMY-R	TATCCTGGGCCTCATCGTCAGTTA		
bla$_{TEM}$	blaTEM-F	GAGTATTCAACATTTTCGT	50	(Maynard, Fairbrother et al. 2003)
	blaTEM-R	ACCAATGCTTAATCAGTGA		
bla$_{TEM}$	MultiTSO-T_for	CATTTCCGTGTCGCCCTTATTC	60	(Dallenne, Da Costa et al. 2010)
	MultiTSO-T_rev	CGTTCATCCATAGTTGCCTGAC		
bla$_{SHV}$	MultiTSO-S_for	AGCCGCTTGAGCAAATTAAAC		
	MultiTSO-S_rev	ATCCCGCAGATAAATCACCAC		
OXA-1, OXA-4 and OXA-30	MultiTSO-O_for	GGCACCAGATTCAACTTTCAAG		
	MultiTSO-O_rev	GACCCCAAGTTTCCTGTAAGTG		
MOX (*bla*$_{MOX}$ or *bla*$_{CMY}$)	MOXMF	GCTGCTCAAGGAGCACAGGAT	64	(Perez-Perez and Hanson 2002)
	MOXMR	CACATTGACATAGGTGTGGTGC		
CIT (*bla*$_{LAT}$ or *bla*$_{CMY2}$)	CITMF	TGGCCAGAACTGACAGGCAAA		

	CITMR	TTTCTCCTGAACGTGGCTGGC
bla$_{DHA}$	DHAMF	AACTTTCACAGGTGTGCTGGGT
	DHAMR	CCGTACGCATACTGGCTTTGC
bla$_{ACC}$	ACCMF	AACAGCCTCAGCAGCCGGTTA
	ACCMR	TTCGCCGCAATCATCCCTAGC
EBC (*bla*$_{MIR1}$ or *bla*$_{ACT1}$)	EBCMF	TCGGTAAAGCCGATGTTGCGG
	EBCMR	CTTCCACTGCGGCTGCCAGTT
bla$_{FOX1-5}$	FOXMF	AACATGGGGTATCAGGGAGATG
	FOXMR	CAAAGCGCGTAACCGGATTGG

Genotypic AMR testing

Previously published PCR protocols were used to evaluate a variety of antimicrobial resistance genes known to be associated with ceftiofur and amoxicillin/clavulanic acid resistance. Table A1 lists each of the protocols used in this study

Results and Discussion

Phenotypic testing of PSEC240 and PSEC1780 using the Sensititre test system and disk diffusion produced highly reproducible results for amoxicillin/clavulanic acid and ampicillin. Both isolates had MICs greater than the maximum concentration tested for each antimicrobial and they repeatedly had zones of clearance which were less than the resistance cut-offs defined by CLSI. Reproducibility of the ceftiofur result for PSEC1780 was not observed. Repeated testing of the strain using the AUSVN2 custom plate and the Sensititre system did give results within one doubling dilution of between 4 and 8 µg/mL. As the MIC breakpoint for ceftiofur is ≥8 µg/mL the results observed with PSEC1780 confirms the isolate demonstrates reduced susceptibility to ceftiofur but fails to absolutely confirm its resistant status. CLSI nor EUCAST have published breakpoints for assessing ceftiofur resistance using disk diffusion assays. Nevertheless breakpoints have been proposed within the literature and they have been applied to this study (Burton, Thornsberry et al. 1996). Using the breakpoints proposed, PSEC1780 would be categorised as having intermediate resistance to ceftiofur.

Testing of PSEC240 and 1780 for the presence of genes commonly associated with resistance to amoxicillin/clavulanic acid and ceftiofur failed to identify the genetic basis of resistance. As bla_{TEM} was considered the gene most likely to be the basis of the amoxicillin/clavulanic acid resistance dual PCR protocols to detect this family of genes were employed. Both PCR protocols did not identify the presence of bla_{TEM}. An amplicon was produced using the primer pair CITMF and CITMF that detect the presence of LAT-1 to LATT4, CMY-2 to CMY-7 and BIL-1. However the amplicon was approximately 100bp larger than predicted and sequencing of the amplicon did not identify sequences consistent with the presence of the above mentioned genes.

Conclusion

This study has not identified the genetic basis of resistance to amoxicillin/clavulanic acid or ceftiofur for either of the *E. coli* isolates investigated. However, additional phenotypic AMR testing has confirmed that

the resistance to amoxicillin/clavulanic acid in the 2013 survey is not an error of the Sensititre testing system. The variability of the ceftiofur MIC for PSEC1780 when tested using the Sensititre system and the absence of bla_{CTX-M} and bla_{CMY} may suggest that the reduced susceptibility is more likely associated with the physiological state of the cell (e.g overexpression of efflux pumps) as opposed to the typical resistance gene determinants. It is proposed that both PSEC240 and PSEC1780 be sequenced and analysed for the presence of genes or other molecular markers that may explain the phenotypic resistances or reduced susceptibilities to antimicrobials observed in the 2013 survey and again in this study.

References

Burton, P. J., C. Thornsberry, et al. (1996). "Interpretive criteria for antimicrobial susceptibility testing of ceftiofur against bacteria associated with swine respiratory disease." Journal of Veterinary Diagnostic Investigation **8**(4): 464-468.

Dallenne, C., A. Da Costa, et al. (2010). "Development of a set of multiplex PCR assays for the detection of genes encoding important beta-lactamases in Enterobacteriaceae." Journal of Antimicrobial Chemotherapy **65**(3): 490-495.

Edelstein, M., M. Pimkin, et al. (2003). "Prevalence and molecular epidemiology of CTX-M extended-spectrum beta-lactamase-producing Escherichia coli and Klebsiella pneumoniae in Russian hospitals." Antimicrobial Agents and Chemotherapy **47**(12): 3724-3732.

Maynard, C., J. M. Fairbrother, et al. (2003). "Antimicrobial resistance genes in enterotoxigenic Escherichia coli O149 : K91 isolates obtained over a 23-year period from pigs." Antimicrobial Agents and Chemotherapy **47**(10): 3214-3221.

Perez-Perez, F. J. and N. D. Hanson (2002). "Detection of plasmid-mediated AmpC beta-lactamase genes in clinical isolates by using multiplex PCR." Journal of Clinical Microbiology **40**(6): 2153-2162.

Sidjabat, H. E., D. L. Paterson, et al. (2009). "Clinical Features and Molecular Epidemiology of CMY-Type beta-Lactamase-Producing Escherichia coli." Clinical Infectious Diseases **48**(6): 739-744.

Overall progress of the project

V.MFS.0333 is a multi-module project with the investigation of previously observed AMR comprising the first module. This Milestone describes the re-evaluation of the resistances initially observed using the Sensititre test system and enables the publishing of all AMR results with confidence. As expected the project will continue as scheduled and will shift its focus to activities relating to non-O157 pSTEC.

Recommendations

Proceed with project modules relating to the subtyping of non-O157 pSTEC isolates and the investigation of the efficiencies of existing and novel pSTEC test systems.

Milestone 3 – Key collaborator meetings – Attend and present pSTEC research findings at IAFP 2015 and VTEC 2015.

Abstract

VTEC 2015 and IAFP 2015 provide an unparalleled opportunity for international research groups to meet and exchange ideas in relation to pathogenic Shiga-toxigenic *Escherichia coli* research (pSTEC). Four CSIRO staff attended the conferences and a total of six presentations were given. Existing and new collaborative research opportunities were discussed with key project partners and a number of new relationships have formed post conference. The use of next generation sequencing (NGS) data to aid in understanding the risk burden of pSTEC was a prominent feature of both IAFP and VTEC. As NGS data is generated and the pan-genome of pSTEC further understood it is becoming apparent that the existing focus on serotypes as a way of categorising pSTEC may be inadequate and a move to the assessment of risk based on genetic attributes has been proposed. For example, the carriage of stx_{2a} and supershedding animals remain as major risk factors contributing to human illness. Despite advances in our understanding of pSTEC, there remains gaps in the knowledge around detection and ecological considerations such as shedding frequency, the impact of slaughter facilities on pSTEC contamination of product and why some cattle never shed pSTEC at all. The priority pSTEC research areas currently of interest to CSIRO and MLA appear in line with international research efforts in the pSTEC space. Ongoing access to the annual IAFP and triennial VTEC conferences should remain a key investment for an Australian red meat food safety program.

Table of Contents

1 Milestone description 46
2 Project objectives 46
3 Success in meeting the milestone 46
4 Discussion 46
4.1 International Association for Food Protection (IAFP 2015) 46
4.1.1 Presentation of MLA/CSIRO co-funded research 46
4.1.2 Collaborator meetings at IAFP 2015 47
4.1.3 Summary of conference 48
4.2 Symposium on Shiga Toxin (Verocytotoxin) Producing *Escherichia coli* Infections (VTEC 2015) 51
4.2.1 Presentation of MLA/CSIRO co-funded research 51
4.2.2 Collaborator meetings at VTEC 2015 52
4.2.3 Summary of conference 52
5 Conclusions/recommendations 60
6 Appendix A 19
6.1 IAFP 2015 and VTEC 2015 poster presentations 19

Milestone description

Milestone 3 – Key collaborator meetings – Attend and present pSTEC research findings at IAFP 2015 and VTEC 2015.

Project objectives

- Determine the relatedness of 'big 6' Shiga toxin-producing *Escherichia coli* strains isolated from Australian cattle with those isolated from Australian human clinical cases
- Genetically characterise pathogenic Shiga toxin-producing *Escherichia coli* strains from Australian beef and cattle samples and assess improvements in specificity of emerging commercial confirmation test systems
- **Engage with leading international researchers, share information with these researchers and others and contribute to international consensus on defining the risk of *E. coli* found in cattle and beef to human health**
- Phenotypically and genetically characterise the antibiotic resistance status of enterococci and *E. coli* isolated from beef cattle at slaughter

Success in meeting the milestone

A total of four CSIRO staff attended and presented pSTEC research findings at IAFP 2015 (2 presentations) and VTEC 2015 (4 presentations). The research findings presented demonstrated a consistency of approach with international research groups and highlighted the ability of Australian researchers to deliver impact in this research space despite discrepancies in budget, research effort applied and the absence of a coordinated national approach for understanding and characterising foodborne pathogens. It is clear that the use of next generation sequencing (NGS) methodologies has become the cornerstone of pSTEC, antimicrobial resistance and bacterial diversity research efforts. The use of NGS data for achieving applied outcomes for the animal production industries and human health was a notable feature of many presentations. In addition to the presentations, CSIRO staff met with a number of project collaborators to discuss ongoing investigations and to detail activities within the current project that require their input. Neogen, Roka Bioscience and ANSES all reiterated their willingness to continue to work closely with CSIRO and MLA in order to provide industry with superior pSTEC detection and confirmation procedures. Discussions were also held with a range of research groups and companies that are looking at potential collaborations either within the current project or subsequent to the current research efforts. Attendance at IAFP 2015 and VTEC 2015 continues to be a priority activity for Australian pSTEC researchers wishing to engage and exchange ideas with leading international researchers.

Discussion

International Association for Food Protection (IAFP 2015)

Presentation of MLA/CSIRO co-funded research

One CSIRO staff member attended and presented at IAFP 2015. Research topics were carefully selected in consultation with MLA to

ensure the most appropriate research findings were presented. Selected topics were presented in the form of posters to highlight key findings from MLA/CSIRO funded projects G.MFS.0285 and G.MFS.0286 (Table 1). In addition, Glen Mellor was invited to present a seminar at Washington State University (WSU) in which he provided an overview of CSIRO's Food Safety and Stability group that covered a number of past and present MLA/CSIRO funded studies on *Escherichia coli*. Following the presentation, discussions were held with WSU researchers to gain insights into the methods they are currently using for pSTEC detection and isolation.

Table 1. Presentations from MLA/CSIRO funded projects

Authors (* presenter)	Title	Type of presentation	MLA project code
Glen E. Mellor, Narelle Fegan, Lesley L. Duffy, Kate E. McMillan, David Jordan, Robert S. Barlow	Comparison of two methods for the isolation of Shiga toxin producing *Escherichia coli* O157 from cattle feces at slaughter	Poster	G.MFS.0286
Robert Barlow, Kate McMillan, Lesley Duffy, Narelle Fegan, David Jordan and Glen Mellor	Prevalence, serovars and antimicrobial resistance of *Salmonella* from Australian cattle populations at slaughter	Poster	G.MFS.0285

Collaborator meetings at IAFP 2015

CSIRO is currently coordinating a research project to compare detection and confirmation systems in an Australian setting. This project aims to provide industry with insights into the effectiveness and applicability of commercially available pSTEC test systems. Discussion around potential collaborations were held with representatives from five major commercial and research providers of pSTEC testing systems (a list of companies and meeting attendees is provided in Table 2). The key details of these discussions are listed below.

- Ian Jenson provided an Australian perspective on pSTEC testing. Representatives from Roka Biosciences were pleased to hear that similar challenges were being faced in Australia that are being faced in the USA.
- Companies were provided with an overview of the project aims and methodology and company representatives were invited to include their proprietary pSTEC test systems in the comparison. The following discussions were held around the distribution and confidentiality of results:
    - CSIRO provided assurance that each company would, as a minimum, get their results compared to BAX and GDS screening plus the outcome of confirmation testing. CSIRO also mentioned that, while they would like to provide participating companies with the results from the complete comparison, this

cannot be guaranteed prior to commencing the project as it ultimately depends on all companies agreeing to share their results.
    - CSIRO provided assurance to each company that they will be given the right of reply before publishing any data.
- Overall, CSIRO/MLA received positive feedback on the proposal and all companies expressed an interest in collaborating.
- Key actions identified included:
    - A need to send a project outline/protocol to each participating organisation.
    - A need to continue discussions with Robert Barlow at VTEC 2015 to assess the logistics and costs associated with each test platform.
    - Where applicable, initiate discussions with Australian representatives of participating companies.

Table 2. Collaborator meetings

Company	Attendees
Roka Biosciences (formal meeting)	Glen Mellor (CSIRO), Mick Becker (Vice President of Research) W. Evan Chaney (Sr. Manager, Scientific Affairs), Bettina Groschel (Scientist II, Assay Lead), Erin Dreyling (Director, Scientific and Government Affairs) and Ian Jenson (MLA).
Neogen Corporation (formal meeting)	Glen Mellor (CSIRO), Edan Hosking (Senior Research Scientist, R&D, Neogen), Joe Heinzelmann (Marketing Development Manager, Neogen), Ian Jenson (MLA)
Pall Corporation (informal discussion)	Glen Mellor (CSIRO) and Sebastien Bouton (R&D Senior Manager, Pall)
ANSES (informal discussion)	Glen Mellor (CSIRO) and Patrich Fach (ANSES)
BioControl (informal discussion)	Lyssa Sakaley (Product Manager, BioControl)

Summary of conference

The IAFP Annual Meeting is the leading meeting concerned with the protection of the worldwide food supply and is arguably the best forum to gain exposure to the latest trends in food safety research. The 2015 IAFP meeting was held at the Oregon convention center, Portland, Oregon, USA from 25 – 28th July, 2015. The conference was attended by more than 3,200 representing 49 nations. Research from MLA and CSIRO has been presented at consecutive annual IAFP meetings since 2011.

A broad range of food safety topics were presented at IAFP 2015, including microbial and non-microbial food safety, food law and regulations, modelling and risk assessment, communication and education, antimicrobials, sanitation and epidemiology. IAFP 2015 also saw a large number of symposia sessions and posters presented around genomics that provided an insight into the growing uptake and rapid pace at which this field is changing the landscape of clinical and food microbiology. Presentations demonstrated how genomics can be applied to characterising food microbiomes, microbial traceability, serotyping/subtyping and antimicrobial resistance. The following summarises the major research findings from these presentations.

Microbial traceability

- Presentations demonstrated how genomics can be used for the surveillance of foodborne pathogens.
  - Greater discriminatory power of whole genome sequencing has enabled the identification of smaller outbreaks and the number of identified outbreaks has increased with technological advancements.
  - The clonal nature of *S.* Enteritidis hinders outbreak investigations using traditional techniques such as PFGE. Whole genome sequencing (WGS) has been used to overcome limitations of traditional methods.
  - FDA Genome Trakr program continues to develop existing genome sequence databases for foodborne outbreak investigation. Efforts are being made to expand databases to include additional organisms that are not yet captured in the Genome Trakr program.
  - WGS used to identify geographic location of watercress seeds that were implicated in an *E. coli* O157 outbreak in England. Implicated strains were not related to previous strains identified in England and WGS facilitated tracking of strains to North American sourced seeds.

Serotyping/subtyping

- The CDC is coordinating a program designed to use genomics to transform public health microbiology.
  - Aim to replace components of traditional microbiology with whole genome sequencing techniques.
  - CDC is working towards creating a consolidated workflow for identification, serotyping, virulence profiling, antimicrobial resistance profiling.
- Online tools have been developed to simplify the generation of descriptive information from genome sequences.
  - The Center for Genomic Epidemiology has developed an online service that allows users to input raw sequence reads for the identification or prediction of:
    - Antibiotic resistance genes
    - Virulence genes and potential pathogens
    - Subtypes – Multi locus sequence types, Plasmid types, serotypes of *E. coli* and *Salmonella*, identification of species
    - Phylogeny

Characterising food microbiomes

- There are limitations to what culture microbiology techniques can achieve when investigating food microbiomes
- Metagenomics can help to answer 'who' and 'how' many organisms are present
  Examples of metagenomics presentations include.
    - Microbiome changes in raw meat during low temperature storage.
        - This study found that microbial successions are associated with chemical profiles of refrigerated pork sausages. Untreated sausage displayed a complex multi-species pattern of successions over the storage trial while treatment with lactate-diacetate yielded a monophasic growth curve of a single species.
    - Microbial ecosystems present in cheese rinds.
        - Identification of microbial profiles found on the same type of cheese produced in different regions.
        - Researchers examined 137 different rind communities from 10 different countries.
        - Formation of communities was correlated with abiotic (e.g environmental) factors but not geographic location of production.
    - Metagenomics application to food safety. Topics included:
        - Detecting and solving outbreaks.
        - Controlling contamination on tomatoes in the US. Metagenomics was used to establish baseline microflora profiles associated with tomatoes that have a high and low risk of *Salmonella* contamination. *Paenibacillus alvei* identified as part of microflora in tomatoes with low *Salmonella* risk. Researchers are investigating its potential use as bio-pesticide for controlling *Salmonella* on tomatoes.
        - Use of metagenomics to determine implication of enrichment bias for pathogen detection. Enriching may dramatically alter the taxonomic profile of samples and may not increase probability of detecting pathogen.
        - Commonly used methods of generating metagenomics data. Considerations for study design and analysis and obstacles associated with generating and handling metagenomics data.
        - How metagenomics/sequence data can be translated into practical outcome that can be adopted by regulators.
            - Audit trail needed if data is to be used to inform regulators. Need to understand the repeatability and extensibility of approaches.

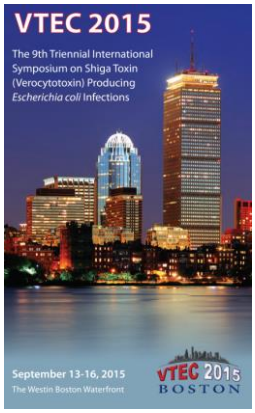
Comparison of commercial test systems for detection of pathogenic STEC

- Prevalence of EHEC in culled dairy cows determined using commercial molecular test systems and culture methods.
    - Enrichments plated onto STEC heart infusion washed blood agar with mitomycin-C, CHROMagar O157 and USMARC STEC agar. Enrichments were also tested by commercially available NeoSEEK and Atlas EG2 Combo assays.
    - EHEC-7 isolates were recovered by culture based methods from 7.0% faeces, 16% hides and 1.0% carcasses. EHEC-7 prevalence's were mostly higher when using NeoSEEK and Atlas systems. Moderate agreement between culture and NeoSEEK was observed for the detection of three EHEC-7

serotypes but no agreement was observed between other NeoSEEK serotypes or test methods.

Symposium on Shiga Toxin (Verocytotoxin) Producing *Escherichia coli* Infections (VTEC 2015)

Presentation of MLA/CSIRO co-funded research



Three CSIRO staff attended the conference and presented posters, talks and chaired sessions. There were 3 posters and one invited oral directly associated with MLA/CSIRO funded work which were presented at VTEC 2015 (Table). In addition, Dr Kari Gobius was invited to present on the geographical characterisation of *E. coli* O157 at the pre-symposium workshop on Next Generation Sequencing and Genomic Evolution and Dissemination. This presentation included work from the MLA/CSIRO funded project on Subtyping of *E. coli* O157 (A.MFS.0236) and work that has continued since the completion of that project which CSIRO has been undertaking as part of its strategic research program.

Table 2. Presentations from MLA/CSIRO funded projects

Authors (* presenter)	Title	Type of presentation	MLA project code
Fegan, N., Besser, T. E., Shringi, S., Baker, K. N. K., Smith, H. V., Jennison, A. V., Gobius, K. S. and Mellor, G. E.	Shiga toxin bacteriophage insertion sites, toxin subtypes and Stx production varies between Australian and U.S. *E. coli* O157 populations.	Poster	A.MFS.0236
Barlow, R. S., McMillan, K. E., Duffy, L. L., Fegan, N., Mellor, G. E., Delannoy, S. and Fach, P.	Genomic characterization of atypical enteropathogenic *E. coli* (aEPEC) strains from Australian cattle	Poster	G.MFS.0286
Mellor, G. E., Fegan, N., Duffy, L. L., McMillan, K. E., Jordan, D. and Barlow, R. S.	Enumeration of pathogenic Shiga toxin-producing *Escherichia coli* in Australian beef cattle feces at slaughter	Poster	G.MFS.0286
Mellor, G. E., Fegan, N., Duffy, L. L., McMillan, K. E.,	National survey of Shiga toxin-producing *Escherichia coli* serotypes O26, O45, O103, O111, O121, O145	Presentation	G.MFS.0286

Jordan, D. and Barlow, R. S. and O157 in Australian beef cattle feces at slaughter

Collaborator meetings at VTEC 2015

Further informal discussions were held with representatives of Neogen, Roka Biosciences and ANSES. These discussions built on earlier discussions that had taken place at IAFP 2015 and were specifically related to upcoming activities within V.MFS.0333. Each of the collaborators expressed a desire to continue to work together and were keen to evaluate their pSTEC test systems/approaches as part of the current project.

Summary of conference

The VTEC 2015 meeting was the 9th Triennial Symposium on Shiga Toxin (Verocytotoxin) Producing *Escherichia coli* Infections and was held at the Westin Hotel, Boston, Massachusetts, USA from the 13 – 16th September 2015. The conference was the most attended VTEC conference to date with delegates representing 23 nations. Research from MLA and CSIRO jointly funded projects has been presented at successive VTEC conference since the 3rd meeting held in 1997.

The following comprises the major areas of significant research findings associated with both the Pre-Symposium workshops and conference. The information has been grouped according to the general areas under which the conference was structured and which follows the format of previous VTEC conferences.

Pre-Symposium I – Food safety from farm and field to plate

- An overview and progress on The STEC CAP Grant activities were provided
    - the CAP Grant is a $25 million project funded by the United States Department of Agriculture – National Institute of Food and Agriculture (USDA-NIC) Coordinated Agricultural Project (Cap) grant scheme, the project is called Shiga-toxigenc *E. coli* in the beef chain: assessing and mitigating the risk by translational science, education and outreach
    - involves many collaborators across universities and state research institutes
    - the project has 3 pillars: 1. live cattle and beef producers, 2. post harvest slaughter and 3. consumers
    - objectives are around STEC detection, biology, interventions, risk analysis and education/outreach
    - investigation of how to control STEC in veal as these are viewed by the FSIS as high risk for carrying STEC
        - various interventions were investigated
    - investigated methods for detection of 8 serogroups (O157 + Big 6 + O104)

        - found similar issues to other researchers, e. g. cattle frequently shed target serogroups (mostly O103, followed by O45, O121 and O26) but very few of the isolates carry *stx*; multiple serogroups can be present in a single sample;
        - magnetic beads are not as specific as researchers would like and in this survey many of the serogroups specific beads captured *E. coli* O103 – unknown as to why this occurs
        - quantitation of serogroups (not STEC) using spiral plating was able to detect lower levels (10^2/10^3 cfu) than real time PCR (> 10^4 cfu)
        - tested 757 rectal contents from slaughtered cattle for *E. coli* O104 and isolated some *E. coli* O104 which carried stx_{1c}, but none carried *eae* therefore cattle seen as low risk for carrying isolates similar to the German O104 outbreak strain
        - in relation to control in cattle, there was no evidence that vaccines against *E. coli* O157 would have any impact on the non-O157 serogroups but there was speculation that maybe probiotic use would have a greater impact
    - investigations were undertaken into control of STEC during veal processing using various interventions
        - for hide of veal calves, hot water wash (82 °C) with lactic acid was found to be the most effective treatment (~ 5 log reduction)
        - for dressed veal, looked at combinations of lactic acid and antimicrobial sprays but these had limited impact (initial wash without antimicrobials was just as effective)
        - when comparing cooking breaded (crumbed) veal and non-breaded veal at 191 °C for STEC reduction, it was found that at least 30 ml of oil was needed for breaded and 15 ml for non-breaded veal to achieve a 5 log reduction after 1.5 min. For veal cordon bleu (containing 2 pieces of veal with ham in between), the cooking time needed to increase to 7 min to achieve a similar reduction in STEC
    - further information can be found at http://www.stecbeefsafety.org/stec-cap-grant
- Various researchers reported that shedding of STEC on cattle farms was more effective when farmers applied practical risk management strategies
    - another finding was contradictory to previous research and indicated that rainfall decreases STEC shedding (unknown as to why other than possibly reducing dust and therefore transmission)
- A multi-state study (Michigan, Nebraska and Washington) of dairy and beef cattle was undertaken to determine if there were particular microflora associated with STEC positive or negative cattle
    - dairy herd prevalence was 13% (from 718 cattle), and beef was 21% (from 378 cattle)
    - no differences found between the microflora (based on metagenomics) between STEC positive and negative cattle
    - STEC population was very diverse and varied across herds

    - first lactation, early days in milk and heat were associated with increased prevalence
    - animals shedding > 3 log STEC provided 3 x greater risk of transmission than those animals shedding less
    - looked into whether *E. coli* O157 shedding might be a heritable trait (followed 1000 animals), some evidence that this was the case (genes associated with the keratin filament, intermediate filament and cytoskeleton may play a role) and was also influenced in this group of animals by the microbiome, some animals never shed *E. coli* O157
- Control of STEC in cattle using bacteriophages is not very successful
    - control on farm was not shown to be effective mostly because of difficulties in keeping phage away from control groups of animals and the variability of the results
    - believe that predator/prey relationship plays a role in why the shedding of *E. coli* O157 from cattle is variable
    - presence of phage can change the PFGE of the *E. coli* O157 in the cattle
    - phage/host relationships appear to be far more complex than originally thought and there needs to be an understanding of phage ecology in far more detail before progress can occur in this area
- Shiga toxin carrying bacteriophages are present in a range of different environments (including waters - human sewage and effluents from farms; soils; human faeces; foods – minced meat and salads) at levels between 1 and 5 log, with those carrying *stx*$_2$ far more common (found in 68% of all samples tested) than those carrying *stx*$_1$ (7.6% of samples). These phages can be long lived in the environment (> 17 days) and numbers decline more rapidly in summer than in winter
    - *stx*$_1$ carrying phages are expressed by induction or by iron depletion, have their own promoter site and can be expressed without lysis
    - *stx*$_2$ carrying phages need to be induced for propagation and this may be why more *stx*$_2$ phages were found in the samples (due to bacteria lysing and releasing the phages)
- Persistence of STEC in the environment was viewed as an important factor in transmission and research was undertaken to try and understand more about survival of *E. coli* O157 in the environment, *rpoS* was found to play an important role in increased environmental persistence
    - biphasic lifestyle of STEC:
        - within the host/reservoir – nutrient availability, optimal temperature, anaerobic, stable or net increase in population
        - in the environment – nutrient limitations, low temperatures, other stresses (UV etc), net decline in population
    - no obvious links between ability to survive and genetic relationships as there was no difference in survival (in manures and soils) between *E. coli* O157 lineages
    - better environmental survival appeared to be linked to metabolism, particularly hydroxyl butyrate, ketobutyrate and propionic acid, also ability to grow on fatty acids differentiated good survivors from poor survivors

    - *rpoS* mutations were present in poor survivors, but good survivors had the wild type *rpoS*
    - human *E. coli* O157 – 75% carry wild type *rpoS* while 25% have mutations in *rpoS* while cattle isolates all carried wild type *rpoS*
        - theory that allelic variation in *rpoS* of *E. coli* O157 provides an advantage in human (and mice) colonisation as it enables higher overall metabolism and scavenging and ability to utilise amino acids, increased expression of LEE, increased curli expression and fitness in the gut
        - cattle *E. coli* O157 carrying the wild type *rpoS* have an advantage in colonising cattle as they are more stress resistant, lower curli expression and persistence in the soil and better biofilm formation
- STEC and produce
    - US survey – 2200 samples of spinach tested annually, 0.5% prevalence of STEC (and never find *Salmonella*)
    - testing of produce can take up to 23 days, but spinach has a shelf life of up to 14 days
    - have been using the ECID microarray[1] to help achieve more rapid testing
    - other interesting facts about STEC and produce were discussed and included:
        - changing trends to the use of polytunnels, intensification, use of garnishes for food may lead to increased risk of produce as transmission pathway
        - water used for irrigation most likely source of contamination
        - suggestions that some plants may be secondary hosts for STEC (in model systems can see colonies forming within plant tissues after 4 days, bacteria grow on nutrients from the phloem), plant species specific, some plants may never be colonised

Pre-Sympoisum II – Next generation sequencing and genomic evolution and dissemination

- Major focus of the symposium revolved around the development and availability of workflows for the rapid analysis of next generation sequencing (NGS) data. The following were discussed:
    - Illumina BaseSpace – cloud based analysis that is directly integrated with Illumina's suite of sequencing platforms. BaseSpace is an app based set up allowing users to access a range of tools that can be linked together to suit their workflow. They also provide six workflow set ups for custom/PCR amplicons, library QC, resequencing, De Novo assembly, metagenomic analysis and small RNA analysis. The system currently has 60 apps, half of which are specifically designed for microbial populations.
    - Applied Maths – have released Bionumerics Seven. This provides a modular approach to the handling of NGS data. Of particular interest is the genome analysis tools module that offers a range of comparative tools for genomics and

[1] http://www.affymetrix.com/estore/catalog/prod810020/AFFY/Minimal+Signature+E+coli+Array+Strip#1_1

metagenomics. There is specific focus of SNPs with the capacity to explore genomes at multiple levels (e.g. whole genome MLST through to hgMLST).
    - The Public Health Agency of Canada detailed the use of Panseq for the analysis of the pan genome of a group of genomic sequences. This has been developed to rapidly assess AMR and virulence factors associated with STEC. Panseq includes a novel region finder (SuperPhy) which will find sequences that are unique to a strain or group of strains with respect to another strain or group of strains making it an ideal platform for predictive genomics.
    - Serotype finder – whole genome based serotyping of *E. coli* which has been added as a component of the publically available tools hosted by the Center for Genomic Epidemiology (www.genomicepidemiology.org). The tool uses at least 9 O-antigen and flagellin genes and was shown to be extremely effective at determining serotypes with 560/569 O-antigens and 504/508 H types correctly identified.

- The use of NGS for STEC was described using the following examples:
    - Public Health England detailed their daily/weekly workflow and compared and contrasted the transition to WGS for the rapid analysis of gastrointestinal pathogens. WGS provides a rapid cost-effective approach to O157 and non-O157 characterisation and outbreak investigation
    - The use of NGS for the characterisation of O157 in geographical distinct environments was discussed. Bovine biased and clinical biased genotypes can be identified in cattle with the latter group typically composed of O157 belonging to specific lineages or carrying specific Shiga toxin genes (notable stx_{2a}. The loci at which stx phage insert into the genomes of O157 also provides a hypothesis as to why different incidences of disease is observed throughout the world.

VTEC 2015 – Epidemiology

- outbreak investigations are now based on whole genome sequencing of STEC and ways of managing all this new data need to be developed
- risk factors associated with sporadic non-O157 STEC infections in the US based on a case control study include:
    - eating a hamburger at a fast food restaurant (particularly a pink hamburger)
    - working of a farm
    - camping
    - swimming in a lake
    - for *E. coli* O26, visiting a zoo, camping and swimming in a lake were the major risk factors
    - most common serotypes of non-O157 responsible for human cases were O103 (24% of cases), O111 (11%) and O26 (32%)
- *E. coli* of serotype O80:H2 have been found in France (resulting in 10 cases, 8 of which were HUS)

- STEC O26 is the most common non-O157 found associated with human cases in the US, most isolates carry *stx*$_1$ but are starting to see an increase in isolates with *stx*$_2$ (paralleling observations from the UK and Europe several years earlier)

VTEC 2015 – Diagnostics, typing and phylogenetics

- Culture Independent Diagnostic Tests (CIDT) are coming, provide the ability to diagnose multiple infections at once
    - some questions were raised around needing to understand more about viability along with infectivity if using CIDT
    - metagenomics will be used in the future for outbreak investigations
- Whole genome sequencing and phylogenetic comparisons have enabled a group in the UK (through comparison of 1,075 clinical and 95 cattle isolates of O157) to estimate global expansion of the 3 lineages of *E. coli* O157 occurred around the year 1840
- *E. coli* O157 can change virulence potential if Stx phages are gained or lost (example below in Genetics and virulence factors of O157 PT32 converting to a more pathogenic version of PT21/28 through acquisition of *stx*$_{2a}$)

VTEC 2015 – Genetics and virulence factors

- Whole genome sequencing (WGS) is the new typing tool of preference for all STEC
    - various groups have been developing tools to deal with the information gathered by WGS
        - SuperPhy
- Variability in pathogenicity between isolates
    - for non-O157 STEC, genes associated with catabolism of aromatic compounds and involved in cell adhesion are more highly expressed in HUS isolates suggesting roles for these genes in disease
- Stx containing bacteriophages were a topic of particular focus
    - *E. coli* O157 PT32 (carrying *stx*$_{2c}$) is converted to PT21/28 through the acquisition of *stx*$_{2a}$ phage (*E. coli* O157 PT21/28 has become predominant in the UK)
    - evidence of cross regulation between bacteriophages
        - from vero cell assays experiments *stx*$_{2a}$ phage repress *stx*$_{2c}$ phage
        - removal of *stx* from the phage does not stop cross regulation (requires removal of the entire phage)
    - evidence of phage genes regulating other genes such as LEE (*cro* deletion reduces LEE expression)
    - induction of Stx phages does not lead to the lysis of the entire STEC population (always around 1,000 cfu remain unaffected) and the reasons for this are unclear. There were several hypothesis put forward around this area including involvement of rpoS expression, different growth phased (some of the population being in log while others in lag) and the possibility of involvement of epigenetics.

VTEC 2015 – Treatments (animal and human)

- Human treatments have focused on
    - inhibiting expression of virulence factors such as Shiga toxin (Detox) and type III secretions systems and effector molecules (Disengage)
    - using chimeric antibodies to block the action of Shiga toxins (SHIGATEC – which is at the clinical trial stage of development)
    - treating HUS with volume expansion (saline injections) to limit tissue damage
- Control in animals has focussed on
        - phage therapy (see pre-symposium I notes on phage treatments which were not found to be effective for reducing shedding in cattle)
        - vaccination – mathematical modelling was used to examine theoretical effectiveness of cattle vaccination which indicated the use of vaccines could be effective at reducing human disease
            - issues were also raised around who should pay for vaccination (as it isn't an animal health issue so why should farmers pay)
            - the effectiveness of the vaccine needs to be considered
            - the model predicted that shedding 100 fold more bacteria in cattle would lead to 8 times more infections in humans (reducing supershedding in animals was therefore seen to be important)

VTEC 2015 – Animal reservoir, food, environment, transmission

- some STEC are more heat tolerant than others and able to survive at 60 °C for 1 hour, they contain a specific protein which is yet to be fully characterised but is thought to be involved in heat resistance
- in the UK, the population of *E. coli* O157 circulating in cattle is believed to have changed over time possibly due to the restocking cattle from Southern Europe and North Africa after the foot and mouth outbreak and BSE incidents
- non-O157 STEC are rarely detected in feedlot cattle from the USA (O26 (1%), O103 (1.6%) and O145 (0.8%) were detected only in summer from over 500 samples)
- higher summer prevalence of *E. coli* O157 in cattle in the USA is believed to be the result of extrinsic (temperature, decreased predation of *E. coli* O157, growth outside the host) due to increased exposure rather than intrinsic factors associated with the animals (endocrine levels, host microbiome, physiology)
    - cattle carriage appears to be highly individual – with different animals carrying different subtypes of *E. coli* O157
    - cattle shedding varies with individual cattle
    - some cattle never shed O157 even when artificially dosed
- Supershedding:
    - investigation of supershedding in cattle by focussing on the transcriptome of the recto – anal junction (RAJ) of colonised cattle indicated some potential association

between colonisation and adaptive immune functions (with adaptive immune responses decreased in animals colonised by *E. coli* O157).
        - this presentation generated much discussion and was thought to be quite preliminary as the cause/effect issue was a little unclear, e. g. whether the lower immune responses in colonised cattle were the result of the colonisation by O157 (with O157 causing the lowered immune response) or the cause of colonisation (with O157 able to colonise these cattle due to the lower immune responses).
        - this will be an area that will be explored in much greater detail with the molecular tools now available to get much more detailed information.
    - *E. coli* O26 may also be supershed (shed in high numbers) from cattle
    - artificial inoculation of calves has indicated that certain phage types of *E. coli* O157 which carry stx_{2a} (PT21/28) are shed in higher numbers than others (PT32)
- Results of an Irish survey using RAJ swabs from 1,317 cattle at slaughter were similar to other published information:
    - *E. coli* O157 was isolated from 3.87% and *E. coli* O26 from 0.68% of cattle
    - supershedding (counts > 4 log cfu/swab) of *E. coli* O157 was observed in 2% of cattle and supershedding of *E. coli* O26 in 0.2% of cattle
- Analysis of data from the UK showed *E. coli* O157 phage types which are shed in higher numbers in cattle (> 1,300/g) match the phage types seen in clinical cases
    - further supporting high shedding animals as a risk for transmission to humans

VTEC 2015 – Pathogenesis, host response, animal models

- Shiga toxin and its role in disease
    - following on from the information in the Genetics and Virulence factors information associated with bacteriophages, experiments using mouse assays and co-intoxication with both stx_{2a} and stx_1 determined that the combination of toxins produced less morbidity than intoxication with stx_{2a} alone
        - administration of stx_1 3 h prior to stx_{2a} increased survival of mice and the ratio of toxin subtype may also play a role, EDL933 (Sakai strain and considered of low virulence) produces 5 times more stx_1 than stx_{2a} while EC2812 (Jack in the Box strain) produces 2 fold more stx_1
        - model is that stx_1 arrives at the kidneys first and will bind and prevent stx_{2a} binding, if both toxins arrive at the same time, stx_1 may change the structure of the receptor (Gb3) and prevent binding of stx_{2a}

Other general comments

There were several breakfast symposia held during the conference, these were more generic presentations that focussed on the human impact and regulatory space rather than the science associated with STEC. Of specific note were:

- Bill Mahler's presentation around the Jack in the Box outbreak and details of the law in relation to food safety incidents. When questions about the litigious nature of the US, Bill

noted that the US legal system developed in a society where there was very poor health care and poor regulation meaning that litigation became the only route for retribution and apparent justice. He believed this would change over time given regulation and health care has improved over time and will continue to do so.

- Presentations from USDA (National Institute of Food and Agriculture – NIFA), FDA and CDC discussing their priorities and successes around STEC research
    - NIFA $40 million in funding, projects can range from $300,000 to $5 million, fund US groups which can partner with international groups. fund the STEC CAP project and are also currently focussing on produce along with novel concentration and purification methods
    - FDA are developing Genome Trakr to assist in analysis of WGS information, keen to have international labs on board with this analysis (already have partners in UK, Ireland and Argentina)
    - CDC are about to celebrate 20 years of PulseNet (in 2016), along with NARMS and FoodNet, have developed a steering committee to look at Culture Independent Detection Tests.

Conclusions/recommendations

The current focus of this MLA project is appropriate based on the latest information coming out of the IAFP and VTEC conferences. The project remains well placed to deliver on agreed objectives within the expected timeframes.

Milestone 4 – Non-O157 subtyping – Draft publication prepared and submitted to MLA – Comparative analysis of up to 100 Australian non-O157 EHEC from cattle and clinical sources

Abstract

Shiga toxin-producing *Escherichia coli* (STEC) are important food-borne pathogens capable of causing a variety of disease symptoms from uncomplicated diarrhoea to haemorrhagic colitis (HC) and haemolytic uremic syndrome (HUS) in humans. A subset of STEC that are defined by the presence of *stx* and *eae* and their association with particular serotypes such as O157, O26 and O111 are known as pathogenic STEC (pSTEC). Whilst all pSTEC share a set of cardinal virulence markers there remains an inability to accurately assess the capacity of individual strains to cause severe human disease. The main objective of this study was therefore to genetically characterise and compare cattle and human non-O157 pSTEC isolates and to determine if a non-random distribution of virulence-associated genotypes are present. A total of 106 predominantly Australian human and cattle sourced non-O157 pSTEC (i.e. O26, O45, O103, O111, O121 and O145) were selected from 170 isolates on the basis of pulsed field gel electrophoresis and *stx* subtype profiles and subsequently characterised using next generation sequencing based analysis for multi-locus variable number tandem repeats (MLVA), comparative genome fingerprinting, pangenome content, Shiga toxin bacteriophage insertion sites (SBI) and single nucleotide polymorphisms (SNPs). Additionally, the production of Shiga toxin by each isolate was assessed. Cattle carriage and the frequency of human clinical disease in Australia is dominated by O26 and O111 with the remaining non-O157 serotypes of regulatory importance to export beef producers present at very low prevalence, if at all. The results of this study demonstrated that there were notable differences were observed between isolates belonging to different serotypes, however no significant differences were observed between cattle and human isolates of the same serotype. Indeed, there was substantial correlation between isolates of cattle and human origin suggesting that cattle are a major reservoir of non-O157 isolates associated with human clinical disease. The plasticity and ability of the *E. coli* genome to undergo substantial recombination events is highlighted within this study with the first description of a stx_{2a} producing O26 from cattle and a hybrid O26/O111 pSTEC as a causative agent of human clinical disease. Virulence-associated genotypes appear to group based on serotype and not by source, and therefore the ability to accurately identify individual isolates of non-O157 pSTEC with the greatest human clinical disease potential remains a significant challenge.

Contents

Abstract 61
Project objectives 63
Success in achieving milestone 63
Introduction 64
Materials and Methods 67
Strain selection 67
Pulsed field gel electrophoresis 67
Shiga toxin gene subtyping 68
Shiga toxin production 68
Sequencing 68
Multi-locus variable number tandem repeat analysis (MLVA) 69
Comparative genome fingerprinting (CGF) 69
Shiga toxin bacteriophage insertion site (SBI) typing 70
Pangenome 70
Single nucleotide polymorphisms (SNPs) 73
Results 73
Strain selection 73
Shiga toxin production 78
Sequencing statistics 79
MLVA 79
Comparative genome fingerprinting (CGF) 79
Shiga toxin bacteriophage insertion site (SBI) typing 82
Pangenome 85
Single nucleotide polymorphisms (SNPs) 88
Discussion 88
Acknowledgements 92
References 93
Appendix 1 97

Project objectives

1. Determine the relatedness of 'big 6' Shiga toxin-producing *Escherichia coli* strains isolated from Australian cattle with those isolated from Australian human clinical cases
2. Genetically characterise pathogenic Shiga toxin-producing *E. coli* strains from Australian beef and cattle samples and assess improvements in specificity of emerging commercial confirmation test systems
3. Engage with leading international researchers, share information with these researchers and others and contribute to international consensus on defining the risk of *E. coli* found in cattle and beef to human health
4. Phenotypically and genetically characterise the antibiotic resistance status of enterococci and *E. coli* isolated from beef cattle at slaughter

Milestone 4 primarily addresses project objective number 1 by determining the relatedness of 'big 6' pSTEC from cattle and human clinical sources. Additionally, the findings of this Milestone will assist the successful completion of project objective number 2 by enhancing our knowledge of the genetic variability of Australian pSTEC isolates and how this might be exploited in pSTEC test systems.

Success in achieving milestone

Milestone 4 details the use of NGS as a basis for comparative analysis of up to 106 predominantly Australian non-O157 EHEC from cattle and clinical sources. It demonstrates that unlike O157 isolates there does not appear to be subgroups of non-O157 isolates within Australian cattle populations that are more associated with human disease. Indeed the correlation of genomic content between isolates from cattle and human clinical sources suggests that cattle are a major reservoir of the non-O157 serotypes O26 and O111 with the remaining regulatory related non-O157 serotypes of O45, O103, O121 and O145 seldom present in either cattle or human clinical cases. Furthermore, the study highlights the evolving nature of pathogenic *E. coli* by describing the first stx_{2a} producing O26 isolate from Australian cattle and hybrid O26/O111 isolates as causative agents of human clinical disease.

Introduction

Shiga toxin-producing *Escherichia coli* (STEC) are important food-borne pathogens capable of causing a variety of disease symptoms from uncomplicated diarrhoea to haemorrhagic colitis (HC) and haemolytic uremic syndrome (HUS) in humans. By definition, STEC are *E. coli* that produce Shiga toxin, however, they may additionally harbor a range of genetic elements such as bacteriophages and pathogenicity islands that affect their ability to cause severe disease. Not all STEC cause disease and the subset that do have been referred to as pathogenic STEC (pSTEC). Internationally, several pSTEC serotypes have been associated with sporadic and outbreak cases of foodborne disease. A number of studies have implicated beef products in cases of foodborne disease and beef cattle have been shown to be a source of pSTEC *(27)*. Whilst initial attention focused on the prototypal pSTEC serotype O157, recent regulatory changes in the USA around the supply of raw, non-intact beef products has broadened the scope to include six additional pSTEC serotypes: O26, O45, O103, O111, O121 and O145. The additional serotypes, known as non-O157 pSTEC, are included based on epidemiological data which demonstrates that these serotypes account for the majority of non-O157 illness in the USA *(9)*. Prevalence estimates for the six additional pSTEC serotypes in Australian cattle groups are low with studies repeatedly demonstrating prevalences of O26 and O111 of less than 1% and an absence of pSTEC belonging to serotypes O45, O103, O121 and O145 *(2, 27)*. Similarly, the burden of non-O157 pSTEC on the Australian public health system would also appear to be low with an annual notification rate of STEC illness estimated at 0.4 cases per 100,000 per year *(40)*. Nevertheless, the non-O157 pSTEC serotypes detected in human illness reflect those detected in cattle populations with O26 and O111 being identified more frequently than the serotypes O45, O103, O121 and O145 which were identified at <1% or not at all in 504 human cases from 2001 to 2009 *(40)*.

Although the prevalence and types of non-O157 pSTEC in animal populations or implicated in human clinical cases can vary throughout the world, the ability to define which particular isolates are more likely to give rise to the more frequent or severe disease scenarios remains problematic. Whilst it is obvious that isolates harbouring specific toxin types (i.e. *stx2a*) and genes that aid in attachment (e.g. *eae*) are most likely to result in human disease, there is no single virulence marker or combination of virulence markers that accurately predicts the likelihood that a specific non-O157 pSTEC strain will cause more severe disease, including HC and HUS. In addition, there are examples of STEC that have caused large outbreaks or are involved routinely in sporadic disease cases that possess hybrid combinations of virulence markers (e.g. O104; *(32)*) or that do not harbor virulence markers (e.g. O113 and O91; *(41)*) that are typically present in isolates of STEC that cause frequent

or severe disease. Virulence profiling of non-O157 pSTEC and the subsequent development of molecular risk assessment (MRA) algorithms for identifying isolates that are a high to moderate risk for severe disease are therefore a challenge.

Whilst applications of MRA to non-O157 pSTEC isolates is a relatively new area of research with conclusions based on relatively low numbers of isolates *(14, 17, 19)*, its application to O157 isolates is well established. Numerous studies have compared isolates from cattle, foods and humans to identify if certain genotypes are over represented in particular sources and therefore indicative as isolates most likely to cause frequent or severe human disease *(8, 18, 28, 37)*. A general finding of these studies was that the genotypes of O157 typically associated with severe human disease comprise only a small percentage of the total cattle pool of isolates. Indeed, O157 isolates in these studies could be characterised as possessing genotypes that were either bovine biased (BBG) or clinically biased (CG) *(28, 37)*. Investigations into O157 isolates from Australian cattle revealed that BBG of lineage I/II clade 7 isolates with stx_1 prophage integrated in the *argW* chromosomal insertion site were most prevalent *(29)*. This finding is in stark contrast to the abundance of CG of lineage I/II clade 8 isolates with stx_2 prophage inserted in *argW* chromosomal insertion sites that are common to countries with elevated HUS rates and may provide evidence for the observed lower prevalence and severity of disease in Australia. As the prevalence and severity of disease associated with non-O157 pSTEC in Australia are also low when compared internationally, it is plausible to suggest that BBG may exist within this group of isolates.

There are a range of methods currently in use to categorise pSTEC*,* each providing differential levels of phylogenetic and phenotypic resolution *(12, 33)*. Methods such as pulsed field gel electrophoresis (PFGE) with internationally established databases and procedures could be considered to be a gold standard for the characterisation of *E. coli*. This approach is particularly applicable to the identification of outbreak sources and epidemiological studies. Other methods such as serotyping, multiple-locus variable number tandem repeat (MLVA), Shiga toxin-encoding bacteriophage insertion (SBI), and lineage specific polymorphism analysis (LSPA-6) provide information on the relationships between different isolates of pSTEC*.* DNA sequence based methods such as multilocus sequence typing (MLST) which can elucidate phylogenetic relationships has been adopted internationally. Multiple MLST databases have been established for the typing of isolates.

The increasing prevalence of next-generation DNA sequencing (NGS) is transforming the identification and typing methods applied to pSTEC and other organisms *(20)*. NGS based analysis methods offer advantages in reduced costs, greater resolution as well as improved reproducibility between facilities. Typically genome sequence data is in the form of millions of reads greater than 150 bp in length (depending upon the chemistry and equipment selected). Contiguous assemblies of sequencing read data are not required for typing, identification, or phylogenetic analysis so finished genomes are not typically generated. Near complete genomic sequence data can be generated far more quickly than any of the other typical laboratory methods used for typing or characterisation. Many time consuming methods such as MLST and MLVA can be accomplished in silico using genomic DNA sequence data. Although less time consuming than traditional analysis methods, NGS data has the additional requirement for computational and data storage resources. International standards are currently being developed for the handling and processing of NGS data so no single methodology has been determined for processing the data. Despite the uncertainty in processing method and the requirement for computational analysis NGS delivers significant benefits. Near complete coverage of the genome provides the highest possible level of resolution for phylogenetic analysis. This data also permits detailed epidemiological and source tracking analyses to be undertaken *(23)*. The main objective of this study was therefore to genetically characterise and compare cattle and human non-O157 pSTEC isolates and to determine if a non-random distribution of genotypes are present. The study aimed to include isolates from each of the six non-O157 pSTEC serotypes, however a general absence of isolates belonging to serotypes O45, O103, O121 and O145 from either cattle or human clinical cases within Australia results in a focus on O26 and O111 isolates for the majority of comparisons.

Materials and Methods

Strain selection

A total of 170 predominantly Australian pathogenic STEC isolates were selected from across CSIRO, Queensland Health, Melbourne Diagnostic Unit and Murdoch Children's Research Institute culture collections (Table 1) for initial characterisation by Pulsed Field Gel Electrophoresis (PFGE) and Shiga toxin gene subtyping (see below for method details). Initial characterisation data was combined with spatial and temporal data to select a subset of 106 isolates for inclusion in subsequent characterisation studies. The 106 isolates were selected on the basis that they broadly represent the diversity in PFGE, *stx* genotypes, sources and years observed across the initial set of 170 isolates. The final group of isolates selected comprised 104 isolates from Australian sources and a further two isolates from Hong Kong and the United States of America of serotypes O121 and O45, respectively.

Table 1. Distribution of cattle and human pSTEC isolates included in initial characterisation by pulsed field gel electrophoresis and Shiga toxin gene subtyping.

	cattle		human	
serotype	*n*	isolation date range	*n*	isolation date range
O26	89	1998-2015	36	2000-2015
O111	19	1995-2015	22	1997-2015
O121			2	2012, unknown
O103			1	2012
O45			1	unknown

Pulsed field gel electrophoresis

PFGE profiles were generated for 170 isolates using the PulseNet one-day standardised protocol for non-O157 STEC *(11)*. Genomic DNA fragments were generated using a single restriction enzyme (XbaI) and separated on 1% SeaKem Gold agarose (Lonza, Rockland, ME, USA) gels using a CHEF DR-III system with the following electrophoresis conditions; initial switch time of 6.76 s, final switch time of 35.38 s, voltage of 6 V/cm, included angle of 120° and a total run time of 19 hours. Gel images were uploaded to BioNumerics (version 7.6; Applied Maths, Austin, TX) for processing and analysis.

Pairwise cluster analysis was calculated using UPGMA with a dice coefficient. Both band matching tolerance and optimisation were set at 1.5% as prescribed by the CDC.

Shiga toxin gene subtyping

The diversity of Shiga toxin gene subtypes in 170 pathogenic STEC isolates was assessed following a PCR based procedure defined by Scheutz *et al. (34)*. Isolates that were confirmed for stx_1 were screened for subtypes stx_{1a}, stx_{1c} and stx_{1d} while those that were confirmed for stx_2 were screened for subtypes stx_{2a}, stx_{2b}, stx_{2c}, stx_{2d}, stx_{2e}, stx_{2f} and stx_{2g}. All PCR amplicons were separated on 1.5% agarose gels and visualised using the T:Genius gel imaging system (Syngene, Cambridge, UK).

Shiga toxin production

Total Shiga toxin production was assessed using an ELISA based method previously described by Mellor *et al. (28)*. In brief, overnight cultures for each of the 106 isolates were diluted (1:21) and aliquoted in 1 ml volumes into separate wells of a 2 ml deep well plate (Sarstedt, SA, Australia). Mitomycin C (Sigma-Aldrich Corp., St. Louis, MO, USA) was added to each well to achieve a final concentration of 0.5 µg ml^{-1} after which plates were sealed and placed in a rotary shaker (250 rpm) for 24 ± 2 h at 37°C. Following overnight induction, cells were first lysed with polymyxin B sulfate (0.5 mg/ml) and then diluted (1:100) in sterile LB broth. A subsequent 1:2 dilution in sample diluent (Premier EHEC; Meridian Bioscience Inc., Cincinnati, Ohio) was prepared to achieve a final dilution (1:200) for use in downstream assays. Total Stx production was measured using the Premier EHEC enzyme-linked immunosorbent assay (ELISA) (Meridian Bioscience Inc.) following the manufacturers guidelines. Absorbance readings were obtained at dual wavelengths (450 nm/630 nm) using a Victor X microtiter plate reader (PerkinElmer, Glen Waverley, Australia). Assays were replicated and mean values were used in subsequent analyses.

Sequencing

DNA was prepared for each isolate using the DNeasy blood and tissue kit (Qiagen) and the concentration determined using the Qubit dsDNA HS (high sensitivity) assay kit (Thermofisher Scientific). DNA quality control, library preparation and sequencing runs were performed at the Ramciotti Centre for Genomics, University of New South Wales or through the Microbiological Diagnostic Unit at the University of Melbourne. DNA libraries were prepared using the Nextera XT DNA library preparation kit (Illumina) and paired-end (2x300bp) reads were generated using the

MiSeq desktop sequencer v3 (Illumina). Illumina MiSeq reads were quality filtered, adapters removed, and clipped using the program ea-tools' command, fastq-mcf, v.1.04.676 (https://code.google.com/archive/p/ea-utils), with a minimum read length of 100 bp; and a quality clip score of 20 required within a sliding window of five bp. Assembly was performed using the program, idba-ud v.1.1.2 (https://github.com/loneknightpy/idba) with kmers from 150 to 300 bp in length in 50 bp steps. Contigs shorter than 1 kbp in length were removed as they are most likely to result from low coverage and are less reliable in our experience.

Multi-locus variable number tandem repeat analysis (MLVA)

A virtual PCR (vPCR) was performed in order to obtain MLVA data from the assembled genomes. Primers designed by Lindstedt *et al. (24)* and Lobersli *et al. (25)* were used to search assemblies and find possible PCR products with the python script, mlva.py (https://github.com/tallnuttcsiro/mlva.py/blob/master/mlva.py). A minimum spanning tree was generated using the following MLVA repeat regions; CVN001, CVN004, CVN007, CVN014, CVN015, CVN016 and CVN017. Isolates that had undetermined values for CVN016 (n = 5) and CVN017 (n = 3) were assigned a value that was equal to the most frequently observed for each corresponding locus in order to allow analysis using the following method, without exclusion of samples with missing values. Repeat values were imported to BioNumerics (version 7.6; Applied Maths, Austin, TX) as character data that was subsequently used to generate a minimum spanning tree (MST). The MST was colour coded by serotype and overlayed with *stx* gene information that corresponds with the isolates in each node.

Comparative genome fingerprinting (CGF)

The Center for Genomic Epidemiology (CGE) virulence gene database (https://cge.cbs.dtu.dk/services/VirulenceFinder/) was used as the basis for identification of virulence genes in the assembled genomes and corresponding Illumina MiSeq reads. The gene content of the *E. coli* virulence gene database is described by Joensen *et. al.*, (2014) *(21)*. Assembled genomes or MiSeq reads were submitted for online identification and a local version of the blast procedure for identification was also performed using a python script (https://github.com/tallnuttcsiro/pathtype.py) with a local copy of the virulence gene database. Genes were identified as present if Blast hits had 95% of higher identity and 80% or more of their length. In addition to the CGE virulence gene database, further isolate information was determined using the CGE's multi locus sequence typing (MLST; https://cge.cbs.dtu.dk/services/pMLST/) and

serotype finder (https://cge.cbs.dtu.dk/services/SerotypeFinder/) schemes. Finally, the distribution of virulence genes on pathogenicity islands (PAIs) or throughout the genome that are not included in the CGE virulence gene database were examined by *in silico* or virtual PCR (vPCR) using primers shown in Table 2 *(13, 22, 30, 36, 39)*. vPCR uses existing primer pair combinations and conducts simulated PCR on assembled genomes. The predicted amplification product size can be determined and compared with the expected size to determine the presence or absence of a gene. The CGE virulence gene database and vPCR were used to identify the presence of 101 *E. coli* virulence genes. A full list of genes is detailed in Appendix 1.

Shiga toxin bacteriophage insertion site (SBI) typing

The insertion sites of *stx*-type prophage were analysed using vPCR (using a python script, https://github.com/tallnuttcsiro/vpcr.py/blob/master/vpcr.py) with primers described in Bonanno *et al (6)*. Two mismatches were allowed between primers and target sequences, with none in the three 3' bases of each priming site. Shiga toxin gene subtyping data (described earlier) was combined with vPCR data on bilateral phage-chromosome junctions for *stx* insertion sites that are common in non-O157 STEC. If in the vPCR either or both bacteriophage-chromosome insertion site locus junctions were detected, the locus was considered occupied. A genotyping code was assigned to each isolate using the characters A, P, S, Ss, T, W, Ye, Y, Z, 1a, 2a to represent the occupation of a prophage at *argW*, *prfC*, *sbcB*, *ssrA*, *torST*, *wrbA*, *yecE*, *yehV*, *Z2577* and the carriage of stx_{1a} and stx_{2a}, respectively.

Pangenome

The pangenome (i.e. all coding sequences found among all isolates) was produced by the following method: 1. coding sequences (CDS) for all isolates were identified using prokka v.1.11 (http://www.vicbioinformatics.com/software.prokka.shtml) in 'fast' mode. 2. All CDS were amalgamated into one multi-fasta file. The Usearch v.8.1.1861 program (http://www.drive5.com/usearch/) cluster_fast command was used to find clusters of CDS with a 60% amino acid identity threshold. 3. The Usearch cluster output was parsed into a pangenome CDS table using a python script (https://github.com/tallnuttcsiro/vpcr.py/blob/master/gene-matrix-from-uclust3.py) and the CDS usearch centroid file was used to obtain annotations for the pangenome table via a Blast search. Pangenome CDS were ranked in order of their high and low frequency in different factors: O26 vs. O111; human vs. cattle origin for O26 isolates; and human vs. cattle origin for O111 isolates. These ranks were determined by sorting the table of all accessory CDS (not

present in all genomes). Each accessory CDS from all the genomes examined was assigned a pangenome score which was calculated by multiplying the highest Usearch percentage identity by the percentage query coverage. Once the pangenome was sorted by the desired parameter, e.g., isolate source, the pangenome scores were summed within the two broad groups. The summed

Table 2. Gene targets and PCR primers for the detection of additional virulence markers

Gene	Genetic support	Primers	Sequence (5' - 3')	Reference
pagC	O-island 122	Z4321-a	ATGAGTGGTTCAAGACTGG	22
		Z4321-b	CCAACTCCAACAGTAAATCC	
terC	OI-43 & OI-48	TerC1	TCCTGGCGCTGAAAGAT	39
		TerC2	GAAACACTCATAAAATAACCTCTT	
aid-1	OI-43 & OI-48	1396-1	ACTGGTTACCAGTACTGCTG	30
		1396-2	ACCAGTCTTCATCGCTGTCA	
nleG2-3	pO157	Primer	GGATGGAACCATACCTGG	13
		Primer 2	CGCAATCAATTGCTAATGC	
nleG6-2	pO157	Primer	CGGGTCAGTGGATGATATGAGC	
		Primer 2	AAGTAGCATCTAGCGGTCGAGG	
nleG5-2	pO157	Primer	TGGAGGCTTTACGTCATGTCG	
		Primer 2	CCGGAACAAAGGGTTCACG	
irp2	HP	irp2 (FP)	AAGGATTCGCTGTTACCGGAC	36
		irp2 (RP)	TCGTCG GGCAGCGTTTCTTCT	
fyuA	HP	fyuA (FP)	GCGACGGGAAGCGATTTA	
		fyuA (RP)	CGCAGTAGGCACGATGTTGTA	
adfO	O-island 57	Forward	TGGTGGCCCGCATACAGC	19
		Reverse	TGCCCAGTCAGCCCAGGTTA	
chuA	Chromosome	Forward	GACGAACCAACGGTCAGGAT	
		Reverse	TGCCGCCAGTACCAAAGACA	
ckf	O-island 57	Forward	ATGCTCGTCACATATAGATTG	
		Reverse	GTTCGTAAGCTGTGAAGACA	
ent/espL2	O-island 122	Forward	GAATAACAATCACTCCTCACC	
		Reverse	TTACAGTGCCCGATTACG	
nleB2	O-island 36	Forward	GTTAATACTAAGCAGCATCC	
		Reverse	CCATATCAAGATAGATACACC	
nleD	O-island 36	Forward	GGTATTACATCAGTCATCAAGG	
		Reverse	TTGTGGAAAACATGGAGC	
nleE	O-island 122	Forward	GTATAACCAGAGGAGTAGC	
		Reverse	GATCTTACAACAAATGTCC	
nleF	O-island 71	Forward	ATGTTACCAACAAGTGGTTCTTC	
		Reverse	ATCCACATTGTAAAGATCCTTTGTT	
nleG2-1	O-island 71	Forward	ACCAGAAACCTGACTTCG	
		Reverse	CAGCATCTTCATATACTACAGC	
nleG5-2	O-island 57	Forward	TGGAGGCTTTACGTCATGTCG	
		Reverse	CCGGAACAAAGGGTTCACG	
nleG9	O-island 71	Forward	GTTCGTGCCCGAATTGTAGC	
		Reverse	CACCAACCAAACGAGAAAATG	
nleH1-1	O-island 36	Forward	GTTACCACCTTAAGTATCC	
		Reverse	GTTTCTCATGAACACTCC	
nleH1-2	O-island 71	Forward	AACGCCTTATATTTTACC	
		Reverse	AGCACAATTATCTCTTCC	
terB	OI-43 & OI-48	Forward	GCCAGGTTGGCCGTTTC	
		Reverse	CCGTCACTCGATACGGCAAT	
TspE4.C2	Chromosome	Forward	GAGTAATGTCGGGGCATTCA	
		Reverse	CGCGCCAACAAAGTATTACG	
ureC	OI-43 & OI-48	Forward	TCTAACGCCACAACCTGTAC	
		Reverse	GAGGAAGGCAGAATATTGGG	
yjaA	Chromosome	Forward	TGAAGTGTCAGGAGACGCTG	
		Reverse	ATGGAGAATGCGTTCCTCAAC	

values were then subtracted to yield a final score used to rank the proteins. Tables were then sorted by the ranking score to determine the top and bottom 50 CDS. The top 50 CDS in each case (i.e. those CDS with the greatest variance in frequency between the two groups being analysed) were submitted to the protein interaction network analysis tool, STRING (http://string-db.org/), which identifies relationships based on numerous different parameters between proteins. It is expected that if CDS that were most present or absent for a particular factor are part of a coherent mechanism or genome feature (e.g. an operon) then they would be clearly shown as an interconnected network by STRING.

Single nucleotide polymorphisms (SNPs)

Three different SNP analyses were performed: one with all isolates vs. the O157, 'Sakai' genome as reference (GenBank BA000007); a second using only O26 isolates with O26 genome NC_013361 as reference; and a third using only O111 isolates with O111 genome NC_013364 as reference. The program, Parsnp v.1.2 was used to find SNPs among the isolates' genome assemblies (https://github.com/marbl/parsnp). In order to reduce the potentially very large number of SNPs to a manageable level, while retaining any structure present in the data, only SNPs which showed a minimum variance of 0.05 among isolates were used for principal components analysis (PCA) and Neighbour-joining phylogenies.

Results

Strain selection

A total of 106 pathogenic STEC isolated from Australian cattle (*n* = 47), Australian human (*n* = 59), Hong Kong cattle (*n* = 1) and USA human (*n* = 1) were selected for inclusion in this study. The strain set broadly represents the PFGE types, sources and isolation dates represented across the 170 isolates that were included in the initial characterisation phase of the study (e.g. those that were screened for PFGE and *stx* gene subtypes). A summary of strain details is provided in Table 3 and Figure 1. Australian cattle isolates were comprised of serotypes O26 (*n* = 36) and O111 (*n* = 11) spanning 20 years and representing 47 distinguishable PFGE profiles while human isolates were comprised of serotypes O26 (n = 33), O111 (n = 22), O121 (n = 1) and O103 (n = 1) spanning 18 years and representing 55 distinguishable PFGE profiles. A further two isolates from Hong Kong and the USA were of serotypes O121 and O45, respectively.

Analysis of isolates for *stx* gene subtypes revealed that only two of the 10 subtypes (stx_{1a} and stx_{2a}) were present across the 106 isolates. Of these, 105 (99%) were shown to possess stx_{1a} and 25 (24%) were shown to possess stx_{2a}. Isolates that possessed stx_{1a} alone (76%) were more common than those that possessed both stx_{1a} and stx_{2a} (23%). Overall, isolates were shown to possess similar *stx* genotypes if they were from the same serotype, regardless of their source of isolation. For instance, the majority of cattle (97%) and human (100%) isolates belonging to serotype O26 were shown to possess stx_{1a} alone while a single cattle isolate was shown to possess stx_{2a} alone. Similarly, the majority of cattle (55%) and human (77%) isolates of serotype O111 were shown to possess both stx_{1a} and stx_{2a} while the remaining isolates were shown to possess stx_{1a} alone. Although no significant difference ($P < 0.05$) in *stx* genotypes was observed between cattle of human isolates from the same serotype (O26 or O111), the occurrence of stx_{1a} and stx_{2a} together was significantly greater in O111 isolates than O26 isolates ($P < 0.05$) and the occurrence of stx_{1a} alone was significantly greater in O26 than O111 isolates ($P < 0.05$).

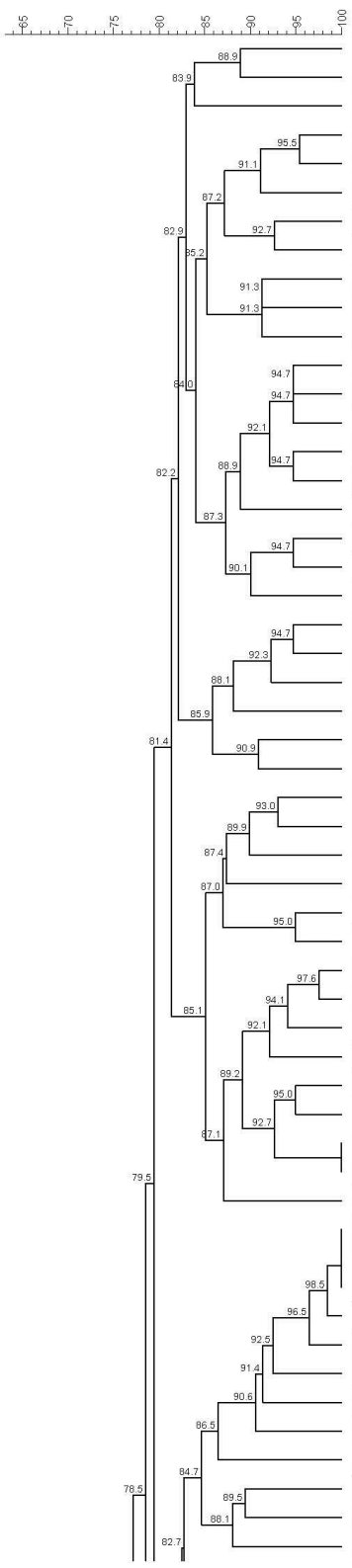
Pairwise analysis of PFGE profiles revealed that 93 of the 106 isolates possessed distinguishable PFGE profiles (Figure 1). A further 13 isolates (12 human and 1 cattle), belonging to five groups, were shown to possess PFGE profiles that were identical to at least one other isolate. Of these, a single cattle isolate was shown to possess an identical PFGE profile with two human isolates. With the exception of a sole O45 isolate, all other isolates grouped according to serotype at a cut-off value of 72%, however, no distinct subgroupings were observed for isolates from different sources.

Table 3. Distribution of pathogenic STEC isolates

	cattle			human		
serotype	n =	*stx* subtypes	distinguishable PFGE profiles	n =	*stx* subtypes	distinguishable PFGE profiles
O26	36	1a (35*); 2a (1)	36	33	1a (33)	32
O111	11	1a (5); 1a 2a (6)	11	22	1a (5); 1a 2a (17)	16
O121	0	-	-	2	1a (2)	2
O103	0		-	1	1a,2a	1

O45	0	-	-	1	1a	1
Total	47	N/A	47	59	N/A	55

* Figures in parenthesis represent the number of isolates possessing each of the *stx* subtype profiles



Key	Serotype	Isolation date	Source
4403	O26	2008	human
4421	O26	2015	human
473b	O26	1997	cattle
4218a	O26	2014	cattle
4414	O26	2014	human
4482a	O26	2004	human
270	O26	1996	cattle
4479a	O26	2002	human
4415	O26	2014	human
4484a	O26	2005	human
4486a	O26	2007	human
4401	O26	2008	human
4489a	O26	2010	human
1113b	O26	1998	catte
3608a	O26	2013	cattle
4490a	O26	2011	human
4164a	O26	2008	human
4492a	O26	2012	human
4162a	O26	2006	human
4422	O26	2015	human
4261a	O26	2014	cattle
4476a	O26	2000	human
4173a	O26	2014	cattle
4264a	O26	2014	cattle
4234b	O26	2014	cattle
4240a	O26	2014	cattle
4047a	O26	2013	cattle
4291a	O26	2015	cattle
673a	O26	1998	cattle
4450	O26	2012	human
4303a	O26	2015	cattle
4166a	O26	2007	human
4329a	O26	2015	cattle
4413	O26	2013	human
640	O26	1997	cattle
4160a	O26	2011	human
4281a	O26	2014	cattle
4407	O111	2009	human
3226	O111	2008	human
4400	O111	2007	human
4042a	O26	2013	cattle
4253a	O26	2014	cattle
4417	O26	2015	human
4423	O26	2008	human
4159a	O26	2010	human
4317a	O26	2015	cattle
4318a	O26	2015	cattle
3430a	O26	2012	cattle
4416	O26	2015	human
4161a	O26	2005	human
4203a	O26	2014	cattle
4326a	O26	2015	cattle
4252a	O26	2014	cattle

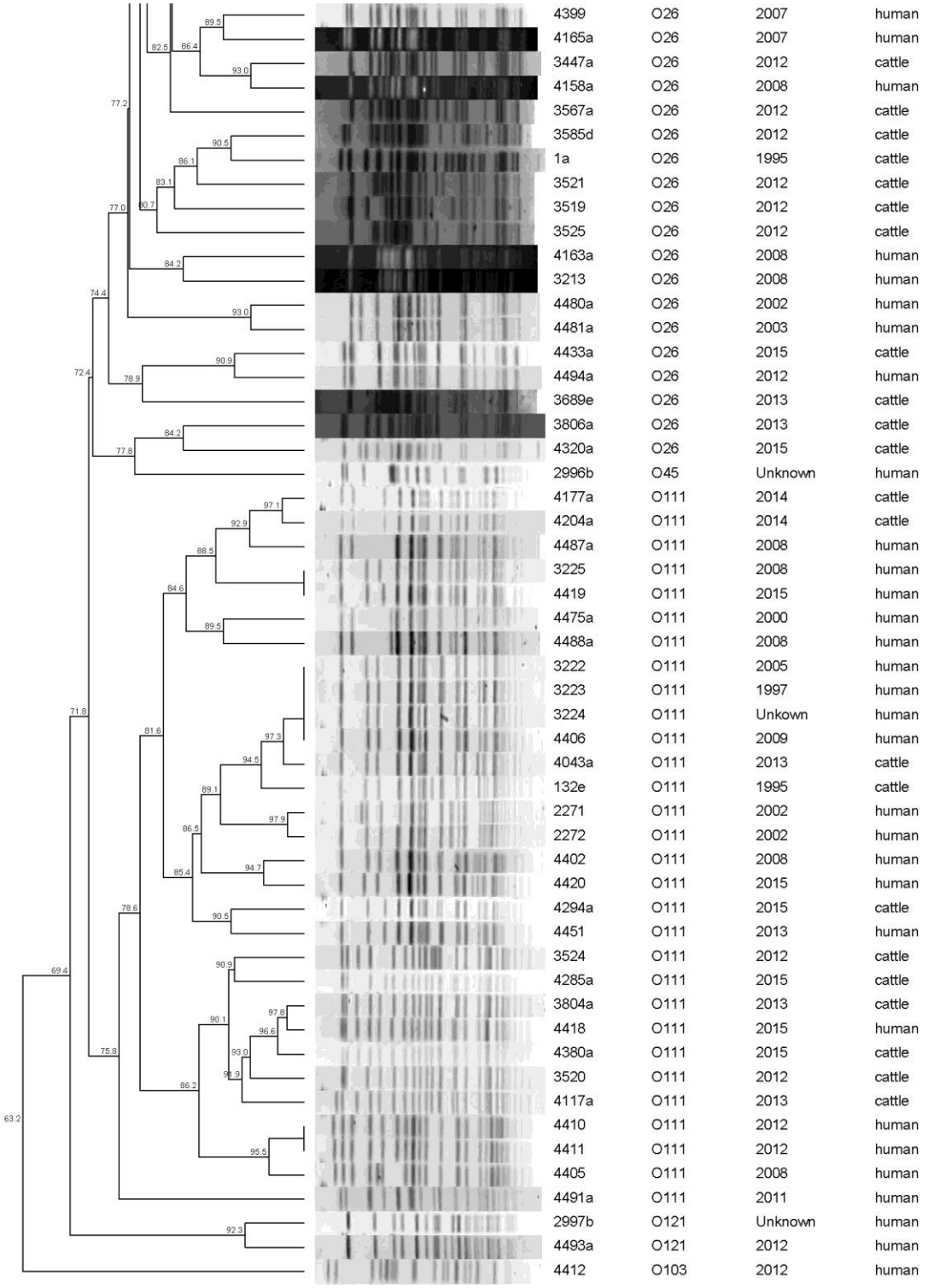


Figure 1: Dendrogram of Australian pathogenic STEC isolates; (key refers to the isolate number within the CSIRO culture collection; PFGE profiles were considered distinguishable if they had a similarity value of less than 100% in the cluster analysis)

Shiga toxin production

Total Shiga toxin production in isolates representing different serotypes and *stx* gene groupings was assessed by ELISA. Overall, the mean production of Stx was greatest in isolates belonging to serotype O111 (1.55 Abs Units) and lowest in those belonging to serotype O26 (0.57 Abs Units). Within each serotype grouping, the mean production of toxin was greater in isolates that possessed stx_{2a} than those that lacked this variant (Figure 2). Excluding groups with less than five isolates, the greatest range in absorbance values was observed in O26 isolates carrying stx_{1a} and O111 isolates carrying stx_{1a} and stx_{2a}. In contrast, O111 isolates that possessed stx_{1a} alone produced relatively low quantities of toxin, spanning a very small range (0.12 Abs Units).

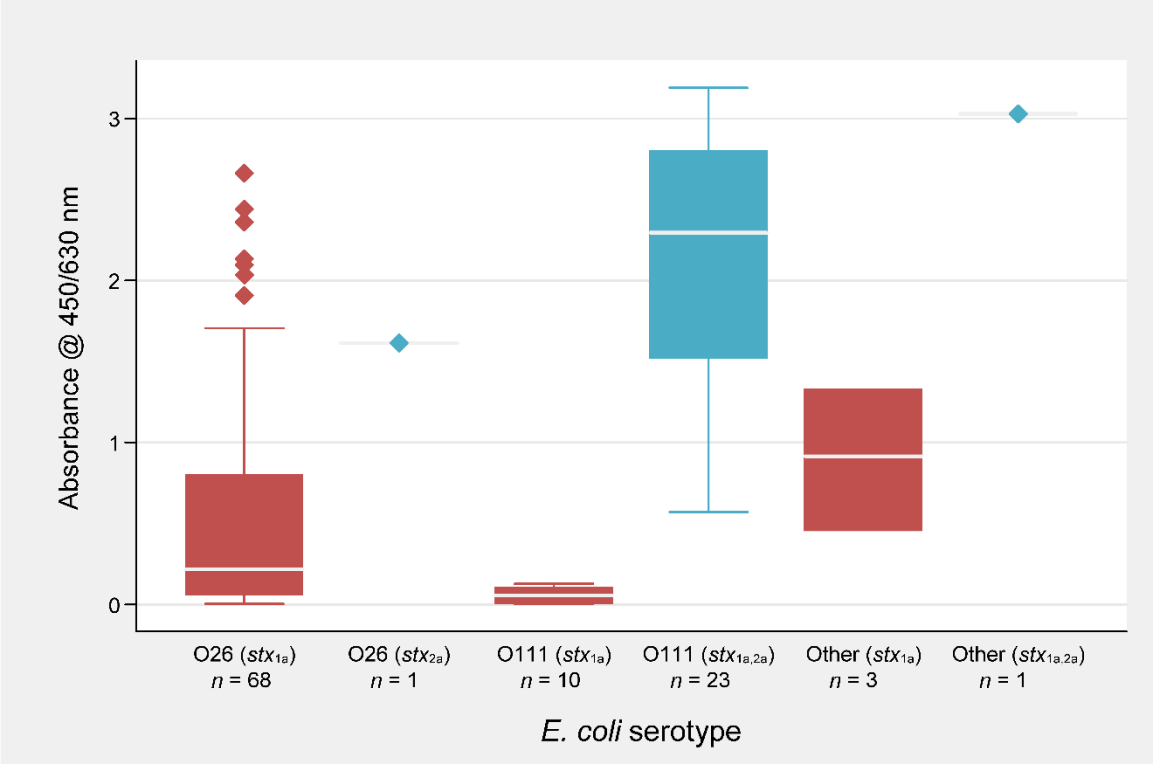


Figure 2. Box-plot of Shiga toxin production of non-O157 isolates grouped by serotype and *stx* subtype. For each box, the lower hinge, upper hinge and inside line represent the 25th (Q1) percentile, the 75th (Q3) percentile and the median, respectively. The interquartile range (IQR) is represented by Q3 - Q1. The upper whisker represents Q3 + (1.5 X IQR), the lower whisker represents Q1 – (1.5 X IQR) and outliers are represented by solid diamonds if data points were above or below Q3 + (1.5 X IQR) and Q1 - (1.5 X IQR), respectively. Similarly, data points in groups containing a single isolate (e.g. O26 (stx_{2a}) are also represented by solid diamonds.

Sequencing statistics

A total of 106 non-O157 EHEC isolates were sequenced during the study. After filtering and clipping, the number of paired reads per genome ranged from 866,205 to 4,221,257 with a mean of 2,096,208, equating to a mean coverage per genome ranging from 37 to 193 with a mean of 91. Assembly of the resulting sequence gave an average of 640 contigs per isolate with a minimum of 106 and a maximum of 2567 contigs. Contig size ranged from 104.5kb to 597.1 kb with an average of 265.8 kb. The average N50 value for all isolates was 88420 bp.

MLVA

A minimum spanning tree (Figure 3) was generated to assess the relationship between isolates with different serotypes and different *stx* gene profiles. Results demonstrate that 80 of 106 isolates possessed unique MLVA types. The 26 isolates that did not possess unique profiles were present across 13 nodes, each representing two isolates. Overall, isolates separated into two main clusters based largely on serotype and *stx* gene profiles though no further separation was observed for isolate source (cattle or human) in either cluster. Interestingly, three O111 isolates (EC3226, EC4400 and EC4407) were shown to group more closely with O26 isolates, as did a small number of other serotypes (O121, O103 and O45). In contrast, no O26 isolates were shown to group with the main O111 cluster.

Comparative genome fingerprinting (CGF)

CGF determined the distribution of *E. coli* associated virulence genes, the predicted serotype, multi-locus sequence type (ST) and the intimin subtype for each isolate. Regardless of source, all pSTEC O26 isolates were H11 and carried the β (beta) intimin subtype. STs were highly conserved amongst the O26 isolates with 67/69 (97%) confirmed as ST21. The remaining two O26 isolates that were not ST21 were ST5429 and were both of human origin. EHEC O111 isolates were more diverse in their H type, ST and intimin profiles. Nineteen of 33 (58%) O111 isolates comprising five of 11 (45%) cattle isolates and 14 of 22 (64%) human isolates were H8, ST294 and carried θ (theta) intimin. The remaining six (55%) cattle isolates and a single (5%) human isolate were O111:H-, ST16 and carried θ intimin. The remaining two O111 profiles were only found in human isolates and were H8, ST16 and θ intimin or H11, ST21 and β intimin with these profiles identified in four (18%) and three (14%) isolates, respectively. The remaining four isolates in the study included two O121:H19 isolates with

ST655 carrying ε (epsilon) intimin, an O103:H25 isolate with ST343 and θ intimin, and an O45:H2 isolate with ST17 and ε intimin.

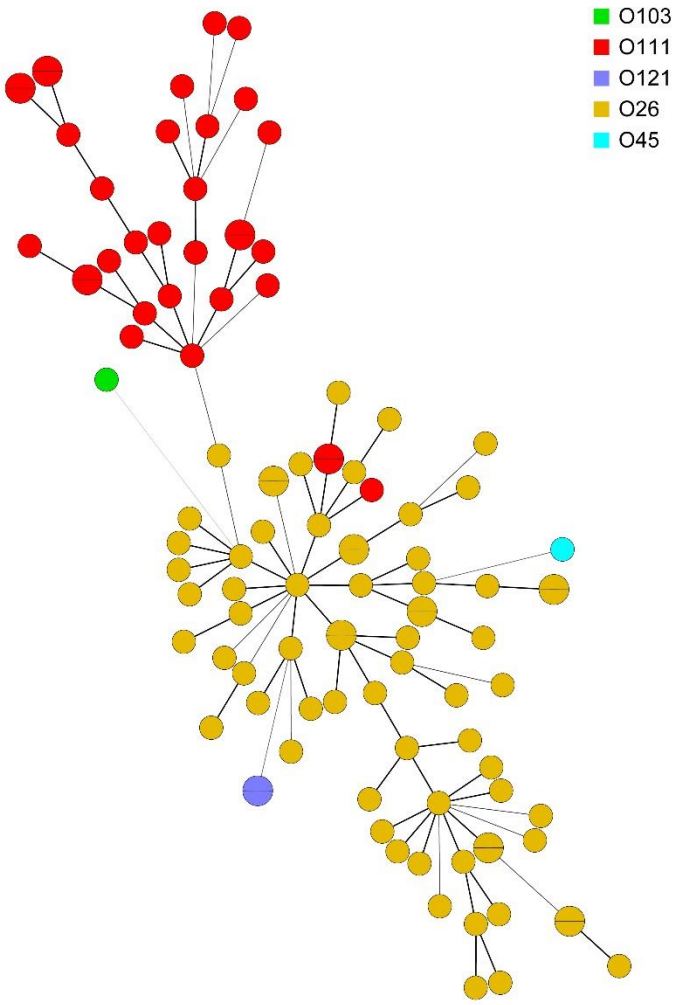


Figure 3. A minimum spanning tree for non-O157 isolates generated using multi-locus variable number-tandem repeats (MLVA). Isolates comprise serotypes O26 (n=69), O111 (n=33), O121 (n=2), O45 (n=1) and O103 (n=1). Each node on the tree represents a unique MLVA type while divisions within nodes represent the number of isolates that contain the same MLVA profile. Isolates are colour coded by serotype as follows; O111, red; O26, orange; O121, purple; O103, green and O45, aqua.

The distribution of virulence genes on the PAIs OI-122 (*ent_espl2, nleE, pagC*), OI-36 (*nleB2, nleD, nleH1-1*), OI43/48 (*aid-1, terB, terC, ureC*), OI-57 (*adfO, ckf, nleG2-3, nleG5-2, nleG6-2*), OI-71 (*nleF, nleG2-1, nleG9, nleH1-2*) and HPI (*fyuA, irp2*) are shown in Figure 4. In general all serotypes demonstrated carriage of each of the PAIs. The exception to this was the carriage of OI-36 by O111 isolates which was not observed in any O111 cattle isolate and only observed in one human O111 isolate (EC4407). The presence of OI-36 associated virulence genes in EC4407 is unsurprising given the other similarities this isolate displays with O26 isolates. On average O26 isolates contained 13.8 of 21 (66%) PAI associated genes with O111 isolates averaging 11.2 (53%) PAI genes. The range of PAI genes in the remaining serotypes was nine (O45) to 15 (O121). No meaningful differences in PAI gene carriage were observed between human and cattle isolates within serogroups. However, differences in gene carriage between serogroups O26 and O111 were evident with O26 isolates significantly ($p<0.05$) more likely to harbor *nleH1-1*, *nleH1-2*, *nleG6-2* and *irp2*. Conversely, O111 isolates were significantly ($p<0.05$) more likely to carry *pagC* than O26 isolates.

Genes involved in secretion (*nleB, nleC, nleA, cif, espF, espJ,* and *espA*) were prominent features of the majority of non-O157 isolates (Figure 5A). The TIR cytoskeleton coupling protein (*tccP*) is also involved in secretion and ranged in prevalence from 24.2% in O26 human isolates to 81.8% in O111 human isolates. The variation in prevalence of *tccP* was also observed in O26 and O111 cattle isolates albeit at less magnitude. The type II secretion gene *etpD* was not detected in O26 or O111 isolates regardless of source but was present in the O45, O103 and both O121 isolates. Aside from Shiga toxin carriage, the presence of genes encoding accessory toxins (*toxB, cba, astA, cma, mchC, mchB, mcmA, mchF* and *celB*) was generally low with only *cba* and *astA* having an overall prevalence >50%. However, differences in toxin gene prevalence between isolates within serogroups of differing source and between serogroups were evident. For example, *toxB* was found in 10 of 33 (30.3%) of O26 isolates of human origin and only one O26 isolate from cattle. Genes encoding colicin B (*cba*) were present in all bar one O111 isolate yet only found in 19 (27.5%) of O26 isolates with no differences observed between cattle and human isolates. Similarly, elevated rates of colicin E2 (*celB*) were observed in O111 isolates when compared to O26 isolates. Although evidence for colicin production in O26 isolates was shown to be lower than in O111 isolates, two O26 isolates of human origin harbored a series of genes (*cma, mchB, mchC, mchF* and *mcmA*) involved in colicin M and H production that were not observed in any O111 isolate.

The prevalence of genes encoding adhesins, fimbrae, serine protease autotransporters (SPATE) and additional miscellaneous *E. coli* virulence markers are shown in Figure 5B. Notable differences were seen between O26 and O111 isolates in relation to secreted protein B (*espB*) and the SPATE (*espP*). Both of these genes were present in at least 93.9% of all O26 isolates but absent from all cattle O111 isolates and only observed in three O111 isolates from humans. Interestingly, the three human O111 isolates (EC3226, EC4400 and EC4407) also possess the typical O26 marker combination of H11, ST21 and β intimin. The adhesin *efa1* and a miscellaneous marker for increased serum survival (*iss*) appeared to be over represented in O111 isolates from humans compared with O111 isolates from cattle. No other notable differences were identified in gene prevalence within serogroups or between sources.

Shiga toxin bacteriophage insertion site (SBI) typing

A Shiga toxin bacteriophage insertion site (SBI) typing scheme was used to investigate the relationship between non-O157 isolates from cattle and human sources. Overall, 17 different SBI types were identified across 106 isolates (Figure 6), however, the bulk of these (*n* = 11) were represented by fewer than five isolates. Despite the increased number of O26 represented in the study than O111, a greater diversity of SBI types were identified in O111 isolates (*n* = 10) than O26 (*n* = 6). The major SBI genotypes identified in O26 were different to those identified for O111, but in most cases were not different between cattle and human isolates of the same serotype. The only notable exception to this was observed in O111 isolates belonging to SBI type A P Ss T Z 1a 2a, in which a greater proportion of human isolates were represented than cattle. Both O26 and O111 isolates displayed insertion of prophage at numerous chromosomal sites, each of which can be considered candidates for *stx* insertion points. For O26, prophage were inserted most frequently at *ssrA* (100%), *z2577* (100%), *torST* (94%) and *yehV* (28%) and least frequently at *sbcB* (6%), *wrbA* (4%) and *argW* (1%). For O111 prophage were inserted at *prfC* (100%) *ssrA* (100%), Z2577 (100%), *torST* (87%), *argW* (83%) and *sbcB* (26%) in isolates carrying stx_{1a} and stx_{2a} and *Z2577* (100%), *ssrA* (100%), *torST* (90%), *prfC* (70%), *argW* (40%) and *wrbA* (10%) in isolates that possessed stx_{1a} alone.

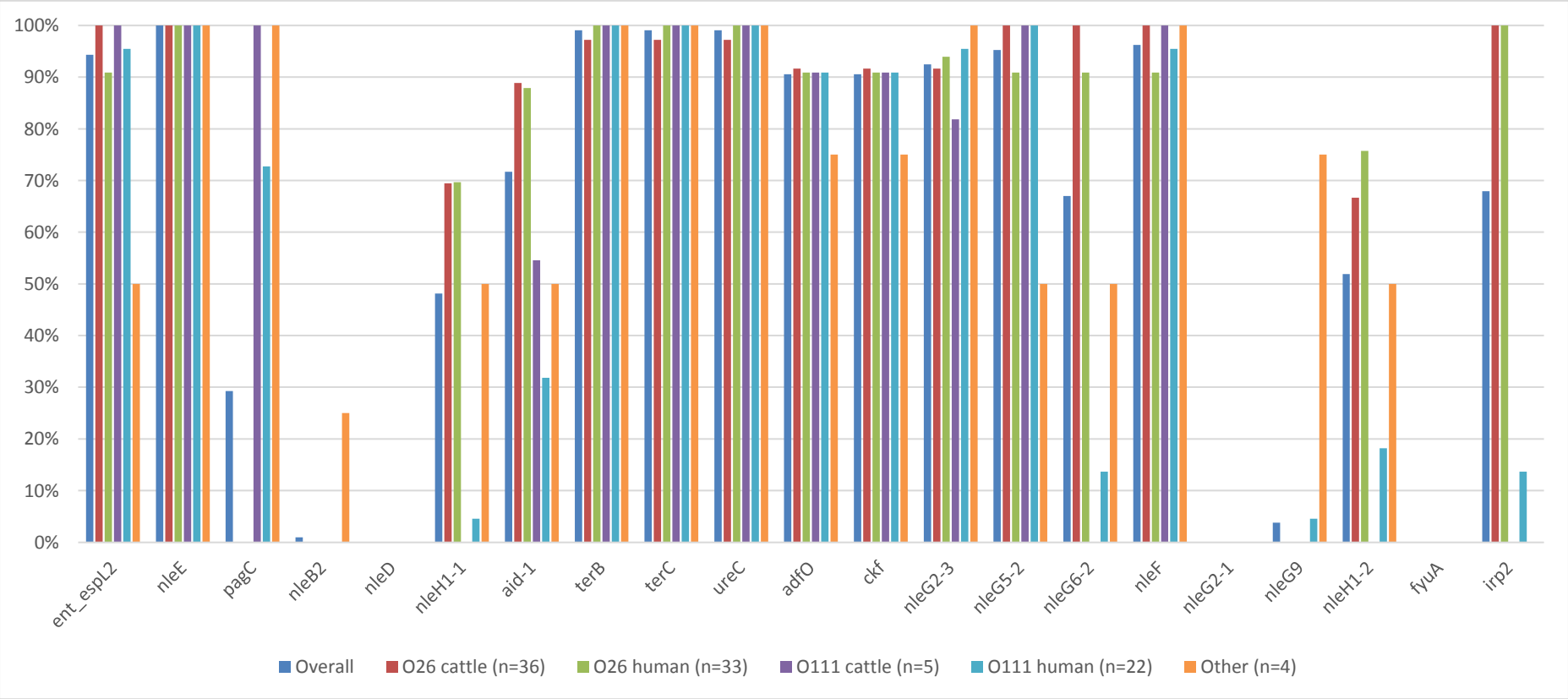
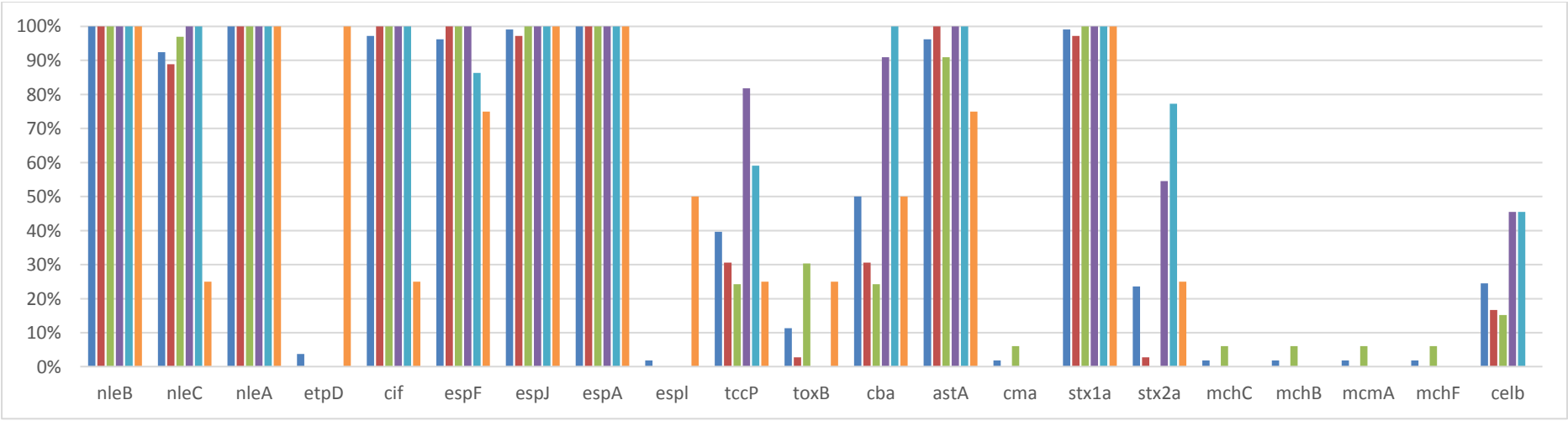
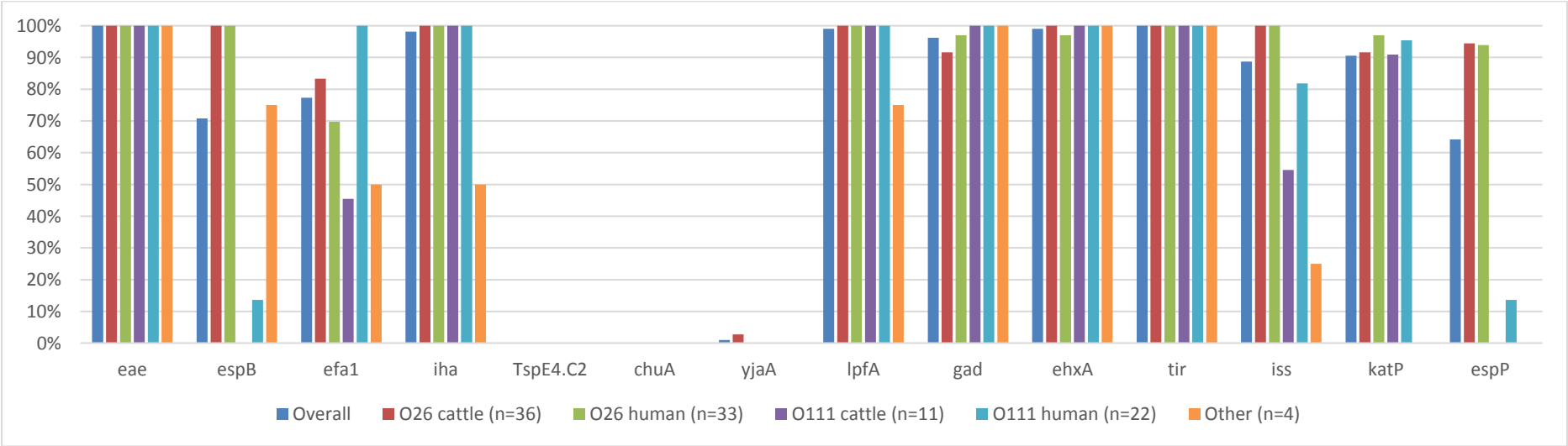


Figure 4. Distribution of PAI associated virulence genes in non-O157 EHEC isolates from human and cattle sources.





Figures 5A (top) & 5B (bottom). Distribution of genes encoding secretion and toxin proteins (Fig. 5A) as well as adhesins, fimbrae, SPATE (serine protease autotransporters of Enterobacteriaceae) and miscellaneous markers (Fig. 5B) in non-O157 pSTEC isolates from human and cattle sources.

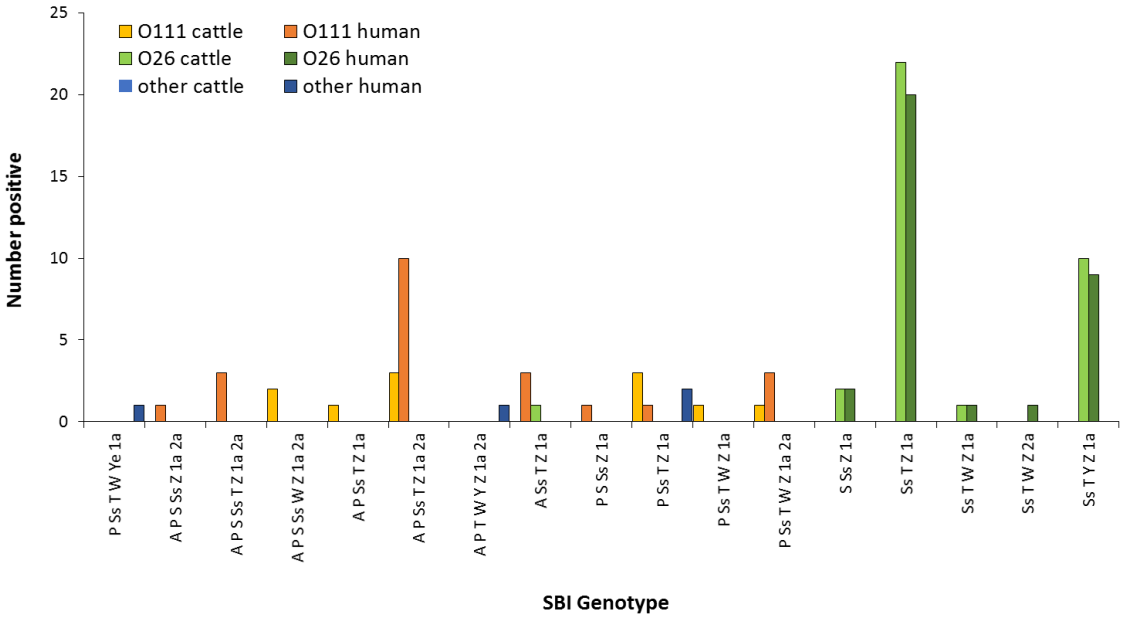


Figure 6. Distribution of Shiga toxin bacteriophage insertion sites present in cattle and human isolates belonging to pathogenic *E. coli* serotypes O26 and O111. Four isolates belonging to serotypes O45, O103 and O121 were grouped together as "other human". Phage insertion sites and Shiga toxin gene subtypes are represented by the following abbreviations; A = *argW*, P = *prfC*, S = *sbcB*, Ss = *ssrA*, T = *torST*, W = *wrbA*, Ye = *yecE*, Y = *yehV*, Z = *Z2577*, stx_{1a} = 1a, stx_{2a} = 2a.

Pangenome

The pangenome content of the isolates in this study was determined to be 7,975 genes. Comparative pangenome analysis using the isolates in this study determined that they have a genome size of 4915 ± 145 genes (mean ± standard deviation). Genome size ranged from 4568 to 5262 genes. It was estimated that 4033 of the genes are present in at least 95% of the genomes analysed regardless of source or serogroup. PCA analysis of the pangenome (Figure 7A & B) revealed clustering by serogroup but not by source with cattle and human isolates interspersed regardless of serogroup cluster. The three human O111 isolates (EC3226, EC4400 and EC4407) that had previously been shown to harbor a range of typical O26 markers continued to demonstrate strong association with the O26 isolates.

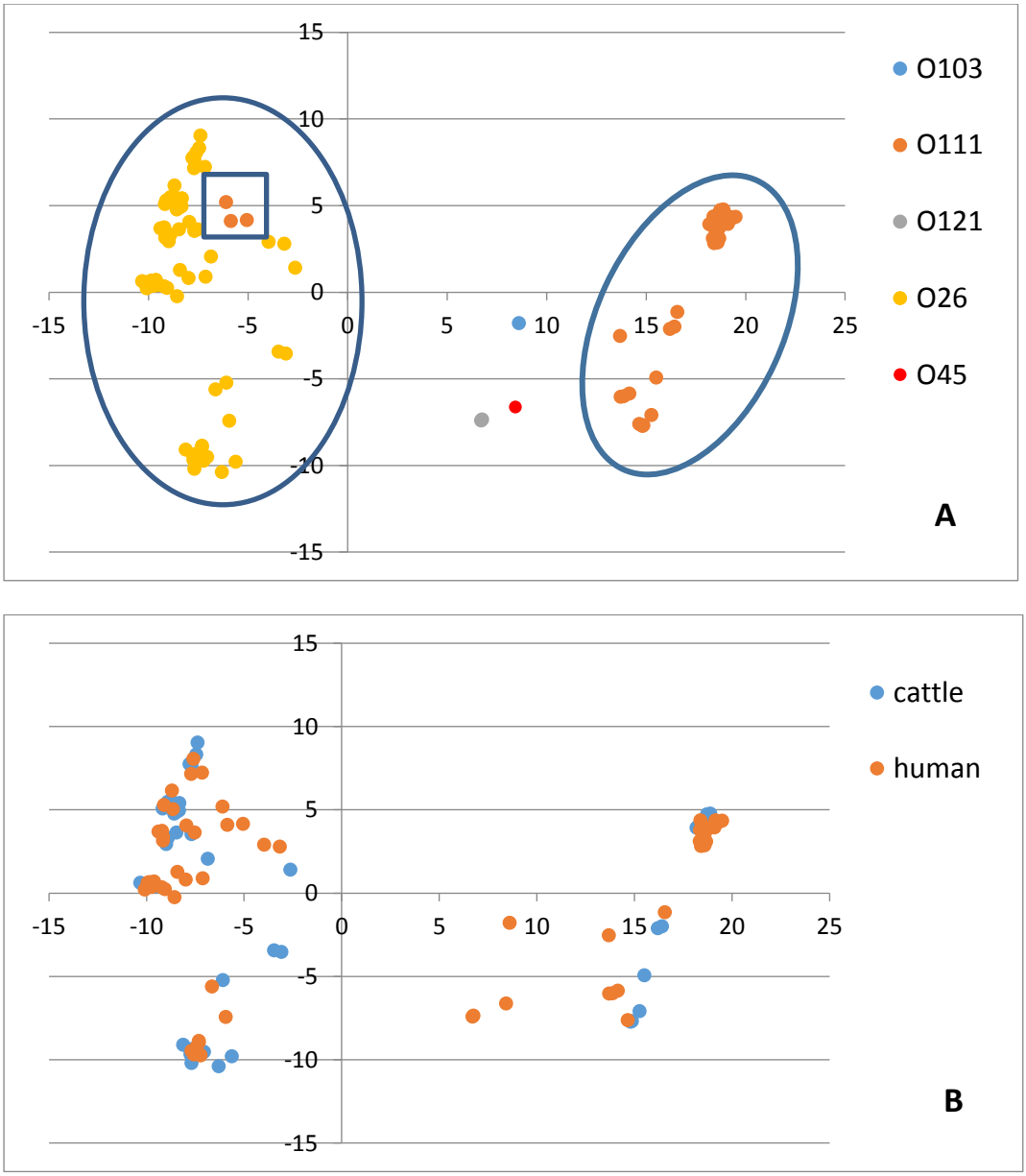


Figure 7A & B. PCA analysis of the pangenome of non-O157. (A) Pangenome content using serogroup as a designator. Isolates enclosed within oval boundaries are either O111 (right hand side) or O26 (left hand side). The three O111 isolates exhibiting O26 characteristics are shown within the left hand oval boundary and are enclosed within the box. (B) Pangenome content using isolate source as the designator.

Functional analysis enabled further interrogation of the pangenome data by determining if the predicted function of proteins encoded by genes with elevated gene prevalence within a source and/or serotype are linked as part of a coherent mechanism or genome feature. Figure 8 shows the protein interaction networks for: A) all O26 v all O111; B) all O111 v all O26; C) O26 human v O26 cattle; D) O111 human v O111 cattle; E) O26 cattle v O26 human; F) O111 cattle v O111 human. The

analysis reveals that O26 isolates are more likely than O111 isolates to harbor genes involved in the sequential transfer of the component sugars of colanic acid (CA) and assembly of the CA polymer (panel A). The predominant feature of O111 isolates on the other hand is the presence of a network of genes that typically comprise the type II secretion system (panel B). Further comparisons of isolates from human or cattle sources within O26 (panels C and E) and O111 (panels D and F) serogroups did not reveal mechanistic or genomic differences of note.

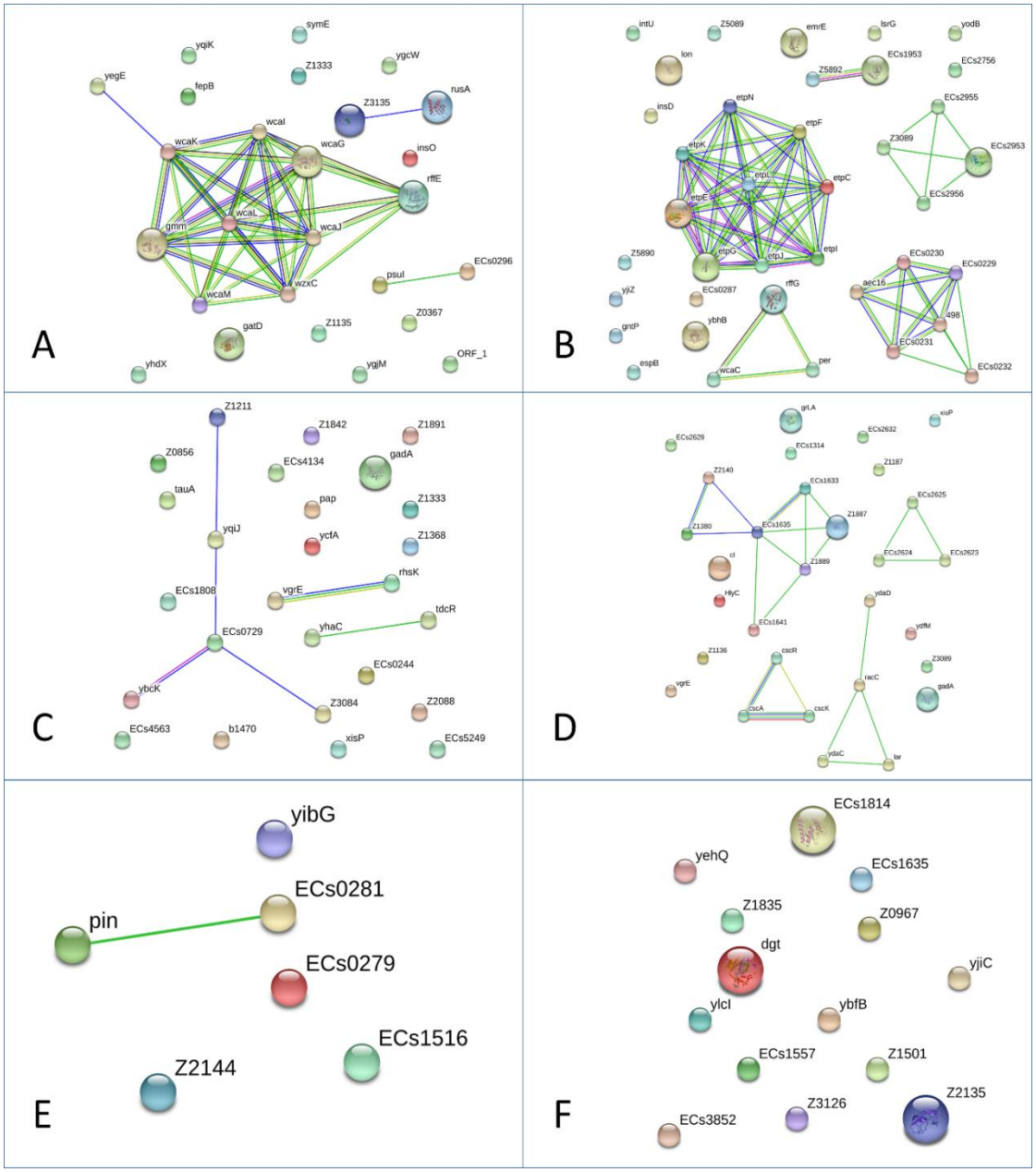


Figure 8. STRING protein interaction networks for non-O157 isolates. A) all O26 vs. all O111; B) all O111 vs. all O26; C) O26 human vs. O26 cattle; D) O111 human vs. O111 cattle; E) O26 cattle vs. O26 human; F) O111 cattle vs. O111 human.

Single nucleotide polymorphisms (SNPs)

SNP analysis identified 94,216 SNPs when using O157 Sakai as the reference; and 4,156 and 8,875 within O26 and O111 serotypes respectively. This indicated that there was less genetic variation in O26 than O111. PCA analysis of SNPs vs. O157 (Figure 9) showed a general clustering of isolates according to serotype, with three O111 isolates being more similar to O26 (4400, 4407 and 3226). No clear clustering of isolates according to their source was evident. Similarly, when O26 and O111 were analysed separately, no clear human or cattle clusters were apparent.

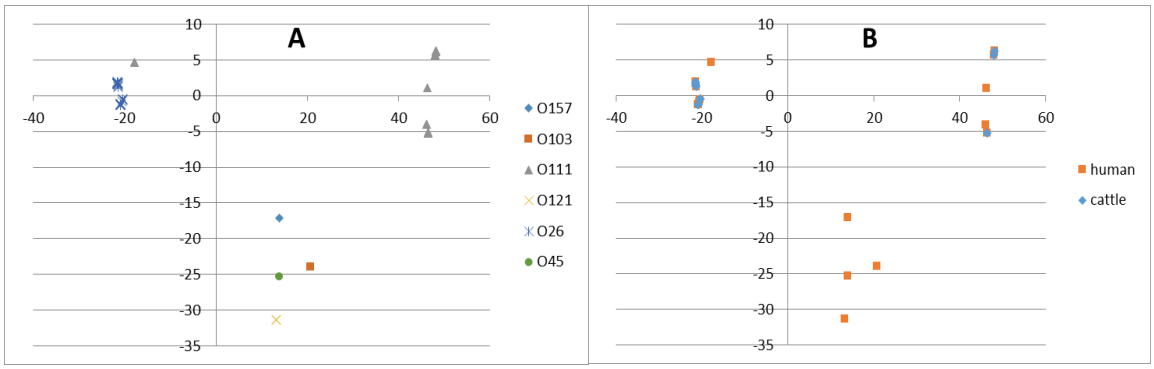


Figure 9. PCA analysis of 5,768 SNPs with variance greater than 0.05 using O157 Sakai (BA0000007) as the reference. A) labelled by serotype; B) labelled by source

Discussion

Shiga toxin-producing *Escherichia coli* (STEC) are important food-borne pathogens capable of causing a variety of disease symptoms from uncomplicated diarrhoea to haemorrhagic colitis (HC) and haemolytic uremic syndrome (HUS) in humans. A subset of STEC that are defined by the presence of *stx* and *eae* and their association with particular serotypes such as O157, O26 and O111 are known as pSTEC. Whilst all pSTEC share a set of cardinal virulence markers there remains an inability to accurately assess the capacity of individual strains to cause severe human disease. The use of next generation sequencing (NGS) enables the virulence characteristics of large numbers of isolates to be rapidly deduced. These findings can then be utilized in a molecular risk assessment or as part of a predictive hazard identification *(19)* to identify isolates with enhanced clinical or epidemic potential. In this study, non-O157 pSTEC isolates sourced from cattle or human sources were subjected to NGS and subsequently analysed using a range of genetic typing techniques. There is limited evidence for the widespread presence of the pSTEC serotypes O45, O103, O121 and O145 in Australian cattle or

as causative agents of clinical disease in Australia. Consequently, the collection of pSTEC isolates examined in this study is dominated by O26 and O111 isolates.

The importance of Shiga toxin subtypes to the virulence potential of pSTEC isolates is well documented *(26, 37)* with carriage of *stx*$_{2a}$ more commonly associated with severe human disease than other *stx*$_2$ or *stx*$_1$ subtypes *(18)*. The presence of *stx*$_{1a}$ is a typical feature of O26 pSTEC from cattle worldwide *(7, 27)* and is again a notable feature of the O26 isolates in this study. However, this study does report the first description of a *stx*$_{2a}$ producing O26 in Australia. Surprisingly this isolate was recovered from a cattle source and is yet to be observed as a causative agent of human disease in Australia. This finding is in contrast to the emergence of *stx*$_{2a}$ producing O26 in Europe over the last 20 years where isolations from food or food animals are rare and instead are more likely to be observed as causative agents of severe human disease *(4, 15)*. The *stx*$_{2a}$ producing O26 identified in this study belongs to clone ST21 and not ST29 and is therefore not related to the highly virulent clones that have recently emerged in Germany and spread throughout Europe *(4)*. Nevertheless, the emergence of *stx*$_{2a}$ producing O26 throughout the world has correlated with increases in associated HUS rates *(10)* and therefore specific investigation of the reservoirs and sources within Australia is warranted.

The link between *stx*$_{2a}$ producing pSTEC and human disease was confirmed by the over representation of *stx*$_{2a}$ producing O111 isolates from human sources in comparison to cattle sources. Whilst the split of *stx*$_{2a}$ producing and *stx*$_{1a}$ producing O111 isolates was reasonably even in cattle sourced isolates, isolates from human sources were three times more likely to be *stx*$_{2a}$ producing than *stx*$_{1a}$-producing alone. Measurement of the levels of Shiga toxin produced by each of the strains in this study further highlights the importance of *stx*$_{2a}$ in the pathogenicity of pSTEC isolates with toxin production greater in isolates possessing *stx* profiles *stx*$_{2a}$ alone or *stx*$_{1a}$ and *stx*$_{2a}$ than those isolates with the *stx*$_{1a}$ profile. These findings are consistent with previous studies that investigated toxin production in Australian O157 isolates *(28)*. Previous studies have demonstrated that *stx*-prophages integrate at specific sites in STEC genomes and that there is a level of conservation across different O serotypes *(6, 31, 38)*. For example, Ogura *et al*. *(31)* identified *stx*-prophage insertion sites in LEE positive isolates occurred at *wrbA* and *ssrA* in O111, *argW* and *prfC* in O103 and *wrbA* in O26. Likewise, Bonanno *et al*. *(6)* identified *wrbA* and *yehV* as common and *yecE* and *sbcB* as less common *stx* integration sites in O26 isolates. In the current study, the frequent insertion of

prophage at ssrA, z2577, torST and *yehV* in O26 and *prfC*, *ssrA*, *z2577*, *torST* and *argW* in O111 suggests that these are potential candidate sites for *stx*-prophage insertion in Australian isolates. However, it is equally possible that *stx*-prophage are present in additional untested sites, such as those described for LEE negative STEC *(38)*. Technical limitations associated with traditional SBI characterisation techniques make it very difficult to categorically link *stx*-prophage with integration sites. Advances in next generation sequences are likely to bridge this gap in the coming years.

Genetic subtyping techniques have for a long period of time been the cornerstone of bacterial subtyping and source tracking investigations. For pSTEC, PFGE could be considered the 'gold standard' subtyping technique. Additional DNA based methods such as MLVA and MLST have provided enhanced resolution and discrimination between strains. The reduction in cost and increase in availability and access to NGS technologies now makes it possible to routinely perform NGS to subtype pSTEC. NGS provides superior resolution to earlier subtyping techniques whilst permitting many of the earlier subtyping techniques to be performed *in silico*. This study utilised PFGE, MLVA and MLST to replicate traditional subtyping approaches and pan genome and SNP analysis as contemporary techniques. Regardless of the technique used, isolates typically grouped based on serotype and there was no discernible segregation of isolates based on their source. These findings contrast those observed with Australian O157 where separation of BBG and CG was observed *(28, 29)*. They do, however, confirm that Australian cattle are likely to be a major reservoir of O26 and O111 pSTEC isolates that are able to cause disease in humans. One exception to the grouping of isolates by serotype was noted with three O111 isolates of human origin grouping with O26 isolates in all subtyping analyses. Further interrogation of those isolates determined that they were H11, ST21 and β intimin, a profile routinely observed in O26 isolates. This observation is noteworthy as it questions the applicability and specificity of pSTEC detection methodologies that employ intimin subtyping *(3)* and rely on the linkage of O111 isolates with theta intimin. Fegan *et al*. *(16)* previously described a recombination event involving the transfer of the O157 somatic antigen to an *E. fergusonii* isolate. The three O111 isolates in this study that display O26 properties would appear to have undergone a similar recombination with the O111 somatic antigen displacing the O26 somatic antigen in this instance. This finding further highlights the complexity faced by scientists and regulators in identifying and defining key virulence markers for use in pSTEC detection, confirmation and characterisation methodologies.

As previously mentioned, NGS data was used to explore the pangenome of all isolates in this study. In addition to the pangenome analysis, a comparative genome fingerprinting (CGF) analysis was conducted by determining the presence of 101 virulence markers in each of the isolates. The molecular profiling of isolates of interest is a useful way to identify groups of isolates with varying pathogenic potential, however it appears to be of most relevance when different pathotypes of *E. coli* are assessed *(1, 17)*. Not surprisingly, the specific focus on pSTEC isolates in this study resulted in decreases in the observed variability in both pangenome and virulence marker content when compared with cross pathotype comparisons. In general, CGF analysis did not identify source associated differences between isolates of the same serotype although some examples were evident. The presence of *toxB* occurred in about one third of human O26 but was only found in one cattle isolate. Ferdous *et al*. *(17)* concluded that the presence of *toxB* was significantly more likely to be associated with isolates recovered from cases of bloody diarrhoea than non-bloody diarrhoea and may be indicative of enhanced virulence. The adhesin *efa1* and a miscellaneous marker for increased serum survival (*iss*) were over represented in O111 isolates from humans than from cattle. In a previous study that attempted to use virulence markers as part of disease screening algorithm *(14)*, *efa1* and *iss* were also over represented amongst isolates causing the most severe human disease albeit not significantly so. Follow up comparisons on isolates causing bloody and non-bloody diarrhoea are consistent with this initial conclusion *(17)*.

Differences were noted when isolates of differing serotypes were compared with each other using CGF. Despite having fewer isolates in the study, O111 pSTEC isolates demonstrated greater variation in ST, H type and intimin subtypes then O26 isolates which were highly clonal in nature. The clonality of O26 isolates is a routinely reported finding of pSTEC characterisation studies *(35)*, particularly prior to the emergence of stx_{2a} producing O26 *(4)*. Nevertheless, CGF can facilitate a greater understanding of the plasticity of the *E. coli* genome with a number of examples evident in this study. The description of the first stx_{2a} producing O26 from Australia and the identification of what appears to be hybrid O26/O111 isolates represent potentially emergent clones of pSTEC that warrant further investigation. Additionally, less obvious examples of genome plasticity or rearrangement can be identified, such as the two O26 isolates of human origin that harbor a series of genes, typically associated with *eae*-negative isolates *(17)* involved in colicin M and H production. Although infrequent in this study and worldwide, the emergence of novel rearrangements in *E. coli* can have disastrous consequences. The 2011 outbreak involving a hybrid *stx* producing enteroaggregative *E. coli* (*stx*-EAEC) clone resulted in the deadliest *E. coli* outbreak in history with

over 4,100 cases, approximately 900 cases of HUS and 54 deaths *(5)*. The *stx*-EAEC clone would not be identified using traditionally pSTEC detection methodologies, however it has reinforced the need for scientists and clinicians to consider the possibility of hybrid organisms when assessing environments, food or clinical cases.

In conclusion, this study is the first to use NGS techniques and analysis pipelines to characterise a collection of Australian non-O157 isolates from human and clinical sources. The use of NGS is less time consuming than traditional analysis methods and despite a level of uncertainty in standardization of processing methods, it provides the highest possible level of resolution for phylogenetic analysis. The collection of isolates established for this study is dominated by O26 and O111 isolates which reflects the relative frequencies that both serotypes are identified in cattle and human clinical samples. The remaining serotypes of regulatory importance (i.e. O45, O103, O121 and O145) to beef producers exporting to the USA are seldom found in cattle or human clinical cases in Australia and are therefore poorly represented here. Unlike previous studies on O157 *(28, 37)* there is little evidence that subgroups of non-O157 exist within Australian cattle populations that are most likely to result in severe human clinical disease. Indeed, there is substantial correlation between isolates of cattle and human origin suggesting that cattle are a major reservoir of non-O157 isolates capable of causing human clinical disease. The plasticity and ability of the *E. coli* genome to undergo substantial recombination events is highlighted within this study and more recently worldwide via the *E. coli* O104 outbreak *(5)* confirming the need to maintain a focus on pSTEC and associated pathotypes as possible vehicles for the development of emerging pathogens.

Acknowledgements

The authors acknowledge Queensland Health (Amy Jennison), Melbourne Diagnostic Unit, University of Melbourne (Mary Valcanis) and the Murdoch Childrens Research Institute (Roy Robins-Browne) for providing pSTEC isolates of human origin.

References

1. Baliere, C., A. Rince, S. Delannoy, P. Fach, and M. Gourmelon. 2016. Molecular profiling of STEC and EPEC strains isolated from French coastal environments. *Appl Environ Microbiol*. 00271-16 (Epub ahead of print).

2. Barlow, R. S., and G. E. Mellor. 2010. Prevalence of enterohemorrhagic *Escherichia coli* serotypes in Australian beef cattle. *Foodborne Pathog Dis*. 7:1239-45.

3. Bibbal, D., E. Loukiadis, M. Kerouredan, C. Peytavin de Garam, F. Ferre, P. Cartier, E. Gay, E. Oswald, F. Auvray, and H. Brugere. 2014. Intimin gene (*eae*) subtype-based real-time PCR strategy for specific detection of Shiga toxin-producing *Escherichia coli* serotypes O157:H7, O26:H11, O103:H2, O111:H8, and O145:H28 in cattle feces. *Appl Environ Microbiol*. 80:1177-84.

4. Bielaszewska, M., A. Mellmann, S. Bletz, W. Zhang, R. Kock, A. Kossow, R. Prager, A. Fruth, D. Orth-Holler, M. Marejkova, S. Morabito, A. Caprioli, D. Pierard, G. Smith, C. Jenkins, K. Curova, and H. Karch. 2013. Enterohemorrhagic *Escherichia coli* O26:H11/H-: a new virulent clone emerges in Europe. *Clin Infect Dis*. 56:1373-81.

5. Boisen, N., A. R. Melton-Celsa, F. Scheutz, A. D. O'Brien, and J. P. Nataro. 2015. Shiga toxin 2a and Enteroaggregative *Escherichia coli*--a deadly combination. *Gut Microbes*. 6:272-8.

6. Bonanno, L., E. Loukiadis, P. Mariani-Kurkdjian, E. Oswald, L. Garnier, V. Michel, and F. Auvray. 2015. Diversity of Shiga Toxin-Producing *Escherichia coli* (STEC) O26:H11 Strains Examined via *stx* Subtypes and Insertion Sites of Stx and EspK Bacteriophages. *Appl Environ Microbiol*. 81:3712-21.

7. Bonardi, S., I. Alpigiani, R. Tozzoli, A. Vismarra, V. Zecca, C. Greppi, C. Bacci, I. Bruini, and F. Brindani. 2015. Shiga toxin-producing *Escherichia coli* O157, O26 and O111 in cattle faeces and hides in Italy. *Vet Rec Open*. 2:e000061.

8. Brandt, S. M., N. King, A. J. Cornelius, A. Premaratne, T. E. Besser, and S. L. On. 2011. Molecular risk assessment and epidemiological typing of Shiga toxin-producing *Escherichia coli* by using a novel PCR binary typing system. *Appl Environ Microbiol*. 77:2458-70.

9. Brooks, J. T., E. G. Sowers, J. G. Wells, K. D. Greene, P. M. Griffin, R. M. Hoekstra, and N. A. Strockbine. 2005. Non-O157 Shiga toxin-producing *Escherichia coli* infections in the United States, 1983-2002. *J Infect Dis*. 192:1422-9.

10. Byrne, L., G. L. Vanstone, N. T. Perry, N. Launders, G. K. Adak, G. Godbole, K. A. Grant, R. Smith, and C. Jenkins. 2014. Epidemiology and microbiology of Shiga toxin-producing *Escherichia coli* other than serogroup O157 in England, 2009-2013. *J Med Microbiol*. 63:1181-8.

11. CDC. 2013. Standard operating procedure for PulseNet PFGE of *Escherichia coli* O157:H7, *Escherichia coli* non-O157 (STEC), *Salmonella* serotypes, *Shigella sonnei* and *Shigella flexneri*. *Centers for Disease Control and prevention, Atlanta, GA*.

12. Clermont, O., D. Gordon, and E. Denamur. 2015. Guide to the various phylogenetic classification schemes for *Escherichia coli* and the correspondence among schemes. *Microbiology*. 161:980-988.

13. Coombes, B. K., M. E. Wickham, M. Mascarenhas, S. Gruenheid, B. B. Finlay, and M. A. Karmali. 2008. Molecular analysis as an aid to assess the public health risk of non-O157 Shiga toxin-producing *Escherichia coli* strains. *Appl Environ Microbiol*. 74:2153-60.

14. de Boer, R. F., M. Ferdous, A. Ott, H. R. Scheper, G. J. Wisselink, M. E. Heck, J. W. Rossen, and A. M. Kooistra-Smid. 2015. Assessing the public health risk of Shiga toxin-producing *Escherichia coli* by use of a rapid diagnostic screening algorithm. *J Clin Microbiol*. 53:1588-98.

15. Delannoy, S., P. Mariani-Kurkdjian, S. Bonacorsi, S. Liguori, and P. Fach. 2015. Characteristics of emerging human-pathogenic *Escherichia coli* O26:H11 strains isolated in France between 2010 and 2013 and carrying the stx_{2d} gene only. *J Clin Microbiol*. 53:486-92.

16. Fegan, N., R. S. Barlow, and K. S. Gobius. 2006. *Escherichia coli* O157 somatic antigen is present in an isolate of *E. fergusonii. Curr Microbiol*. 52:482-6.

17. Ferdous, M., A. W. Friedrich, H. Grundmann, R. F. de Boer, P. D. Croughs, M. A. Islam, M. F. Kluytmans-van den Bergh, A. M. Kooistra-Smid, and J. W. Rossen. 2016. Molecular characterization and phylogeny of Shiga toxin-producing *E. coli* (STEC) isolates obtained from two Dutch regions using whole genome sequencing. *Clin Microbiol Infect*.

18. Franz, E., A. H. van Hoek, F. J. van der Wal, A. de Boer, A. Zwartkruis-Nahuis, K. van der Zwaluw, H. J. Aarts, and A. E. Heuvelink. 2012. Genetic features differentiating bovine, food, and human isolates of shiga toxin-producing *Escherichia coli* O157 in The Netherlands. *J Clin Microbiol*. 50:772-80.

19. Franz, E., A. H. van Hoek, M. Wuite, F. J. van der Wal, A. G. de Boer, E. I. Bouw, and H. J. Aarts. 2015. Molecular hazard identification of non-O157 Shiga toxin-producing *Escherichia coli* (STEC). *PLoS One*. 10:e0120353.

20. Fricke, W. F., and D. A. Rasko. 2014. Bacterial genome sequencing in the clinic: bioinformatic challenges and solutions. *Nat Rev Genet*. 15:49-55.

21. Joensen, K. G., F. Scheutz, O. Lund, H. Hasman, R. S. Kaas, E. M. Nielsen, and F. M. Aarestrup. 2014. Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic *Escherichia coli. J Clin Microbiol*. 52:1501-10.

22. Karmali, M. A., M. Mascarenhas, S. Shen, K. Ziebell, S. Johnson, R. Reid-Smith, J. Isaac-Renton, C. Clark, K. Rahn, and J. B. Kaper. 2003. Association of genomic O island 122 of *Escherichia coli* EDL 933 with verocytotoxin-producing *Escherichia coli* seropathotypes that are linked to epidemic and/or serious disease. *J Clin Microbiol*. 41:4930-40.

23. Köser, C. U., M. T. Holden, M. J. Ellington, E. J. Cartwright, N. M. Brown, A. L. Ogilvy-Stuart, L. Y. Hsu, C. Chewapreecha, N. J. Croucher, S. R. Harris, M. Sanders, M. C. Enright, G. Dougan, S. D. Bentley, J. Parkhill, L. J. Fraser, J. R. Betley, O. B. Schulz-Trieglaff, G. P. Smith, and S. J. Peacock. 2012. Rapid whole-genome sequencing for investigation of a neonatal MRSA outbreak. *N Engl J Med*. 366:2267-75.

24. Lindstedt, B. A., M. Torpdahl, E. M. Nielsen, T. Vardund, L. Aas, and G. Kapperud. 2007. Harmonization of the multiple-locus variable-number tandem repeat analysis method between Denmark and Norway for typing *Salmonella* Typhimurium isolates and closer examination of the VNTR loci. *J Appl Microbiol*. 102:728-35.

25. Lobersli, I., K. Haugum, and B. A. Lindstedt. 2012. Rapid and high resolution genotyping of all *Escherichia coli* serotypes using 10 genomic repeat-containing loci. *J Microbiol Methods*. 88:134-9.

26. Luna-Gierke, R. E., P. M. Griffin, L. H. Gould, K. Herman, C. A. Bopp, N. Strockbine, and R. K. Mody. 2014. Outbreaks of non-O157 Shiga toxin-producing *Escherichia coli* infection: USA. *Epidemiol Infect*. 142:2270-80.

27. Mellor, G. E., N. Fegan, L. L. Duffy, K. E. McMillan, D. Jordan, and R. S. Barlow. 2016. National survey of Shiga toxin-producing *Escherichia coli* serotypes O26, O45, O103, O111, O121, O145 and O157 in Australian beef cattle feces. *Journal of Food Protection*:*In Press*.

28. Mellor, G. E., N. Fegan, K. S. Gobius, H. V. Smith, A. V. Jennison, B. A. D'Astek, M. Rivas, S. Shringi, K. N. Baker, and T. E. Besser. 2015. Geographically distinct *Escherichia coli* O157 isolates differ by lineage, Shiga toxin genotype, and total Shiga toxin production. *J Clin Microbiol*. 53:579-86.

29. Mellor, G. E., E. M. Sim, R. S. Barlow, B. A. D'Astek, L. Galli, I. Chinen, M. Rivas, and K. S. Gobius. 2012. Phylogenetically related Argentinean and Australian *Escherichia coli* O157 isolates are distinguished by virulence clades and alternative Shiga toxin 1 and 2 prophages. *Appl Environ Microbiol*. 78:4724-31.

30. Nakano, M., T. Iida, M. Ohnishi, K. Kurokawa, A. Takahashi, T. Tsukamoto, T. Yasunaga, T. Hayashi, and T. Honda. 2001. Association of the urease gene with enterohemorrhagic *Escherichia coli* strains irrespective of their serogroups. *J Clin Microbiol*. 39:4541-3.

31. Ogura, Y., T. Ooka, A. Iguchi, H. Toh, M. Asadulghani, K. Oshima, T. Kodama, H. Abe, K. Nakayama, K. Kurokawa, T. Tobe, M. Hattori, and T. Hayashi. 2009. Comparative genomics reveal the

mechanism of the parallel evolution of O157 and non-O157 enterohemorrhagic *Escherichia coli. Proc Natl Acad Sci U S A*. 106:17939-44.

32. Prager, R., C. Lang, P. Aurass, A. Fruth, E. Tietze, and A. Flieger. 2014. Two novel EHEC/EAEC hybrid strains isolated from human infections. *PLoS One*. 9:e95379.

33. Rivas, L., E. G. Mellor, K. Gobius, and N. Fegan. 2015. Typing and Subtyping Methods for Pathogenic *Escherichia coli.* p. 67-99. *In*, Detection and Typing Strategies for Pathogenic *Escherichia coli* Springer New York, New York, NY.

34. Scheutz, F., L. D. Teel, L. Beutin, D. Pierard, G. Buvens, H. Karch, A. Mellmann, A. Caprioli, R. Tozzoli, S. Morabito, N. A. Strockbine, A. R. Melton-Celsa, M. Sanchez, S. Persson, and A. D. O'Brien. 2012. Multicenter evaluation of a sequence-based protocol for subtyping Shiga toxins and standardizing Stx nomenclature. *J Clin Microbiol*. 50:2951-63.

35. Schmidt, H., C. Geitz, P. I. Tarr, M. Frosch, and H. Karch. 1999. Non-O157:H7 pathogenic Shiga toxin-producing *Escherichia coli*: phenotypic and genetic profiling of virulence traits and evidence for clonality. *J Infect Dis*. 179:115-23.

36. Schubert, S., A. Rakin, H. Karch, E. Carniel, and J. Heesemann. 1998. Prevalence of the "high-pathogenicity island" of Yersinia species among *Escherichia coli* strains that are pathogenic to humans. *Infect Immun*. 66:480-5.

37. Shringi, S., C. Schmidt, K. Katherine, K. A. Brayton, D. D. Hancock, and T. E. Besser. 2012. Carriage of stx2a differentiates clinical and bovine-biased strains of *Escherichia coli* O157. *PLoS One*. 7:e51572.

38. Steyert, S. R., J. W. Sahl, C. M. Fraser, L. D. Teel, F. Scheutz, and D. A. Rasko. 2012. Comparative genomics and stx phage characterization of LEE-negative Shiga toxin-producing *Escherichia coli. Front Cell Infect Microbiol*. 2:133.

39. Taylor, D. E., M. Rooker, M. Keelan, L. K. Ng, I. Martin, N. T. Perna, N. T. Burland, and F. R. Blattner. 2002. Genomic variability of O islands encoding tellurite resistance in enterohemorrhagic *Escherichia coli* O157:H7 isolates. *J Bacteriol*. 184:4690-8.

40. Vally, H., G. Hall, A. Dyda, J. Raupach, K. Knope, B. Combs, and P. Desmarchelier. 2012. Epidemiology of Shiga toxin producing *Escherichia coli* in Australia, 2000-2010. *BMC Public Health*. 12:63.

41. Wang, J., K. Stanford, T. A. McAllister, R. P. Johnson, J. Chen, H. Hou, G. Zhang, and Y. D. Niu. 2016. Biofilm Formation, Virulence Gene Profiles, and Antimicrobial Resistance of Nine Serogroups of Non-O157 Shiga Toxin-Producing *Escherichia coli. Foodborne Pathog Dis*. Epub ahead of print.

Appendix 1

Table A1. List of virulence genes used for CGF

astA	*fim41a*	*sfaS*	*nleE*
bfpA	*gad*	*sigA*	*nleF*
cba	*hlyE*	*sta1*	*nleG2-1*
ccI	*iha*	*stb*	*nleG5-2*
cdtB	*ipaD*	*stx1A*	*nleG9*
celb	*ipaH9.8*	*stx1B*	*nleH1-1*
cfa_c	*ireA*	*stx2A*	*nleH1-2*
cif	*iroN*	*stx2B*	*terB*
cma	*iss*	*subA*	*TspE4.C2*
cnf1	*K88ab*	*saa*	*ureC*
cofA	*katP*	*tccP*	*yjaA*
eae	*lngA*	*tir*	
eatA	*lpfA*	*toxB*	
efa1	*ltcA*	*tsh*	
ehxA	*mchB*	*vat*	
epeA	*mchC*	*virF*	
espA	*mchF*	*pagC*	
espB	*mcmA*	*terC*	
espC	*nfaE*	*aid-1*	
espF	*nleA*	*nleG2-3*	
espI	*nleB*	*nleG6-2*	
espJ	*nleC*	*nleG5-2*	
espP	*perA*	*irp2*	
etpD	*pet*	*fyuA*	
f17A	*pic*	*adfO*	
f17G	*prfB*	*chuA*	
fanA	*rpeA*	*ckf*	
fasA	*sat*	*ent/espL2*	
fedA	*senB*	*nleB2*	
fedF	*sepA*	*nleD*	

Milestone 5 – Comparison of pSTEC test systems and metagenomics profiling of manufacturing beef enrichment broths

Executive summary

Shiga toxin-producing *Escherichia coli* (STEC) are important foodborne pathogens capable of causing a variety of disease symptoms from uncomplicated diarrhoea to haemorrhagic colitis (HC) and haemolytic uremic syndrome (HUS) in humans. In 2012, the addition of a further six *E. coli* serogroups known as the 'Big6' (O26, O45, O103, O111, O121 and O145) to the STEC testing program prompted a shift in STEC detection technology to methods that detect multiple genetic targets found in strains of STEC associated with human clinical disease. The aim of this study was to compare the performance of commercially available STEC test systems on Australian manufacturing beef enrichment broths and to conduct metagenomics analysis of the broths to further understand their complexity.

STEC test systems can be broadly classified into three groups: classical, advanced, and confirmatory. Classical screening test systems typically detect *stx*, *eae* and O-serogroups, advanced test systems use additional or alternative genetic markers, and the confirmatory test system uses polymerase chain reaction (PCR) coupled with mass spectrometry-based multiplexing to develop a genetic profile of STEC in a sample. Nine STEC test systems comprising five classical screening (BAX, RapidFinder, Qiagen, Biotecon and FSIS), three advanced screening (GDS, Roka and PALL), and one confirmatory (NeoSeek) test system were assessed using 100 potential positive broths generated as part of routine STEC testing. The number of samples testing positive in each test system ranged in total from 39 (Roka) to 85 (FSIS). Systems belonging to the classical test category produced the highest number of PP results with between 64 and 85 samples being categorised in this way. The advanced methods produced substantially fewer PP's with the Roka, PALL and GDS test systems producing 39, 42 and 56 PP results, respectively. The reductions in PPs did not affect the ability to identify those samples that yielded a Big6 STEC during confirmation with Roka and PALL matching the performance of the traditional screening systems in that regard. The GDS system did show a slight reduction with 10 of 12 confirmed positives identified. NeoSeek was the only confirmatory test system assessed in this study and it determined that 16 samples were positive for Big6 STEC. Of the 16 samples deemed positive for Big6 STEC by NeoSeek, 11 were samples that had confirmed for Big6 STEC.

Detecting the genes associated with STEC represents only the first phase of STEC testing and it is necessary to attempt to isolate the STEC strain in order to obtain a confirmed result. Isolation is a challenging process and there are many enrichment broths, selective agars, and isolation aids such as immunomagnetic separation (IMS) available to assist. Understanding the microbial community profiles present in manufacturing beef samples may identify groups of organisms (who is there) or genetic markers (what is there) that hinder or aid isolation. In turn, this knowledge may assist in designing enrichment protocols that specifically select for STEC of interest. Metagenomic sequence reads were obtained from 20 enrichment broths using two HiSeq 2500 rapid runs of 151bp paired end (PE). Metagenomic analysis determined that manufacturing beef enrichment broths are often dominated by non-*E. coli* organisms such as *Clostridium perfringens*, with *E. coli* often comprising less than 10% of the enrichment. Furthermore, the target serogroups are seldom the dominant serogroups in each broth and can be present at ratios lower than 1 in a 1000 *E. coli*. Attempts were made to identify novel genetic markers, away from *stx*, *eae* and O-serogroups, that may be suitable for incorporation into novel screening test systems that are best suited for Australian conditions. The sequencing depth achieved in this study was insufficient to identify additional markers and may need to be revisited in coming years as sequencing technology improves further.

The results of this study indicate that the test systems being currently used in Australia's STEC testing program are comparable to other available test systems that belong to the same test kit category (e.g. classical or advanced). Although this study did not evaluate the cost, capital investment and labour intensity of each test system, it is clear that systems that utilise additional or alternative genetic markers may substantially reduce the number of PP samples requiring confirmation. The use of such systems could reduce testing costs but more importantly would substantially reduce the costs of holding product prior to export. In addition, the use of PCR-based confirmation as opposed to culture confirmation is being assessed internationally for a range of foodborne pathogens. The results of this study do provide encouragement for the industry to continue to explore the implementation of STEC confirmation systems such as NeoSeek STEC as a means of reducing the costs and timeframes associated with the STEC testing program.

Contents

Executive summary 98

Project objectives 101

Success in achieving milestone 101

Introduction 103

Materials and Methods 105

- Sample collection 105
- Initial STEC screening 105
- Additional STEC testing 105
- Quantitative PCR 107
- Bioinformatic methods 107

Results and Discussion 111

Methods comparison 111

- Sample summary 111
- Detection of confirmed samples 111
- Concentration of *eae* and O-antigen genes 113

Metagenomic assessment 116

- Taxonomic abundance 116
- *E. coli* virulence markers and O antigen genes 118
- Mapping of *eae*, stx_1 and stx_2 122

Conclusions 124

- Acknowledgements 126

References 127

Appendix 1 129

Appendix2 136

Appendix 3 140

Project objectives

5. Determine the relatedness of 'Big 6' Shiga toxin-producing *Escherichia coli* strains isolated from Australian cattle with those isolated from Australian human clinical cases
6. Genetically characterise pathogenic Shiga toxin-producing *E. coli* strains from Australian beef and cattle samples and assess improvements in specificity of emerging commercial confirmation test systems
7. Engage with leading international researchers, share information with these researchers and others and contribute to international consensus on defining the risk of *E. coli* found in cattle and beef to human health
8. Phenotypically and genetically characterise the antibiotic resistance status of enterococci and *E. coli* isolated from beef cattle at slaughter

Milestone 5 primarily addresses project objective number 2 by comparing the effectiveness of pSTEC test systems currently used in the Australian beef industry with a range of alternative test systems with varying degrees of similarity and complexity. The challenges associated with isolating non-O157 pSTEC were also investigated by exploring the genotypic and phenotypic diversity of manufacturing beef enrichment broths.

Success in achieving milestone

Milestone 5 details the performance of a variety of STEC test systems available for use in Australia on 100 potential positive (PP) broths generated as part of the Australian STEC testing program for raw non-intact beef products. It demonstrated that test systems belonging to the same test category are comparable to each other and that reductions in PPs can be achieved by using test systems that incorporate additional or alternative STEC genetic markers. One PCR-based confirmation method was assessed and its performance was comparable, though not equal to the culture confirmation outcomes and consequently it may be appropriate for the industry to consider further evaluation of novel STEC test systems. In addition to comparing the performance of STEC test systems, metagenomics analysis of manufacturing beef enrichment broths was conducted. Metagenomic analysis determined that broths are routinely dominated by *C. perfringens* and *E. coli* is often present at concentrations below 10%. Furthermore, target serogroups are seldom the main *E. coli* serogroup present and suggests that ratios between all *E. coli* and target *E. coli* often exceed 1000:1. Analysis of the genes present in enrichment broths did not provide additional insights that could assist in distinguishing broths that confirm from those that do not, however there are many *E. coli*-

associated virulence genes present in enrichment broths that test systems must overcome. In general, the outcomes of this study confirm that the STEC testing program utilised in Australia is appropriate, however the industry should continue to focus on changes to screening or confirmation methods that could result in savings associated with holding product and testing costs.

Introduction

Shiga toxin-producing *Escherichia coli* (STEC) are important foodborne pathogens capable of causing a variety of disease symptoms from uncomplicated diarrhoea to haemorrhagic colitis (HC) and haemolytic uremic syndrome (HUS) in humans. Internationally, several pSTEC serotypes have been associated with sporadic and outbreak cases of foodborne disease. A number of studies have implicated beef products in cases of foodborne disease and beef cattle have been shown to be a source of STEC *(10)*. Epidemiological investigations have demonstrated that alongside STEC serogroup O157, an additional six serogroups: O26, O45, O103, O111, O121 and O145, known as non-O157 STEC or 'Big6', account for the majority of non-O157 illness in the USA *(3)*. As a consequence, Australian beef exporters wishing to send raw non-intact beef products to the USA or Canada must undertake testing for STEC (O157 and Big6).

The addition of the Big6 serogroups to the STEC testing program in 2012 prompted a shift in STEC detection technology to methods that detect multiple genetic targets found in strains of STEC associated with human clinical disease. The Australian beef industry typically uses either the BAX System Real-Time PCR STEC Suite (Hygiena) or the Assurance GDS MPX STEC assays (Biocontrol) for STEC testing. These systems are two examples of STEC test systems that rely of the detection of three genes: two genes (*stx* and *eae*) that have been shown to be important in the development of human clinical disease and an O-antigen gene belonging to O157 or Big6. Despite this, some manufacturing beef samples may be potentially positive for STEC (from a screening test) without containing a culturable O157 or Big6 STEC. In this scenario, beef exporters face unwanted distribution delays and economic impacts. Novel approaches to STEC detection may incorporate additional genetic targets that enhance the specificity of the detection system and subsequently reduce the impact on beef producers. Rapid advances in the development of whole genome sequencing technologies is enhancing our knowledge of STEC and permits the identification of novel targets for STEC detection. Refinements of STEC testing systems have focused on the inclusion of additional or alternative genetic markers to more accurately determine the likelihood of STEC presence in a sample. The Atlas STEC EG2 Combo Detection assay (Roka Bioscience), Neoseek STEC (Neogen), and GeneDisc System (PALL) employ additional or alternative genetic markers as part of their STEC testing strategy, although they are yet to be used commercially within the Australian beef industry. Similarly, Dellanoy and colleagues *(6)* have proposed the use of *espK* and *espV* for more reliable detection of STEC. Whilst the commercial performance of the refined STEC test systems remains largely unknown, there is demand from industry for methods that reduce the ratio of

potential positive to confirmed positives. Therefore studies that compare their performance against existing STEC test systems are required.

Detecting the genes associated with STEC represents only the first phase of STEC testing and it is necessary to attempt to isolate the STEC strain in order to obtain a confirmed result. Isolation of STEC is an extremely laborious process that often will be unable to isolate an O157 or Big6 STEC. Reasons for the low isolation rate are plentiful, however most STEC strains present in food, animals and the environment are *eae* negative, most are present in low concentrations even after enrichment, and the majority of STEC strains show phenotypic properties very similar to commensal *E. coli* and other Enterobacteriaceae *(1)*. These are considerable factors for isolation to overcome and consequently there are many enrichment broths and selective agars available to assist in isolating STEC. Understanding the microbial community profiles present in manufacturing beef samples may identify groups of organisms or genetic markers that hinder or aid isolation. Chopyk and colleagues, *(4)* have demonstrated that cattle hides with low bacterial diversity correlated with increased prevalence of STEC serogroups and it is hypothesised that low bacterial diversity may correlate with the ability to isolate STEC during the confirmation process.

Screening and isolation are major components of the STEC testing process. Each of which presents substantial challenges to the Australian beef industry. This study will compare the performance of commercially available STEC test systems on Australian manufacturing beef enrichment broths. In addition it will undertake a metagenomics analysis of manufacturing beef enrichment broths to generate additional understanding of the bacterial community and its genetic composition such that improvements to the STEC testing process can be recommended.

Materials and Methods

Sample collection

A total of 100 manufacturing beef enrichment broths collected between July 2016 and January 2017 were included in the study. Samples were included in the study if they were deemed potentially positive (PP) for non-O157 STEC of serogroups O26, O45, O103, O111, O121 or O145 and had been tested using a Department of Agriculture and Water Resources approved STEC confirmation process. De-identified aliquots (~25 mL) of manufacturing beef enrichment broths were provided to CSIRO on a weekly basis for the purposes of conducting all additional STEC testing. Details of which samples had confirmed (i.e. an STEC had been isolated) and the serogroup to which the isolate belonged were provided to CSIRO. No attempts were made by CSIRO to isolate STEC from any of the broths received.

Initial STEC screening

The majority of STEC screening in Australia is presently conducted using either the BAX System Real-Time PCR STEC Suite (Hygiena) or the Assurance GDS MPX Top 7 STEC (BioControl). Enrichment and screening of manufacturing beef enrichment broths for STEC was conducted as described at http://www.agriculture.gov.au/SiteCollectionDocuments/biosecurity/export/meat/elmer-3/approved-methods-microbiological-testing.pdf. Briefly, BAX samples were generated by enriching 375 g of manufacturing beef in 1.5 L of pre-warmed (45-46°C) MP enrichment broth (Hygiena) at 39-42°C for 12-24 h. GDS samples were generated by enriching 375 g of manufacturing beef in 1.5 L of pre-warmed (42°C) mEHEC medium (BioControl) at 42°C for a minimum of 10 hours. Enrichment and screening can be conducted at on plant laboratories (abattoir) or at centralised commercial testing laboratories. All samples included in this study were initially screened and determined to be PP for non-O157 using either the BAX or GDS test systems. Prior to the samples being made available to CSIRO, the commercial laboratory conducted a secondary screening test using the other test system to that which generated the original PP result (i.e. samples that were determined to be PP by BAX were subsequently tested using the GDS system and vice versa).

Additional STEC testing

In addition to the initial screening, all samples were tested using seven additional STEC test systems. Brief details of each system including their manufacturer, test system category and DNA preparation method are shown in Table 1. Testing using the BAX, RapidFinder, Qiagen, Biotecon and FSIS systems

was carried out at CSIRO laboratories in Brisbane, Australia. Testing using the Roka and Neogen systems took place at their company headquarters in San Diego, USA (Roka) and Lincoln, USA (Neogen). PALL GeneDisc testing was conducted by ANSES, France. All testing was conducted as per the manufacturer's recommendations with one exception. Re-suspended RapidFinder reaction mixes were transferred to a MicroAmp fast optical 96-well reaction plate (Applied Biosystems) prior to PCR.

Table 1. STEC test system details

Manufacturer	Test category	Name	Sample preparation method
Hygiena	Classical	BAX System Real-Time PCR STEC Suite.	Protease based lysis (included in Bax Suite)
Thermo Fisher Scientific	Classical	RapidFinder STEC	PrepSeq Rapid Spin Sample Preparation Kit
Qiagen	Classical	Mericon *E. coli* STEC O-Type	Qiagen DNeasy Blood & Tissue Kit
Biotecon Diagnostics	Classical	Foodproof STEC LyoKit	Foodproof StarPrep One Kit
FSIS	Classical	Detection and Isolation of non-O157 Shiga Toxin-Producing Escherichia coli (STEC) from Meat Products and Carcass and Environmental Sponges	Qiagen DNeasy Blood & Tissue Kit
BioControl	Advanced	Assurance GDS MPX for Top 6 or 7 STEC	PickPen IMS
Roka Bioscience	Advanced	Atlas STEC EG2 Combo Detection Assay	G2 Sample transfer tubes
PALL	Advanced	GeneDisc System	Extraction Pack Food 1
Neogen	Confirmatory	NeoSeek STEC	Boiled cell lysis

Quantitative PCR

All PCR-based test systems will have a limit of detection below which a negative result will be obtained. In general, PCR has a theoretical limit of detection of 4.00 $\log_{10}$ copies/mL and consequently enrichments broths containing target genes at or below this concentration are a challenge for test systems. Test systems may vary in their PCR efficiency and specificity which may affect the ability to identify some gene targets which may ultimately influence the outcome of testing. The concentration of *eae* and O-antigen genes was determined using the primer and probes outlined in MLG5B Appendix 1.01 (https://www.fsis.usda.gov/wps/wcm/connect/0330211c-81ab-4e97-a9f3-d425f5759ee1/MLG_5B_Appendix_1_01.pdf?MOD=AJPERES). Standard curves were prepared from CSIRO culture collection isolates and obtained by serially diluting boiled cell lysates (10 min at 100°C) of each isolate. Standard curves were generated using duplicate CT measurements for all dilutions.

Bioinformatic methods

Understanding the microbial community profiles present in manufacturing beef samples may identify groups of organisms (who is there) or genetic markers (what is there) that hinder or aid isolation. In turn, this knowledge may assist in designing enrichment protocols that specifically select for STEC of interest. The samples selected for analysis and the bioinformatics analysis conducted are described below. All command lines and custom scripts used in bioinformatic analyses are provided in:

- https://github.com/bioinformatics-deakin/056_CSIRO_metagenomics_analysis/blob/master/056analysis_pipeline.sh

and the repository:

- https://github.com/bioinformatics-deakin/056_CSIRO_metagenomics_analysis

Metagenomic samples

A total of 20 samples were selected for metagenomics analysis (Table 2). Nine of the 20 samples are confirmed positives comprising eight O26 and a single O157 positive. The remaining 11 samples did not confirm during the culture confirmation process. DNA was prepared using the DNeasy blood and tissue kit (Qiagen) and sequencing libraries were generated using TruSeq Nano DNA Library prep kit

(Illumina). Libraries were run across two HiSeq 2500 rapid runs of 151bp PE at the Ramaciotti Centre for Genomics, University of New South Wales.

Table 2. Samples selected for metagenomics analysis

MC	Date_collected	Confirmed	Serotype, if confirmed
8	27/07/2016	No	
11	3/08/2016	Yes	O26
13	11/08/2016	No	
17	11/08/2016	No	
22	17/08/2016	Yes	O26
28	7/09/2016	Yes	O26
30	14/09/2016	Yes	O26
32	14/09/2016	No	
38	28/09/2016	No	
42	5/10/2016	No	
49	12/10/2016	No	
54	12/10/2016	Yes	O26
56	19/10/2016	No	
63	2/11/2016	Yes	O26
64	2/11/2016	Yes	O26
75	22/11/2016	No	
78	30/11/2016	No	
79	30/11/2016	No	
82	7/12/2016	Yes	O26

83	7/12/2016	Yes	O157

Read quality control and clipping

Illumina 150 bp, paired end reads from 20 samples were quality clipped using Trimmomatic v0.36 *(2)*. Clipping was performed with a sliding window of 5 bp and a minimum quality Phred score of 25. Trimmomatic also removed any remaining Illumina adapters. Reads were then interleaved and converted to fasta format for further analyses, except for mapping, which used clipped paired fastq data.

Taxonomic assignment and abundance

Metaphlan2 provided a taxonomic abundance analysis of shotgun metagenomic data by mapping reads to a database of 'elite' microbial genes that have been selected for their phylogenetic utility *(13)*. Taxonomic assignment is more reliable for longer reads, therefore paired reads were merged where possible using USEARCH v8.1.1861 *(8)*. To normalise abundances across samples, the number of reads mapped using Metaphlan was rarefied to the lowest number of paired reads observed for a sample (753903 for sample MC38).

Assembly

The metagenome assembler, IDBA_UD v1.1.3 *(12)* was used to assemble sequences prior to searches for specific genes and genome features. IDBA iterates assembly over mulitple kmer lengths. We used sizes of 50, 100 and 150 bp with a minimum contig size corresponding to read length (150 bp), which prevented the loss of any read data due to it not being assembled.

Metagenome binning

In order to compare sample taxonomic composition to the merged read-based Metaphlan approach, MaxBin v2.2.1 *(14)* was used to assign assembled contigs to bins based on kmer frequency, abundance and GC content with default settings. The taxonomy of bins for each sample was assigned using Metaphlan and a taxonomy table was produced. The genomic composition of bins was visually examined by BLAST alignment to a selection of genomes from the taxonomy table, followed by making a circular diagram of the alignments with the python script, circles1.4.py

(https://github.com/bioinformatics-deakin/056_CSIRO_metagenomics_analysis/blob/master/circles1.4.py)

Gene searches

We used a locally running python script (pathtype.py): a version of the online bacterial gene search tools hosted at: https://cge.cbs.dtu.dk/. These tools used a BLAST (https://blast.ncbi.nlm.nih.gov) search of the assemblies to identify the presence of *E. coli* virulence genes and O-type genes. Search parameters were set to require at least a 90% sequence identity match to at least 10% of the length of the reference gene.

Gene mapping

All clipped reads were mapped to *E. coli* Shiga toxin genes, *stx*1 and *stx*2, and intimin gene, *eae*, using the mapping program, bbmap v36.11 (https://sourceforge.net/projects/bbmap/). The number of reads mapped to each bp of the genes was plotted in bar charts in order to display the distribution of read alignment.

Results and Discussion

Methods comparison

Sample summary

Of the 100 PP samples included in the study, 64 were tested at on site laboratories and 36 were tested at a commercial laboratory. In total, 61 PP's were generated using BAX and 39 by GDS. Agreement between the BAX and GDS test systems was low when BAX was used as the initial screening test with only 17/61 (27.9%) samples subsequently confirming as a PP using the GDS system. This difference likely reflects the ability of the immunomagnetic separation step in GDS to remove *stx* and/or *eae* containing *E. coli* that don't belong to a non-O157 serogroup. In contrast, when GDS was used as the initial screening test, the level of agreement with the BAX system was much higher with 29/39 (74.4%) samples confirming as a PP.

Detection of potential positives

The number of samples deemed to be potential positives is summarized in Table 3 and test scores for each samples shown in Appendix 1. The number of samples testing positive in each test system ranged in total from 39 (Roka) to 85 (FSIS). Systems belonging to the classical test category produced the highest number of PP results with between 64 and 85 samples being categorised in this way. The Qiagen and FSIS tests had the most PP calls with 82 and 85 positives, respectively. Both the Qiagen and FSIS tested DNA prepared using the Qiagen DNeasy blood & tissue kit. This kit produces DNA of very high quality and is often used for genomic and metagenomic studies. It is plausible to suggest that this has contributed to the larger number of positive calls. The advanced methods produced substantially fewer PP's with the Roka, PALL and GDS test systems producing 39, 42 and 56 PP results, respectively. A software malfunction resulted in the PALL system providing analysis on 94 of the 100 samples in the study.

Detection of confirmed samples

The likelihood that a potential positive will give rise to a confirmed sample remains low. In this study 12 of the 100 samples that were originally determined to be PP yielded an STEC belonging to the Big6. All 12 samples yielded O26 STEC. This study specifically targeted samples that were PP for Big6, however there were many samples that had been sent for confirmation for O157 and the Big6. Six of these samples yielded O157 isolates. Isolation of O157 and O26 from manufacturing beef

enrichment broths, almost to the exclusion of all other regulated STEC serogroups is typical for Australia and the overall STEC confirmation rate of 18% is consistent with historical trends.

The performance of each of the STEC test systems is shown in Table 4. The majority of tests systems detected 11/12 (91.7%) of Big6 confirmed samples, the exception being GDS which detected 10 (83.3%) of the 12 confirmed samples. The Qiagen and PALL detected all confirmed samples for which they generated a test result. Samples MC28 and MC102 were the samples not categorised as PP by some of the test methods. Analysis of the quantitative data for *eae* and O-antigens confirms that both of these samples had concentrations of target genes at the lower end of what PCR would typically be expected to detect. Sample MC28 had an *eae* and O26 concentration of 4.17 and 3.45 $\log_{10}$ copies/mL, respectively. Similarly, sample MC102 had an O26 concentration of 4.37 $\log_{10}$ copies/mL. The lack of detection of *stx* in MC28 and MC102 by many of the test systems was also a notable feature. Although quantitative data for *stx* was not generated as part of the study, it is possible that *stx* concentrations were also at or below the limit of detection in these samples.

Table 3. Potential positives identified by each STEC test system.

Test system	Test category	Positives
FSIS	Classical	85
QIAGEN	Classical	82
BAX	Classical	67
RAPIDFINDER	Classical	64
BIOTECON	Classical	64
GDS	Advanced	56
PALL	Advanced	42/94*
ROKA	Advanced	39
NEOSEEK	Confirmatory	16

*A PALL software malfunction resulted in six samples not being analysed.

The NeoSeek test system was the only STEC confirmation method evaluated as part of this study. In total NeoSeek concluded that 21 samples were confirmed positive for either O157 or Big6 STEC. Of these, 14 were O26, five were O157 and one sample was O45. A further sample was confirmed

positive for both O26 and O157. When Big6 STEC are considered, NeoSeek concluded that 16 of the 100 samples included in this study were positive. Of the 16 samples identified, 11 were confirmed for Big6 STEC (all O26 STEC) during culture confirmation. Sample MC28 was confirmed as containing O26 STEC during culture confirmation, however the NeoSeek STEC system did not detect *eae* in this sample and it was subsequently deemed to be negative for O26 STEC. Sample MC34 was confirmed as both O26 and O157 STEC positive using Neoseek STEC though only O157 STEC was isolated during culture confirmation. Four additional samples, MC20, 66, 88 and 91 were deemed to be positive using NeoSeek STEC, however these samples did not yield STEC isolates during culture confirmation.

Table 4. Ability of STEC test systems to detect samples that were confirmed for Big6

Test system	Big6 positives detected	Test result for not detected sample
BAX	11/12	MC28 – *eae*, O45
Rapid Finder	11/12	MC28 – *eae*
Qiagen	12/12	
Biotecon	11/12	MC28 – *stx*, *eae*
FSIS	11/12	MC102 – *eae*, O26, O45
Roka	11/12	MC102 – negative
GDS	10/12	MC28 – *stx*; MC102 - negative
PALL	10/10*	No result for MC101, MC102
Neogen	11/12	MC28 – *stx*, O26

*A PALL software malfunction resulted in two confirmed samples not being analysed.

Concentration of *eae* and O-antigen genes

The enrichment of manufacturing beef samples is a critical component of the overall STEC testing approach designed to increase the concentration of STEC and subsequently aid detection and isolation strategies. It must be noted that all samples included in this study were enriched in either mEHEC medium or MP enrichment broth that are the recommended enrichment broths for the

Assurance GDS and BAX test system. STEC test system manufacturers will typically recommend enrichment broths along with time and temperatures for incubation of samples. For example, RapidFinder recommends the use of trypticase soy broth with incubation at 42°C for a minimum of 8 h whereas the PALL system promotes the use of buffered peptone water at 41.5°C for a minimum of 10 hours, to name just a couple. Inclusion of all broth and enrichment combinations within this study was beyond scope, however it is acknowledged that this may have impacted the results produced by the test system, particularly in instances where concentration of the target genes are at or below the theoretical limit of detection (approximately 4.00 $\log_{10}$ copies/mL).

The concentration of *eae* and O antigen genes was produced for all samples tested using the FSIS PCR conditions. The percentage of tests systems that called each sample a potential positive was plotted against the concentration of each target gene and shown in Figure 1. Plots are not available for O111 or O145 as these were not detected in any sample using the FSIS system. Weak positive correlations between increasing test system percentages and an increased concentration of target genes were observed for *eae* (0.42) and O26 (0.29). All remaining target genes had correlation values >0.20. The mean concentration of each target gene for samples with overall test system percentages above or below 70% is shown in Table 5. With the exception of O45, mean concentrations of all targets were at least 3.89 fold greater in samples with overall test system percentages >70% compared to those with scores <70%. The mean concentrations reinforce the challenge associated with the STEC culture confirmation process. Typical enrichment broths have total aerobic counts of between 8.00 and 9.00 $\log_{10}$ CFU/mL. In samples where concentrations of O antigen genes are 5.00 $\log_{10}$ CFU/mL the relative concentration is 1 in 1000 or 0.1% of the enrichment broth. Similarly at 4.00 $\log_{10}$ CFU/mL the relative concentration is 1 in 10,000 or 0.01%.

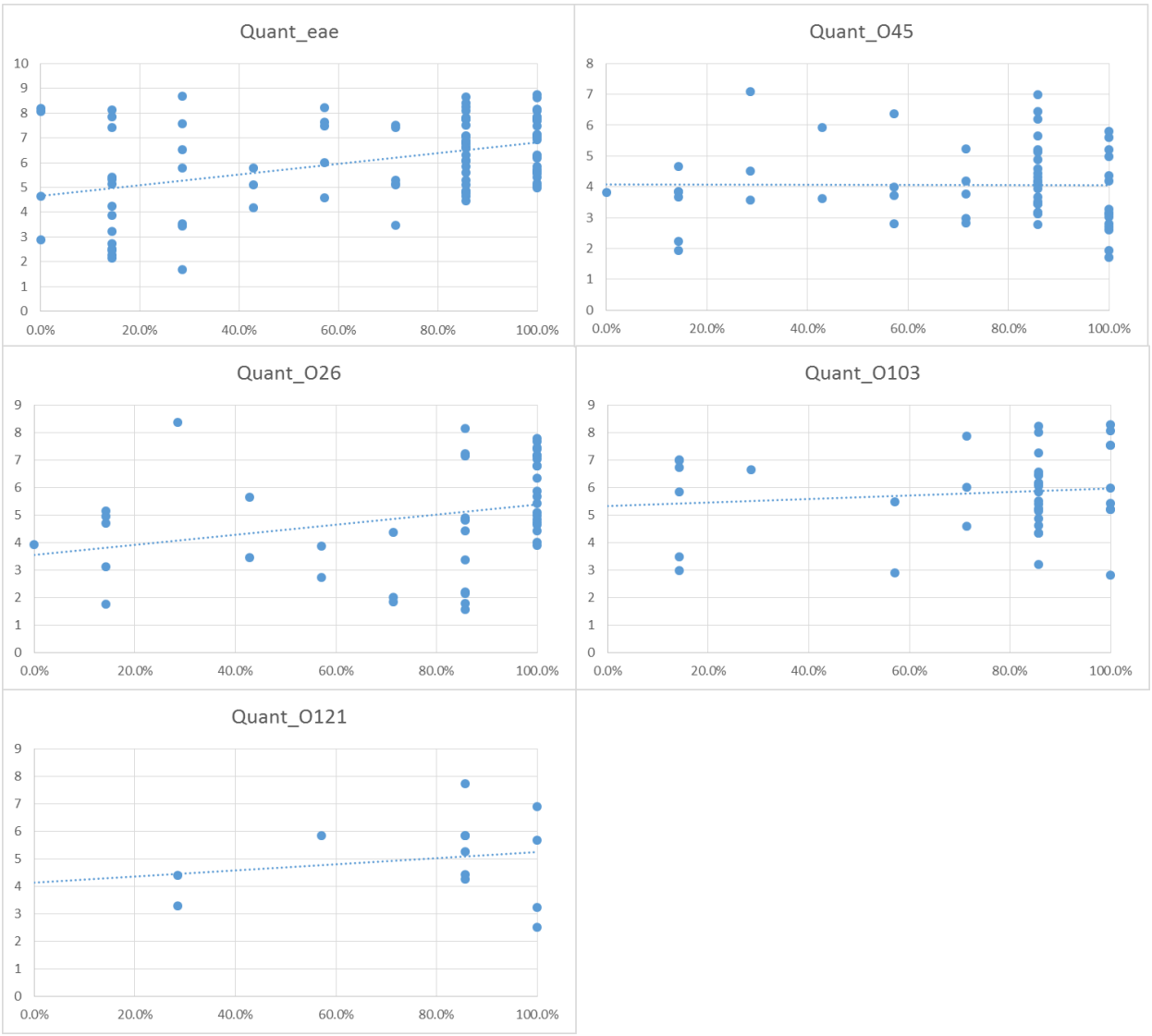


Figure 1. Scatter plots of overall test system percentage versus concentration of target genes. (values on the *x* axis are the overall test system percentage values and values on the *y* axis are $\log_{10}$ copies/mL).

Table 5. Mean concentrations of target genes in samples with overall test system percentages above or below 70%

Target	>70%	<70%	Fold difference
Eae	6.60*	5.23	23.4
O26	5.17	4.34	6.76
O45	4.05	4.10	0.89
O103	5.93	5.34	3.89
O121	5.17	4.51	4.57

*All counts are $\log_{10}$ copies/mL

Metagenomic assessment

Taxonomic abundance

The bacterial diversity of 20 manufacturing beef enrichment broths were assessed to determine if broths with lower diversity are more likely to yield STEC during the culture confirmation process than those with higher bacterial diversity. The taxanomic abundance within nine confirmed samples and 11 PP's that did not culture confirm were determined using Metaphlan2. The composition of each broth at the species level is shown in Figure 2 and Appendix 2. Manufacturing beef enrichment broths were dominated by the presence of *Clostridium perfringens* and *Escherichia coli* with a mean abundance across all samples of 36.3% and 19.7%, respectively. Interestingly, the mean *C. perfringens* abundance was higher in samples that confirmed (42.1%) compared to those that didn't (31.8%). Conversely, the mean *E. coli* abundance was lower in confirmed samples (10.94%) than samples that did not confirm (26.87%). Furthermore, in samples that confirmed for O26 STEC, the mean *E. coli* abundance was 6.95%. Principal component analysis of the profiles of each sample was conducted at genus and species level to determine if notable differences exist between samples that had culturally confirmed and those that did not (Figure 3). The bacterial diversity within enrichment broths that confirmed for O26 or O157 STEC was similar to those from those that did not, indicating that this did not have a major impact on the ability to confirm STEC.

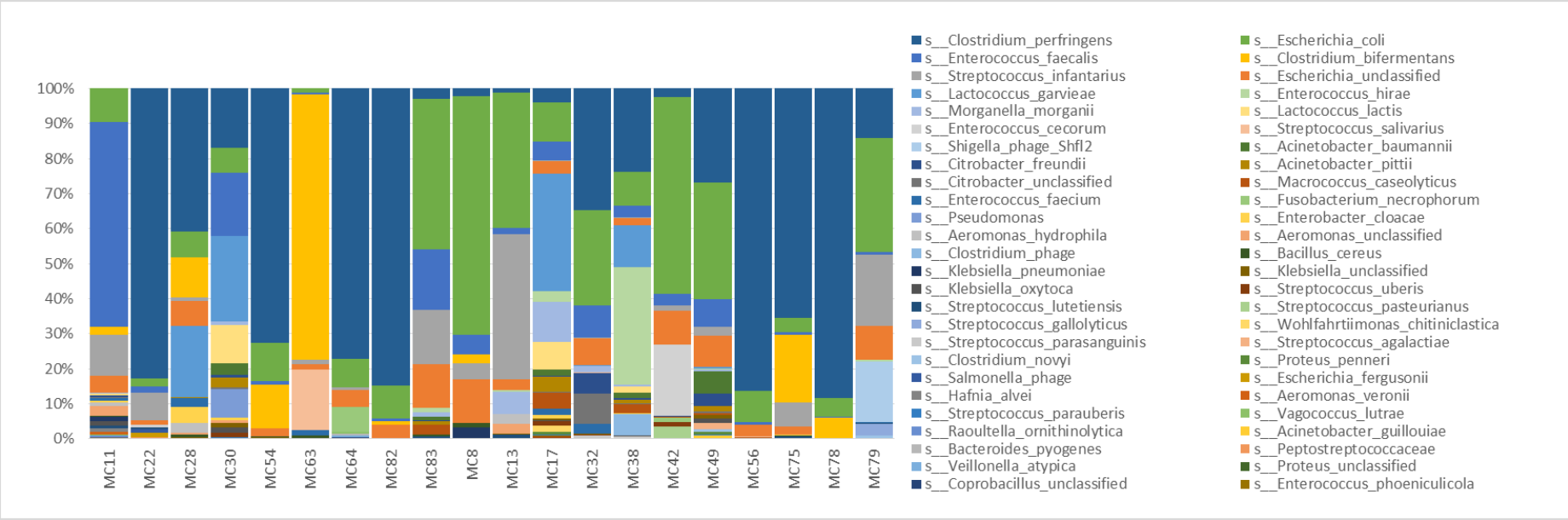


Figure 2. Species composition of manufacturing beef enrichment broths. Samples MC11 to MC82 confirmed for O26, sample MC83 confirmed for O157, and samples MC8 to MC79 did not culture confirm. Each bar represents the proportion (*x* axis) of total metagenomics reads that can be assigned to each species

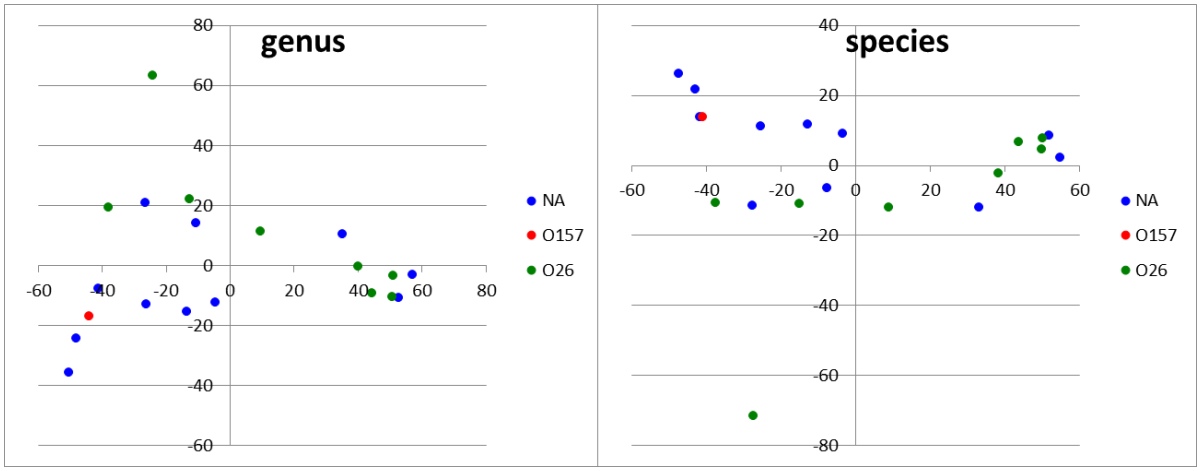


Figure 3. PCA comparison of taxonomic content of manufacturing beef enrichment samples at genus and species level.

E. coli virulence markers and O antigen genes

The development of future STEC test systems will rely on both an understanding of the genetic composition of STEC isolates as well as the total genetic load of all organisms present in manufacturing beef enrichments broths. The challenge associated with developing STEC test systems relates specifically to the ability to identify genes or genetic markers that are most associated with STEC and which do not routinely occur in non-pathogenic strains of *E. coli* or other microflora. Comparing the presence of *E. coli* virulence markers and O antigen genes between broths that confirmed and those that didn't can provide genetic information that can be used to develop STEC test system specificity. The presence of *E. coli* associated virulence markers (Appendix 3) and O antigen genes were determined bioinformatically and are shown in Figure 4. Genes associated with increased serum survival (*iss*) and long polar fimbrae (*lpfA*) were present in all confirmed broths, however they were also present in >80% of non-confirmed samples. Another adherence associated gene (*iha*) had the greatest variation in prevalence between sample groups with 72.7% of non-confirmed and 22.2% of confirmed samples possessing this gene. PCA did not identify correlations between virulence markers and sample groups (Figure 5). However, cross referencing with quantitative data generated for these samples would suggest that only genes present at >1% of the total population are likely to have been sufficiently sequenced in this study. This depth of sequencing is therefore unlikely to have captured the presence of the STEC organisms in their respective enrichments and the results must be considered accordingly.

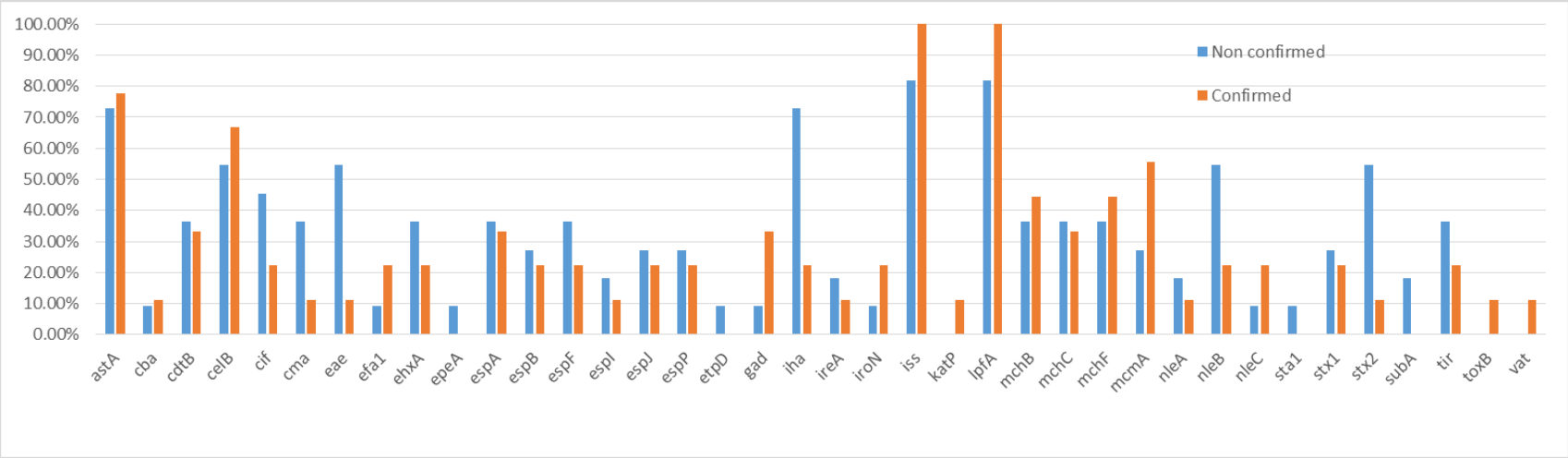


Figure 4. *E. coli* associated virulence genes present in confirmed and non-confirmed samples.

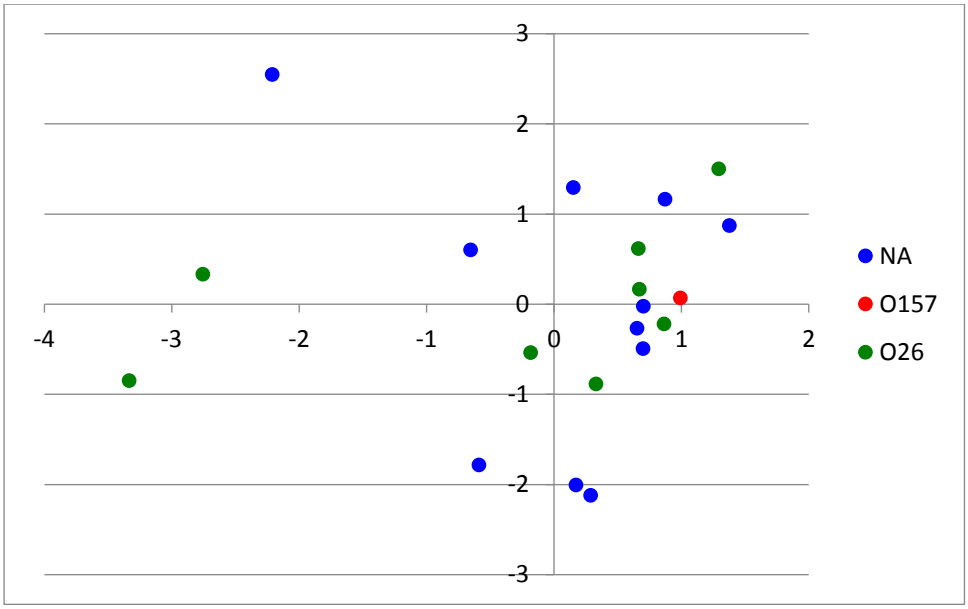


Figure 5. PCA of *E. coli* associated virulence genes for confirmed and non-confirmed samples

The variety of O-antigens present in each broth was assessed by blasting assemblies against a database of *wzx* /wzy and *wzm*/*wzt* gene sequences. In total, 74 different serogroups and subtypes were shown to be present across the 20 samples examined. Up to nine *E. coli* serogroups could be present in some samples and in most cases these were in addition to O157 or Big6 serogroups. A summary of the serogroups determined to be present through bioinformatics analysis and the serogroups identified by the STEC test systems is shown in Table 6. Whilst *wzx*/*wzy* and *wzm*/*wzt* genes have been shown to be highly predictive of O serogroup, a recent study has proposed that 11 combinations of O serogroups should be merged based on immunological and nucleotide sequence similarity, with an additional 10 combinations being virtually indistinguishable at the nucleotide level *(5)*. Further characterisation of the O-antigen clusters identified in this study is required in order to determine the overall variability of O serogroups present in manufacturing beef enrichments. Nonetheless, this study suggests that *E. coli* belonging to O157 or Big6 serogroups are seldom likely to be present in enrichment broths as the only *E. coli* and most often they will be present at concentrations well below many other serogroups. Finally, it may appear as if there is an inability of the bioinformatic analysis to identify those serogroups detected by the STEC test systems. It must be noted that the depth of sequencing achieved in this study is not comparable with PCR and therefore caution is required when directly comparing the results.

Table 6. *E. coli* O serogroups present in enrichment broths as determined bioinformatically and via STEC test systems.

Sample	Serogroups determined bioinformatically	Serogroups determined by STEC test systems
MC11	O45, O91	O26, O45
MC22	O26	O26
MC28	O24, O120, O175	O26, O45
MC30	O8, O21, O93	O26, O45
MC54	O7, O8, O17/77	O26
MC63	No O type present	O26
MC64	O112ab, O116	O26
MC82	O9, O19, O88, O117, O172, O175	O26, O45, O103
MC83	O9, O10, O23, O86, O103	O45, O103
MC8	O8, O9, O86, O157, O185	O26
MC13	O8, O28ac, O86, O156	O26, O45, O103
MC17	O6, O8, O10, O17/44, O17/77, O26, O40, O96, O136	O26, O45
MC32	O8, O11, O45, O75, O88, O93, O96, O103, O141ac	O45, O103
MC38	O150, O153/178, O162	O26, O45, O103
MC42	O38, O103, O175	O45, O103
MC49	O8, O17/44, O17/77, O22, O59, O134	O45, O103
MC56	O9, O25, O48, O175, O182	O45, O103
MC75	O6, O8, O159	O26, O45, O103
MC78	O8, O51, O103	O45, O103
MC79	O6, O8, O22, O89, O117, O146, O148	O26, O45, O103

Mapping of *eae*, *stx*$_1$ and *stx*$_2$

The relative abundance of STEC virulence markers such as *stx* and *eae* can be determined by mapping each metagenomics read to a reference sequence. Virulence markers of greater abundance will have more reads mapped to them than those of lower abundance Phylogenetic typing of STEC based on the presence of *E. coli* associated virulence markers is an approach that facilitates predictive hazard identification of STEC strains. Virulence markers such as *eae* and *stx*, particularly *stx*$_{2a}$ are proposed as key determinants of isolates with the greatest pathogenic potential *(9)*. Similarly, STEC test systems that attempt to link *eae* subtypes to specific O serogroups, such as PALL and NeoSeek, appear to demonstrate enhanced screening/confirmation specificity. Understanding the variability of *eae* and *stx* subtypes in enrichment broths may assist in further developing the algorithms used by STEC systems to classify samples as PP or confirmed. Attempts were made to map sequence reads to the four intimin subtypes known to be associated with O157 and Big6 STEC (Figure 6). The four samples presented in Figure 6 represent two samples that confirmed (MC 22 and MC 64) and two samples that did not culture confirm. In all four samples it is possible to identify the presence of multiple *eae* subtypes which would suggest that discriminating samples based on *eae* subtypes is problematic. However, *eae* subtypes are known to be highly conserved at the 5' end of the gene and least conserved at the 3' end. Therefore read maps that demonstrate even gene coverage across the entire subtype (e.g epsilon in MC 78 and theta in MC13) or which elevate at the 3' end of the gene (e.g beta in MC22 and MC64) are the best predictors for the calling of *eae* subtypes.

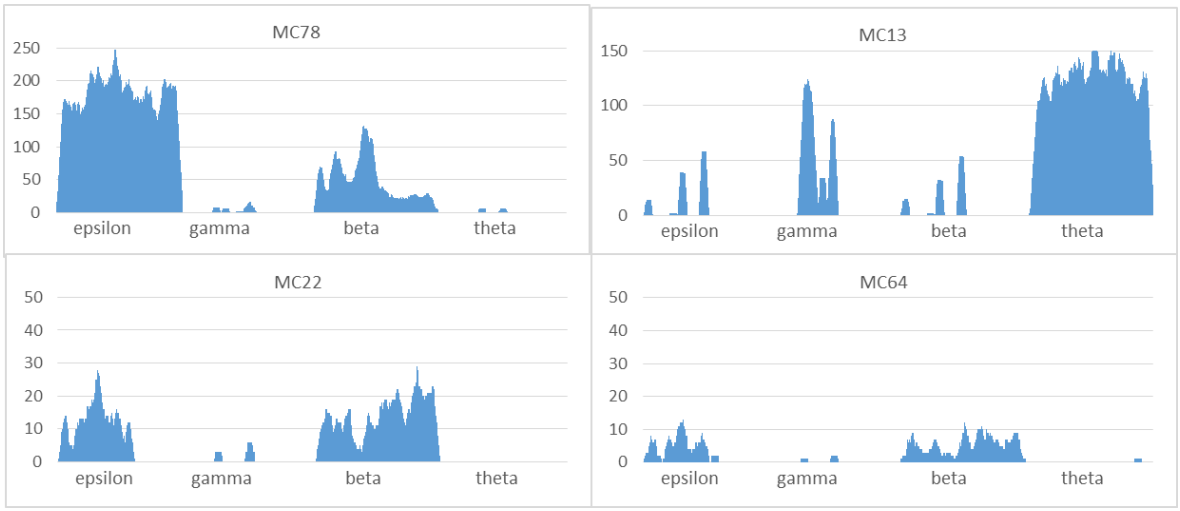


Figure 6. Read mapping of *eae* subtypes using bbmap (values on the *x* axis represent the number of metagenomics reads that align with each of four *eae* subtypes (*y* axis) associated with Big6 STEC).

As previously mentioned, specific *stx* subtypes (e.g stx_{2a}) are known to be more closely associated with severe human clinical disease than others. Mellor and colleagues *(11)* demonstrated that Australian O157 STEC typically do not harbor stx_{2a}, the Shiga toxin most commonly associated with haemolytic uremic syndrome. In addition, Australian Big6 isolates, with the exception of O111 STEC do not harbor stx_{2a}. This is of particular significance to the public health relevance of STEC in Australia as globally there have been shifts in the *stx* profiles of some serogroups (e.g O26) from stx_{1a} to stx_{2a} *(7)*. Read mapping of *stx* genes (Figure 7) confirmed the presence of stx_1, stx_{2a} and stx_{2c} often at high concentrations in many samples. As *stx* phages are highly mobile genetic elements, there remains potential for *stx*-negative *E. coli* and/or STEC to acquire *stx* phage that may increase their predicted pathogenic potential. The results of this study confirm the potential for uptake of phage carrying stx_{2a} exists and ongoing monitoring of *stx* profiles of Australian STEC isolates should be considered a priority.

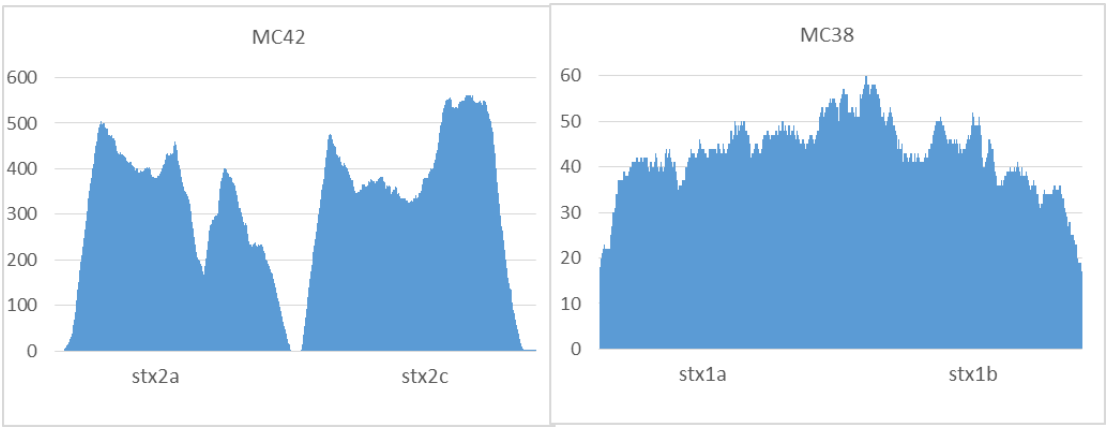


Figure 7. Read mapping of *stx* subtypes using bbmap (values on the *x* axis represent the number of metagenomics reads that align with a variety of *stx* subtypes (*y* axis)).

Conclusions

June 2017 will mark five years since the introduction of testing regulations in the USA for Big6 STEC. Presently, the Australian beef industry generally conducts STEC testing through the application of either BAX or GDS test systems for screening followed by confirmation at an approved commercial testing laboratory. During the last five years there have been significant increases in our understanding of STEC resulting from the application of next generation sequencing strategies. Similarly there have been increases in both the numbers of companies manufacturing STEC screening kits as well as the complexity of testing schemes. This study was designed to evaluate the performance of a range of STEC test systems of varying complexity and to use metagenomics analysis to further understand the complexity of manufacturing enrichment broths and the challenges facing STEC test method development.

This study evaluated a range of STEC test systems that were broadly categorised as traditional, advanced, or confirmatory for their ability to categorise 100 samples as PP or confirmed STEC positives. Whilst there were differences observed between test systems, their overall performance was consistent with their nominated test category with traditional systems producing more PP's than advanced systems. Within test category differences were observed and these may be attributable to the DNA preparation method and PCR efficiency of each test system particularly as quantitative data on *eae* and O-antigen concentrations in enriched broths would suggest that many samples have concentration of these gene targets close to the theoretical limit of detection of PCR (approximately 4.00 $\log_{10}$ copies/mL). It must also be noted that many of the STEC test systems were used on enrichment media for which they are not optimised. Variation in results would be expected if a repeat survey was conducted using different enrichment media.

Advances in our understanding of STEC permits the evolution of STEC screening test systems and directly addresses the desire of industry to reduce the ratio of PP's to confirmed positives. The PALL GeneDisc and Roka test systems utilise additional STEC virulence markers in their test systems and in this study a reduction in PP calls of >50% was observed. The GDS test system employs an immunomagnetic separation step prior to PCR testing for *stx* and *eae* and would therefore be categorised as an advanced method. Similar reduction in PPs to the PALL and ROKA systems were observed with the GDS system. All of the advanced systems generated reductions in PP call rates without impacting on their ability to identify samples that ultimately were confirmed during the

culture confirmation process. The final test system assessed in this study was the NeoSeek STEC. This test system is a confirmatory test system that uses PCR coupled with mass spectrometry to assess the presence of more than 86 specific genetic markers and subsequently compares the PCR profile to the known genetic makeup of reference STEC strains. The use of the NeoSeek STEC system in this study resulted in a >80% reduction in Big6 positives whilst remaining consistent with all other test systems in their ability to detect the presence of STEC in culturally confirmed samples.

Metagenomic assessment of manufacturing beef enrichment broths coupled with quantitative PCR data reinforces the challenge facing those who conduct STEC screening or confirmation. This study demonstrated that the diversity of microflora present in enrichment broths is relatively high and is often dominated by non-*E. coli* organisms such as *C. perfringens*. Furthermore, *E. coli* often comprises less than 10% of the enrichment broth and regularly the target serogroups are not the dominant serogroups in each broth. The outcome of these findings is that confirmation of STEC is routinely attempting to isolate organisms that are present in ratios of 1:1000 or lower. Nonetheless, all indications from this study and those previously funded by MLA and CSIRO (e.g. G.MFS.0282) confirm that there is high likelihood of isolation from enrichment broths and consequently false negatives remain low. The ratio of PPs to confirmed positives remains high regardless of sample source or cattle type. Comparison of STEC testing systems suggests that approaches that use additional or alternative targets can reduce the number of PPs without affecting the performance of the STEC testing program. Attempts were made to analyse the metagenomic reads for the presence of genetic signatures that are more likely to be present in enrichment broths that confirm than those that do not. A variety of *E. coli*-associated virulence markers were identified, however there were very few differences between broths. It is the case though that the depth of sequencing in this study was insufficient to capture the genomic contribution of the actual STEC isolate and it is possible that the results presented for both confirmed and non-confirmed samples represent typical background genetic material. Further advances in sequencing technology and analysis of increased numbers of enrichment broths may permit greater resolution of the genetic content of manufacturing beef enrichment broths.

The evolution of STEC test systems will continue and refinements to screening and isolation methods will aid the Australian beef industry in maintaining access to export markets with STEC regulations. The Australian industry typically uses BAX or GDS test systems and in this study they have both

performed in equivalent ways to comparable test systems. One of the concerns that industry continues to discuss is the ratio of PPs to confirmed samples. This study has determined that STEC test systems that use additional or alternative test systems may substantially reduce the number of PP samples requiring confirmation. The use of such systems could reduce testing costs but more importantly would substantially reduce the costs of holding product prior to export. Culture confirmation remains an integral part of the STEC testing program in Australia and PCR-based confirmation methods are yet to be widely adopted for STEC testing. This study did include a PCR-based confirmation test system (NeoSeek) and its overall performance was comparable, though not equal to culture confirmation. The results do provide encouragement for the industry to continue to explore the implementation of novel STEC test systems such as NeoSeek as a means of reducing the costs and timeframes associated with the STEC testing program.

Acknowledgements

The authors acknowledge the support of the Australian beef industry in providing samples for the study and Symbio Alliance for coordination sample collection and provision of samples to CSIRO. Many thanks go to all of the STEC testing system manufacturers that voluntarily participated in the study and in particular PALL (via Patrick Fach at ANSES), Neogen and Roka who conducted the tests in their laboratories.

References

1. Beutin, L., and P. Fach. 2014. Detection of Shiga Toxin-Producing *Escherichia coli* from Nonhuman Sources and Strain Typing. *Microbiol Spectr*. 2.

2. Bolger, A. M., M. Lohse, and B. Usadel. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 30:2114-20.

3. Brooks, J. T., E. G. Sowers, J. G. Wells, K. D. Greene, P. M. Griffin, R. M. Hoekstra, and N. A. Strockbine. 2005. Non-O157 Shiga toxin-producing *Escherichia coli* infections in the United States, 1983-2002. *J Infect Dis*. 192:1422-9.

4. Chopyk, J., R. M. Moore, Z. DiSpirito, Z. R. Stromberg, G. L. Lewis, D. G. Renter, N. Cernicchiaro, R. A. Moxley, and K. E. Wommack. 2016. Presence of pathogenic *Escherichia coli* is correlated with bacterial community diversity and composition on pre-harvest cattle hides. *Microbiome*. 4:9.

5. DebRoy, C., P. M. Fratamico, X. Yan, G. Baranzoni, Y. Liu, D. S. Needleman, R. Tebbs, C. D. O'Connell, A. Allred, M. Swimley, M. Mwangi, V. Kapur, J. A. Raygoza Garay, E. L. Roberts, and R. Katani. 2016. Comparison of O-Antigen Gene Clusters of All O-Serogroups of *Escherichia coli* and Proposal for Adopting a New Nomenclature for O-Typing. *PLoS One*. 11:e0147434.

6. Delannoy, S., B. D. Chaves, S. A. Ison, H. E. Webb, L. Beutin, J. Delaval, I. Billet, and P. Fach. 2016. Revisiting the STEC Testing Approach: Using espK and espV to Make Enterohemorrhagic *Escherichia coli* (EHEC) Detection More Reliable in Beef. *Front Microbiol*. 7:1.

7. Delannoy, S., P. Mariani-Kurkdjian, S. Bonacorsi, S. Liguori, and P. Fach. 2015. Characteristics of emerging human-pathogenic *Escherichia coli* O26:H11 strains isolated in France between 2010 and 2013 and carrying the stx2d gene only. *J Clin Microbiol*. 53:486-92.

8. Edgar, R. C. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*. 26:2460-1.

9. Franz, E., A. H. van Hoek, M. Wuite, F. J. van der Wal, A. G. de Boer, E. I. Bouw, and H. J. Aarts. 2015. Molecular hazard identification of non-O157 Shiga toxin-producing *Escherichia coli* (STEC). *PLoS One*. 10:e0120353.

10. Mellor, G. E., N. Fegan, L. L. Duffy, K. E. McMillan, D. Jordan, and R. S. Barlow. 2016. National survey of Shiga toxin-producing *Escherichia coli* serotypes O26, O45, O103, O111, O121, O145 and O157 in Australian beef cattle feces. *Journal of Food Protection*. 79:1868-1874.

11. Mellor, G. E., N. Fegan, K. S. Gobius, H. V. Smith, A. V. Jennison, B. A. D'Astek, M. Rivas, S. Shringi, K. N. Baker, and T. E. Besser. 2015. Geographically distinct *Escherichia coli* O157 isolates differ by lineage, Shiga toxin genotype, and total shiga toxin production. *J Clin Microbiol*. 53:579-86.

12. Peng, Y., H. C. Leung, S. M. Yiu, and F. Y. Chin. 2012. IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics*. 28:1420-8.
13. Segata, N., L. Waldron, A. Ballarini, V. Narasimhan, O. Jousson, and C. Huttenhower. 2012. Metagenomic microbial community profiling using unique clade-specific marker genes. *Nat Methods*. 9:811-4.
14. Wu, Y. W., B. A. Simmons, and S. W. Singer. 2016. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics*. 32:605-7.

Appendix 1

Table A1: Summary table of all STEC test system results for 100 manufacturing beef enrichment broths

MC#	S_Bax	S_GDS	C_Bax	RapidFinder	Qiagen	Biotecon	FSIS	Roka	PALL	NeoSeek	Confirmed
1	1*	1	0	0	1	0	1	0	0	0	0
2	1	0	0	0	0	0	1	0	0	0	0
3	1	1	1	1	1	1	1	1	1	0	0
4	1	0	0	1	1	1	1	0	1	0	0
5	1	0	0	0	1	0	1	0	0	0	0
6	0	1	0	0	0	0	0	0	0	0	0
7	1	1	1	1	1	1	1	0	0	0	0
8	1	1	1	1	1	1	1	0	1	0	0
9	1	1	1	1	1	1	1	1	0	0	0
10	1	1	1	0	1	1	1	0	0	0	0
11	1	1	1	1	1	1	1	1	1	1	O26
13	1	1	1	1	1	1	1	0	0	0	0
14	1	1	0	0	1	0	0	0	0	0	0
15	1	0	0	0	1	1	0	1	1	0	0

MC#	S_Bax	S_GDS	C_Bax	RapidFinder	Qiagen	Biotecon	FSIS	Roka	PALL	NeoSeek	Confirmed
16	1	0	1	1	1	1	1	0	0	0	0
17	1	0	1	1	1	1	1	0	1	0	0
18	1	0	1	1	1	1	1	0	1	0	0
19	1	0	1	1	1	1	1	0	1	0	0
20	1	1	1	1	0	1	1	1	1	1	0
21	1	1	0	0	1	0	0	1	1	0	0
22	1	1	1	1	1	1	1	1	1	1	O26
23	1	1	1	1	1	1	1	0	1	0	0
24	1	0	0	0	0	0	0	1	0	0	0
25	0	1	1	1	1	1	1	0	0	0	0
26	0	1	0	0	1	0	1	0	0	0	0
27	0	1	0	0	0	0	0	1	1	1	O157
28	1	0	0	0	1	0	1	1	1	0	O26
29	1	1	1	1	1	1	1	0	0	0	0
30	1	1	1	1	1	1	1	1	1	1	O26
31	1	1	1	1	1	1	1	1	1	0	0

MC#	S_Bax	S_GDS	C_Bax	RapidFinder	Qiagen	Biotecon	FSIS	Roka	PALL	NeoSeek	Confirmed
32	1	0	1	1	1	1	1	1	0	0	0
33	1	0	0	0	0	0	1	0	0	0	0
34	1	1	1	1	1	1	1	1	1	1	O157
37	1	0	1	0	1	1	1	1	0	0	0
38	1	0	1	1	1	0	1	0	1	0	0
39	1	0	1	1	1	1	1	0	1	0	0
40	0	1	0	0	0	0	1	0	0	0	0
41	1	0	0	0	0	0	0	0	0	0	0
42	1	0	1	1	1	1	1	1	1	0	0
43	1	0	1	1	1	1	1	0	0	0	0
44	1	0	0	0	0	0	0	0	0	0	0
45	1	0	0	0	0	0	1	1	0	0	0
47	1	1	1	1	1	1	1	1	0	0	0
48	1	1	1	1	1	1	1	1	0	0	0
49	1	0	1	1	1	1	1	1	0	0	0
50	1	0	1	1	1	1	1	1	0	0	0

MC#	S_Bax	S_GDS	C_Bax	RapidFinder	Qiagen	Biotecon	FSIS	Roka	PALL	NeoSeek	Confirmed
51	1	0	0	0	0	0	1	0	0	0	0
52	1	0	0	0	1	0	1	0	0	0	0
53	1	0	0	0	0	0	1	0	0	0	0
54	1	1	1	1	1	1	1	1	1	1	O26
55	1	1	1	1	1	1	1	0	1	0	0
56	1	0	1	1	1	1	1	1	0	0	0
57	1	0	0	0	1	0	0	0	0	0	0
59	0	1	0	0	1	0	1	0	1	0	0
60	0	1	1	1	1	1	1	0	0	0	0
61	0	1	0	0	1	0	1	0	0	0	0
62	1	1	1	1	1	1	1	0	0	0	0
63	1	1	1	1	1	1	1	1	1	1	O26
64	1	1	1	1	1	1	1	1	1	1	O26
65	1	1	1	1	1	1	1	1	1	1	O157
66	1	1	1	1	1	0	1	1	1	1	0
67	1	1	0	0	1	0	0	1	0	0	0

MC#	S_Bax	S_GDS	C_Bax	RapidFinder	Qiagen	Biotecon	FSIS	Roka	PALL	NeoSeek	Confirmed
68	1	0	1	1	1	1	1	0	0	0	0
69	1	0	1	1	1	0	1	0	0	0	0
70	1	1	1	1	1	1	1	1	1	1	O157
71	1	0	0	0	0	0	1	0	0	0	0
72	1	0	1	0	1	1	1	0	0	0	0
73	1	0	1	1	1	1	1	0	0	0	0
74	1	0	1	1	1	0	1	1	1	0	0
75	0	1	1	1	1	1	1	0	0	0	0
76	1	1	0	0	0	0	0	0	0	0	0
78	1	0	1	1	1	1	1	1	1	0	0
79	1	1	1	1	1	1	1	0	1	0	0
80	1	1	1	1	1	1	1	0	1	0	0
81	1	0	0	0	0	0	1	0	0	0	0
82	1	1	1	1	1	1	1	1	1	1	O26
83	1	1	1	1	1	1	1	1	1	1	O157
84	1	0	1	1	1	1	1	1	1	1	O157

MC#	S_Bax	S_GDS	C_Bax	RapidFinder	Qiagen	Biotecon	FSIS	Roka	PALL	NeoSeek	Confirmed
86	1	1	1	1	1	1	1	0	1	0	0
87	1	1	1	1	1	1	1	0	0	0	0
88	1	1	1	1	1	1	1	1	1	1	0
89	1	1	1	1	1	1	1	1	1	1	O26
90	0	1	0	0	1	0	0	0	0	0	0
91	1	1	1	0	1	1	1	1	1	1	0
92	1	0	0	1	0	0	1	0	0	0	0
93	1	1	1	1	1	1	1	0	1	0	0
94	1	0	0	0	0	0	1	0	0	0	0
95	1	1	1	1	1	1	1	1	1	1	O26
96	1	0	1	1	1	1	1	0	0	0	0
97	1	1	0	0	1	0	1	0	0	0	0
98	1	0	1	1	1	0	1	0	NT	0	0
99	1	1	1	0	1	1	1	0	NT	0	0
100	1	0	0	0	0	0	0	0	NT	0	0
101	1	1	1	1	1	1	1	1	NT	1	O26

MC#	S_Bax	S_GDS	C_Bax	RapidFinder	Qiagen	Biotecon	FSIS	Roka	PALL	NeoSeek	Confirmed
102	1	0	1	1	1	1	0	0	NT	1	O26
103	1	1	1	1	1	1	1	0	NT	0	0
104	1	1	1	1	1	1	1	0	1	0	0
105	1	1	1	1	1	1	1	0	0	0	0
106	1	0	0	0	1	0	0	0	0	0	0
107	1	0	1	1	1	1	1	0	0	0	0

*A score of 0 indicates a negative result; A score of 1 indicates that the sample was PP or confirmed.

Appendix2

	MC 11	MC 22	MC 28	MC 30	MC 54	MC 63	MC 64	M 82	MC 83	MC 8	MC 13	MC 17	MC 32	MC 38	MC 42	MC 49	MC 56	MC 75	MC 78	MC 79
s__Clostridium_perfringens	0.0	82.9	40.8	16.9	72.7	0.0	77.3	84.8	3.0	2.2	1.1	3.8	34.7	23.8	2.4	26.8	86.5	65.6	88.4	14.1
s__Escherichia_coli	9.6	2.1	7.4	7.2	10.7	1.2	8.1	9.3	42.9	68.2	38.7	11.4	27.3	9.6	56.3	33.4	8.9	4.0	5.4	32.4
s__Enterococcus_faecalis	58.5	1.9	0.1	17.9	1.2	0.3	0.0	0.9	17.2	5.5	1.7	5.3	9.2	3.3	3.2	7.7	0.7	0.7	0.2	1.0
s__Clostridium_bifermentans	2.3	0.0	11.4	0.0	12.5	75.8	0.0	1.1	0.0	2.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	19.3	6.0	0.0
s__Streptococcus_infantarius	11.7	7.9	1.0	0.0	0.0	1.3	0.6	0.0	15.7	4.7	41.5	0.1	0.1	0.2	1.6	2.7	0.0	7.0	0.1	20.3
s__Escherichia_unclassified	4.9	1.4	7.2	0.0	2.1	1.5	5.0	3.9	12.2	12.4	3.0	3.6	7.7	2.2	9.5	8.8	3.4	2.3	0.0	9.5
s__Lactococcus_garvieae	0.0	0.0	20.2	24.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	33.7	0.3	12.0	0.0	0.7	0.0	0.0	0.0	0.0
s__Enterococcus_hirae	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.5	0.0	0.6	2.9	0.0	33.6	0.0	0.2	0.0	0.0	0.0	0.5
s__Morganella_morganii	0.1	0.0	0.0	1.1	0.0	0.0	0.0	0.0	1.3	0.0	6.3	11.4	1.8	0.6	0.0	0.4	0.0	0.0	0.0	0.0
s__Lactococcus_lactis	0.6	0.0	0.1	10.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	8.1	0.0	1.7	0.0	0.1	0.0	0.0	0.0	0.0
s__Enterococcus_cecorum	0.0	0.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	20.5	0.0	0.0	0.0	0.0	0.0
s__Streptococcus_salivarius	0.0	0.0	0.0	0.0	0.0	17.4	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.0	0.1	0.0	0.2	0.0	0.0	0.0
s__Shigella_phage_Shfl2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	17.3
s__Acinetobacter_baumannii	0.0	0.0	0.1	3.3	0.0	0.0	0.0	0.0	0.8	0.0	0.0	1.6	0.0	1.5	0.0	6.2	0.0	0.0	0.0	0.0
s__Citrobacter_freundii	0.5	0.0	0.0	0.8	0.0	0.0	0.0	0.0	0.3	0.0	0.0	0.4	5.9	0.4	0.1	3.5	0.0	0.0	0.0	0.0
s__Acinetobacter_pittii	0.0	0.0	0.2	2.8	0.0	0.0	0.0	0.0	0.9	0.0	0.0	4.3	0.0	1.1	0.0	1.5	0.0	0.0	0.0	0.0
s__Citrobacter_unclassified	0.3	0.0	0.0	0.4	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.3	8.6	0.1	0.1	0.7	0.0	0.0	0.0	0.0
s__Macrococcus_caseolyticus	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.7	0.0	0.0	4.4	0.0	2.6	0.2	0.4	0.2	0.0	0.0	0.0

	MC 11	MC 22	MC 28	MC 30	MC 54	MC 63	MC 64	M 82	MC 83	MC 8	MC 13	MC 17	MC 32	MC 38	MC 42	MC 49	MC 56	MC 75	MC 78	MC 79
s__Enterococcus_faecium	0.7	0.0	2.5	0.0	0.0	1.6	0.0	0.0	0.0	0.0	0.0	1.8	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.3
s__Fusobacterium_necrophorum	0.0	0.0	0.0	0.0	0.0	0.0	7.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.4	0.0	0.0	0.0	0.0	0.0
s__Pseudomonas	0.0	0.0	0.0	8.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
s__Enterobacter_cloacae	0.6	0.0	4.5	0.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.2	0.0	0.2	0.0	0.0	0.0	0.1	0.0	0.0
s__Aeromonas_hydrophila	1.0	0.0	3.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
s__Aeromonas_unclassified	2.6	0.0	0.5	0.7	0.0	0.0	0.0	0.0	0.0	0.0	2.9	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0
s__Clostridium_phage	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.9	0.0	0.0	0.0	0.0	0.0	0.0
s__Bacillus_cereus	0.3	0.0	0.7	0.0	0.4	0.8	0.0	0.0	0.6	1.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.4	0.0	0.0
s__Klebsiella_pneumoniae	1.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
s__Klebsiella_unclassified	0.0	0.0	0.4	1.3	0.4	0.0	0.0	0.0	0.0	0.0	0.2	0.3	0.2	0.2	0.0	1.1	0.1	0.0	0.0	0.0
s__Klebsiella_oxytoca	1.2	0.0	0.1	1.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.1	0.1	0.0	1.2	0.0	0.0	0.0	0.0
s__Streptococcus_uberis	0.1	0.0	0.0	1.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.2	0.1	0.2	1.1	0.0	0.0	0.0	0.0	0.0
s__Streptococcus_lutetiensis	0.8	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.4	0.0	1.1	0.0	0.0	0.0	0.1	0.1	0.0	0.6	0.0	0.3
s__Streptococcus_pasteurianus	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	3.2	0.0	0.0	0.0	0.0	0.0
s__Streptococcus_gallolyticus	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.3
s__Wohlfahrtiimonas_chitiniclastica	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
s__Streptococcus_parasanguinis	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.9	0.6	0.0	0.0	0.0	0.0	0.0	0.0
s__Streptococcus_agalactiae	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.7	0.0	0.0	0.0	0.0
s__Clostridium_novyi	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.8	0.0	0.0	0.0	0.9

	MC 11	MC 22	MC 28	MC 30	MC 54	MC 63	MC 64	M 82	MC 83	MC 8	MC 13	MC 17	MC 32	MC 38	MC 42	MC 49	MC 56	MC 75	MC 78	MC 79
s__Proteus_penneri	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.8	0.0	0.0	0.0	0.0
s__Salmonella_phage	0.0	1.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
s__Escherichia_fergusonii	0.0	1.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
s__Hafnia_alvei	0.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
s__Aeromonas_veronii	0.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
s__Streptococcus_parauberis	0.0	0.0	0.0	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0
s__Vagococcus_lutrae	0.2	0.0	0.0	0.0	0.0	0.0	0.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0
s__Raoultella_ornithinolytica	0.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
s__Acinetobacter_guillouiae	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.7	0.0	0.0	0.0	0.0
s__Bacteroides_pyogenes	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0
s__Peptostreptococcaceae	0.0	0.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
s__Veillonella_atypica	0.0	0.0	0.0	0.0	0.0	0.0	0.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0
s__Proteus_unclassified	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.3	0.0	0.0	0.0	0.0
s__Coprobacillus_unclassified	0.0	0.0	0.0	0.0	0.0	0.0	0.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
s__Enterococcus_phoeniculicola	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
s__Enterobacteriaceae_bacterium	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
s__Streptococcus_mitis_oralis_pneumoniae	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
s__Lactobacillus_mucosae	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
s__Enterococcus_durans	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

	MC 11	MC 22	MC 28	MC 30	MC 54	MC 63	MC 64	M 82	MC 83	MC 8	MC 13	MC 17	MC 32	MC 38	MC 42	MC 49	MC 56	MC 75	MC 78	MC 79
s__Comamonas_unclassified	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
s__Veillonella_unclassified	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
s__Pseudomonas_fragi	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
s__Aeromonas_salmonicida	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
s__Aeromonas_media	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Appendix 3

Table A3. List of *E. coli* associated virulence genes

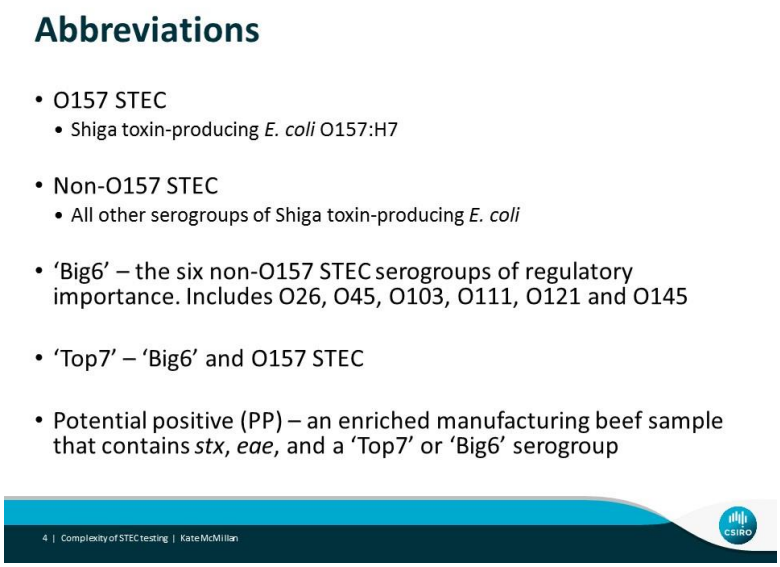
astA	*fim41a*	*sfaS*	*nleE*
bfpA	*gad*	*sigA*	*nleF*
cba	*hlyE*	*sta1*	*nleG2-1*
ccI	*iha*	*stb*	*nleG5-2*
cdtB	*ipaD*	*stx1A*	*nleG9*
celb	*ipaH9.8*	*stx1B*	*nleH1-1*
cfa_c	*ireA*	*stx2A*	*nleH1-2*
cif	*iroN*	*stx2B*	*terB*
cma	*iss*	*subA*	*TspE4.C2*
cnf1	*K88ab*	*saa*	*ureC*
cofA	*katP*	*tccP*	*yjaA*
eae	*lngA*	*tir*	
eatA	*lpfA*	*toxB*	
efa1	*ltcA*	*tsh*	
ehxA	*mchB*	*vat*	
epeA	*mchC*	*virF*	
espA	*mchF*	*pagC*	
espB	*mcmA*	*terC*	
espC	*nfaE*	*aid-1*	
espF	*nleA*	*nleG2-3*	
espI	*nleB*	*nleG6-2*	
espJ	*nleC*	*nleG5-2*	
espP	*perA*	*irp2*	
etpD	*pet*	*fyuA*	
f17A	*pic*	*adfO*	
f17G	*prfB*	*chuA*	
fanA	*rpeA*	*ckf*	
fasA	*sat*	*ent/espL2*	
fedA	*senB*	*nleB2*	
fedF	*sepA*	*nleD*	

Industry workshop – Shiga toxin-producing *Escherichia coli* in manufacturing beef: Where have we been? Where should we be going?

The complexity of STEC testing – Kate McMillan

Background

The majority of *E. coli* that humans or animals carry are harmless, however some carry genes that enable them to cause disease. *E. coli* that produce Shiga toxins (stx) are termed Shiga toxigenic *E. coli* (STEC). Some strains of STEC appear to have greater potential to cause human disease than others. This subset includes STEC belonging to certain serogroups (e.g. O157, O26, O111) and have additional virulence mechanisms (e.g. *E. coli* attaching and effacing gene; eae). In 2012, FSIS expanded its regulations from just testing for O157 to include an additional six serogroups, O26, O45, O103, O111, O121 and O145 which are colloquially known as the 'Big6' or non-O157 STEC. Companies exporting beef for grinding to countries with STEC regulations maybe required to conduct pre-export testing for STEC.



Australian perspective

There are many STEC test systems commercially available. The Australian beef industry typically uses two systems:

- BAX system real-time PCR STEC Suite (Hygiena)
- Assurance GDS MPX STEC assays (BioControl)

Samples that test positive using these systems are classified as 'potential positives' (PP) and subsequently proceed for culture confirmation at a Department of Agriculture and Water Resources approved laboratory.

In Australia, samples that are PP for O157 are more often confirmed than samples that are PP for non-O157 STEC.

STEC testing is now more complex

Prior to 2012 – testing for O157 only was fairly straight forward

- O157 does not ferment sorbitol so easy to identify on plates
- Most O157 strains are likely to have *stx* and *eae*
- Only looking for one serogroup – easier to detect, isolate and confirm

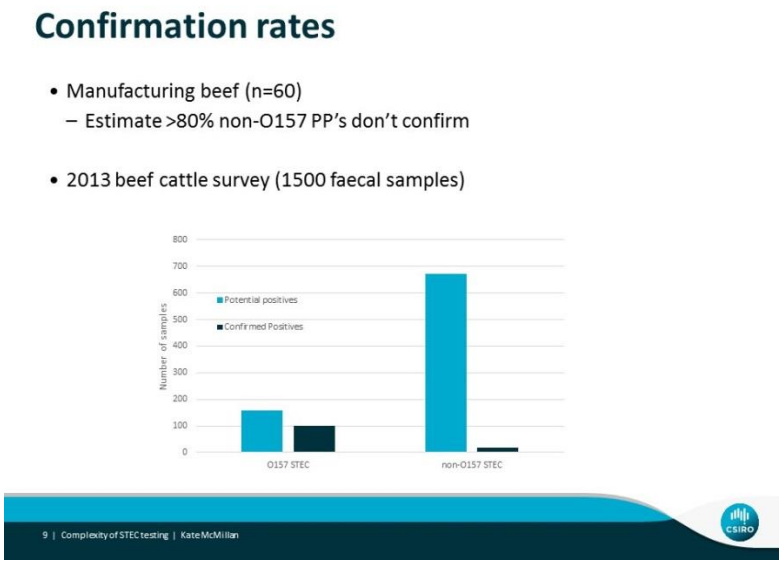
Post 2012 – testing for O157 and non-O157 STEC

- Non-O157 have no distinguishing features to exploit e.g sorbitol
- Now looking for multiple serogroups not just O157
- Not all strains have *stx* and *eae*
- Very hard to distinguish from harmless *E. coli* during culture confirmation

Low confirmation rates

Most STEC testing protocols look for *stx*, *eae* and O serogroups. A positive screening test therefore only indicates that these genetic targets are present in the sample, it can't tell us if they are in the same cell or if that cell is an *E. coli*.

A survey of STEC in Australian cattle faeces conducted in 2013 had a low conversion rate of PP to confirmed positives[2]. Of the 1,500 samples tested, 44.5% were PP for non-O157 STEC but only 1.3% were culture confirmed as non-O157 STEC.



Samples that don't culture confirm

Low conversion rates of PP's to confirmed positives can give rise to the following questions:

[2] https://www.mla.com.au/research-and-development/search-rd-reports/final-report-details/Product-Integrity/Understanding-confirmation-test-failures-for-detecting-pathogenic-E-coli/1167

- Was there an error with the screening test (i.e. a false positive) or
- Did the confirming lab miss the STEC?

The answer to at least the first question is most likely to be no.

PP samples likely contain a variety of *E. coli* that in combination carry *stx*, *eae* and belong to one of the targeted serogroups. Therefore, a PP that does not confirm positive is not necessarily a 'false' positive as the test correctly identified the presence of the right combination of targets. Table 1 shows the variety of *E. coli* possessing STEC markers associated with samples that were PP but did not confirm.

Table 1. STEC virulence marker combinations in *E. coli* recovered from potential positive manufacturing beef enrichment broths (of the broths tested none confirmed positive for the targeted STEC).

***E. coli* isolates with STEC markers**	**Prevalence (n=93)**
STEC (*eae* and *stx*)	0 (0.0%)
stx only	40 (43.0%)
eae only	26 (28.0%)
stx & non-O157 serogroup	2 (2.2%)
eae & non-O157 serogroup	19 (20.4%)
Non-O157 serogroup only	26 (28.0%)

Conclusions

- STEC screening systems detect genetic markers to identify potential positive samples. They do not tell us if the genetic markers are in the same *E. coli.*
- Conversion rates of PP's to confirmed positives are low for non-O157 PP's as most often the genetic targets identified by the STEC screening systems are present in different isolates of *E. coli.*
- Culture confirmation of non-O157 STEC is a laborious lengthy procedure as it attempts to identify a small group of *E. coli* that appear similar to harmless *E. coli*.

Comparison of STEC detection systems – Robert Barlow

Background

Australian beef exporters have been conducting pre-export testing of manufacturing beef lots destined to the US since the expansion of the STEC regulations in June, 2012. In general, Australian exporters use one of two test systems (BAX or GDS) to initially screen lots for the presence of STEC, with screen positives being subsequently culture confirmed at a Department of Agriculture and Water Resources laboratory. This approach has served the Australian beef industry well and assists in maintaining access into markets, such as the USA, that have regulations relating to the presence of STEC in beef destined for grinding. Our understanding of STEC is increasing due to advances in analytical technologies (genomics). From a STEC testing perspective this has supported the development of more sensitive and specific testing systems. Some of these systems employ

detection strategies identifying the three markers commonly used to define STEC (i.e. *stx*, *eae* and O serogroup) whereas other systems are using additional or alternative markers to enhance the specificity of the test system in an attempt to reduce the numbers of PP's that are sent for culture confirmation. Additionally, there are STEC test systems that remove the need for culture confirmation completely by assaying a sample for large numbers of genetic targets that are then aligned with known STEC profiles. Assessing the performance of these systems in an Australian context will enable the effectiveness of currently used systems to be determined and may identify those systems that can reduce the number of PP's without compromising the ability to identify positive lots.

What does a good test look like?

STEC test systems can be broadly categorised as classical (few targets), advanced (more targets) or confirmatory (lots of targets). Test systems with most value to the industry are those that are able to reduce the number of PP samples while still able to identify samples that actually contain STEC.

Study design

- 100 manufacturing beef enrichment samples that were PP for non-O157 STEC
- Tested using the following STEC screening systems:
    - BAX system real-time PCR STEC suite (Hygiena)
    - RapidFinder STEC (ThermoFisher Scientific)
    - Mericon *E. coli* STEC O-type (Qiagen)
    - Foodproof STEC Lyokit (Biotecon Diagnostics)
    - Non-O157 STEC from meat products (FSIS)
    - Assurance GDS MPX (BioControl)
    - Atlas STEC EG2 combo detection assay (Roka Bioscience)
    - GeneDisc system (PALL)
- Tested using the following STEC confirmation system
    - NeoSeek STEC (Neogen)
- Performance measured by:
    - Ability to detect samples that were culture confirmed, and
    - Total number of PP's

Manufacturer	Test type	Name
Hygiena	Classical	BAX System Real-Time PCR STEC Suite.
Thermo Fisher Scientific	Classical	RapidFinder STEC
Qiagen	Classical	Mericon E. coli STEC O-Type
Biotecon Diagnostics	Classical	Foodproof STEC LyoKit
FSIS*	Classical	Detection and Isolation of non-O157 Shiga Toxin-Producing Escherichia coli (STEC) from Meat Products and Carcass and Environmental Sponges
BioControl	Advanced	Assurance GDS MPX for Top 6 or 7 STEC
Roka Bioscience	Advanced	Atlas STEC EG2 Combo Detection Assay
PALL	Advanced	GeneDisc System
Neogen	Confirmatory	NeoSeek STEC

stx
eae
O serogroup

Classical with additions

* Primers & probes – MLG 5B, Appendix 1.01

5 | STEC Methods | Robert Barlow

CSIRO

Results

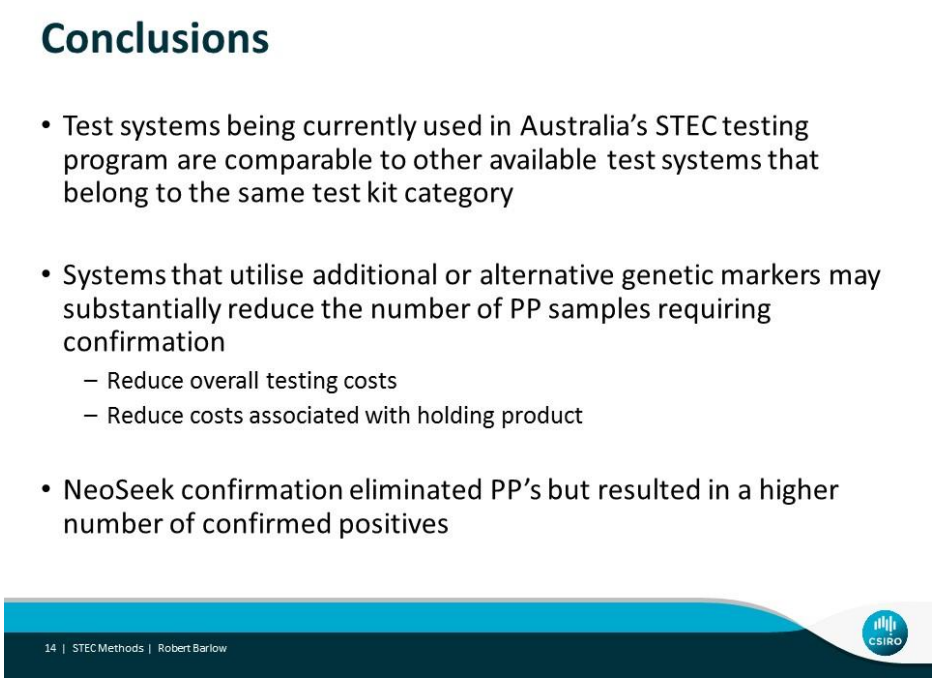
- 100 non-O157 PP samples collected between July 2016 and January 2017 – 61 generated by BAX and 39 by GDS
- 12 samples culture confirmed as O26
- The majority of STEC screening systems detected 11 of the 12 culture confirmed samples, the exception being GDS which detected 10 of the 12 confirmed samples. The Qiagen and PALL systems detected all confirmed samples for which they generated a test result (Table 1). All systems detected 10 of the 12 culture confirmed samples with variable results obtained for the remaining two culture confirmed samples.
- The advanced test systems of Roka, GDS, and PALL target additional or alternative genetic markers during screening. The use of these systems reduced the number of PP's without affecting the ability to detect culture confirmed samples (Table 1).
- This study used enrichments broths recommended by the GDS or BAX test systems. When comparing performance of test systems in this study it is necessary to consider:
    - Recommended enrichment media were not used for all tests
    - Recommended enrichment protocols were not used for all tests
    - Enrichment broths may change over time affecting what can be detected
- NeoSeek STEC was the only non-culture confirmation method evaluated. Using NeoSeek 16 samples were identified as positive for non-O157 STEC, this included 11 of the 12 culture confirmed samples.

Table 1. Detection of culture confirmed positives and overall positives by STEC test systems.

Test system	Test category	Non-O157 confirmed positives detected	Positives
FSIS	Classical	11/12	85
QIAGEN	Classical	12/12	82
BAX	Classical	11/12	67
RAPIDFINDER	Classical	11/12	64
BIOTECON	Classical	11/12	64
GDS	Advanced	10/12	56
PALL	Advanced	10/10*	42/94*
ROKA	Advanced	11/12	39
NEOSEEK	Confirmatory	11/12	16

* A software malfunction resulted in no result being generated for six samples, two of which culture confirmed.

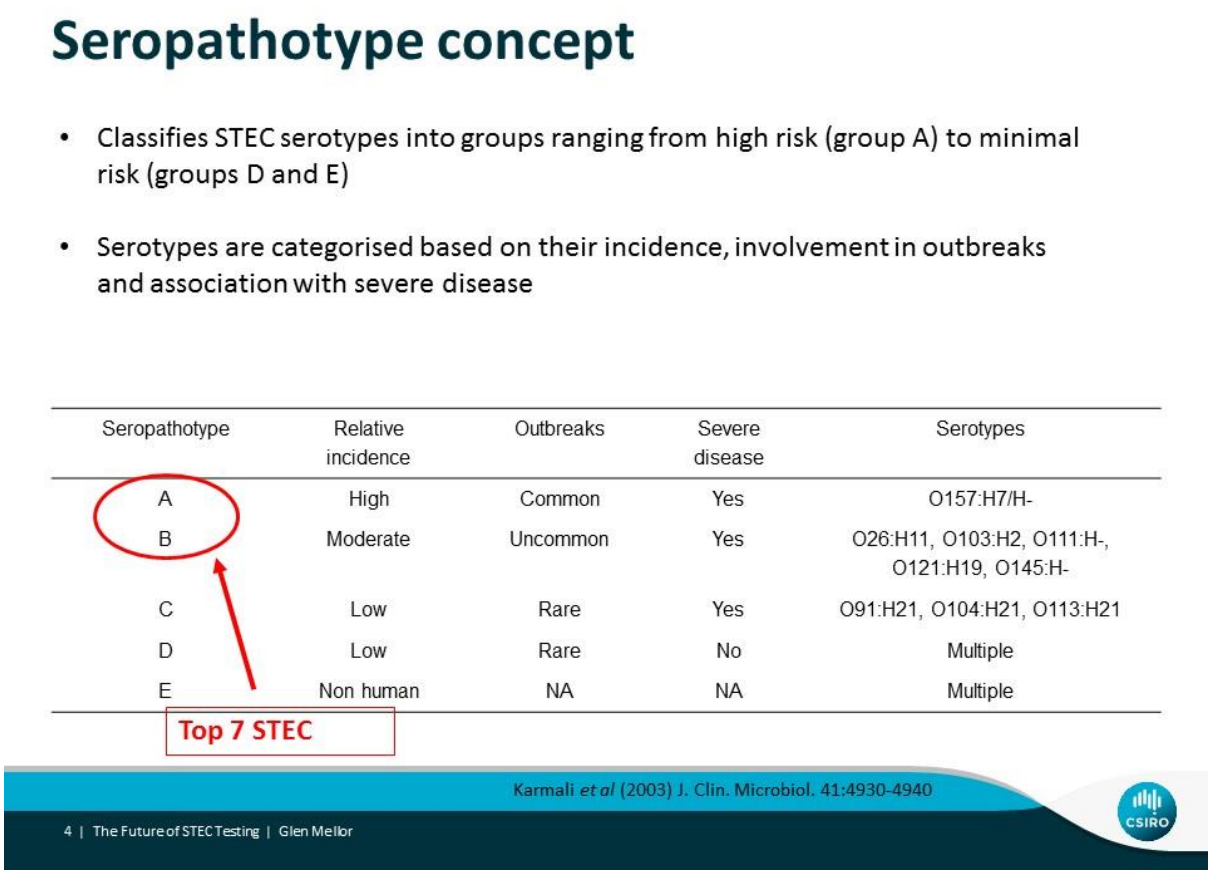
Conclusions



The future of STEC testing – Glen Mellor

Current concept

The addition of the non-O157 serogroups to the STEC testing program in 2012 was a response to human illness data that demonstrated that these serogroups were responsible for the majority of non-O157 STEC related disease. Human illness data from the USA in 2013 supported the regulatory response with 48.5% of STEC-associated illness attributable to O157 and 44.6% attributable to the non-O157 serogroups. Identification of the specific serogroups for inclusion in the STEC testing program followed on from an earlier classification concept known as the seropathotype concept, where serogroups are categorised based on their incidence, involvement in outbreaks and association with disease. The current STEC regulations assume that all STEC belonging to a particular serogroup have the same disease causing potential. However, there is evidence to suggest that within serogroups STEC may have differing ability to cause severe human disease.



Not all STEC are equal

The advent of genomic sequencing is enabling relationships between STEC to be further understood. For example, by analysing small variations in the genetic composition of O157 isolates they can be grouped into very specific groups or Clades. Some of these groups correlate highly with human disease and outbreaks (hypervirulent) and others do not. Indeed, some groups of isolates appear

unlikely to cause disease in humans[3]. The genetic differences between isolates that are highly associated with human disease and those that aren't can be defined and tested for.

Molecular risk assessment?

The disease potential that a STEC has is governed by the virulence genes it carries and not by its serogroup. Molecular risk assessment has evolved as our understanding of exactly what is required to cause severe human disease has increased. That is, defining risk based on the presence of genetic markers and not on a STEC's affiliation with a particular serogroup as was previously the case.

Future testing systems

- NeoSeek – highly adaptable i.e. capable of rapidly integrating new genetic targets. Measures PCR amplicon size based on mass therefore avoiding the issues of using probes as in real-time PCR applications.
- Droplet digital PCR – partitions the samples into 1000's of droplets (single cells) and tests each droplet for genetic targets. Would allow genetic targets to be linked i.e. have confidence that *stx* and *eae* are in the same *E. coli.*
- Desktop sequencers – USB connected device that is rapid and requires minimal hands-on effort. Suitable for analysing 100's to 1000's of genes.

Conclusions

- Comparisons of STEC that cause human disease with those that generally do not allows us to identify the genetic factors that contribute most to human disease.
- Categorising STEC based on molecular risk will likely see a shift away from serogroup focused testing.
- Future testing platforms will increase the speed of testing primarily by removing the need for culture confirmation. Reductions in the costs of sequencing systems and the simplification of conducting these tests will aid the integration of future test systems into food production businesses.

Future typing methods – here now – P. Scott Chandry

Background

Technological and computational advances in the sequencing of DNA has transformed most of the biological sciences, particularly microbiology. Since the first commercial next generation sequencing (NGS) equipment became available (~2007), whole genome sequencing (WGS) has become a standard application in most microbial research laboratories. These advances have not been limited to the realm of research, NGS is rapidly becoming the "gold standard" technology for public health and food regulatory agencies around the world. The recent proliferation in the use of WGS for typing

[3] https://www.mla.com.au/research-and-development/search-rd-reports/final-report-details/Product-Integrity/E-coli-subtyping-data-collection/106

bacterial pathogens involved in food borne disease outbreaks in the USA, Canada, Europe and the UK indicates that it will become the standard technology for disease investigation globally.

Applications of NGS technology

- Tracking and identification of bacterial isolates using techniques like:
    - Single nucleotide polymorphism (SNP) analysis – in which every difference between the isolate strain and a reference strain are determined
    - Whole genome multi-locus sequence typing (wgMLST) – in which a reference database of gene types across 1000's of genes is established and all isolates are scored against these references. NOTE – June 8, 2017 PulseNet published a review paper suggesting that wgMLST is their preferred method to replace PFGE (Eurosurveillance Vol. 22, Issue 23, 2017)
- Predict functions e.g., antimicrobial resistance
- Analyse large microbial community (determine who is there without culturing)
- Numerous other applications + research tools

Industry adoption

This technology will replace commonly used methods such as Pulsed Field Gel Electrophoresis (PFGE), Multi Locus Sequence Typing (MLST) and Multi Locus Variable number tandem repeat Analysis (MLVA). The adoption of NGS based methods to the typing and testing of foodborne microbes is certain, only the extent of the disruption to current testing regimes and regulations remains to be determined. The Australian Red Meat Industry will need to be aware of the potential issues and benefits that the adoption of a new technology will bring.

Issues caused by NGS/WGS

- The end of serotyping
    - Classification systems will need to be revamped
    - New regulations will need to be discussed
- The end of PFGE – PulseNet
    - Now transitioning to WGS
    - Higher level of discrimination with WGS
        - New definitions of "relatedness"
        - Better understanding of the biogeographic variability
- New standards
    - New regulation, accreditation and standards needed
        - Laboratory data generation
        - Computational analysis (statistical, bioinformatic, phylogenetic)

Local Issues

- Australia lags behind the US, Canada, Europe on NGS implementation
- Limited baseline data for Australian food pathogens – may impact assumptions on isolate origin

- Date handling and availability
    - Who will access the data
    - How/where will it be stored (off-shore cloud?)
- WGS methods are moving forward for health applications – will the food sector have a voice in what is developed?
-

Benefits of WGS / NGS

- More certainty on accuracy of source tracking / typing data
    - False positive PFGE should end
    - Regional differences likely to be detectable
- Faster identifications and analysis
    - Sequence data can be transported electronically
    - Analysis can be automated
- Early detection of emerging food-borne pathogens

Next steps

NGS based methods represent the next logical step in the development of typing methodologies. Initially, typing was dominated by culture based methods that examined biochemical or physiological characteristics. This was followed by methods such as serotyping that examined the nature of important surface molecules on the cells. Then methods that used the genetic composition of the cells for typing were deployed such as PFGE, MLVA, and MLST. Technological changes have simply permitted a greater quantity of genetic information to be examined; so the current NGS based methods can be equated to an extremely high resolution version of PFGE. Although the research community has a myriad of applications for NGS, the public health community appears to be adopting a slow and steady approach toward applying NGS to the development of extremely accurate typing systems. Coincident with this, several older technologies such as PFGE will no longer be used. Methods such as serotyping will cease to be used in the very near future and necessitate some significant changes in the way microbes are typed. This will in turn lead to the requirement for some sweeping changes to regulations and standards.