Final report

Molecular assessment and characterisation of Australian Shiga toxin-producing E. coli (STEC)

Project code: V.MFS.0440

Prepared by: Glen Mellor
CSIRO Agriculture and Food

Project team: Narelle Fegan, Kate McMillan, Sean Moore, Lesley Duffy, Stanley Chen, Robert Barlow, Myintzu Hlaing, Derek Benson, Scott Chandry

Date published: 28 April 2023

PUBLISHED BY
Meat & Livestock Australia Limited
PO Box 1961
NORTH SYDNEY NSW 2059

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

Abstract

Shiga toxin producing *Escherichia coli (STEC)* are traditionally classified by serogroups, and of the hundreds of types isolated to date from cattle, only a small percentage are known to be pathogenic. The US considers seven serogroups (Top 7) to be adulterants of manufacturing beef. Expert groups (JEMRA, NACMCF) have recently proposed alternative risk schemes for STEC based on their disease potential, whose adoption could impact the Australian red meat industry. To estimate this impact, STEC isolates from Australian red meat sources were sequenced, analysed, and classified into the new schemes from low (group 5) to high (group 1) health risk. A genomics workflow was designed in-house for rapid, high-throughput and flexible risk profiling of STEC. Using this workflow, Australian isolates were shown to mostly fall into risk levels associated with medium to low human health risk. The JEMRA scheme was more likely to classify STEC into higher risk categories than NACMCF or FSIS risk schemes. This evidence-based data will prepare industry to make informed risk management decisions and to meet future export market requirements that may arise from the adoption of the new risk schemes.

Executive summary

Background

The Australian red meat industry routinely monitors manufacturing beef destined for export to the United States (US), for Shiga toxin-producing *Escherichia coli* (STEC) to meet current Food Safety and Inspection Service (FSIS) requirements and to minimise the reputational, trade and public health risks associated with contaminated beef entering the US market. The FSIS uses a serogroup-based approach to manage the risk of STEC, which considers the top seven serogroups (Top 7 STEC) associated with disease to be adulterants of manufacturing beef. This approach mostly excludes additional factors that might enhance or reduce the potential of STEC to cause disease. To account for this limitation, expert groups have proposed revised risk schemes that incorporate additional criteria to predict the likelihood of STEC to cause mild or severe disease outcomes in humans.

Here, we attempt to place two internationally proposed risk frameworks into a domestic context:
1) a risk framework proposed by the Food and Agriculture Organization (FAO)/World Health Organization (WHO) Joint Expert Meetings on Microbiological Risk Assessment (JEMRA), and
2) a risk framework proposed by the National Advisory Committee on Microbiological Criteria for Foods (NACMCF).

We believe that the adoption of such schemes by policy makers or overseas customers would greatly impact the Australian red meat industry; thus, it is imperative that industry develops a strong, scientifically robust knowledge base to engage with customers and regulators and to make informed decisions on managing the risk of STEC in the red meat industry.

Objectives

The project had three core objectives: (i) to describe the virulence profiles of historical Australian STEC (from CSIROs culture collection) according to internationally proposed risk classification schemes (ii) to conduct a survey of beef cattle faeces and manufacturing beef enrichment broths for the presence of *E. coli* harbouring priority virulence marker combinations as proposed by JEMRA and NACMCF and (iii) to characterise *E. coli* harbouring virulence markers of clinical relevance that were isolated from the survey of beef cattle faeces and manufacturing beef enrichment broths. The project achieved all three objectives, except for the survey sample target for calves, which was lowered due to challenges in sourcing samples from this animal class.

Methodology

- A national survey of STEC in Australian cattle (faeces: 2021-2022) and manufacturing beef enrichment broths (2021-2022) was undertaken to determine the prevalence of STEC, and to generate contemporary cattle/red meat isolates for inclusion in the risk characterisation component of the project.
- Additional STEC were sourced from historical samples (1987 – 2019 from the CSIRO culture collection) to capture Australian cattle, goat, sheep, and human sources.
- Development of a toolkit and data analytics workflow was developed for high throughput analysis of WGS data to determine the presence of virulence markers.
- The genomes of historical and contemporary STEC were sequenced and analysed and the results were used to classify isolates into the risk levels described in the two proposed risk schemes: JEMRA and NACMCF.

Results/key findings

National survey of STEC

Beef cattle faeces and manufacturing beef enrichment broths were screened for the presence of *E. coli* harbouring priority virulence marker combinations (*stx*$_1$, *stx*$_2$, *eae* and *aggR* gene markers) as proposed by JEMRA and NACMCF. A total of 40,950 isolates were recovered from across 910 samples: 710 cattle faeces and 200 manufacturing beef enrichment broths, with 32.7% of cattle faeces and 39.0% of beef enrichment broths yielding isolates possessing one or both *stx* genes (1 or 2) either alone or in combination with *eae*. Of the 310 samples that yielded STEC, 20% contained STEC with multiple different virulence profiles, yielding a total of 387 different virulotypes. All STEC with unique virulence gene profiles from each sample were characterised as below.

STEC risk characterisation

Two prominent international schemes (JEMRA and NACMCF) have been proposed to risk assess STEC into classes based on their estimated potential to cause disease. These two schemes plus the current FSIS definition for adulterant STEC were used to classify STEC recovered from both surveys into risk categories.

National survey of cattle faeces

Using the FSIS definition, 3.0% of samples contained STEC that were classified as adulterants while 27.9% were shown to contain non-adulterant STEC. By comparison, using the JEMRA system, 8.5% of samples contained STEC belonging to levels 1, 2 or 3 which have the highest potential for severe disease and 22.4% contained STEC belonging to levels 4 or 5 which have lower potential to cause haemolytic uremic syndrome (HUS) but may cause diarrhoea or bloody diarrhoea. Using the NACMCF risk scheme, no isolates were assigned to category 1 (highest health risk), 3.0% of samples were assigned to levels 2 and 3 (equivalent to current FSIS definitions) and 27.9% were assigned to risk levels 4 and 5 (lowest health risks).

Survey of beef trim enrichments

Using the FSIS definition, 5.5% of samples contained STEC that were classified as adulterants, while 33.0% contained STEC that were deemed non-adulterant. By comparison, using the JEMRA risk scheme, 14.0% of samples were assigned to levels 1, 2 or 3 which contain STEC that have the highest potential for severe disease, 24.5% were assigned to levels 4 or 5 which contain STEC with reduced potential to cause HUS but may cause diarrhoea or bloody diarrhoea. Using the NACMCF risk scheme, 0% of samples were assigned to level 1 (containing STEC of highest health risk), 5.5% were assigned to risk levels 2 and 3 (contain STEC that conform to current FSIS definition for adulterants) and 33.0% were assigned to risk levels 4 and 5 (STEC of lowest health risk).

Benefits to industry

The project generated a substantial database of whole genome sequences from a diverse set of Australian STEC, that will act as a valuable resource for genomic investigations into disease potential. The project also developed new tool kits that can be adapted and modified for the high-throughput whole genome sequence characterisation of STEC. We anticipate that risk classification data generated in this project will allow industry to assess the risk of STEC that are likely to be isolated through red meat testing, rapidly respond to changes to risk classifications and regulatory requirements from overseas customers and potentially lobby to reduce testing requirements and regulations around “low risk” STEC.

Future research and recommendations

Additional national surveys to collect STEC from animal groups and food sources that were underrepresented in this project should be undertaken. The further development of capability in the detection, isolation, and characterisation of STEC is recommended to support industry to evaluate/manage risk and meet future export market requirements of global customers of Australian red meat products. Collaboration with public health labs is recommended to gain access to data for comparing predicted disease potential of isolates with patient symptoms.

Table of contents

1. Background

The ability to export raw beef products to international markets requires certification that products have been tested and deemed free of certain pathogens, including Shiga-toxin producing *Escherichia coli* (STEC). Currently, STEC are broadly defined by importing countries as belonging to specific serogroups and carrying genes encoding for Shiga toxin (*stx*) and factors associated with gut attachment (*eae*, *aggR*). However, a greater understanding of STEC and the diseases they cause has provided new information around serogroups and Shiga toxins that may be used to redefine the types of STEC that are most likely to cause severe disease.

The United States (US) Food Safety and Inspection Service (FSIS) currently uses serogroups as a basis for managing the risk of STEC in beef, which they achieve through targeting the seven serogroups that account for most of the human disease cases in the US. These seven serogroups are collectively referred to as Top 7 STEC, and their detection in beef at the point of entry in the US, results in the product being rejected and consequently exported, destroyed or converted into feed within 45 days. The Top 7 STEC are defined in USDA FSIS Microbiology Laboratory Guidebook (1) as STEC containing *eae*, and belonging to one of the Top 7 serogroups (Table 1) The importance of serotyping as a tool for understanding the risk potential of STEC and other bacterial species is well established. The observation that certain serogroups are commonly associated with disease has allowed clinicians, researchers, and policy makers to group STEC into high and low risk groups. While serogroups are useful for broad epidemiological investigations, the antigens do no not confer virulence and are not the basis for disease. Rather, it is the genetic features conserved or commonly present in the genomes of isolates that comprise a serogroup that govern disease potential. However, not all isolates within a serogroup have the same disease potential and genetic variation can exist within a serogroup that alters their capacity to cause disease. Thus, there are many dimensions to consider when assessing the risk of STEC and while serogroups can provide insight into risk, there are knowledge gaps that limit the usefulness of serogroup-based risk schemes.

Table 1**.** US Food Safety and Inspection Service definition for adulterant (Top 7) STEC

LEVEL	VIRULENCE FACTORS
ADULTERANT	*stx* (any type) AND *eae* AND O26 OR O111 OR O103 OR O121 OR O45 OR O145 OR O157
NON-ADULTERANT	STEC with any other virulence gene combinations

Identifying the specific factors involved in pathogenesis of STEC is complex. Patients infected with STEC can show no symptoms (asymptomatic), or present with symptoms ranging from mild diarrhea to bloody diarrhea or haemolytic uraemic syndrome (HUS), a condition that can lead to hospitalisation, kidney failure and in extreme cases, death. To cause disease, STEC must have the capacity to produce a type of toxin, known as Shiga toxin (Stx). The frequency and severity of human disease has been linked, among other things, to the types of Shiga toxin(s) produced by STEC. Shiga toxin can be categorised into two broad classes, Stx1 and Stx2, each of which can be further divided into three Stx1 subtypes; Stx1a, Stx1c and Stx1d and seven Stx2 subtypes; Stx2a, Stx2b, Stx2c, Stx2d, Stx2e, Stx2f and Stx2g. While new subtypes have recently been described, their distribution in animals, the environment and their role in human disease is comparatively poorly understood. While Stx is necessary for disease, it has been suggested that STEC that only possess *stx* without additional genes that enable attachment to intestinal cells, are less likely to cause severe disease, though there are some exceptions. Consequently, high risk STEC are often, but not always, defined by the carriage of additional genes that enable them to adhere to intestinal cells, the most common of which is *eae* and the least common is *aggR*.

The ability to associate clinical outcome with toxin types and other such factors, has led to renewed interest in refining risk schemes for STEC. In 2018, an expert group from the FAO/WHO (JEMRA) proposed a risk classification scheme that classifies STEC into risk groups based on their potential to cause severe disease (2). The JEMRA scheme departs entirely from serogroup-based definitions, which may be due to a general agreement by the expert authors that all STEC have the capacity to cause disease, regardless of the serogroup to which they belong. The proposed scheme uses combinations of Stx subtypes and genes known to play a role in adherence, to classify STEC into five risk categories ranging from category 1 (highest potential for severe disease) to category 5 (lowest potential to cause severe disease) (Table 1).

Table 1. FAO/WHO (JEMRA) risk classification scheme – the estimated potential of STEC in each level to cause diarrhoea (D), bloody diarrhoea (BD) or haemolytic uraemic syndrom (HUS) is based on the combination of virulence factors they possess (2).

LEVEL	VIRULENCE FACTORS	POTENTIAL FOR[1]:
1	*stx*$_{2a}$ AND *eae* OR *aggR*	D/BD/HUS
2	*stx*$_{2d}$	D/BD/HUS[2]
3	*stx*$_{2c}$ AND *eae*	D/BD[3]
4	*Stx*$_{1a}$ AND *eae*	D/BD[3]
5	Any other *stx* type	D[3]

[1] Depending on host susceptibility or other factors; e.g. antibiotic treatment

[2] Association with HUS dependent on *stx*$_{2d}$ variant and strain background

[3] Some subtypes have been reported to cause BD, and on rare occasions HUS

A second, similar scheme, was proposed by the National Advisory Committee on Microbiological Criteria for Foods (NACMCF) in 2019 (3). Unlike the JEMRA scheme, the NACMCF scheme ranks STEC risk primarily by serogroup with secondary consideration given to the carriage of *stx* subtype(s) and gene(s) important for colonisation and adherence to the human gastrointestinal tract (*aggR* or *eae*). The scheme categorises STEC into five risk levels ranging from highest health risk (level 1) to lowest health risk (level 5), with sub-rankings to capture what the authors deem to be the *stx* subtypes presenting the greatest risk at each level. An adaptation of the risk diagram, originally published in the NACMCF work is shown in Table 2. Importantly, the authors of the report acknowledge that *eae*-positive STEC that are not part of the Top 7 STEC serogroups (risk level 4), may be of equal virulence to Top 7 serogroups (risk levels 2 and 3), however, their pathogenic potential is harder to assess due to their relatively low occurrence in disease.

Table 2. An adaptation of the NACMCF risk classification scheme (3). Human health risk is estimated to be high or low based on serogroup and virulence gene combinations associated with STEC of each risk level[a].

LEVEL	VIRULENCE FACTORS	HEALTH RISK
1	*stx*$_{2a}$ & *aggR*[b]	Highest risk
2	*stx* (any type) & *eae* & O157 *stx*$_{2a}$ > *stx*$_{2c}$ > *stx*$_{2a}$ + *stx*$_{1a}$ > *stx*$_{1a}$	↕
3	*stx* (any type) & *eae* & O26, O111, O103, O121, O45, O145 *stx*$_{2a}$ > *stx*$_{2d}$[a] > *stx*$_{2c}$ > *stx*$_{1a}$	
4	*stx* (any type) & *eae* & all other serogroups *stx*$_{2a}$ > *stx*$_{2d}$[1] > *stx*$_{2c}$ > *stx*$_{1a}$	
5	*stx* (any type) *stx*$_{2a}$ > *stx*$_{2d}$[1] > *stx*$_{2c}$ > *stx*$_{1a}$	Lowest risk

[a]Note that risk levels (1 to 5) have been added to simplify descriptions of levels in the report, and that these were not presented in the original risk diagram.
[b]Risk level 1 of the original diagram list *stx*$_{2a}$ + EAEC. EAEC refers to Enteroaggregative *Escherichia coli* (EAEC) which have, on rare occasions, been shown to carry *stx*, making them a hybrid of EAEC and STEC pathotypes. For simplicity, the EAEC has been replaced with *aggR* gene, which is a defining feature of EAEC.

This study was undertaken to place the JEMRA and NACMCF internationally proposed risk frameworks into a domestic context, and to estimate the potential impact that adoption of these schemes by policy makers would have on the Australian industry. To achieve this, we used Next Generation Sequence tools to obtain whole genome sequences of historical and contemporary sets of Australian STEC. Whole genome sequencing is strongly recommended for typing STEC (4), and is widely adopted globally, particularly for pathogen surveillance. We believe it is the most suitable method of collecting information on Australian STEC to support industry to make informed, evidence-based decisions. A comprehensive database of information about STEC, generated from genome sequences, will provide a strong, scientifically robust knowledge base for the red meat industry to engage with customers and regulators and to make informed domestic decisions on risk management approaches, rapidly respond to changes to risk classifications and regulatory requirements from overseas customers and potentially lobby to reduce testing requirements and regulations around “low risk” STEC.

2. Objectives

The objectives of this project were to:

- Describe virulence profile of historic Australian *E. coli* (from CSIRO culture collection) according to internationally proposed risk classification schemes – JEMRA and NACMCF.

- Conduct a survey of beef cattle faeces and manufacturing beef enrichment broths for the presence of *E. coli* harbouring priority virulence marker combinations as proposed by JEMRA and NACMCF.

- Characterise *E. coli* harbouring virulence markers of clinical relevance that were isolated from a survey of beef cattle faeces and manufacturing beef enrichment broths.

The objectives of this project were overall met and are described in the following sections. The target of 1000 samples collected as part of national surveys was not achieved with only 910 samples being collected, this was due to the difficulties in obtaining faecal samples from calves with slaughtered animals proving difficult to source

during the sampling period. The virulence profiles of historical Australian STEC, and those isolated from a national survey of cattle faeces and manufacturing beef enrichments were characterised to determine their risk in relation to proposed JEMRA and NACMCF schemes.

3. Methodology

3.1 Selection of historical Australian STEC

A total of 579 historical STEC isolates were selected from the CSIRO culture collection to represent a variety of sources, years, and strain types. Isolates were selected based on the following criteria:

- Must possess either *stx*$_1$, *stx*$_2$ or both based on PCR testing, with priority given to isolates carrying *eae*
- must be directly associated with red meat animal or human sources (e.g., no isolates from environmental sources or from dairy animals unless destined for red meat)
- selected to represent a range of serogroups
- selected to represent a wide range of years based on the date of isolation
- where possible, isolates were chosen to represent a broad range of animal groups tested (e.g., limit choosing isolates obtained from the same animal cohort)

The majority of STEC were isolated between 1996 and 2000 (covering early MLA/CSIRO co-funded projects and CSIRO funded projects) and 2012 and 2018 (covering MLA/CSIRO national surveys for pathogenic and Top 7 STEC in cattle and sheep faeces) (Figure 1).

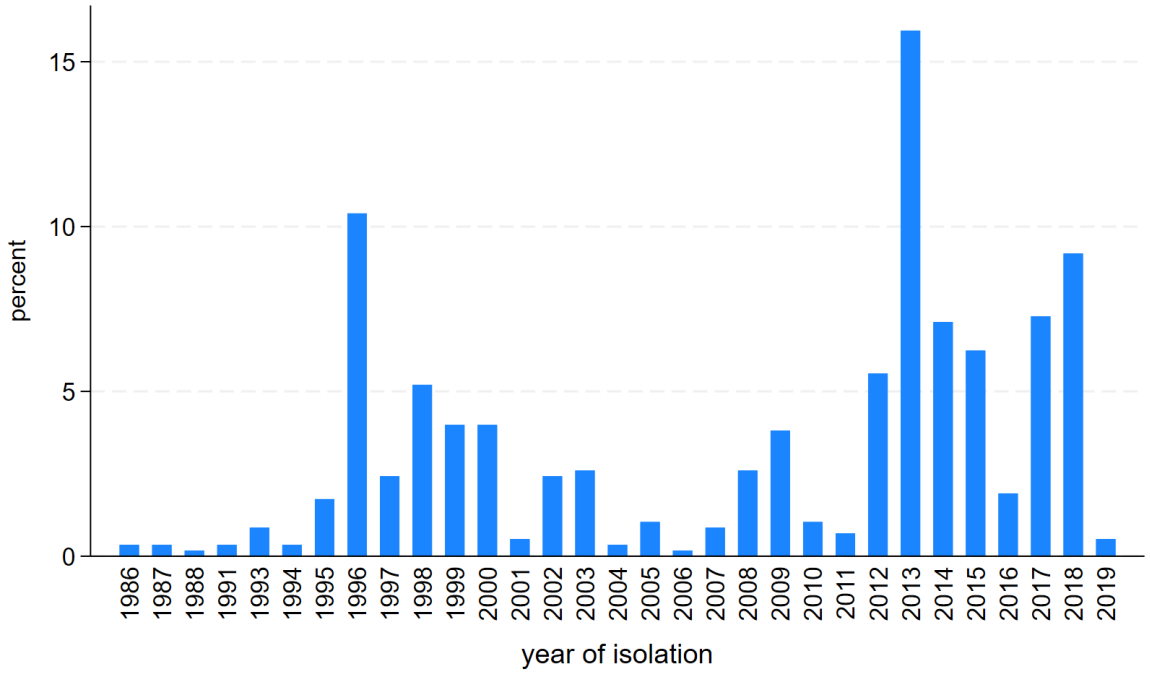

Figure 1. Percent of STEC selected for whole genome sequencing (WGS) showing the years from which they were isolated.

Most STEC isolates (56%) belonged to three serogroups: O157, O26 and O111. A further 34% of isolates belonged to 34 different serogroups (grouped as 'other') and a smaller proportion of isolates (10%) had an undetermined serogroup status (Figure 2).

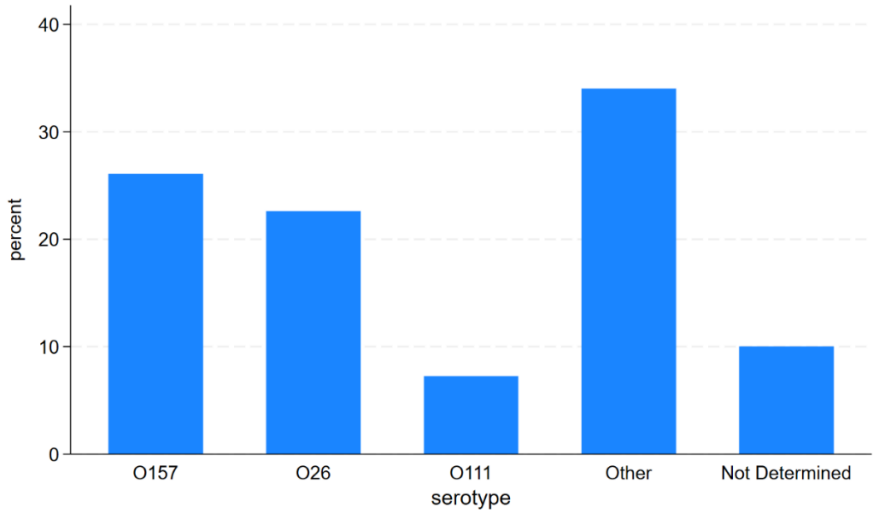

Figure 2. Serogroups of 579 STEC isolates used for this study. 'Other' represents 34 different serogroups.

The majority of STEC were from cattle (56%) as this represented the largest group of STEC isolates held in the collection, followed by sheep (24%), human (17%) and goat (3%) (Figure 3). STEC were obtained from a variety of sample types including faeces, beef trim (manufacturing beef), carcases, hide and fleece as well as clinical human samples (Figure 3). The 17% of isolates that were derived from clinical sources (humans) were sourced from Australian State public health laboratories, or publicly available datasets.

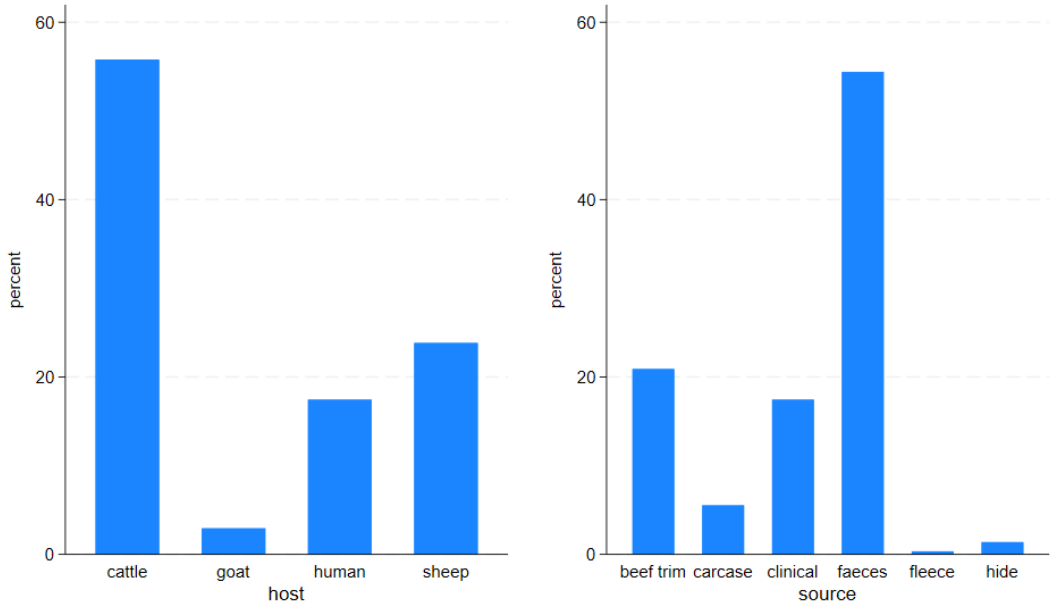

Figure 3. Percent of STEC isolated from different hosts (left) and sources (right).

STEC isolates were selected to provide information around their potential risks based on JEMRA and NACMCF criteria, of which *eae* and *stx* type are key criteria in each scheme. A large selection (53%) of isolates were chosen based on the presence of *eae* in combination with one or both Stx genes (Figure 4) as these isolates are of potential higher risk to humans. A further breakdown of STEC based on source along with virulence markers shows that those from cattle represent the greatest proportion which carry *eae* and *stx*$_2$ (Figure 4), this reflects the large numbers of isolates collected from cattle throughout the years along with a lower prevalence of *eae* containing isolates obtained from sheep.

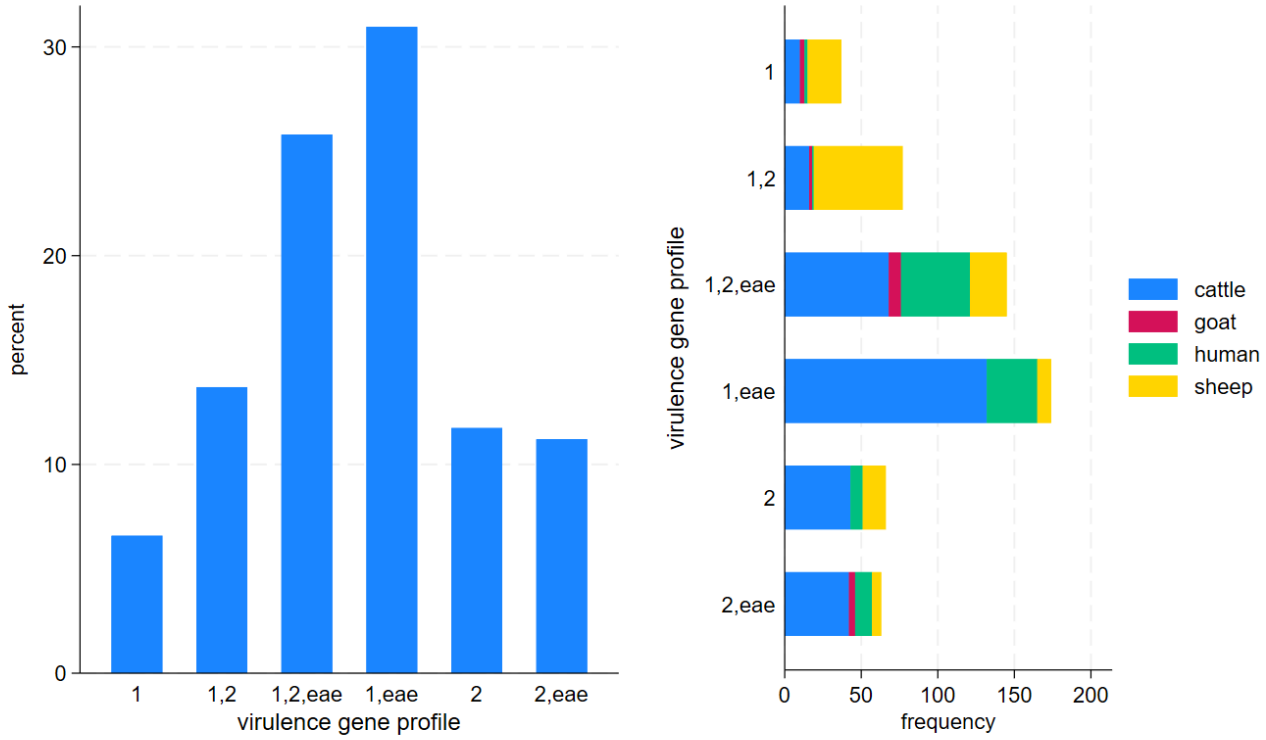

Figure 4. Percent of STEC carrying various virulence gene marker combinations (1– *stx*$_1$; 2 – *stx*$_2$; *eae* – *eae*) (left). The number of STEC carrying various virulence gene markers (1– *stx*$_1$; 2 – *stx*$_2$; *eae* – *eae*) from different hosts (right).

3.2 National survey of STEC in cattle and manufacturing beef

A national survey of STEC in Australian cattle (faeces) and manufacturing beef enrichment broths was undertaken to determine the prevalence of STEC using methods that are designed to capture all STEC.

3.2.1 Survey of beef cattle (faeces) at slaughter

Invitation letter/questionnaire

Australian Meat Industry Council (AMIC) members were invited to participate in a national survey of STEC in cattle faeces at slaughter. A summary of participating establishments and details of the sampling approach are discussed in the results.

Establishment participation

A formal survey invitation (Appendix 8.1 Survey invitation letter) and questionnaire (Appendix 8.2 Survey Questionnaire) was developed in consultation with Meat & Livestock Australia (MLA) and the Australian Meat Industry Council (AMIC). Survey invitations and questionnaires were distributed to AMIC members through AMIC processor group circular No: 15-21: 2021 and to non-members by email. Establishments were provided with background information on the purpose of the project, the benefits and risks of participating, privacy and confidentiality considerations and the commitment required from establishments. The survey was reviewed and approved by CSIRO's human ethics (clearance number 139/20) and privacy governance teams prior to commencement. To help in planning the survey, establishments were asked to complete a brief one-page questionnaire about their facility that included questions on current processing volumes of the three animal classes of interest: beef cattle, dairy cattle, and veal calves. To achieve national representation of slaughter

animals, attempts were made to capture industry participation to the extent that the collective production volume of participating plants was at least 50% of the total yearly Australian production.

A sampling plan, consistent with previous beef and sheep pathogen surveys, was developed for estimating the prevalence of STEC in Australian beef cattle. The sampling plan had a target of 800 samples, to be collected from across three animal classes: beef cattle, dairy cattle, and veal calves. A collection target of 500 beef cattle, 150 dairy cattle and 150 veal calves was defined for each animal class. Animal classes, feed type and production system definitions are provided in Table 3. A stratified sampling approach was employed to ensure the number of animals sampled at each processor was proportional to the slaughter volumes for each animal class and total plant production (probability proportional to size sampling approach). Samples were collected on 3 occasions across 11 months, from May 2021 to April 2022. To minimise the effect of clustering and maximise the validity of inferences that can be made from the analysis, establishments were asked to space sample collection across 2-3 days on each of the 3 sampling occasions and to allow a 30 min interval between collection of consecutive samples.

Table 3. Target animal classes, feed type and production system definitions

Animal class / feed type		Definition
Animal class	Beef Cattle	Defined as per Aus-Meat beef meat language[a]
	Dairy Cattle	Beef Cattle derived from a dedicated dairy operation
	Calves	Defined as per Aus-Meat beef meat language[a]
Feed type/production system	Grass-fed	Raised solely on pasture
	Grain-assisted grass-fed	Raised on a diet of pasture and supplemented feed (grain, oats etc)
	Grain-fed	Animals that have been through a NFAS feedlot

[a]Aus-Meat (2005) Handbook of Australian meat, 7th ed, AUS-MEAT Limited, Brisbane, Australia

Sample collection and enrichment

Participating establishments were asked to collect faecal samples from target animal's post-evisceration by cutting the intestine approximately 30-50cm from the bagged rectum and squeezing the faecal content into a sterile jar. Establishments were asked to complete a sample datasheet to capture relevant metadata such as animal class, feed, date, and time of sample etc. Samples were kept chilled and returned to the CSIRO's Coopers Plains laboratory by overnight courier using chiller boxes designed to maintain a temperature of 2-8°C during transport. On arrival, a single faecal slurry was prepared for each sample by diluting 25 g of faeces (1 in 10) in buffered peptone water (BPW; Oxoid, UK). Each slurry was stomached for 60 s, transferred to filter bags, and enriched at 41.5 ± 1°C for 18 ± 2 h.

STEC Isolation

Attempts were made to isolate STEC from all faecal enrichments. An enrichment/isolation approach, largely consistent with ISO/TS 13136 (5), was used for the isolation of STEC with plating media tested using Australian STEC (Appendix 8.3 STEC growth characteristics). BPW enrichments (described in sample collection and enrichment) were serially diluted and spiral plated onto three different media: modified Rainbow™ agar O157 (Biolog, Hayward, CA, USA) supplemented with 5.0 mg/l sodium novobiocin, 0.05 mg/l cefixime trihydrate, and 0.15 mg/l potassium tellurite (mRBA), Tryptone Bile X-Glucuronide agar (TBX; Oxoid, UK) and CHROMagar™ STEC

agar (CHROMagar, Springfield, NJ, USA). The easySpiral Dilute was used to automate dilution and plating onto 90 mL Petri dishes using the exponential plating mode to maximise colony separation for ease of downstream isolation/purification. Spiral plates were incubated at 37 ± 1°C for 24 ± 2 h after which 45 colonies were picked into BPW from across the 3 plates: 10 from mRBA, 10 from CHROMagar STEC and 25 from TBX. Colonies picked into BPW were incubated at 37 ± 1°C for 18 ± 2 h. Colonies from mRBA plates were selected based on their distinct morphologies. Following incubation, whole cell suspensions (WCS) were prepared by transferring 20 µL of BPW enrichments into 180 µL of sterile water. A 5 µl volume of each WCS was used as DNA template in real-time PCR for the detection of *stx*$_1$, *stx*$_2$, *stx*$_{2f}$, *eae* and *aggR* gene markers using the primer and probe combinations described in *PCR screening of STEC*. For each sample, WCS's that tested positive for *stx* with or without *eae* and *aggR* were plated onto Tryptone Soya Agar (TSA; Oxoid). Following incubation, single colonies were picked into 500 µL of sterile water and screened again for *stx*$_1$, *stx*$_2$, *stx*$_{2f}$, *eae* and *aggR* gene markers to ensure isolate purity.

A subset of 103 isolates (Appendix 8.5) that tested positive were subsequently screened for *stx* gene subtypes following the Statens Serum Institut method for the identification of *stx*$_1$ and *stx*$_2$ subtypes – used in the international external quality assessment (EQA) scheme (EQA-10) for typing of STEC**.** EQA subtyping was used as a 'gold standard' reference for evaluating whole genome sequence typing pipelines. Isolates that tested positive for *stx* were stored at -80°C using bacterial preservation cryobeads (Technical Service Consultants, United Kingdom).

PCR screening of E. coli *for virulence markers*

A custom 5-plex real-time PCR assay was developed for high throughput screening of STEC for the presence of *stx*$_1$, *stx*$_2$, *stx*$_{2f}$, *eae* and *aggR* gene markers. The multiplex PCR was designed using the *stx*$_1$, *stx*$_2$ and *eae* primers outlined in appendix 4 of FSIS guidebook 5C.03 (1), the *stx*$_{2f}$ primers described by Holmes *et al*. (6) and *aggR* primers described in Method 05 Rev 1 of the European Union (EU) Reference Laboratory for *E. coli* (7). The fluorophores used for each probe are detailed in Table 4. Unique fluorophores were used to differentiate *stx*$_1$ from *stx*$_2$, which contrasts with the US Department of Agriculture (USDA) FSIS approach of using an identical fluorophore for both *stx* classes. The USDA FSIS primer/probe sequences are identical to those used in ISO/TS 13136 and capture *stx* subtypes: *stx*$_{1a}$, *stx*$_{1c}$, *stx*$_{1d}$, *stx*$_{2a}$, *stx*$_{2b}$, *stx*$_{2c}$, *stx*$_{2d}$, *stx*$_{2e}$ and *stx*$_{2g}$. As the FSIS/ISO *stx* primers fail to capture the *stx*$_{2f}$ subtype, which varies significantly from the other *stx*$_2$ subtypes, an additional set of primers specific to this subtype were incorporated in the multiplex PCR. A cycle threshold (CT) of 35 was used as an arbitrary cut-off value for positive samples (i.e. samples with CT values above 35 cycles were deemed to be negative).

Table 4. Primer/probe combinations used in the 5-plex real-time PCR.

Gene target	Primer/probe (5'-3')	Primer sequence (5'-3')	Reference
stx$_1$	Forward Primer	TTTGTYACTGTSACAGCWGAAGCYTTACG	U.S. Dep. Of Agriculture 2019 (1)
	Reverse Primer	CCCCAGTTCARWGTRAGRTCMACRTC	
	Probe stx1	ATTO550/CTGGATGATCTCAGTGGGCGTTCTTATGTAA-IBRQ	
stx$_2$	Forward Primer	TTTGTYACTGTSACAGCWGAAGCYTTACG	U.S. Dep. Of Agriculture 2019 (1)
	Reverse Primer	CCCCAGTTCARWGTRAGRTCMACRTC	
	Probe stx2	6-FAM/ZEN/ TCGTCAGGCACTGTCTGAAACTGCTCC-IBFQ	
eae	Forward Primer	CATTGATCAGGATTTTTCTGGTGATA	U.S. Dep. Of Agriculture 2019 (1)
	Reverse Primer	CTCATGCGGAAATAGCCGTTA	
	Probe eae	SUN[a]/ZEN/ATAGTCTCGCCAGTATTCGCCACCAATACC-IBFQ	
stx$_{2f}$	Forward Primer	TTGTCACAGTGATAGCAGAAGCTCTG	Holmes et al. 2018 (2)
	Reverse Primer	CAGTTCAGGGTAAGGTCAACATCC	

	Probe stx2f	6-FAM/ZEN/CGCTGTCTGAGGCATCTCCGCTTTATAC-IBFQ	
aggR	Forward Primer Reverse Primer Probe aggR	GAATCGTCAGCATCAGCTACA CCTAAAGGATGCCCTGATGA Cy5/CGGACAACTGCAAGCATCTA-IBFQ	EU Reference Laboratory for E. coli 2013 Method 05 Rev 1 (3)

[a]The SUN fluorophore is identical to the VIC fluorophore recommended for labelling stx_1/stx_2 in MLG 5C.

3.2.2 Survey of manufacturing beef enrichments

Sample collection and processing

Australian manufacturing beef enrichment broths were collected from across 5 commercial laboratories that collectively service most of the Australian beef producing establishments. Commercial labs were asked to supply enrichment broths that were PCR screen positive for Shiga toxin gene markers, regardless of whether they were deemed potentially positive for Top 7 STEC. De-identified aliquots of manufacturing beef enrichment broths (~25 mL) were provided by commercial laboratories to CSIRO for STEC testing - typically on a fortnightly basis until a collection target of 200 broths was achieved. Sample metadata collection was limited to remove the risk of reidentification of establishments. Sample data included: date of collection, sample type (raw meat/broth), test system, potential positive and confirmed status (i.e., an STEC had been isolated), along with the serogroup and the gene profile of confirmed isolates. Upon arrival at CSIRO, attempts were made to isolate STEC from enrichments using the method described for *STEC isolation*.

3.3 Whole genome sequence analysis workflow

3.3.1 Sequencing

A total of 959 STEC isolates, comprising 579 from historical and 380 from contemporary strain collections, were sequenced and characterised for the priority virulence markers outlined in the risk classification schemes proposed by JEMRA and NACMCF. Automated genomic DNA (gDNA) sample preparation was performed using a Qiagen QIACube following the protocol for gram-negative bacteria from the DNeasy Blood and Tissue Kit (Qiagen). Preliminary quality control was performed on gDNA extractions using a Qubit Fluorometer with the dsDNA HS Assay Kit (Invitrogen). All extractions were screened for stx_1, stx_2, *eae* and *aggR* gene markers by real-time PCR to confirm isolate purity and identity prior to sequencing. Further quality control of gDNA, library preparation and sequencing runs were performed at the Ramaciotti Centre for Genomics, University of New South Wales (historical set) or the Australian Centre for Ecogenomics (ACE: contemporary set). DNA libraries were prepared using the Nextera DNA Flex library preparation kit (Illumina) and paired end (2x150bp) reads were generated using the NovaSeq 6000 system (Illumina).

3.3.2 Workflow development

Whole genome sequence data was transferred from service provider facilities to a secure server within CSIRO's Advanced Scientific Computing facility for storage and analysis. The Galaxy biomedical platform workflow was used for the computational analysis of sequences. Galaxy is a workflow management tool that allows for scalable, sharable, and reproducible, version controlled computational workflows. CSIRO maintains its own instance of Galaxy that interacts with CSIRO's big data storage and high-performance computing facilities which enables the rapid and high throughput risk profiling of STEC isolates. Some computational analysis was performed outside of the Galaxy workflow management system, due to the absence of available tools within the Galaxy ecosystem. An overview of the workflow steps is presented in Figure 5. Note that the workflow contains multiple tools and

options and not all of these were required for the delivery of the milestone and therefore do not appear in the report.

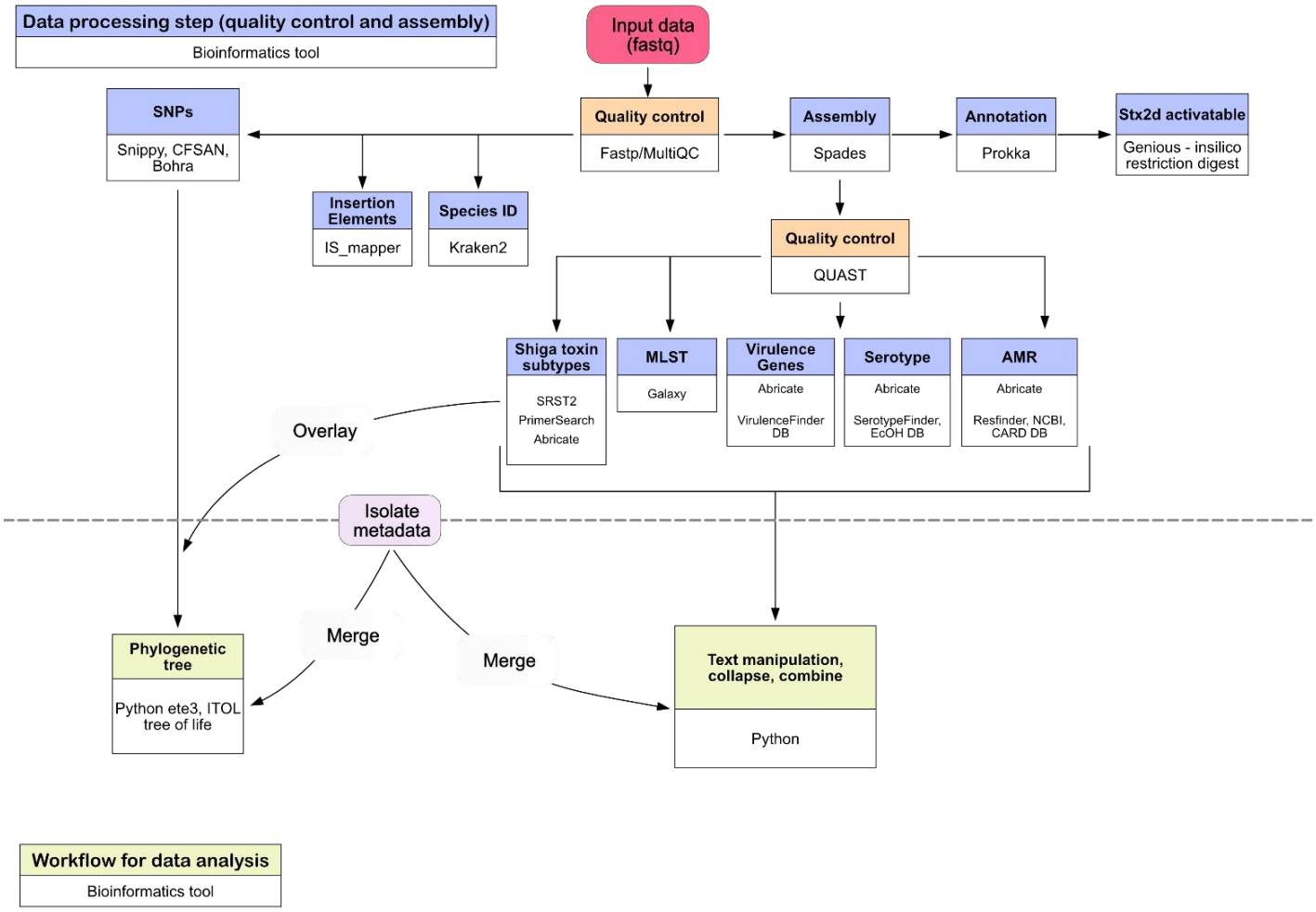

Figure 5. Galaxy workflow showing data processing (violet), filtering of sequences that failed quality control and incorporation of sequence metadata (orange), analysis of sequences (green) and reporting (pink). Some workflow steps were performed outside of Galaxy due to absence of available tools within the ecosystem.

The processing and subsequent analysis of sequences is computationally intensive and time consuming when done at scale. To significantly increase the computational power available for sequence analysis, the CSIRO Galaxy service was used to submit jobs to the CSIRO High Performance Computing Resource systems (Figure 6). The Galaxy workflow shown in Figure automates what is computationally demanding task and enables it to be completed with very little user interaction.

Figure 6. CSIRO's High-Performance Computing Resource (400 server nodes with greater than 25,000 cores and 235TB of system memory).

In silico detection of serogroup and stx genes

Sequence read files were concatenated where required. Illumina adaptors were removed as were low quality bases (clipping) using Fastp v0.20.1 (8) and outputs were visualised in MultiQC v1.9 (9). Reads were De Novo assembled using Spades version 3.12.0 (10) with kmers of 21, 33, 55, 77. Constructed assemblies were quality assessed with QUAST version 5.0.2 (11). Spades assembles short sequence reads into a series of contiguous DNA segments (contigs) of various sizes. Each of these contigs was screened for virulence genes using the Abricate application against the Centre for Genomic Epidemiology (CGE) database (accessed 30-09-2019: https://cge.cbs.dtu.dk/services/VirulenceFinder/). Serogroups were identified using the CGE SerotypeFinder database 1.0.0 (accessed 27-02-2019: https://cge.cbs.dtu.dk/services/SerotypeFinder/) and the EcOH database (12). Serogroup and virulence gene outputs were filtered to achieve >95% base identity over >90% read length match with database entries. Any sequences not meeting these requirements were dropped from the analysis. Multiple gene identity and coverage cut-offs were tested and only those that yielded the highest concordance between *in silico* and expected results were used in analyses (results not shown).

Tools used to assess performance of workflow

Additional tools and pipelines were compared to Abricate to assess their performance to accurately call virulence genes. These included the European Union Reference Laboratory (EURL) VTEC WGS Pipeline (13), Short Read Sequence Typing tool version 0.2.0 (SRST2) (14), and the PrimerSearch tool from the EMBOSS suite (15). The EURL pipeline was run on default settings. The SRST2 tool was run with 10 allowed mismatches per read, a minimum of 98% coverage, 1% divergence cut-off and greater than 5 x coverage of the gene. Raw reads were mapped directly to three separate databases for O-antigen genes, Locus of Enterocyte Effacement (LEE) genes, and Shiga toxin genes (*stx*), accessed through the SRST2 git repository (srst2/data at master · katholt/srst2 · GitHub). The divergence cut-off was set low (1%) due to the use of a large database containing many genes and alleles. PrimerSearch was used to search the following primer pair gene targets against the sequence assemblies: $stx_{1a,c,d}$ and $stx_{2a\text{-}f,}$ *eae* and *aggR* with 2% allowable mismatches.

Assessment of Insertion Elements (IS elements)

Raw sequences were also assessed for the presence of insertion elements using the ISMapper tool (16). Outputs tables were analysed to identify the presence of insertion elements (IS1203, IS1203v, IS1203e and IS629) at Shiga toxin gene locations, using the *E. coli* O157 Sakai strain genome for reference mapping. The four insertion elements were chosen on the basis that they have previously reported to cause insertional inactivation of the *stx* gene.

Data processing of workflow and genomic tool outputs

All workflow outputs were processed in Python version 3.10.9 to clean and transform data prior to the automating the classification isolates into JEMRA and NACMCF risk classification levels in addition to investigating other attributes.

3.3.3 Assessment of workflow performance

A subset of 103 STEC sequences were used to assess the ability of Galaxy workflow tools to correctly call *stx* subtypes (Appendix 8.5). The subset of sequences was selected to represent the diversity of STEC serogroups and virulotypes reported in MS2. All isolates were typed using conventional PCR following the "gold standard" method for detecting *stx* gene subtypes: stx_{1a}, stx_{1c}, stx_{1d}, stx_{2a}, stx_{2b}, stx_{2c}, stx_{2d}, stx_{2e} and stx_{2g} (EQA-10)(17). The concordance between in silico typing tools and the "gold standard" EQA method are reported in Section 4.2.1.

3.4 Shiga toxin expression assay

3.4.1 Shiga toxin induction

Shiga toxin induction was performed using a method adapted from Shringi et al. (18) . In brief, cultures were prepared with a single colony in 5 ml sterile Luria- Bertani Miller broth (LB broth; Oxoid, Basingstoke, United Kingdom) in 50-ml centrifuge tubes (Corning, Australia). Cultures were incubated at 37°C with rotary shaking (160 rpm) for 18±2 h. Cultures were diluted 1:200 in 20 mL fresh LB broth in 50 mL falcon tube and incubated the cultures for 30 min (to reach an absorbance of 0.3-0.4 at A_{600} nm). The cultures were then induced with mitomycin C at a final concentration of 0.5 µg/ml for 7 h at 37°C with rotary shaking (160 rpm).

Cells were harvested hourly by centrifugation, washed twice, and then resuspended with sterile PBS. Enumeration of viable mitomycin-induced cells was evaluated by direct viable count from serial dilutions of the washed samples (in peptone saline recovery solution) on Nutrient agar after aerobic incubation for 18±2 h at 37 °C. Viable counts were expressed as CFU/mL. The cell viability and cellular membrane integrity measurement of the mitomycin-induced cells was determined using Flow cytometry. For downstream transcriptomic experiments, one volume of the cell suspension in PBS was immediately mixed in two volumes of RNA protect Bacteria Reagent (Qiagen, Australia) and incubated for 5 min at room temperature (22 °C). Following treatment, the cells were collected by centrifugation and the pellets were stored frozen at − 20 °C for up to 2 weeks prior to RNA isolation.

3.4.2 Fluorescent-based analysis of cellular viability

A rapid counting of live/dead bacteria and membrane integrity assay (BD™ Cell Viability Kit) was used to evaluate the extent of membrane integrity of treated cells according to the supplier. A 3 µL of each dye component (the propidium iodide, PI and thiazole orange (TO) dye components provided with the kits) was added to each 500 µL of the washed cell suspension and then incubated in the dark for at least 5 minutes at room temperature. For each analysis, additional tubes of an unstained control sample and TO alone stained cells were included to determine the viable cell gate. Samples were analysed on the BD Accuri C6 flow cytometer equipped with BD Accuri™ C6 Software by setting up an SSC threshold for microbial cells. TO stained cells fluoresce primarily in FL1 and FL2 and PI-stained cells fluoresce primarily in FL3. Therefore, the best discrimination of live/dead populations and membrane injured cells was performed on an FL1 vs FL3 plot.

3.4.3 Transcriptomic analysis /targeted gene expression using droplet digital PCR

Frozen cell suspensions were thawed at room temperature (22 °C), and the biomass was harvested by centrifugation. Total RNA was isolated from each of the bacterial cell samples using the RNeasy® Mini Kit (Qiagen, Australia) following the procedures described by the manufacturer. The RNA concentration was measured using a Qubit Fluorometer and Qubit RNA HS assay Kits (Life Technologies, Australia). The RNA quality was evaluated with

an Agilent 2200 TapeStation system using the Agilent RNA ScreenTape assay. The extracted RNA was stored at -80 °C until cDNA synthesis for real-time PCR.

Equal amounts of total RNA extracted from the treated samples and from untreated control samples (two biological replicates each from different sample preparations) were reverse transcribed to cDNA using iScript™ cDNA Synthesis Kit (Bio-Rad, Australia) according to the manufacturer's instructions.

The cDNA samples were used for absolute quantification of nucleic acid target sequence-specific gene expression analysis using the QX200 ddPCR system (Bio-Rad, Australia) with ddPCR EvaGreen Supermix Kit (Bio-Rad, Australia) according to the manufacturer's instructions. Fluorescent probes used for PCR are listed in Table 5. Stx primer/probe sequences were obtained from USDA-FSIS MLG 5 while oligonucleotide primer and probe sequences for the reference gene (*tufA*) were designed in-house using NCBI's primer tool software (19). Emulsified reaction droplets were generated using a QX100 Droplet generator (Bio-Rad) by loading 20 µl of each reaction mixture into a sample well of DG8 cartridge (Bio-Rad, Cat. # 186-4008) followed by 70 µl of ddPCR droplet generation oil for Probes (Bio-Rad. Cat. #186-3005). The 40 µl samples of the generated droplet emulsions were transferred to 96-well PCR plates, which were heat-sealed using foil sheets. Target DNA amplification was performed by thermal cycling the droplet emulsions and the fluorescence of each thermal cycled droplet was measured using the QX100 droplet reader (Bio-Rad). Data were analysed using QuantaSoft software (Bio-Rad) after threshold setting on fluorescence of negative controls. The expression level of each gene was calculated based on absolute copies number/microliter of the reaction/nanogram of cDNA and normalized using that of the housekeeping gene. The mean fold-change (Log_2 ratio) expression level for each gene between each treatment group and the control untreated group was calculated.

Table 5. Oligonucleotides primers and probes used in ddPCR toxin expression experiments.

Oligonucleotide ID	Sequence	Description
TufAF	5'-AAT GTT CCG CAA ACT GCT GG-3'	Primers for housekeeping reference gene
TufAR	5'-CAG TAC CTG ACC ACG TTC GA-3'	
StxF	5'-TTT GTY ACT GTS ACA GCW GAA GCY TTA CG-3'	Primers for *stx*
StxR	5'-CCC CAG TTC ARW GTR AGR TCM ACD TC-3'	
TufAP	5'-/5SUN/CTG CTG CGT/ZEN/GGT ATC AAA CG/31AkFQ/-3'	Probe for housekeeping reference gene
stx_1 probe	5'-/56-FAM/CTG GAT GAT/ZEN/CTC AGT GGG CGT TCT TAT GTA A/31ABkFQ/-3'	Probe for *stx_1* gene
stx_2 probe	5'-/56-FAM/TCG TCA GGC/ZEN/ACT GTC TGA AAC TGC TCC /31ABkFQ/-3'	Probe for *stx_2* gene

3.4.4 Shiga toxin production assays

In addition to expression studies, a total of 93 isolates, selected to represent different risk groups and Shiga toxin genotypes, were assessed for their capacity to produce Shiga toxins 1 and 2 using a rapid membrane enzyme immunoassay (Quik Chek). Shiga toxin is visible as blue lines on a lateral flow card (Figure 7). The intensity of the blue line was used to roughly estimate relative amount of toxin produced across samples.

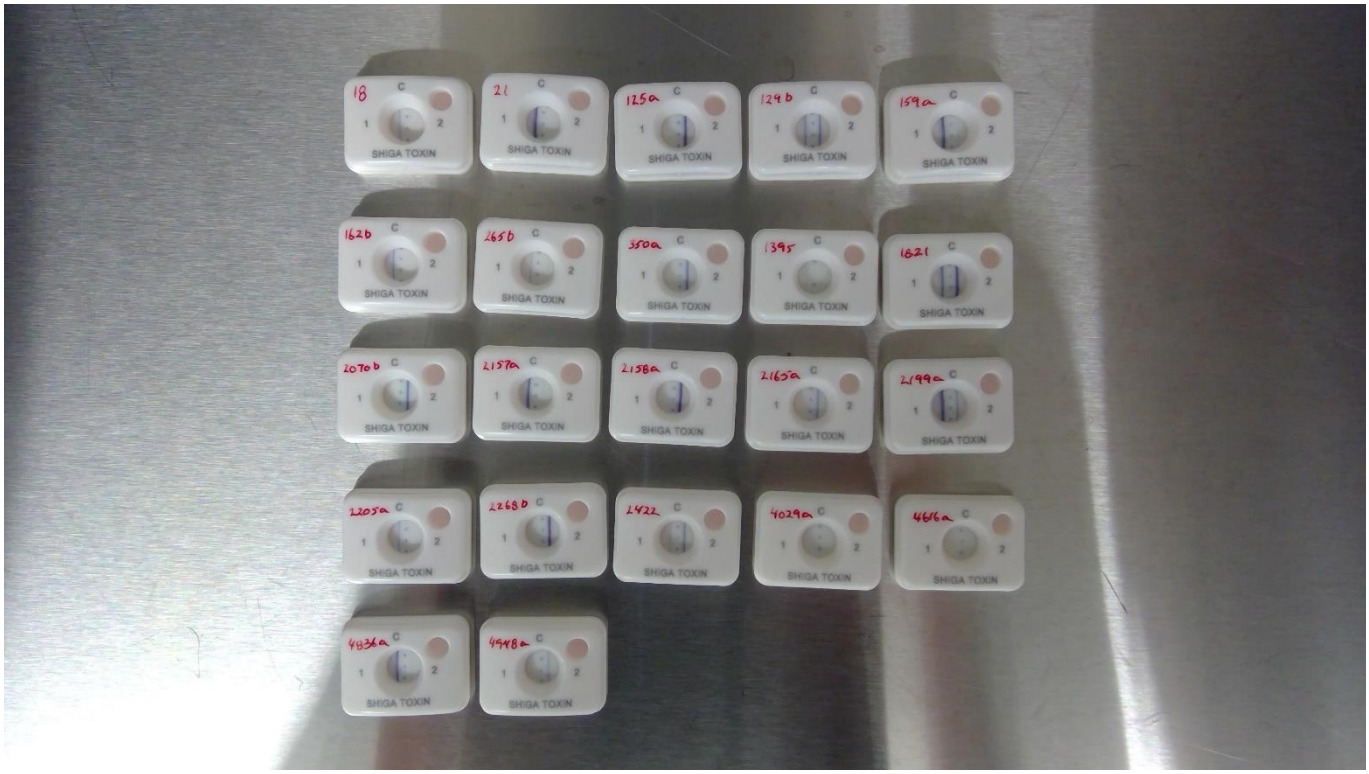

Figure 7. Example of Shiga toxin QuikChek immunoassay results. The production of Stx1 and Stx2 toxin appears as a visible blue line.

4. Results

4.1 Survey of beef cattle (faeces) at slaughter

4.1.1 Establishment participation

A total of 22 establishments participated in the project, collectively representing 57% and 6% of the 2021 forecast production of beef and veal, respectively (Table 6). All establishments indicated that they process beef cattle and 7 of the 22 also indicated that they process dairy cattle in addition to beef cattle. Only one establishment indicated that they were slaughtering calves during the survey period.

Table 6. Number of animals processed by the 22 participating establishments as a percent of total Australian beef and calf slaughterings.

Animal class	Combined yearly slaughtering's from all 22 participants	Total yearly Australian slaughterings[a]	Percent of total Australian slaughtering's represented by the 22 participants
Beef cattle[b]	3,949,092	6,914,000	57
Calves	24,000[c]	419,000	6

[a]MLA Industry Projections 2021 (Meat and Livestock Australia (2021). Includes export and domestic production.
[b]Includes dairy cattle slaughterings.
[c]Only one of the 22 participating establishments were processing calves during the survey period.

4.1.2 Sample collection

With the exception of veal/calf samples, the survey target quota for all sample types, was achieved or close to achieved (Table 7). Specifically, 99.8% of the target quota was achieved for beef, 97.3% of the target quota was achieved for dairy and 100% of the target quota was achieved for enrichment broths. Due to challenges in identifying and recruiting veal producers during the sampling period, the survey did not achieve the target quota for veal/calves (43.3% of target number achieved) or achieve national representation of this animal class.

In total, 710 beef faecal samples were collected across three windows spanning from May to June 2021 (window 1, n=190), September to December 2021 (window 2, n=304) and February to April 2022 (window 3, n=216). Samples included in the survey were derived from beef cattle (70%), dairy cattle (21%) and calves (9%) from across the major beef producing States: QLD (44.4%), NSW (21.2%), VIC (14.0%), TAS (10.3%), SA (5.7%), NT (0.5%) and WA (0.5%). A small percentage of samples were listed as unknown origin (3.8%) where source information was not provided.

A total of 200 enrichment broth samples were collected from commercial test labs continuously across a single seven-month window between September 2021 and February 2022. According to the metadata provided by the commercial labs, 198 of the 200 (99%) enrichment broths were potentially positive for *stx* either with or without *eae*. The remaining 2 (1%) samples were screen positive for *eae* only – these were processed despite not meeting the minimum requirements for testing (potentially positive for *stx*). A total of 160 (81%) of the 198 *stx*-positive samples were potentially positive for *stx* and *eae*, while 38 (19%) of 198 were potentially positive for *stx* alone.

Table 7. Survey sample target quota versus quota achieved

Sample type	Initial collection target for each sample type	Number of samples collected in the survey
Beef faeces	500	499
Dairy faeces	150	146
Calf faeces	150	65
Enrichment broths	200	200
Total	1000	910

4.1.3 Isolation of STEC

A total of 40,950 isolates (45 isolates per sample from the three different media) were recovered from across 910 samples: 710 beef faeces and 200 enrichment broths. Isolates were screened by real-time PCR for stx_1, stx_2, *eae* and *aggR* gene markers. Of the 40,950 isolates, 2,071 (5.1%) were confirmed to possess one or both *stx* genes (1 or 2) either alone or in combination with *eae* (Appendix 8.4 Number of isolates with distinct virulence profiles). None of the 40,950 isolates tested positive for *aggR*.

4.1.4 Prevalence of STEC

In total, 310 of 910 samples (34.1%) were confirmed to contain at least one *stx*-containing isolate (Table 8). The number of enrichment broth samples likely to contain STEC was slightly higher at 39.0% than faecal samples (32.7%). The percent of samples containing isolates with distinct virulotypes was similar across the different sample types (faecal vs enrichment broth). A detailed breakdown of virulotype prevalence is provided in Table 8.

Of the 310 samples, the majority (247/79.7%) contained a single virulotype (different *stx*$_1$, *stx*$_2$ and eae gene combinations) while 62 (20%) contained multiple virulotypes: 5 samples had 4 different virulotypes, 8 samples had 3 different virulotypes and 49 samples had 2 different virulotypes (Appendix 8.4 - number of isolates with distinct virulence profiles). This resulted in a total of 387 virulotypes comprising *stx*$_2$ (14.5% prevalence), *stx*$_1$ (9.1% prevalence), *stx*$_1$, *eae* (7.7% prevalence), *stx*$_1$, *stx*$_2$ (6.6% prevalence), *stx*$_2$ *eae* (2.6% prevalence), and *stx*$_1$, *stx*$_2$, *eae* (2.0% prevalence) (Table 8).

Table 8. Number and percent of samples that contain isolates of each virulotype.

Virulotype (virulence gene profile)	Number positive enrichment broth (%)	Number positive faecal samples (%)	Total number enrichment + faeces
stx$_1$	19 (9.5)	64 (9.0)	83 (9.1)
stx$_1$, *stx*$_2$	24 (12.0)	36 (5.1)	60 (6.6)
stx$_1$, *stx*$_2$, *eae*	6 (3.0)	12 (1.7)	18 (2.0)
stx$_1$, *eae*	10 (5.0)	60 (8.5)	70 (7.7)
stx$_2$	34 (17.0)	98 (13.8)	132 (14.5)
stx$_2$, *eae*	4 (2.0)	20 (2.8)	24 (2.6)
Total samples that were *stx*-positive	78 (39.0)	232 (32.7)	310 (34.1)
Total samples containing *eae*-negative STEC	66 (33.0)	177 (24.9)	243 (26.7)
Total samples containing *eae*-positive STEC	18 (9.0)	85 (12.0)	103 (11.3)
Total number of samples tested	200	710	910

4.2 WGS analysis workflow

A total of 923 of the 959 whole genome sequences were included in WGS analyses. Sequences that failed quality control or were negative for *stx* by *in silico* methods were excluded.

4.2.1 Assessment of WGS pipeline for *stx* gene subtype determination

To establish the most relevant pipeline for *stx* determination, a subset of 103 STEC of varying O-types and *stx* subtypes (Appendix 8.5) were used to assess the performance of four silico typing tools (Methods A, B, C and D) employed in the CSIRO Galaxy workflow (Table 9). A “gold standard” gene typing method, used to verify lab performance in international external quality assurance schemes, was used to confirm the *stx* subtypes carried by each strain included in the panel. For *stx* subtyping, the *in silico* method C outperformed all other tools with 97% match to expected results. In silico methods A and C correctly called 84.5% and 83.5% of strains, while method D had the lowest correct calls at 78.6%. Additional analysis of *stx* gene class was performed on a larger set of 765 isolates using the two highest performing *in silico* methods (A and C) and revealed the opposite result, with in silico method A outperforming method C.

Table 9. Concordance between 'gold standard' (PCR) and *in silico* methods of Shiga toxin and eae gene typing of Australian STEC.

Shiga toxin gene	Number of isolates tested	Percent concordance between 'gold standard' and *in silico* methods			
		In silico (Method A)[1]	*In silico* (Method B)[2]	*In silico* (Method C)[3]	*In silico* (Method D)[4]
Subtype (1a, 1c, 1d, 2a, 2b, 2c, 2e, 2f, 2g, eae)	103	84.5%	83.3%	97.1%	78.6%
Class (1, 2, eae)	765	91.4%	Not Tested	86.8%	Not Tested

[1]Method A was performed using the Abricate tool with 90% identity and 80% coverage thresholds.
[2]Method B was performed using the SRST2 tool.
[3]Method C was performed using the PrimerSearch tool with 5% allowed mismatches.
[4]Method D was performed using the European Union Reference Laboratory pipeline with default settings.

An additional analysis for O-antigen serogroup revealed strong concordance between traditional (serological/PCR) and *in silico* serotyping, with 96.4 % concordance (390/402 isolates) (Table 10).

Table 10. Concordance between traditional typing (serological or PCR) and *in silico* genotypes of STEC isolates.

Antigen	Traditional typing (Serological or PCR)	*In silico* genotype match	*In silico* genotype discordant	Percent concordance between traditional and *in silico*
O-antigen	402	390	14	96.5%

The discordance between traditional and *in silico* results may be due to:

I. Limitations of traditional serotyping methods (i.e., lack of antisera for identifying novel antigens)
II. Absence of novel serogroup genes within *in silico* databases
III. Laboratory errors (tendency for errors when handling 1000s of samples - mislabelling, cross-contamination etc)
IV. Older methods generating less reliable results - approaches that were considered 'gold standard' 2 decades ago may be inaccurate by today's standards
V. Lack of standardisation of workflow approaches (from gDNA extraction to analysis)
VI. Control issues, inside and outside of our control (i.e., sequence service providers)

With the scope and scale of the current project, we were able to achieve 96.5% concordance for serotyping and 91.4% concordance for *stx* typing. While some *in silico* methods performed better than others, differences are likely to be the result of a range of factors that would require a much deeper analysis to solve than was possible in the current study. However, we are confident that discordance could be greatly improved in future iterations of the workflow, particularly through the global standardisation of all parts of *in silico* workflows, from preparation of DNA extracts to sequencing and downstream analysis. It will be important to ensure that any future risk schemes based on genomic information have a clear and accurate pipeline and analysis flow to ensure accuracy and reproducibility of information.

4.2.2 Shiga toxin gene profiles

In total, 45 unique *stx* gene profiles were detected among the 923 isolates included in the study, with all stx_1 and stx_2 subtypes detected. The occurrence of isolates possessing each subtype, in order of frequency, was: stx_{1a}

(67%), *stx*$_{2c}$ (29%), *stx*$_{2a}$ (18%), *stx*$_{2b}$ (13%), *stx*$_{1c}$ (10%), *stx*$_{2d}$ (4%), *stx*$_{1d}$ (1%), *stx*$_{2e}$ (1%), *stx*$_{2g}$ (1%). A breakdown of the diversity and frequency of *stx* profiles observed in this study is provided in Appendix 8.6.

4.2.3 Serogroups

Across all isolates (n=923) included in this study, a total of 98 serogroups were detected (Appendix 8.6). STEC belonging to serogroups O157 (19%), O26 (14.5%) and O111 (5.3%) (Top 3 STEC) accounted for 358/923 (38.8%) of isolates. Non-Top 3 STEC accounted for 565/923 (61.2%) isolates and comprised 93 different serogroups, with the most common being O130 (4.2), O91 (4.1%), O174 (2.8%). A breakdown of the diversity and frequency of serogroups is provided in Appendix 8.7.

4.2.4 Risk classification schemes

Two prominent international schemes (JEMRA and NACMCF) have been proposed to risk assess STEC into classes based on their estimated potential to cause disease. The JEMRA scheme classifies STEC into risk levels based on their estimated potential to cause mild (risk level 5) through to severe (risk level 1) disease. Similarly, the NACMCF scheme classifies STEC into risk levels based on the human health risk they pose from low health risk (level 5) to high health risk (level 1). Here we report on the distribution of Australian STEC across the risk levels of each scheme. Isolates used in this analysis were recovered as part of a national survey of cattle faeces and beef trim enrichment broths and only the highest-level isolate from each sample has been used in the risk tables.

Risk schemes – faecal survey isolates

Using the FSIS definition, 3.0% of samples were classified as containing an adulterant STEC while 27.9% were shown to contain non-adulterant STEC (Table 11). In comparison, using the JEMRA risk scheme, 8.5% of samples contained STEC that were assigned to the highest risk groups (levels 1, 2 or 3) (Table 12). Isolates in these three levels are predicted to have the highest potential for severe disease. Most samples (22.4%) contained STEC that were assigned to JEMRA levels 4 or 5 which have lower potential to cause HUS but may cause diarrhoea or bloody diarrhoea. Using the NACMCF risk scheme, 0% of samples were assigned to level 1 (highest health risk), 3.0% were assigned to risk levels 2 and 3 (equivalent to the current FSIS definition for adulterants) and 27.9% were assigned to risk levels 4 and 5 (lowest health risks) (Table 13).

Table 11. FSIS level assignment – number (%) in each animal class. Highest level isolate selected from each sample.

FSIS	Beef Cattle Faeces (n=499)	Dairy Cattle Faeces (n=146)	Veal Faeces (n=65)	Total (n=710)
Adulterant	11 (2.2)	1 (0.7)	9 (13.8)	21 (3.0)
Non-adulterant	128 (25.7)	39 (26.7)	31 (47.7)	198 (27.9)

Table 12. JEMRA level assignment – number (%) in each animal class. Highest level isolate selected from each sample

JEMRA level	Virulence factors	Potential for:	Beef Cattle Faeces (n=499)	Dairy Cattle Faeces (n=146)	Veal Faeces (n=65)	Total (710)
1	*stx*$_{2a}$ AND *eae* OR *aggR*	D/BD/HUS	6 (1.2)	4 (2.7)	4 (6.2)	14 (2.0)
2	*stx*$_{2d}$	D/BD/HUS	21 (4.2)	2 (1.4)	14 (21.5)	37 (5.2)

3	*stx*$_{2c}$ AND *eae*	D/BD	4 (0.8)	1 (0.7)	4 (6.2)	9 (1.3)
4	*stx*$_{1a}$ AND *eae*	D/BD	24 (4.8)	4 (2.7)	10 (15.4)	38 (5.4)
5	Any other *stx* type	D	84 (16.8)	29 (19.9)	8 (12.3)	121 (17.0)

Table 13. NACMCF level assignment – number (%) in each animal class. Highest level isolate selected from each sample.

NACMCF Level	Virulence factors	Beef cattle faeces (n=499)	Dairy Cattle faeces (n=146)	Veal faeces (N=65)	Total (n=710)
1	*stx*$_{2a}$ & *aggR*	0	0	0	0
2	*stx* (any type) & *eae* & O157 *stx*$_{2a}$ > *stx*$_{2c}$ > *stx*$_{2a}$ + *stx*$_{1a}$ > *stx*$_{1a}$	6 (1.2)	1 (0.7)	3 (4.6)	10 (1.4)
3	stx (any type) & eae & O26, O111, O103, O121, O45, O145 *stx*$_{2a}$ > *stx*$_{2d}$ > *stx*$_{2c}$ > *stx*$_{1a}$	5 (1.0)	0	6 (9.2)	11 (1.5)
4	*stx* (any type) & *eae* & all other serotypes *stx*$_{2a}$ > *stx*$_{2d}$ > *stx*$_{2c}$ > *stx*$_{1a}$	29 (5.8)	8 (5.5)	12 (18.5)	49 (6.9)
5	stx (any type) *stx*$_{2a}$ > *stx*$_{2d}$ > *stx*$_{2c}$ > *stx*$_{1a}$	99 (19.8)	31 (21.2)	19 (29.2)	149 (21.0)

Risk schemes – beef enrichment survey

Using the FSIS definition, 5.5% of samples contained STEC that were classified as adulterants, while 33.0% contained STEC that were deemed non-adulterant (Table 14). By comparison, using the JEMRA risk scheme, 14.0% of samples were assigned to levels 1, 2 or 3 which contain STEC with the highest potential for severe disease, 24.5% were assigned to levels 4 or 5 which contain STEC that are unlikely to cause HUS but may cause diarrhoea or bloody diarrhoea (Table 15). Using the NACMCF risk scheme, 0% of samples were assigned to level 1 (highest health risk), 5.5% were assigned to risk levels 2 and 3 (equivalent to current FSIS definition for adulterants), 33.0% were assigned to risk levels 4 and 5 (lowest health risks) (Table 16).

In a hypothetical scenario, where risk management systems targeted levels 1, 2 and 3 of JEMRA and NACMCF schemes, the percentage of 'high risk STEC' in beef trim would be 5.5% for FSIS and NACMCF schemes and 14.0% for JEMRA. In this scenario, using the methods of isolation employed in the current study, industry could expect lower confirmed positives from NACMCF than JEMRA schemes.

A high proportion of STEC isolated from cattle faeces and beef trim enrichments were assigned to risk level 2 of the JEMRA scheme. Isolates belonging to risk group 2 must contain *stx*$_{2d}$ with or without *eae*, however, disease severity of *stx*$_{2d}$ *E. coli* is largely dependent on *stx*$_{2d}$ variant (20). Those that possess *stx*$_{2d}$-activatable (*stx*2d-act) toxin types produce higher levels of toxin (200X greater) and are more likely to cause severe disease than those that possess the non-activatable form of the toxin (21, 22). Similarly, anecdotal evidence exists to support the presence of virulent types of *stx*$_{2d}$-positive / eae-negative STEC, that have been associated with severe disease in Australia, however, the source of these isolates is unknown, and further characterisation is necessary to understand the true health risk posed by Stx2d subtypes from Australian red meat sources.

Table 14. FSIS level assignment – number (%) in each sample type 2022 survey. Highest level isolate selected from each sample.

FSIS level	Beef Enrichments (n=200)
Adulterant	11 (5.5)
Non-adulterant	66 (33.0)

Table 15. JEMRA level assignment – number (%) in each sample type 2022 survey. Highest level isolate selected from each sample.

JEMRA Level	Virulence factors	Potential for:	Beef Enrichments (n=200)
1	*stx*$_{2a}$ AND *eae* OR *aggR*	D/BD/HUS	3 (1.5%)
2	*stx*$_{2d}$	D/BD/HUS	22 (11%)
3	*stx*$_{2c}$ AND *eae*	D/BD	3 (1.5%)
4	*stx*$_{1a}$ AND *eae*	D/BD	4 (2.0%)
5	Any other *stx* type	D	45 (22.5%)

Table 16. NACMCF level assignment – number (%) in each sample type 2022 survey. Highest level isolate selected from each sample.

NACMCF Level	Virulence factors	Beef Enrichments (n=200)
1	*stx*$_{2a}$ & *aggR*	0
2	*stx* (any type) & *eae* & O157 *stx*$_{2a}$ > *stx*$_{2c}$ > *stx*$_{2a}$ + *stx*$_{1a}$ > *stx*$_{1a}$	5 (2.5%)
3	*stx* (any type) & *eae* & O26, O111, O103, O121, O45, O145 *stx*$_{2a}$ > *stx*$_{2d}$ > *stx*$_{2c}$ > *stx*$_{1a}$	6 (3.0%)
4	*stx* (any type) & *eae* & all other serotypes *stx*$_{2a}$ > *stx*$_{2d}$ > *stx*$_{2c}$ > *stx*$_{1a}$	3 (1.5%)
5	stx (any type) *stx*$_{2a}$ > *stx*$_{2d}$ > *stx*$_{2c}$ > *stx*$_{1a}$	63 (31.5%)

4.2.5 Additional genetic factors for informing disease potential

Australian STEC genomes were assessed for additional characteristics previously reported to play a role in pathogenesis or associated with STEC that cause disease. Due to resource constraints, these were limited to two

areas: (i) screening for insertion elements for disruption of the *stx* gene, and (ii) characterisation of isolates belonging to serogroup O91 for *stx* type and H types.

Insertion Elements

Mobile genetic elements (MGE) are characterised as genetic material capable of moving from one location to another within a genome or from one bacterium to another. MGE's such as phages, integrons, transposons, plasmids, gene cassettes and insertion elements can carry virulence and antimicrobial resistant genes that can have implications for the pathogenicity of an isolate.

Insertion elements (IS elements) are a type of MGE that are widely distributed in bacteria. They are characteristically small DNA sequences that insert at different locations within a genome. When inserted into virulence gene sequences, IS elements inactivate the gene through a process termed insertional inactivation. Several studies have described the insertional inactivation of Shiga toxin genes, including an Australian study that demonstrated insertional inactivation in an outbreak strain that caused mild disease (23) .

Using ISMapper, we were able to generate preliminary data on the presence of a small number of IS elements, previously reported to cause insertional inactivation in Shiga toxin genes. Results suggest that only a small proportion of Australian isolates are likely to possess insertional inactivation within a *stx* gene, with 8/923 (1%) isolates likely to contain insertional inactivation at the stx_2 region (Table 17). However, follow-up investigations are needed to unequivocally demonstrate insertional inactivation of Shiga toxin expression.

Table 17. Detection of insertion elements in Australia STEC.

IS element	Number of isolates	Number of isolates with IS elements (gene insertion) within an *stx* gene
IS1203v	923	8 (stx2)
IS1203	923	7 (stx2)
IS629	923	8 (stx2)
IS1203E	923	8 (stx2)

O91 characterisation

Subtypes of *eae*-negative STEC, belonging to serogroup O91, are reported to have different disease potentials. For instance, strains belonging to O91:H21 have been reported to cause severe infections in humans in the US, while the majority of those belonging to O91:H10 or O91:H14 appear to lack the capacity to cause severe disease (24). Most of the O91 STEC strains investigated in this study belong to O91 types (O91:H14) that are less likely to result in severe disease in humans, while a small number were shown to belong to O91:H21 (Table 18). Clinical strains belonging to STEC O91 have been reported in Australia (25) and follow-up investigations could be undertaken to better understand the clinical significance of animal derived O91 strains.

Table 18. Characterisation of eae-negative STEC belonging to serogroup O91.

O-antigen:H-antigen	*stx* genotype	Animal	Source	Count	Percent of total O91	Percent of survey samples (n=910)	Percent of all study isolates (n=923)
O91:H10	2a	human	unknown	1	2.6	0	0.1
O91:H14	1c, 2b	sheep	carcase	1	2.6	0	0.1
O91:H14	1a, 2b	cattle	carcase	1	2.6	0	0.1
O91:H14	1a	goat	carcase	1	2.6	0	0.1
O91:H14	1a	sheep	faeces	2	5.3	0	0.2

O91:H14	2b	sheep	faeces	3	7.9	0	0.3
O91:H14	1a, 2b	sheep	carcase	6	15.8	0	0.6
O91:H14	1a, 2b	sheep	faeces	14	36.8	0	1.4
O91:H21	2b	cattle	faeces	2	5.3	0.2	0.2
O91:H21	2a	cattle	faeces	5	13.1	0.5	0.5
O91:H21	1a, 2b	cattle	faeces	1	2.6	0	0.1
O91:H21	1a, 2a	cattle	faeces	1	2.6	0.1	0.1
Total	-	-	-	38	100	100	100

4.3 Shiga toxin expression

In addition to whole genome sequence analyses, a subset of STEC were assessed for their capacity to express or produce Shiga toxin under induced conditions, to gain additional insight into the disease potential of Australian STEC.

4.3.1 Determination of suitable Mitomycin C exposure time for toxin expression experiments

Chemicals that cause stress and cell damage, such as antimicrobials, have been shown to induce synthesis of Shiga toxin. One such chemical, Mitomycin C, is a chemotherapeutic agent proven to induce toxin production. However, the induction of toxin also induces the lytic cycle of phage which can lead to cell death. The following experiments were undertaken to: (i) assess the viability of STEC when exposed to Mitomycin C (ii) identify a Mitomycin C exposure time that leads to an intermediate live/dead population state (membrane injured cells) (ii) confirm through droplet digital PCR, that Mitomycin C exposure concentration and time leads to *stx* expression.

Viability assessment after exposure to Mitomycin C

The viability of *E. coli* O157 strain (ec242) was assessed following exposure to Mitomycin C at a final concentration of 0.5 µg/ml for 7 h (Figure 1). The data shows a loss of approximately 2 log CFU/mL after 6 & 7 h, with the steepest loss in viability occurring between 2 and 4 h.

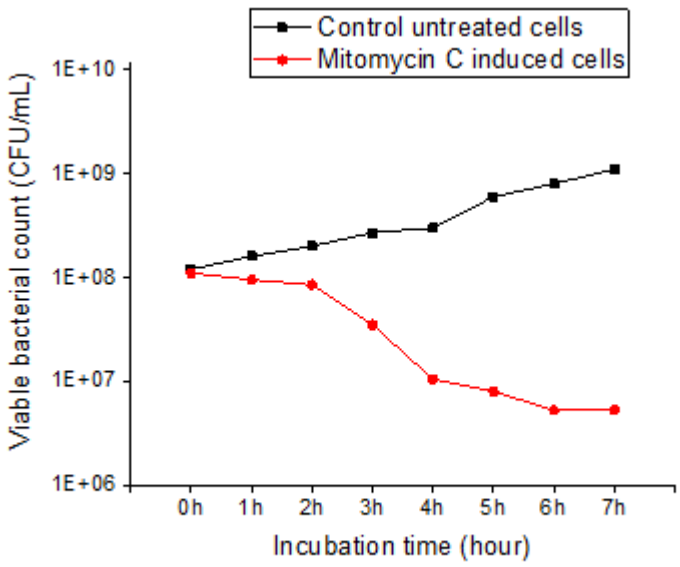

Figure 1. Viability of E. coli cells after Mitomycin C treatment for 7 h.

Viability assessment using cell viability staining in combination with flow cytometry analysis

Flow cytometry provides a rapid and reliable method to quantify viable cells in eukaryotic and prokaryotic cell suspensions. Figure 2 demonstrates simultaneous double-staining used in this work which could allow characterization of an intermediate state where cells show fluorescence with both dyes. Example dot plots of cells to assess the effects of different preservative treatments are shown in Figure 2. Dual parameters with two discriminated dyes was applied where TO stain all cells in FL1 and PI stain dead cells in FL3. Four different behaviours were detected such as quadrant Q1-UL (FL1-, FL3+), quadrant Q1-UR (FL1+, FL3+), quadrant Q1-LL (FL1-, FL3-) and quadrant Q1-LR (FL1+, FL3-).

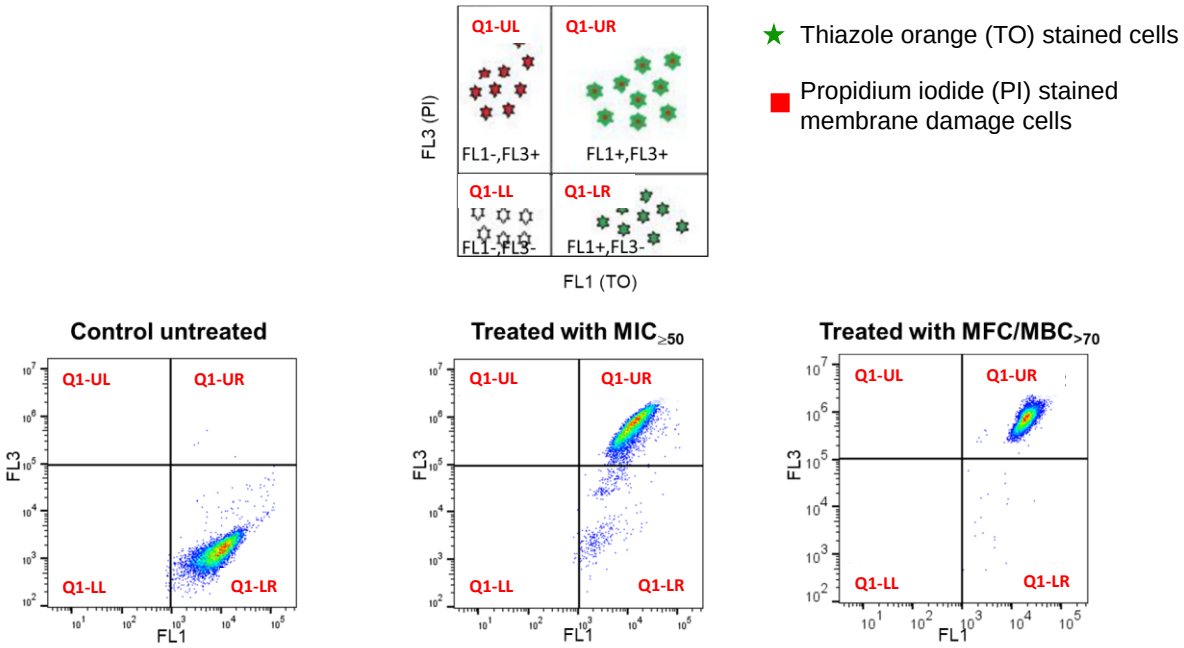

Figure 2. Dual parameter Flow cytometry dotplots. Dual parameters with two discriminated dyes: TO stained cells in FL1 and PI stained dead cells in FL3 and four different behaviours representing quadrant Q1-UL (FL1-, FL3+), quadrant Q1-UR (FL1+, FL3+), quadrant Q1-LL (FL1-, FL3-) and quadrant Q1-LR (FL1+, FL3-) are indicated.

As is typical for the co-stain TO/PI, the healthy viable microbial population demonstrated strong TO stained cells (green fluorescence) while a membrane permeabilized population showed the population shift due to strong PI stained cells (red fluorescence) in the cells. In the process of cells becoming permeabilized, the cells cluster on the two-dimensional dot plot moved in a distinctive shift from strong green and weak red fluorescence intensity (region Q1-LR) to increased green and red fluorescence intensity (Q1-UR). This pattern strongly suggests that intermediate states are occurring, which are characterized by different intracellular concentrations of TO and PI. Therefore, this initial movement of the microbial cluster can be attributed solely to higher intracellular TO concentrations and was not affected by intracellular PI levels. The results demonstrate that harvesting cells 7 h post mitomycin C treatment is suitable for downstream transcription analysis.

Transcriptomic analysis/targeted gene expression using droplet digital PCR

The relative expression level of each *stx* gene was calculated based on the changes in absolute copy number/microliter of the reaction/nanogram of cDNA and normalized using that of the housekeeping gene in response to Mitomycin treatments (e.g., control versus treated cells) after normalising by that of the housekeeping gene. The mean fold-change (Log_2 ratio) expression level for each gene between treated and the control untreated samples was calculated. The transcriptional response of these targeted genes in the Mitomycin C treated *E. coli* cells are shown in Figure 3. The upregulation of stx_1 and stx_2 in response to the Mitomycin C treatment was observed in 6 h and 7 h treated cells compared to the control untreated cells and was shown to be suitable for subsequent toxin expression analysis of STEC. Assessment of expression levels among a varied selection of STEC will commence in subsequent phases of the project.

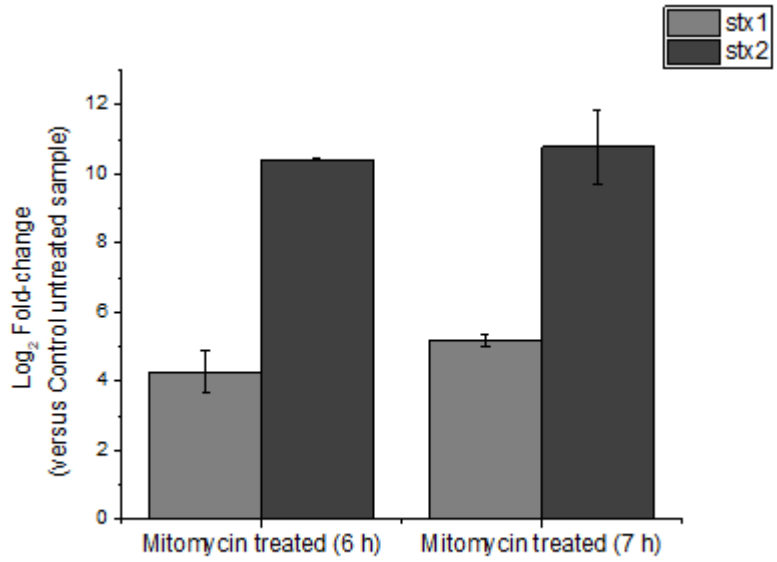

Figure 3. The abundance levels of *E. coil stx1* and *stx*2 gene transcripts after normalization using *tufA* gene as the internal reference. Fold expression levels of Mitomycin treated cells are presented relative to control untreated cells. The error bars indicate standard deviations within each group.

4.3.2 Assessment of toxin expression among a varied selection of STEC

To determine if each STEC isolate could be induced to express different levels of *stx* transcript, mid-log phase cultures (i.e., overnight cultures of each STEC isolates were diluted 1:200 into fresh LB broth and grown to an OD600 of ~0.3-0.4) were incubated at 37°C with shaking for 7 h either in the presence or absence of Mitomycin C (a final concentration of 0.5 µg/mL), and *stx* gene expression was determined by ddPCR.

The expression of stx_1 and stx_2 were measured separately in 22 STEC isolates carrying different Shiga toxin gene combinations. Levels of *stx* transcripts in induced cultures, normalized to *stx* mRNA levels with housekeeping gene, are shown in Figure 4. The most highly induced stx_1 gene was observed in Ec21 strain, where the level of induction was over 8 x Log_2 Fold change greater than that observed for control (untreated) cells. The highest induced stx_2 gene was observed in Ec2268b strain, where the level of induction was over 13 times Log_2 Fold change greater than that observed for control untreated cells and all other treated cells.

Only two, Ec1821 and Ec4029a, of the eight isolates shown to harbor both stx_1 and stx_2 genes expressed both toxin types. Most strains (59%) with stx_2 expressed minimal or no toxin, while most strains (66%) with stx_1 produced toxin. It is important to note that toxin may be produced under *in vivo* conditions. However, due to the absence of animal models that simulate human physiology, the study of toxin production and its impact on clinical outcomes is not an area amenable to experimentation.

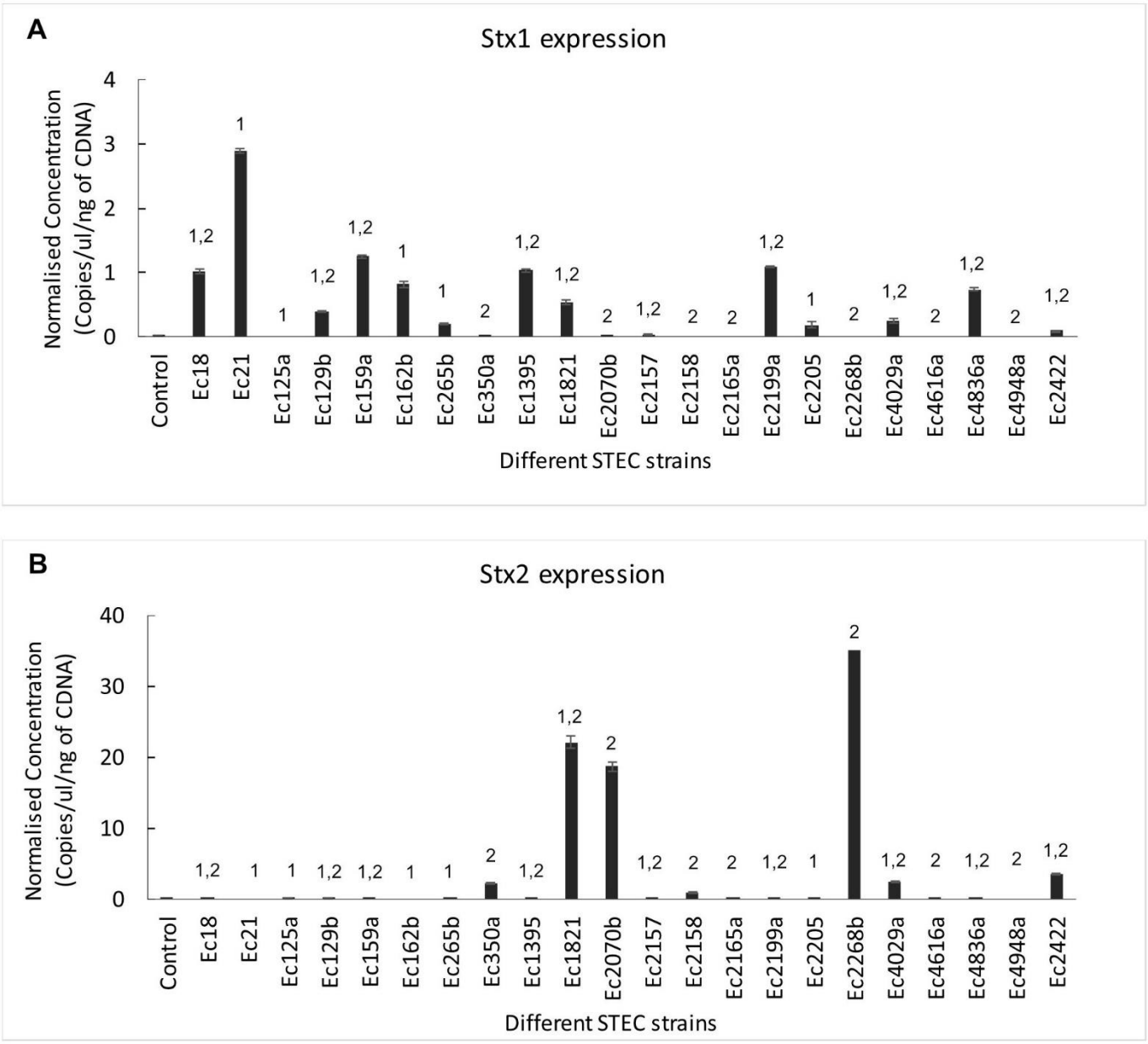

Figure 4. The abundance levels of *E. coil stx*$_1$ and *stx*$_2$ gene transcripts of different strains after normalization using *tufA* gene as the internal reference. Expression levels of 7 h-Mitomycin treated cells are presented as normalised absolute concentration (Copies/µL/ng of cDNA). The error bars indicate standard deviations within each group. Numbers above each bar indicate the *stx* genotype of each isolate tested.

4.3.3 Shiga toxin production assay

In total, 71/93 (76.3%) isolates showed alignment between phenotypic (toxin production), and genotypic (presence of toxin genes) results and the amount of toxin produced varied from very low (almost undetectable bands) to very strong (highly visible band).

Understanding the potential for an isolate to produce toxin is critical for understanding disease potential, as isolates that are incapable of producing toxin may need to be moved from high to low-risk categories. While investigating the genotypic basis for phenotype/genotype discordance was beyond the scope of this study, the data highlights that the presence of a gene does not guarantee that it will result in the production of toxin, nor does it confirm that an isolate will cause disease. Future development of WGS methods may enable the prediction of toxin production capacity, which would be a valuable addition to proposed risk assessment schemes.

5. Conclusion

Australian exporters must meet the regulatory food safety requirements of destination markets. Manufactured beef is a highly regulated category in the US and is subject to testing for seven types of STEC which are considered adulterants. Recently, expert groups have proposed that regulatory frameworks associated with STEC incorporate risk-based approaches proposed by JEMRA (EU) and NACMCF (US). These approaches incorporate genes that are indicative of the disease-causing potential of STEC. If such schemes are adopted in export jurisdictions, the Australian red meat industry is now well placed to evaluate and meet these new requirements through the capability and data developed in this project.

The results from the current project suggest that the majority of Australian STEC isolated from red meat sources are unlikely to fall within the highest risk categories of the internationally proposed risk schemes. The adoption of these schemes by overseas countries or customers is most likely to impact the industry in relation to the development and evaluation of methods to meet market access requirements for the detection, isolation, and characterisation of STEC. To our knowledge, there are no commercial test kits currently approved in Australia for screening of the priority virulence marker combinations proposed in the JEMRA and NACMCF schemes in red meat samples (e.g. Stx 1 and Stx 2 subtypes). Current methods used by industry to test for STEC are unlikely to be suitable and new approaches may be required to detect and isolate STEC belonging to the risk management level targeted by regulators, which could hypothetically be levels 1, 2 and/or 3.

This project included whole genome sequencing of historical (n=579) and contemporary (n=380) Australian isolates, providing a substantial database of diverse STEC genomes, that will act as a valuable resource for genomic investigations into disease potential and source attribution. An extension of the present work could include addressing the identified remaining knowledge gaps which include:

- Establishing the risks associated with products underrepresented in this study such as sheep, manufacturing beef and STEC collected from hides and carcasses,
- Investigating the association between the predicted disease potential of STEC and clinical disease and symptoms,
- Understanding the role of additional traits, not captured by proposed schemes, in prediction of disease potential (e.g. Stx production),
- The underlying basis for the small discordance between expected results and *in silico* results.

There are a few limitations of the present study that must be considered when interpreting the data and when applying new knowledge to direct risk management decisions. Strain selection was limited to the isolates within the CSIRO collection and while this is the largest collection of STEC from red meat in Australia, there are underrepresented sample types/animal groups that limit our ability to make firm conclusions on STEC risk for these types. While the use of WGS has been strongly recommended for STEC strain characterisation, and in many labs, it has replaced traditional typing approaches to become the 'gold standard' method, it is not without limitations as shown by the discordance in the methods comparison. Therefore, further understanding around concordance between different methods would be important to ensure robustness of testing schemes.

Overall, the extensive genomics data collected in this study, representing a major portion of the STEC isolates from the Australian red meat industry, complemented with the developed genomics workflow capability, will serve as an asset in the risk assessment and risk management of STEC in the Australian and global supply chain. The outcomes of this project provide foresight into the impact of the adoption of newly proposed risk classification schemes into the regulatory frameworks of export markets and thus reduce the market access risks for the red meat industry.

5.1 Key findings

- The majority of Australian STEC from red meat production animals were shown to fall into levels of proposed risk schemes that are associated with medium to low human health risk.
    - A percentage of STEC were assigned to risk level 2 in the JEMRA scheme due to the carriage of the *stx*$_{2d}$ subtype. Further characterisation is required to understand the disease potential of these isolates.

- The JEMRA scheme was more likely to classify STEC into higher risk categories than NACMCF or FSIS risk schemes.
    - For example, in a hypothetical scenario, where risk management systems targeted levels 1, 2 and 3 of JEMRA and NACMCF schemes, the percentage of 'high risk STEC' in beef trim would be 5.5% for FSIS and NACMCF schemes and 14.0% for JEMRA. In this scenario, using the methods of isolation employed in the current study, industry could expect lower confirmed positives from the NACMCF than the JEMRA scheme.

- Characterisation of STEC for a small number of additional traits suggests that genotypic and phenotypic data, not captured in JEMRA and NACMCF schemes, may be used to improve prediction of disease potential of Australian isolates. For example:
    - Most of the O91 STEC strains investigated in this study belong to O91 types (O91:H14) that are less likely to result in severe disease in humans, while a small number were shown to belong to O91:H21 that have been associated with disease overseas.
    - A small proportion of Australian isolates show evidence of insertional inactivation within a *stx* gene, with 8/923 (1%) isolates likely to contain insertional inactivation at the *stx*$_2$ region.
    - Discordance between phenotypic (toxin production or gene expression), and genotypic (presence of toxin genes) results was observed, and when produced/expressed, toxin levels varied considerably between isolates.

- Results from a national survey of STEC in cattle faeces (n=710) and manufacturing beef (n=200) indicate that STEC is likely to be common in these sample types (32-39% positive). Most STEC positive samples are expected to contain 1 type of STEC while 20% of STEC positive samples are expected to contain 2 or more STEC types. The number of enrichment broth samples likely to contain STEC was slightly higher at 38.5% than faecal samples (32.7%).

5.2 Benefits to industry

The key benefits to industry from this project include:

- Development of capability and generation of resource of sequence data that can be used to support industry to meet future needs associated with molecular risk-based schemes.

- Data can be used by industry to evaluate and manage the risks associated with STEC in beef.

- Data can be used to inform risk management practices that in turn minimise the reputational, trade and public health risks associated with contaminated beef entering domestic and export markets.

- This project has developed a tool kit in the form of software analysis workflows for the rapid characterisation of STEC that can be adapted or modified to evaluate new risk schemes across a range of pathogens, including *Salmonella* and STEC. The flexibility of the platform allows for new genetic targets to be added to analysis workflows as needed. The large database of sequences generated in this project acts as a valuable resource that can be interrogated rapidly to provide updated information on risk, as molecular risk approaches evolve.

- Risk classification data will allow the red meat industry to assess the disease potential of STEC that are likely to be isolated through industry testing, rapidly evaluate the impact of changes to risk classifications and regulatory requirements from overseas customers and implement management practices that best address STEC associated risks.

- Data on the risk potential of STEC could be used to inform customers on the risk level of STEC believed to be associated with Australian product and engage in discussion with regulators using evidence-based data on risk potential. In scenarios where risk is assessed to be low, the red meat industry could lobby for reduced stringency of testing.

6. Future research and recommendations

This study has developed research capability that allows for high-throughput risk profiling of bacterial isolates using whole genome sequence methods. This capability can be used to support industry to meet future market access requirements that could result from the adoption of proposed risk schemes for STEC or other foodborne pathogens by trading partners. Further work to build on this capability and address knowledge gaps should be undertaken to provide industry with tools to address future food safety based technical barriers to trade.

Priority recommendations include*:*

- Collection of isolates from underrepresented groups to better understand and manage risk

- Continue support of the genomics capability to ensure industry can meet future changes to market access requirements and better manage risk

- Partner with Australian public health laboratories to leverage information and tools established as part of a national surveillance networks - FoodTrakka and AusPathogen to investigate whether predicted disease potentials are accurate and match Australian patient symptoms.

- Improve risk characterisation through mining genomics data for attributes that inform pathogen potential and source attribution and undertake follow-up work to identify the basis of observed discrepancies between genotyping methods**.**

- STEC from sheep and manufacturing beef were underrepresented, as were samples from specific collection sites such as hides and carcases.
    - Sheep/lamb are Australia's second largest red meat export. Most of the sheep derived isolates were sourced from past studies, with minimal representation of contemporary isolates.

Undertaking a national survey of STEC in sheep, using a similar approach to this study, would generate a temporally relevant set of sheep isolates for risk characterisation.
    - Further collection and analysis of enrichment broths from commercial laboratories would generate a larger collection of manufacturing beef strains that would be highly relevant to understanding the diversity and risk levels of STEC in export product.

- Further developing capability in the detection, isolation, and characterisation of STEC will support industry to evaluate risk and meet the changing quality assurance and food safety needs of global customers of Australian red meat products.
    - This project has developed a tool kit in the form of a software analysis workflow that can be adapted or modified for the high-throughput characterisation of STEC to assess new risk schemes across a range of pathogens, including *Salmonella* and STEC. Until such time that international methods are standardised, work should be undertaken to increase confidence in results from *in silico* analyses. If internationally harmonised workflows are developed, these should be adopted and evaluated against 'gold standard' methods.
    - Results from the survey component of the project should be further interrogated to assess the recovery of STEC, along with the morphology and growth characteristics of STEC on different growth media. This is a small additional body of work that we believe will provide results that can be used to direct future survey study design and can inform industry and commercial testing laboratories of the most appropriate growth media to use for Australian STEC isolation, potentially reducing the costs associated with testing.
    - Work from this study should be disseminated through presentation at international conferences and through publication in peer reviewed journals. The publication of work will improve the transparency and credibility of the study, while providing public access to sequence data to support domestic and international efforts to improve risk schemes.
    - It is our opinion that the cost to undertake risk characterisation work is orders of magnitude less than the costs associated with public health, trade, and reputational risks, and continued support of the work will provide a valuable resource for industry to evaluate and manage risk.

- Australian public health laboratories have committed significant resources toward developing nationally harmonised genomics platforms for pathogen surveillance (AusTrakka and AusPathogen). The industry should seek to leverage knowledge from such platforms to compare predicted disease potential to human disease data, and to better understand the disease burden of STEC associated with red meat. There is some anecdotal evidence for the asymptomatic carriage of STEC in Australia and working with public health labs to better understand this population of STEC may help identify STEC that pose low health risk and inform industry risk management decisions.

- Whole genome sequencing produces an enormous amount of data that can be mined to provide insight into the pathogenic potential of bacteria. This project has focussed on a small subset of genes believed to be relevant to understanding risk. Further interrogation of WGS data is likely to identify additional factors that are important for risk assessment. WGS data should be further mined for information that can be used to predict, among other things: (i) the ability of an STEC to produce toxin, (ii) the toxin amounts produced, and (iii) the contribution of STEC from animal and food sources to the overall disease burden in Australia.

- Further understanding around concordance between different methods would be important to ensure robustness of testing schemes. For some isolates, different genotyping methods (PCR vs whole genome sequence results) and different *in silico* methods produced different results. As such, it is recommended that repeat sequencing and PCR be performed to establish the basis of such discrepancies.

- Further characterisation of the *stx*$_{2d}$ subtype carried by some Australian STEC is likely to demonstrate that this is not the *stx*$_{2d}$ variant associated with severe disease. Such an outcome would reduce the percentage of Australian STEC assigned to risk level 2 in the JEMRA scheme.

7. References

1. U.S. Department of Agriculture FSIS. Detection, isolation and identification of top seven Shiga toxin-producing *Escherichia coli* (STEC) from meat products and carcass and environmental sponges MLG 5C.03. *In* The Microbiology Laboratory Guidebook (MLG). 2023;Available at: https://www.fsis.usda.gov/sites/default/files/media_file/documents/MLG-5C.03.pdf Accessed 25 June 2023.
2. Joint FAO/WHO expert meetings on microbiological risk assessment (JEMRA). Shiga toxin-producing *Escherichia coli* (STEC) and food: attribution, characterisation and monitoring 2018.
3. National Advisory Committee On Microbiological Criteria For Foods. Response to Questions Posed by the Food and Drug Administration Regarding Virulence Factors and Attributes that Define Foodborne Shiga Toxin-Producing Escherichia coli (STEC) as Severe Human Pathogens (dagger). J Food Prot. 2019;82(5):724-67.
4. Koutsoumanis K, Allende A, Alvarez-Ordonez A, Bover-Cid S, Chemaly M, Davies R, et al. Pathogenicity assessment of Shiga toxin-producing Escherichia coli (STEC) and the public health risk posed by contamination of food with STEC. Efsa Journal. 2020;18(1).
5. International Organization for Standardization. Microbiology of food and animal feed - Real-time polymerase chain reaction (PCR)-based method for the detection of food-borne pathogens - Horizontal method for the detection of Shiga Toxin-producing *Escherichia coli* (STEC) and the determination of O157, O111, O26, O103 and O104 serogroups. 2012;ISO/TS 13136.
6. Holmes A, Dallman TJ, Shabaan S, Hanson M, Allison L. Validation of Whole-Genome Sequencing for Identification and Characterization of Shiga Toxin-Producing Escherichia coli To Produce Standardized Data To Enable Data Sharing. J Clin Microbiol. 2018;56(3).
7. EU Reference Laboratory for *E. coli*. Detection of Enteroaggregative Escherichia coli in food by Real Time PCR amplification of the aggR and aaiC genes. EU RL_Method_05_Rev 1. 2013.
8. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018;34(17):i884-i90.
9. Ewels P, Magnusson M, Lundin S, Kaller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics. 2016;32(19):3047-8.
10. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012;19(5):455-77.
11. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. Bioinformatics. 2013;29(8):1072-5.
12. Ingle DJ, Valcanis M, Kuzevski A, Tauschek M, Inouye M, Stinear T, et al. In silico serotyping of E. coli from short read data identifies limited novel O-loci but extensive diversity of O:H serotype combinations within and between pathogenic lineages. Microb Genomics. 2016;2(7).
13. Knijin A, Michelacci V, Massimiliano O, Morabito S. Advanced Research Infrastructure for Experimentation in genomicS (ARIES): A lustrum of Galaxy experience. BioRxiv. 2020;2020.05.14.095901.
14. Inouye M, Dashnow H, Raven LA, Schultz MB, Pope BJ, Tomita T, et al. SRST2: Rapid genomic surveillance for public health and hospital microbiology labs. Genome Med. 2014;6(11):90.
15. Rice P, Longden I, Bleasby A. EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet. 2000;16(6):276-7.
16. Hawkey J, Hamidian M, Wick RR, Edwards DJ, Billman-Jacobe H, Hall RM, et al. ISMapper: identifying transposase insertion sites in bacterial genomes from short read sequence data. BMC Genomics. 2015;16:667.
17. European Centre for Disease Prevention and Control (ECDC). External quality assurance scheme for typing of verocytotoxin-producing *E. coli* (VTEC). 2020;Available from, https://en.ssi.dk/surveillance-and-preparedness/national-reference-laboratories/escherichia-shigella-og-klebsiella.
18. Shringi S, Garcia A, Lahmers KK, Potter KA, Muthupalani S, Swennes AG, et al. Differential virulence of clinical and bovine-biased enterohemorrhagic *Escherichia coli* O157:H7 genotypes in piglet and Dutch belted rabbit models. Infection and Immunity. 2012;80(1):369-80.
19. http://www.ncbi.nlm.nih.gov/tools/primer-blast. [

20. Bielaszewska M, Friedrich AW, Aldick T, Schurk-Bulgrin R, Karch H. Shiga toxin activatable by intestinal mucus in Escherichia coli isolated from humans: predictor for a severe clinical outcome. Clin Infect Dis. 2006;43(9):1160-7.
21. Melton-Celsa AR, Darnell SC, O'Brien AD. Activation of Shiga-like toxins by mouse and human intestinal mucus correlates with virulence of enterohemorrhagic *Escherichia coli* O91:H21 isolates in orally infected, streptomycin-treated mice. Infection and Immunity. 1996;64(5):1569-76.
22. Fuller CA, Pellino CA, Flagler MJ, Strasser JE, Weiss AA. Shiga toxin subtypes display dramatic differences in potency. Infection and Immunity. 2011;79(3):1329-37.
23. Vasant BR, Stafford RJ, Jennison AV, Bennett SM, Bell RJ, Doyle CJ, et al. Mild Illness during Outbreak of Shiga Toxin-Producing Escherichia coli O157 Infections Associated with Agricultural Show, Australia. Emerg Infect Dis. 2017;23(10):1686-9.
24. Feng PCH, Delannoy S, Lacher DW, Bosilevac JM, Fach P, Beutin L. Shiga Toxin-Producing Serogroup O91 Escherichia coli Strains Isolated from Food and Environmental Samples. Appl Environ Microbiol. 2017;83(18).
25. Ingle DJ, Goncalves da Silva A, Valcanis M, Ballard SA, Seemann T, Jennison AV, et al. Emergence and divergence of major lineages of Shiga-toxin-producing Escherichia coli in Australia. Microb Genom. 2019;5(5).

8. Appendix

8.1 Survey invitation letter

2021 Beef Meat Pathogen Survey

Project (study) overview

You are invited to take part in a research study to determine the public health significance of Shiga toxin-producing *E. coli* (STEC) in beef cattle. The study is being carried out by staff from CSIRO's Food program. The study is jointly funded by Meat and Livestock Australia (MLA) and CSIRO and is supported by the Australian Meat Processor Corporation (AMPC) and the Australian Meat Industry Council (AMIC).

STEC are a subset of *E. coli* that produce potentially lethal toxins, referred to as Shiga toxins. Of the many different types of STEC that exist, only a small percentage are known to cause illness in humans and create risk for industry. These important STEC are currently classified based on the serogroups that most frequently cause illness in humans, such as O157, O26, O111, etc. Together these form internationally recognised high-risk *E. coli* groupings, "Big 6" and "Top 7", which are considered adulterants of manufacturing beef in the USA. While serotyping has been a convenient way to risk group STEC, not all STEC within high-risk serogroups possess the same potential to cause human disease.

Through Whole Genome Sequencing (WGS) and other molecular methods, international researchers have identified genetic markers that improve on current, serogroup-based methods of predicting STEC disease potential. International scientific experts in risk assessment have used these genetic markers to propose a risk classification scheme that ranks STEC from highest to lowest potential to cause severe disease in humans.

The purpose of this study is to use WGS technologies to build knowledge around the public health risk of STEC in Australian beef. STEC will be isolated from beef cattle and characterised using WGS tools to understand how they fit within the newly proposed risk schemes and determine the proportion of Australian isolates that fall into high and low risk groups. The outcomes will be used to assess the impacts and opportunities that global developments in STEC risk assessment have on the Australian beef industry and to support policy decisions and activities related to the risk assessment of STEC. As a longer-term objective, improved risk classification schemes could lead to targeted management of high-risk pathogens and cost savings associated with changes in how contaminated meat is managed by processors.

What does participation involve?

Participation in this study will involve completing a questionnaire and collecting samples for analysis by CSIRO.

Questionnaire (attached)

The questionnaire will be used to gather information on animal production type and slaughter volumes that will help guide the development of a robust sampling plan. The questionnaire is attached to the email and can be completed at a time that is convenient to you. It will take approximately 5 min to complete.

Sample collection (sample data sheet attached)

You will be required to provide a staff member to collect faecal samples from the gut of randomly selected animals post evisceration on allocated sampling days.

- For most establishments, samples will not exceed 50.
- Samples will be collected between April 2021 and March 2022.
- The number of samples and sampling days allocated to each plant will depend on the production output of the plant (large plants will be expected to collect more than small plants) and the total number of participants.
- You will be required to complete a sampling data sheet for each animal sampled.
- Samples will need to be stored at 4°C until they are ready to be sent to CSIRO.
- All sampling materials, parcels for returning samples and freight costs will be met by CSIRO.

How can you participate?

Please confirm your participation by taking a few minutes to complete the brief one page questionnaire that accompanies this letter.

Please return completed questionnaires by 31st March, 2021.

Risks and benefits

Aside from giving up your time, there are no foreseeable risks associated with participating in this study.

Withdrawal from the research study

Participation in this study is completely voluntary and you do not have to take part. Your decision whether to participate will not affect your current or future relationship with the CSIRO researchers or MLA. You are free to stop participation in the survey at any time. In this case, any data will be erased and the information you have provided will not be included in the study results. You may withdraw from this study at any time up until publication of the final outputs.

Confidentiality

All information provided by you will be treated confidentially. All data collected in this study will be coded and subsequently analysed and reported in such a way that responses will not be able to be linked to any individual establishment. Any data collected as part of this study will be securely stored as per CSIRO's Recordkeeping Procedure.

How will my information be handled?

Privacy Statement:

Your personal information is protected by the *Privacy Act 1988* (Cth) (Privacy Act) and CSIRO will handle your personal information in accordance with the Privacy Act and the NH&MRC National Statement on Ethical Conduct in Human Research (2007, updated 2018).

Your personal information, including your name, contact details, company/establishment/plant number, and your opinions and answers to the questionnaire and sample data sheet, is being collected by CSIRO for the purposes of conducting the study and related scientific research. If you do not provide your personal information, you will be unable to participate in the study.

Results from the study will be de-identified and published/presented in a variety of forums, including the preparation of MLA reports, and communication materials such as scientific papers and seminar presentations. Establishments will not be identified nor will their individual data be discussed in public without seeking prior approval. Approval will also be sought from MLA prior to publishing material other than MLA reports.

De-identified information may be shared with other researchers for the purposes of verifying published results or advancing other research on this topic.

The CSIRO Privacy Policy available at https://www.csiro.au/en/About/Access-to-information/Privacy outlines how your personal information will be handled, including details about how you can seek access or correction of the personal information we hold about you, how you can lodge a complaint about a breach of the Australian Privacy Principles (APPs) and how CSIRO will deal with the complaint. If you require further information on how your personal information will be handled, please contact privacy@csiro.au.

Ethics clearance and contacts

This study has been approved by CSIRO's Social Science Human Research Ethics Committee in accordance with the National Statement on Ethical Conduct in Human Research 2007 (Updated 2018). If you have any questions concerning your participation in the study please contact the researchers via their contact details provided. Alternatively, any concerns or complaints about the conduct of this study can be raised with the Executive Manager of Social Responsibility and Ethics on +61 7 3833 5693 or by email at csshrec@csiro.au.

Glen Mellor
Project Leader
CSIRO Agriculture and Food
39 Kessels Road
Coopers Plains QLD 4108
Ph: (07) 3214 2038
Email: glen.mellor@csiro.au

Dr Ian Jenson
Manager
Market Access Science and Technology
Level 1, 40 Mount Street
North Sydney NSW 2060
Ph: +61 (2) 9463 9264
Email: ijenson@mla.com.au

Consent:

By participating in this study and by completing and returning the questionnaire and data sampling sheet to CSIRO, you agree to the collection, use and disclosure of your personal information, in the ways described in this document.

Thank you for taking the time to help with this research study. Please keep this sheet for your information.

8.2 Survey Questionnaire

2021 BEEF MEAT PATHOGEN SURVEY: ESTABLISHMENT PARTICIPATION QUESTIONNAIRE

Company /establishment number: _____

Contact name & position (for arranging sampling): _____

Email: _____

Phone number: _____ Fax number (if applicable): _____

Would you be willing to support beef meat exports by participating in this survey?	Yes ☐ | No ☐

We will be requesting samples on two occasions between April and Dec 2021. When would you prefer to collect and transport samples?	2021: Apr ☐ | May ☐ | Jun ☐| July ☐| Aug ☐| Sep ☐| Oct ☐| Nov ☐ |Dec ☐ 2022: Jan ☐ | Feb ☐ | Mar ☐ Any of the above ☐ Suggest other suitable dates:

To help us plan the survey please indicate in the table below the animal classes being processed.

Beef category	Processed (Yes/No)	Average monthly kill	Slaughter months (if relevant)*
Beef Cattle			
Dairy Cattle			
Veal			
Do you slaughter species other than cattle (if so, what)?			
Comments (if relevant)			

*If slaughter volumes are confined to particular times of the year then please indicate here

Which of the following information fields can you provide for each sampled animal?	Response
Animal type (Beef Cattle; Dairy Cattle; Veal):	Yes ☐ | No ☐
Carcase weight (Kg):	Yes ☐ | No ☐
Feed Type/Production system (feed lot; grain-assisted grass-fed; grass-fed; unknown):	Yes ☐ | No ☐
Lot size:	Yes ☐ | No ☐
Lot number:	Yes ☐ | No ☐
Origin of animal (state /region or postcode):	Yes ☐ | No ☐
Animal ID/NLIS ID Tag number:	Yes ☐ | No ☐

Please provide any additional comments below

_____

Please return completed questionnaires via post (address in header) or email (glen.mellor@csiro.au) by 31st March, 2021. Thank you for this information, further information on the project will be sent to participating abattoirs early April 2021.

Survey data collection sheet

Sample Data Sheet (2021 Slaughter Cattle Faecal Pathogen Survey).
Please complete for all samples collected. **Please use definitions on page 2 to categorize animal types.**

COMPANY/ ESTABLISHMENT NUMBER: _____ NAME OF SAMPLER: _____

	Compulsory fields							**Optional fields** (the greatest benefit to plants and to industry will be achieved by incorporating as much information on samples as possible)		
Sample number	**Date collected**	**Time collected** (approx.)	**Lot number**	**Animal type** Beef - *B*; Dairy - *D*; Veal - *V*	**Feed Type /Production system** Grain-fed *Grain*; Grain-assisted grass-fed *GA*; Grass-fed *Grass*; unknown *U*	**Origin of animal** (State /region or postcode)	**Carcass weight (Kg)**	**Lot size**	**Animal ID/NLIS ID Tag**	**Notes**
example	23.10.17	13:00	321	B	Grain	NSW/Riverina	275	200	QGBI0288DBM02441	State/region is best estimate
1										
2										
3										
4										
5										
6										
7										
8										
9										
10										

8.3 STEC growth characteristics

Pre-survey experiments were performed to assess the ability of isolation media: mRBA, CHROMagar STEC and TBX to support the growth of 26 STEC of varying serogroups and virulotypes (virulence gene profiles). STEC were grown for 18±2 h in BPW, serially diluted and spiral plated onto each media type. Spiral plates were incubated at 37 ± 1°C for 24 ± 2 h after which growth characteristics and colony colour were visually assessed and recorded. For each isolate, the degree of growth on selective media was graded relative to growth on non-selective media (TBX), which served as the control medium. The degree of growth of selective media was qualitatively assessed and reported as: no growth (-), weak growth (+), medium growth (++) or strong growth (+++). TBX agar supported the growth of all 26 isolates, mRBA supported the growth of 23 isolates (88%) and CHROMagar STEC supported the growth of 7 (27%) of isolates (Table 13). The degree of growth varied on mRBA and CHROMagar, ranging from weak (+) to strong (+++) when compared to growth on TBX. Guided by these results, a decision was made to use a combination of selective (mRBA/CHROMagar) and non-selective (TBX) agar in the STEC survey isolation approach. This isolation approach is largely consistent with recommendations provided in the draft ISO/TS 13136 (unpublished). However, for ease of workflow, 45 isolates were screened per sample (as opposed to 50 recommended in ISO/TS 13136) in the following ratio: TBX (n=25), mRBA (n=10) or CHROMagar STEC (n=10).

Table 19. Growth characteristics of STEC strains on selective (CHROMagar STEC and mRBA) and non-selective isolation media.

Serogroup	Strain ID	Virulotype (virulence gene profile)	CHROMagar		mRBA		TBX	
			Growth[a]	Colour	Growth	Colour	Growth	Colour
O10	2118a	1a eae	+++	purple	+++	purple	+++	green
O103	3221a	2a eae	+	purple	++	purple	+++	green
O108	162b	1a	-		+++	purple	+++	green
O128	1624c	1c 2b	-		+	purple	+++	green
O130	3233a	2a	-		+++	purple	+++	green
O136	424a	1a	-		++	white	+++	green
O142	3164a	1d	-		+++	white	+++	white
O157	3419a	2c eae	++	purple	+++	grey	+++	white
O163	2615a	2a	+++	purple	+++	purple	+++	green
O165	309b	1a 2c eae	-		++	purple	+++	green
O166	3234a	2a	-		++	purple	+++	green
O168	129b	1a 2g	-		+++	purple	+++	green
O17/O77	3241a	2d	-		+++	purple	+++	green
O171	2154a	2c	-		-		+++	green
O174	18a	1c 2b	-		+++	purple	+++	green
O175	2199a	1a 2a	-		+++	purple	+++	green
O177	2070b	2c eae	+++	purple	+++	white	+++	white
O3	1256a	1a	-		+++	purple	+++	green
O5	265b	1c	-		+++	purple	+++	green
O6	94a	1c	-		+	white	+++	green
O76	2402a	1c	-		-		+++	green
O8	395a	1a 2d	-		+	purple	+++	green
O8/O96	1395a	1a 2b	-		-		+++	green
O84	21a	1a eae	+++	purple	+++	grey	+++	white
O91	206a	1c 2b	-		+++	purple	+++	green
O98	2158a	2a eae	+++	purple	+++	purple	+++	green

[a]Estimated degree of growth for each serogroup: no growth (-), weak growth (+), medium growth (++), strong growth (+++).

8.4 Representative number of isolates with distinct virulence profiles (based on *stx*$_1$, *stx*$_2$ and *eae* typing) and total number of positive picks from each sample

Sample ID	Number of distinct isolates (isolates with unique stx profiles)	Total number of positive picks from the 45 isolates tested per sample	Virulence profiles of isolates recovered
184	4	13	2 eae (3); 1 eae (2); 1 (7); 2 (1)
240	4	28	1 (17); 1 2 eae (1); 1 eae (3); 2 (7)
247	4	24	1 2 eae (1); 1 eae (5); 2 (16); 2 eae (2)
364	4	20	1 (2); 1 2 (10); 1 2 eae (1); 2 eae (7)
543	4	22	1 2 eae (1); 1 eae (12); 2 (8); 2 eae (1)
145	3	5	1 2 (2); 1 2 eae (1); 1 eae (2)
302	3	20	1 2 (7); 1 eae (3); 2 (10)
328	3	20	1 (5); 1 eae (3); 2 (12)
536	3	10	1 2 (1); 1 2 eae (1); 1 eae (8)
540	3	16	1 (8); 1 eae (1); 2 (7)
598	3	3	1 (1); 1 eae (1); 2 (1)
SYM038	3	4	1 (1); 1 2 eae (2); 2 (1)
SYM176	3	14	1 (2); 2 (11); 2 eae (1)
1	2	32	1 2 eae (10); 1 2 (22)
9	2	32	2 eae (1); 1 2 (31)
20	2	7	1 (5); 2 (2)
89	2	14	1 (12); 1 2 (2)
90	2	14	1 2 (13); 2 (1)
115	2	28	1 (1); 1 2 (27)
146	2	6	1 (3); 2 eae (3)
191	2	12	1 (3); 2 eae (9)
213	2	3	1 2 (2); 2 (1)
219	2	5	1 eae (1); 2 (4)
237	2	13	1 (1); 1 eae (12)
245	2	6	1 (4); 1 2 (2)
250	2	5	1 (4); 1 eae (1)
252	2	2	1 (1); 1 2 (1)
266	2	3	1 (1); 1 eae (2)
342	2	21	1 2 (1); 1 2 eae (20)
372	2	11	1 (1); 2 (10)
389	2	2	1 eae (1); 2 (1)
413	2	8	1 (7); 1 eae (3)
493	2	3	1 2 (2); 2 (1)
511	2	9	1 2 eae (1); 1 eae (8)
530	2	18	2 (15)
535	2	13	1 2 (3); 1 eae (10)
541	2	4	1 2 (2); 2 eae (2)
544	2	8	1 (6); 2 (2)
603	2	34	1 2 (32); 2 eae (2)
605	2	2	1 (1); 1 eae (1)
655	2	9	1 (7); 2 (2)

665	2	10	1 (2); 2 eae (8)
668	2	2	1 (1); 2 (1)
693	2	13	1 eae (11); 1 2 eae (2)
701	2	2	2 eae (1); 2 (1)
707	2	3	1 eae (2); 2 (1)
708	2	2	1 2 eae (1); 2 (1)
709	2	5	1 (4); 2 (1)
710	2	8	1 (1); 2 eae (7)
SYM034	2	14	1 (13); 1 2 (1)
SYM065	2	4	1 2 eae (2); 2 (2)
SYM075	2	21	1 (18); 1 2 (3)
SYM077	2	35	1 (25); 1 2 (10)
SYM115	2	4	1 2 (3); 2 (1)
SYM116	2	6	1 2 (5); 2 (1)
SYM117	2	7	1 2 (6); 2 (1)
SYM130	2	13	1 2 eae (2); 2 (11)
SYM147	2	3	1 (2); 2 (1)
SYM158	2	3	1 eae (1); 2 (2)
SYM167	2	16	1 (6); 2 (10)
SYM177	2	6	1 2 (2); 1 2 eae (4)
SYM194	2	3	1 2 (1); 2 (2)
2	1	1	2 eae (1)
3	1	33	1 eae (33)
4	1	26	1 (26)
5	1	2	2 (2)
11	1	5	1 (5)
12	1	1	2 (1)
15	1	8	1 (8)
18	1	1	1 (1)
23	1	1	1 (1)
27	1	3	1 (3)
28	1	3	2 (3)
36	1	1	2 (1)
38	1	8	2 (8)
43	1	4	1 eae (4)
45	1	1	1 eae (1)
51	1	1	1 2 eae (1)
58	1	1	1 2 (1)
87	1	2	1 2 (2)
88	1	19	1 2 (19)
94	1	3	1 2 (3)
103	1	1	2 (1)
106	1	8	1 (8)
107	1	6	2 (6)
110	1	1	2 eae (1)
117	1	31	1 2 (31)
121	1	4	1 (4)
130	1	24	1 2 (24)

132	1	8	2 eae (8)
135	1	6	1 (6)
139	1	16	2 (16)
150	1	1	2 (1)
154	1	1	2 (1)
156	1	2	2 (2)
158	1	2	2 (2)
161	1	3	2 (3)
162	1	1	2 eae (1)
165	1	5	1 eae (5)
178	1	6	1 (6)
180	1	1	1 (1)
193	1	5	2 (5)
194	1	3	2 (3)
196	1	9	2 (9)
198	1	15	1 2 (15)
200	1	7	2 (7)
204	1	4	1 (4)
205	1	8	2 (8)
206	1	3	2 (3)
207	1	13	1 (13)
209	1	8	1 eae (8)
211	1	14	1 eae (14)
214	1	3	2 (3)
215	1	2	2 (2)
216	1	1	1 2 (1)
220	1	9	2 (9)
223	1	1	2 (1)
224	1	10	2 (10)
232	1	13	1 eae (13)
233	1	3	1 (3)
236	1	4	1 eae (4)
239	1	2	1 eae (2)
243	1	3	1 eae (3)
248	1	1	1 eae (1)
249	1	21	1 (21)
253	1	10	1 2 (10)
254	1	3	1 eae (3)
256	1	10	1 eae (10)
257	1	1	1 eae (1)
258	1	11	1 2 (11)
259	1	12	2 (12)
260	1	3	1 (3)
265	1	2	2 (2)
269	1	4	1 2 (4)
273	1	13	2 (13)
274	1	1	2 (1)
281	1	4	1 eae (4)

283	1	9	1 eae (9)
284	1	1	1 eae (1)
294	1	5	2 (5)
301	1	2	2 (2)
305	1	1	2 (1)
311	1	3	2 (3)
312	1	11	2 (11)
314	1	8	1 eae (8)
319	1	1	1 2 (1)
320	1	17	2 (17)
323	1	1	2 (1)
324	1	8	2 (8)
330	1	1	1 (1)
340	1	6	2 (6)
341	1	1	1 eae (1)
343	1	1	2 (1)
355	1	1	1 (1)
363	1	10	1 (10)
382	1	12	2 (12)
385	1	1	1 eae (1)
386	1	1	2 (1)
388	1	2	1 (2)
391	1	17	2 (17)
392	1	1	1 (1)
393	1	5	2 (5)
395	1	1	1 eae (1)
399	1	1	1 (1)
401	1	1	2 (1)
402	1	6	2 (6)
406	1	3	1 (3)
409	1	7	1 eae (7)
410	1	2	1 (2)
411	1	1	1 eae (1)
414	1	3	1 eae (3)
415	1	1	1 eae (1)
416	1	7	1 eae (7)
420	1	1	1 (1)
421	1	3	2 (3)
431	1	1	1 2 eae (1)
436	1	1	1 eae (1)
438	1	1	2 (1)
439	1	2	2 (2)
440	1	1	1 eae (1)
444	1	4	2 (4)
445	1	3	2 (3)
447	1	1	1 2 (1)
454	1	1	1 (1)
455	1	2	2 (2)

461	1	15	2 (15)
467	1	5	2 (5)
470	1	2	2 (2)
473	1	1	1 eae (1)
476	1	10	1 eae (10)
477	1	1	1 eae (1)
483	1	1	1 eae (1)
485	1	1	1 2 (1)
486	1	1	2 eae (1)
488	1	8	1 eae (8)
489	1	3	1 eae (3)
491	1	12	2 eae (12)
503	1	5	2 (5)
518	1	1	2 (1)
523	1	5	2 (5)
524	1	6	2 (6)
525	1	1	2 (1)
526	1	7	2 (7)
527	1	1	1 2 (1)
528	1	3	2 (3)
531	1	1	2 eae (1)
534	1	11	2 (11)
537	1	1	1 2 (1)
538	1	9	2 (9)
542	1	1	1 2 (2);
549	1	2	2 (2)
558	1	1	2 (1)
559	1	5	1 2 (5)
561	1	1	1 (1)
566	1	1	1 (1)
569	1	1	1 (1)
576	1	4	2 (4)
578	1	8	2 (8)
579	1	1	1 (1)
580	1	2	2 (2)
581	1	3	1 eae (3)
596	1	1	1 (1)
599	1	2	2 (2)
602	1	2	2 (2)
606	1	2	1 (2)
607	1	1	1 2 eae (1)
619	1	1	2 (1)
624	1	2	2 (2)
625	1	2	2 (2)
626	1	2	1 eae (2)
629	1	2	2 (2)
656	1	5	1 (5)
659	1	2	1 (2)

660	1	2	1 2 (2)
661	1	2	1 (2)
677	1	11	1 eae (11)
678	1	10	1 (10)
681	1	16	2 (16)
691	1	1	1 eae (1)
692	1	1	1 (1)
695	1	1	2 eae (1)
696	1	4	1 eae (4)
699	1	1	2 (1)
704	1	23	1 (23)
705	1	2	1 eae (2)
706	1	11	1 2 eae (11)
SYM002	1	1	2 (1)
SYM005	1	1	2 eae (1)
SYM008	1	1	1 eae (1)
SYM028	1	4	1 (4)
SYM032	1	1	2 (1)
SYM033	1	15	1 (15)
SYM035	1	16	1 (16)
SYM041	1	1	1 2 eae (1)
SYM043	1	5	1 2 (5)
SYM044	1	5	2 (5)
SYM045	1	36	1 2 (36)
SYM046	1	13	2 (13)
SYM047	1	7	2 (7)
SYM048	1	12	1 2 (12)
SYM052	1	2	2 (2)
SYM053	1	1	1 2 (1)
SYM054	1	3	1 (3)
SYM058	1	1	1 (1)
SYM059	1	2	1 eae (2)
SYM063	1	4	2 (4)
SYM064	1	7	1 2 (7)
SYM066	1	5	1 2 (5)
SYM068	1	11	2 (11)
SYM071	1	11	2 (11)
SYM074	1	22	2 (22)
SYM076	1	2	1 (2)
SYM080	1	11	1 eae (11)
SYM083	1	3	1 (3)
SYM093	1	1	1 eae (1)
SYM096	1	3	1 eae (3)
SYM099	1	1	1 (1)
SYM100	1	2	1 eae (2)
SYM101	1	4	1 eae (4)
SYM108	1	1	1 (1)
SYM109	1	2	2 (2)

SYM110	1	1	1 eae (1)
SYM112	1	9	1 2 (9)
SYM114	1	1	1 (1)
SYM118	1	5	1 2 (5)
SYM120	1	2	2 (2)
SYM121	1	16	1 2 (16)
SYM122	1	16	2 (16)
SYM123	1	9	1 2 (9)
SYM131	1	35	1 2 (35)
SYM132	1	4	1 2 (4)
SYM133	1	12	2 (12)
SYM136	1	4	2 (4)
SYM148	1	1	1 (1)
SYM157	1	12	1 (12)
SYM161	1	2	1 2 (2)
SYM163	1	3	2 (3)
SYM165	1	2	1 2 (2)
SYM166	1	8	2 (8)
SYM175	1	2	2 (2)
SYM178	1	13	1 2 eae (13)
SYM179	1	30	2 (30)
SYM180	1	20	2 (20)
SYM188	1	14	2 (14)
SYM189	1	1	1 2 (1)
SYM192	1	3	1 eae (3)
SYM193	1	2	2 (2)
SYM198	1	3	1 2 (2)
SYM199	1	1	2 (1)
Total	387	2,071	n/a

8.5 Relative performance of *in silico* typing methods against the EQA *stx* subtyping "gold standard" method

Isolate #	EQA stx subtype "gold standard"	Number and percent of isolates with correct virulence profile in each in silico method			
		Abricate 95% identity, 90% coverage	SRST2	PrimerSearch (5% mismatch)	EURL (default)
18	1c 2b	1c 2b[a]	1c 2b	1c 2b	1c 2b
21	1a	1a	1a	1a	1a
61	1c 2b	1c 2b	1c 2b	1c 2b	1c 2b
73	1a 1c 2b	1a 1c 2b	1a 1c	1a 1c 2b	1a 1c 2b
94	1c	1c	1c	1c	1c

100	1a 1c 2b	1a 1c 2b	1a 1c 2b	1a 1c 2b	1a 1c 2b
125	2d	2d	2d	2d	2d
130	2a	2a 2d	2a	2a	2c
131	2d	2d	2d	2d	2d
156	2a	2a 2c	2a	2a	2c
159	1a 2c	1a 2a 2d	1a 2a 2c 2d	1a 2c	1a 2d
206	1c 2b	1c 2b	1c 2b	1c 2b	1c 2b
259	2b 2d	2b 2d	2b 2d	2b 2d	2
262	1c 2b	1c 2b	1c 2b	1c 2b	1c 2b
287	2b 2d	2b 2d	2d	2b 2d	2
333	1a 2a	1a 2a	1a 2a	1a 2a	1a 2a
338	2a 2c	2c	2a 2c 2d	2a 2c	2
350	2b	2b	2b	2b	2
395	1a 2d	1a 2d	1a 2d	1a 2d	1a 2d
412	1a 2d	1a 2d	1a 2d	1a 2d	1a 2d
424	1a	1a	1a	1a	1a
428	1c	1c	1c	1c	1c
782	2b 2d	2b 2c 2d	2b	2b 2d	2
1054	1c	1c	1c	1c	1c
1197	1a 2b	1a 2b	1a 2b	1a 2b	1a 2b
1256	1a	1a	1a	1a	1a
1261	1a 2c	1a 2c 2d	1a 2c	1a 2c	1a 2c
1395	1a 2b	1a 2b	1a 2b	1a 2b	1a 2b
1616	1a 2b	1a 2b	1a 2b	1a 2b	1a 2b
1730	1d	1d	1d	1d	1d
1821	1a 2a	1a 2a	1a 2a	1a 2a	1a 2a
1823	1a	1a	1a	1a	1a
1922	2d	2a 2d	2c 2d	2d	2
2118	1a	1a	1a	1a	1a
2138	2b	2b	2b	2b	2b
2143	2d	2c 2d	2a 2c 2d	2d	2d
2152	2a 2d	2a 2d	2a 2c 2d	2a 2d	2
2154	2c	2c	2c	2c	2c
2157	1a 2d	1a 2d	1a 2c 2d	1a 2d	1a 2d
2158	2a	2a	2a	2a	2a
2165	2a	2a	2a	2a	2a
2174	2a 2c	2a 2c	2a 2c	2a 2c	2a 2c
2199	1a 2a	1a 2a	1a 2a	1a 2a	1a 2a
2205	1d	1d	1d	1d	1d
2253	1a 2a 2c	1a 2c	1a 2a 2c	1a 2a 2c	1a 2a 2c
2268	2a 2c	2a	2a 2c	2a 2c	2
2384	1a	1a	1a	1a	1a
2402	1c	1c	1c	1c	1c
2915	2c	2c	2c	2c	2c
3212	2a 2d	2a 2d	2a	2d	2
3220	2a	2a	2a	2a	2a
3221	2a	2a	2a	2a	2a
3225	1a 2a	1a 2a	1a 2a	1a 2a	1a 2a

3233	2a	2a	2a	2a	2a
3234	2a	2a	2a	2a	2a
3241	2d	2d	2d	2d	2d
3419	2c	2c	2c	2c	2c
3456	1a	1a	1a	1a	1a
3517	1a	1a	1a	1a	1a
3526	1a	1a	1a	1a	1a
3532	2c	2c	2c	2c	2c
3546	1a 2a	1a 2a	1a 2a	1a 2a	1a 2a
3550	1a 2d	1a 2d	1a 2d	1a 2d	1a 2d
3572	1a 2a	1a 2a	1a 2a	1a 2a	1a 2a
3588	2a	2a	2a	2a	2a
3687	1a 2c	1a 2c	1a 2c	1a 2c	1a 2c
3707	2a 2d 2g	2d 2e 2g	2g	2a 2g	2a 2g
3709	1a	1a	1a	1a	1a
3714	1a 2c	1a 2c	1a 2c	1a 2c	1a 2c
3733	2c	2c	2c	2c	2c
3792	1d	1d	1d	1d	1d
3800	1a 2a	1a 2a	1a 2a	1a 2a	1a 2
3831	2b	2b	2b	2b	2b
3925	2a 2d	2a	2a	2a	2
3982	2a 2d	2a 2d	2a	2a 2d	2
4140	1a 2g	1a 2g	1a 2g	1a 2g	1a 2g
4182	1a 2a	1a 2a	1a 2a	1a 2a	1a 2a
4230	1a 2c	1a 2c 2d	1a 2c	1a 2c	1a 2c
4380	1a	1a	1a	1a	1a
4586	2d	2a 2d	2d	2d	2d
4616	2e	2e	2e	2e	2e
4742	1a 2a	1a 2a	1a 2a	1a 2a	1a 2a
4752	1a 2b	1a 2b	1a 2b	1a 2b	1a 2b
4761	1a 2b	1a 2b	1a 2b	1a 2b	1a 2b
4786	1a 1c 2b	1a 1c 2b	1a 1c 2b	1a 1c 2b	1a 1c 2b
4789	2b	2b	2b	2b	2
4790	2b	2b	2b	2b	2b
4836	1c 2d	1c 2d	1c 2c 2d	1c 2d	1c 2b 2d
4885	2e	2e	2e	2e	2e
4948	2g	2g	2g	2g	2g
5054	2a	2a	2a	2a	2a
5108	1a 2c	1a 2c	1a 2c	1a 2c	1a 2c
129b	1a 2g	1a 2g	1a 2g	1a 2g	1a 2g
1624c	1c 2b	1c 2b	1c 2b	1c 2b	1c 2b
162b	1a	1a	1a	1a	1a
173b	2b 2d	2b 2d	2b 2d	2b 2d	2
2070b	2c	2c	2c 2d	2c	2c
2268b	2c	2c	2c	2c	2c
265b	1c	1c	1c	1c	1c
288c	2b 2d	2b 2d	2d	2b 2d	2
309b	1a 2a 2c	1a 2c	1a 2c	1a 2a 2c	1a 2c

3566d	2c 2d	2c	2c 2d	2c 2d	2
984b	1a 1c 2b	1a 1c	1a 1c 2b	1a 1c 2b	1a 1c 2b
Total correct calls		87	86	100	81
Total correct calls %		84.5	83.5	97.1	78.6

[a]Green shaded cells indicate agreement with the EQA "gold standard" method of subtyping.

8.6 Diversity of STEC virulence gene profiles

Virulence gene profile	No. of isolates
1a, eae	216
1a, 2c, eae	118
1a	78
2a	64
1a, 2a	56
2c, eae	52
1a, 2a, eae	47
1c, 2b	28
2a, eae	27
2d	27
1a, 2d	26
1c	25
1a, 2b	23
2g	22
2b	17
2c	13
1d	10
1a, 2g	7
2a, 2d	6
2b, 2d	6
1a, 2c	5
2a, 2g	5
2d, 2e, 2g	5
1a, 2a, 2d	4
1a, 1c, 2b	3
1a, 2a, 2c	3
1a, 2d, eae	3
2a, 2c	3
1a, 2c, 2d, eae	2
1a, 2d, 2e, 2g	2
1c, 2d	2
1d, 2a	2
2d, 2e	2
2e	2
2f, eae	2
1a, 1c	1
1a, 1d, 2c	1
1a, 2a, 2g	1
1a, 2b, 2d	1
1a, 2e, eae	1
1d, 2g	1
2b, 2c, 2d	1

2c, 2d	1
2c, 2d, eae	1
2d, eae	1
All	923

8.7 Diversity of STEC serogroups

Serogroup	count	percent
O157	175	19.0
O26	134	14.5
O111	49	5.3
O130	39	4.2
O91	38	4.1
O174	26	2.8
Onovel17	24	2.6
O182	23	2.5
O2	22	2.4
O84	21	2.3
O5	21	2.3
O177	19	2.1
O8	14	1.5
O168	14	1.5
Onovel21	13	1.4
O76	12	1.3
O3	12	1.3
O171	11	1.2
O113	11	1.2
O165	10	1.1
O128	10	1.1
O6	9	1.0
O112	9	1.0
O28ab	8	0.9
Onovel4	7	0.8
O185	7	0.8
O176	7	0.8
O136	7	0.8
O75	6	0.7
O179	6	0.7
OgN12	5	0.5
O22	5	0.5
O163	5	0.5
O159	5	0.5
O134, O46	5	0.5
O116	5	0.5
O103	5	0.5
Onovel7	4	0.4
O74	4	0.4
O37	4	0.4
O150	4	0.4
O121	4	0.4
O109	4	0.4
Onovel1	3	0.3

O98	3	0.3
O96	3	0.3
O87	3	0.3
O81	3	0.3
O55	3	0.3
O146	3	0.3
O142	3	0.3
O108	3	0.3
O104	3	0.3
O28ac/O42, O28ac-O42-Gp2, O28ac, O42-Gp2	3	0.3
Onovel8, O41	2	0.2
Onovel24	2	0.2
Onovel20	2	0.2
OgN9	2	0.2
O88	2	0.2
O79	2	0.2
O63	2	0.2
O59	2	0.2
O45	2	0.2
O39	2	0.2
O38	2	0.2
O172	2	0.2
O17-O77-Gp9, O17/O44, O17-O44-O77-Gp9, O17/O77	2	0.2
O156	2	0.2
O153var2, O8,	2	0.2
O153-O178-Gp11, O153/O178	2	0.2
O149	2	0.2
O124var1, O8	2	0.2
O123, O123/O186, O123-O186-Gp5, O123-Gp5	2	0.2
O115	2	0.2
O110	2	0.2
O10	2	0.2
O169, O169-Gp16, O169/O183	2	0.2
Onovel29	1	0.1
Onovel19	1	0.1
Onovel13	1	0.1
O96, O8	1	0.1
O93	1	0.1
O82	1	0.1
O40, O8	1	0.1
O21	1	0.1
O181	1	0.1
O175	1	0.1
O166	1	0.1
O153-O178-Gp11, O178-Gp11	1	0.1

O153	1	0.1
O152	1	0.1
O15	1	0.1
O148	1	0.1
O145	1	0.1
O137-Gp1, O20	1	0.1
O137	1	0.1
O125,	1	0.1
O108var1	1	0.1
Total	923	100

8.8 Shiga toxin production assessed via rapid membrane enzyme immunoassay (QuikChek)

Isolate ID	Serogroup	*stx* gene profile	Stx production detection	Concordance	Stx1 production amount[a]	Stx2 production amount
ec125a	O174	2d	2	1	-	+++++
ec129b	O168	1a, 2g	1, 2	1	+++	+++
ec1395a	O96, O8	1a, 2b	1	0	+	-
ec162b	O108	1a	1	1	+++	-
ec18	O174	1c, 2b	1	0	+++	-
ec1821	O157	1a, 2c, eae	1, 2	1	++++	+++++
ec2070b	O177	2c, eae	2	1	-	+++++
ec21	O84	1a, eae	1	1	+++++	-
ec2158a	O98	2a, eae	2	1	-	+++++
ec2165a	O163	2a	2	1	-	+++
ec2199a	O175	1a, 2a	1, 2	1	+++++	+++
ec2205a	O2	1d	1	1	+++	-
ec2268b	O171	2c	2	1	-	+++++
ec265b	O5	1c	1	1	+++	-
ec350a	O91	2b	2	1	-	+++
ec4616a	O159	2e	neg	0	-	-
ec4836a	O176	1c, 2d	1, 2	1	+++++	+
ec4948a	Onovel17	2g	2	1	-	+++
ec5122b	O157	1a, 2c, eae	1, 2	1	+++	++++
ec5126b	Onovel20	2d	2	1	-	++
ec5129a	Onovel17	1a, 2g	neg	0	-	-
ec5140a	O157	1a, 2a, eae	1, 2	1	++++	+++++
ec5142b	O113	1a, 2d	1, 2	1	+++++	+++++
ec5145a	O113	1a, 2d	1	0	++++	-
ec5154b	O113	2d	2	1	-	+++++
ec5157c	O157	1a, 2c, eae	1,2	1	+++	+++
ec5164a	O10	2a, eae	2	1	-	+++++
ec5169b	O182	2a, eae	2	1	-	+++++
ec5186a	O168	2d	neg	0	-	-
ec5191b	O5	1a, 2a, eae	1	0	+++	-
ec5195a	O182	1a, 2a, eae	neg	0	-	-
ec5196a	O26	1a, eae	1	1	+++++	-
ec5202b	O157	1a, 2c, eae	1, 2	1	++	+++
ec5203a	O157	1a, 2c, eae	1, 2	1	+++	+++++
ec5205a	O116	2d	neg	0	-	-
ec5206b	O130	1a, 2a	1, 2	1	+++++	++++
ec5207a	O157	2c, eae	neg	0	-	-
ec5207c	O157	2c, eae	2	1	-	+
ec5207d	O26	1a, eae	1	1	+++++	-

ec5213a	O26	1a, eae	1	1	+++++	-
ec5218b	O26	1a, eae	1	1	+++++	-
ec5236a	O149	1a, 2d	1, 2	1	+++++	+
ec5237c	O157	1a, 2a, eae	1, 2	1	++++	+++++
ec5242b	O84	1a, eae	1	1	+++++	-
ec5242f	O172	1a, 2d	1, 2	1	+++++	+++++
ec5254b	O157	1a, 2c, eae	1, 2	1	+++	++++
ec5257a	Onovel17	2g	2	1	-	+
ec5267b	O26	1a, 2d, eae	1	0	+++++	-
ec5270b	O130	1a, 2d	1, 2	1	+++++	++++
ec5281a	O177	2d, eae	neg	0	-	-
ec5290a	O104	1a	1	1	+++++	-
ec5291c	O108	2a, eae	2	1	-	+++++
ec5301a	O182	1a, eae	1	1	+++++	-
ec5311b	O26	1a, eae	1	1	+++++	-
ec5315a	O111	1a, eae	1	1	+++++	-
ec5322a	O179	1a, 2d	1, 2	1	+++++	+++++
ec5325b	O177	2c, 2d, eae	2	1	-	+++++
ec5344a	O157	2c, eae	2	1	-	+
ec5345a	O157	1a, 2c, eae	1, 2	1	+++++	+++++
ec5349a	O182	1a, 2a, eae	1	0	+++++	-
ec5351a	O45	1a, eae	1	1	+++++	-
ec5352a	O84	1a, eae	1	1	+++++	-
ec5355a	O84	1a, 2d, eae	1	0	+++++	-
ec5361a	O157	1a, 2c, eae	1, 2	1	++	+++
ec5363c	O165	2c, eae	2	1	-	+++++
ec5364a	O113	1a, 2a, 2c	2	0	-	+++++
ec5367a	O174	2d	2	1	-	+++++
ec5369a	O26	1a, 2a, eae	1	0	+++++	-
ec5375b	O182	1a, 2a, eae	1	0	+++++	-
ec5375c	O157	1a, 2a, eae	1, 2	1	++++	+++++
ec5383a	O177	2a, eae	2	1	-	+++++
ec5385b	O26	1a, eae	1	1	+++++	-
ec5386a	O111	1a, eae	1	1	++++	-
ec5388a	O111	1a, eae	1	1	++++	-
ec5389a	O157	2c, eae	2	1	-	+
ec5390a	O134, O46	1a, 2d	1, 2	1	++++	+++++
ec5391a	O111	1a, eae	1	1	++++	-
ec5398a	O74	1a, eae	1	1	+++	-
ec5407b	O115	2a, eae	2	1	-	+++++
ec5417a	O115	1a, 2a, eae	2	0	-	+++++
ec5419a	O150	1a, eae	1	1	++++	-
ec5424a	O112	2d	2	1	-	+++++
ec5425a	O112	2d	2	1	-	+++++
ec5427a	O157	1a, 2a, eae	1, 2	1	++++	+++++
ec5433a	O5	1a, eae	1	1	+++++	-

ec5435a	O5	1a, eae	1	1	+++++	-
ec5435c	O111	1a, 2a, eae	1, 2	1	++++	+++++
ec5436a	O157	1a, 2c, eae	2	0	-	+
ec5437a	O156	1a, eae	1	1	++	-
ec5442a	O157	1a, 2c, eae	2	0	-	+
ec5443c	O137	2d	neg	0	-	-
ec5444b	O165	1a, 2c, eae	1, 2	1	+++++	+++++
ec5446a	O98	2a, eae	1, 2	0	+++	+++

[a]A rough estimate of toxin level produced was recorded based on the relative intensity of band(s) that appeared on rapid membrane enzyme immunoassay (Quik Chek) cards. Toxin level was evaluated to range from very low (+) to very high (+++++)