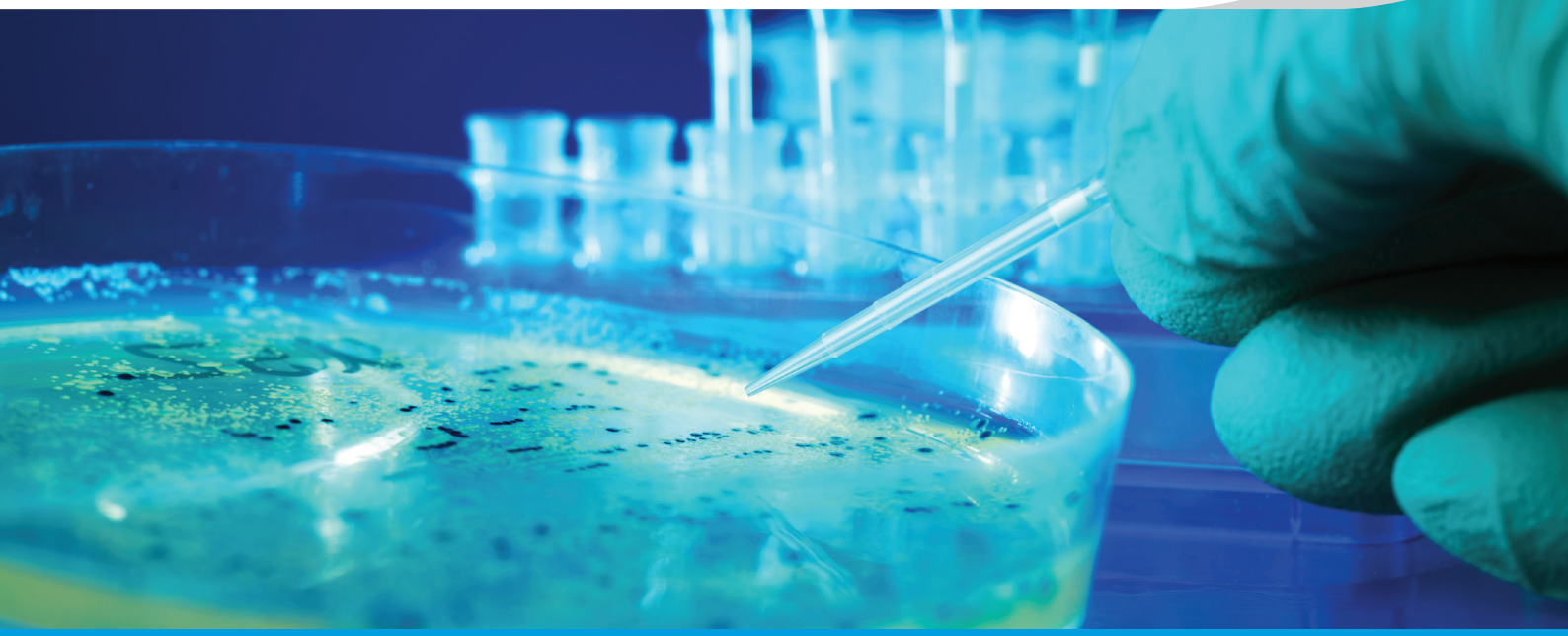
Water Research AUSTRALIA

Research

Discussion Report

Project 1101



Discussion Paper Identification and Management of Environmental *E. coli* Blooms

Collaborate Innovate Impact

Discussion Paper
Identification and Management of Environmental *E. coli* Blooms

Prepared by Dr Martha Sinclair
School of Public Health and Preventive Medicine, Monash University

on behalf of Water Research Australia

for submission to the Drinking Water Quality Advisory Committee
of the National Health and Medical Research Council, Australia

WaterRA Project #1101

Contents

1

Introduction

For many decades it was thought that the bacterium *Escherichia coli* could grow only within human or warm-blooded animal hosts, and that it died quickly in the natural environment. The presence of *E. coli* in water was therefore regarded as an indicator of recent faecal pollution. However, since the 1980s, evidence has emerged to confirm that some *E. coli* strains can grow in soil and freshwater in a range of climatic zones. These environmental *E. coli* strains show significant genetic differences from the commensal and pathogenic *E. coli* strains derived from human or animal faeces.

The available evidence suggests that environmental *E. coli* strains are often present in low numbers in freshwater bodies, however high density 'blooms' of these organisms have been observed in several Australian water storages, including some which supply water for drinking. Although recurrent bloom events have been observed in a small number of reservoirs, this phenomenon currently remains rare overall. The published literature on environmental *E. coli* blooms is scant, and the conditions which favour bloom formation are poorly understood. Rapid and specific methods for identifying bloom strains have been lacking, and the effectiveness of common water treatment processes for bloom strains has not been specifically tested. In addition, the relative rarity of bloom events has sometimes led to the erosion of corporate memory because of staff turnover within water supply organisations and health regulatory agencies. This situation poses significant challenges for organisations to mount appropriate management responses when environmental *E. coli* blooms are first encountered.

The presence of environmental *E. coli* strains in water triggers the same operational response as the detection of *E. coli* derived from faecal contamination. Blooms rarely result in breakthrough of *E. coli* into treated drinking water, however under the proposed Health Based Targets system soon to be incorporated into the Australian Drinking Water Guidelines, maximum levels of *E. coli* in raw water will be used to classify catchment pathogen risks. Therefore, it is critical that water suppliers are able to identify environmental *E. coli* blooms and demonstrate that they do not arise from faecal contamination, in order to exclude such events from the risk assessment dataset and ensure accurate catchment classifications.

In response to this industry need, Water Research Australia (WaterRA) has undertaken research to fill knowledge gaps about environmental *E. coli* blooms, and provide information and communication resources to assist in identifying and managing such events. The components of this project were:

- Rapid Identification Methodology (Prof. David Gordon, Australian National University)

  Development of a rapid PCR-based identification methodology to identify Group 1 capsule genes (characteristic of known bloom strains) in *E. coli* isolates, and to differentiate capsule types.

- Predictive Modelling (Dr Edoardo Bertone, Griffith University)

  Determination of environmental conditions that may predict the onset of an environmental *E. coli* bloom.

- Water Treatment Validation (Dr Paul Monis & Dr Brendon King, Australian Water Quality Centre)

  Evaluation of the efficacy of current treatment processes to remove, inactivate or reduce environmental *E. coli* bacteria in water.

- Communication and Decision Support (Dr Martha Sinclair, Monash University)

  Documentation of Australian water industry experiences and health regulator perspectives, and development of a decision-making framework that informs a risk-based response to blooms, as well as communication material for industry, regulators and the public.

This Discussion Paper has been prepared to assist the NHMRC Water Quality Advisory Committee by summarising information on the current state of knowledge on environmental *E. coli* strains and blooms in water storages. The contents include:

- A brief review of *E. coli* genetics to provide context for the discussion of environmental strains.
- A summary of the published literature on environmental *E. coli* strains and blooms in water storages.
- A brief description of the methods and outcomes of the WRA research project.
- Appendices comprising:
  - The Fact Sheet for public communication.
  - The Utility Response Protocol developed from the collective experiences of Australian water organisations and the perspectives of health regulators.

The WaterRA Project Report (Management of Environmental *E. coli*: Project 1101-16) will also be provided to the WQAC in confidence in order not to compromise planned publication of the project results in the scientific literature.

2

Genetics of *E. coli*

2.1 Methods for differentiation of strains

Prior to the advent of molecular biology techniques, the identification and taxonomic classification of *E. coli* and other bacterial species was based on phenotypic properties such as cell morphology and structure, and biochemical tests to detect enzymatic activities. Classification algorithms were developed based on the most common biochemical profile for a given species. Strains within a species could be distinguished by serotyping, bacteriophage typing or occasionally by differences in their biochemical profile. In the case of pathogens, clinical features were also important in determining the identity of a microorganism.

Subsequently, techniques were developed to enable a finer level of discrimination between strains by examining genetic differences. These methods included multilocus enzyme electrophoresis (MLEE) to detect differences in the electrophoretic mobility of intracellular enzymes. The observed mobility differences are caused by changes in the net electrical charge of the protein molecule as a consequence of variations in the amino acid sequence. In turn, the protein changes reflect differences in the DNA sequence of individual genes within strains. Use of MLEE to analyse several intracellular enzymes involved in cell functions which are shared by most *E. coli* strains (so-called 'housekeeping' functions) permitted differentiation of individual strains.

As molecular biology techniques developed, MLEE was superseded by multilocus sequence typing (MLST). With this methodology, the DNA sequence of specific regions within selected genes is determined. Each unique sequence (300-700 base pairs depending on the gene) is regarded as an allele, and each combination of alleles for seven selected genes (or eight genes, depending on the typing scheme being used) carried by individual strains corresponds to a sequence type (ST). Currently, three different MLST typing systems for housekeeping genes are recognised, with each having its own set of target genes, with only one gene shared by all systems. Each MLST system has a separate database of strain STs against which newly characterised isolates can be compared (Clermont, Gordon *et al.* 2015). MLST analysis of housekeeping genes provides a more detailed level of discrimination between strains than MLEE.

New isolates are often compared to the ECOR (*E. coli* Reference) collection using MLST. The ECOR collection was established in the early 1980s, and comprises 72 strains isolated from a variety of hosts and geographical locations. It was intended to broadly reflect known levels of genetic variation among *E. coli* strains, as revealed by MLEE. This collection does not contain members of the cryptic clades described in Section 2.4.

2.2 Genome size and diversity

The genome of a typical *E. coli* strain is about 5 Mb in size and contains about 5,000 coding genes, however, only about two-thirds of these are part of the 'core' genome (defined as genes found in at least 95% of *E. coli* strains). The remainder of the genes in any given strain are drawn from a large pool of 'accessory' genes. By early 2015, whole genome sequencing of 2,085 *E. coli* strains had identified about 85,000 accessory genes, of which around one-third were represented only once in the strain collection (Land, Hauser *et al.* 2015). The genome size of different *E. coli* strains may vary by more than 1Mb, or roughly 1000 genes (Gordon 2010). This high degree of genetic variability is thought to play a role in the diversity of characteristics displayed by different *E. coli* strains, ranging from commensal organisms in a range of different hosts through to facultative and obligate pathogens, and the recently recognised environmental strains.

2.3 *E. coli* phylogroups

E. coli is comprised of several phylogroups (or subspecies). Initially, four main phylogroups were identified using MLEE to characterise intracellular enzymes for several housekeeping genes. The genetic relationships revealed by this analysis were supported by independent studies of polymorphisms in ribosomal RNA gene regions. The four main phylogroups were designated A, B1, B2 and D. Increasingly sophisticated methods of genetic analysis have continued to provide greater insights into the evolutionary relationships within the species, and seven *E. coli* phylogroups are now recognised (A, B1, B2, C, D, E and F).

In order to assign a new isolate to a phylogroup without performing extensive DNA sequencing, a method targeting a small number of genetic markers which are not shared by all phylogroups was developed (Clermont, Bonacorsi *et al.* 2000). With this approach, PCR techniques are used to detect the presence or absence of three DNA segments (two genes and one sequence of unknown function). The pattern of results enables classification into one of the four original phylogroups A, B1, B2 and D. Each strain is also given a 3-digit code, with each digit corresponding to the presence (1) or absence (0) of a specific target sequence (e.g. A-010). Following the discovery of additional phylogroups and cryptic clades, this triplex PCR method was changed by addition of a fourth genetic marker, modification of some primers and addition of confirmatory tests. The resulting quadruplex PCR system was reported to discriminate all seven *E. coli* phylogroups and five cryptic clades with few strains remaining untypable (Clermont, Christenson *et al.* 2013). However, subsequent use of this method in several strain collections has shown large differences in the proportions being reclassified from one phylogroup to another, and significant percentages of unclassifiable isolates, suggesting that more work is needed to refine the new system (Starčič Erjavec, Predojević *et al.* 2017).

2.4 Cryptic clades

Five 'cryptic' clades of *E. coli* were discovered about a decade ago following detailed genetic analysis of 37 isolates which were generally phenotypically indistinguishable from classical *E. coli* on commonly used biochemical tests, but notably divergent with respect to nucleotide sequences of core genes (Walk, Alm *et al.* 2009). These strains were derived from collections from several different countries, and most had been omitted from publications because of their 'unusual' status. Most of these strains (19 of 37) were described as being isolated from environmental or water locations, seven were from birds, six were from mammals other than humans, and five were from human sources. Analysis of 22 conserved genes revealed that these strains did not map within the known *E. coli* phylogroups, but instead fell into five groupings which were designated as clades CI, CII, CIII, CIV, CV.

Genetic comparison of these strains with information held in three major *E. coli* databases showed that very few representatives of the cryptic clades had previously been characterised. Whole genome sequencing of representative strains from the cryptic clades has confirmed that these clades span the phylogenetic tree between *E. coli* and *E. albertii*, with CI being most closely related to *E. coli*, while the other clades are more distant (Luo, Walk *et al.* 2011). Information about these cryptic clades is currently very limited, however these findings suggest that environmentally adapted *E. coli* lineages are widespread in the environment, but have been rarely detected because of the strong historical focus of research on commensal and pathogenic *E. coli* strains from human sources and, to a lesser extent, from domesticated mammals.

2.5 Pathogenic *E. coli* and Shigella

Six categories of pathogenic *E. coli* which cause diarrhoeal disease are currently recognised: enterotoxic (ETEC), enteropathogenic (EPEC), enterohaemorrhagic (EHEC), enteroinvasive (EIEC), enteroaggregative (EAEC) and diffusely adherent strains (DAEC) (Croxen, Law *et al.* 2013). In addition, some strains of *E. coli* can cause infections in other parts of the body (e.g. urinary tract infections, neonatal meningitis, septicaemia, and bacteraemia), and these have been collectively termed extraintestinal pathogenic *E. coli* (exPEC) (Russo and Johnson 2000). Phylogenetic analysis has shown that pathogenic strains associated with diarrhoeal disease are scattered among the phylogroups, while those associated with extraintestinal infections tend to be concentrated in phylogroups B2 and D (Gordon 2013). Over 60 virulence factors have been identified as playing a role in the pathogenicity of *E. coli* (Kaper, Nataro *et al.* 2004), and several such genes are required in combination to confer a pathogenic phenotype.

Bacterial phylogenetic studies using DNA sequencing data have demonstrated that the four bacterial species comprising the genus Shigella, are in fact members of the *E. coli* species (Gordon 2013). Analysis of several strains by MLST showed that *S. sonnei*, *S. boydii* and *S. flexneri* are closely related to each other and map between members of *E. coli* phylogroups A and C, while *S. dysenteriae* maps within phylogroup E.

2.6 *E. coli* capsules

Some *E. coli* strains are capable of producing extracellular polysaccharide capsules. These capsules are classified into four types (Group 1, 2, 3 and 4) based on their distinctive genetic and biosynthetic characteristics (Whitfield and Roberts 1999). Over 80 serotypes of capsular antigens (termed K-antigens) are known to exist. The extracellular capsule can sometimes physically mask the O-antigens on the cell surface, preventing serotyping using antisera. Production of a capsule may confer a number of characteristics including resistance to desiccation, adherence to biotic and abiotic surfaces, bacteriophage resistance, resistance to nonspecific host immunity, and resistance to specific host immunity (Taylor and Roberts 2005). Some of these properties (e.g. adherence to surfaces, bacteriophage resistance and resistance to desiccation) are relevant to environmental survival as well as interactions with human and other hosts. Commensal *E. coli* and pathogenic enteric strains that cause diarrhoeal disease are generally not encapsulated, while many extraintestinal pathogenic strains produce a Group 2 capsule (Yang, Xi *et al.* 2018).

The Group 1 capsule gene cluster of *E. coli* is closely related in terms of DNA sequence and structural arrangement to the capsule gene cluster in *Klebsiella pneumoniae*, while the Group 2, 3 and 4 capsule gene clusters are not (Whitfield and Roberts 1999). In *E. coli* strains which produce a Group 1 capsule, the capsule biosynthesis genes (*cps* gene cluster) are located between the *galF* gene and the serotype O8/O9 antigen synthesis genes (*rfb* genes). It is believed that a DNA segment encoding the Group 1 capsule genes, the O8/O9 antigen genes and other genes for synthesis of capsule precursor molecules was acquired from *K. pneumoniae* during the evolution of some *E. coli* strains (Samuel and Reeves 2003). Expression of the *galF* gene is believed to enhance capsule synthesis by boosting intracellular levels of precursor molecules, but is not strictly required for capsule production. The length of the Group 1 capsule gene region is variable from strain to strain, and the region often contains insertion sequences which are a legacy of past translocation events.

3

Environmental *E. coli* Literature Summary

3.1 Recognition of environmental *E. coli*

Escherichia coli is a bacterial species found in the gastrointestinal tract of humans and warm-blooded animals (mammals and birds). Discovered in 1884, and initially named *Bacterium coli commune*, *E. coli* was for many decades regarded solely as an enteric microorganism that was unable to grow in the natural environment. Studies on the survival of *E. coli* strains isolated from human and animal faeces showed that their viability rapidly declined in water, and this property was a key factor in the adoption of *E. coli* as an indicator of faecal pollution in water supplies.

During the 1970s and 1980s, there were several reports that high numbers of *E. coli* could be detected in soil and water in pristine tropical environments where faecal inputs were low (Hazen and Toranzos 1990). Further studies confirmed the bacteria were growing, not merely exhibiting prolonged survival in the environment. Most of these studies focused on *E. coli* in soil or sediment samples, and the presence of identical strains in water was attributed to run off after rainfall or resuspension of sediments. It was hypothesised either that these *E. coli* strains had originated from faecal sources, but had been able to adapt to growth in tropical environments because of the warm temperatures and stable nutrient levels, or that *E. coli* had always been part of the natural microbiota of such environments (Alm, Walk *et al.* 2011). The status of *E. coli* as a faecal indicator in tropical waters was therefore thrown into doubt, but it was still regarded as a reliable indicator in other climatic zones (Hazen and Toranzos 1990).

However, similar observations on 'naturalised' (or environmental) *E. coli* populations were later reported from several subtropical and temperate regions of the world, including some northern states of the USA where winter temperatures fall below 0°C. A notable feature of these environmental *E. coli* populations was the relatively low diversity of strains at any particular location, in contrast to the high diversity associated with *E. coli* strains derived from human or animal faecal sources:

- A study of six freshwater beaches in a lake and river system in Michigan, USA used biochemical typing, multilocus enzyme electrophoresis (MLEE) and multilocus sequence typing (MLST) to examine 190 *E. coli* isolates collected over a 20 month period (Walk, Alm *et al.* 2007). Sand was sampled at four depths up to 20 cm below the surface, with a total of 205 samples collected. Biochemical tests (21 tests) revealed between 10 and 14 biotypes at each beach, however one biotype was present in 45.3% of all samples and was present at multiple time points at all six beaches. Collectively, the top four biotypes accounted for 70% of strains isolated. MLEE and MLST characterisation showed many different variants (143 electrophoretic types and 130 sequence types, respectively) but most occurred only once, and the *E. coli* populations at each site were dominated by relatively few strains which persisted over time. The results of this study showed strong multilocus disequilibrium, suggesting strong selection in favour of certain *E. coli* genotypes, particularly those of the B1 phylogroup (which accounted for 56% of all isolates).
- Similar observations were made in a study of *E. coli* in soil from the shores of Lake Superior in Minnesota (Ishii, Ksoll *et al.* 2006). Samples of the top 10 cm of soil were collected on 10 occasions from multiple locations at three sites between October 2003 and October 2004. A total of 586 *E. coli* soil isolates were analysed using a PCR DNA fingerprinting technique which identified 32, 84, and 49 unique *E. coli* strains, respectively, from the three sampling sites. Comparison of these DNA fingerprints with DNA libraries for faecal *E. coli* collections from 346 local animal and bird species, and from more than 1,500 *E. coli* strains from humans and 12 animal species around the state, showed that less than 20% of isolates from each site matched those from known faecal sources. Among the non-faecal strains, a single fingerprint was present at all sites throughout the one year sampling period. This isolate accounted for between 14% and 28% of all isolates detected, depending on the site.

3.2 Environmental *E. coli* blooms

In Australia, environmental *E. coli* strains have been observed to form high density 'blooms' in several freshwater bodies including some reservoirs which supply water for drinking, or which are used for recreation. Several drinking water supply organisations have experienced multiple blooms in particular reservoirs over several decades, however there are only a few publications relating to these events in the scientific literature (Mackay and Ridley 1983, Ashbolt, Dorsch *et al.* 1997, Power, Littlefield-Wyer *et al.* 2005). Attempts to locate internal reports of investigations undertaken by water companies during early bloom events have been unsuccessful, and it is presumed that these have been lost with the passage of time.

There is little evidence that environmental *E. coli* blooms have been recognised in drinking water storages in other countries. According to Ashbolt *et al.* (1997) levels of up to 80,000 coliforms per 100 mL were documented in the San Joaquin Reservoir in an internal report produced in 1983 by the Orange County Water District in California. No other reports of apparent *E. coli* or coliform bloom events of the magnitude experienced in Australia were located in the literature during preparation of this report.

3.2.1 Bloom strains and water quality monitoring methods

Water quality monitoring methods for faecal indicator bacteria have changed over the decades, with biochemical methods based on fermentation of lactose now being replaced by defined substrate technology techniques such as Colilert™ which detect the activity of the enzymes β-galactosidase (common to coliform bacteria) and β-glucuronidase (specific to *E. coli*). Regardless of the particular method used, the specific identification of *E. coli* by biochemical means relies on algorithms that define the 'typical' properties of this species. These algorithms have been compiled by testing bacterial strain collections that are likely to be heavily biased towards clinical and commensal faecal strains. Even so, most algorithms only specify that the large majority (often ≥95%) of strains will conform to this phenotype. It is now known that two of the three environmental *E. coli* strains responsible for blooms in the East coast states of Australia produce 'atypical' results on some biochemical tests and one fails to fluoresce in Colilert™ medium (D. Gordon unpublished results). In addition, one strain is unable to grow at 44.5°C. Therefore, these strains may not have been identified as *E. coli* at the time of the original investigations, and may instead have been classified as other members of the thermotolerant coliform group.

3.2.2 Blooms in Lake Burragorang

Multiple blooms of thermotolerant coliform bacteria in Lake Burragorang, NSW (the impoundment of the Warragamba Dam) have been reported. This lake provides the primary drinking water supply for the city of Sydney, NSW. According to a review covering the period from 1961 to 1981, blooms began soon after the dam was completed in 1961 (Mackay and Ridley 1983). Blooms occurred every year for seven years after the initial filling of the dam (1961-67), then were absent for six years (1968-73), then returned for eight consecutive years (1974-81). The majority of blooms (10 of 15) occurred during the period from November to January. This pattern over the twenty years was said by the authors to be "without a corresponding known botanical cycle or weather cycle".

Graphs showing thermotolerant coliform numbers indicate that background levels in the reservoir were in the order of 10 CFU/100 mL or less, while levels ranging from 1,000 CFU /100 mL to more than 100,000 CFU /100 mL were found at various water sampling sites during bloom events. Not all sites were affected during every event, however the water sampling site at the offtake point for the drinking water supply recorded bloom levels during 11 of the 15 events. Two bloom events (Events 14 and 15) were investigated in detail, including biochemical testing and serotyping of bacterial isolates that produced a positive result on the Eijkman test at 44°C (acid and gas production in lactose peptone water). This test was considered to have a high selectivity for *E. coli*, but other members of the thermotolerant coliform group may sometimes give positive results. Isolates were further characterised using the API 20E test kit for Enterobacteriaceae.

Event 14 (December 1978/January 1979)

- This event was preceded by record heavy rainfall in March and smaller falls in July and December. The rainfall was associated with a rise in water turbidity and phosphorus levels. There was also a bloom of green algae (predominantly *Volvox* sp.) prior to this *E. coli* bloom.
- Thermotolerant coliform numbers near the dam wall were as high as 16,000 CFU/100 mL on the first sampling date (19 December 1978) and had declined to 220 CFU/100 mL at this site by the final sample date (21 January 1979), but remained higher at two distal sites (each located about 20 km from the dam wall, but in different branches of the lake). The duration of the bloom was said to be six to eight weeks.
- The composition of the bloom changed over time. Initial samples (collected on 19 December) were dominated by *Citrobacter freundii* with a uniform API 20E profile, but *E. coli* were also present. By 27 December, *E. coli* had become predominant (70/82 isolates) and *Citrobacter freundii* had declined. By 21 January, all isolates (161/161) were identified as *E. coli*. Biochemical testing of 309 Eijkman-positive isolates showed significant deviations from the 'typical' *E. coli* pattern.
- A subsequent publication reported that the 70 *E. coli* strains isolated from the 27 December samples were all of the same API 20E profile (Ashbolt, Dorsch *et al.* 1997).
- Twenty-two of 24 *E. coli* isolates sent for somatic serotyping were serotype O8. Two samples were tested for their flagellar serotype, both were H10.
- Tests for faecal streptococci were negative.
- No mention is made of *E. coli* colony morphology in the description of this event.

In discussing this bloom, Mackay and Ridley state that an 'almost pure culture of *E. coli* O8:H10 was present over the whole lake at 0.5 m and also at 42 m depth' (Mackay and Ridley 1983). They postulate that the preceding algal bloom may have provided nutrients that enabled *E. coli* to grow, and that elevated turbidity persisting after the heavy rainfall (8.5 NTU recorded in November, compared to usual levels around 1 NTU) may have had a role in prolonging *E. coli* survival.

Event 15 (December 1981)

- There was no significant rainfall prior to this event, water turbidity was low and no major algal blooms were detected. The only notable weather phenomenon was period of five days of intense heat during the first week of December, prior to the bloom event.
- Water sampling results for one location near the dam wall and eight dates from 9 November to 29 December 1981 are presented. Thermotolerant coliform numbers were moderately increased until 30 November (maximum 236 CFU/100 mL), but peaked on 14 December at 296,000 to 298,000 CFU/100 mL from the surface down to at least 6 metres in depth. Counts fell to 5,500 to 6,100 CFU /100 mL through the top 6 metres of the water column by the final sampling date on 29 December.
- Some 73% of the 282 tested isolates gave positive results on the Eijkman test, but it is not stated whether all of these were confirmed as *E. coli*. Again, it was noted that the results of biochemical tests deviated from the 'typical' *E. coli* profile for several characteristics.
- A later publication reports that this bloom was dominated by two *E. coli* strains, one of which was an indole-negative strain and the other was of the same API E20 type as the most common *E. coli* strain in the 1978/79 bloom (Ashbolt, Dorsch *et al.* 1997).
- Attempts to serotype 40 isolates were hampered because they were mucoid in nature, but where this was successful (number of isolates not stated), they were reported to be O8 or related serotypes. Phage typing of 33 isolates showed 32 to have the same phage type.
- Faecal streptococci numbers were low (less than 10 CFU /100 mL) near the dam wall when *E. coli* levels were high on 18 and 21 December, but faecal streptococci were detected at the inflow of the Monkey Creek tributary (88 to 100 CFU /100mL) at this time.

In this event, high *E. coli* counts were detected only at two sampling sites (near the dam wall and about 4 km from the dam), however the same strain (O8 serotype) was also detected at lower concentrations at three other sites. The detection of faecal streptococci in the Monkey Creek area was thought to have been attributable to animal farming and lack of a reticulated sewerage system in this area. Mackay and Ridley state that "the results of the sub-species grouping indicated that the high counts of *E. coli* were not generated by inflowing faecal pollution and the regrowth hypothesis is proven."

A 1997 review assessed data from a 20-year period (1972 to 1992) for Lake Burragorang and also described previously unpublished observations from the earlier investigation (Ashbolt, Dorsch *et al.* 1997). This analysis yielded the following observations:

- Blooms of thermotolerant coliforms were associated with warm surface water temperatures (above about 18°C) and the establishment of water stratification during springtime.
- The data for rainfall and thermotolerant coliform numbers are presented as monthly means and percentiles averaged over the 20-year interval. Therefore, it is not possible to determine the date of occurrence of any individual blooms within the data set.
- In the averaged data, two peaks in thermotolerant coliform levels (for both monthly mean and monthly 95th percentile measures) were evident, with the largest in December, and a smaller peak in April-May. High average rainfall (mean 95th percentile) occurred about 1 to 2 months before the thermotolerant coliform peaks.
- Other coliform bacteria including *Enterobacter cloacae* (concentration not stated) were also detected during the 1978/79 and 1981 *E. coli* blooms reported by Mackay and Ridley. *Aeromonas hydrophila* was found at levels up to 1,600 CFU/L (not stated whether this relates to one or both of these blooms), but baseline levels for this organism were not reported. In addition, a bloom of *Enterobacter cloacae* was observed in the summer of 1994/95.
- The authors do not explicitly state that tests for faecal streptococci returned negative results during *E. coli* blooms, but comment on "the absence of faecal streptococci" in the context of calculations about expected ratios of thermotolerant coliforms to faecal streptococci in faecal pollution events. They estimate that a faecal streptococci concentration of about 8,000 CFU /100 mL would have been expected if the 1978/79 bloom was attributable to faecal pollution. This calculation assumed a ratio of thermotolerant coliforms to faecal streptococci in the range of 5-35 for sewage or stormwaters, and the peak *E. coli* level close to 300,000 CFU/ 100 mL.
- The amount of faecal material required to generate the high levels of *E. coli* throughout the large volume of affected water seen in the 1981 bloom is also calculated. This was estimated at 300 tonnes of human faeces, or 12,900 tonnes of cow faeces or 296,000 tonnes of horse, pigeon or sparrow faeces.
- DNA sequencing of the 16S ribosomal genes was carried out for some *E. coli* bloom stains and some differences were detected compared to a clinical strain. It was suggested that further work might identify a suitable probe to differentiate between environmental and faecal *E. coli* strains.

E. coli bloom events have continued to occur in Lake Burragorang, although perhaps not as frequently as during the 1960s and 70s. A former staff member of Sydney Water's in-house laboratory estimates that there were three or four blooms investigated between 1984 and 1998. More recently, two blooms in this lake between 2008 and 2018 were reported when data were gathered for the WRA research project.

3.2.3 Blooms in Lake Burley Griffin

Lake Burley Griffin in the ACT has reportedly experienced annual *E. coli* blooms for 30 years prior to 2005 (Power, Littlefield-Wyer *et al.* 2005). This artificial ornamental lake located within the city of Canberra was first filled in 1964. Primary and secondary contact recreational water activities are seasonally permitted in the lake, but closures due to poor water quality from stormwater inflows and cyanobacterial blooms are not uncommon.

E. coli blooms have occurred from mid-summer to early autumn when water temperatures were 18°C or higher. Water sampling results suggested that the *E. coli* strains responsible for bloom formation are present at low numbers at other times of year in the lake, but under favourable conditions cell numbers can increase to densities of over 10,000 CFU/100 mL within one week. Once again, calculation of the amount of human faeces needed to produce the number of *E. coli* present during a bloom showed that faecal pollution was not a plausible source of contamination. It was estimated that raw sewage from more than 470,000 people (greater than the population of Canberra at that time) would have needed to enter the lake daily for 7 days to produce the observed levels of *E. coli* (Power, Littlefield-Wyer *et al.* 2005).

4

Phenotypic and Genotypic Analysis of *E. Coli* Bloom Strains

4.1 Characterisation of East coast bloom strains

The characteristics of two *E. coli* isolates from blooms in Lake Burragorang and six from Lake Burley Griffin have been described (Power, Littlefield-Wyer *et al.* 2005). The Lake Burragarong isolates both originated from a bloom that occurred in February 2003, while the Lake Burley Griffin isolates came from blooms in April 2002 (2 isolates), February 2003 (1 isolate), April 2003 (2 isolates) and February 2004 (1 isolate).

All eight isolates were found to exhibit mucoid colony morphology when grown on MacConkey agar, and all produced large extracellular capsules which could be visualised by negative staining and differential interference microscopy. ECOR typing using the method of Clermont *et al.* (2000) showed that five of the isolates were members of phylogroup A, and three were members of phylogroup B1. Biochemical typing using 16 traditional tests, plus two defined substrate tests for the presence of the enzymes β-D-galactosidase (ONPG test) and β-glucuronidase (MUG test) revealed that the eight isolates comprised three strains with distinct biochemical profiles. The phenotypic characteristics of the three strains are summarised in Table 4.1.

Table 4.1 Phenotypic characteristics of E. coli bloom strains

Strain designation (ECOR group-genotype)	Biochemical profile	Growth at 44.5°C	Presence in blooms				
			Lake Burley Griffin				Lake Burragorang
			Apr 02	Feb 03	Apr 03	Feb 04	Feb 03
A-000	*Typical E. coli*	yes	+		+		+
A-010	*MUG negative adonitol positive*	yes	+				+
B1-001	*MUG negative indole negative lysine negative melibiose negative*	no		+	+	+	

+ signifies the strain was detected in the bloom

Strain A-000 had a profile typical of *E. coli*, strain A-010 differed in two properties, but was still identified as *E. coli* by standard biochemical test systems. Strain B1-001 diverged significantly in biochemical profile and also showed reduced thermotolerance. This strain was identified as *Citrobacter youngii* with the API20E test system, and either *E. coli* or *Shigella sonnei* with the BBL Enteric/non-fermentor system, depending on whether individual colonies were able to utilise arabinose. These observations support the view that some of the isolates identified as other thermotolerant coliforms in early studies, may in fact have been atypical *E. coli*.

Serotyping by PCR amplification of regions of the genes involved in O-antigen biosynthesis showed that strains A-000 and B1-001 were of serotype O9, while tests on strain A-010 failed to amplify any gene product for either the O8 or O9 serotypes. Subsequent DNA sequencing indicated that strain A-010 had suffered a large deletion affecting the genes encoding O-antigen production.

The three strains were screened for 21 genes associated with virulence. The A-010 strain was found to carry the fimH gene (Type 1 fimbrin D-mannose specific adhesion). This gene is involved in the regulation of type 1 fimbriae, but is not of itself sufficient to cause production of fimbriae or to confer pathogenicity.

DNA sequencing of the capsule-encoding genes (cps gene cluster) in the three strains confirmed that all produced a Group 1 capsule, but their cps DNA sequences were not identical. The strains could be readily distinguished from one another by digestion of a 1200 bp PCR product from the *wzi* gene (capsule assembly protein) using the restriction enzymes *Hinf1* and *Taq*.

Comparison of the three bloom strains with a collection of 435 strains from a range of vertebrates showed that none of the vertebrate isolates matched the phenotypic or genotypic profiles of the bloom strains (Power, Littlefield-Wyer *et al.* 2005).

Further studies have shown that these three strains, in various combinations, have dominated all the environmental *E. coli* blooms in the eastern states (ACT, NSW, and Qld) where bacterial isolates have been characterised (Nanayakkara, O'Brien *et al.* 2018) (D. Gordon, unpublished data). They have therefore been dubbed the 'East coast bloom strains'. In Lake Burley Griffin, strain B1-001 has been present in all 10 blooms that have been investigated, while strain A-000 was detected in only one of these blooms and strain A-010 in two blooms. In addition to Lake Burragorang and Lake Burley Griffin, environmental *E. coli* blooms have been documented in two reservoirs in the ACT, three in NSW, and two in Queensland since 2002.

Overall, four different capsule types have been observed among isolates from East coast blooms, with three types (KL16, KL49 and KL113) occurring in phylogroup A isolates, and only one type (KL53) seen in phylogroup B1 isolates (Nanayakkara, O'Brien *et al.* 2018).

Examination of *E. coli* strains collected from soil, water or sediments showed that an isolate matching the B1-001 strain was present in Lake Burley Griffin for several months during non-bloom period in 2003 and also in the Tallowa Dam, NSW in the absence of a bloom (Power, Littlefield-Wyer *et al.* 2005). Isolates matching strain A-010 were also detected in Lake Ginninderra, ACT when no blooms were present. Therefore, it is likely that bloom-forming strains persist in low numbers for prolonged periods in water bodies between bloom events. The mucoid nature of bloom isolates suggests that they may be well adapted for growth in biofilms.

4.2 Characterisation of West coast bloom strains

Only one environmental *E. coli* bloom in a water storage has thus far been documented in Western Australia. Five distinct *E. coli* strains were isolated from this 2015 event (Nanayakkara, O'Brien *et al.* 2018). Four strains were classified as phylogroup A and one strain as phylogroup B1, but none matched the ECOR types of the East coast strains. All five strains exhibited a typical *E. coli* phenotype on standard identification tests and Colilert™ tests (D. Gordon, unpublished results).

All of the West coast strains produced a Type 1 capsule. When the capsule genes were sequenced, five different capsule types were seen in the West coast strains. One strain had a capsule type (KL53) that matched the capsule in East coast strain B1 isolates, although the West coast strain belonged to phylogroup A. The remaining capsule types (KL60, KL63 and KL101) have not been detected in East coast bloom isolates (Nanayakkara, O'Brien *et al.* 2018).

5

Methods and Outcomes of WaterRA Project

A brief outline of the Water Research Australia project is presented here, and the Project Report (Management of Environmental *E. coli*: Project 1101-16) will be made available to the WQAC to assist in their consideration of this issue.

5.1 Rapid Identification Methodology

5.1.1 Aims

- To determine if a PCR-based approach can be developed to identify *E. coli* bloom strains by searching for unique capsule genes among strains carrying the Group 1 capsule gene cluster.
- To develop a rapid multiplex PCR identification methodology for bloom strains to permit their detection and differentiation from each other.
- To estimate the frequency of Group 1 capsules among *E. coli* strains detected in water by testing an existing collection of over 1600 isolates from Queensland and NSW. These strains were chosen to represent all REP-PCR fingerprint types for *E. coli* detected in water samples.

5.1.2 Methods

- Analyse existing whole genome sequence (WGS) data for more than 1200 *E. coli* isolates from a variety of vertebrate species and environmental samples for DNA sequences similar to the *Klebsiella* capsule (Group 1) genes. Then compare capsule-positive strains to determine which gene(s) were present in all capsule types and which gene(s) were unique to a particular capsule type.
- Identify genes that are unique to individual known bloom strains, then identify single nucleotide polymorphisms that distinguish the bloom strains from each other.
- Develop PCR primers to detect these differences, then test multiplex PCR primer sets to allow:
  - detection of all Group 1 capsule positive strains.
  - discrimination of all previously characterized *E. coli* bloom strains by differences in the size and pattern of the PCR product(s) from capsule genes.

5.1.3 Outcomes

- WGS data for over 1200 *E. coli* strains showed that less than 10% had a Group 1 capsule genes. All phylogroups were represented in this collection, but strains with Group 1 capsules fell into only phylogroups A, B1, and C, with no such isolates detected in phylogroups B2, D, E or F.
- Two sets (pools) of PCR primers were developed to characterise bloom strains of *E. coli*.
  - PCR Pool 1 - detects the galF gene (common to all strains) and distinguishes between the three East coast bloom strains on the basis of capsule type.
  - PCR Pool 2 - detects the galF gene (common to all strains) and distinguishes between the five West coast bloom strains on the basis of capsule type.
- Less than 5% of over 1600 *E. coli* strains from freshwater sources had a Group 1 capsule. Of these, less than half had capsule types matching those of known bloom strains. Extrapolating these results to include all *E. coli* phylogroups in the denominator, suggests that around 2% of *E. coli* isolates would be expected to produce Group 1 capsules.

5.2 Predictive Modelling

5.2.1 Aims

- To determine if there are any conditions that may predict the onset of an environmental *E. coli* bloom.

5.2.2 Methods

- Conduct a literature review to identify factors possibly influencing *E. coli* numbers in water.
- Collect available water quality, land use, weather and other data potentially relating to bloom events from participating water utilities and other sources.
- Analyse the data using methods including time series analysis, linear and non-linear regression and self-organising maps to look at relationships between measured/observed parameters and bloom events.

5.2.3 Outcomes

- Data sets were highly variable in terms of frequency of observations (hourly to yearly), duration (less than 5 years to over 20 years) and units of measurement (MPN vs CFU for *E. coli* counts) and parameter measured (algal counts vs chl-a levels). Data relating to 12 blooms in five dams in four states were included in the analysis.
- Bushfires/planned burns, dust storms and algal blooms were identified as predisposing events for blooms. Any one of these factors, when combined with water temperatures above 18°C and a rapidly falling dam level, was associated with an increased probability of bloom occurrence in the following months. It is postulated that these predisposing events provide a nutrient trigger for blooms, but the specific nutrient(s) involved are unknown.
- A Bayesian Network model was developed to predict the likelihood of bloom occurrence using historical data inputs for predisposing events and reservoir levels. This model could be used to trigger targeted monitoring programs to better characterise factors associated with *E. coli* blooms.

5.3 Water Treatment Validation

5.3.1 Aims

- Investigate the efficacy of current treatment processes to remove, inactivate or reduce environmental *E. coli.*

5.3.2 Methods

- Conduct a series of laboratory-scale tests comparing the responses of 3 *E. coli* bloom strains (from East and West coast blooms) and 3 comparison strains (from faecal sources) to chlorination, chloramination, UV disinfection, coagulation and filtration.
- Tests were conducted in triplicate under a range of conditions as appropriate for each treatment (e.g. disinfectant doses, UV doses, log-phase and starved cells, reservoir water types).

5.3.3 Outcomes

- Some variation was observed between strains in all types of tests, but no consistent differences for bloom versus non-bloom *E. coli* strains.
- Current treatment processes are efficacious for the removal, inactivation or reduction environmental *E. coli* strains.

5.4 Communication and Decision Support

5.4.1 Aims

- Document Australian water industry experiences and health regulator perspectives.
- Develop a decision-making framework that informs a risk-based response to environmental *E. coli* blooms.
- Develop communication materials for industry, regulators and the public.

5.4.2 Methods

- Collect information on management responses from six water utilities in four states which had experienced environmental *E. coli* blooms in the last 10 years. Also canvas public health concerns with health regulators in the same states.
- Identify public health priorities and common elements of management responses, then develop a Utility Response Protocol to promote a consistent, evidence-based response to *E. coli* blooms.
- Develop a Fact Sheet for public communication, and a Discussion Paper (this document) for submission to the NHMRC WQAC.

5.4.3 Outcomes

Documents were developed:

- This Discussion Paper.
- Fact Sheet (Appendix 1) summarising knowledge about bloom strains and bloom events.
- Utility Response Protocol (Appendix 2). This outlines a basic level of response, plus a range of optional investigations that can be undertaken to further characterise blooms, if desired.

6

Summary and Discussion

The existence of *E. coli* strains that are adapted to environmental growth in soil and water is now supported by a substantive body of scientific evidence. Australian experience has shown that some of these strains are capable of growing to high densities (blooms) in water bodies. *E. coli* blooms have been documented in the East coast states of Australia for several decades, but only one bloom has been observed in Western Australia. To date, eight distinct *E. coli* strains have been identified in bloom events in Australian water storages. Production of a Group 1 extracellular capsule is shared feature of these strains, and is thought to be a key factor in bloom formation. The available evidence indicates that these bacteria do not pose a public health risk in drinking water sources. Exclusion of bloom events from catchment pathogen risk assessments under the ADWG health-based targets system is therefore justified, provided that appropriate investigations to rule out faecal contamination events have been undertaken.

Water Research Australia Project 1101-16 has provided new tools for the rapid identification and differentiation of bloom strains, and verified that conventional drinking water treatment and disinfection methods are efficient for removal of these strains. Predisposing events and reservoir conditions favouring blooms have been characterised, but the suspected nutrient trigger for bloom formation remains unknown. Apart from a few reservoirs with a history of multiple blooms, these events currently appear to be rare. However, it is possible that blooms have been occurring unobserved in some reservoirs, and have been missed because of the infrequency of raw water testing. Given that water suppliers are now undertaking more intensive raw water monitoring in anticipation of the implementation of health-based targets in the Australian Drinking Water Guidelines, the identification of environmental *E. coli* blooms may become a more frequent occurrence.

It is notable that the three predisposing factors for bloom formation (bushfires/burns, dust storms and algal blooms) are all likely to increase in frequency as a consequence of climate change. Similarly, the other contributing factors (warm water temperatures and falling dam levels) will be commonly encountered. The Utility Management Protocol developed from the experiences of Australian water organisations provides guidance on an appropriate risk-based response to *E. coli* bloom events, while safeguarding public health. Inclusion of information about environmental *E. coli* blooms in the ADWG will also ensure that industry awareness and knowledge will be maintained into the future.

7

References

Alm, E. W., S. T. Walk and D. Gordon (2011). The niche of *Escherichia coli*. Population Genetics of Bacteria: a Tribute to Thomas S. Whittam. S. T. a. F. Walk, P. C. H. Washington, DC, ASM Press: 69-91.

Ashbolt, N., M. Dorsch, P. Cox and P. Banens (1997). Blooming *E. coli*, what do they mean? Coliforms and *E. coli*. Problem or Solution? D. Kay and C. Fricker. Cambridge, Royal Society of Chemistry Special Publication. **191** 78-85.

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

Clermont, O., J. K. Christenson, E. Denamur and D. M. Gordon (2013). "The C lermont E scherichia coli phylo-typing method revisited: improvement of specificity and detection of new phylo-groups." Environmental microbiology reports **5**(1): 58-65.

Clermont, O., D. Gordon and E. Denamur (2015). "Guide to the various phylogenetic classification schemes for *Escherichia coli* and the correspondence among schemes." Microbiology **161**(5): 980-988.

Croxen, M. A., R. J. Law, R. Scholz, K. M. Keeney, M. Wlodarska and B. B. Finlay (2013). "Recent advances in understanding enteric pathogenic *Escherichia coli*." Clinical microbiology reviews **26**(4): 822-880.

Gordon, D. M. (2010). "Strain typing and the ecological structure of *Escherichia coli*." Journal of AOAC International **93**(3): 974-984.

Gordon, D. M. (2013). The ecology of Escherichia coli. Escherichia coli: Pathotypes and Principles of Pathogenesis. M. Donnenberg, Elsevier: 3-20.

Hazen, T. C. and G. A. Toranzos (1990). Tropical Source Water. Drinking Water Microbiology. G. A. McFeters. New York, Springer-Verlag: 32-53.

Ishii, S., W. B. Ksoll, R. E. Hicks and M. J. Sadowsky (2006). "Presence and growth of naturalized *Escherichia coli* in temperate soils from Lake Superior watersheds." Applied and environmental microbiology **72**(1): 612-621.

Kaper, J. B., J. P. Nataro and H. L. Mobley (2004). "Pathogenic *escherichia coli*." Nature reviews microbiology **2**(2): 123.

Land, M., L. Hauser, S.-R. Jun, I. Nookaew, M. R. Leuze, T.-H. Ahn, T. Karpinets, O. Lund, G. Kora and T. Wassenaar (2015). "Insights from 20 years of bacterial genome sequencing." Functional & integrative genomics **15**(2): 141-161.

Luo, C., S. T. Walk, D. M. Gordon, M. Feldgarden, J. M. Tiedje and K. T. Konstantinidis (2011). "Genome sequencing of environmental *Escherichia coli* expands understanding of the ecology and speciation of the model bacterial species." Proceedings of the National Academy of Sciences: 201015622.

Mackay, S. and J. Ridley (1983). Survival and regrowth of Escherichia coli in Lake Burragorang. Proceedings Australian Water & Wastewater Association Convention 1983.

Nanayakkara, B. S., C. L. O'Brien and D. M. Gordon (2018). "Diversity and distribution of Klebsiella capsules in *Escherichia coli*." Environmental Microbiology Reports.

Power, M. L., J. Littlefield-Wyer, D. M. Gordon, D. A. Veal and M. B. Slade (2005). "Phenotypic and genotypic characterization of encapsulated *Escherichia coli* isolated from blooms in two Australian lakes." Environmental Microbiology **7**(5): 631-640.

Russo, T. A. and J. R. Johnson (2000). "Proposal for a new inclusive designation for extraintestinal pathogenic isolates of *Escherichia coli*: ExPEC." The Journal of Infectious Diseases **181**(5): 1753-1754.

Samuel, G. and P. Reeves (2003). "Biosynthesis of O-antigens: genes and pathways involved in nucleotide sugar precursor synthesis and O-antigen assembly." Carbohydrate Research **338**(23): 2503-2519.

Starčič Erjavec, M., L. Predojević and D. Žgur-Bertok (2017). "Commentary: Comparative Analysis of Phylogenetic Assignment of Human and Avian ExPEC and Fecal Commensal *Escherichia coli* Using the (Previous and Revised) Clermont Phylogenetic Typing Methods and its Impact on Avian Pathogenic *Escherichia coli* (APEC) Classification." Frontiers in microbiology **8**: 1904.

Taylor, C. M. and I. S. Roberts (2005). Capsular polysaccharides and their role in virulence. Concepts in Bacterial Virulence, Karger Publishers. **12**: 55-66.

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

Walk, S. T., E. W. Alm, D. M. Gordon, J. L. Ram, G. A. Toranzos, J. M. Tiedje and T. S. Whittam (2009). "Cryptic lineages of the genus Escherichia." Applied and environmental microbiology **75**(20): 6534-6544.

Whitfield, C. and I. S. Roberts (1999). "Structure, assembly and regulation of expression of capsules in *Escherichia coli*." Molecular microbiology **31**(5): 1307-1319.

Yang, S., D. Xi, F. Jing, D. Kong, J. Wu, L. Feng, B. Cao and L. Wang (2018). "Genetic diversity of K-antigen gene clusters of *Escherichia coli* and their molecular typing using a suspension array." Canadian journal of microbiology **64**(4): 231-241.

Appendix 1 Fact Sheet

Factsheet **Project 1101**

Management of Environmental *E. coli* Blooms

Recognition of environmental *E. coli* strains

Escherichia coli is a bacterial species which is excreted in high numbers from the intestinal tract of humans, other mammals and birds. The vast majority of *E. coli* strains are harmless, and these bacteria are considered to be a beneficial inhabitant of the healthy human digestive system, however a few types are able to cause disease in humans (pathogenic *E. coli*). For many decades it was thought that *E. coli* could grow only within human or animal hosts, and that it died quickly in the natural environment. Therefore, the presence of *E. coli* in water was regarded as an indicator of recent faecal pollution, with the possibility that faecal pathogens could also be present. In recent years, it has been demonstrated that some strains of *E. coli* have adapted to live and grow in the environment, meaning that the presence of this bacterium in water may not always indicate faecal pollution. These environmental strains are genetically distinct from the *E. coli* strains derived from human or animal faeces, and evidence indicates that they do not carry the genes needed to cause disease.

Environmental *E. coli* and Health Based Targets

Environmental *E. coli* strains cannot be distinguished from faecal strains by routine tests, and their detection in water triggers the same operational response as the detection of *E. coli* derived from faecal contamination. Environmental *E. coli* usually occur in very low numbers in water bodies, and therefore do not pose a problem for water quality management. However, some environmental *E. coli* strains form high density 'blooms' in water storages, including those used to supply drinking water. Under the proposed health based targets system for the Australian Drinking Water Guidelines, the maximum levels of *E. coli* detected in untreated water will be used to classify the level of potential risk from pathogenic faecal microorganisms in each water catchment. The risk classification will determine the degree of water treatment needed to produce drinking water. In order to ensure accurate assessment, it is important that water suppliers are able to identify environmental *E. coli* blooms and demonstrate that they do not arise from faecal contamination, so that these events can be excluded from the risk classification process.

Environmental *E. coli* blooms in Australia

Environmental *E. coli* blooms have been documented in Lake Burragorang, NSW (the impoundment for the Warragamba Dam) since at least 1975, and in Lake Burley Griffin, ACT (the impoundment for the Scrivener Dam) since the late 1990s. Although recurrent bloom events have been observed in a small number of reservoirs, this phenomenon remains rare overall. It is believed that in these lakes a sudden change in the growth rate of bloom-forming *E. coli* strains is triggered by an increase in nutrient levels in combination with warm water temperatures, however, the specific nutrient factors involved in this phenomenon are unknown.

Eight bloom-forming *E. coli* strains have been characterised, and all produce a Group 1 extracellular polysaccharide capsule. Production of this type of capsule is rare among *E. coli* strains. Three bloom-forming strains have been associated with bloom events in six reservoirs located in the ACT, NSW and Queensland (designated as East coast strains). Only one bloom has been documented in WA, and five strains which were different from those seen in the eastern states were isolated from this event (designated as West coast strains).

Water industry experience of environmental *E. coli* blooms

Information about water quality and environmental factors associated with bloom events has been collected from six Australian water organisations and analysed. The affected reservoirs varied from 1 gigalitre to more than 2,000 gigalitres in capacity (1 gigalitre = 1,000,000,000 litres). Catchment types ranged from mixed rural and urban uses (some with primary contact recreational water activities), through to a well-protected catchment with no public access to the water storage. The key findings were:

- Bushfires or planned burns in the catchment, dust storms, and algal blooms were identified as predisposing events for blooms. Any one of these events, in combination with a rapidly falling dam level, significantly increases the likelihood of a bloom within the following few months.
- Bloom events generally occurred in spring or early summer when weather conditions were dry, and water inflows to the reservoir were low or absent. Water temperatures were 18°C or higher, but in most instances it was uncertain if thermal stratification was present in the water body.
- Blooms were not associated with noticeable changes in water quality parameters such as turbidity.
- Levels of enterococci bacteria remained normal during *E. coli* blooms. Enterococci are a different group of bacteria sometimes used as an indicator of faecal contamination.
- For most events, the first sign of a bloom was an *E. coli* reading in untreated water that was about 100-times higher than levels previously seen during routine monitoring. Peak concentrations up to 10,000-times higher than normal were recorded in some events. Levels as high as 298,000 *E. coli* colony forming units /100 mL have been observed.
- Where water sampling was carried out at different depths and locations across reservoirs, high *E. coli* counts were found over large areas and at depths down to at least several metres.
- Bloom duration was usually three to four weeks, although durations up to three months have been seen in recent years in one major reservoir.

Appendix 1 Fact Sheet

Factsheet **Project 1101**

Management of Environmental *E. coli* Blooms

Identification and typing of *E. coli* bloom strains

The genetic diversity of *E. coli* isolates from a suspected bloom may be characterised using established methods to look at variations in genes for basic cell functions that are carried by all *E. coli* bacteria (so-called 'housekeeping' genes). If contamination of water has arisen from faecal sources, then a high degree of strain diversity will be seen when the *E. coli* isolates are tested in this way. However during bloom events, it has been found that only a few strains will account for the majority of isolates. Another characteristic of bloom-forming strains is their mucoid colony appearance, caused by capsule production, when they are grown on agar plates. In addition, the genes for capsule formation can be detected and characterised using specific DNA probes. Variations in the capsule genes can be detected, and this allows the strains to be compared to known bloom-forming strains.

Effectiveness of water treatment and disinfection processes

The responses of bloom-forming environmental *E. coli* strains to common water treatment and disinfection processes have been tested and compared to *E. coli* strains from human and bird faeces. The processes assessed were coagulation (with alum or ferric iron compounds), filtration, chlorination, chloramination and UV disinfection. These experiments showed that the responses of bloom isolates were not significantly different from those of faecal *E. coli* strains. Therefore, commonly used water treatment and disinfection processes are considered to be effective to remove bloom-forming *E. coli* strains from water.

Utility Response Protocol

A Utility Response Protocol has been developed (see below) from the collective experience of Australian water organisations that have encountered environmental *E. coli* blooms over the last decade, and the perspectives of health regulatory agencies. This protocol enables water suppliers to recognise blooms in drinking water storages, and provides guidance on appropriate management measures.

References

Alm, E. W., S. T. Walk and D. Gordon (2011). The niche of *Escherichia coli*. Chapter 6 in *Population Genetics of Bacteria: a Tribute to Thomas S. Whittam*. S. T. a. F. Walk, P. C. H. Washington, DC, ASM Press: 69-91.

Power, M. L., J. Littlefield-Wyer, D. M. Gordon, D. A. Veal and M. B. Slade (2005). Phenotypic and genotypic characterization of encapsulated *Escherichia coli* isolated from blooms in two Australian lakes. *Environmental Microbiology* 7(5): 631-640.

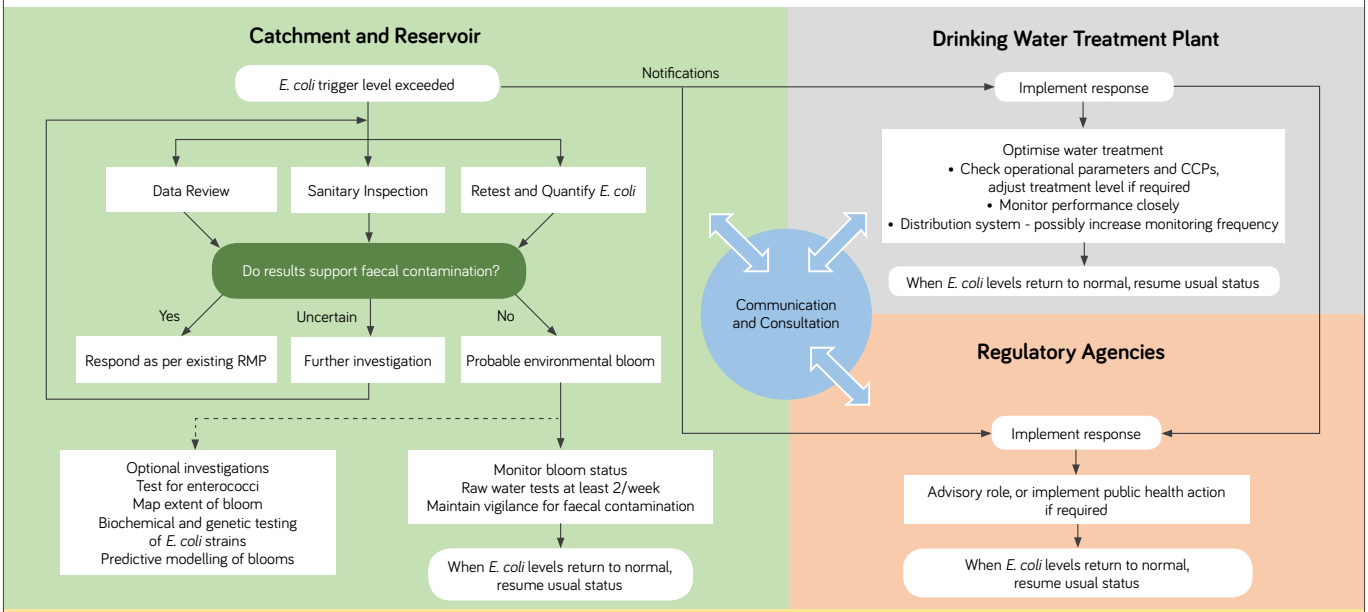
Nanayakkara, B.S., C.L. O'Brien and D. Gordon (2019). Diversity and distribution of Klebsiella capsules in *Escherichia coli*. *Environmental Microbiology* Reports. doi:10.1111/1758-2229.12710

Utility Response Protocol for Management of Environmental *E. coli*. Water Research Australia (2019)

Stay ahead of the curve

Access the Protocols and our extensive Knowledge Library of topic-specific resources
waterra.com.au/publications/fact-sheets/

Utility Response Protocol



Appendix 2 Utility Response Protocol

Utility Response Protocol for Management of Environmental *E. coli* Blooms

This Protocol has been developed to assist the water industry's response to blooms of environmental *E. coli* strains in water storages used for drinking water supply. This document should be read in conjunction with the Fact Sheet – Management of Environmental *E. coli* Blooms. Accurate identification of such blooms is essential to ensure correct classification of the catchment risk level under the proposed health-based targets system for the Australian Drinking Water Guidelines.

The Protocol was developed from information provided by six Australian water utilities that have experienced environmental *E. coli* blooms within the last 10 years. Perspectives from three state health regulators about public health concerns relating to the management of environmental blooms have also been incorporated.

The Protocol outlines a basic level of response to rule out faecal contamination as the source of elevated *E. coli* counts, and ensure drinking water safety is not compromised. In addition, there are optional elements which may be added if a more extensive investigation is desired for the first event, or if subsequent blooms occur in the same reservoir. The use of optional elements may also be determined by the type of evidence required by health regulatory agencies to make a determination that the event is environmental in origin.

The Utility Response Protocol assumes that water utilities are already operating under the principles embodied in the Australian Drinking Water Guidelines Framework for Management of Drinking Water Quality, and that procedures exist for risk assessment and management of individual water supply systems. It is also assumed that in situations where different organisations are responsible for management of the catchment/reservoir and the drinking water treatment plant (DWTP), a cooperative relationship exists which facilitates communication and investigation of the reasons for changes in the microbial quality of raw water. This protocol, in effect, provides an additional 'branch pathway' to be added to existing protocols for investigation of elevated *E. coli* numbers detected by raw water monitoring programs.

A Bayesian Network Model has been developed to predict the likelihood of blooms based on historical data for water quality, weather conditions, occurrence of predisposing events (bushfires/ planned burns, dust storms or algal blooms) and the rate of change in reservoir levels. This model can be used to run scenarios based on catchment/reservoir conditions to predict the likelihood of the occurrence of a bloom, provided that sufficient data are available. If a high likelihood is predicted, a more intensive sampling campaigns focused on the identified predictors could be triggered, in order to collect more information and to improve the understanding and modelling of future environmental *E. coli* blooms.

References

Management of Environmental *E. coli* Blooms - Fact Sheet (2019) Water Research Australia.

http://www.waterra.com.au/publications/fact-sheets/

Bayesian Network Model. Contact Water RA for further information

Basic Protocol

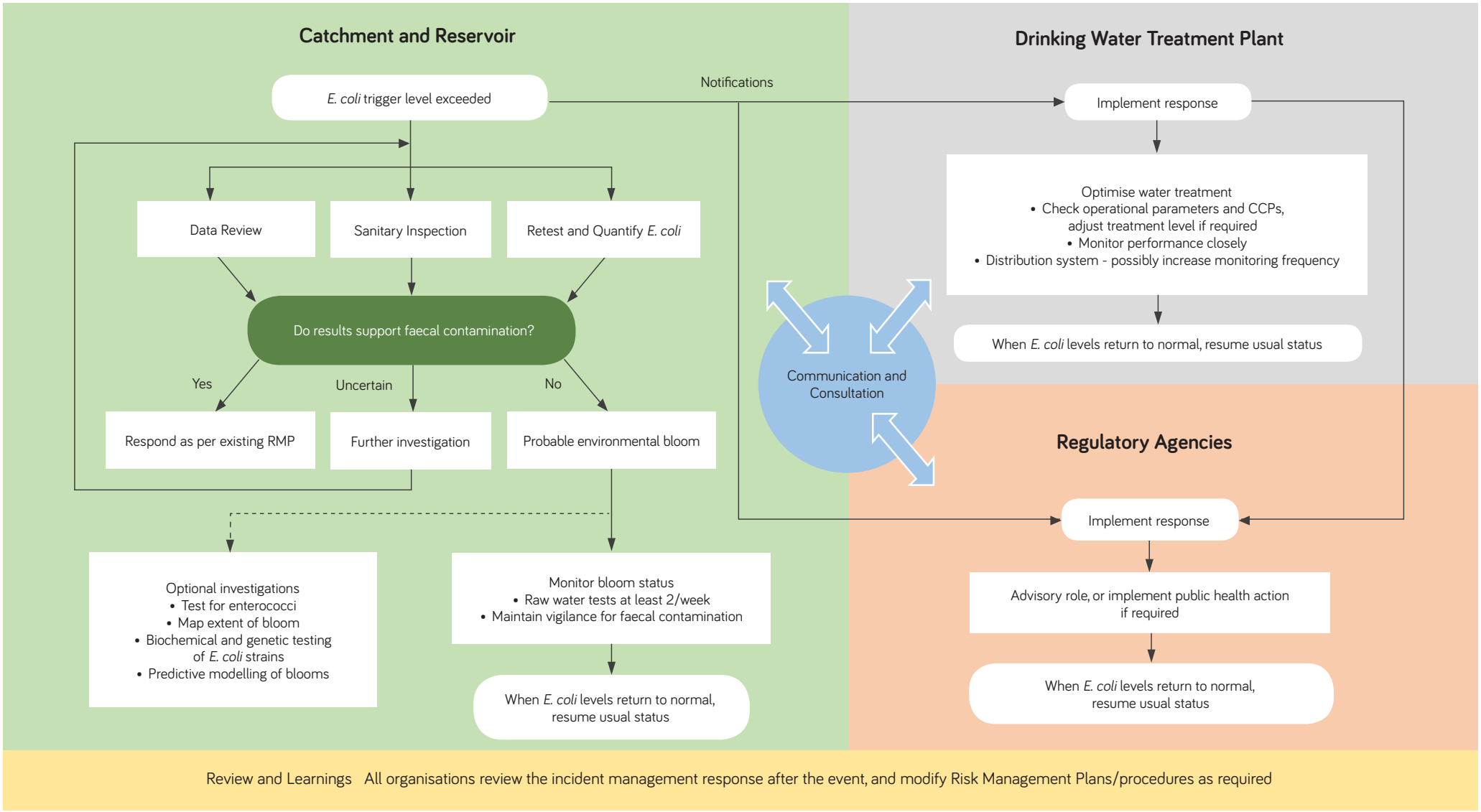
Element	Description
***E. coli* trigger level**	A value that would trigger investigation of presumed faecal contamination should be defined for the *E. coli* level in raw water entering the treatment plant (i.e. samples collected at the water offtake or the raw water tank at the treatment plant). If a trigger level does not already exist, considerations for setting such a level might include: • a value based on the 'high end' of routine *E. coli* monitoring results • a value based on the assessed log reduction capability of the drinking water treatment plant taking into account potential bacterial levels in raw water, from historic monitoring results • a boundary value for health-based-target catchment classifications Depending on the existing routine monitoring program, it may also be desired to define trigger levels for locations in the reservoir.
Data review and sanitary inspection	Established procedures for investigation of potential faecal contamination events should be implemented. These are likely to include a review of recent water quality monitoring results and weather data, and sanitary inspection of the catchment and reservoir.
Resampling to confirm and quantify *E. coli* level	Samples should be collected at the site(s) where the initial high *E. coli* reading(s) were recorded. Additional samples should be taken at other locations as part of normal investigation procedures and at any suspect locations identified by the data review and/or sanitary inspection. If laboratory testing for *E. coli* does not currently include quantification of high concentrations when all test wells are positive, then this should be requested for the confirmatory samples. Consideration should also be given to instructing the laboratory to retain the incubated test trays and/or plates to permit strain characterisation, should this be desired at a later stage.
Decision point. **Do the results support faecal contamination?**	If the high *E. coli* count is confirmed and the data review and sanitary inspection reveal a likely source of faecal contamination, the investigation and operational response should proceed according to existing risk management plans for the water supply. If the results are conflicting (e.g. high *E. coli* count not confirmed, but data review and sanitary inspection indicate a likely source of faecal contamination) then further sampling and investigations should be undertaken. If the high *E. coli* count is confirmed but the data review and sanitary inspection do not show any evidence of a source or event that could have caused faecal contamination, this should be taken as preliminary evidence of an environmental bloom.
Notifications and communication	Notifications about unusually high *E. coli* levels in raw water entering the treatment plant should be made under existing legislative obligations, MOUs or other formal agreements with regulatory agencies and other organisations. Depending on usual practices, contact may also be made with regulatory agencies and other organisations to advise them of the situation, even if notification is not mandatory. The health regulator should always be advised if there is doubt about the capability of the DWTP to provide safe water. This is a legislated requirement under some state and territory laws. Communication and consultation between all involved organisations should be maintained throughout the incident.
Drinking Water Treatment Processes (DWTP)	The operational parameters and critical control points of the DWTP should be checked and performance optimised. Depending on judgement regarding the performance of the plant, consideration may be given to enhancing coagulation and/or increasing disinfection above usual levels. The operation of drinking water treatment processes should be closely monitored until *E. coli* levels in raw water return to the normal range. For conventional treatment this would include monitoring compliance with targets for individual filter bed turbidity and for disinfectant levels. If the log reduction capability of the DWTP is considered inadequate or borderline, consideration should be given to cease harvesting the raw water source temporarily, or relocating the raw water offtake location or level to avoid areas with high *E. coli* counts, if this is feasible.

Monitoring in distribution system	Additional monitoring in the distribution system for disinfection residual and other parameters (e.g. turbidity for unfiltered systems) may be advisable. If contamination of the distribution system is suspected, consideration for issuing a boil water advisory to eliminate other bacterial pathogens potentially present may be needed. In this situation, the health regulator should be consulted.
Monitor bloom status	Raw water monitoring (at least for water entering the treatment plant) should be increased to twice per week or more. Normal levels of vigilance for faecal contamination events in the catchment should be maintained throughout the bloom event.
Return to normal status	When *E. coli* levels return to the normal range, return to usual operational status.
Review and learnings	All involved organisations should review their incident management response after the event. If necessary, modifications should be made to Risk Management Plans to incorporate/refine bloom response information, trigger levels, communication protocols etc., in order to optimise the response to any future bloom events.

Optional Bloom Investigations

Investigation type	Description
Enterococci testing	Water samples may be tested for enterococci as well as *E. coli* to provide a comparative marker for faecal contamination. Addition of this test to the routine *E. coli* monitoring program for raw water entering the drinking water treatment plant should be considered if a second bloom event occurs in a reservoir.
Map extent of bloom	The extent of the bloom within the reservoir may be mapped by sampling at different depths in several different locations and testing for *E. coli*. Counts should be accurately quantified (this may require requesting that laboratory samples be diluted). This information will allow calculation of the amount of sewage effluent or animal waste that would have been required to produce the observed volume of contaminated water. This may provide evidence that a contamination event of this magnitude is not feasible, or at least extremely unlikely to have occurred in the absence of evidence from the sanitary inspection and the data review.
Genetic and/or biochemical testing of *E. coli* isolates **Some of these tests can be performed by extracting DNA (potentially a mix of strains from multiple E. coli cells in the water sample) directly from the wells on test trays, while others require plating of strains to obtain pure colonies for testing.**	Tests for strain diversity or strain origin: • genetic and/or biochemical typing of several isolates to characteristic strain diversity. These tests detect the DNA sequence or biochemical activity of a number of genes used for basic metabolic functions shared by most *E. coli* strains. During environmental *E. coli* blooms a small number of strains will dominate the population, and diversity will therefore be low. However, if faecal *E. coli* contamination has occurred, many strains will be present, and high diversity will be seen. Note that *E. coli* bloom strains may produce atypical results on some biochemical tests (e.g. chromogenic media) and some do not grow at the elevated incubation temperature (44.2°C) used in some tests. • microbial source tracking may provide evidence that isolates are not of faecal origin. Tests to identify bloom-forming strains: • membrane filtration enumeration or streak plating to observe mucoid colony morphology consistent with capsule production. • specific multiplex PCR tests to confirm the presence of the Group 1 capsule genes and determine whether the isolates have capsule types matching those found in previously documented blooms.
Predictive modelling of blooms	To gain a better understanding of the factors which cause blooms, the Bayesian Network Model could be used to define a likelihood level that would trigger more intensive water quality monitoring and/or testing for additional parameters.

Utility Response Protocol



Water Research Australia Limited
ACN 127 974 261
Level 2, 250 Victoria Square, Adelaide SA 5000
GPO BOX 1751, Adelaide SA 5001
+61 8 7424 2445
info@waterra.com.au
waterra.com.au

Water Research AUSTRALIA