CHAPTER 2: LITERATURE REVIEW

2.1 Water Quality and Sanitation

Water is a natural resource that has an essential role in all aspects of life. Water resources include lakes, rivers and underground aquifers that may be used for human consumption, recreation or agricultural applications. Maintaining and protecting the world's water sources has been a big challenge for many countries. Accessibility of safe drinking water, particularly among the low-income communities is still a problem in developing countries including Malaysia. Human interactions with aquatic systems increase the cause for concern of potential health related risks. Poor waste disposal mechanisms in both urban and rural areas contribute in the pollution of water sources. Thus, in order to improve the nation's water quality, a closer examination of the water resources, sources of contamination, and methods to prevent contamination is needed (Malakoff, 2002; Simpson *et al*., 2002).

2.2 Microbial Contamination

The most common and deadly pollutants in the water in developing countries are of biological origin. WHO (2004) states that the infectious diseases caused by pathogenic bacteria, viruses and protozoa are the most common and widespread health risk associated with drinking water. Some of the major emerging and re-emerging water-borne agents are *Vibrio cholerae* responsible for cholera, *E. coli* 0157:H7 responsible for most dysentery cases even in developed countries and *Salmonella typhi* which is

responsible for typhoid fever. Waterborne infectious diseases are transmitted primarily through contamination of the water sources with excreta of humans and animals who are either active cases or carriers of the disease. Use of such water for drinking or cooking, contact with it during washing or bathing, or even inhalation of its fine droplets as aerosols, may result in infection.

2.3 Water quality indicator

The concentration of the water borne pathogens in groundwater and natural streams from fecal contamination is generally small and the number of different disease causing organisms that could be present is large. It is not practical to test for all of these pathogens. Consequently, the presence of pathogens is determined with indirect evidence by testing for an indicator organism such as coliform bacteria.

Total coliform which may originate from many sources has been used as an indicator of water contamination (Strauss *et al*., 2001). The traditional role of indicator parameters in drinking water was as an index of fecal pollution. The original microbial parameters were all bacteria that were derived from fecal contamination. Fecal indicator bacteria should fulfill certain criteria to give meaningful results. They should be universally present in high numbers in the feces of humans and other warm-blooded animals, should be readily detectable by simple methods and should not grow in natural water.

Analysis for fecal indicator bacteria provides a sensitive indication of pollution of drinking-water supplies. However this is not the most rapid detection as the growth

medium, the conditions of incubation and the nature of the water sample can influence the species isolated and the count of colonies. Conflicting reports of the comparison of fecal indicators and pathogenic organisms lead to a closer examination on effects on human and animal health (Meays *et al*., 2004). The basic idea behind the use of traditional fecal indicator parameters is when they are absent, the pathogens are absent. It is not universally valid because of the low accuracy but is still applied and useful today if the parameter is chosen correctly.

2.4 *Escherichia coli*

Escherichia coli first described by Theoder Escherich in 1885, is a member of family *Enterobacteriaceae* which is present as normal flora in the lower intestine of both humans and animals (Ewing, 1986). *E. coli* is a gram negative, lactose positive, mostly non-pathogenic, rod shaped and usually motile by peritrichous flagella. It produces different kinds of fimbriae that vary structurally among strains, which are important during the adhesion of host cells (Percival *et al*., 2004).

E. coli normally colonizes an infant's gastrointestinal tract within 40 hours of birth, arriving with food or with the individuals handling the child. In the bowel, it adheres to the mucus of the large intestine (Percival *et al*., 2004). It is the primary facultative organism of the human gastrointestinal tract. As long as these bacteria do not acquire genetic elements encoding for virulence factors, they remain harmless.

Different strains of *E. coli* are often host-specific, making it possible to determine the source of fecal contamination in environmental samples (Sussman, 1985). Depending on which *E. coli* strains are present in a water sample, assumptions can be made about whether the contamination originated from a human, mammal or bird source. New strains of *E. coli* evolve through the natural biological process of mutation, and some strains develop traits that can be harmful to a host animal. Pathogenic *E. coli* strains typically cause no more than diarrhea in healthy adult humans. However, virulent strains such as O157:H7 or O111:B4, can cause serious illness or death in the elderly, the very young or the immunocompromised.

2.5 Pathogenic *E. coli*

De *et al*. (1956) was the first to identify enterotoxins secreted by *E. coli* which was able to mediate increased fluid secretion in gut leading to diarrhea. Nataro *et al*. (1998) have classified diarrhea causing *E. coli* into six groups based on virulence properties of *E. coli*, their difference in epidemiology and distinct O:H serotypes. The six groups are Enterotoxigenic *E. coli* (ETEC), Enteropathogenic *E. coli* (EPEC), Enteroinvasive *E. coli* (EIEC), Enterohaemorrhagic *E. coli* (EHEC) or Verocytotoxigenic *E. coli* (VTEC), Enteroaggregative *E. coli* (EAEC) and diffusely adherent *E. coli* (DAEC).

The ETEC strains produce plasmid mediated enterotoxins which bind to their specific receptors on the gut epithelium and by a complex interplay of biological mediators cause diarrhea. Two main classes of enterotoxins are identified as heat-labile (LT) enterotoxin and heat-stable (ST) enterotoxin. There are two major subtypes of LT, LT1 and LT2

which do not cross react immunologically. LT1 is expressed by *E. coli* strains that are pathogenic for both humans and animals. LT2 is found primarily in animal *E. coli* isolates and rarely in human isolates but in neither animals nor humans has it been associated with disease.

Heat stable enterotoxin has two subtypes, STa and STb. STa toxins are produced by ETEC and several other gram-negative bacteria including *Yersinia enterocolitica* and *V. cholerae* non-O1. STb has been found only in ETEC (Levine, 1987). ETEC strains were first recognized as causes of diarrhoeal disease in piglets, where the disease continues to cause lethal infection in newborn animals (Alexander, 1994). The first descriptions of ETEC in humans reported that certain *E. coli* isolates from the stools of children with diarrhea elicited fluid secretion in ligated rabbit intestinal loops (Taylor *et al.*, 1961).

Konowalchuk *et al*. (1977) identified a distinct group of *E. coli* named as Verocytotoxigenic *E. coli* (VTEC), which had the ability to produce a toxin with profound and irreversible effect on vero cells. The toxin differed from heat-labile (LT) and heat-stable (ST) enterotoxins of *E. coli* and was cytotoxic only for vero cells. The term shiga-like toxin (SLT) was also applied to verotoxin (VT) as it has similar structure and biological activity. Therefore, SLT and VT nomenclature systems have been used interchangeably. Thus, verotoxin-producing *E. coli* is also termed as shiga-like toxin producing *E. coli* (SLTEC) or shiga-toxin producing *E. coli* (STEC).

2.6 Molecular Approach

The conventional microbial parameter detection involves sampling and filtration followed by cultivation of the chosen microorganism on selective media and colony counting. It is a process that can take 24 to 72 hours and may not pick the desired microorganisms. The development of molecular biology shorten the time to less than 8 hours. This resulted in techniques, such as polymerase chain reaction (PCR), for the rapid, sensitive and specific detection of indicator microorganisms and pathogens. Several researchers have developed PCR techniques for the rapid detection of *E. coli* and coliforms, which make detection possible within several hours (Fricker and Fricker, 1994).

One of the challenges for molecular methods is to assess the viability of the detected microorganisms. They detect the presence of a nucleic acid sequence, which may have originated from a dead organism or even from DNA that has not been decomposed in the aquatic environment. Culture techniques provide this information as only viable microorganisms are detected. Several techniques or combinations of techniques are now available to overcome the viability problem. For examples, the use of inducible mRNA as target for reverse transcriptase PCR (RT-PCR).

2.7 Detection of pathogenic strains

2.7.1 Phenotypic Assays

There are no suitable biochemical markers for identification of ETEC and distinguished from other pathogenic *E. coli* strains. Detection of ETEC has long relied on detection of the enterotoxins LT and/or ST. ST was initially detected in a rabbit ileal loop assay (Evans *et al.*, 1973) but the expenses and lack of standardization caused this test to be replaced by the suckling mouse assay (Gianella, 1976).

Several immunoassays like RIA, ELISA have been developed for detection of ST. The traditional bioassay for detection of LT involves the use of cell culture, either the Y1 adrenal cell assay or the Chinese hamster ovary (CHO) cell assay. Immunological assays for LT include Biken test, ELISA, latex agglutination and reversed passive latex agglutination test (Nataro *et al*., 1998). One of the most useful phenotypic assays for the diagnosis of diarrhoeagenic *E. coli* is the HEp-2 adherence assay (Donnenberg *et al*., 1995). This assay provides the best ability to differentiate among all three adherent diarrhoeagenic categories of EPEC, EAEC, and DAEC (Vial *et al.*, 1988).

Bettelheim *et al*. (2003) have stated that VTEC can be detected by detection of verotoxins produced by the bacteria or of the genes associated with VT (verotoxin) production. Vero cell toxicity of verotoxins can be detected by observation of cytopathic effects on vero cells (Konowalchuk *et al.*, 1977). Serological (ELISA,

immunoblot and reversed passive latex agglutination) tests were developed for identification of verotoxins produced by bacteria.

2.7.2 Molecular Assays

Pathogenic *E. coli* strains were among the first pathogens for which molecular diagnostic methods were developed. Indeed, molecular methods remain the most popular and most reliable techniques for differentiating pathogenic strains from nonpathogenic members. Progress has been made both in the development of nucleic acid based probe technologies as well as PCR methods (Nataro *et al*., 1998). The use of DNA probes for detection of LT and ST enterotoxins in ETEC revolutionized the study of these organisms, replacing cumbersome and costly animal models of toxin detection (Moseley *et al.*, 1982). Since then, gene probes have been introduced for all diarrhoeagenic *E. coli* categories.

PCR primers have been developed successfully for several categories of diarrhoeagenic *E. coli*. ETEC strains were among the first pathogenic micro-organism for which molecular diagnostic techniques were developed. Moseley *et al.* (1982) found DNA probes useful in the detection of LT and ST encoding genes in stool and environmental samples. Since then, several advances in ETEC detection have been made but genetic techniques continue to attract the most attention and use.

A useful adaptation of PCR is the multiplex PCR assay (mPCR) (Stacy-Phipps *et al.*, 1995). In this method several PCR primers are combined with the aim of detecting one

of several different diarrhoeagenic *E. coli* pathotypes in a single reaction. Lang *et al.* (1994) developed a triplex PCR method to simultaneously amplify a heat-labile toxin sequence (LT), a shiga-like toxin 1 sequence (SLT-1) and a shiga-like toxin 2 sequence (SLT-2) from toxigenic strains of *E. coli*. While, Chattopadhyay *et al.* (2001) isolated and characterized STEC strains from animal, human and food products.

2.8 Microbial source tracking

Microbial source tracking is the use of chemical, microbiological, molecular, and other methods to determine the source of fecal contamination (Meays *et al*., 2004). Testing for the presence of chemicals, such as detergents and caffeine, has been used to determine if groundwater is contaminated by human sources. However, the concentrations of these must be very high to be effective and testing must be very close to the source (Meays *et al*., 2004). Species specific indicators are groups of bacterial strains that tend to be more prevalent in certain animal species. Differing nutrients and biological requirements within the host animal or environment cause changes in the bacterial population, and the use of source tracking technology will develop the link between have been used in tracking the source of bacterial contamination of water, soil, food products, and other resources (Griffith *et al*., 2003).

2.8.1 Phenotypic source tracking

The phenotypic methods have been in use for many years due to developing technology and an understanding of physical interactions. These methods examine the

physical attributes, biochemical products, and chemical requirements of microorganisms (Leung *et al*., 2004). Fatty acid analysis examines the composition of bacterial membranes and ratio comparisons are based on host animal flora. These were used to assess human and non-human contamination. However, some data shows that survival rates of these bacteria differ, and the ratio method does not provide an accurate account (Simpson *et al*., 2002; Meays *et al*., 2004).

The two most common biochemical source tracking methods are antibiotic resistance analysis (ARA) and carbon utilization (Griffith *et al*., 2003). ARA is based on the specific resistance patterns of microorganisms present in animal populations that are exposed to or inoculated with different antibiotics (Meays *et al*., 2004). Carbon utilization patterns develop as microorganisms adapt to the different food sources of their host animal (Stoeckel *et al*., 2004). Both of these require a library base of known patterns for the hosts, microorganisms, and strains of interest.

2.8.1.1 Antibiotic resistance analysis (ARA)

Antibiotic resistant in bacteria are a natural evolutionary phenomenon and bacteria with intrinsic resistance to antibiotics are found in nature (Ash *et al*., 2002). Only those bacteria with the ability to adapt are those with the ability to survive (White *et al*., 2001). Koplin *et al.* (2002) suggested that the rate at which pathogenic bacteria develop resistance to antibiotics is affected by even low-level concentrations of antibiotic residues present in the environment. Thus antibiotic residues could result

in serious threats to public health as more bacterial infections become resistant to treatments using presently known antibiotics (Hirsch *et al*., 1999).

Repeated and continuous use of antibiotics also creates selection pressures that favor the growth of antibiotic resistant mutants (Dizidic *et al*., 2003). The increased frequency and spectrum of antibiotic resistance has also been attributed to social and technical changes that increase the transmission rate of resistant organisms. These changes include an increase in the use and accessibility of antimicrobial agents and medically invasive procedures (Dizidic *et al*., 2003).

2.8.2 DNA-based source tracking

Several different molecular methods have been used in microbial source tracking, with varying success (Griffith *et al*., 2003; Leung *et al*., 2003). The library-based methods include Amplified Fragment Length Polymorphisms (AFLP), Repetitive Extragenic Palindromic-PCR (REP-PCR), ribotyping, and Pulsed-Field Gel Electrophoresis (PFGE).

The AFLP procedure utilizes species specific adapter and restriction site sequences that serve as primer target sites (Vos *et al*., 1995). Rep-PCR examines the fragments located between repeated sequences (REP, BOX, or ERIC) in genomic DNA (Meays *et al*., 2004). Ribotyping with one or two specific restriction enzymes examines fragment variances in the 16S ribosomal sequence. While PFGE is a more sensitive

and intricate method that observes fragment variances in the whole genome (Stoeckel *et al*., 2004).

The library-independent methods include host-specific PCR and t-RFLP which are specific to the host population (Simpson *et al*., 2002). Host-specific PCR relies on the length differences of host-specific genetic markers in genomic DNA or 16S rDNA of intestinal microorganisms, while t- RFLP focuses on the sizes of terminal end fragments (Meays *et al*., 2004). PCR-based fingerprinting methods have been used for several years. However many problems have arisen such as slight temperature changes, DNA and primer concentration, DNA quality, and type of polymerase all can affect the efficiency and reproducibility (Janssen *et al*., 1996).

2.9 Rep-PCR

The rep-PCR method utilizes repeated, conserved, natural sequences within the genome of bacteria to provide strain-specific fingerprints (Dombek *et al*., 2000). Three of these repeated sequences have been identified and employed in molecular methods, including repetitive extragenic palindromic (REP), enterobacterial repetitive intergenic consensus (ERIC), and BOX sequences (Versalovic *et al*., 1994). Primers for the repeated sequences are not specific for a given organism, and no information is needed about the genome of interest. When compared to ribotyping, the two-step process of rep-PCR and gel electrophoresis allows for relatively non-complex and quick results (Carson *et al*., 2003). Rep-PCR is also the preferred method due to reproducibility of products and superior banding patterns (Carson *et al*., 2003). Rep-PCR is valuable when a large

collection of possible sources is available for comparison. However, this also limits the technique to use in small localities (McLellan *et al*., 2003).

BOX-PCR has produced better results than REP-PCR when comparing various animal- and human source bacterial fingerprints (Dombek *et al*., 2000). Kon *et al*. 2007 utilized REP-PCR to characterize the genetic diversity of environmental bacteria from shores of Lake Huron, Ontario Canada based on fingerprint libraries. Rep-PCR has also been used to accurately cluster the highly pathogenic O157:H7 strain of *E. coli* from various contaminated processed meat sources (Hahm *et al*., 2003). The equipment, personnel, and time resources required for AFLP and REP-PCR has been utilized in the attempt to identify *E. coli* from cattle sources in environmental water samples.

2.9.1 Advantages of rep-PCR Analysis

Rep-PCR genomic fingerprinting has been found to be extremely reliable, reproducible, rapid and highly discriminatory (Versalovic *et al*., 1994). The rep-PCR analysis has been shown to be as sensitive as other genotyping methods such as ribotyping and pulsed-field gel electrophoresis (Jordens *et al.,* 1995). The presence of multiple copies of highly conserved rep sequences in DNA has been identified in a number of microorganisms. The technique of rep-PCR takes advantage of the fact that multiple copies of these rep sequences may be found randomly distributed throughout the genomes of strains of a bacterial species. Thus, by employing a portion of a highly conserved rep sequence as a primer, it is possible to amplify DNA sequences that are

located between closely situated pairs of correctly oriented rep elements. Rep-PCR has been successfully used to distinguish between strains of *E. coli* (Dombek *et al*., 2000).