



# Journal of Applied Sciences

ISSN 1812-5654

**science**  
alert

**ANSI***net*  
an open access publisher  
<http://ansinet.com>

## Metagenomic Analysis of 16S rRNA Sequences from Selected Rivers in Johor Malaysia

<sup>1,2</sup>Topik Hidayat, <sup>1</sup>Mohd. Aszuan Abdul Samat, <sup>1</sup>Muhamad Atiq bin Elias and <sup>3</sup>Tony Hadibarata

<sup>1</sup>Department of Biological Sciences, Faculty Bioscience and Bioengineering,  
Universiti Teknologi Malaysia, Johor Bahru, Malaysia

<sup>2</sup>Department of Biology Education, Indonesia University of Education, Bandung, Indonesia

<sup>3</sup>Institute of Environmental and Water Research Management, Universiti Teknologi Malaysia, Malaysia

**Abstract:** Index for water quality is solely based upon the conditions of physical and chemical which provide quantitative data on the presence and level of aquatic pollution. However, the parameters for the assessment do not represent the environmental stress and ecological health of the river which affect the microbial diversity. In this study, we introduced a new approach of assessing river water from metagenomic view by constructing phenetic tree which show relation between microbial community and the level of pollution. Four river waters, of which one river is categorised as unpolluted water, located around Johor Bahru (Malaysia) were examined and compared. Genomic DNA of these uncultured community of microorganisms was directly extracted and 16S rRNA gene was amplified by PCR using a set of primer pairs to generate cloned libraries. In total 24 isolates which are consisted of 18 isolates from polluted rivers and six from unpolluted one, were sequenced. Along these sequences, six of 16S rRNA sequences of colliform bacteria obtained from genbank were placed in Operational Taxonomic Units (OTUs). Phenetic analysis revealed that the river water used in this study were classified into two groups, representing the polluted and unpolluted. Since the tree can clearly distinct unpolluted river water from that of polluted one, we tried to develop a putative sequences motif for both condition. In practical purpose, the sequences motif can be used to screen river water quality in certain river water environment.

**Key words:** Metagenomic, microbial community, molecular screening, phenetic analysis, sequence motif

### INTRODUCTION

The ecological distress due to the contaminated freshwater sources by human activity leads to the importance of monitoring the freshwater status. Water quality assessments are usually carried out by doing the physico-chemical analysis, represented by Water Quality Index (WQI) (Agarwal and Rajwar, 2010; Nasab *et al.*, 2010) and sometimes include the microbiological study (Trivedy and Goel, 1984). However, this physical and chemical measurement does not reflect the extent of environmental stress reaching the living organisms or the subsequent effects of the stress towards the organisms in that environment (Kerkhof, 2009; Maznah and Omar, 2010; Oparaku *et al.*, 2011). The use of metagenomics approach has offered in comprehensive environmental assessment of microbial community in water environment (Cottrell *et al.*, 2005; Marshall *et al.*, 2008), reflecting of microbial diversity under environmental stress.

Metagenomics, or also known as environmental genomics, is study of collective genomes of microorganisms. By using metagenomics, we are able to

perform genomic analysis of microorganisms by direct extraction and cloning of DNA from an assemblage of microorganisms, collected from various types of environment such as soil, river, feces and marine (Handelsman, 2004; Bhuiyan *et al.*, 2011). This is therefore unnecessary to collect and culture the microorganisms and allows us to avoid the bias imposed by culturing and led to the discovery of vast new lineages of microbial life (Daniel, 2004; DeLong *et al.*, 2006). It is now widely accepted that the application of standard microbiological methods for the recovery of microorganisms from the environment has had limited success in providing access to the true extent of microbial biodiversity. It follows that much of the extant microbial genetic diversity remains unexploited, an issue of considerable relevance to a wider understanding of microbial communities. The recent development of molecular biology designed to access this wealth of genetic information through environmental nucleic acid extraction has provided a means of avoiding the limitations of culture-dependent genetic exploitation (Cowan *et al.*, 2005).

**Corresponding Author:** Topik Hidayat, Department of Biological Sciences, Faculty Bioscience and Bioengineering, Universiti Teknologi Malaysia, 81310, Skudai, Johor Bahru, Malaysia  
Tel: +(6)07-5532715 Fax: +(6)07-5531112

In this study, we conducted metagenomic analysis of 16S rRNA gene, in particular, to evaluate utility of the sequences information in assessing river water quality. With its high level of conservative sequences, the 16S rRNA gene has been the main source of genetic material used to study comparative genomic analysis since long (Weisburg *et al.*, 1991; Devereux and Mundfrom, 1994; Farrelly *et al.*, 1995). Thus, we compared metagenomic sequences from different site of river water (polluted and unpolluted) and subjected them to phenetic analysis. Phenetic relationships groups the organisms together based on overall similarity of their phenotypic (nucleotides) characteristic which are observable. Molecular approach to study water quality can widen our perspective in assessing pollution level of the river, for instance, by reconstructing a sequences motif and oligonucleotide primers/probes which have the potential utility to assess rapidly the environmental health condition and status in the future application.

#### **MATERIALS AND METHODS**

Four rivers, namely, Kota Tinggi (CN), Sayong (SY), Skudai (SK) and Kempas (KM) were selected. This selection was made based upon the information issued by local government (Department of Environment/DOE and Department of Irrigation and Drainage/DOIAD, Johor, Malaysia). According to this report, CN is subjected to unpolluted, whereas the rest of three rivers are polluted. A water sample was collected from a depth of 50 cm in July 2011.

Genomic DNA from the selected rivers was extracted by using the CTAB method developed by Marshall *et al.* (2008) with a little modification. About 5 L water samples were filtered by using aspirator pump. The samples were filtered through 0.2 mm microfibre filter paper that was placed in 2 mL CTAB extraction buffer before incubate them in water bath for 70°C for at least 1 h. Then, the filter papers were removed from the buffer. Each 1 mL of the buffer was transferred to the 2 mL tube. Then, 1 mL 24:1 (v/v) Chloroform:Isoamyl alcohol was added, the tube mix by inverted several times before centrifuge for 10 min at 13,200 rpm. The supernatant was transferred into the new 1.5 mL tube, 0.7 volumes of 100% cold 2-Propanol were added, the tube mixed by inverting several times. The tube was subsequently centrifuged for 30 min at 13,200 rpm at room temperature. The supernatant were decanted, whereas the pellet was kept and leave air-dried by opened the tube about 10-15 min or until the pellet dried. The pellet was subsequently rehydrated in 25 µL TE buffer pH 8.0 and stored in -20°C for long term usage. In some cases, purification was performed with a HiYield™ Gel/PCR Mini Kit following the manufacturer's instructions.

The amplification of the 16S rRNA gene was carried out using the primer pairs pAf and pHr (Edwards *et al.*, 1989; Bruce *et al.*, 1992), forward and reverse, respectively. Each reaction included 5 µL 10x Standard Taq reaction buffer, 1 µL of each 10 µM forward and reverse primers, 1 µL 10 µM dNTPs, 39.75 µL sterile deionized water, 0.25 µL Taq DNA polymerase and 2 µL template genomic DNA. The PCR profile consisted of an initial 2 min premelt at 94°C and 30 cycles of 1 min at 94°C (denaturation), 1 min at 45°C (annealing) and 2 min at 72°C (extension), followed by a final 2 min extension at 72°C. PCR products were visualized by 0.8% of agarose gel electrophoresis.

Ligation was set up by mixing 5 µL of PCR product with 5 µL of 2 X rapid ligation buffer, 1 µL of T4 DNA ligase and 1 µL pGEM-T vector. The reaction were mixed by pipetting and incubated for overnight at 4°C in refrigerator. Before starting the transformation procedure, the incubator shaker was preheated at 37°C and water bath at 42°C. Ice also was prepared. The ligation mixtures were removed from the refrigerator and equilibrated to room temperature for 1 min and centrifuged briefly. Then 6 µL of each ligation mixture were added to 1.5 mL tube which has been pre cooled on ice. Then the DH5α Competent cells were removed from freezer and placed in a 50% ice/50% distilled water bath for 5 min. Then 150 µL of competent cells were added to the tubes on ice. The mixture then leaved on ice for 20 min. Then, heatshocked the cells for 45 sec at 42°C and returned to ice for 2 min. 300 µL of LB media were added to the tube and mixed by flicking gently. After that, the tube was closed completely [airtight] and put in incubator shaker at 37°C [optimum temperature for bacteria growing] for 90 min. During the incubation process, the LAIX plates were prepared and put in 37°C incubator oven to dry out for 30 min. After the plates were dried and transformation culture incubated, the plates then were taken to the laminar flow hood. Then, 150 µL of the transformation cultures were transferred onto the plates and spread by using hockeystick. The plates were sealed with parafilm and placed in 37°C oven for 20 h. After the LAIX plates incubation of 20 h, the colonies were screened to select white colonies instead of blue (possible positive colonies that have an insert of gene interest).

About 50% from total white colonies were selected and labeled. By using sterile pipette tips, each single white colony was picked and put a drop on the labeled grid plate (for PCR library construction) and then dipped in the 5 mL LB medium (for plasmid extraction). The grid plates then were incubated for 20 h and 37°C before stored at 4°C. Plasmid extraction was conducted using QIAmp Spin Miniprep following the manufacturer's instructions.

The restriction enzyme digestion was conducted using *EcoRI* to reconfirm whether the plasmid has the insert desired or not. Digestion mixture were added in 1.5 mL tube and mixed together by pipetting before adding *EcoRI*. The mixtures were gently mixed by pipetting and incubated for 4 h at 37°C. As many as 24 isolates/clones which consist of 18 isolates/clones from polluted river (SY, SK dan KM) and six from unpolluted one (CN) were sent to 1stBase for sequencing.

DNA sequences obtained were edited and assembled using CodonCode Aligner (<http://www.codoncode.com/aligner/>). Multiple alignments were conducted using ClustalX ([www.clustal.org/](http://www.clustal.org/)). In total 30 OTUs (of which 24 obtained from this study plus six sequences retrieved from Genebank) was subjected to phenetic analysis. The 30 aligned sequences were used to construct the phenetic tree with MEGA version 4 ([www.megasoftware.net](http://www.megasoftware.net)). After the construction of the phenetic tree, a putative DNA sequences motifs for both unpolluted and polluted group was developed from the consensus sequences which were identified by GENEIOUS ([www.geneous.com/](http://www.geneous.com/)).

## RESULTS and DISCUSSION

The aligned 16S RNA gene comprised 1,544 characters. Phenetic analysis using UPGMA (Unweighted Pair Group Method with Arithmetic Mean) clustering method resulted in a single tree (Phenogram) as depicted in Fig. 1. The tree demonstrated that the OTUs comprises two groups which represent two conditions of river, unpolluted and polluted group. The polluted group is housed by mixtures among three examined rivers (SY, SK and KM) and the existence of the coliform bacteria (*Klebsiella* spp., *Hafnia* spp., *Serratia* spp., *Enterobacter* spp., *Citrobacter* spp. and *Escherichia coli*) supported the grouping.

The term “sequences motif” here refers to DNA sequences of the 16S rRNA gene present in the sample that is unique and can be used to capture and record environmental condition under ecological surrounding in particular time. This term can be compared with other term such as “metagenomic profile”. As the tree can

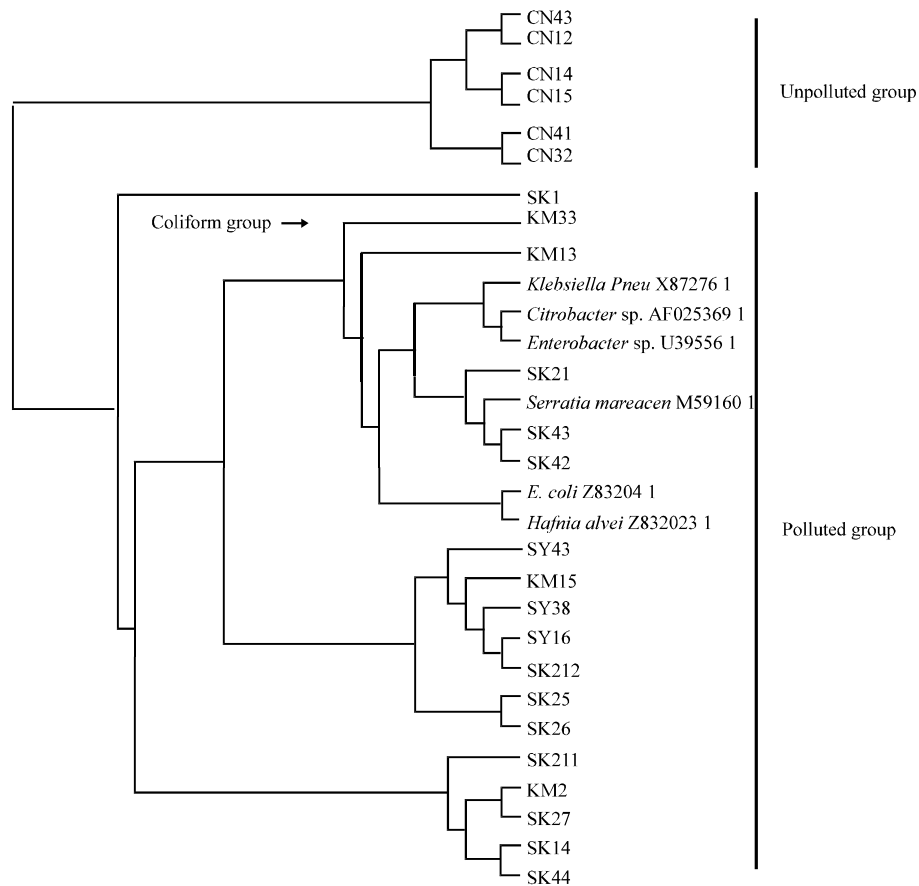


Fig. 1: Phenetic tree derived from 24 isolates which were consisted of 18 from polluted rivers (SY, SK and KM) and six from unpolluted one (CN). Six coliform bacteria were embedded in polluted group

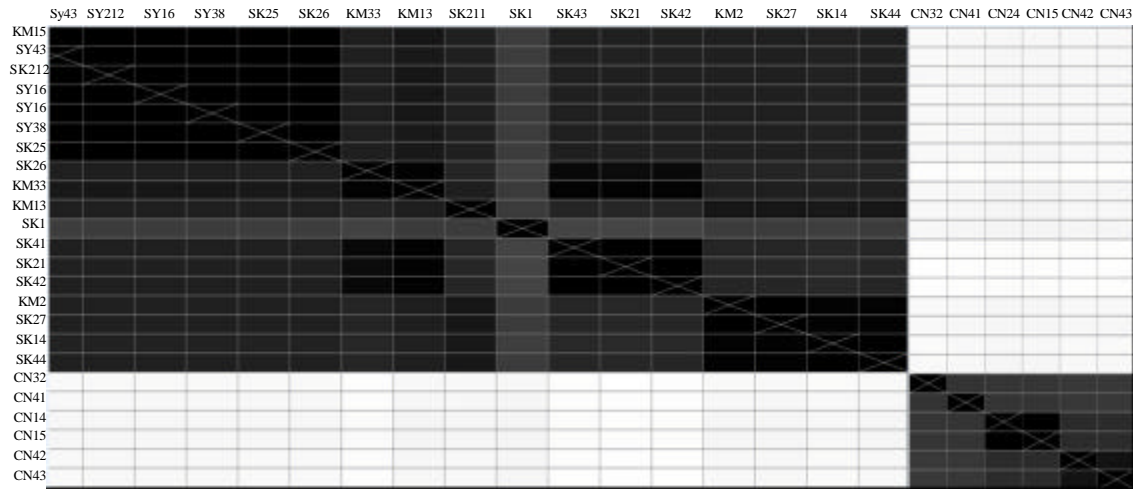


Fig. 2: The heat map of 24 isolates generated by GENEIOUS software. Distances are represented by shades of gray. A lighter gray means further apart while darker means closer together. The longest distance in the tree is white and the shortest distance is black (may not be zero)

clearly distinguish unpolluted river water from that of polluted one, a putative sequence motif was made, as can be seen in Fig. 3 and 4.

This study focused on determining the overall similarity of microbial communities among different level of river water pollution. They were compared with each other and discerned a significance differences at molecular level. The phenetic tree (Fig. 1) separated distinctly unpolluted bacterial groups from those polluted group. This grouping is consistent with what has been reported by local government (DOE and DOIAD) that three rivers, notably, SY, SK and KM are polluted, whereas CN is unpolluted.

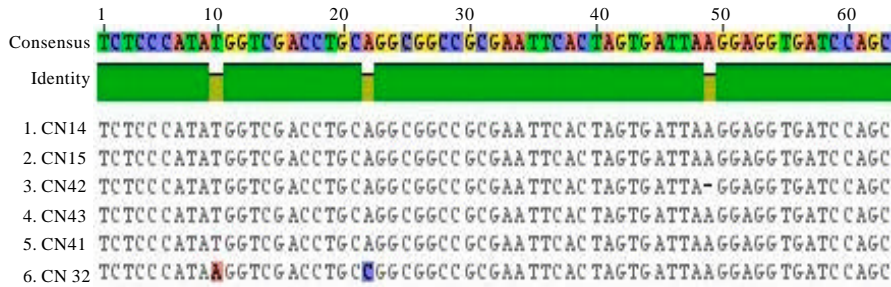
To confirm this, we included the 16S rRNA sequences of the coliform bacteria from GeneBank in the phenetic tree reconstruction. As result, all the coliform species used were housed in polluted group. The reason why we included the gene sequence from coliforms is because the coliforms always refer as indicator organisms when assessing water quality (Kovacs, 1992; Ghandhari and Alavi Moghaddam, 2011; Durai and Rajasimman, 2011). The coliform bacteria have been served as bioindicator for potential presence of disease-causing bacteria in water. This group of bacteria is common in soil and surface water as it also abundance in waste from humans and animals. The presence of coliform indicates that the water were polluted either cause of water discharge from farm, drainage or industrial waste.

Phenetic relationship shows that the organisms that share the same characteristics will group together. In this case, perhaps, the organisms have the same characteristics at molecular level which will help them to adapt at certain environment (Herrera and Cockell, 2007).

The genetic similarity between OTUs was also determined by converting the results to the heat map (Fig. 2) and distance matrix (Table 1). The patristic distances were calculated by summing together the lengths of the branch between each pair of OTUs. The value of the distance matrix based on fixation index which the value range is in between 0 to 1. The value 0 indicates that the isolates were genetically identical while the value 1 indicates that the isolates were completely separated or different. Based on Fig. 3, it is clearly showed that the samples representative of unpolluted and polluted group were significantly different with each other due to the range value of unpolluted to polluted group were more than 0.9. The isolates which have the range within  $0.1 \pm 0.05$  can be considered in the same group (Kwoon, 2000). This result suggested that each of the microbial community which represent unpolluted and polluted group having different set genetic data of certain environment condition, either the microbial community have particular site of nucleotide or amino acid which enable them to survive in polluted condition or the deletion/insertion of certain DNA location may absence/presence due to the environmental stress.

Table 1: The patristic distance matrix of 24 isolates

Isolates	KM15	SY43	SK212	SY16	SY38	SK25	SK26	KM33	KM13	SK211	SK1	SK43	SK21	SK42	KM2	SK27	SK14	SK44	CN32	CN41	CN14	CN15	CN42	CN43
KM15	0.000	0.002	0.001	0.001	0.003	0.003	0.003	0.147	0.123	0.152	0.247	0.149	0.153	0.150	0.153	0.151	0.160	0.152	0.961	0.956	0.952	0.952	0.957	0.957
SY43	0.002	0.000	0.002	0.002	0.004	0.004	0.004	0.148	0.124	0.153	0.248	0.150	0.154	0.151	0.154	0.152	0.161	0.153	0.962	0.958	0.953	0.953	0.958	0.958
SK212	0.001	0.002	0.000	0.000	0.003	0.003	0.003	0.146	0.123	0.151	0.246	0.148	0.153	0.150	0.152	0.151	0.159	0.152	0.960	0.956	0.951	0.951	0.956	0.957
SY16	0.001	0.002	0.000	0.000	0.001	0.003	0.003	0.146	0.123	0.151	0.246	0.148	0.153	0.150	0.152	0.151	0.159	0.152	0.960	0.956	0.951	0.951	0.956	0.957
SY38	0.001	0.002	0.001	0.000	0.003	0.003	0.003	0.146	0.123	0.151	0.247	0.149	0.153	0.150	0.152	0.151	0.159	0.152	0.960	0.956	0.951	0.951	0.957	0.957
SK25	0.003	0.004	0.003	0.003	0.000	0.000	0.000	0.147	0.124	0.152	0.248	0.150	0.154	0.151	0.153	0.152	0.160	0.153	0.961	0.957	0.952	0.952	0.958	0.958
SK26	0.003	0.004	0.003	0.003	0.000	0.000	0.000	0.147	0.124	0.152	0.248	0.150	0.154	0.151	0.153	0.152	0.160	0.153	0.961	0.957	0.952	0.952	0.958	0.958
KM33	0.147	0.148	0.146	0.146	0.147	0.147	0.147	0.000	0.052	0.169	0.264	0.077	0.082	0.079	0.170	0.168	0.177	0.169	0.978	0.973	0.968	0.968	0.974	0.974
KM13	0.123	0.124	0.123	0.123	0.124	0.124	0.124	0.052	0.000	0.145	0.241	0.039	0.043	0.040	0.146	0.145	0.153	0.146	0.954	0.950	0.945	0.945	0.951	0.951
SK211	0.152	0.153	0.151	0.151	0.152	0.152	0.152	0.169	0.145	0.000	0.244	0.171	0.175	0.172	0.118	0.117	0.125	0.118	0.957	0.953	0.948	0.948	0.954	0.954
SK1	0.247	0.248	0.246	0.246	0.247	0.248	0.248	0.264	0.241	0.244	0.000	0.267	0.271	0.268	0.245	0.243	0.252	0.244	0.953	0.948	0.943	0.943	0.949	0.949
SK43	0.149	0.150	0.148	0.148	0.149	0.150	0.150	0.077	0.039	0.171	0.267	0.000	0.005	0.002	0.172	0.171	0.179	0.172	0.980	0.976	0.971	0.971	0.977	0.977
SK21	0.153	0.154	0.153	0.153	0.154	0.154	0.154	0.082	0.043	0.175	0.271	0.005	0.000	0.003	0.176	0.175	0.184	0.176	0.984	0.980	0.975	0.975	0.981	0.981
SK42	0.150	0.151	0.150	0.150	0.151	0.151	0.151	0.079	0.040	0.172	0.268	0.002	0.003	0.000	0.173	0.172	0.180	0.173	0.981	0.977	0.972	0.972	0.978	0.978
KM2	0.153	0.154	0.152	0.152	0.153	0.153	0.153	0.170	0.146	0.118	0.245	0.172	0.176	0.173	0.000	0.006	0.052	0.044	0.958	0.954	0.949	0.949	0.955	0.955
SK27	0.151	0.152	0.151	0.151	0.152	0.152	0.152	0.168	0.145	0.117	0.243	0.171	0.175	0.172	0.006	0.000	0.050	0.043	0.957	0.953	0.948	0.948	0.953	0.953
SK14	0.160	0.161	0.159	0.159	0.160	0.160	0.160	0.177	0.153	0.125	0.252	0.179	0.184	0.180	0.052	0.05	0.000	0.008	0.965	0.961	0.956	0.956	0.962	0.962
SK44	0.152	0.153	0.152	0.152	0.153	0.153	0.153	0.169	0.146	0.118	0.244	0.172	0.176	0.173	0.044	0.043	0.008	0.000	0.958	0.953	0.949	0.949	0.954	0.954
CN32	0.961	0.962	0.960	0.960	0.961	0.961	0.961	0.978	0.954	0.957	0.953	0.980	0.984	0.981	0.958	0.957	0.965	0.958	0.000	0.217	0.227	0.227	0.232	0.232
CN41	0.956	0.958	0.956	0.956	0.957	0.957	0.957	0.973	0.950	0.953	0.948	0.976	0.98	0.977	0.954	0.953	0.961	0.953	0.000	0.217	0.222	0.222	0.228	0.228
CN14	0.952	0.953	0.951	0.951	0.952	0.952	0.952	0.968	0.945	0.948	0.943	0.971	0.975	0.972	0.949	0.948	0.956	0.949	0.958	0.000	0.222	0.000	0.000	0.190
CN15	0.952	0.953	0.951	0.951	0.952	0.952	0.952	0.968	0.945	0.948	0.943	0.971	0.975	0.972	0.949	0.948	0.956	0.949	0.958	0.000	0.222	0.000	0.000	0.190
CN42	0.957	0.958	0.956	0.956	0.957	0.957	0.957	0.974	0.951	0.954	0.949	0.977	0.981	0.978	0.955	0.953	0.962	0.954	0.958	0.000	0.222	0.190	0.000	0.108
CN43	0.957	0.958	0.957	0.957	0.958	0.958	0.958	0.974	0.951	0.954	0.949	0.977	0.981	0.978	0.955	0.953	0.962	0.954	0.958	0.000	0.222	0.190	0.000	0.108



```

TCTCCCATATGGTCGACCTGCAGGCGGCCGCGAATTCAC TAGT GATTAAGGAGGTGATCCAGCAGCCASSTTCCCCTA
CGGCTACCTTGTACGACTTACCCCACTACTAAYCCTACCGTGGBCGGCGTCTCCTTGC GGTTAGACTACCGAC
TTCDGGYATGACCARCTCCCATGGTGTGACGGGCGGTGTGTACAAGGCCCGGGAACGTATTCACCGTGGCATGCTGA
TCCACGATTACTAGCGATTCCAACTTCACGCAGBCGAGTTGCAGACTRCGATCCGAACTGAGACNGBTTTTGGGA
TTAGCTCCACCTCGCGGTBTBGCWDCWCWYTGTC HCVGCTTGTAGCAGCTGTGTAGCCCTAACCRTAAGGGCCA
TGAGGACTTGACGTATCCCCACCTTCTCCGGTTTRTACCGGCAGTCTCTAGAGTGCCWAAATGTRGCAACTA
RAAACAGGGTTGCGCTCGTTGCGGGACTTAACCCAACTCTACGACACGAGCTGACGACAGCCATGCAGCACCTG
TGTTCVGYCCNKAACDGCACAGHTCCATCTCTGGTGCCTTCBGGACATGTCAAGGGTAGGTAAGGTTCTTCGGC
TTGACTCGAATAAACACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCCTTGAAGTTTAAATCTTGCACCG
TACTCCCCAGGCGGSTRCTTAAYGCGTTAGCTGCGYYACYRAGMNAGTVAABYCCRACARCTAGTASCCATCGTT
TACGGCGTGGACTACCAGGGTATCTAATCCTGTTTGCTCCCCACGCTTTCGTGCCTSAAGCGTCAGTATYGGCCAGT
GRGCGCTTCGCCAYYGGTGTTCYTCGAATATCTACGCATTTCCACCGCTACACTCGGAATTCACYCACCTCTGC
CGTACTCTAGATATVCAGHTCAAAGGCAAGTTCCGRGGTTGAGCCCCGGGATTTACCTCTGACTTAHCTAMCCGCC
TRCGCAGCTTTACGCCAGTAATCCGAA YAACGCTTGCAACCTACGATTACC GCGGCTGCTGGCACGTAGTTAG
CCGGTGCTTATCTTCAGGTACCGTCAATWTCGTCCCTGACAAAAGAGSTTTACAACCCGAAGGCCTTCTCMCTCA
CGCGGCATGGCTGGATCAGGCTTGCGCCATTGTCCAAAATCCCCACTGCTGCCTCCCGTAGGAGTCTGGGCCGTG
TCTCAGTCCCAGTGTGGCTGATCATCCTCTCAGACCAGTACGGATCGTCSCTTGGTRGGCCGTTACCCCAACCAAC
TASCTAATCCGACGCGGGCCSTCTAADGGCGVDAAYCTTTCCCYCRGRCGATGCGGTATTAGCWCAAGTTT
CCCTGWGTRTCCCCACCHCTAGGCAGGTTCCACGTTACTACCCGTYCGCCACTGAAGTATCCSTACTCCCG
TTGACTTGCATGTTWARGCATGCCGCCAGCGTTCRTTCTGAGCCAGGATCAAACCTCTAATCGAATTCGCCGGCC
    
```

Fig. 3: Putative sequence motif for unpolluted river water sample derived from consensus sequences

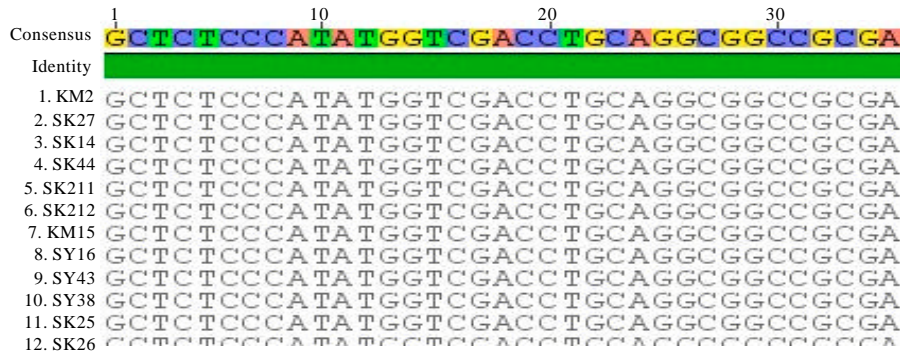
Pollution caused by pollutants will result in environmental stress. This situation will inevitably affect the genome structure of microbial community. Miller (1993) and Al-Sheikh and Fathi (2010) reported that environmental stress will increase the genetic alteration by mutation. Thus, there is potential and chances of the mutation within a specific condition (e.g., polluted) environment such as aquatic ecosystem (Praveena *et al.*, 2011; Ahmed, 2012).

Exchanges genetic material between natural microbial communities appears to be common phenomenon (Ford, 1994; Ahmed, 2012). The exchanging will mediated by several microbial processes either by conjugation, transformation and transduction. Microbial diversity adaptation towards environmental stress involve either by selective species enrichment, induction/derepression of enzymes or by genetic changes. Bacteria have the ability to rapidly tolerate and adapt toxic substances in their environment (Ford, 1994; Amann *et al.*, 1995; Ahmed, 2012).

The data and information obtained from the alignment can be used to develop a putative sequence motif for polluted and unpolluted group which is derived from consensus sequences (Fig. 4). These putative motifs are unique for each group. With their term “metagenomic profile”, Marshall *et al.* (2008) suggested that this “profile” can be used to characterize ecosystems with some difficulties in implementing it in the field. The difference between our study and that of Marshall *et al.* (2008) is that this study are trying to facilitate river quality assessment using molecular screening through the sequences motifs reconstruction.

This samples number were small but present study suggest that even with the small number of samples, the molecular assessment of aquatic ecosystem can be conducted to characterize the ecological health situation. This is because our phenetic tree obviously shows the two main alliances representing the polluted and unpolluted group. Even the coliform groups were embedded in the polluted group. This proved that





GCTCTCCCATATGGTTCGACCTGCAGGCGGCGCGGAATTCAGTACTAGTATTAGAGTTTGATCCTGGCTCAGATTGAACGC  
 TGCGGYYAGGCYTAACACATGCAAGTCAACGGCAGCACAGAGGAGCTTGCTCCTTGGGTGGCGAGTGGCGGACGGGT  
 GAGKAATACATCGGAATCCTCCTTTTCGTGGGGATAACGTAGGGAACTTACGCTAATACCGCATACGACCTACGGG  
 TGAAAGCAGGGGATCTTCGGACCTTTCGCGGATTGGATGAGCCGATGTCGGATTAGCTAGTTGGTGGGGTAAAGGCCA  
 CCAAGGCGACGATCCGTAGCTGGTCTGAGAGGATGAYCAGCCACACTGGAAGTGAAGACACGGTCCAGACTCCTACGGG  
 AGGCAGCAGTGGGGAATATTGGACAATGGGCGCAAGCCTGATCCAGCCATGCCGCGTGGGTGAAGAAAGGCCTTCGGGT  
 TGTAAAGCCCTTTTGTGGGAARGAAATCSATCCGGTTAATACCTGGTTSGGATGACGGTACCCAAAGAATAAGCACC  
 GGCTAACTTCGTGCCAGCAGCCGCGGTAATACGAAGGGTGAAGCGTTAATCGGAATTAAGTGGCGTAAAGCGTGCY  
 AGGCGGTATTTAAGTCAAGTGTGAAATCCCCGGCTCAACCTGGGAAGTGCATTTGATACTGGATGGCTAGAGTGT  
 GGTAGAGGGGGTGAATTCCTGGTGTAGCAGTGAATGCGTAGAGATCAGGAGGAACACCGATGGCGAAGGCAGCCC  
 CCTGGACCAACAATGACGCTCAGGCACGAAAGCGTGGGAGCAACAGGATTAGATACCCTGGTAGTCCACGCCCTAA  
 ACGATGTCGAACGGATGTTGGGTCTTTGGGACGAGTATCGAAGCTAACCGGTTAAGTTGACCGCCTGGGGAGTAC  
 GGCCGCAAGGTTGAAACTCAAAGGAATTGACGGGGGCCGCAACAGCGGTGGAGTATGTGGTTAATTGATGCAACG  
 CGAAGAACCTTACCTACCCTTGACATGTMGAGAATTTCCAGAGATGGATTGGTGCCTTCGGGAACCTCGAACACAGGT  
 GCTGCATGGCTGTCGTGAGCTCCTGTCGTGAGATGTTGGTTAAGTCCCGCAACGAGCGCAACCCCTGTCCTTAGTTG  
 CCAGCAGTAAGGGTGGGAACCTAAGGAGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAGTCATCAT  
 GGCCCTACGGGTAGGGCTACACACGTAACAATGGYAGGTACAGAGGGCTGCAAGCCCGGAGGGTAAAGCCAATCC  
 CATAAASCCTRTCGTAGTCCGGATTGGAGTCTGCAACTCGACTCCATGAAGTCCGAATCGTAGTAATCGCAGATCAG  
 CATGCTGCCGGTGAATACGTTCCCGGCTTGTACACACCGCCGTCACCATGGGAGTGGGTTGACCCAGAAAGTAGG  
 TAGCTTAACCTTCGGGAGGGCGCTTACCACGGTGTGGTTCATGACTGGGGTGAAGTCGTAACAAGGTAGCCGTATCGG  
 AAGGTGCGGTGGATCACCTCCTTAATCGAATCCCGCGGCCCATGGCGG

Fig. 4: Putative sequence motif for polluted river water sample derived from consensus sequences

although in small number of samples, the microbial community still capable to shows their varied tolerance ranges, in this case their adaptability towards the pollution level.

Although, the study is considered to be preliminary, our results suggest that the metagenomic analysis of microbial communities have the potential to characterize ecosystem which are differ from each other. Microbial communities can provide useful information about ecological health status which cannot be told by chemical and physiological indicators. Information such as microbial interactions within community and putative sequences motif which are unique for each environment situation can be retrieved. Although there are lot of studies must be conducted about to optimize the methodology, this study shows that this technology can be useful tools for research, environmental agencies and industry. However, further analysis with greater study sites is desirable to establish more robust conclusion.

**ACKNOWLEDMENT**

This study was funded by Research University Grant from the Universiti Teknologi Malaysia (Vote No. 77553) which is gratefully acknowledged.

**REFERENCES**

Agarwal, A.K. and G.S. Rajwar, 2010. Physico-chemical and microbiological study of tehri dam reservoir, Garhwal Himalaya, India. *J. Am. Soc.*, 6: 65-71.  
 Ahmed, Z., 2012. Microbial community in nutrient-removing membrane bioreactors: A review. *J. Environ. Sci. Technol.*, 5: 16-28.  
 Al-Sheikh, H. and A.A. Fathi, 2010. Ecological studies on lake Al-Asfar (Al\_Hassa Saudi Arabia) with special references to the sediments. *Res. J. Environ. Sci.*, 4: 13-22.



- Amann, R.L., W. Ludwig and K.H. Schleifer, 1995. Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. *Microbiol. Rev.*, 59: 143-169.
- Bhuiyan, F.A., S. Nagata and K. Ohnishi, 2011. Novel chitinase genes from metagenomic DNA prepared from marine sediments in Southwest Japan. *Pak. J. Biol. Sci.*, 14: 204-211.
- Bruce, K.D., W.D. Hiorns, J.L. Hobman, A.M. Osborn, P. Strike and D.A. Ritchie, 1992. Amplification of DNA from native populations of soil bacteria by using the polymerase chain reaction. *Applied Environ. Microbiol.*, 58: 3413-3416.
- Cottrell, T.M., L.A. Waidner, L. Yu and D.L. Kirchman, 2005. Bacterial diversity of metagenomic and PCR libraries from the Delaware River. *Environ. Microbiol.*, 7: 1883-1895.
- Cowan, D., Q. Meyer, W. Stafford, S. Muyanga, R. Cameron and P. Wittwer, 2005. Metagenomic gene discovery: Past, present and future. *Trends Biotech.*, 23: 321-329.
- Daniel, R., 2004. The soil metagenome-a rich resource for the discovery of novel natural products. *Curr. Opin. Biotechnol.*, 15: 199-204.
- DeLong, E.F., C.M. Preston, T. Mincer, V. Rich and S.J. Hallam *et al.*, 2006. Community genomics among stratified microbial assemblages in the ocean's interior. *Science*, 311: 496-503.
- Devereux, R. and G.W. Mundfrom, 1994. A phylogenetic tree of 16S rRNA sequences from sulfate-reducing bacteria in a sandy marine sediment. *Applied Environ. Microbiol.*, 60: 3437-3439.
- Durai, G. and M. Rajasimman, 2011. Biological treatment of tannery wastewater. *J. Environ. Sci. Technol.*, 4: 1-17.
- Edwards, U., T. Rogall, H. Blocker, M. Emde and E.C. Bottger, 1989. Isolation and direct complete nucleotide determination of entire genes. Characterization of a gene coding for 16S ribosomal RNA. *Nucleic Acids Res.*, 17: 7843-7853.
- Farrelly, V., F.A. Rainey and E. Stackebrandt, 1995. Effect of genome size and *rrn* gene copy number on PCR amplification of 16S rRNA genes from a mixture of bacterial species. *Applied Environ. Microbiol.*, 61: 2798-2801.
- Ford, T., 1994. Pollutant effects on the microbial ecosystem. *Environ. Health Perspect.*, 102: 45-48.
- Ghandhari, A. and S.M.R. Alavi Moghaddam, 2011. Water balance principle: A review of studies on five watersheds in Iran. *J. Environ. Sci. Technol.*, 4: 465-479.
- Handelsman, J., 2004. Metagenomics: Application of genomics to uncultured microorganisms. *Microbiol. Mol. Biol. Rev.*, 68: 669-685.
- Herrera, A. and C.S. Cockell, 2007. Exploring microbial diversity in volcanic environments: A review of methods in DNA extraction. *J. Microbiol. Method*, 70: 1-12.
- Kerckhof, L.J., 2009. Ocean microbial genomics. *Deep Sea Res.*, 56: 1824-1829.
- Kovacs, M., 1992. Biological Indicators of Environmental Pollution. In: *Biological Indicators in Environmental Protection*, Kovacs, M. (Ed.). Ellis Horwood, New York.
- Kwon, O.S., 2000. Characterization of isolated *Lactobacillus* spp and classification by RAPD-PCR analysis. *J. Microbiol.*, 38: 137-144.
- Marshall, M.M., R.N. Amos, V.C. Henrich and P.A. Rublee, 2008. Developing SSU rDNA metagenomic profiles of aquatic microbial communities for environmental assessments. *Ecol. Indic.*, 8: 442-453.
- Maznah, W. and W. Omar, 2010. Perspectives on the use of algae as biological indicators for monitoring and protecting aquatic environments, with special reference to Malaysian freshwater ecosystems. *Trop. Life Sci. Res.*, 21: 51-67.
- Miller, R.V., 1993. Genetic Stability of Genetically Engineered Microorganisms in the Aquatic Environment. In: *Aquatic Microbiology: An Ecological Approach*, Ford, T.E. (Ed.). Blackwell, Boston, USA., pp: 483-511.
- Nasab, S.B., A. Bavi, S. Karami and M. Albaji, 2010. Evaluation of the wastewater-related problem of Shoteit river Shushtar (Southwest Iran). *Res. J. Environ. Sci.*, 4: 23-32.
- Oparaku, N.F., B.O. Mgbenka and C.N. Ibeto, 2011. Waste water disinfectant utilizing ultraviolet light. *J. Environ. Sci. Technol.*, 4: 73-78.
- Praveena, S.M., S.S. Siraj, A.K. Suleiman and A.Z. Aris, 2011. A brush up on water quality studies of Port Dickson Malaysia. *Res. J. Environ. Sci.*, 12: 841-849.
- Trivedy, R.K. and P.K. Goel, 1984. *Chemical and Biological Methods for Water Pollution Studies*. Environmental Publications, Karad, India.
- Weisburg, W.G., S.M. Barns, D.A. Pelletier and D.J. Lane, 1991. 16S ribosomal DNA amplification for phylogenetic study. *J. Bacteriol.*, 173: 697-703.