EUROPEAN ACADEMIC RESEARCH
Vol. I, Issue 12/ March 2014

ISSN 2286-4822
www.euacademic.org

Impact Factor: 3.1 (UIF)
DRJI Value: 5.9 (B+)

In-silico Designing of SYBR Green Based Real-Time PCR Array for the Quantification of *Salmonellae* and Enterotoxigenic *Escherichia coli* in Water

RAJESH SINGH TOMAR[1]
ANURAG JYOTI
RAGHVENDRA K MISHRA
VIKAS SHRIVASTAVA
SHUCHI KAUSHIK
Amity Institute of Biotechnology
Amity University, Madhya Pradesh, Gwalior
India

Abstract:
Rapid and accurate quantification of Salmonellae and enterotoxigenic Escherichia coli (ETEC) is critical for management of water-borne diseases. This needs careful computation of primers for amplification in single run. In this study, the potential of in-silico amplification of target genes of the designed SYBR Green real-time PCR array was determined to detect target pathogens. Salmonellae and ETEC signature virulent gene sequences invA and LT1 were retrieved from NCBI's GenBank database. The primers were designed using Primer-BLAST and analysed for specificity. In-silico PCR simulations of these primers were done for validation. The computed primers exhibited melting temperatures in the narrow range of 58.75-59.89^{0}C, making them suitable in array format. The primers were highly specific towards target genes on the basis of BLAST results against known microbial genomes. In-silico PCR simulation of these assays generated amplicons for invA (115 bp) and LT1 (90 bp) genes. Genomic equivalent standards (10^6-1 CFU/PCR) of both the target genes on 96-well PCR plate were formulated. These results indicate the computational exercises for designing and validation of primers for

[1] Corresponding author

array will reduce the time and effort in optimizing on-bench assays and can be used for the management of water quality and outbreaks caused by Salmonellae and ETEC.

Key words: SYBR Green, Real Time PCR, Array, water, virulent genes**.**

1. Introduction

Water-borne pathogens pose severe health and economic burdens worldwide. The monitoring and management of water resources is of prime importance as surface and subsurface water is used for many purposes, including drinking, cooking and recreational activities. The consumption and exposure to unsafe water is responsible for the outbreaks of infectious diseases like diarrhea, typhoid, cholera etc. causing serious loss to human health. Approximately 94% of the diarrheal burden of disease is due to the environment, and associated with risk factors such as unsafe drinking water, lack of sanitation and poor hygiene (Prüss and Corvalán 2006; WHO 2008). *Salmonella* and *E. coli* are among the most prevalent water-borne pathogenic bacteria in recent decades.

Salmonella is an important food and water-borne pathogen. Each year, approximately, 93.8 million human cases of gastroenteritis and 155, 000 deaths occur due to *Salmonella* infection around the world (Majowicz *et al*. 2010). The typhoid caused by *Salmonella enterica* serotype Typhi remains an important public health problem in developing countries. Further, Salmonellosis causes substantial medical and economic burdens worldwide (Voetsch *et al*. 2004). Countries like India, Indonesia, Bangladesh and Pakistan have been identified as high risk sites for infections caused by *Salmonella* spp. (Bahl *et al*. 2004; Brooks *et al*. 2005; Ochiai *et al*. 2008). Previous studies have indicated the prevalence of *Salmonellae* in surface and potable waters in India (Jyoti *et al*. 2010; 2011).

Enterotoxigenic *Escherichia coli* (ETEC), a potential pathovar of *E. coli* is regarded as a major cause of diarrhea worldwide in humans, mainly affecting children and travelers. The contamination of drinking or recreational waters with ETEC has been associated with waterborne disease outbreaks. Diarrhea due to ETEC is caused by the consumption of contaminated water (Qadri *et al*. 2005, Ram *et al*. 2008). In case of improper sanitation and hygiene, the ETEC is a major cause of diarrhea. A few studies report the prevalence of ETEC in surface waters and in macrophytes (Ram *et al*. 2008; Singh *et al*. 2010).

The conventional methods for detection of pathogenic bacteria involve selective enrichment and biochemical identification. These methods are time consuming, less sensitive and generally require at least three to five days for identification and confirmation. The immunoassay based diagnostics have disadvantages including nonspecific binding and maintaining physical conditions during the procedures. The rapid detection of these pathogens is extremely important since delayed diagnosis causes morbidity and mortality.

Efficiency of the detection of microbial pathogens present in water depends on a number of factors. The advancements in the area of genomics have led to the development of DNA-based diagnostic methods such as polymerase chain reaction (PCR) amplification techniques which overcome problems like culturing of microorganisms, false positives and low sensitivity in pathogen detection. Real Time PCR technique has advantages over conventional methods in terms of rapidity, high specificity and sensitivity. Further, the SYBR Green real-time PCR arrays exhibit good reproducibility among different users, PCR instruments and test sites. In addition, the SYBR Green PCR arrays have the highest concordance with TaqMan PCR, and a high level of concordance with other quantitative methods and microarrays (Arikawa *et al*. 2008). Rapid advances in genome sequencing

have resulted in sequencing of thousands of microorganisms in recent years. Thousands of microbial genomes have been sequenced and many are in progress (NCBI Genome database, http://www.ncbi.nlm.nih.gov/genomes/lproks.cgi). The resulting sequence database has made it possible to analyze microbial pathogens at the molecular level. Each pathogen contains unique nucleotide sequences which can be used to differentiate from other organisms.

Oligonucleotides computed to amplify target genes in the developed real-time PCR array require the on bench validation. The laboratory validation of oligonucleotides is not only time consuming and expensive but also prone to failures due to non-specificity (Patel *et al.* 2011). Therefore, in this study real-time PCR array was designed and the oligonucleotides for the virulent genes of *Salmonellae* and *Escherichia coli*, were computed and validated *in silico*.

2. Materials and Methods

2.1 Retrieval of gene sequences from database

The two pathogenic members of Enterobacteriaceae family, *Salmonellae* and enterotoxigenic *Escherichia coli*, were selected to design a SYBR Green based real-time PCR array for their detection in surface and potable water. Nucleotide sequences of virulent genes of *Salmonellae* and ETEC were retrieved from the GenBank (www.ncbi.nlm.nih.gov/genbank/), National Centre for Biotechnology Information, Bethesda, MD, USA (Table 1). The target gene for *Salmonellae* was virulent gene-*invA*. Similarly, the virulent gene *LT1* was chosen for the detection of ETEC.

Microorganisms/ Genes	**Accession number of gene sequence**
Salmonellae	
invA	M90846, EU348368, EU348366, EU348369, U43271, DQ644615
Enterotoxigenic *Escherichia coli*	
LT1	AY342056, JX504011, FJ407053, S60731, AB011677

Table 1. GenBank sequence identification numbers of bacterial genes used for BLAST analysis to determine conserved sequences for strain specific primers.

2.2 Multiple Sequence Alignment

Multiple sequence alignments of retrieved sequences of genes selected for the present study were carried out using ClustalW (http://www. ebi.ac.uk/clustalW) to find out conserved contigs. Five nucleotide sequences of *invA* gene and five nucleotide sequences of *LT1* gene were aligned separately in themselves to create the conserved sequences of these.

2.3 Primer designing

The conserved regions, unique to a target pathogen, obtained from ClustalW were used to design primers using dedicated web based software Primer-BLAST (http://www.ncbi.nlm.nih.gov/tools/primer-blast/). The primer designing parameters such as melting temperature, GC content, amplicon length, etc were taken into consideration.

2.4 Specificity of designed primers

The specificity of computed primers of *invA* of *Salmonellae* and *LT1* of ETEC were determined against the known microbial genomes and sequences by BLAST (Basic Local Alignment Search Tool) programme of NCBI to ensure no homology found in other genera or species.

2.5 *In-Silico* analysis of primers

The primer-pairs were used for *in-silico* analysis using, in-silico PCR (http://insilico.ehu.es/PCR/) as described by (Ram et al.

2008) to ensure that each primer pair amplified the targeted sequences of the genes in the conserved region of the selected gene.

2.6 Designing of PCR Array for the detection of *Salmonellae* and *Escherichia coli*

The array was designed to detect *Salmonellae* and *Escherichia coli* in reference strains as well as in water samples. For this purpose the *invA* and *LT1* genes were selected. The *invA* gene, lies at centisome 63 region of the Salmonella chromosome which constitutes a pathogenicity island and encodes an essential component of the invasion associated specialized type Ø protein secretion apparatus (Galan, 1996). A wide range of Salmonella serotypes including all subspecies exhibit *invA* gene which is absent in other bacterial species and genera (Malorny *et al*. 2003; Hadjinicolaou *et al*. 2009). The virulent *LT1* gene was chosen which is exclusively present in ETEC. To address these pathogens in the environmental samples, two domains of aquatic systems were selected, viz. potable water and surface water.

3. Results

3.1 Computation and specificity of primers for detection of *Salmonellae* and Enterotoxigenic *Escherichia coli*

Highly specific real-time PCR primers exhibiting melting temperature 58.75-59.89^{0}C and amplicon sizes- 115 bp (for *invA* gene of *Salmonellae*) and 90 bp (for *LT1* gene of ETEC) were computed using dedicated web based primer designing tool, Primer-BLAST (Table 2). The specificity of computed oligonucleotides was determined against the known microbial genomes by NCBI-BLAST. The BLAST analysis showed no homology in other genera or species. Hence, the computed primers were highly specific towards their respective targets.

Genes	Primers Sequences (5'-3')	Length (bp)	Tm (°C)	GC (%)	Product size (bp)
Salmonellae					
invA	**F:** GAGGGCCTGGACGATAACAG	20	59.89	60	115
	R: AGGACACGACTTCATCGGAA	20	58.75	50	
Enterotoxigenic *Escherichia coli*					
	F: CATTTAAGAGCGGCGCAACA	20	59.83	50.00	90
	R: TGTCCTTCATCCTTTCAATGGC	22	58.91	45.45	

Table 2. Primers used in SYBR Green real-time PCR array

3.2 *In-Silico* validation of computed primers

In silico PCR amplification of the genes selected for *Salmonellae* (*invA*) and ETEC (*LT1*) were performed using the respective computed primers. In the case of *Salmonellae*, *in-silico* PCR product of 115 bp was observed for the *invA* gene (Figure 1). Similarly, *in silico* amplification of the *LT1* gene selected to quantify ETEC exhibited an amplification of the product of 90 bp (Figure 2).

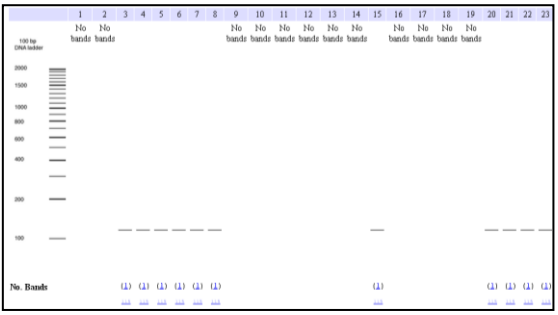

Figure 1. Graphical representation of *in silico* PCR amplification for *invA* gene of *Salmonellae*.

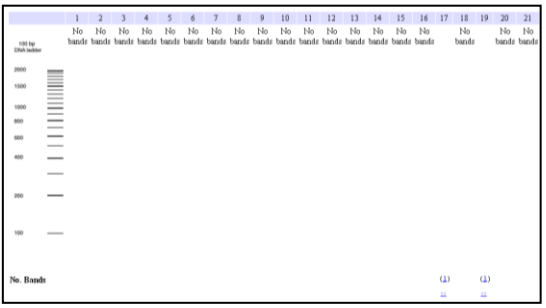

Figure 2. Graphical representation of *in silico* PCR amplification for *LT1* gene of ETEC.

3.3 Designing of PCR Array for the detection of *Salmonellae* and *Escherichia coli*

The array was designed in such a manner that serially ten-fold diluted standard, ranging from 10^6 CFU/PCR down to 1CFU/PCR can be run in doublets for each gene. Here wells A1 to G4 contained standard ten-fold dilution of reference strains. Wells H1 to H4 contained no template control (NTC). Wells A5 to H12 (64 wells) were dedicated to environmental samples. In the present array two environmental samples of each sampling type targeting one gene can be assessed at a time. Wells A5 to H6 and A7 to H8 contained DNA templates from surface and potable waters with primer pairs for *invA* gene of *Salmonellae*. Similarly wells A9 to H10 and A11 to H12 contained DNA templates from surface and potable waters with primer pairs for *LT1* gene of and ETEC (Figure 3).

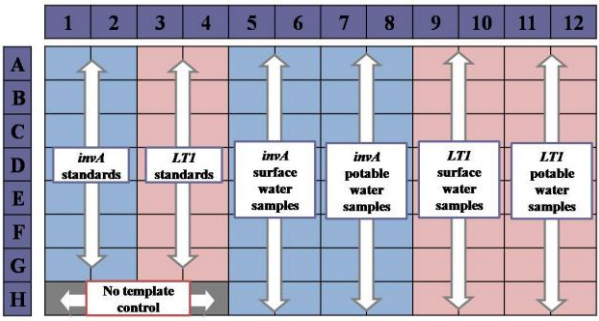

Figure 3. Schematics of a real-time PCR array in 96 well format for quantification of *Salmonellae* and *Escherichia coli*.

4. Discussion

This study has led to the computation and *in-silico* validation of a quantitative culture independent PCR array in SYBR Green

format for quantification of *Salmonellae* and ETEC bacterial cells exhibiting *invA* and *LT1* genes respectively in water samples.

The essential virulent genes are primarily located in the pathogenic islands of bacteria which are required for the survival and pathogenicity of bacteria (Dobrindt *et al.* 2004). These essential genes are evolutionary conserved than non-essential genes in bacteria (Jordan *et al.* 2002). In the present study *Salmonellae* and enterotoxigenic *Escherichia coli,* belonging to family Enterobacteriaceae were chosen due to their prevalence in surface and potable water. The virulence factor, *invA* is the potential virulence factor and is present in almost all the serovars of *Salmonella* (Malorny *et al.* 2003; Malorny *et al.* 2004). Similarly, *LT1* gene expressing the enterotoxin LT in ETEC is directly related to its pathogenicity (Nataro *et al.* 1998).

The detection of these organisms needs development of methods that are sensitive and specific. Real time PCR has revolutionized PCR technique by application of specific chemistries and instrumentation. The chemistries consist of DNA-binding dyes like SYBR Green I and hybridization probes like molecular beacons, scorpion primers and TaqMan. SYBR Green I is minor groove binder (MGB) dye which intercalate to double stranded DNA and emit 1,000 fold greater fluorescence as compared to unbound state (Zipper *et al.* 2004). The real-time PCR methods are potential tools for diagnostics, environmental monitoring and risk assessment of the microbiological quality of water due to increase in detection specificity and reduction in analysis time (Kubista *et al.* 2006). The potential of real time PCR was exploited to quantitatively enumerate the target pathogens in an efficient manner. To provide higher specificity to the assay, signature virulence genes of *Salmonellae* and Enterotoxigenic *Escherichia coli* were chosen. The attempt has been made to detect multiple genes simultaneously by applying multiplexing (Lata *et al.* 2009).

However, there are limitations in multiplexing of genes. Hence, the limitations of multiplexed assays were overcome by development of SYBR Green based high throughput PCR array format enabling simultaneous detection of 64 environmental samples.

The computed primers were highly specific towards their respective targets as confirmed from the BLAST results. Various parameters such as E-values (minimum), max score (maximum) and query coverage (maximum) were analysed for the specificity of oligonucleotides. The BLAST analyses showed lower E-values, this result in more quality hits as the chance of non-specific hits expected when searching a database was low. The performance of computed primers was evaluated by web based *in silico* PCR programme. The computed simulation showed the length and the electrophoretic mobility of amplicon. Further, no amplification was observed *in silico* when primers specific to one gene were used for amplification with genes of the organism in closely related genera (Fig. 1 and 2). This approach helped in selection of the desired primer pairs minimizing possible non-targeted products. This not only improves the specificity but also minimizes time needed for *in vitro* verification of work. Therefore, the computed primers of selected genes would serve as the signature for bacterial identification.

Real-time PCR array can be successfully used in the monitoring of surface, ground and potable and wastewater specifically when viable but non-culturable pathogenic bacteria are present. Detection of stressed or injured pathogens is not possible by culture based protocols. *In- silico* amplification of the designed primers for genes specific to *Salmonellae* and ETEC will reduce the bench optimization time. Thus, highly specific and rapid Real-time PCR array based molecular methods will be an important tool for culture-free detection and quantification of the target pathogens. This can be used for monitoring of contamination in sewage impacted natural water

bodies and civic potable water distribution system for reducing the chances of waterborne diseases outbreaks. Therefore, the detection of *Salmonellae* and ETEC by Real Time PCR array molecular methods became specific and rapid obviating the need for cultivation.

It can be concluded from this study that *in-silico* amplification of the target genes prior to bench optimization of the designed real-time PCR array for culture-independent quantitative enumeration of *Salmonellae* and ETEC have the potential for improving specificity and minimizing time and cost needed for *in vitro* verification of work. Such validated economically viable SYBR Green real-time PCR array could be used for the management of water quality, outbreaks and tracking non-point sources of water pollution.

Acknowledgment

We wish to express our sincere acknowledgement to Dr. Ashok Kumar Chauhan, President, RBEF parent organization of Amity University Madhya Pradesh (AUMP), Mr. Aseem Chauhan, Additional President, RBEF and chairman of AUMP; Lt. Gen. V.K. Sharma, AVSM (Retd.), Vice Chancellor of AUMP, Gwalior, for providing their valuable support, necessary facilities and encouragement throughout the work. Both RST and AJ have contributed equally to this paper.

BIBLIOGRAPHY

Arikawa, E., Y. Sun, J. Wang, Q. Zhou, B. Ning, S. Dial, L. Guo, and J. Yang. 2008. “Cross-platform comparison of SYBR Green real-time PCR with TaqMan PCR, microarrays and other gene expression measurement technologies evaluated in the MicroArray Quality Control (MAQC) study.” *BMC Genomics* 9: 328.

Bahl, R, A. Sinha, C. Poulos, D. Whittington, S. Sazawal, R. Kumar, D. Mahalanabis, C.J. Acosta, J.D. Clemens, and

M.K. Bhan. 2004. "Costs of illness due to typhoid fever in an Indian urban slum community: implications for vaccination policy." *J. Heal. Popul. Nutr.* 22: 304-310.

Brooks, W.A., A. Hossain, D. Goswami, K. Nahar, K. Alam, and N. Ahmed. 2005. "Bacteremic typhoid fever in children in an urban slum, Bangladesh." *Emerg. Infect. Dis.* 11: 326-329.

Dobrindt, U., B. Hochhut, U. Hentschel and J. Hacker. 2004. "Genomic islands in pathogenic and environmental microorganisms." *Nat. Rev. Microbiol.* 2: 414–424.

Galan, J.E. 1996. "Molecular bases of Salmonella entry into host cells." *Mol Microbiol.* 20: 263-271.

Hadjinicolaou, A.V., V.L. Demetriou, M.A. Emmanuel, C.K. Kakoyiannis, and L.G. Kostrikis. 2009. "Molecular beacon-based real-time PCR detection of primary isolates of *Salmonella* Typhimurium and *Salmonella* Enteritidis in environmental and clinical samples." *BMC Microbiol.* 9: 97.

Jordan, I.K., I.B. Rogozin, Y.I. Wolf, and E.V. Koonin. 2002. "Essential genes are more evolutionarily conserved than are nonessential genes in bacteria." *Genome Res.* 12: 962-968.

Jyoti, A., S. Ram, P. Vajpayee, G. Singh, P.D. Dwivedi, S.K. Jain, and R. Shanker. 2010. "Contamination of surface and potable water in South Asia by *Salmonellae:* Culture-Independent quantification with molecular beacon real-time PCR." *Sci. Tot. Environ.* 408: 1256-1263.

Jyoti, A., P. Vajpayee, G. Singh, C.B., Patel, K.C. Gupta, and R. Shanker. 2011. "Identification of environmental reservoirs of non-typhoidal salmonellosis: aptamer assisted bio-concentration and subsequent detection of *Salmonella typhimurium* by quantitative Polymerase Chain Reaction." *Environ. Sci. Techn.* 45: 8996-9002.

Kubista, M., J.M. Andrade, M. Bengtsson, A. Forootan, J.

Jonak, K. Lind, R. Sindelka, R. Sjoback, B. Sjogreen, L. Strombom, A. Stahlberg, and N. Zoric. 2006. "The real-time polymerase chain reaction." *Mol. Asp. Med.* 27: 95-125.

Lata, P., S. Ram, M. Agrawal and R. Shanker. 2009. "Real Time PCR for the rapid detection of *vanA* gene in surface waters and aquatic macrophyte by molecular beacon probe." *Environ. Sci. Technol.* 43**:** 3343-3348.

Majowicz, S.E., J. Musto, E. Scallan, F.J. Angulo, M. Kirk, S.J. O'Brien, T.F. Jones, A. Fazil, and R.M. Hoekstra. 2010. "The global burden of nontyphoidal Salmonella gastroenteritis." *Clin. Infect. Dis.* 50: 882-889.

Malorny, B., J. Hoorfar, C. Bunge, and R. Helmuth. 2003. "Multicenter Validation of the Analytical Accuracy of *Salmonella* PCR: towards an International Standard." *Appl. Environ. Microbiol.* 69: 290-296.

Malorny, B., E. Paccassoni, P. Fach, C. Bunge, A. Martin, and R. Helmuth. 2004. "Diagnostic Real-Time PCR for detection of *Salmonella* in food." *Appl. Environ. Microbiol.* 70: 7046-7052.

Nataro, J.P. and J.B. Kaper. 1998. "Diarrheagenic *Escherichia coli*." *Clin. Microbiol. Rev.* 1: 142-201.

Ochiai, R.L., C.J. Acosta, M.C.D. Holliday, D. Baiqing, S.K. Bhattacharya, M.D. Agtini, Z.A. Bhutta, D.G. Canh, M. Alim, S. Shin, J. Wain, A.L. Page, M.J. Albert, J. Farrar, R.A. Elyazeed, T. Pang, C.M. Galindo, L. Seidlein, and J.D. Clemens, the Domi Typhoid Study Group. 2008. "A study of typhoid fever in five Asian countries: disease burden and implications for controls." *Bull. World Heal. Organ.* 86: 260-268.

Patel, C.B., P. Vajpayee, G. Singh, A. Jyoti, R.S. Upadhyay, and R. Shanker. 2011. "Computation and *in-silico* validation of a Real-Time PCR array for quantitative detection of fecal coliforms in surface and potable water." *Int. J. Biosci. Biotechnol.* 1: 184-195.

Prüss-Üstün, A. and C. Corvalán. 2006. "Preventing disease through healthy environments: toward an estimate of the environmental burden of disease. Geneva: The World Health Organization."

Qadri, F., A.M. Svennerholm, A.S. Faruque, and R.B. Sack. 2005. '"Enterotoxigenic *Escherichia coli* in developing countries: epidemiology, microbiology, clinical features, treatment, and prevention." *Clin Microbiol. Rev.* 18: 465-483.

Ram, S., P. Vajpayee, and R. Shanker. 2008. "Rapid culture-independent quantitative detection of enterotoxigenic *Escherichia coli* in surface waters by Real-Time PCR with molecular beacon." *Environ. Sci. Technol.* 42: 4577-4582.

Singh, G., P. Vajpayee, S. Ram, and R. Shanker. 2010. "Environmental reservoirs for enterotoxigenic *Escherichia coli* in south Asian gangetic riverine system." *Environ. Sci. Technol.* 44: 6475–6480.

Voetsch, A.C., T.J.V. Gilder, F.J. Angulo, M.M. Farley, S. Shallow, and R. Marcus. 2004. "Food Net estimate of the burden of illness caused by non-typhoidal *Salmonella* infections in the United States." *Clin. Infect. Dis.* 38: S127-134.

WHO: Water Quality Interventions to Prevent Diarrhoea: Cost and Cost-Effectiveness. 2008**.**

Zipper, H., H. Brunner, J. Bernhagen, and F. Vitzthum. 2004. "Investigations on DNA intercalation and surface binding by SYBR Green I, its structure determination and methodological implications." *Nucl. Acids Res.* 32: e103 1-10.