Songklanakarin J. Sci. Technol.
39 (3), 373-381, May - Jun. 2017

http://www.sjst.psu.ac.th

Original Article

Comparative analysis of fliC Gene from *Salmonella enterica* sub-species for biosensor probe design and phylogenetic tree construction

Sunderamurthe Satisvar[1], Subash Chandra Bose Gopinath[1,2*], Mohd Khairuddin Md Arshad[1], Dee Vie Ong[1], Thangavel Lakshmipriya[1], Chun Hong Voon[1], Uda Hashim[1], Abdul Rahim Ruslinda[1], and Suresh Venkata Chinni[3]

[1] *Institute of Nano Electronic Engineering (INEE), Universiti Malaysia Perlis, Kangar, Perlis, 01000 Malaysia*

[2] *School of Bioprocess Engineering, Universiti Malaysia Perlis, Arau, Perlis, 02600 Malaysia*

[3] *Department of Biotechnology, Faculty of Applied Sciences, Asian Institute of Medical, Science and Technology (AIMST) University, Bedong, 08100 Malaysia*

Received: 10 January 2016; Revised: 16 May 2016; Accepted: 17 June 2016

Abstract

In tandem to the Salmonellosis infection worldwide, a study was conducted to determine the conserved and non-conserved region in *Salmonella enterica* sub-species so that it can be used to design probes in biosensors for the detection of *Salmonella enterica* as a species or sub-species distinctively. The region was selected for this study through *fliC* gene that is present in all *Salmonella* sub-species, encodes the *Salmonella* flagella or flagellin determines the serotype due to its H antigen and provides virulence to the bacteria. All sub-species were analyzed in a group of six and then analyzed individually in groups of two with all possible combinations to determine the overlapping regions. Based on the study, the predominant conserved region for *S. enterica* sub-species is between 103 and 158, whereas the non-conserved region is from 1245 to 1285. A phylogenetic tree was constructed for the *S. enterica* sub-species to determine the evolution of the *Salmonella* sub-species.

Keywords: *Salmonella enterica*, sub-species, *fliC* gene, flagellin, H antigen, biosensor

1. Introduction

Salmonella are facultative anaerobes from the *Enterobacteriaceae* family which do not develop spores and are gram negative rod structured gamma-proteobacteria as well as parasites commonly occupying the intestinal zones of humans, cattle, avian, insects and reptiles (Andino & Hanning, 2015; Mcquiston *et al.*, 2008). The common characteristics that all *Salmonella* share are *Salmonella* require minimal medium for growth and grows ideally in a pH range of 7-7.5, under a temperature scope of 35-43°C (Finn *et al.*, 2013; Mattick *et al.*, 2001). The span of the bacteria is usually 2.5 microns long with a width of 0.5-1.5 microns and the genome proportions vary among strains within bounds of 4,460 to 4,857 kb (Andino & Hanning, 2015).

The nomenclature of *Salmonella* has been in dispute for a period of time as there are two systems in circulation. The first was suggested by Le Minor and Popoff (1987) and does not adhere to the rules established in the Bacteriological Code and other system conforms to the rules of the Bacterio-

* Corresponding author.
Email address: subash@unimap.edu.my

logical Code. However, it must be noted that although the first system does not follow the Bacteriological Code, it is more recognized and commonly used. Due to the confusion amid researchers worldwide, the Judicial Commission of the International Commission on the Systematics of Prokaryotes has declared the Opinion 80 which was restricted to nomenclature and not taxonomy. Later on, Le Minor and Popoff discovered *S. enterica* which has several sub-species and thus by relating commission ruling with the Le Minor and Popoff system, a bridge was fashioned between the two nomenclature systems (Tindall *et al.*, 2005).

Salmonella is broadly split into two distinct species namely *S. enterica* and *S. bongori* which was formerly acknowledged as sub-species VII and now documented as a separate species (Mcquiston *et al.*, 2008). The *Salmonella* is differentiated in the Kauffmann-White serological scheme of classification as Kauffmann propositioned on the recognition through antigens O and H in *Salmonella* (Brenner *et al.*, 2000). *S. enterica* is separated into six sub-types namely *S. enterica* subsp. *enterica* (sub sp I), *S. enterica* subsp. *salamae* (subsp II), *S. enterica* subsp. *arizonae* (subsp IIIa), *S. enterica* subsp. *diarizonae* (subsp IIIb), *S. enterica* subsp. *houtenae* (subsp IV) and *S. enterica* subsp. *indica* (subsp VI) (Mcquiston *et al.*, 2008).

At present, there are 2463 serovars of *Salmonella* which comprises of 46 groups of O antigen and 114 groups of H antigen. The O antigens exists as several groups namely A, B, C1, C2, D and E whereas the H antigens can be separated into two flagellin loci (*fliC* and *fliB*). This loci gene function unfettered of each other and is controlled by the *hin* switch mechanism. However, not all *Salmonella* has this switching mechanism as some *Salmonella* such as *S. enterica* subsp. *arizonae* has only *fliC* gene (Andrews-polymenis *et al.*, 2010; Brenner *et al.*, 2000; Mcquiston *et al.*, 2008; Mermin *et al.*, 2004). The surface antigens influences the virulence of the bacteria, provides motility for the bacteria through the flagella and illicit inflammation response (Coburn *et al.*, 2007; Mechanisms, 1995). The O and H antigens are also the chief antigenic composition in *Salmonella* (Nalbantsoy *et al.*, 2010).

The virulent region on the *Salmonella* is located on the pathogenicity islands (SPI 1 and SPI 2) which are actually extended regions of DNA. SPI 1 allows bacterial breach whereas SPI 2 assists intracellular endurance (Ly & Casanova, 2007). The breach into host cells rely on type III secretion systems (T3SS1 and T3SS2) encoded by SPI which permits bacterial replication and causes rapid macrophage pyroptosis (Fink & Cookson, 2007; Ibarra & Steele-mortimer, 2009). Salmonellosis is a medical condition with varying indications caused by the infection of *Salmonella* bacteria (Maki-Yonekura *et al.*, 2010; Nithya *et al.*, 2015). *Salmonella* serovars can be split into three classes as typhoid fever, bacteraemia and enteritis which instigate separate symptoms. Typhoid fever occurs due to *S. typhi* and *S. paratyphi* but it must be noted that the illness is less severe if triggered by *S. paratyphi A* and more mild if through *S. paratyphi B* (Schotmulleri) or *S. paratyphi* C (Hirshfeldii) (Bhutta, 2006; Connor & Schwartz, 2005; Everest *et al.*, 2001; Pang *et al.*, 1995). Severe disease symptoms of typhoid fever includes septic shock, intestinal bleeding, intestinal perforation and neuropsychiatric complications (Pang *et al.*, 1998).

Salmonella have now developed resistance to several constantly used antibiotics such as chloramphenicol, trimethoprim, ampicillin and tetracycline (Pang *et al.*, 1998). The resistance genes that are presently identified are positioned on mobile genetic elements called transposons, gene cassettes, plasmids and genomic islands (Brenner *et al.*, 2006). It is unfeasible to forecast the emergence of a new serovar or the degree of threat a new serovar poses to human and animals. Due to this factor, some antibiotics cannot be regarded as the first line drug for treating salmonellosis anymore factors that affect resistance to infection include rapid stomach emptying, carcinoma, lymphoma, sickle cell anaemia, parasitism, leukaemia, diabetes, antibiotic treatment and altered gut flora (Lax *et al.*, 1995).

The objective of this work is to determine the conserved as well as the non–conserved region in the *fliC* gene for biosensor probe designing and to construct the phylogenetic tree of *S. enterica* subspecies. Previous studies on *S. enterica* sub-species have been mainly focused towards genetic and functional aspects, whereas current study is more towards finding minimal regions for sensing applications. Other downstream applications include knocking down and imaging analysis using this fragment. One of the reasons that *fliC* gene is used to design the probe is because the flagella which the *fliC* gene codes is used to determine the serotype between sub-species which makes identification faster and easier as salmonellosis is a global concern. *Salmonella* flagella provide antigenicity by antigen H present on the surface of the flagella and are also a virulence factor in *Salmonella*. The fact that SPI is present on plasmids makes them highly mutable, hence flagella which has less polymorphisms is a better choice for probe design. The flagella also initiate the immune response and this region can also be used to create vaccines (Coburn *et al.*, 2007). The fact that the motility of S*almonella* through flagella which allows invasion into host cells and the H antigen bind to phagosome signals gene expression can be utilized to disrupt the functions of *Salmonella* flagella by identifying the non-conserved region (Mechanisms, 1995). The *fliC* gene is used as there a higher number of available H antigens than O antigens and fliB are not present in all *Salmonella* sub-species (Mcquiston *et al.*, 2008).

2. Materials and Methods

The region which determines serotype difference and provides virulence to *Salmonella* was determined by different studies. The terminologies have been considered for the current study design are, *S. enterica* nomenclature, *S. enterica* serotyping, *S. enterica* salmonellosis, *S. enterica* phylogeny, *S. enterica* typhoidal salmonellosis, *S. enterica*

non-typhoidal salmonellosis, *S. enterica* mechanisms of infection and *S. enterica* vaccines and treatment. The region that was chosen for current investigation is after thorough literature search and it was the *fliC* gene. The sequence for the *fliC* gene for six different *S. enterica* subspecies with no respect to the strain variation was then obtained from the National Centre for Biotechnology Information (NCBI). Only *fliC* gene with complete coding sequence was selected for the alignment.

The alignment of sequences was done for all the gene sequence in CLUSTAL X and then viewed in GeneDoc. The sequences were compared and contrasted to determine the region with the longest conserved region and longest non-conserved region in the six different *S. enterica* subspecies. The alignment procedure was then repeated by using two different *Salmonella* subspecies of all combinations possible using the six *Salmonella* subtypes. A phylogenetic tree of the six different *Salmonella* sub-species was then created using the software Splits Tree (Splitstree4, version 4.13.1.0) to determine the evolution of *Salmonella* sub-species and calculated with the application Bootstrap (Huson & Bryant, 2006). The diagram of the protein was then obtained by searching the terms *Salmonella* flagella in Protein Database (RCSB PDB; http://www.rcsb.org/pdb/home/home.do). The highest antigenic region in the flagellin protein was determined by reviewing scientific literature exceeding the span of 20 years (He *et al.*, 1994). To analyze the structural information Pymol software (version 5.10.130.0; Install shield Software Corporation) has been used.

The details of accession are as follows;

- "Salmonella Enterica Subsp. Arizonae Strain CDC_99_350 Flagellin (fliC) Gene, Complete Cds," February 8, 2011. http://www.ncbi.nlm.nih.gov/nuccore/HM142026.1 (Accession No. HM142026).
- "Salmonella Enterica Subsp. Diarizonae Strain CDC_Ar90 Flagellin (fliC) Gene, Complete Cds," February 8, 2011. http://www.ncbi.nlm.nih.gov/nuccore/HM142031.1. (Accession No. HM142031)
- "Salmonella Enterica Subsp. Enterica Strain CDC_2038 Flagellin (fliC) Gene, Complete Cds," February 8, 2011. http://www.ncbi.nlm.nih.gov/nuccore/HM142000.1. (Accession No. HM142000).
- "Salmonella Enterica Subsp. Houtenae Strain CDC_CNM2556 Flagellin (fliC) Gene, Complete Cds," February 8, 2011. http://www.ncbi.nlm.nih.gov/nuccore/HM142047.1. (Accession No. HM142047).
- "Salmonella Enterica Subsp. Indica Strain CDC_1415 Flagellin (fliC) Gene, Complete Cds," February 8, 2011. http://www.ncbi.nlm.nih.gov/nuccore/HM142052.1. (Accession No. HM142052).
- "Salmonella Enterica Subsp. Salamae Strain CDC_365 Flagellin (fliC) Gene, Complete Cds," February 8, 2011. http://www.ncbi.nlm.nih.gov/nuccore/HM142012.1. (Accession No. HM142012).

3. Results and Discussion

Salmonellosis occurs mainly through faecal-oral such as household contact for *S. typhi*, by contaminated food for *S. paratyphi* and typhoid fever is linked to travellers while non-typhoid fever which is elicited by the commensal, non-typhoid *Salmonella* are food-borne (Butaye *et al.*, 2006; Cheong *et al.*, 2007; Fangtham & Wilde, 2008). The vaccines available currently for use are parenteral Vi polysaccharide vaccine and Ty21a with the suggestion that both vaccines be used simultaneously to illicit a stronger immune response (Engels *et al.*, 1998; Fangtham & Wilde, 2008; Pang *et al.*, 1995; 1998). Towards vaccine developments, designing sensing strategies and appropriate probe/receptor to find necessary binding affinities are mandatory as shown previously (Anbu *et al.*, 2006; Balakrishnan *et al.*, 2016; Gopinath, 2007, 2008; Gopinath *et al.*, 2006, 2007, 2008, 2009, 2012a, 2012b, 2013; Kumarevel *et al.*, 2004, 2008; Perumal *et al.*, 2015). In this study, the bioinformatics analysis has been carried out to find right sequence from *fliC* gene for the development of probe to be for sensing applications. Six sub-species were analyzed and all possible combinations were considered to determine the overlapping regions. The criteria that was utilized for the selection of those particular sequences was random selection to avoid biases over specific selection of a particular strain of *Salmonella enterica* without any particular interests in order to cover a wide range of conserved region and non-conserved region in *S. enterica* to serve as targets for probe designing for biosensors.

The longest conserved region for all six *S. enterica* subspecies when aligned is between 103 and 158 base pairs which consist of 56 base pairs (Figure 2). This region can be utilized to design a probe to detect all *S. enterica* sub-species without distinction. This region is completely aligned and does not even have a single base pair which is unaligned, which makes it the best to detect all six different subspecies as *S. enterica*. The longest non-conserved region for all six *S. enterica* subspecies is between 1,245 and 1,285 base pairs

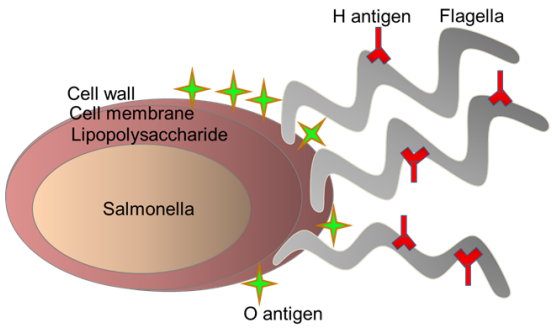

Figure 1. Image of Salmonella. Flagellin determines the serotype due to its H antigen and provides virulence to the bacteria. *fliC* gene is specifically focused in this study and it is from flagella.

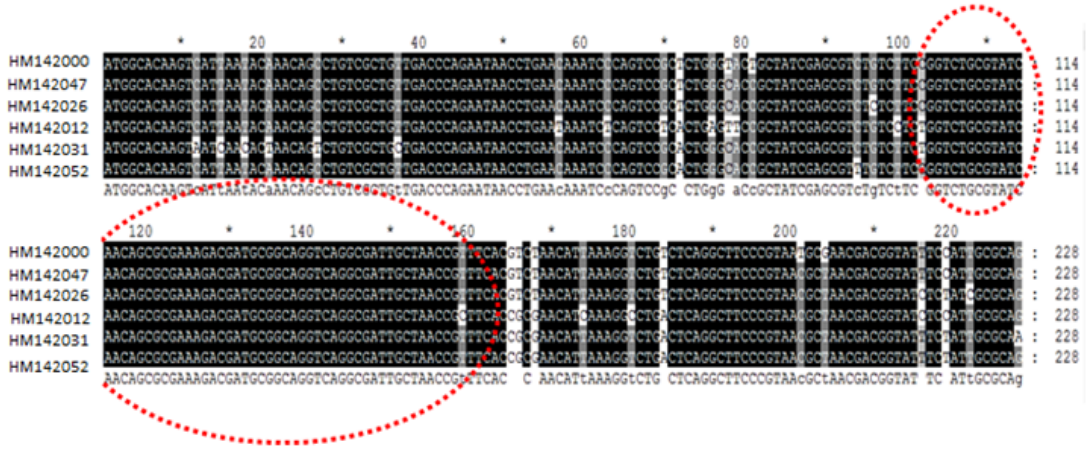

Figure 2. Conserved region among all six *S. enterica* subspecies. Aligned between 103 and 158 bases which consist of 56 bases. Conserved regions from all species are shaded in black and non-conserved regions are in grey. Major conserved region is encircled. *Enterica* (Accession No. HM142000); *enterica-houtenae* (Accession No. HM142047); *enterica* Subsp. *arizonae* (Accession No. HM 142026); *enterica-salamae* (Accession No. HM142012); *enterica diarizonae* (Accession No. HM142031); *enterica indica* (Accession No. HM142052).

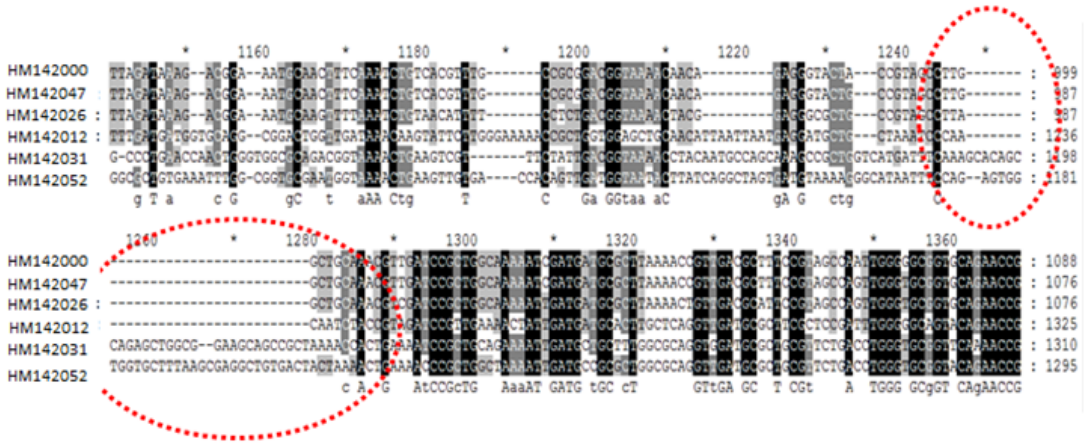

Figure 3. Longest non-conserved region among six *S. enterica* sub-species. Between 1245 and 1285 bases which consist of 41 bases. Conserved regions from all species are shaded in black and non-conserved regions are in grey. Major non-conserved region is encircled. *Enterica* (Accession No. HM142000); *enterica-houtenae* (Accession No. HM142047); *enterica* Subsp. *arizonae* (Accession No. HM142026); *enterica-salamae* (Accession No. HM142012); *enterica diarizonae* (Accession No. HM142031); *enterica indica* (Accession No. HM142052).

which consist of 41 base pairs (Figure 3). This region can be used to manufacture a probe to detect the different sub-species of *S. enterica*. This region is unaligned with very few minor alignments between few sub-species at places. Even though there are several minor alignments towards the end, this region is still the best non-conserved region for probe design as it can differentiate between *S. enterica* sub-species effectively.

When the *fliC* regions was aligned using two different *S. enterica* sub-species at a time with all 15 possible combinations, the results varied among various combination. The conserved regions for the sub-species are 1,456 to 1,518 for sub-species *enterica* and *salamae*, 322 to 440 for sub-species *enterica* and *arizonae*, 103 to 164 for sub-species *enterica* and *diarizonae*, 845 to 986 for sub-species *enterica* and *houtenae*, 1 to 71 for sub-species *enterica* and *indica*, 103 to 158 for sub-species *salamae* and *arizonae*, 100 to 158 for sub-species *salamae* and *diarizonae*, 1 to 56 and 103-158 for sub-species *salamae* and *houtenae*, 1,438 to 1,521 for sub-species *salamae* and *indica*, 103 to 164 for sub-species *arizonae* and *diarizonae*, 97 to 215 for *arizonae* and *houtenae*, 1 to 71 for sub-species *arizonae* and *indica*, 103 to 164 for sub-species *diarizonae* and *houtenae*, 103 to 227 for sub-species *diarizonae* and *indica* and 1 to 71 for sub-species *houtenae* and *indica* (Table 1).

The non-conserved regions for the sub-species are 679 to 726 for *enterica* and *salamae*, 790 to 793 and 807 to 810 for *enterica* and *arizonae*, 1,178 to 1,234 for sub-species *enterica* and *diarizonae*, 692 to 697 for sub-species *enterica* and *houtenae*, 1,140 to 1,193 for sub-species *enterica* and

Table 1. Comparative analysis among *Salmonella* species.

Species	*S. enterica* subsp *enterica*	*S. enterica* subsp *salamae*	*S. enterica* subsp *arizonae*	*S. enterica* subsp *diarizonae*	*S. enterica* subsp *houtenae*	*S. enterica* subsp *indica*
S.enterica subsp *enterica*	-	1456-1518(CR)/ 679-726(NCR)	322-440(CR)/ 790-793&807-810 (NCR)	103-164(CR)/ 1178-1234(NCR)	845-986(CR)/ 692-697(NCR)	1-71(CR)/ 1140-1193(NCR)
S.enterica subsp *salamae*	1456-1518(CR)/ 679-726(NCR)	-	103-158(CR)/ 896-943(NCR)	100-158(CR)/ 952-963(NCR)	1-56&103-158 (CR)/986-1018 (NCR)	1438-1521(CR)/ 724-732(NCR)
S.enterica subsp *arizonae*	322-440(CR)/ 790-793& 807-810(NCR)	103-158(CR)/ 896-943(NCR)	-	103-164(CR)/ 510-551(NCR)	97-215(CR)/ 693-697(NCR)	1-71(CR)/ 1141-1173(NCR)
S.enterica subsp *diarizonae*	103-164(CR)/ 1178-1234 (NCR)	100-158(CR)/ 952-963(NCR)	103-164(CR)/ 510-551(NCR)	-	103-164(CR)/ 652-700(NCR)	103-227(CR)/ 708-722(NCR)
S.enterica subsp *houtenae*	845-986(CR)/ 692-697(NCR)	1-56&103-158 (CR)/986-1018 (NCR)	97-215(CR)/ 693-697(NCR)	103-164(CR)/ 652-700(NCR)	-	1-71(CR)/ 1111-1172(NCR)
S.enterica subsp *indica*	1-71(CR)/ 1140-1193 (NCR)	1438-1521(CR)/ 724-732(NCR)	1-71(CR)/ 1141-1173(NCR)	103-227(CR)/ 708-722(NCR)	1-71(CR)/ 1111-1172(NCR)	-

CR – Complementary region; NCR – Non-complementary region

"Salmonella Enterica Subsp. Arizonae Strain CDC_99_350 Flagellin (fliC) Gene, Complete Cds," February 8, 2011. http://www.ncbi.nlm.nih.gov/nuccore/HM142026.1 (Accession No. HM142026)

"Salmonella Enterica Subsp. Diarizonae Strain CDC_Ar90 Flagellin (fliC) Gene, Complete Cds," February 8, 2011. http://www.ncbi.nlm.nih.gov/nuccore/HM142031.1. (Accession No. HM142031)

"Salmonella Enterica Subsp. Enterica Strain CDC_2038 Flagellin (fliC) Gene, Complete Cds," February 8, 2011. http://www.ncbi.nlm.nih.gov/nuccore/HM142000.1. (Accession No. HM142000)

"Salmonella Enterica Subsp. Houtenae Strain CDC_CNM2556 Flagellin (fliC) Gene, Complete Cds," February 8, 2011. http://www.ncbi.nlm.nih.gov/nuccore/HM142047.1. (Accession No. HM142047)

"Salmonella Enterica Subsp. Indica Strain CDC_1415 Flagellin (fliC) Gene, Complete Cds," February 8, 2011. http://www.ncbi.nlm.nih.gov/nuccore/HM142052.1. (Accession No. HM142052)

"Salmonella Enterica Subsp. Salamae Strain CDC_365 Flagellin (fliC) Gene, Complete Cds," February 8, 2011. http://www.ncbi.nlm.nih.gov/nuccore/HM142012.1. (Accession No. HM142012)

indica, 896 to 943 for sub-species *salamae* and *arizonae*, 952 to 963 for sub-species *salamae* and *diarizonae*, 986 to 1018 for sub-species *salamae* and *houtenae*, 724 to 732 for sub-species *salamae* and *indica*, 510 to 551 for sub-species *arizonae* and *diarizonae*, 693 to 697 for sub-species *arizonae* and *houtenae*, 1,141 to 1,173 for sub-species *arizonae* and *indica*, 652 to 700 for sub-species *diarizonae* and *houtenae*, 708 to 722 for subspecies *diarizonae* and *indica* and 1,111 to 1,172 for sub-species *houtenae* and *indica* (Table 1). The alignment for the all the combinations was also done vice versa to check the accuracy of the results with the initially obtained result.

The majority of the conserved regions for the sub-species were overlapped with each other showing that these regions did not vary much with sub-species as time progressed. There are mainly three regions which the *S. enterica* sub-species conserved regions overlapped (Figure 4). The first region is from 1 to 56 with the sub-species comparison of sub-species *houtenae-indica*, *arizonae-indica*, *enterica-indica* and *salamae-houtenae*. The second region is from 103 to 158 with the sub-species comparison of sub-species *diarizonae-indica*, *diarizonae-houtenae*, *arizonae-diarizonae*, *enterica-diarizonae*, *salamae-houtenae*, *salamae-arizonae*, *salamae-diarizonae* and *arizonae-houtenae*. The third region is from 1,456 to 1,518 with the sub-species comparison of sub-species *salamae-indica* and *enterica-salamae*. The regions are for probe designing for specific combination of subspecies targets as opposed to probe designing for the entire *Salmonella* subspecies. The regions also indicate that there are more overlapping in conserved region rather than non-conserved regions which indicate the possibility of different types of probes to be designed for various types of detection in the biosensor.

Unlike the conserved regions, the bulk of the non-conserved region for the sub-species is scattered across. However, there are five minor regions which have some over-

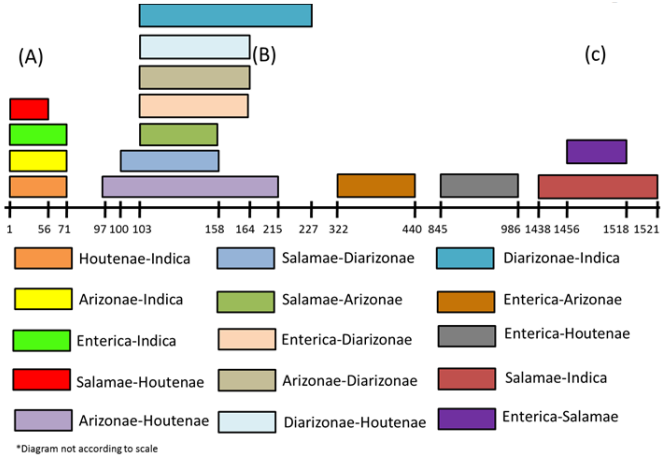

Figure 4. Three main conserved regions overlapped in *S. enterica* sub-species. First region is from 1 to 56 (A) with the sub-species comparison (*houtenae-indica*, *arizonae-indica*, *enterica-indica* and *salamae-houtenae)*. Second region is from 103 to 158 (B) with the sub-species comparison (*diarizonae-indica*, *diarizonae-houtenae*, *arizonae-diarizonae*, *enterica-diarizonae*, *salamae-houtenae*, *salamae-arizonae*, *salamae-diarizonae* and *arizonae-houtenae)*. The third region (C) is from 1,456 to 1,518 with the sub-species comparison (*salamae-indica* and *enterica-salamae)*.

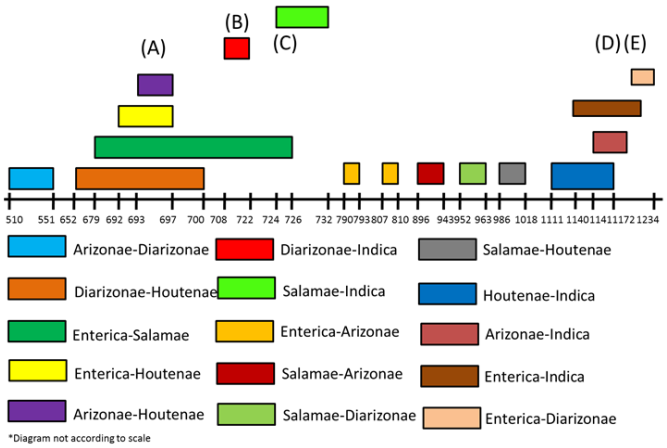

Figure 5. Five minor non-conserved regions have some overlap of *S. enterica* subspecies. First region is from 693 to 697 with the sub-species comparison (*arizonae-houtenae*, *enterica-houtenae*, *enterica-salamae* and *diarizonae-houtenae)*. Second region is from 708 to 722 with the sub-species comparison (*enterica-salamae* and *diarizonae-indica)*. Third region is from 724 to 726 with the sub-species comparison (*salamae-indica* and *enterica-salamae)*. The fourth region is from 1,141 to 1,172 with the sub-species comparison (*houtenae-indica*, *arizonae-indica* and *enterica-indica)*. Fifth region is from 1,178 to 1,193 with the sub-species comparison (*enterica-indica* and *enterica-diarizonae)*.

lap of *S. enterica* subspecies non-conserved region (Figure 5). The first region is from 693 to 697 with the sub-species comparison of sub-species *arizonae-houtenae*, *enterica-houtenae*, *enterica-salamae* and *diarizonae-houtenae*. The second region is from 708 to 722 with the sub-species comparison of sub-species *enterica-salamae* and *diarizonae-indica*. The third region is from 724 to 726 with the sub-species comparison of sub-species *salamae-indica* and *enterica-salamae*. The fourth region is from 1,141 to 1,172 with the sub-species comparison of sub-species *houtenae-indica*, *arizonae-indica* and *enterica-indica*. The fifth region is from 1,178 to 1,193 with the sub-species comparison of sub-species *enterica-indica* and *enterica-diarizonae*. The overall comparison is shown in the Table 1.

The phylogenetic tree was constructed with a parameter that the strains of the sub-species do not explicitly represent the particular sub-species. The phylogenetic tree is constructed to obtain general information on the evolutionary divergence of the *S. enterica* sub-species (Figure 6). The phylogenetic tree generated does not have a definite root for the *S. enterica* sub-species. The node circled in red is the common ancestor of the sub-species labeled a, b, c, d, e and f. The node circled in green is the common ancestor of b, c, d, e and f. The node circled in blue is the common ancestor of c, d, e and f. The node circled in orange is the common ancestor of d, e and f. The node circled in purple is the common ancestor of e and f. The sub-species a, b and c are diverged quickly as a separate sub-species as

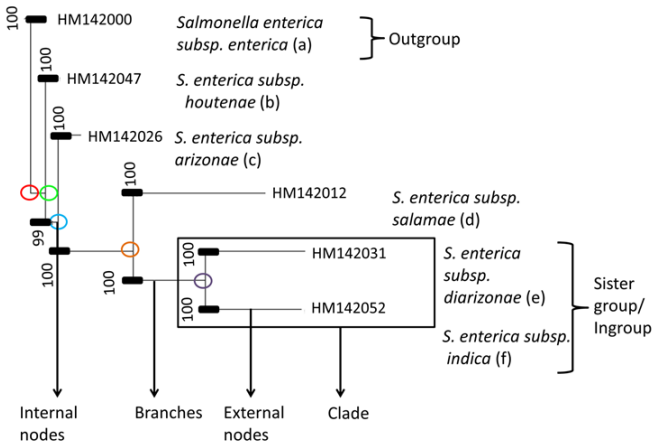

Figure 6. Phylogenetic tree constructed. Color circles are to represent the branching out of the evolution tree between different subspecies in a clearer manner. Software Splitstree4, version 4.13.1.0 was used to determine the evolution of *Salmonella* sub-species and calculated with the application Bootstrap.

denoted by the short branches whereas the sub-species d, e and f diverged after a period of longer time as denoted by the relatively longer branches. Sub-species 'a' is an out-group, while sub-species 'e' and 'f' sister groups. The sub-species (d) was determined with amino acids at 229 to 230 to be an important and was marked as red spheres on R type straight flagellar filament made of full-length flagellin (Figure 7).

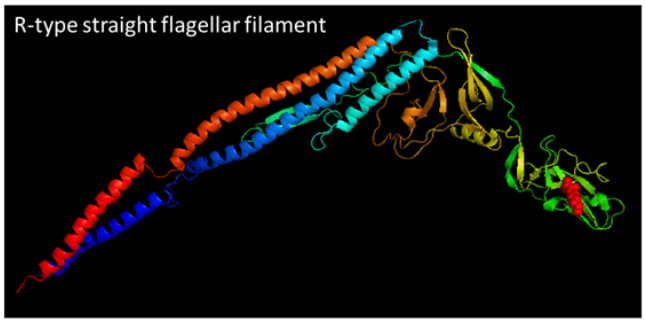

Figure 7. R type straight flagellar filament of full-length flagellin. Pymol online software has been used to create this structure. Amino acids are at 229 to 230 to be an important and marked as red spheres.

4. Conclusion

In summary, *Salmonella* is a bacterium of global interest as it carries about salmonellosis through food or other methods to both humans and animals. The increasing figures of worldwide salmonellosis rate have also brought us to address these bacteria in a highly serious manner. The development of new strains of *Salmonella* which increasingly have more and more antibiotic resistance poses a threat especially in areas with poor sanitation. Currently available detection methods are extremely time consuming and therefore the designing probes for biosensors is crucial in the upcoming battle against salmonellosis with the divergence of the sub-species with time. Based on the obtained results, 50 bases are maximum length and on sensing surface this length can reduce until 20 bases with in this length (50 bases). The strategy is shown here will aid to design probes to diagnose different pathogens.

References

Anbu, P., Gopinath, S. C. B., Hilda, A., Mathivanan, N., & Annadurai, G. (2006). Secretion of keratinolytic enzymes and keratinolysis by Scopulariopsis brevicaulis and Trichophyton mentagrophytes: Regression analysis. *Canadian Journal of Microbiology, 52*, 1060-1069.

Andino, A., & Hanning, I. (2015). Salmonella enterica/: Survival, colonization, and virulence differences among serovars. *The Scientific World Journal, 2015,* ID 520179

Andrews-Polymenis, H. L., Baumler, A., J., McCormick, B. A., & Fang, F. C. (2010). Taming the elephant: Salmonella biology, pathogenesis, and prevention. *Infection and Immunity, 78*, 2356-236.

Balakrishnan, S. R., Hashim, U., Gopinath, S. C. B., Poopalan, P., Ramayya, H. R., Veeradasan, P., Haarindraprasad, R., & Ruslinda, A. R. (2016). Polysilicon nanogap lab-on-chip facilitates multiplex analyses with single analyte. *Biosensors and Bioelectronoics, 84,* 44-52.

Brenner, F. W., Villar, R. G., Angulo, F. J., Tauxe, R., & Swaminathan, B. (2000). Guest Commentary. *Journal of Clinical Microbiology, 38*, 2465-2467.

Brenner, G., Butaye, P., Cloeckaert, A., & Schwarz, S. (2006). Genes and mutations conferring antimicrobial resistance in Salmonella/: An update. *Microbes and Infection, 8*, 1898-1914.

Butaye, P., Michael, G. B., Schwarz, S., Barrett, T. J., Brisabois, A., & White, D. G. (2006). The clonal spread of multidrug-resistant non-typhi Salmonella serotypes.

Microbes Infection, 8, 1891-1897.

Cheong, H. J., Lee, Y. J., Hwang, I. S., Kee, S. Y., Cheong, H. W., Song, J. Y., Kim, J. M., Park, Y. H., Jung, J., & Kim, W. J. (2007). Characteristics of non-typhoidal salmonella isolates from human and broiler-chickens in Southwestern Seoul, Korea, *Journal of Korean Medical Science,* 773-778.

Coburn, B., Grassl, G. A., & Finlay, B. B. (2007). Salmonella, the host and disease/: A brief review. *Immunology and Cell Biology, 85,* 112-118.

Connor, B. A., & Schwartz, E. (2005). Typhoid and paratyphoid fever in travellers. *The Lancet Infectious Diseases, 5*, 623-628.

Engels, E. A., Falagas, M. E., Lau, J., & Bennish, M. L. (1998). Typhoid fever vaccines/: a meta-analysis of studies on efficacy and toxicity. *British Medical Journal*, *316,* 110-116.

Everest, P., Wain, J., Roberts, M., Rook, G., & Dougan, G. (2001). The molecular mechanisms of severe typhoid fever. *Trends in Microbiology, 9,* 316-320.

Fangtham, M., & Wilde, H. (2008). Emergence of *Salmonella paratyphi A* as a major cause of enteric fever: Need for early detection, preventive measures, and effective vaccines. *Journal of Travel Medicine, 15*, 344-350.

Fink, S. L., & Cookson, B. T. (2007). Microreview Pyroptosis and host cell death responses during Salmonella infection. *Cellular Microbiology, 9*, 2562-2570.

Finn, S., Condell, O., Mcclure, P., Amézquita, A., & Fanning, S. (2013). Mechanisms of survival, responses, and sources of Salmonella in low-moisture environments. *Frontiers Microbiology, 4*, 1-15.

Gopinath, S. C. B. (2008). Anti-coagulant aptamers. *Thrombosis Research, 122*, 838-847.

Gopinath, S. C. B. (2007). Antiviral aptamers. *Archives of Virology, 152*, 2137-2157.

Gopinath, S. C. B., Anbu, P., Lakshmipriya, T., & Hilda, A. (2013). Strategies to characterize fungal lipases for applications in medicine and dairy industry. *Biomed Research International, 2013*, 31-34.

Gopinath, S. C. B., Awazu, K., Fons, P., Tominaga, J., & Kumar, P. K. R. (2009). A sensitive multilayered structure suitable for biosensing on the BioDVD platform. *Analytical Chemistry, 81*, 4963-4970.

Gopinath, S. C. B., Awazu, K., & Fujimaki, M. (2012). Waveguide-mode sensors as aptasensors. *Sensors, 12*, 2136-2151.

Gopinath, S. C. B., Awazu, K., Fujimaki, M., Sugimoto, K., Ohki, Y., Komatsubara, T., ...Kumar, P. K. R. (2008). Influence of nanometric holes on the sensitivity of a waveguide-mode sensor: Label-free nanosensor for the analysis of RNA aptamer-ligand interactions. *Analytical Chemistry, 80*, 6602-6609.

Gopinath, S. C. B., Balasundaresan, D., Akitomi, J. & Mizuno, H. (2006). An RNA aptamer that discriminates bovine factor IX from human factor IX. *Journal of Biochemistry, 140*, 667-676.

Gopinath, S. C. B., Hayashi, K., & Kumar, P. K. R. (2012). Aptamer that binds to the gD protein of herpes simplex virus 1 and efficiently inhibits viral entry. *Journal of Virology, 86*, 6732-6744.

Gopinath, S. C. B., Shikamoto, Y., Mizuno, H., & Kumar, P. K. R. (2007). Snake-venom-derived factor IX-binding protein specifically blocks the gamma-carboxyglutamic acid-rich-domain-mediated membrane binding of human factors IX and X. *Biochemical Journal, 405*, 351-357.

He, X. S., Rivkina, M., Stocker, B. D., & Robinson, W. S. (1994). Hypervariable region IV of Salmonella gene *fliC*(d) encodes a dominant surface epitope and a stabilizing factor for functional flagella. *Journal of Bacteriology, 176*, 2406-2414.

Huson, D. H., & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution, 23,* 254-267.

Ibarra, J. A., & Steele-mortimer, O. (2009). Microreview salmonella – the ultimate insider. Salmonella virulence factors that modulate intracellular survival. *Cellular Microbiology, 11*, 1579-1586.

Kumarevel, T., Tanaka, T., Nishio, M., Gopinath, S. C. B., Takio, K., Shinkai, A., ...Yokoyama, S. (2008). Crystal structure of the MarR family regulatory protein, ST1710, from Sulfolobus tokodaii strain 7. *Journal of Structural Biology, 161*, 9-17.

Kumarevel, T. S., Gopinath, S. C. B., Nishikawa, S., Mizuno, H., & Kumar, P. K. R. (2004). Identification of important chemical groups of the hut mRNA for HutP interactions that regulate the hut operon in Bacillus subtilis. *Nucleic Acids Research, 32*, 3904-3912.

Le Minor, L., & Popoff, M. Y. (1987). Request for an opinion. Designation of Salmonella enterica sp. nov., nom. rev., as the type and only species of the genus Salmonella. *International Journal of Systematic Bacteriology, 37*, 465-468.

Ly, K. T., & Casanova, J. E. (2007). Microreview Mechanisms of Salmonella entry into host cells. *Cellular Microbiology, 9,* 2103-2111.

Maki-Yonekura, S., Yonekura, K., & Namba, K. (2010). Conformational change of flagellin for polymorphic supercoiling of the flagellar filament. *Nature Structural and Molecular Biololgy, 17*, 417-422.

Mattick, K. L., Jørgensen, F., Wang, P., Pound, J., Vandeven, M. H., Ward, L. R., ...Humphrey, T. J. (2001). Effect of challenge temperature and solute type on heat tolerance of Salmonella Serovars at low water activity. *Applied and Environmental Microbiology, 67*, 4128-4136.

Mcquiston, J. R., Wertheim, B. C., Doyle, J., Fields, P. I., Tauxe, R. V., & Logsdon, J. M. (2008). Molecular phylogeny of the Salmonellae/: Relationships among Salmonella species and subspecies determined from four housekeeping genes and evidence of lateral gene transfer events. *Journal of Bacteriolgy, 190*, 7060-7067.

Mechanisms, C. (1995). Bacterial pathogenesis of plants and animals; molecular and cellular mechanisms. *General Pharmacology the Vascular System, 26*, 1795.

Mermin, J., Hutwagner, L., Vugia, D., Shallow, S., Daily, P., Bender, J., ...Group, W. (2004).Reptiles, amphibians, and human Salmonella infection/: A population-based, case-control study. *Clinical Infectious Diseases, 30333*, 253–261.

Nalbantsoy, A., Karaboz, I., & Ivanova, R. (2010). Isolation and purification of O and H antigens from Salmonella Enteritidis as diagnostic tool. *Annals of Microbiology, 60*, 565-571.

Nithya, R., Ahmed, S. A., Hoe, C., Gopinath, S. C. B., Citartan, M., Chinni, S. V., ...Tang, T. H. (2015). Non-protein coding RNA genes as the novel diagnostic markers for the discrimination of Salmonella species using PCR. *Plos One, 10*, e0118668.

Pang, T., Bhutta, Z. A., Finlay, B. B., & Altwegg, M. (1995). Typhoid fever and other salmonellosis: A continuing challenge. *Trends in Microbiology, 3,* 253-255.

Pang, T., Levine, M. M., Ivanoff, B., Wain, J., & Finlay, B. B. (1998). Typhoid fever – important issues still remain. *Trends in Microbiology, 6*, 131–133.

Perumal, V., Hashim, U., Gopinath, S. C. B., Haarindraprasad, R., Poopalan, P., Liu, W.-W., ...Ruslinda, A. R. (2015). A new nano-worm structure from gold-nanoparticle mediated random curving of zinc oxide nanorods. *Biosensors and Bioelectronoics, 78,* 14-22.

Tindall, B. J., Grimont, P. A. D., Garrity, G. M., & Euze′by, J. (2005). Nomenclature and taxonomy of the genus Salmonella. *International Journal of Systematic and Evolutionary Microbiology, 55*, 521-524.