Phylogenetic Delineation of Salmonella serovars using Multiple Locus Sequence Typing (MLST)

Balasubramanian Suryalakshmi[1], Mareeswaran Hema[2], Velayutham Rajapria[2]

Abstract— Multiple Locus Sequence Typing (MLST) based on *fliB*, *fliD*, *metR*, *ompS*, *phoQ*, *rfaF* and *yehU* genes were developed for differentiation of s*almonella* serovars. The complete genome sequences were available for 55 different s*almonella* serovars at the National Centre for Biotechnology Information (NCBI) database. Single locus gene markers like rnpB gene, operons of 16s rRNA and 23s rRNA were unable to distinguish all the 55 serovars of *Salmonella.* Hence the seven novel genes which include both virulent and housekeeping genes were used in MLST technique and the discriminatory power of MLST was found to be 0.998.

Index Terms— , MLST, *Salmonella* Serovars, Phylogenetic Analysis, Single Locus Gene Markers, Housekeeping Gene, Discriminatory power.

---

1 Introduction

SALMONELLA belongs to the family Enterobacteriales. It is classified into two species namely *S. enterica* and *S. bongori. S. enterica*is further classified into six main subspecies: *S. enterica (I)*, *S. salamae* (II), *S. arizonae* (IIIa), *S. diarizonae* (IIIb), *S. houtenae* (IV) and *S. indica* (VI). Currently, around 2463 serotypes of *Salmonella* have been identified [1].

Non-typhoidal Salmonellae (NTS) cause gastroenteritis after ingestion of contaminated food or water. Other symptoms include acute onset of fever, abdominal cramps, nausea, diarrhoea and vomiting [2]. Patients with HIV, AIDS, diabetes mellitus, rheumatic diseases and people who have undergone transplantation (i.e. with severe immunosuppression) are more susceptible to NTS [3]. Human adapted pathogens *S. enterica* serotype Typhi and *S. enterica* serotype Paratyphi A, B, and C cause enteric fever. *Salmonella* colonizes the reproductive tract of hens and hence the egg also gets contaminated with *Salmonella* infection [4].Non-typhoidal Salmonellae caused around 1.4 million infections in the United States annually [5].Typhoid fever caused an estimated 21.7 million illnesses and 217,000 deaths and paratyphoid fever caused an estimated 5.4 million illnesses worldwide in 2000 [2]. The greatest burden of illness was experienced by infants, children and adolescents in south central and south eastern Asia [6].

Due to the increasing incidence of *Salmonella* infection in recent years, the need for the development of rapid and accurate detection methods have been intensified. The traditional methods for diagnosis of *Salmonella* involve blood culture test and Widal test. The microbiological method of culturing the organism from blood is time consuming and has poor sensitivity [7]. The Widal test is not suitable for early diagnosis of the disease [7]. Hence quicker approaches have been searched for, mainly at the DNA level. Recent genome based methods for typing and sub-typing of *Salmonella* involves Pulsed Field Gel Electrophoresis (PFGE) [8], Multiple Loci Variable Number Tandem Repeat Analysis (MLVA) [9], Multiple Locus Sequence Typing (MLST) [10] and multiplex- PCR based methods [11].

In the present study, seven gene markers which are able to distinguish the various *Salmonella* serovars were found and MLST technique was performed. The characteristics of the different gene loci which include length of the gene, gene description, number of alleles, number of synonymous sites (dS), number of non-synonymous sites (dN) and dN/dS ratio were found. The discriminatory power of MLST was also calculated using Simpson's Index of Diversity.

2 Methods

2.1 Bacterial Strains Used

Fifty five different *Salmonella* serovars which had their complete genome sequence at the National Centre for Biotechnology Information (NCBI) database (http://www.ncbi.nlm.nih.gov/) were used for the study (Appendix A).*Citrobacter rodentium* str. ICC168, which has 60-65% genome identity with *Salmonella* was used as outgroup to study the phylogeny.

3 Multiple Locus Sequence Typing

The technique used was Multiple Locus Sequence Typing (MLST). The MLST measures the DNA sequence variations in a set of genes (both housekeeping and virulent) and characterizes the bacterial strains by their unique allelic profiles. The workflow of MLST involves the following steps: 1) data collection 2) data analysis and 3) multilocus sequence analysis.

3.1 Data Collection

In data collection, well defined identification of variation is obtained by nucleotide sequence determination of genes. The seven *Salmonella* loci selected for MLST analysis were the genes encoding flagellin methylation protein (*fliB*), flagellar capping protein (*fliD*), outer membrane protein (*ompS*), sensor protein (*phoQ*), lipopolysaccharide heptosyltransferaseII (*rfaF*), two component system sensor kinase (*yehU*) and transcriptional regulator (*metR*). The characteristics of seven gene loci are listed in Table 2. The sequences of the above genes were

obtained from NCBI database.

3.2 Data Analysis

In data analyses, all unique sequences were assigned allele numbers, combined into an allelic profile and assigned a sequence type (ST) number (Table 1).

3.3 Concatenation of genes

All the seven genes were concatenated using UNION, a tool to concatenate multiple FASTA files in Emboss union (http://www.bioinformatics.nl/cgi-bin/emboss/union).

3.4 Order for concatenation

The seven genes were concatenated in different orders to check whether the phylogenetic tree obtained was consistent. There were nearly 7 factorial combinations for concatenation of these seven genes. Two random combinations include

1. *phoQ-fliB-fliD-ompS-rfaF-yehU-metR*

2. *rfaF-yehU-fliB-phoQ-ompS-fliD-metR*

3.5 Optimization of gene order for concatenation

To optimize the gene order for concatenation and to find whether the tree is reliable or not, bootstrapping was performed.Bootstrapping is a commonly used method to check the reliability in phylogenetic analysis. It is used to assess the repeatability of a given result, Bootstrap values that correspond to a probability of ≥ 95% shows that the corresponding clade is real [18].The genes were optimized in step by step procedure. To start with, using seven genes, 42 different combinations were performed by selecting and concatenating two genes at a time and the bootstrap test of phylogeny was performed to confirm the phylogenetically reliable tree. Then the third gene was concatenated to the optimized two genes and so on. Ultimately the best gene order with high bootstrap value was arrived.

4 Multilocus Data Analysis

4.1 Phylogeny

The phylogenetic analysis was performed using MEGA software. The Molecular Evolutionary Genetics Analysis (MEGA) software was developed for the statistical analyses of evolutionary relationships from the DNA and protein sequence data. MEGA is an integrated tool for performing sequence alignment, inferring phylogenetic trees, estimating the number of synonymous and non-synonymous polymorphic sites. MEGA now provides the bootstrap test for nucleotide and amino acid alignments to determine the reliability of evolutionary trees [19]. The Neighbour Joining (NJ) method with maximum composite likelihood was used to construct phylogenetic trees.

4.2 Simpson's index of diversity

The Discriminatory Power (D) that can be expressed by the formula of Simpson's index of diversity given by Hunter and Gaston [20] is:

$$D = 1 - \frac{1}{N(N-1)}\sum_{j=1}^{s} xj(xj-1)$$

D – The index of discriminatory power
N - The number of unrelated strains tested
S - The number of different types
x_j - The number of strains belonging to the
j^{th} type, assuming that strains will be classified into mutually exclusive categories. Thus this formula can be used to determine the discriminatory power of MLST.

5 Results and Discussion

5.1 *Salmonella* serovars studied

The salmonella serovars having complete genome sequences available in the NCBI database were studied and tabulated in Appendix A. The allele numbers were assigned to all the unique sequences and a sequence type was determined by combining with allelic profile. This was tabulated below.

- *Balasubramanian Suryalakshmi, Author, Assistant Professor, Department of Biotechnology, P.S.R. Engineering College, Sivakasi E-mail: rbasurya@gmail.com*
- *Mareeswaran Hema & Velayutham RajaPria, Co-Authors,Final year Biotechnology,P.S.R. Engineering College, Sivakasi*

IJSER

Typhimurium D23580	10	10	8	11	10	6	9	34
Typhimurium 140285	9	10	9	10	4	6	9	35
Typhimurium T000240	9	10	8	10	4	9	9	36
Typhimurium ST4/74:SL1344	9	10	8	10	4	10	9	37
Typhimurium UK-1	9	10	8	10	4	6	9	38
Typhimurium CDC 2011K-0870	9	10	8	10	4	9	9	39
Typhimurium DT2	9	10	9	10	4	6	9	40
Typhimurium U288	9	10	9	10	4	9	10	41
Typhimurium USDA-ARS-USMARC-1899	9	10	9	10	4	6	10	42
Typhimurium CFSAN001921	9	10	9	10	4	6	10	43
Bongori	11	11	10	12	11	11	10	44
Bongori RKS3044	11	11	10	12	11	11	10	45
Gallinarum 287/91	12	12	11	13	12	12	11	46
Gallinarum /Pullorum RKS5078	13	13	11	14	13	13	12	47
Pullorum	13	13	11	14	13	13	12	48
Enteritidis	14	14	11	15	12	12	11	49
Dublin	15	14	12	16	14	14	13	50
Agona SL483	16	15	13	17	15	15	14	51
Agona 460004 2-1	16	15	13	17	15	15	14	52
Agona24249	16	15	13	17	15	15	14	53
Arizonae RKS2983	17	16	14	18	16	16	15	54
Arizonae62:z4,z231	17	16	14	18	16	16	15	55

Abaetetuba	1	1	1	2	2	1	1	8
Abony	1	1	1	2	2	1	1	9
Javiana	2	2	2	3	3	2	2	10
Montevideo507440-20	2	2	2	3	3	2	2	11
Montevideo USDA-ARS-USMARC-1903	2	2	2	3	3	2	2	12
Montevideo USDA-ARS-USMARC-1921	2	2	2	3	3	2	2	`13
Schwarzengrund	2	2	2	3	3	2	2	14
ParatyphiB	3	3	3	4	4	3	3	15
NewportSL254	4	4	4	5	5	4	4	16
NewportCVM21538	4	4	4	5	5	4	4	17
NewportCVM21550	4	4	4	5	5	4	4	18
NewportCVM21554	4	4	4	5	5	4	4	19
NewportUSDA-ARS-USMARC- 1927	4	4	4	5	5	4	4	20
NewportCVMN18486	4	4	4	5	5	4	4	21
TyphiCT18	5	5	5	6	6	5	5	22
TyphiTy2	5	6	5	6	6	5	6	23
Typhi P-stx-12	5	6	5	6	6	5	6	24
Tennessee	5	6	5	6	6	5	6	25
HeidelbergSL476	6	7	6	7	7	6	7	26
HeidelbergB182	6	7	6	7	7	6	7	27
Heidelberg CFSAN002069	6	7	6	7	7	6	7	28
Heidelberg41578	6	7	6	7	7	6	7	29
Choleraesuis	7	8	7	8	8	7	8	30
ParatyphiC	8	9	7	9	9	7	8	31
TyphimuriumLT2	9	10	8	10	4	8	9	32
Typhimurium 798	9	10	8	10	4	8	9	33
TyphimuriumD23580	10	10	8	11	10	6	9	34
Typhimurium140285	9	10	9	10	4	6	9	35
TyphimuriumT000240	9	10	8	10	4	9	9	36
TyphimuriumST4/74:SL1344	9	10	8	10	4	10	9	37
TyphimuriumUK-1	9	10	8	10	4	6	9	38
Typhimurium CDC 2011K-0870	9	10	8	10	4	9	9	39
Typhimurium DT2	9	10	9	10	4	6	9	40
Typhimurium U288	9	10	9	10	4	9	10	41
Typhimurium USDA-ARS-USMARC-1899	9	10	9	10	4	6	10	42
Typhimurium CFSAN001921	9	10	9	10	4	6	10	43
Bongori	11	11	10	12	11	11	10	44
Bongori RKS3044	11	11	10	12	11	11	10	45
Gallinarum 287/91	12	12	11	13	12	12	11	46
Gallinarum /Pullorum RKS5078	13	13	11	14	13	13	12	47
Pullorum	13	13	11	14	13	13	12	48
Enteritidis	14	14	11	15	12	12	11	49
Dublin	15	14	12	16	14	14	13	50
Agona SL483	16	15	13	17	15	15	14	51
Agona460004 2-1	16	15	13	17	15	15	14	52
Agona24249	16	15	13	17	15	15	14	53
Arizonae RKS2983	17	16	14	18	16	16	15	54
Arizonae62:z4,z231	17	16	14	18	16	16	15	55

5.2 Diversity among the candidate Loci analyzed

fliB, *fliD*, *metR*, *ompS*, *phoQ*, *rfaF* and *yehU* genes were analysed for study based on the abundance of nucleotide substitutions found in these regions. The sizes of the genes analysed ranged between 954 bp (*metR*) to 1690 bp (*yehU*). The locus with highest diversity was *ompS* with 18 different alleles found among the 55 different serovars studied. The ratio of non-synonymous (dN) to synonymous substitution(dS) were greater than 1, thus suggesting a high flexibility for amino acid changes (Table 2).

TABLE 2

CHARACTERISTICS OF THE SEVEN GENE LOCI

Genes	Length (bp)	No of Alleles	No of Synonymous polymorphic sites (dS)	No of Non-Synonymous polymorphic sites (dN)	dN/dS ratio
fliB	1210	17	238	911	3.83
fliD	1400	16	316	963	3.05
metR	954	14	240	711	2.96
ompS	1180	18	187	545	2.91
phoQ	1460	16	342	1092	3.19
rfaF	1047	16	251	763	3.04
yehU	1690	15	426	1254	2.94

5.3 Phylogenetic Analysis

The seven genes when concatenated were able to distinguish all the 55 *Salmonella* serovars. But the Neighbour Joining trees obtained (using MEGA software) by changing the order of seven genes during concatenation were found to be inconsistent. The phylogenetic distances varied while comparing both the trees in Figure 1 and Figure 2. Hence it was concluded that the gene order for concatenation plays a role in phylogeny.After the optimisation of gene order (i.e. *rfaF-yehU-fliB-phoQ-ompS-fliD-metR*), the Neighbour Joining tree. (Fig. 3) was able to discriminate all the fifty five *Salmonella* serovars with considerable phylogenetic distances and better bootstrap values.

IJSER

5.2.1. Concatenation of gene order *phoQ-fliB-fliD-ompS-rfaF-yehU-metR*

The genes were concatenated using UNION tool in the order *phoQ-fliB-fliD-ompS-rfaF-yehU-metR* and the tree was drawn using MEGA software. The phylogenetic distance was calculated using Neighbour Joining method and maximum composite likelihood and is indicated on the tree branches

5.2.2. Concatenation of gene order *rfaF-yehU-fliB-phoQ-ompS-fliD-metR*

The genes were concatenated using UNION tool in the order *rfaF-yehU-fliB-phoQ-ompS-fliD-metR* and the tree was drawn using MEGA software. The phylogenetic distance was calculated using Neighbour Joining method and maximum composite likelihood and is indicated on the tree branches.

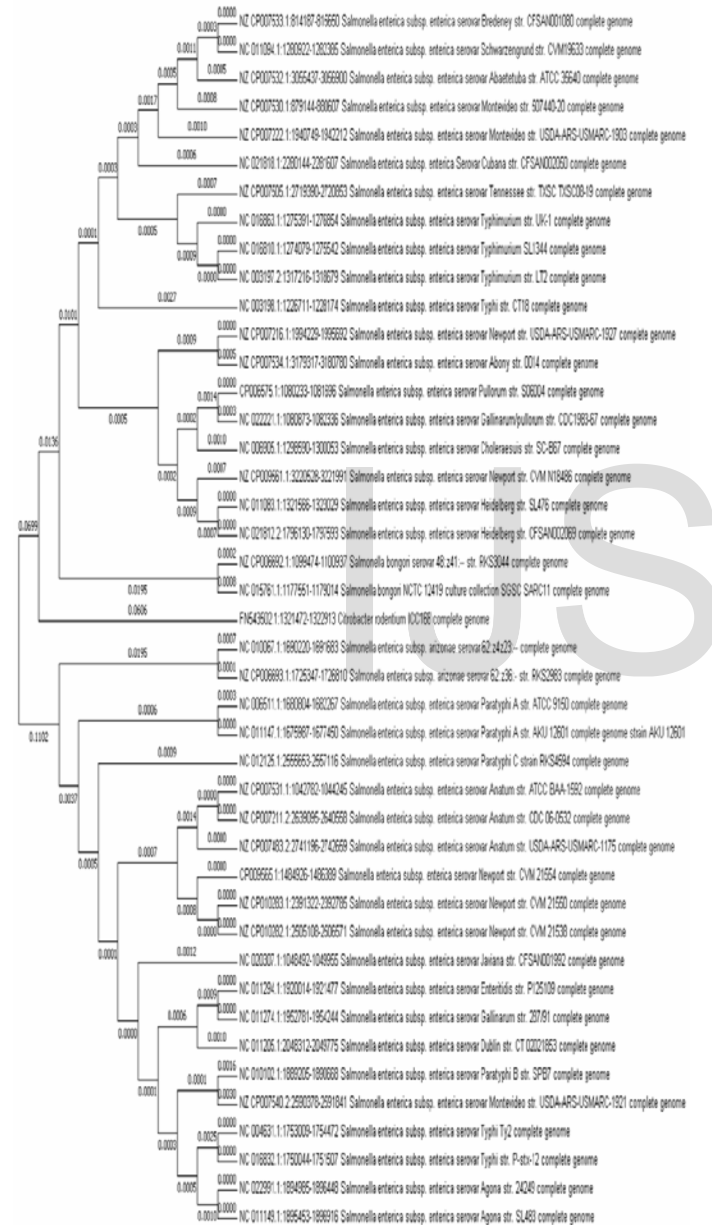

Fig. 1. The phylogenetic tree obtained for the gene order *phoQ-fliB-fliD-ompS-rfaF-yehU-metR*.

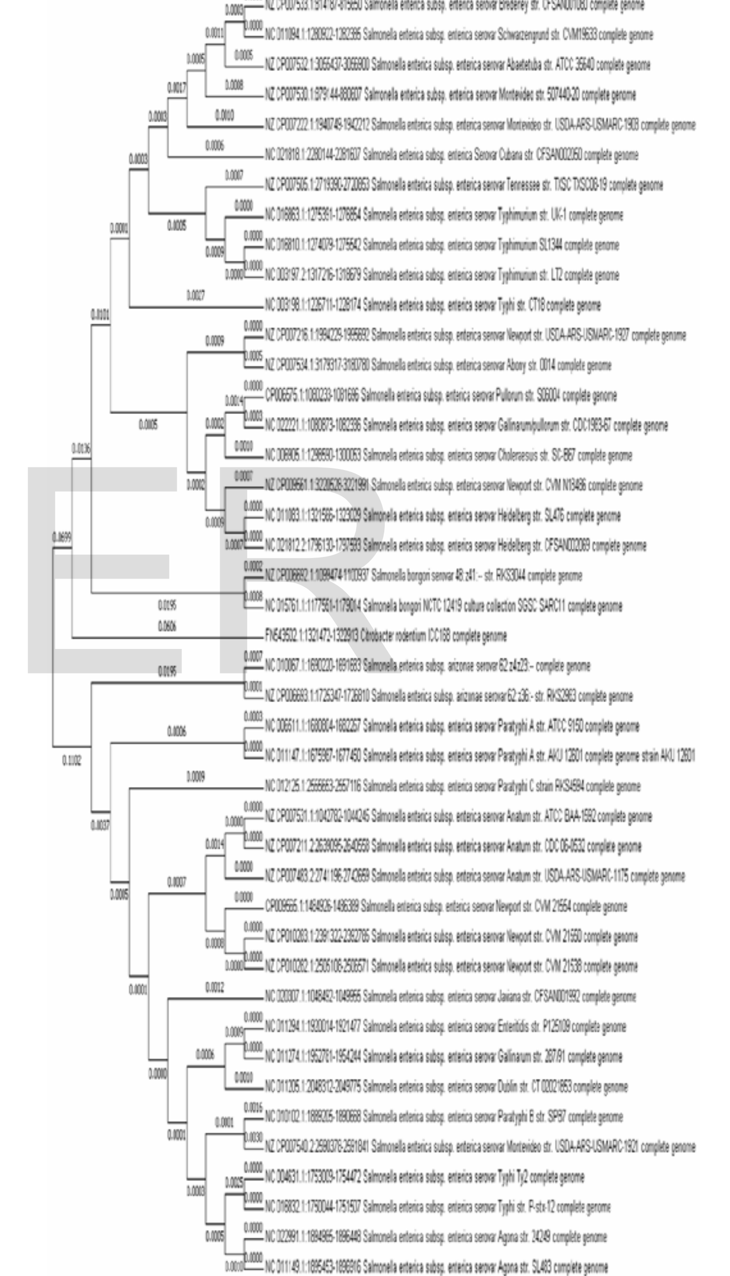

Fig. 2. The phylogenetic tree obtained for the gene order *rfaF-yehU-fliB-phoQ-ompS-fliD-metR*.

IJSER

5.2.3. Concatenation of gene order *phoQ-fliB-fliD-ompS-rfaF-yehU-metR*

The gene order was optimized using bootstrap test of phylogeny (in MEGA), which predicts the reliable phylogenetic tree. The greater the bootstrap value, the more reliable the tree is. *Citrobacter rodentium* ICC168 was used as an outgroup for phylogenetic study. The phylogenetic distance was calculated using Neighbour Joining method and maximum composite likelihood and is indicated on the tree branches. The bootstrap values are also indicated.

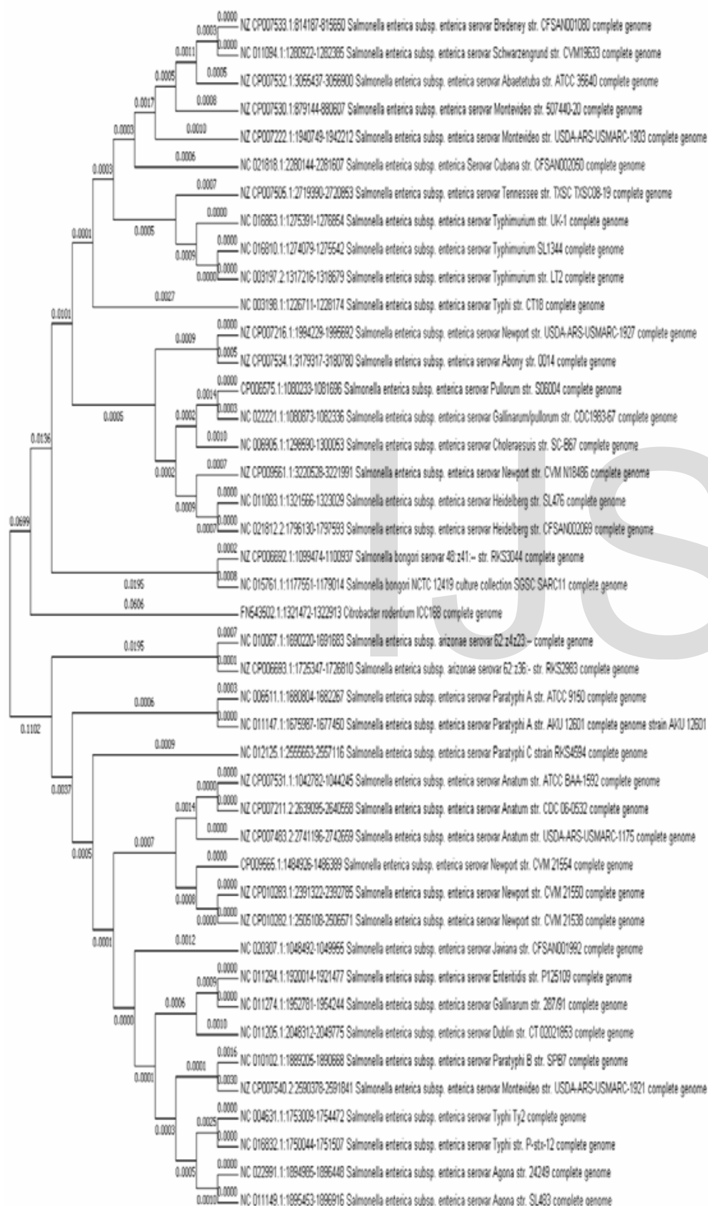

Fig. 3. The phylogenetic tree obtained for the optimized gene order *phoQ-fliB-fliD-ompS-rfaF-yehU-metR*.

5.4 Simpson's Index of Diversity

The seven loci scheme differentiates 55 serotypes of *Salmonella*. As the number of genes included for analysis increases, the discriminatory power of MLST also increases. The discriminatory power obtained by using the seven genes namely *fliB*, *fliD*, *metR*, *ompS*, *phoQ*, *rfaF* and *yehU* was 0.998

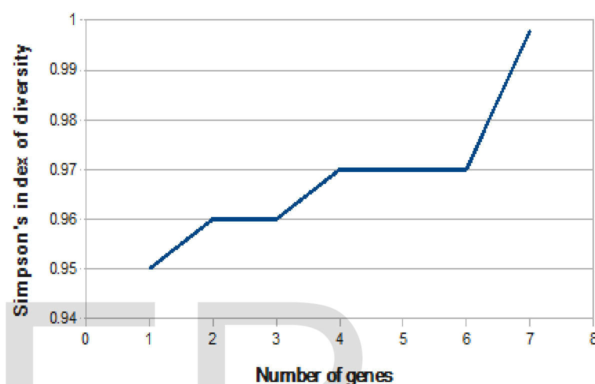

Fig. 4. Relation between the Simpson's index of diversity and the number of genes used to differentiate *Salmonella* serovars

5.5 Appendices

The salmonella serovars whose complete genome sequences were available in NCBI database were updated below.

IJSER

APPENDIX A
SALMONELLA SEROVARS STUDIED

Species	Serovar	Strain	NCBI Reference no
*S.enterica*subsp.enterica	Agona	SL483	NC_011149.1
*S.enterica*subsp.enterica	Agona	4600042-1	NZ_CP011259.1
*S.enterica*subsp.enterica	Agona	24249	NC_022991.1
*S.enterica*subsp.enterica	Abaetetuba	ATCC 35640	NZ_CP007532.1
*S.enterica*subsp.enterica	Abony	0014	NZ_CP007534.1
*S.enterica*subsp.enterica	Anatum	CDC 06-0532	NZ_CP007211.2
*S.enterica*subsp.enterica	Anatum	ATCC BAA-1592	NZ_CP007531.1
*S.enterica*subsp.enterica	Anatum	USDA-ARS- USMARC-1175	NZ_CP007483.2
*S.enterica*subsp.enterica	Bredeney	CFSAN001080	NZ_CP007533.1
*S.enterica*subsp.enterica	Choleraesuis	SC-B67	NC_006905.1
*S.enterica*subsp.enterica	Cubana	CFSAN002050	NC_021818.1
*S.enterica*subsp.enterica	Dublin	CT_02021853	NC_011205.1
*S.enterica*subsp.enterica	Enteritidis	P125109	NC_011294.1
*S.enterica*subsp.enterica	Gallinarum	287/91	NC_011274.1
*S.enterica*subsp.enterica	Gallinarum/pullorum	RKS5078	NC_016831.1
*S.enterica*subsp.enterica	Heidelberg	SL476	NC_011083.1
*S.enterica*subsp.enterica	Heidelberg	41578	NC_021810.1
*S.enterica*subsp.enterica	Heidelberg	B182	NC_017623.1
*S.enterica*subsp.enterica	Heidelberg	CFSAN002069	NC_021812.2
*S.enterica*subsp.enterica	Javiana	CFSAN001992	NC_020307.1
*S.enterica*subsp.enterica	Montevideo	507440-20	NZ_CP007530.1
*S.enterica*subsp.enterica	Montevideo	USDA-ARS-USMARC-1903	NZ_CP007222.1
*S.enterica*subsp.enterica	Montevideo	USDA-ARS-USMARC-1921	NZ_CP007540.2
*S.enterica*subsp.enterica	Newport	CVM 21550	NZ_CP01028.3
*S.enterica*subsp.enterica	Newport	CVM 21554	CVM N1848.6
*S.enterica*subsp.enterica	Newport	USDA-ARS-USMARC-1927	NZ_CP007216.1
*S.enterica*subsp.enterica	Newport	CVM N18486	NZ_CP009561.1
*S.enterica*subsp.enterica	Newport	SL254	NC_011080.1
*S.enterica*subsp.enterica	Paratyphi A	AKU_12601	NC_011147.1
*S.enterica*subsp.enterica	Paratyphi A	ATCC 9150	NC_006511.1
*S.enterica*subsp.enterica	Paratyphi B	SPB7	NC_010102.1
*S.enterica*subsp.enterica	Paratyphi C	RKS4594	NC_012125.1
*S.enterica*subsp.enterica	Pullorum	S06004	CP006575.1
*S.enterica*subsp.enterica	Schwarzengrund	CVM19633	NC_011094.1
*S.enterica*subsp.enterica	Tennessee	TXSC_TXSC08-19	NZ_CP007505.1
*S.enterica*subsp.enterica	Typhi	CT18	NC_003198.1
*S.enterica*subsp.enterica	Typhi	P-stx-12	NC_016832.1
*S.enterica*subsp.enterica	Typhi	Ty2	NC_004631.1
*S.enterica*subsp.enterica	Typhimurium	LT2	NC_003197.1
*S.enterica*subsp.enterica	Typhimurium	T000240	NC_016860.1
*S.enterica*subsp.enterica	Typhimurium	14028S	NC_016856.1
*S.enterica*subsp.enterica	Typhimurium	D23580	NC_016854.1
*S.enterica*subsp.enterica	Typhimurium	ST4/74	NC_016857.1
*S.enterica*subsp.enterica	Typhimurium	SL1344	NC_016810.1
*S.enterica*subsp.enterica	Typhimurium	UK-1	NC_016863.1
*S.enterica*subsp.enterica	Typhimurium	798	NC_017046.1
*S.enterica*subsp.enterica	Typhimurium	CDC 2011K-0870	NZ_CP007523.1
*S.enterica*subsp.enterica	Typhimurium	DT2	NC_022544.1
*S.enterica*subsp.enterica	Typhimurium	L-3553 DNA	NZ_AP014565.1
*S.enterica*subsp.enterica	Typhimurium	U288	NC_021151.1
*S.enterica*subsp.enterica	Typhimurium	CFSAN001921	NC_021814.1
*S.enterica*subsp. arizonae	62:z4,z23:--	RSK2980	NC_010067.1
S.enterica subsp. arizonae	62:z36	RKS2983	NZ_CP006693.1
Salmonella bongori	--	NCTC 12419	NC_015761.1

6 CONCLUSION

Single Locus Gene markers were not able to distinguish some of the *Salmonella* Serovars. On Considering the Housekeeping genes and Virulence genes in the identification of *Salmonella* serovars, the seven genes employed Say fliB, fliD, metR, ompS, phoQ, rfaF and yehU were able to distinguish all the 55 *Salmonella* serovars. Hence, the genotyping of *Salmonellae* using MLST differentiated the serovars more accurately than the serotyping method. Also, MLST can be used as a tool to classify and identify *Salmonella* serovars using the optimized concatenation gene order (rfaF-yehU-fliB-phoQ-ompS-fliD-metR) for which the discriminatory power was calculated as 0.998.

ACKNOWLEDGMENT

First and foremost, we bow down in front of the almighty who spelled the rain with blessings and showered us with good health and peace of mind for the successful completion of our research. We would also like to thank our family for carrying us through this journey showing us the right path at all times. We must sincerely acknowledge with deep appreciation the indispensable help and encouragement received from them.

REFERENCES

[1] J. Bockemuhl, M.Y. Popoff and F.W. Brenner, "Supplement 1998 (no.42) to Kauffmann-White Scheme", Research in Microbiology, no. 151, pp. 63–65, 2008.

[2] E.L. Hohmann,"Non-Typhoidal Salmonellosis", Clinical Infectious Diseases, no. 32; pp. 263–269, 2001.

[3] Y.F. Yen, Y.C. Lin, T.L. Chen, Y.Y. Chen and M.L. Lin, "Non-Typhoidal *Salmonella* Bacteremia in Adults", Journal of Microbiology, Immunology and Infection, no. 40, pp. 227–233, 2007.

[4] I. Gantois, R. Ducatelle, F. Pasmans, F. Haesebrouck and R. Gast,"Mechanisms of Egg Contamination by *Salmonella enteritidis*", FEMS Microbiology Reviews, no. 33; pp. 718-738, 2009.

[5] A. Voetsch, T. Gilder, F. Angulo, M. Farley, S. Shallow,"Food Net Estimate of the Burden of Illness Caused by Non Typhoidal Salmonella Infections in the United States", Clinical Infectious Diseases, no. 38; pp. 127-134, 2004.

[6] J.A. Crump, S.P. Luby, E.D. Mintz,"The Global Burden of Typhoid Fever", Bulletin of the World Health Organization, no. 82, pp. 346–353, 2004.

[7] A. Haque, J. Ahmed,J. Qureshi,"Early Detection of Typhoid by Polymerase Chain Reaction", Annals of Saudi Medicine no. 19, pp. 337-340, 1999.

[8] D.G. White,H. Harbottle, P.F. McDermott, R.D. Walker, S. Zhao," Comparison of Multilocus Sequence Typing, Pulsed - Field Gel Electrophoresis and Antimicrobial Susceptibility Typing for Characterization Of *Salmonella enterica* serotype Newport Isolates", Journal of Clinical Microbiology 44; 2449-2457.

[9] B.A. Lindstedt, T. Vardund, L. Aas, G. Kapperud," Multiple-Locus Variable-Number Tandem-Repeats Analysis of *Salmonella enterica* subsp. *enterica* Serovar Typhimurium Using PCR Multiplexing and Multicolor Capillary Electrophoresis", Journal of Microbiological Methods, no. 59, pp. 163-172, 2004.

[10] B. Curran, D. Jonas, H. Grundmann, T. Pitt, C. Dowson," Development of a Multilocus Sequence Typing Scheme for the Opportunistic Pathogen *Pseudomonas aeruginosa*", Journal of Clinical Microbiology, no. 42, pp. 5644–5649, 2004.

[11] S. Sukhnanand, S. Alcaine, L. Warnick, W. Su, J. Hof,"DNA Sequence - Based Subtyping and Evolutionary Analysis of Selected *Salmonella enterica* Serotypes", Journal of Clinical Microbiology, no. 43, pp. 3688-3698, 2005.

[12] W.H. Ewing, P.R. Edwards,"The Principal Divisions and Groups of Enterobacteriaceae and their Differentiation", International bulletin of bacteriological nomenclature and taxonomy, no. 10, pp. 1-12, 1960.

[13] N.K. Petty, R. Bulgin, V.F. Crepin, C. Tarraga, G.N. Schroeder,"The *Citrobacter rodentium* Genome Sequence Reveals Convergent Evolution With Human Pathogenic *E. coli",*Journal of Bacteriology, no. 192, pp. 525-538, 2010.

[14] E.J. Richardson, B. Limaye, H. Inamdar, A. Datta, K.S. Manjari,"Genome Sequences of *Salmonella* e*nterica* SerovarTyphimurium, Choleraesuis, Dublin and Gallinarum Strains of Well-defined Virulence in Food - Producing Animals", Journal of Bacteriology, no. 193, pp. 3162-3163, 2011.

[15] N.R. Thomson, D.J. Clayton, D. Windhorst, G. Vernikos, S. Davidson,"Comparative Genome Analysis of *Salmonella enteritidis* PT4 and *Salmonella Gallinarum* 287/91 Provides Insights into Evolutionary and Host Adaptation Pathways",Genome Research, no. 18, pp. 1624-1637, 2008.

[16] D.M. Hillis, J.J. Bull,"An Empirical Test of Bootstrapping as a Method For Assessing Confidence in Phylogenetic Analysis", Systematic Biology, no. 42, pp. 182-192, 1993.

[17] W.F. Fricke, M.K. Mammel, P.F. McDermott, C. Tartera, D.G. White,"Comparative Genomics of 28 *Salmonella enterica* Isolates: Evidence for CRISPR-Mediated Adaptive Sublineage Evolution",Journal of Bacteriology, no. 193, pp. 3556-3568, 2011.

[18] K. Tamura, J. Dudley, M. Nei, S. Kumar,"MEGA 4 : Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0.",Molecular Biology and Evolution, no. 24, pp. 1596-1599, 2007.

[19] P. Hunter, M. Gaston,"Numerical Index of the Discriminatory Ability of Typing Systems: An Application of Simpson's Index of Diversity", Journal of Clinical Microbiology, no. 26, pp. 2465-2466, 1988.

[20] Doyle, McQuiston, H. Leon, Wertheim, Fields, Tauxe and Logsdon,"Molecular Phylogeny of *Salmonella*: Relationships Among *Salmonella* Species and Subspecies Determined from Four Housekeeping Genes and Evidence of Lateral Gene Transfer Events",Journal of Bacteriology, no. 190, pp. 7060-7067, 2008.

[21] M. McCleland, steffen Porwollik, rita Mei-yi Wong, Michael,"Evolutionary Genomics of Salmonella: Gene Acquisitions Revealed by Microarray Analysis",Proceedings of the National Academy of Sciences of the United States of America, Vol 99, issue 13, pp. 8956-8961, 2002.

[22] Samuel I. Miller, Anne M. Kukral and John J. Mekalanos," A Two Component Regulatory System (PhoQ PhoP) Controls Salmonella Typhimurium Virulence", Proceedings of the National Academy of Sciences of the United States of America, Vol 86, pp. 5054-5058, 1989.

[23] I. Ross and M. Heuzenroeder,"Discrimination Within Phenotypically Closely Related Definitive Types of *Salmonella enterica* Serovar Typhimurium by Multiple Amplification of Phage Locus typing Technique", Journal of Clinical microbiology, no. 43, pp. 1604-1611, 2004.

[24] D. Sirisena, R. Maclachlan, A. Hessel and K. Sanderson," Molecular Analysis of rfaD Gene for Heptose Synthesis and rfaF gene for Heptose Transfer in Lipopolysaccharide Synthesis in *Salmonella* Typhimurium", Journal of Bacteriology, no. 176, pp. 2379-2385, 1994.

[25] K. Tamura, D. Peterson, N. Peterson, Stecher, M. Nei and S. Kumar,"MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance and Maximum Parsimony Methods:, Molecular Biology and Evolution, no. 28, pp. 2731-2739, 2011.

IJSER