

## Genetic dynamics and molecular epidemiology of *Salmonella typhi* S9-1

K.L. Thong, S. Nair, G. Subramaniam, S. Puthucheary<sup>1</sup>, R. Yassin<sup>2</sup>, Y.M. Cheong<sup>3</sup>, S.L. Liu<sup>4</sup>, K.E. Sanderson<sup>4</sup>, A.M. Cordano<sup>5</sup>, T. Pang and the International Multicentre Study Group on the Molecular Epidemiology & Genetic Diversity of *S. typhi* (P. Sudarmono, E. Soewandjo, I. Handojo, M. Padmidewi, S. Sarasombath, M. Passey, B. Combs, R. Chaudhry)

### Abstrak

Telah dilakukan suatu analisa secara Pulse-field gel electrophoresis (PFGE) dengan restriksi endonuklease XbaI, SpeI, AvrII dan ICeul pada isolat *S. typhi* dari berbagai wilayah geografis yang diperoleh secara sporadik, kejadian luar biasa (KLB), lingkungan, demikian pula isolat dari kasus fatal dan non-fatal demam tifoid. Kesimpulan utama yang diperoleh menunjukkan bahwa: (1) terdapat keragaman genetik yang cukup bermakna pada isolat-isolat yang diperiksa; (2) isolat dari KLB biasanya lebih homogen dibandingkan dengan isolat yang diperoleh secara sporadik; (3) terdapat pergerakan tipe PFGE tertentu di negara-negara Asia Tenggara; (4) isolat yang diperoleh dari lingkungan juga menunjukkan keragaman, dan pada isolat-isolat ini terlihat tipe PFGE yang sama dengan isolat yang diperoleh dari kasus klinis; (5) terdapat korelasi pola PFGE pada isolat yang diperoleh dari kasus yang fatal; (6) terdapat keragaman yang terbatas di daerah di mana demam tifoid baru dikenal; (7) terdapat perbedaan bermakna pada ukuran total dari genom; (8) telah terjadi perubahan genom pada operon *rrn* dari berbagai isolat yang diperiksa. Hasil penelitian ini menunjukkan bahwa terdapat plastisitas dari genom *S. typhi*, yang kemungkinan besar berhubungan dengan potensi patogen dan keganasan dari organisme ini.

### Abstract

A large number of *S. typhi* isolates from various geographic regions, representing sporadic, outbreak and environmental isolates, as well as isolates from fatal and non-fatal cases of typhoid fever have been analysed by pulsed-field gel electrophoresis (PFGE) with the restriction endonucleases XbaI, SpeI, AvrII, and ICeul. The main conclusions are: (1) significant genetic diversity exists among these isolates; (2) outbreak isolates are more homogeneous compared to sporadic isolates; (3) there is movement of certain PFGE types among the countries of Southeast Asia; (4) environmental isolates also show diversity and share PFGE types with clinical strains; there appear to be a correlation between PFGE pattern and isolates from fatal disease; (6) limited diversity exists in a region where typhoid fever is a recent introduction; (7) there is significant diversity in total genome size; (8) genomic rearrangements within the *rrn* operon has occurred among many of these strains. Our findings emphasize the genomic plasticity of *S. typhi* and it is likely that this is closely related to the pathogenic potential and virulence of this organism.

### INTRODUCTION

Based on biochemical studies of enzyme polymorphisms using multilocus enzyme electrophoresis (MLEE), *Salmonella typhi* has been assumed to be a homogeneous and highly clonal pathogen<sup>1</sup>. A study of more than 300 *S. typhi* strains grouped them into only two clones worldwide<sup>1</sup>. Homogeneity among

strains has also been observed in membrane protein and plasmid profiles. However, field observations of typhoid fever suggest that significant diversity and possible strain variations may exist. For example, differing disease severity has been observed in various endemic regions. In Indonesia, for example, severe disease with serious complications and significant mortality is frequently seen while in other parts of Asia, the disease manifestations are milder. Variable vaccine efficacy with the Ty21a oral typhoid vaccine has also been reported with good efficacy noted in studies in Chile and Egypt, while poor efficacy was seen in Indonesian trials, where only around 50% of immunized persons were protected. Additionally, there is also anecdotal evidence of antigenic variation as detected during serological diagnostic testing for typhoid fever. We thus formulated a working hypothesis stating that significant genetic diversity exists among strains/isolates of *S. typhi* which may be

<sup>1</sup>Institute of Postgraduate Studies & Research and Dept. of Medical Microbiology, University of Malaya, Kuala Lumpur, Malaysia;

<sup>2</sup>Div. of Bacteriology, Institute for Medical Research, Kuala Lumpur, Malaysia;

<sup>3</sup>Pfizer Malaysia, Kuala Lumpur, Malaysia;

<sup>4</sup>Dept. of Biological Sciences, University of Calgary, Calgary, Alberta, Canada;

<sup>5</sup>Institute of Public Health, Santiago, Chile.

related to their biological behaviour, including the ability to cause disease.

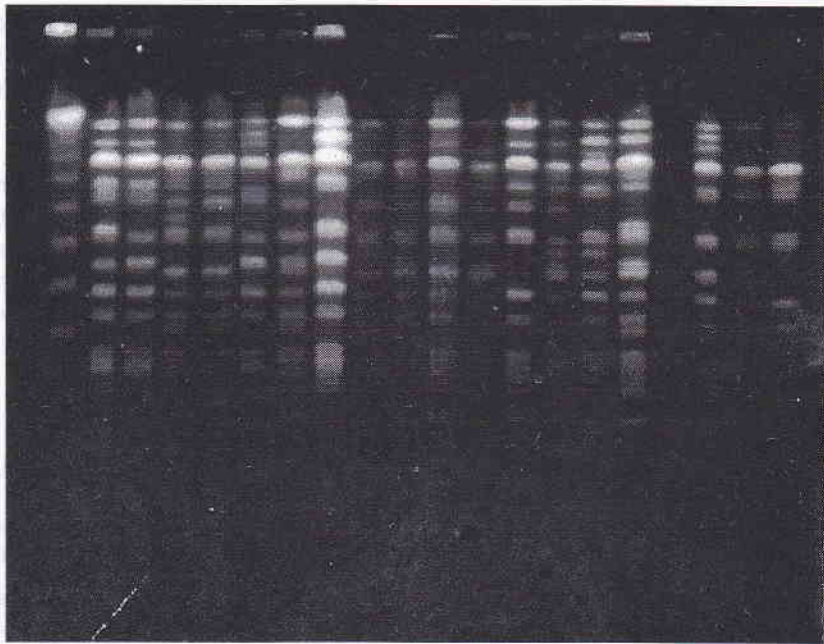
## METHODOLOGY

Our experimental approach was to assemble a collection of *S. typhi* isolates from various endemic regions of the world and performing molecular analysis on these isolates, with a view to assessing the extent of genetic diversity. Various methods were utilized but emphasis was placed on macrorestriction DNA analysis using the technique of pulsed-field gel electrophoresis (PFGE)<sup>2-6</sup>. In this technique, genomic *S. typhi* DNA fragments, after digestion with infrequently-cutting restriction endonucleases (primarily *Xba*I, *Spe*I and *Avr*II), were separated using PFGE which, after standard staining of the gel with ethidium bromide, provided a pattern of DNA fragment distribution representing the entire genome. These patterns were then subjected to comparative analysis, either visually or using RFLP (restriction fragment length polymorphism) pattern analysis software (GelCompar, USA), in order to assess diversity. The objective measure of similarity/dissimilarity between strains was the F value (coefficient of similarity) which calculates the number of shared bands between 2 isolates relative to the total number of bands. The formula used to calculate this coefficient was  $F = \frac{nx+ny}{2nxy}$  where  $nx$  = the number of bands in isolate x;  $ny$  = the number of bands in isolate y and  $nxy$  = the number of bands shared between the 2 isolates. By this measurement, an F value of 1.0 indicates identity of 2 isolates and a value of 0 indicates total dissimilarity. The advantages of PFGE include relatively few DNA bands (ideally between 20-30), the fact that changes noted are those distributed over the entire genome, technically straightforward (no requirement for blots, hybridizations) and the possibility of objective, quantitative analysis of the DNA band patterns obtained. In addition to PFGE, Southern blot hybridization with selected gene probes and DNA sequencing of PCR-amplified 16S rRNA genes were also performed using standard methodologies.

## RESULTS

The main results of the study are as follows:

1. Significant diversity was observed among 400 isolates of *S. typhi* obtained from various endemic regions including Malaysia, Indonesia, Thailand, India, Chile and Papua New Guinea. The F values of these isolates from sporadic cases of typhoid fever ranged from approximately 0.6-1.0<sup>2-4</sup>. In contrast, analysis of isolates from several Malaysian outbreaks of typhoid fever showed more limited diversity and greater homogeneity of PFGE patterns ( $F = 0.8-1.0$ ) indicating that these isolates belonged to a single clone or several closely related clones<sup>2</sup>. There appeared to be no correlation between phage types and PFGE patterns.
2. By comparative analysis of PFGE patterns, it was established that similar patterns were seen among isolates from Malaysia, Indonesia and Thailand. This would suggest the movement of these strains of *S. typhi* in the Southeast Asian region<sup>3</sup>, although the possibility of these strains arising independently in the various regions cannot be discounted.
3. Similar diversity was noted among environmental isolates of *S. typhi* isolated from the Mapocho river in Santiago, Chile (Figure 1). Comparison of the molecular profiles of these isolates with clinical isolates from the same area and at the same time period showed several similarities (Figure 1), thus proving the link of transmission between sewage-contaminated water supplies and human disease<sup>4</sup>.
4. Certain PFGE molecular types were associated with the ability to cause fatal disease<sup>5</sup>. Of 53 isolates tested, all 12 isolates from fatal cases of typhoid had the PFGE pattern X1S1A1, whereas 41 isolates from non-fatal cases had various other PFGE pattern combinations, the most common being X2S1A2. It was also shown that in Papua New Guinea, where typhoid fever was introduced relatively recently (10-15 years ago), limited genetic diversity ( $F=0.78-1.0$ ) was seen even among the sporadic isolates of *S. typhi*<sup>5</sup>. This is in contrast to what was observed in Southeast Asia<sup>2,3</sup>.
5. Analysis of multiple isolates from the same patient with relapsing typhoid fever indicates that the PFGE pattern remained stable, even under selection pressure of antibiotic treatment and the host immune response.
6. Detailed analysis of genome sizes of 17 *S. typhi* isolates from different parts of the world showed significant genome size variation. Genome sizes differed by as much as 959 kbp (representing about 20% of the total genome size and containing sufficient information for approximately 600 genes), ranging from 3,964 to 4,923 kbp with a mean genome size of 4,528 kbp<sup>6</sup>. These observations were obtained following digestion of genomic DNA with the three restriction endonucleases, *Xba*I, *Avr*II and *Spe*I.
7. It was also shown that major genomic rearrangements has occurred in the *S. typhi* genome and



**Figure 1.** PFGE pattern profile of Chilean *S. typhi* isolates after digestion with XbaI.

Lane 1 = lambda concatemer markers; Lanes 2-14: environmental isolates; Lanes 15-19: human isolates (lane 17 is empty); Lane 20: *S. typhi* ATCC strain.

that a degree of genomic plasticity is clearly present<sup>7</sup>. The rearrangements have occurred within the *rrn* operon and involved the various fragments defined after digestion with the restriction endonuclease ICeu-I. In a study of more than 100 isolates of *S. typhi* from diverse geographical regions it was shown that these rearrangements have also occurred amongst recent clinical isolates from patients with typhoid fever. This is in contrast to the conservation in genome structure seen in other salmonellae and gram-negative bacteria.

8. In the attempt to determine where the diversity resides within the *S. typhi* genome, restriction digests were also probed with a variety of gene probes. Polymorphisms were detected among isolates when *fliC*, *groEL*, 16S rRNA and *rfbS* gene probes were used to probe restricted genomic DNA. In contrast, homogeneity in hybridization patterns was noted when *ompC* and *viaB* gene probes were used.
9. We were also interested in the taxonomic and phylogenetic relationships of *S. typhi* with other salmonellae. Determination of the complete 16S rRNA gene sequences from 15 Salmonella serovars indicated that the percentage similarity of sequences among these serovars was very high (97.4 to 99.2%). Base substitutions occurred

throughout the gene with two apparent hotspots, one of which corresponded to the V6 variable region.

## DISCUSSION AND CONCLUSIONS

The results of our studies clearly shows that significant diversity exists among sporadic isolates of *S. typhi* from diverse geographical regions and that this diversity was more limited among isolates obtained during outbreaks of typhoid fever. In addition to its value in strain differentiation and molecular typing for epidemiological studies, the PFGE methodology used to perform the molecular analysis of genomic DNA from *S. typhi* was also shown to be of value in further studies to determine where the genetic variation occurs within the genome, tracking the movement of certain molecular types in a particular geographic region, in establishing the link between environmental and human isolates, and in investigating correlations between molecular profiles and disease severity. The evidence gathered would seem to validate our hypothesis that significant genetic diversity exists at the DNA level among natural human and environmental isolates of *S. typhi*.

The PFGE approach is a measurement of diversity within restriction endonuclease recognition sites distributed throughout the entire genome which may

have arisen as a result of: (1) mutations which remove or create a recognition site; (2) insertions, deletions, duplications or inversions; (3) mobile genetic elements (e.g. prophage, transposons, etc). Liu & Sanderson<sup>7</sup> suggested that lateral transfer of DNA from an external source is the most important mechanism for producing genetic changes in *S. typhi*. The observed genetic diversity among these *S. typhi* isolates is in good agreement with work of Sanderson et al<sup>7</sup> who showed that the *S. typhi* genome had undergone major rearrangements not seen in other *Salmonella* spp. and which strongly suggests the genomic plasticity of the *S. typhi* genome.

Although the biological consequences of the genetic diversity and genome size variation noted among *S. typhi* in the present study are unclear, the observed results has important implications for virulence and biological behaviour of *S. typhi* and the rational design of control strategies. Given the epidemiology of typhoid fever, genome diversity as a result of genomic flexibility/plasticity is compatible with the fact that *S. typhi* is normally confronted with a wide range of growth conditions in the course of its transmission, pathogenesis and survival in the environment. The organism needs to survive in contaminated food and/or water, in different bodily sites with differing microenvironments (e.g. the gastrointestinal tract, intracellular milieu of mononuclear cells, blood during bacteremia, bile, urine, feces, etc.) and in environmental reservoirs (e.g. raw sewage, river water). Given this potentially wide range of growth conditions, a degree of flexibility in rearranging relevant genes and in genome size would clearly be an advantage for the organism's survival. At present, we do not know of any biological traits or phenotypes which are associated with the various molecular types or with genome sizes of *S. typhi* isolates although this is clearly a subject for further investigation. Ultimately, however, an understanding of the nature and extent of genetic differences among natural human isolates of *S. typhi* may provide important

insights into the basis of geographic variations in disease symptoms and severity and has important implications for diagnostic detection of *S. typhi* and development of vaccines.

In the broader context of bacterial population ecology, Levin et al. have stated that the genetic dynamics and flexibility of a wide variety of pathogens, from influenza and HIV to parasites, have major implications for the dynamics of control, the evolution of resistance, and the emergence of new pathogens<sup>8</sup>.

## REFERENCES

1. Selander RK, Beltran P, Smith NH, Helmuth R, Rubin FA, Kopecko DJ, et al. Evolutionary genetic relationships of clones of *Salmonella* serovars causing human typhoid and other enteric fevers. *Infect Immun* 1990; 58: 2262-75.
2. Thong KL, Cheong YM, Puthuchery SD, Koh CL, Pang T. Epidemiologic analysis of sporadic *Salmonella typhi* isolates and those from outbreaks by pulsed-field gel electrophoresis. *J Clin Microbiol* 1994; 32: 1135-41.
3. Thong KL, Puthuchery SD, Yassin RM, Sudarmono P, Padmidewi M, Soewandjojo E, et al. Molecular analysis of *S. typhi* from Southeast Asia using PFGE. *J Clin Microbiol* 1995; 33: 1938-41.
4. Thong KL, Cordano AM, Yassin RM, Pang T. Molecular analysis of environmental and human isolates of *Salmonella typhi*. *Appl Environ Microbiol* 1996; 62: 271-4.
5. Thong KL, Passey M, Clegg A, Combs BG, Yassin RM, Pang T. Molecular analysis of isolates of *Salmonella typhi* obtained from fatal, mild and sporadic cases of typhoid fever. *J Clin Microbiol* 1996; 34: 1029-33.
6. Thong KL, Puthuchery SD, Pang T. Genome size variations among recent human isolates of *Salmonella typhi*. *Res Microbiol* 1997; 148: 229-35.
7. Liu SL, Sanderson KE. Rearrangement in the genome of the bacterium *Salmonella typhi*. *Proc Natl Acad Sci USA* 1995; 92: 1018-22.
8. Levin SA, Grenfell B, Hastings A, Perelson AS. Mathematical and computational challenges in population biology and ecosystems science. *Science* 1997; 275: 334-43.