**Editorial**                                                                                                    **Open Access**

# Comparative Genomics of Salmonella Could Reveal Key Features of Adaptation

Gerardo M Nava[*] and Yajaira Esquivel-Hernandez

*Autonomous University of Queretaro, Mexico*

[*]**Corresponding author:** Gerardo M Nava, Faculty of Chemistry, Department of Research and Graduate Studies in Food, Autonomous University of Queretaro, Mexico, Tel: +52-55-56567871; E-mail: gerardomnava@gmail.com

## Editorial

Worldwide, *Salmonella enterica* remains an important health threat. A recent study by the World Health Organization estimated that non-typhoidal *S. enterica* causes ~ 230,000 deaths annually [1]. The biology of this pathogen has been studied for almost a century [2]; until recently, however, we have started to elucidate genomic features of adaptation to its hosts. Now, it is know that *S. enterica* has evolved to establish sympatric (generalists) and allopatric (specialists) and associations with its host. For example, *S. enterica* subspecies *enterica*, serotypes Choleraesuis, Dublin and Gallinarum have established allopatric association with porcine, bovine and avian species, respectively. In contrast, serotype Enteritidis and Typhimurium have adapted a sympatric strategy to colonize the intestinal tract of a broad number of avian and mammalian species [3]. This progress in knowledge has been accomplished with the aid of molecular microbiology.

The implementation of high-throughput sequencing, bioinformatics, data mining and comparative genomics approaches has uncover key genomic adaptations employed by *S. enterica* to define bacterial ecotypes or subpopulations suited to inhabit specific niches [4]. Nowadays, it is possible to identify specific genes responsible of ecological adaptation. It is a golden opportunity to redefine the biology, diversity and host-specificity of *S. enterica*, and assigning unique taxonomic designations (i.e., biotypes, serotypes, ecotypes, ribotypes, etc.) to particular diseases.

Fortunately, there is a massive amount of genomic data waiting to be explored to uncover new genes responsible of adaptive and diversifying process, as well as reductive evolution in *S. enterica*. For example, as of January 2016, the Integrated Microbial Genomes (IMG) system [5] contains ~ 1,000 Salmonella whole-genomes, covering at least 138 serovars. We carried out a bioinformatics analysis of this genomic dataset (936 genomes) and identified important features of genomic variation. Briefly, it was found that Salmonella genome size varies considerably, ranging between 3.39 and 5.59 Mbp. Also, the average gene count was 4,742 (ranging from 3,969 to 9,898 genes) (Figure 1A). Remarkably, 11% of the genomes (106/936) comprise more than 5,000 genes, and 15% (143/936) comprise less than 4,500 genes (Figure 1B). Also, a correlation analysis showed no linear relationship (R2 = 0.0027) between genome size and gene count (Figure 2). Interestingly, eleven genomes of *Salmonella typhimurium* and *typhi* have an average genome size of 4.47 Mbp but an average gene count of 8,310. This discrepancy between genome size and gene number could reveal an important host-adaptation process in this pathogen.
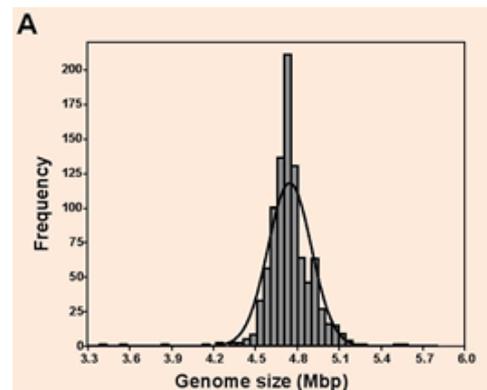


**Figure 1A:** Differences in genome size and gene count in Salmonella spp. Whole-genomes (n = 936) covering at least 138 serovars were analyzed. Genomes were obtained from the Integrated Microbial Genomes (IMG) system. Histograms showing A. Salmonella genome size (range between 3.39 and 5.59 Mbp).
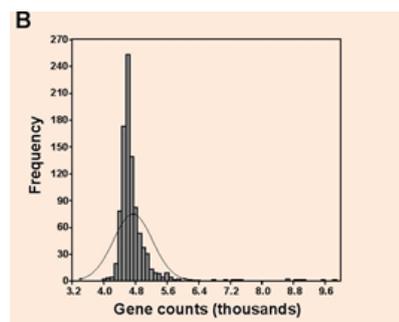


**Figure 1B:** Differences in genome size and gene count in Salmonella spp. Whole-genomes (n = 936) covering at least 138 serovars were analyzed. Genomes were obtained from the Integrated Microbial Genomes (IMG) system. Histograms showing B. Gene count (range between 3,969 and 9,898 genes). Black line depicts a normal distribution.
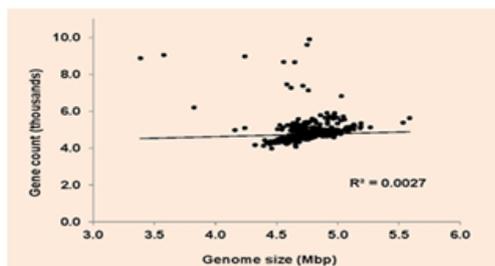
**Figure 2:** Gene count vs Genome size.

These genomic analyses with *Salmonella* spp. point out important biological questions that need to be addressed via datamining and comparative genomic analysis. This opportunity is unique due to the large amount of genomic data available and current advancements in bioinformatics. Moreover, it is an opportunity for colleges around the word to outline and renew academic curricula and improve undergraduate and graduate programs focused on analysis and interpretation of comparative genomic data. Thus, we are facing golden years for molecular microbiology.

## References

1.  Havelaar AH, Kirk MD, Torgerson PR, Gibb HJ, Hald T, et al. (2015) World Health Organization Global Estimates and Regional Comparisons of the Burden of Foodborne Disease in 2010. PLoS Med 12: e1001923.

2.  Savage WG, White PB (1923) Rats and Salmonella Group Bacilli. J Hyg (Lond) 21: 258-261.

3.  Foley SL, Johnson TJ, Ricke SC, Nayak R, Danzeisen J (2013) Salmonella pathogenicity and host adaptation in chicken-associated serovars. Microbiol Mol Biol Rev 77: 582-607.

4.  den Bakker HC, Moreno SAI, Govoni G, Cummings CA, Ranieri ML, et al. (2011) Genome sequencing reveals diversification of virulence factor content and possible host adaptation in distinct subpopulations of *Salmonella enterica*. BMC Genomics 12: 425.

5.  Markowitz VM, Chen IM, Palaniappan K, Chu K, Szeto E, et al. (2014) IMG 4 version of the integrated microbial genomes comparative analysis system. Nucleic Acids Res 42: D560-D567.