**Ciprofloxacin induced antibiotic resistance in *Salmonella* *Typhimurium* mutants and genome analysis**

**Aarti S. Kakatkara, Anubrata Dasc, Ravindranath Shashidharab#**

**Food Technology Division, Bhabha Atomic Research Centre, Mumbai, Indiaa**

**Life Sciences, Homi Bhabha National Institute (Deemed to be University), Mumbai, Indiab**

**Bio Organic Division, Bhabha Atomic Research Centre, Mumbai, Indiac**

**#Corresponding author: Ravindranath Shashidhar, Food Technology Division, Bhabha Atomic Research Centre, Trombay, India- 400085, (O)+91 22-25593961**

**Address correspondence to** [***shashi@barc.gov.in***](mailto:shashi@barc.gov.in)

**Journal name Archives of Microbiology**

**Supplementary Information Materials and Methods**

**Whole Genome Sequencing (WGS) and Bioinformatic tools used**

WGSwas outsourced to Genotypic Solutions Pvt Ltd Bangalore India. Reference based whole genome sequencing of *Salmonella* *Typhimurium* was carried out using reference genome available at NCBI for of *S*. *Typhimurium*, strain LT2. More than 1000 X for all the samples, of sequencing coverage was achieved for the genome of approximate size 4.8 MB. More than 90% of the reference genome was covered at 1X and 20X by good quality data which confirms the choice of reference and the sufficiency of the data for reference based WGS. The consensus sequence resulted from the analysis was compared with the reference sequence in order to know genome significant rearrangements if any. The data obtained from the sequencing analysis was analyzed in Jupyter notebook with python 3.7 backend in the Anaconda package version 2-2.4.0. The site wise mutation data was compiled to a list of mutations per gene. The sequence data was uploaded in the public server usegalaxy.org and annotated by Prokka genome annontation package (Afgan et al. 2018). The circos plot was made by compiling the chromosome coordinates and annotates mutations and plotting the data with circos-0.69-9 (Krzywinski et al. 2009). The distribution of mutations per gene was calculated using the Pandas library in Anaconda and selected genes with mutations at higher frequencies were chosen to study the pattern of distribution in these individual genes. For the proteins analyzed further, the protein sequences obtained after annotation were submitted to Phyre2 web portal for protein modeling, prediction and analysis to obtain homology models use the normal mode (Kelley et al. 2015). The homology models from PHYRE2 server showing 100 % confidence were used for further structure visualization and analyse by PyMol package version 2). The genes with non-synonymous mutations were compiled and the annotated list of proteins were obtained from galaxy server. The Uniprot accession numbers were appended to this list of proteins and they were submitted to the Panther database public server (Mi et al. 2019) and DAVID database (Huang et al. 2009) for gene function classification. After obtaining the pathway classification from the databases the classified genes were sorted by number of genes mutated in each pathway and this was plotted in Pandas using Matplotlib library. The protein-protein interaction of the DNA metabolic genes and resistance genes were obtained from STRING v11 public server (Szklarczyk et al. 2019). To find antibiotic resistance genes, we submitted our assembled contig for all mutants at ResFinder 4.1 public server (Zankari et al. 2012) Comprehensive Antibiotic Resistance Database public server (Alcock *et al.*, 2020) and ARGminer Database public server (Arango-Argoty et al. 2020). The servers returned a list of genes known for resistance to fluoroquinolones and we studied these gene sequences for non-synonymous substitutions to confirm the findings. The relationship between *iroC*, *recD*, *sbcC*, and *gyrA* were obtained by intensive modelling at PHYRE 2 server and string database (Szklarczyk et al. 2019)

**References**

Alcock B, Rapheynya AR, Lau TTY, et al (2020) CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. Nucleic Acids Res 48:D517–D525. <https://doi.org/10.1093/nar/gkz935>

Afgan E, Baker D, Batut B, et al (2018) The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. Nucleic Acids Res 46:W537–W544. https://doi.org/10.1093/nar/gky379

Arango-Argoty GA, Guron GKP, Garner E, et al (2020) ARGminer: a web platform for the crowdsourcing-based curation of antibiotic resistance genes. Bioinformatics 36:2966–2973. https://doi.org/10.1093/bioinformatics/btaa095

Huang DW, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4:44–57. https://doi.org/10.1038/nprot.2008.211

Kelley LA, Mezulis S, Yates CM, et al (2015) The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc 10:845–858. https://doi.org/10.1038/nprot.2015.053

Krzywinski M, Schein J, Birol I, et al (2009) Circos: an information aesthetic for comparative genomics. Genome Res 19:1639–1645. https://doi.org/10.1101/gr.092759.109

Mi H, Muruganujan A, Ebert D, et al (2019) PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. Nucleic Acids Res 47:D419–D426. https://doi.org/10.1093/nar/gky1038

Szklarczyk D, Gable AL, Lyon D, et al (2019) STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res 47:D607–D613. https://doi.org/10.1093/nar/gky1131

Zankari E, Hasman H, Cosentino S, et al (2012) Identification of acquired antimicrobial resistance genes. J Antimicrob Chemother 67:2640–2644. https://doi.org/10.1093/jac/dks261