Sujay Chattopadhyay, PhD
Associate Professor
Centre for Health Science and Technology
JIS Institute of Advanced Studies and Research Kolkata
Email: chatsujay@gmail.com
Google scholar link: https://tinyurl.com/y74bnqfb
Peer-reviewed publications link: https://tinyurl.com/yy8un8ed

Dr. Chattopadhyay received his MTech in Biotechnology from Indian Institute of Technology, Kharagpur in 1998, and completed his PhD in Bioinformatics and Computational Biology from the Department of Theoretical Physics at Indian Association for the Cultivation of Science, Kolkata in 2003. Afterwards, he pursued his post-doctoral research, followed by Research Assistant Professorship in the Department of Microbiology, University of Washington, Seattle, USA. His area of research is computational and evolutionary genomics of microbes, with a special interest in microbial virulence and antibiotic resistance.

Since 2016, he has been a consultant of ID Genomics SPC in Seattle USA for detecting high-resolution genotyping markers to aid clonal diagnostics, holding the following patents:

(a) Received US patent no. 9,970,063 (May 15, 2018): Compositions and methods for identifying bacterial clonotypes and detecting antibiotic susceptibility. Inventor: Sujay Chattopadhyay. Applicant: ID Genomics SPC, Seattle, USA;

(b) Filed US patent no. 62/668,042 (May 7, 2018): Methods and tools for determining clonal relatedness and predicting clonal traits. Inventors: Sujay Chattopadhyay, Veronika Tchesnokova, Elena Rechina. Applicant: ID Genomics SPC, Seattle, USA.

On his return to India in late 2017, Dr. Chattopadhyay worked as an Associate Professor at the School of Biotechnology, Amrita Vishwa Vidyapeetham in Kerala. In March 2019, he has joined the Centre for Health Science and Technology at JIS Institute of Advanced Studies & Research Kolkata.

Current research areas:

(A) Adaptive Co-evolution of Genes in Microbial Pathogens.

Similar to genome-wide association studies in humans, given a large sample size, the co-evolved loci or adaptive mutations in microbes (pathogens in particular) can in effect be predicted for their association to specific host-compartments, geographical locations, epidemic/endemic outbreaks, or disease phenotypes in hosts. The present pace of genome sequencing indicates that, by using affordable and rapid sequencing technologies, tens of thousands of microbial genomes will be sequenced during this decade, thereby enabling us perform such association studies in near future. This project plans to develop an analytical tool to detect co-evolution of genes across the genome, to be used to assess phylogenetic congruence for the entire tree (i.e. involving all isolates) or for any sub-tree (i.e. across specific phylogenetic clades) for a given species. This information will provide important insights to create genomic network of adaptive loci functioning within a particular bacterial lineage or across multiple lineages in parallel.

An important extension of within-species co-evolution studies would be to study cross-species interplay of such adaptive forces in a given habitat. His earlier work on Escherichia coli and Salmonella enterica subspecies I core genes demonstrated that there was a significant overlap in the functional trajectories of adaptive evolution in two species. Recent studies showed that specific virulence factors in S. typhimurium stimulate strong host inflammatory response, and eventually help the pathogen gain an advantage in its growth competition with the resident microbiota. Therefore, it would be important to

study the role of co-evolving metabolic pathways in the interactions/competitions of microbiota in host-compartments, e.g. in the inflamed gut. Such a study can offer the possibility to identify new targets for intervention.

Sample relevant publications:

1. A. Thomas, S. Preetha, A. Omanakuttan, L. Vidyullata, A. Ashokan, V. Rajachandran, **S. Chattopadhyay**. 2019. Mutational convergence acts as a major player in adaptive parallel evolution of *Shigella* spp. ***Scientific Reports***, 9: 3252.

2. **S. Chattopadhyay**, P. B. Chi, V. N. Minin, D.E. Berg, E. V. Sokurenko. 2018. Recombination-independent rapid convergent evolution of the gastric pathogen *Helicobacter pylori*. ***BMC Genomics***, 19: 835.

3. D. I. Kisiela, **S. Chattopadhyay**, V. Tchesnokova, S. Paul, S. J. Weissman, I. Medenica, S. Clegg, E. V. Sokuernko. 2013. Evolutionary analysis points to divergent physiological roles of type 1 fimbriae in *Salmonella* and *E. coli*. ***mBio***, 4: e00625-12.

(B) Role of Truncation Mutations in Virulence Evolution.

Occurrence of pseudogene formation via truncation mutation and gene deletion is a common phenomenon in bacterial world, especially in the evolution of the host-adapted/host-restricted bacterial pathogens. A general belief is that pseudogene formation and gene deletion are results of reductive evolution, following a 'use-or-lose' dynamics which suggests purging of traits that are of no use in the organism. Based on the preliminary studies on Salmonella, however, Dr. Chattopadhyay's lab hypothesizes that accumulation of truncation mutations leading to pseudogene formation can often be result of adaptive evolution. We anticipate that such events rather follow a 'die-or-lose' dynamics indicating purging of traits that are otherwise deleterious to the organism.

The goal of this project is to understand the role of gene inactivation via truncation mutations in the (patho)adaptive evolution of bacteria. Events leading to evolutionary convergence are often adaptive and positively selected. Based on the detection of recent non-random convergent events of truncation mutations, the lab proposes a novel approach to distinguish adaptive truncation mutations from reductive ones. The primary focus in this study will be on Salmonella, along with other pathogens, for developing the analytical tool to decipher the adaptive significance of gene truncation mutations leading to the loss of protein function.

Sample relevant publications:

1. S. Paul, A. Bhardwaj, S. K. Bag, E. V. Sokurenko, **S. Chattopadhyay**. 2015. PanCoreGen – profiling, detecting, annotating protein-coding genes in microbial genomes. ***Genomics***, 106:367-372.

2. **S. Chattopadhyay**, S. Paul, D. E. Dykhuizen, E. V. Sokurenko. 2013. Tracking recent adaptive evolution in microbial species using TimeZone. ***Nature Protocols***, 8: 652-665.

3. D. I. Kisiela, **S. Chattopadhyay**, S. J. Libby, V. Tchesnokova, J. J. Kramer, E. Wheeler, R. I. Mackie, S. Clegg, E. V. Sokurenko. 2012. Convergent evolution of invasive serovars of *Salmonella enterica* via point mutations in the type 1 fimbrial adhesin FimH. ***PLoS Pathogens***, 8: e1002733.

(C) Clonal Evolution of Virulence and Antibiotic Resistance.

Most bacterial species, pathogens or commensals, are clonal in nature, represented by the strains with distinct phenotypes circulating as a limited number of genetically related (i.e. clonal) lineages. The stability of such (adapted) clonal lineages has been demonstrated to be strong enough, both temporally

and spatially, to decipher consistent clonal association with important traits like specific virulence potentials or antibiotic resistance profiles.

Multilocus sequence typing (MLST) is presently the method of choice for determining the clonal structure of a bacterial species, and for numerous important pathogens the MLST schemes have been standardized. However, since the STs are discriminated based on the genetic relatedness of a set of housekeeping genes, they are neither uniform nor fine-tuned with respect to the pathotypes and resistance/susceptibility profiles of their representative strains. For example, in E. coli, ST73 includes pathogenic strains like CFT073 that causes pyelonephritis in humans, as well as commensal strains like ABU83972 and Nissle1917 that have been used as probiotics in humans. Also, MLST requires involvement of 7 loci, limiting its efficiency in terms cost and time. This work aims to detect potential candidate genes and mutations therein as high-resolution clonal markers for selected bacterial pathogens to associate specific virulence and/or multidrug resistance properties of interest.

Sample relevant publications:

1. V. Tchesnokova, M. Billig, **S. Chattopadhyay**, E. Linardopoulou, P. Aprikian, P. L. Roberts, V. Skrivankova, B. Johnston, A. Gileva, I. Igusheva, A. Tolland, K. Riddell, P. Rogers, X. Qin, S. Butler-Wu, B. T. Cookson, F. C. Fang, B. Kahl, L. B. Price, S. J. Weissman, A. Limaye, D. Scholes, J. R. Johnson, E. V. Sokurenko. 2013. Potential for predictive diagnostics of *Escherichia coli* infections based on the clonal association of antimicrobial resistance and clinical outcome. ***Journal of Clinical Microbiology***, 51: 2991-2999.

2. S. Paul, S. Million-Weaver, **S. Chattopadhyay**, E. V. Sokurenko, H. Merrikh. 2013. Replication-transcription conflicts increase the rate of evolution in specific genes. ***Nature***, 495: 512-515.

3. S. J. Weissman, **S. Chattopadhyay**, P. Aprikian, M. Obata-Yasuoka, Y. Yarova-Yarovaya, A. Stapleton, W. Ba-Thein, D. Dykhuizen, J. R. Johnson and E. V. Sokurenko. 2006. Clonal analysis reveals high rate of structural mutations in fimbrial adhesins of extraintestinal pathogenic *Escherichia coli*. ***Molecular Microbiology***, 59: 975-988.

(D) Microbial Species-specific Variome Databases.

Variome as a term refers to the sum of genetic variations in a species population. Microbial Variome Databases would stand for species-specific genomic resource databases. The species-specific approach will be aimed at pathogens important in environmental, food and infectious diseases research (primarily in Indian context) such as *Helicobacter pylori*, *Streptococcus pneumoniae*, *Klebsiella pneumoniae*, *Mycobacterium tuberculosis*, *Escherichia coli*/*Shigella* spp., *Salmonella enterica*, *Pseudomonas aeruginosa* and *Staphylococcus aureus*. For each of these species, the plan would be to map all the variations at pan-genomic population level and perform genome-wide analysis to study (patho)adaptive co-evolutionary network of genes/proteins and their association to specific host-compartments, geographical locations, epidemic/endemic outbreaks, or disease phenotypes in hosts. This would eventually allow the shift of bacterial genomics to the level of population genomics, involving:

(a) databases of genetic variations with the predictions of possible functional and adaptive ones for the pathogens (linked to specific disease phenotypes, origins of isolation, etc.);

(b) genome-wide information on potential targets for vaccines, antibiotics and other therapeutics; and

(c) nation-wide as well as global surveillance system that enables rapid determination of newly-emerging or re-emerging pathogenic clones and of the genetic mechanisms behind the emergence.

Sample relevant publications:

1. **S. Chattopadhyay**, F. Taub, S. Paul, E. V. Sokurenko. 2013. Microbial Variome Database: point mutations, adaptive or not, in bacterial core genomes. ***Molecular Biology and Evolution***, 30: 1465-1470.

2. **S. Chattopadhyay**, S. Paul, D. I. Kisiela, E. Linardopoulou, E. V. Sokurenko. 2012. Convergent molecular evolution of genomic cores in *Salmonella enterica* and *Escherichia coli*. ***Journal of Bacteriology***, 194: 5002-5011.

3. **S. Chattopadhyay**, S. J. Weissman, V. N. Minin, T. A. Russo, D. E. Dykhuizen, E. V. Sokurenko. 2009. High frequency of hotspot mutations in core genes of *Escherichia coli* due to short-term positive selection. ***Proceedings of the National Academy of Sciences USA***, 106: 12412-12417.