Hospital acquired infections are major cause of mortality and morbidity in the hospitalized patients, neonates, and immunocompromised patients. *S. maltophilia* and nonmeloid *Burkholderia* like *Burkholderia cepacia* complex are globally emerged important opportunistic pathogens. These groups of were relatively less explored than other prominent non-fermenter opportunistic pathogen like *Acinetobacter baumannii* and *Pseudomonas aeruginosa*. In India *S. maltophilia* and members of Bcc represents most frequent non-fermenter opportunistic pathogen after *A. baumannii* and *P. aeruginosa*. Phylogenetic position of the *S. maltophilia* makes this bacterium as a model to study adaptation of environmental bacterium to human host as its phylogenetic neighbors are environmental bacterium and well known plant pathogens. Revolution in the sequencing technologies transformed the field of Microbiology. Here, I took this opportunity to study phylogeny and evolution of emerging opportunistic pathogens *S. maltophilia* and *Burkholderia cepacia* complex by using high-throughput genomic approaches.

6.1 Taxonogenomics based insights into the phylogeny and evolution of clinical isolates of the S. maltophilia

The genus *Stenotrophomonas* is a taxonomically challenging genus with numerous reclassifications and misclassified species. Thus, to study the phylogeny first whole genome sequencing of all type strains of the species belonging to *Stenotrophomonas* and misclassified species associated with *Stenotrophomonas* was performed. Type strains plays important role in assigning the species status in bacterial and archaea and discovery of the novel species. The analysis by using modern genome based taxonomic methods coupled with type strains revealed later synonyms of *S. maltophilia* and members of Smc (*P. genicualata, P. beteli, S. aftricana*, and *P. hibisicola)* represent the separate distinct species of Smc.

Further, to study the genomic diversity among clinical isolates, I sequenced twenty-seven clinical isolates from hospitalized patients at tertiary care hospital of postgraduate institute of medical education and research, Chandigarh. After integrating the genomes sequences of type strains belonging to genus *Stenotrophomonas* with genomes of clinical isolates and employed modern genome based taxonomic methods were employed. Phylogenomic analysis revealed that clinical isolated are distributed over *Stenotrophomonas maltophilia* complex (Smc) within the genus *Stenotrophomonas*. Further, Genome-based taxonomy coupled with the genomes of type strain of genus *Stenotrophomonas* allowed identification of five cryptic and novel genomospecies associated with clinical isolates of *S. maltophilia*, which are putative novel

species. Pan-genome analysis suggested small core genome size of Smc due to high level of genetic diversity within isolates. In order to understand reason for diversification of members of Smc recombination analysis was performed. It suggested that mutations were dominant over recombination, while the impact of recombination is more than mutation in diversification of members of Smc. As *S. maltophilia* is well known for the resistance genes, we performed the resistome analysis, which suggested that all novel genomospecies harbour well-characterised efflux pumps genes from *S. maltophilia*. Except one novel genomospecies, well-characterised antibiotic resistance genes from *S. maltophilia* were found to be present in all genomospecies. The results of systematic studies using the genomes of type strains of genus *Stenotrophomonas* supports that multiple cryptic novel genomespecies associated with clinical isolates of *S. maltophilia*. Our study also highlights the importance of genomic-based approaches to delineate bacterial species and discovery of putative novel species. The genome sequences of type strains of medically and biotechnologically important genus *Stenotrophomonas* will further help to identify new species of rapidly growing for genus and for comparative genomics studies. This study also gives suggestion to clinician that clinical isolates of *S. maltophilia* are too diverse and not to be considered as single species.

6.2 Transcriptomics basis for understanding of S. maltophilia adaptation to human body temperature

S. maltophilia, an environmental origin opportunistic pathogen is only species in genus *Stenotrophomonas* that cause infections to humans. Interestingly, the non-pathogenic *S. rhizophilia,* closest phylogenetic neighbour of *S. maltophilia* in genus *Stenotrophomonas* do not grow at 37 °C. Previous genome comparison of *S. maltophilia* and *S. rhizophilia* revealed that *S. rhizophilia* lacks heat shock genes, crucial virulence factors and exclusively presence of genes involved in suicidal mechanism at higher temperature, thus *S. rhizophilia* is unable to growth at human body temperature 37°C. During transition from the environment to human body, first major stress posed by environmental bacterium is the physiological temperature of human body (37°C) that is higher than that of environmental niches (22-30°C). In pathogenic bacterium, genes involved in virulence, antibiotic resistance and adaptation to human body are upregulated at 37°C. Therefore, to understand the transcriptional machinery of *S. maltophilia* in response to 37°C, I performed comparative transcriptome analysis of *S.maltophilia* at 28°C as a temperature representative for the environmental niches and 37 °C as human body by using RNA sequencing.

Our results of comparative transcriptome analysis revealed that human body temperature is not heat stress for *S. maltophilia* as genes involved in heat shock response were not upregulated at 37 °C. This suggests that *S.maltophilia* has evolved to thrive at human body temperature. The genes induced at 37 °C include genes for Type IV secretion system, aerotaxis, and cation diffusion facilitator family transporter suggests its potential role in adaptation and virulence of *S. maltophilia*. Interestingly, genes for flagella and fimbria along with other cell motility genes were downregulated. The flagella and fimbria was shown to be a major virulence factor of *S. maltophilia* despite that it was significantly downregulated at 37 °C, suggesting that it is an adaptive mechanism by which *S. maltophilia* avoid the host recognition and subsequent host innate immune response during infection. Rest of downregulated genes at 37 °C includes the genes for cell motility, energy generation and metabolism, lipid metabolism, translation, amino acid metabolism and transport, replication and repair, inorganic ion and transport metabolism lipid metabolism, coenzyme metabolism. The findings enhanced our understanding of the strategies employed by emerging opportunistic pathogen *S. maltophilia* during adaptation towards human body temperature. This will be helpful to discriminate pathogenic and non-pathogenic *Stenotrophomonas,* which will be important for their biotechnological exploitation.

6.3 Genomic characterization of nosocomial outbreak isolates of B.cenocepacia and identification of actively excising novel genomic island

Members of Bcc are responsible for nosocomial outbreaks as they have vey high intrinsically resistance to most of antibiotics and antiseptics as well as their survival ability in nutrition-limited condition for long time. In the present study, I sequenced and analyzed the genomes of seven nosocomial Bcc isolates, five of which were isolated from bloodstream infections and two isolates were recovered from hospital setting surveillance during hospital outbreak. Using state of art genome-based species identification coupled with the genome of type strain, confirmed the species of Bcc isolates as *B. cenocepacia.* SNP analysis of isolates from the suspected source of outbreak i.e. amikacin vial cap and isolates from patient's blood cultures revealed that they were clonal with few notable SNPs among them. SNP analysis also discriminated one non-outbreak isolate among seven sequenced isolates, which is non-clonal with other isolates and that might have been acquired by the patient from some other source. Phylogenetic analysis revealed that phylogenetically distinct clone of *B. cenocepacia* was associated with nosocomial outbreaks. Comparative genomics distinctly revealed the larger genome size of six clonal isolates as well as presence of a novel 107kb genomic island named

as BcenGI15, which encodes putative pathogenicity-associated genes. By using set of PCR experimentations, we demonstrated that BcenGI15 was actively excising from the genome and forming an extrachromosomal circular form suggesting it is conjugationally active element. Interestingly, a homologue of BcenGI15, which is also present in the genome of a clinical isolate named *Burkholderia pseudomallei* strain EY1. This novel genetic element was present only in the variants of *B. cenocepacia* and *B. pseudomallei* isolates suggesting its interspecies existence in the main pathogenic species of the genus *Burkholderia*. Whole genome analysis of the genomically distinct *B. cenocepacia* clinical isolates has advanced our understanding of the epidemiology and evolution of this important nosocomial pathogen as well as its relatives.

6.4 Genomic diversity of the Burkholderia cepacia complex clinical isolates from India

Population structure analysis by using multilocus sequence typing and analysis revealed that high level of genetic diversity and existence of two lineage among the clinical isolates from India. Among two lineages, major lineage belongs to *B. cenocepacia* and minor one to *B. cepacia*. Further, in order to understand and to get insights in to the genomic diversity, confirmation of species status by modern taxonomic criterions, variation by horizontal gene transfer, candidate virulence loci and its resistome, I aimed to carry out whole genome sequencing. Whole genome sequencing of seventeen representative isolates (Blood; n=14, Respiratory specimen; n=3), including three isolates from cystic fibrosis patients was carried out. Taxonogenomics confirmed the species identity of isolates as *B. cenocepacia* (n=12) and *B. cepacia* (n=5) which is as similar by MLSA except one isolate Bcc1236 that is taxonomic outlier of *B. cenocepacia*. Interestingly, *B. cepacia* Bcc8947, which is cabel pilus gene *cblA* positive and four other isolates belongs to *B. cenocepacia*, Bcc1125, Bcc22565, Bcc9500, Bcc7142 harboured *Burkholderia cepacia* epidemic strain marker gene thus belongs to the ET12 lineage. Three isolates of *B. cenocepacia* (Bcc25980, Bcc30380 and Bcc30379) harbours the BcenGI15 pathogenic island identified in nosocomial outbreak isolates of *B. cenocepacia*. The virulence genes belongs to iron uptake, surface polysaccharides (lipopolysaccharides, exopolysaccharides) flagella, proteases, super oxidase, amidases, porin are present in all isolates of Bcc under study. However, the virulence factor belongs to the quorum sensing, cable pilus, O-antigen and pathogenic islands have limited distribution across the sequenced isolates. Resistome includes the efflux pump complex or subunit conferring antibiotic resistance or regulation of efflux, resistance gens for the antibiotic amino glycosidase, Beta-lactam, fluroquinilone, tetracyclins, isonizid, ethioamide, polymyxin, chloramphenicol, timulin,

sufonamide, mupirocin and disinfectant triclosan. Genomes of the sequenced isolated were compared to the *B. cenocepacia* J2315, which revels that several genomic islands namely BcenGI2, BcenGI3, BcenGI9, BcenGI10, BcenGI13 were absent whereas BcenGI5, BcenGI6, BcenGI7, BcenGI8, BcenGI11, BcenGI12, BcenGI14 having limited distribution among Bcc isolates under study. Overall, this study of genome-based characterization of clinical Bcc isolates from India having important implication in treatment, global epidemiology and management of this important group of pathogen.

Overall, my work using genomics approaches like whole genome and transcriptome sequencing provided major insights in to the phylogeny and evolution of emerging opportunistic pathogens like *S. maltophilia* and members of *Burkholderia cepacia complex*. Further, my studies also supports the use of *S. maltophilia* as model to study the adaptation of environmental origin opportunistic pathogens to human body.