The introduction of NGS technologies and the resulting genome assembly at economical cost have transformed microbiological research radically. Here we have utilized high-throughput sequencing technologies to unravel the genome level diversity of order *Myxococcales* members (second chapter) followed by comparative *in silico* analysis to understand their social behaviors (third chapter).

The second chapter discusses the culturing, genome assembly and taxonomy of diverse myxobacteria. We suggested some probable reasons for genome expansion as compared to *Deltaproteobacteria* and also discussed the pan and core genome for available myxobacterial genomes. In the first objective of this study, we have procured several myxobacterial strains from DSMZ, and we have cultured them in our lab and isolated good quality genomic DNA as required for respective NGS technologies. This DNA was used to obtain NGS data which was further used to assemble them into the draft or complete genomes. We have assembled nine myxobacterial genomes which include five complete genomes and four high-quality draft assemblies. These genomes have provided diversity to the currently available myxobacterial genome datasets. In the second objective, we studied the differences using different annotation programs and found RAST to be a reasonable annotation platform to use for our studies (Sharma et al., 2016b). All twenty-five myxobacterial genomes under investigation were annotated using RAST server at the same time. The third objective was the phylogeny and taxonomic studies of sequenced myxobacterial genomes. Based on 16S rRNA, concatenated housekeeping genes and genome-genome distances all sequenced genomes were confirmed belonging to their previously classified taxonomy (Garcia et al., 2010; Goldman et al., 2006; Sharma et al., 2016a; Sharma et al., 2016b). For *Myxococcus hansupus* which was found as a contaminant on a plate of *Chondromyces apiculatus* obtained from DSMZ, we confirmed this strain as a novel species in genus *Myxococcus* belonging to family *Myxococcaceae* (Sharma et al., 2016b). *Enhygromyxa salina* was also suggested to be a close relative of family *Nannocystaceae members such as Plesiocystis pacifica* (*Pp*) and *Nannocystis* species. The fourth objective of this study explored the possible reasons for genome expansion within myxobacteria. We found many Pfam families and clans to be overrepresented within myxobacterial genomes as compared to the non-Myxococcales *Deltaproteobacteria* genomes suggesting the gene duplication as a primary reason for genome expansion. Protein sequence-based clustering is also in agreement with this conclusion. These large numbers of

duplicated proteins probably help myxobacteria to adapt to its diverse habitats and accomplishing its complex life cycle (Bratlie et al., 2010; Kondrashov, 2012). The focus of the fifth objective was to identify the pan, core, dispensable and unique genomes. This study helped to understand the extent of diversity within the order *Myxococcales* members. The pan-genome for 25 myxobacterial genomes is open which suggests that addition of more diverse genomes will allow it to increase further. The core genome of 257 genes is mostly involved in housekeeping gene functions. The presence of a large number genes in the dispensable and unique genomes reveal the strain-specific diversity within these myxobacterial genomes.

We performed *in silico* studies to identify the proteins involved in social behavior activities such as lignocellulose degradation, gliding motility, sporulation, and signal transduction and compared them among all myxobacteria. The first objective of this study was to identify the carbohydrate active enzyme repertoire and their evolutionary relationships. Myxobacteria have evolved within class *Deltaproteobacteria* which are not known as lignocellulose degraders. The evolution of proteins involved in starch, cellulose, agar, and chitin degradation is an important feature within myxobacteria to thrive in soil, bark, dung, and other lignocellulose-rich places. Homologs of almost all proteins involved in the degradation of these substrates are rarely present in *Deltaproteobacteria*. Our study revealed that these proteins had been acquired from diverse bacterial phyla such as *Actinobacteria*, *Alphaproteobacteria*, *Firmicutes*, *Betaproteobacteria* and others via horizontal gene transfer events during evolution. The presence of multiple proteins involved in the degradation of these substrates as compared to *Deltaproteobacteria* also suggests another reason for genome expansion within myxobacteria apart from gene duplication. We have also used *in silico* methods to compare the proteins involved in gliding motility and fruiting body formation. Till now the proteins involved in these functions have been studied only in *M. xanthus* DK1622. Based on comparative genomic studies across all myxobacteria, 15 and 40 proteins out of 50 A-motility and 109 S-motility proteins respectively were found conserved. The most important proteins MglA and MglB, known to communicate between both motilities are well conserved across all myxobacterial genomes. In the case of fruiting body formation, we found the conservation of about 40 proteins out of the total 95 proteins as reported previously (Huntley et al., 2011a). The NFS cluster, known to be essential in *M. xanthus* development (Muller et al., 2010), was found conserved only

across family *Myxococcaceae* and *Cystobacteraceae* members. The absence of the majority of proteins in suborder *Nannocystineae* and *Sorangiineae* suggest that there may be other genus/family/suborder specific proteins involved in gliding motility and sporulation. We also explored if there is any link between genome size and the number of myxobacterial signal transduction proteins, which are known to coordinate behaviors such as nutrition sensing, motility, sporulation, secondary metabolite biosynthesis, prey hunting etc. (Capra and Laub, 2012; Szurmant and Ordal, 2004; Whitworth and Cock, 2008, 2009). We identified the TCS repertoire for 25 myxobacterial genomes and showed that there is no strong correlation between genome size and TCS proteins unlike that reported for other bacteria (Borland et al., 2015b; Galperin, 2005; Ulrich et al., 2005). TCS protein count has also been known to be related to organism complexity (Galperin et al., 2010; Galperin and Koonin, 2010). Similarly, we found no strong correlation between genome size and chemosensory systems, but the expansion in TCS and CSS in myxobacteria is likely governed by the habitat and environmental stimuli. Based on CheA phylogeny and modular architecture this study identified six novel ECSS apart from the previously known Che1 to Che8. Che5 and Che8 systems might be regulating their adjacent genes involved in ATP-dependent protein degradation and riboflavin synthesis as determined by the conserved synteny across diverse myxobacteria. Phylogenetic studies revealed monophyletic nature of Che1-Che8 systems to *Deltaproteobacteria* whereas the ECSS were more diverged as they were sharing clades with *Acidobacteria*, *Betaproteobacteria*, *Gammaproteobacteria*, cyanobacteria, and *Firmicutes*. CheB based phylogenetic analysis revealed four distinct clusters with the unique modular organization, classified as ACSS.