7. Summary

The eventual objective of all biochemical and molecular studies related to bacterial degradation of xenobiotic chemical pollutant is to develop a viable bioremediation process, which could be used for decontamination/remediation under natural environments. With this aim, the present dissertation work was designed to deals with the whole genome sequencing assembly and annotation of *Burkholderia* sp. strain SJ98. From the genome annotation and manual curation, various putative ORFs have been identified and characterized. These ORFs were predicted to be involved in the degradation of *p*-nitrophenol and other nitroaromatic compounds.

Whole genome sequencing of *Burkholderia* sp. strain SJ98 was performed on two independent sequencing platforms Roche's- 454 GS FLX titanium and Illumina Genome analyzer GA IIX. Roche's – 454 short reads assembled into larger contigs by using the Newbler v2.5.03 software, total 79 contigs and 17 scaffolds of size 7.89-Mb with the 1,315,287 of N's. Illumina's sequencing short reads were assembled into 132 scaffolds by SOAPdenovo software with total size of 7.48-Mb and 12527 of N's. Further, gaps (N's) removed from the SOAPdenovo assembled data by Gapcloser v1.10. Seventeen scaffolds obtained from the Roche's – 454 sequencing and Newbler assembly were selected for the finishing and removing the gaps using the Gapcloser using the short reads and contigs from the Illumina sequencing and SOAP*denovo* assembly respectively. Attempts were made to remove the gaps by PCR sequencing using the primer walking. However, only one gap was filled between the contigs 5 and 15. Finally, after filling the gaps 14 contigs (7.87-Mb without N's) were obtained successfully from the sequencing data obtained from two platforms.

Genome annotation of the strain SJ98 was performed by Rapid Annotation Server Using Subsystem Technologies (RAST, server) and NCBI Prokaryotic Genome Annotation Pipeline (PGAAP) and RNAmmer v1.0. RAST server has predicted 7414 coding regions (CDSs) including 52 tRNAs & 3 rRNAs genes. PGAAP pipeline has predicted 7268 CDSs, 52 tRNA and 3 rRNA genes. Gene ontology terms were identified by using Uniprot database mapping. Phylogenomics results suggested that a recently sequenced genome of *Burkholderia* sp. strain YI23 is the closest neighbor of strain SJ98.

Various genes involved in the degradation/transformation of nitroaromatic and aromatic compounds were identified and some of them were selected for further functional characterization.

Complete gene cluster of *p*-nitrophenol (PNP) degradation was identified from the genome annotation data of strain SJ98. PNP degradation genes in the genome of strain SJ98 are present in two distinctly situated gene clusters. One gene cluster contains- PNP 4-monooxygenase (pnpA1) and *p*-benzoquinone reductase (pnpB), while hydroquinone dioxygenase (*pnpE1* and *pnpE2*), 4-hydroxumuconic semialdehyde dehydrogenase (pnpF), 1,2,4-benzenetriol dioxygenase (pnpC) and maleylacetate reductase (pnpD) constitutes the second gene cluster. Additionally, one homolog of *pnpA1* was found in the contig 12 and termed as *pnpA2*. These results suggested a different type of evolution of PNP degradation genes in train SJ98. Strain SJ98 can be further used as a model system to study the evolution of PNP degradation genes.

Putative ORFs involved in the degradation of *p*-nitrophenol were selected for cloning, expression and purification. PNP 4-monooxygenase (*pnpA1* and *pnpA2*), *p*-benzoquinone reductase (*pnpB*) and hydroquinone reductase (two subunits- *pnpE1* and *pnpE2*) were functionally characterized. His-6-PnpA1 was found as a functional *p*-nitrophenol 4-monooxygenase enzyme, which transforms *p*-nitrophenol into *p*-benzoquinone, while His-6-PnpA2 was inactive in *in-vitro* assays. Kinetic parameters were determined for the purified His-6-PnpA1 and found in accordance with the previously characterized PNP-4-monooxygenases. PnpA1 was also found to transform other nitroaromatic compounds (i.e., 2C4NP, 4NC, 2,4-DNP, 3M4NP, 2,6-DCP, ONP and PCP) suggesting its broad substrate specificity. The molecular mechanisms behind this broad range specificity could be elucidated by structural and functional studies of PnpA1.

In-vitro assays of His-6-PnpB confirmed its identity as *p*-benzoquinone reductase, which transforms *p*-benzoquinone to hydroquinone with the utilization of NADPH as electron donor. Biochemical and kinetic parameters were identified for the purified His-6-PnpB enzyme. This enzyme performs the second step of the *p*-nitrophenol degradation pathway. Two subunits of putative hydroquinone dioxygenase gene (*i.e., pnpE1 and pnpE2*) cloned, expressed and purified from the *E. coli* BL21.AI. His-6-PnpE1 was purified under the denaturing condition, refolded by the dilution method. His-6-PnpE2 was purified under the native condition. Further the enzyme assay and kinetic properties

of hydroquinone dioxygenase enzyme was assessed by mixing the two subunits His-6-PnpE1 and His-6-PnpE2. Purified His-6-PnpE1 and His-6-PnpE2 together form a functional hydroquinone dioxygenase (HqD) enzyme. Molar stoichiometry of functional HqD was found as $PnpE1_2PnpE2_2$. RT-PCR has supported the result as *pnpE1* and *pnpE2* subunits are selectively induced during the growth of strain SJ98 on *p*-nitrophenol as the sole carbon, nitrogen and energy source. Past reports on studies for *p*-nitrophenol degradation in strain SJ98 was found to proceed via the formation of 4-NC and BT. Together with the earlier reports and present study, we have found that the genes for both the pathways for PNP degradation are functional in the strain SJ98.

Gene *xenobiotic reductase A* (*xenA*), *catechol 1,2-dioxygenase* (*C12D*), *catechol 2,3-dioxygenase* (*C23D*) , *phenol 2-monooxygenase* (*p2m*) and *pentachlorophenol 4-monooxygenase* (*pcpm*) were selected for the functional characterization. Gene *xenA* cloned, expressed and protein was purified as His-6-XenA. *In-vitro* assay of His-6-XenA has ascertained its identity as functional reductase transforming the Coumarin into 4-hydrocoumarin. Further, the kinetic parameters were determined. Substrate spasticity of the His-6-XenA showed significant enzymatic activity on 2,5-dinitrophenol, 2,6-dinitrophenol, however it showed relatively weaker activity on compounds like pentachlorophenol, *p*-nitrophenol and 2-chloro 4-nitrophenol.

Catechol is the intermediate compound which is formed during the degradation of various aromatic compounds. The ring cleavage of catechol is directed by the enzyme catechol 1,2-dioxygenase (C12D) or catechol 2,3-dioxygenase (C23D) . Two ORFs encoding for the C12D and C23D were cloned, expressed and purified. His-6-C12D was inactive on catechol or other catechol isomers, whereas His-6-C23D was found to be an active dioxygenase enzyme that transforms catechol into 2-hydroxymuconic semialdehyde. Furthermore, C23D was found to be active on other compounds i.e., 4-chlorocatechol and homoprotocatechuate.

Present work of whole genome sequencing, assembly and annotation of *Burkholderia* sp. strain SJ98 has revealed the genes and gene clusters involved in the degradation of various nitroaromatic and aromatic compounds. Furthermore, *in-vivo* genetic aspects could also be explored to study the molecular mechanisms involved in the degradation of nitroaromatic compounds such as study of molecular mechanism for chemotaxis towards nitroaromatic compounds by strain SJ98.