Abstract:

The *Alphaproteobacteria* represent a biologically diverse group of bacteria with members like *Brucella, Bartonella, Agrobacterium* and *Ochrobactrum* that are capable of interacting with eukaryotic cells. The genus *Ochrobactrum* is a Gram-negative, capsulating, aerobic bacilli having closest phylogenticgenetic relative ogenus *Brucella* as evidenced by protein profiling, western blot, immunoelectrophoresis, amplified fragment length polymorphism, 16S rRNA gene and *RecA* gene sequence based studies. In earlier study, several bacteria other than *Helicobacter pylori* have been detected earlier in gastric biopsies in individuls diagnosed with non-ulcer dyspepsia. A unique observation was the presence of severe fibrosis in the lamina propria of the gastric mucosa revealed during histological examination of the gastric antral biopsy. Whether this fibrosis was caused either partially or totally by *Ochrobactrum* was not clear. *O. anthropi* have also been associated along with *H. pylori* with mild gastritis in squirrel monkeys. Importantly, both *H. pylori* and *Ochrobactrum* produce urease, and thus the detection of *H. pylori* by urease test in the presence of *Ochrobactrum* may be confounded. The role of *Ochrobactrum* in gastric pathology remains uncertain and requires detailed pathologic, microbiological and genetic investigations in order to evaluate the link between *H. pylori* and *Ochrobactrum* in the gastric niche. This thesis attempts to determine the population structure of *Ochrobactrum* sp. By development of new multilocus sequence analysis scheme and to gain whole-genome based insights into the putative gene determinants of *Ochrobactrum* for survival in the highly acidic stomach lumen environment.

In all, 75 *Ochrobactrum* spp. strains were recovered from patients suffering with gastroduodenal diseases. These isolates were identified by biochemical characteristics and 16S rRNA gene sequencing and phylogenetic affiliation studies indicated that they indeed belong to genus *Ochrobactrum*. The new multilocus sequence analysis method comprised of four different gene loci

from each chromosome was developed for and applied to characterize the genomic relatedness of *Ochrobactrum* sp. All sequences were aligned by Clustal W software. Phylogenetic analysis was carried out in Molecular Evolutionary Genetic analysis (MEGA) version 5. The results reveals formation of two large groups, the first group included *O. intermedium* specific clade and the second group included strains related *to O. anthropi*. The multilocus sequence analysis involving seven genes (*dnaK*, *recA*, *rpoB*, *trpE, aroC,* and *gap*) was conducted to determine the population structure of Ochrobactrum sp. of clinical and environmental origin.

The draft genome sequence of strain M86 and strain 229E was determined by Ion Torrent Personal Genome Machine (PGM™) sequencer using a 316 chip with 200-bp single-end shotgun sequencing. A total of 2,602,696 reads were obtained. PGM sequencing resulted in about 67X genome coverage with 148 contigs. The *de novo* approach was applied to finalize the unclosed draft genome using MIRA 3.4.0 version using default parameters. Prediction and annotation of genes were done using RAST server with SEED database and ISGA pipeline. The data were further validated using gene prediction tools such as Glimmer. Functional annotation was also performed by PGAAP using public database of National Centre for Biotechnology Information (NCBI).The *Ochrobactrum intermedium* strain M86 whole genome shotgun (WGS) project was submitted to the GenBank and has the project accession AOGE00000000 and consists of sequences AOGE01000001-AOGE01000148.).The *Ochrobactrum intermedium* strain 229E whole genome shotgun (WGS) project was submitted to the GenBank and has the project accession ASXJ00000000.

Genome of *O. intermedium* strain M86 was sequenced on the IonTorrent Personal Genome Machine (PGM™) using 316 chip that resulted in 2,602,696 total reads with a mean read length of 155 bp. *de-novo* assembly using the MIRA assembler v3.4.0 with default parameters yielded ~67X coverage. A total of 148 contigs with >500 bp length were obtained. The unclosed draft genome sequence of strain M86 is of 5,188,688 bps and 5043

predicted coding DNA sequences (CDSs) and 66 RNA genes with mean G + C content of 57.9 %. RAST server based annotation of the whole genome, showed the presence of 437 subsystems.Genome of *O. intermedium* strain 229E was sequenced on the IonTorrent Personal Genome Machine (PGM™) using 316 chip that resulted 7,50,279 total reads with a mean read length of 185 bp. *de-novo* assembly using the MIRA assembler v3.4.0 with default parameters yielded ~18X coverage. A total of 128 contigs with >500 bp length were obtained. The unclosed draft genome sequence of strain M86 is of 4,808,223 bps and 5641 predicted coding DNA sequences (CDSs) and 66 RNA genes with mean G + C content of 57.7 %. RAST server based annotation of the whole genome, showed the presence of 437 subsystems.

H. pylori have several genes for biosynthesis of cytosolic urease for its survival in the acidic environment of stomach lumen. Genome of strain M86 and strain 229E contains urease gene cluster out of which, *UreA*, beta subunit, *UreB* gamma subunit, *UreC*, alpha subunit, are part of core Urease enzyme, While four accessory proteins: *UreD*, *UreE*, *UreF* and *UreG* play important role in Ni^{2+} uptake and insertion into active site of apo-enzyme. A complete operon encoding the *VirB* gene involved in conjugative transfer is present in strain M86 and strain 229E. Genes encoding osmotic stress, oxidative stress *HPIIb*, cold shock *GrpE*, heat shock *DnaK*, periplasmic stress *DegQ* and protection from reactive oxygen species, *sod* are found. Genes predicated to encode flagellar biosynthesis protein *FlhA and FlhB* has been identified in genome of strain M86 and strain 229E which are likely elementary to adaptation of new lifestyle. Enterobactin synthesis clusters of *entA, entB1, entB2, entC, entD, entE, entF, entG, entH* genes were also observed in the genome sequence of strain M86 suggesting its ability of iron acquisition by siderophore production. Presence of membrane transport machinery with dominance of Dipeptide-binding ABC transporter, periplasmic substrate-binding component was detected in the genome of strain M86.It has been found that all the clinical isolates and the type strains of *Ochrobactrum* were highly resistant to all forms of ß-lactams except imipenem. This resistance profile is consistent with the expression of the *AmpC* beta-lactamase

characterized in *O. anthropi*. The genome of strain M86 and strain 229E shows the presence of *AmpC* beta-lactamase gene which supports its resistance to ß-lactams antibiotic observed earlier. Prophages and putative phage like elements in the genome were identified using prophage-predicting PHAST Web server. Regions identified algorithmically as “intact” by PHAST, as well as regions sharing a high degree of sequence similarity and conserved synteny with predicted “intact” prophages, were identified as prophages A study involving genome sequencing and comparison of the strains isolated from non-ulcer dyspeptic individuals helps to understand the genomic features of its survival in the acidic condition of the stomach.

The genomic properties of *Ochrobactrum* are poorly characterized and, as a consequence, their role in human health and disease remains unclear. Elucidation of the physiological properties and identification of genes putatively involved in the various metabolic pathways may lead to a better understanding of the survival of *Ochrobactrum* in acidic condition. Further studies involving large scale genome sequencing and comparison of the *Ochrobactrum* strains isolated from several other non-ulcer dyspeptic individuals will help us apprehend the genomic features of its survival in the acidic condition of the stomach.