Genome Sequence Analysis and Identification of Genes Associated to Pesticide Degradation from *Enterobacter cloacae* Strain MR2

M.V. Parakhia[*], R.S. Tomar, H. Dalal, V.V. Kothari, V.M. Rathod and B.A. Golakiya

Department of Biotechnology, Junagadh Agricultural University, Junagdah-362001, Gujarat, India

*Corresponding author

A B S T R A C T

Today's burning problem in the world is pesticide residues in foods. To overcome this problem, nineteen chlorpyrifos-degrading bacteria were isolated from soil of adjoining area of pesticide manufacturing industries located in Gujarat, India. The strain CPD-12 (MR2) degraded highest chlorpyrifos among the other strains isolated from different sites, i.e. Up to 500 ppm in 30 hrs. And hence was selected further for whole genome sequencing. This strain showed maximum similarity to members of the order *Enterobacteriales* and was closest to *Enterobacter cloacae* of this group. The genome sequence of strain *Enterobacter cloacae* MR2 consisted of a circular 4,758,062bp chromosome with a 55.1% G +C value, 5571 protein coding genes, 16rRNA and 72 tRNAs. The genome annotation and functional characterization of the strain MR2 provided insights into various genetic processes involved in the degradation of several pesticides and detoxification of toxic compounds. The genome of MR2 was also compared with *Enterobacter cloacae* subsp *cloacae* ATCC 13047 and *Enterobacter* sp. 638 which showed the presence of genes for the pesticide degradation as in ATCC 13047 and also had genes to promote plant growth as in *Enterobacter* sp. 638.

Introduction

Organophosphates (OPs) pesticides are highly toxic chemical pesticide that exhibit broad-spectrum activity against insects and accounts for about 38% of the total pesticides used globally for agricultural crops. Continuous and excessive use of OPs has caused not only nerve (this class of pesticide has acute neurotoxicity due to their ability to suppress acetyl- choline-esterase) and muscular diseases in human and animals but also have

contaminated ecosystems in different parts of the world (Zhang *et al.,* 2008).

Chlorpyrifos (O, O-diethyl O-3, 5, 6-trichloropyridin-2-yl phosphorothioate) as an active ingredients a broad spectrum organophosphorus insecticide, most widely used for pest control (Cho *et al.,* 2002). It has been widely used for aerial application to control surface feeding insects (Dhawan and Simwat, 1996; Gupta *et al.,* 2001; Sasikala *et al.,* 2012) and also applied to soil for root

damaging insect larvae (Bhatnagar and Gupta, 1992; Rouchaud *et al.,* 1991; Davis *et al.,* 1976). Pesticides and their degradation products generally get accumulated in the soil and influence not only the population of various groups of microbes, but also their biochemical activities like nitrification, ammonification, decomposition of organic matter and nitrogen fixation (Agnihotri *et al.,* 1981; Faldu *et al.,* 2014). In soil, chlorpyrifos may remain biologically active for periods ranging from days to months. Dosage rates, soil type, soil moisture and organic matter, content, temperature and insecticide formulation are among the factors which influence the biological persistence (Read 1976; Tashiro *et al.,* 1978) it is moderately persistent in nature as its residues were detected in soil even after 3 months of application and hence causes potential environmental hazards (Chapman *et al.,* 1984).

Microorganisms play an important role in degrading synthetic chemicals in soil (Alexander, 1981). They have the broad capacity to utilize almost all natural and some synthetic compounds as their sole carbon and energy source. Chlorpyrifos degrading bacteria can be used either directly or indirectly, for the bioremediation of chlorpyrifos contaminated soils. Till now, various genes, such as *opd* (organophosphate-degrading) and *mpd* (methyl parathion degrading) and several enzyme systems have been identified which were found to be involved in degradation of certain organophosphates (Serdar, 1982; Mulbry *et al.,* 1986; Horne *et al.,* 2002; Yang *et al.,* 2006; Cui *et al.,* 2001; Parakhia *et al.,* 2014).

In the present study, chlorpyrifos degrading bacteria were isolated from various pesticide contaminated sites and were screened for their chlorpyrifos degradation capability through High Performance Liquid Chromatography (HPLC). The most efficient bacterium *Enterobacter cloacae* stain MR2 was sequenced for complete genome. The genome of *E. cloacae* MR2 was characterized for identification of genes responsible for the degradation of chlorpyrifos and was also compared with *Enterobacter cloacae* subsp *cloacae* ATCC 13047 and *Enterobacter* sp. 638 for synteny.

Materials and Methods

Isolation and screening of chlorpyrifos degraders

Soil samples were collected from five different sites which were contaminated regularly with the pesticides from Gujarat, India (Table 1). Out of 45 strains initially isolated, 19 were screened out with the ability to degrade 50-500 ppm chlorpyrifos by Shake flask method and were quantified by HPLC among 19 strains, CPD-12 was found to be most efficient degrader with the ability to degrade 500 ppm within 30 hrs was selected for the genome sequence analysis.

Genome sequencing

For genome sequencing, DNA of *Enterobacter cloacae* stain MR2 was isolated using Phenol-Chloroform method (Sambrook *et al.,* 1989). The DNA concentration and purity was determined using Picodrop PET01 (Picodrop Ltd., Cambridge, U.K).

The DNA was enzymatically fragmented to construct a library of 260 bp, which was further used for template preparation. Sequencing was carried out using Ion Torrent Personal Genome Machine (PGM[TM]) from Life Technologies, at Department of Biotechnology, Junagadh Agricultural University, Junagadh, India as per the manufacture's guidelines.

Gene prediction and annotation

Raw reads of the sequence were processed for the quality control through default plug-in in Ion Torrent Software Server (FastQC). The quality reads were assembled in MIRA v 3.4.1 by using Smith-Waterman algorithm (Chevreux *et al.,* 200). Contings were ordered through the tool Mummer (Kurtz *et al.,* 2004) and were aligned with reference genome *E. cloacae* ATCC 13047 and *Enterobacter* sp. 638 using Mauva (Darling *et al.,* 2010) software. Putative coding sequences (CDS) were initially identified by RAST automated annotation software (Aziz *et al.,* 2008; Overbeek *et al.,* 2014) followed by Magnifying Genome annotation platform (MaGe) (http://www.genoscope.cns.fr/agc/mage/). All CDS identified were manually reviewed, and false CDS were flagged as ''artifact''. The remaining CDS were then submitted to automatic functional annotation *via* BLAST searches against the UniProt databank in order to determine significant homology. Circular chromosomal map of *E. cloacae* MR2 with annotated genes/CDS was constructed using CGView (Stein *et al.,* 2001). Core and Pan Genome analysis of *E. cloacae* MR2 with *E. cloacae* ATCC 13047 and *Enterobacter* sp. 638 was analyzed by MaGe-Microscope Pan Genome Analysis interface (Vallenet et at., 2006).

Results and Discussion

Characterization of bacterial strain

Soil samples collected from five chlorpyrifos contaminated sites of Gujarat, India resulted in the isolation of 45 stains. Out of 45, 19 strains were able to degrade chlorpyrifos (Table 2) and among them, CPD-12 was found to be most efficient degrade with the ability to degrade 500 ppm within 30 hrs.

Genome *de Novo* assembly

Whole genome sequencing of *Enterobacter cloacae* MR2 was carried out using Ion Torrent (PGM) whole genome sequencer (Life Technologies) at the Department of Biotechnology, JAU, Junagadh. A total of 549,959 reads with an average length of 176 bp and have coverage of the 18.06X.Initial quality check of raw data was performed through FASTQC and reads were filtered based on base quality and length (Fig. 1). Quality reads were assembled by MIRA which resulted in 230contigswith longest contig of 177,145bp and N95 of 7,065 bp (Table 3-4).

Assembled genome was submitted to an automated annotation tool RAST (Rapid Annotation using Subsystem Technology), which provides high quality genome annotations for bacterial and archaea genomes. RAST indicated *E. cloacae* subsp cloacae ATCC 13047 (score 500) and *Enterobacter* sp. 638 (score 452) as the closest members of *E. cloacae* MR2. The 'neighbor' score in RAST was estimated *via*" quick and dirty" ad hoc heuristic method which is based on the number of times that the 'neighbor' genome was the top hit in BLAST against the candidate (in this case *E. cloacae*MR2) from the set of "unique" genes within the query genome. A higher score suggested that the two genomes are likely to be metabolically similar. The comparative profile of the MR2 RAST distribution (Fig. 2a) covers 4068 subsystems compare to 3909 of ATCC 13047 and 3564 of strain 638. MR2 codes highest for the metabolism of carbohydrates and amino acids followed by membrane transport mechanism and production of cofactor, vitamins, prosthetic groups, pigments. Complete genome of *Enterobacter cloacae* subsp *cloacae* ATCC 13047 (ref: NC_014121.1) and *Enterobacter* sp. 638 genome from NCBI (ref:

NC_009436.1) were used for comparative and synteny analysis (Fig. 3). In the synteny map the pattern of the arrangement of the blocks indicated variation with reference genome. The space between two blocks indicates the gap region, which was not found similar in the referred genome, may be probable horizontally transferred regions.

Genome annotation

Genome annotation is the process of attaching biological information to sequences. It includes predicting genes function, structure, coding regions and ORFs. Genome annotation of *E. cloacae* MR2 predicted various genes involved in various stress response(s) as well as genes involved in resistance to antibiotics and toxic compounds as indicated by RAST analysis (Fig. 5).

Total 64 genes that are involved in multi-drug resistance, multiple antibiotic resistance, resistance to Fluoro-quinolones, copper homeostasis and tolerate to heavy metals like Copper, Cobalt, Arsenic, Zinc, Cadmium etc. Were identified. Above this, annotation also revealed 32 stress responsive genes that coded for universal stress response protein family, phage shock proteins and genes involved in various processes like carbon starvation, sugar phosphate stress regulation etc. Along with these, genes that take part in other stress responses were also identified which included 59 genes for oxidative stress, 9 for desiccation stress and 24 for osmotic stress / Osmo-regulation. Thirty one gene that is responsible for detoxification like Nudix proteins (nucleoside triphosphate hydrolases) which are activated in plant defense response, a family of versatile, widely distributed housekeeping enzymes, housekeeping nucleoside triphosphate pyrophosphatases, genes involved in tellurite resistance and chromosomal determinants etc. Were identified. The annotation also indicated the presence of genes for phosphorus metabolism,

sulfur metabolism, metabolism of aromatic compounds, nitrogen metabolism, protein metabolism, potassium metabolism and iron acquisition and metabolism in the genome of strain MR2.

Enzymes responsible for catabolism of organophosphate compound such as Inorganic Pyrophosphatase (EC 3.6.1.1), Phosphonoacetaldehyde hydrolase (EC 3.11.1.1), 3-ketoacyl-CoA thiolase (EC 2.3.1.16), Salicylate hydroxylase (EC 1.14.13.1), Catechol 1,2-dioxygenase (EC 1.13.11.1), 1H-3-hydroxy-4-oxoquinaldine 2,4-dioxygenase, Catechol 2,3-dioxygenase (EC 1.13.11.2), Gentisate 1,2-dioxygenase (EC 1.13.11.4) and Monoamine oxidase (EC 1.4.3.4) were also identified during the process of annotation. Metabolic Reconstruction of *Enterobacter cloacae* MR2 and *Enterobacter* sp. 638 allowed the comparison of functioning parts of two organisms (Table.5; Fig.2b).

It provided a list of all genes which were associated with a subsystem in the respective organism. Genes for stress responsive (142), phosphorus metabolism (41) and sulphur metabolism (54) were found to be common in both genomes. While comparing with *Enterobacter cloacae* subsp *cloacae* ATCC 13047, MR2 cluster of orthologous genes (COG) categories indicated highest distribution for the general function prediction (718), Amino acid transport and metabolism (657), Carbohydrate transport and metabolism (596) while ATCC 13047 revealed 626,500 and 463 CDS respectively (Table 3b).

The BLAST map of *Enterobacter cloacae* MR2 with the other genomes present in the microscope platform software indicated *E. cloacae* ATCC 13047 as the nearest genome (Fig. 4), followed by *E. hormaechei* ATCC 49162, *E. concerogenus* ATCC 35316, *Enterobacter* sp. 638 and *E. aerogenes* KTCC 2190.

Table.1 Locations of screening of chlorpyrifos degrading bacteria

Sr. No.	Name of CPD isolate	Industry	GPS Coordinates	Area	District
1	CPD-1	Near Pioneer Agro Industry	Latitude : 23.070887 | Longitude 72.671289	Ahmedabad G.I.D.C.	**Ahmedabad**
2	CPD-2				
3	CPD-3				
4	CPD-4				
5	CPD-5	United Phosphorus Limited	Latitude : 21.618039 | Longitude 73.022817	Ankleshwar G.I.D.C.	**Bharuch**
6	CPD-6				
7	CPD-7				
8	CPD-8				
9	CPD-9				
10	CPD-10				
12	CPD-12				
13	CPD-13				
14	CPD-14	Near GIDC, Kadi, Gujarat, India	Latitude : 23.29042 | Longitude : 72.36219	Kalol G.I.D.C.	**Ahmedabad**
15	CPD-15				
16	CPD-16				
17	CPD-17				
18	CPD-18	Field Collection Ivnagar	Latitude : 21.477184 | Longitude : 70.43203	Field Collection Ivnagar &Vadla	**Junagadh**
19	**CPD-19**	**Field Collection Vadla**	**Latitude : 21.477991 | Longitude : 70.40041**		

Table.2 Concentration of standard chlorpyrifos at different incubation period

Sr. No.	Name of CPD isolates	Concentration of standard chlorpyrifos at different incubation period (mg/l)			
		6 hr	**12 hr**	**18 hr**	**24 hr**
1	CPD-1	248	195.47	139.86	90
2	CPD-2	264.18	223.99	188.99	151.43
3	CPD-3	238.5	189.04	149.46	124.15
4	CPD-4	241.72	186.07	102	66.56
5	CPD-5	261.5	217.94	187.09	145.24
6	CPD-6	254.29	214.33	159.55	123.68
7	CPD-7	252.75	203.6	154.04	110.79
8	CPD-8	262.1	229.2	173.25	131.1
9	CPD-9	246.13	193.02	162.92	102
10	CPD-10	245.24	191.57	110.25	75.64
11	CPD-11	243.15	179.05	130.94	87.71
12	**CPD-12**	**234.17**	**163.66**	**80.26**	**22**
13	CPD-13	254.37	206.96	145.15	110.33
14	CPD-14	261.96	215.17	186.33	155.11
15	CPD-15	240.13	179.02	82.35	58
16	CPD-16	253.02	203.31	163.26	115.27
17	CPD-17	269.02	232.03	210.24	164.72
18	CPD-18	236.04	171.04	172.26	49.4
19	CPD-19	247.91	192.02	161.13	104.4
20	Control	300	298.29	297.24	296.17

Table.3a Genome information of *Enterobacter Cloacae* Mr2

	MR2
DNA, total number of bases	4758062
DNA coding number of bases	4224707
DNA G+C number of bases	2621395
Genes	5571
Protein coding genes	5404
RNA genes	167
rRNA genes	16
5S rRNA	8
16S rRNA	4
23S rRNA	4
tRNA genes	72
Protein coding genes with function prediction	4573
without function prediction	831
Protein coding genes with enzymes	1115
w/o enzymes but with candidate KO based enzymes	646
Protein coding genes connected to Transporter Classification	992
Protein coding genes connected to KEGG pathways3	1266
not connected to KEGG pathways	4138
Protein coding genes connected to KEGG Orthology (KO)	2423
not connected to KEGG Orthology (KO)	2981
Protein coding genes connected to MetaCyc pathways	1085
not connected to MetaCyc pathways	4319
Protein coding genes with COGs	2998
in paralog clusters	3861
in Chromosomal Cassette	5571
Biosynthetic Clusters	20
Genes in Biosynthetic Clusters	199
Fused Protein coding genes	99
Protein coding genes coding signal peptides	446
Protein coding genes coding transmembrane proteins	1334
COG clusters	1811
KOG clusters	757
Pfam clusters	2505
TIGRfam clusters	1413

Table.3b COG categories *of Enterobacter cloacae* MR2 and *Enterobacter cloacae* subsp. cloacae ATCC 13047

		MR2		ATCC13047	
Class ID	Description	CDS			
A	RNA processing and modification	1		1	
W	Extracellular structures	2		4	
D	Cell cycle control, cell division, chromosome partitioning	43		50	
V	Defense mechanisms	88		59	
F	Nucleotide transport and metabolism	114		81	
Q	Secondary metabolites biosynthesis, transport and catabolism	145		114	
N	Cell motility	154		177	
I	Lipid transport and metabolism	162		131	
U	Intracellular trafficking, secretion, and vesicular transport	164		161	
O	Posttranslational modification, protein turnover, chaperones	175		158	
H	Coenzyme transport and metabolism	198		165	
L	Replication, recombination and repair	208		289	
J	Translation, ribosomal structure and biogenesis	227		206	
T	Signal transduction mechanisms	282		258	
M	Cell wall/membrane/envelope biogenesis	318		282	
C	Energy production and conversion	320		253	
S	Function unknown	390		383	
P	Inorganic ion transport and metabolism	431		351	
K	Transcription	470		437	
G	Carbohydrate transport and metabolism	596		463	
E	Amino acid transport and metabolism	657		500	
R	General function prediction only	718		626	

Table.4 Assembly statistics genome sequence of *Enterobacter cloacae* MR2

Sr. No.	**Assembly Statistics**	
1	Total number of reads	549,959
2	Assembled Reads	507,383
3	Coverage	18.06 X
4	Number of Contigs	230
5	Consensus Length	4,758,062bp
6	Largest Contig	177,145bp
7	N50	36,082bp
8	N90	9,841bp
9	N95	7,065bp

Table.5 RAST distribution of *Enterobacter cloacae* MR2 compare with *Enterobacter cloacae* subsp. cloacae ATCC 13047 and *Enterobacter* sp. 638

	Enterobacter Cloacae		
	MR2	ATCC 13047	638
Cofactors, Vitamins, Prosthetic Groups, Pigments	9.41	6.63	7.04
Cell Wall and Capsule	5.58	6.22	6.87
Virulence, Disease and Defense	2.85	3.71	2.64
Potassium metabolism	0.98	0.84	0.87
Photosynthesis	0.00	0.00	0.00
Miscellaneous	0.81	0.92	0.90
Phages, Prophages, Transposable elements, Plasmids	0.71	1.84	1.46
Membrane Transport	5.70	5.07	4.41
Iron acquisition and metabolism	2.02	2.23	1.40
RNA Metabolism	4.35	4.73	5.95
Nucleosides and Nucleotides	3.10	3.12	3.40
Protein Metabolism	6.29	6.32	8.02
Cell Division and Cell Cycle	0.69	0.72	1.07
Motility and Chemotaxis	2.93	3.89	2.61
Regulation and Cell signaling	3.83	3.61	3.84
Secondary Metabolism	0.12	0.13	0.11
DNA Metabolism	2.95	3.20	3.42
Fatty Acids, Lipids, and Isoprenoids	3.27	3.48	3.68
Nitrogen Metabolism	1.57	1.23	1.32
Dormancy and Sporulation	0.07	0.08	0.11
Respiration	3.98	3.97	3.96
Stress Response	4.50	4.89	4.97
Metabolism of Aromatic Compounds	1.35	0.97	0.14
Amino Acids and Derivatives	12.27	12.28	11.95
Sulfur Metabolism	1.60	1.59	1.85
Phosphorus Metabolism	1.30	1.20	1.46
Carbohydrates	17.75	17.17	16.55

Table.6 Core and Pan genome analysis of *Enterobacter cloacae* MR2 with other *Enterobacter* genus spp

Organism	CDS	Pan CDS	Core CDS	Var CDS	Strain specific CDS
E. cloacae **MR2**	5703	5697	2089	3608	1876
Enterobacter sp. **638**	4396	4394	2056	2338	1062
E. cloacae **ENHKU01**	4570	4566	2055	2511	681
E.cloacae subsp. dissolvens **SP1**	4682	4682	2052	2630	1848
E. cloacae subsp. dissolvens **SDM**	4758	4752	2055	2697	459
E. cloacae **ATCC 13047**	5707	5704	2061	3643	1311

Fig.1 Per base sequence quality scores before (A) and after pre-processing and filtering (B)
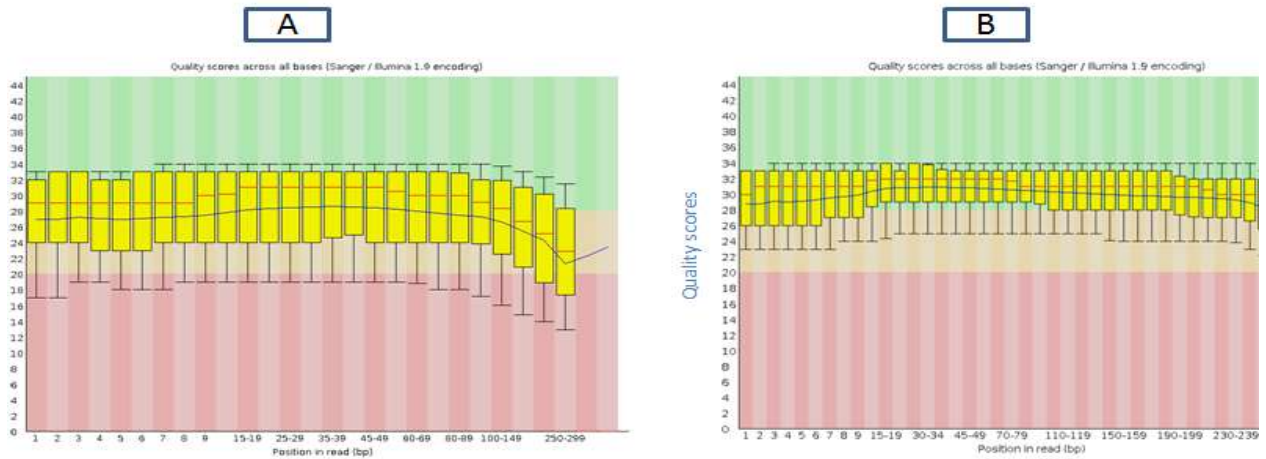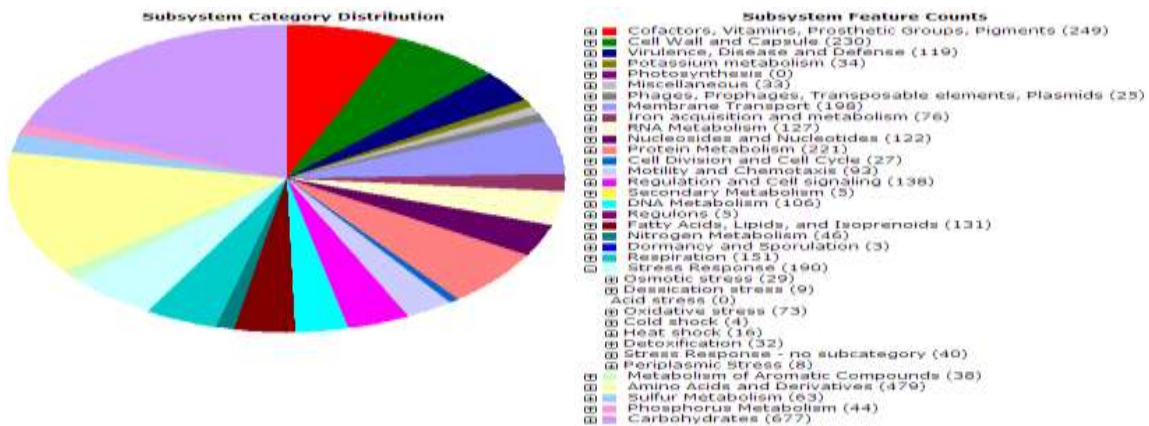

Fig.2a Genes connected to the subsystems and their distribution in different categories

Annotation indicated 2907 features (genes or CDS) within 458 Subsystems and 4507 Coding Sequences

Fig.2b RAST distribution comparison *Enterobacter cloacae* MR2 with *Enterobacter cloacae* subsp *cloacae* ATCC 13047 and another neighbor genome *Enterobacter sp.* 638
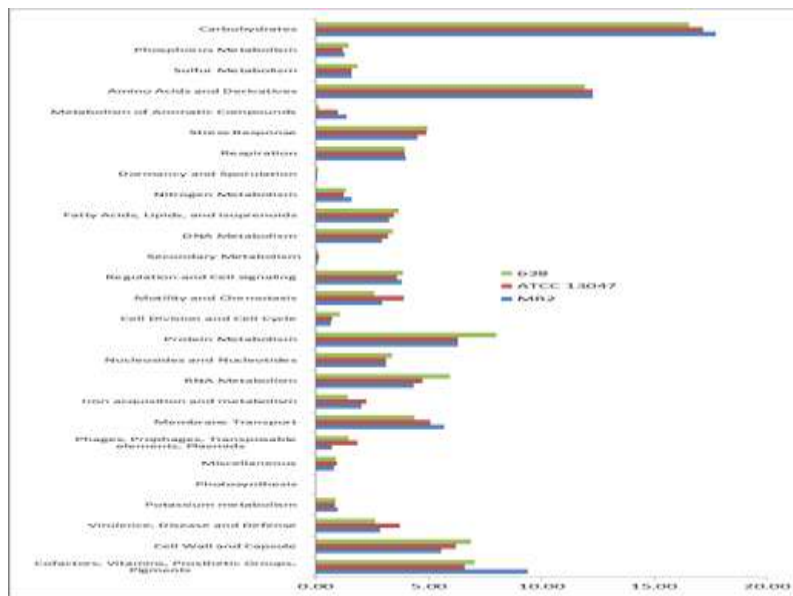
Fig.3 Alignment of *Enterobacter cloacae* MR2 ordered contigs with the reference genome *Enterobacter cloacae subspp. cloacae* ATCC 13047
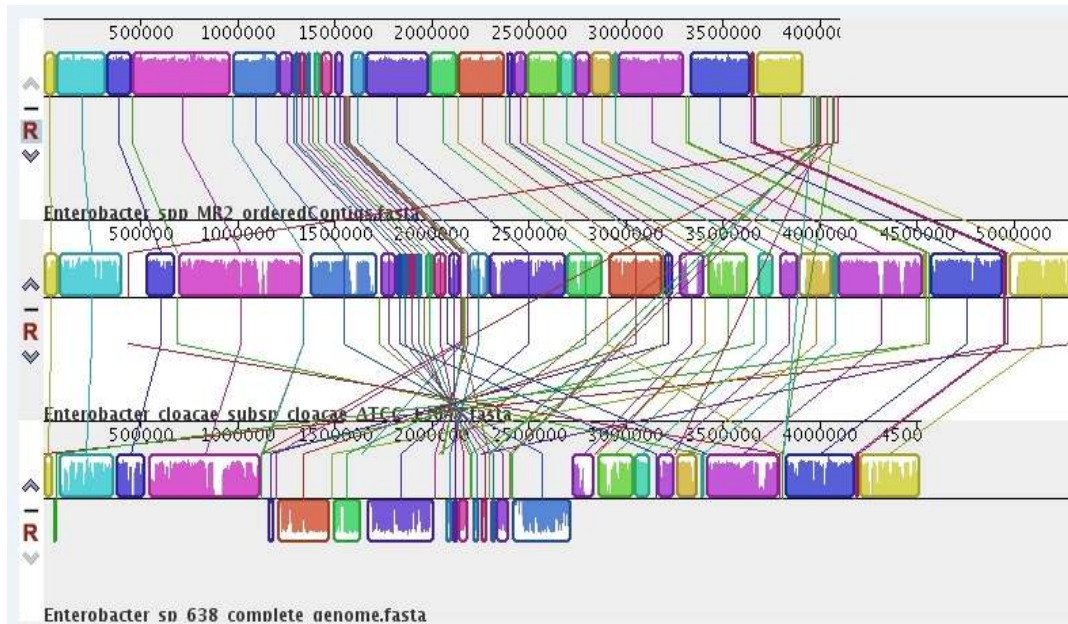
The alignment display is organized into one horizontal "panel" per input genome sequence. Each genome's panel contains the name of the genome sequence, a scale showing the sequence coordinates for that genome, and a single black horizontal center line. The regions of sequence with homology in the other two genome are indicated by colored blocks. The lines joining the blocks between three genomes trace each orthologous Locally Collinear Blocks (LCB) through every genome In this case, Row1: *Enterobacter cloacae* MR2 ordered contigs, Row2: *Enterobacter cloacae* subspp. *cloacae* ATCC 13047 genome, Row3: *Enterobacter* spp. 638 genome.

Fig.4 *Enterobacter cloacae* MR2 graphical representation in MaGe's genome browser and synteny maps
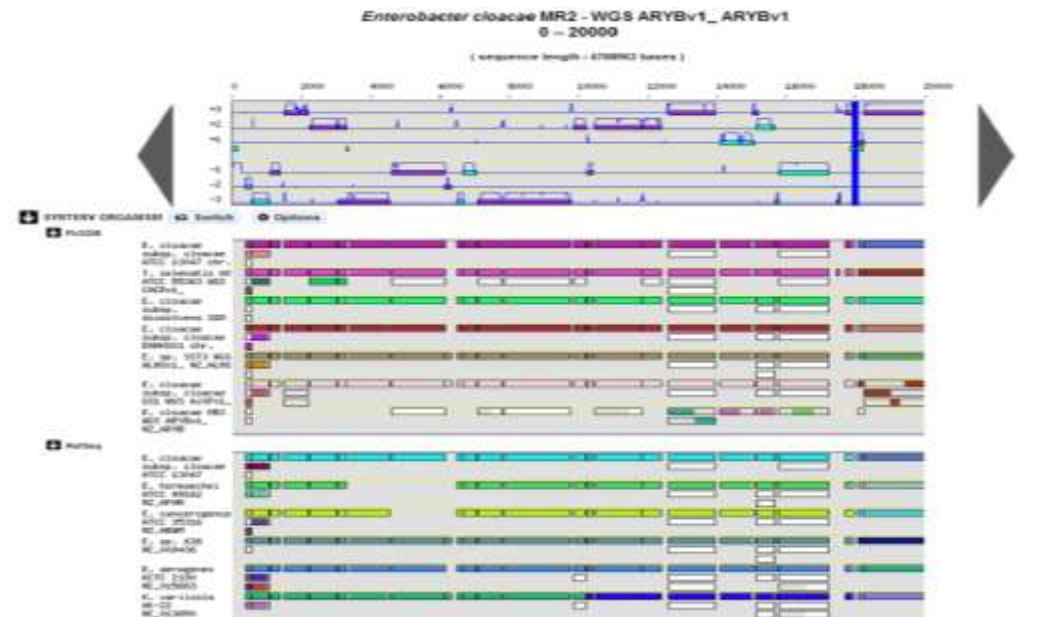
Fig.5 Circular representation of the *Enterobacter cloacae* MR2 genome
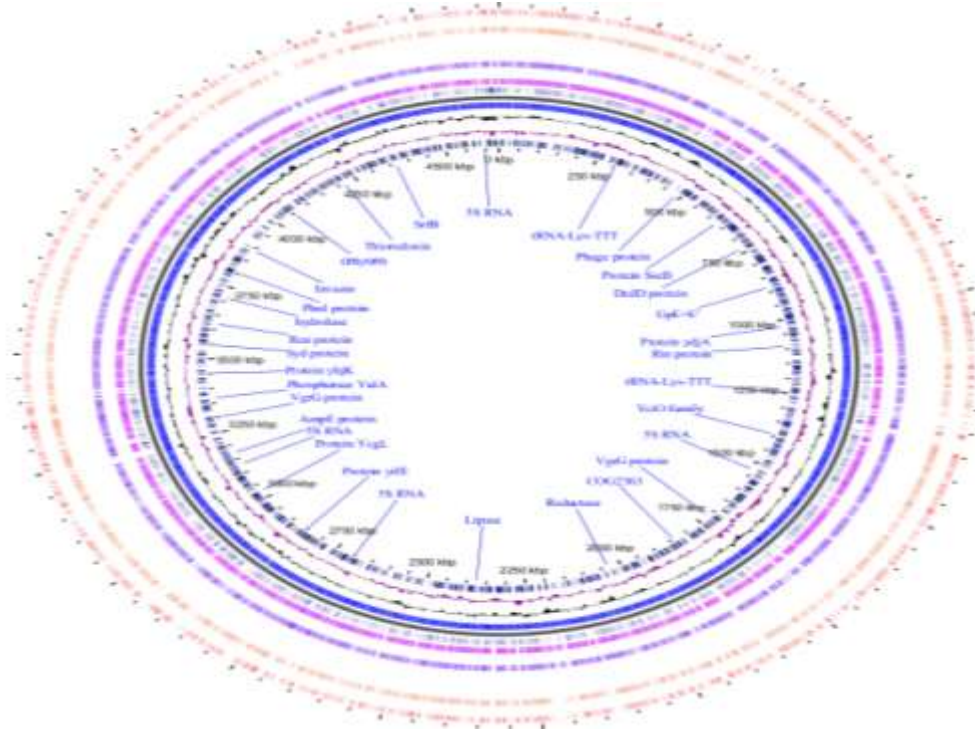
Circles display (from the outside): (1) GC percent deviation (GC window - mean GC) in a 1000-bp window. (2) Predicted CDSs transcribed in the clockwise direction. (3) Predicted CDSs transcribed in the counterclockwise direction. Genes displayed in (2) and (3) are color-coded according different categories: red and blue: MaGe validated annotations. orange: MicroScope automatic annotation with a reference genome. purple: Primary/Automatic annotations. (4) GC skew (G+C/G-C) in a 1000-bp window. (5) rRNA (blue), tRNA (green), misc_RNA (orange), Transposable elements (pink) and pseudogenes (grey).

Fig.6 Pan/Core Genome Analysis of *Enterobacter cloacae* MR2, its reference *Enterobacter cloacae* subsp *cloacae* ATCC 13047 and another neighbor genome *Enterobacter sp.* 638
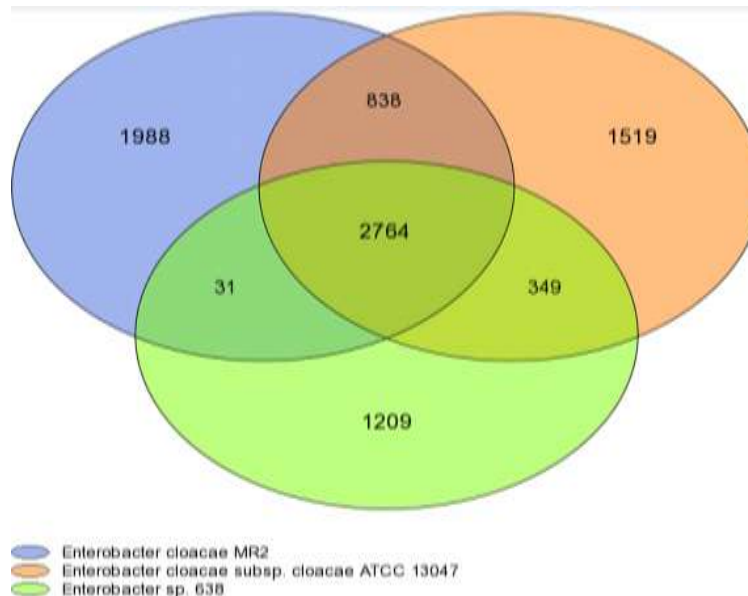

Core-genome, variable-genome and strain specific sizes are represented with a Venn diagram. Values on diagram represent the numbers of MICFAM families for each organism intersections

Clusters of orthologous genes and CDS have been highlighted in the map with the color intensity.

Comparative genome analysis

Comparative whole genome alignment is the key to find important regions to determine gene functions and establishing evolutionary events. Genome annotation identified plentiful of interesting genes that were involved in organo-phosphate degradation, detoxification and stress responses.

The BLAST based comparison of *E. cloacae* MR2 genome (Fig. 5) allowed sequence feature information to be visualized in context to reference sequence. Since this approach is reference based, only the regions present in the reference sequence or absent in the query sequence could be visualized. It also indicated that some of the features and sequences of the draft genome, which are missing from the reference genomes differed in significantly in terms of gene content. It shared most of the features in the reference genome, but there are regions and parts which might have deletions or missing of sequence parts.

Mauve software was used to construct genome-wide pair wise DNA alignments between *Enterobacter cloacae* MR2 and *Enterobacter* sp. 638. Synteny map was used for visualization of comparative analysis of complete genome assemblies at different levels of resolution, ranging from genome-scale comparison of chromosomes to comparisons of individual regions of alignment at the nucleotide level (Fig. 3).

The position of the blocks indicating the MR2 genome has different synteny as compare to other strains of the *Enterobacter*. The gap region between the block is also have the difference in the all three genomes.

Pan/core genome analysis

The Pan Genome analysis allowed to determine the common and variable genome proportion for each genome involve in the analysis. It also extracts core-genome, variable-genome and strain specific coding sequences. The Pan/Core genome analysis was carried out using the genome of *Enterobact*er sp. 638, *E. cloacae* subsp. cloacae ENHKU01, *E. cloacae* subsp. dissolvens SP1, *E. cloacae* subsp. dissolvens SDM, *E. cloacae* subsp. cloacae ATCC 13047 along with MR2. The core and pan genome interpretation is represented by Venn diagram which represents the numbers of genes/CDS in cluster of eight groups, out of which majority of the genes are shared (2764) among the three genomes (Fig. 6). MR2, ATCC 13047 and *Enterobact*er sp. 638 has 2089, 2061 and 2056 CDS respectively, for the core genome while MR2 has strain specific 1876 CDS compare to 1311 and 1062 CDS in ATCC 13047 and *Enterobact*er sp. 638 CDS (Table 6).

The Indian agriculture sector has been growing rapidly over the years. In order to control the insect from attacking their crops, pesticides are applied to lower the damages on crops and forestry products. Application of pesticides increases the cost of cultivation; on the other hand it greatly reduces the losses caused by pests and diseases, creating great economic benefits. However, pesticide residues can adversely affect ecosystems and human health and also cause serious environmental pollution.

Organophosphorus compounds cause short and long term environmental hazards and health problems. These pesticides affect the nervous system by disrupting the enzyme that regulates acetylcholine, a neurotransmitter. Organophosphorus pesticide is used to control a variety of sucking, chewing and boring

insects, spider mites, aphids and pests that attack crops like cotton, sugarcane, peanuts, tobacco, vegetables, fruits and ornamentals. Among the insecticides, monocrotophos, quinalphos and chlorpyrifos top the list of organophosphorus insecticides in the Indian market (Singh *et al.,* 2003; Singh *et al.,* 2004).

In this study, Nineteen isolates were initially screened which were found to be highly efficient in degrading chlorpyrifos and environmental stress tolerance. The *mpd* gene specific marker was amplified in all chlorpyrifos degrading isolates, except 2 isolates which indicated the presence of new or unidentified gene. The bacterial strains were tested for their ability to degrade chlorpyrifos by High Performance Liquid Chromatography (HPLC) and the results of these works have been published (Parmar *et al.,* 2014). It was found that CPD-12 named as *E. cloacae* strain MR2 based on 16s rDNA analysis, was able to degrade up to 500 ppm of chlorpyrifos in 30 h. Exploring and analyzing the genome of *E. cloacae* strain MR2 have further paved our way for understanding the mechanisms of pesticide degradation and the genes, biochemical pathways, and metabolites involved in organophosphate degradation. RAST annotation provided us with the most closely related strains in respect to our bacterial strain which are *E. cloacae* ATCC 13047 and *Enterobacter sp.* 638. RAST distribution also indicates the highest subsystem distribution as compare to *E. cloacae* ATCC 13047 and *Enterobacter sp.* 638.

Mapping of MR2 with *E. cloacae* ATCC 13047 and *Enterobacter sp.* 638 mapped most of the contigs from our draft genome with the reference genome *E. cloacae* subsp. ATCC 13047. The regions mapped were variations in the location and had gap regions, indicating the genomic island's presence in the genome.

The annotation also revealed the presence of genes for phosphorus metabolism, sulfur metabolism, metabolism of aromatic compounds, nitrogen metabolism, protein metabolism, potassium metabolism, iron acquisition and metabolism in the genome of strain MR2 which indicated the capability of the stain to promote plant growth and related activity.

The pan-genome describes the full complement of genes in a list of organisms. It is the union of all the gene families and specific genes of all the strains. It includes the core-genome containing gene families shared by all the organisms (intersection of gene families) and the variable-genome containing genes families shared by two or more organisms and strain specific genes. MR2 have higher strain specific genes as compared to *E. cloacae* ATCC 13047 and *Enterobacter sp.* 638. (Ren *et al.,* 2010) characterized the genome of *E. cloacae* ATCC 13047 and showed that the chromosome carries seven operons involved in toxic heavy-metal resistance, including two *sil* operons, three *ars* operons, a *mer* operon, and a *cop* operon. MR2 genome when compared to the same through synteny, it indicated the presence of the same gene clusters genome.

The *Enterobacter* sp. 638 genome was sequenced and analyzed by Taghavi *et al.,* (2010) and explained the plant growth promoting properties of the strain. While comparing genome of MR2 to sp.638, it also showed the presence of the gene clusters for the siderophores, Indol 3-acetic acid and phosphate solubilization. Hence the *Enterobacter cloacae* MR2 found during the course of this study, if used in the soil will not only degrade pesticide but will also increase the plant growth.

Nucleotide sequence accession numbers

This Whole Genome Shotgun project has been deposited at DDBJ / EMBL /GenBank under the accession ARYB00000000. Bio-project registered under Accession: PRJNA203096 ID: 203096.

Acknowledgements

References

Agnihotri, N., Pandey, H., Jain, S. and Srivastava, K., Persistence, leaching and movement of chlorfenvinphos, chlorpyrifos, disulfoton, fensulfothion, monocrotophos and tetrachlorvinphos in soil. *Ind J Agr Chemis*. 24, 27-31(1981)

Alexander, M., Non-biodegradable and other recalcitrant molecules. *Biotechnol and Bioengin*. 15, 611-47 (1981).

Aziz, R.K., *et al.,* The RAST Server: rapid annotations using subsystems technology. BMC *Genom*. 9, 75(2008).

Bhatnagar, A. and Gupta, A., Persistence of chlorpyriphos residues in soil and groundnut seed. *Internat Arac Newsletter*. 11, 16-17 (1992)

Chapman, R., Harris, C., Svec, H. and Robinson, J., Resistance and mobility of granular insecticides in an organic soil following furrow application for onion maggot. *J Env Sci Heal*. 19, 259-70 (1984).

Chevreux, B., Pfisterer, T. and Suhai, S., Automatic Assembly and Editing of Genomic Sequences. Genomics and Proteomics - Functional and Computational Aspects. Kluwer Academic/Plenum Publishers, New York, chap. 5, 51-65 (2000).

Cho, C. M., Mulchandani, A. and Chen, W., Bacterial cell surface display of organophosphorus hydrolase for selective screening of improved hydrolysis of organophosphate nerve agent. *Appl Environ Microb*. 68, 2026–30 (2002).

Cui, Z., Li, S. and Fu, C. Isolation of Methyl Parathion-Degrading Strain M6 and Cloning of the Methyl Parathion Hydrolase Gene. *Appl Environ Microbe*. 67, 4922-25 (2001).

Darling, A. E., Mau, B., Perna and N.T., Progressive Mauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One*. 5(6), e11147 (2010).

Davis, A. and Kuhr, R., 1976. Dissipation of chlorpyriphos from muck soil and onions. *J Econ Entom*., 69, 665-66 (1976).

Dhawan, A. and Simwat, G., Pattern of insecticides use on cotton by farmers in Punjab. *Pestol*. 20, 30-34 (1996).

Faldu, P.R,, Kothari, V.V., Kothari, C.R., Rawal, C.M., Domadia, K.K., Patel, P.A., Bhimani, H.D., Raval, V.H., Parmar, N.R., Nathani, N.M., Koringa, P.G., Joshi, C.G. and Kothari, R.K., Draft Genome Sequence of Textile Azo Dye-Decolorizing and -Degrading *Pseudomonas aeruginosa Strain PFK10*, Isolated from the Common Effluent Treatment Plant of the Ankleshwar Industrial Area of Gujarat, India. *Genome Announcement*, (2014)

Gupta, A., Parihar, N. and Bhatnagar, A.,. Lindane, chlorpyriphos and quinalphos residues in mustard seed and oil. *Bulletin of Environmental Contamination and Toxicology*. 67, 122-25 (2001).

Horne, I., Sutherland, T. and Harcourt, R., Identification of an opd

(Organophosphate Degradation) Gene in an *Agrobacterium* Isolate. *Appl Environ Microb*. 68, 3371-3376 (2002).

Kurtz, S., Phillippy, A. and Delcher, A.L., Versatile and open software for comparing large genomes. Genome Biol. 5, R12 (2004).

Mulbry, W.W., Karns, J.S., Kearney, P.C., *et al.,* Identification of a plasmid-borne parathion hydrolase gene from *Flavobacterium* sp. By Southern hybridization with opd from Pseudomonas diminuta. *Appl Environ Microbiol*. 51, 926–30 (1986).

Overbeek, R., Olson, R., Pusch, G.D., *et al.,* The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res*. 42, 206-14 (2014).

Parakhia, M.V., Tomar, R.S., Malaviya, B.J., Dhingani, R.M., Rathod, V.M., Thakkar, J.R. and Golakiya, B.A., Draft Genome Sequence of the Endophytic Bacterium *Enterobacter sp.* MR1, Isolated from Drought Tolerant Plant *(Butea monosperma)*. *Indian J Microbiol*. 54, 118–9 (2014).

Parmar, K.J., Tomar, R.S., Parakhia, M.V., Malviya, B.J., Rathod, V.M., Thakkar, J.R., Kothari, V.V., Bhatt, A.J., Bhalara, R.L. and Golakiya, B.A., Isolation and bio-analytical characterization of chloropyrifos degrading bacteria. *J. Cell and Tis. Res*. 14, 4641-46 (2014).

Read, D., Comparison of residual toxicities of twenty four registered or candidate pesticides applied to field microplots of soil by different methods. *J Econ Entom*. 69, 429-37(1976).

Ren, Yan., Ren, Yi., Zhou, Zhemin., Guo, Xi.Li., Yayue., Feng., Lu. and Wang, Lei., Complete genome sequence of *Enterobacter cloacae* subsp. cloacae type strain ATCC 13047. *J Bacteriol*.

192, 2463–64 (2010).

Rouchaud, J., Gustin, F. and Vanparys, L., Influence of slurry and ammonium nitrate fertilizations on soil and plant metabolism of chlorpyriphos in field cauliflower. *Bull Env Cont and Toxi*. 46,705-12 (1991).

Sambrook, J., Fritsch, E. and Maniatis, T., Molecular Cloning: A Laboratory Manual, 2nd edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NewYork, (1989).

Sasikala, C., Jiwal, S., Rout, P. and Ramya, M., Biodegradation of Chlorpyrifos by *Bacterial Consortium* Isolated from Agriculture Soil. *World J Microbiol Biotechnol*. 28, 1301–8 (2012).

Serdar, C.M., Gibson, D.T., Munnecke, D.M. and Lancaster, J.H., Plasmid involvement in parathion hydrolysis by *Pseudomonas diminuta*. *Appl Environ Microb*. 44, 246–49 (1982).

Singh, B.K., Walker, A., Morgan, J. and Wright, D.J., Biodegradation of Chlorpyrifos by *Enterobacter Strain B-14* and Its Use in Bioremediation of Contaminated Soils. *Appl Environ Microbiol*. 70, 4855–63(2004).

Singh, B.K., Walker, A., Morgan, J. and Wright, D.J., Effects of Soil pH on the biodegradation of chlorpyrifos and isolation of a chlorpyrifos-degrading bacterium. *Appl Environ Microb*. 69, 5198–206 (2003).

Stein, L., Genome annotation: from sequence to biology. *Nat Rev Gen*. 7, 493–503(2001).

Taghavi, S., van der Lelie, D., Hoffman, A., *et al.,* Genome sequence of the plant growth promoting endophytic bacterium *Enterobacter sp.* 638. *PLoS Genet*. 6, e1000943(2010).

Tashiro, H. and Kuhr, R., Some factors influencing the toxicity of soil applications of chlorpyriphos and diazinon to European chafer grubs. J

Econ Entom. 71, 903-07(1978).

Vallenet, D., Labarre, L., Rouy, Z., Barbe, V., Bocs, S., Cruveiller, S., Lajus, A., Pascal, G., Scarpelli, C. and C Médigue., MaGe: a microbial genome annotation system supported by synteny results. *Nucleic Acids Res.* 34(1), 53-65(2006).

Yang, C., Liu, N., Guo, X. and Qiao, C., Clonning of mpd gene from achlorpyrifos-degrading bacterium and use of this strain in bioremediation of contaminated soil. *Fed Europ Microb Soc.* 265, 118-25 (2006)

Zhang, H., Yang, C., Zhao, Q. and Qiao, C., Development of an auto fluorescent organophosphates-degrading *Stenotrophomonas sp.* with dehalogenase activity for the biodegradation of hexachlo-rocyclohexane (HCH). Bioresou Technol. 100, 3199–204 (2008).