Journal of Zhejiang University-SCIENCE B (Biomedicine & Biotechnology)
ISSN 1673-1581 (Print); ISSN 1862-1783 (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

Isolation and characterization of a crude oil degrading bacteria from formation water: comparative genomic analysis of environmental *Ochrobactrum intermedium* isolate versus clinical strains*

Lu-jun CHAI[§1], Xia-wei JIANG[§2], Fan ZHANG[1], Bei-wen ZHENG[†‡2], Fu-chang SHU[3], Zheng-liang WANG[3], Qing-feng CUI[4], Han-ping DONG[4], Zhong-zhi ZHANG[5], Du-jie HOU[1], Yue-hui SHE[†‡3]

([1]*Key Laboratory of Marine Reservoir Evolution and Hydrocarbon Accumulation Mechanism, School of Energy Resources, China University of Geosciences, Beijing 100083, China*)

([2]*State Key Laboratory for Diagnosis and Treatment of Infectious Diseases, the First Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou 310003, China*)

([3]*College of Chemistry and Environmental Engineering, Yangtze University, Jingzhou 434023, China*)

([4]*Institute of Porous Flow & Fluid Mechanics, Chinese Academy of Sciences, Langfang 065007, China*)

([5]*State Key Laboratory of Heavy Oil Processing, China University of Petroleum, Beijing 257061, China*)

[†]E-mail: zhengbw@zju.edu.cn; sheyuehui@163.com

Received Jan. 26, 2015; Revision accepted July 6, 2015; Crosschecked Sept. 17, 2015

Abstract: In this study, we isolated an environmental clone of *Ochrobactrum intermedium*, strain 2745-2, from the formation water of Changqing oilfield in Shanxi, China, which can degrade crude oil. Strain 2745-2 is aerobic and rod-shaped with optimum growth at 42 °C and pH 5.5. We sequenced the genome and found a single chromosome of 4 800 175 bp, with a G+C content of 57.63%. Sixty RNAs and 4737 protein-coding genes were identified: many of the genes are responsible for the degradation, emulsification, and metabolizing of crude oil. A comparative genomic analysis with related clinical strains (M86, 229E, and LMG3301^T) showed that genes involved in virulence, disease, defense, phages, prophages, transposable elements, plasmids, and antibiotic resistance are also present in strain 2745-2.

Key words: Comparative genome, *Ochrobactrum intermedium*, Oil degradation, Pathogen
doi:10.1631/jzus.B1500029 **Document code:** A **CLC number:** X172

1 Introduction

Oil contamination is a worldwide problem, which is growing more serious with economic development; its effects are long lasting and remediation is difficult. Several methods of oil degradation have been developed, the method with the longest history being land farming, which is "low-tech" but time-consuming (Genouw *et al.*, 1994). Physical methods such as surface heating are more efficient but energy-consuming (Edelstein *et al.*, 1994). Microbial oil degradation shows promise of being sustainable and environmentally friendly, and the screening of potential oil degrading microorganisms is becoming increasingly important. Bacteria from different habitats, such as soil (Jesubunmi, 2014; Kumar *et al.*, 2014; Pham *et al.*, 2014) and the ocean (Hazen *et al.*, 2010; Hassanshahian *et al.*, 2014), are screened for their oil degrading properties. These bacteria are then

‡ Corresponding authors
§ The two authors contributed equally to this work
* Project supported by the National High-Tech R & D Program (863) of China (No. 2013AA064402), the National Natural Science Foundation of China (Nos. 81301461 and 51474034), the Zhejiang Provincial Natural Science Foundation of China (No. LQ13H190002), and the Scientific Research Foundation of Zhejiang Provincial Health Bureau (No. 2012KYB083), China
ORCID: Bei-wen ZHENG, http://orcid.org/0000-0002-7690-372X

used individually or in a mixture (Creencia *et al.*, 2014; Silva *et al.*, 2015).

In our previous studies, several strains of bacteria, which have the ability to degrade crude oil, were isolated from the formation water of Chinese oilfields (She *et al.*, 2011; 2014; Zhang *et al.*, 2012; 2014; Zheng *et al.*, 2014). In this study, we isolated from the Changqing oilfield a strain which has rarely been isolated from an oilfield, *Ochrobactrum intermedium* strain 2745-2. *O. intermedium* was first described in 1998 with five strains formerly known as members of *Ochrobactrum anthropi* (Holmes *et al.*, 1988; Velasco *et al.*, 1998). The name, *O. intermedium*, indicates an intermediate position between *O. anthropi* and *Brucella* spp. (Velasco *et al.*, 1998). *O. anthropi* is an emerging opportunistic pathogen in immunocompromised patients (Mudshingkar *et al.*, 2013) and members of *Brucella* are pathogens causing brucellosis which is a common zoonotic infection globally (Dean *et al.*, 2012). Strains of *O. intermedium* are associated with both human beings and the environment. Some strains are pathogens which cause infection (Möller *et al.*, 1999; Apisarnthanarak *et al.*, 2005); some live in environments polluted by chromium (Kavita and Keharia, 2012), lead (Waranusantigul *et al.*, 2011), and tobacco waste (Yuan *et al.*, 2007), etc.

As a human pathogen and environmental bacterium, *O. intermedium* attracts a lot of interest. From a database survey, we found three draft genome sequences within *O. intermedium*, two of which (strains M86 and 229E) have been published (Kulkarni *et al.*, 2013; 2014). Strains M86, 229E, and LMG3301^T were isolated from a stomach biopsy and blood taken from a non-ulcer dyspeptic individual from India. Thus, all of these three strains are associated with humans; no environmental strain had been sequenced before our study. Comparative genomic analysis is needed between human and environmental isolates of *O. intermedium* to give us a better understanding of the mechanisms by which it adapts to its environment.

Here, we describe the classification and features of *O. intermedium* strain 2745-2, together with its genome sequence and the comparative genomic study we conducted with its clinical relatives, strains LMG3301^T, 229E, and M86. The aims of this work are to investigate the oil-degrading genes of strain 2745-2 and to find the distinction and similarities among the genomes and genes that reflect adaptation to specific environments.

2 Materials and methods

2.1 Sampling and isolation of oil degrading bacteria

A water sample was collected from an oil-producing well in Changqing oilfield, Shanxi Province, China, in 2012. The sample was stored immediately at 4 °C. Oil degrading bacteria were isolated using sterile crude oil as the medium. After incubation, the culture was spread on LB agar plates containing 5.0 g/L yeast extract (Difco, USA), 10.0 g/L NaCl, 10.0 g/L tryptone, and 20.0 g/L agar (Difco, USA) to select the single clones. One strain (2745-2) was further characterized. It was cultured in LB medium and genomic DNA was extracted using QIAamp DNA Mini Kit (Qiagen, Germany) following the manufacturer's instructions. 16S ribosomal RNA (rRNA) was amplified by polymerase chain reaction (PCR) using the primers as follows: 27F (5'-AGA GTT TGA TCC TGG CTC AG-3') and 1492R (5'-GGT TAC CTT GTT ACG ACT T-3').

2.2 Phylogenetic tree construction

16S rRNA nucleotide sequence analysis was conducted using the BLASTN program against the national center for biotechnology information (NCBI)-nucleotide collection (nr/nt) database. Sequences were aligned by the CLUSTALW (Larkin *et al.*, 2007). A Neighbor-Joining phylogenetic tree based on the Tamura-Nei model was constructed using MEGA6 software (Tamura *et al.*, 2013).

2.3 Characterization of strain 2745-2

Cell morphology of strain 2745-2 was examined using a scanning electron micrograph (Quanta 200, FEI Co., USA). The temperature range, pH range, and NaCl range for growth were determined using methods described before (Cheng *et al.*, 2015). Gram-reaction was carried out according to Bergey's manual (Holt *et al.*, 1994). Tests for H_2S production and indole production were conducted using the method described by Mata *et al*. (2002). Hydrolase of starch, gelatin, and casein were tested. Single carbon source utilization tests were performed using

D-glucose, maltose, lactose, D-galactose, rhamnose, raffinose, sorbitol, glycerol, cellobiose, sucrose, tetradecane, and hexadecane. Resistance to ampicillin, erythromycin, tetracycline, kanamycin, and gentamicin were tested.

2.4 Whole genome sequencing

Strain 2745-2 was cultivated aerobically in LB medium, pH 5.5 at 42 °C overnight. Genomic DNA was extracted using the method described by Marmur and Doty (1962). The resulting genomic DNA was then measured using gel electrophoresis 0.7% (7 g/L) agarose with λ-Hind III digest DNA as the marker (TaKaRa, Dalian, China). The concentration of the genomic DNA was measured by NanoDrop™ 1000 spectrophotometer (Thermo Fisher Scientific Inc., USA). Genomic DNA sequencing was performed using Illumina HiSeq2000 with Solexa paired-end sequencing strategy. One DNA library (500 bp insert size with Illumina adapter at both ends) was generated and detected by Agilent DNA analyzer 2100 (Agilent Technologies, USA).

2.5 Sequence assembly and annotation

Clean reads were assembled into scaffolds using the Velvet version 1.2.07 (Zerbino and Birney, 2008). We then used PAGIT flow (Swain *et al.*, 2012) to prolong the initial contigs and correct sequencing errors.

The transfer RNAs (tRNAs) and rRNAs were identified using tRNAscan-SE (Lowe and Eddy, 1997), RNAmmer (Lagesen *et al.*, 2007), and Rfam database (Griffiths-Jones *et al.*, 2003; Burge *et al.*, 2012). The genome annotation was predicted using the RAST server online (Aziz *et al.*, 2008). Predicted genes were blast against the Clusters of Orthologous Groups (COGs) database (Tatusov *et al.*, 2000; 2001). We applied the PHAST program to predict the prophages and putative phage-like elements in the genome (Zhou *et al.*, 2011).

2.6 Comparative genomic analysis

The genome sequences of *O. intermedium* M86, 229E, and LMG3301^T were downloaded from the NCBI database under the accession numbers AOGE00000000.1, ASXJ00000000.1, and ACQA 00000000.1, respectively (Kulkarni *et al.*, 2013; 2014). All these genomes were annotated by the RAST on-line server, which was also used for subsystem annotations (Aziz *et al.*, 2008). Contigs were re-ordered using the Mauve program (Darling *et al.*, 2010). Blasts of the three genomes together with strain 2745-2 were performed using the BLAST+ program (Camacho *et al.*, 2009). The BLAST Ring Image Generator (BRIG) was used for genome alignment visualization (Alikhan *et al.*, 2011).

2.7 Nucleotide sequence accession number

The genome sequence of *O. intermedium* 2745-2 has been deposited in GenBank with the accession number JFHY00000000.1.

3 Results

3.1 Phylogenetic analysis and characterization of strain 2745-2

Neighbor-Joining phylogenetic analysis indicated the taxonomic status of 2745-2, which is clearly classified into the same branch as *O. intermedium* LMG3301^T. *Rhizobium leguminosarum* IAM 12609^T was used as an out group (Fig. 1).

Strain 2745-2 was capable of growing at 15–45 °C and pH 5.5–9.0 with optimum conditions being 42 °C and pH 5.5. Cells are straight rods, 0.6–0.9 μm in diameter and 1.7–5.3 μm long (Fig. 2). Colonies grown at 42 °C on LB agar plate are gray, circular, and convex. H_2S and indole are produced. Gelatin and casein are hydrolyzed, but not starch. Lactose, rhamnose, tetradecane, and hexadecane are used as the carbon source, while D-glucose, maltose, D-galactose, raffinose, sorbitol, glycerol, cellobiose, and sucrose are not used. An antimicrobial susceptibility test showed that strain 2745-2 is resistant to ampicillin, erythromycin, tetracycline, kanamycin, and gentamicin.

3.2 Genome features

The draft genome size of *O. intermedium* 2745-2 was 4 800 175 bp with a G+C content of 57.62%. The draft genome contains 4737 coding sequences (CDSs) and 60 RNAs including two complete rRNA operons. Detailed information on the genome is summarized in Table 1. A total of 4285 genes were categorized into COGs functional groups (Fig. 3). Five prophage regions have been identified (Fig. 4), including one intact, two incomplete, and two questionable regions (Table 2).

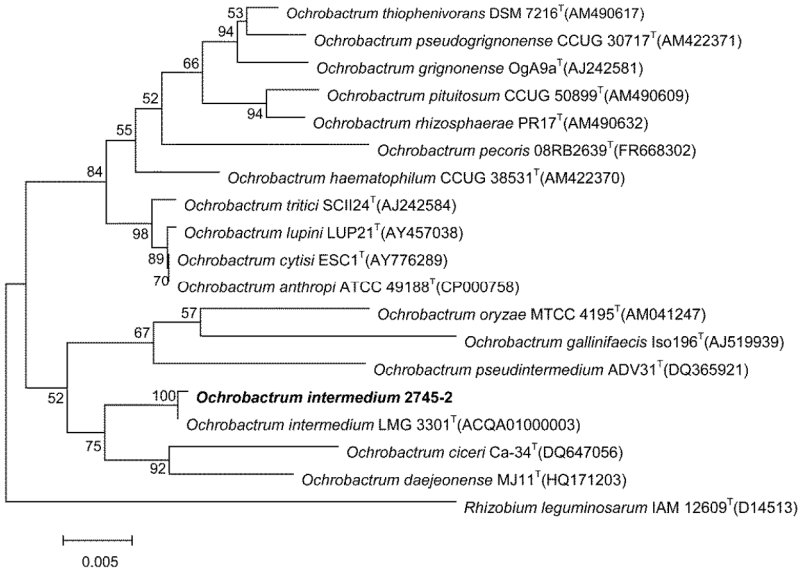

Fig. 1 Phylogenetic tree presenting the position of *Ochrobactrum intermedium* strain 2745-2

GenBank accession numbers for 16S rRNA genes of the strains used in this phylogenetic tree are shown following the organism names. The bootstrap values on the branching nodes were calculated on 1000 replications. *Rhizobium leguminosarum* IAM 12609^T was used as an out group. The scale bar indicated 0.005 substitutions per nucleotide position

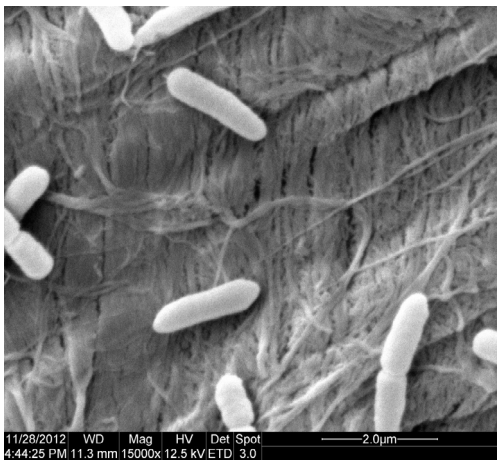

Fig. 2 Scanning electron micrograph of cells of *Ochrobactrum intermedium* strain 2745-2

Cells were grown at 42 °C in LB broth for 16 h (about early stationary phase)

3.3 Comparative genomic analysis

Comparative genomic analyses of *O. intermedium* M86, 229E, LMG3301^T, and 2745-2 were conducted. The isolation source and genomic statistics are shown in Table 3. Comparisons of subsystem features between the four genomes revealed some distinctions between 2745-2 and the other three strains (Table 4). The numbers of genes involved in virulence, disease, and defense were significantly lower in 2745-2 than in the other strains. Genes involved in secondary metabolism were higher in 2745-2. This difference in gene numbers may be due to the different habitat of 2745-2. BLAST of nucleotide sequence between 2745-2 and the other three strains was performed and the identities were visualized (Fig. 5).

Table 1 Detailed information of the draft genome sequence of *Ochrobactrum intermedium* strain 2745-2

Size (bp)	G+C content (bp)	Coding region (bp)	Gene number			
			Total	RNA	Protein-coding	COGs
4 800 175	2 766 132	4 145 190	4797	60	4737	4285

Table 2 Summary of prophage regions in strain 2745-2

Region	Region length (kb)	Completeness	CDS	Specific keyword
1	51.0	Intact	47	Terminase, portal, plate, tail
2	10.3	Incomplete	16	Tail
3	35.3	Questionable	32	Integrase, terminase, portal, head, capsid
4	40.1	Questionable	54	Terminase, portal, head, capsid, tail
5	15.5	Incomplete	20	Terminase, capsid, head, tail

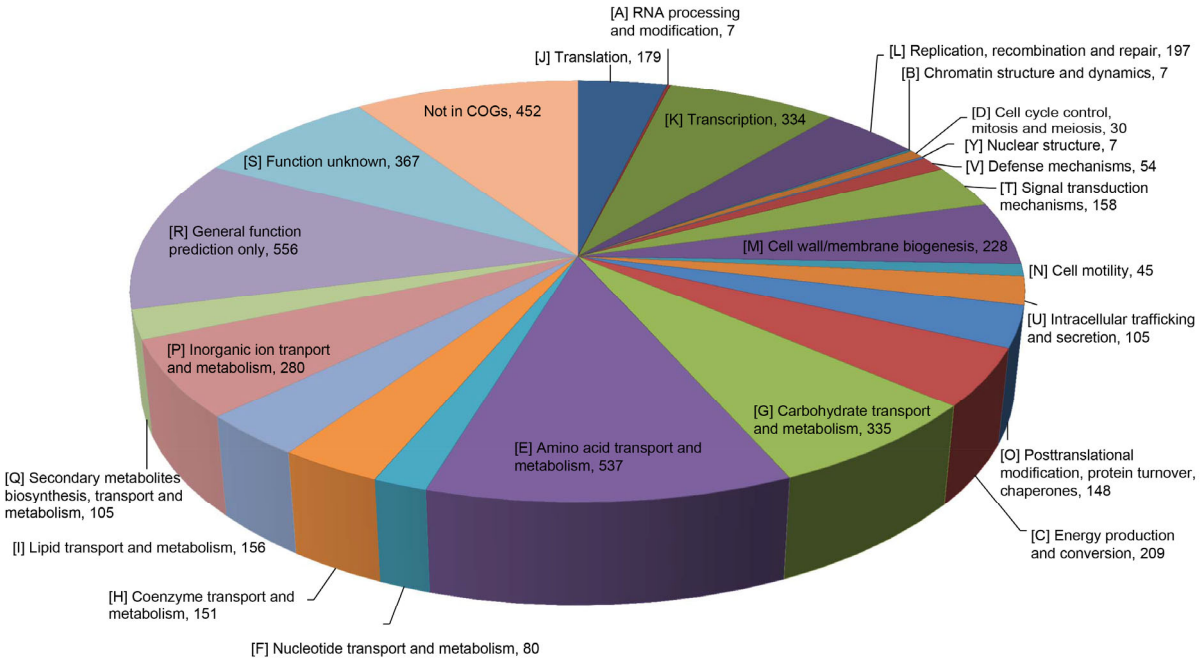

Fig. 3 Distribution of the genes associated with COG functional categories in *Ochrobactrum intermedium* strain 2745-2

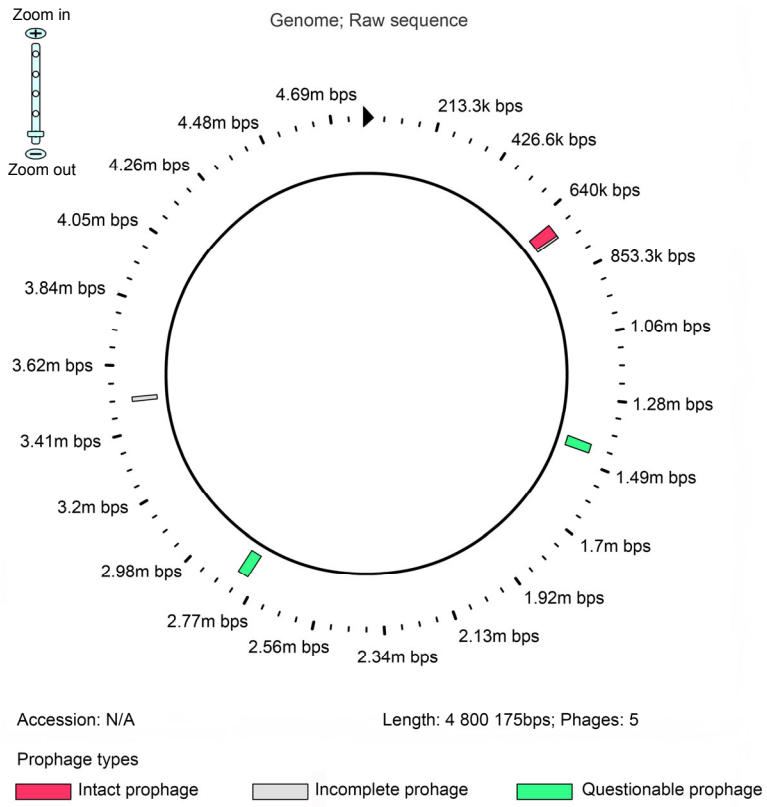

Fig. 4 Genomic view of prophage regions identified in the genome of strain 2745-2

Table 3 Comparison between the genome statistics of four strains of *Ochrobactrum intermedium*

Strain	Isolation source	Accession No.	Size (Mb)	GC (%)	Number			
					Contig	Subsystem	CDS	RNA
2745-2	Formation water	JFHY00000000.1	4.80	57.6	95	440	4737	60
M86	Gastric biopsy	AOGE00000000.1	5.19	57.9	148	480	5473	66
229E	Stomach biopsy	ASXJ00000000.1	4.81	57.9	378	468	5610	58
LMG3301^T	Blood	ACQA00000000.1	4.73	57.7	4	474	4723	70

Table 4 Comparisons between subsystem features of four strains of *Ochrobactrum intermedium*

Subsystem feature	Number of genes			
	2745-2	M86	229E	LMG3301^T
Cofactors, vitamins, prosthetic groups, pigments	224	276	305	275
Cell wall and capsule	111	107	134	111
Virulence, disease, and defense	77	102	112	94
Potassium metabolism	15	18	24	17
Photosynthesis	0	0	0	0
Miscellaneous	55	63	76	63
Phages, prophages, transposable elements, plasmids	19	38	0	15
Membrane transport	180	254	281	226
Iron acquisition and metabolism	41	55	80	56
RNA metabolism	127	145	189	157
Nucleosides and nucleotides	100	109	121	110
Protein metabolism	207	273	305	265
Cell division and cell cycle	16	30	30	28
Motility and chemotaxis	77	84	78	88
Regulation and cell signaling	70	82	84	75
Secondary metabolism	9	4	4	4
DNA metabolism	98	123	124	107
Fatty acids, lipids, and isoprenoids	113	153	169	153
Nitrogen metabolism	35	38	51	38
Dormancy and sporulation	2	2	3	1
Respiration	126	134	157	134
Stress response	119	131	146	128
Metabolism of aromatic compounds	30	32	42	32
Amino acids and derivatives	409	522	618	503
Sulfur metabolism	11	54	65	54
Phosphorus metabolism	43	53	57	49
Carbohydrates	440	496	580	484

4 Discussion

4.1 Crude oil degradation related genes

Crude oil is a mixture of hydrocarbons of various molecular weights. Many bacteria in nature have been found to be capable of degrading crude oil (Dawar and Aggarwal, 2015; Lincoln *et al.*, 2015), using it as their sole carbon source. Strain 2745-2 is one of them. Experiments showed that this strain can use tetradecane and hexadecane, the main compounds in crude oil, as its carbon source. Tetradecane and hexadecane are alkanes which can be oxidized by alkane hydroxylases, such as AlkB and P450. Two *alkB* genes and one P450 gene were found in the genome of 2745-2. *alkB* encodes a protein named alkane 1-monooxygenase, which is the key enzyme in the degradation of alkanes. Genes encoding for 2-polyprenylphenol hydroxylase

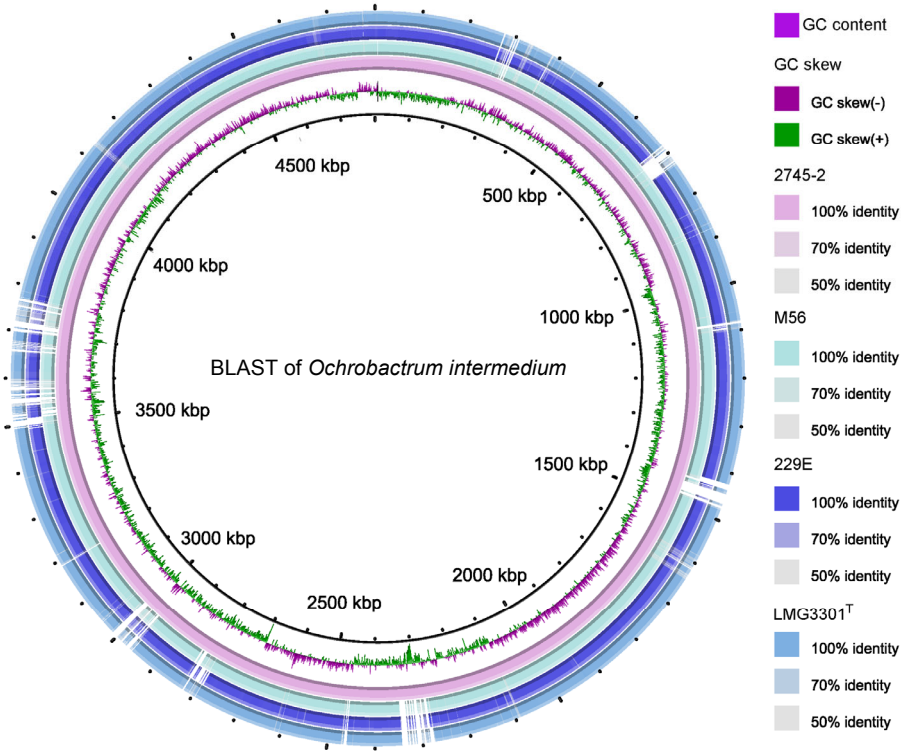

Fig. 5 BLAST visualization of *Ochrobactrum intermedium* genomes
The rings illustrate a shared identity with the four strains

and alkaline phosphatase, which are responsible for degrading aminobenzoate, were found in the genome. Genes of 2-haloalkanoic acid dehalogenase and alcohol dehydrogenase for chloroalkane and chloroalkene degrading exist in the genome. Furthermore, 2745-2 also contains all genes involved in the assembly of flagella which allows the bacterium to move to the oil-water interface the degradation process. Many other oil-degrading bacteria also have suits of flagella assembly-related genes and it is believed that these genes can also benefit emulsification of the hydrocarbon in crude oil (Das *et al.*, 2015). Strain 2745-2 also contains genes encoding for phosphomannomutase, acyl transferase, glycosyl transferase, rhamnosyltransferase, glucose-1-phosphate thymidylyltransferase, dTDP-glucose 4,6-dehydratase, dTDP-4-dehydrorhamnose 3,5-epimerase, dTDP-4-dehydrorhamnose reductase, and *N*-acyl-L-homoserine lactone synthase. These enzymes are involved in the synthesis of rhamnolipids, a class of glycolipid, which work as bacterial surfactants by reducing the surface tension, critical micelle concentration, and interfacial tension, and increasing the emulsification and solubility of hydrocarbons in mixtures such as crude oil (Das *et al.*, 2015). All these genes reflect the ability of strain 2745-2 to degrade crude oil.

4.2 Pathogen potential of strain 2745-2

Previous studies on the genome of two strains, *O. intermedium* M86 and 229E, identified many gene clusters related to virulence (Kulkarni *et al.*, 2013; 2014). In our study, the genomes of three strains of *O. intermedium*, which are available in the public database (M86, 229E, and LMG3301^T), were compared with the genome of strain 2745-2. All the strains have an average genomic size of 4.8 Mb except M86 (Table 3). 2745-2 is an environmental strain and the other three are clinical. There are fewer genes involved in virulence, disease, and defense in strain 2745-2 compared with the others (Table 4). However,

2745-2 contains several genes that are related to phages, prophages, transposable elements, and plasmids. Furthermore, we identified five prophage regions in 2745-2. One region is an intact phage with a length of 51 kb, which encodes for phage-like and hypothetical proteins (Fig. 4). It is believed that the phage-like sequences improve the cell adhesion and the ability to acquire antibiotic resistance, properties that can enable bacteria to survive in new environments and become pathogens (Casjens, 2003; Zhou *et al.*, 2011). The clinical isolates of *O. intermedium*, which are related to pathogens such as *O. anthropi* and *Brucella* spp., display a high level of resistance to forms of β-lactam antibiotics (Teyssier *et al.*, 2005) and are considered to be pathogens. Although 2745-2 was isolated from a non-clinical environment, it was found to be resistant to ampicillin and the genomic annotation results showed the presence of several β-lactamase genes that provide resistance to β-lactam antibiotics. With all these properties, strain 2745-2 may have the potential to be a pathogen.

5 Conclusions

As the first environmentally-derived strain of *O. intermedium* whose genome has been sequenced, strain 2745-2 is giving us a new perspective on its adaption to the environment. Genes involved in crude oil degradation are annotated in its genome, reflecting its ability to degrade crude oil. Further comparative genomic studies between 2745-2 and strains isolated from clinical samples will give us a better understanding of the adaption and evolution of environmental bacteria into host pathogens.

Compliance with ethics guidelines

Lu-jun CHAI, Xia-wei JIANG, Fan ZHANG, Bei-wen ZHENG, Fu-chang SHU, Zheng-liang WANG, Qing-feng CUI, Han-ping DONG, Zhong-zhi ZHANG, Du-jie HOU, and Yue-hui SHE declare that they have no conflict of interest.

This article does not contain any studies with human or animal subjects performed by any of the authors.

References

Alikhan, N.F., Petty, N.K., Zakour, N.L.B., *et al.*, 2011. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics*, **12**(1):402. [doi:10.1186/1471-2164-12-402]

Apisarnthanarak, A., Kiratisin, P., Mundy, L.M., 2005. Evaluation of *Ochrobactrum intermedium* bacteremia in a patient with bladder cancer. *Diagn. Micr. Infec. Dis.*, **53**(2):153-155. [doi:10.1016/j.diagmicrobio.2005.05.014]

Aziz, R.K., Bartels, D., Best, A.A., *et al.*, 2008. The RAST Server: rapid annotations using subsystems technology. *BMC Genomics*, **9**(1):75. [doi:10.1186/1471-2164-9-75]

Burge, S.W., Daub, J., Eberhardt, R., *et al.*, 2012. Rfam 11.0: 10 years of RNA families. *Nucleic Acids Res.*, **41**(D1): D226-D232. [doi:10.1093/nar/gks1005]

Camacho, C., Coulouris, G., Avagyan, V., *et al.*, 2009. BLAST+: architecture and applications. *BMC Bioinformatics*, **10**(1): 421. [doi:10.1186/1471-2105-10-421]

Casjens, S., 2003. Prophages and bacterial genomics: what have we learned so far? *Mol. Microbiol.*, **49**(2):277-300. [doi:10.1046/j.1365-2958.2003.03580.x]

Cheng, H., Zhang, S., Huo, Y.Y., *et al.*, 2015. *Gilvimarinus polysaccharolyticus* sp. nov., an agar-digesting bacterium isolated from seaweed, and emended description of the genus *Gilvimarinus*. *Int. J. Syst. Evol. Microbiol.*, **65**(Pt 2): 562-569. [doi:10.1099/ijs.0.065078-0]

Creencia, A.R., Mendoza, B.C., Migo, V.P., *et al.*, 2014. Degradation of residual jatropha oil by a promising lipase-producing bacterial consortium. *Philipp. J. Sci.*, **143**(1):73-78.

Darling, A.E., Mau, B., Perna, N.T., 2010. ProgressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS ONE*, **5**(6):e11147. [doi:10.1371/journal.pone.0011147]

Das, D., Baruah, R., Roy, A.S., *et al.*, 2015. Complete genome sequence analysis of *Pseudomonas aeruginosa* N002 reveals its genetic adaptation for crude oil degradation. *Genomics*, **105**(3):182-190. [doi:10.1016/j.ygeno.2014.12.006]

Dawar, C., Aggarwal, R.K., 2015. Draft genome sequence of hydrocarbon-degrading *Pseudomonas putida* strain KG-4, isolated from soil samples collected from Krishna-Godavari Basin in India. *Genome Announc.*, **3**(3): e00590-e00615. [doi:10.1128/genomeA.00590-15]

Dean, A.S., Crump, L., Greter, H., *et al.*, 2012. Global burden of human brucellosis: a systematic review of disease frequency. *PLoS Negl. Trop. Dis.*, **6**(10):e1865. [doi:10.1371/journal.pntd.0001865]

Edelstein, W., Iben, I., Mueller, O., *et al.*, 1994. Radiofrequency ground heating for soil remediation: science and engineering. *Environ. Prog.*, **13**(4):247-252. [doi:10.1002/ep.670130413]

Genouw, G., de Naeyer, F., van Meenen, P., *et al.*, 1994. Degradation of oil sludge by landfarming—a case-study at the Ghent harbour. *Biodegradation*, **5**(1):37-46. [doi:10.1007/BF00695212]

Griffiths-Jones, S., Bateman, A., Marshall, M., *et al.*, 2003. Rfam: an RNA family database. *Nucleic Acids Res.*, **31**(1):

439-441. [doi:10.1093/nar/gkg006]

Hassanshahian, M., Zeynalipour, M.S., Musa, F.H., 2014. Isolation and characterization of crude oil degrading bacteria from the Persian Gulf (Khorramshahr provenance). *Mar. Pollut. Bull.*, **82**(1-2):39-44. [doi:10.1016/j.marpolbul.2014.03.027]

Hazen, T.C., Dubinsky, E.A., DeSantis, T.Z., *et al.*, 2010. Deep-sea oil plume enriches indigenous oil-degrading bacteria. *Science*, **330**(6001):204-208. [doi:10.1126/science.1195979]

Holmes, B., Popoff, M., Kiredjian, M., *et al.*, 1988. *Ochrobactrum anthropi* gen. nov., sp. nov. from human clinical specimens and previously known as group Vd. *Int. J. Syst. Bacteriol.*, **38**(4):406-416. [doi:10.1099/00207713-38-4-406]

Holt, J.G., Krieg, N.R., Sneath, P.H., *et al.*, 1994. Bergey's Manual of Determinative Bacteriology, 9th Ed. Williams and Wilkins, Baltimore.

Jesubunmi, C.O., 2014. Isolation of oil-degrading microorganisms in spent engine oil-contaminated soil. *J. Biol. Agric. Healthcare*, **4**(25):191-195.

Kavita, B., Keharia, H., 2012. Reduction of hexavalent chromium by *Ochrobactrum intermedium* BCR400 isolated from a chromium-contaminated soil. *3 Biotech*, **2**(1):79-87. [doi:10.1007/s13205-011-0038-0]

Kulkarni, G., Dhotre, D., Dharne, M., *et al.*, 2013. Draft genome of *Ochrobactrum intermedium* strain M86 isolated from non-ulcer dyspeptic individual from India. *Gut Pathog.*, **5**:7. [doi:10.1186/1757-4749-5-7]

Kulkarni, G., Shetty, S., Dharne, M., *et al.*, 2014. Genome sequencing analysis reveals virulence-related gene content of *Ochrobactrum intermedium* strain 229E, a urease-positive strain isolated from the human gastric niche. *FEMS Microbiol. Lett.*, **359**(1):12-15. [doi:10.1111/1574-6968.12549]

Kumar, V., Singh, S., Manhas, A., *et al.*, 2014. Bioremediation of petroleum hydrocarbon by using *Pseudomonas* species isolated from petroleum contaminated soil. *Analysis*, **30**(4):1771-1776. [doi:10.13005/ojc/300436]

Lagesen, K., Hallin, P., Rødland, E.A., *et al.*, 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.*, **35**(9):3100-3108. [doi:10.1093/nar/gkm160]

Larkin, M.A., Blackshields, G., Brown, N., *et al.*, 2007. Clustal W and Clustal X version 2.0. *Bioinformatics*, **23**(21):2947-2948. [doi:10.1093/bioinformatics/btm404]

Lincoln, S.A., Hamilton, T.L., Juárez, A.G.V., *et al.*, 2015. Draft genome sequence of the piezotolerant and crude oil-degrading bacterium *Rhodococcus qingshengii* strain TUHH-12. *Genome Announc.*, **3**(2):e00268-e00315. [doi:10.1128/genomeA.00268-15]

Lowe, T.M., Eddy, S.R., 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.*, **25**(5):955-964. [doi:10.1093/nar/25.5.0955]

Marmur, J., Doty, P., 1962. Determination of the base composition of deoxyribonucleic acid from its thermal denaturation temperature. *J. Mol. Biol.*, **5**(1):109-118. [doi:10.1016/S0022-2836(62)80066-7]

Mata, J.A., Martínez-Cánovas, J., Quesada, E., *et al.*, 2002. A detailed phenotypic characterisation of the type strains of *Halomonas* species. *Syst. Appl. Microbiol.*, **25**(3):360-375. [doi:10.1078/0723-2020-00122]

Möller, L.V., Arends, J.P., Harmsen, H.J., *et al.*, 1999. *Ochrobactrum intermedium* infection after liver transplantation. *J. Clin. Microbiol.*, **37**(1):241-244.

Mudshingkar, S., Choure, A., Palewar, M., *et al.*, 2013. *Ochrobactrum anthropi*: an unusual pathogen: are we missing them? *Indian J. Med. Microbiol.*, **31**(3):306-308. [doi:10.4103/0255-0857.115664]

Pham, V.H., Kim, J., Jeong, S.W., 2014. Enhanced isolation and culture of highly efficient psychrophilic oil-degrading bacteria from oil-contaminated soils in South Korea. *J. Environ. Biol.*, **35**(6):1145-1149.

She, Y.H., Zhang, F., Xia, J.J., *et al.*, 2011. Investigation of biosurfactant-producing indigenous microorganisms that enhance residue oil recovery in an oil reservoir after polymer flooding. *Appl. Biochem. Biotech.*, **163**(2):223-234. [doi:10.1007/s12010-010-9032-y]

She, Y.H., Wu, W.Q., Hang, C.C., *et al.*, 2014. Genome sequence of *Brevibacillus agri* strain 5-2, isolated from the formation water of petroleum reservoir. *Mar. Genomics*, **18**:123-125. [doi:10.1016/j.margen.2014.08.006]

Silva, D.S.P., de Lima Cavalcanti, D., de Melo, E.J.V., *et al.*, 2015. Bio-removal of diesel oil through a microbial consortium isolated from a polluted environment. *Int. Biodeter. Biodegr.*, **97**:85-89. [doi:10.1016/j.ibiod.2014.09.021]

Swain, M.T., Tsai, I.J., Assefa, S.A., *et al.*, 2012. A post-assembly genome-improvement toolkit (PAGIT) to obtain annotated genomes from contigs. *Nat. Protoc.*, **7**(7):1260-1284. [doi:10.1038/nprot.2012.068]

Tamura, K., Stecher, G., Peterson, D., *et al.*, 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.*, **30**(12):2725-2729. [doi:10.1093/molbev/mst197]

Tatusov, R.L., Galperin, M.Y., Natale, D.A., *et al.*, 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.*, **28**(1):33-36. [doi:10.1093/nar/28.1.33]

Tatusov, R.L., Natale, D.A., Garkavtsev, I.V., *et al.*, 2001. The COG database: new developments in phylogenetic classification of proteins from complete genomes. *Nucleic Acids Res.*, **29**(1):22-28. [doi:10.1093/nar/29.1.22]

Teyssier, C., Marchandin, H., Jean-Pierre, H., *et al.*, 2005. Molecular and phenotypic features for identification of the opportunistic pathogens *Ochrobactrum* spp. *J. Med.*

Microbiol., **54**(10):945-953. [doi:10.1099/jmm.0.46116-0]

Velasco, J., Romero, C., López-Goñi, I., *et al.*, 1998. Evaluation of the relatedness of *Brucella* spp. and *Ochrobactrum anthropi* and description of *Ochrobactrum intermedium* sp. nov., a new species with a closer relationship to *Brucella* spp. *Int. J. Syst. Bacteriol.*, **48**(3):759-768. [doi:10.1099/00207713-48-3-759]

Waranusantigul, P., Lee, H., Kruatrachue, M., *et al.*, 2011. Isolation and characterization of lead-tolerant *Ochrobactrum intermedium* and its role in enhancing lead accumulation by *Eucalyptus camaldulensis*. *Chemosphere*, **85**(4):584-590. [doi:10.1016/j.chemosphere.2011.06.086]

Yuan, Y., Lu, Z., Huang, L., *et al.*, 2007. Biodegradation of nicotine from tobacco waste extract by *Ochrobactrum intermedium* DN2. *J. Ind. Microbiol. Biotechnol.*, **34**(8): 567-570. [doi:10.1007/s10295-007-0212-x]

Zerbino, D.R., Birney, E., 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Res.*, **18**(5):821-829. [doi:10.1101/gr.074492.107]

Zhang, F., She, Y., Chai, L., *et al.*, 2012. Microbial diversity in long-term water-flooded oil reservoirs with different *in situ* temperatures in China. *Sci. Rep.*, **2**:760. [doi:10.1038/srep00760]

Zhang, F., Jiang, X., Chai, L., *et al.*, 2014. Permanent draft genome sequence of *Bacillus flexus* strain T6186-2, a multidrug-resistant bacterium isolated from a deep-subsurface oil reservoir. *Mar. Genomics*, **18**:135-137. [doi:10.1016/j.margen.2014.09.007]

Zheng, B., Zhang, F., Chai, L., *et al.*, 2014. Permanent draft genome sequence of *Geobacillus thermocatenulatus* strain GS-1. *Mar. Genomics*, **18**:129-131. [doi:10.1016/j.margen.2014.09.005]

Zhou, Y., Liang, Y., Lynch, K.H., *et al.*, 2011. PHAST: a fast phage search tool. *Nucleic Acids Res.*, **39**(Suppl. 2): W347-W352. [doi:10.1093/nar/gkr485]

中文概要

题　目：一株分离自地层水的石油降解菌的特性研究：*Ochrobactrum intermedium* 环境分离菌株与临床分离菌株的比较基因组分析

目　的：对一株地层水分离的石油降解菌 *Ochrobactrum intermedium* 2745-2 进行生理生化特性的研究、全基因组测序以及比较基因组研究。

创新点：首次对一株分离自地层水的石油降解菌 *O. intermedium* 2745-2 进行了生理生化特性研究以及基因组测序，从基因组角度解释菌株 2745-2 对石油的降解能力。通过菌株 2745-2 与同种其他临床分离菌株的比较基因组学分析，表明 2745-2 仍具有多种与致病性相关的基因。

方　法：通过微生物富集培养的方法从油井的地层水中分离石油降解微生物，通过聚合酶链反应（PCR）扩增 16S 核糖体 RNA（rRNA）序列进行比较和分析确定菌株的分类地位属于 *O. intermedium*（图 1）。采用 Illumina HiSeq2000 对菌株 2745-2 进行高通量测序，采用 Velvet 1.2.07 和 RAST server 分别进行数据组装和注释（表 1）。PHAST 寻找基因组中的噬菌体相关序列（图 4 和表 2）。通过 BLAST+和 BRIG 对环境分离菌株（2745-2）和临床分离菌株（M86、229E 和 LMG3301^T）进行基因组比较（表 3、表 4 和图 5）。

结　论：首次对一株环境分离的 *O. intermedium*（2745-2）进行全基因组测序，揭示具有多种与石油降解相关的基因。通过环境分离菌株（2745-2）与临床分离菌株（M86, 229E 和 LMG3301^T）的基因组比较分析，表明 2745-2 仍具有多种致病性相关的基因。

关键词：*Ochrobactrum intermedium*；石油降解；比较基因组；病原菌