Archives of Microbiology
https://doi.org/10.1007/s00203-019-01622-2

ORIGINALPAPER

Genome sequence of a spore-laccase forming, BPA-degrading *Bacillus* sp. GZB isolated from an electronic-waste recycling site reveals insights into BPA degradation pathways

Ranjit Das[1,2] · Zhishu Liang[1,2] · Guiying Li[2] · Bixian Mai[1] · Taicheng An[2]

Received: 10 September 2018 / Revised: 18 January 2019 / Accepted: 22 January 2019

Abstract

Bisphenol A (BPA) is a synthetic chemical with known deleterious effects on biota. A genome sequencing project is an important starting point for designing a suitable BPA bioremediation process, because it provides valuable genomic information about the physiological, metabolic, and genetic potential of the microbes used for the treatment. This study explored genomic insights provided by the BPA-degrading strain *Bacillus* sp. GZB, previously isolated from electronic-waste-dismantling site. The GZB genome is a circular chromosome, comprised of a total of 4,077,007 bp with G+C content comprising 46.2%. Genome contained 23 contigs encoded by 3881 protein-coding genes with nine rRNA and 53 tRNA genes. A comparative study demonstrated that strain GZB bloomed with some potential features as compared to other *Bacillus* species. In addition, strain GZB developed spore cells and displayed laccase activity while growing at elevated stress levels. Most importantly, strain GZB contained many protein-coding genes associated with BPA degradation, as well as the degradation of several other compounds. The protein-coding genes in the genome revealed the genetic mechanisms associated with the BPA degradation by strain GZB. This study predicts four possible degradation pathways for BPA, contributing to the possible use of strain GZB to remediate different polluted environments in the future.

Keywords Bisphenol A · Bioremediation · *Bacillus* sp. GZB · Electronic-waste · Genome

Introduction

Bisphenol A [BPA; 2, 2-bis (4-hydroxyphenyl) propane] is a synthetic organic compound, produced at high volumes worldwide (Huang et al. 2012; Im et al. 2016). BPA is an essential constituent in the manufacturing of polycarbonate plastics, epoxy resins, and other applied materials; these materials are found in many every day products, such as baby bottles, food containers, thermal papers, lacquers, water pipes, toys, medical equipment, and electronics (Rochester et al. 2015; Larsson et al. 2017). Broad applications and human activities have made BPA ubiquitous in water, sediment/soil, the atmosphere, and in living flora and fauna worldwide (Grignard et al. 2012; Li et al. 2012; Chen et al. 2016). BPA is an endocrine disruptor, genotoxic, neurotoxin, and estrogenic compounds that poses significant risk to all living organisms, including human health (Grignard et al. 2012; Rochester et al. 2015; Wu et al. 2017). This makes it important to protect the environment and human beings from the impact of BPA by developing an efficient remediation system. Implementing indigenous microbes to

Communicated by Erko Stackebrandt.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00203-019-01622-2) contains supplementary material, which is available to authorized users.

✉ Guiying Li
ligy1999@gdut.edu.cn

[1] State Key Laboratory of Organic Geochemistry, Guangzhou Institute of Geochemistry, Chinese Academy of Sciences, Guangzhou 510640, China

[2] Guangzhou Key Laboratory Environmental Catalysis and Pollution Control, Guangdong Key Laboratory of Environmental Catalysis and Health Risk Control, School of Environmental Science and Engineering, Institute of Environmental Health and Pollution Control, Guangdong University of Technology, Guangzhou 510006, China

Published online: 12 February 2019

remediate BPA-exposed environments is the most crucial step for developing a bioremediation strategy. Thus, studying the microbial genome is a significant step forward in advancing our genomic knowledge, particularly with respect to the physiological, metabolic, and genetic potential of the microorganisms. This would advance the design of a precise bioremediation technique for polluted environments.

Currently, there are many members of the genus *Bacillus* in the environment, with a total of ~336 species and seven subspecies (List of Prokaryotic names with Standing in Nomenclature; http://www.bacterio.net/bacillus.html). The *Bacillus* species have developed physiological abilities and adaptability in different environmental situations, which explain their presence in different environments, including anaerobic, temperate, acidic, neutral, and alkaline soils, sediment, and water (Nicholson et al. 2000; Li et al. 2012; García-Díaz et al. 2013; Das and Kazy 2014; Logan et al. 2015). Most *Bacillus* species have flourished, conducting many biochemical activities. For instance, they include antibiotic producers, heterotrophic nitrifiers, denitrifiers, nitrogen fixers, iron precipitators, selenium oxidizers, and manganese oxidizers or reducers. They also degrade different substrates derived from both natural and human origins.

Most of the known *Bacillus* species have been applied to several agricultural, industrial, medical, and biotechnological processes (García-Díaz et al. 2013; Das and Kazy 2014; Logan et al. 2015; Liang et al. 2016). *Bacillus* species can produce laccase, an oxidoreductase enzyme, which can oxidize a wide range of substrates, including phenols, polyphenols, methoxy phenols, aromatic amines, diamines, and other recalcitrant environmental pollutants (Held et al. 2005; Lu et al. 2012; Hautphenne et al. 2016). In our previous studies, we isolated the strain *Bacillus* sp. GZB from creek sediment at an electronic-waste recycling site and used it to degrade BPA (Li et al. 2012). The strain GZB was also characterized as a spore-laccase producer, and the enzyme was further applied to degrade BPA (Das et al. 2018). However, to further apply strain GZB, it was essential to investigate their genomic information to extend our knowledge with respect to the environmental remediation of BPA.

This study classified and established a set of features by describing the genome sequencing and annotation of *Bacillus* sp. GZB; the genomic features of the species were also compared with other members of *Bacillus* sp. The genomic study revealed previously unknown metabolic physiognomies of strain GZB, increasing our understanding about their metabolic plasticity. The GZB strain generated spore cells in a nutrient sporulation medium (NSM) and showed laccase activity. Here, we measured laccase activity during spore development in the NSM broth. The genome sequencing of strain GZB provided important genomic data, linked with the degradation of different xenobiotic compounds, especially BPA. This allowed the prediction of possible BPA degradation routes. The genes present in the GZB genome flourished in their potentiality to degrade BPA or other relevant compounds. This information could be further applied to bioremediate many other contaminated environments.

Materials and methods

Media formulation, growth, and genomic DNA preparation

The *Bacillus* sp. GZB used in this study had been previously isolated from the sludge of an electronic-waste recycling site in Guiyu, China. The GZB strain was aerobically cultivated at 37 °C for 12 h in a Luria–Bertani (LB) broth. Strain GZB had a strong ability to degrade BPA in mineral salt medium (MSM) or 0.1 M McIlvaine buffer (pH 6.8) (Li et al. 2012; Das et al. 2018); we further clarified it to grow on an MSM agar plate, supplemented with 10 mg L^{-1} of BPA. The medium composition and growth conditions for strain GZB were previously established by Li et al. (2012). For genome sequencing, the LB-broth culture was harvested using centrifugation at 12,000 rpm for 10 min. The genomic DNA was extracted using a Promega Wizard Genomic DNA Purification Kit (Promega, Madison, USA), according to the manufacturer's instructions. The extracted genomic DNA was checked using 1.0% agarose gel electrophoresis and quantified by spectrophotometry using a UV–Vis spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). To develop the GZB spore cells, we used a nutrient sporulation medium (NSM) broth.

Laccase activity of GZB during spore-cell growth in NSM

The laccase activity of strain GZB was assayed spectrophotometrically using spore cells. The spore cells from GZB were developed in NSM broth at 37 °C for 4 days (Das et al. 2018). The cells were collected at different time intervals and harvested using centrifugation at 8000 rpm for 10 min. The cells were then re-suspended in sterile deionized water with 0.1 mg mL^{-1} of lysozyme (incubated at 37 °C for 10 min) to lyse the remaining vegetative cells (Lu et al. 2012). Then, the cells were purified by washing them with 1 M NaCl, followed by 0.14 M NaCl, 0.1% (w/v) SDS and deionized water (Das et al. 2018). The laccase activities of the spore suspension were estimated using 2, 2′-Azino-bis (3-ethylbenzothiazoline-6-sulfonate) (ABTS) and syringaldazine (SGZ) as the substrates, respectively. A standard reaction mixture in a 0.1 M McIlvaine buffer (pH 3.0 for ABTS and pH 6.8 for SGZ) contained either 0.1 mM of ABTS or SGZ and a purified spore-cell suspension. The reaction mixtures were incubated at 50 °C for 5 min in a

water bath and then placed on ice for one minute to stop the reactions. ABTS and SGZ oxidation were determined at 420 nm (ε_{420} = 36,000 M^{-1} cm^{-1}) and 525 nm (ε_{525} = 65 000 M^{-1} cm^{-1}), respectively (Lu et al. 2012).

Genome project history

Bacillus sp. GZB was selected for the genome sequencing because of its ability to degrade BPA under both aerobic and anaerobic conditions; genome sequencing was conducted to identify the detailed metabolic mechanisms during BPA degradation. Sangon Biotech (Shanghai, China) previously completed genome sequencing for strain GZB. A summary of the sequencing project and its association with MIGS version 2.0 (Field et al. 2008) compliance are shown in Table S1.

Genome sequencing and assembly

The GZB genome was sequenced using Illumina Miseq platform with a 2 × 300 pair-end sequencing strategy; the average length of the paired-end sequence was 301 bp. The paired-end sequencing library was constructed using 3,910,132 reads from a total of 1,176,949,732 bp. It was further assembled using SPAdes version 3.5.0 (Bankevich et al. 2012). The final genome contained 23 contigs with N50 of 1,002,563 bp. The longest contig was 1,181,109 bp, where the library presented an approximately 268.78 × coverage of the genome.

Genome analysis

After genome sequencing, the data were analyzed using the NCBI Prokaryotic Genomes Automatic Annotation Pipeline (PGAAP) (Angiuoli et al. 2008); the Open Reading Frames (ORFs) were generated by PROKKA (Prokka et al. 2014). The predicted ORFs were translated using EMBOSS Transeq program and were searched in the NCBI non-redundant database, including the Cluster of Orthologous Groups (COG), Conserved Domain Database (CDD), PFAM, and Kyoto Encyclopedia of Genes and Genomes (KEGG) database (Finn et al. 2015). All data sources from the database were combined to assert a product description for each of the predicted proteins. The KEGG database was further used to construct complete metabolic pathways (Kanehisa et al. 2016). The ribosomal RNAs and Transfer RNAs were predicted using the RNAmmer and tRNAscan-SE programs (Lagesen et al. 2007). The TMHMM server v.2.0 (Krogh et al. 2001) and SignalP server v.4.1 (Petersen et al. 2011) were applied to predict genes with signal peptides and transmembrane helices, respectively. In addition, the Genomic Islands were recognized using Island viewer4 online tool (Bertelli et al. 2017).

Phylogenetic analysis and average nucleotide identity calculation

The phylogenetic analysis using the 16S rRNA gene of GZB with other sequences retrieved from the GenBank database was conducted by aligning them with ClustalW (Thompson et al. 1994). The phylogenetic tree was constructed using MEGA 6.0 (Tamura et al. 2013), followed by neighbor-joining methods incorporating a Jukes–Cantor distance correction (Jukes and Cantor 1969). The sequence of *Clostridium botulinum* DSM1734 (NR_036786.1) was used as an outgroup; a bootstrap analysis was performed with 100 re-samplings to assign confidence levels to the nodes in the phylogenetic tree. The average nucleotide identity (ANI) between the genome sequences was calculated using the ANI calculator (Yoon et al. 2017). For the ANI study, the most closely phylogenetically related strains were applied to predict the specific taxonomic affiliation of the strain GZB.

Nucleotide sequence accession number

The genome project was submitted in GOLD (Genomes Online Database) with the identity number of Gp0206491. The Whole Genome Shotgun (WGS) project of strain GZB was deposited to the DDBJ/EMBL/GenBank under the accession number MTQG00000000; the first version described in this paper was MTQG01000001. The additional functional prediction was accomplished using the RAST server with the SEED database (Aziz et al. 2008) under the accession number 936599.3.

Results and discussion

General features and phylogenetic analysis of strain *Bacillus* sp. GZB

Bacillus sp. GZB was isolated from sediment collected at an electronic-waste recycling site in Guiyu, China. The strain is best known for its strong ability to degrade BPA; the BPA became a unique source of carbon and energy under aerobic and anaerobic conditions (Li et al. 2012). Strain GZB grew readily on LB agar at 37 °C and produced circular white pigmented colonies with an undulating margin after 24 h of incubation (Fig. 1Sa). The strain also showed visible colonies on a mineral salt medium (MSM) agar plate supplemented with BPA (10 mg L^{-1}) (Fig. 1Sb). A gram-staining examination indicated a Gram-positive, rod-shaped bacterium (Fig. 1Sc). Strain GZB is a motile, spore-forming, and facultative anaerobic bacterium, and displayed growth at a wide temperature range (10–50 °C) and pH range (4.0–9.0). The optimal temperature and pH were 37 °C and 7.2, respectively. The GZB strain exhibited ample growth at a salinity

range of 0−3.0% (Table S2). The scanning electron microscopy (SEM; ZEISS Ultra 55, Carl Zeiss, Germany) showed that strain GZB cells had different shapes, with both spore and vegetative structures. Most of the vegetative cells were rod-shaped, approximately 1 µm long and 0.5 µm in diameter (Fig. 1Sd).

The sequence analysis of the 16S rRNA gene revealed that the GZB strain shared more than 99% of its identity with other *Bacillus* strains, such as *Bacillus amyloliquefaciens* FZB42^T, *Bacillus velezensis* M75^T, *Bacillus subtilis* ATCC19217^T, and *Bacillus siamensis* KCTC13613^T. These were identified as a result of a BLAST search against the GenBank database (Altschul et al. 1990). Specifically, the GZB strain belongs to the family *Bacillaceae*, in the order *Bacillales* and class *Bacilli*, with the phylum *Firmicutes*. Furthermore, strain GZB was applied to construct a phylogenetic tree based on the 16S rRNA gene sequence analyses, along with sequences represented by the other members of the genus *Bacillus* (Fig. 1).

The phylogenetic analysis indicated that the GZB strain is closely related to the *Bacillus amyloliquefaciens* strains. The analysis was further clarified using an average nucleotide identity (ANI) assay of the GZB strain genome, using genomes of strictly phylogenetically related *Bacillus* species. The results indicated that the GZB strain showed a high degree of ANI similarity to *Bacillus amyloliquefaciens* strain FZB42 (98.5%), *Bacillus amyloliquefaciens* strain CC178 (98.5%), and *Bacillus amyloliquefaciens* strain UCMB5036 (98.3%). Similarly, strain GZB exhibited 97.7% and 94.4% ANI values with *Bacillus velezensis* M75 and *Bacillus siamensis* KCTC13613, respectively. However, it showed only 67.9% of ANI similarity with strain *Bacillus* sp. GZT (Table S3). In general, the ANI values between the genomes in the same species exceed 95% (Rodriguez and Konstantinidis 2014); this indicates that *Bacillus* sp. GZB is closely related to *Bacillus amyloliquefaciens*.

Growth and laccase activity assays of strain *Bacillus* sp. GZB in NSM broth

Bacillus sp. GZB exhibited laccase activity after the development of the spore stage in the NSM broth. The spore-associated laccase with a 63 kDa protein fragment showed significant action and stability under different physicochemical conditions. The laccase completely degraded BPA within the 30 h timeframe previously reported (Das et al. 2018). Here, laccase activity was characterized at different time intervals during spore growth in the NSM broth using both ABTS and SGZ as substrates. The reaction mixtures showed a green color complex with ABTS, and were pink with SGZ (Fig. 2a). The GZB strain developed a maximum number of spore cells $(3.17 \pm 0.29) \times 10^8$, after 5 days of incubation in NSM broth. The estimated laccase activities of the spore cells at the 5-day point were 0.219 and 0.127 µmol min^{-1} mL^{-1} using the substrates ABTS (at pH 3.0) and SGZ (at pH 6.8), respectively.

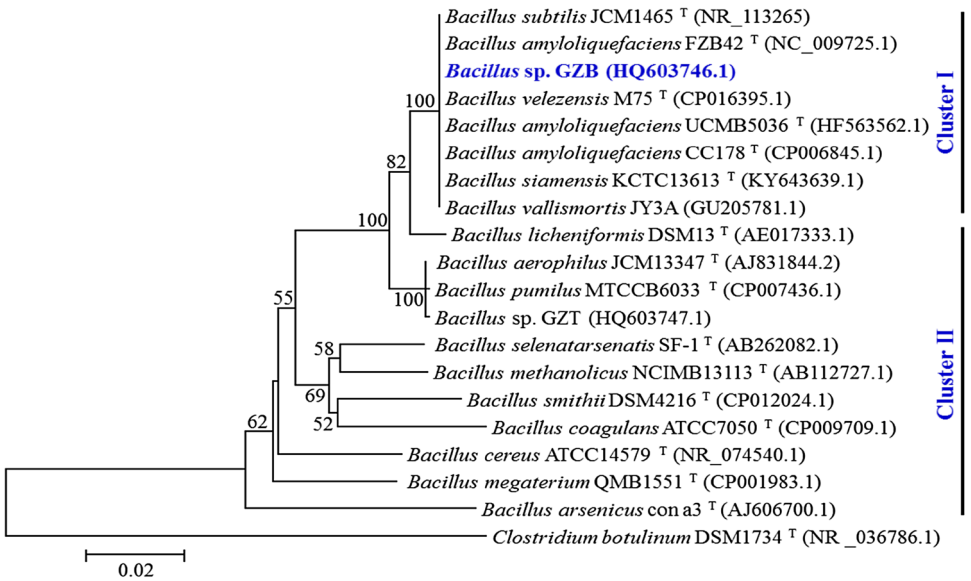

Fig. 1 Phylogenetic tree based on 16S rRNA gene sequences, highlighting that the position of *Bacillus* sp. GZB was relative to other species within the genus *Bacillus*. A tree was constructed using the neighbor-joining method incorporating Jukes–Cantor distance corrections. The type strains are labeled with superscript T, and the corresponding GenBank accession numbers are indicated parentheses. The cluster I indicated strain GZB assigned to the strains with NCBI blast match sequences and cluster II related to other relevant strains within genus *Bacillus*. The sequence of *Clostridium botulinum* DSM1734 used as an outgroup. One thousand bootstrap analyses were conducted, and bootstrap values > 50% were indicated at the nodes. Scale bar = 0.02 change per nucleotide position

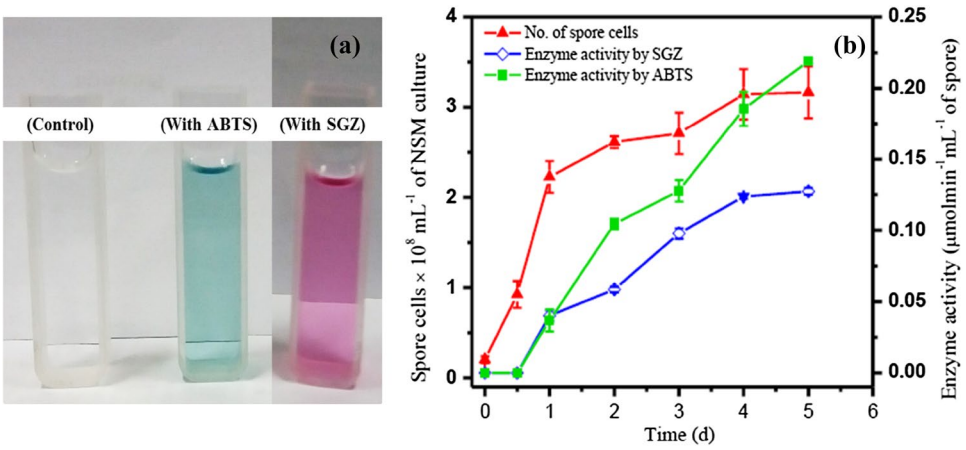

Fig. 2 **a** Laccase assays using ABTS and SGZ as the substrates. In the reactions, laccase produced a green color complex with ABTS and pink color complex with SGZ compared to enzyme-blank (control); **b** spore-cell development in NSM medium and laccase activity assays using ABTS and SGZ as the substrates

Figure 2b shows the time-dependent spore-cell development and laccase activity of strain GZB. Spore-associated laccase production is an essential feature of the *Bacillus* species (Held et al. 2005; Lu et al. 2012). Laccase is an oxidoreductase enzyme within the largest protein superfamily multi-copper oxidase, which catalyzes the oxidation of a wide variety of organic and inorganic substrates, coupled with reducing molecular oxygen to water (Dwivedi et al. 2011). Due to excellent oxidation potential, it has been extensively applied in many industrial and biotechnological processes. For example, it has been frequently used to biodegrade or transform different phenolics, polyphenolics, methoxy phenolics, aromatic amines, diamines, and other recalcitrant environmental pollutants (Telke et al. 2009; Lu et al. 2012; Hautphenne et al. 2016; Das et al. 2018).

Genome analysis

Genome properties

A total of 372 genome sequencing projects (finished and draft) associated with the *Bacillus* species have been submitted to the NCBI database (https://www.ncbi.nlm.nih.gov/genome/genomes). However, there have been limited genome analyses of *Bacillus* sp. isolated from adverse environments, such as the significantly polluted environment of the electronic-waste recycling site. This study analyzed the genome of *Bacillus* sp. GZB, a BPA-degrading strain and explored the genetic basis for BPA degradation. The genome of the GZB strain was comprised of a circular chromosome consisting of 4,077,007 bp, with an average G+C content of 46.2% (Fig. 3). Of the 4057 predicted genes, 3881 (95.66%) were protein-coding genes. There were 105 predicted RNA genes including 82 tRNAs, 18 rRNAs (9×5S, 7×16S, and 2×23S) and five ncRNA genes. The genome contained a total of 71 pseudogenes and no evidence of a CRISPR repeat. Table 1 summarizes the genome properties and the GZB strain statistics. The RAST database analysis of the GZB strain genome found that the closest neighbors were the *Bacillus amyloliquefaciens* LL3 (score, 503), *Bacillus amyloliquefaciens* FZB42 (score, 448), and *Bacillus subtilis* QB928 (score, 348).

COG analysis

The core gene functions of the *Bacillus* sp. GZB were estimated using a COG analysis. A total of 3881 protein-coding sequences (CDSs) in the GZB genome were identified; of these, 2877 (70.91%) protein-coding genes were assigned to COG functional categories. However, most genes in the GZB strain demonstrated a common subsystem structure associated with different functions (Table 2 and S6). We found high abundances of genes in the genome for transcription (K, 6.36%), followed by amino acid transport and metabolism (E, 6.11%); carbohydrate transport and metabolism (G, 5.46%); cell wall/membrane/envelope biogenesis (M, 4.38%); energy production and conversion (C, 4.20%); translation, ribosomal structure, and biogenesis (J, 4.17%); inorganic ion transport and metabolism (P, 4.10%); replication, recombination, and repair (L, 3.20%); signal transduction mechanisms (T, 3.14%); coenzyme transport and metabolism (H, 3.07%); post-translational modification, protein turnover, chaperones (O, 2.24%); lipid transport and metabolism (I, 2.19%); secondary metabolites biosynthesis, transport and catabolism (Q, 2.19%); nucleotide transport and metabolism (F, 2.06%); defense mechanisms (V, 1.37%); intracellular trafficking, secretion, and vesicular transport (U, 1.19%); and cell cycle control, cell division, chromosome partitioning (D, 0.93%), and cell motility (N, 0.59%).

Some genes were also recognized as having unknown functions (S, 7.70%); other genes were applied to predict only general functions (R, 9.48%). The COG categories were used to distinguish between different functional characteristics of strain GZB and other *Bacillus* species (Fig. 4). The COG analysis suggested that the GZB strain contained genes that may have evolved from harsh environmental conditions. The

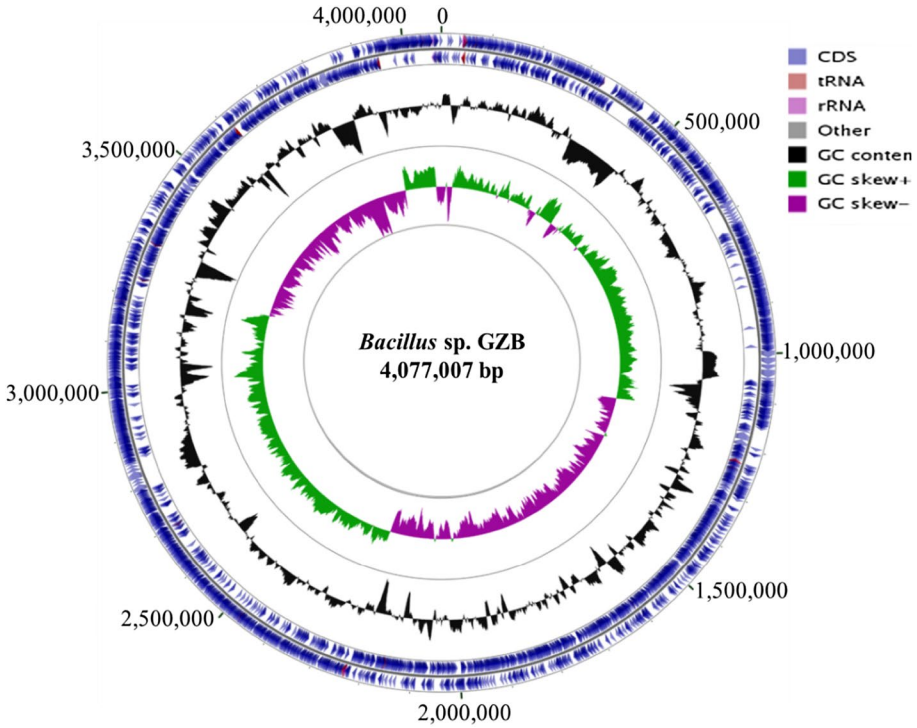

Fig. 3 Graphical circular map of the *Bacillus* sp. GZB genome displayed relevant genome features. Circles representing the following (from center to outside): (G+C) skew [green, positive (G+C) skew; purple, negative (G+C) skew]; (G+C) content (black); and coding DNA sequence (CDS, blue). The figure was build using CG viewer server

Table 1 Summary of the genome of *Bacillus* sp. GZB

Attribute	Value	% of total
Genome size (bp)	4,077,007	100
DNA coding (bp)	3,646,365	89.44
DNA (G+C) (bp)	1,883,578	46.2
Total contigs	23	100
Total Genes	4057	100
Protein-coding genes	3881	95.66
RNA genes	105	2.59
Pseudogenes	71	1.75
Genes assigned to COGs	2877	70.91
Genes with Pfam	3205	79.0
Genes with signal peptides	239	5.89
Genes with transmembrane helices	893	22.01
CRISPR repeats	0	0

**COGs* cluster of orthologous groups

external environment and several ecological factors are directly correlated with strain survival in the harsh conditions; these conditions influenced changes in the bacteria's genetic structure (Madigan and Martinko 1997; Yang et al. 2018).

Genetic mechanisms of BPA degradation and pathways prediction

As noted previously, *Bacillus* sp. GZB can degrade BPA, making it a unique source of carbon and energy (Li et al. 2012). BPA degradation occurs due to contribution and coordination of many protein-coding genes present in the GZB genome. We identified the gene-encoding proteins that are primarily involved in degradation of BPA and other xenobiotic or aromatic compounds. Based on the genome analysis and previous reports, we designed four probable BPA degradation pathways for strain GZB (Fig. 5). In this genome analysis, we characterized one of the most important enzyme complexes for BPA degradation: a cytochrome p450 monooxygenase system. This system consists of cytochrome p450, ferredoxin, and ferredoxin reductase, as reported by Sakai et al. (2008). Initial BPA hydroxylation at the methyl group and the quaternary carbon begins with the cooperation of cytochrome P450, ferredoxin, and ferredoxin reductase enzymes complex. This results in the catalysis of two different intermediates: 1,2-bis(4-hydroxyphenyl)-2-propanol and 2,2-bis(4-hydroxyphenyl)-1-propanol (Sasaki et al. 2008). In this study, we found five copies of genes encoded for cytochrome P450 monooxygenase (BXO87_13955, BXO87_02590, BXO87_04980, BXO87_05585, and BXO87_01290) and one copy of both ferredoxin (BXO87_14345) and ferredoxin reductase (BXO87_06325) (Tables S4 and S5).

The amino acid sequence alignment data associated with cytochrome p450 and ferredoxin from GZB showed significant homologies with the cytochrome p450 monooxygenase and ferredoxin sequences in *Bacillus* sp., and *Sphingomonas bisphenolicum* AO1. These are known to be BPA degraders (Sasaki et al. 2005). Alignment data (Fig. S2) further

Table 2 Number of genes associated with general COG functional categories

Code	Value	% age	Description
J	162	4.17	Translation, ribosomal structure, and biogenesis
A	0	0	RNA processing and modification
K	247	6.36	Transcription
L	124	3.20	Replication, recombination, and repair
B	0	0	Chromatin structure and dynamics
D	36	0.93	Cell cycle control, cell division, chromosome partitioning
V	53	1.37	Defense mechanisms
T	122	3.14	Signal transduction mechanisms
M	170	4.38	Cell wall/membrane biogenesis
N	23	0.59	Cell motility
U	46	1.19	Intracellular trafficking, secretion, and vesicular transport
O	87	2.24	Posttranslational modification, protein turnover, chaperones
C	163	4.20	Energy production and conversion
G	212	5.46	Carbohydrate transport and metabolism
E	237	6.11	Amino acid transport and metabolism
F	80	2.06	Nucleotide transport and metabolism
H	119	3.07	Coenzyme transport and metabolism
I	85	2.19	Lipid transport and metabolism
P	159	4.10	Inorganic ion transport and metabolism
Q	85	2.19	Secondary metabolites biosynthesis, transport, and catabolism
R	368	9.48	General function prediction only
S	299	7.70	Function unknown
-	1004	25.87	Not in COGs

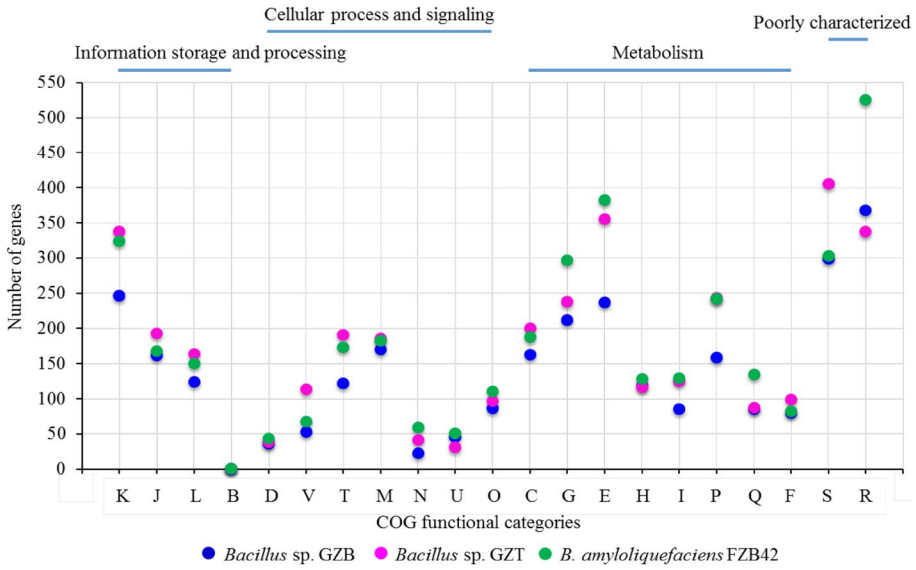

Fig. 4 Functional annotations of genes within *Bacillus* species. Four functional groups were categorized based on the function of each protein-coding gene under the COG assignments. The colors of these circles show different *Bacillus* species

suggested that the cytochrome p450 and ferredoxin from GZB may have similar functions as the *Bacillus* sp., *Rhodococcus* sp. P14, and *Sphingomonas bisphenolicum* AO1. Moreover, the sequence alignment had a heme-binding region with a strictly conserved cysteine residue in all cytochrome p450, at a position of 334 with the consensus sequence of FGXGXHXCXG region. These anchor the heme at the active site of cytochrome p450 (Degtyarenko 1995). In addition, the oxygen-binding consensus regions with two conserved threonine residues at the positions 230 and 231 play a critical role in the monooxygenation reactions (Imai et al. 1989).

The amino acid sequences of ferredoxin are similar to the conserved cysteine residues, which are mainly associated

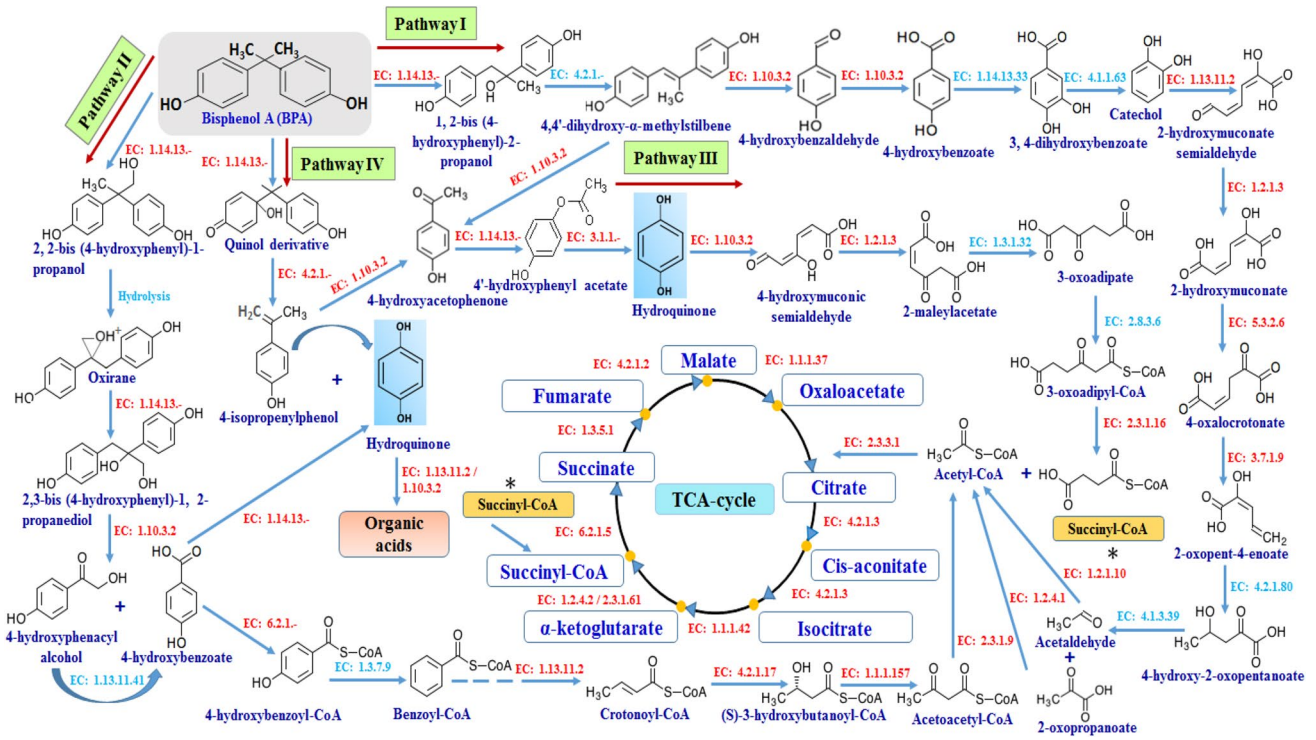

Fig. 5 Proposed degradation pathways of BPA based on the genome analysis of *Bacillus* sp. GZB. The Enzyme Commission (EC) numbers highlighted in red were proteins identified from the genome of GZB, whereas highlighted in blue obtained from the reference pathways of the KEGG database

with electron transport to a corresponding cytochrome P450 with two Fe ions (Sasaki et al. 2005). In addition, the GZB strain was capable of producing a spore-laccase enzyme, which was imitated by a gene encoding for copper binding oxidoreductase enzyme (BXO87_19270) (Table S4). It is primarily a multi-copper oxidase (laccase), which can be applied to different biotechnological and industrial processes. Recently, laccases from *Bacillus* species were studied to assess the biotransformation of many phenols and synthetic dyes (Held et al. 2005). Previously, we isolated the spore-laccase from the GZB strain and applied it to degrade BPA (Das et al. 2018). In this genome context, the presence of cytochrome P450, ferredoxin, ferredoxin reductase, and laccase genes sparked further interest in analyzing other relevant genes involved in different steps of BPA degradation pathways. All the genes and their respective functions, the Enzyme Commission (EC) numbers, and expected isolation sources are described in Table S4.

In pathway I, BPA degradation evolves by forming a major hydroxylation product 1,2-bis(4-hydroxyphenyl)-2-propanol through the enzyme complex cytochrome p450 (EC 1.14.13.-), ferredoxin, and ferredoxin reductase (EC 1.18.1.2). It is then further dehydrated to 4,4′-dihydroxy-alpha-methylstilbene. An oxidative cleavage on 4,4′-dihydroxy-alpha-methylstilbene produces two intermediates of 4-hydroxybenzaldehyde and 4-hydroxyacetophenone. The reactions may be achieved by an oxidoreductase enzyme such as laccase (EC 1.10.3.2) (Das et al. 2018; Zhang et al. 2013). Further oxidation of 4-hydroxybenzaldehyde makes 4-hydroxybenzoic acid, which is a primary end product from most BPA degradation pathways.

A previous study identified a spore-laccase from *Bacillus* sp. GZB, which was applied to degrade BPA. During BPA degradation, we detected both benzaldehyde and benzoic acid derivatives, followed by several organic acids as the intermediates (Das et al. 2018). Here, we identified an extradiol dioxygenase (EC 1.13.11.2) (BXO87_10735), which may react on catechol, a possible transformation product of 4-hydroxybenzoic acid. This further indicates a *meta*-ring cleavage degradation pathway. Furthermore, this study also identified several other enzymes present in the GZB genome and illustrated their possible roles during BPA degradation. These enzymes are aldehyde dehydrogenase (EC 1.2.1.3) (BXO87_10215 and BXO87_18805), 4-oxalocrotonate tautomerase (EC 5.3.2.6) (BXO87_09140), and pyruvate dehydrogenase (EC 1.2.4.1) (BXO87_03660 and BXO87_18430). Together, these result in the formation of final products, including pyruvate and acetaldehyde, which are further processed in the TCA cycle (Fig. 5).

Pathway II of BPA degradation starts with the hydroxylation product of 2,2-bis(4-hydroxyphenyl)-1-propanol generated by the enzymes complex; the product is further transformed to 2,3-bis(4-hydroxyphenyl)-1,2-propanediol either by the monooxygenase (EC 1.14.13.-) (BXO87_07120) or by cytochrome p450 enzymes complex (Sakai et al. 2005). Concurrently, 2,2-bis(4-hydroxyphenyl)-1-propanol are oxidized and rearranges to form oxirane; further hydrolysis of oxirane produces 2,3-bis(4-hydroxyphenyl)-1,2-propanediol (Spivack et al. 1994). Previous studies reported that organic compounds were rearranged or re-aromatized by short-chain dehydrogenase and cytochrome P450 monooxygenase (Li et al. 2016).

This study found four copies of short-chain dehydrogenase (BXO87_10185, BXO87_06795, BXO87_06195, and BXO87_19000) and five copies of cytochrome P450 monooxygenase. Furthermore, an oxidative cleavage occurs on 2,3-bis(4-hydroxyphenyl)-1,2-propanediol by either laccase or other oxidoreductases (Table S4). This converts 4-hydroxyphenacyl alcohol and 4-hydroxybenzoate. A previous study used GC–MS analysis to identify both these compounds as BPA-degrading intermediates (Li et al. 2012; Das et al. 2018). Furthermore, 4-hydroxybenzoate is converted to 4-hydroxybenzoyl-CoA by acyl-CoA synthetase (EC 6.2.1.-) (BXO87_01010); then, a dehydroxylation reaction occurs on 4-hydroxybenzoyl-CoA to form benzoyl-CoA (Glockler et al. 1989).

A similar kind of BPA degradation pathway for *Bacillus thuringiensis* was noted by Li et al. (2018). Benzoyl-CoA was successively cleaved by aromatic ring-cleaving dioxygenase (EC 1.13.11.2) to produce crotonoyl-CoA. This genome study identified three copies of ring-cleaving dioxygenase (BXO87_02800, BXO87_15460, and BXO87_17705). The crotonoyl-CoA is further converted to (S)-3-hydroxybutanoyl-CoA by the enoyl-CoA hydratase (EC 4.2.1.17). This study found five copies (BXO87_13965, BXO87_04945, BXO87_05635, BXO87_00765, and BXO87_12075) of enoyl-CoA hydratase encoding gene in the GZB genome. Then, (*S*)-3-hydroxybutanoyl-CoA is oxidized to acetoacetyl-CoA by the 3-hydroxybutyryl-CoA dehydrogenase (EC 1.1.1.157) (BXO87_13705). Furthermore, acetoacetyl-CoA is cleaved by acetyl-CoA acetyltransferase (EC 2.3.1.9) (BXO87_06705 BXO87_01005, and BXO87_13700) to transform acetyl-CoA, which is metabolized by the TCA cycle (Fig. 5).

In pathway III, 4-hydroxyacetophenone, an oxidative intermediate of 4,4′-dihydroxy-alpha-methylstilbene, is converted to 4-hydroxyphenyl acetate by the enzyme monooxygenase (EC 1.14.13.-). Then, 4-hydroxyphenyl acetate is transformed to hydroquinone by several esterases. This study detected four copies of esterase in the GZB genome (EC 3.1.1-) (BXO87_17630, BXO87_18115, BXO87_01320, and BXO87_06110), and two copies of carboxylesterase (BXO87_15465, and BXO87_07210) (EC 3.1.1.1). These may act on 4-hydroxyphenyl acetate to produce hydroquinone. Hydroquinone is further oxidized by either laccase or by several oxidoreductases to form 4-hydroxymuconic semialdehyde. In addition, this genome study found three copies of aldehyde dehydrogenase (EC 1.2.1.3) (BXO87_10215 and BXO87_18805, and BXO87_11570). These may catalyze the oxidation of the aldehyde byproduct to a corresponding acidic derivative such as 2-maleylacetate (Sophos and Vasiliou 2003). The 2-maleylacetate then degrades to produce 3-oxoadipyl-CoA. The 3-oxoadipyl-CoA is further catalyzed to succinyl-CoA and acetyl-CoA by the 3-ketoacyl-CoA thiolase (EC 2.3.1.16) (BXO87_01005), and is further metabolized through the TCA cycle (Fig. 5).

In pathway IV, BPA is degraded by forming a quinol intermediate through the addition of molecular oxygen into BPA by the NADPH and FAD-aided monooxygenase; this approach was previously reported in *Sphingomonas* sp. TTNP3. After that, the C−C bond is cleaved between the phenolic moiety and isopropyl group of BPA. This results in the formation of 4-isoprophenylphenol and hydroquinone (Fig. 5) (Kolvenbach et al. 2007). Our previous study detected this hydroquinone, a major intermediate of BPA degraded by laccase (Das et al. 2018). Telke et al. (2009) also reported that hydroquinone was an essential BPA intermediate when treated with laccase derived from *Pseudomonas* sp. LBC1. Further oxidation of hydroquinone by several ring-cleaving dioxygenases (EC 1.13.11.2), or oxidoreductases like laccase or other oxidoreductases present in the GZB genome, produces organic acids. These acids include 2-hydroxypropanoic acid, 2-methylbutanoic acid, and 3-methylbutanoic acid (Das et al. 2018). The GZB strain contains a complete TCA-cycle pathway; this may help process the resulting intermediates from BPA degradation (Fig. S3 and Table S4).

Genes of *Bacillus* sp. GZB are involved in degrading other xenobiotic compounds

The GZB genome contained several protein-coding genes known to be involved in different pathways for degrading hydrocarbons and many xenobiotic compounds. Many genes in the GZB strain are encoded for hydrocarbon and/or xenobiotic compounds degradation; these include alkanesulfonate monooxygenase (BXO87_00270); biphenyl 2, 3-dioxygenase (BXO87_11265); nitroalkane dioxygenase (BXO87_05890); aromatic ring hydroxylase (BXO87_17830); extradiol dioxygenase (BXO87_10735); 1,4-dihydroxy-2-naphthoyl-CoA synthase (BXO87_11040); 4-hydroxy-3-methylbut-2-enyl diphosphate reductase (BXO87_13205); 2-amino-3-carboxymuconate-6-semialdehyde decarboxylase (BXO87_00250); *p*-nitrobenzyl esterase (BXO87_07610); 2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylic-acid

synthase (BXO87_11030); 2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase (BXO87_11035); enoyl-CoA hydratase (BXO87_13965, BXO87_04945, BXO87_05635, BXO87_00765, and BXO87_12075); and butanediol dehydrogenase (BXO87_19290). (Tables S4 and S5) (Li et al. 2016; McLeod et al. 2006; Giebel et al. 2016). A previous study found that many polyaromatic hydrocarbons (PAHs), such as pyrene and phenanthrene, are aerobically degraded through cleavage on the benzene ring by dioxygenase or monooxygenase (Jouanneau et al. 2016; Kim et al. 2007; Peng et al. 2015).

In the GZB genome, we found three copies of aromatic ring-cleaving dioxygenase and one copy of monooxygenase; these may be degraded different PAHs. However, many researchers have found that the anaerobic degradation of PAHs occurred through the carboxylation of compounds by carboxylase (Mouttaki et al. 2012). In anaerobic environments, acetoacetate decarboxylase played an important role in degrading acetoacetate compounds to acetone and carbon dioxide. We identified two copies of carboxylase coding genes (BXO87_01320, BXO87_17630) and one copy of an acetoacetate decarboxylase (BXO87_06765) gene in the GZB genome. However, the alcohol and aldehyde intermediates, originally derived from different hydrocarbon degradations, were further converted to acidic byproducts (McLeod et al. 2006).

The previous reports suggested that alcohol dehydrogenase and aldehyde dehydrogenases were the main catalyzers transforming alcohols to aldehydes and aldehydes to acidic derivatives (Sophos and Vasiliou 2003). In this study, we identified six copies of alcohol dehydrogenase (BXO87_15820, BXO87_05450, BXO87_16415, BXO87_17805, BXO87_18130, and BXO87_05790), two NADH-dependent alcohol dehydrogenase (BXO87_05960 and BXO87_05965), one NADP-dependent alcohol dehydrogenase (BXO87_19345), one zinc-dependent alcohol dehydrogenase (BXO87_05205), and two copies of aldehyde dehydrogenases (BXO87_10215 and BXO87_18805) (Tables S4 and S5). Most hydrocarbon and xenobiotic compounds were degraded through benzoic acid as the major intermediate (Li et al. 2012; Das et al. 2018; Zhang et al. 2013).

The GZB genome contained different genes encoded for benzoic acid or benzoic acid derivatives degradations, including 2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase (BXO87_06270); 2,3-dihydroxybenzoate–AMP ligase (BXO87_06260); acetyl-CoA acetyltransferase (BXO87_06705 and BXO87_13700); 3-ketoacyl-CoA thiolase (BXO87_01005); *O*-succinyl benzoate synthase (BXO87_11050); 2-succinyl benzoate-CoA ligase (BXO87_11045); 4-carboxymuconolactone decarboxylase (BXO87_10645); 4-oxalocrotonate tautomerase (BXO87_09140); and 3-hydroxybutyryl-CoA dehydrogenase (BXO87_13705) (Tables S4 and S5). The genes acylphosphatase (BXO87_18670) and amidase (BXO87_08195) have been identified as involved in aminobenzoate degradation. They are also engaged in the degradation of styrene and 2,4-dichloroalkane, respectively (Sudtachat et al. 2009).

A different study found that the cytochrome P450 monooxygenase, which includes a ferredoxin and ferredoxin reductase system, may catalyze many other reactions. These reactions could include the hydroxylation, epoxidation, sulfoxidation, or dealkylation of a wide range of hydrocarbons and xenobiotic compounds, including drugs, perfumes, and pesticides (Sasaki et al. 2005; Luo et al. 2016). Most importantly, cytochrome P450 monooxygenase with ferredoxin and ferredoxin reductase system played a vital role during *n*-alkanes oxidation (Hong et al. 2017). A previous study found that *n*-alkane oxidation produced fatty acids as major intermediates; these were further processed through β-oxidation to generate acetyl-CoA (Rojo 2009). The GZB genome contained five copies of cytochrome P450 monooxygenase gene, one copy of both ferredoxin and ferredoxin, and an intact β-oxidation pathway.

We found that the GZB genome also contained one copy of haloacid dehalogenase (BXO87_11885) belonging to the superfamily of hydrolases. These hydrolases may degrade different chloroalkanes and chloroalkenes, along with chlorocyclohexanes and chlorobenzenes. As such, the presence of haloacid dehalogenase was resolved to degrade the highly toxic halogenated organic compounds from the environment (Wang et al. 2015; Yang et al. 2018). Previously, *Bacillus* sp. GZT, isolated from the same sample, was involved in degrading 2,4,6-tribromophenol by synthesizing a novel dehalogenase (Liang et al. 2017). Furthermore, the GZB genome contained a copy of copper oxidase (BXO87_19270), recognized as laccase. This laccase was involved in many oxidation processes associated with aromatic compounds, including pulp delignification, biobleaching, detoxification of phenolics from industrial effluents, and textile dye decolorization (Lu et al. 2012; Hautphenne et al. 2016; Zhang et al. 2013).

The RAST subsystem analysis revealed several xenobiotic compounds that the strain GZB degraded and metabolized. The gene-encoding proteins in the GZB genome were primarily responsible for degrading and metabolizing different compounds; they bloomed significantly in KEGG pathways. Some of the essential pathways for strain GZB include benzoate, fluorobenzoate, 2,4-dichlorobenzoate degradation; naphthalene, anthracene, and fluorene degradation; 1,4-dichlorobenzene, atrazine, biphenyl degradation; and toluene, xylene, and ethylbenzene degradation. Some metabolic pathways for carbazole, 1,2-dichloroethane, styrene, gamma-hexachlorocyclohexane, and 1,1,1-trichloro-2,2-bis (4-chlorophenyl) ethane (DDT) also flourished (Fig. S4). This highlights the importance of this study in assessing

the ability of strain GZB to degrade and detoxify several environmental contaminants.

Regulation systems involved in BPA degradation

BPA accessibility to cells as a unique source of carbon and energy facilitated the BPA degradation process, and it could occur through resistance, transport, and transformation associated with cellular regulation, signaling, and metabolic functions (Li et al. 2018). Several genes were involved in these regulatory functions of BPA degradation. The KEGG analysis identified 107 genes as stress response genes; these genes helped cells survive when first exposed to BPA. At this phase, cell motility and chemotaxis regulating genes played a vital role in BPA reorganization and mobilization. A total of 88 genes were involved in this cell motility and chemotaxis regulation (Fig. S5). The transport systems of strain GZB included 1023 genes (26.37% of total CDS) recognized by COG functional categories, which helped transport aromatic compounds, amino acids, carbohydrates, lipids, nucleotides, coenzymes, metabolites, and inorganic ions (Tables 2 and S6).

Most importantly, 61 ABC transporter ATP-binding proteins and 68 ABC transporter permease which were mostly responsible for importing BPA and several other organic or inorganic compounds into cells. Besides, 62 encoded genes were Major Facilitator Superfamily (MFS), engaged in resisting multidrug and solute transport. The GZB genome, including genes encoding ABC phosphate (8), sulfate (1), nitrate and nitrite (4), ABC iron (9), ABC molybdate (2), ABC zinc (1), and potassium (6) transporter systems, played a significant role in nutrient uptake and in the detoxification of harsh compounds in different environments (Pal et al. 2017; Yang et al. 2018).

The COG analysis also demonstrated that 247 CDSs in the GZB genome belong to the transcription category (K, 6.36%). Of these, 17 proteins are LysR-type transcriptional regulators (LTTRs) (COG0583); these proteins have been reported to be regulatory genes involved in aromatic compound catabolism (Binnewies et al. 2006). Another 15 proteins belong to TetR/AcrR family regulators (COG1309) and played a regulatory role over the genes involved in alkanes catabolism (Yang et al. 2018). Twenty-five other proteins belonged to the multiple antibiotic resistance (MarR) family regulators (COG1846), and another eight proteins were recognized as MerR family transcriptional regulator (COG0789). Nine proteins under GntR group of transcriptional regulators (COG1802) were also involved with some aromatic compound degradation (Gerischer et al. 2002). Two proteins belonged to AraC family regulators (COG2207), and one protein was identified as the IclR family of the transcriptional regulator (COG1414). This protein may control the expression of dioxygenase and cytochrome p450 genes during hydrocarbon degradation (Yang et al. 2018).

Recent studies have suggested that some upregulated ribosomal proteins, including RplP, RpsS, RplO, RpsP, RplF, and InfA and down-regulated proteins Tkt and RpoA in the *Bacillus thuringiensis* strain, directly interacted during BPA degradation (Table S6). Results also showed that the upregulated proteins, e.g., RplP, RpsS, RplO, RpsP, and RplF, were increasingly expressed under BPA stress and assisted BPA degradation. The upregulation and down-regulation of these proteins also elucidated the regulation of the expression of different functional genes after BPA degradation (Li et al. 2018). In the GZB genome, we found one copy of each of the upregulated proteins RplP (COG0197), RpsS (COG0185), RpsP (COG0228), RplO (COG0200), InfA (COG0361), and RplF (COG0097) and the down-regulated proteins Tkt (COG0021) and RpoA (COG0202). These benefit BPA degradation in *Bacillus* sp. GZB, like *Bacillus thuringiensis*.

However, the expression of some ribonuclease proteins was affected by BPA, which directly interfered with the translation process. This meant that BPA posed a risk to protein folding and translation (Liang and Deutscher 2013). Thiol-disulfide oxidoreductase is a crucial enzyme for protein folding and function (Feng and Coulombe 2015). The enzyme promotes enzyme activity and stability (Lu et al. 2014); its regulation during BPA degradation influenced enzyme activity, cellular biological process, and cellular components (Li et al. 2018). The GZB genome contained two copies of thiol-disulfide oxidoreductase (BXO87_06030 and BXO87_14305), which played a vital role in protein stability and functions during BPA degradation. Another essential protein was DsbA, containing a thioredoxin domain. This domain is useful in the oxidative folding of the bacterial protein in the natural environment (Crow et al. 2009). In our genome study, we obtained five copies of thioredoxin genes (BXO87_11425, BXO87_12095, BXO87_16730, BXO87_17695, and BXO87_06675) (Table S4); these may have added some benefits to strains that could be applied for remediation in the natural environments. In this study, the functions of many proteins were not identified. Further research is needed to confirm the precise role of the proteins and their regulatory mechanisms during BPA degradation.

Sporulation and CotA expression of *Bacillus* sp. GZB

Most of the *Bacillus* species produce spores under stressful physical and chemical conditions (Nicholson et al. 2000), including *Bacillus* sp. GZB (Das et al. 2018). Sporulation is a crucial step for strain GZB, because during this condition, cells express CotA, a spore-associated protein that can oxidize many aromatic compounds (Das et al. 2018). Spore resistance under certain limitations and stable under different

adverse situations because of an assemblage of several protective structures: cortex, coat, and exosporium (Bressuire-Isoard et al. 2016). The genome analysis identified 132 protein-coding genes in the GZB genome (Table S7). Of these, 11 were alpha/beta/gamma-type small, acid-soluble spore proteins (SasP) that bind to DNA in spores and protect them from chemical and enzymatic degradation under stressful conditions.

A different study found that the SasP-1 (major beta-type SasP) is down-regulated during BPA degradation. Risk occurs on the spore DNA due to cleavage. This suggests that the SasP-1 can be a potential biomarker in pollutant stress (Li et al. 2018). Other genes also play a pivotal role during GZB spore development; these genes include spore coat proteins (16), several sporulation proteins involved in different stages of spore development (22), and the YjcZ family sporulation protein (7). The assembly of the coat and exosporium occurs in the cell under the governance of RNA polymerase sigma factors (McKenney et al. 2013). The GZB genome contains five copies of RNA polymerase sigma factors, which are involved in the synthesis of several proteins. The CotE is a morphogenetic protein, and acts as a critical determinant in assembling the spore outer coat under the control of a sigma factor (σ^E) (Bressuire-Isoard et al. 2016).

The GZB genome was found to carry one copy of the CotE gene. In addition, the CotA protein is a significant component of the outer coat layer and the CotA assembly is significantly monitored by the CotE protein (Little and Driks 2001). Many researchers have tried to clarify the exact function of CotA within the spore coat; however, it is not fully understood (Bressuire-Isoard et al. 2016; Enguita et al. 2003). Most notably, the assembly of CotA within the spore coat is essential for the full complement of spore resistance properties (Enguita et al. 2003). Different types of spore germination proteins (15) were also found in the GZB genome; these proteins help germinate the spore to transition to a vegetative phase when the cells are in favorable conditions.

Genomic islands and horizontal gene transfer

The strain *Bacillus* sp. GZB was isolated from an electronic-waste recycling site. The presence of the strain in that environment indicated its ability to survive under extreme pollutant-driven environmental stresses. Ecological stresses are critical factors that impact horizontal gene transfer (HGT) mechanisms. HGT is a universal microbial process that frequently occurs among microbes in the environment. Genomic islands (GIs) are clusters of genes present in bacterial genomes that have been acquired by horizontal origins (Hacker and Kaper 2000). This study identified several genes obtained through HGT from different bacterial species. A total of 20 GIs were identified in the genome of *Bacillus* sp. GZB using the integrated mode of Island Viewer 4 (Bertelli et al. 2017). Figure S6 shows the localization of predicted GIs. Of 314,216 bp GIs encoded by 338 genes, 84 genes were functioning predicted and 254 were hypothetical proteins (Table S8).

Most of the highlighted proteins present in GIs are mainly involved in active transport, spore coat synthesis, cell division, flagella formation, and DNA replication and transcriptional regulation. Some proteins are involved in both BPA and hydrocarbon degradation; these proteins include enoyl-CoA hydratase, acyl carrier protein, polyketide synthase, short-chain dehydrogenase, and aldehyde oxidase. The presence of flagellin protein and several phage-associated proteins could indicate effective HGT or genetic alterations using bacteriophages in the oligotrophic environment. A large number of proteins were discovered that have unknown functions; these functions may be involved in different cellular and biological processes.

Comparative study of *Bacillus* sp. GZB with other genomes

To conduct a comparative study, we used *Bacillus* sp. GZT (LVVJ00000000), a 2,4,6-tribromophenol degrading strain (Liang et al. 2016) previously isolated from the same sample, and a *Bacillus amyloliquefaciens* FZB42^T (NC_009725), a phylogenetically closed model bacterium with *Bacillus* sp. GZB isolated from a non-polluted environmental sample (Chen et al. 2007). Table 3 shows the data comparing the GZB genome with other *Bacillus* species. The strains *Bacillus* sp. GZB and *Bacillus* sp. GZT were obtained from the same sample; however, the GZB strain showed a higher G+C content than GZT (35.1%). In contrast, *B. amyloliquefaciens* FZB42^T was quite similar to 46.0% of G+C.

A recent study reported that G+C content did not vary more than 1% within the species (Meier-Kolthoff et al. 2014); this indicates that strain GZB is close to *B. amyloliquefaciens* FZB42^T. Moreover, the protein-coding sequences in GZB genome exceeded those of *B. amyloliquefaciens* FZB42^T (3881 vs. 3693), but were smaller than *Bacillus* sp. GZT (3881 vs. 5252). A total of 105 RNA genes was found in the GZB genome. This was high compared to *Bacillus* sp. GZT (52; RNA genes), but was close to *B. amyloliquefaciens* FZB42^T (121; RNA genes).

The COG analysis indicated that the abundance of proteins responsible for carbohydrate transport and metabolism in both strains of *Bacillus* sp. GZB (G; 212) and *Bacillus* sp. GZT (G; 238) were lower than *B. amyloliquefaciens* FZB42^T (G; 297). This indicated that there was a lower level of available carbohydrates in the electronic-waste sludge sample from which the strains were isolated. The average number of GZB genes responsible for defense mechanisms; the transport and metabolism of carbohydrates, amino acids, lipids, and inorganic ions; secondary metabolites synthesis; and

Table 3 Genomic comparison of *Bacillus* sp. GZB with the genomes of *Bacillus* sp. GZT, isolated from the same sample and *B. amyloliquefaciens* FZB42 was closely related strain observed in the NCBI blast match as well as RAST analysis

Attribute	*Bacillus* sp. GZB	*Bacillus* sp. GZT	*B. amyloliquefaciens* FZB42
Genome size (bp)	4,077,007	5,188,788	3,918,589
GC %	46.2	35.1	46.4
Total genes	4057	5431	3892
Protein-coding genes	3881	5252	3687
RNA genes	105	52	121
rRNA genes	18	11	29
tRNA genes	82	36	89
ncRNAs	5	5	4
Pseudo genes	71	127	84
Genes in COGs	2877	3605	3305
Accession no.	MTQG00000000	LVVJ00000000	NC_009725
Isolation source	Electronic-waste sludge	Electronic-waste sludge	Plant root
References	This study	Liang et al. (2016)	Chen et al. (2007)

transcription were nearly similar to *Bacillus* sp. GZT. This indicated that both strains were metabolically active and could survive in harsh environmental conditions.

Figure 4 shows that the distribution of genes into COG functional categories was not identical in all *Bacillus* genomes. The GZB genome contained a total of 299 genes (7.70% of the COGs), which were annotated with unknown functions. These genes may be elaborated with several functions, similar to other proteins in transport, metabolism, and regulation systems. Further study is needed to characterize the functions of these proteins in the GZB genome. However, the presence of large numbers of functional genes in the GZB genome indicates they can adapt and metabolize in the harsh environments. Moreover, the previous studies found that most reported genome sequences of BPA-degrading bacteria belonged to the genus *Sphingobium*. As such, we compared the genomic features of *Bacillus* sp. GZB with other BPA-degrading strains to clarify their basic genomic differences. All the data are shown in Table S9.

Conclusion

This study identified the overall ~ 4.08 Mb genome structure and genome-based functional characteristics of an environmentally relevant and biotechnologically important *Bacillus* sp. GZB. Information about the genome identified the genetic basis of several metabolic pathways for BPA degradation and many other cellular functions of the GZB strain. The GZB genome included a large number of protein-coding genes, including cytochrome P450 monooxygenase, ferredoxin, and ferredoxin reductase complex, laccase, and many others. These genes are involved in BPA and other xenobiotic compound degradation. The protein-coding genes in the GZB genome were particularly helpful in predicting BPA degradation pathways. The comparative genomics suggested that GZB is a disparate category within the *Bacillus* species.

This study is the first known systematic and detailed elucidation of BPA degradation pathways associated with the genus *Bacillus* using genome sequencing. The genomic data provide new insights into the BPA degradation mechanisms; these insights should help to design a future systematic and comprehensive metabolic pathway of the GZB strain for bioremediating diverse BPA-polluted environments.

Acknowledgements We gratefully acknowledge Sangon Biotech, Shanghai, China for genome analysis. This study was financially supported by the National Natural Science Foundation of China (41877363, 41425015 and 41373103) and the Science and Technology Program of Guangzhou, China (201704020185).

Compliance with ethical standards

Conflict of interest The authors declare that there is no conflict of affairs concerning the work published in this paper.

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Angiuoli SV, Gussman A, Klimke W, Cochrane G, Field D, Garrity G, Kodira CD, Kyrpides N, Madupu R, Markowitz V, Tatusova T, Thomson N, White O (2008) Toward an online repository of standard operating procedures (SOPs) for (Meta) genomic annotation. Omics J Integrative Bio 12:137–141

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppiq JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G (2000) Gene ontology: tool for the unification of biology. Nat Genet 25:25–29

Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil KL, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O,

Vonstein V, Wilke A, Zagnitko O (2008) The RAST server: rapid annotations using subsystems technology. BMC Genom 9:75–89

Bertelli C, Laird MR, Williams KP, Fraser S, Lau BY, Hoad G, Winsor GL, Brinkman FSL (2017) IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. Nucleic Acids Res 45:W30–W35

Binnewies TT, Motro Y, Hallin PF, Lund O, Dunn D, La T, Hampson DJ, Bellgard M, Wassenaar TM, Ussery DW (2006) Ten years of bacterial genome sequencing: comparative-genomics-based discoveries. Funct Integr Genomics 6:165–185

Bressuire-Isoard C, Bornard I, Henriques AO, Carlin F, Broussolle V (2016) Sporulation temperature reveals a requirement for CotE in the assembly of both the coat and exosporium layers of *Bacillus cereus* spores. Appl Environ Microbiol 82:232–243

Chen XH, Koumoutsi A, Scholz R, Eisenreich A, Schneider K, Heinemeyer I, Morgenstern B, Voss B, Hess WR, Reva O, Junge H, Voigt B, Jungblut PR, Vater J, Süssmuth R, Liesegang H, Strittmatter A, Gottschalk G, Borriss R (2007) Comparative analysis of the complete genome sequence of the plant growth-promoting bacterium *Bacillus amyloliquefaciens* FZB42. Nat Biotechnol 25:1007

Chen D, Kannan K, Tan H, Zheng Z, Feng YL, Wu Y, Widelka M (2016) Bisphenol analogs other than BPA: environmental occurrence, human exposure, and toxicity—a review. Environ Sci Technol 50:5438–5453

Crow A, Lewin A, Hecht O, Carlsson Moller M, Moore GR, Hederstedt L, Le Brun NE (2009) Crystal structure and biophysical properties of Bacillus subtilis BdbD. An oxidizing thiol:disulfide oxidoreductase containing a novel metal site. J Biol Chem 284:23719–23733

Das R, Kazy SK (2014) Microbial diversity, community composition and metabolic potential in hydrocarbon contaminated oily sludge: prospects for in situ bioremediation. Environ Sci Pollut Res 21:7369–7389

Das R, Li G, Mai B, An T (2018)) Spore cells from BPA degrading bacteria *Bacillus* sp. GZB displaying high laccase activity and stability for BPA degradation. Sci Total Environ 640–641:798–806

Degtyarenko KN (1995) Structural domains of P450-containing monooxygenase systems. Protein Eng 8:737–747

Dwivedi UN, Singh P, Pandey VP, Kumar A (2011) Structure–function relationship among bacterial, fungal and plant laccases. J Mol Catal B: Enzyme 68:117–128

Enguita FJ, Martins LO, Henriques AO, Carrondo MA (2003) Crystal structure of a bacterial endospore coat component. A laccase with enhanced thermostability properties. J Biol Chem 278:19416–19425

Feng X, Coulombe PA (2015) A role for disulfide bonding in keratin intermediate filament organization and dynamics in skin keratinocytes. J Cell Biol 209:59–72

Field D, Garrity G, Gray T, Morrison N, Selengut J, Sterk P, Tatusova T, Thomson N, Allen MJ, Angiuoli SV, Ashburner M, Axelrod N, Baldauf S, Ballard S, Boore J, Cochrane G, Cole J, Dawyndt P, De Vos P, dePamphilis C, Edwards R, Faruque N, Feldman R, Gilbert J, Gilna P, Glöckner FO, Goldstein P, Guralnick R, Haft D, Hancock D, Hermjakob H, Hertz-Fowler C, Hugenholtz P, Joint I, Kagan L, Kane M, Kennedy J, Kowalchuk G, Kottmann R, Kolker E, Kravitz S, Kyrpides N, Leebens-Mack J, Lewis SE, Li K, Lister AL, Lord P, Maltsev N, Markowitz V, Martiny J, Methe B, Mizrachi I, Moxon R, Nelson K, Parkhill J, Proctor L, White O, Sansone SA, Spiers A, Stevens R, Swift P, Taylor C, Tateno Y, Tett A, Turner S, Ussery D, Vaughan B, Ward N, Whetzel T, Gil IS, Wilson G, Wipat A (2008) The minimum information about a genome sequence (MIGS) specification. Nat Biotechnol 26:541–547

Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, Salazar GA, Tate J, Bateman A (2015) The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res 44:D279–D285

García-Díaz C, Ponce-Noyola MT, Esparza-García F, Rivera-Orduña F, Barrera-Cortés J (2013) PAH removal of high molecular weight by characterized bacterial strains from different organic sources. Int Biodeter Biodegr 85:311–322

Gerischer U (2002) Specific and global regulation of genes associated with the degradation of aromatic compounds in bacteria. J Mol Microbiol Biotechnol 4:111–121

Giebel HA, Klotz F, Voget S, Poehlein A, Grosser K, Teske A, Brinkhoff T (2016) Draft genome sequence of the marine *Rhodobacteraceae* strain O3.65, cultivated from oil-polluted seawater of the Deepwater Horizon oil spill. Stand Genomic Sci 11:81

Glockler R, Tschech A, Fuchs G (1989) Reductive dehydroxylation of 4-hydroxybenzoyl-CoA to benzoyl-CoA in a denitrifying, phenol-degrading *Pseudomonas* species. FEBS Lett 251:237–240

Grignard E, Lapenna S, Bremer S (2012) Weak estrogenic transcriptional activities of Bisphenol A and Bisphenol S. Toxicol In Vitro 26:727–731

Hacker J, Kaper JB (2000) Pathogenicity islands and the evolution of microbes. Ann Rev Microbiol 54:641–679

Hautphenne C, Penninckx M, Debaste F (2016) Product formation from phenolic compounds removal by laccases: a review. Environ Technol Innovation 5:250–266

Held C, Kandelbauer A, Schroeder M, Cavaco-Paulo A, Gübitz GM (2005) Biotransformation of phenolics with laccase containing bacterial spores. Environ Chem Lett 3:74–77

Hong YH, Ye CC, Zhou QZ, Wu XY, Yuan JP, Peng J, Deng H, Wang JH (2017) Genome sequencing reveals the potential of *Achromobacter* sp. HZ01 for bioremediation. Front Microbiol 7:1507

Huang YQ, Wong CKC, Zheng JS, Bouwman H, Barra R, Wahlström B, Neretin L, Wong MH (2012) Bisphenol A (BPA) in China: a review of sources, environmental levels, and potential human health impacts. Environ Int 42:91–99

Im J, Loffler FE (2016) Fate of bisphenol A in terrestrial and aquatic environments. Environ Sci Technol 50:8403–8416

Imai M, Shimada H, Watanabe Y, Matsushima-Hibiya Y, Makino R, Koga H, Horiuchi T, Ishimura Y (1989) Uncoupling of the cytochrome P-450cam monooxygenase reaction by a single mutation, threonine-252 to alanine or valine: possible role of the hydroxy amino acid in oxygen activation. Proc Natl Acad Sci USA 86:7823–7827

Jouanneau Y, Meyer C, Duraffourg N (2016) Dihydroxylation of four- and five-ring aromatic hydrocarbons by the naphthalene dioxygenase from *Sphingomonas* CHY-1. Appl Microbiol Biotechnol 100:1253–1263

Jukes TH, Cantor CR (1969) Evolution of protein molecules. In: Munro HN (ed) Mammalian protein metabolism. Academic Press., New York, pp 21–132

Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M (2016) KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res 44:D457–D462

Kim SJ, Kweon O, Jones RC, Freeman JP, Edmondson RD, Cerniglia CE (2007) Complete and integrated pyrene degradation pathway in *Mycobacterium vanbaalenii* PYR-1 based on systems biology. J Bacteriol 189:464–472

Kolvenbach B, Schlaich N, Raoui Z, Prell J, Zuhlke S, Schaffer A, Guengerich FP, Corvini PFX (2007) Degradation pathway of bisphenol A: Does *ipso* substitution apply to phenols containing a quaternary α-carbon structure in the *para* position? Appl Environ Microbiol 73:4776–4784

Krogh A, Larsson B, Von Heijne G, Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 305:567–580

Lagesen K, Hallin P, Rødland EA, Stærfeldt HH, Rognes T, Ussery DW (2007) RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res 35:3100–3108

Larsson K, Lindh CH, Jönsson BAG, Giovanoulis G, Bibi M, Bottai M, Bergström A, Berglund M (2017) Phthalates, non-phthalate plasticizers, and bisphenols in Swedish preschool dust in relation to children's exposure. Environ Int 102:114–124

Li G, Zu L, Wong PK, Hui X, Lu Y, Xiong J, An T (2012) Biodegradation and detoxification of Bisphenol A with one newly-isolated strain *Bacillus* sp. GZB: kinetics, mechanism, and estrogenic transition. Bioresour Technol 114:224–230

Li ZY, Wu YH, Huo YY, Cheng H, Wang CS, Xu XW (2016) Complete genome sequence of a benzo[a]pyrene-degrading bacterium *Altererythrobacter epoxidivorans* CGMCC 1.7731^{T}. Mar Genomics 25:39–41

Li C, Lu Q, Ye J, Qin H, Long Y, Wang L, Ou H (2018) Metabolic and proteomic mechanism of bisphenol A degradation by *Bacillus thuringiensis*. Sci Total Environ 640–641:714–725

Liang WX, Deutscher MP (2013) Ribosomes regulate the stability and action of the exoribonuclease RNase R. J Biol Chem 288:34791–34798

Liang Z, Li G, An T, Das R (2016) Draft genome sequence of *Bacillus* sp. GZT, a 2,4,6-tribromophenol degrading strain isolated from the river sludge of an electronic waste-dismantling region. Genome Announc 3:e00474–e00416

Liang Z, Li G, An T (2017) Purifying, cloning and characterizing a novel dehalogenase from *Bacillus* sp. GZT to enhance the biodegradation of 2,4,6-tribromophenol in water. Environ Pollut 225:104–111

Little S, Driks A (2001) Functional analysis of the Bacillus subtilis morphogenetic spore coat protein CotE. Mol Microbiol 42:1107–1120

Logan NA, Vos PD (2015) Bacillus. In: Whitman WB (ed) Bergey's manual of systematics of archaea and bacteria. Wiley, Hoboken

Lu L, Zhao M, Wang TN, Zhao LY, Du MH, Li TL, Li DB (2012) Characterization and dye decolorization ability of an alkaline resistant and organic solvents tolerant laccase from *Bacillus licheniformis* LS04. Bioresour Technol 115:35–40

Lu A, Peng Q, Ling E (2014) Formation of disulfide bonds in insect prophenoloxidase enhances immunity through improving enzyme activity and stability. Dev Comp Immunol 44:351–358

Luo A, Wu YR, Xu Y, Kan J, Qiao J, Liang L, Huang T, Hu Z (2016) Characterization of a cytochrome P450 monooxygenase capable of high molecular weight PAHs oxidization from *Rhodococcus* sp. P14. Process Biochem 51:2127–2133

Madigan MT, Martinko J (1997) Brock biology of microorganisms. Pearson Prentice Hall Inc., In

McKenney PT, Driks A, Eichenberger P (2013) The *Bacillus subtilis* endospore: assembly and functions of the multilayered coat. Nat Rev Microbiol 11:33–44

McLeod MP, Warren RL, Hsiao WW, Araki N, Myhre M, Fernandes C, Miyazawa D, Wong W, Lillquist AL, Wang D, Dosanjh M, Hara H, Petrescu A, Morin RD, Yang G, Stott JM, Schein JE, Shin H, Smailus D, Siddiqui AS, Marra MA, Jones SJ, Holt R, Brinkman FS, Miyauchi K, Fukuda M, Davies JE, Mohn WW, Eltis LD (2006) The complete genome of *Rhodococcus* sp. RHA1 provides insights into a catabolic powerhouse. Proc Natl Acad Sci USA 103:15582–15587

Meier-Kolthoff JP, Klenk HP, Göker M (2014) Taxonomic use of DNA G + C content and DNA-DNA hybridization in the genomic age. Int J Syst Evol Microbiol 64:352–356

Mouttaki H, Johannes J, Meckenstock RU (2012) Identification of naphthalene carboxylase as a prototype for the anaerobic activation of non-substituted aromatic hydrocarbons. Environ Microbiol 14:2770–2774

Nicholson WL, Munakata N, Horneck G, Melosh HJ, Setlow P (2000) Resistance of *Bacillus* endospores to extreme terrestrial and extraterrestrial environments. Microbiol Mol Biol Rev 64:548–572

Pal S, Kundu A, Das Banerjee T, Mohapatra B, Roy A, Manna R, Sar P, Kazy SK (2017) Genome analysis of crude oil degrading *Franconibacter pulveris* strain DJ34 revealed its genetic basis for hydrocarbon degradation and survival in oil contaminated environment. Genomics 109:374–382

Peng R, Shi B, Fu X, Tian Y, Zhao W, Zhu B, Xu J, Han H, Yuan Z, Liu K, Yao Q (2015) Improving rice ability to degrade different polycyclic aromatic hydrocarbons through multigene transformation of a hybrid dioxygenase system. Plant Mol Biol Report 33:1030–1041

Petersen TN, Brunak S, von Heijne G, Nielsen H (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Med 8:785–786

Prokka ST (2014) Rrapid prokaryotic genome annotation. Bioinformatics 30:2068–2069

Rochester JR, Bolden AL (2015) Bisphenol S and F: a systematic review and comparison of the hormonal activity of Bisphenol A substitutes. Environ Health Perspect 123:643–650

Rodriguez -RLM, Konstantinidis KT (2014) Bypassing cultivation to identify bacterial species. Microbe 9:111–118

Rojo F (2009) Degradation of alkanes by bacteria: mini review. Environ Microbiol 11:2477–2490

Sasaki M, Akahira A, Tsuchido T, Matsumura Y (2005) Purification of cytochrome P450 and ferredoxin involved in BPA degradation by *Sphingomonas* sp. strain AO1. Appl Environ Microbiol 171:8024–8030

Sasaki M, Tsuchido T, Matsumura Y (2008) Molecular cloning and characterization of cytochrome P450 and ferredoxin genes involved in bisphenol A degradation in *Sphingomonas bisphenolicum* strain AO1. Appl Microbiol 105:1158–1169

Sophos NA, Vasiliou V (2003) Aldehyde dehydrogenase gene superfamily: the 2002 update. Chem Biol Interact 143–144:5–22

Spivack J, Leib TK, Lobos JH (1994) Novel pathway for bacterial metabolism of bisphenol A. Rearrangements and stilbene cleavage in bisphenol A metabolism. J Biol Chem 269:7323–7329

Sudtachat N, Ito N, Itakura M, Masuda S, Eda S, Mitsui H, Kawaharada Y, Minamisawa K (2009) Aerobic vanillate degradation and C1 compound metabolism in *Bradyrhizobium japonicum*. Appl Environ Microbiol 75:5012–5017

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol 30:2725–2729

Telke AA, Kalyani DC, Jadhav UU, Parshetti GK, Govindwar SP (2009) Purification and characterization of an extracellular laccase from a *Pseudomonas* sp. LBC1 and its application for the removal of bisphenol A. J Mol Catal B-Enzym 61:252–260

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties, and weight matrix choice. Nucleic Acids Res 22:4673–4680

Wang Y, Xin Y, Cao X, Xue S (2015) Enhancement of L-2-haloacid dehalogenase expression in *Pseudomonas stutzeri* DEH138 based on the different substrate specificity between dehalogenase-producing bacteria and their dehalogenases. World J Microbiol Biotechnol 31:669–673

Wu M, Pan C, Chen Z, Jiang L, Lei P, Yang M (2017) Bioconcentration pattern and induced apoptosis of Bisphenol A in zebrafish embryos at environmentally relevant concentrations. Environ Sci Pollut R 24:6611–6621

Yang R, Liu G, Chen T, Li S, An L, Zhang G, Li G, Chang S, Zhang W, Chen X, Wu X, Zhang B (2018) Characterization of the genome of a *Nocardia* strain isolated from soils in the Qinghai–Tibetan Plateau that specifically degrades crude oil and of

this biodegradation. Genomics. https://doi.org/10.1016/j.ygeno.2018.02.010

Yoon SH, Ha SM, Lim JM, Kwon SJ, Chun J (2017) A large-scale evaluation of algorithms to calculate average nucleotide identity. Antonie Van Leeuwenhoek 110:1281–1286

Zhang C, Zhang S, Diao H, Zhao H, Zhu X, Lu F, Lu Z (2013) Purification and characterization of a temperature-and pH-stable laccase from the spores of *Bacillus vallismortis* fmb-103 and its application in the degradation of malachite green. J Agric Food Chem 61:5468–5473