




# Genomic analysis of a functional haloacid-degrading gene of *Bacillus megaterium* strain BHS1 isolated from Blue Lake (Mavi Gölü, Turkey)

Batool Hazim Abdul Wahhab<sup>1,2</sup>, Nurul Hidayah Samsulrizal<sup>3\*</sup>, Mohamed Faraj Edbeib<sup>4</sup>, Roswanira Ab. Wahab<sup>5</sup>, Marwan S. M. Al-Nimer<sup>6</sup>, Azzmer Azzar Abdul Hamid<sup>7</sup>, Habeebat Adekilekun Oyewusi<sup>2</sup>, Yilmaz Kaya<sup>8,9</sup>, Kin Israel R. Notarte<sup>10</sup>, Amir Husni Mohd Shariff<sup>11</sup> and Fahrul Huyop<sup>2\*</sup> 

## Abstract

**Purpose:** *Bacillus megaterium* strain BHS1, isolated from an alkaline water sample taken from Mavi Gölü (Blue Lake, Turkey), can grow on minimal medium containing 2,2-dichloropropionic acid. We characterized this bacterium at the genomic level.

**Methods:** The HiSeq platform was used to carry out genome sequencing, de novo assembly, and scaffolding with strain BHS1. Next, genome data were analyzed to demarcate DNA regions containing protein-coding genes and determine the function of certain BHS1 genes. Finally, results from a colorimetric chloride ion–release assay demonstrated that strain BHS1 produces dehalogenase.

**Results:** De novo assembly of the BHS1 genomic sequence revealed a genome size of ~ 5.37 Mb with an average G+C content of 38%. The predicted nuclear genome harbors 5509 protein-coding genes, 1353 tRNA genes, 67 rRNA genes, and 6 non-coding (mRNA) genes. Genomic mapping of strain BHS1 revealed its amenability to synthesize two families of dehalogenases (Cof-type haloacid dehalogenase IIB family hydrolase and haloacid dehalogenase type II), suggesting that these enzymes can participate in the catabolism of halogenated organic acids. The mapping identified seven Na<sup>+</sup>/H<sup>+</sup> antiporter subunits that are vital for adaptation of the bacterium to an alkaline environment. Apart from a pairwise analysis to the well-established L-2-haloacid dehalogenases, whole-cell analysis strongly suggested that the haloacid dehalogenase type II might act stereospecifically on L-2-chloropropionic acid, D,L-2-chloropropionic acid, and 2,2-dichloropropionic acid. Whole-cell studies confirmed the utilization of these three substrates and the gene's role in dehalogenation.

**Conclusions:** To our knowledge, this is the first report of the full genome sequence for strain BHS1, which enabled the characterization of selected genes having specific metabolic activities and their roles in the biodegradation of halogenated compounds.

**Keywords:** Genome, *Bacillus megaterium* BHS1, Dehalogenase gene, Operon, Alkaliphile, Blue Lake, Dehalogenase regulator gene, 2,2DCP

\* Correspondence: [hidayahsamsulrizal@iiu.edu.my](mailto:hidayahsamsulrizal@iiu.edu.my); [fahrul@utm.my](mailto:fahrul@utm.my)

<sup>3</sup>Department of Plant Science, Kulliyah of Science, International Islamic University Malaysia, 25200 Kuantan, Pahang, Malaysia

<sup>2</sup>Department of Biosciences, Faculty of Science, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Malaysia

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## Highlights

- The genome of the *Bacillus megaterium* strain BHS1 was fully sequenced.
- Genes regulating metabolic activities to degrade halogenated compounds were characterized
- Alkaliphiles can be reengineered into effective bioremediation system to treat halogen contaminated sites
- Strain BHS1 genomic data offers the NCBI database with new genomic information.
- This permits comparative genomic analysis studies with other *B. megaterium* species

## Introduction

The soda lake, Mavi Gölü (Blue Lake), in Turkey is a famous tourist destination known for its uniquely beautiful turquoise water that is retained in the lakebed from June to December. It is formed in the Göksu Creek, the only carbonated water that flows into the Black Sea. Given the lake's unique and rugged landscape, the identification/characterization of the lake's culturable microbial communities is essential for expanding the repertoire of alkaliphilic microbes in fortifying available databases for alkaliphilic microbes. The data are particularly pertinent for basic research that elucidates various applications of alkaliphiles, such as in food industries, bioremediation, and medicine (Batumalaie et al. 2018; Bagherbaigi et al. 2013; Neelam et al. 2019; Kevbrin, 2019). Despite being unpolluted, Mavi Gölü may contain organobromines, which are produced naturally in the lake. The organobromine naturally produced by an array of the lake's biological and chemical processes. The processes might trigger certain bacteria to produce dehalogenases (Gribble, 2000). Hence, we investigated the properties of dehalogenases produced by bacteria in heavily polluted areas as well as natural ecosystems that produce environmentally detrimental halogenated compounds.

Many bacteria utilize halogenated compounds as an energy source with the aid of dehalogenases (Abel et al. 2012; Adamu et al. 2016; Edbeib et al. 2016; Edbeib et al. 2017; Akcay and Kaya, 2019; Heidarrezaei et al. 2020; Oyewusi et al. 2020; Muslem et al. 2020). These bacteria generate dehalogenases that catalyze the cleavage of carbon-halogen bonds, which generally yields less toxic intermediates (Kurihara and Esaki, 2008). In particular, *Rhizobium* sp. RC1 remains to be the only bacterium known to produce three different dehalogenases, namely DehD, DehE, and DehL, each of which has unique substrate specificities (Adamu et al. 2016; Hamid et al. 2011; Huyop and Cooper, 2011; Huyop and Sudi, 2012; Sudi et al. 2014).

Because BHS1 produces a dehalogenase, it can utilize 2,2-dichloropropionic acid (2,2DCP) as a sole carbon source (Wahhab et al. 2020). Therefore, we sequenced the genome of strain BHS1 to understand the regulatory mechanisms for overexpression of dehalogenase genes and how dehalogenases contribute to BHS1 survival in an alkali-laden ecosystem. Previous literature reported that *B. megaterium* produces proteins of unknown function (Korneli et al. 2013) that could potentially be used for bioremediation. Thus, to consider its appropriate applications and understand its adaptability to alkali-laden environments such as soda lakes, the acquisition of a full genomic sequence for BSH1 is of interest.

## Materials and methods

### Sampling

During June and July 2019, water samples were collected at ~ 30 cm below the lakewater surface of Mavi Gölü (Dereli), Giresun, Black Sea Region, Turkey (latitude 40°38'08''N north, longitude 38°22'56''E east).

### Preparation of pure genomic DNA

*B. megaterium* strain BHS1 was isolated from water samples taken from Mavi Gölü and grown aerobically in minimal medium containing 20 mM 2,2-DCP (pH 9.0) at 30 °C with rotation. After 18 h, the sample was harvested via centrifugation (13,000×g, 20 min, 4 °C). The QIAamp DNA Minikit (Qiagen, Germany) was used to extract and purify genomic DNA, the quality of which was validated using a Qubit 2.0. fluorometer (Apical Scientific, Malaysia). The purity of DNA samples ( $UVA_{260nm}/A_{280nm}$ ) was assessed using a NanoDrop spectrophotometer (ThermoFisher).

### Analysis of the 16S rRNA gene

*Amplification of the 16S rRNA gene.* Genomic DNA from strain BHS1 was extracted from bacterial cultures grown on minimal media containing 20 mM 2,2-DCP using the Wizard® Genomic DNA Purification kit. PCR was used to amplify the target DNA fragments using the universal primers fP1 (5'-AGAGTTTGATCCTGGC TCAG-3') and rP1 (5'-ACGGTCATACCTTGTTAC GACTT-3') (Fulton and Cooper, 2005). PCR was carried out by 30 cycles of denaturation at 94 °C for 1 min and annealing at 55 °C for 1 min, with final extension at 72 °C for 10 min. Amplicons were purified using the QIAquick PCR purification kit (Qiagen, Germany) and sequenced by the 1st Base Laboratories Sdn Bhd. (Malaysia).

Evolutionary distances were computed using the Maximum Composite Likelihood method (Tamura et al. 2013), with units being the number of base substitutions per site. This analysis involved 15 sequences for bacterial 16S rDNA.

### Genome sequencing, de novo assembly, and scaffolding

The genome of strain BHS1 was sequenced using the HiSeq platform. The insert library with  $\leq 500$  bp was prepared using standard protocols as stated by New England Biolabs Inc. and sequenced using the NGS-Illumina HiSeq TM 2000 Platform (Apical Scientific, Malaysia). All sequences were uploaded to the UCLA CNSI Hoffman 2 computer cluster lab located at the Malaysia Genome Institute for gene assembly. The read maps were based on the Pipeline of Bioinformatics Analysis afforded by Apical Scientific. Data processing, involving the original optical density data obtained by high-throughput sequencing (Illumina platform), were transformed into raw sequenced reads by CASAVA (Hosseini, et al. 2010) and stored in FASTQ format for both the read sequences and the corresponding sequencing-quality information. Quality control was performed, and low-quality sequences were removed.

*De novo* genome assembly was accomplished using Velvet assembler v1.2.10 at the Malaysia Genome Institute (Zerbino and Birney, 2008). High-quality, short paired-end Illumina reads were assembled using VelvetOptimiser.pl v2.2.5. An optimal parameter value of k-mer lengths was identified, and the best assembly was chosen based on N50, number of contigs, and assembly size values. Scaffolding was performed using SSPACE-Standard v3.0 (Boetzer et al. 2010). SSPACE revealed distance information for the paired-end Illumina reads to assess the order, distance, and orientation of the assembled contigs, which were combined into scaffolds. The assembled genome was further refined using the multi-draft-based scaffolder MeDuSa v1.6 (Bosi et al. 2015). This tool allows for the use of multiple reference genomes during scaffolding.

The statistics for final assembly were successfully obtained as one continuous scaffold with 5.37 Mb of total genome size, GC content 38%, and 0.05% gaps (contains N bases). For data accessibility, the *B. megaterium* BHS1 (L1) genome was deposited into NCBI under BioProject PRJNA637885. The Fasta genome submission was under accession number CP058255 (<https://submit.ncbi.nlm.nih.gov>).

### Genome data analysis

Following genome assembly, genetic prediction or annotation was performed to identify DNA regions containing protein-coding genes. To determine the function of genes in the *B. megaterium* BHS1 genome, suitable software and databases, such as BLAST (Basic Local Alignment Search Tool), InterProScan, KEGG, and Blast2Go, were used. BLAST is a program that compares protein or DNA sequences with sequences listed inside various databases, i.e., nr, ref-seq, or SWISS-PROT.

The BLAST search tools used in this study were Blast2GO high-performance cloud server, BLAST Search,

and the stand-alone NCBI-BLAST+ with DIAMOND high-performance analysis with a local database. BLAST analysis aided in finding sequence similarity with proteins in the non-redundant database GenBank with the default parameter set BLOSUM62. Automated annotation using Blast2GO version 5.2.5. (2018) [<http://www.blast2go.com>] was conducted to characterize the coding regions in *B. megaterium* strain BHS1. OmicsBox version 1.3 (<https://www.biobam.com/omicsbox>) was used for mapping, annotation, and visualization as well as the quantitative and statistical analyses for the scaffold of strain BHS1. This program was utilized because of its advanced features compared with Blast2GO (Götz et al., 2008).

Depending on the Kyoto Encyclopedia of Genes and Genomes (KEGG) (Galperin et al. 2014; Zhang et al. 2019) with BLASTP, several databases were utilized for genome annotation and gene function prediction, including Gene Ontology, Clusters of Orthologous Groups (COG), and “shn.” Java version 1.8.0\_152 (2019), which is the OS version for Java 10.0 (<http://java.sun.com>), was used for accessing most databases. Comparison against other reference genomes was performed using a progressive Mauve tool; Mauve is a system for constructing multiple genome alignments that provide a basis for comparative genomics and the study of genome-wide evolutionary dynamics (Darling et al. 2010).

### Whole-cell analysis of maximum chloride released from 2,2DCP, D,L-2-chloropropionic acid (D,L-2CP), L-2-chloropropionic acid (L-2CP), and D-2CP

*B. megaterium* BHS1 was incubated with individual selected substrates (2,2-DCP, D,L-2CP, L-2CP, and D-2CP) under specific growth conditions as previously described (Wahhab et al. 2020). A pure grade of the substrates above were purchased from Sigma-Aldrich or Merck (USA). The chloride ion released was determined colorimetrically for each substrate (Bergmann and Sanik, 1957). The color was allowed to develop for 10 min at 20 °C and measured at  $A_{460 \text{ nm}}$ . After incubation at 20 °C for 10 min, chloride-ion release for each substrate was determined colorimetrically ( $A_{460 \text{ nm}}$ ) (Bergmann and Sanik, 1957).

## Results and discussion

### Organism information

The general features of strain BHS1 have been described by Wahhab et al. (2020). Its full 16S rRNA gene sequence was deposited in GenBank under accession number MT883351. The 16S rRNA gene sequence of BHS1 (GenBank accession number MT883351) was analyzed to determine its evolutionary relationships with other bacterial species. The final dataset contained 1591 bases. Evolutionary analyses were conducted with MEGA X

(Kumar et al. 2018). As shown in Figure 1, BHS1 is closely related to *B. megaterium* strains NCT-2 and WSH-002. Hence, it was confirmed that BHS1 belongs to the species *B. megaterium* (homotypic synonym of *Priestia megaterium*) (Gupta et al. 2020) (Fig. 1).

#### Comparison with other complete reference genome sequences

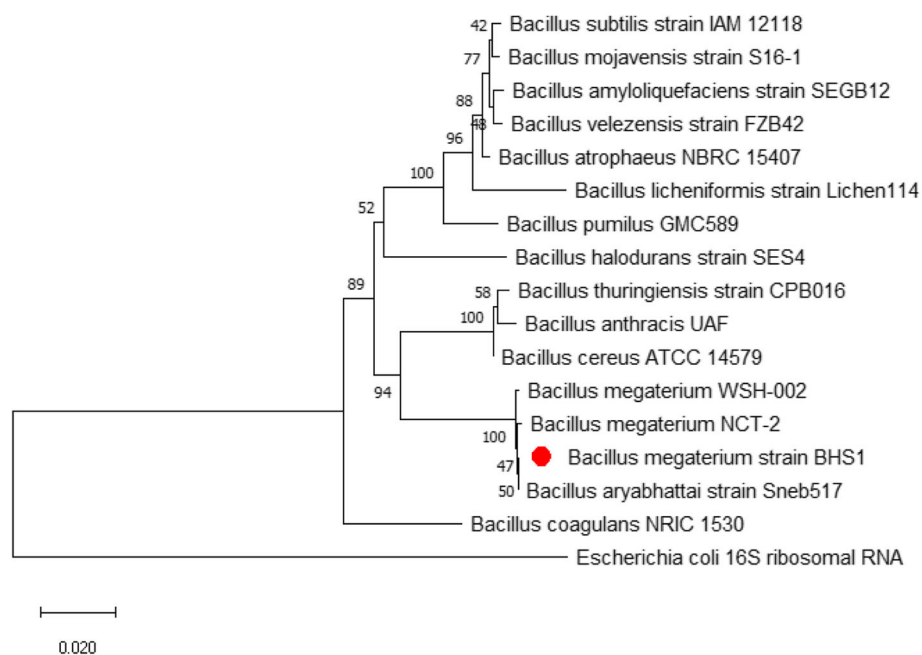
A draft genome for BHS1 was generated based on a comparison with complete reference genomes of *B. megaterium* strains available in the NCBI database. The reference strains WSH-002 (NC\_01738) and NCT-2 (NZ\_CP032527) agreed with the 16S rRNA gene analysis. Figure 2 shows the alignment of the whole genome, where each genome is laid out horizontally. Homologous genome segments in the BHS1 draft genome are shown as colored blocks connected by lines to similarly colored blocks of the reference genomes for the other two strains. Regions that are entirely white could not be aligned and probably contain sequence elements that were recently acquired and thus are unique to the particular strain.

#### Whole-genome sequencing information for strain BHS1

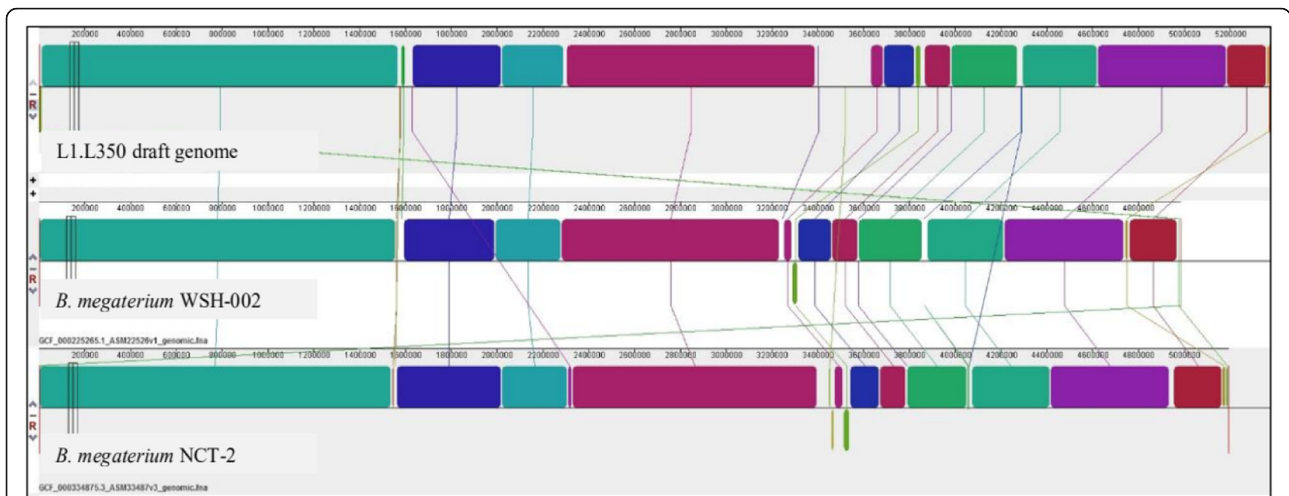
Analysis of one scaffold revealed that *B. megaterium* strain BHS1 has a 5,376,285-bp circular chromosome with an average G+C content of 38%. Analysis of the

nuclear genome of strain BHS1 predicted ~ 5509 protein-coding genes, 79 tRNA genes, 67 rRNA genes, and 6 non-coding (mRNA) genes. Figure 3 shows the circular map of the genome, which was generated using CGview based on open reading frames with COG information. Table 1 summarizes other general features of the genome, and Table 2 lists the functional categories in COG for the full genome sequence of strain BHS1. The main categories are cell wall/membrane/envelope biogenesis (M) ~ 28.3%, transcription (K) 8.0%, translation, ribosomal structure and biogenesis (J) 3.6%, and transport and metabolism of nucleotides (F) ~3%.

With respect to potential biotechnology applications on the industrial aspect, the genome for BHS1 harbors several important industrial enzymes such as  $\alpha$ -2 and  $\beta$ -amylase 1. Moreover, although plants, algae, fungi, and yeast have been the traditional primary sources for urease, this enzyme is also produced by strain BHS1. Other biotechnologically relevant enzymes encoded in the BHS1 genome include thirteen proteases, nine ureases, seven aminopeptidases, five lipases, four aminomutases, three serine proteases, three enzymes for L-lysine synthesis, two glucose dehydrogenases, two pullulanases, and one mutarotase. Thus, it is clear that alkaliphilic *Bacillus* strains are quite important and exciting to explore. This includes the production of biotechnologically important enzymes such as the annotated extracellular hydrolases.



**Fig. 1** Phylogenetic tree highlighting the relative position of the *Bacillus* genus with another closely related genus/species and strains. The evolutionary history was inferred using the Neighbor-Joining method (Saitou and Nei, 1987), where the optimal tree with the sum of branch length is shown to be 0.287. The percentages of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) were sited next to the branches (Felsenstein, 1985). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree



**Fig. 2** Comparison of assembled strain BHS1 draft genome (Sample L1.L350) against two complete reference genomes. Line 1: Strain BHS1 draft genome (Sample L1.L350), Line 2: *B. megaterium* strain WSH-002, and Line 3: *B. megaterium* strain NCT-2

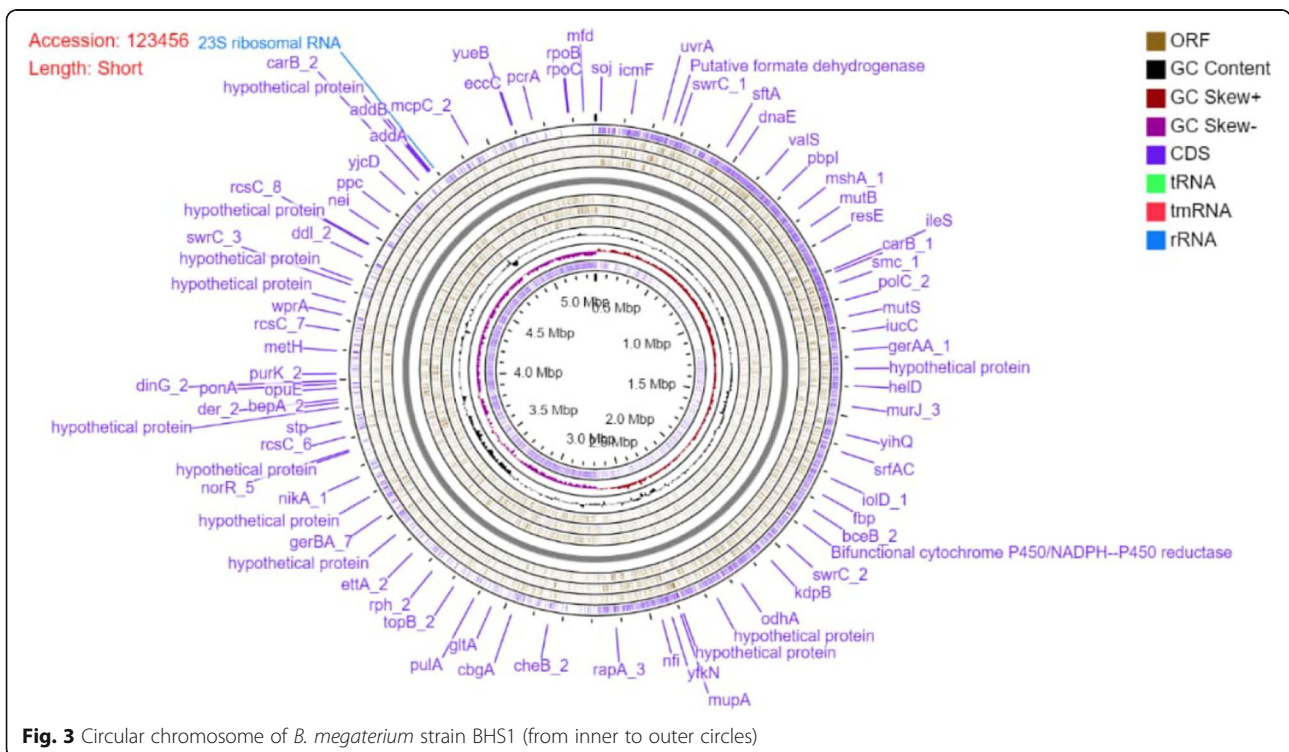
**Dehalogenase genes and dehalogenase subunits for adaptation of strain BHS1 to an alkaline environment**

The germane finding of the genomic annotation for strain BHS1 was the identification of a metabolic pathway that includes two dehalogenase families that were inferred based on the ability of the cells to optimally grow on minimal medium containing 2,2-DCP as the sole carbon source, as reported by Wahhab et al. (2020).

Thus, genes linked to these two dehalogenases with a Locs tag were found in the genome of BHS1: the

first gene encodes a Cof-type haloacid dehalogenase IIB family hydrolase (Nemati et al. 2013), and the second encodes a haloacid dehalogenase type II (Table 3).

The existence of more than a single dehalogenase gene in a bacterial genome has been reported only for *Rhizobium* sp. RC1, and this bacterium has the ability to grow solely on 2,2-DCP and D,L-2-chloropropionate as the only sources of both carbon and energy (Allison et al. 1983; Leigh et al. 1986). To date, no study has explored the possibility that dehalogenase genes could contribute



**Fig. 3** Circular chromosome of *B. megaterium* strain BHS1 (from inner to outer circles)

**Table 1** Genome features of *B. megaterium* strain BHS1

Features	Chromosome (value)	% of total
Genome size (bp)	5,376,285 bp (~ 5.37 Mb)	100%
Protein-coding genes (CDS).	5509	100%
Number of bases for both strand genes (CDS features) without a pseudo count.	4,484,403	81.4%
G+ C content (%)	-	38%
tRNA genes (1353)	79	1.4%
Number of non-coding (mRNA genes)	6	0.1%
Number of rRNA genes	67	1.2%
DNA scaffolds	1	-
Number of gaps	39	0.7%
Repeats regions	22	0.399% ~ 0.4%
Pseudo genes	9	0.16%
Coding percentage (including introns)	83.4	1.5%

**Table 2** Number of genes associated with general COG functional categories of the full genome sequences of *B. megaterium* strain BHS1

Code	Functional annotation	Strain BHS1 (number of genes)	% *
J	Translation, ribosomal structure, and biogenesis	200	3.6%
K	Transcription	443	8.0%
L	Replication, recombination, and repair	139	2.5%
D	Cell cycle control, cell division, chromosome partitioning	59	1.0%
O	Posttranslational modification, protein turnover, chaperones	57	1.0%
M	Cell wall/membrane/envelope biogenesis	1559	28.29%~ 28.3%
N	Cell motility	17	0.3%
P	Inorganic ion transport and metabolism	48	0.87 %
T	Signal transduction mechanisms	117	2.1%
C	Energy production and conversion	8	0.1%
G	Carbohydrate transport and metabolism	16	0.29%~0.3 %
E	Amino acid transport and metabolism	32	0.58 %~0.6 %
F	Nucleotide transport and metabolism	164	2.97 %~3%
H	Coenzyme transport and metabolism	16	0.29%~0.3%
I	Lipid transport and metabolism	46	0.8%
Q	Secondary metabolites biosynthesis, transport, and catabolism	47	0.85%
S	Function unknown	1902	34.5%
-	No. IPS (not in COGs).	469	8.5%
A	RNA processing and modifications	116	2.1%
B	Chromatin structure and dynamics	0	0%
V	Defense mechanism	1	0.018%
U	Intracellular trafficking	0	0%
<b>Total</b>	Total number of genes	5509	

\*The total based on the total number of genes in the full genome

**Table 3** The genes code for enzymes of dehalogenase degradation pathway in strain BHS1

Enzyme	Locs tag
Cof-type HAD-IIB family hydrolase	Bacillus_3231. WP_013082899.1
Haloacid dehalogenase type II	Bacillus_4574. WP_045293204.1

to adaptation to an alkali-laden environment (Ismail et al. 2017).

The 2-haloacid dehalogenases act on 2-haloacids, liberating a halide ion(s) and the corresponding 2-hydroxy acid. The 2-haloacid dehalogenases have been categorized phylogenetically into two groups, namely I and II. Group II, or L-type enzymes, include L-2-haloacid dehalogenases (i.e., L-DEXs) that explicitly target L-2-haloacids. The group II enzymes are more common than group I dehalogenases, which catalyze on both D and L-haloacids (Wang et al. 2018).

#### Co-transport mechanisms

Many cation/proton antiporters have been identified in bacteria (Krulwich et al. 2009; Tsujii et al. 2020). Cation/proton antiporters contribute to maintaining the appropriate intracellular pH homeostasis particularly in alkaline environments that require import of hydrogen ions from a comparatively proton-poor external environment. Three families of cation/proton antiporters of the Gene Ontology Membrane Transporter Group have been identified in bacteria and perform significant roles in maintaining cation equilibrium in the cytoplasm in extreme environmental settings. For instance, monovalent cation/proton antiporters are integral membrane proteins that facilitate the efflux of cytoplasmic sodium, potassium, or lithium ions in exchange for external hydrogen ions (protons) (Morino et al. 2014). The genome of strain BHS1 which was isolated from the alkaline lake, encodes all seven cation/proton antiporter subunits (A-G) in a single gene. These seven subunits with their Locs tag are shown in Table 4. Na<sup>+</sup>/H<sup>+</sup> antiporters have numerous functions, namely the creation of an essential electrochemical gradient of Na<sup>+</sup> across the plasma

**Table 4** The genetic codes for enzymes of the Na<sup>+</sup>/H<sup>+</sup> adaptation pathway of strain BHS1 in an alkaline environment

Enzyme	Locs tag
Na <sup>+</sup> /H <sup>+</sup> antiporter subunit A	Bacillus_354. WP_050688978.1
Na <sup>+</sup> /H <sup>+</sup> antiporter subunit B	Bacillus_353. WP_013085301.1
Na <sup>+</sup> /H <sup>+</sup> antiporter subunit C	Bacillus_352. WP_013059592.1
Na <sup>+</sup> /H <sup>+</sup> antiporter subunit D	Bacillus_351. WP_028412608.1
Na <sup>+</sup> /H <sup>+</sup> antiporter subunit E	Bacillus_350. WP_013059594.1
Na <sup>+</sup> /H <sup>+</sup> antiporter subunit F1	Bacillus_349. WP_013059595.1
Na <sup>+</sup> /H <sup>+</sup> antiporter subunit G	Bacillus_348. WP_013085304.1

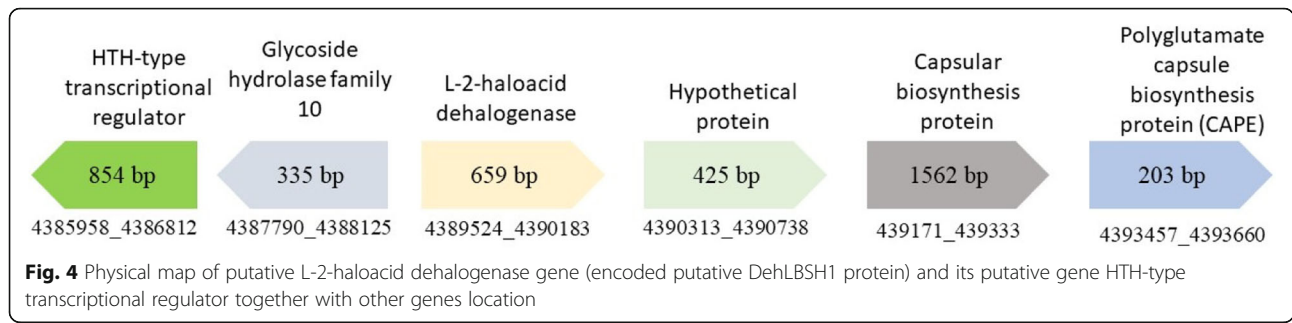
membrane for the purpose Na<sup>+</sup>-driven flagellar rotation, the efflux of Na<sup>+</sup> and Li<sup>+</sup> (both of which are toxic if accumulated at high concentrations within cells), and regulation of intracellular pH under alkaline conditions (Wai Liew et al. 2007). Interestingly, strain BHS1 encodes several types of Na<sup>+</sup>/H<sup>+</sup> antiporters (Table 4) that might be vital for maintaining intracellular pH during growth in alkaline environments, e.g., pH 10.5, similar to the previously reported Na<sup>+</sup>/H<sup>+</sup> antiporter gene (*g1-nhaC*) of alkaliphilic *Bacillus* sp. G1 when expressed in *Escherichia coli* (Wai Liew et al. 2007). The full genome sequence of *Bacillus* may not only reveal individual bacilli functions but also play a role in identifying unique properties that are essential for the survival and adaptation of bacilli in alkaline ecosystems.

#### Organization of dehalogenase genes and a putative regulatory gene

The regulation of dehalogenase gene expression is poorly understood from the perspective of whole-genome function (Huyop and Cooper, 2011). Hence, analysis of the full genome of BHS1 may shed light on this aspect via the identification of an operon consisting of a functional gene cluster, i.e., regulatory gene and structural genes, the expression of which depends on their location within the operon. A putative haloacid dehalogenase operon comprised of L-type dehalogenase (group II haloacid dehalogenase or L-specific dehalogenases) was identified in the BHS1 genome [4,389,524\_4,390,183; 659 bp], with downstream genes consisting of a hypothetical protein [4,390,313\_4,390,738; 425 bp], a capsular biosynthesis protein [439,171\_439,333; 1562 bp], and the polyglutamate capsule biosynthesis protein CAPE [4,393,457\_4,393,660; 203 bp]. Conversely, the upstream region of the operon encodes the glycoside hydrolase family 10 [4,387,790\_4,388,125; 335 bp] and a helix-turn-helix (HTH)-type transcriptional regulator [4,385,958\_4,386,812; 854 bp]. Thus, the acquisition of the full genomic sequence for BHS1 is crucial for determining the most appropriate applications of these gene clusters for biotechnology purposes (Fig. 4). Toward that end, we are currently using cloning analysis to decipher the regulatory mechanisms for these genes, especially given the existence of a HTH-type transcriptional regulator gene.

#### Analysis of a putative L-type dehalogenase (group II haloacid dehalogenase) and HTH-type transcriptional regulator

The size of L-type dehalogenase was 657 bp equivalent to 219 amino acid residues. The L-type dehalogenase of strain BHS1 was later designated as DehLBHS1. The program ProtParam (Gasteiger et al. 2005) indicated that DehLBHS1 has a molecular mass of 25,636.47 Da and a



theoretical pI of 6.66, i.e., close to neutral. The sequence is comprised of 28% negatively charged residues (Asp, Glu) and 27% positively charged residues (Arg, Lys). The grand average of hydropathicity index (GRAVY) (Kyte and Doolittle, 1981) of  $-0.42$  suggests that, overall, DehLBHS1 is a hydrophilic protein. Amino-acid sequence alignment of DehLBHS1 and L-DEX YL (Hisano et al. 1996; Li et al. 1998) revealed three key residues, namely D (Asp), R (Arg), and S (Ser), that may play key roles in the catalytic mechanism of DehLBHS1. The functional residues of L-DEX YL were determined to be D10, R41, S118, and D180, which are comparatively equivalent to the BHS1 residues D10, R41, S127, and D180, respectively (Fig. 5). Although the amino-acid sequences of the L-haloacid dehalogenases from BHS1 and *Pseudomonas* sp. YL (L-DEX YL) are only 46% identical, the key residues are 100% identical and thus may have similar functions. Analysis of the HTH-type transcriptional regulator, designated as DehRBHS1 (a BHS1 strain, dehalogenase regulator protein) revealed very low sequence identity ( $< 17%$ ) compared with the well-established regulator protein (DehR) of *Rhizobium* sp. RC1 (data not shown). This suggests that DehRBHS1 may have a similar regulatory function in the dehalogenase operon system that controls dehalogenases.

**Functional characterization of dehalogenase-producing strain BHS1**

Dehalogenase-producing *B. megaterium* strain BHS1 was incubated with 5 mM of each substrate (2,2-DCP, D,L-2CP, L-2CP, and D-2CP) under proper growth conditions, and the maximum amount of chloride ion released was monitored. In the presence of 2,2-DCP or D,L-2CP, 50% of the chloride was released (Fig. 6). In the presence of L-2CP alone, essentially all chloride ions were released ( $4.5 \mu\text{mol Cl}^-/\text{ml}$ ), whereas, with D-2CP, no chloride ion was detected, as expected, suggesting the stereospecificity of the enzyme. A control experiment, L-2CP without any cells, showed no dehalogenation reaction, affirming the substrate’s non-autodegradation (Fig. 6).

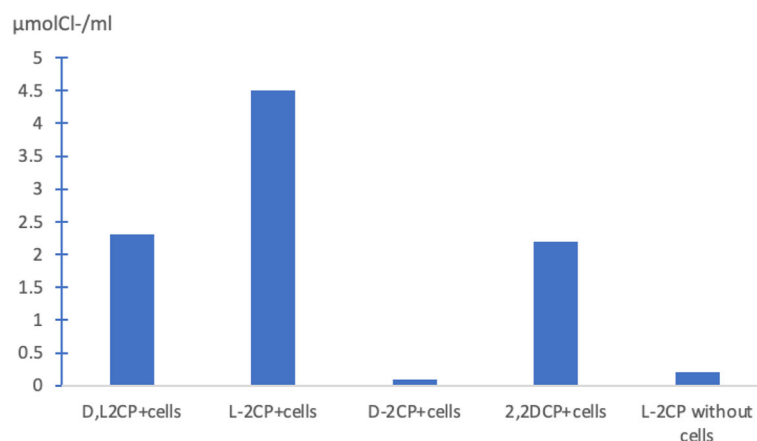
**Conclusions**

The study performed here is the first full genome sequence analysis of the functional haloacid-degrading gene from the alkaliphilic *B. megaterium* strain BHS1 from Mavi Gölü with the potential to degrade haloalkanoic acid in an alkaline environment that amenable to converting haloacids for carbon and energy sources. The genomic data acquired during this study may facilitate the discovery of new dehalogenases and its regulatory

DehLBHS1	001	MKF-RTILFDAYGTLFDVHTVIETCDKLYPERGESISHLW <b>RLKQIEYAMQ</b>	049
L-DEXYL	001	MDYIKGIAFDLYGTLFDVHVVGRGDEAFPPGRGREISALW <b>RQKQLEYTWL</b>	050
DehLBHS1	050	YQLMGRYVDFYTLTNQSLKYAAEANDIDLTEHEEKMLMSAYMKLNVDYDEV	099
L-DEXYL	051	RSLMNRYVNFQQATEDALRFTCRHLGLDLDARTRSTLCCDAYLRLAPFSEV	100
DehLBHS1	100	KEVLAYLKEKGYRLAIFTNGPKNMIDPLVSYHNMNHLFEDVISVDEIKQY	149
L-DEXYL	101	PDSLRELKRRGLKLAILSNGSPQSIDAVVSHAGLRDGFHLLSVDPPQVY	150
DehLBHS1	150	KPTMASYHYAKNMNAKREEVFLFLSSNTWDIAGAKNYGFATAWVNRKGNV	199
L-DEXYL	151	KPDNRVYELAEQALGLDRSAILFVSSNAWDATGARYFGFPFCWINRTGNV	200
DehLBHS1	200	AEHKELQPNVVIKSLNQLIK*-----	219
L-DEXYL	201	FEEMGQTPDWEVTSLRVVVELFETAAGKAEG*	232

**Fig. 5** The pairwise sequence alignment results of DehLBHS1 and L-DEX YL from *Pseudomonas* sp. strain YL (Nardi-Dei et al., 1997). Mark in bold were three key amino acids (D,Asp; R,Arg; S,Ser), which play an important role in the catalytic function of L-DEX YL haloacid dehalogenase. The three key amino acids (in bold) also were located in DehLBHS1 in the same position except for S(Ser). \* symbol indicates a stop codon





**Fig. 6** Resolution of racemic D,L-2-chloropropionic acid, L-2-CP, D-2-CP by *B. megaterium* strain BHS1 expressing stereospecific DehLBHS1

functions. The knowledge may also provide insights into the bacterium's genetic and metabolic regulatory pathways, of which will be vital for potential application of BHS1 to the bioremediation of contaminated alkaline ecosystems.

#### Acknowledgements

The authors would like to acknowledge the Scientific Technological Research Council of Turkey (TUBITAK), Program 2221-Fellowship for Visiting Scientists and Scientists on Sabbatical Leave with Grant number 1059B11800242. FH and BHAW thanks the Fundamental Research Grant Scheme (FRGS)-Ministry of Education RJ130000.7854.5F189/FRGS/1/2019/STG05/UTM/01/1 and QJ130000.2414.08G59 for partial financial assistance. Finally, AHMS, BHAW, and HAO also thank Universiti Malaysia Sabah (UMS) for a joint university research project between the Faculty of Science UTM and the Faculty of Food Science and Nutrition UMS.

#### Authors' contributions

All authors contributed equally to the work. The first author is a Ph.D. candidate and responsible for the data collections. FH is the main supervisor and principal investigator for this research. All materials collected from YK, the principal researcher at Ondokuz Mayıs University, Turkey, were initially analyzed and prepared by MFE. NHS expert in genomic data analysis and HAO contributed partially to data analysis. KIRN and RAW read the overview of the manuscript and data organization. AAAH helped to perform the phylogenetic analysis. AHMS and MSMAN approved special funding related to the research work and publication. All authors read and approved the final manuscript.

#### Funding

Program 2221-Fellowship for Visiting Scientists and Scientists on Sabbatical Leave with Grant number 1059B11800242; (FRGS)-Ministry of Education RJ130000.7854.5F189/FRGS/1/2019/STG05/UTM/01/1; RUGQJ130000.2414.08G59; University Malaysia Sabah Special Grants.

#### Availability of data and materials

Strictly for research without profit-making interest.

#### Ethics approval and consent to participate

Ethics approval and consent to participate: This article does not contain any studies with human participants or animals performed by any authors. Informed consent was obtained from all individual participants included in the study.

#### Consent for publication

All authors agreed to participate and agreed on the submission.

#### Competing interests

The authors declare that they have no competing interests.

#### Author details

<sup>1</sup>Department of Microbiology, Faculty of Medicine, Al-Mustansiriyah University, Baghdad, Iraq. <sup>2</sup>Department of Biosciences, Faculty of Science, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Malaysia. <sup>3</sup>Department of Plant Science, Kulliyah of Science, International Islamic University Malaysia, 25200 Kuantan, Pahang, Malaysia. <sup>4</sup>Department of Medical Laboratories, Faculty of Medical Technology, Bani Walid University, Bani Walid, Libya. <sup>5</sup>Department of Chemistry, Faculty of Science, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Malaysia. <sup>6</sup>Department of Pharmacology, Faculty of Pharmacy, University Medical Hawler, Erbil, Iraq. <sup>7</sup>Department of Biotechnology, Kulliyah of Science, International Islamic University Malaysia, 25200 Kuantan, Pahang, Malaysia. <sup>8</sup>Department of Biology, Faculty of Science, Kyrgyz-Turkish Manas University, Bishkek, Kyrgyzstan. <sup>9</sup>Department of Agricultural Biotechnology, Faculty of Agriculture, Ondokuz Mayıs University, Samsun, Turkey. <sup>10</sup>Faculty of Medicine and Surgery, University of Santo Tomas España, 1015 Manila, Philippines. <sup>11</sup>Faculty of Food Science and Nutrition, Universiti Malaysia Sabah, Jalan UMS, 88400 Kota Kinabalu, Sabah, Malaysia.

Received: 9 December 2020 Accepted: 17 February 2021

Published online: 27 February 2021

#### References

- Abel E, Pakingking RV Jr, Gregoria G, Wint MT, Huyop F (2012) Characteristics of dehalogenase from bacteria isolated from the Gut of Pond-reared Rohu (*Labeo rohita*) Juveniles in Myanmar. *Adv Biosci Biotechnol* 3(4):353–361 <https://doi.org/10.4236/abb.2012.34051>
- Adamu A, Wahab RA, Huyop F (2016) L-2-haloacid dehalogenase (DehL) from *Rhizobium* sp. RC1. *SpringerPlus* 5(1):695 <https://doi.org/10.1186/s40064-016-2328-9>
- Akcay K, Kaya Y (2019) Isolation, characterization and molecular identification of a halotolerant *Bacillus megaterium* CTBmeg1 able to grow on halogenated compounds. *Biotechnol Biotechnol Equip* 33(1):945–953 <https://doi.org/10.1080/13102818.2019.1631717>
- Allison N, Skinner AJ, Cooper RA (1983) The dehalogenases of a 2,2-dichloropropionate-degrading bacterium. *J Gen Microbiol* 129:1283–1293
- Bagherbaigi S, Gicana RG, Lamis RJ, Nemat M, Huyop F (2013) Characterisation of *Arthrobacter* sp. S1 that can degrade  $\alpha$  and  $\beta$ -haloalkanoic acids isolated from contaminated soil. *Ann Microbiol* 63:1363–1369 <https://doi.org/10.1007/s13213-012-0595-4>
- Batumalaie K, Edbeib MF, Mahat NA, Huyop F, Wahab RA (2018) *In silico* and empirical approaches toward understanding the structural adaptation of the alkaline-stable lipase KV1 from *Acinetobacter haemolyticus*. *J Biomol Struct Dynamics* 36(12):3077–3093 <https://doi.org/10.1080/07391102.2017.1377635>
- Bergmann J, Sanik JJ (1957) Determination of trace amounts of chlorine in naphtha. *Anal Chem* 29(2):241–243

- Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W (2010) Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* 27(4):578–579 <https://doi.org/10.1093/bioinformatics/btq683>
- Bosi E, Donati B, Galardini M, Brunetti S, Sagot M-F, Lió P, Crescenzi P, Fani R, Fondi M (2015) MeDuSa: a multi-draft based scaffolder. *Bioinformatics* 31(15):2443–2451 <https://doi.org/10.1093/bioinformatics/btv171>
- Darling AE, Mau B, Perna NT (2010) Progressive Mauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* 5(6):e11147 <https://doi.org/10.1371/journal.pone.0011147>
- Edbeib MF, Wahab RA, Huyop F (2016) Characterization of an  $\alpha$ -haloalkanoic acid-degrading *Pseudomonas aeruginosa* MX1 isolated from contaminated seawater. *Bioremed J* 20(2):89–97 <https://doi.org/10.1080/10889868.2015.1114464>
- Edbeib MF, Wahab RA, Kaya Y, Huyop F (2017) *In silico* characterization of a novel dehalogenase (DehHX) from the halophile *Pseudomonas halophila* HX isolated from Tuz Gölü Lake, Turkey: insights into a hypersaline-adapted dehalogenase. *Ann Microbiol* 67(5):371–382 <https://doi.org/10.1007/s13213-017-1266-2>
- Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39(4):783–791 <https://doi.org/10.1111/j.1558-5646.1985.tb00420.x>
- Fulton CK, Cooper RA (2005) Catabolism of sulfamate by *Mycobacterium* sp. CF1. *Environ Microbiol* 7:378–381 <https://doi.org/10.1111/j.1462-2920.2004.00719.x>
- Galperin MY, Makarova KS, Wolf YI, Koonin EV (2014) Expanded microbial genome coverage and improved protein family annotation in the COG database. *Nucleic Acids Res* 43(D1):D261–D269 <https://doi.org/10.1093/nar/gku1223>
- Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A (2005) Protein identification and analysis tools on the ExpASY Server. In: Walker JM (ed) *The Proteomics Protocols Handbook*. Humana Press, pp 571–607
- Götz S, Garcia-Gomez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, Robles M, Talon M, Dopazo J, Conesa A (2008) High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic acids Res* 36(10):3420–3435
- Gribble GW (2000) The natural production of organobromine compounds. *Environ Sci Poll Res* 7(1):37–49 <https://doi.org/10.1065/espr199910.002>
- Gupta RS, Patel S, Saini N, Chen S (2020) Robust demarcation of 17 distinct bacillus species clades, proposed as novel bacillaceae genera, by phylogenomics and comparative genomic analyses: description of *Robertmurraya kyonggiensis* sp. Nov. And Proposal for an Emended Genus *Bacillus* Limiting It Only to the Members of the *Subtilis* and *Cereus* Clades of Species. *Int J Syst Evol Microbiol* 70(11):5753–5798 <https://doi.org/10.1099/ijsem.0.004475>
- Hamid THTA, Hamid AAA, Huyop F (2011) A review on non-stereospecific haloalkanoic acid dehalogenases. *Afr J Biotechnol* 10(48):9725–9736 <https://doi.org/10.5897/AJB11.934>
- Heidarrezaei, M., Shokravi, H., Huyop, F., Rahimian Koloor, S.S., Petru, M. (2020) Isolation and characterization of a novel bacterium from the marine environment for trichloroacetic acid bioremediation. *Appl Sci*, 10(13):4593. <https://doi.org/10.3390/app10134593>
- Hisano T, Hata Y, Fujii T, Liu J-Q, Kurihara T, Esaki N, Soda K (1996) Crystal structure of L-2-haloacid dehalogenase from *Pseudomonas* sp. YL AN  $\alpha/\beta$  hydrolase structure that is different from the  $\alpha/\beta$  hydrolase fold. *J Biol Chem* 271(34):20322–20330 <https://doi.org/10.1074/jbc.271.34.20322>
- Hosseini P, Tremblay A, Matthews BF, Alkharouf NW (2010) An efficient annotation and gene-expression derivation tool for Illumina Solexa datasets. *BMC Res notes* 3:183 <https://doi.org/10.1186/1756-0500-3-183>
- Huyop F, Cooper RA (2011) Regulation of dehalogenase E (Dehe) and expression of dehalogenase regulator gene (Dehr) from *Rhizobium* sp. RC1 in *E. coli*. *Biotechnol Biotechnol Equip* 25(1):2237–2242 <https://doi.org/10.5504/BBEQ.2011.0009>
- Huyop F, Sudi IY (2012) D-specific dehalogenases, a review. *Biotechnol Biotechnol Equip* 26(2):2817–2822 <https://doi.org/10.5504/BBEQ.2011.0143>
- Ismail SNF, Wahab RA, Huyop F (2017) Microbial isolation and degradation of selected haloalkanoic aliphatic acids by locally isolated bacteria: A review. *Malaysian J Microbiol* 13(3):261–272
- Kevbrin W (2019) Isolation and Cultivation of Alkaliphiles. In: Mamo G, Mattiasson B (eds) *Alkaliphiles in Biotechnology*. Advances in Biochemical Engineering/ Biotechnology, vol 172. Springer, Cham, pp 195–220 [https://doi.org/10.1007/10\\_2018\\_84](https://doi.org/10.1007/10_2018_84)
- Korneli C, David F, Biedendieck R, John D, Wittmann C (2013) Getting the big beast to work—Systems biotechnology of *Bacillus megaterium* for novel high-value proteins. *J Biotechnol* 163(2):87–96 <https://doi.org/10.1016/j.jbiotec.2012.06.018>
- Krulwich TA, Hicks DB, Ito M (2009) Cation/proton antiporter complements of bacteria: why so large and diverse? *Mol Microbiol* 74(2):257–260 <https://doi.org/10.1111/j.1365-2958.2009.06842.x>
- Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol* 35(6):1547–1549 <https://doi.org/10.1093/molbev/msy096>
- Kurihara T, Esaki N (2008) Bacterial hydrolytic dehalogenases and related enzymes: occurrences, reaction mechanisms, and applications. *Chem Record* 8(2):67–74. <https://doi.org/10.1002/tcr.20141>
- Leigh JA, Skinner AJ, Cooper RA (1986) Isolation and partial characterization of dehalogenase-deficient mutants of a *Rhizobium* sp. *FEMS Microbiol Letters* 36:136–166
- Li Y-F, Hata Y, Fujii T, Hisano T, Nishihara M, Kurihara T, Esaki N (1998) Crystal structures of reaction intermediates of L-2-haloacid dehalogenase and implications for the reaction mechanism. *J Biol Chem* 273(24):15035–15044 <https://doi.org/10.1074/jbc.273.24.15035>
- Morino M, Suzuki T, Ito M, Krulwich TA (2014) Purification and functional reconstitution of a seven-subunit Mrp-Type Na<sup>+</sup>/H<sup>+</sup> Antiporter. *J Bacteriol* 196(1):28–35 <https://doi.org/10.1128/jb.01029-13>
- Muslem WH, Edbeib MF, Aksoy HM, Kaya Y, Hamid AAA, Hood MHM, Wahab RA, Huyop F (2020) Biodegradation of 3-chloropropionic acid (3-CP) by *Bacillus cereus* WH2 and its *in silico* enzyme-substrate docking analysis. *J Biomol Struct Dynamics* 38(11):3432–3441 <https://doi.org/10.1080/07391102.2019.1655482>
- Nardi-Dei V, Kurihara T, Chung P, Esaki N, Soda K (1997) Bacterial D,L-2-haloacid dehalogenase from *Pseudomonas* sp. strain 113: Gene cloning and structural comparison with D and L-2 haloacid dehalogenases. *J Bacteriol* 179(3):4232–4238. <https://doi.org/10.1128/jb.179.13.4232-4238.1997>
- Neelam DK, Agrawal A, Tomer AK, Bandyopadhyaya S, Sharma A, Jagannadham MV, Mandal CC, Dadheech PK (2019) A *Piscibacillus* sp. isolated from a soda lake exhibits anticancer activity against breast cancer MDA-MB-231 cells. *Microorganisms* 7(2):34 <https://doi.org/10.3390/microorganisms7020034>
- Nemati, M., Edbeib, M.F., Hamid, A.A.A., Gicana, R.G., Ibrahim, N., Huyop, F. (2013) Identification of putative Cof-like hydrolase associated with dehalogenase in *Enterobacter cloacae* MN1 isolated from the contaminated sea-side area of the Philippines. *Malaysian J Microbiol*, 9(3):253-259. <https://doi.org/10.21161/mjm.51113>
- Oyewusi HA, Wahab RA, Kaya Y, Edbeib MF, Huyop F (2020) Alternative bioremediation agents against haloacids, haloacetates and chlorpyrifos using novel halogen-degrading bacterial isolates from the hypersaline lake Tuz. *Catalysts* 10(6):651 <https://doi.org/10.3390/catal10060651>
- Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406–425 <https://doi.org/10.1093/oxfordjournals.molbev.a040454>
- Sudi IY, Hamid AAA, Shamsir MS, Jamaluddin H, Wahab RA, Huyop F (2014) Insights into the stereospecificity of the D-specific dehalogenase from *Rhizobium* sp. RC1 toward d- and l-2-chloropropionate. *Biotechnol Biotechnol Equip* 28(4):608–615 <https://doi.org/10.1080/13102818.2014.937907>
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 30(12):2725–2729 <https://doi.org/10.1093/molbev/mst197>
- Tsujii M, Tanudjaja E, Uzumi N (2020) Diverse Physiological Functions of Cation Proton Antiporters across Bacteria and Plant Cells. *Int J Mol Sci* 21(12):4566 <https://doi.org/10.3390/ijms21124566>
- Wahhab BHA, Anuar NFSK, Wahab RA, Al Nimer MS, Samsulrizal NH, Hamid AAA, Edbeib MF, Kaya Y, Huyop F (2020) Identification and characterization of a 2, 2-dichloropropionic acid (2, 2-DCP) degrading alkalotolerant bacterium strain BHS1 isolated from Blue Lake, Turkey. *J Trop Life Sci* 10(3):245–252 <https://doi.org/10.11594/jtls.10.03.08>
- Wai Liew C, Illias RM, Muhammad Mahadi N, Najimudin N (2007) Expression of the Na<sup>+</sup>/H<sup>+</sup> antiporter gene (g1-nhaC) of alkaliphilic *Bacillus* sp. G1 in *Escherichia coli*. *FEMS Microbiol Letters* 276(1):114–122 <https://doi.org/10.1111/j.1574-6968.2007.00925.x>
- Wang Y, Feng Y, Cao X, Liu Y, Xue S (2018) Insights into the molecular mechanism of dehalogenation catalyzed by D-2-haloacid dehalogenase from crystal structures. *Sci Rep* 8(1):1454 <https://doi.org/10.1038/s41598-017-19050-x>

- Zerbino DR, Birney E (2008) Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* 18(5):821–829 <https://doi.org/10.1101/gr.074492.107>
- Zhang H, Dong S, Lou T, Wang S (2019) Complete genome sequence unveiled cellulose degradation enzymes and secondary metabolic potentials in *Streptomyces* sp. CC0208. *J Basic Microbiol* 59(3):267–276 <https://doi.org/10.1002/jobm.201800563>

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

