

Research Article

## Genomic Analysis of *Mesorhizobium loti* Strain TONO Reveals Dehalogenases for Bioremediation

Sefatullah Zakary<sup>1,2</sup>, Habeebat Adekilekun Oyewusi<sup>1,3</sup>, Fahrul Huyop<sup>1\*</sup>

<sup>1</sup> Department of Biosciences, Faculty of Sciences, Universiti Teknologi Malaysia, Johor Bahru 81310, Malaysia

<sup>2</sup> Department of Botany, Faculty of Biology, Kabul University, Dehberi, Kabul 1006, Afghanistan.

<sup>3</sup> Department of Biochemistry, School of Science and Computer Studies, Federal Polytechnic Ado Ekiti, Ado Ekiti PMB 5351, Ekiti State, Nigeria

*Article history:*

Submission December 2020

Revised December 2020

Accepted January 2021

\*Corresponding author:

E-mail: [fahrul@utm.my](mailto:fahrul@utm.my)

### ABSTRACT

Halogenated compounds are extensively utilized in different industrial applications such as pesticides and herbicides and cause severe environmental problems because of their toxicity and persistence. Degradation of these compounds by the biological method is a significant method to reduce these recalcitrant. *Mesorhizobium loti* is important for nitrogen fixation in legume roots. Up to now, there is no report to indicate *M. loti* can produce dehalogenase enzymes. Thus, a total of twenty-five genomes of *M. loti* strains from the National Center for Biotechnology Information (NCBI) were analyzed. These strains notably carry dehalogenase genes and were further investigated. The relative ratio of haloalkane and haloacid dehalogenase type II or L-type from all twenty-five genomes was 26% and 74%, respectively, suggesting type II dehalogenase is common. Surprisingly, only *M. loti* strain TONO carries four dehalogenases and therefore it was further characterized. The chromosome of *M. loti* strain TONO contains four haloacid dehalogenase type II genes namely, *dehLt1* (MLTONO\_2099), *dehLt2* (MLTONO\_3660), *dehLt3* (MLTONO\_4143), and *dehLt4* (MLTONO\_6945), and their corresponding enzymes were designated as DehLt1, DehLt2, DehLt3, and DehLt4, respectively. The only haloalkane dehalogenase gene (MLTONO\_4828) was located upstream of the *dehLt3* gene and its amino acid share 88% identity with DmlA of *Mesorhizobium japonicum* MAFF 303099. The putative haloacid permease gene designated as *dehrPt* (MLTONO\_0284) was located downstream of the *dehLt1* and its amino acids show 69% identity with haloacid permease of *Rhizobium* sp. RC1. The gene encoding helix-turn-helix (HTH) motif family DNA-binding protein regulator and LysR family transcriptional regulator genes were also identified, possibly for regulatory functions. The genomic studies as such, have good potential to be screened for new type of dehalogenases based on basic molecular structure and functions analysis.

*Keywords:* Genomic analysis, Haloacid dehalogenase, Haloacid permease, Halogenated organic compounds, *Mesorhizobium loti*

### Introduction

A significant group of halogenated organic compounds such as chlorinated hydrocarbons was produced and extensively utilized in agriculture and various industrial applications. This resulted in severe environmental dilemmas due to their toxicity and persistence [1]. Fortunately, many bacterial species produce dehalogenases that could breakdown these toxic compounds and convert them into harmless products [2]. Most of these

bacteria species used organohalides as their sole source of carbon and energy [3]. Dehalogenase catalyzes halogenated organic pollutants' by cleaving of carbon–halogen bonds. Many dehalogenases were identified, for example haloalkane dehalogenase [4], fluoroacetate dehalogenase [5], haloalkanoic acid dehalogenase [6-8], and 4-chlorobenzoyl-CoA dehalogenase [9]. Haloalkanoic acid dehalogenase catalyzes haloalkanoic acids

*How to cite:*

Zakary S, Oyewusi HA, Huyop F (2021) Genomic Analysis of *Mesorhizobium loti* Strain TONO Reveals Dehalogenases for Bioremediation. Journal of Tropical Life Science 11 (1): 67 – 77. doi: 10.11594/jtls.11.01.09.

and produces hydroxy-acid and a halide [10].

The 2-haloacid dehalogenases were classified into three groups based on the substrate specificities [11], whereas the genetic approach by Hill *et al.* [10] classified dehalogenases into groups I and II. Group II was stereoselective dehalogenases acting only on L-2-haloacids. While group I dehalogenases were non stereospecific dehalogenases can react with both D and L form of substrates producing L- and D-hydroxy acids, respectively or can retain the chemical's isomeric pattern [10]. D-specific dehalogenase was categorised as group I as well. Slater and colleagues [8] showed that a dehalogenase crude extract from *Pseudomonas putida* PP3 grown on 2-chloropropionate, dehalogenate only a small amount 3 chloropropionate. Recently, these types of dehalogenases have received more attention since not many dehalogenases that remove chloride from the  $\beta$ -carbon of chloroalkanoic acids [12, 13].

The genetic organization of *Rhizobial* dehalogenases and their operon was proposed based on gene cloning analysis. The location of *dehE* and *dehD/dehL* were far from clear. However, it was strongly believed that a single regulator gene (*dehR*) controls all three dehalogenases [2]. To understand how dehalogenases is regulated, soil bacteria like *Mesorhizobium* species is studied due to the fact that its full genome sequence readily available in the genebank [14]. Therefore, soil bacteria which is closely related to *Rhizobium* sp. RC1 like *Mesorhizobium* will be further investigated. There were several other bacteria species with the ability to degrade pollutants had been reported [15-18]. However, not many studies reported on the use of whole-genome sequencing to reveal their genetic sources of pollutant degradation potential [14]. The current study is to screen several *Mesorhizobium* strains for the presence of specific dehalogenases. By nature, the *Mesorhizobium* isolated from soil, regarded as rhizosphere bacterial of agronomic significance due to their ability to form nitrogen-fixing interactions with leguminous plants [19-21]. Analyzing dehalogenase genes from this bacterium will predict other possible functions aside from nitrogen-fixation.

Herein, we partially analyze the complete genome sequence of *Mesorhizobium loti* strains in view of predicting their pollutant's degradation ability just by focusing on the dehalogenase enzymes family. The metabolic novelty of *M. loti* strain TONO and its potential for use in bioremediation provided the motivation for complete genome sequence analysis. Here, we describe bioinformatics and phenotypic analysis that reveals

diverse metabolic capabilities of strain TONO that may be useful for bioremediation of halocetate and haloacid pollutants [20]. The dehalogenase genes arrangement including their upstream and downstream regions and their non-coding sequences for the possible regulator and promoter binding sites in the complete genome sequence were investigated. The newly identified dehalogenase genes were aligned using multiple sequence comparison and phylogenetic tree. We also highlighted on the regulatory element that might present and predicted to be potential in dehalogenase regulatory mechanism for *M. loti* strain TONO. This approach will provide an integrative knowledge to understand the complete bioremediation process.

## Material and Methods

### Whole genome sequence retrieval

Altogether, twenty-five genomes of *M. loti* strains were retrieved, and the annotated genomes were downloaded from <http://www.ncbi.nlm.nih.gov/genome> (GenBank accession number listed in Table 1). The dehalogenase genes within these genomes were searched, using the dehalogenase as a keyword search. Then the number and types of dehalogenase genes were recorded.

### Phylogenetic analysis of dehalogenase amino acid sequences

Phylogenetic analysis is a highly reliable bioinformatics approach to find out the differences and identities among genes, proteins, taxa and organisms [22]. The phylogenetic analyses of dehalogenases from *M. loti* strain TONO with other dehalogenases, L-DEX (*Pseudomonas* sp. Strain YL; AAB32245), DehH109 (*P. putida* 109; BAA04474), DehII (*P. putida* PP3; CAB38090), DehCI (*Pseudomonas* sp. CBS3; AAA63640), DehIVa (*Pseudomonas cepacia* MBA4; CAA46976), DehL (*Rhizobium* sp. RC1; CAA63794) were carried out. The phylogenetic tree was constructed using the CLUSTAL-W and MEGA7 software package [23, 24]. The sequence percentage identified among dehalogenases was then auto-calculated from Clustal 2.1 Identity Matrix [25].

### Multiple sequence alignment and conserved amino acid residues

The amino acid sequence of selected dehalogenases from the genome sequence of *M. loti* strain TONO (GenBank accession: AP017605 to AP017608) [20] were retrieved. The amino acid residues of all type II dehalogenases were aligned

Table 1. Complete genome features of *M. loti* from various strains. TypeII and HKD (Haloalkane Dehalogenases) showing number of dehalogenase genes

Bacteria strains	Size (MB)	G+C%	CDS	Dehalogenase		Accession number
				TypeII	HKD	
TONO	8.452	62.5	7831	4	1	NZ_AP017605.1
NZP2037	7.481	62.7	6954	3	1	NZ_CP016079.1
DSM 2626	7.451	62.4	6987	1	0	NZ_QGGH01000001.1
LMG 6125	4.879	62.6	4489	1	0	NZ_PYJO01000018.1
NZP2042	6.842	63.1	6403	1	1	NZ_LYTK00000000.1
NZP2037	7.462	62.8	6952	3	1	NZ_KB913026.1
R7ANS::ICEMISym2042	6.544	62.9	6089	3	1	NZ_LZTJ01000001.1
R7ANS::ICEMcbSym1271	6.508	62.9	6084	3	1	NZ_LZTK01000001.1
SU343	7.182	62.9	6765	3	1	NZ_LYTL01000001.1
R88b	7.195	62.4	6709	2	1	NZ_KI912159.1
R7ANS::ICEMISym343	6.566	62.9	6113	3	1	NZ_LZTL01000001.1
R7ANSxCC1192	6.458	62.9	6009	3	1	NZ_NSFO01000001.1
R7ANS::ICEMISym2014	6.525	62.9	6078	3	1	NZ_LZTI01000001.1
R7ANS::ICEMISym2037	6.586	62.9	6125	3	1	NZ_LZTH01000001.1
R7ANSx1497	6.486	63	6041	3	1	NZ_NSFS01000001.1
R7ANSxAA22	6.835	62.8	6271	3	1	NZ_NSFP01000001.1
NZP2014	6.577	62.9	6109	2	0	NZ_LYTJ01000001.1
R7ANSx3876	6.480	63	6043	3	1	NZ_NSFQ01000001.1
R7ANSx3859	6.436	63	6015	3	1	NZ_NSFR01000001.1
LU	6.399	61.5	5951	1	0	NZ_PJRO01000001.1
CJ3sym	7.563	62.2	6954	2	1	NZ_AXAL01000001.1
UFLA 01-765	7.464	63	5958	2	1	NZ_LPWA01000001.1
NZP2042	6.870	63.1	6418	1	1	NZ_CP033334.1
R88b	7.207	62.4	6698	2	1	NZ_CP033367.1
SU343	7.200	62.8	6782	3	1	NZ_CP033368.1

using the multiple sequence alignment method by Multalin ([multalin.toulouse.inra.fr/multalin/](http://multalin.toulouse.inra.fr/multalin/)) [26] and was enhanced by the ESPRIPT program [27] to look for conserved amino acid residues.

### Locus prediction and genetic analysis of dehalogenase genes

To define the potential operon of the genetic arrangement, the FGENESH program available in Softberry tools (<http://www.softberry.com/>) was utilised to check for the presence of the multiple genes in genomic DNA sequences. The software also predicts operon on ORFs and its neighboring genes of the bacterial genome, promoter and terminator prediction [28]. The InterPro (<https://www.ebi.ac.uk/interpro/>) was used for the classification of protein families.

## Results and Discussions

### Screening of dehalogenase genes

An advent of genome sequencing coupled with advances in bioinformatics promises invaluable insights into genomic analysis, leading to new knowledge in gene evolution, ecology, and the design of related therapeutic interventions. Pre-

viously, about 1,800 bacterial genomes have been fully sequenced, and these cover most of the major bacterial pathogens [29]. Genome sequencing can help in identifying an organism. When multiple genomes of the same or closely related species present, the 'pan-genome' can be designed and analysed as a single comprehensive catalog of all the sequences in the population [30]. In current study, 25 genome sequences of various *M. loti* strains were screened for dehalogenases. The dehalogenase genes were identified within these genomes alongside of size, G+C content (%) and protein-coding sequences (CDS) as summarised in Table 1. Haloacid dehalogenase type II (typeII) and haloalkane dehalogenase (HKD) were mostly identified (Table 1). The total percentage of the proportion of haloalkane and haloacid dehalogenase type II was 26% and 74% from the total of all dehalogenase genes, respectively. It was a common feature that many soil bacteria produce more than one dehalogenases. Many of these dehalogenases were identified in their chromosomes, although some were in the plasmids. However, no dehalogenase gene was detected in the plasmids of these strains. The presence of multiple dehaloge-

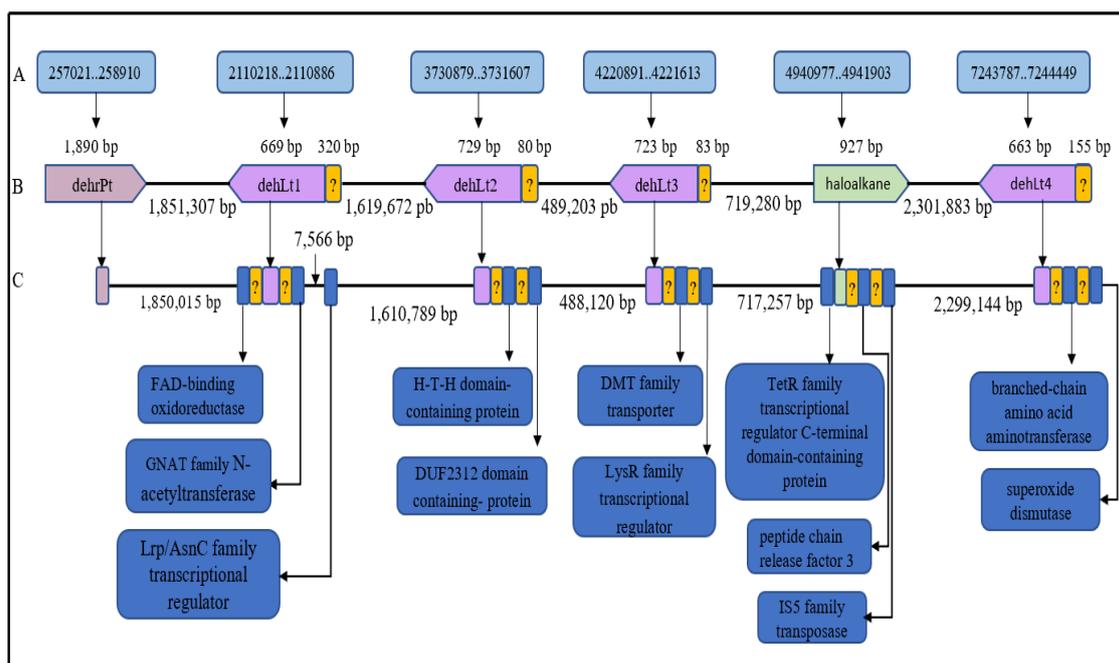


Figure 1. Genetic organization of dehalogenase genes of *M. loti* TONO. (A) The gene location within the genome. (B) Arrows indicate the direction of transcription, *dehrPt* (MLTONO\_0284) represents the haloacid uptake gene. The dehalogenases are designated as *dehLt1* (MLTONO\_2099), *dehLt2* (MLTONO\_3660), *dehLt3* (MLTONO\_4143), *dehLt4* (MLTONO\_6945). The question mark (?) indicates non-coding regions. (C) The upstream and downstream genes of dehalogenases.

nases were not fully understood. Haloacid dehalogenase type II was common compared to D-specific dehalogenases since most naturally occurring halogenated compounds exist in the L-form [15]. This is the first report on *M. loti* strain TONO produced four haloacid dehalogenases apart from *Rhizobium* sp. RC1 [31, 32]. These facts raised a fascinating question as to why *M. loti* strain TONO produces four haloacid dehalogenase type II. It was suggested that the ability to produce all four dehalogenases may afford the *M. loti* strain TONO the ability to degrade or utilize high amount of halogenated organic pollutants thus, suggesting that it might be a good candidate used for bioremediation of halogenated organic compounds or used in the synthesis of chiral active industrial chemical intermediates.

### Genetic organisation of dehalogenase genes of *M. loti* TONO

The full chromosome of *M. loti* strain TONO sequence was further analysed to search for regulatory gene functions in relation to the control of all dehalogenases from the perspective of genome studies. *M. loti*, also known as *Rhizobium loti*, is fast-growing Rhizobia that lives in symbiosis with the genus *Lotus* sp. legumes. These bacteria infect leguminous plants to form root nodules, resulting

in the exchange of fixed nitrogen and carbon between rhizobia and plants from the atmosphere [33]. The genetic organization of *M. loti* strain TONO related to dehalogenase has not been studied so far from the whole genome sequence [20]. The genome consists of a chromosome (7,856,088 bp; 62.8% G+C content) and three plasmids (80,491 bp; 60.4% G+C, 294,703 bp; 59.5% G+C and 220,869 bp; 58.9% G+C content (%)). The genome sequence was deposited in DDBJ/EMBL/GenBank under the accession number AP017605 to AP017608 and annotated by the NCBI Prokaryotic Genome Annotation Pipeline [20].

The location of dehalogenase accessory genes was investigated such as regulatory and an uptake gene. In this study, a putative uptake gene (MLTONO\_0284) that encodes for Major Facilitator Superfamily protein (designated as *DehrPt*) was identified. The gene is possibly responsible for haloacid uptake protein. All four haloacid dehalogenase type II were present and designated as *dehLt1* (MLTONO\_2099), *dehLt2* (MLTONO\_3660), *dehLt3* (MLTONO\_4143) and *dehLt4* (MLTONO\_6945), which were 669 bp, 729 bp, 723 bp, and 663 bp in size, respectively (Figure 1). One haloalkane dehalogenase gene (MLTONO\_4828) was also present with 927 base pairs (bp) in size (Figure 1). The distance between

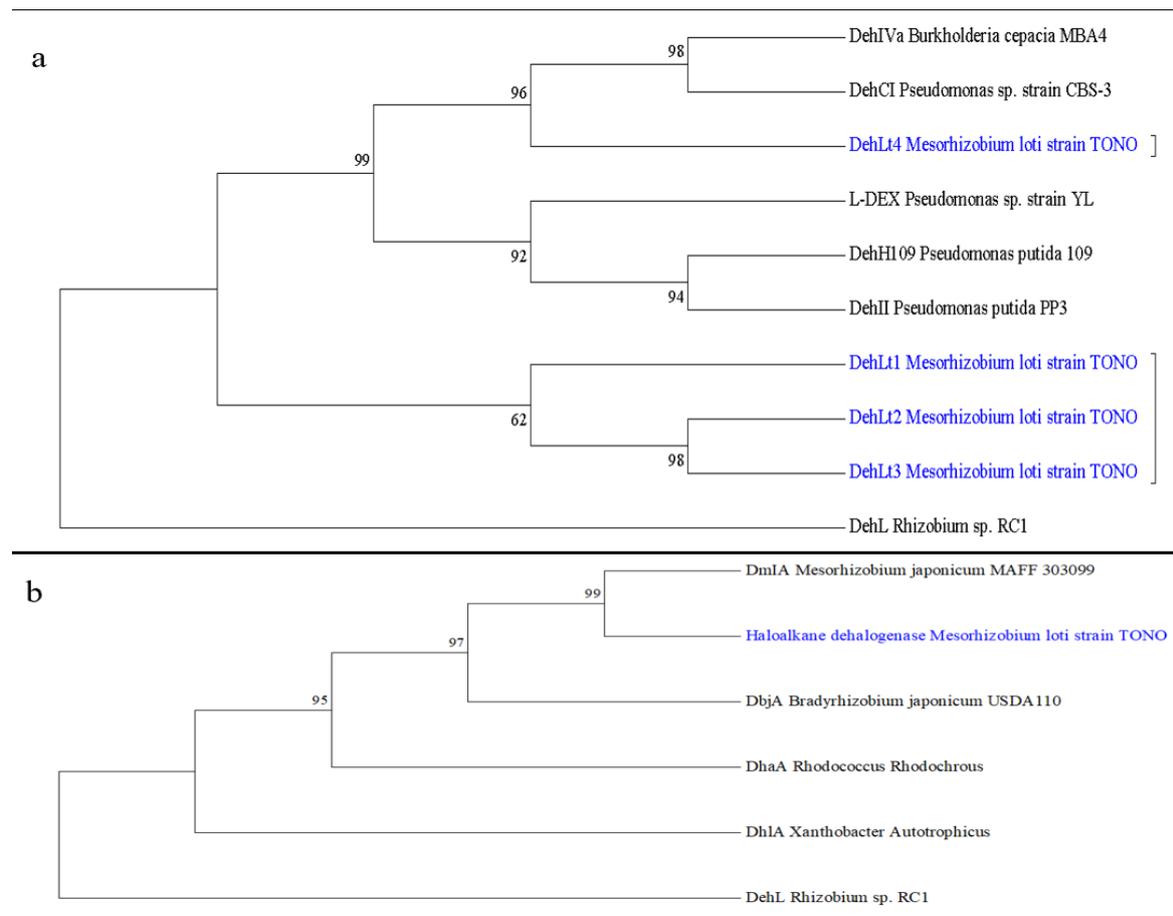


Figure 2. The Phylogenetic tree constructed by using the Maximum Likelihood method [38]. Bootstrap replications were 1000. The dehalogenases of the current study are indicated in blue. The analysis involved 10 haloacid and 6 haloalkane dehalogenases. The evolutionary analyses were conducted in MEGA7 [23].

*dehLt1*, *dehLt2*, *dehLt3*, haloalkane dehalogenase gene and *dehLt4* were 1,619,992 bp, 489,283 bp, 719,363 bp and 2,301,883 bp, respectively (Figure 1). The genes were not located near to each other in the genome. The sequences at the upstream position of each *dehLt1*, *dehLt2*, *dehLt3*, and *dehLt4* were 320 bp, 80 bp, 83 bp, and 155 bp, respectively, and were not encoded for any functional proteins and possibly consist of specialised promoters for each dehalogenase genes (Figure 1).

The genome analysis by the FGENESH server showed that these dehalogenase genes were not located in a single operon to have a single regulator gene. Unlike dehalogenases from *Rhizobium* sp. RC1, all dehalogenases were located as one operon and regulated by a single regulator gene [2, 15]. Herein, we hypothesized that dehalogenases of *M. loti* strain TONO may not possibly be controlled by one regulator gene since they are not located near to each other in the genome. Therefore, further analyses of all these dehalogenases, such

as site-directed mutagenesis (SDM) is needed to clearly state their functions.

The *dehLt1* forms an operon with the gene (MLTONO\_2098) encoding FAD-binding oxidoreductase. The adjacent upstream of *dehLt1* is a gene encoding GNAT family N acetyltransferase. There is no regulator gene adjacent to these genes except the Lrp/AsnC family transcriptional regulator (MLTONO\_2107), located six genes (7,566 bp) downstream of the GNAT family N-acetyltransferase gene (Figure 1C). The *dehLt2*, a gene (MLTONO\_3661) encoding helix-turn-helix domain-containing protein and a gene (MLTONO\_3662) encoding DUF2312 domain containing protein were located in an operon (Figure 1C). Analysis by InterPro showed the DUF2312 domain-containing protein as a hypothetical protein. Besides, helix-turn-helix domain-containing protein was classified as helix-turn-helix (HTH) motif family DNA-binding protein regulator and possibly regulates *dehLt2*. The *dehLt3*

and a gene (MLTONO\_4144) encoding the DMT family transporter were organized as an operon. While upstream of the latter gene, a gene (MLTONO\_4145) encoding LysR family transcriptional regulator exists, and it is not included in any operon (Figure 1C). The regulation of the dehalogenase gene, *hdhA* in *Sinorhizobium meliloti* strain 1021 by LysR family transcriptional regulator, HdhR, is previously reported, which is located upstream of the *hdhA*. HdhR represses the transcription of *hdhA* via direct binding of HdhR to the *hdhA* promoter [34]. It should be mentioned here that, the finding of this study corroborated with previous studies in which most dehalogenase regulators are located adjacent upstream of dehalogenase structural genes [15, 34-37].

The *dehLt4* gene is not organized in any operon. In the annotated genome, the two genes upstream of *dehLt4* are genes encoding branched-chain amino acid aminotransferase and superoxide dismutase (Figure 1C). Haloalkane dehalogenase (HKD) is organized in an operon with the TetR family transcriptional regulator C-terminal domain-containing protein gene (MLTONO\_4827), peptide chain release factor 3 gene (MLTONO\_4829), and IS5 family transposase gene (MLTONO\_4830) (Figure 1C).

### Phylogenetic analysis of dehalogenases from *M. loti* strain TONO

Sequence comparison of the dehalogenase amino acid from *M. loti* strain TONO with established dehalogenase amino acids suggested that all four dehalogenases were highly identical to group II. Therefore, multiple sequence alignment by CLUSTAL-W and Maximum Likelihood techniques [38] were used to identify the evolutionary relationships of all dehalogenases from *M. loti* strain TONO compared to the well-established group II dehalogenase as reported by Hill *et al.* [10]. In Figure 2a, DehLt4 is closely related to DehCI and DehIVa that possesses 49 % and 48 % identity, respectively and has a common ancestor. The L-DEX, DehH109, and DehII are closely related and share a common ancestor. The DehLt1, DehLt2 and DehLt3 have a common ancestor in which DehLt2 and DehLt3 share 50 % identity while DehLt1 have 28% identity with both DehLt2 and DehLt3 (Figure 2a). DehL from *Rhizobium* sp. RC1 is not closely related to all members of group II dehalogenase and therefore, it was outgroup in the phylogenetic analysis. Haloalkane dehalogenase is closely related to DmlA of *Mesorhizobium japonicum* MAFF 303099, sharing 88% identity (Figure 2b).

### Multiple sequence alignment

Multiple sequence alignments were used to describe the level of conserved regions within the dehalogenases of *M. loti* strain TONO (Figure 3). The sequence-based identification of the important amino acid residues is sometimes difficult due to the low percentage of conserved amino acids in such enzymes. Notably, catalytically important amino acids are highly conserved in dehalogenase type II [15]. The nine catalytic important residues in L-DEX, DhIB and DehIVa have been previously reported. Asp11 in DehIVa acts as a nucleophilic by attacking carbon no. 2 of the substrate during catalysis [39]. The corresponding residues in L-DEX (Asp10) [40] and DhIB (Asp8) [41] were reported to have a similar function. Ser176 in DehIVa and its corresponding residue Ser175 in L-DEX, formed a hydrogen bond with nucleophile Asp to probably maintain the orientation of its carboxyl group in a way suitable to attack the carbon no. 2 of the substrate [39, 42]. The role of Thr14 in L-DEX is not yet identified, however its corresponding residue in DhIB (Thr12) together with Ser171 and Asn173 were reported to firmly hold the Asp8 in a position that favors the nucleophilic attack [41]. In L-DEX, Lys151 and Asp180 interact strongly with Asp10, L-2-chloropropionate and catalytic water. Lys151 stabilises the substrate orientation and balance the charges around the active site. Asp180 stabilises the Asp10 rotation, prevents Lys151 from moving toward Asp10 and fixes the catalytic water in proximity to Asp10. Asp180 may independently activate the catalytic water or in collaboration with Lys151, Ser175, and Asn177 [43]. Mutation of all these functional amino acids by site-directed mutagenesis significantly reduced the enzyme activity [44]. Arg41 in L-DEX accepts the released chloride ion [45] and its corresponding residue in DehIVa (Arg42) is important in the substrate “lock down” mechanism; and also acts a member of the halide- binding cradle together with Asn120 and Trp180 [39]. The alteration of Asp11 and Asn178 residues in DehIVa turn the protein to inactive [46].

In this study, it was possible to predict a conserved catalytic residue of dehalogenases in *M. loti* strain TONO by referring to DehIVa and L-DEX (Figure 3). The conserved amino acid residues of DehLt1, DehLt2, DehLt3 and DehLt4 were summarised in Table 2. The percentage of sequence identities of DehIVa with dehalogenases from *M. loti* strain TONO were: 21.00% (DehLt1), 23.64% (DehLt2), 22.37% (DehLt3) and 48.18% (DehLt4), whereas with L-DEX were 23.62%



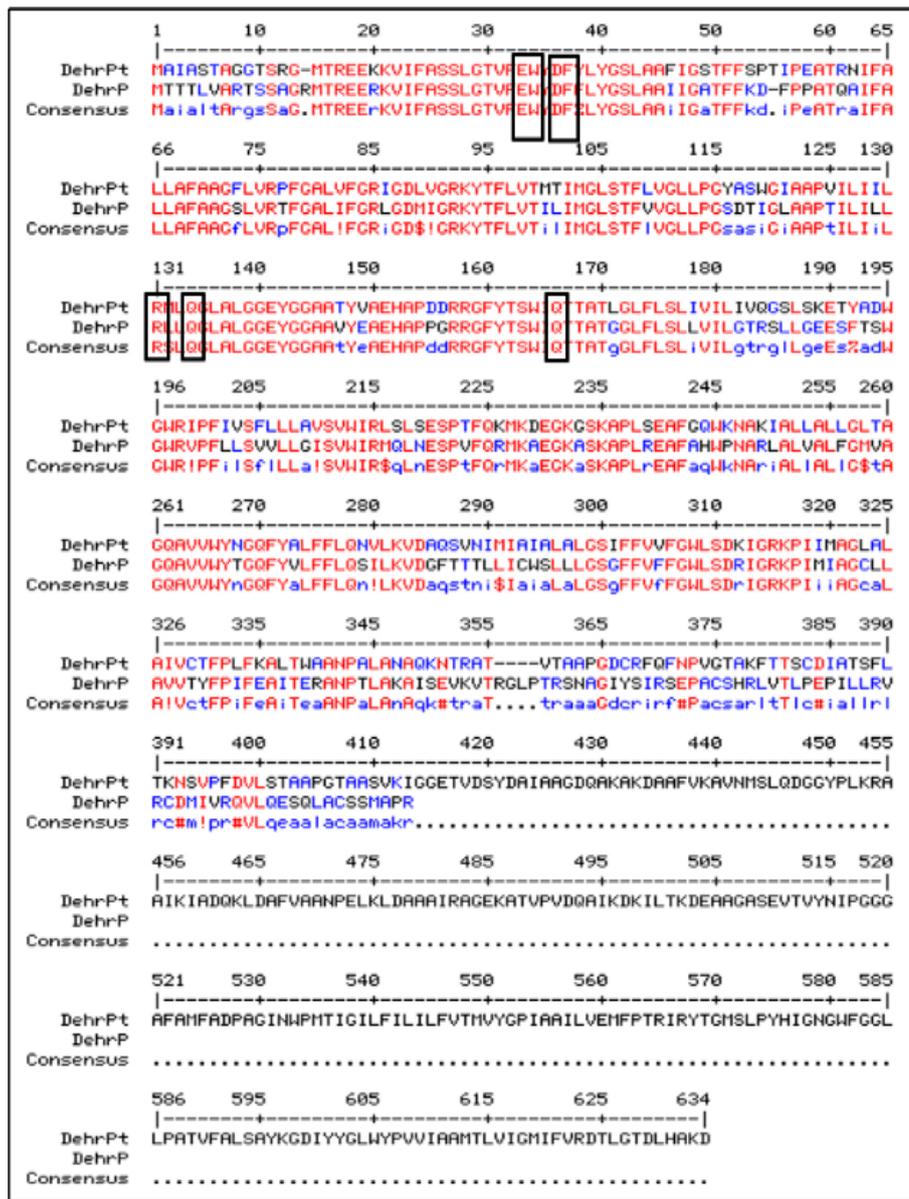


Figure 4. Pairwise sequence alignment of DehrPt (strain TONO) and DehrP (*Rhizobium* sp. RC1; accession number CAJ98618.1). Amino acids with high consensus values were highlighted in red; In box: conserved amino acid residues.

(DehLt1), 21.00% (DehLt2), 22.94% (DehLt3) and 42.73% (DehLt4).

The sequence alignment of four haloacid dehalogenases from strain TONO were also compared with group I dehalogenases. According to Hill *et al.* [10] five essential catalytic residues (Thr65, Glu69, Asn117, Tyr120, and Asp194) reported by Nardi dei *et al.* [47] in DL-DEX were supposed to be conserved in almost all group I dehalogenases. However, the multiple sequence alignment neither show any significant identity or similarity. Therefore, all four haloacid dehalogenase sequences found in *M. loti* strain TONO were

proposed belongs to type II.

#### Analysis of an uptake protein of DehrPt from *M. loti* strain TONO

According to the FGESH, *dehrPt* gene (MLTONO\_0284) is not organize in any operon. Pairwise analysis of DehrPt and DehrP from *Rhizobium* sp. RC1 [48] showed 69% amino acid sequence identity. Asp36(Asp35), Arg130(Arg130) and Gln133(Gln133) which were H<sup>+</sup>-binding sites were conserved to each other. The interacting amino acid residues of DehrP (*Rhizobium* sp. RC1) and DehrPt (strain TONO) depicted as

Glu33(Glu32), Trp34 (Trp33), Phe37(Phe36), Gln165(Gln165) (Figure 4). The sequence analysis suggested that DehrPt is equivalent to DehrP function as haloacid uptake protein in *M. loti* strain TONO.

## Conclusion

L-2-Haloacid dehalogenases (type II) are common and this is consistent with the current study. Most dehalogenase genes were not organized as a single operon and scattered within its chromosomes and their expression is regulated differently referring to helix-turn-helix (HTH) motif family DNA-binding protein regulator and LysR family transcriptional regulator. The dehalogenase genes sequence is subjected to further analysis using bioinformatic tools for their protein structure and functions.

## Acknowledgment

The authors would like to express their gratitude to Geran Universiti Penyelidikan-UTM Number Q.J130000.2414.08G59 & FRGS R.J130000.7854.5F189 for financial assistance. SZ thanks the Afghanistan Higher Education Ministry for MSc. studentship.

## References

- Edbeib MF, Wahab RA, Huyop FZ et al. (2020) Further Analysis of *Burkholderia pseudomallei* MF2 and Identification of Putative Dehalogenase Gene by PCR. Indonesian Journal of Chemistry 20 (2): 86 – 394. doi: 10.22146/ijc.43262.
- Adamu A, Wahab RA, Aliyu F et al. (2020) Haloacid dehalogenases of *Rhizobium* sp. and related enzymes: catalytic properties and mechanistic analysis. Process Biochemistry 92: 437 – 446. doi: 10.1016/j.procbio.2020.02.002.
- Heidarrezaei M, Shokravi H, Huyop F et al. (2020) Isolation and Characterization of a Novel Bacterium from the Marine Environment for Trichloroacetic Acid Bioremediation. Applied Sciences 10 (13): p. 4593. doi: 10.3390/app10134593.
- Pries F, Kingma J, Pentenga M et al. (1994) Site-Directed Mutagenesis and Oxygen Isotope Incorporation Studies of the Nucleophilic Aspartate of Haloalkane Dehalogenase. Biochemistry 33 (5): 1242-1247. doi: 10.1021/bi00171a026.
- Zhang Y, Li ZS, Wu JY et al. (2004) Homology modeling and SN2 displacement reaction of fluoroacetate dehalogenase from *Burkholderia* sp. FA1. Biochemical and Biophysical Research Communications 325 (2): 414-420. doi: 10.1016/j.bbrc.2004.10.044.
- Allison N, Skinner AJ, Cooper RA (1983) The dehalogenases of a 2, 2-dichloropropionate-degrading bacterium. Microbiology 129 (5): 1283-1293. doi: 10.1099/00221287-129-5-1283.
- Slater JH, Bull AT, Hardman DJ (1996) Microbial dehalogenation of halogenated alkanic acids, alcohols and alkanes. Advances in Microbial Physiology 38: 133-176. doi: 10.1016/S0065-2911(08)60157-5.
- Slater JH, Lovatt D, Weightman AJ et al. (1979) The growth of *Pseudomonas putida* on chlorinated aliphatic acids and its dehalogenase activity. Microbiology 114 (1): 125-136. doi: 10.1099/00221287-114-1-125.
- Yang G, Liang PH, Dunaway-Mariano D (1994) Evidence for Nucleophilic Catalysis in the Aromatic Substitution Reaction Catalyzed by (4-Chlorobenzoyl) coenzyme A Dehalogenase. Biochemistry 33 (28): 8527-8531. doi: 10.1021/bi00194a018.
- Hill KE, Marchesi JR, Weightman AJ (1999) Investigation of Two Evolutionarily Unrelated Halocarboxylic Acid Dehalogenase Gene Families. Journal of Bacteriology, 181 (8): 2535-2547. doi: 10.1128/JB.181.8.2535-2547.1999.
- Slater JH, Bull AT, Hardman DJ (1995) Microbial dehalogenation. Biodegradation 6 (3): 181-189. doi: 10.1007/BF00700456.
- Muslem WH, Edbeib MF, Aksoy HM et al. (2020) Biodegradation of 3-chloropropionic acid (3-CP) by *Bacillus cereus* WH2 and its in-silico enzyme-substrate docking analysis. Journal of Biomolecular Structure and Dynamics 38 (11): 3432-3441. doi: 10.1080/07391102.2019.1655482.
- Mesri S, Wahab RA, Huyop F (2009) Degradation of 3-chloropropionic acid (3CP) by *Pseudomonas* sp. B6P isolated from a rice paddy field. Annals of Microbiology, 59 (3): 447-451. doi: 10.1007/bf03175129.
- Chan WY, Wong M, Guthrie J et al. (2010) Sequence- and activity-based screening of microbial genomes for novel dehalogenases. Microbial Biotechnology 3 (1): 107-120. doi: 10.1111/j.1751-7915.2009.00155.x.
- Adamu A, Wahab RA, Huyop F (2016) L-2-Haloacid dehalogenase (DehL) from *Rhizobium* sp. RC1. SpringerPlus 5 (1): 1-17.
- Oyewusi HA, Huyop F, Wahab RA (2020) Molecular docking and molecular dynamics simulation of *Bacillus thuringiensis* dehalogenase against haloacids, haloacetates and chlorpyrifos. Journal of Biomolecular Structure and Dynamics: 1-16. doi: 10.1080/07391102.2020.1835727.
- Wahhab BH, Anuar NFSK, Wahab RA et al. (2020) Characterization of a 2,2-dichloropropionic acid (2,2-DCP) degrading alkalotolerant *Bacillus megaterium* strain BHS1 isolated from Blue Lake in Turkey. Journal of Tropical Life Science 10 (3): 245 – 252. doi: 10.11594/jtls.10.03.08.
- Zang H, Wang H, Miao L et al. (2020) Carboxylesterase, a de-esterification enzyme, catalyzes the degradation of chlorimuron-ethyl in *Rhodococcus erythropolis* D310-1. Journal of Hazardous Materials, 387: p. 121684. doi: 10.1016/j.jhazmat.2019.121684.
- Azam S, Parthasarathy S, Singh C et al. (2019) Genome organization and adaptive potential of archetypal organophosphate degrading *Sphingobium fuliginis* ATCC 27551. Genome Biology and Evolution 11 (9): 2557-2562. doi: 10.1093/gbe/evz189.
- Shimoda Y, Hirakawa H, Sato S et al. (2016) Whole-genome sequence of the nitrogen-fixing symbiotic *Rhizobium Mesorhizobium loti* strain TONO. Genome Announcements 4 (5). doi: 10.1128/genomeA.01016-16.
- Yamaya-Ito H, Shimoda Y, Hakoyama T et al. (2018) Loss-of-function of Aspartic peptidase nodule-induced 1

- (APN 1) in *Lotus japonicus* restricts efficient nitrogen-fixing symbiosis with specific *Mesorhizobium loti* strains. *The Plant Journal* 93 (1): 5-16. doi: 10.1111/tpj.13759.
22. Roy SS, Dasgupta R, Bagchi A (2014) A review on phylogenetic analysis: a journey through modern era. *Computational Molecular Bioscience* 4 (03): 39. doi: 10.4236/cmb.2014.43005.
  23. Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution* 33 (7): 1870-1874. doi: 10.1093/molbev/msw054.
  24. Kumar S, Stecher G, Li M et al. (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution* 35 (6): 1547-1549. doi: 10.1093/molbev/msy096.
  25. Madeira F, Park YM, Lee J et al. (2019) The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Research* 47 (W1): W636-W641. doi: 10.1093/nar/gkz268.
  26. Corpet F (1988) Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Research* 16 (22): 10881-10890. doi: 10.1093/nar/16.22.10881.
  27. Robert X, Gouet P (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research* 42 (W1): W320-W324. doi: 10.1093/nar/gku316.
  28. Salamov VSA, Solovyevand A (2011) Automatic annotation of microbial genomes and metagenomic sequences. In: *Metagenomics and its Applications in Agriculture, Biomedicine and Environmental Studies*. Hauppauge, Nova Science Publishers. 61-78.
  29. Ribeiro FJ, Przybylski D, Yin S et al. (2012) Finished bacterial genomes from shotgun sequence data. *Genome Research* 22 (11): 2270 - 2277. doi: 10.1101/gr.141515.112.
  30. Marcus S, Lee H, Schatz MC (2014) SplitMEM: a graphical algorithm for pan-genome analysis with suffix skips. *Bioinformatics* 30 (24): 3476-3483. doi: 10.1093/bioinformatics/btu756.
  31. Cairns SS, Cornish A, Cooper RA (1996) Cloning, sequencing and expression in *Escherichia coli* of two *Rhizobium* sp. genes encoding haloalkanoate dehalogenases of opposite stereospecificity. *European Journal of Biochemistry* 235 (3): 744-749. doi: 10.1111/j.1432-1033.1996.t01-1-00744.x.
  32. Stringfellow JM, Cairns SS, Cornish A, Cooper RA (1997) Haloalkanoate dehalogenase II (DehE) of a *Rhizobium* sp.—Molecular analysis of the gene and formation of carbon monoxide from trihaloacetate by the enzyme. *European Journal of Biochemistry* 250 (3): 789-793. doi: 10.1111/j.1432-1033.1997.00789.x.
  33. Kasai-Maita H, Hirakawa H, Nakamura Y et al. (2013) Commonalities and differences among symbiosis islands of three *Mesorhizobium loti* strains. *Microbes and Environments* 28 (2): 275-278. doi: 10.1264/jsme2.ME12201.
  34. Sallabhan R, Kerdwong J, Dubbs JM et al. (2013) The *hdhA* gene encodes a haloacid dehalogenase that is regulated by the LysR-type regulator, *HdhR*, in *Sinorhizobium meliloti*. *Molecular Biotechnology* 54 (2): 148-157. doi: 10.1007/s12033-012-9556-1.
  35. Krasper L, Lilie H, Kublik A et al. (2016) The MarR-type regulator *Rdh2R* regulates *rdh* gene transcription in *Dehalococcoides mccartyi* strain CBDB1. *Journal of Bacteriology* 198 (23): 3130-3141. doi: 10.1128/JB.00419-16.
  36. Wagner A, Segler L, Kleinstueber S et al. (2013) Regulation of reductive dehalogenase gene transcription in *Dehalococcoides mccartyi*. *Philosophical Transactions of the Royal Society B: Biological Sciences* 368 (1616): 20120317. doi: 10.1098/rstb.2012.0317.
  37. La Roche SD, Leisinger T (1991) Identification of *DcmR*, the regulatory gene governing expression of dichloromethane dehalogenase in *Methylobacterium* sp. strain DM4. *Journal of Bacteriology* 173 (21): 6714-6721. doi: 10.1128/jb.173.21.6714-6721.1991.
  38. Whelan S, Goldman N (2001) A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Molecular Biology and Evolution* 18 (5): 691-699. doi: 10.1093/oxfordjournals.molbev.a003851.
  39. Schmidberger JW, Wilce JA, Tsang JSH, Wilce MCJ (2007) Crystal structures of the substrate free-enzyme, and reaction intermediate of the HAD superfamily member, haloacid dehalogenase *DehIVa* from *Burkholderia cepacia* MBA4. *Journal of Molecular Biology* 368 (3): 706-717. doi: 10.1016/j.jmb.2007.02.015.
  40. Liu JQ, Kurihara T, Miyagi M et al. (1995) Reaction mechanism of L-2-haloacid dehalogenase of *Pseudomonas* sp. YL: identification of Asp10 as the active site nucleophile by 18O incorporation experiments. *Journal of Biological Chemistry* 270 (31): 18309-18312. doi: 10.1074/jbc.270.31.18309.
  41. Ridder IS, Rozeboom HJ, Kalk KH et al. (1997) Three-dimensional structure of L-2-haloacid dehalogenase from *Xanthobacter autotrophicus* GJ10 complexed with the substrate-analogue formate. *Journal of Biological Chemistry* 272 (52): 33015-33022. doi: 10.1074/jbc.272.52.33015.
  42. Hisano T, Hata Y, Fujii T et al. (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. *Journal of Biological Chemistry* 271 (34): 20322-20330. doi: 10.1074/jbc.271.34.20322.
  43. Nakamura T, Yamaguchi A, Kondo H et al. (2009) Roles of K151 and D180 in L-2-haloacid dehalogenase from *Pseudomonas* sp. YL: Analysis by molecular dynamics and ab initio fragment molecular orbital calculations. *Journal of Computational Chemistry* 30 (16): 2625-2634. doi: 10.1002/jcc.21273.
  44. Kurihara T, Liu JQ, Nardi-Dei V et al. (1995) Comprehensive site-directed mutagenesis of L-2-halo acid dehalogenase to probe catalytic amino acid residues. *The Journal of Biochemistry* 117 (6): 1317-1322. doi: 10.1093/oxfordjournals.jbchem.a124861.
  45. Kondo H, Nakamura T, Tanaka S (2014) A significant role of Arg41 residue in the enzymatic reaction of haloacid dehalogenase L-DEX YL studied by QM/MM method. *Journal of Molecular Catalysis B: Enzymatic* 110: 23-31. doi: 10.1016/j.molcatb.2014.09.006.
  46. Pang BCM, Tsang JSH (2001) Mutagenic analysis of the conserved residues in dehalogenase IVa of *Burkholderia cepacia* MBA4. *FEMS Microbiology Letters* 204 (1): 135-140. doi: 10.1111/j.1574-6968.2001.tb10876.x.
  47. Nardi-Dei V, Kurihara T, Park C et al. (1997) Bacterial DL-2-haloacid dehalogenase from *Pseudomonas* sp.

- strain 113: gene cloning and structural comparison with D-and L-2-haloacid dehalogenases. Journal of Bacteriology 179 (13): 4232-4238. doi: 10.1128/jb.179.13.4232-4238.1997.
48. Musa MA, Wahab RA, Huyop F (2018) Homology modelling and in silico substrate-binding analysis of a *Rhizobium* sp. RC1 haloalkanoic acid permease. Biotechnology & Biotechnological Equipment 32 (2): 339-349. doi: 10.1080/13102818.2018.1432417.

This page is intentionally left blank.