J. Sci. Inform. Soc. 2023, Vol. 1, No. 1
ISSN XXXX-XXXX

Exploration of Gene Encoded Thermostable Enzymes by Using Random PCR from Natural Sample of Domas Crater

Suharti[1*], Saniya Almira[1], Akhmaloka[2]

[1]Department of Chemistry, Universitas Pertamina, Jakarta 12220, Indonesia
[2]Biochemistry Research Group, Institut Teknologi Bandung, Jl. Ganesha 10 Bandung, Indonesia
*Corresponding Author: suharti.s@universitaspertamina.ac.id
Received: 27 December 2022
Accepted: 6 January 2023

Abstracts

Thermophilic microorganisms can survive and live in extremely high-temperature habitats. This special trait made thermophilic microorganisms have heat-resistant proteins such as thermostable enzymes and have been used by the industry as biocatalysts. Some advantages of using thermostable enzymes in the industry are reducing operating costs, increasing the reaction rate, increasing productivity, and being environmentally friendly. This research aims to isolate the thermostable enzyme encoding gene and determine homology, differences in the sequence, and the closest relative of samples. The metagenomic approaches using random PCR techniques were used as a methodology in this research. The substances successfully obtained from random PCR of the metagenome samples were cloned in the cloning vector and then analyzed nucleotide sequence (sequencing). The highest protein homology in samples 23 and 24 are Short-chain Dehydrogenase Reductase (SDR) oxidoreductase [*Caldivirga maquilingensis*] of 58.58% and 58.02%, respectively. The 3-dehydroquinate dehydratase Type I [*Metallosphaera*] is the highest protein homology in sample 45, with a similarity of 79.88%. The phylogenetic tree analysis of sample 23 has the closest relationship with sample 24. Meanwhile, sample 45 has the closest relationship with 3-dehydroquinate dehydratase Type I from Metallospaera species. The three samples formed different branches with other amino acid sequences, showing that the three samples are new enzymes that are different from the previous ones.

Keywords: Thermostable enzymes, metagenomic approaches, protein homology, phylogenetics.

Introduction

Thermophiles have evolved to strive in extreme conditions such as high temperatures, highly alkaline or acidic conditions, and other conditions [1]. The reason behind this, they have the potential to produce valuable enzymes which can be active under extreme conditions, called extremozymes. Extremozymes are categorized into various groups, such as thermozymes (activity at high temperatures), halozymes (tolerance at high levels of salt), and alkalozymes (activity at alkaline conditions)[2].

Thermophiles can be a source of thermostable enzymes that catalyze specific reactions at high temperatures. In addition to providing stability, thermophiles generally increased resistance to denaturation and proteolysis. This advantage can be used in industrial processes. Furthermore, enzymes acquired from thermophiles have high thermostability and optimal activity above 70℃. Therefore, carrying out the process at a higher temperature would be feasible and reduce the risk of microbial combinations [1, 2].

One of the thermophile habitats in Indonesia is the Domas crater. Domas crater was located in Indonesia West Java. It has a temperature of around 93℃ - 95℃, allowing the potential of thermophilic microorganisms [3].

This research aims to isolate potential genetic sequences encoding thermostable enzymes from Domas crater, then determine protein homology using bioinformatics, the relationship of each sample using phylogenetic tree constructions, and distinguish amino acid sequences of each sample.

Material and Methods

Material

Materials used in this research are microbial genomic DNA from Domas Crater [4], aquabidest, agarose (Thermoscientific top vision agarose R0491), tryptone (HIMEDIA CR014-500G), yeast extract (HIMEDIA RM0270-500G), NaCl (MERCK 1.08122.055), tris acetate-*EDTA* (TAE), bacterial agar, CaCl2 (EMSURE1.04936.1000), ampicillin, tetracycline, EtBr, NaOH, SDS, isopropanol, EtOH, nuclease-free water (NFW), rapid disruption solution (KCl, EDTA, SDS, NaOH, Sucrose, BPB), Go-taq hot start green master mix (Promega ADM7122 000037110111), Kit CloneJET PCR Cloning Kit (Thermo Fisher Scientific), Primer Fjet and Rjet, Buffer, dNTP, MgCl2, Taq polymerase, and *E. coli* TOP10F' (Thermo Fisher Scientific).

Methods

DNA Amplification

Microbial DNA from Domas Crater was amplified using Polymerase Chain Reaction with two different types of primers, TM and T1509. These are nucleic acid sequences of both primers.

TMF: 5'- CTC TCG CTA GCT GGA AGC ATG TAA AAC AGG T -3'
TMR: 5'- GAG AGG TCG ACT CAT ATT TCC ACC CCT TCG ATC TTG -3,
T1509F: 5'- GTG TGG CTA GCT ATT TAA CAA GCC AGC GAT C -3'
T1509R: 5'- ACA CGG TCG ACT CAG ATC AGC TTT ATG TAT G -3'

PCR was performed using *thermal cycler* Bio-RadTM T100 and GoTaq Promega as reagent kits. The protocol used in this experiment is 94 oC, 2:00; 46 oC, 0:30; 72 oC, 1:15, 30 cycles; 72 oC, 10:00; 12 oC, 10:00. Further, samples were electrophoresed to observe the result.

Ligation of Genetically Amplified Sample to Vector

The ligation process was carried out using the CloneJET1.2/blut PCR Cloning Kit Thermo Fisher Scientific protocol.

Transformation and Isolation of DNA Plasmid

Samples were selected by size screening method where samples with the higher band during electrophoresis were selected. Selected samples were then inoculated in Luria Bertani liquid growth medium. To isolate the plasmid, sample cultures were centrifuged, and pellets were obtained. The solution I (glucose, Tris-Cl, EDTA) was added to the pellets. Furthermore, solution II was added (NaOH, 1% SDS). Lastly, the pellets were added to cold solution III (potassium acetate, acetic acid, and H2O). All the samples were centrifuged to gain the supernatant. The DNA sample should be separated from the vector for analyzing the plasmid. This process is called restriction and was conducted using a restriction enzyme, which in this experiment is *Bam*HI. Samples were electrophoresed to observe the results. The chosen results were sequenced by Macrogene inc, Korea, and performed analysis.

Bioinformatics Analysis

All of processes of bioinformatics analysis were performed using NCBI site (https://blast.ncbi.nlm.nih.gov/Blast.cgi), BioEdit, MEGA X, clustal x, and Genedoc.

Results and Discussion

The amplicon of DNA

DNA amplification is a process of copying multiple DNA sequences employing PCR. The initial PCR process was operated after extracting microorganism DNA from the sample environment. PCR will purify the desired target gene by amplifying it using a specific primer [5]. T1509 and TM. The result shows in Figure 1.

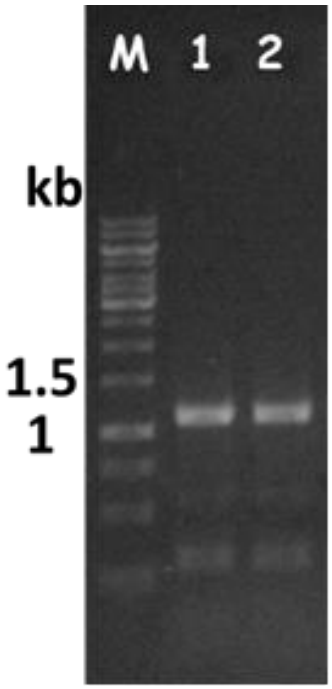


Figure 1. The electropherogram of the amplicon of samples visualized on agarose gel. M is DNA leader, 1 is the amplicon using T1509 primer, and 2 is the amplicon using TM primer.

After the amplification process, the next step is ligation. Ligation is inserting the target DNA sequence (DNA amplicon) into a vector using a restriction enzyme as a cutter, then glue back using DNA ligase [6]. Vectors are DNA molecules that act as vehicles to carry foreign genetic material to another cell, where it can be replicated or expressed (e.g., plasmid) [7].

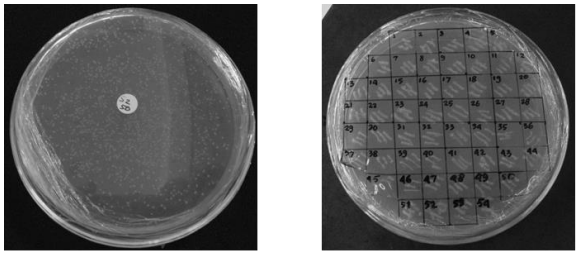
A plasmid inserted with DNA is called recombinant DNA. The results of the ligation can be vectors without DNA insertions, vectors with target DNA inserts, and vectors with unwanted DNA insertions. The ligation process was undertaken using the CloneJET PCR Cloning Kit Thermo Fisher Scientific protocol.

The transformants and the plasmids

Transformation is the stage after ligation. At this stage, the ligation results were inserted into competent cells. Entail more explanatory sentences. Competent cells are host cells that are modified to take up DNA. Producing competent cells can be taken by using a cold salt solution in CaCl2 form. Although the mechanism is not fully understood, $CaCl_2$ causes DNA to bind to the outer bacterial cell wall [8]. Salts may be responsible for specific changes in the bacterial cell wall that increase DNA binding [6].

Treatment using $CaCl_2$ did not affect DNA outside the cell to enter the cell. Hence, The heat shock method was executed by inserting DNA from the outside into the cell. Transference of the plasmid DNA into the competent cells when heat shock treatment was applied. This process would be incubated for several minutes. An increase of temperature by 42℃ for a short duration affects the motion of DNA to enter inside of the cell from outside.

The transformant in the form of a single colony was then transferred to numbered petri dishes (Fig. 2).



A B

Figure 2. The transformant colonies on selective agar media. A is all transformants, B is the transferred colonies on numbered petri dishes.

The single colonies that grew on selective agar media were screened with electrophoresis to determine the size of the plasmid. Based on the result, plasmid DNA has different sizes. Its case depends on the inserted DNA's condition in the plasmid. Colonies with a plasmid size of about 3000 base pairs will be selected, and re-selection follows. Colonies of about 1000 bp were taken during the selection steps.

Plasmid isolation purified plasmid DNA from host cells and separated it from mixed molecules such as protein, DNA, and RNA. The

yields of the isolation of the plasmid formed a band indicating the plasmid was successfully isolated from the host cell. More than one band represents the different conformational forms of the plasmid. Plasmid DNA usually has a supercoiled form. However, the supercoiled structure of plasmid DNA is susceptible to heat, friction, and freeze-thaw processes that cause DNA strand breaks. Thus, changing the supercoiled plasmid into another conformation [9].

From the results of plasmid isolation, a band closer to the well of gel electrophoresis was chosen, indicating that the plasmid had a larger size (Fig. 3).

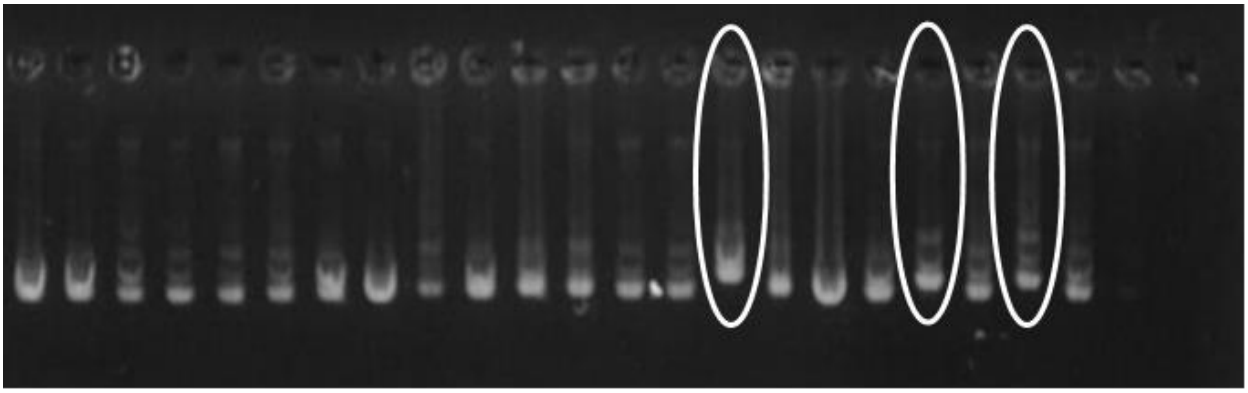


Figure 3. Electropherogram of transformants plasmid visualised on agarose gel. The white labels are indicates the plasmids selected in this research.

The white circle label indicates a sample with a larger plasmid size. After the plasmid isolation stage is complete, it can be confirmed using restriction enzymes or the PCR method to determine the presence of the inserted gene in the isolated plasmid. Restriction enzymes are used to cut the plasmid with DNA insertion, then visualized by agarose gel electrophoresis (Fig. 4).

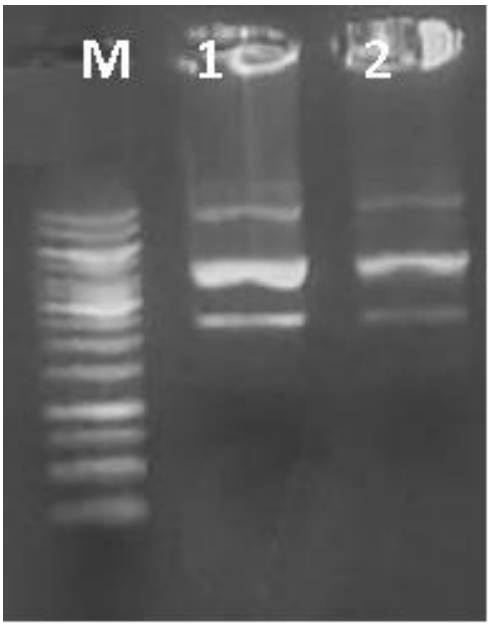


Figure 4. The electropherogram of plasmid cut using *Bam*HI.

The restriction enzyme used was BamHI. There is still more than one band visible in Fig. 4 because BamHI does not have a cleavage site region in the pJet 1.2/blunt vector. On the other hand, the insert does not have a cleavage site of BamHI. Thus, the plasmid is not cut off. Uncut plasmid DNA shows three possible plasmid conformations, namely supercoiled circular DNA, open circular, and linearized plasmid DNA. Three samples were successfully sequenced: samples 23 and 24 (TM primers) and sample number 45 (T1509) primers.

The bioinformatic of the samples

The yields of the sequencing are the nucleotide sequences of each sample. Based on the results, Blast at NCBI website can identify homology (similarity through common ancestry) as seen in Table 1.

The results of samples 23 and 24 resembled the SDR class Oxidoreductase (Caldivirga maquilingensis) of 58.58% and 58.02%, respectively. Sample number 45 has the highest protein homology with 3-dehydroquinate dehydratase Type I (Metallosphaera), with a similarity level of 79.88%.

Table 1. BLAST alignment results of samples

Sample number 23			
Protein Homology	**Species**	**NCBI Code**	**Similarity (%)**
SDR Family Oxidoreductase	*Caldivirga maquilingensis*	WP_012185583.1	58.58
SDR Family Oxidoreductase	*Caldivirga sp.*	HEV64393.1	53.29
Oxidoreductase NAD(P)-dependent	*Caldivirga sp. MU80*	WP_066792788.1	53.79
SDR Family Oxidoreductase	*Sulfolobus sp. A20*	WP_069283465.1	44.71
SDR Family Oxidoreductase	*Sulfolobus acidocaldarius*	WP_011277956.1	45.18
Dehydrogenase	*Sulfolobus acidocaldarius SUSAZ*	AHC51404.1	42.77
Sample number 24			
Protein Homology	**Species**	**NCBI Code**	**Similarity (%)**
SDR Family Oxidoreductase	*Caldivirga maquilingensis*	WP_012185583.1	58.02
Oxidoreductase NAD(P)-dependent	*Caldivirga sp. MU80*	WP_066792788.1	54.26
SDR Family Oxidoreductase	*Caldivirga sp.*	HEV64393.1	51.94
Oxidoreductase NAD(P)-dependent	*Caldivirga sp.*	NAZ29280.1	53.15
SDR Family Oxidoreductase	*Sulfolobus acidocaldarius*	WP_069283465.1	41.67
Sample number 45			
Protein Homology	**Species**	**NCBI Code**	**Similarity (%)**
MULTISPECIES Type I 3-dehydroquinate dehydratase	*Metallosphaera*	WP_01221797.1	79.88
Type I 3-dehydroquinate dehydratase	*Metallosphaera hakonensis*	WP_110369394.1	68.90
Type I 3-dehydroquinate dehydratase	*Metallosphaera cuprina*	WP_148230947.1	63.41
3-dehydroquinate dehydratase	*Metallosphaera cuprina*	AEB94521.1	63.41
Type I 3-dehydroquinate dehydratase	*Metallosphaera tengchongensis*	WP_174632462.1	57.93
3-dehydroquinate dehydratase	*Metallosphaera yellowstonensis*	EHP68820.1	53.37
Type I 3-dehydroquinate dehydratase	*Metallosphaera yellowstonensis*	WP_155812468.1	53.37
Type I 3-dehydroquinate dehydratase	*Acidianus manzaensis*	WP_148692138.1	49.38
Type I 3-dehydroquinate dehydratase	*Acidianus sulfidovorans*	WP_110380531.1	50.00
Type I 3-dehydroquinate dehydratase	*Acidilobales archaeon*	HHQ55762.1	49.08

Another bioinformatic analysis is to determine phylogenetics. Phylogenetics will classify organisms based on ancestral descent relationships. Phylogenetics is generally described as a tree or network where each tree branch represents one organism or gene [10, 11]. Sample number 23 has the closest relationship with number 24. Fig. 5 shows the two samples in order with the same phylogenetic tree construction because both

originate from the oxidoreductase group. The phylogenetic tree shows that sample 45 has the closest relationship with 3-dehydroquinate dehydratase Type I from Metallospaera species. Likewise, the results of the analysis of protein homology have a value of 79.88%.

The phylogenetic tree of all samples no. 23, 24, and 45 have formed separate branches from other amino acid sequences. It was suggested that the three amino acid sequences are new sequences different from the other ones. Thus, the samples are new proteins. Changes in the sequence of amino acids can be discovered by the amino acid sequence of the sample aligned with the amino acid sequences of the closest proteins.

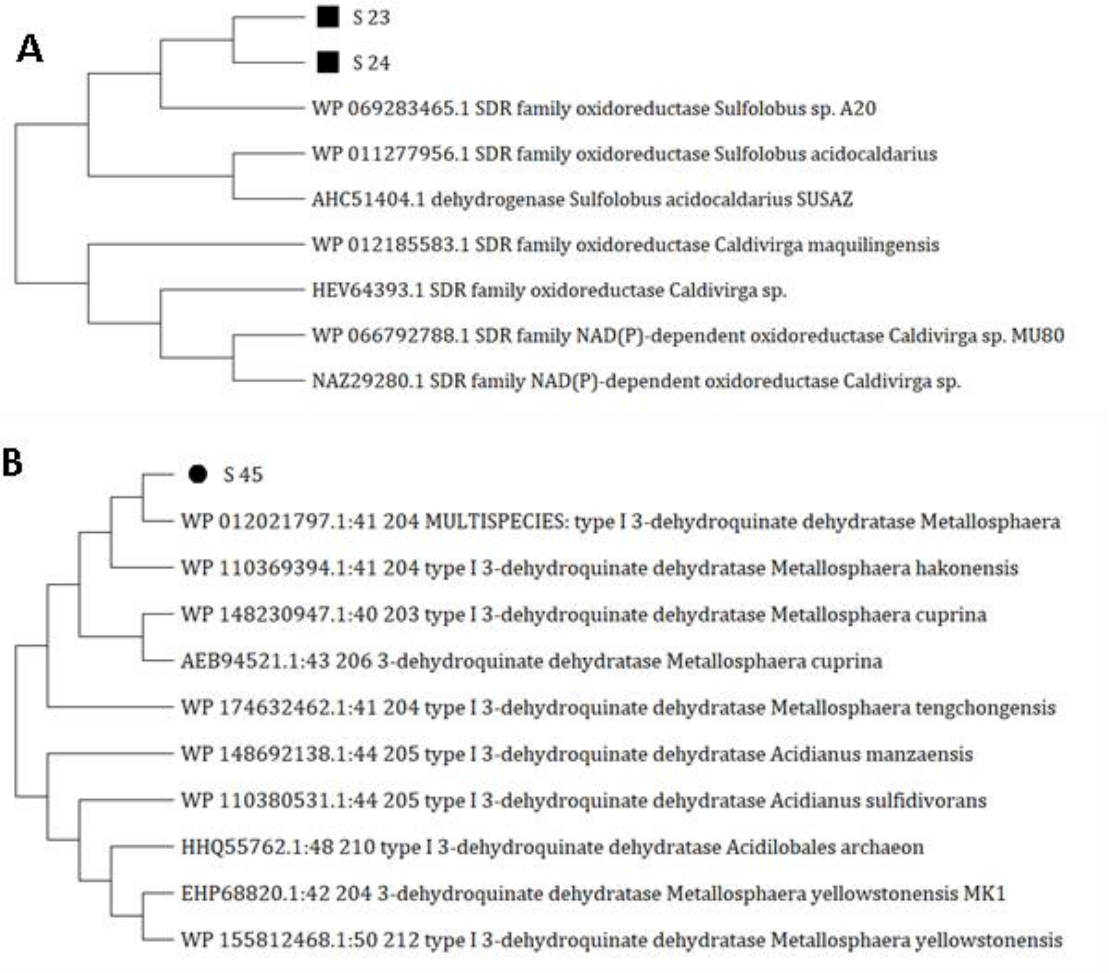


Figure 5. The phylogenetic tree of the samples. A is the phylogenetic tree for samples no 23 and 24, B is the phylogenetic tree for sample no 45.

The analysis of amino acid differences was performed using the Genedoc program. This program is used to visualize and analyze the sequence alignment from nucleic acid and protein in an evolutionary context [11]. Samples 23 and 24 were compared with the protein homology of oxidoreductase groups represented in Fig. 6 (A/B). There are some sequence differences in samples 23 and 24, marked with red circles. Sequences from sample 45 were aligned with protein homology (Fig. 6 A/B). Sample number 45 has the slightest sequence difference compared to samples 23 and 24. Therefore, sample 45 has the most significant homology similarity percentage, 79.88%.

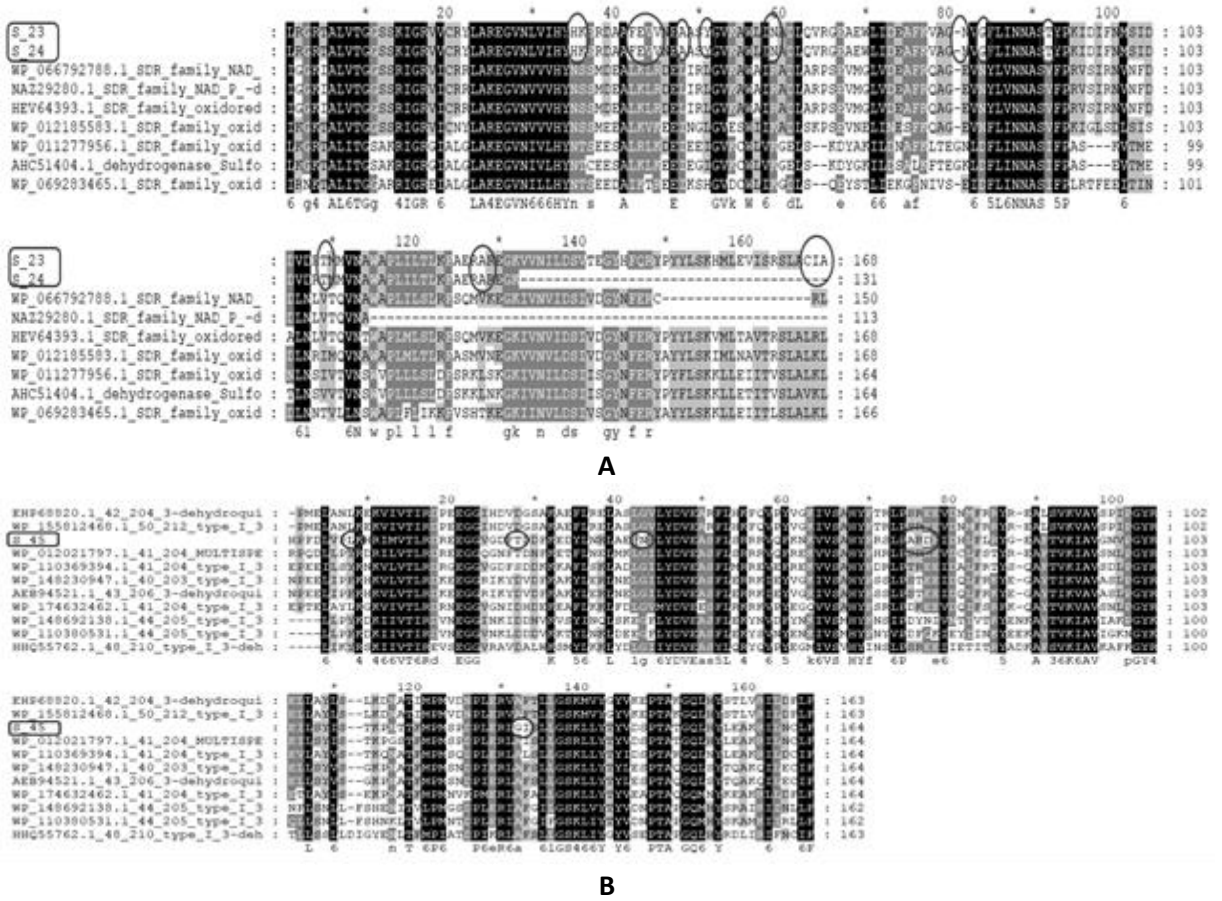


Figure 6. The alignment of amino acid sequences of the samples with the amino acid sequences of the closest protein using clustal x, visualized on genedoc. The circles shown the differences of amino acid of the samples to the others. A is the alignment of sample no 23 and 24, B is the alignment of sample no 45.

Conclusion

DNA was successfully isolated from Domas Crater by primers TM and T1509, which had similarities with the SDR oxidoreductase group. And sample 45 with 3-dehydroquinate dehydratase Type I. For phylogenetic tree analysis, sample 23 has the closest relationship with sample 24. However, sample 45 has the closest relationship with 3-dehydroquinate from *Metallospaera* species, and sample 45 has the slightest amino acid sequence difference from their homology protein compared to samples 23 and 24. A metagenomic approach has been successfully carried out, exploring new genes encoding thermostable proteins or enzymes from extreme environments. This research discovered two new types of proteins from the archaea group that have never been found.

Acknowledgment

References

[1] F. Rigoldi, S. Donini, A. Redaelli, E. Parisini, and A. Gautieri, “Review: Engineering of thermostable enzymes for industrial applications,” APL Bioeng., vol. 2, no. 1, pp. 1–17, 2018, doi: 10.1063/1.4997367.

[2] S. Sharma, S. Vaid, B. Bhat, S. Singh, and B. K. Bajaj, Thermostable enzymes for industrial biotechnology. Elsevier B.V., 2019.

[3] Suharti, N. W. Meray, S. Nurbaiti, and Akhmaloka, “Cloning and expression of a novel gene encoded thermostable archaeal aldolase class II from natural Sample,” Biosci. Biotechnol. Res. Asia, vol. 12, no. 2, pp. 1023–1029, 2015, doi: 10.13005/bbra/1754.

[4] Suharti, Rukman Hertadi, Fida Madayanti, Santi Nurbaity, and Akhmaloka, “Diveristy of gene encoded crenarchaeal DNA polymerase B from natural sample”, International Journal of Integrative Biology, vol. 15, no. 22, pp. 44-48, 2014.

[5] T. K. Johny and S. G. Bhat, “PCR in metagenomics,” Methods Mol. Biol., vol. 1620, pp. 249–265, 2017, doi: 10.1007/978-1-4939-7060-5_17. R. J. Suharti, Rukman Hertadi, Fida Madayanti, Santi Nurbaity, and Akhmaloka, “Diveristy of gene encoded crenarchaeal DNA polymerase B from natural sample”, International Journal of Integrative Biology, vol. 15, no. 22, pp. 44-48, 2014.

[6] T. A. Brown, Gene Cloning & DNA Analysis an Introduction, 6th ed. West Sussex, UK: John Wiley & Sons, Ltd., 2010.
[7] M. Carter and J. Shieh, "Molecular Cloning and Recombinant DNA Technology." Elsevier Inc., 2010, doi: 10.1016/B978-0-12-800511-8.00010-1.
[8] Bio-Rad, "Ligation and Transformation Module." https://www.bio-rad.com/webroot/web/pdf/lse/literature/10011370A.pdf (accessed Jan. 24, 2021).
[9] C. H. Lin, Y. C. Chen, and T. M. Pan, "Quantification bias caused by plasmid DNA conformation in quantitative real-time PCR assay," PLoS One, vol. 6, no. 12, 2011, doi: 10.1371/journal.pone.0029101.
[10] M. Charleston, "Phylogeny," Brenner's Encycl. Genet. Second Ed., vol. 3, pp. 324–325, 2013, doi: 10.1016/B978-0-12-374984-0.01160-8.
[11] N. Ziemert and P. R. Jensen, Phylogenetic approaches to natural product structure prediction, Methods Enzymol., 2012;517:161-82.doi: 10.1016/B978-0-12-404634-4.00008-5.
[12] T. J. Jankun-Kelly, A. D. Lindeman, and S. M. Bridges, "Exploratory visual analysis of conserved domains on multiple sequence alignments," BMC Bioinformatics, vol. 10, no. SUPPL. 11, pp. 3–11, 2009, doi: 10.1186/1471-2105-10-S11-S7.