The study of Heat shock protein 70 (Hsp70) gene expression on Ageratum conyzoides L. and Synedrella nodiflora L. in Universitas Indonesia, Depok and Cibodas Botanical Garden

C K Windarti, A E Maryanto, A Salamah and N Andayani
Department of Biology, Faculty of Mathematics and Natural Sciences (FMIPA), Universitas Indonesia, Depok 16424, Indonesia

Corresponding author’s email: ae.maryanto@sci.ui.ac.id

Abstract. Asteraceae is one of the invasive plant family. Invasive plants have more ability in tolerating high temperatures with the role of the Hsp70 mechanism. The Hsp70 gene is conserved in cells as a form of adaptation to some environmental stresses, one of which is temperature stress. The difference of average temperature between Universitas Indonesia-Depok (28.6 °C) and Cibodas Botanical Garden (20.06 °C) become the basis of this research to see the level of Hsp70 gene expression in Ageratum conyzoides and Synedrella nodiflora. RNA isolation from the leaf tissues was carried out using modified CTAB method followed by cDNA synthesis. The cDNA obtained is then amplified by PCR using a Hsp70 primer from Arabidopsis thaliana. The results of this research using in silico techniques showed a partial amplification of the Hsp70 gene and the presence of nucleotide variations in Universitas Indonesia-Depok and Cibodas Botanical Garden at the positions number 110 (Ageratum conyzoides) and 108 (Synedrella nodiflora). These variations resulted in different amino acids but did not change the protein structure prediction of the samples.

Keywords: Asteraceae, temperature, protein, invasive

1. Introduction

Asteraceae is the largest invasive plant family amongst other plants [1]. Invasive species are the dominating species or the new species within an ecosystem which are highly adaptive to compete with other species. Invasive species can disrupt the abundance and diversity of other surrounding species, leading to extinction in local species or endangered species [2]. Some Asteraceae species have adapted and become heat tolerant by having a protein Heat shock protein (Hsp) that make the plants survive in hot conditions [3]. Ageratum conyzoides and Synedrella nodiflora were chosen to be the subjects of this study because in Universitas Indonesia, they had the highest number of individuals with the highest Important Value Index (IVI) [4].

Heat shock protein (Hsp) gene is a molecular mechanism that plants do to deal with high temperature stress. The heat exposure in nature can trigger a stress response in plant cells, characterized by rapid formation of Hsp molecules. Hsp is a widespread protein in both plant and animal cells. Hsp can be induced by heat stress, cold temperature stress, UV exposure and other biotic stresses [5].

Heat shock protein 70 (Hsp70) is the most prominent Hsp molecule that is visible when there is an environment temperature increment or other stresses that makes damages for proteins. The Hsp70 gene
is a group of multifunctional proteins that act as molecular chaperones. Molecular chaperones are a large group of proteins that works in stabilizing proteins, ensuring protein folding, and refolding process on damaged proteins. The chaperone group belong to a conserved protein group in cells with high homology levels in its primary structure. Chaperones also regulate cellular processes such as protein trafficking, protein degradation, complex protein formation, and functional protein regulation. The expression of the Hsp70 gene positively correlates with the ability to tolerate heat stress. In the event of stress, Hsp70 provides protection to the cells by refolding and disaggregating substrate proteins [6-9].

2. Method

Isolation of RNA is the first step and this was performed using modified Chang’s CTAB (Cetyltrimethylammonium bromide) method [7]. The modification was done with the reduction of centrifugation speed from 30,000 g to 10,000 g related to the capacity of the centrifuge machine. The RNA is then transferred into a corning tube and stored in liquid nitrogen as or a freezer at -80 °C. DNA contaminant was then removed from RNA with RN1-free DNase RQ1 kit based on Promega kit protocol.

After the total RNA was obtained, visualization was performed by gel electrophoresis. The 1 % agarose was used and the machine was run on 100 V for 30 minutes. The results from gel electrophoresis are two RNA bands (28s RNA and 18s RNA). The visualization results from electrophoresis can be seen using Gel Doc [Bio-Rad Universal HOOd II]. The concentration and purity of RNA was then measured by using spectrophotometer [BioDrop].

The complementary DNA (cDNA) synthesis was performed using the GoScriptTM Reverse Transcription System [Promega] kit. RNA, primer, and RNAse-free water were mixed before heating to 70 °C for 5 minutes. We used the primer for amplification Hsp70 from Arabidopsis thaliana with forward sequence TCAAGCGGATAAGAGTCACT (CG258F) and reverse sequence CTCGTCCGGGTTAATGCT (CG259R) [6]. As a model organism, A. thaliana served in several of the most comprehensive studies for responses to biotic and abiotic stresses using differential relative and absolute quantitative strategies [10]. The PCR machine was run with regulating the temperature and cycle according to the protocol kit. The cDNA can be stored in a cooling machine with a temperature of -20 °C. Amplification with Polymerase Chain Reaction (PCR) was done using GoTaq® Green Master Mix kit [Promega]. The PCR result can be seen directly by using gel electrophoresis and stored in cooling machine with temperature -20 °C.

Sequencing was done with the help from a service [Macrogen]. The sequencing results can be seen in various type of files. The chromatograms from samples viewed from DNABaser, then the sequences aligned with the sequence from Arabidopsis thaliana to show the attach position from the samples and variations of nucleotides that could occur with Geneious ver 11.1.4 . The amino acid sequence analysis was performed by ORFfinder application on NCBI. Then, with Phyre2 website (http://www.sbg.bio.ic.ac.uk/phyre2/) the secondary structure composition, protein indication, and modeling of three-dimensional structure of protein were obtained.

3. Results and discussion

The measurement of some abiotic parameters during the sample collections can be seen from the table 1. Based on observation data (table 1), there is a 10.7 °C temperature difference between Universitas Indonesia and Cibodas Botanical Garden. A hot dan dry environment can somehow disrupt photosynthetic and transpiration rate also damage the enzymes and proteins [11]. The amplification of Hsp70 gene were seen in samples from Cibodas Botanical Garden with average temperature 23.9 °C meanwhile Barua and Heckathorn [12], stated that the expression of Hsp70 gene starting to induced at 33 °C. This condition could happen because the temperature at the location is very different in day and
night. The temperature can be very hot during the day and cooler at night. Cibodas Botanical Garden had minimum temperature of 17.04 °C and maximum temperature of 26.44 °C [13].

The PCR amplification results in figure 1. show one band in the same position on all samples, which is about 200–300 bp. The amplification bands resulted from the PCR process for both Ageratum conyzoides and Synedrella nodiflora samples are not in line with the expected bands from Sung (2001) [6] literature (650 bp) or the NCBI database (861 bp). This condition appeared because the primer used is not specific enough for the species we used. Another assumption is the complementary DNA (cDNA) from cDNA synthesis is only partially amplified. To find out whether the amplified band of the PCR is a targeted gene sequence or not, sequencing is necessary.

The result from sequencing is nucleotide bases sequence of the Hsp70 gene which is then aligned with Hsp70 from Arabidopsis thaliana using Geneious software version 11.1.4. in order to ensure that the amplified sequences are the amplification result of the Hsp70 gene. The result in figure 2 shows the presence of basic sequence similarity in all four samples at position 933–1,121 with Arabidopsis thaliana Hsp70 gene. The black bars in the sequence alignment of the four samples indicate that the Hsp70 gene is successfully amplified but produces a partial product [14].

Based on the results of sequence alignment with BioEdit software, there is a difference of one nucleotide base sequence from all four samples at the 1050th position (figure 3). The sequences from Ageratum conyzoides and Synedrella nodiflora samples obtained from Universitas Indonesia have a cytosine base (C) meanwhile thymine base (T) was found in samples from the location of Cibodas Botanical Gardens.

| No. | Location                         | Average temperature (°C) | Air humidity | Soil pH | Light intensity (100 x 200.000 Lux) |
|-----|---------------------------------|--------------------------|--------------|---------|-----------------------------------|
| 1.  | Universitas Indonesia          | 34.6                     | 59.417 %     | 6.63    | 337.13                            |
| 2.  | Cibodas Botanical Garden       | 23.9                     | 75 %         | 6.95    | 67.5                              |

**Figure 1.** The Electrophoresis Visualization of Hsp70 Amplification on A. conyzoides and S. nodiflora in Universitas Indonesia and Cibodas Botanical Garden. Electrophoresis gel is performed using 1 % agarose 100 V in 30 minutes with GelRed. (1) AC Cibodas Oligo(dT) Primer, (2) AC Cibodas Hsp70 Primer, (3) AC UI Oligo(dT) Primer, (4) AC UI Hsp70 Primer, (5) SN Cibodas Oligo(dT) Primer, (6) SN Cibodas Hsp70 Primer, (7) SN UI Oligo(dT) Primer, (8) SN UI Hsp70 Primer, (M) DNA Marker 1kb [Thermoscientific].
Figure 2. Prediction of amplified Hsp70 gene position on Ageratum conyzoides and Synedrella nodiflora with Geneious Software.

Figure 3. Sequences alignment of Hsp70 gene from Ageratum conyzoides and Synedrella nodiflora with Arabidopsis thaliana in Bioedit Software.

The different base in the four samples sequences can affect the code of amino acid. Ageratum conyzoides and Synedrella nodiflora samples in Universitas Indonesia encode Proline (P) in CCG nucleotide base sequences. Point mutations (C → T) in Ageratum conyzoides and Synedrella nodiflora samples located at the Cibodas Botanical Garden resulted in a sequence CUG which coding to Leusin (L) (figure 4).

The difference in nucleotide sequence of the four samples sequence allegedly happened because of point mutation [15]. Point mutations in nucleotide base sequences can lead to differences readings in amino acid, namely nonsense, silent mutation, and missense. Point mutation on a single nucleotide base which changes the triplet readings of amino acids to a stop codon is called a nonsense mutation. Silent mutation occurs when a point mutation on a single nucleotide base does not change the type of amino acid even though it changes the codon triplet. Missense mutation occurs when the change affects the type of amino acid that is formed [16]. The result from this research show the mutations that occur in the sample is a missense mutation that alters the readings of the amino acids that are formed.
Differences in the amino acids caused by point mutations did not affect the location of proteins in the cell. This is assumed to be caused by the fact that both the Leucine (L) and Proline (P) belong to the non-polar amino acid group so it does not affect the bond of the secondary structures [17]. The polarity of an amino acid plays a role in determining the location of a protein. Proteins with non-polar properties usually prefer to be in a non-aqueous environment. Leucine is a hydrophobic amino acid that tends to be in the protein hydrophobic cores in the α-helix form. Leucine plays a role in the introduction of substrates such as lipids. Proline is a unique amino acid that has a side chain connected to the backbone protein twice. Proline tends to be found on the protein surface and has a very non-reactive side chain so proline is rarely involved in the activity of protein or as a binding site [17].

Based on the analysis results with Phobius site (figure 5), it was found that all protein predictions performed both in Ageratum conyzoides and Synedrella nodiflora from the Universitas Indonesia and Cibodas Botanical Garden predicted to be in the cell but not at the cytoplasm (non-cytoplasmic). The predicted results are different from those described in Sung et al. that Hsp70 is a cytosolic protein. The prediction is thought that maybe the protein has the possibility of being at the cell organelle [6]. The difference may occur because the sample we used only encodes a partial sequence of Hsp70 coding gene. The composition of different amino acids can perform different protein properties and different predictions [18].

The predictions from Phyre2 site shows in figure 6 all samples have the same structure predictions. Both samples of Ageratum conyzoides and Synedrella nodiflora from Universitas Indonesia and Cibodas Botanical Garden are predicted as transport protein/chaperone based on Protein Database (PBD). Hsp70 itself is one of the molecular chaperones group that function in general in stabilizing the proteins, assisting in membrane translocation, and regulating in response to heat shock [19].

**Figure 4.** Amino acid sequence prediction on Ageratum conyzoides and Synedrella nodiflora samples with Proter Software (a) Prediction to A. conyzoides dan S. nodiflora sequences from Universitas Indonesia, (b) Prediction to A. conyzoides dan S. nodiflora sequences from Cibodas Botanical Garden.
Figure 5. Prediction of protein location from *Ageratum conyzoides* and *Synedrella nodifloram* samples with Phobius Software.

Figure 6. Three dimensional structure prediction of Hsp70 on *Ageratum conyzoides* and *Synedrella nodiflora* with Phyre2.
4. Conclusion
We found from 188 bp partial gene of Hsp70 from Ageratum conyzoides and Synedrella nodiflora there is a conserved sequenced that different between two species. Ageratum conyzoides had a single base point mutation at 110th base and Synedrella nodiflora at 108th base. These mutations express different amino acid as a plasticity adaptation to temperature stress, but still same secondary structure of Heat shock protein 70.

Acknowledgments
This work was financially supported by Kementrian Riset, Teknologi, dan Pendidikan Tinggi Republik Indonesia under PUPT grant with contract number 2710/UN2.R3.1/HKP05.00/2017.

References
[1] Sunaryo S, Uji T and Tihurua E F 2012 Berk. Panel. Hayati 17 147-52
[2] Solfiyeni S, Chairul C and Marpaung M 2016 Proc. Biol. Edu. Conf. 13 743-7
[3] Nash G and Waterman J 2011 Homesteading in the 21st Century: How One Family Created a More Sustainable, Self-sufficient, and Statifying Life (Newtown: Taunton Press)
[4] Agassi R 2017 Struktur Komunitas Tumbuhan Famili Asteraceae di Kampus Universitas Indonesia, Depok B.Sc Projects (Depok: FMIPA Universitas Indonesia)
[5] Park C-J and Seo Y-S 2015 Plant Pathol. J. 31 323-33
[6] Sung D Y, Vierling E and Guy C L 2001 Plant Physiol. 126 789-800
[7] Zeng X C, Bhasin S, Wu X, Lee J G, Maffi S, Nichols C J, Lee K J, Taylor J P, Greene L E and Einsberg E 2004 J. Cell. Sci. 117 4991-5000
[8] Manitsavevic S, Dunderski J, Matic G and Tucic B 2007 Plant Cell Environ. 30 1-11
[9] Baurle I 2016 F1000Res 18 1-5
[10] Wienkoop S, Baginsky S and Weckwerth W 2010 J. Proteomics 73 2239-48
[11] Hasanuzzaman M, Nahar K, Alam M M, Roychowdhury R and Fujita M 2013 Int. J. Mol. Sci. 14 9643-84
[12] Barua D and Heckathorn S A 2006 Am. J. Bot. 9 102-9
[13] Setiawati P 2012 Pengaruh Ruang Terbuka Hijau Terhadap Iklim Mikro BSc Projects (Bogor: Departemen Arsitektur Lanskap, Fakultas Pertanian, IPB University)
[14] Saibil H 2013 Nat. Rev. Mol. Cell. Biol. 14 630-42
[15] Mayer M P and Bukau B 2005 Cell. Mol. Life Sci. 62 670-84
[16] Rosalind 2019 Point Mutation available at http://rosalind.info/glossary/point-mutation/
[17] Betts M J and Russel R B 2007 Bioinformatics for Geneticists: A Bioinformatics Primer for the Analysis of Genetic Data, 2nd edition (Germany: John Wiley & Sons)
[18] Berg J M, Tymoczko J L and Stryer L 2002 Biochemistry 5th edition (New York: W H Freeman)
[19] Zerebecki R A and Sorte C J 2011 PLoS One 6 e14806