Bioinformatics Study of The Expressed Sequence Tags of Salt Tolerance Genes from Mangrove Plant *Rhizophora stylosa*

M Basyuni1*, S Baba2, R Wati1 and H Oku3

1 Department of Forestry, Faculty of Forestry, Universitas Sumatera Utara, Jl. Tri Dharma Ujung No. 1 Medan, North Sumatera 20155, Indonesia
2 International Society for Mangrove Ecosystems, Faculty of Agriculture, University of the Ryukyus, 1 Senbaru, Nishihara, Okinawa 903-0213, Japan
3 Molecular Biotechnology Group, Tropical Biosphere Research Center, University of the Ryukyus, 1 Senbaru, Nishihara, Okinawa 903-0213, Japan

*E-mail: m.basyuni@usu.ac.id

Abstract. The present study describes the bioinformatics analysis of 13 expressed sequence tags (ESTs) of salt tolerance genes in mangrove plant, *Rhizophora stylosa* as well as homology, phylogenetic, sequence analysis, potential peptide, and subcellular localization. The DNA sequence among the ESTs from *R. stylosa* exhibited 50-97% homology between themselves. The target peptide value of chloroplast and mitochondrial varied from 0.071-0.800, 0.053 to 0.254, respectively, indicated it was possible to exist. Sub-cellular of the fragment genes mostly was in the plasma membrane and endoplasmatic reticulum. On the other hand, a few genes restored in golgi bodies, vacuole, and lysosome. These results suggested the importance of understanding the function of properties of the probably salt tolerance genes in *R. stylosa* genomic library. To clarify the relationship among the ESTs in *R. stylosa*, a phylogenetic tree was constructed. The phylogenetic tree depicts that there are three branches, the first branch contained one EST, the second cluster consists of 9 genes, in which the majority ESTs resides and the last group comprised of 3 ESTs. The present study, therefore, suggested the diversity of salt tolerance genes form discrete clusters in the phylogenetic tree.

Keywords: chloroplast, plasma membrane, *R. stylosa*, salt tolerance gene, triterpenoid

1. Introduction

Mangrove plants are widespread in intertidal tropical and subtropical regions. The remarkable ability of mangrove species is to grow in varying degree of salinity ranging from fresh water to the levels found in seawater. Mangrove plants can be categorized into two groups according to their morphological characteristics in salt management [1]. The first group is the salt-secreting species that have either salt glands or salt hairs to remove excess salt. The second is non-secreting species that do not have such morphological characteristics for excretion of excess salt [1]. *Rhizophora stylosa* (Rhizophoraceae) is a common mangrove species belongs to the latter group, and it is distributed in the more coastal region compared to the habitat of other species of Rhizophoraceae [2]. This study demonstrated that *R. stylosa* is more tolerant to salinity than the other species. *R. stylosa*, therefore, is an ideal representation for studying the mechanisms of salinity tolerance in species level [2].

Despite the fact that the stress-tolerance mechanisms of the mangrove plants were well documented [3,4], however, the physiological roles remain obscure. Our previous study has shown that 157 salt
tolerance genes involved in a broad spectrum of biological pathways [5]. With remaining, 83 genes were as unclassified or unclear classification or no hits and were considered as new gene fragments [5]. In this context, it is noteworthy to analyze the salt tolerance genes from mangrove using bioinformatics method. However, the bioinformatics analysis of salt tolerance genes from mangrove plants is limited. Recently the bioinformatics study on fifteen salt tolerance genes from *R. stylosa* has been described [6]. The present study, therefore, extends our previous work and reports the bioinformatics analysis of 13 expressed sequence tags (ESTs) of probably salt tolerance genes in mangrove plant, *R. stylosa* as well as homology, phylogenetic, sequence analysis, potential peptide, and subcellular localization.

2. Materials and methods

2.1. Sample collection

A total of 13 expressed sequence tags (ESTs) of *Rhizophora stylosa* deposited officially in the DDBJ/EMBL/GenBank were investigated. The DDBJ/EMBL/GenBank accession numbers of the DNA sequence used in this analysis are as follows: FS997158 (Rs2), FS997160 (Rs4), FS997169 (Rs13), FS997242 (RS86), FS997243 (Rs187), FS997247 (Rs91), FS997248 (Rs102), FS997264 (Rs108), FS997271 (Rs115), FS997274 (Rs118), FS997275 (Rs119), and FS997177 (Rs121).

2.2. Homology and phylogenetic analysis of 13 ESTs

The DNA sequences of EST were aligned, and similarity scores were obtained using the FASTA version 3.426 [7] of the DNA Data Bank of Japan (Mishima, Shizuoka, Japan). The best score of results is shown in Table 1. Phylogenetic analysis of 13 ESTs DNA sequences was conducted with CLUSTAL W version 1.83 [8] of the DNA Data Bank of Japan followed by drawing with TreeView, ver. 1.6.6 [9] based on a neighbor-joining method. Bootstrap analysis with 1000 replications was used to assess the strength of the nodes in the tree [10]. The DDBJ/GenBank/EMBL accession numbers of the DNA sequence of using this analysis is described in the sample subsection.

2.3. DNA Sequence analysis

The functional assignment of DNA sequences of 13 ESTs was based on a similarity search of the sequences against the Genbank non-redundant (nr) peptide database of NCBI (https://blast.ncbi.nlm.nih.gov/Blast.cgi) using BLASTX [11].

2.4. Possibility of the potential peptide and subcellular localization in 13 ESTs

The targetP 1.1 Server online (www.cbs.dtu.dk/services/targetp/) was used for transit peptide prediction as previously described [6]. The site assignment is based on the predicted existence of any of the N-terminal pre-sequences chloroplast transit peptide (cTP), mitochondrial targeting peptide (mTP) and secretory pathway signal peptide (SP). PSORT Prediction online (psort.hgc.jp/form.html) was used to analyze the subcellular localization of 13 ESTs from *R. stylosa*.

3. Results

The results will be discussed in three subsections; they are homology and phylogenetic analysis of ESTs, DNA sequence analysis, and the possibility of the potential peptide and subcellular localization in *R. stylosa*.

3.1. Homology and phylogenetic analysis of ESTs from *R. stylosa*

The DNA sequence between 13 ESTs from *R. stylosa* shared 50-97% among themselves (Table 1). The DNA sequences of clone Rs13 (no hits) was showed significant similarity (97 %) to clone Rs87 (no hits). Clone Rs13 (no its) is close to clone Rs115 (Probable lactoylglutathione lyase, chloroplastic from Nelumbo nucifera) and showed relatively high similarity (97%). On the other hand, clone Rs87 (no hits) also exhibited significantly with Rs115. Among the probably salt tolerant genes analyzed,
clone Rs2 (no hits) and clone Rs86 (no hits) showed the lowest similarity (50%) of DNA sequence (table 1).

To shed light on the relationship among the candidate of salt-tolerant genes in *R. stylosa*, a phylogenetic tree was created. The phylogenetic tree shows that there are three clusters in the tree. The first branch consists of Rs121 clone only, hypothetical protein CISIN_1g037109mg, partial from *Citrus sinensis* (table 2), the second branch includes the typical genes and last group contained three clones namely Rs86 (no hits), Rs91 (hypothetical protein JCGZ_06052 from *Jatropha curcas*), and Rs118 (no hits).

![Phylogenetic tree of 13 ESTs from *R. stylosa*. Phylogenetic of DNA sequences were constructed with the neighbor-joining method of the CLUSTAL W [8]. The indicated scale corresponds to 0.1 DNA sequence substitutions per site. Numbers indicate bootstrap value from 1000 replicates. The accession numbers of the DNA sequence from DDBJ/GenBank/EMBL of using this analysis is described in the Materials Section.](image)

**Figure 1.** Phylogenetic tree of 13 ESTs from *R. stylosa*. Phylogenetic of DNA sequences were constructed with the neighbor-joining method of the CLUSTAL W [8]. The indicated scale corresponds to 0.1 DNA sequence substitutions per site. Numbers indicate bootstrap value from 1000 replicates. The accession numbers of the DNA sequence from DDBJ/GenBank/EMBL of using this analysis is described in the Materials Section.
Table 1. DNA sequence similarity between ESTs in *R. stylosa*

| Clone ID | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 9   | 10  | 11  | 12  | 13  |
|----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Rs2      | 100 |     |     |     |     |     |     |     |     |     |     |     |     |
| Rs4      | 72  | 100 |     |     |     |     |     |     |     |     |     |     |     |
| Rs13     | 94  | 76  | 100 |     |     |     |     |     |     |     |     |     |     |
| Rs86     | 50  | 67  | 52  | 100 |     |     |     |     |     |     |     |     |     |
| Rs87     | 87  | 72  | 97  | 52  | 100 |     |     |     |     |     |     |     |     |
| Rs91     | 61  | 69  | 62  | 94  | 61  | 100 |     |     |     |     |     |     |     |
| Rs92     | 73  | 55  | 76  | 76  | 72  | 66  | 100 |     |     |     |     |     |     |
| Rs102    | 72  | 52  | 77  | 70  | 81  | 66  | 51  | 100 |     |     |     |     |     |
| Rs108    | 75  | 55  | 82  | 73  | 81  | 68  | 51  | 57  | 100 |     |     |     |     |
| Rs115    | 86  | 81  | 97  | 76  | 97  | 79  | 52  | 48  | 48  | 100 |     |     |     |
| Rs118    | 83  | 64  | 83  | 81  | 78  | 80  | 61  | 60  | 60  | 60  | 100 |     |     |
| Rs119    | 87  | 59  | 85  | 79  | 84  | 80  | 58  | 56  | 56  | 55  | 62  | 100 |     |
| Rs121    | 72  | 54  | 73  | 77  | 68  | 69  | 98  | 52  | 52  | 53  | 63  | 62  | 100 |

3.2. DNA analysis

Table 2 shows the distribution of the putative function of ESTs from *R. stylosa* analyzed by Blastx. About 6 ESTs (46%) showed significant homologies to the sequence in the GenBank *nr* database, with the lasting 54% classified as no hits.

Table 2. Distribution of putative function of ESTs from *R. stylosa* generated by Blastx

| Clone ID | Length (bp) | Functional distribution of the top Blastx hits                                                                 |
|----------|-------------|-------------------------------------------------------------------------------------------------------------|
| Rs2      | 72          | No hits                                                                                                   |
| Rs4      | 279         | Probable tRNA N6-adenosine threonyl carbamoyltransferase, mitochondrial isoform X2 from *Nicotiana tabacum* (ID: XP_016462295.1) |
| Rs13     | 67          | No hits                                                                                                   |
| Rs86     | 96          | No hits                                                                                                   |
| Rs87     | 70          | No hits                                                                                                   |
| Rs91     | 115         | Hypothetical protein JCGZ_06052 from *Jatropha curcas* (ID: KDP36996.1)                                     |
| Rs92     | 290         | Probable tRNA N6-adenosine threonyl carbamoyltransferase, mitochondrial isoform X2 from *N. tabacum* (ID: XP_016462295.1) |
| Rs102    | 287         | No hits                                                                                                   |
| Rs108    | 321         | Pentatricopeptide repeat-containing protein At2g01390 from *Ananas comosus* (ID: XP_020103869.1)            |
| Rs115    | 324         | Probable lactoylglutathione lyase, chloroplastic from *Nelumbo nucifera* (ID: XP_010254032.1)              |
| Rs118    | 437         | No hits                                                                                                   |
| Rs119    | 364         | No hits                                                                                                   |
| Rs121    | 279         | Hypothetical protein CISIN_1g037109mg, partial from *Citrus sinensis* (ID: KDO45458.1)                     |
3.3. The potential peptide and subcellular localization of salt tolerance gene
Table 3 shows the possibility of the potential transit peptide in 13 EST clones from R. stylosa. There are three options: chloroplast transit peptide, mitochondrial target peptide and signal peptide of secretory pathway along with the prediction probability. The target chloroplast varied from 0.051 to 0.738, with the highest values of chloroplast belongs to clone Rs108 (0.738), indicated that chloroplast transit peptide present in the candidate genes of salt tolerance. It is noteworthy that target peptide value of mitochondria was less compared with chloroplast transit or mitochondrial peptide. The highest signal peptide of the secretory pathway was Rs92 (tRNA N6-adenosine threonyl carbamoyltransferase). Reliability prediction value of 5 (77%) dominated in the among the EST genes.

Table 4 shows subcellular localization of EST fragments in R. stylosa. The subcellular localization of these genes was mostly stored in the endoplasmic reticulum, plasma membrane, and outside. On the other hand, a few genes restored in Golgi bodies, vacuole, and lysosome.

Table 3. Possibility of the potential transit peptide in 13 ESTs

| Clone ID | Chloroplast transit peptide | Mitochondrial target peptide | Signal peptide of secretory pathway | Reliability prediction |
|----------|-----------------------------|-------------------------------|-------------------------------------|-----------------------|
| Rs2      | 0.221                       | 0.254                         | 0.105                               | 5                     |
| Rs4      | 0.071                       | 0.098                         | 0.091                               | 3                     |
| Rs13     | 0.155                       | 0.232                         | 0.263                               | 5                     |
| Rs86     | 0.413                       | 0.067                         | 0.034                               | 5                     |
| Rs87     | 0.263                       | 0.158                         | 0.236                               | 5                     |
| Rs91     | 0.259                       | 0.080                         | 0.114                               | 5                     |
| Rs92     | 0.051                       | 0.063                         | 0.354                               | 5                     |
| Rs102    | 0.140                       | 0.083                         | 0.114                               | 5                     |
| Rs108    | 0.738                       | 0.158                         | 0.039                               | 3                     |
| Rs115    | 0.166                       | 0.179                         | 0.060                               | 5                     |
| Rs118    | 0.080                       | 0.037                         | 0.722                               | 3                     |
| Rs119    | 0.214                       | 0.183                         | 0.124                               | 5                     |
| Rs121    | 0.052                       | 0.064                         | 0.365                               | 5                     |

4. Discussion
The enzymatic reaction products of a nominee of salt tolerance genes diverged from those of their neighbor clones in the tree. The data indicates that the relationships only in the phylogenetic tree have limited importance in predicting the product profile of the ESTs fragment genes [12]. The confirmation of functional expression of ESTs clones should be performed in bacteria, yeast, or other organisms. Furthermore, it might be the presence of an additional protein domain to control the product profile of genes [12].

Table 2 shows the distribution of the putative function of ESTs from R. stylosa analyzed by Blastx. About 6 ESTs (46%) showed significant homologies to the sequence in the GenBank nr database, with the lasting 54% classified as no hits. These unknown identified gene fragments might play a vital function for mangrove plants in the changing to abiotic stress tolerance [6]. Table 3 shows the possibility of the potential transit peptide in 13 EST clones from R. stylosa. Based on the results, the present study supported the previous finding on the similar pattern of salt tolerance genes in R. stylosa [6].

Table 4 depicts subcellular localization of EST fragments in R. stylosa. Recently, it has been reported that the expression of triterpenoid synthase genes enhanced the triterpenoid content of whole
cell body and plasma membrane fractions [13]. Table 4 shows that Rs92 and Rs118, which have the highest value were placed on the plasma membrane, supported previous results on their subcellular localization of triterpene genes located in the plasma membrane [13]. The reliability and functionality of the plasma membrane are therefore a key factor for salt tolerance mechanism in plants [14].

| Clone ID | Endoplasmic reticulum | Plasma membrane | Vacuole | Outside | Lysosome | Golgi body |
|----------|------------------------|-----------------|---------|---------|----------|------------|
| Rs2      | 0.100                  | 0.190           | nd      | 0.370   | nd       | nd         |
| Rs4      | 0.550                  | 0.190           | nd      | 0.100   | nd       | nd         |
| Rs13     | 0.100                  | 0.190           | 0.900   | 0.370   | nd       | nd         |
| Rs 86    | 0.550                  | 0.190           | nd      | 0.100   | nd       | nd         |
| Rs87     | 0.100                  | nd              | 0.900   | 0.370   | nd       | nd         |
| Rs91     | 0.550                  | nd              | nd      | 0.100   | nd       | 0.100      |
| Rs92     | 0.100                  | 0.640           | nd      | 0.380   | nd       | nd         |
| Rs102    | 0.100                  | 0.190           | nd      | 0.523   | nd       | nd         |
| Rs108    | 0.550                  | 0.190           | nd      | 0.100   | nd       | nd         |
| Rs115    | 0.550                  | nd              | nd      | 0.100   | nd       | 0.100      |
| Rs118    | 0.100                  | 0.514           | nd      | 0.100   | nd       | nd         |
| Rs119    | 0.550                  | nd              | nd      | 0.100   | 0.190    | nd         |
| Rs121    | 0.100                  | nd              | nd      | 0.380   | 0.100    | nd         |

nd= not detected

5. Conclusion
There are two clones: Rs92 and Rs118, which have the highest value on the plasma membrane, supported previous results on their subcellular localization of triterpene genes located in the plasma membrane. The present study suggested the importance of understanding the function of properties of the ESTs that probably salt tolerance genes in R. stylosa genomic library.

Acknowledgments
This study was supported by an International Research Collaboration Grant (No. 003/SP2H/LT/DRPM/IV/2017 to MB) from the Directorate for Research and Community Service, Ministry of Research, Technology, and Higher Education, Republic of Indonesia.

References
[1] Tomlinson P B 1986. The Botany of Mangroves (London: Cambridge University Press) 419 pp.
[2] Basyuni M, Baba S, Kinjo Y, Oku H 2012 Aquat. Bot. 97 17–23
[3] Flowers T J, Colmer T D 2008 New Phytol. 179 945–963
[4] Parida A K, Jha B 2010 Trees 24 199–217
[5] Basyuni M, Kinjo Y, Baba S, Shinzato N, Iwasaki H, Siregar E B M, Oku H 2011 Plant Mol. Biol. Rep. 29 533–543.
[6] Basyuni M, Sumardi 2017 J. Phys. : Conf. Ser. 801 012012.
[7] Pearson W R, Lipman DJ 1988 Proc. Natl. Acad. Sci. USA 85 2444–2448.
[8] Thompson J D, Higgins D G, Gibson T J 1994 Nucleic Acid Res. 22 4673–4680.
[9] Page R D 1996 Comput. Appl. Biosci. 12 357–358.
[10] Felsenstein J 1985 *Evolution* **39** 783–791.
[11] Altschul S F, Madden T L, Zhang J H, Zhang Z, Miller W, Lipman DJ 1997 *Nucleic Acid Res.* **25** 3389–3402.
[12] Basyuni M, Oku H, Tsujimoto E, Kinjo K, Baba S, Takara K 2007 *FEBS J.* **274** 5028–5042.
[13] Inafuku M, Basyuni M, Oku H 2018 *Saudi J. Biol. Sci.* **25** 1–9.
[14] Mansour M M F, Salama K H A 2004 *Environ. Exp. Bot.* **52** 113–122.