Bioinformatics approach of three partial polyprenol reductase genes in *Kandelia obovata*

M Basyuni*, R Wati¹, H Sagami², H Oku³, S Baba⁴

¹Department of Forestry, Faculty of Forestry, Universitas Sumatera Utara, Jl. Tri Dharma Ujung No. 1 Medan, North Sumatera 20155, Indonesia
²Institute of Multidisciplinary Research for Advanced Material, Tohoku University, Sendai, 980-8577, Japan
³Molecular Biotechnology Group, Tropical Biosphere Research Center, University of the Ryukyus, 1 Senbaru, Nishihara, Okinawa 903-0213, Japan
⁴International Society for Mangrove Ecosystems, Faculty of Agriculture, University of the Ryukyus, 1 Senbaru, Nishihara, Okinawa 903-0213, Japan

*Email: m.basyuni@usu.ac.id

Abstract. This present study describes the bioinformatics approach to analyze three partial polyprenol reductase genes from mangrove plant, *Kandelia obovata* as well as predicted physical and chemical properties, potential peptide, subcellular localization, and phylogenetic. The diversity was noted in the physical and chemical properties of three partial polyprenol reductase genes. The values of chloroplast were relatively high, showed that chloroplast transit peptide occurred in mangrove polyprenol reductase. The target peptide value of mitochondria varied from 0.088 to 0.198 indicated it was possible to be present. These results suggested the importance of understanding the diversity of physicochemical properties of the different amino acids in polyprenol reductase. The subcellular localization of two partial genes located in the plasma membrane. To confirm the homology among the polyprenol reductase in the database, a dendrogram was drawn. The phylogenetic tree depicts that there are three clusters, the partial genes of *K. obovata* joined the largest one: C23157 was close to *Ricinus communis* polyprenol reductase. Whereas, C23901 and C24171 were grouped with *Ipomoea nil* polyprenol reductase, suggested that these polyprenol reductase genes form distinct separation into tropical habitat plants.

1. Introduction

Mangrove plants are well known a source of secondary metabolites mostly derived from isoprenoids [1]. Polysisoprenoid alcohols are linear five unit polymers that are present in almost all living organisms. Long chain polyisoprenoids have occurred in various plant tissues [2]. These studies revealed the presence two types of polyisoprenoids regarding the stereochemistry: polyprenol (α-unsaturated isoprenoid alcohols) and dolichol (α-saturated isoprenoid alcohols). In plant photosynthetic tissues, polyrenols are usually abundantly detected in comparison to dolichols [2-4]. On the other hand, the sources for dolichols mainly are derived from animals (livers) [5], plant roots [6-8], and yeast cells [9].

Recently, dolichols but not polyrenols have been reported as the dominant polyisoprenoids alcohols of mangrove and coastal plants [6-8]. The occurrence of prospective dolichols in the leaves of mangroves and coastal plants indicate that the enzyme of polyrenol reductase might be active to
converse the polyprenols to dolichols [6-8]. In this context, it is essential to get more insight into the polyprenol reductase genes in mangrove plants. Nonetheless, the information on the polyprenol reductase has not been previously available in mangroves. The present study, therefore, aimed to analyze three predicted polyprenols genes in mangrove plant, \textit{Kandelia obovata} using the bioinformatics approach.

2. Materials and method

2.1. Materials
A total three partial polyprenol reductase genes from \textit{Kandelia obovata} namely c23157, c23901, and c24171 were studied. These genes mainly were derived from genome sequence of \textit{K. obvata} using a Blast search. These genes showed high homology to the plant polyprenol reductase genes.

2.2. Physicochemical properties of the polyprenol reductase gene
Protparam online (web.expasy.org/protparam/) was used to analyze the composition, physical and chemical properties of three polyprenol reductase genes. The computed parameters describe the molecular weight, theoretical isoelectric point values, amino acid composition, atomic composition, extinction coefficient, estimated half-life, instability index, fat coefficient, and average hydrophilicity as previously reported [10].

2.3. Potential transit of peptide and subcellular localization of OSC gene
The target P1.1 server online (www.cbs.dtu.dk/services/targetp/) was used to predict transit peptide. The location is based on the predicted occurrence of any of the \textit{N}-terminal pre-sequences chloroplast transit peptide (cTP), mitochondrial targeting peptide (mTP) and secretory pathway signal peptide (SP). Furthermore, PSORT Prediction online (psort.hgc.jp/form.html) was used to determine the subcellular localization of polyprenol reductase genes as previously described [11].

2.4. Phylogenetic analysis of three polyprenol reductase
The amino acid sequences were aligned using the FASTA version 3.4.26 [12] of the DNA Data Bank of Japan (Mishima, Shizuoka, Japan). Phylogenetic analysis of deduced amino acid alignment from three partial genes in addition to twenty one predicted polyprenol reductase was analyzed using CLUSTAL W ver. 1.83 [13] of the DNA Data Bank of Japan followed by drawing with TreeView ver. 1.6.6 [14] based on a neighbour-joining method. Bootstrap analysis with 1000 replicates was used to measure the strength of the nodes in the tree [15]. The GenBank accession numbers of the DNA and amino acid sequence of plant polyprenol reductase genes used this analysis as follows, \textit{Arachis duranensis} polyprenol reductase 2 (XM_016081788), \textit{A. ipaensis} polyprenol reductase 2 (XM_021109944), \textit{Cajanus cajan} polyprenol reductase 2 (XM_020359734), \textit{Elaeis guineensis} polyprenol reductase 1 (XM_019849699), \textit{Eucalyptus grandis} polyprenol reductase 2 (XM_018861897), \textit{Glycine max} polyprenol reductase 2 (XM_003537780), \textit{Gossypium arboreum} polyprenol reductase 2-like (XM_017765134), \textit{G. hirsutum} polyprenol reductase 2-like (XM_016848239), \textit{G. raimondii} polyprenol reductase 2-like (XM_012601789), \textit{Ipomoea nil} polyprenol reductase 2-like (XM_019302562), \textit{Jatropha curcas} polyprenol reductase 2 (XM_012210100), \textit{Juglans regia} polyprenol reductase 2-like (XM_018985852), \textit{Malus x domestica} polyprenol reductase 2-like (XM_017330714), \textit{Phoenix dactylifera} polyprenol reductase 1 (XM_017840866), \textit{Populus euphratica} polyprenol reductase 2-like (XM_011038928), \textit{Prunus persica} polyprenol reductase 2 (XM_007200320), \textit{Pyrus x breitneri} polyprenol reductase 2-like (XM_018642260), \textit{Ricinus communis} polyprenol reductase 2 (XM_015715302), \textit{Theobroma cacao} polyprenol reductase 2 (XM_018128982), \textit{Vigna radiata} var. \textit{radiata} polyprenol reductase 2 (XM_014668834), and \textit{Ziziphus jujuba} polyprenolreductase 2 (XM_016036804).
3. Results and Discussions

3.1. Physical and chemical properties of the polypropenol reductase gene
Table 1 shows the start codon of the partial genes was not detected, where the stop codons were found only in C2390 (917TGA) and C24171 (504TGA). The open reading frame length was not identified due to the partial genes. Some encoded amino acids were 166 to 306. The more range of genes had the more values of the physicochemical parameters such as in molecular mass, the total number of atoms, extinction coefficient, Fat coefficient, and overall average hydrophilicity (Table 1). No stable proteins were found in predicted polypropenol reductase genes. The stability protein coefficient was generally below 40 as previously reported in BgBAS and RsCAS [10] of oxidosqualene cyclase genes and KcAct1, BgAct1, and RsAct1 as members of plant actin genes [11]. The diversity was noted in the physicochemical properties of three partial polypropenol reductase genes.

Table 1. Physical and chemical properties of the polypropenol reductase in K. obovata

| Nucleotide name | C23157 | C23901 | C24171 |
|-----------------|--------|--------|--------|
| Length of genes/bp | 700    | 960    | 531    |
| Open reading frame length/bp | nd     | nd     | nd     |
| Start site and codon | nd     | nd     | nd     |
| Stop site and codon | nd     | 917TGA | 504TGA |
| Number of encoded amino acids | 228    | 306    | 166    |
| Relative molecular mass | 59214.59 | 79063.22 | 44281.16 |
| Theoretical isoelectric point values | 5.15    | 4.98   | 5.17   |
| Total number of atoms | 7618    | 10023  | 5672   |
| Extinction coefficient | 0.148   | 0.183  | 0.175  |
| Half-life period | 1.2h    | 7.2h   | 30h    |
| Instability coefficient | 48.67   | 47.77  | 56.94  |
| Fat coefficient | 22.46   | 28.81  | 34.40  |
| Overall average hydrophilicity | 0.576   | 0.870  | 0.962  |

nd= not detected

3.2. Potential transit of peptide and subcellular localization of polypropenol reductase gene
Table 2 depicts the possibility of the possible transit peptide in K. obovata polypropenol reductase. There are four reliabilities concerning the possibility of the transit peptide namely chloroplast transit peptide, mitochondrial target peptide, signal peptide of secretory pathway, and reliability prediction (Table 2). The values of chloroplast were relatively high, showed the occurrence of chloroplast transit peptide. The contents of propenols and dolichols were enhanced during the life-course of a tissue or organ [6, 16].The target peptide value of mitochondria varied from 0.088 to 0.198 indicated it was possible to be present. The low signal peptide of the secretory pathway has been described in mangrove OSC genes and actin genes [10-11], indicating no signal peptide in the polypropenol reductase of K. obovata.

Table 2. Possibility of the potential transit peptide in K. obovata polypropenol reductase

| Nucleotide ID | Chloroplast transit peptide | Mitochondrial target peptide | Signal peptide of secretory pathway | Reliability prediction |
|---------------|-----------------------------|-------------------------------|-------------------------------------|-----------------------|
| C23157        | 0.236                       | 0.182                         | 0.031                               | 4                     |
| C23901        | 0.204                       | 0.198                         | 0.009                               | 4                     |
| C24171        | 0.214                       | 0.088                         | 0.076                               | 5                     |
Figure 1. Phylogenetic tree of plant polyprenol reductase including three partial polyprenols reductase from *K. obovata*. Phylogenetic tree of deduced amino acid sequences was constructed with the neighbour-joining method of the CLUSTAL W [13]. The indicated scale represents 0.1 amino acid substitutions per site. Numbers indicate bootstrap value from 1000 replicates. The GenBank accession numbers of the amino acid sequence of using this analysis are shown in the Materials subsection.

Table 3 displays subcellular localization of polyprenol reductase genes in *K. obovata*. The subcellular localization of these genes was mostly in the endoplasmic reticulum (membrane and lumen). The C23901 has also located in Golgi bodies and plasma membrane, the C23157 was found in outside and lysosome. The C24171 was detected in the plasma membrane and outside. Recently, it has been shown that the gene expression of two triterpenoid synthases, *BgbAS* and *RsM1* enhanced the triterpenoid content of plasma membrane fractions as well as found in the plasma membrane [9, 16]. Furthermore, it has been established views that the plasma membrane is the first defence against changes in physicochemical variables of the altered environment [18-19]. Several salt tolerance genes
from *Rhizophora stylosa* located in the plasma membrane supported the previous study that on the importance of layer for abiotic stress tolerance including salt stress [20].

| Nucleotide ID | Golgi bodies | Plasma membrane | Endoplasmic reticulum (membrane) | Endoplasmic reticulum (lumen) | Outside | Lysosome |
|---------------|--------------|-----------------|---------------------------------|-------------------------------|---------|---------|
| C23901        | 0.460        | 0.640           | 0.370                           | 0.100                         | nd      | nd      |
| C23157        | nd           | nd              | 0.550                           | 0.100                         | 0.100   | 0.317   |
| C24171        | nd           | 0.717           | 0.100                           | 0.100                         | 0.437   | nd      |

3.3. Phylogenetic analysis of polyprenol reductase gene

To confirm the homology among the polyprenol reductase gene in *K. obovata* with plant polyprenol reductases, a phylogenetic tree was constructed by their amino acid sequences (Figure 1). The phylogenetic tree forms three clusters; the first branch consists of three genes: *T. cacao*, *G. arboretum*, and *G. raimondii*. The second group comprises the largest members where the partial polyprenol genes joined in the tree. The C23157 was close to *Ricinus communis* polyprenol reductase. Whereas, C23901 and C24171 were grouped with *Ipomoea nil* polyprenol reductase, suggested that these polyprenol reductase genes form distinct separation into tropical habitat plants.

4. Conclusions

The present study confirmed the chloroplast transit peptide occurred in *K. obovata* polyprenol reductase. The cluster analysis indicated that polyprenol reductase from *K. obovata*, *R. communis*, and *I. nil* grouped distinct separation into tropical habitat plants.

Acknowledgment

This study was partly supported by an International Research Collaboration and Scientific Publication 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] Basyuni M, Oku H, Baba S, Takara K and Iwasaki H 2007 Isoprenoids of Okinawan mangroves as lipid input into estuarine ecosystem *J. Oceanography* 63 601–608.

[2] Swiezewska E and Danikiewicz W 2005 Polyisoprenoids: Structure, biosynthesis, and function. *Prog. Lipid Res.* 44 235–258.

[3] Swiezewska E, Sasak W, Mankowski T, Jankowski W, Vogtman T, Krajewska I, Hertel J, Skoczylas E and Chojnacki T 1994 The search for plant polyprenols *Acta Biochim. Pol.* 41 221–260.

[4] Arifiyanto D, Basyuni M, Sumardi, Putri LAP, Siregar ES, Risnashari I and Syahputra I 2017 Occurrence and cluster analysis of palm oil (*Elaeis guineensis*) fruit type using two-dimensional thin layer chromatography *Biodiversitas* 18 1487–1492.

[5] Ishiguro T, Morita-Fujimira Y, Shidoji Y and Sagami H 2014 Dolichol biosynthesis: The occurrence of epoxy dolichol in skipjack tuna liver *Biochem. Biophys. Res. Comm.* 451 277–281.

[6] Basyuni M, Sagami H, Baba S, Iwasaki H and Oku H 2016 Diversity of polyisoprenoids in ten Okinawan mangroves *Dendrobiology* 75 167–175.

[7] Basyuni M, Sagami H, Baba S and Oku H 2017 Distribution, occurrence, and cluster analysis of new polyprenyl acetones and other polyisoprenoids from North Sumatran mangroves *Dendrobiology* 78 18–31.
[8] Basyuni M, Wati R, Sagami H, Sumardi, Baba S and Oku H 2018 Diversity and abundance of polyisoprenoid composition in coastal plant species from North Sumatra, Indonesia Biodiversitas 19 1–11.

[9] Grabinska K and Palamarczyk G 2002 Dolichol biosynthesis in the yeast Saccharomyces cerevisiae: an insight into the regulatory role of farnesyl diphosphate synthase FEBS Lett. 2 259–265.

[10] Basyuni M and R Wati 2017 Bioinformatics analysis of the oxidosqualene gene and the amino acid sequence in mangrove plants J. Phys. : Conf. Ser. 801 012011.

[11] Basyuni M, Wasilah M and Sumardi 2017 Bioinformatics study of the mangrove actin genes J. Phys. : Conf. Ser. 801 012013.

[12] Pearson WR and Lipman DJ 1988 Improved tools for biological sequence comparison Proc. Natl. Acad. Sci. USA 85 2444–2448.

[13] Thompson JD, Higgins DG and Gibson TJ 1994 CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties, and weight matrix choice Nucleic Acid Res. 22 4673–4680.

[14] Page RD 1996 TreeView: an application to display phylogenetic trees on personal computers Comput. Appl. Biosci. 12 357–358.

[15] Felsenstein J 1985 Confidence limits on phylogenies: an approach using the bootstrap Evolution 39 783–791.

[16] Tateyama S, Wititsuwannakul R, Wititsuwannakul D, Sagami H and Ogura K 1999 Dolichols of rubber plant, ginkgo and pine Phytochemistry 51 11–15.

[17] Inafuku M, Basyuni M and Oku H 2016 Triterpenoid modulates the salt tolerance of lanosterol synthase deficient Saccharomyces cerevisiae, GIL77 Saudi J. Biol. Sci. http://dx.doi.org/10.1016/j.sjbj.2016.10.009.

[18] Liu J, Zhu Y, Du G, Zhou J and Chen J 2013 Exogenousergosterol protects Saccharomyces cerevisiae from D-limonenestress J. Appl. Microbiol. 114 482–491.

[19] Turk M, Plemenitas A and Gunde-Cimerman N 2011 Extremophilicyeasts: plasma-membrane fluidity as determinant of stress tolerance Fungal Biol. 115 950–958.

[20] Basyuni M and Sumardi 2017 Bioinformatics approach of salt tolerance gene in mangrove plant Rhizophora stylosa J. Phys. : Conf. Ser. 801 012012.