The bioinformatics method of polyprenol reductase genes in *Elaeis guineensis*

M Basyuni, R Hayati, D. Chalil

Department of Forestry, Faculty of Forestry, Universitas Sumatera Utara, Jl. Tri Dharma Ujung No. 1 Medan, North Sumatera 20155, Indonesia
Graduate School of Agricultural Sciences, Faculty of Agriculture, Universitas Sumatera Utara, Medan, North Sumatera 20155, Indonesia
Department of Agribusiness, Faculty of Agriculture, Universitas Sumatera Utara, Medan, North Sumatera 20155, Indonesia

Email: m.basyuni@usu.ac.id

Abstract. Even though led to minor components, *Elaeis guineensis* is well known to contain vitamin E, carotenoids, fatty acid, triterpenoid, and phytosterols. Recently the occurrence and pattern of long-chain polyisoprenoids from oil palm has been reported. The current study assesses the bioinformatics approaches to assay four probably polyprenol reductase genes from *E. guineensis* on NCBI database along with predicted the physical chemical, subcellular echolocation, and phylogenetic of polyprenol reductase. Several parameters of physicochemical polyprenol reductase in *E. guineensis* were varied among the genes observed. Results also showed the implication of understanding the variation and role of physical and chemical characteristics of the changed amino acids in oil palm polyprenol reductase genes. The subcellular location of these three genes has stored in the plasma membrane, Golgi body, endoplasmic reticulum (membrane and lumen). It is surprising to note that all the genes had the same values. The existence polyprenol reductase genes of *E. guineensis* in the clustering was supported by the dominated of carbon chain length of dolichols in the monocot plants.

1. Introduction

Oil palm plantation (*Elaeis guineensis* Jacq, Arecaceae) is the significant commodities in Indonesia to yield crude palm oil from the fruit mesocarp and palm kernel oil from the seed or kernel [1]. Even though produced minor components, *E. guineensis* is well known to contain vitamin E (tocopherols, tocotrienols), carotenoids, fatty acid, triterpenoid, and phytosterols [1-3]. Recently the occurrence and pattern of long-chain polyisoprenoids from oil palm has been reported [4-5]. Polyisoprenoids have two kinds of polyisoprenoids concerning the α-isoprene structure: polyprenol or dehydrodolichol (α-unsaturated isoprenoid alcohols) and dolichol (α-saturated isoprenoid alcohols) [4].

It has been shown that the distribution of polyisoprenoids in the *E. guineensis* displayed Type-II only, the presence of mutually polyprenol and dolichols. Type-I, showing predominance dolichol over polyprenol or Type-III, displaying a dominating polyprenol over dolichol was not detected [4-5]. However, in the NCBI database, three predicted polyprenol reductase had been reported. This enzyme converted polyprenol to dolichols in the dolichol biosynthetic [6]. Recently, predicted polyprenol reductase and its bioinformatics information have been depicting from *Kandelia obovata* [7-8].
Study on bioinformatics of the plant polyprenol reductase from *E. guineensis* is limited. The current report thus intended to analyse three expected polyprenols genes in *E. guineensis* by means of the bioinformatics analysis.

2. Materials and method

2.1. Materials

Three predicted polyprenol reductases from *E. guineensis* deposited in NCBI were collected. The NCBI reference sequence of the DNA and amino acid employed in this research in this way: XM_010920852 and XP_010919154 (*E. guineensis* polyprenol reductase 1, transcript variant X1), XM_010920853 and XP_010919155 (*E. guineensis* polyprenol reductase 1, transcript variant X2), and XM_019849699 and XP_019705258 (*E. guineensis* polyprenol reductase 1, transcript variant X3).

2.2. Physical and chemical features of the polyprenol reductase gene

The structure, physicochemical characteristics of three polyprenol reductase genes was determined by ProtParam online (web.expasy.org/protparam/). The calculated factors refer to the molecular mass, theoretical isoelectric of the DNA and amino acid employed in this research and extinction coefficient, probable half-life, instability index, aliphatic index, and mean hydropathicity as earlier used [9].

2.3. Possibility transit of peptide and subcellular localisation

To calculate transfer peptide, the target P1.1 server online (www.cbs.dtu.dk/services/targetp/) was employed. The site is according to the projected existence of any of the N-terminal pre-sequences chloroplast transit peptide (cTP), mitochondrial targeting peptide (mTP) and secretory pathway signal peptide (SP). Moreover, PSORT Prediction online (psort.hgc.jp/form.html) was applied to decide the subcellular localisation of polyprenol reductase genes as foregoing reported [9].

2.4. Phylogenetic analysis of monocots polyprenol reductase

The NCBI accession numbers of the monocots sequence applied in this analysis are following: *Brachypodium distachyon* X2 (XM_024455440), *B. distachyon* X1 (XM_010241988), *Zea mays* X2 (XM_008665246), *Z. mays* X1 (XM_008665245), *Setaria italica* (XM_004976504), *Sorghum bicolor* (XM_002488977), *Dendrobium catenatum* X4 (XM_020842138), *D. catenatum* X2 (XM_020842137), *D. catenatum* X1 (XM_020842136), *Phalaenopsis equestris* X3 (XM_020717201), *P. equestris* X2 (XM_020717200), *Asparagus officinalis* X3 (XM_020418745), *A. officinalis* X2(XM_020418744), *A. officinalis* X1(XM_020418743), *A. officinalis* (M_020394737), *Aegilops tauschii* subsp. *Tauschii* (M_020323691), *Ananas comosus* X5 (XM_02029570), *A. comosus* X4 (XM_020229569), *A. comosus* X3 (XM_020229568), *A. comosus* X2 (XM_020229567), *A. comosus* X1 (XM_020229566), *E. guineensis* X3 (XM_019849699), *E. guineensis* X2 (XM_010920853), *E. guineensis* X1 (XM_010920852), *Musa acuminate* subsp. *Malaccensis* X3 (XM_009402198), *M. acuminate* subsp. *Malaccensis* X2 (XM_009402197), *M. acuminate* subsp. *Malaccensis* X1 (XM_009402195), *Phoenix dactylifera* X4 (XM_017840866), *P. dactylifera* X3 (XM_008781658), *P. dactylifera* X2 (XM_008781654), *P. dactylifera* X1 (XM_008781648), *Oryza brachyantha* (XM_006652612), *O. sativa* Japonica Group (XM_015790678), *O. sativa* Japonica X2(XM_015778025), *O. sativa* Japonica Group X1(XM_015778024), and *Z. mays* uncharacterized (NM_001139278). Phylogenetic analysis of deduced amino acid alignment from monocots polyprenol reductase genes was conducted with CLUSTAL W ver. 1.83 [10] of the DNA Data Bank of Japan (Mishima, Shizuoka, Japan) accompanied by depiction with TreeView, ver. 1.6.6 [11] with neighbor-joining approach. Bootstrap analysis with 1000 duplications was employed to measure the ratio of the knots in the tree [12].
3. Results and Discussions

3.1. Physicochemical characteristics
Table 1 depicts a number line factors of physicochemical polyprenol reductase in *E. guineensis*. The polyprenol reductase consists of three variants: X1, X2, and X3. The length of the genes was varied with the genes determined. Coded amino acids were 223 to 337. It is remarkable the heterogeneity of comparative molecular weight, theoretical isoelectric factor worth, the total atomic numbering, scattering coefficient, instability coefficient, and general mean hydropathicity along with the analysed genes (Table 1).

The calculated half-life time interval was precisely similar along the genes (1.2 h) and was much lower that oxidosqualene cyclase genes [9], mangrove actin genes [13], *Kandelia obovata* polyprenol reductase gene [7]. However, the predictable half-life of polyprenol reductase in this study was similar to polyprenols reductase from other majorities of plant species [8].

Based on stability coefficients, all genes were non-stable proteins. A few genes have been shown as stable genes from plant polyprenol reductase genes, for example, *Glycine max*, *G. arboretum*, and *G. raimondii* [11]. Recognized genes also have been described in *Bruguiera gymnorrhiza* β-amyrin synthase (*BgβAS*) and *Rhizophora stylosa* cycloartenol synthase (*RsCAS*) [9], mangrove actin genes: *B. gymnorrhiza* *BgAct1*, *KcAct1* from *K. candel*, and *RsAct1* from *R. stylosa* [13]. These results indicated the significant knowledge for variation and role of physical and chemical characteristics of the distinguishable amino acids in *E. guineensis* polyprenol reductase genes [8].

| Table 1. Physicochemical properties of polyprenol reductase in *E. guineensis* | E. guineensis ID | Variant X1 | Variant X2 | Variant X3 |
|--------------------------------|-----------------|------------|------------|------------|
| Length of genes/bp            | 1571            | 1440       | 1229       |
| Number of encoded amino acids | 337             | 293        | 223        |
| Molecular weight              | 132333.71       | 121005.27  | 103348.42  |
| Theoretical isoelectric point values | 4.97          | 4.99      | 5.02       |
| Total number of atoms         | 16888           | 15409      | 13154      |
| Extinction coefficient        | 22000           | 20625      | 17750      |
| Half-life period              | 1.2h            | 1.2h       | 1.2h       |
| Instability coefficient       | 41.49           | 41.03      | 40.19      |
| Aliphatic index               | 22.80           | 22.67      | 22.95      |
| Average of hydropathicity     | 0.661           | 0.676      | 0.688      |

3.2. Potential transport of peptide and subcellular localisation
Table 2 shows the prospect of the potential transfer peptide in the probability of the likely transfer peptide in *E. guineensis* polyprenol reductase genes. Four consistencies were analysed: chloroplast transit peptide, mitochondrial target peptide, the signal peptide of the secretory pathway, and the likelihood view. The controls of chloroplast peptide but not indication peptide were relatively medium value, had that infrequent chloroplast transit peptide or somewhat in elevation display peptide of secretion pathway in oil palm polyprenol reductase genes. It is prominent that mitochondria target peptide value differentiated from 0.057 to 0.063, representing that is expected to be an occurrence.

| Table 2. The promising of the potential transit peptide of polyprenol reductase in *E. guineensis* | E. guineensis ID | Reliability |
|--------------------------------|-----------------|-------------|
| Chloroplast transit peptide  | Mitochondrial target peptide | Signal peptide of secretory pathway | Reliability prediction |
| Variant X1                    | 0.127           | 0.057       | 0.144       | 5           |
| Variant X2                    | 0.138           | 0.063       | 0.170       | 5           |
| Variant X3                    | 0.127           | 0.057       | 0.144       | 5           |
The utmost mitochondrial target peptide was X2 variant, an *E. guineensis* polyprenol reductase gene. The low value of mitochondrial target peptide was contrasting to previous findings on the concentrated mitochondrial target peptide amongst the phytosterol genes from mangrove tree species [9], actin gene from mangrove Rhizophoraceae [10], *Kandelia obovata* polyprenol reductase [7-8], and salt tolerance gene from *Rhizophora stylosa* [14]. However, the low value was similar results detected in plant polyprenol reductase [15]. It is noteworthy that reliability prediction (5) was higher value comparing to oxidoqualene cyclase genes from mangrove plants (2-4) [9], mangrove actin genes (2-3) [13], *K. obovata* polyprenol reductase genes (4-5) [7-8], and most salt tolerance genes from *R. stylosa* (3-4) [14].

Table 3 shows the subcellular location of polyprenol reductase genes in *E. guineensis*. The subcellular localisation of these three genes has existed in the plasma membrane, Golgi body, endoplasmic reticulum (membrane and lumen). It is surprising to note that all the genes had the same values (Table 3). Very newly to be reported that the expressed gene of two triterpene synthases, *BgBAS* and *RsM1* accelerated the triterpene concentration of plasma membrane portions along with traced in the plasma membrane localisation [16]. Similarly, it has been accepted understandings that the plasma membrane is the major defence counter to alterations in physicochemical adaptable to the changing externals [17]. A number of salt tolerance genes from *R. stylosa* localised in the plasma membrane sustained the earlier documents on the prominence of the lipid bilayer for abiotic stress tolerance along with salinity stress [16].

| Table 3. Subcellular localization of polyprenol reductase in *E. guineensis* |
|------------------------|-----------------|---------------------|------------------------|
| *E. guineensis* ID     | Plasma Membrane | Golgy Body          | endoplasmic reticulum  |
|                        |                 |                     | (membrane)             |
| Variant X1             | 0.640           | 0.460               | 0.370                  |
| Variant X2             | 0.640           | 0.460               | 0.370                  |
| Variant X3             | 0.640           | 0.460               | 0.370                  |

3.3. Phylogenetic analysis of polyprenol reductase gene

To clarify the similarity between the polyprenol reductase genes in monocots plant family, a dendogram was drawn. The polyprenol reductase 1 (LOC105043343) of *E. guineensis*, variants X1, X2, and X3 sequences were 1571, 1440, and 1229 bp, correspondingly. These DNA sequences encoded 227, 293, and 223 amino acids residues for three circumstances, respectively. Variant X1 and X2 shared 87% identities in their amino acid sequences and 77% between X1 and X3 in their amino acid sequences. X2 and X3 displayed high homologies in amino acid (98 %). To clarify the similarity amongst the polyprenol reductase gene in *E. guineensis* with monocot plant polyprenol reductases, a clustering tree was created by their amino acid sequences (Figure 1). The phylogenetic tree characterizes several branches; where *E. guineensis* joined with other monocots such as *Asparagus officinalis* (4 genes), *Aegilops tauschii*, *Oryza brachyanta*, and *Zea mays*.

The presence of polyprenol reductase was parallel with our previous reports that dolichols were predominated in true mangrove monocots of *Nypa fruticans* leaves (C75-C90) and roots (C75-C90) [18]. Similarly, monocot species of coastal plants, *Pandanus odoratatissima* also contain dolichols only. In the leaves, dolichols with the chain length of C75-C95, whereas dolichols C80-C90 was detected in the roots [19]. The occurrence of dolichols only and did not contain polyprenols in these species was supported by the previous study on mangrove and coastal plants [18-19]. This polyprenol reductase enzyme changed polyprenol to dolichols in the dolichol biosynthetic [7-8].

Very newly it has been shown that salt stress changes the polyisoprenoid concentrations in mangrove plants [20]. The shift of polyisoprenoids containing polyprenols, dolichols, and bombiprenone as well supported the accepted views on the accumulation of polyisoprenoid upon abiotic and biotic tolerance [18-20].
Figure 1. Clustering tree of monocot plant polyprenol reductase was comprising three polyprenols reductase from *E. guineensis*. Dendrogram of deduced amino acid sequences was done with the neighbour-joining approach of the CLUSTAL W [10]. The specified scale signifies 0.1 amino acid switch per place. Numbers show bootstrap value from 1000 measurements. The NCBI accession numbers of the amino acid sequence applied this examination are shown in the Materials section.

4. Conclusions
The present-day study confirmed study clarified the variation and role of physical and chemical characteristics of the *E. guineensis* polyprenol reductase genes. Our report promoted preceding studies on the subcellular location of salt tolerance as well as oxidosqualene genes positioned in the plasma membrane. The existence polyprenol reductase genes of *E. guineensis* in the dendrogram was supported by the dominated of carbon chain length of dolichols in the monocot plants.

Acknowledgement
This work was supported in part by a Penelitian Strategis Nasional Institusi (PSN Institusi 2018 to MB) from the Directorate for Research and Community Service, Ministry of Research, Technology and Higher Education, Republic of Indonesia. The authors are grateful to the University of Sumatera Utara.

References
[1] Basyuni M, Amri N, Putri LA, Syahputra I, Arifiyanto D 2017 Characteristics of fresh fruit bunch yield and the physicochemical qualities of palm oil during storage in North Sumatra, Indonesia *Indones. J. Chem.* 17 182–190.
[2] Sundram K, Sambanthamurthi R, Tan YA 2003 Palm fruit chemistry and nutrition Asia Pac. J. Clin. Nutr. 12 355–362.

[3] Sambanthamurthi R, Sundram K, Tan YA 2000 Chemistry and biochemistry of palm oil Prog. Lipid Res. 39 507–558

[4] Arifiyanto D, Basyuni M, Sumardi, Putri LAP, Siregar ES, Risnasari I, 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] Basyuni M, Wati R, Deni I, Tia AR, Slamet B, Siregar ES, Syahputra I 2018 Cluster analysis of polyisoprenoid in oil palm (Elaeis guineensis) leaves in different land-uses to find the possible cause of yield gap from planting materials Biodiversitas 19 1492–1501.

[6] Basyuni M, Wati R 2018 Bioinformatics analysis of the predicted polyphenol reductase gene in higher plants J. Phys.: Conf. Ser. 978 012050.

[7] Basyuni M, Wati R, Sagami H, Oku H, Baba S 2018 Bioinformatics approach of three partial polyphenol reductase genes in Kandelia obovata J. Phys.: Conf. Ser. 978 012044.

[8] Basyuni M, Sagami H, Baba S, Oku H 2018 Genome sequence analysis of predicted polyphenol reductase gene from mangrove plant Kandelia obovata IOP Conf. Ser. : Earth Environ. Sci. 130 012039.

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

[10] Thompson JD, Higgins DG, 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.

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

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

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

[14] Basyuni M, Sumardi 2018 Bioinformatics approach of salt tolerance gene in mangrove plant Rhizophora stylosa J. Phys. : Conf. Ser. 801 012012.

[15] Basyuni M, Wati R 2018 Bioinformatics analysis of the predicted polyphenol reductase genes in higher plants J. Phys. : Conf. Ser. 978 012050.

[16] Inafuku M, Basyuni M, Oku H 2018 Triterpenoid modulates the salt tolerance of lanosterol synthase deficient Saccharomyces cerevisiae, GIL77 Saudi J. Biol. Sci. 25 1-9.

[17] Henry G, Deleu M, Jourdan E, Thonart P, Ongena M 2011 The bacterial lipopeptide surfactin targets the lipid fraction of the plant plasma membrane to trigger immune - related defence responses Cell. Microbiol. 13 1824-1837.

[18] Basyuni M, Sagami H, Baba S, Oku H 2017 Distribution, occurrence, and cluster analysis of new polypropenyl acetones and other polyisoprenoids from North Sumatran mangroves Dendrobiology 78 18–31.

[19] Basyuni M, Wati R, Sagami H, Sumardi, Baba S, Oku H 2018 Diversity and abundance of polyisoprenoid composition in coconut plant species from North Sumatra, Indonesia Biodiversitas 19 1–11.

[20] Basyuni M, Sagami H, Baba S, Putri LAP, Wati R, Oku H 2017 Salinity alters the polyisoprenoid alcohol content and composition of both salt-secreting and non-salt-secreting mangrove seedlings HAYATTI J. Biosci. 24 206-2014.