Bioinformatics study of the mangrove actin genes

M Basyuni[1*], M Wasilah[1], Sumardi[2]

[1]Department of Forestry, Faculty of Forestry, University of Sumatera Utara, Jl. Tri Dharma Ujung No. 1 Medan, North Sumatera 20155, Indonesia
[2]Faculty of Pharmacy, Tjut Nyak Dhien University, Jl. Rasmi No. 28 Medan 20123, Indonesia

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

Abstract. This study describes the bioinformatics methods to analyze eight actin genes from mangrove plants on DDBJ/EMBL/GenBank as well as predicted the structure, composition, subcellular localization, similarity, and phylogenetic. The physical and chemical properties of eight mangroves showed variation among the genes. The percentage of the secondary structure of eight mangrove actin genes followed the order of α helix > random coil > extended chain structure for *BgAct1, KcAct1, RsAct1*, and *A. corniculatum Act*. In contrast to this observation, the remaining actin genes were random coil > extended chain structure > α helix. This study, therefore, shown the prediction of secondary structure was performed for necessary structural information. The values of chloroplast or signal peptide or mitochondrial target were too small, indicated that no chloroplast or mitochondrial transit peptide or signal peptide of secretion pathway in mangrove actin genes. These results suggested the importance of understanding the diversity and functional of properties of the different amino acids in mangrove actin genes. To clarify the relationship among the mangrove actin gene, a phylogenetic tree was constructed. Three groups of mangrove actin genes were formed, the first group contains *B. gymnorrhiza BgAct* and *R. stylosa RsAct1*. The second cluster which consists of 5 actin genes, the largest group, and the last branch consist of one gene, *B. sexagula* Act. The present study, therefore, supported the previous results that plant actin genes form distinct clusters in the tree.

1. Introduction

Higher plants including mangrove plants contain families of actin-encoding genes which are divergent and differently expressed and constitutively involved in basic housekeeping function required for cell maintenance [1]. They are used as endogenous internal references for normalizing gene expression studies. Actin gene, therefore, plays a significant role in simple eukaryotic cellular processes such as motility, cell growth regulation, cell differentiation and stability structure [2].

The identification of plant actin genes will help the researcher to choose the appropriate internal standard with stable expression under particular conditions, either biotic or abiotic stress [1]. The actin genes of dicot and monocot plants are well conserved and more distinct from each other. Therefore, actin genes are more strongly linked than duplications, represents that actin gene is purely from gene family of ancient times [1].

Mangrove is halophytes that are defined ecologically by their location in upper inter-tidal zones of tropical and sub-tropical climates and physiologically by their ability to withstand high concentrations of salt or low levels of soil aeration. A number cDNA cloning of actin genes have been reported from mangrove plants, *Bruguiera gymnorrhiza* [3], *Kandelia candel* [3], *Rhizophora stylosa* [4], *Aegiceras*

corniculatum [5], several others have been deposited in the DDBJ/GenBank/EMBL. This information can be an excellent molecular tool and allow us used the bioinformatics methods to analyze eight mangrove actin genes as well as predicted the structure, composition, similarity, subcellular localization and phylogenetic to provide references for the actin genes and to further understand the diversity and functional of properties of amino acids in mangrove actin genes.

2. Materials and method

2.1. Materials collection

A total of eight mangrove actin genes officially deposited in the DDBJ/EMBL/GenBank were collected. The DDBJ/GenBank/EMBL accession numbers of the DNA sequence and amino acid sequence of used this analysis are as follows: AB491932, BAI39620 (*Bruguiera gymnorrhiza BgAct1*), AB491931, BAI39619 (*Kandelia candel KcAct1*), AB573024, BAK52268 (*Rhizophora stylosa RsAct1*), DQ831135, ABI144668 (*Aegiceras corniculatum Act*), JX679722, AFV91359 (*B. sexangula Act*), JX679721, AFV91358 (*B. gymnorrhiza Act*), JX679720, AFV91357 (*B. cylindrica Act*), and JX679723, AFV91360 (*K. candel Act*)

2.2. Physicochemical properties and secondary structure of the actin gene

The composition, physical and chemical properties of DNA and amino acid sequences of actin genes were analyzed using Protparam online analysis (web.expasy.org/protparam/). The computed parameters include the molecular weight, theoretical isoelectric point values, amino acid composition, atomic composition, extinction coefficient, estimated half-life, instability index, fat coefficient, and grand average hydrophilicity. HNN Secondary structure prediction method online (https://npsa-prabi.ibcp.fr) was used to analyze the secondary structure of mangrove actin genes.

2.3. Potential transit peptide in the mangrove actin gene

The targetP 1.1 Server online (www.cbs.dtu.dk/services) was used for transit peptide prediction of mangrove actin genes. The location assignment is based on the predicted presence of any of the N-terminal pre-sequences chloroplast transit peptide (cTP), mitochondrial targeting peptide (mTP) and secretory pathway signal peptide (SP).

2.4. Similarity and phylogenetic analysis of mangrove actin gene

The amino acid sequences were aligned, and similarity scores were obtained using the FASTA ver. 3.4t26 [6] of the DNA Data Bank of Japan (Mishima, Shizuoka, Japan). The best score of results is shown in Table 4. Phylogenetic analysis of deduced amino acid alignment from 8 mangrove actin genes was conducted with CLUSTAL W ver. 1.83 [7] of the DNA Data Bank of Japan (Mishima, Shizuoka, Japan) followed by drawing with TreeView, ver. 1.6.6 [8] based on a neighbor-joining method. Bootstrap analysis with 1000 replications was used to assess the strength of the nodes in the tree [9]. The DDBJ/GenBank/EMBL accession numbers of the DNA sequence and amino acid sequence of used this analysis is described in subsection Materials collection.

3. Results and Discussions

The results will be discussed in three subsections; physicochemical properties and secondary structure of the actin gene, potential transit peptide in the salt tolerance gene, and similarity and phylogenetic analysis of mangrove actin gene.

3.1. Physicochemical properties and secondary structure of the actin gene

Table 1 shows the physical and chemical properties of mangrove actin genes. The initiation codon ATG of actin gene only found in *A. corniculatum* actin (DQ 831135), where the stop codons were TAA, AAG, and GTA. No open reading frame length found, all are partial actin genes. Some encoded amino acids were 33 to 297. *K. candel Act* had the smallest relative molecular mass; *K. candel KcAct1*

had the largest one. Mangrove actin genes had variation theoretical isoelectric point value (5 to 10). *K. candel Act* had a minimum total number of atoms; *B. gymnorrhiza BgAct1* had a maximum total number of atoms. The half-life period varied among actin gene. Based on stability coefficients, *BgAct1, KcAct1, RsAct1*, and *A. corniculatum Act* were stable proteins, the remaining were non-stable proteins. These results suggested the importance of understanding the diversity and functional of properties of the different amino acids in mangrove actin genes [1].

Tabel 1. Physical and chemical properties of the actin genes cDNA in mangrove plants

Nucleotide accession number	AB491932	AB491931	AB573024	DQ831135	JX679722	JX679721	JX679720	JX679723
Length of genes/bp	894	894	894	930	221	209	203	204
Open reading frame length/bp	nd	nd	nd	nd	nd	nd	nd	nd
Start site and codon	nd	nd	nd	55ATG	nd	nd	nd	nd
Stop site and codon	894 TAA	894 TAA	894 TAA	930 AAG	99 TAA	87 TAA	90 AAG	84 GTA
Number of encoded amino acids	297	297	297	292	33	29	30	28
Relative molecular mass	33377.30	33459.35	33274.23	32331.91	3606.24	3203.78	3334.97	3075.60
Theoretical isoelectric point values	5.37	5.27	5.27	5.17	10.01	10.00	10.00	9.70
Positively charged residues	31	30	30	30	4	4	4	3
Negatively charged residues	40	40	40	42	1	1	1	1
Total number of atoms	4684	4681	4670	4529	520	467	484	446
Extinction coefficient	1.288	1.285	1.247	0.897	1.938	2.182	2.096	2.273
Half life period	2.8 h	2.8 h	2.8 h	30 h	20 h	1.3 h	30 h	20 h
Instability coeffiient	37.81	35.85	38.64	30.08	49.74	41.94	40.88	46.11
Fat coeffiient	86.73	83.43	87.37	85.89	91.52	104.14	100.67	107.86
Overall average hydrophilicity	-0.179	-0.207	-0.146	-0.213	0.061	0.114	0.173	0.257

The secondary structure of OSC genes consist of α helix, extended chain structure, and random coil with their relative proportion was depicted in Table 2. Table 2 shows that the percentage of the secondary structure of eight mangrove actin genes followed the order of α helix > random coil > extended chain structure for *BgAct1, KcAct1, RsAct1*, and *A. corniculatum Act*. In contrast to this observation, the remaining actin genes were random coil> extended chain structure > α helix. This study, therefore, indicated the prediction of secondary structure was performed for necessary structural information. This study also suggested the diversity of secondary structure among the mangrove actin genes.

Table 2. Secondary structure analysis of mangrove actin gene

Protein registry number	**α helix number**	**α helix ratio/ %**	**Extended chain structure number**	**Extended chain structures ratio/ %**	**Random coil number**	**Random coil ratio/ %**
BAI39620	121	40.74	56	18.86	120	40.40
BAI39619	120	40.40	58	19.53	119	40.07
BAK52268	122	41.08	57	19.19	118	39.73
ABI14668	105	35.96	61	20.89	126	43.15
AFV91359	5	15.15	12	36.36	16	48.48
AFV91358	5	17.24	10	34.48	14	48.28
AFV91357	5	16.67	11	36.67	14	46.67
AFV91360	5	17.86	9	32.14	14	50.00

3.2. Potential transit peptide in the salt tolerance gene

Table 3 shows the possibility of the potential transit peptide in mangrove actin genes. There are three possibilities: chloroplast transit peptide, mitochondrial target peptide and signal peptide of secretory pathway along with the prediction probability. The values of chloroplast or signal peptide or signal

peptide secretory pathway were too small, indicated that no chloroplast transit peptide or mitochondrial peptide or signal peptide of secretion pathway in eight actin genes.

At primary structural level, transit peptide sequence are highly divergent, showing that the transit peptide contains multiple domains that provide either distinct or overlapping functions [10]. Therefore it will provide an update on the limited structural information of some transit peptide and the evolution [10]

Table 3. Possibility of the potential transit peptide in salt tolerance gene

Protein accession number	Reliability			
	Chloroplast transit peptide	Mitochondrial target peptide	Signal peptide of secretory pathway	Reliability prediction
BAI39620	0.111	0.023	0.103	2
BAI39619	0.136	0.021	0.087	2
BAK52268	0.124	0.022	0.092	2
ABI14668	0.173	0.109	0.097	2
AFV91359	0.127	0.154	0.173	3
AFV91358	0.092	0.229	0.139	3
AFV91357	0.080	0.238	0.156	3
AFV91360	0.100	0.212	0.147	3
ABI51282	0.177	0.040	0.083	2

3.3. Similarity and phylogenetic analysis of mangrove actin gene

Eight mangrove actin sequences in this study were aligned each other as displayed in Table 4. The highest similarity (100%) was shown by *B. gymnorrhiza* BgAct1, *K. candel* KcAct1 and *R. stylosa* RsAct1 with four short DNA sequences of *B. sexangula* Act, *B. gymnoorhiza* Act, *B. cylindrica* Act, and *K. candel* Act (221-209 bp). Among of these short sequences also show identical sequences (100% similarity). The clustering of the mangrove actin genes suggests that these sequences may have involved in the evolution of plant dicots and they are functionally conserved [1]. This study makes the actin gene useful for phylogenetic analysis, especially using neighbor joining.

Tabel 4. The similarity of amino acid sequences between mangrove actin gene

Clones	1	2	3	4	5	6	7	8
1. *B. gymnorhiza* BgAct1	100							
2. *K. candel* KcAct1	97	100						
3. *R. stylosa* RsAct1	98	96	100					
4. *A. corniculatum* Act	69	68	69	100				
5. *B. sexangula* Act	100	100	100	18	100			
6. *B. gymnorhiza* Act	100	100	100	20	100	100		
7. *B. cylindrica* Act	100	100	100	20	100	100	100	
8. *K. candel* Act	100	100	100	21	100	100	100	100

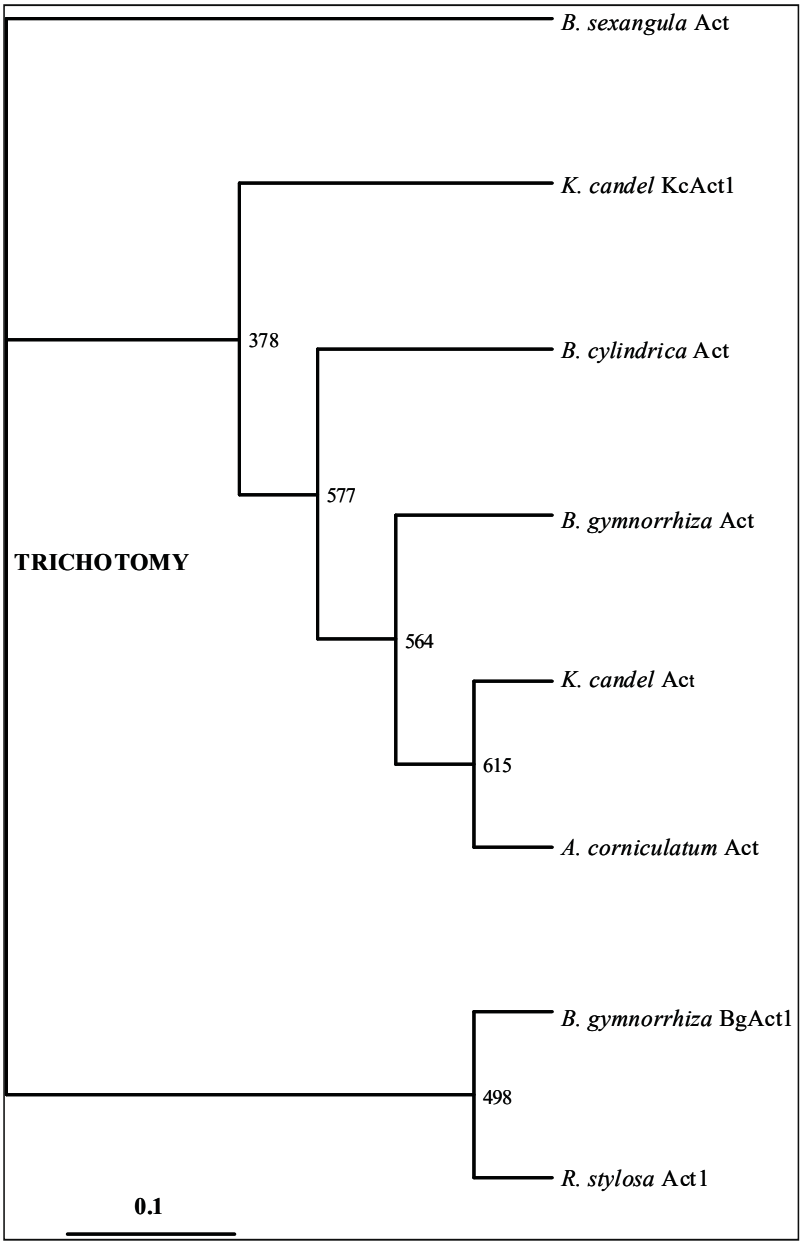

Figure 1. Phylogenetic tree of mangrove actin genes. Phylogenetic tree of deduced amino acid sequences was constructed with the neighbor-joining method of the CLUSTAL W [7]. The indicated scale represents 0.1 amino acid substitutions per site. Numbers indicate bootstrap value from 1000 replicates. The DDBJ/GenBank/EMBL accession numbers of the amino acid sequence of used this analysis are reported in the Materials subsection.

Figure 1 shows there are three groups of mangrove actin genes; the first group contains *B. gymnorrhiza BgAct* and *R. stylosa RsAct1*. The second cluster which consists of 5 actin genes, the largest group and the last branch consist of one gene, *B. sexagula Act.* The clustering of the mangrove actin genes suggests that these sequences may have involved in the evolution of plant dicots and they are functionally conserved

4. Conclusions

The present study demonstrated the importance of understanding the diversity and functional of properties of the different amino acids in eight mangrove actin genes. Furthermore, our data improve our understanding of the structure of actin genes in the mangrove.

Acknowledgment

This study was partly supported by an International Research Collaboration and Scientific Publication Grant (No. 117/SP2H/LT/DRPM/II/2016 to MB) from the Directorate for Research and Community Service, Ministry of Research, Technology and Higher Education, Republic of Indonesia.

References

[1] Slajcherova K, Fiserova J, Fischer L, Schwarzerova K 2012 Multiple actin isotypes in plants: diverse genes for diverse roles? *Frontiers Plant Sci.* **3**:1–8.

[2] Wery J, Dalderup MJM, Linde JT, Boekhout T, van Ooyen AJJ 1996 Structural and phylogenetic analysis of the actin gene from the yeast *Phaffia rhodozyma Yeast* **12**: 641–651.

[3] Basyuni M, Baba S, Inafuku M, Iwasaki H, Kinjo K, Oku H 2009 Expression of terpenoid synthase mRNA and terpenoid content in salt-stressed mangrove *J. Plant Physiol*. **166**: 1786–1800.

[4] Basyuni M, Kinjo Y, Baba S, Shinzato N, Iwasaki H, Siregar EBM, Oku H 2011 Isolation of salt stress tolerance genes from roots of mangrove plant, *Rhizophora stylosa* Griff., using PCR-based suppression subtractive hybridization *Plant Mol. Biol. Rep.* **29**: 533–543.

[5] Ya-Lan Peng, You-Shao Wang, Ji-Dong Gu 2015 Identification of suitable reference genes in mangrove *Aegiceras corniculatum* under abiotic stresses. Ecotoxicology (2015) 24:1714–1721

[6] Pearson WR, Lipman DJ 1988 Improved tools for biological sequence comparison *Proc. Natl. Acad. Sci. USA* 85: 2444–2448.

[7] 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.

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

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

[10] Bruce BD 2000 Chloroplast transit peptides: structure, function and evolution *Trends Cell Biol* **10**: 440–447.