PLACEMENT OF *Syzygium boerlagei* (Merr.) Govaerts (MYRTACEAE) CONFIRMED WITH ATPB-RBCL INTERGENIC SPACER

PUDJI WIDODO[1*], TATIK CHIKMAWATI[2] AND YAYAN WAHYU CANDRA KUSUMA[3]

[1]*Faculty of Biology, Universitas Jenderal Soedirman, Purwokerto 53122, Indonesia*
[2]*Department of Biology, Faculty of Mathematics and Natural Sciences, Institut Pertanian Bogor, Bogor 16680, Indonesia*
[3]*PKT Kebun Raya Bogor – Indonesian Institute of Sciences (LIPI), Bogor 16003, Indonesia*

Received 06 March 2017 / Accepted 15 November 2017

ABSTRACT

A molecular analysis was conducted to determine whether *Eugenia boerlagei* Merr. (Myrtaceae) belongs to genus *Eugenia* or *Syzygium* based on sequences of cpDNA fragments namely atpB-rbcL intergenic spacer. The study used seven specimens of *Syzygium* sect. Jambosa, three of *Syzygium* sect. *Syzygium*, two of *Eugenia s.s.* and one of *Eugenia boerlagei* Merr. with *Baeckea ovalifolia* and *B. tuberculata* as the outgroup. The results showed that *Eugenia boerlagei* is appropriately placed under the genus *Syzygium*.

Keywords: atpB-rbcL intergenic spacer, chloroplast DNA, *Eugenia*, Myrtaceae, *Syzygium*

INTRODUCTION

Eugenia boerlagei Merr. (type: Robinson 1872, lost; iso: BO, L, elsewhere? Moluccas, Amboina, Liang) (Myrtaceae) is a shrub or small tree characterized by lateral and terminal, slender, 3-flowered inflorescences, long pedicels, and a long, narrowed calyx-tube, which, with sepals and petals is glandular-punctate. The species was dedicated to J.G. Boerlage who contracted a fever while conducting a botanical exploration of Amboina in the year 1900, which resulted in his untimely death (Cox & Merrill 1916). Govaerts *et al.* (2008) transferred it to *Syzygium boerlagei* Govaerts. In this study, a molecular analysis was done to verify whether *Eugenia boerlagei* should be transferred to *Syzygium* or retained in *Eugenia* L.

Although it is rare, at least in L there is only an isotype, but then there are also c. 55 boxes of *Eugenia* and *Syzygium* indet. *Eugenia boerlagei* was selected because it was just transferred in 2008, some months before the authors started working, and because of the availability of the living materials in Bogor.

*Corresponding author: pwidodo@unsoed.ac.id

Syzygium is the largest genus of the Myrtaceae, comprising c. 1200 species in the Old World (Craven *et al.* 2006) or approximately 1040 species (Govaerts *et al.* 2008). The current concept of *Syzygium* includes species with and without an inter-cotyledonary inclusion, inflorescence either solitary, axillary or terminal, calyx either calyptrate or free (Craven *et al.* 2006). Recent revision of *Syzygium s.l.* (sensu Hyland 1983 and Biffin *et al.* 2006) is based on a sub-generic arrangement that distinguishes *Syzygium s.s.* from the traditionally associated taxa by the presence of indistinct calyx lobes and coherent petals.

The generic taxonomy of *Syzygium* has long been associated with *Eugenia*, from which it is only weakly distinguished by macro-morphological data. Anatomical and molecular evidences now suggest that these two groups are in fact quite distantly related (Biffin *et al.* 2006). Most of the species of *Eugenia* in the Old World have been transferred to *Syzygium*.

This study used atpB-rbcL intergenic spacer because it is a highly conserved cytoplasmic molecule inherited clonally (without recombination), which has been shown to be a powerful tool to document the parentage of

polyploids and the phylogenetic relationships between distinct polyploid taxa in polyploid complexes. Furthermore, atpB-rbcL has been used successfully to support the phylogeny of the moss (Chiang *et al.* 2000).

In plants, the mitochondrial and chloroplast genomes are evolving too slowly to provide enough variation. Thus, for taxonomists, the current strategy is to sequence several DNA regions (Taberlet *et al.* 2007). The chloroplast genome that we used was the atpB-rbcL spacer which is a noncoding region of the genome that has been used in phylogenetic studies of Angiosperms (Manen *et al.* 1994; Manen & Natali 1995). In this case rbcL and atpB are transcribed in opposite directions.

Plant molecular systematics and DNA barcoding techniques rely heavily on the use of chloroplast gene sequences (Dong *et al.* 2012). The plastid locus most commonly sequenced by plant systematists for phylogenetic purposes is rbcL, followed by the trnL-F intergenic spacer, matK, ndhF, and atpB (Shaw *et al.* 2015). The sequence data from the atpB-rbcL intergenic spacer region has been used successfully to support DNA barcode loci to distinguish one species from the others (Costion *et al.* 2016).

This study presents a molecular analysis of some species of *Syzygium* and *Eugenia* based on the sequence of the atpB-rbcL intergenic spacer from representative samples of *Syzygium* and *Eugenia*. The objectives are: (1) to determine the placement of *Eugenia boerlagei*, whether in *Syzygium* or in *Eugenia*; (2) to provide a better understanding of the relationships between *Eugenia* and *Syzygium* which are slightly morphologically different.

MATERIALS AND METHODS

Sample Preparation

Samples were obtained from living plants growing in the Bogor Botanic Garden and its vicinity. The ingroup represents a sampling of morphological diversity within *Syzygium*. Ten types of *Syzygium* comprising six specimens of sect. Jambosa, four of sect. *Syzygium*, two of *Eugenia s.s.*, two of *Baeckea ovalifolia* and *B. tuberculata* from GenBank and one of *Eugenia boerlagei* (*Syzygium boerlagei*) were examined (Table 1). Voucher specimens have been stored in the Herbarium Bogoriense (BO and Herbarium Fakultas Biologi Unsoed, PUNS). The sequences of *Eugenia* and *Syzygium* were submitted to GenBank on 22 November 2017 and are now waiting for their accession number.

DNA Extraction, Amplification, Sequencing and Alignment

Total DNA was extracted from fresh material following the standard hexadecyl trimethyl ammonium bromide (CTAB) extraction methods (Doyle & Doyle 1990). Double stranded DNA was directly amplified by Polymerase Chain Reaction (PCR) for all loci. The reaction volumes were 25 µL and contained 2.5 µL PCR buffer, 1 µL dNTPs, 0.1 µL in each of the 10 mM primers, 1.5 µL 25 mM $MgCl_2$, 0.1 µL TaqPol and 15.2 µL ddH_2O. Approximately 4.5 µL genomic DNA was added to the PCR mixture of *Eugenia*. The primers used in this study for atpB-rbcL intergenic spacers are: atpB-1: 5'- ACATCKARTACKGGACC AATAA-3' and reverse rbcL-1: 5'- AACACCAGCTTTRAATCCAA-3' (Chiang *et al.* 1998). A non coding cpDNA fragment namely atpB-rbcL spacer was amplified.

PCR was performed with 4 min at 94ºC for the activation of the polymerase, followed by 35 cycles of 45 sec at 94ºC, 45 sec at 55ºC, 2 min at 42ºC, with a final extension period of 10 min at 72ºC. Following the manufacturers' protocol prior to sequencing, the PCR product was checked on 1% agarose gel, and was purified using a purification kit of Wizard SV Gel and PCR clean up system (PROMEGA). The DNA concentration was measured with the nanodrop. Cycle sequencing was performed by MACROGEN Korea. The sequences were edited manually and sequently manually adjusted using Sequencher 4.6 and MEGA 6.0 (Tamura *et al.* 2013).

Table 1 List of examined voucher specimens of *Baeckea*, *Eugenia*, and *Syzygium*

Taxa	Voucher detail	Localities	Accession
Eugenia boerlagei Merr. – *Syzygium boerlagei* (Merr.) Govaerts	*Widodo 143*	Mollucas, Ambon (KRB*)	MG669291
Eugenia pyriformis Cambess	*Widodo 142*	Brazil, Indonesia (KRB)	MH191262
Eugenia uniflora L.	*Widodo 141*	Java Bogor (IPB**)	SAMN08056079
Syzygium aqueum (Burm. f.) Alston	*Widodo 132*	Java Bogor (IPB)	MH191263
Syzygium aromaticum (L.) Merr. & L.M. Perry	*Widodo 137*	Java Bogor (IPB)	MH191264
Syzygium littorale (Blume) Amshoff	*Widodo 135*	Borneo (KRB)	MH191265
Syzygium polyanthum (Wight) Walp.	*Widodo 139*	Java Bogor (KRB)	MH191266
Syzygium polycephalum (Miq.) Merr. & L.M. Perry	*Widodo 136*	Java Bogor (KRB)	MH191267
Syzygium samarangense (Blume) Merr. & L.M. Perry	*Widodo 131*	Java Bogor (IPB)	MH191268

Note: **KRB = cultivated in Kebun Raya Bogor (Bogor Botanic Gardens)
**IPB = cultivated in Institut Pertanian Bogor (Bogor Agricultural University)

Phylogenetic Analysis

Cladistic analyses of the atpB-rbcL IGS sequence data were performed by using a MEGA 6.0 maximum parsimony criterion (Tamura *et al.* 2013). The methods produced phylogenetic trees that provided insights concerning major general evolutionary trends in the *Eugenia* and *Syzygium*. Notable findings were: (i) *Eugenia boerlagei* is a sister species to *Syzygium aqueum* (Burm. f.) Alston; (ii) the two *Eugenia* samples are distantly related to all *Syzygium*.

The fit of character data on phylogenetic hypotheses (Swofford 1998) was evaluated by the consistency index, CI (Kluge & Farris 1969), and the retention index, RI (Archie 1989; Farris 1989). The statistical significance of the CI was determined according to the method of Klassen *et al.* (1991). Confidence in the clades was tested by bootstrapping (Effron 1982; Felsenstein 1985) with 100 replicates of heuristic searches on the 50% majority rule trees. The nodes with bootstrap values >0.70, as a rule of thumb, were considered significantly supported with 395% probability (Hillis & Bull 1993).

RESULTS AND DISCUSSION

DNA Sequencing and Alignment

The data for some *Syzygium*, the length of atpB-rbcL intergenic spacer varied from 903 to 962 base pairs within Myrtaceae. The shortest is *Syzygium lineatum* (903 bp), followed by *Syzygium astronioides* (912 bp), *Syzygium samarangense* (916 bp), and *Syzygium aqueum* (920 bp). While the longest is *Syzygium cumini* (962 bp), followed by *Syzygium malaccense* (955 bp), *Syzygium aromaticum* (942 bp) etc. (Table 1). The position of *Eugenia uniflora* and *Eugenia pyriformis* is in between those of *Syzygium*. Thus, the length of atpB-rbcL spacer does not determine the differences between *Eugenia* and *Syzygium*.

Table 2 Length variation, AT and GC content of *atpB-rbcL* intergenic spacer in *Baeckea, Eugenia,* and *Syzygium*

Taxa	Sequence length (bp)	AT* content	GC** content
E_boerlagei	941	106	16
E_pyriformis	923	106	**18**
E_uniflora	925	107	**18**
S_aqueum	920	107	16
S_aromaticum	942	111	16
S_littorale	925	107	16
S_polyanthum	936	105	16
S_polycephalum	921	108	16
S_samarangense	916	107	16

Note: **AT = Adenine Thymine
**GC = Guanine Cytosine

In general, *Syzygium* is AT-rich, where the AT content of the spacer ranges from 105 – 111. Though the GC content ranges from 16 - 18, the AT content in both *Eugenia* is 106 and 107, in between all *Syzygium*. Thus, the AT content can not be used to determine the differences between *Eugenia* and *Syzygium*. However, the GC content of both *Eugenia s.s.* is the richest (18) compared to *Syzygium*. Thus, the GC content of the atpB-rbcL intergenic spacer may indicate the differences between *Eugenia* and *Syzygium* (Fig. 1). This shows that *Eugenia boerlagei* should be named as *Syzygium boerlagei*.

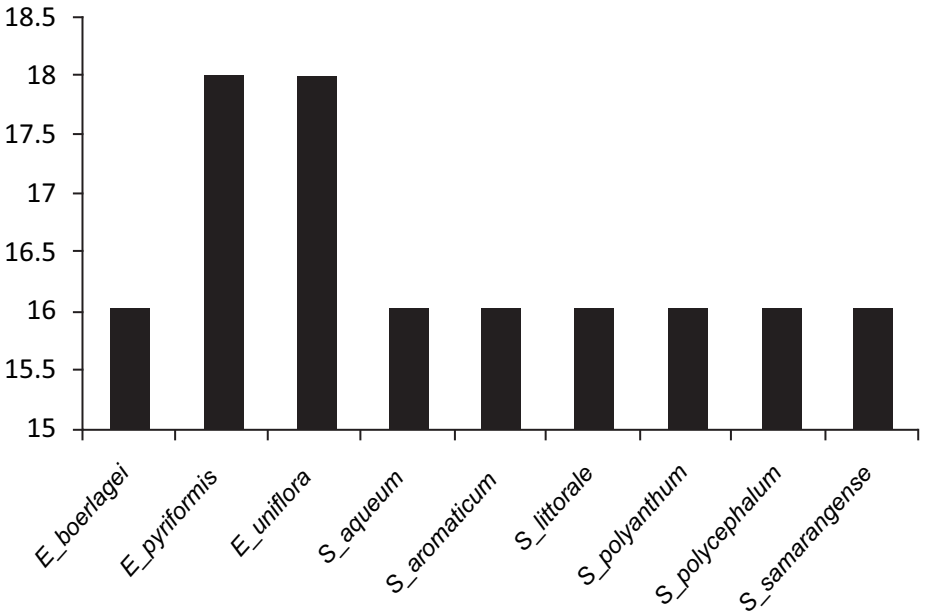

Figure 1 GC content of *atpB-rbcL* spacer sequences; in *Eugenia* and *Syzygium*

Most of the variation was due to insertions, deletions and substitutions in atpB-rbcL IGS (Table 3). When aligned, the sequences have 1060 sites for atpB-rbcL IGS. For the two fragments, there are 77 variable characters with parsimony informative sites for atpB-rbcL. The most parsimonious analysis generated six most parsimonius trees with CI = 0.861842, RI = 0.798077, RCI = 0.687816 for all sites), iCI = 0.771739, iRI = 0.798077, iRCI = 0.615907 (for parsimony informative sites).

The molecular evolution of the chloroplast noncoding region between atpB-rbcL genes in both *Eugenia* and some *Syzygium* showed that most variations amongst *Syzygium* were contributed by insertion and only a few nucleotide substitutions were found. Remarkable findings are as follows: (i) The main characters distinguishing *Eugenia s.s.* from *Syzygium* are the substitutions. *Eugenia s.s.* is characterized by the high number of substitutions namely ca. 33 of 1060 or around 3%. On the other hand, *Syzygium* is characterized by the low number of substitutions where the average is 0.4%.

Based on Unweighted Pair Group Method with Arithmetic Mean (UPGMA) analysis of taxa, the two morphologically distinct taxa, *Eugenia* and *Syzygium* are distantly related and clearly separated. *Syzygium boerlagei* is closely related to *S. aqueum* and *S. samarangense* (Fig. 2). *Eugenia boerlagei* is better placed in *Syzygium*, so the correct name becomes *Syzygium boerlagei*.

Table 3 Insertion, deletion and substitution on DNA sequence of each taxa

No	Taxa	Insertion	Deletion	Substitution
1	*Eugenia uniflora*	24	137	33
2	*Eugenia pyriformis*	20	138	33
3	*Eugenia* or *Syzygium boerlagei*	25	19	8
4	*Syzygium aqueum*	5	131	3
5	*Syzygium aromaticum*	30	18	2
6	*Syzygium littorale*	10	130	5
7	*Syzygium polyanthum*	31	124	1
8	*Syzygium polycephalum*	8	139	6
9	*Syzygium samarangense*	1	144	3

UPGMA Analysis of Taxa

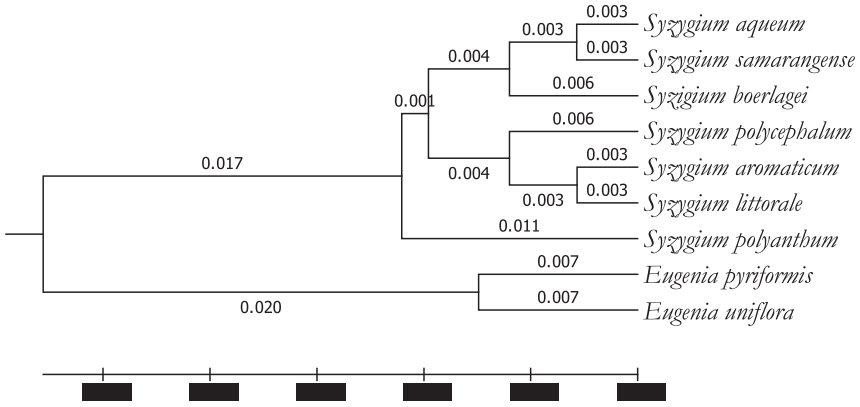

Figure 2 UPGMA tree of *Syzygium* and *Eugenia* based on *atpB-rbcL* intergenic spacer sequence; *Eugenia boerlagei* is nested in *Syzygium*

Based on maximum parsimony (MP) analysis, *Eugenia* and *Syzygium* are distantly related and clearly separated. *Syzygium boerlagei* is closely related to *S. samarangense* (Fig. 3). Hence, *Eugenia boerlagei* is better placed in *Syzygium* becoming *Syzygium boerlagei*.

The evolutionary history was inferred using the Maximum Parsimony (MP) method. Tree #1 out of 4 most parsimonious trees (length = 93) is shown. The consistency index (CI) is 0.946237 (0.903846), the retention index (RI) is 0.918033 (0.918033), and the composite index (CI) is 0.868676 (0.829760) for all sites and parsimony-informative sites (in parentheses). The MP tree was obtained using the Subtree-Pruning-Regrafting (SPR) algorithm (Nei & Kumar 2000) with search level 0 in which the initial trees were obtained by the random addition of sequences (10 replicates). The analysis involved 9 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 886 positions in the final dataset. Evolutionary analyses were conducted in MEGA6 (Tamura *et al.* 2013).

Morphologically, *Eugenia boerlagei* is closer to *Syzygium* than to *Eugenia* s.s. because it is characterized by: 1) shoot sylleptic (not proleptic); 2) leaf bud smooth (not papillous); 3) inflorescence panicle (not solitary and clustered at nodes); and 4) fruits with 1-2 seeds (not many). Either morphologically or molecularly, *E. boerlagei* is very much closer to *Syzygium* than to *Eugenia*. Thus, its transfer to *Syzygium* by Govaerts *et al.* (2008) is acceptable. These are strongly supported by the facts that on one hand, the leaf buds are smooth and not papillose, it has a low number of substitutions (<15) compared to the "real" *Eugenia* which has >30 substitutions in terms of DNA sequences.

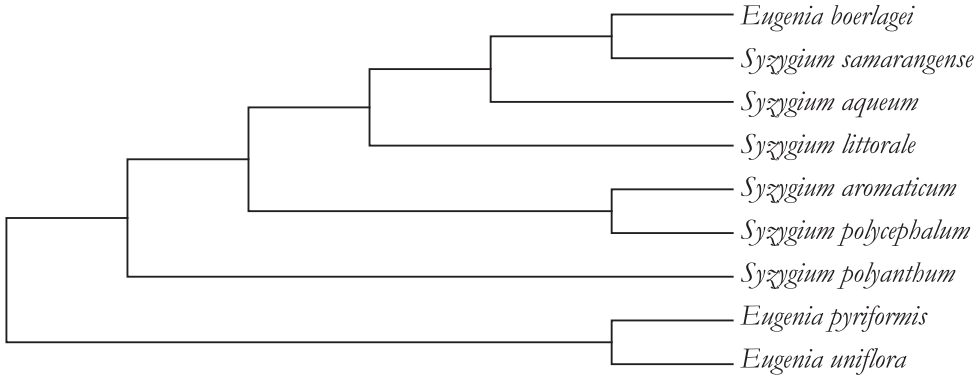

Figure 3 MP tree of *Syzygium* and *Eugenia* based on *atpB-rbcL* intergenic spacer sequence; *Eugenia boerlagei* is nested in *Syzygium* and closely related to *S. samarangense*

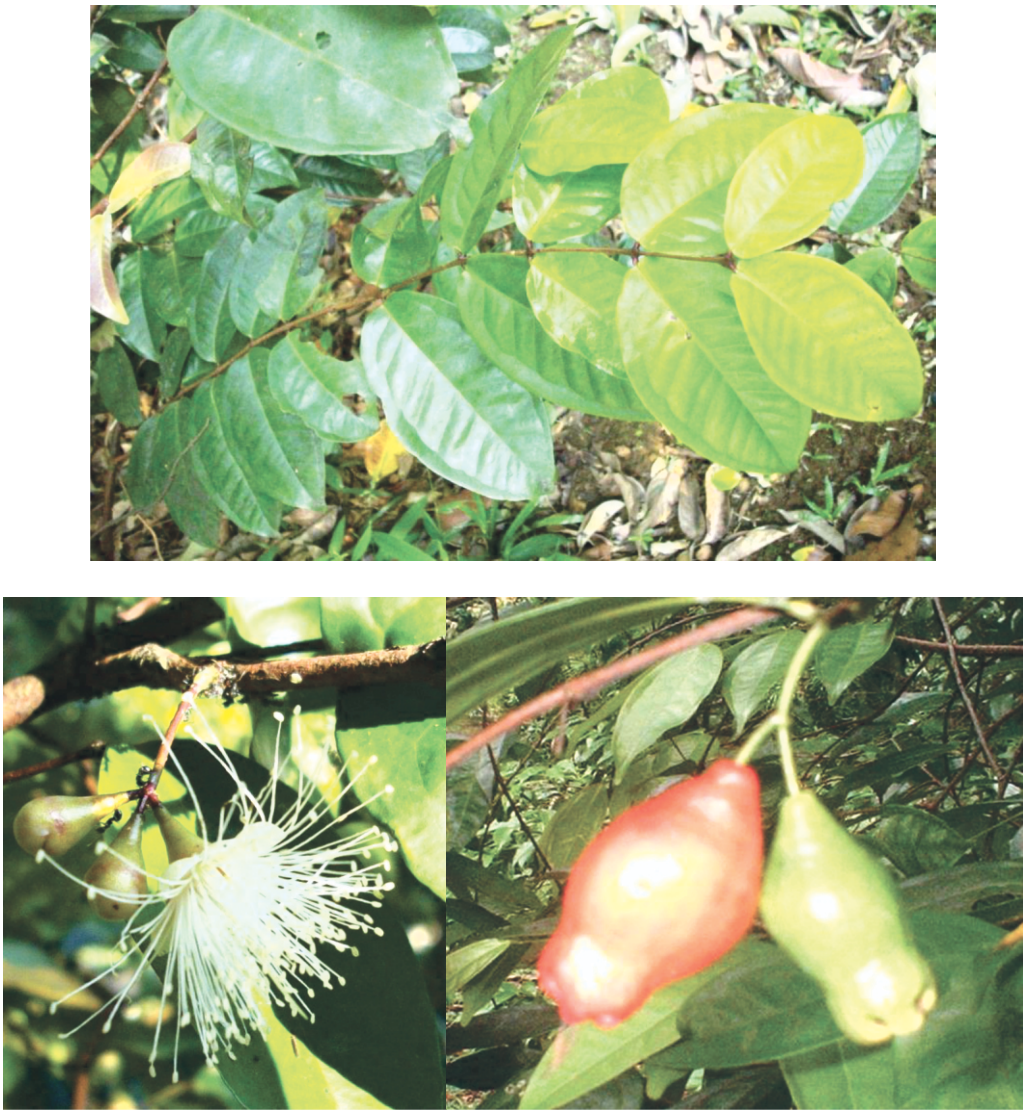

Figure 4 *Syzygium boerlagei* leaves (above), flower (left below) and fruits (right below)

Based on atpB-rbcL data, *S. cumini* and *S. polyanthum* are closely related. Morphologically, both plants have similar bark patterns that are whitish, and close to each other. *S. lineatum* is closely related to *S. cumini* because they have the same number of GC content namely 17. The *Eugenia* group is separated from *Syzygium* because *Eugenia* has much more substitutions or mutations in some sites than *Syzygium*.

CONCLUSION

UPGMA analysis and the maximum parsimony analysis of the two species of *Eugenia s.s.* as the outgroup shows evidence that *Eugenia boerlagei* is nested in *Syzygium*, so it should be transferred to *Syzygium* as was done by Govaerts *et al.* (2008). The two samples of *Eugeni pyriformis* and *Eugenia uniflora* are distantly related to all *Syzygium*.

ACKNOWLEDGEMENTS

The authors wish to thank the following for their roles in the conduct of this study: former head of the Bogor Botanical Garden, Dr Irawati and current head of the Bogor Botanical Garden, Ir Mustaid Siregar for the permission to observe and sample the collections; Mr Yayan Wahyu Chandra Kusuma for providing the materials and assistance; Dr Barbara Gravendeel and Marcel Eurlings of the Van der Klaauw Laboratory Leiden University, The Netherlands for the assistance, chemicals, software, and

facilities; Dr Tim Fulcher of Jodrell Laboratory of the Royal Botanic Gardens, Kew, UK for the assistance and facilities; and Dr J.F. Veldkamp (L) for critically reading the draft.

REFERENCES

Archie JW. 1989. Homoplasy excess ratio: New indices for measuring levels of homoplasy in phylogenetic systematics and a critique of the consistency index. Syst Zool 38:253-69.

Biffin E, Craven LA, Crisp MD, Gadek PA. 2006. Molecular systematics of *Syzygium* and allied genera (Myrtaceae): Evidence from the chloroplast genome. Taxon 55:79-94.

Chiang TY, Schaal BA, Peng CI. 1998. Universal Primers for amplification and sequencing a noncoding spacer between atpB and rbcL genes of chloroplast DNA. Bot Bull Acad Sin 39:245-50.

Chiang TY, Schaal BA. 2000. Molecular evolution of the atpB-rbcL noncoding spacer of chloroplast DNA in the moss family Hylocomiaceae. Bot Bull Acad Sin 41:85-92.

Costion CM, Kress WJ, Crayn DM. 2016. DNA barcodes confirm the taxonomic and conservation status of a species of tree on the brink of extinction in the Pacific. PloS ONE 11(6):e0155118.

Cox AJ, Merrill ED. 1916. Myrtaceae. Philipp J Sci 11:296-7.

Craven LA, Biffin E, Ashton PS. 2006. Acmena, Acmenosperma, Cleistocalyx, Piliocalyx and Waterhousea formally transferred to *Syzygium* (Myrtaceae). Blumea 51:131-42.

Dong W, Liu J, Yu J, Wang L, Zhou S. 2012. Highly variable chloroplast markers for evaluating plant phylogeny at low taxonomic levels and for DNA barcoding. PloS ONE 7(4):e35071.

Doyle JJ, Doyle JL. 1990. The isolation of plant DNA from fresh tissue. Focus 12:13-5.

Effron B. 1982. The jackknife, the bootstrap, and other resampling plans. Conf Board Math Sci Soc Ind Appl Math 38:1-92.

Farris JS. 1989. The retention index and homoplasy excess. Syst Zool 38:406-7.

Felsenstein J. 1985. Confidence limits on phylogenetics: An approach using the bootstrap. Evolution 39:783-91.

Govaerts R, Sobral M, Ashton P, Barrie F, Holst BK, Landrum L, … Lucas E. 2008. World checklist of Myrtaceae. Kew (UK): The Board of Trustees of the Royal Botanic Gardens, Kew.

Hillis DM, Bull JJ. 1993. An empirical test of bootstrapping as a method assessing confidence in phylogenetic analysis. Syst Biol 42:182-92.

Klassen GJ, Moy RD, Locke A. 1991. Consistency indices and random data. Syst Zool 40:446-57.

Kluge AG, Farris JS. 1969. Quantitative phyletics and the evolution of anurans. Syst Zool 18:1-32.

Manen J, Savolainen V, Simon P. 1994. The atpB and rbcL promoters in plastid DNAs of a wide dicot range. J Mol Evol 38:577-82.

Manen J, Natali A. 1995. Comparison of the evolution of ribulose-1, 5-biphosphate carboxylase (rbcL) and atpB-rbcL spacer sequences in a recent plant group, the tribe Rubieae (Rubiaceae). J Mol Biol 41:920-27.

Nei M, Kumar S. 2000. Molecular Evolution and Phylogenetics. New York (US): Oxford University Press.

Shaw J, Lickey EB, Beck JT, Farmer SB, Liu W, Miller J, … Small RL. 2005. The tortoise and the hare II: Relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis 1. Am J Bot 92:142-66.

Swofford DL. 1998. PAUP*: Phylogenetic analysis using parsimony (* and other methods) version 4.08. Massachusetts (US): Sinauer Associates.

Taberlet P, Coissac E, Pompanon F, Gielly L, Miquel C, Valentini A, … Willerslev E. 2007. Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. Nucleic Acid Res 35(3):e14.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30:2725-9.