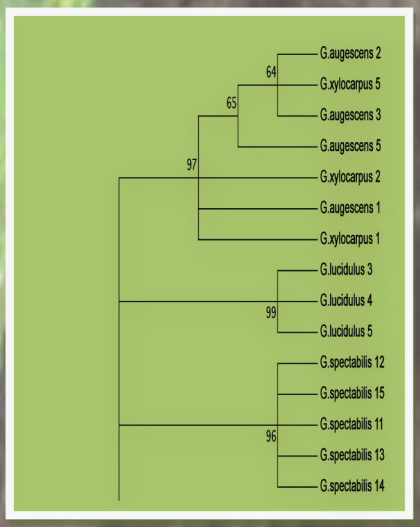
DEVELOPMENT OF DNA SEQUENCE DATABASE OF RAMIN FOR DNA-BASED SPECIES IDENTIFICATION



BIBIAN DIWAY
LAI PEI SING
KEVIN NG KIT SIONG
TNAH LEE HONG
NG CHIN HONG
LEE SOON LEONG
LUCY CHONG

DEVELOPMENT OF DNA SEQUENCE DATABASE OF RAMIN FOR DNA-BASED SPECIES IDENTIFICATION

Bibian Diway
Lai Pei Sing
Kevin Ng Kit Siong
Tnah Lee Hong
Ng Chin Hong
Lee Soon Leong
Lucy Chong

ITTO-CITES Activity: Use of DNA for Identification of *Gonystylus* species and Timber Geographical Origin in Sarawak

First published in August 2014

All enquires should be forwarded to:
Chief Executive Officer
SARAWAK FORESTRY Corporation
Lot 218, KCLD, Jalan Tapang, Kota Sentosa
93250 Kuching, Sarawak, Malaysia
Tel: +6082-610088 Fax: 6082-610099
Website: www.sarawakforestry.com
Email:info@sarawakforestry.com

Development of DNA sequence database of Ramin for DNA- based species identification/Bibian Diway, Lai Pei Sing, Kevin Ng Kit Siong, Tnah Lee Hong, Ng Chin Hong, Lee Soon Leong & Lucy Chong.
ISBN 978- 967 – 12855 – 2 - 7

This project was funded by a grant from ITTO under the collaborative program with CITES.

Photo credits: Rantai Jawa.
Printed in Kuching, Sarawak, Malaysia.

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Abbreviations

DNA	Deoxyribonucleic acid
PCR	Polymerase chain reaction
nrDNA	Nuclear DNA
cpDNA	Chloroplast DNA
$MgCl_2$	Magnesium chloride
dNTPs	Deoxynucleotides
NJ	Neighbor-Joining
K2P	Kimura 2-paramter
BLAST	Basic Local Alignment Search Tool
NCBI	Center for Biotechnology Information

Acknowledgements

This work was made possible by a grant from ITTO under its collaborative program with CITES “Support to ITTO:CITES Implementation for Tree Species and Trade/Market Transparency (TMT)”. Donors to this collaborative program include the EU (primary donor), the USA, Germany, the Netherlands and Norway. The project was implemented by SARAWAK FORESTRY.

We would like to thank the Director of Forests, Sarawak and the Chief Executive Officer of SARAWAK FORESTRY for granting us approval and trust to carry out this project. Many staff members from the Botanical Research Centre, SARAWAK FORESTRY were involved in the fieldwork throughout Sarawak to collect the Ramin samples in particular, Rantai Jawa, Tinjan Kuda, Army Kapi, Paul Sawai, Dino Jengka, Peter Libu and Roy Uning. We are also grateful to the staff of Finance Section in the management and administration of the project fund. We would also like to thank Dr. Lee Chai Ting and Mdm. Nurul Farhana Zakaria from the Forest Research Institute of Malaysia for assisting in the sequencing analysis and Dr. Yayoi Takeuchi from the National Institute for Environmental Studies, Japan for her comments and for editing this report.

Summary

A molecular genetic tool for species identification was developed for *Gonystylus* species. The development of this tool is urgently needed as 15 *Gonystylus* species (including *G. bancanus*) are listed as vulnerable in the IUCN Red List of Threatened Species and all these 15 species are also listed in Appendix II of CITES. The objective of this study is to establish a database for *Gonystylus* species identification using selected chloroplast DNA and nuclear DNA markers. A total of 21 *Gonystylus* species was collected and studied. For sequence data analysis, four regions (trnL, trnH-psbA, trnF-trnE and ITS2) were included. The sequences of these regions were combined and showed 90.48%, 85.71% and 80.95% of resolution in the analysis of the phylogenetic tree, genetic gap distance and BLAST, respectively. Overall, 85.71% of the species can be resolved according to the database generated in this study. The result obtained validated the feasibility of using the molecular genetic tool for species identification in *Gonystylus*.

1.0 Introduction

Gonystylus, locally known as Ramin, is a commercial timber but its habitat has been over-exploited in the South East Asian countries. It has resulted in approximately half of the genus (15 species including *G. bancanus*) being classified as vulnerable in the IUCN Red List of Threatened Species and all these 15 species are also listed in Appendix II of CITES. This is not only due to the extensive logging but also due to the natural slow regeneration of these species. The average annual increment in height is 100 cm and 0.79 cm in diameter (CITES, 2008). Although law enforcement in accordance with CITES regulations has controlled the trade for these species, illegal logging operations and laundering activities still cannot be stopped. The source and identity of the Ramin logs could be falsified along the supply chain. Therefore, a tool which is able to confirm the identity of the timber and overrule the questionable certificate or documentation is highly needed for the trade of legal timber and the development of conservation strategies for Ramin species.

Usually, timber species are identified by wood anatomy expert through light microscopy, but it only goes down to the genus level for Ramin (Gasson, 2011). Moreover, the reference database for species identification is inadequate since not every species in the taxa is well described and wood anatomy varies in different position on the tree and the age of the wood material (Wheeler and Baas, 1998). However, precise identification to species level is needed for the detection of the threatened timber species particularly those listed in Appendix II of CITES. DNA barcodes which are universally amplified, short, and highly variable DNA markers, are believed to

be useful for discrimination among closely related species and yet able to differentiate genetically distant species (Gonzalez *et al*., 2009). In animals *cytochrome c oxidase I* (*CO1*) is the standard marker for DNA barcoding, however, the standard barcoding marker for the vast plant kingdom is still controversial. Several proposals have been made on chloroplast DNA, such as *trnH-psbA* (Kress and Erickson, 2007), *matK* (Sun *et al*., 2012; Lahaye *et al*., 2008), *trnL* (Kress and Erickson, 2007), as well as nuclear DNA, such as *ITS2* (Chen *et al*., 2010), or a combination of several regions.

cpDNA is a highly conserved region, but interspecific introgression can occur in closely related species which cause limitation in the species identification down to species level (Gonzalez *et al.*, 2009). While on the contrary, nrDNA possesses high species discrimination ability, the universal primer for nrDNA is not available (Duminil *et al.*, 2010). As both of these regions exhibit properties needed in species identification, the comparison of these regions may show better insight in this analysis (Madesis *et al.*, 2012). Furthermore, it is claimed that the ITS2 region was capable of discriminating plant taxa at the genus and species level, and was able to discriminate 92.70% of the medicinal plants (Chen *et al.*, 2010). The inclusion of this region for Ramin species identification is beneficial for this study as the discrimination to species level is needed and amplification efficiency of this region was poor in certain families (Liu et al., 2011).

The application of DNA barcoding for timber tracking requires the establishment of a database system which works as a foundation for future diagnostic efforts but also offer quick, cheap, high-volume processing and expressed statistical certainty of results in timber

tracking (Lowe and Cross, 2011). At the same time, it is difficult or impossible to falsify the DNA profile of individual logs (Lowe *et al*., 2010).

The objective of this study is to establish a DNA sequence database for the identification of *Gonystylus* species using selected chloroplast DNA and nuclear DNA markers.

2.0 Materials and Method

2.1 Plant samples

A total of 80 samples comprising 21 species, were collected from 13 locations in Sarawak. Each species was represented by two to five samples according to the availability of the samples (*n* in Appendix A). The samples were first identified to species based on the morphological characters at the Sarawak Herbarium at the Forest Research Cenre, Kuching, Sarawak. Fresh leaf tissue was preserved in silica gel in the field and DNA extraction was carried out right after the samples were brought back to the laboratory. Vouchers specimens for each of the samples are kept at the working herbarium of the Botanical Research Centre (BRC), Kuching, Sarawak.

2.2 DNA extraction and purification

DNA was extracted according to the method as stated in the report "Standard DNA Extraction protocol *for Gonystylus* species", (Diway *et al.,* 2014). All DNA samples were further purified with the High

Pure PCR Template Preparation Kit (Roche, USA) according to the manufacturer's protocol. The DNA was quantified by NanoDrop (Thermo Fisher Scientific Inc., USA) and diluted to 20 ng / µl.

2.3 Polymerase Chain Reaction (PCR)

The extracted DNA underwent PCR amplification using Type-it® Microsatellite PCR kit (Qiagen, Germany). A total of one nuclear (nrDNA) (internal transcribed spacer 2 - ITS2) and three chloroplast (cpDNA) (*trnF-trnE, trnL* and *trnH-psbA*) DNA regions were selected for the species identification (Appendix B). The total PCR reaction volume of 20 µl contained 40 ng of genomic DNA, HotStarTaq®*Plus* DNA Polymerase, Type-it Microsatellite PCR buffer with 6 mM $MgCl_2$ and dNTPs. PCR reactions were carried under the thermal cycling profile as follows: 95˚C for 5 minutes, followed by 28 cycles of denaturation at 95˚C for 30 seconds, annealing at 65˚C for 90 seconds and extension at 72˚C for 60 seconds, with final extension at 60˚C for 30 minutes.

2.4 DNA sequencing

The PCR products were purified using USB®ExoSAP-IT PCR Product cleanup (Affymetrix, USA) and sequenced in both directions using the BigDye®Terminator Sequencing Kit (Applied Biosystems, USA) based on the standard dideoxy-mediated chain termination method. Sequencing reactions were purified using ethanol precipitation and sequences were resolved on an ABI3130xl Genetic

Analyzer (Applied Biosystems, USA). The sequences were assembled and edited using Sequencher® version 5.0 (Gene Codes Corporation, USA) sequence analysis software. The sequence from each sample was trimmed to the same length.

2.5 Data analysis

The informative variable sites, both substitutions and insertions and deletions (indels), which distinguish Ramin species were extracted (Appendix A). The sequences of four regions were combined to form a long sequence and analyzed in three different analyses to verify the resolution of species identification. First, sequences were aligned and a Neighbor-Joining (NJ) tree was constructed using MEGA5 (Tamura *et al*., 2011). The NJ tree was constructed using Kimura 2-parameter (K2P) model with 1000 bootstrap replicates. Then, the pairwise K2P distance was calculated. The highest interspecies value and lowest intraspecies value for each species were plotted using Microsoft Excel. In the final analysis, each individual sequence was blasted to the local database which was created using the sequences generated from this project. This analysis was carried using a stand alone Basic Local Alignment Search Tool (BLAST) from the National Center for Biotechnology Information (NCBI)(www.ncbi.nlm.nih.gov) (United States of America). The blast match which has 100 - 99.80% identity was filtered and applied in the analysis.

3.0 Results and discussion

3.1 DNA sequence variation among species

A total of two nrDNA and eight cpDNA regions were tested in the current study for species identification of Ramin. However, we found only four (one nrDNA and three cpDNA) that were variable among Ramin species. The selected regions consisted of trnL (235-241bp), trnH-psbA (293-405 bp), trnF-trnE (375-424 bp) and ITS2 (333-336 bp). The alignment of the sequences for trnH-psbA required addition of gaps, whereas the others were straightforward. Similar with the present study, trnH-psbA intergenic spacer also showed significant length variation in the study of *Carex* species by Starr *et al*.(2009).

3.2 Neighbor-Joining tree

The sequences of the four regions were combined in the sequence of trnL - trnH-psbA - trnF-trnE - ITS2. The total aligned length for these regions are 1430 bp. According to the generated phylogenetic tree (Figure 1), 19 out of 21 (90.48%) Ramin species that were analyzed have been resolved. The representative individuals for each of these species were clustered in the same clade. However, all of the representative individuals of *G. augescens* and *G. xylocarpus* were grouped in the same cluster, showing that these two species could not to be resolved.

Based on the constructed phylogenetic tree, two major clusters were formed. In a previous Ramin species study by Bogor (2010), *G. brunnescens*, *G. consanguineus* and *G. velutinus* were clustered

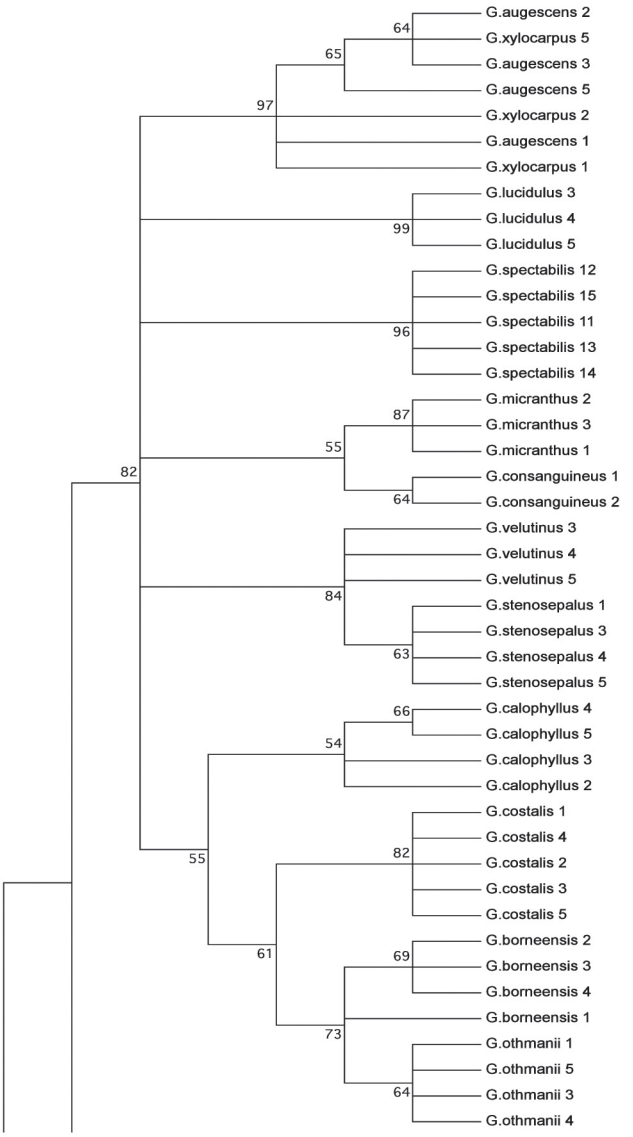


Figure 1: Phylogenetic tree of Ramin constructed using 4 regions (trnL, trnH, trnF and ITS2).

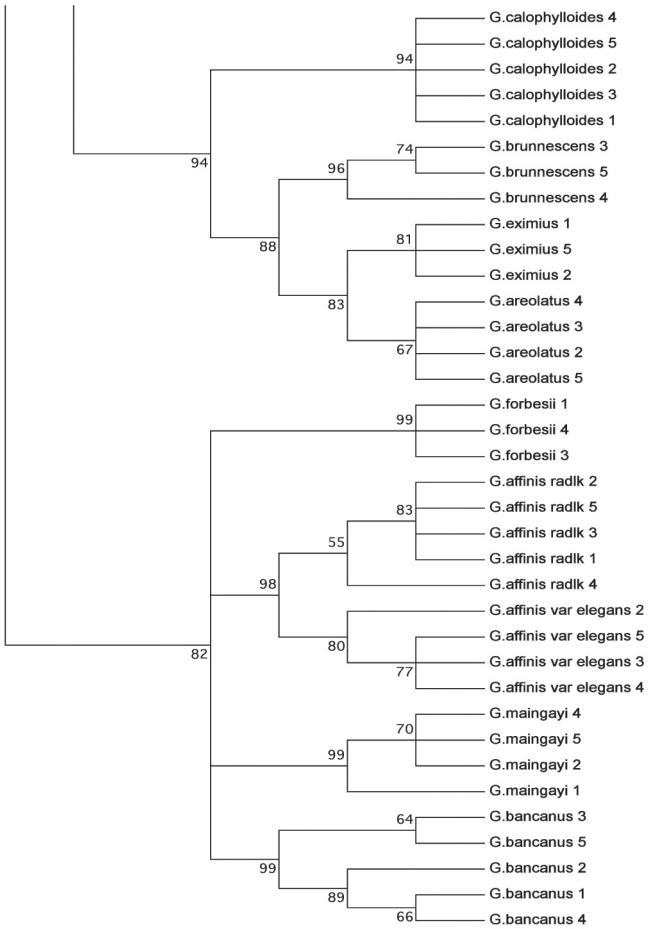


Figure 1: Phylogenetic tree of Ramin constructed using 4 regions (trnL, trnH, trnF and ITS2 2) (continued).

together while *G. bancanus* was in another cluster. Our results are in agreement with his study where *G. bancanus* was separated from the other three species in another major cluster.

3.3 Genetic gap distance

The genetic gap distance was calculated using MEGA 6 and the comparison of inter-species (vertical axis) and intra-species (horizontal axis) gap distance were plotted (Figure 2). A suitable barcode must exhibit high sequence variation between species which can be discriminated from one another and low within species which create a significant gap between intra- and inter-specific genetic variation (Lahaye *et al*., 2008; Meyer and Paulay, 2005). Hence, the species which are plotted above the plotted linear line are defined as resolved, whereas, those on or below the linear line are determined as unresolved. The result generated showed that 85.71% (18/21) of Ramin species, which have higher inter-specific distance compared to intra-specific distance, were resolved. The two unresolved species (intra-specific > inter-specific) are *G. augescens* and *G. xylocarpus*, whereas *G. borneensis* is on the linear line (intraspecific = interspecific).

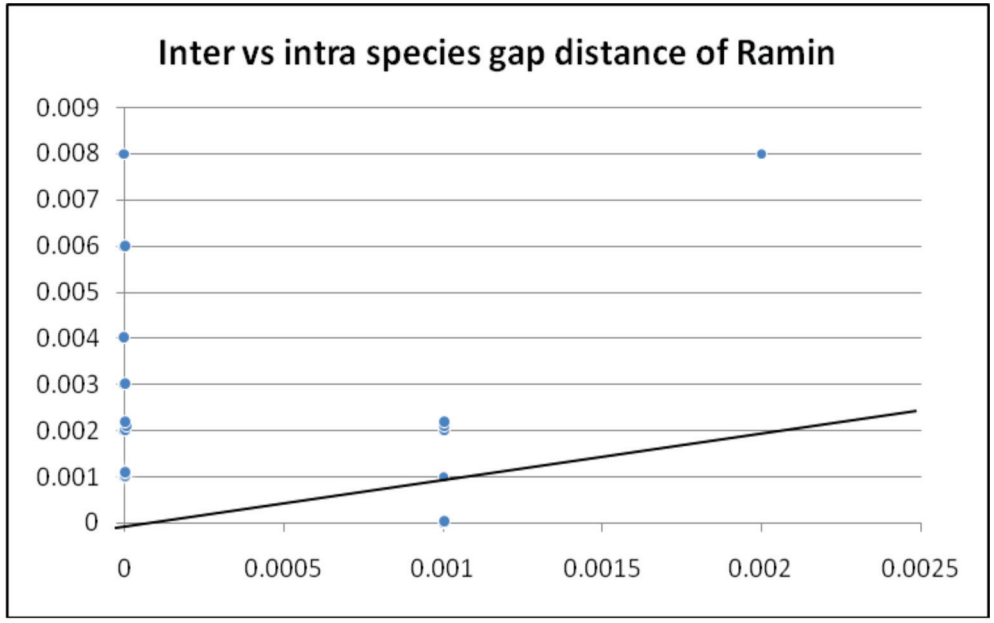


Figure 2: The comparison of inter-species and intra-species gap distance.

3.4 BLAST

The sequences generated were compiled and a local database was constructed. Each of the sequence was then blasted individually to the local database. From the result obtained, the blast match which has 100% to 99.80% identity were considered as resolved species. The result showed that 17 out of 21 (80.95%) Ramin species were correctly assigned back to the same species and the remaining species which were not correctly assigned are *G. augescens*, *G. calophyllus*, *G.velutinus* and *G. xylocarpus*.

3.5 Resolution of DNA-based species identification in Ramin

Overall, by combining the percentage of resolution from each analysis, approximately 85.71% of Ramin species can be resolved (Table 1). This percentages of resolution is comparable to the results obtained by Burgess *et al*.(2011), Seberg and Petersen (2009), Lahaye *et al*. (2008) and Kress and Erickson (2007) where the percentage of resolution were approximately 90% or greater. The two species which could not be resolved in the three analyses were *G. augescens* and *G. xylocarpus*. *G. bancanus* was resolved in all three analyses and this would make this database valuable for the timber tracking of this threatened species.

Table 1: Average resolution percentage of Ramin species resolved.

Analysis	Resolution (%)
Phylogenetic tree	90.48
Genetic gaps distance	85.71
BLAST	80.95
Average	85.71

3.6 Variable sites for species identification

Notwithstanding the most accurate way to identify the species is by blasting the whole generated sequence with the database, the list of the informative variable sites (Appendix A) provided a clue on the sequence variations for each species. There are two informative variable sites which distinguish Ramin species found in the trnL region; seven in the trnF-trnE region; three in the trnH-psbA; and 17 in ITS2 region. There are more variable sites found in the ITS2 because this region possesses high interspecific divergence (Chen *et al*., 2010). Most of the species can be distinguished by referring to one variable site of the regions, whereas *G. areolatus* requires six variable sites to be distinguished from the other species. Moreover, *G.augescens* and *G xylocarpus* were not resolved in the three analyses mentioned above as they appeared genetically identical, and thus, these species were not distinguishable in this study. The two species were either genetically close to each other or errors could have occurred in the species identification using morphological characteristics. If the morphological identification

errors were ruled out, more or different cpDNA regions would be required to further resolve the identification of these two species.

Additionally, *G. bancanus*, which is listed as vulnerable in The IUCN Red List of Threatened Species and Appendix II of the CITES, was found to contain several unique sites in the ITS2 region. These specific sites are leads to identify this species genetically. Hence, unknown samples or samples which are suspected to be *G. bancanus* could be screened using this database to solve queries on their identities.

4.0 Conclusion

A total of 17 Ramin species, including *G. bancanus*, were genetically distinguishable using the four variable regions in this study using the DNA. A DNA database for Ramin species was established and the capability of molecular genetics in the identification of Ramin species in Sarawak was proven. However, more Ramin species and samples from other locations in Sarawak should be analyzed to build a more comprehensive and complete database which is necessary for creating an effective and validated forensic identification tool for Ramin species identification.

References

Bogor. (2010). Genetic relationship between species of *Gonystylus* spp.. Indonesia: AYPBC Widyatmoko.

Burgess, K.S., Fazekas, A.J., Kesanakurti, P.R., Graham, S.W., Husband, B.C., Newmaster, S.G., Percy, D.M., Hajibabaei, M. and Barrett, S.C.H. (2011). Discriminating plant species in a local temperate flora using the rbcL + matK DNA barcode.*Methods in Ecology and Evolution*, **2**, 333-340. DOI: 10.111/j.2041-210X.2011.00092.x.

Chen, S., Yao, H., Han, J., Liu, C., Song, J., Shi, L., Zhu, Y., Ma, X., Gao, T., Pang, X., Luo, K., Li, Y., Li, X., Jia, X., Lin, Y. and Leon, C. (2010). Validation of the ITS2 region as a novel DNA barcode for identifying medicinal planr species.*PLoS ONE*, **5(1)**, e8613. DOI: 10.1371/journal.pone.0008613.

CITES. (2008). Summary report on the non-detriment findings for Ramin (*Gonystylus* spp.) for Malaysia in 2008.Seventeenth meeting of the Plant Committee Geneva (Switzerland), 15-19 April 2008.

Diway, B., Lai, P.S. and Chong L. (2014) Standard DNA Extraction Protocol for *Gonystylus* species. In press.

Duminil, J., Heuertz. M., Doucet, J.L., Bourland, N., Cruaud, C., Gavory, F., Doumenge, C., Navascues, M. and Hardy, O.J. (2010). CpDNA-based species identification and phylogeography: application to African tropical tree species. Molecular Biology, **19**, 5469-5483. DOI: 10.1111/j.1365-294X.2010.04917.x.

Gasson, P. (2011). How precise can wood identification be? Wood anatomy's role in support of the legal timber trade, especially CITES. IAWA Journal, **32(2),** 137-154.

Gonzalez, M.A., Baraloto, C., Engel, J., Mori, S.A., Pétronelli, P., Riéra, B., Roger, A., Thénaud, C. and Chave, J. (2009). Identification of Amazonian trees with DNA barcodes.*PLoS ONE*, **4(10)**, e7438. DOI: 10.1371/journal.pone.0007483.

Kress, W.J. and Erikson, D.L. (2007). A two-locus global DNA barcode for land plants: the coding rbcL gene complements the non-coding trnH-psbA spacer region. PLoS ONE, **2(6)**, e508. DOI: 10.1371/journal.pone.0000508.

Lahaye, R., van der Bank, M., Bogarin, D., Warner, J., Pupulin, F., Gigot, G., Maurin, O., Duthoit, S., Barraclough, T.G. and Savolainen, V. (2008). DNA barcoding the floras of biodiversity hotspots.*Proceedings of the National Academy of Sciences of the United States of America*, **105(8)**, 2923-2928. DOI: 10.1073/pnas.0709936105.

Liu, Z., Chen, S.L., Song, J.Y., Zhang, S.J. and Chen, K.L. (2011). Application of deoxyribonucleic acid barcoding in lauraceae plants.*Pharmacognosy Magazine*, **8(29)**, 4-11. DOI: 10.4103/0973-1296.93301.

Lowe, A.J. and Cross, H.B. (2011). The application of DNA methods to timber tracking and origin verification.*IAWA Journal*,**32(2),** 251-262.

Lowe, A.J., Wong, K.N., Tiong, Y.S., Iyerh, S. and Chew, F.T. (2010). A DNA method to verify the integrity of timber supply chains; confirming the legal sourcing of Merbau timber from logging concession to sawmill.SilvaeGenetica, **59**, 263-268.

Madesis, P., Ganopoulos, I., Ralli, P. and Tsaftaris, A. (2012). Barcoding the major Mediterranean leguminous crops by combining universal chloroplast and nuclear DNA sequence targets. *Genetic and Molecular Research*, **11(3)**, 2548-2558. DOI: 10.4238/2012.July.10.10.

Meyer, C.P. and Paulay, G. (2005). DNA barcoding: error rates based on comprehensive sampling. *PLoS Biology*, **3(12)**, e422. DOI: 10.1371/journal.pbio.0030422.

Seberg, O. and Petersen, G. (2009). How many loci does it take to DNA barcode a crocus? *PLos ONE*, **4(2)**, e4598. DOI: 10.1371/journal.pone.0004598.

Starr, J.R., Naczi, R.F.C. and Chouinard, B.N. (2009). Plant DNA barcodes and species resolution in sedges (*Carex*, *Cyperaceae*).*Molecular Ecology Resources*, **9(1)**, 151-163. DOI: 10.1111/j.1755-0998.2009.02640.x.

Sun, X.Q., Zhu, Y.J., Guo, J.L., Peng, P., Bai, M.M. and Hang, Y.Y. (2012). DNA barcoding the *Dioscorea* in China, a vital group in the evolution of monocotyledon: use of *matK* gene for species discrimination. *PLoS ONE*, **7(2)**, e32057. DOI: 10.1371/journal.pone.0032057.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., & Kumar, S. (2011). MEGA5: Molecular Evolutionary Genetics Analysis using maximum likelihood, evolutionary distance, and, maximum parsimony methods. *Molecular Biology and Evolution*, **28(10)**, 2731-2739. DOI: 10.1093/molbev/msr121.

Wheeler, E.A. and Baas, P. (1998). Wood identification – a review.*IAWA Journal*, **19(3)**, 241-264.

Appendix A

DNA database for Ramin species from trnL, trnF-trnE, trnH-psbA and ITS2.

Position	n	trnL		trnF-trnE							trnH-psbA		
		82	184	111-118	119-133	134-149	158	343-347	354	406	374	390	434
G. affinisradlk	5	A	A	TTAAAATA	:	AAAGTAAAAAAAAGTA	G	:	A	A	G	C	G
G. areolatus	4	A	A	TTAATATT	AAAATA	AAAGTCAAAAAAAGTA	G	:	A	A	G	C	G
G. augescens	4	A	A	TTAAAATA	:	AAAGTCAAAAAAAGTA	G	:	A	A	G	C	T
G. bancanus	5	A	A	TTAAAATA	:	AAAGTAAAAAAAAGTA	G	:	A	A	G	C	G
G. borneensis	4	A	A	TTAAAATA	:	AAAGTAAAAAAAAGTA	G	:	A	A	G	C	T
G. brunnescens	3	A	A	TTAATATT	AAAATA	AAAGTCAAAAAAAGTA	G	:	C	A	G	C	G
G. calophyllus	4	A	A	:	:	:	G	:	A	A	G	C	T
G. consanguinues	2	A	A	TTAAAATA	:	AAAGTAAAAAAAAGTA	G	:	A	A	G	C	T
G. forbesii	3	A	A	TTAAAATA	:	AAAGTAAAAAAAAGTA	G	:	A	A	G	C	G
G. lucidulus	3	A	A	TTAAAATA	AAAGAATA	AAAGTAAAAAAAAGTA	G	:	A	C	G	T	T
G. micranthus	3	A	A	TTAAAATA	:	AAAGTAAAAAAAAGTA	G	:	A	A	G	C	T
G. othmanii	4	A	A	TTAAAATA	:	AAAGTAAAAACAAGTA	G	:	A	A	G	C	T
G. velutinus	3	A	A	TTAAAATA	AAATAAAAAAAAAAA	AAAGTAAAAAAAAGTA	G	:	A	A	G	C	T
G. xylocarpus	3	A	A	TTAAAATA	:	AAAGTCAAAAAAAGTA	G	:	A	A	G	C	T
G. calophylloides	5	A	A	TTAATATT	AAAATA	AAAGTCAAAAAAAGTA	G	:	A	A	G	C	G
G. affinisvarelegans	4	A	A	TTAAAATA	:	AAAGTAAAAAAAAGTA	G	:	A	A	G	C	G
G. stenosepalus	4	A	A	TTAAAATA	:	AAAGTAAAAAAAAGTA	G	AAAGA	A	A	A	C	T
G. eximius	3	A	A	TTAATATT	AAAATA	AAAGTCAAAAAAAGTA	G	:	A	A	G	C	G
G. maingayi	4	C	A	TTAAAATA	:	AAAGTAAAAAAAAGTA	G	:	A	A	G	C	G
G. costalis	5	A	A	TTAAAAAA	:	AAAGTAAAAAAAAGTA	G	:	A	A	G	C	T
G. spectabilis	5	A	C	TTAAAATA	:	AAAGTAAAAAAAAGTA	A	:	A	A	G	C	T

DNA database for Ramin species from trnL, trnF-trnE, trnH-psbA and ITS2. (continued)

Position	ITS2																
	58	67	68-73	89-96	100-103	114	122	142	171-172	191-192	197	203	230	257-259	265-266	270-275	280
G. affinisradlk	A	T	ACC	GGGGG	ACGA	G	A	:	TA	TC	A	:	G	TAA	AC	TATCTA	T
G. areolatus	A	T	ACC	GGGGG	ATGA	G	A	:	TA	TC	A	:	G	TAA	GC	TATCTA	T
G. augescens	A	T	ACC	GGGGG	GTGA	C	A	:	TA	TC	A	:	T	TAA	GC	TATCTA	T
G. bancanus	A	T	ACC	AGGGG	ACGA	G	A	:	A	TT	T	:	G	TAA	GC	CATCTG	T
G. borneensis	A	T	ACC	GGGGG	ATGA	G	A	G	TA	TC	A	:	G	TAA	GT	TATATA	T
G. brunnescens	A	T	ACC	GGGGG	ATGA	A	A	:	TG	TC	A	:	G	TAA	GC	TATCTA	T
G. calophyllus	A	T	GCCACC	GGGGG	ACGA	G	A	:	TA	TC	A	:	G	TAA	GT	TATCTA	T
G. consanguinues	A	T	ACC	GGGGG	ATGA	G	A	:	TA	TC	A	T	G	TAA	GC	TGTCTA	T
G. forbesii	A	T	GCC	GGGGG	ACGA	G	A	:	TA	TC	A	:	G	TAA	GC	TATCTA	T
G. lucidulus	A	A	ACC	GGGGGGGG	ATGA	A	A	:	TA	TC	A	:	G	CAA	GC	TATCTA	T
G. micranthus	G	T	ACC	GGGGG	ATGA	G	A	:	TA	TC	A	:	G	TAA	GC	TATCTA	T
G. othmanii	A	T	ACC	GGGGG	ATGA	G	A	G	TA	TC	A	:	G	TAA	GT	TATATA	T
G. velutinus	A	T	ACC	GGGGG	ATGA	G	A	:	TG	TC	A	:	G	TAA	GC	TATCTA	T
G. xylocarpus	A	T	ACC	GGGGG	GTGA	C	A	:	TA	TC	A	:	T	TAA	GC	TATCTA	T
G. calophylloides	A	T	ACC	GGGGGG	ATGA	G	A	G	TA	TC	A	:	G	TAA	GT	TATCTA	T
G. affinisvarelegans	A	T	ACC	GGGGG	ACGG	G	A	:	TA	TC	A	:	G	TAG	GC	TATCTA	T
G. stenosepalus	A	T	ACC	GGGGG	ATGA	G	A	:	TG	TC	A	:	G	TAA	GC	TATCTA	T
G. eximius	A	T	ACC	GGGGG	ATGA	G	A	:	TA	CC	A	:	G	TAA	GC	TATCTA	T
G. maingayi	A	T	ACC	GGGGG	ATGA	A	G	:	TA	TC	A	:	G	TAA	GC	TATSTA	A
G. costalis	A	T	ACC	GGGGG	ATGA	G	A	G	TA	TC	A	:	G	TAA	GT	TATCTA	T
G. spectabilis	A	T	ACC	GGGGG	ATGA	G	A	:	TA	TC	A	:	G	TAA	GC	TATCTA	T

Appendix B

CpDNA and nrDNA Primers Information

Primer	Sequence	Fragment size (bp)
trnF-trnE	trnF (F): 5'-ATTTGAACTGGTGACACGAG-3' trnE (R): 5'-GGTTCAAGTCCCTCTATCCC-3'	449
trnL	trnL (F): 5'-CGAAATCGGTAGACGCTACG-3' trnL (R): 5'-GGGGATAGAGGGACTTGAAC-3'	243
ITS 2	ITS2 (F): 5'-ATGCGATACTTGGTGTGAAT-3' ITS2 (R): 5'-GACGCTTCTCCAGACTACAAT-3'	341
trnH-psbA	trnH (F): 5’-CGCGCATGGTGGATTCACAATCC-3’ psbA (R): 5’-GTTATGCATGAACGTAATGCTC-3’	445

ISBN 978-967-12855-2-7