

**PHYLOGENETICS OF BACTROCERA FRUIT FLIES
IN MALAYSIA WITH REFERENCE TO
SELECTED HOST FRUITS**

JOEY EE ULI

**DISSERTATION SUBMITTED IN FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE**

**INSTITUTE OF BIOLOGICAL SCIENCES
FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR**

2013

ABSTRACT

Bactrocera carambolae and *Bactrocera papayae* fruit flies, also known as the carambola and papaya fruit fly respectively, are significant pest fruit flies in Malaysia. They are capable of causing losses in the agricultural industry by infesting host fruits and making them unable to be sold. These two species are grouped in the *Bactrocera dorsalis* species complex, and are oftentimes difficult to distinguish from one another due to similar and intermediate morphological characteristics. A precise method for identifying pest fruit flies is important to properly monitor the infestation of host fruits, for the purpose of quarantine management. The aims of this research are to determine the effect of locations and host fruits in the phylogenetics of *Bactrocera carambolae* and *Bactrocera papayae*, to determine the phylogenetic relationships between *Bactrocera carambolae* and *Bactrocera papayae*, and to determine the ability of the selected molecular markers (COI, COII, and *cytb*) in distinguishing *Bactrocera carambolae* and *Bactrocera papayae*. To determine the taxonomic position between *Bactrocera carambolae* and *Bactrocera papayae*, three molecular markers, COI, COII, and *cytb*, were utilized. Infested host fruits were collected from two locations in Peninsular and East Malaysia (Serdang and Sarawak), and fruit fly specimens were hatched and identified based on morphological characteristics. Molecular phylogenetic analyses using maximum likelihood, Bayesian Inference, Neighbor-Joining, and haplotype network reconstruction based on COI, COII, *cytb*, and the combination of the three molecular markers, were not able to differentiate *Bactrocera carambolae* and *Bactrocera papayae* as two distinct species as they tend to group together within the same clade. *Bactrocera carambolae* and *Bactrocera papayae* specimens also tend to group together within the same clade as other members of the *Bactrocera dorsalis* complex. This suggests that *Bactrocera carambolae* and *Bactrocera papayae* could

possibly belong to the same species. The fruit flies hatched from different host fruits collected from Serdang could not be distinguished using the three mitochondrial DNA markers, however, the fruit flies hatched from host fruits collected from Sarawak tend to group separately from the Serdang specimens. This suggests that the fruit flies collected from Sarawak were genetically different from the fruit flies collected from Serdang.

ABSTRAK

Lalat buah *Bactrocera carambolae* dan *Bactrocera papayae*, lebih dikenali masing-masing sebagai lalat buah carambola dan lalat buah betik, adalah lalat buah perosak yang penting di Malaysia. Lalat-lalat buah ini berkeupayaan untuk menyerang dan merosakkan buah-buahan perumahannya. Akibatnya, buah-buahan ini tidak dapat dijual dan industri pertanian akan menghadapi kerugian dalam jualan dan eksport buah-buahan. Kedua-dua spesies lalat buah ini adalah diklasifikasikan di dalam kompleks spesies *Bactrocera dorsalis*, dan lazimnya sukar untuk dibezakan antara satu sama lain disebabkan oleh sifat-sifat morfologi yang mirip dan berada di perantaraan. Kaedah yang jitu untuk mengenalpasti identiti spesies lalat buah perosak adalah amat penting untuk pemantauan serangan perosak terhadap buah-buahan perumahannya, dan juga untuk pengurusan kuarantin buah-buahan yang telah diserang oleh perosak. Objektif-objektif kajian ini adalah untuk mengenalpasti pengaruh lokasi dan buah perumah terhadap filogenetik *Bactrocera carambolae* dan *Bactrocera papayae*, untuk mengenalpasti hubungan filogenetik di antara *Bactrocera carambolae* dan *Bactrocera papayae*, dan untuk mengenalpasti keberkesanan penanda molekular yang dipilih (COI, COII, dan *cytb*) dalam membezakan antara *Bactrocera carambolae* dan *Bactrocera papayae*. Bagi mengenalpasti kedudukan taksonomi antara *Bactrocera carambolae* dan *Bactrocera papayae*, tiga penanda molekular, COI, COII, dan *cytb* digunakan. Buah-buahan yang telah diserang dikumpulkan daripada dua lokasi di Semenanjung Malaysia dan Malaysia Timur (Serdang dan Sarawak). Setelah lalat buah menetas daripada buah-buahan perumahannya, spesimen-spesimen dikenalpasti spesiesnya berdasarkan sifat-sifat morfologinya. Analisa filogenetik molekular dengan kaedah *maximum likelihood*, *Bayesian Inference*, *Neighbor-Joining*, dan pembinaan semula rangkaian *haplotype* berdasarkan COI, COII, *cytb*, dan kombinasi ketiga-tiga penanda molekular COI, COII,

dan *cytb* tidak dapat membezakan *Bactrocera carambolae* dan *Bactrocera papayae* sebagai two spesies yang mutlak. Kedua-dua spesies ini berkecenderungan untuk berkumpul dalam klad yang sama. Spesimen-spesimen *Bactrocera carambolae* dan *Bactrocera papayae* juga berkecenderungan untuk berkumpul dalam klad yang sama dengan ahli-ahli lain kompleks spesies *Bactrocera dorsalis*. Ini mencadangkan bahawa *Bactrocera carambolae* dan *Bactrocera papayae* berkemungkinan tergolong dalam spesies yang sama. Lalat-lalat buah yang menetas daripada buah-buahan perumah yang dikumpul daripada Serdang tidak dapat dibezakan dengan menggunakan ketiga-tiga penanda DNA mitokondria, bagaimanapun, lalat-lalat buah yang menetas daripada buah-buahan perumah yang dikumpul daripada Sarawak berkecenderungan untuk berkumpul secara berasingan daripada spesimen-spesimen Serdang. Ini mencadangkan bahawa lalat-lalat buah yang berasal daripada Sarawak adalah berlainan secara genetik dengan lalat-lalat buah yang berasal daripada Serdang.

ACKNOWLEDGEMENTS

First and foremost, I would like to extend my sincerest and most heartfelt gratitude to my supervisor, Professor Dato' Dr. Mohd. Sofian Azirun for giving me the opportunity to pursue this project as part of my Master's degree. Your guidance and initiative have greatly assisted me in the completion of this dissertation. I would also like to express my deepest thanks to my co-supervisor, Dr. Lim Phaik-Eem, for your relentless counsel and feedback, which have helped me learn and understand this scientific field a lot more than when I started out. My sincerest gratitude to University of Malaya for providing the SBUM scholarship, as well as UPGP for the funding of my research (PS259/2009C).

I want to extend my thanks to Professor Emeritus Yong Hoi Sen for his sage advice in regards to many matters. Your wise words have helped me push forward with my project. My deepest gratitude to Encik Fadzlee Hamzah for all the pep talk, and to my lab-mate Daicus Belabut for all the tea and titbits. A million thanks to Tan Ji, for all the technical advice in regards to my project. Not to forget, friends from my cohort, Gary Sing, Yu Chew Hock, Karen Chia, and Lucas Low. Thank you for your friendship and for being there when I needed some assistance. I'd also like to thank my buddies from UPM, Chong Chou Min, Chew Weiyun, Ooi Yue Shin, Bartholomew Tu, and Ang Siok Yee for keeping me inspired and encouraging me to complete my dissertation.

Last but not least, I'd like to give my deepest and sincerest thanks to my parents. Without their self-less help, their diamond advice, their constant nagging, their endless prayers, and utmost support, I couldn't possibly imagine myself ploughing through the daily grind, trials, and tribulations in order to complete my dissertation.

TABLE OF CONTENTS

	Page
ABSTRACT	ii
ABSTRAK	iv
ACKNOWLEDGEMENTS	vi
TABLE OF CONTENTS	vii
LIST OF FIGURES	xi
LIST OF TABLES	xv
LIST OF SYMBOLS AND ABBREVIATIONS	xvi
CHAPTER 1 INTRODUCTION	1
CHAPTER 2 LITERATURE REVIEW	7
2.1 Taxonomic Hierarchy	7
2.2 Dacinae fruit flies	8
2.3 <i>Bactrocera dorsalis</i> species complex	9
2.4 Economic and Agricultural Importance	10
2.5 Life Cycle	12
2.6 Species of <i>Bactrocera carambolae</i> and <i>Bactrocera papayae</i>	14
2.7 Problems with Distinguishing <i>B. carambolae</i> and <i>B. papayae</i>	17
2.8 Molecular Markers	17
2.9 Mitochondrial DNA (mtDNA)	19
2.10 Molecular Phylogenetic Studies on <i>Bactrocera carambolae</i> and <i>Bactrocera papayae</i>	21
2.11 Polymerase Chain Reaction	22
2.12 Electrophoresis	23

2.13 DNA Sequencing and Phylogenetics	24
2.14 Sequence Alignment	25
2.15 Maximum Likelihood	26
2.16 Bayesian Inference	27
2.17 Neighbor-Joining	28
2.18 Haplotype Network	28
CHAPTER 3 MATERIALS AND METHODS	30
3.1 Host Fruit Collection	30
3.2 Fruit Fly Hatching, Storage and Identification	31
3.3 DNA Extraction	31
3.3.1 i-Genomic CTB DNA Extraction Mini Kit	31
3.4 PCR Amplification	33
3.4.1 PCR Amplification of Cytochrome oxidase subunit I	34
3.4.2 PCR Amplification of Cytochrome oxidase subunit II	34
3.4.3 PCR Amplification of Cytochrome-b	35
3.5 Electrophoresis	36
3.5.1 Preparation of Agarose Gel	36
3.5.2 Electrophoresis of PCR Products	36
3.6 DNA Purification and Sequencing	37
3.7 Phylogenetic Analysis	37
3.8 GenBank DNA Sequences	38
3.9 Haplotype Network Reconstruction	40

CHAPTER 4	RESULTS	41
4.1	<i>Bactrocera carambolae</i> and <i>Bactrocera papayae</i> Specimens Collected	41
4.2	Polymerase Chain Reaction	42
4.2.1	Primer Temperature Gradient Screening	42
4.2.2	Cytochrome Oxidase I Based Primer	44
4.2.3	Cytochrome Oxidase II Based Primer	47
4.2.4	Cytochrome-b Based Primer	50
4.3	Sequence Alignment	54
4.4	Phylogeny Based On Cytochrome Oxidase Subunit I Gene	55
4.4.1	Maximum-Likelihood	55
4.4.2	Bayesian Inference	57
4.4.3	Neighbor-Joining	59
4.5	Phylogeny Based On Cytochrome Oxidase Subunit II Gene	61
4.5.1	Maximum-Likelihood	61
4.5.2	Bayesian Inference	63
4.5.3	Neighbor-Joining	65
4.6	Phylogeny Based On Cytochrome-b Gene	67
4.6.1	Maximum-Likelihood	67
4.6.2	Bayesian Inference	69
4.6.3	Neighbor-Joining	71
4.7	Phylogeny Based On Combined COI, COII, and <i>cytb</i> Genes	73
4.7.1	Maximum-Likelihood	73
4.7.2	Bayesian Inference	75

4.7.3 Neighbor-Joining	77
4.8 Haplotype Network Reconstruction	79
4.8.1. Haplotype Network Reconstruction Based On Cytochrome Oxidase Subunit I Gene	81
4.8.2 Haplotype Network Reconstruction Based On Cytochrome Oxidase Subunit II Gene	83
4.8.3 Haplotype Network Reconstruction Based On Cytochrome-b Gene	85
4.8.4 Haplotype Network Reconstruction Based On Combined COI, COII, and <i>cytb</i> Genes	88
CHAPTER 5 DISCUSSION	91
5.1 Species Identification	91
5.2 DNA Extraction	92
5.3 PCR Primers	93
5.4 PCR Optimization	93
5.5 PCR Reaction Mixture	95
5.6 Phylogeny Based On Cytochrome Oxidase Subunit I Gene	96
5.7 Phylogeny Based On Cytochrome Oxidase Subunit II Gene	98
5.8 Phylogeny Based On Cytochrome-b Gene	99
5.9 Phylogeny Based On Combined COI, COII, and <i>cytb</i> Genes	100
5.10 Haplotype Network Reconstruction	101
5.11 Overview	102
CHAPTER 6 CONCLUSION	106
REFERENCES	109

LIST OF FIGURES

		Page
Figure 2.1	Morphology of <i>Bactrocera carambolae</i>	10
Figure 2.2	The life cycle of a fruit fly	13
Figure 2.3	Morphological features of <i>Bactrocera carambolae</i>	15
Figure 2.4	Morphological features of <i>Bactrocera papayae</i>	16
Figure 4.1	COI temperature gradient optimization. Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 1-6: products of amplification at a temperature gradient of 48-53°C. Lane 3 at 50°C was selected for amplification.	42
Figure 4.2	COII temperature gradient optimization. Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 1-6: products of amplification at a temperature gradient of 41-46°C. Lane 4 at 44°C was selected for amplification.	43
Figure 4.3	<i>Cytb</i> temperature gradient optimization. Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 1-6: products of amplification at a temperature gradient of 43-48°C. Lane 3 at 45°C was selected for amplification.	43
Figure 4.4	COI banding profile of fruit fly samples from Serdang, collected from papaya host fruit. Amplification of COI fragment was expected to be around 690 base pairs in length. Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 1-4: <i>Bactrocera carambolae</i> samples. Lanes 5-8: <i>Bactrocera papayae</i> samples.	44
Figure 4.5	COI banding profile of fruit fly samples from Serdang, collected from guava host fruit. Amplification of COI fragment was expected to be around 690 base pairs in length. Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 1-4: <i>Bactrocera carambolae</i> samples. Lanes 5-8: <i>Bactrocera papayae</i> samples.	45
Figure 4.6	COI banding profile of fruit fly samples from Serdang, collected from jambu madu host fruit. Amplification of COI fragment was expected to be around 690 base pairs in length. Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 1-4: <i>Bactrocera carambolae</i> samples. Lanes 5-8: <i>Bactrocera papayae</i> samples.	45
Figure 4.7	COI banding profile of fruit fly samples from Serdang, collected from jambu air host fruit. Amplification of COI fragment was expected to be around 690 base pairs in length. Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 1-4: <i>Bactrocera carambolae</i> samples. Lanes 5-8: <i>Bactrocera papayae</i> samples.	46
Figure 4.8	COI banding profile of fruit fly samples from Sarawak, collected from starfruit host fruit. Amplification of COI fragment was expected to be around 690 base pairs in length.	46

	Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 1-4: <i>Bactrocera carambolae</i> samples. Lanes 5-6: Null amplifications. Lanes 7-8: <i>Bactrocera papayae</i> samples.	
Figure 4.9	COI banding profile of fruit fly samples from Sarawak, collected from starfruit host fruit. Amplification of COI fragment was expected to be around 690 base pairs in length. Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 5-6: <i>Bactrocera papayae</i> samples.	47
Figure 4.10	COII banding profile of fruit fly samples from Serdang, collected from papaya host fruit. Amplification of COII fragment was expected to be around 374 base pairs in length. Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 1-4: <i>Bactrocera carambolae</i> samples. Lanes 5-7: <i>Bactrocera papayae</i> samples.	48
Figure 4.11	COII banding profile of fruit fly samples from Serdang, collected from guava host fruit. Amplification of COII fragment was expected to be around 374 base pairs in length. Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 1-4: <i>Bactrocera carambolae</i> samples. Lanes 5-7: <i>Bactrocera papayae</i> samples.	48
Figure 4.12	COII banding profile of fruit fly samples from Sarawak, collected from starfruit host fruit. Amplification of COII fragment was expected to be around 374 base pairs in length. Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 1-4: <i>Bactrocera carambolae</i> samples. Lanes 5-7: <i>Bactrocera papayae</i> samples.	49
Figure 4.13	COII banding profile of fruit fly samples from Serdang, collected from jambu madu host fruit. Amplification of COII fragment was expected to be around 374 base pairs in length. Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 1-4: <i>Bactrocera carambolae</i> samples. Lanes 5-7: <i>Bactrocera papayae</i> samples.	49
Figure 4.14	COII banding profile of fruit fly samples from Serdang, collected from jambu air host fruit. Amplification of COII fragment was expected to be around 374 base pairs in length. Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 1-4: <i>Bactrocera carambolae</i> samples. Lanes 5-7: <i>Bactrocera papayae</i> samples.	50
Figure 4.15	<i>Cytb</i> banding profile of fruit fly samples from Serdang, collected from papaya host fruit. Amplification of COI fragment was expected to be around 484 base pairs in length. Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 1-4: <i>Bactrocera carambolae</i> samples. Lanes 5-8: <i>Bactrocera papayae</i> samples.	51
Figure 4.16	<i>Cytb</i> banding profile of fruit fly samples from Serdang, collected from guava host fruit. Amplification of COI fragment was expected to be around 484 base pairs in length. Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 1-4: <i>Bactrocera carambolae</i> samples. Lanes 5-8: <i>Bactrocera papayae</i> samples.	51

Figure 4.17	<i>Cytb</i> banding profile of fruit fly samples from Sarawak, collected from starfruit host fruit. Amplification of COI fragment was expected to be around 484 base pairs in length. Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 1-4: <i>Bactrocera carambolae</i> samples.	52
Figure 4.18	<i>Cytb</i> banding profile of fruit fly samples from Sarawak, collected from starfruit host fruit. Amplification of COI fragment was expected to be around 484 base pairs in length. Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 1-4: <i>Bactrocera papayae</i> samples.	52
Figure 4.19	<i>Cytb</i> banding profile of fruit fly samples from Serdang, collected from jambu madu host fruit. Amplification of COI fragment was expected to be around 484 base pairs in length. Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 1-4: <i>Bactrocera carambolae</i> samples. Lanes 5-8: <i>Bactrocera papayae</i> samples.	53
Figure 4.20	<i>Cytb</i> banding profile of fruit fly samples from Serdang, collected from jambu air host fruit. Amplification of COI fragment was expected to be around 484 base pairs in length. Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 1-4: <i>Bactrocera carambolae</i> samples. Lanes 5-8: <i>Bactrocera papayae</i> samples.	53
Figure 4.21	The 50% majority-rule consensus tree generated by maximum likelihood analysis based on COI DNA sequences. Numbers at the nodes indicate bootstrap values.	56
Figure 4.22	Bayesian phylogenetic analysis depicting the various categorizing of <i>Bactrocera</i> fruit fly variants based on COI. Numbers at the nodes indicate percentage of posterior probabilities over a generation number of two million generations.	58
Figure 4.23	Phylogenetic tree generated by Neighbor-Joining analysis based on COI gene. Numbers at the nodes indicate bootstrap values.	60
Figure 4.24	The 50% majority-rule consensus tree generated by maximum likelihood analysis based on COII DNA sequences. Numbers at the nodes indicate bootstrap values.	62
Figure 4.25	Bayesian phylogenetic analysis depicting the various categorizing of <i>Bactrocera</i> fruit fly variants based on COII. Numbers at the nodes indicate percentage of posterior probabilities over a generation number of two million generations.	64
Figure 4.26	Phylogenetic tree generated by Neighbor-Joining analysis based on COII gene. Numbers at the nodes indicate bootstrap values.	66
Figure 4.27	The 50% majority-rule consensus tree generated by maximum likelihood analysis based on <i>cytb</i> DNA sequences. Numbers at the nodes indicate bootstrap values.	68
Figure 4.28	Bayesian phylogenetic analysis depicting the various categorizing of <i>Bactrocera</i> fruit fly variants based on <i>cytb</i> . Numbers at the nodes indicate percentage of posterior probabilities over a generation number of two million	70

	generations.	
Figure 4.29	Phylogenetic tree generated by Neighbor-Joining analysis based on <i>cytb</i> gene. Numbers at the nodes indicate bootstrap values.	72
Figure 4.30	The 50% majority-rule consensus tree generated by maximum likelihood analysis based on combined COI, COII, and <i>cytb</i> DNA sequences. Numbers at the nodes indicate bootstrap values.	74
Figure 4.31	Bayesian phylogenetic analysis depicting the various categorizing of <i>Bactrocera</i> fruit fly variants based on combined COI, COII, and <i>cytb</i> DNA sequences. Numbers at the nodes indicate percentage of posterior probabilities over a generation number of two million generations.	76
Figure 4.32	Phylogenetic tree generated by Neighbor-Joining analysis based on combined COI, COII, and <i>cytb</i> DNA sequences. Numbers at the nodes indicate bootstrap values.	78
Figure 4.33	Statistical parsimony networks for COI of <i>Bactrocera carambolae</i> and <i>Bactrocera papayae</i> . Size of squares and ovals are proportional to the haplotype frequency. Lines represent parsimonious connections between haplotype groups while the clear circles indicate hypothetical missing haplotype.	82
Figure 4.34	Statistical parsimony networks for COII of <i>Bactrocera carambolae</i> and <i>Bactrocera papayae</i> . Size of squares and ovals are proportional to the haplotype frequency. Lines represent parsimonious connections between haplotype groups while the clear circles indicate hypothetical missing haplotype.	84
Figure 4.35	Statistical parsimony networks for <i>cytb</i> of <i>Bactrocera carambolae</i> and <i>Bactrocera papayae</i> . Size of squares and ovals are proportional to the haplotype frequency. Lines represent parsimonious connections between haplotype groups while the clear circles indicate hypothetical missing haplotype.	87
Figure 4.36	Statistical parsimony networks for combined COI, COII, and <i>cytb</i> of <i>Bactrocera carambolae</i> and <i>Bactrocera papayae</i> . Size of squares and ovals are proportional to the haplotype frequency. Lines represent parsimonious connections between haplotype groups while the clear circles indicate hypothetical missing haplotype.	90

LIST OF TABLES

		Page
Table 2.1	Distinguishing morphological features between <i>Bactrocera carambolae</i> and <i>Bactrocera papayae</i> .	16
Table 3.1	Host fruit collection, date of collection, location of origin, and species of host plant.	30
Table 3.2	Oligonucleotide primers used for polymerase chain reaction (PCR) amplification.	33
Table 3.3	Species used for bioinformatics analysis.	39
Table 4.1	Specimen list and host fruits origin.	41
Table 4.2	Haplotype grouping and their identities based on COI, COII, <i>cytb</i> , and combined COI, COII, and <i>cytb</i> molecular markers.	79

LIST OF SYMBOLS AND ABBREVIATIONS

AIC	Akaike Information Criterion
AFLP	Amplified Fragment Length Polymorphism
AW-IPM	Area-wide integrated pest management
BAT	Baited annihilation trapping
bp	Base pair
BI	Bayesian Inference
BIC	Bayesian Information Criterion
COI	Cytochrome oxidase subunit I
COII	Cytochrome oxidase subunit II
<i>cytb</i>	Cytochrome-b
°C	Degree Celsius
DNA	Deoxyribonucleic Acid
dNTP	Deoxyribonucleotide triphosphate
ddNTP	Dideoxynucleotide
<i>et al.</i>	et alia (“and others”)
EDTA	Ethylenediaminetetraacetic acid
F81	Felsenstein 1981 model
GTR	General time-reversible DNA substitution model
GIS	Geographical information system
JC	Jukes-Cantor model
K2P	Kimura 2-parameter model
MgCl ₂	Magnesium chloride
MAT	Male annihilation trapping
MCMC	Markov chain Monte Carlo

ML	Maximum likelihood
μl	Microliter
μM	Micromolar
ml	Milliliter
mM	Millimolar
mtDNA	Mitochondrial DNA
>	More than
ng	Nanogram
-	Negative
NJ	Neighbor-Joining
%	Percent
PCR	Polymerase Chain Reaction
PCR-RFLP	Polymerase Chain Reaction Restriction Fragment Length Polymorphism
+	Positive
PP	Posterior probability
RAPD	Random Amplification of Polymorphic DNA
RFLP	Restriction Fragment Length Polymorphism
rpm	Revolutions per minute
RNA	Ribonucleic Acid
rRNA	Ribosomal RNA
<i>s. s.</i>	Sensu stricto
STR	Short tandem repeat
SNP	Single nucleotide polymorphism
SIT	Sterile insect technique
tRNA	Transfer RNA
TAE	Tris-acetate-EDTA

UHQ	Ultra high quality
UPGMA	Unweighted Pair Group Method with Arithmetic Mean
v	Version
V	Volts

CHAPTER 1

INTRODUCTION

The fruit fly, genus *Bactrocera*, comprises of about 500 described species and is grouped in the subfamily Dacinae (Drew, 1989a; Drew and Hancock, 2000). Several species of fruit flies from this genus are pests to economically important fruits in the agricultural industry; for example in Malaysia, the melon fly, *Bactrocera cucurbitae*; the Carambola fruit fly, *Bactrocera carambolae*; and the papaya fruit fly, *Bactrocera papayae*. These fruit flies have a wide host range, making a lot of fruits and vegetables vulnerable to be attacked from not only by one, but several different species of fruit flies. The *Bactrocera dorsalis* species complex comprises of at least 52 described species in the Asia-Pacific region (Shearman *et al.*, 2006). One of the most significant groups of fruit flies to the agricultural industry is the Dacinae fruit flies. They are key pest groups of Asia and the Pacific (Waterhouse, 1993; Waterhouse, 1997), and these fruit flies are frugivorous on a wide range of fruits and vegetables (Allwood *et al.*, 1999).

In Malaysia, these fruit flies are considered serious quarantine pests as they inflict irrefutable losses in field productions of fruits and vegetables, and they also cause difficulties in fresh horticultural exports due to infestations. If left unchecked, a single fruit may be completely damaged by fruit flies (Vijaysegaran, 1983). Damaged fruits caused by fruit flies raise the cost of fruit production as well as reduce the overall production of fruits in the field. Methods to curb fruit fly infestations are known to be labour intensive and costly, and these too raises the overall cost of fruit production. Fruit flies attack ripe fruits and this limits export of fruit produce to other countries. Exported produce that is infested by fruit flies will have to undergo quarantine

disinfections, or total eradication of the produce and this causes major losses for the exporter country. Countries such as Japan, Europe, and the United States of America take quarantine pests such as fruit flies very seriously due to the fear of introducing these pest species into the country (Vijaysegaran, 1996).

Two of the most significantly important fruit flies in Malaysia are the carambola fruit fly, *Bactrocera carambolae*, and the papaya fruit fly, *Bactrocera papayae*. These two fruit flies have been considered as major pest species in Malaysia due to their ability to infest a wide range of host fruits (Shi *et al.*, 2009). *Bactrocera carambolae* and *Bactrocera papayae* belong in the *Bactrocera dorsalis* species complex and despite being classified as two separate species, both species share very similar morphological characters, with minor variations in wing pattern bandings and abdominal markings (Ebina and Ohto, 2006; Chua *et al.*, 2009). Hybridizations between *Bactrocera carambolae* and *Bactrocera papayae* are known to occur and this gives rise to fertile offspring that have intermediate morphological features, making identification of the fruit flies even more daunting and complicated. Identification of fruit flies based on morphological features alone is a difficult task (Clarke *et al.*, 2005) due to morphological similarities between members of the genus. Misidentifications of fruit flies have occurred in the past, in that the *Bactrocera carambolae* and *Bactrocera papayae* fruit flies were once classified as *Dacus dorsalis* (Hendel) (Hardy and Adachi, 1959; White and Elson-Harris, 1992), but revisions of the *Bactrocera dorsalis* species complex by Drew and Hancock (1994) resulted in the *Bactrocera carambolae* and *Bactrocera papayae* fruit flies separated as two distinct species. In addition to that, identification of fruit flies in egg and larval stages are also difficult, if not impossible to be carried out (Baliraine *et al.*, 2002). There are no distinguishing features between the two species of fruit flies at the developmental stages, and this makes quarantine management of infested produce difficult.

To facilitate with fruit fly identification, a reliable method of identification that is not limited by polymorphism and stage development of a target species is required (Asokan *et al.*, 2007). Molecular markers are an essential tool in differentiating species that are not easily separated through morphological methods (Yu *et al.*, 2000). It is very important to obtain molecular data in order to resolve and establish phylogenetic relationships of fruit flies, particularly of species that are of economic significance, such as *Bactrocera carambolae* and *Bactrocera papayae*. Not only will the molecular data provide insight to identification and quarantine management of economically significant fruit flies, it will also provide important taxonomical data for classification of fruit flies and to further improve the taxonomy status of the *Bactrocera* genus. Identification of species of fruit flies in fruit produce will undoubtedly help in the quarantine management of exported produce.

A very useful application of molecular markers is the identification of a particular species of organism. A short segment of an organism's DNA is compared with a database in order to determine the identity of an unknown organism. This method is best utilized when morphological methods of identification are unable to determine the species of an organism. An example is the identification of insects during their larval stages – insect larvae tend to have fewer diagnostic characteristics compared to their adult stage, thus making identification impossible (Caterino and Tishechkin, 2006; Tang *et al.*, 2010). Using molecular markers, it would be a simple task of identifying unknown species of organisms regardless of the lack of morphological distinction between life cycle stages and/or diagnostic characteristics. Ball and Armstrong (2006) have demonstrated the usage of molecular markers to identify tussock moth species in their larval stages. Lefort *et al.* (2012) have also demonstrated noninvasive molecular methods in identifying live scarab larvae using molecular markers. Routine identification of species of insects oftentimes requires highly specialized knowledge and

can be time consuming, usually proving to be limiting factors for ecological or biodiversity studies (Floyd *et al.*, 2002; Hajibabaei *et al.*, 2007). Raupach *et al.* (2010) have utilized molecular markers to identify species of ground beetles, and Piffaretti *et al.* (2012) have revealed the existence of two sibling species of aphids in the *Brachycaudus helichrysi* species using molecular markers.

In line with identification of unknown species, molecular markers are also utilized to study evolutionary relationships of an organism. The relatedness of a group of organisms are studied by examining the molecular differences of the organisms' DNA sequences and a phylogenetic tree is inferred to determine which organism is more closely related to another and how they are grouped together according to the differences of their DNA sequences. The large size of the *Bactrocera* genus warrants a phylogenetic study in order to further organize the taxonomy of the fruit flies, particularly the *Bactrocera dorsalis* complex which comprises of sibling species that are closely related (Yong, 1995). Furthermore, this complex consists of species that have remarkably similar morphological characteristics. Thus, it is important that a reliable method of identification is established to properly identify the many morphologically similar species of fruit flies, as well as establishing a more concrete taxonomic classification for this group of organisms. Another aspect addressed in this project is the host specificity of the *Bactrocera carambolae* and *Bactrocera papayae* fruit flies – what are the phylogenetic relationships between the two species of fruit flies that infest different host fruits? Are the fruit flies that infest one particular fruit different than fruit flies that infest other fruits? Molecular markers can certainly be utilized to ascertain the phylogenetic relationships among fruit flies that infest different types of fruits.

Therefore, in this project, we would like to determine:

1. What are the effects of locations and host fruits in the phylogenetics of *Bactrocera carambolae* and *Bactrocera papayae*?
2. What are the phylogenetic relationships between *Bactrocera carambolae* and *Bactrocera papayae*?
3. Will the selected molecular markers (COI, COII, and *cytb*) be able to distinguish between *Bactrocera carambolae* and *Bactrocera papayae*?

This project encompasses the molecular differentiation and phylogenetics of the *Bactrocera dorsalis* complex, with special reference to the *Bactrocera carambolae* and *Bactrocera papayae* species. Within the *Bactrocera dorsalis* complex are sibling species that share common morphological features, and thus makes identification of fruit fly species within the complex difficult. Hybridizations between sibling species give rise to fruit flies with intermediate morphological features that cause distinguishing between one species with another difficult to conduct. Therefore, it is important to develop molecular markers that are capable and reliable in distinguishing between species of fruit flies within the *Bactrocera dorsalis* complex. Rapid identification of fruit flies may facilitate in quarantine, as well as pest management to reduce crop destruction caused by fruit flies. The molecular markers of choice for this project are mitochondrial-encoded. The molecular markers that have been widely used in phylogenetic studies of fruit flies (Nakahara and Muraji, 2008; Han and Ro, 2009; Zhang *et al.*, 2010; Lim *et al.*, 2012) and have assisted in resolving many *Bactrocera* relationships from different taxa.

What is expected from this project is mainly to be able to distinguish between the *Bactrocera carambolae* and *Bactrocera papayae* fruit flies based on the selected molecular markers. At the same time, the phylogenetic relationship between the two species can be determined, also based on the selected molecular markers. In part with collecting host fruits from different locations in Malaysia, the effect of geographical locations and host fruits in the phylogenetics of *Bactrocera carambolae* and *Bactrocera papayaecan* can be determined as well in this study.

The objectives of this study include:

1. To determine the effects of locations and host fruits in the phylogenetics of *Bactrocera carambolae* and *Bactrocera papayae*.
2. To determine the phylogenetic relationships between *Bactrocera carambolae* and *Bactrocera papayae*.
3. To determine the ability of selected molecular markers (COI, COII, and *cytb*) in distinguishing between *Bactrocera carambolae* and *Bactrocera papayae*.

CHAPTER 2

LITERATURE REVIEW

2.1 Taxonomic Hierarchy

Kingdom: Animalia

Phylum: Arthropoda

Class: Insecta

Order: Diptera

Family: Tephritidae, Newman 1834

Subfamily: Dacinae

Tribe: Dacini

Genus: *Bactrocera*, Macquart 1835

Subgenus: *Bactrocera* (*Bactrocera*), Macquart 1835

Species: *Bactrocera carambolae*, Drew and Hancock, 1994; *Bactrocera papayae*, Drew and Hancock, 1994

(Source: Norrbom *et al.*, 1998)

2.2 Dacinae fruit flies

The subfamily Dacinae consists of four genera of fruit flies; two minor genera *Ichneumonopsis* Hardy and *Monacrostichus* Bezzi; and two large genera, the *Bactrocera* Macquart and *Dacus* Fabricius (Drew and Hancock, 2000). They are found predominantly in tropical and subtropical regions. They are distributed from the continent of Africa, across the Indian subcontinent, through the Southeast Asian region and across the southern Pacific zone (Tsuruta and White, 2001; Drew, 2004). Within South East Asia and the Pacific, the Dacinae species of fruit flies are found throughout Malaysia, Indonesia, Papua New Guinea, New Caledonia, and Vanautu. The genus *Bactrocera* Macquart itself is the largest genera not only within the subfamily Dacinae, but within the Tephritidae family as well, consisting of 500 described species and arranged in 28 subgenera (Drew, 1989a; Drew and Hancock 2000).

Classification of the Dacinae at subfamily level is primarily based on antennal segment 3 elongate, elongated apical lobe extension on cell Cu, dense mitotrichia in males, reduced chaetotaxy on the head and thorax, tergum V containing a pair of shining spots usually with presence of cilia on posterolateral margins of tergum III in males, and females with a pair of coiled spermathecae (Hardy, 1973, 1974).

2.3 *Bactrocera dorsalis* species complex

The *Bactrocera dorsalis* forms a species complex consisting of sibling species belonging in the genus *Bactrocera*, formerly genus *Dacus* (Drew, 1989b). Up to 75 species have been described in Asia (Clarke *et al.*, 2005). Species within the complex share common morphological features such as the wings and thorax, and are oftentimes difficult to distinguish morphologically. Some of the species have morphological characteristics that fall within an intermediate range and they tend to segregate throughout a population (Yong, 1995; Iwahashi, 1999; Clarke *et al.*, 2005). Once, many species in this complex, such as *Bactrocera carambolae*, *Bactrocera papayae*, and *Bactrocera dorsalis* were misidentified or classified as one species, *Bactrocera dorsalis* (*Dacus dorsalis*) (Hardy, 1969; White and Elson-Harris, 1992).

Based on morphological features, Drew and Hancock (1994) have revised 12 species of fruit flies within the complex and multiple sibling species were recognised, including *Bactrocera carambolae* and *Bactrocera papayae*. Though these sibling species had very similar morphological features, it is with a set of findings that the distinction between these sibling species were recognized – findings include host range, geographical data, pheromone analysis, as well as allozyme analysis (Perkins *et al.*, 1990; Ooi, 1991).

As described by Drew and Hancock (1994), the *Bactrocera dorsalis* complex is distinguished as having clear wing membranes with dark, narrow coastal bands not reaching R4+5. Scutum is mostly black, while the scutellum is yellow with a narrow dark basal band. The abdominal tergites T3-T5 have a distinct medial longitudinal black “T”-shaped mark that varies from species to species within the complex. Figure 2.1 shows the general morphology of a *Bactrocera carambolae* fruit fly.

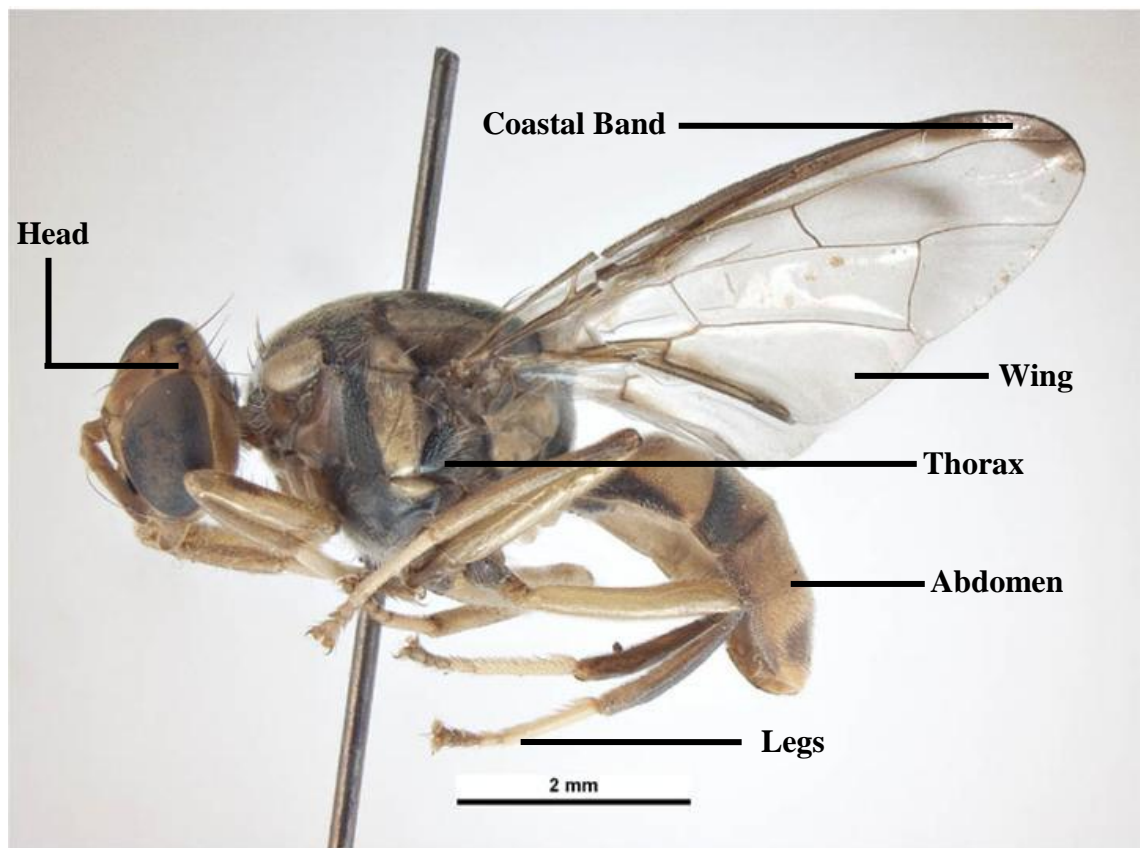


Figure 2.1: Morphology of *Bactrocera carambolae*

(Modified from Walker, 2005)

2.4 Economic and Agricultural Importance

Out of the number of species found within the complex, a few species have been classified as economically significant (Drew and Hancock, 1994; Drew and Romig, 1997; Clarke *et al.*, 2005). In Malaysia, *Bactrocera carambolae* and *Bactrocera papayae* both are the dominant tephritid fruit flies (Chua, 1991) and are significantly important in the agricultural industry. These fruit flies lay eggs in fruits or other horticultural products and the young larvae tunnel and feed inside the fruit. Puncture of the fruit causes discolouration, oozing, and also an increase in fermentation and decomposition of the fruit due to secondary infection of the fruit by other

microorganisms. This causes the fruit to be unsalable in the market due to the poor quality of the fruit and also causes restriction of export to other countries. Quarantine of infested fruits results in loss of potential markets. Countries such as Japan and the United States of America take quarantine pests such as fruit flies very seriously (Vijaysegaran, 1996).

Due to infestations, control measures have been taken to reduce the fly populations. Examples of control measures include bagging of individual fruits with paper to prevent infestation, and also by the usage of insecticides. Bagging reduces damage and infestation of individual fruits, but it is laborious and time consuming. Insecticides are a fast and effective method to control fruit flies, but it is harmful to the environment. Other species of insects are indirectly killed; such as pollinators and natural predators of other fruit flies. If used without restraint, insecticides are detrimental to the environment as water supplies may be contaminated with insecticides (Liess and Schulz, 1998).

Area-wide integrated pest management (AW-IPM) is an effective and environmentally friendly method for controlling fruit fly pests. It generally involves pest management techniques such as protein baited annihilation trapping (BAT), male annihilation trapping (MAT), wild host cutting, deployment and augmentation of natural predators, and geographical information system (GIS). Application of the AW-IPM method is concerned with controlling entire pest populations, including fruit orchards, domestic gardens, as well as areas where wild hosts are found (Hendrichs *et al.*, 2007; Lindquist 2000). An example of a successful application of the AW-IPM method includes the control of the Mediterranean fruit flies (Hendrichs *et al.*, 2007). Another alternative biological control method of pest fruit flies is the sterile insect technique (SIT) which involves the mass release of sterile insects into the wild. In the wild, these sterile wild insects are utilized to compete and win the overall competition

with the fertile insects. SIT is believed to be the most target-species specific and the least destructive pest control technique (Enkerlin *et al.*, 2003) utilized as a means to control the population of pest insects. However, the application of AW-IPM-SIT depends on the competitive mating between the released sterile flies and their wild fertile counterparts. The level of sterility of the released sterile flies is also of concern as there are no proper methods to detect the accidental release of fertile flies into the wild (Aketerawong *et al.*, 2011). Another factor that needs to be taken into account is the possible pre-mating reproductive isolation between the released sterile flies and the wild fertile flies (Krafsur, 2005).

2.5 Life Cycle

In unfavourable conditions for breeding, adult fruit flies enter a stage of facultative reproductive diapause where they shelter and remain sexually inactive. During this time, adult female fruit flies with eggs and developing follicles in the ovaries are resorbed and resources mobilised to reserve energy and increase the chances of survival (Fletcher, 1989).

The female fruit fly lays its eggs below the skin of the host fruit and they hatch within 1-2 days under tropical conditions. The larvae start feeding on the fruit's flesh and they develop inside the fruit. The larvae undergo three larval developmental stages (instar) before they finish feeding and enter the pupa stage. This development phase takes around 6-9 days. Once the larvae have reached the third instar, they burrow into the soil and form a barrel-shaped, tanned brown and hard shell known as a puparium, when the fruit reaches maturity and drops to the ground. Within the puparium, the

larvae develop into an adult fruit fly. After 10 to 14 days, the adult fruit fly emerges from the puparium (Narayanan and Batra, 1960).

Adult fruit flies may live for 1-3 months after emergence depending on the temperature (Christenson and Foote, 1960). After emerging from the puparium, the adult fruit flies are sexually immature and must forage for resources in order to survive and sexually mature (Raghu, 2003). The adult fruit flies only start mating after 8-12 days. The females are capable of laying 1,200 to 1,500 eggs per individual in its lifetime in field conditions, and they lay their eggs in readily available host fruits. The fruit flies may have more than one generation of offspring depending on the availability of host fruits. Figure 2.2 shows the life cycle of a typical fruit fly.

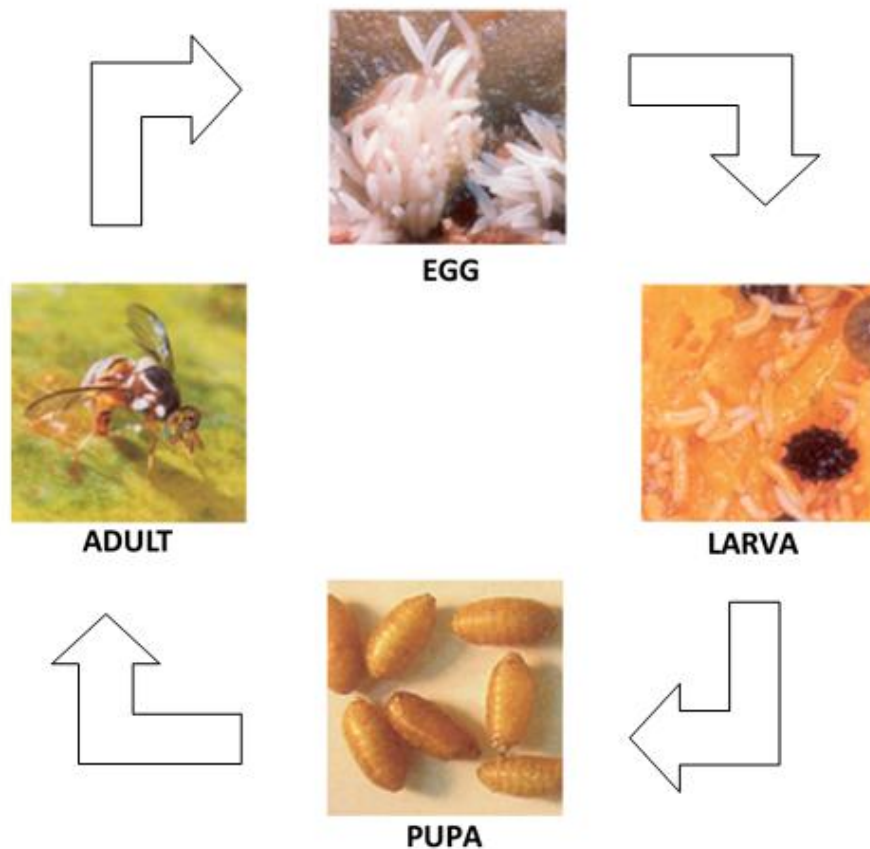


Figure 2.2: The life cycle of a fruit fly

(Source: www.extento.hawaii.edu)

2.6 Species of *Bactrocera carambolae* and *Bactrocera papayae*

The *Bactrocera carambolae* and *Bactrocera papayae* fruit flies are known for being extremely polyphagous. *Bactrocera carambolae* has a host range of 77 host species across 27 families, while *Bactrocera papayae* has a host range of 209 host species across 51 plant families (Drew, 1989b; Drew and Hancock, 1994; Drew and Raghu, 2002).

Bactrocera carambolae, also known as the carambola fruit fly, is a polyphagous pest fruit fly that is capable of infesting many different types of hosts. Its host range includes the carambola, mango, avocado, guava, jackfruit, and orange, to name a few. The carambola fruit fly is known to be a very serious pest of the carambola fruit, in that it attacks fruits while they are still very young. Within Southeast Asia, they are distributed in Western Indonesia, Southern Thailand, Peninsular and East Malaysia, the Andaman Islands (India), Singapore, and Brunei (Drew and Hancock, 1994; White, 1996). *Bactrocera carambolae* is native to Indonesia and Malaysia, but is known to disperse to other areas through methods of fruit importation, as with the introduction of *Bactrocera carambolae* fruit flies to places like Suriname, South America (van Sauer-Muller, 1991). The *Bactrocera carambolae* is nearly similar to *Bactrocera papayae* morphologically, except that the *Bactrocera carambolae* fruit flies have deep coastal bands on their wings, and intermediate abdominal markings.

Bactrocera papayae, also known as the papaya fruit fly, is also part of the *Bactrocera dorsalis* complex, just like the *Bactrocera carambolae*. They are also serious polyphagous pest fruit flies and have an overlapping range of host fruits with *Bactrocera carambolae* fruit flies. Their range of host fruits includes mango, papaya, carambola, guava, and banana to name a few. *Bactrocera papayae* are native to

Southeast Asia within Thailand, Peninsular Malaysia, East Malaysia, Singapore, Indonesia, and Kalimantan (Drew and Hancock, 1994). They share similar morphological characters with the *Bactrocera carambolae*, except that the *Bactrocera papayae* fruit flies have narrow coastal bands on their wings, and narrow abdominal markings (Drew and Hancock, 1994). Figures 2.3 and 2.4 show the morphological features of *Bactrocera carambolae* and *Bactrocera papayae* respectively. Table 2.1 lists the distinguishing morphological features between the *Bactrocera carambolae* and *Bactrocera papayae* fruit flies.



Figure 2.3: Morphological features of *Bactrocera carambolae*

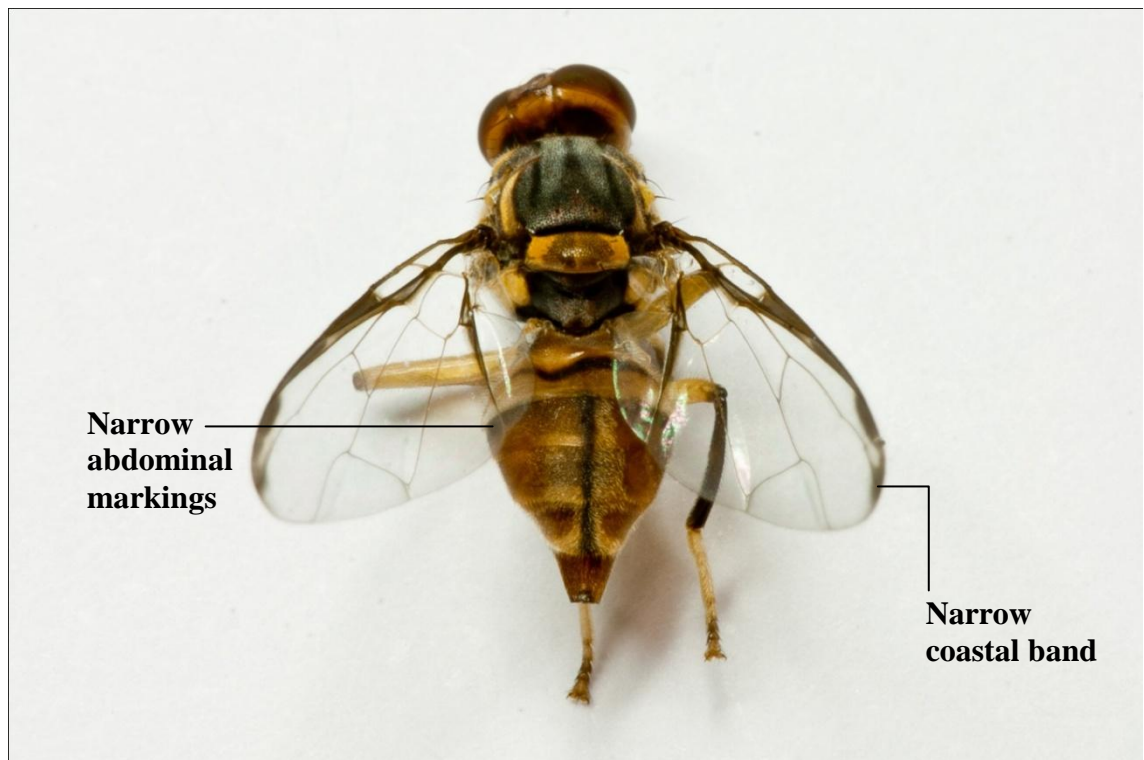


Figure 2.4: Morphological features of *Bactrocera papayae*

Table 2.1: Distinguishing morphological features between *Bactrocera carambolae* and *Bactrocera papayae*.

Characteristic Feature	<i>Bactrocera carambolae</i>	<i>Bactrocera papayae</i>
Aculeus length	Short	Long
Coastal band	Deep	Narrow
Abdominal markings	Intermediate	Narrow

2.7 Problems with Distinguishing *B. carambolae* and *B. papayae*

Despite being classified as two separate species, researchers have had prior difficulties in distinguishing *B. carambolae* and *B. papayae*. This is greatly attributed by the two species having similar morphological features; hence prior to Drew and Hancock's (1994) revision, *B. carambolae* and *B. papayae* were once classified as one species, *Dacus dorsalis* (Hendel). Morphological methods for identifying *B. carambolae* and *B. papayae* primarily involves the abdominal markings and coastal bands on the wings, however, these morphological features are polymorphic and range in intermediate forms that segregate within the species (Iwahashi, 2001). Behavioural research has shown that *B. carambolae* and *B. papayae* have the capability to cross-breed in laboratory conditions (Yong, 1995; Tan, 2003), which may lead to hybrids with intermediate morphological features. Ebina and Ohto (2006) have shown that hybrids of *B. carambolae* and *B. papayae* have intermediate morphological features. Genetic methods have also had varying results in attempting to distinguish *B. carambolae* and *B. papayae* (Muraji and Nakahara, 2002; Chua *et al.*, 2009; Krosch *et al.*, 2012a).

2.8 Molecular Markers

Rapid development in the genetics field has led to the development of a variety of techniques to analyze genetic variation (Karp *et al.*, 1996, 1997a, b; Parker *et al.*, 1998; Schlötterer, 2004). Molecular markers are now mainly used to investigate life history and evolutionary relationships of organisms, and to a certain extent, their behaviour as well. Most of the time, molecular markers are used alongside information

from other various fields such as comparative morphology, ecology, systematics, paleontology, and ethology to gain a better understanding of the molecular data (Awise, 1994).

Molecular markers differ with respect to their important features, such as level of polymorphism detected, specificity of locus, reproducibility, genomic abundance, cost, and technical requirements. The appropriate molecular marker depends on its application; hence, no molecular marker is superior to the other.

Usage of molecular markers is advantageous in that molecular markers do not exhibit phenotypic plasticity and are better at providing homologous traits while morphological and biochemical markers are subject to environmental conditions and can vary depending on the environment. Data from molecular markers are more easily scored as discrete states of alleles or DNA base pairs compared to morphological and biochemical parameters where the data must be scored as continuous variable characters, limiting its usage in analytical methods. The ease in scoring allele states or DNA base pairs leads to better classification of species with very similar morphological features (Muraji and Nakahara, 2001; Wanwisa *et al.*, 2003). Molecular markers are also selectively neutral and an abundant of independent molecular markers are available for research (Spooner *et al.*, 2005).

Some common types of molecular markers and their recent applications in fruit fly and insect research include PCR-RFLP (Nakahara *et al.*, 2008; Chua *et al.*, 2009), RAPD (Segura *et al.*, 2008, Zahran *et al.*, 2009), AFLP (Kakouli-Duarte *et al.*, 2001; Sadeghi *et al.*, 2010), microsatellite (Aketarawong *et al.*, 2011; Wu *et al.*, 2011; Shi *et al.*, 2012), and mitochondrial DNA.

2.9 Mitochondrial DNA (mtDNA)

Mitochondrial DNA (mtDNA) is localized within the matrix of the organelle mitochondrion. As opposed to the linearly shaped nuclear DNA, mitochondrial DNAs are covalently closed circular molecules. Mitochondrial DNAs run the length of about 16 to 20 kilobases long. Animal mitochondrial DNAs have 37 tightly packed genes which comprises of 13 protein genes; two ribosomal RNA (rRNA) genes; and 22 transfer RNA (tRNA) genes. Introns are absent within the mitochondrial DNA. A “control” region is also present and its sequences are responsible for initiating mitochondrial DNA replication and RNA transcription. The control region in insects is rich in adenine and thymine (Rand and Harrison, 1986) and is roughly 0.8 kilobases long. The mitochondrion genome of fruit flies consists of a circular DNA molecule which is approximately 16,000 base pairs. The complete mitochondrial genome of the *Drosophila melanogaster* fruit fly has recently been sequenced, and its genome consists of 15,915 base pairs that encode 37 genes generally found in animal mitochondrial genomes (Yu *et al.*, 2007).

Mitochondrial DNA sequences have been extensively used for phylogenetic studies (Lunt *et al.*, 1996). Their advantages include (1) they are maternally inherited (Avice and Lansman, 1983; Avice, 1986); (2) mitochondrial DNA are present and well distributed among a wide variety of organisms (Avice *et al.*, 1987); (3) mitochondrial DNA evolves at a faster rate than nuclear DNA (Brown *et al.*, 1982; Moriyama and Powell, 1997); (4) mitochondrial DNA are highly conserved among different phyla (Morlais and Severson, 2002); and (5) mitochondrial DNA are haploid and sequences are easily obtainable without carrying out any DNA cloning (Hurst and Jiggins, 2005).

Evolution of mitochondrial DNA nucleotide sequence is 5 to 10 times faster than that of nuclear DNA (Brown *et al.*, 1982), most of which occur through base substitution, additions and/or deletions of nucleotides, and differences in mitochondrial DNA length. Nuclear DNA markers require a significant amount of time to refine primers for a target species. Steps involved include the sampling of genes with appropriate evolutionary rates, and once the correct genes are located, alleles from heterozygous individuals need to be separated through cloning before any DNA sequencing can be performed (Hurst and Jiggins, 2005).

In this study, three mitochondrial DNA genes are utilized for phylogenetic studies, namely cytochrome c oxidase subunit I (COI), cytochrome c oxidase subunit II (COII), and cytochrome b (*cytb*).

COI and COII are part of the cytochrome c oxidase complex, with COI being the main subunit. The COI gene is about 1400 base pairs (Schroeder *et al.*, 2003) and considered to be highly conserved among protein-coding genes in the animal mitochondrion genome (Brown, 1985). This characteristic of COI genes makes it beneficial for phylogenetic studies at the species level, and thus, many studies have been carried out using the COI gene as a molecular marker (Jamnongluk *et al.*, 2003; Lewis *et al.*, 2005; Nakahara and Muraji, 2008, Zhang *et al.*, 2010). The COII gene is also widely used in various phylogenetic studies (Crozier *et al.*, 1989; Simon *et al.*, 1994; Ito *et al.*, 2010; Ruiz-Garcia and Pinedo-Castro, 2010). Besides phylogenetic studies, cytochrome c oxidase sequences have been used to study genetic diversity, population structure, and origination of a particular species (Zhang and Hewitt 1997; Shi *et al.*, 2009; Li *et al.*, 2011; Prabhakar *et al.*, 2012).

Cytb is the main subunit of the complexes cytochrome bc1 and b6f, and is a component of the mitochondrial oxidative phosphorylation system (Hatefi, 1985). Although the *cytb* gene has a slow evolution rate due to regions of the gene being more conserved (Meyer 1994), the silent regions of the gene has a relatively fast evolution rate as a result of nucleotide transversions (Irwin *et al.*, 1991). The dual nature of the gene's variability and conservability warrants the gene's potential for population and phylogenetic studies (Meyer, 1994). Various fruit flies have been studied by utilizing the *cytb* gene as a molecular marker (Zhu *et al.*, 2005a, b, Wan *et al.*, 2011), as well as various other organisms (Cook *et al.*, 1999; Su *et al.*, 1999; Segura *et al.*, 2006; Nishikawa *et al.*, 2012).

2.10 Molecular Phylogenetic Studies on *Bactrocera carambolae* and *Bactrocera papayae*

Over the years, studies pertaining to the molecular phylogeny of the *Bactrocera dorsalis* species complex have been conducted to determine the taxonomic status of members of the species complex. PCR-RFLP methods have been used to discriminate between *Bactrocera* pest species (including *Bactrocera carambolae* and *Bactrocera papayae*) (Muraji and Nakahara, 2002; Chua *et al.*, 2009), and also to investigate interspecific hybrids between *Bactrocera carambolae* and *Bactrocera papayae* (Ebina and Ohto, 2006). Molecular phylogeny of fruit flies on higher taxonomic levels which include the *Bactrocera* fruit flies were also studied using mitochondrial DNA markers, particularly on the family level (Han and Ro, 2009), and tribe level (Smith *et al.*, 2002; Zhang *et al.*, 2010; Krosch *et al.*, 2012a). Studies of fruit flies within the genus *Bactrocera* have also been conducted which attempts to discriminate members of

different groups of fruit flies contained within the *Bactrocera* genus, such as *Bactrocera*, *Zeugodacus*, *Austrodacus*, and so forth (Jamnongluk *et al.*, 2003; Smith *et al.*, 2003; Nakahara and Muraji, 2008, Zhang *et al.*, 2010). Within all the different levels of taxonomic studies, members of the *Bactrocera dorsalis* species complex, including *Bactrocera carambolae* and *Bactrocera papayae*, were included as well in their studies. Comparative studies between morphological and genetic data have been conducted in attempts to distinguish between *Bactrocera dorsalis* and *Bactrocera papayae* (Kroschet *et al.*, 2012b; Schutze *et al.*, 2012).

2.11 Polymerase Chain Reaction

Polymerase Chain Reaction (PCR) is an enzymatic DNA amplification method whereby the template DNA is multiplied million-folds by a set of cycles. Products from the previous cycle will be used as a template for the following cycles which causes the amount of DNA produced to double with each successive cycle. The cycles in a PCR run include DNA denaturing, primer annealing onto the DNA template, and primer extension along the DNA template (Newton and Graham, 1997).

Three main steps make up a typical PCR amplification, that is, the denaturation phase, the annealing phase, and the extension or polymerization phase. An initial pre-denaturation phase, which is prior to the denaturation phase, weakens the hydrogen bonds of the template DNA and it is usually carried out at 94°C. The pre-denaturation phase facilitates the denaturation of template DNA during the denaturation phase. In the next phase, the denaturation phase causes DNA templates to open up as single stranded templates for the annealing of complementary primers. Once the complementary sequences are located, the primers anneal onto their respective locations on the DNA

template. This is called the annealing phase, and the process of DNA replication begins. In the extension phase, dNTPs bind to the primers and this causes an extension that forms a new DNA template.

PCR reactions have been widely used for many different types of studies, and it involves a huge array of molecular markers, such as restriction fragment length polymorphism (RFLP), random amplification of polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), microsatellites, single nucleotide polymorphism (SNP) and short tandem repeat (STR). One of the methods of choice for the studies of taxonomy, population, and evolution of animals is the mitochondrial DNA marker (Lunt *et al.*, 1996).

2.12 Electrophoresis

Electrophoresis is a technique used to separate protein and fragments of DNA and RNA chains that differ in size, charge, and orientation. In molecular genetics research, this is a priceless and very important technique. Electrophoresis refers to the resolution of a charged molecule through a restrictive matrix under the influence of an electrical force. Charged molecules such as protein, DNA or RNA fragments are placed in an electric field and they migrate either towards the positive or negative pole according to their charge. The molecule with the greatest net charge will migrate more rapidly toward the pole of opposite polarity, even if the two molecules have approximately the same shape and mass (Klug *et al.*, 2006). During gel electrophoresis, molecules that are larger in size migrate slower due to resistance between the molecule and the gel matrix. The molecule's size can also be determined by the distance in which the molecule travels across the gel matrix (Blankenship, 2007).

When performing electrophoresis, the PCR products are loaded on the gel at the cathode (-) side of the gel because DNA is negatively charged. The opposite anode (+) end of the gel will cause the DNA to migrate towards it due to differences in charge. The distance the molecule travels depends on the molecular weight of the PCR product loaded onto the gel. Once electrophoresis is completed, the bands that represent the variously sized molecules are observed under autoradiography or by using a fluorescent dye.

2.13 DNA Sequencing and Phylogenetics

DNA sequencing refers to the sequencing of DNA nucleotide bases, which are adenine, guanine, cytosine, and thymine. Two methods have been developed for DNA sequencing; the Sanger sequencing method (Sanger *et al.*, 1977) with chain-termination using dideoxynucleotides (ddNTPs), and the Maxam-Gilbert sequencing method (Maxam and Gilbert, 1977) which involves chemical degradation of radio-labelled DNA fragments. Both the methods utilize high resolution polyacrylamide gel electrophoresis to separate the labelled fragments according to size and are read in a ladder-like fashion to determine the nucleotide order of a designated nucleotide sequence. The most common method used for routine DNA sequencing work is the Sanger sequencing method because it is an easy, fast, and reliable method (Graham and Hill, 2001).

Some of the most important applications of the DNA sequencing method include the sequencing of the human genome, in the Human Genome Project. Detailed knowledge of the sequences of genes and proteins are useful in advancing the medical and biotechnological field. Another useful application of gene and protein sequences

lies in the evolution field, whereby DNA sequencing and morphological data is combined to study evolutionary relatedness between different groups of organisms. This study of evolutionary relation among different groups of organisms through the use of molecular sequencing is called phylogenetics (Edwards and Cavalli-Sforza, 1964). The similarity and differences in DNA sequences can be used to infer evolutionary relationships among different organisms. It can be assumed that organisms with similar DNA sequences are more closely related compared to organisms that have different DNA sequences (Hedrick, 2011). The availability of a database of DNA sequences can be used to determine phylogenetic relationships between species or other taxa that are not clear based on other traits such as morphology.

Phylogenetic studies utilizing DNA sequencing have been carried out on fruit flies, especially utilizing the mitochondrial DNA region. Smith *et al.* (2003) carried out a phylogenetic relationship study among selected species of *Bactrocera* and *Dacus* fruit flies using mitochondrial DNA sequences and cladistic analysis. Yu *et al.* (2007) sequenced the complete mitochondrial genome of the oriental fruit fly, *Bactrocera dorsalis*.

2.14 Sequence Alignment

Sequence alignment is a method to identify similar regions in DNA, RNA, or protein sequences by comparing two or more biological sequences. Regions of similarity can be used to infer homology in function, structure, or even common evolutionary relationships between the sequences (Rosenberg, 2009). The importance of this technique encompasses the profiling of genetic diseases (Dreses-Werringloer *et al.*, 2008; Cheng *et al.*, 2009; Choi *et al.*, 2009), phylogenetic analysis (Han and Ro, 2009;

Zhang *et al.*, 2010), and identification and quantification of conserved regions (Kirkness *et al.*, 2003). Sequences from homologous molecules are arranged and lined up to maximise the similarity or to minimise the number of changes among the sequences. Alignments are easily made for individuals and coding genes that are closely related, for example sibling species (Hatadani *et al.*, 2009), however alignment becomes increasingly difficult with increasingly distant related taxa or from non-coding gene regions (Sinclair *et al.*, 2005). Up until 1989, sequence alignments were usually done manually by hand due to computational restrictions. However, with the introduction of “progressive sequence alignment”, it is possible to conduct multiple sequence alignments using a computer (Higgins and Sharp, 1988).

Phylogenetics and sequence alignment are closely related fields, in that phylogenetics makes use of aligned sequences in the construction of phylogenetic trees. The alignment of nucleotide or amino acid sequences implies that the individuals share a common ancestor (Pevsner, 2009). Based on the homology of the sequences, phylogenetic trees can be inferred and constructed.

2.15 Maximum Likelihood

Maximum likelihood (ML) is a parametric statistical method for inferring phylogenetic relationships by utilizing models of character evolution (in this case, nucleotide substitution). In ML, the correct model of nucleotide substitution for a particular set of queried nucleotide sequence alignments are established based on several assumptions, for example (1) all nucleotides are substituted equally likely, or (2) transversions and transitions of different nucleotide segments have different rates. Within the suitable substitution model and based on the queried sequence alignment,

ML determines the likelihood of every possible phylogenetic tree. The tree (topology) and branch length with the highest maximum likelihood is selected as the final phylogenetic tree (Felsenstein, 1981). When utilized with a suitable model, ML is theoretically immune to the long-branch attraction occasionally seen in maximum parsimony method when nucleotide character substitution rates are not the same (Felsenstein, 1978).

Some of the models employed in ML calculation include the general time-reversible DNA substitution model (GTR) (Lanave *et al.*, 1984), the Jukes-Cantor model (JC) (Jukes and Cantor, 1969), the Kimura 2-parameter model (K2P) (Kimura, 1980), and Felsenstein 1981 model (F81) (Felsenstein, 1981).

2.16 Bayesian Inference

Bayesian Inference (BI) is derived off ML (Huelsenbeck *et al.*, 2001) and utilizes algorithms to infer phylogeny. Utilizing queried sequence alignment and a nucleotide substitution model, along with a computational algorithm called the Markov chain Monte Carlo (MCMC) (Gilks *et al.*, 1996), an approximation is generated as the posterior probability (PP) of a given hypothetical phylogenetic tree. PP is the probability that a phylogenetic tree is correct, and is used to infer phylogenetic relationships within the queried group. The advantage of BI is that the method is able to manage large data sets with relative ease and a faster manner compared to the ML method.

2.17 Neighbor-Joining

Neighbor-Joining (NJ), a derivative of the UPGMA algorithm (Huson *et al.*, 2010), is a clustering method for deriving evolutionary trees by grouping together a set of taxa based on a matrix of pairwise evolutionary distances (Gascuel and Steel, 2006). Developed by Saitou and Nei (1987), the NJ algorithm is based on the minimum-evolution criterion of Sattath and Tversky (1977). Using the neighbourly methods of Sattath and Tversky (1977), only tree topologies are generated. With NJ, not only are the tree topologies generated, the branch lengths of the final tree are generated as well (Saitou and Nei, 1987). NJ is widely accepted as a tool for preliminary phylogenetic analysis (Zaslavsky and Tatusova, 2008) as it is a fast method for generating NJ trees even with large sets containing many haplotypes (Templeton, 2006), and is oftentimes used alongside ML and BI for phylogenetic analyses (Yang, 2006). NJ has had its fair share of usage for phylogenetic studies involving *Bactrocera* fruit flies (Hu *et al.*, 2008; Zhang *et al.*, 2010; Lim *et al.*, 2012; Liu *et al.*, 2012; Wan *et al.*, 2012)

2.18 Haplotype Network

A haplotype network is a phylogenetic network that is unrooted and consists of nodes which represent different haplotypes. Haplotypes which are closely related are joined by branches based on the degree of genetic differences. Haplotype networks are useful for visualizing genetic differences of groups of haplotypes based on the differences in DNA sequences. Haplotype networks are also able to identify ancestral haplotypes from which the other sequences most likely originated from. A popular

method to construct haplotype networks is the TCS program (Clement *et al.*, 2000) which is based on the concept of statistical parsimony (Huson *et al.*, 2010). TCS is widely used to study the genetic structure and diversity of organisms of different locations.

CHAPTER 3

MATERIALS AND METHODS

3.1 Host Fruit Collection

Host fruits were collected from two locations in Peninsular and East Malaysia. Fruits collected from the field were fruits that were observed to have puncture wounds and discolouration on the skin were taken back to the insect culture room in Institute of Biological Sciences, Universiti Malaya for hatching. A variety of fruits were collected for the purpose of the study of host fruit specificity, particularly fruits that are known to be host fruit targets for *Bactrocera carambolae* and *Bactrocera papayae*. Table 3.1 lists the locations of host fruit sampling.

Table 3.1: Host fruit collection, date of collection, location of origin, and species of host plant.

Location	Date of collection	Type of fruit collected	Species of Host Plant
Gua Sengkeli, Sarawak	13/12/10	Carambola	<i>Averrhoa carambola</i>
	10/01/11	Papaya	<i>Carica papaya</i>
Serdang, Selangor	09/03/11	Guava	<i>Psidium guajava</i>
	05/05/11	Jambu air	<i>Syzygium samarangense</i>
	05/05/11	Jambu madu	<i>Syzygium aquem</i>

Two related species of *Bactrocera* fruit flies, *Bactrocera umdbrosa* and *Bactrocera tau*, were trapped using cue-lure (4-[4-(acetyloxy) phenyl]-2-butanone) sex-attractant. These two species of fruit flies were collected from Petaling Jaya by wiping the sex-attractant onto the upper surface of a leaf. The fruit flies were then captured using plastic bags and specimen tubes, and were brought back to Universiti Malaya for

identification, freezing and storage. The *Bactrocera umbrosa* and *Bactrocera tau* specimens were collected to be utilized as outgroup specimens for the phylogenetic analyses.

3.2 Fruit Fly Hatching, Storage and Identification

Host fruits were placed in cages that were filled with moist soil. Each cage was filled with a single type of host fruit, and placed in the culture room. Emergence of larvae from host fruit and pupation was observed. After two weeks when the fruit flies have emerged from their puparium, they were captured using specimen tubes and frozen in a -20°C freezer for storage and identification.

3.3 DNA Extraction

3.3.1 i-Genomic CTB DNA Extraction Mini Kit

DNA extraction was performed using the i-Genomic CTB DNA Extraction Kit (iNtRON Biotechnology, Inc, South Korea) with several modifications to the standard protocol.

Two legs from each individual fruit fly were used. They were placed in a sterilized 1.5 ml Eppendorf tube and suspended in 50 µl of Buffer CG. Using a micropestle, the fruit fly leg samples were disrupted and homogenized for 20 minutes. Once the samples have been homogenized, 150 µl of Buffer CG, 3 µl of RNase A Solution, and 10 µl of Proteinase K were added into the sample tube and vortexed

rigorously to enable the solutions and sample to mix. The sample tube was then placed in a pre-heated waterbath at 65°C for one hour for the cell lysis step. To further assist lysis of the sample, inversion of the sample tube every two minutes is carried out. After the lysis step is completed, the sample tube was centrifuged to remove unlysed tissue particles and 150-180 µl of supernatant was carefully transferred into a new 1.5 ml Eppendorf tube.

For the DNA binding step, 250 µl of Buffer CB was added into the lysate and gently mixed by inverting the Eppendorf tube six times. After inversion, the mixture was spun down in a centrifuge to remove drops of mixture from the lid of the Eppendorf tube. Next, 250 µl of 80% ethanol was added to the lysate and gently mixed by inverting the Eppendorf tube six times. After inversion, the mixture was spun down. The whole mixture was then carefully pipetted into a spin column that has been inserted into a 2 ml collection tube, without wetting the rim. The spin column and collection tube was then centrifuged at 13,000 rpm for two minutes. The flow through and collection tube was discarded after centrifuge.

Prior to using the Buffer CW for the first time, 40 ml of ethanol was added into the solution. The spin column was then placed into a new 2.0 ml collection tube, and 700 µl of Buffer CW was added into the spin column and centrifuged at 13,000 rpm for two minutes. The flow through was discarded, and the collection tube was reused. The spin column was centrifuged again at 13,000 rpm for another two minutes. The flow through and collection tube were then discarded.

For the first elution step, the spin column was placed into a sterilized 1.5 ml Eppendorf tube and 50 µl of Buffer CE was added directly into the membrane of the spin column. The spin column was left to sit in room temperature for three minutes, and then it was centrifuged at 13,000 rpm for two minutes to elute. The first elution step

yields a higher final DNA concentration. For the second elution, the elution step was repeated again with the same spin column from the first elution step to obtain a lower final DNA concentration. Two sets of higher and lower concentration final DNA was obtained from the two elution steps. The DNA samples were then kept in a -20°C freezer.

3.4 PCR Amplification

PCR amplification of DNA samples were carried out in Applied Biosystems Veriti 96 Well Thermal Cycler (Applied Biosystems, USA). Forward and reverse primer pairs used in this study are listed in Table 3.2.

Table 3.2: Oligonucleotide primers used for polymerase chain reaction (PCR) amplification.

No.	Primer Name	Type of Primer	Sequence	Gene	Source
1.	UEA7	Forward	5'- TACAGTTGGAATAGACGTT GATAC-3'	Cytochrome oxidase I	Lunt <i>et al.</i> , 1996
2.	UEA10	Reverse	5'- TCCAATGCACTAATCTGCC ATATTA-3'	Cytochrome oxidase I	Lunt <i>et al.</i> , 1996
3.	C2-J-3549 (alias C2KD-F)	Forward	5'- CAAATTCGAATTTTAGTAA CAGC-3'	Cytochrome oxidase II	Simon <i>et al.</i> , 1994
4.	TD-N-3884 (alias C2KD-R)	Reverse	5'- TTAGTTTGACAWACTAATG TTAT-3'	Cytochrome oxidase II	Simon <i>et al.</i> , 1994
5.	CB-J-10933 (alias CB1)	Forward	5'- TATGTACTACCATGAGGAC AAATATC-3'	Cytochrome-b	Simon <i>et al.</i> , 1994
6.	CB-N-11367 (alias CB2)	Reverse	5'- ATTACACCTCCTAATTTAT TAGGAAT-3'	Cytochrome-b	Simon <i>et al.</i> , 1994

3.4.1 PCR Amplification of Cytochrome oxidase subunit I

For each PCR reaction, a 40 µl reaction volume was prepared, containing 4.0 µl of 10X PCR buffer (iNtRON Biotechnology, Inc, South Korea), 5.0 µl of 25mM MgCl₂ (iNtRON Biotechnology, Inc, South Korea), 1.0 µl of 10mM dNTP mixture (iNtRON Biotechnology, Inc, South Korea), 1.0 µl of 10 µM forward and reverse primers, 0.4 µl of i-Taq DNA polymerase (iNtRON Biotechnology, Inc, South Korea), 25-40 ng of DNA template and UHQ (ultra high quality) water was added until the final volume of 40 µl.

The thermal cycling program begins with the initial denaturation step at 94°C for three minutes, followed by 40 cycles at the following parameters: 95°C for one minute, 50°C for one minute, and 72°C for one minute and 30 seconds. The final extension step was at 72°C for seven minutes, and finally the hold step at 4°C.

3.4.2 PCR Amplification of Cytochrome oxidase subunit II

For each PCR reaction, a 40 µl reaction volume was prepared, containing 4.0 µl of 10X PCR buffer (iNtRON Biotechnology, Inc, South Korea), 5.0 µl of 25mM MgCl₂ (iNtRON Biotechnology, Inc, South Korea), 1.0 µl of 10mM dNTP mixture (iNtRON Biotechnology, Inc, South Korea), 1.0 µl of 10 µM forward and reverse primers, 0.4 µl of i-Taq DNA polymerase (iNtRON Biotechnology, Inc, South Korea), 25-40 ng of DNA template and UHQ (ultra high quality) water was added until the final volume of 40 µl.

The thermal cycling program begins with the initial denaturation step at 95°C for two minutes, followed by 40 cycles at the following parameters: 94°C for 45 seconds, 44°C for 45 seconds, and 72°C for 45 seconds. The final extension step was at 72°C for seven minutes, and finally the hold step at 4°C.

3.4.3 PCR Amplification of Cytochrome-b

For each PCR reaction, a 40 µl reaction volume was prepared, containing 4.0 µl of 10X PCR buffer (iNtRON Biotechnology, Inc, South Korea), 5.0 µl of 25mM MgCl₂ (iNtRON Biotechnology, Inc, South Korea), 1.0 µl of 10mM dNTP mixture (iNtRON Biotechnology, Inc, South Korea), 1.0 µl of 10 µM forward and reverse primers, 0.4 µl of i-Taq DNA polymerase (iNtRON Biotechnology, Inc, South Korea), 25-40 ng of DNA template and UHQ (ultra high quality) water was added until the final volume of 40 µl.

The thermal cycling program begins with the initial denaturation step at 94°C for 2 minutes, followed by 40 cycles at the following parameters: 94°C for 1 minute, 45°C for 1 minute, and 72°C for 1 minute and 30 seconds. The final extension step was at 72°C for 5 minutes, and finally the hold step at 4°C.

3.5 Electrophoresis

3.5.1 Preparation of Agarose Gel

1% agarose gel was prepared by adding powdered agarose to 1X TAE buffer. SYBR Safe DNA Gel Stain (Invitrogen, USA) was then added in a ratio of 1:10,000 into the 1X TAE buffer to stain the agarose gel. The mixture was then heated in a microwave for 1 minute and then poured into a casting tray affixed with a comb. The agarose gel was left to harden at room temperature for 30 minutes and was then placed in a buffer chamber, where it was then submerged in 1X TAE buffer.

3.5.2 Electrophoresis of PCR Products

Amplification of PCR products were confirmed using standard horizontal submarine gel electrophoresis. Four μl of PCR product was mixed with 1 μl of 6X loading buffer. The mixture was then loaded into a 1% agarose gel that was submerged in 1X TAE buffer. A 100 base pair extended molecular weight size marker (Bioron, Germany) was added as the DNA size standard and to track the migration of the PCR products. The agarose gel was subjected to electrophoresis at 70V for 45 minutes or until the blue dye has migrated a distance judged to be enough for separation of DNA fragments. The agarose gel was then visualized using DigiDoc-It Imaging System (UVP, LLC, USA) and banding patterns were photographed for further analysis and as a permanent record.

3.6 DNA Purification and Sequencing

Purification was carried out using LaboPass™ PCR purification kit (CosmoGenetech, South Korea) and sequencing of purified PCR products were sent to and carried out by First Base Laboratories Sdn Bhd.

3.7 Phylogenetic Analysis

The raw DNA sequences were edited using ChromasPro version 1.42 (Technelysium Pty Ltd, Australia). The sequences were then preliminarily aligned using ClustalX version 2.0.8 (Larkin *et al.*, 2007) and subsequently manually aligned.

The aligned sequences were then subjected to neighbour-joining (NJ) analysis using PAUP* 4.0b10 (Swofford, 2002). To determine the NJ bootstrap values, the Kimura's two-parameter model of substitution (K2P distance) evolution model with 1000 replications was utilized.

To determine the best fit nucleotide substitution model for maximum likelihood (ML) and bayesian inference (BI), Kakusan v.3 (Tanabe, 2007) was utilized to generate suitable output model files for use in their respective analyses.

Maximum likelihood (ML) analysis was carried out using Treefinder version October 2008 (Jobb *et al.*, 2004), with 1000 bootstrap replicates and utilizing the corrected Akaike Information Criterion (AIC) (Akaike, 1973; Shono, 2000).

Bayesian inference (BI) was performed using MrBayes version 3.1.2 (Huelsenbeck and Ronquist, 2001). Best fit models were evaluated using the Bayesian

Information Criterion (BIC) and phylogeny inferred using the GTR model with gamma distribution. Two million Markov chain Monte Carlo (MCMC) generations were run; while stabilization of the log likelihood scored was monitored by calculating the convergence diagnostics every 1000th generation. A 50% majority rule consensus tree was generated from the sampled trees after discarding the first 20%. Log likelihood values reached a plateau before 400 trees were sampled for all four analyses on the three molecular markers.

Phylogenetic analyses were conducted separately on the aligned COI, COII, and *cytb* DNA sequences datasets. A combined dataset of COI, COII, and *cytb* DNA sequences was also analyzed using phylogenetic methods.

3.8 GenBank DNA Sequences

DNA sequences of members of the *Bactrocera dorsalis* complex as well as other *Bactrocera* species were downloaded from GenBank in order to compare their genetic diversity with the DNA sequences obtained from this study. The DNA sequence of an outgroup organism was also downloaded to root the phylogenetic trees. Table 3.3 lists the the downloaded GenBank sequences by molecular marker.

Table 3.3: Species used for bioinformatics analysis.

Molecular Marker	Species	GenBank Accession Numbers	Country of Origin	
COI	<i>Bactrocera papayae</i>	AY398756	Thailand	
	<i>Bactrocera papayae</i>	AY053513	Unknown	
	<i>Bactrocera papayae</i>	AB192436	Bandung	
	<i>Bactrocera dorsalis</i>	EU076665	Taiwan	
	<i>Bactrocera carambolae</i>	DQ006872	Unknown	
	<i>Bactrocera dorsalis</i>	AY398752	Malaysia	
	<i>Bactrocera dorsalis</i>	JN644036	China	
	<i>Bactrocera dorsalis</i>	AY053507	Unknown	
	<i>Bactrocera papayae</i>	DQ917578	Malaysia	
	<i>Bactrocera philippinensis</i>	AB192439	Japan	
	<i>Bactrocera carambolae</i>	AY053509	Unknown	
	<i>Bactrocera kanchanaburi</i>	AY274169	Unknown	
	<i>Bactrocera carambolae</i>	FJ903495	Malaysia	
	<i>Bactrocera carambolae</i>	AB192420	Bandung	
	<i>Bactrocera occipitalis</i>	AY398754	Philippines	
	<i>Bactrocera pyrifoliae</i>	AY053514	Unknown	
	<i>Bactrocera kandiensis</i>	AB192431	Sri Lanka	
	<i>Bactrocera tryoni</i>	AB192442	Australia	
	<i>Bactrocera arcae</i>	AY053508	Unknown	
	COII	<i>Bactrocera papayae</i>	DQ917578	Malaysia
		<i>Bactrocera philippinensis</i>	NC_009771	Philippines
		<i>Bactrocera papayae</i>	NC_009770	Malaysia
		<i>Bactrocera carambolae</i>	EF014414	Japan
		<i>Bactrocera dorsalis</i>	NC_008748	China
<i>Bactrocera carambolae</i>		NC_009772	Japan	
<i>Bactrocera dorsalis</i>		DQ845759	China	
<i>Bactrocera philippinensis</i>		DQ995281	Philippines	
<i>Bactrocera dorsalis</i>		HQ260727	USA	
<i>Bactrocera dorsalis</i>		HQ260726	USA	
<i>Bactrocera dorsalis</i>		JQ671182	Unknown	
<i>Bactrocera dorsalis</i>		EU926791	USA	
<i>Bactrocera dorsalis</i>		AB090271	Taiwan	
<i>Bactrocera dorsalis</i>		DQ917577	China	
<i>Bactrocera dorsalis</i>		FJ172048	Thailand	
<i>Bactrocera cacuminata</i>		AY037413	Australia	
<i>Bactrocera cacuminata</i>		JQ671162	Unknown	
<i>Bactrocera dorsalis</i>		AB090272	Taiwan	
<i>Bactrocera dorsalis</i>		JN578415	Unknown	
<i>Bactrocera carambolae</i>		AY037410	Malaysia	
<i>Bactrocera occipitalis</i>		AY037429	Philippines	
<i>Bactrocera papayae</i>		AY037420	Australia	
<i>Bactrocera dorsalis</i>		AY037415	USA	
<i>Bactrocera cognata</i>		AY037426	Philippines	
<i>Bactrocera caryeae</i>		AY037427	India	
<i>Bactrocera kandiensis</i>		AY037428	Sri Lanka	
<i>Bactrocera arecae</i>		AY037411	Thailand	
<i>Bactrocera tryoni</i>		GQ255823	Australia	
<i>Bactrocera cucurbitae</i>		FJ172050	Bangladesh	
<i>Bactrocera endiandrae</i>		JQ671169	Unknown	
cytb		<i>Bactrocera carambolae</i>	AF033911	Malaysia
		<i>Bactrocera papayae</i>	AF033912	Malaysia
		<i>Bactrocera papayae</i>	AF033198	Australia
	<i>Bactrocera dorsalis</i>	AF033914	Tahiti	
	<i>Bactrocera dorsalis</i>	JF521028	China	
	<i>Bactrocera dorsalis</i>	JF521037	China	
	<i>Bactrocera papayae</i>	DQ006903	Unknown	
	<i>Bactrocera papayae</i>	DQ006902	Unknown	
	<i>Bactrocera oleae</i>	GU108463	Pakistan	
	COI, COII, cytb	<i>Drosophila melanogaster</i>	NC_001709	Unknown

3.9 Haplotype Network Reconstruction

TCS version 1.21 (Clement *et al.*, 2000) was used to reconstruct the haplotype network of aligned COI, COII, and *cytb* DNA sequences. The fix connection limit option of the TCS software was used to determine the number of mutational steps required to link the sequences with >95% confidence. Four separate haplotype networks were generated based on four datasets: 1) all the COI DNA sequences obtained in this study; 2) all the COII DNA sequences obtained in this study; 3) all the *cytb* DNA sequences obtained in this study; and 4) combined sequences of COI, COII, and *cytb* obtained from this study.

CHAPTER 4

RESULTS

4.1 *Bactrocera carambolae* and *Bactrocera papayae* Specimens Collected

The two main fruit flies, *Bactrocera carambolae* and *Bactrocera papayae* were identified based on their wing and abdominal patterns, as described by Drew and Hancock (1994). Table 4.1 lists the specimen labels, species, and the host fruits.

Table 4.1: Specimen list and host fruits origin.

Species	Host Fruit	Specimen
<i>Bactrocera carambolae</i>	Carambola	SS1, SS2, SS3, SS4
	Guava	GS1, GS2, GS3, GS4
	Papaya	PS1, PS2, PS3, PS4
	Jambu Madu	JS11, JS12, JS13, JS14
	Jambu Air	JS21, JS22, JS23, JS24
<i>Bactrocera papayae</i>	Carambola	SS5, SS6, SS7, SS8
	Guava	GS5, GS6, GS7, GS8
	Papaya	PS5, PS6, PS7, PS8
	Jambu Madu	JS15, JS16, JS17, JS18
	Jambu Air	JS25, JS26, JS27, JS28

4.2 Polymerase Chain Reaction

4.2.1 Primer Temperature Gradient Screening

The three primer pairs selected for phylogenetic analysis – namely UEA7 and UEA10; C2KD-F and C2KD-R; and CB1 and CB2 – were first relegated to a temperature gradient screening to determine the clearest and brightest band to be used as the annealing temperature. Six replicates with an annealing temperature gradient were used, each containing a reaction volume of 20 μ l. The selected annealing temperatures for each primer set were: UEA7-UEA10, 50°C; C2KD-F-C2KD-R, 44°C; and CB1-CB2, 45°C. Figures 4.1 to 4.3 shows the primer temperature gradient screening for all three primer pairs used in this study.

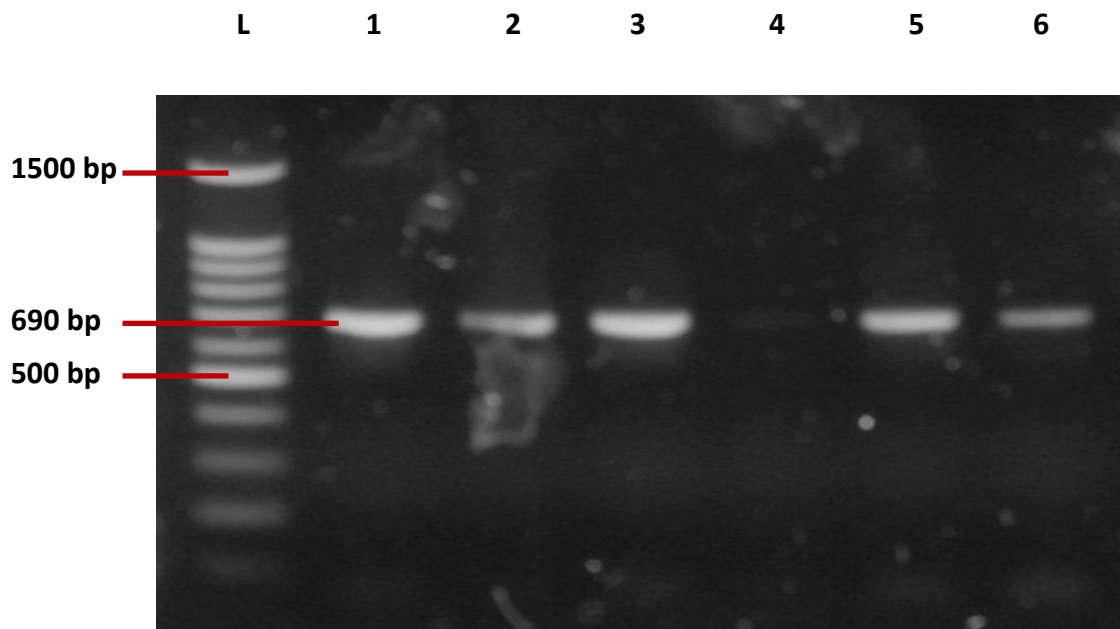


Figure 4.1: COI temperature gradient optimization. Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 1-6: products of amplification at a temperature gradient of 48-53°C. Lane 3 at 50°C was selected for amplification.

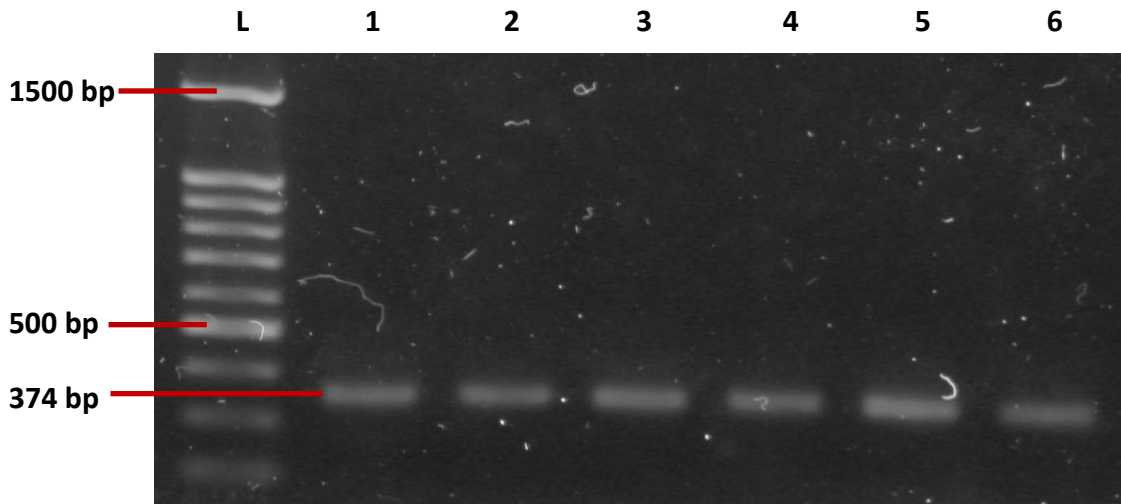


Figure 4.2: COII temperature gradient optimization. Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 1-6: products of amplification at a temperature gradient of 41-46°C. Lane 4 at 44°C was selected for amplification.

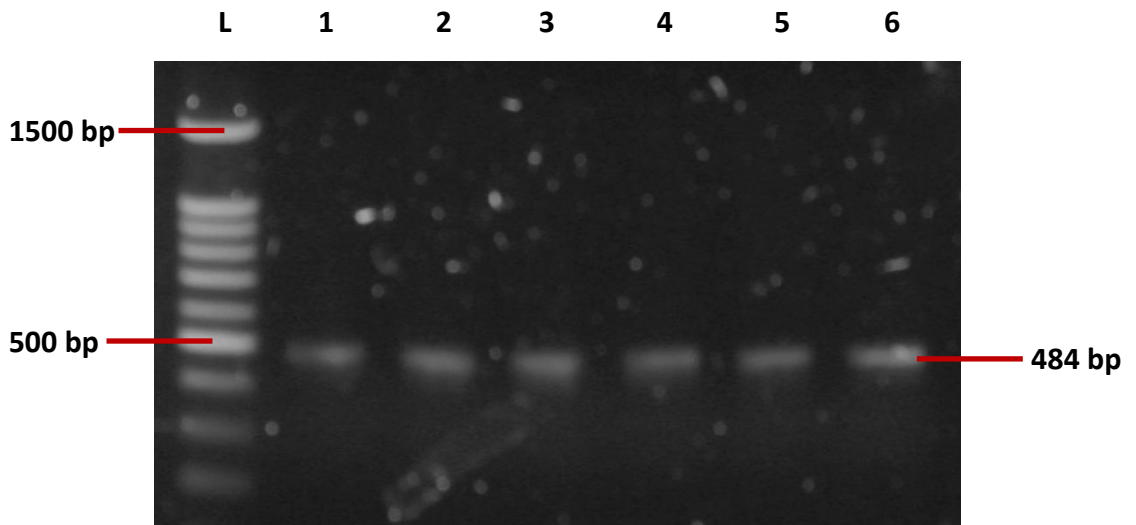


Figure 4.3: *Cytb* temperature gradient optimization. Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 1-6: products of amplification at a temperature gradient of 43-48°C. Lane 3 at 45°C was selected for amplification.

4.2.2 Cytochrome Oxidase I Based Primer

Primers utilized for amplification of the cytochrome c oxidase subunit I (COI) gene were UEA7 and UEA10 (Luntet *al.*, 1996). Identification of COI bands was determined with reference to a 100 base pair extended molecular weight size marker (Bioron, Germany) and by comparing the expected base pair size for the primer. The primer was expected to amplify a 690 base pair fragment of the cytochrome c oxidase subunit I gene.

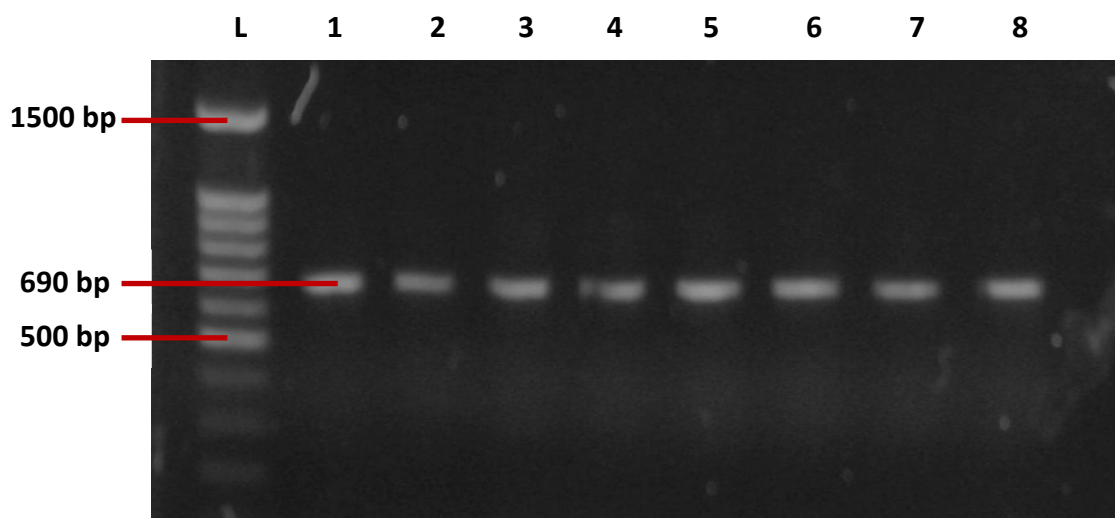


Figure 4.4: COI banding profile of fruit fly samples from Serdang, collected from papaya host fruit. Amplification of COI fragment was expected to be around 690 base pairs in length. Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 1-4: *Bactrocera carambolae* samples. Lanes 5-8: *Bactrocera papayae* samples.

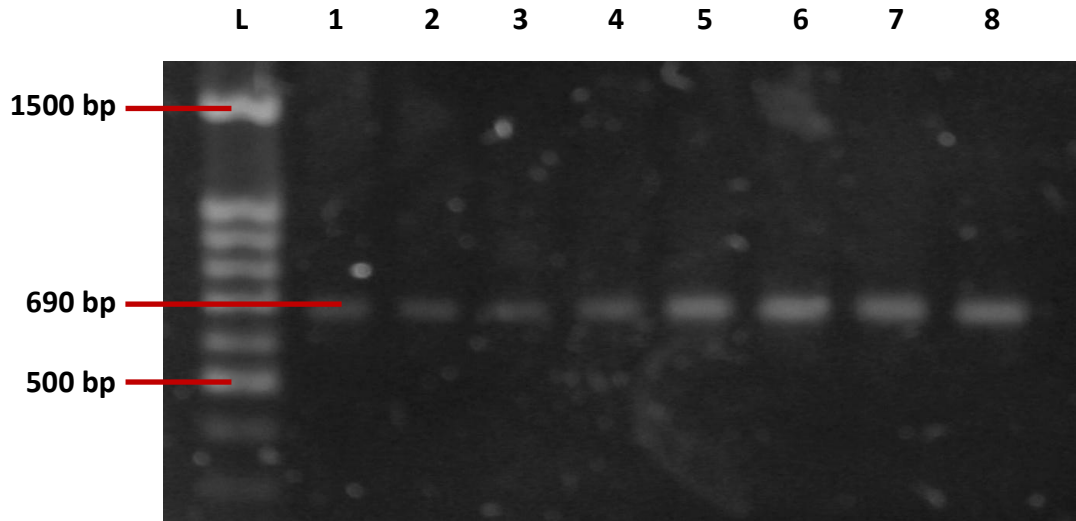


Figure 4.5: COI banding profile of fruit fly samples from Serdang, collected from guava host fruit. Amplification of COI fragment was expected to be around 690 base pairs in length. Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 1-4: *Bactrocera carambolae* samples. Lanes 5-8: *Bactrocera papayae* samples.

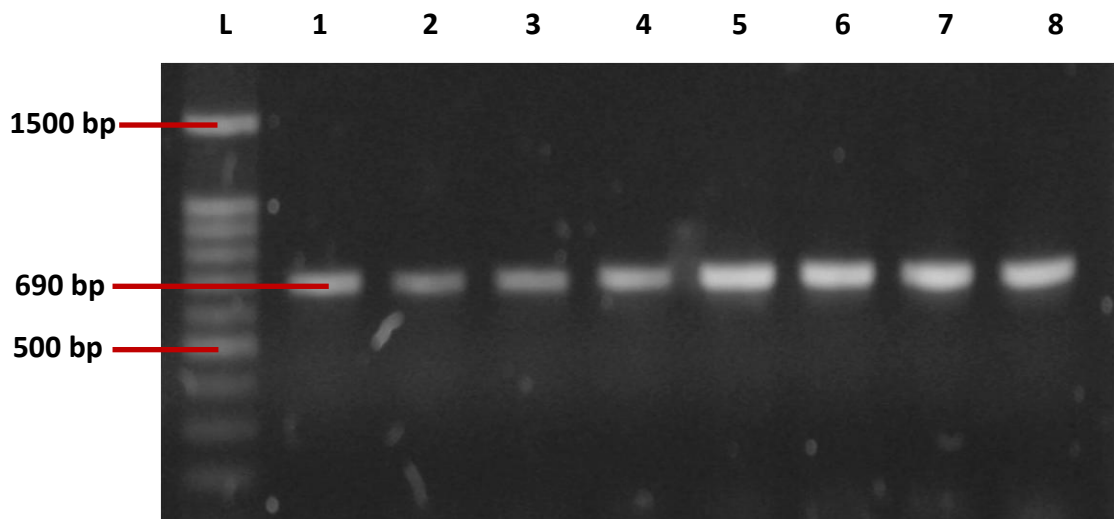


Figure 4.6: COI banding profile of fruit fly samples from Serdang, collected from jambu madu host fruit. Amplification of COI fragment was expected to be around 690 base pairs in length. Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 1-4: *Bactrocera carambolae* samples. Lanes 5-8: *Bactrocera papayae* samples.

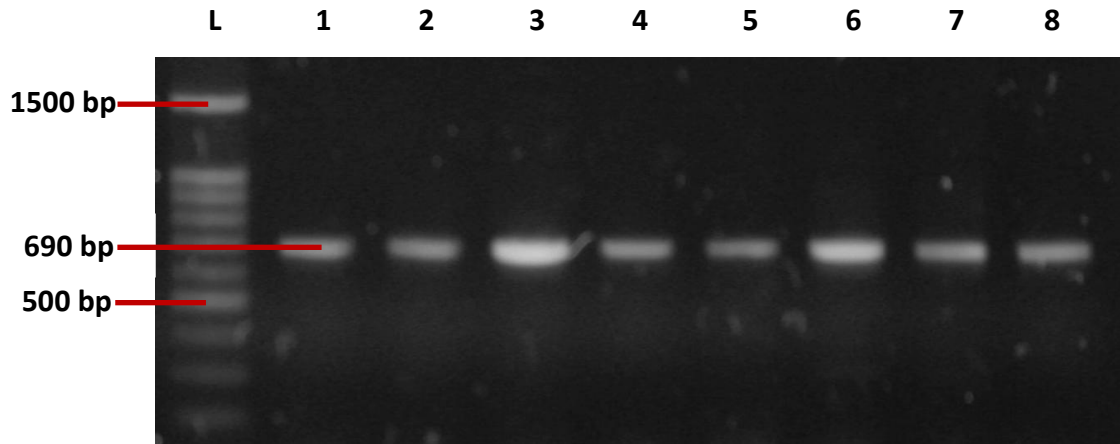


Figure 4.7: COI banding profile of fruit fly samples from Serdang, collected from jambu air host fruit. Amplification of COI fragment was expected to be around 690 base pairs in length. Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 1-4: *Bactrocera carambolae* samples. Lanes 5-8: *Bactrocera papayae* samples.

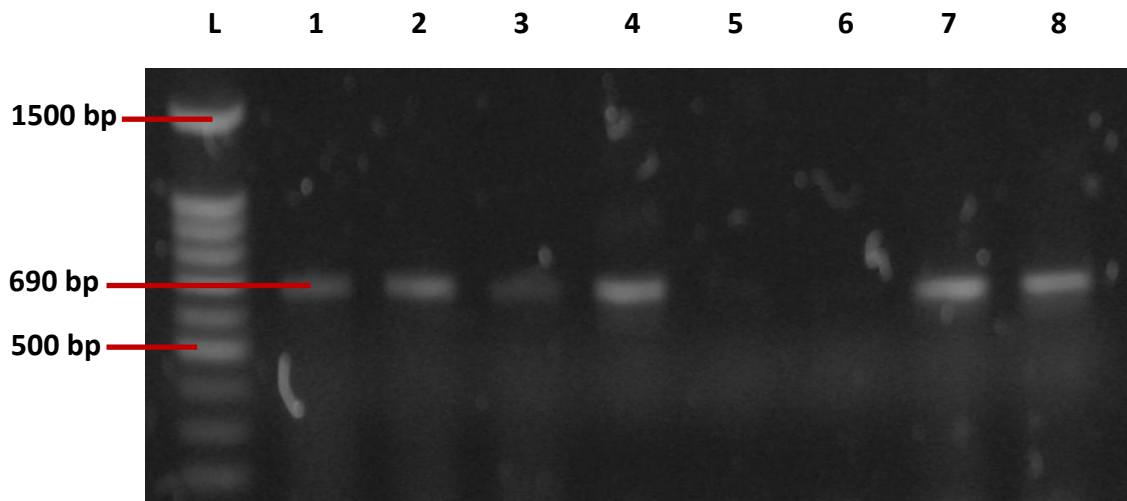


Figure 4.8: COI banding profile of fruit fly samples from Sarawak, collected from starfruit host fruit. Amplification of COI fragment was expected to be around 690 base pairs in length. Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 1-4: *Bactrocera carambolae* samples. Lanes 5-6: Null amplifications. Lanes 7-8: *Bactrocera papayae* samples.

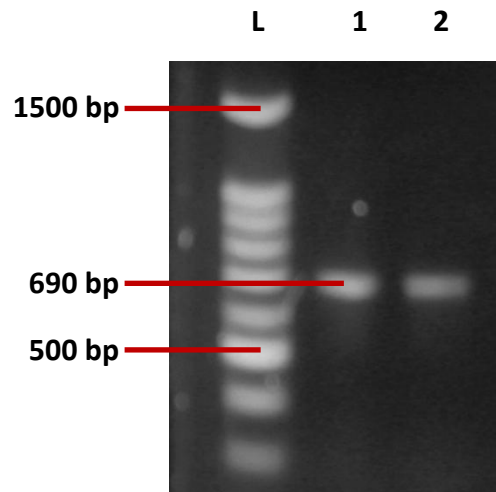


Figure 4.9: COI banding profile of fruit fly samples from Sarawak, collected from starfruit host fruit. Amplification of COI fragment was expected to be around 690 base pairs in length. Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 5-6: *Bactrocera papayae* samples.

4.2.3 Cytochrome Oxidase II Based Primer

Primers utilized for amplification of the cytochrome c oxidase subunit II (COII) gene were C2-J-3549 and TD-N-3884 (Simon *et al.*, 1994). Identification of COII bands was determined with reference to a 100 base pair extended molecular weight size marker (Bioron, Germany) and by comparing the expected base pair size for the primer. The primer was expected to amplify a 374 base pair fragment of the cytochrome c oxidase subunit II gene.

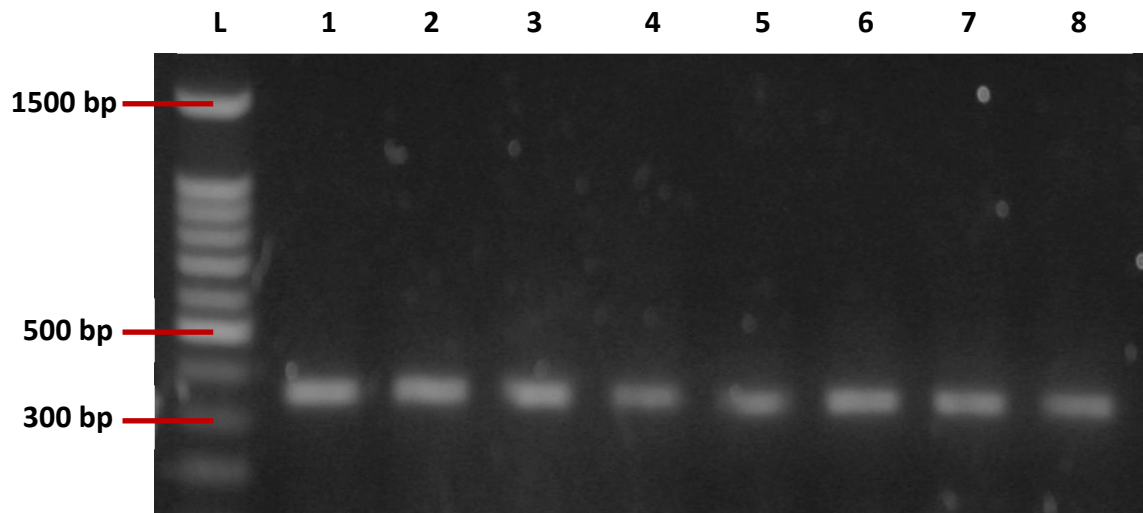


Figure 4.10: COII banding profile of fruit fly samples from Serdang, collected from papaya host fruit. Amplification of COII fragment was expected to be around 374 base pairs in length. Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 1-4: *Bactrocera carambolae* samples. Lanes 5-7: *Bactrocera papayae* samples.

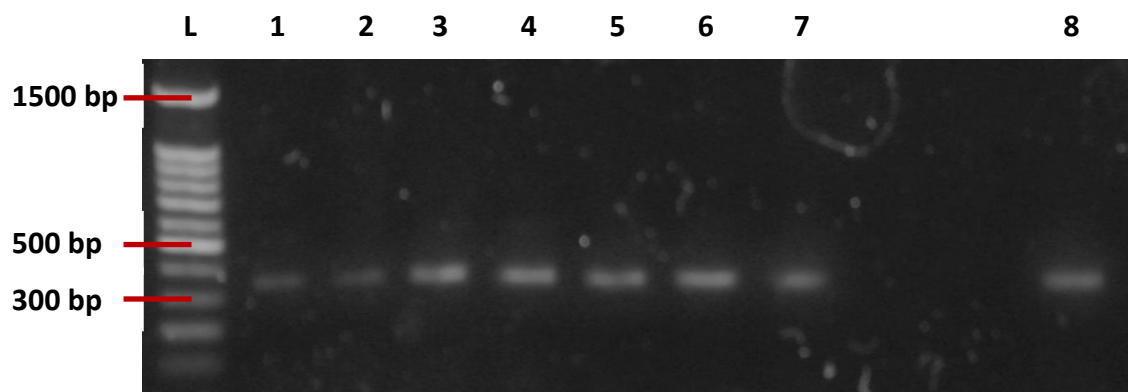


Figure 4.11: COII banding profile of fruit fly samples from Serdang, collected from guava host fruit. Amplification of COII fragment was expected to be around 374 base pairs in length. Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 1-4: *Bactrocera carambolae* samples. Lanes 5-7: *Bactrocera papayae* samples.

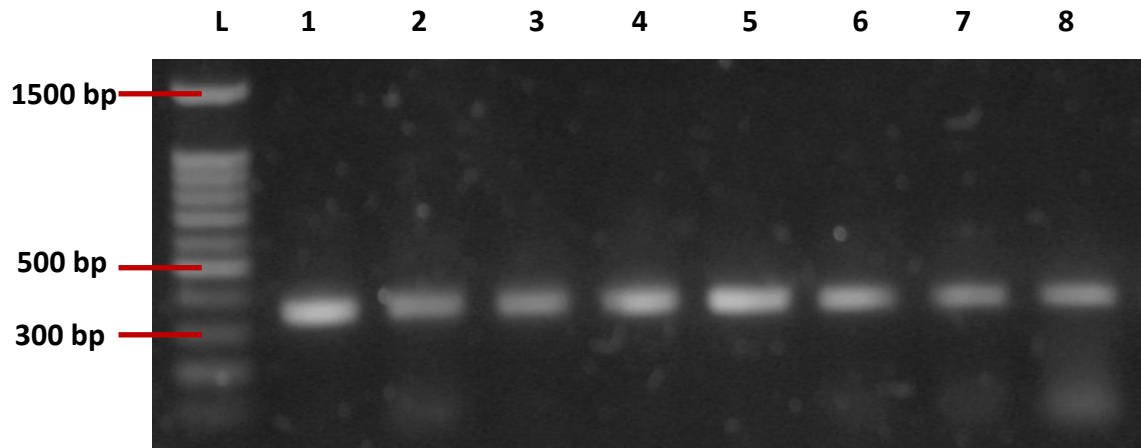


Figure 4.12: COII banding profile of fruit fly samples from Sarawak, collected from starfruit host fruit. Amplification of COII fragment was expected to be around 374 base pairs in length. Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 1-4: *Bactrocera carambolae* samples. Lanes 5-7: *Bactrocera papayae* samples.

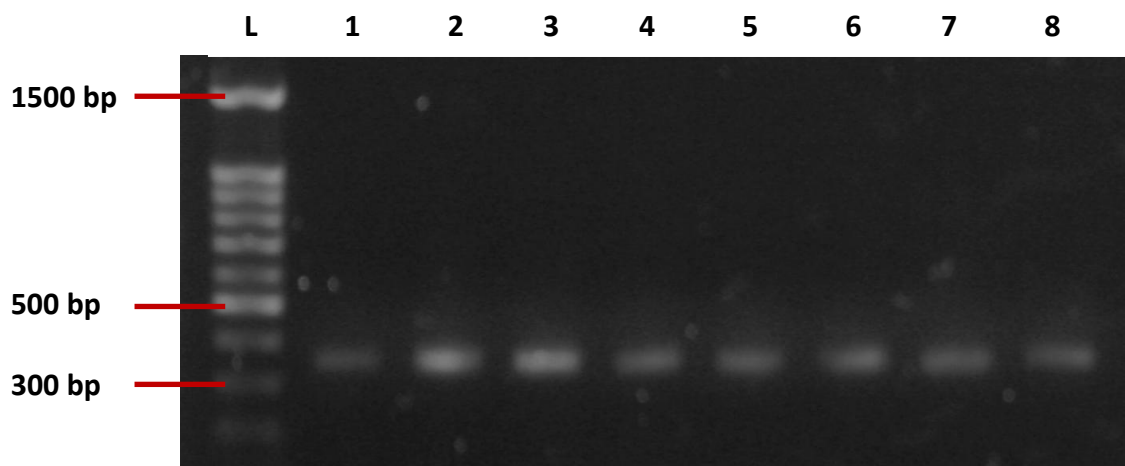


Figure 4.13: COII banding profile of fruit fly samples from Serdang, collected from jambu madu host fruit. Amplification of COII fragment was expected to be around 374 base pairs in length. Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 1-4: *Bactrocera carambolae* samples. Lanes 5-7: *Bactrocera papayae* samples.

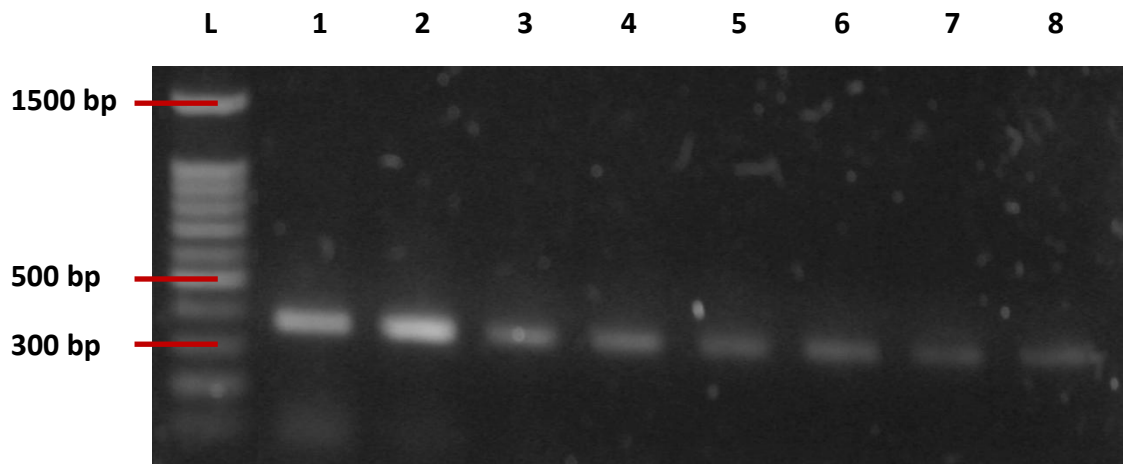


Figure 4.14: COII banding profile of fruit fly samples from Serdang, collected from jambu air host fruit. Amplification of COII fragment was expected to be around 374 base pairs in length. Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 1-4: *Bactrocera carambolae* samples. Lanes 5-7: *Bactrocera papayae* samples.

4.2.4 Cytochrome-b Based Primer

Primers utilized for amplification of the cytochrome-b (*cytb*) gene were CB-J-10933 and CB-N-11367 (Simon *et al.*, 1994). Identification of *cytb* bands was determined with reference to a 100 base pair extended molecular weight size marker (Bioron, Germany) and by comparing the expected base pair size for the primer. The primer was expected to amplify a 484 base pair fragment of the cytochrome-b gene.

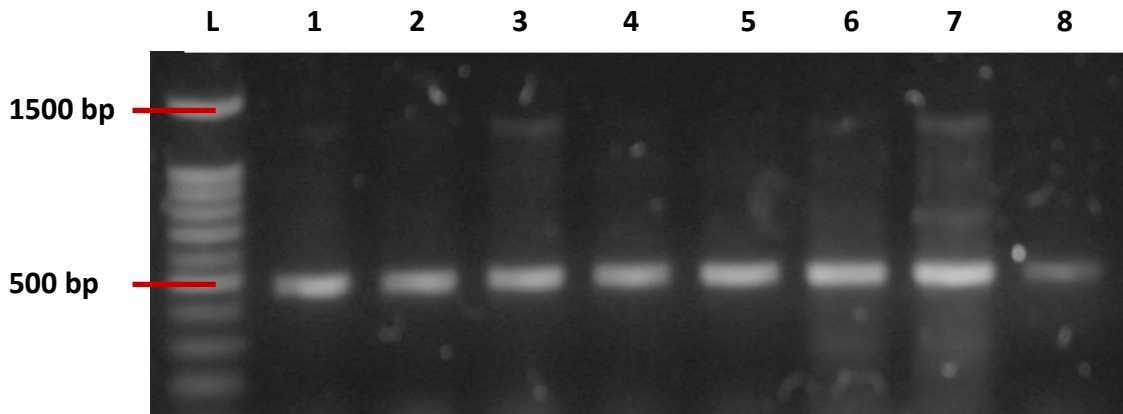


Figure 4.15: *Cytb* banding profile of fruit fly samples from Serdang, collected from papaya host fruit. Amplification of COI fragment was expected to be around 484 base pairs in length. Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 1-4: *Bactrocera carambolae* samples. Lanes 5-8: *Bactrocera papayae* samples.

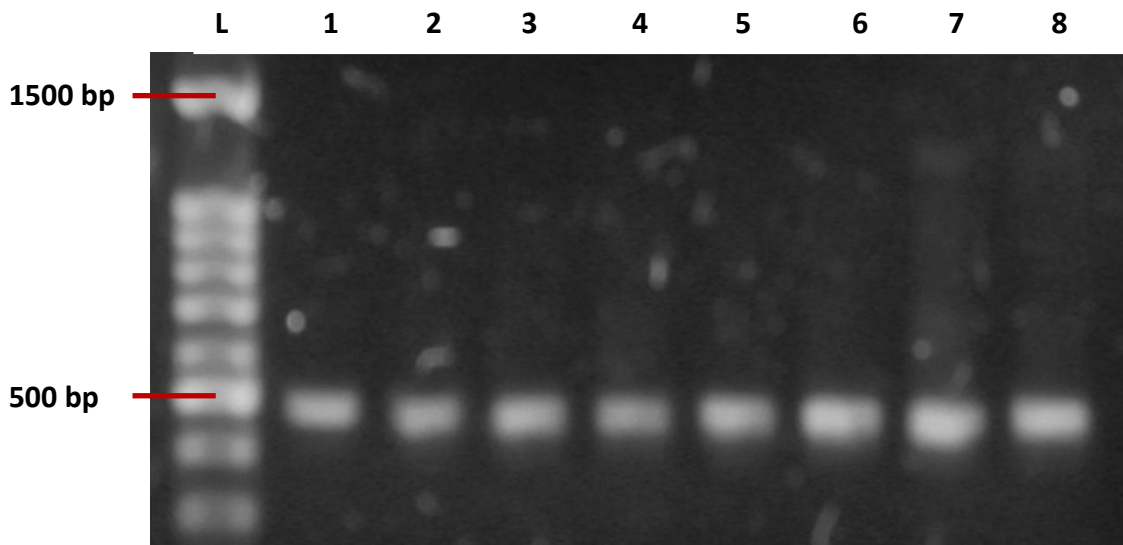


Figure 4.16: *Cytb* banding profile of fruit fly samples from Serdang, collected from guava host fruit. Amplification of COI fragment was expected to be around 484 base pairs in length. Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 1-4: *Bactrocera carambolae* samples. Lanes 5-8: *Bactrocera papayae* samples.

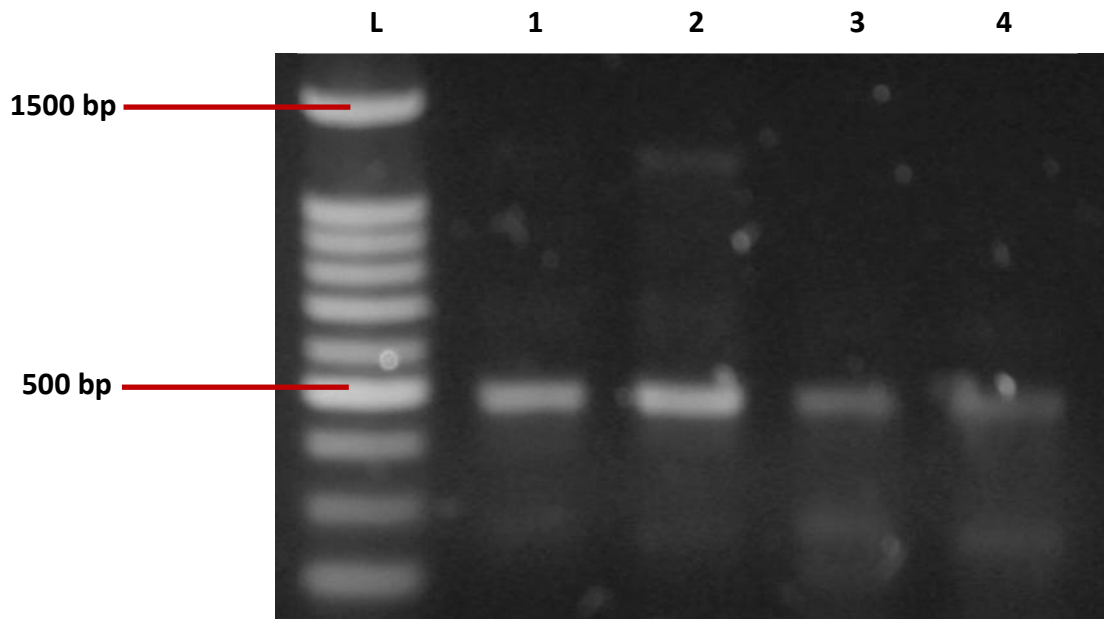


Figure 4.17: *Cytb* banding profile of fruit fly samples from Sarawak, collected from starfruit host fruit. Amplification of COI fragment was expected to be around 484 base pairs in length. Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 1-4: *Bactrocera carambolae* samples.

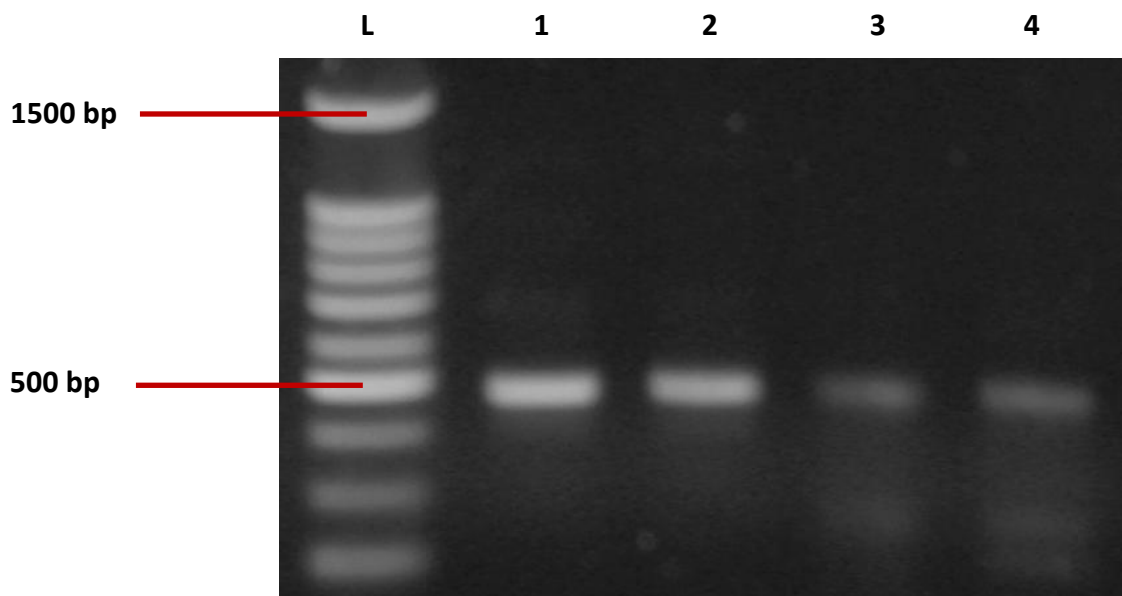


Figure 4.18: *Cytb* banding profile of fruit fly samples from Sarawak, collected from starfruit host fruit. Amplification of COI fragment was expected to be around 484 base pairs in length. Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 1-4: *Bactrocera papayae* samples.

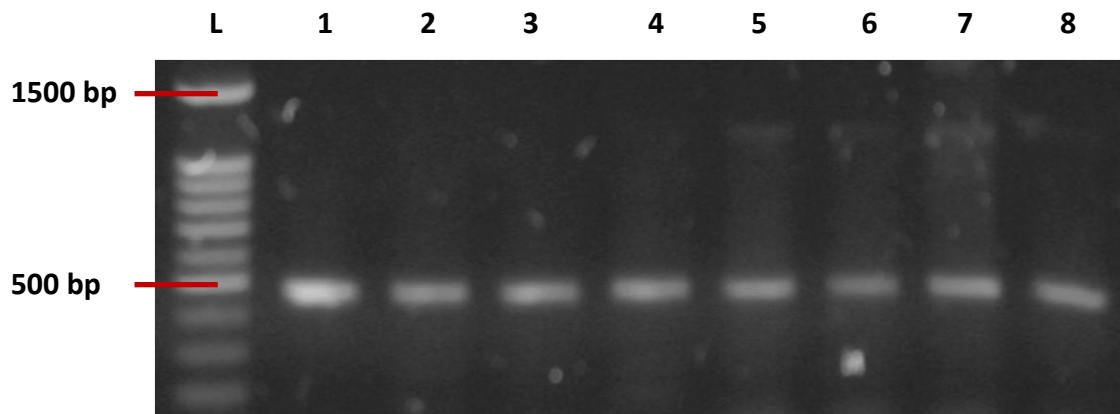


Figure 4.19: *Cytb* banding profile of fruit fly samples from Serdang, collected from jambu madu host fruit. Amplification of COI fragment was expected to be around 484 base pairs in length. Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 1-4: *Bactrocera carambolae* samples. Lanes 5-8: *Bactrocera papayae* samples.

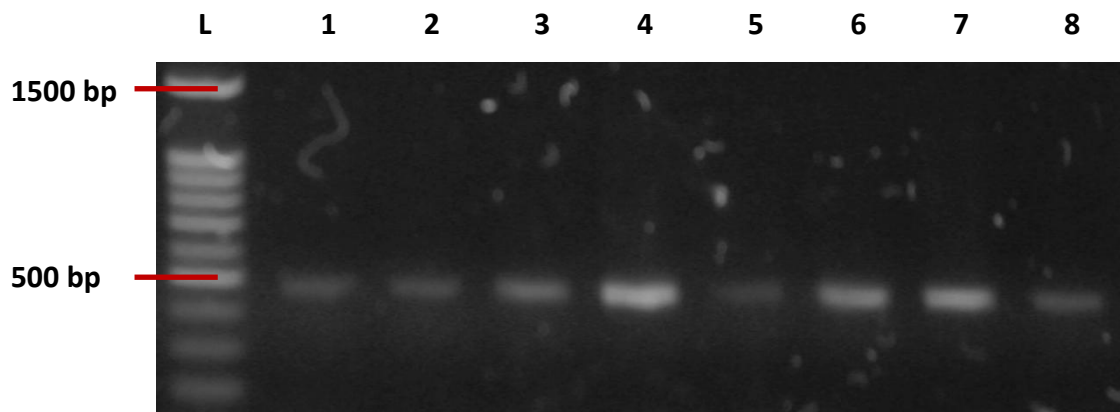


Figure 4.20: *Cytb* banding profile of fruit fly samples from Serdang, collected from jambu air host fruit. Amplification of COI fragment was expected to be around 484 base pairs in length. Lane L: 100 base pair extended molecular weight size marker (Bioron, Germany). Lanes 1-4: *Bactrocera carambolae* samples. Lanes 5-8: *Bactrocera papayae* samples.

4.3 Sequence Alignment

The COI data set consisted of an aligned data set of 536 base pairs; with 79 variable characters and 110 phylogenetically informative characters. The COII data set consisted of an aligned data set of 246 base pairs; with 31 variable characters and 48 phylogenetically informative characters. The *cytb* data set consisted of an aligned data set of 474 base pairs; with 94 variable characters and 141 phylogenetically informative characters. The combined sequences of COI, COII, and *cytb* data set consisted of an aligned data set of 1256 base pairs; with 222 variable characters and 178 phylogenetically informative characters.

4.4 Phylogeny Based On Cytochrome Oxidase Subunit I Gene

4.4.1 Maximum-Likelihood

Based on the COI ML tree (Figure 4.21), three major groups were observed. The first group consisted of *B. carambolae* and *B. papayae* and other *B. dorsalis* complex species such as *B. dorsalis*, *B. philippinensis*, *B. kanchanaburi*, and *B. occipitalis* and *B. pyrifoliae* as the basal species. The second group consisted of other *B. dorsalis* complex species such as *B. kandiensis*, and *B. tryoni*. The third group consisted of *B. umbrosa*, *B. arecae*, and *B. tau*.

The first group consisted of two subgroups; the first subgroup consisted of fruit flies hatched from host fruits collected from Serdang which is supported with moderate bootstrap value of 76.1%. The first subgroup also consisted of clades that contain mixtures of *B. carambolae* and *B. papayae*, as well as *B. dorsalis*, *B. philippinensis*, and *B. kanchanaburi*. No distinct clades were formed based on the host fruits the fruit flies were hatched from. The second subgroup contained three fruit flies hatched from carambola host fruits collected from Sarawak and *B. occipitalis* as the basal species which is supported with high bootstrap value of 93.5%

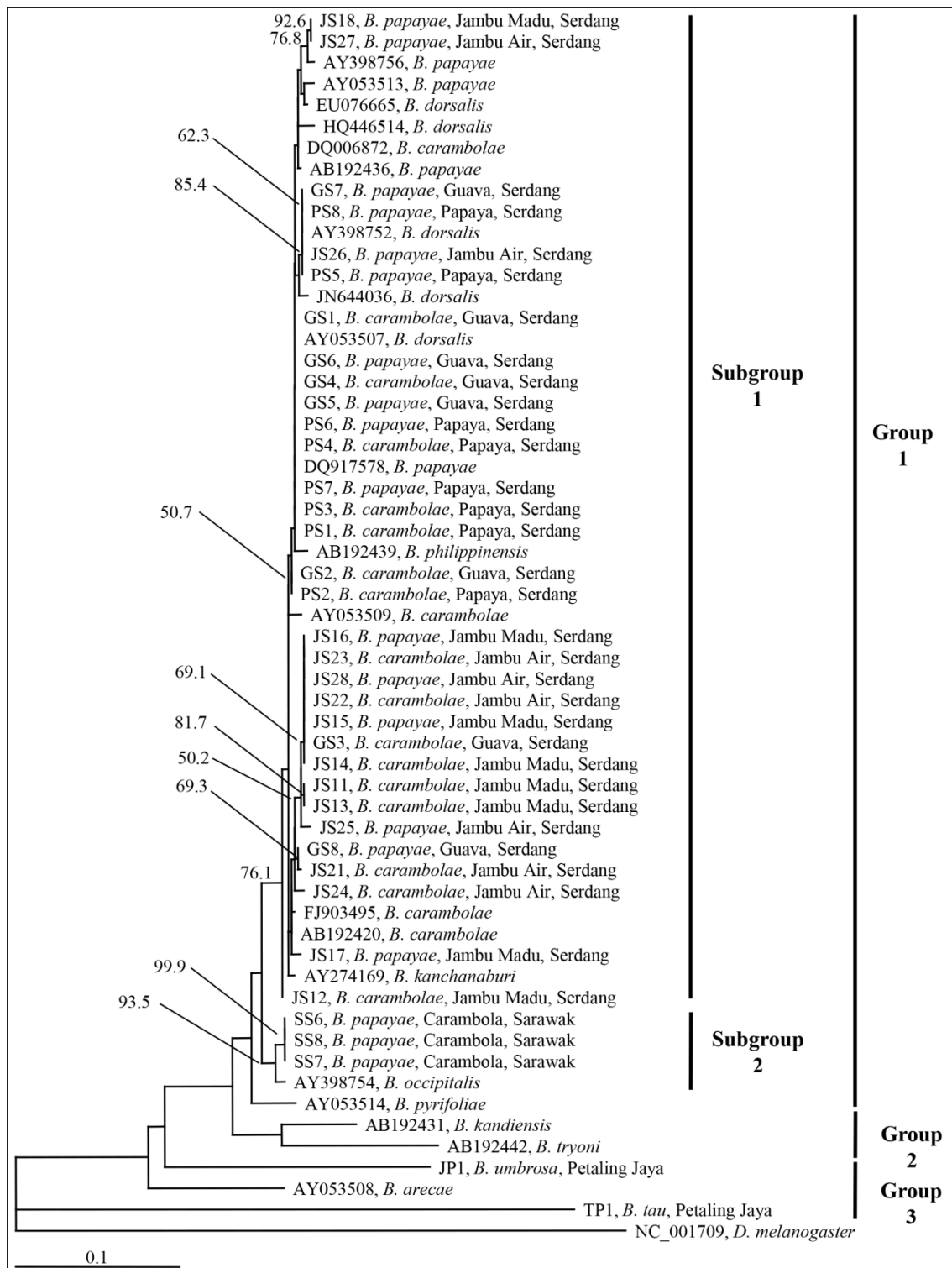


Figure 4.21: The 50% majority-rule consensus tree generated by maximum likelihood analysis based on COI DNA sequences. Numbers at the nodes indicate bootstrap values.

4.4.2 Bayesian Inference

Based on the COI BI tree (Figure 4.22), three major groups were observed. The first group consisted of *B. carambolae* and *B. papayae*, as well as other members of the *B. dorsalis* species complex, such as *B. occipitalis*, *B. pyrifoliae*, *B. kanchanaburi*, *B. philippinensis*, and *B. dorsalis*. The second group consisted of *B. kandiensis* and *B. tryoni*. The third group consisted of *B. arecae*; and other *Bactrocera* species such as, *B. umbrosa*, and *B. tau*.

The first group consisted of *B. carambolae* and *B. papayae* hatched from host fruits collected from Serdang as well as Sarawak. However, one clade within the first group contained specimens hatched from carambola host fruit collected from Sarawak with a posterior probability percentage of 89%, with *B. occipitalis* as the sister species. The other specimens collected from Serdang did not form any distinct clades and were grouped together with *B. carambolae*, *B. papayae*, *B. kanchanaburi*, *B. philippinensis*, and *B. dorsalis* specimens obtained from GenBank.

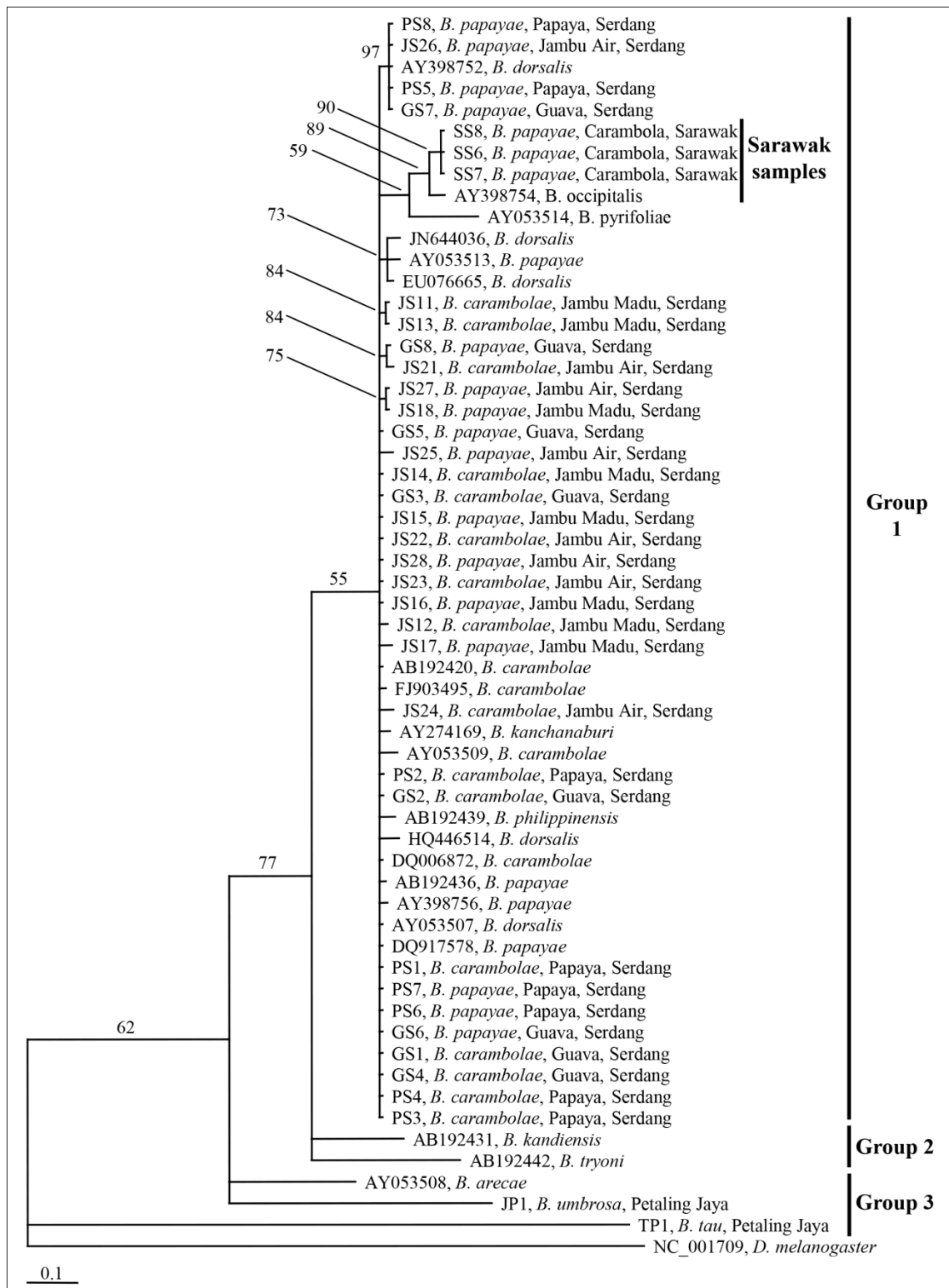


Figure 4.22: Bayesian phylogenetic analysis depicting the various categorizing of *Bactrocera* fruit fly variants based on COI. Numbers at the nodes indicate percentage of posterior probabilities over a generation number of two million generations.

4.4.3 Neighbor-Joining

Based on the COI NJ tree (Figure 4.23), three major groups were observed. The first group consisted of a mixture of *B. carambolae*, *B. papayae*, as well as other members of the *B. dorsalis* species complex, such as *B. dorsalis*, *B. philippinensis*, *B. kanchanaburi*, *B. occipitalis*, and with *B. pyrifoliae* as the basal species supported by a bootstrap value of 93%. The second group consisted of *B. kandiensis* and *B. tryoni*. The third group consisted of *B. arecae*; and other *Bactrocera* species such as, *B. umbrosa*, and *B. tau*.

Within the first group, two subgroups were observed. The first subgroup consisted of *B. carambolae* and *B. papayae* hatched from host fruits collected from Serdang, as well as other GenBank specimens such as *B. carambolae*, *B. papayae*, *B. dorsalis*, *B. philippinensis*, and *B. kanchanaburi*. The first subgroup was supported by a high bootstrap value of 90%. The specimens hatched from host fruits did not form distinct clades and tended to mix with specimens obtained from GenBank. The second subgroup consisted of *B. papayae* specimens hatched from carambola host fruits collected from Sarawak with a high bootstrap value of 96% and with *B. occipitalis* as the sister species.

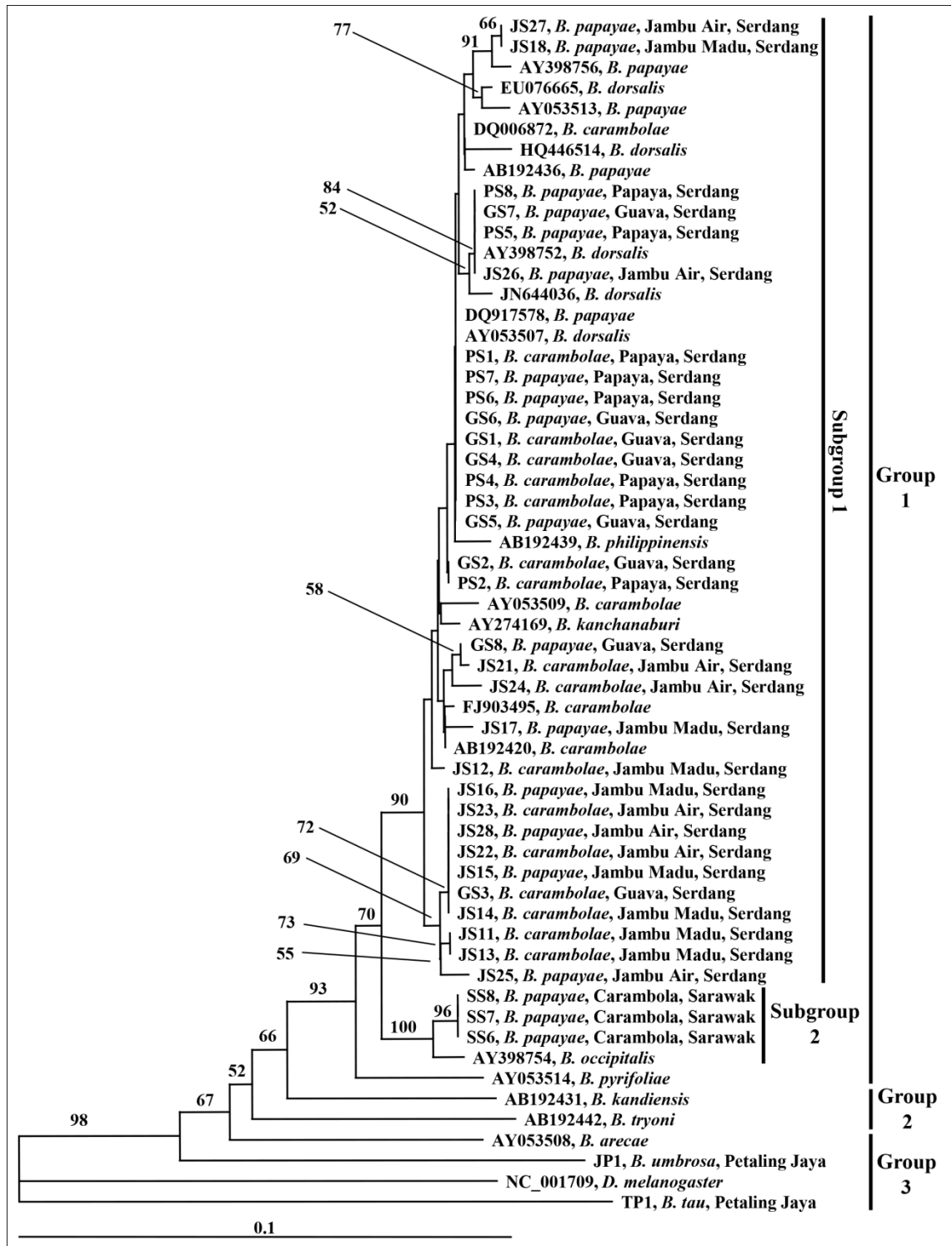


Figure 4.23: Phylogenetic tree generated by Neighbor-Joining analysis based on COI gene. Numbers at the nodes indicate bootstrap values.

4.5 Phylogeny Based On Cytochrome Oxidase Subunit II Gene

4.5.1 Maximum-Likelihood

Based on the COII ML tree (Figure 4.24), two major groups were observed. The first group consisted of *B. carambolae*, *B. papayae*, and other members of the *B. dorsalis* species complex such as *B. philippinensis*, *B. dorsalis*, *B. cacuminata*, *B. occipitalis*, *B. cognata*, *B. caryeae*, and *B. kandiensis* with no bootstrap support. The second major group consisted of *B. arecae*, *B. umbrosa*, *B. tryoni*, *B. tau*, *B. cucubita*, and *B. endiandrae*.

Within the first group, two subgroups were observed. The first subgroup consisted of *B. carambolae* and *B. papayae* specimens hatched from host fruits collected from Serdang and Sarawak, as well as other specimens obtained from GenBank such as *B. carambolae*, *B. papayae*, *B. philippinensis*, *B. dorsalis*, *B. cacuminata*, *B. occipitalis*, and *B. cognata*. The first subgroup consisted of clades that contain a mixture of *B. carambolae* and *B. papayae* specimens hatched from host fruits collected from Serdang as well as *B. carambolae*, *B. papayae*, *B. philippinensis*, *B. dorsalis*, *B. cacuminata*, and *B. occipitalis* specimens obtained from GenBank. However, one clade consisted of *B. carambolae* and *B. papayae* specimens hatched from carambola host fruits collected from Sarawak, with a high bootstrap value of 97.1%, and with *B. cognata* as the sister species. The second subgroup consisted of members of the *B. dorsalis* complex, including *B. caryeae*, and *B. kandiensis* supported by a moderate bootstrap value of 72.7%.

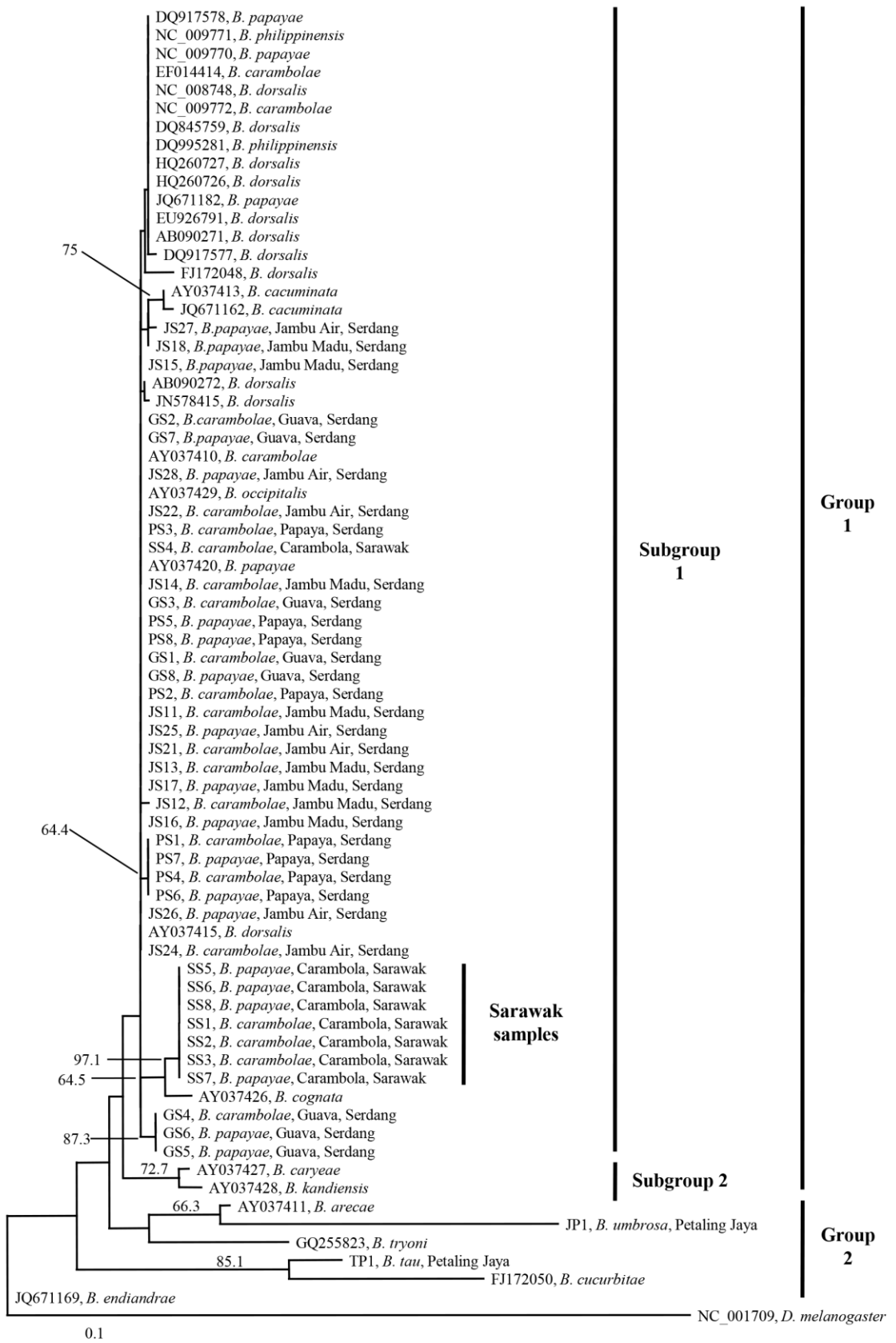


Figure 4.24: The 50% majority-rule consensus tree generated by maximum likelihood analysis based on COII DNA sequences. Numbers at the nodes indicate bootstrap values.

4.5.2 Bayesian Inference

Based on the BI tree (Figure 4.25), two major groups were observed. Group 1 consisted of *B. carambolae*, *B. papayae*, *B. umbrosa*, *B. tryoni*, and other members of the *B. dorsalis* species complex such as *B. cognata*, *B. arecae*, *B. cacuminata*, *B. caryeae*, *B. kandiensis*, *B. dorsalis*, *B. endiandrae*, *B. occipitalis*, and *B. philippinensis*, and is supported by a moderate posterior probability percentage of 65%. Group 2 consisted of *B. cucurbitae* and *B. tau*, and is supported by a high posterior probability percentage of 99%.

Within group 1, seven subgroups were observed. The first subgroup consisted of *B. carambolae* and *B. papayae* specimens hatched from carambola host fruit collected from Sarawak, supported by a moderate posterior probability percentage of 78%, with *B. cognata* as the basal species. The second subgroup consisted of *B. carambolae* and *B. papayae* specimens hatched from papaya host fruit collected from Serdang, supported by a moderate posterior probability percentage of 77%. The third subgroup consisted of *B. carambolae* and *B. papayae* specimens hatched from guava host fruit collected from Serdang, supported by a moderate posterior probability percentage of 76%. The fourth subgroup, supported by a moderate posterior probability percentage of 62%, consisted of *B. umbrosa*, *B. arecae*, and *B. tryoni*. The fifth subgroup consisted of *B. cacuminata* specimens, supported by a high posterior probability percentage of 97%. The sixth subgroup consisted of *B. caryeae* and *B. kandiensis*, supported by a high posterior probability percentage of 92%. The seventh subgroup consisted of *B. carambolae* and *B. papayae* specimens hatched from host fruits collected from Serdang, and GenBank specimens such as *B. carambolae*, *B. papayae*, *B. dorsalis*, *B. endiandrae*, *B. occipitalis*, and *B. philippinensis*.

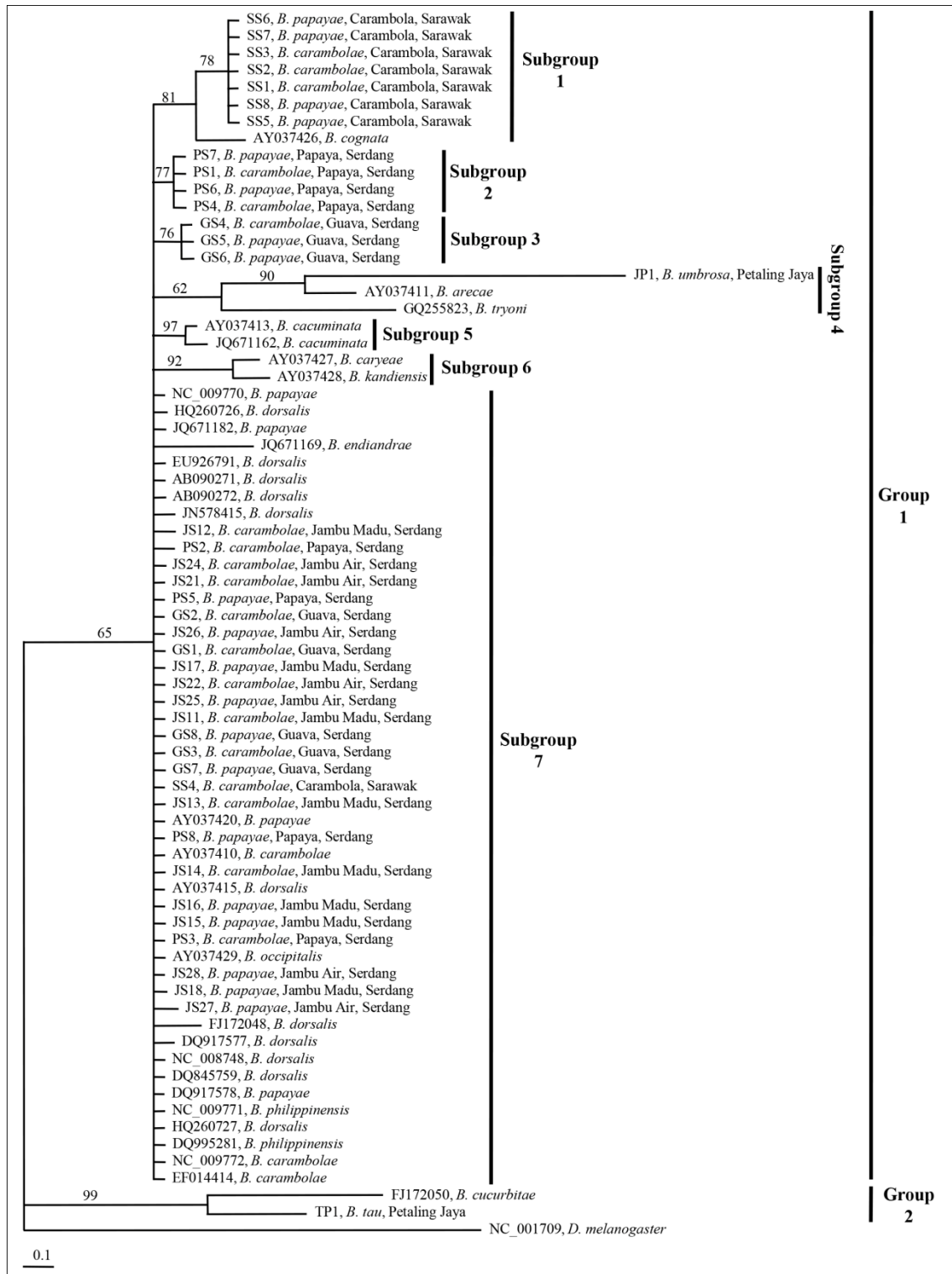


Figure 4.25: Bayesian phylogenetic analysis depicting the various categorizing of *Bactrocera* fruit fly variants based on COII. Numbers at the nodes indicate percentage of posterior probabilities over a generation number of two million generations.

4.5.3 Neighbor-Joining

Based on the NJ tree (Figure 4.26), two major groups were observed. The first group consisted of *B. carambolae*, *B. papayae*, and other members of the *B. dorsalis* species complex such as *B. dorsalis*, *B. philippinensis*, *B. occipitalis*, *B. cacuminata*, *B. cognata*, *B. caryeae*, and *B. kandiensis*. The second group consisted of *B. umbrosa*, *B. tryoni*, *B. cucurbitae*, *B. tau*, and two members of the *B. dorsalis* species complex, including *B. arecae* and *B. endiandrae*.

Within the first group, three subgroups were observed. The first subgroup consisted of a mixture of *B. carambolae* and *B. papayae* specimens hatched from host fruits collected from Serdang, as well as other GenBank specimens such as *B. carambolae*, *B. papayae*, *B. dorsalis*, *B. philippinensis*, *B. occipitalis*, and *B. cacuminata*. Within the first subgroup, two clades were observed. Subgroup 1a consisted of *B. carambolae* and *B. papayae* specimens hatched from papaya host fruits, while subgroup 1b consisted of *B. carambolae* and *B. papayae* specimens hatched from guava host fruits collected from Serdang.

The second subgroup consisted of *B. carambolae* and *B. papayae* specimens hatched from carambolae host fruits collected from Sarawak supported by a high bootstrap value of 89% with *B. cognata* as the sister species. The third subgroup consisted of *B. caryeae*, and *B. kandiensis*, supported by a high bootstrap value of 90%. Within the first subgroup, a clade was observed which consisted of *B. carambolae* and *B. papayae* specimens hatched from papaya host fruits collected from Serdang, supported by a moderate bootstrap value of 64%.

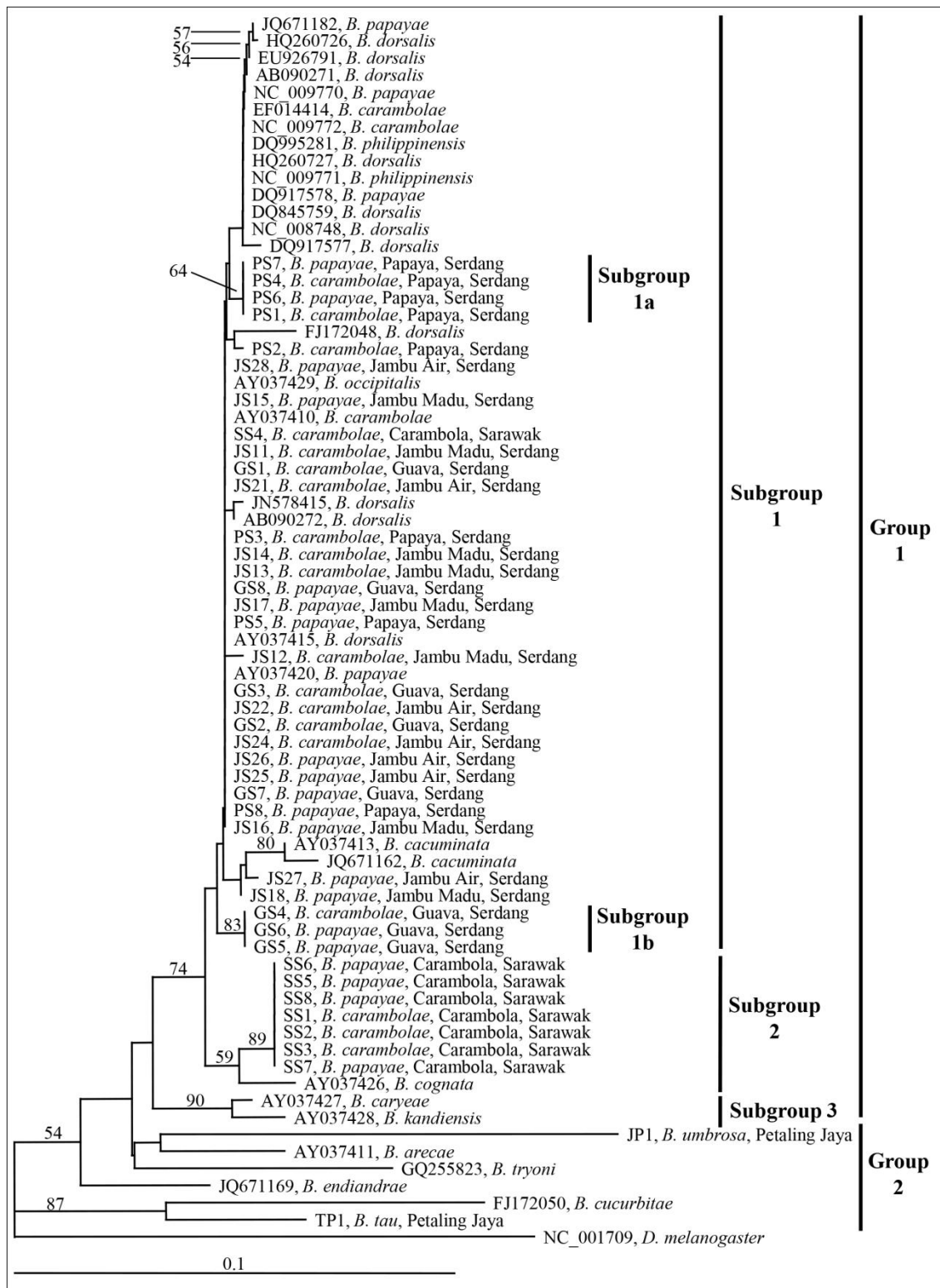


Figure 4.26: Phylogenetic tree generated by Neighbor-Joining analysis based on COII gene. Numbers at the nodes indicate bootstrap values.

4.6 Phylogeny Based On Cytochrome-b Gene

4.6.1 Maximum-Likelihood

Based on the *cytb* ML tree (Figure 4.27), three major groups were observed. The first group consisted of *B. carambolae* and *B. papayae* specimens hatched from host fruits collected from Serdang, as well as one *B. carambolae* specimen hatched from carambola host fruit collected from Sarawak. The remaining specimens within the first group were *B. dorsalis*, *B. carambolae*, and *B. papayae* specimens obtained from GenBank. A few subclades were observed among the first main group but there is no distinct or clear separation on the species or the host where they were originated. The second group consisted of *B. carambolae* and *B. papayae* hatched from carambola host fruit from Sarawak and supported with a moderate bootstrap value of 78.9%. A clade consisting of *B. umbrosa* was placed in between the first and second group. The third group consisted of *B. oleae* and *B. tau*.

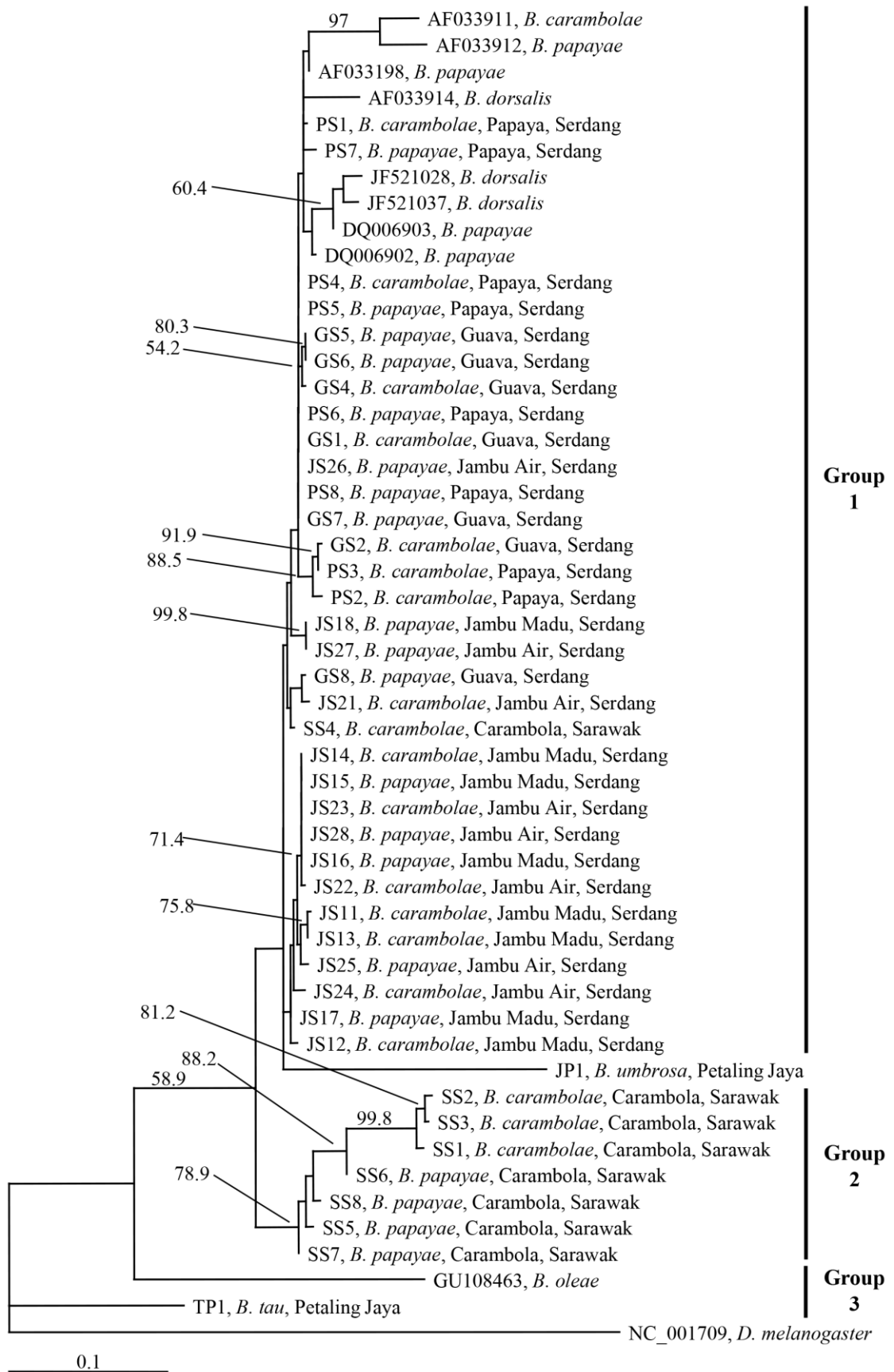


Figure 4.27: The 50% majority-rule consensus tree generated by maximum likelihood analysis based on *cytb* DNA sequences. Numbers at the nodes indicate bootstrap values.

4.6.2 Bayesian Inference

Based on the BI tree (Figure 4.28), two major groups were observed. The first group consisted of *B. carambolae*, *B. papayae*, *B. dorsalis*, and *B. umbrosa*, which is supported by a posterior probability percentage of 69%. The second group consisted of *B. oleae* and *B. tau*.

Within the first group, two subgroups were observed. Subgroup 1 consisted of *B. carambolae* and *B. papayae* specimens hatched from carambola host fruits collected from Sarawak, supported by a high posterior probability percentage of 98%. Subgroup 2 consisted of *B. carambolae* and *B. papayae* specimens hatched from various host fruits collected from Serdang, *B. umbrosa* specimen collected from Petaling Jaya, as well as other *B. carambolae*, *B. papayae*, and *B. dorsalis* specimens obtained from GenBank.

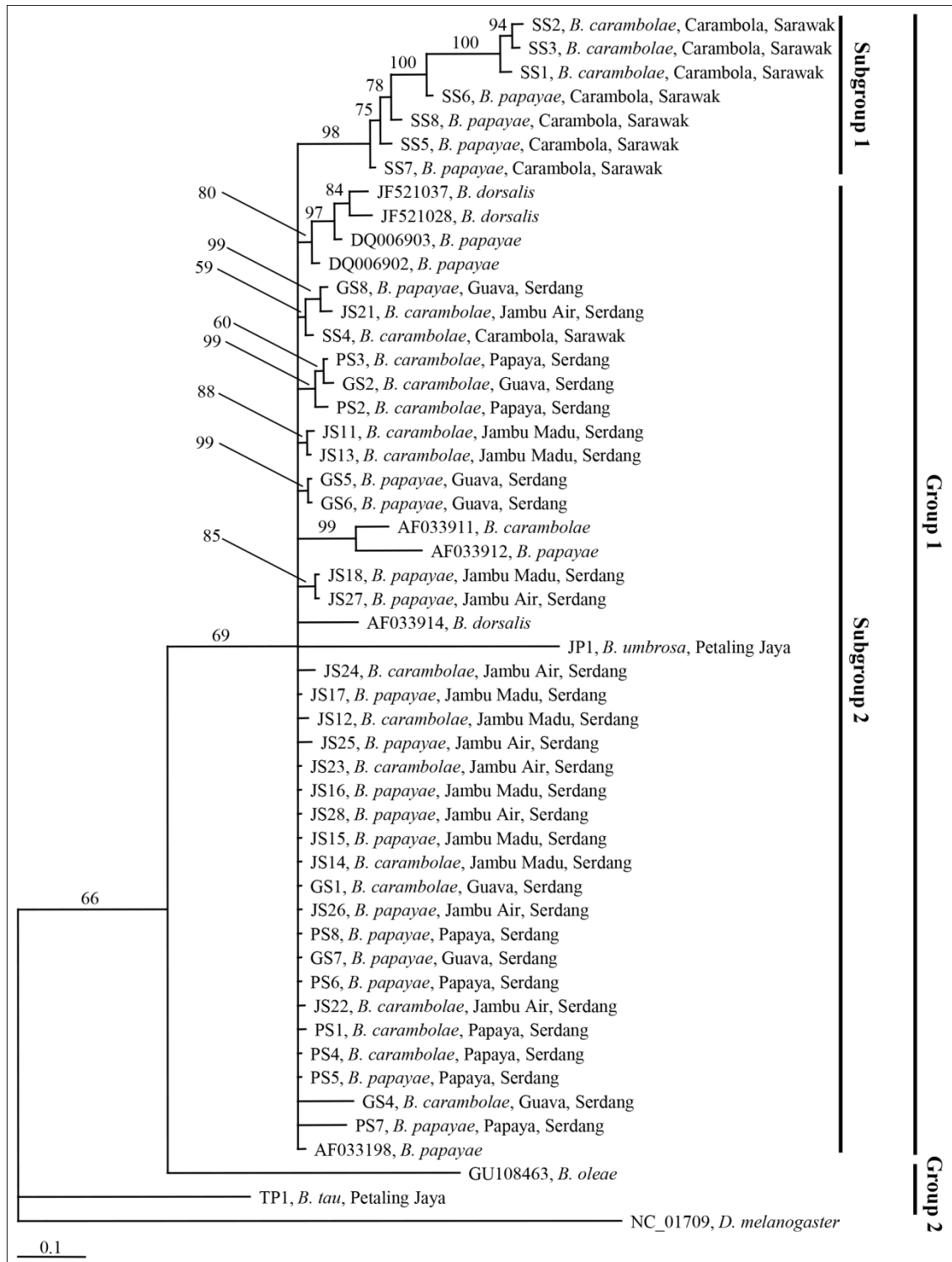


Figure 4.28: Bayesian phylogenetic analysis depicting the various categorizing of *Bactrocera* fruit fly variants based on *cytb*. Numbers at the nodes indicate percentage of posterior probabilities over a generation number of two million generations.

4.6.3 Neighbor-Joining

Based on the NJ tree (Figure 4.29), three major groups were observed. The first group consisted of *B. carambolae* and *B. papayae* specimens hatched from host fruits collected from Serdang, as well as one *B. carambolae* specimen hatched from carambola host fruits collected from Sarawak, and is supported by a high bootstrap value of 96%. The first group also consisted of specimens obtained from GenBank, such as *B. carambolae*, *B. papayae*, and *B. dorsalis*. The second group consisted of *B. carambolae* and *B. papayae* specimens hatched from carambola host fruits collected from Sarawak, supported by a high bootstrap value of 96%. The third group consisted of *B. oleae* and *B. tau*.

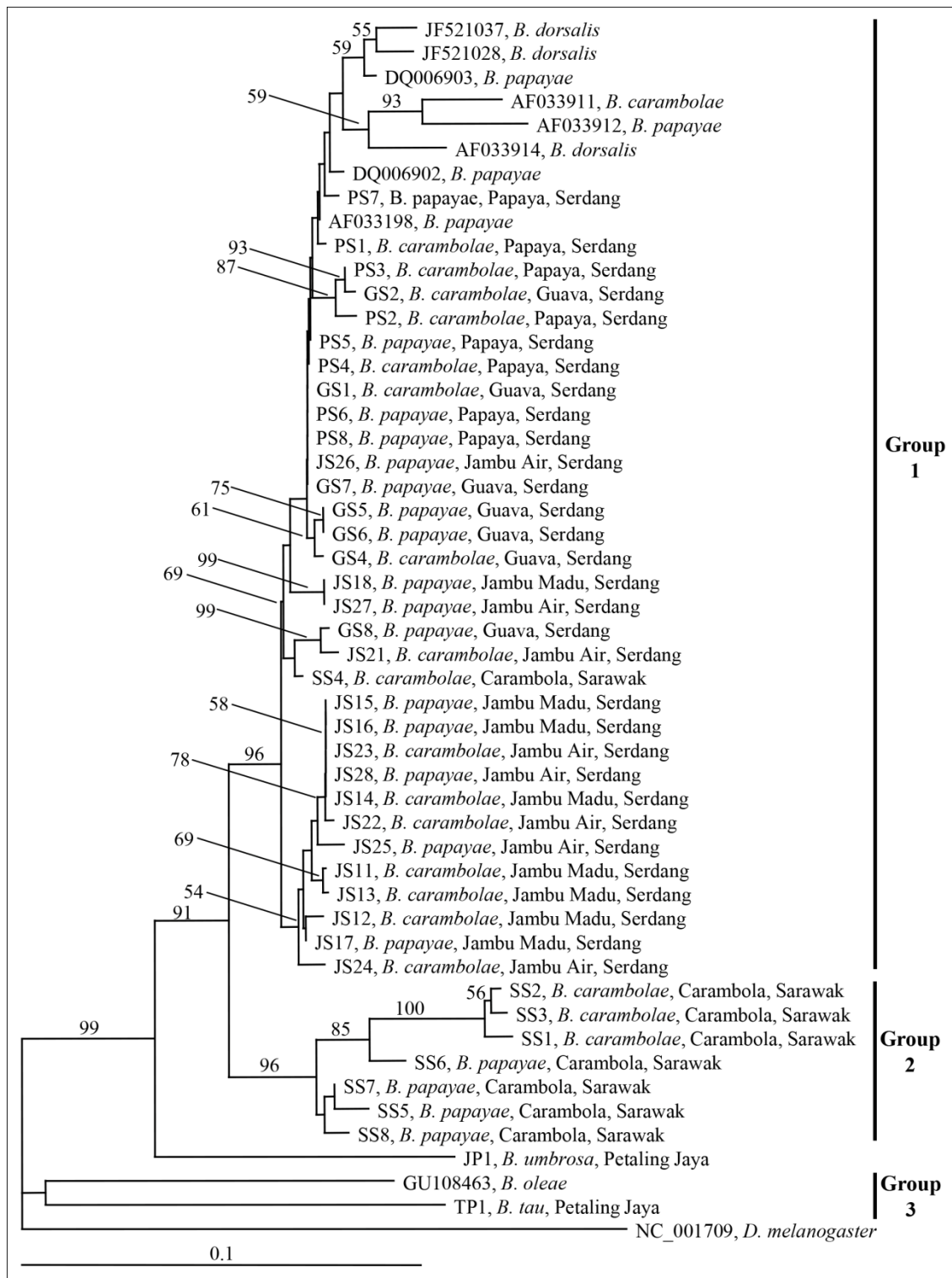


Figure 4.29: Phylogenetic tree generated by Neighbor-Joining analysis based on *cytb* gene. Numbers at the nodes indicate bootstrap values.

4.7 Phylogeny Based On Combined COI, COII, and *cytb* Genes

4.7.1 Maximum-Likelihood

Based on the ML tree (Figure 4.30), two major groups were observed. The first group consisted of *B. carambolae* and *B. papayae* specimens hatched from guava, papaya, jambu madu, and jambu air host fruits collected from Serdang, supported by a moderate bootstrap value of 53.4%. Within this group, no distinct clades were formed based on species or the host fruits they originated from. The second group consisted of *B. papayae* specimens hatched from carambola host fruit collected from Sarawak, supported by a high bootstrap value of 97.6%.

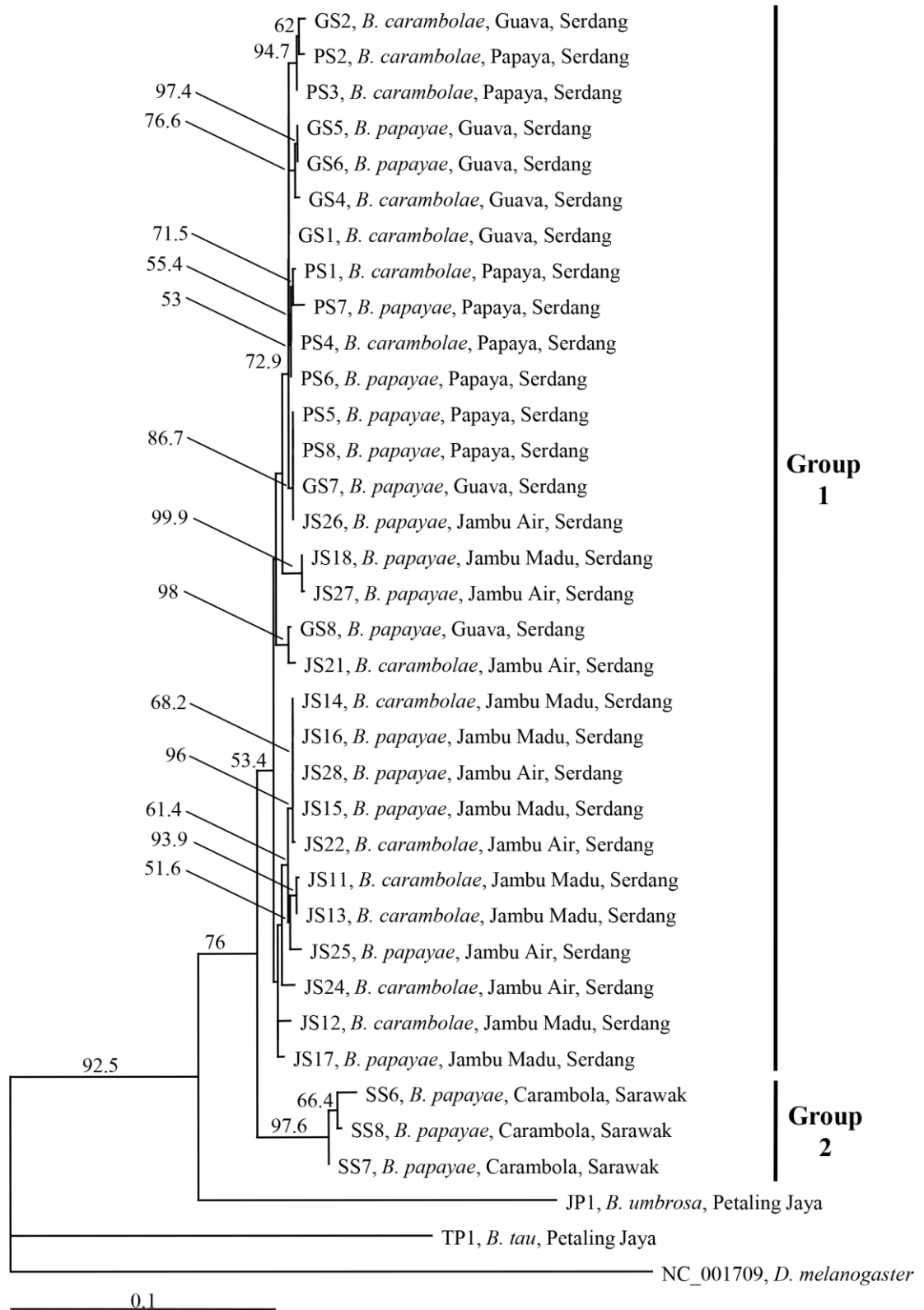


Figure 4.30: The 50% majority-rule consensus tree generated by maximum likelihood analysis based on combined COI, COII, and *cytb* DNA sequences. Numbers at the nodes indicate bootstrap values.

4.7.2 Bayesian Inference

Based on the BI tree (Figure 4.31), three major groupings were observed. The first group, supported by a moderate posterior probability value of 56%, consisted of *B. carambolae* and *B. papayae* specimens hatched from guava and papaya host fruits collected from Serdang, as well as one *B. papayae* specimen hatched from jambu air host fruit collected from Serdang. The second group, supported by a moderate posterior probability value of 59%, consisted of *B. carambolae* and *B. papayae* specimens hatched from jambu madu and jambu air host fruits collected from Serdang, as well as *B. papayae* specimens hatched from carambola host fruit collected from Sarawak. The third group consisted of *B. carambolae* and *B. papayae* specimens hatched from jambu madu and jambu air host fruits collected from Serdang, as well as one *B. papayae* specimen hatched from guava host fruit collected from Serdang. The third group showed no distinct grouping of specimens based on host fruit origin.

The first group consisted of four clades, however only two clades showed distinct grouping based on host fruit. Clade A, supported by a high posterior probability value of 96%, consisted of *B. carambolae* and *B. papayae* specimens hatched from papaya host fruits collected from Serdang. Clade B, supported by a high posterior probability value of 92%, consisted of *B. carambolae* and *B. papayae* specimens hatched from guava host fruit collected from Serdang. The second group consisted of three clades, however only one clade showed distinct grouping based on host fruit origin. Clade C consisted of *B. papayae* specimens hatched from carambola host fruit collected from Sarawak, which is supported by a high posterior probability value of 100%.

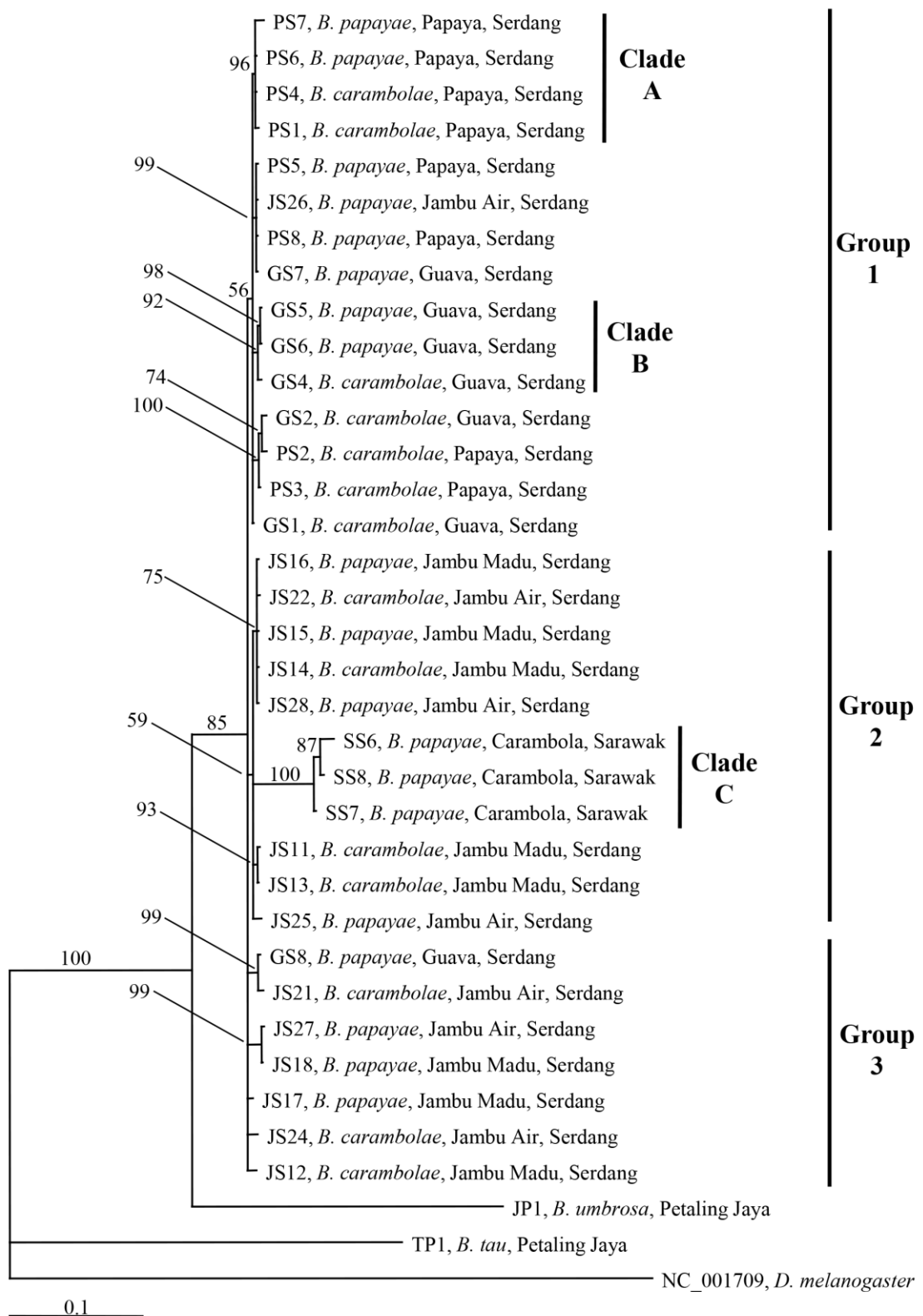


Figure 4.31: Bayesian phylogenetic analysis depicting the various categorizing of *Bactrocera* fruit fly variants based on combined COI, COII, and *cytb* DNA sequences. Numbers at the nodes indicate percentage of posterior probabilities over a generation number of two million generations.

4.7.3 Neighbor-Joining

Based on the NJ tree (Figure 4.32), two major groups were observed. The first group consisted of *B. carambolae* and *B. papayae* specimens hatched from guava, papaya, jambu air, and jambu madu host fruits collected from Serdang. The second group consisted of *B. papayae* specimens hatched from carambola host fruit collected from Sarawak, supported by a high bootstrap value of 100%.

Within the first group, two subgroups were observed. The first subgroup, supported by a moderate bootstrap value of 54%, consisted of mainly *B. carambolae* and *B. papayae* specimens hatched from papaya and host fruits collected from Serdang, as well as two *B. papayae* specimens hatched from jambu madu and jambu air host fruits collected from Serdang. The second subgroup consisted of mainly *B. carambolae* and *B. papayae* specimens hatched from jambu madu and jambu air host fruits collected from Serdang, as well as one *B. papayae* specimen hatched from guava host fruit collected from Serdang.

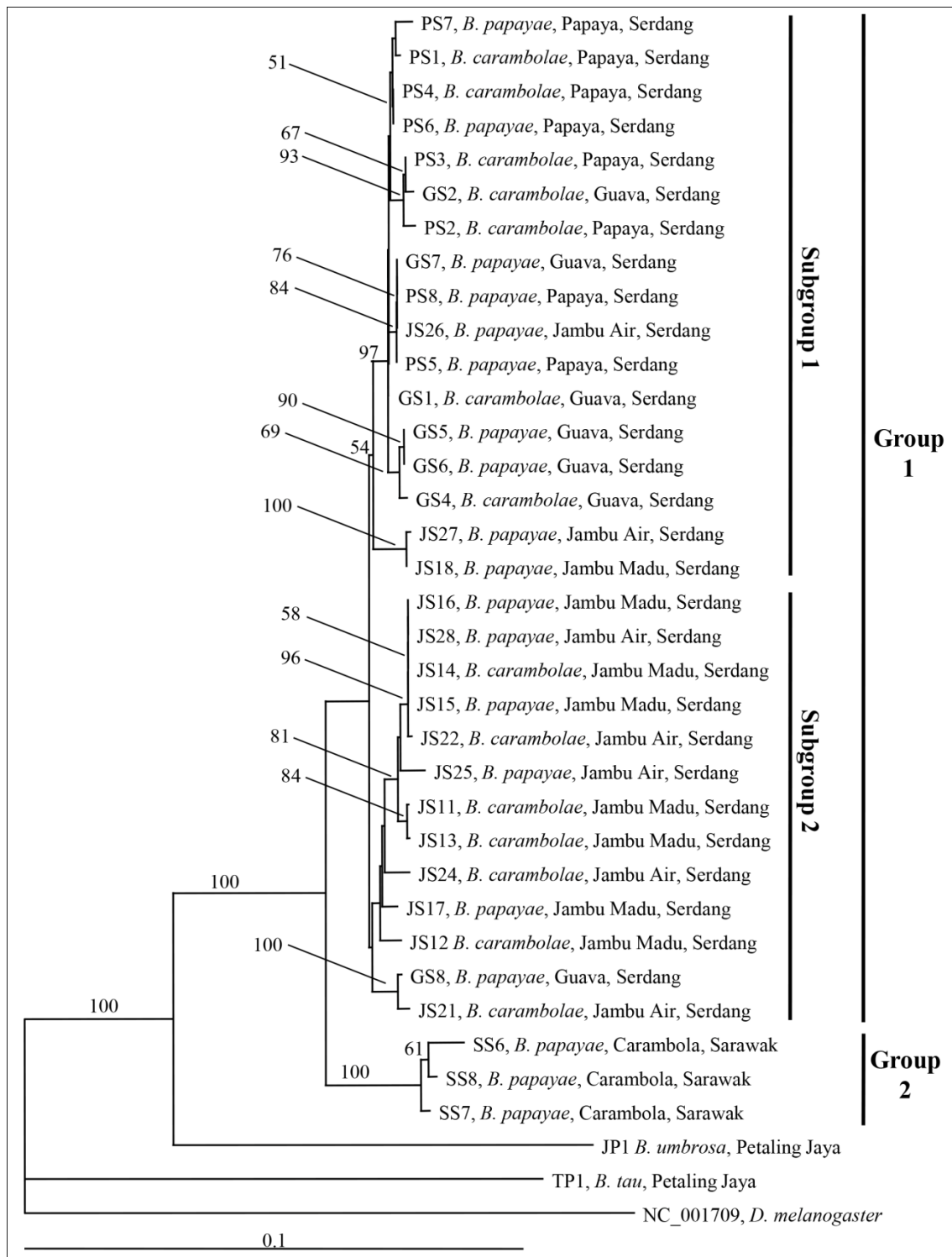


Figure 4.32: Phylogenetic tree generated by Neighbor-Joining analysis based on combined COI, COII, and *cytb* DNA sequences. Numbers at the nodes indicate bootstrap values.

4.8 Haplotype Network Reconstruction

Haplotype networks for all *B. carambolae* and *B. papayae* specimens were constructed based on aligned DNA sequences of COI, COII, *cytb*, and combined COI, COII, and *cytb* molecular markers. Table 4.2 shows the identities of all haplotype groups generated.

Table 4.2: Haplotype grouping and their identities based on COI, COII, *cytb*, and combined COI, COII, and *cytb* molecular markers.

Molecular Marker	Haplotype	Specimens
COI	COI-A	GS1, GS4, GS5, GS6, PS1, PS3, PS4, PS6, PS7
	COI-B	GS3, JS14, JS15, JS16, JS22, JS23, JS28
	COI-C	GS7, JS26, PS5, PS8
	COI-D	SS6, SS7, SS8
	COI-E	GS2, PS2
	COI-F	JS18, JS27
	COI-G	JS11, JS13
	COI-H	JS12
	COI-I	JS17
	COI-J	GS8
	COI-K	JS21
	COI-L	JS24
	COI-M	JS25
COII	COII-A	GS1, GS2, GS3, GS7, GS8, JS11, JS13, JS14, JS15, JS16, JS17, JS21, JS22, JS24, JS25, JS26, JS28, PS2, PS3, PS5, PS8, SS4
	COII-B	SS1, SS2, SS3, SS5, SS6, SS7, SS8
	COII-C	PS1, PS4, PS6, PS7
	COII-D	GS4, GS5, GS6
	COII-E	JS12
	COII-F	JS27
	COII-G	JS18
<i>cytb</i>	<i>cytb</i> -A	GS1, GS7, JS26, PS4, PS5, PS6, PS8
	<i>cytb</i> -B	JS14, JS15, JS16, JS23, JS28
	<i>cytb</i> -C	JS18, JS27
	<i>cytb</i> -D	GS5, GS6
	<i>cytb</i> -E	GS2, PS3
	<i>cytb</i> -F	JS22
	<i>cytb</i> -G	JS25
	<i>cytb</i> -H	JS11

Table 4.1, continued

<i>cytb</i>	<i>cytb-I</i>	JS13
	<i>cytb-J</i>	JS12
	<i>cytb-K</i>	JS17
	<i>cytb-L</i>	JS24
	<i>cytb-M</i>	GS8
	<i>cytb-N</i>	JS21
	<i>cytb-O</i>	SS4
	<i>cytb-P</i>	GS4
	<i>cytb-Q</i>	PS2
	<i>cytb-R</i>	PS1
	<i>cytb-S</i>	SS7
	<i>cytb-T</i>	SS5
	<i>cytb-U</i>	SS8
	<i>cytb-V</i>	SS2
	<i>cytb-W</i>	SS3
	<i>cytb-X</i>	SS1
	<i>cytb-Y</i>	SS6
Combined COI, COII, and <i>cytb</i>	A	GS1
	B	JS14, JS15, JS16, JS22, JS28
	C	GS7, JS26, PS5, PS8
	D	PS1, PS4, PS6
	E	GS5, GS6
	F	JS25
	G	JS11
	H	JS13
	I	JS24
	J	JS17
	K	JS12
	L	GS8
	M	JS21
	N	JS27
	O	JS18
	P	GS2
	Q	PS2
R	PS3	
S	PS7	
T	GS4	
U	SS6	
V	SS8	
W	SS7	

4.8.1. Haplotype Network Reconstruction Based On Cytochrome Oxidase Subunit I Gene

The aligned sequences of COI consisted of 536 characters and formed 13 haplotype groups, with haplotype COI-A inferred as the basal haplotype. A minimum of 12 mutational steps was required to link these haplotype groups. Seven haplotypes (COI-A, COI-B, COI-C, COI-D, COI-E, COI-F, COI-G) were shared by at least two individuals, while the remaining six haplotypes (COI-H, COI-I, COI-J, COI-K, COI-L, COI-M) were unique. No distinct haplogroups were observed. Figure 4.33 shows the haplotype network generated for all *B. carambolae* and *B. papayae* specimens using COI gene.

COI-A consisted of *B. carambolae* and *B. papayae* hatched from guava and papaya host fruits. COI-B consisted of *B. carambolae* and *B. papayae* hatched from guava, jambu madu, and jambu air host fruits, and differs from COI-A by seven mutation steps. COI-C consisted of *B. papayae* hatched from guava, papaya, and jambu air host fruits, and differs from COI-A by two mutation steps. COI-D consisted of *B. papayae* hatched from carambola host fruits, and differs from COI-A by 19 mutation steps. COI-E consisted of *B. carambolae* hatched from guava and papaya host fruits, and differs from COI-A by one mutation step. COI-F consisted of *B. papayae* hatched from jambu madu and jambu air host fruits, and differs from COI-A by five mutation steps. COI-G consisted of *B. carambolae* hatched from jambu madu host fruit, and differs from COI-A by seven mutation steps. COI-H consisted of *B. carambolae* hatched from jambu madu host fruit, and differs from COI-A by four mutation steps. COI-I consisted of *B. papayae* hatched from jambu madu host fruit, and differs from COI-A by six mutation steps. COI-J consisted of *B. papayae* hatched from guava host

fruit, and differs from COI-A by five mutation steps. COI-K consisted of *B. carambolae* hatched from jambu air host fruit, and differs from COI-A by six mutation steps. COI-L consisted of *B. carambolae* hatched from jambu air host fruit, and differs from COI-A by seven mutation steps. COI-M consisted of *B. papayae* hatched from jambu air host fruit, and differs from COI-A by nine mutation steps.

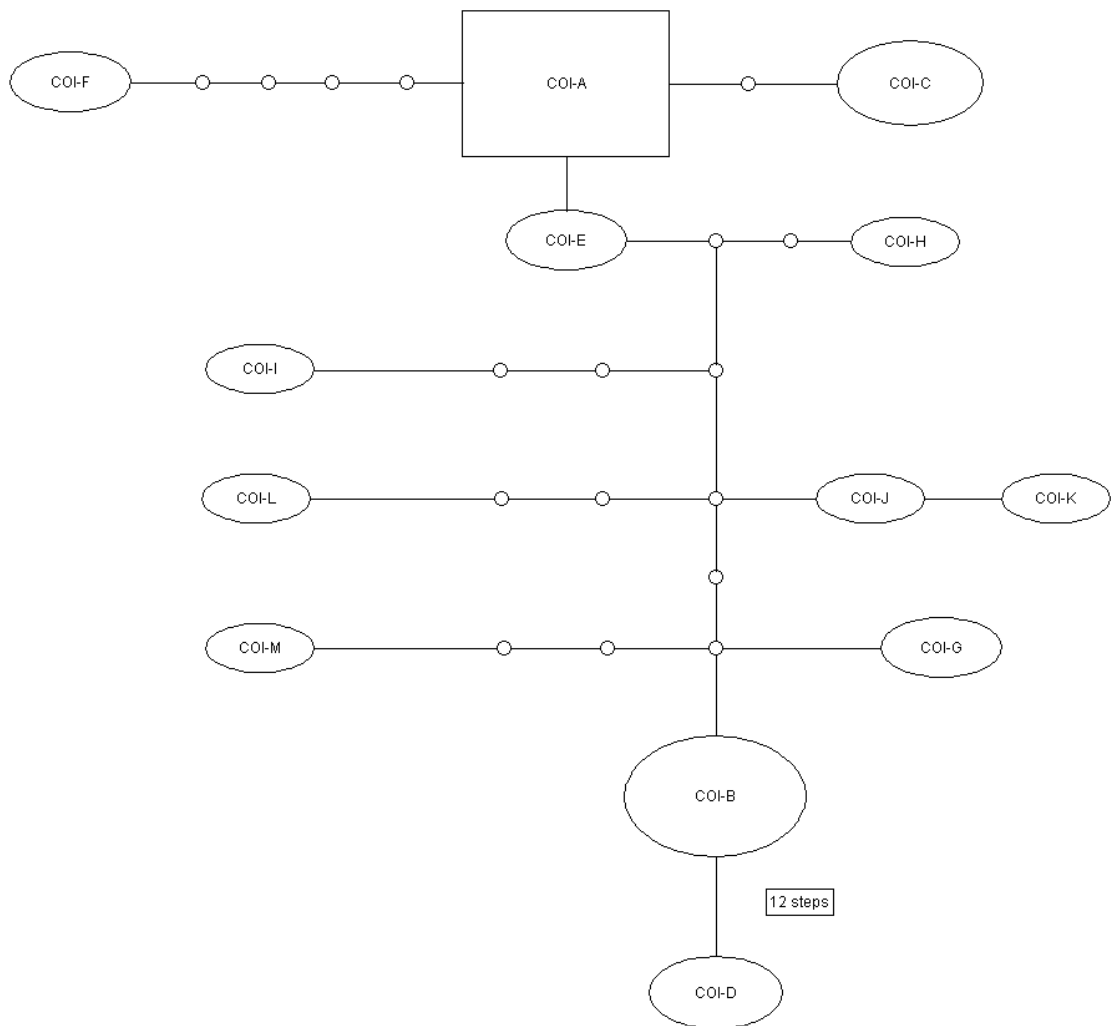


Figure 4.33: Statistical parsimony networks for COI of *Bactrocera carambolae* and *Bactrocera papayae*. Size of squares and ovals are proportional to the haplotype frequency. Lines represent parsimonious connections between haplotype groups while the clear circles indicate hypothetical missing haplotype.

4.8.2 Haplotype Network Reconstruction Based On Cytochrome Oxidase Subunit II Gene

The aligned sequences of COII consisted of 246 characters and formed 7 haplotype groups, with COII-A inferred as the basal haplotype. Four haplotypes (COII-A, COII-B, COII-C, COII-D) were shared by at least two individuals, while the remaining three haplotypes (COII-E, COII-F, COII-G) were unique. Figure 4.34 shows the haplotype network generated for all *B. carambolae* and *B. papayae* specimens using COII gene.

COII-A consisted of *B. carambolae* and *B. papayae* hatched from guava, papaya, jambu madu, jambu air, and carambola host fruits. COII-B consisted of *B. carambolae* and *B. papayae* hatched from carambola host fruit, and differs from COII-A by five mutation steps. COII-C consisted of *B. carambolae* and *B. papayae* hatched from papaya host fruit, and differs from COII-A by one mutation step. COII-D consisted of *B. carambolae* and *B. papayae* hatched from guava host fruit, and differs from COII-A by two mutation steps. COII-E consisted of *B. carambolae* hatched from jambu madu host fruit, and differs from COII-A by one mutation step. COII-F consisted of *B. papayae* hatched from jambu air host fruit, and differs from COII-A by two mutation steps. COII-G consisted of *B. papayae* hatched from jambu madu host fruit, and differs from COII-A by one mutation step.

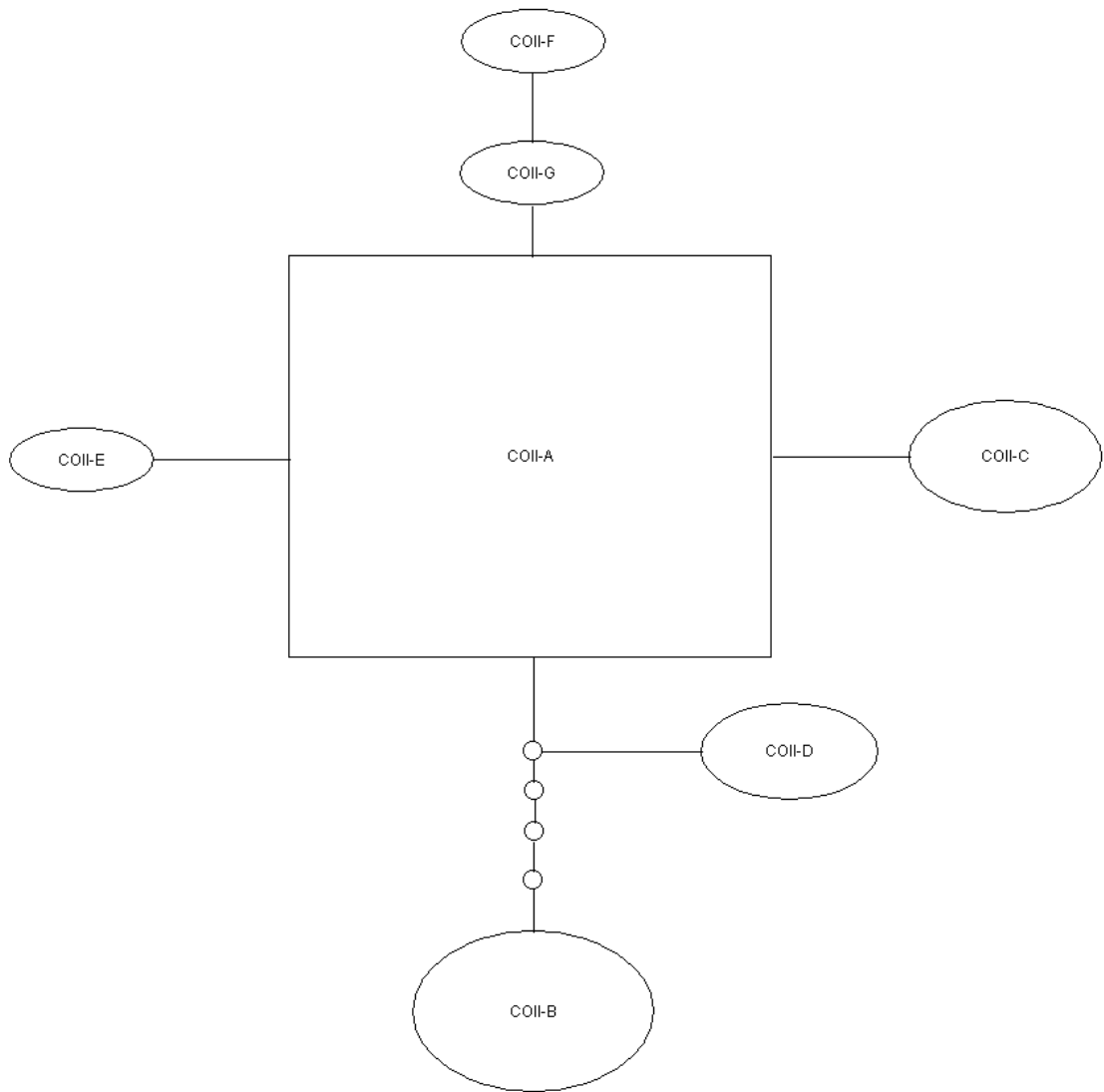


Figure 4.34: Statistical parsimony networks for COII of *Bactrocera carambolae* and *Bactrocera papayae*. Size of squares and ovals are proportional to the haplotype frequency. Lines represent parsimonious connections between haplotype groups while the clear circles indicate hypothetical missing haplotype.

4.8.3 Haplotype Network Reconstruction Based On Cytochrome-b Gene

The aligned sequences of *cytb* consisted of 463 characters and formed 25 haplotype groups, with *cytb*-A inferred as the basal haplotype. A minimum of 19 mutational steps was required to link these haplotype groups. It is observed that *cytb* produced the most haplotypes out of the three molecular markers utilized in this study. Five haplotypes (*cytb*-A, *cytb*-B, *cytb*-C, *cytb*-D, *cytb*-E) were shared by at least two individuals, while the remaining 20 haplotypes (*cytb*-F, *cytb*-G, *cytb*-H, *cytb*-I, *cytb*-J, *cytb*-K, *cytb*-L, *cytb*-M, *cytb*-N, *cytb*-O, *cytb*-P, *cytb*-Q, *cytb*-R, *cytb*-S, *cytb*-T, *cytb*-U, *cytb*-V, *cytb*-W, *cytb*-X, *cytb*-Y) were unique. A loop was observed within the network which involved haplotypes *cytb*-B, *cytb*-K, and *cytb*-L. Figure 4.35 shows the haplotype network generated for all *B. carambolae* and *B. papayae* specimens using *cytb* gene.

cytb-A consisted of *B. carambolae* and *B. papayae* hatched from guava, papaya, and jambu air host fruit. *cytb*-B consisted of *B. carambolae* and *B. papayae* hatched from jambu madu, and jambu air host fruits, and differs from *cytb*-A by 11 mutation steps. *cytb*-C consisted of *B. papayae* hatched from jambu madu and jambu air host fruits, and differs from *cytb*-A by six mutation steps. *cytb*-D consisted of *B. papayae* hatched from guava host fruit, and differs from *cytb*-A by two mutation steps. *cytb*-E consisted of *B. carambolae* hatched from guava and papaya host fruits, and differs from *cytb*-A by three mutation steps. *cytb*-F consisted of *B. carambolae* hatched from jambu air host fruit, and differs from *cytb*-A by 12 mutation steps. *cytb*-G consisted of *B. papayae* hatched from jambu air host fruits, and differs from *cytb*-A by 11 mutation steps. *cytb*-H consisted of *B. carambolae* hatched from jambu madu host fruit, and differs from *cytb*-A by ten mutation steps. *cytb*-I consisted of *B. carambolae* hatched from jambu madu host fruit, and differs from *cytb*-A by nine mutation steps. *cytb*-

Jconsisted of *B. carambolae* hatched from jambu madu host fruit, and differs from *cytb*-A by eight mutation steps. *cytb*-K consisted of *B. papayae* hatched from jambu madu host fruit, and differs from *cytb*-A by six mutation steps. *cytb*-L consisted of *B. carambolae* hatched from jambu air host fruit, and differs from *cytb*-A by ten mutation steps. *cytb*-M consisted of *B. papayae* hatched from guava host fruit, and differs from *cytb*-A by eight mutation steps. *cytb*-N consisted of *B. carambolae* hatched from jambu air host fruit, and differs from *cytb*-A by nine mutation steps. *cytb*-O consisted of *B. carambolae* hatched from carambola host fruit, and differs from *cytb*-A by five mutation steps. *cytb*-P consisted of *B. carambolae* hatched from guava host fruit, and differs from *cytb*-A by two mutation steps. *cytb*-Q consisted of *B. carambolae* hatched from papaya host fruit, and differs from *cytb*-A by five mutation steps. *cytb*-R consisted of *B. carambolae* hatched from papaya host fruit, and differs from *cytb*-A by two mutation steps. *cytb*-S consisted of *B. papayae* hatched from carambola host fruit, and differs from *cytb*-A by 20 mutation steps. *cytb*-T consisted of *B. papayae* hatched from carambola host fruit, and differs from *cytb*-A by 24 mutation steps. *cytb*-U consisted of *B. papayae* hatched from carambola host fruit, and differs from *cytb*-A by 26 mutation steps. *cytb*-V consisted of *B. carambolae* hatched from carambola host fruit, and differs from *cytb*-A by 54 mutation steps. *cytb*-W consisted of *B. carambolae* hatched from carambola host fruit, and differs from *cytb*-A by 52 mutation steps. *cytb*-X consisted of *B. carambolae* hatched from carambola host fruit, and differs from *cytb*-A by 50 mutation steps. *cytb*-Y consisted of *B. papayae* hatched from carambola host fruit, and differs from *cytb*-A by 33 mutation steps.

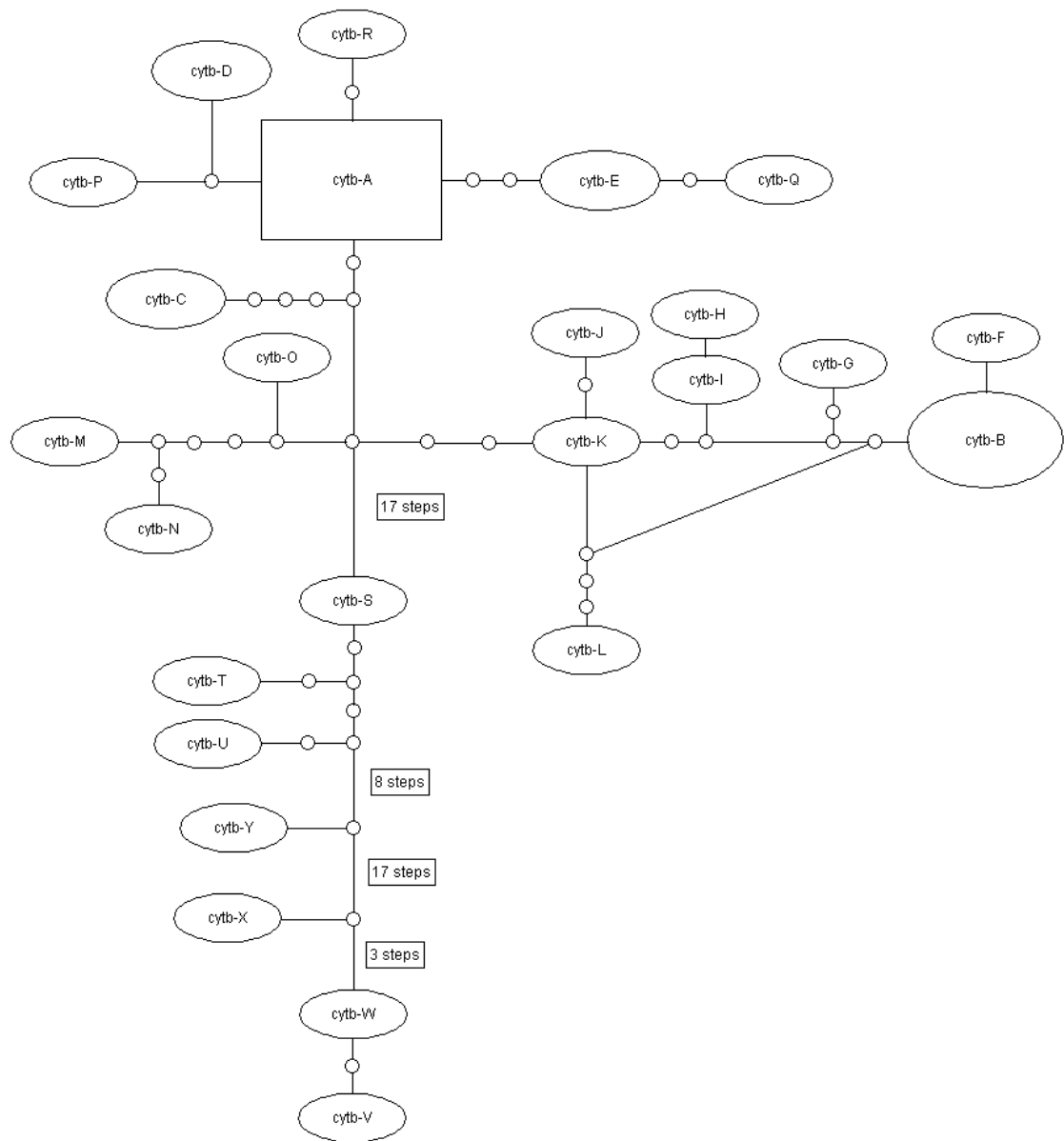


Figure 4.35: Statistical parsimony networks for *cytb* of *Bactrocera carambolae* and *Bactrocera papayae*. Size of squares and ovals are proportional to the haplotype frequency. Lines represent parsimonious connections between haplotype groups while the clear circles indicate hypothetical missing haplotype.

4.8.4 Haplotype Network Reconstruction Based On Combined COI, COII, and *cytb* Genes

The combined aligned sequences of COI, COII, and *cytb* consisted of 1214 characters and formed 23 haplotype groups, with haplotype A inferred as the basal haplotype. A minimum of 39 mutational steps was required to link these haplotype groups. Four haplotypes (B, C, D, and E) were shared by at least two individuals, while the remaining haplotypes (A, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, and W) were unique. Figure 4.36 shows the haplotype network generated for all *B. carambolae* and *B. papayae* specimens using combined COI, COII, and *cytb* genes.

A consisted of *B. carambolae* hatched from guava host fruit. B consisted of *B. carambolae* and *B. papayae* hatched from jambu madu, and jambu air host fruits, and differs from A by 15 mutation steps. C consisted of *B. papayae* hatched from guava, jambu air, and papaya host fruits, and differs from A by two mutation steps. D consisted of *B. carambolae* and *B. papayae* hatched from papaya host fruit, and differs from A by one mutation step. E consisted of *B. papayae* hatched from guava host fruit, and differs from A by four mutation steps. F consisted of *B. papayae* hatched from jambu air host fruit, and differs from A by 19 mutation steps. G consisted of *B. carambolae* hatched from jambu madu host fruit, and differs from A by 18 mutation steps. H consisted of *B. carambolae* hatched from jambu madu host fruit, and differs from A by 17 mutation steps. I consisted of *B. carambolae* hatched from jambu air host fruit, and differs from A by 16 mutation steps. J consisted of *B. papayae* hatched from jambu madu host fruit, and differs from A by 11 mutation steps. K consisted of *B. carambolae* hatched from jambu madu host fruit, and differs from A by 12 mutation steps. L consisted of *B. papayae* hatched from guava host fruit, and differs from A by 12 mutation steps. M

consisted of *B. carambolae* hatched from jambu air host fruit, and differs from A by 14 mutation steps. N consisted of *B. papayae* hatched from jambu air host fruit, and differs from A by 12 mutation steps. O consisted of *B. papayae* hatched from jambu madu host fruit, and differs from A by 11 mutation steps. P consisted of *B. carambolae* hatched from guava host fruit, and differs from A by four mutation steps. Q consisted of *B. carambolae* hatched from papaya host fruit, and differs from A by six mutation steps. R consisted of *B. carambolae* hatched from papaya host fruit, and differs from A by three mutation steps. S consisted of *B. papayae* hatched from papaya host fruit, and differs from A by 42 mutation steps. T consisted of *B. carambolae* hatched from guava host fruit, and differs from A by 33 mutation steps. U consisted of *B. papayae* hatched from carambola host fruit, and differs from A by 66 mutation steps. V consisted of *B. papayae* hatched from carambola host fruit, and differs from A by 59 mutation steps. W consisted of *B. papayae* hatched from carambola host fruit, and differs from A by 57 mutation steps.

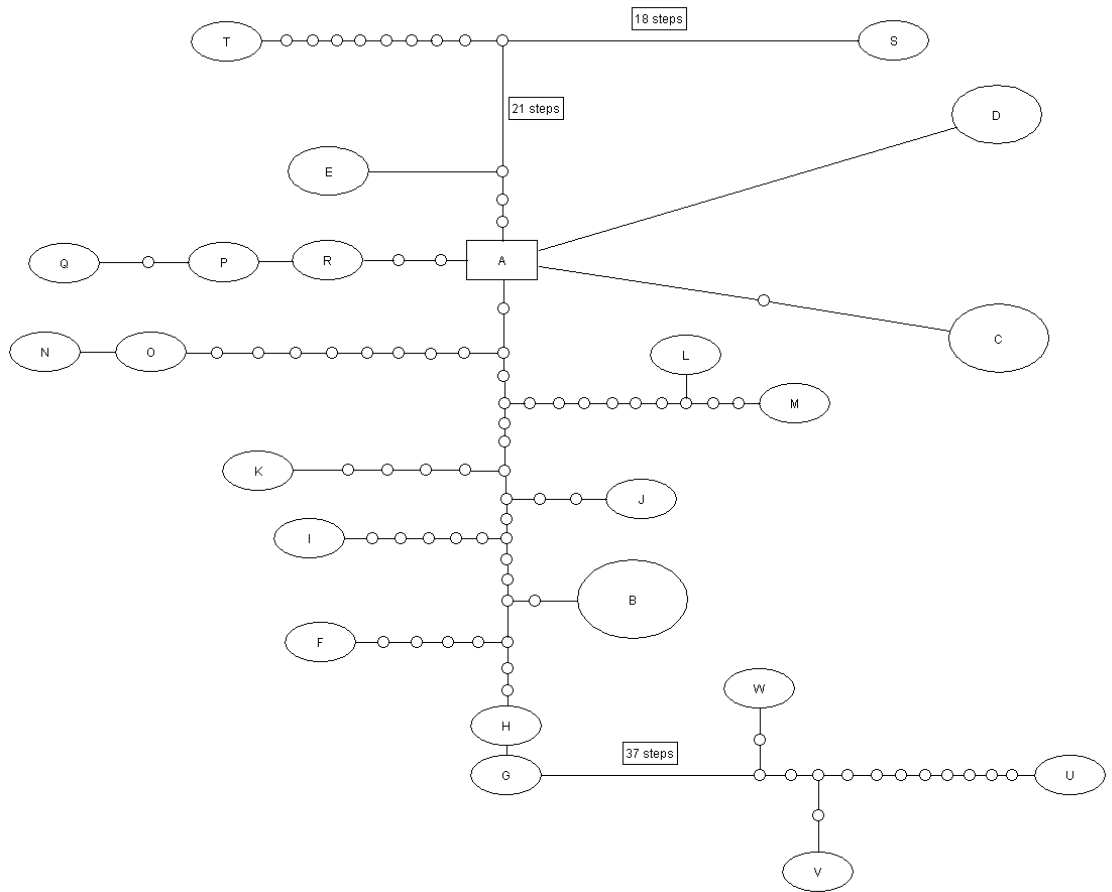


Figure 4.36: Statistical parsimony networks for combined COI, COII, and *cytb* of *Bactrocera carambolae* and *Bactrocera papayae*. Size of squares and ovals are proportional to the haplotype frequency. Lines represent parsimonious connections between haplotype groups while the clear circles indicate hypothetical missing haplotype.

CHAPTER 5

DISCUSSION

5.1 Species Identification

Identification of *B. carambolae* and *B. papayae* were based on morphological methods established by Drew and Hancock (1994). It was observed that all five host fruits had *B. carambolae* and *B. papayae* fruit flies emerged from them; however the number of *B. carambolae* fruit flies identified was usually higher than *B. papayae* fruit flies (results not shown). Identification of *B. carambolae* and *B. papayae* was based on the wing and abdominal patterns; however, there were several instances where the morphological features of the unidentified fruit flies appear to be an intermediate between *B. carambolae* and *B. papayae*, and were difficult to discriminate. Hence, specimens that were ambiguous in their identities were not selected for DNA extraction and phylogenetic studies. The emergence of fruit flies with intermediate morphological features confer with findings published by Iwahashi (2001), stating that the morphological features of *B. carambolae* and *B. papayae* range in intermediate forms that segregate within the species.

At the same time, two other related species of *Bactrocera* fruit flies (*B. umbrosa* and *B. tau*) were used in the phylogenetic studies as a potential outgroup to *B. carambolae* and *B. papayae*. In phylogenetic studies by Smith *et al.* (2002) and Smith *et al.* (2003), it was observed that *B. umbrosa* and *B. tau* grouped separately from *B. carambolae* and *B. papayae*, but were still closely related enough to infer phylogeny.

5.2 DNA Extraction

The i-Genomic CTB DNA Extraction Kit (iNtRON Biotechnology, Inc, South Korea) was found to give good quality and quantity of DNA for PCR amplification with some minor modifications from the manufacturer's instructions. However, in order to obtain satisfactory DNA yield, two legs from each fruit fly was crushed and homogenized in a microcentrifuge tube using a micropestle for up to 20 minutes. Results are not shown here, but it showed that by using one leg, the DNA's yield was too low and was not able to give PCR amplifications.

At the end of the DNA extraction, during the elution step, instead of using 100 μ l of Buffer CE to elute the DNA pellets, 50 μ l of Buffer CE was used instead to elute the first batch of DNA pellets. Eluting the DNA pellets with half the volume increases the DNA concentration of the first batch of DNA samples. Once the first elution was completed, a second elution was carried out on the DNA pellets. The second batch, however, would contain a lower concentration of DNA, but the overall yield of 100 μ l of DNA samples still remains the same. The differing concentrations of DNA samples would prove to be useful during the optimization of PCR parameters in the PCR stage of this study.

5.3 PCR Primers

The three pairs of published PCR primers utilized for this study were able to amplify the three mitochondrial encoded genes without much problem. The COI primer pair UEA7 and UEA 10 has been found to be able to amplify *Bactrocera philippinensis* and *Bactrocera occipitalis* (Yu *et al.*, 2005), the COII primer pair C2-J-3549 (alias C2KD-F) and TD-N-3884 (alias C2KD-R) have been utilized to amplify various *Bactrocera* fruit flies, such as *Bactrocera dorsalis*, *Bactrocera oleae*, and *Bactrocera tryoni* (Smith *et al.*, 2003), and finally the *cytb* primer pair CB1 and CB2 have been utilized to amplify *Bactrocera cucurbitae*, *Bactrocera correcta*, and *Bactrocera scutellata* (Zhu *et al.*, 2005).

5.4 PCR Optimization

Prior to PCR amplification of target genes, the three primer pairs were subjected to optimization. Optimization was carried out to ensure that the DNA bands amplified were clear, as well as to eliminate the presence of unspecific DNA products. PCR amplifications that contained unspecific DNA products were usually discarded and re-optimization was attempted as they would affect the DNA sequencing portion of the PCR product. For the primer pairs utilized to amplify COI and COII genes, only the target DNA fragments were observed after optimization. However, for the primer pair utilized to amplify *cytb* gene, despite numerous optimizations, unspecific bands were observed for several individuals, hence only the target DNA fragments were excised and used for DNA sequencing.

All three primer pairs used three different thermal cycling programs for DNA amplification. For PCR amplification of COI, the initial denaturation step was 94°C for three minutes, followed by 40 cycles of the following three steps: denaturation step at 95°C for one minute, annealing step at 50°C for one minute, and extension step at 72°C for one minute and 30 seconds. The final extension step was at 72°C for seven minutes, and finally the hold step at 4°C. For PCR amplification of COII, the initial denaturation step was 95°C for two minutes, followed by 40 cycles of the following three steps: denaturation step at 94°C for 45 seconds, annealing step at 44°C for 45 seconds, and extension step at 72°C for 45 seconds. The final extension step was at 72°C for seven minutes, and finally the hold step at 4°C. For PCR amplification of *cytb*, the initial denaturation step was 94°C for two minutes, followed by 40 cycles of the following three steps: denaturation step at 94°C for one minute, annealing step at 45°C for one minute, and extension step at 72°C for one minute and 30 seconds. The final extension step was at 72°C for five minutes, and finally the hold step at 4°C. The thermal cycling programmes used in this study to amplify the target genes were adapted from published literature (Jamnongluk *et al.*, 2003; Smith *et al.*, 2003; Zhu *et al.*, 2005) however, the final optimized thermal cycling programmes for each primer pair differed from their original source.

The initial denaturation step which is varied at 94 to 95°C for up to three minutes for all three primer pairs facilitates the denaturation of the DNA template prior to the actual denaturation step. During the denaturation step, the DNA template was denatured at 94 to 95°C for up to one minute for all three primer pairs. The amount of time at which the denaturation step is subjected to should not exceed one minute as the *Taq* polymerase would reduce in activity with prolonged exposure to high temperature. The denaturation step temperature and length should also not be too low as the DNA

template may fail to denature and ultimately causes failure of PCR amplification (McPherson and Moller, 2006).

The annealing step, whereby the primer pairs are annealed to the target DNA loci. The annealing step temperature depends on the melting point of the primer pairs, thus different primer pairs have different annealing temperatures. The annealing temperatures for primer pairs UEA7 and UEA10, C2KD-F and C2KD-R, and CB1 and CB2 were 50°C, 44°C, and 45°C respectively. The annealing step length was also taken into consideration as the longer the interval length was, the higher the possibility of the primer pairs to anneal at untargeted loci and forming unspecific amplification products. For all three primer pairs, the annealing steps varied from 45 seconds to one minute.

The three preceding steps were subjected to a cycle of 40 repeats in order to obtain brighter and sharper DNA bands. The increased repetition of the three steps enables the *Taq* polymerase to amplify more target DNA fragments and thus increases the concentration of the yielded PCR product (McPherson and Moller, 2006).

5.5 PCR Reaction Mixture

Optimization of the reaction mixture was also important in obtaining bright and sharp DNA bands. Whenever unspecific PCR products were observed, the amount of MgCl₂ in the reaction mixture was adjusted until the desired results were observed. The MgCl₂ reagent provides the cofactor necessary for the action of the *Taq* polymerase, however too much MgCl₂ inhibits the action of the *Taq* polymerase. At times, if the target DNA bands observed were not bright enough, the concentration of the DNA template was reduced. If the amount of DNA template is too high, it may inhibit the

action of the other PCR reagents and no amplification will occur (Altshuler, 2006). Optimization of the reaction mixture goes hand-in-hand with optimization of the thermal cycling program to ensure the most optimum products were yielded.

5.6 Phylogeny Based On Cytochrome Oxidase Subunit I Gene

All three phylogenetic analyses based on COI gene show similar topology for the placement of clades, and only differed at poorly supported nodes. Based on all three phylogenetic trees generated using maximum-likelihood, Bayesian Inference, and Neighbor-Joining analyses, no distinct clades were formed that could clearly distinguish between *B. carambolae* and *B. papayae*. A mixture of *B. carambolae* and *B. papayae* specimens tend to group together within the same clade. This suggests that COI molecular marker was unable to distinguish *B. carambolae* and *B. papayae* on a species-level. Other members of the *B. dorsalis* species complex also tended to form within the same clades as *B. carambolae* and *B. papayae*, such as *B. dorsalis*, *B. philippinensis*, *B. kanchanaburi*, and *B. occipitalis*. This further suggests that COI molecular marker was also unable to distinguish *B. carambolae* and *B. papayae* on the *B. dorsalis* species complex level.

In all three phylogenetic analyses based on COI gene, *B. papayae* specimens that were hatched from carambola host fruits collected from Sarawak grouped separately from *B. carambolae* and *B. papayae* specimens hatched from host fruits collected from Serdang. This brings to light the potential genetic difference between these two populations of fruit flies. The question lies, however in whether the genetic difference between the two populations was due to host fruit specificity or geographical factors. At the same time, phylogenetic analyses using COI showed that *B.*

carambolae and *B. papayae* hatched from host fruits collected from Serdang showed no distinct grouping based on host fruits. This suggests that, at least among guava, jambu madu, jambu air, and papaya host fruits, that *B. carambolae* and *B. papayae* have no preference for host fruit to or not showing fruit specificity.

Moreover, the usage of COI provides insight into the phylogenetic relationship of the members of the *B. dorsalis* complex. Based on the three phylogenetic analyses using COI gene, *B. carambolae*, *B. papayae*, *B. philippinensis*, *B. dorsalis*, *B. kanchanaburi* and *B. occipitalis* tended to group together within the same complex. All three phylogenetic analyses using COI gene showed that *B. occipitalis* was the sister to species to the group of *B. papayae* specimens hatched from the Sarawak collected host fruits. *B. dorsalis* was also observed to group together with *B. papayae* specimens in all three phylogenetic analyses. In both ML and NJ analyses, *B. philippinensis* was observed to be sister species to *B. carambolae* and *B. papayae* specimens hatched from Serdang host fruits. This grouping of *B. dorsalis* complex members was shown to cluster separately from three other members of the *B. dorsalis* complex, namely *B. pyrifoliae*, *B. kandiensis*, and *B. arcae*, which suggests that *B. pyrifoliae*, *B. kandiensis*, and *B. arcae* are not as closely related to the other members of the *B. dorsalis* species complex.

Contrasting with the findings of Zhang *et al.* (2010), it was shown that *B. papayae* tended to group separately from *B. carambolae*, suggesting that they are separate species. It was also observed that members of the *B. dorsalis* complex (*B. carambolae*, *B. philippinensis*, *B. dorsalis*, *B. occipitalis*, *B. papayae*) tended to group together within the same clade, which is a similar result shown in this present study. However, Zhang *et al.* (2010) utilizes the maximum parsimony phylogenetic analysis to construct their phylogenetic tree, a method not utilized in this present study.

5.7 Phylogeny Based On Cytochrome Oxidase Subunit II Gene

All three phylogenetic analyses based on COII gene show similar topology for the placement of clades, and only differed at poorly supported nodes. Based on all three phylogenetic trees generated using maximum-likelihood, Bayesian Inference, and Neighbor-Joining analyses, no distinct clades were formed that could clearly distinguish between *B. carambolae* and *B. papayae*. Similar with the phylogenetic analyses based on COI gene, a mixture of *B. carambolae* and *B. papayae* specimens group together within the same clade, suggesting that the COII molecular marker was also unable to distinguish *B. carambolae* and *B. papayae* on a species-level. Other members of the *B. dorsalis* species complex were also observed to group together with *B. carambolae* and *B. papayae* specimens, such as *B. philippinensis*, *B. dorsalis*, *B. cacuminata*, *B. occipitalis*, and *B. cognata*. This further suggests that the COII molecular marker was also unable to distinguish *B. carambolae* and *B. papayae* on the *B. dorsalis* species complex level.

As with the COI gene, phylogenetic analyses based on COII gene also showed that *B. carambolae* and *B. papayae* specimens hatched from carambola host fruits collected from Sarawak clustered separately from *B. carambolae* and *B. papayae* specimens hatched from the Serdang collected host fruits. All three phylogenetic analyses also showed *B. cognata* to be the sister species of the *B. carambolae* and *B. papayae* specimens from Sarawak. The COII molecular marker was also able to distinguish two groups of *B. carambolae* and *B. papayae* specimens, that is one group consisting of specimens hatched from papaya host fruit (PS1, PS4, PS6, and PS7), and another group consisting of specimens hatched from guava host fruit (GS4, GS5, and GS6). These two groups of individuals were grouped in separate clades within one

larger group of Serdang specimens. However, other individuals from the same papaya and guava host fruits were also observed to group together in different clades with other *B. carambolae* and *B. papayae* specimens from jambu madu and jambu air host fruits.

With COII, members of the *B. dorsalis* species complex that were observed to group together include *B. carambolae*, *B. papayae*, *B. dorsalis*, *B. philippinensis*, *B. occipitalis*, *B. cacuminata* and *B. cognata*. Members of the *B. dorsalis* species complex that were observed to group separately based on COII were *B. caryeae*, *B. arcae*, *B. endiandrae*, and *B. kandiensis* which was previously shown to group together with *B. carambolae* and *B. papayae* using COI molecular marker. However, BI analysis showed that all the members of the *B. dorsalis* species included in this analysis grouped together within the same group, including the ambiguous placement of *B. umbrosa*, *B. arecae*, and *B. tryoni* (Subgroup four, Figure 4.23) within the main group which was not observed in the ML and NJ analyses.

5.8 Phylogeny Based On Cytochrome-b Gene

All three phylogenetic analyses based on *cytb* gene show similar topology for the placement of clades, and only differed at poorly supported nodes. As with phylogenetic analyses carried out based on COI and COII genes, no distinct clades were observed that could clearly distinguish between *B. carambolae* and *B. papayae*. Both *B. carambolae* and *B. papayae* were observed to group together within the same clade, which suggests that the *cytb* molecular marker could not distinguish *B. carambolae* and *B. papayae* on the species-level. *B. dorsalis* was also observed to group together within the same clade as *B. carambolae* and *B. papayae*. This also suggests that *cytb* could not distinguish *B. carambolae* and *B. papayae* on a species complex level.

For *cytb*, *B. carambolae* and *B. papayae* specimens hatched from carambola host fruits collected from Sarawak grouped separately from *B. carambolae* and *B. papayae* hatched from Serdang collected host fruits. This was also observed in phylogenetic trees generated from COI and COII datasets. *B. carambolae* and *B. papayae* specimens from Serdang showed no distinct grouping according to host specificity, however, one clade showed grouping of *B. carambolae* and *B. papayae* specimens (GS4, GS5, and GS6) hatched from guava host fruit. Incidentally, the grouping of these three specimens was also observed in COII phylogenetic analyses. Other *B. carambolae* and *B. papayae* specimens hatched from guava host fruit were also observed to group together with other specimens hatched from jambu madu, jambu air, and papaya host fruits. This suggests that *B. carambolae* and *B. papayae* have no distinct host specificity.

5.9 Phylogeny Based On Combined COI, COII, and *cytb* Genes

All three phylogenetic analyses based on combined COI, COII, and *cytb* genes show similar topology for the placement of clades, and only differed at poorly supported nodes. Using ML, BI, and NJ phylogenetic analyses, combined COI, COII, and *cytb* genes were also unable to clearly distinguish between *B. carambolae* and *B. papayae*. No distinct clades were observed and *B. carambolae* and *B. papayae* were observed to group together within the same group. Despite attempting phylogenetic analyses using COI, COII, and *cytb* genes separately, combining the three molecular markers also could not distinguish *B. carambolae* and *B. papayae* on the species-level.

B. papayae specimens hatched from carambola host fruit collected from Sarawak were observed to group separately from *B. carambolae* and *B.*

papayaes specimens hatched from host fruits collected from Serdang. This particular grouping was also observed when utilizing COI, COII, and *cytb* genes separately for phylogenetic analyses. We can surmise that the three molecular markers were able to distinguish between the two populations of fruit flies. However, the *B. carambolae* and *B. papayae* specimens hatched from host fruits collected from Serdang could not be significantly distinguished based on host fruit specificity. Just like with COII, two clades were observed to contain only one type of host fruit each; papaya (PS1, PS4, PS6, PS7) (Figure 4.31 and Figure 4.32), and guava (GS4, GS5, and GS6) (Figure 4.31 and Figure 4.32), but other *B. carambolae* and *B. papayae* specimens from the same host fruits were also observed to group together with other specimens hatched from jambu madu and jambu air host fruits. Based on the combined COI, COII, and *cytb* genes, this suggests that *B. carambolae* and *B. papayae* have no distinct host specificity.

5.10 Haplotype Network Reconstruction

Haplotype analysis for COI, COII, *cytb*, as well as combined COI, COII and *cytb* genes showed that no distinct grouping was observed that clearly distinguishes *B. carambolae* and *B. papayae*. Basal haplotypes for COI, COII, and *cytb* genes show a mixture of *B. carambolae* and *B. papayae* specimens. No distinct haplogroups were formed to distinguish between *B. carambolae* and *B. papayae*, as well. Haplotype analyses using COI, COII, *cytb*, and combined COI, COII, and *cytb* genes were unable to distinguish *B. carambolae* and *B. papayae* on a species-level.

For all four haplotype analyses, it was observed that *B. carambolae* and *B. papayae* specimens hatched from host fruits collected from Sarawak grouped together within the same haplotype groups in their respective analysis. For COI, the Sarawak

haplotype group differed from the basal group by 19 mutational steps (Figure 4.33). For COII, the Sarawak haplotype group differed from the basal group by five mutational steps (Figure 4.34). For *cytb*, the Sarawak haplotype group differed from the basal group by 20 to 54 mutational steps (Figure 4.35). The results showed that among the three molecular markers utilized, *cytb* is the most variable. For the combined COI, COII, and *cytb* genes, the Sarawak haplotype group differed from the basal group by 57 to 66 mutational steps (Figure 4.36). For *cytb*, all the specimens collected from Sarawak were unique haplotypes, however within the network, they were observed to cluster together away from *B. carambolae* and *B. papayae* specimens collected from Serdang. As for the specimens collected from Serdang, no distinct haplogroup was observed that would suggest any form of host fruit specificity for *B. carambolae* and *B. papayae*. The haplotype analysis for COI, COII, *cytb*, and combined COI, COII, and *cytb* genes further suggests that *B. carambolae* and *B. papayae* show no distinct host fruit specificity.

5.11 Overview

Based on all the phylogenetic trees generated using various phylogenetic analyses based on COI, COII, and *cytb* genes, it is clear that *B. carambolae* and *B. papayae* could not be clearly distinguished as two separate species. The results have shown that COI, COII, and *cytb* were unable to resolve *B. carambolae* and *B. papayae* into separate clades. Regardless of the origin of the two species, whether it is host fruit or location, no distinct grouping was observed that shows *B. carambolae* and *B. papayae* as two separate species. A haplotype network reconstruction showed that *B. carambolae* and *B. papayae* also could not be distinguished. Haplotypes were formed

that contain both *B. carambolae* and *B. papayae* in them, and it was also observed that a mixture of *B. carambolae* and *B. papayae* specimens made up basal haplotype groups for several genes.

The only distinct grouping observed from the results was not species-based, but a grouping based on location. All phylogenetic trees have shown that *B. carambolae* and *B. papayae* specimens that originated from Sarawak tended to form a separate clade from *B. carambolae* and *B. papayae* specimens that originated from Serdang. Haplotype network reconstruction also confirmed that specimens from Sarawak formed haplogroups separate from specimens from Serdang, for all molecular markers used in this study. To further solidify the findings, the three molecular markers, COI, COII, and *cytb*, were combined and analyzed using the various phylogenetic analyses employed for this study. The combined markers also show the same results – *B. carambolae* and *B. papayae* could not be differentiated. The separate grouping of Sarawak specimens from Serdang specimens were observed as well. Haplotype network reconstruction also showed Sarawak specimens as separate haplogroups from the Serdang specimens.

It was observed that *B. carambolae* and *B. papayae* also did not form distinct clades based on host fruit specificity. *B. carambolae* and *B. papayae* hatched from host fruits collected from Serdang grouped together within the same clade, as shown by the phylogenetic analyses and haplotype network reconstruction. The only distinct grouping based on host fruit was observed by *B. carambolae* and *B. papayae* specimens hatched from host fruits collected from Sarawak, which was shown by the phylogenetic analyses and haplotype network reconstruction. In this case, the genetic variation observed between these two groups could either be attributed to host fruit specificity or geographical factors. However, a study by Shi *et al.* (2012) on *B. dorsalis* noted that their genetic structure was not affected by the host plant species, and that molecular variation mostly occurred within populations. In addition to that, all five host fruits

collected for this study were observed to have both *B. carambolae* and *B. papayae* emerging from them. Suffice to say, we can conclude that the separate grouping of *B. carambolae* and *B. papayae* specimens collected from Sarawak from those collected from Serdang was due to geographical factors. In a study conducted by Liu *et al.* (2007), they showed that one population of *B. dorsalis* fruit flies differed from other *B. dorsalis* populations collected in their study due to the existence of natural geographical barriers. The natural geographical barrier between Peninsular Malaysia and Sarawak, namely the South China Sea, may have played a part in the separate grouping of *B. carambolae* and *B. papayae* specimens collected from Sarawak from those collected from Serdang.

It can be concluded from this current work that *B. carambolae* and *B. papayae* could possibly belong to the same species. The ability of the species to cross-breed and produce hybrids with intermediate features (Ebina and Ohto, 2006) also highly suggests that the two species are actually one species. The mixed grouping of *B. carambolae* and *B. papayae* could be attributed to the *B. carambolae* and *B. papayae*'s capability to cross-breed (Iwaizumi *et al.*, 1997), whereby groupings such as observed in the COII phylogenetic analyses trees could be attributed to *B. carambolae* and *B. papayae* hybrids (Nakahara and Muraji, 2008). Several studies have shown difficulty in distinguishing between *B. carambolae* and *B. papayae* based on genetic and morphological methods (Iwahashi, 1999; Iwahashi, 2001; Muraji and Nakahara, 2002; Armstrong and Ball, 2005; Nakahara and Muraji, 2008).

One other aspect that should be addressed is the grouping of *B. dorsalis* within the same clades as *B. carambolae* and *B. papayae*. A study carried out by Krosch *et al.*, (2012b), employing the usage of morphological and genetic methods, found that two regions that were believed to exclusively contain *B. papayae* and *B. dorsalis*. *s.* consisted of *B. papayae* and *B. dorsalis*. *s.* that have continuous morphological (wing

shape and aedeagus length) and genetic variation between the two populations. Another study by Schutze *et al.*, (2012), also employing a suite of morphological and genetic methods, have inferred that *B. dorsalis* s., *B. papayae*, and *B. philippinensis* populations from various regions in South East Asia are one species structured around the region of South China Sea. These findings suggest that members of the *B. dorsalis* species complex could be one biological species with continuous morphological features and genetic variation. Schutze *et al.*, (2012) surmises that based on his study's inability to distinguish *B. dorsalis* and *B. papayae* using genetic methods, it is possible that *B. dorsalis* and *B. papayae* are one biological species, rather than two separate species. This study further expands on the findings of Schutze *et al.*, (2012), in that *B. papayae* and *B. carambolae* both are undistinguishable from *B. dorsalis* despite utilizing three molecular markers.

Seeing how the set of molecular markers were unable to distinguish between *B. carambolae* and *B. papayae*, the correlation of the morphological data with molecular data could not be determined. While identification of *B. carambolae* and *B. papayae* was based on morphological methods as described by Drew and Hancock (1994), identification based on molecular markers is only as accurate as the specificity of the marker itself. The phylogenetic analyses and haplotype network reconstruction based on COI, COII, and *cytb* have both shown that *B. carambolae* and *B. papayae* could not be clearly distinguished.

CHAPTER 6

CONCLUSION

This study set out with a goal to determine the phylogenetic relationships between *Bactrocera carambolae* and *Bactrocera papayae*. From the collection and hatching of host fruits, it was observed that both *B. carambolae* and *B. papayae* were capable of infesting all the host fruits collected in this study, namely carambola, guava, jambu madu, jambu air, and papaya. It was observed that all the collected host fruits had both *B. carambolae* and *B. papayae* emerging from them, confirming that these two pest fruit flies were indeed polyphagous and pose detrimental problems to the agricultural industry if not kept in check.

Utilizing three mitochondrial DNA molecular markers, COI, COII, and *cytb*, phylogenetic analyses have shown that *B. carambolae* and *B. papayae* were indeed very closely related. The various phylogenetic trees generated based on maximum-likelihood, Bayesian Inference, and Neighbor-Joining analyses showed that the two species of fruit flies tend to group together within the same clade. Phylogenetic analyses based on all three molecular markers, as well as an analysis involving a combination of all three molecular markers, all show similar grouping of *B. carambolae* and *B. papayae*. In other words, all the phylogenetic analyses conducted showed that no distinct clades were formed that could clearly distinguish *B. carambolae* and *B. papayae*. The same thing can be said with the haplotype analysis dataset as well, which also support the phylogenetic analyses dataset in not being able to clearly distinguish between *B. carambolae* and *B. papayae*. Haplotype network reconstruction analysis has shown that *B. carambolae* and *B. papayae* also tend to group within the same haplotype groups. These findings suggest that *B. carambolae* and *B. papayae* could belong to the

same species. The molecular markers also could not distinguish *B. carambolae* and *B. papayae* according to host fruit specificity. *B. carambolae* and *B. papayae* specimens, regardless of their host fruit origin, grouped together within the same clade with no distinct formation of clades according to host fruit specificity. The grouping of *B. carambolae* and *B. papayae* specimens from Sarawak separately from the Serdang specimens was concluded to be based on geographical factors, and not due to host fruit specificity.

Large scale population studies of *B. carambolae* and *B. papayae* from different locations is important to understand the taxonomy status of these two species of fruit flies. More variety of host fruits should be collected from East Malaysia or Borneo to better understand the host fruit specificity as well as to compare the genetic structure of *B. carambolae* and *B. papayae* in that region with that of Peninsular Malaysia. Traditional morphological methods should be utilized together with molecular methods to better understand the continuous morphological and genetic variance of *B. carambolae* and *B. papayae* on a wider geographic scale. The amalgamation of genetic and morphological data will give a better understanding as to the taxonomic status of *B. carambolae* and *B. papayae*, and to determine if a taxonomic revision of the *B. dorsalis* species complex is required or otherwise. Such a revision of the taxonomic status of *B. carambolae* and *B. papayae* plays an important role in the management of pests and fruit quarantine measures.

In conclusion, *B. carambolae* and *B. papayae* could not be clearly distinguished using COI, COII, and *cytb* molecular markers. *B. carambolae* and *B. papayae* have been shown to have no host fruit specificity. *B. carambolae* and *B. papayae* hatched from host fruits collected from Sarawak have been shown to be genetically different from *B. carambolae* and *B. papayae* hatched from host fruits collected from Serdang. Molecular methods alone cannot be utilized to identify *B. carambolae* and *B. papayae*, rather they

should go hand-in-hand with traditional morphological methods in order to accurately ascertain the identities of the two fruit flies.

REFERENCES

- Akaike, H. (1973). Information theory and an extension of the maximum likelihood principal. In Petrov, B. N. & Caski, F. (eds.), *Second International Symposium on Information Theory* (pp: 267-281). Budapest: Akademia Kiado.
- Aketerawong, N., Chinvinijkul, S., Orankanok, W., Guglielmino, C. R., Franz, G., Malacrida, A. R. & Thanaphum, S. (2011). The utility of microsatellite DNA markers for the evaluation of area-wide integrated pest management using SIT for the fruit fly, *Bactrocera dorsalis* (Hendel), control programs in Thailand. *Genetica*, 139: 129–140.
- Allwood, A. J., Chinajariyawong, A., Drew, R. A. I., Hamacek, E. L., Hancock, D. L., Hengsawad, C., Jipanin, J. C., Jirasurat, M., Kong Krong, C., Kritsaneepaiboon, S., Leong, C. T. S. & Vijaysegaran, S. (1999). Host plant records for fruit flies (Diptera: Tephritidae). In *South East Asia. The Raffles Bulletin of Zoology Supplement No. 7*.
- Altshuler, M. L. (2006). *PCR troubleshooting: The Essential Guide*. Caister Academic Press.
- Armstrong, K. F. & Ball, S. L. (2005). DNA barcodes for biosecurity: invasive species identification. *Philosophical Transactions of the Royal Society B*, 360: 1813-1823.
- Asokan, R., Krishna Kumar, N. K. & Verghese, A. (2007). Molecular identification of fruit flies, *Bactrocera* spp. (Diptera: Tephritidae) using mitochondrial cytochrome oxidase I. *Current Science*, 93(12): 1668-1669.
- Avise, J. C. & Lansman, R. A. (1983). Polymorphism of mitochondrial DNA in populations of higher animals. In Nei, M. & Koehn, R. K. (eds.), *Evolution of Genes and Proteins* (pp: 147-164). Sunderland: Sinauer.
- Avise, J. C. (1986). Mitochondrial DNA and the evolutionary genetics of higher animals. *Philosophical Transactions of the Royal Society of London Series B*, 312:325-342.
- Avise, J. C. (1994). *Molecular markers, natural history, and evolution*. New York: Chapman & Hall.

- Awise, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J. E., Reeb, C. A. & Saunders, N. C. (1987). Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematic. *Annual Review of Ecology and Systematics*, 18: 489-522.
- Ball, S. L. & Armstrong, K. F. (2006). DNA barcodes for insect pest identification: a test case with tussock moths (Lepidoptera: Lymantriidae). *Canadian Journal of Forest Research*, 36(2): 337-350.
- Baliraine, F. N., Bonizzoni, M., Osir, E. O., Lux, S. A., Mulaa, F. J., Quilici, S., Gomulski, L. M., Gasperi, G. & Malacrida, A. R. (2002) Medfly microsatellite markers for species diagnosis and population genetic analysis in three tephritid fruit fly species. In *Proceedings of the 6th International symposium on fruit flies of economic importance*, 6-10 May 2002, Stellenbosch, South Africa.
- Blankenship, J. E. (2007). Electrophoresis.
http://www.accessmylibrary.com/coms2/summary_0193-9034_ITM. Accessed on 28th February 2011.
- Brown, W. M. (1985). The mitochondrial genome of animals. In MacIntyre, R. J. (ed.), *Molecular Evolutionary Genetics* (pp: 95-130). New York: Plenum Press.
- Brown, W. M., Prager, E. M., Wang, A. & Wilson A. C. (1982). Mitochondrial DNA sequences of primates: Tempo and mode of evolution. *Journal of Molecular Evolution*, 18: 225-239.
- Caterino, M. S. & Tishechkin, A. K. (2006). DNA identification and morphological description of the first confirmed larvae of Heteriinae (Coleoptera: Histeridae). *Systematic Entomology*, 31: 405-418.
- Cheng, A., Zhang, M., Okubo, M., Omichi, K. & Saltiel, A. R. (2009). Distinct mutations in the glycogen debranching enzyme found in glycogen storage disease type III lead to impairment in diverse cellular functions. *Human Molecular Genetics*, 18(11): 2045-2052.
- Choi, M., Scholl, U. I., Ji, W., Liu, T., Tikhonova, I. R., Zumbo, P., Nayir, A., Bakkaloglu, A., Özen, S., Sanjad, S., Nelson-Williams, C., Farhi, A., Mane, S. & Lifton, R. P. (2009). Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proceedings of the National Academy of Sciences of the United States of America*, 106(45): 19,096-19,101.

- Christenson, L. D. & Foote, R. H. (1960). Biology of fruit flies. *Annual Review of Entomology*, 5: 171- 192.
- Chua, T. H. (1991). Comparison of demographic parameters in wild *Bactrocera* sp. (Malaysian A) (Diptera: Tephritidae) from different hosts. *Journal of Plant Protection in the Tropic*, 8(3): 161-166.
- Chua, T. H., Chong, Y. V. & Lim, S. H. (2009). Species determination of Malaysian *Bactrocera* pests using PCR-RFLP analyses (Diptera: Tephritidae). *Pest Management Science*, 66: 379–384.
- Clarke, A. R., Armstrong, K. F., Carmichael, A. E., Milne, J. R., Raghu, S., Roderick, G. K. & Yeates, D. K. (2005). Invasive phytophagous pests arising through a recent tropical evolutionary radiation: the *Bactrocera dorsalis* complex of fruit flies. *Annual Review of Entomology*, 50: 293-319.
- Clement, M., Posada, D. & Crandall, K. A. (2000). TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, 9(10): 1657-1660.
- Cook, C. E., Wang, Y. & Sensabaugh, G. (1999). A mitochondrial control region and cytochrome b phylogeny of sika deer (*Cervus nippon*) and report of tandem repeats in the control region. *Molecular Phylogenetics and Evolution*, 12: 47-56.
- Crozier, R. H., Crozier, Y. G. & Mackinlay, A. G. (1989). The CO-I and CO-II region of honeybee mitochondrial DNA: evidence for variation in insect mitochondrial evolutionary rates. *Molecular Biology and Evolution*, 6: 399-411.
- Dreses-Werringloer, U., Lambert, J-C., Vingtdoux, V., Zhao, H., Vais, H., Siebert, A., Jain, A., Koppel, J., Rovelet-Lecrux, A., Hannequin, D., Pasquier, F., Galimberti, D., Scarpini, E., Mann, D., Lendon, C., Campion, D., Amouyel, P., Davies, P., Foskett, J. K., Campagne, F. & Marambaud, P. (2008). A polymorphism in CALHM1 influences Ca²⁺ homeostasis, A β levels, and Alzheimer's disease risk. *Cell*, 133(7): 1149-1161.
- Drew, R. A. I. (1989a). The taxonomy and distribution of tropical and subtropical Dacinae (Diptera: Tephritidae). In Robinson A. S. & Hooper, G. (eds.), *Fruit flies, their biology, natural enemies and control Vol. 3A* (pp: 9-14). Amsterdam: Elsevier.

- Drew, R. A. I. (1989b). The tropical fruit flies (Diptera: Tephritidae: Dacinae) of the Australasian and Oceanian regions. *Memoirs of the Queensland Museum*, 26: 1-521.
- Drew, R.A.I. & Hancock, D.L. (2000). Phylogeny of the tribe Dacini (Dacinae) based on morphological, distributional, and biological data. In Aluja, M. & Norrbom, A.L. (eds.), *Fruit flies (Tephritidae): phylogeny and evolution of behavior* (pp: 491-504). Boca Raton, Florida: CRC Press.
- Drew, R. A. I. (2004). Biogeography and speciation in the Dacini (Diptera: Tephritidae: Dacinae). In Evenhuis N. L. & Kaneshiro, K. Y. (eds.), *D. Elmo Hardy Memorial Volume. Contributions To the Systematics and Evolution of Diptera. Bishop Museum Bulletin in Entomology*, 12: 165-178.
- Drew, R. A. I. & Hancock, D. L. (1994). The *Bactrocera dorsalis* complex of fruit flies (Diptera: Tephritidae: Dacinae) in Asia. *Bulletin of Entomological Research Supplementary 2*: 1-68.
- Drew, R. A. I. & Raghu, S. (2002). The fruit fly fauna (Diptera: Tephritidae: Dacinae) of the rainforest habitat of the Western Ghats, India. *Raffles Bulletin of Zoology*, 50: 327-352.
- Drew, R. A. I. & Romig, M. C. (1997). Overview – Tephritidae in the Pacific and Southeast Asia. In Allwood, A. J. & Drew, R. A. I. (eds.), *Management of fruit flies in the pacific. ACIAR Proceedings 76* (pp: 46-53). Canberra: Australian Centre for International Agricultural Research
- Ebina, T. & Ohto, K. (2006). Morphological characters and PCR-RFLP markers in the interspecific hybrids between *Bactrocera carambolae* and *B. papayae* of the *B. dorsalis* species complex (Diptera: Tephritidae). *Research Bulletin of the Plant Protection Service Japan*, 42: 23-34.
- Edwards, A. W. F. & Cavalli-Sforza, L. L. (1964). Reconstruction of evolutionary trees. In Heywood, W. H. & McNeill, J. (eds.), *Phenetic and phylogenetic classification* (pp: 67-76). London: Systematics Association Publication No. 6.

- Enkerlin, W. R., Bakri, A., Cáceres, C., Cayol, J-P., Franz, G., Parker, A., Robinson, A., Vreysen, M. & Hendrichs, J. (2003). Insect pest intervention using the sterile insect technique: current status on research and on operational programs in the world. In, *Recent trends on sterile insect technique and area-wide integrated pest management. Economic feasibility, control projects, farmer organization and Bactrocera dorsalis complex control study* (pp: 11-24). Okinawa: Research Institute for Subtropics.
- Felsenstein, J. (1978). Cases in which parsimony or compatibility methods will be positively misleading. *Systematic Zoology*, 27: 401-410.
- Felsenstein, J. (1981). Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of Molecular Evolution*, 17(6): 368-376.
- Fletcher, B. S. (1989). Life history strategies of the fruit flies. In Robinson, A. S. & Hooper, G. (eds.), *Fruit flies, their biology, natural enemies and control Volume 3B* (pp: 195-208). Amsterdam: Elsevier.
- Floyd, R., Abebe, E., Papart, A. & Baxter, M. (2002). Molecular barcodes for soil nematode identification. *Molecular Ecology*, 11(4): 839-850.
- Gascuel, O. & Steel, M. (2006). Neighbor-joining revealed. *Molecular Biology and Evolution*, 23(11): 1997-2000.
- Gilks, W. R., Richardson, S. & Spiegelhater, D. J. (1996). *Markov Chain Monte Carlo in Practice*. London: Chapman & Hall.
- Graham, C. A. & Hill, A. J. M. (2001). Introduction to DNA sequencing. In Graham, C. A. & Hill, A. J. M. (eds.), *Methods in molecular biology, vol. 167: DNA sequencing protocols, 2nd edition*. Totowa, New Jersey: Humana Press Inc.
- Hajibabaei, M., Singer, G. A. C., Clare, E. L. & Hebert, P. D. N. (2007). Design and applicability of DNA arrays and DNA barcodes in biodiversity monitoring. *BMC Biology*, 5:24.
- Han, H-Y. & Ro, K-E. (2009). Molecular phylogeny of the family Tephritidae (Insecta: Diptera): new insight from combined analysis of the mitochondrial 12S, 16S, and COII genes. *Molecules and Cells*, 27: 55-66.

- Hardy, D. E. (1969). Taxonomy and distribution of the oriental fruit fly and related species (Tephritidae-Diptera). *Proceedings of the Hawaiian Entomological Society*, 20: 395-428.
- Hardy, D. E. (1973). The fruit flies (Tephritidae-Diptera) of Thailand and bordering countries. *Pacific Insects Monograph*, 31: 1-353.
- Hardy, D. E. (1974). The fruit flies of the Philippines (Diptera: Tephritidae). *Pacific Insects Monograph*, 32: 1-266.
- Hardy, D. E. & Adachi, M. (1956). Insects of Micronesia Diptera: Tephritidae. *Insects of Micronesia*, 14(1): 1-28.
- Hatadani, L. M., McInerney, J. O., de Medeiros, H. F., Junqueira, A. C. M., de Azeredo-Espin, A. & Klaczko, L. B. (2009). Molecular phylogeny of the *Drosophila tripunctata* and closely related species groups (Diptera: Drosophilidae). *Molecular Phylogenetics and Evolution*, 51(3): 595-600.
- Hatefi, Y. (1985). The mitochondrial electron transport and oxidative phosphorylation system. *Annual Review of Biochemistry*, 54: 1015-1069.
- Hawaii Area-Wide Fruit Fly Integrated Pest Management (2002). Fruit Fly Identification and Life Cycle. www.extento.hawaii.edu Accessed on 23rd August 2009.
- Hedrick, P. W. (2011). *Genetics of Populations*. Sudbury, Massachusetts: Jones & Bartlett Learning.
- Hendrichs, J., Kenmore, P., Robinson, A. S. & Vreysen, M. J. B. (2007). Area-wide integrated pest management (AW-IPM): principle, practice and prospects. In Vreysen, M. J. B., Hendrichs, J. & Robinson, A. S. (eds.), *Area-wide control of insect pests: from research to field implementation* (pp: 3-33). Dordrecht: Springer.
- Higgins, D. G. & Sharp, P. M. (1988). CLUSTAL: a package for performing multiple sequence alignment on a computer. *Gene*, 73: 237-244.

- Huelsenbeck, J. P. & Ronquist, F. (2001). MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics Application Note*, 17(8): 754-755.
- Huelsenbeck, J. P., Ronquist, F., Nielsen, R. & Bollback, J. P. (2001). Bayesian Inference of Phylogeny and Its Impact on Evolutionary Biology. *Science*, 294(5550): 2310-2314.
- Hurst, G. D. D. & Jiggins, F. M. (2005). Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. *Proceedings of the Royal Society B*, 272: 1525–1534.
- Huson, D. H., Rupp, R. & Scornavacca, C. (2010). *Phylogenetic networks: concepts, algorithms, and applications*. Cambridge: Cambridge University Press.
- Hu, J., Zhang, J. L., Nardi, F. & Zhang, R. J. (2008). Population genetic structure of the melon fly, *Bactrocera cucurbitae* (Diptera: Tephritidae), from China and Southeast Asia. *Genetica*, 134: 319-324.
- Irwin, D.M., Kocher, T. D. & Wilson, A. C. (1991) Evolution of the cytochrome b gene in mammals. *Journal of Molecular Evolution*, 32: 128-144.
- Ito, M., Jiang W., Sato, J. J., Zhen, Q., Jiao, W., Goto, K., Sato, H., Ishiwata, K., Oku, Y., Chai, J-J. & Kamiya, H. (2010). Molecular phylogeny of the subfamily Gerbillinae (Muridae, Rodentia) with emphasis on species living in the Xinjiang-Uyghur autonomous region of China and based on the mitochondrial cytochrome b and cytochrome c oxidase subunit II genes. *Zoological Science*, 27: 269-278.
- Iwahashi, O. (1999). Distinguishing between the two sympatric species *Bactrocera carambolae* and *B. papayae* (Diptera: Tephritidae) based on aedeagal length. *Annals of the Entomological Society of America*, 92(5): 639-643.
- Iwahashi, O. (2001). Aedeagal length of the oriental fruit fly, *Bactrocera dorsalis* (Hendel) (Diptera: Tephritidae), and its sympatric species in Thailand and the evolution of a longer and shorter aedeagus in the parapatric species of *B. dorsalis*. *Applied Entomology and Zoology*, 36: 289-297.

- Iwaizumi, R., Kaneda, M. & Iwahashi, O. (1997). Correlation of length of terminalia of males and females among nine species of *Bactrocera* (Diptera: Tephritidae) and differences among sympatric species of the *B. dorsalis* complex. *Annals of the Entomological Society of America*, 90: 664-666.
- Jamnongluk, W., Baimai, V. & Kittayapong, P. (2003). Molecular evolution of tephritid fruit flies in the genus *Bactrocera* based on the cytochrome oxidase I gene. *Genetica*, 119: 19-25.
- Jobb, G., von Haeseler, A. & Strimmer, K. (2004). Treefinder: a powerful graphical analysis environment for molecular phylogenetics. *BMC Evolutionary Biology*, 4: 18.
- Jukes, T. H. & Cantor, C. R. (1969). Evolution of protein molecules. In Munro, M. N. (ed.), *Mammalian protein metabolism* (pp: 21-132). New York: Academic Press.
- Kakouli-Duarte, T., Casey, D. G. & Burnell, A. M. (2001). Development of a Diagnostic DNA Probe for the Fruit Flies *Ceratitis capitata* and *Ceratitis rosa* (Diptera: Tephritidae) Using Amplified Fragment-Length Polymorphism. *Journal of Economic Entomology*, 94(4): 989-997.
- Karp, A., Seberg, O. & Buiatti, M. (1996). Molecular techniques in the assessment of botanical diversity. *Annals of Botany*, 78: 143-149.
- Karp, A., Edwards, K. J., Bruford, M., Funk, S., Vosman, B., Morgante, M., Seberge, O., Kremer, A., Boursot, P., Arctander, P., Tautz, D. & Hewitt, G. M. (1997a). Molecular technologies for biodiversity evaluation: opportunities and challenges. *Nature Biotechnology*, 15: 625-628.
- Karp, A., Kresovich, C., Bhat, K. V., Ayad, W. G. & Hodgkin, T. (1997b). Molecular tools in plant genetic resources conservation: a guide to the technologies. *IPGRI Technical Bulletin No. 2*. Rome: International Plant Genetic Resource Institute.
- Kimura, M. (1980). A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, 16(2): 111-120.

- Kirkness, E. F., Bafna, V., Halpern, H. L., Levy, S., Remington, K., Rusch, D. B., Delcher, A. L., Pop, M., Wang, W., Fraser, C. M. & Venter, J. C. (2003). The dog genome: survey sequencing and comparative analysis. *Science*, 301(5641): 1898-1903.
- Klug, W. S., Cummings, M. R. & Spencer C. A. (2006). *Concepts of Genetics* (8th ed.), (pp: 300, 567-568). Upper Saddle River, New Jersey: Pearson Education International.
- Krafsur, E. S. (2005). Role of population genetics in the sterile insect technique. In Dyck, V. A., Hendrichs, J. & Robinson, A. S. (eds.), *Sterile insect technique: principles and practice in area-wide integrated pest management* (pp: 389-406). Dordrecht: Springer.
- Krosch, M. N., Schutze, M. K., Armstrong, K. F., Graham, G. C., Yeates, D. K. & Clarke, A. R. (2012a). A molecular phylogeny for the Tribe Dacini (Diptera: Tephritidae): Systematic and biogeographic implications. *Molecular Phylogenetics and Evolution*, 64(3): 513-523.
- Krosch, M. N., Schutze, M. K., Armstrong, K. F., Boontop, Y., Boykin, L. M., Chapman, T. A., Englezou, A., Cameron, S. L. & Clarke, A. R. (2012b). Piecing together an integrative taxonomic puzzle: microsatellite, wing shape and aedeagus length analyses of *Bactrocera dorsalis* s.l. (Diptera: Tephritidae) find no evidence of multiple lineages in a proposed contact zone along the Thai/Malay Peninsula. *Systematic Entomology*. doi: 10.1111/j.1365-3113.2012.00643.x
- Lanave, C., Preparata, G., Sacone, C. & Serio, G. (1984). A new method for calculating evolutionary substitution rates. *Journal of Molecular Evolution*, 20(1): 86-93.
- Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J. & Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics*, 23(21): 2947-2948.
- Lefort, M-C., Boyer, S., Worner, S. P. & Armstrong K. (2012). Noninvasive molecular methods to identify live scarab larvae: an example of sympatric pest and nonpest species in New Zealand. *Molecular Ecology Resources*, 12(3): 389-395.

- Lewis, R. L., Beckenbach, A. T. & Mooers, A. Ø. (2005). The phylogeny of the subgroups within the melanogaster species group: likelihood tests on COI and COII sequences and a Bayesian estimate of phylogeny. *Molecular Phylogenetics and Evolution*, 37: 15-24.
- Li, Y., Wu, Y., Chen, H., Wu, J. & Li, Z. (2011). Population structure and colonization of *Bactrocera dorsalis* (Diptera: Tephritidae) in China, inferred from mtDNA COI sequences. *Journal of Applied Entomology*, 136: 241-251.
- Liess, M. & Schulz, R. (1998). Linking insecticide contamination and population response in an agricultural stream. *Environmental Toxicology and Chemistry*, 18(9): 1948-1955.
- Lim, P-E., Tan, J., Suana, I. W., Eamsobhana, P., Yong, H. S. (2012). Distinct genetic lineages of *Bactrocera caudata* (Insecta: Tephritidae) revealed by COI and 16S DNA sequences. *PLoS ONE*, 7(5): e37276. doi:10.1371/journal.pone.0037276
- Lindquist, D.A. (2000). Pest management strategies: area-wide and conventional. In Tan, K. H. (ed.), *Area-wide control of fruit flies and other insect pests* (pp: 13-19). Penang: Penerbit Universiti Sains Malaysia.
- Liu, J., Shi, W. & Ye, H. (2007). Population genetics analysis of the origin of the oriental fruit fly, *Bactrocera dorsalis* Hendel (Diptera: Tephritidae), in northern Yunnan province, China. *Entomological Science*, 10:11-19.
- Liu, S-S., Zhang, G-F., Wu, Q., Zhang, A-B., Wang, J-J. & Wan, F-H. (2012). Establishment and application of DNA barcoding technology for identification of the immature and adult debris of *Bactrocera dorsalis* (Hendel) (Diptera: Tephritidae). *Acta Entomologica Sinica*, 55(3): 336-343.
- Lunt, D. H., Zhang, D. X., Szymura, J. M. & Hewitt, G. M. (1996). The insect cytochrome oxidase I gene: evolutionary patterns and conserved primers for phylogenetic studies. *Insect Molecular Biology*, 5: 153-165.
- Maxam, A. M. & Gilbert, W. (1977). A method for sequencing DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 74: 560-564.
- McPherson, M. J. & Moller, S. G. (2006). *PCR* (2nd Edition). England: Taylor & Francis Group.

- Meyer, A. (1994). Shortcomings of the cytochrome b gene as a molecular marker. *TREE*, 9: 278-280.
- Moriyama, E. N. & Powell, J. R. (1997). Synonymous substitution rates in *Drosophila*: mitochondrial versus nuclear genes. *Journal of Molecular Evolution*, 45: 378-391.
- Morlais, I. & Severson, D. W. (2002). Complete mitochondrial DNA sequence and amino acid analysis of the cytochrome c oxidase subunit I (COI) from *Aedes aegypti*. *DNA sequence*, 13(2): 123-127.
- Muraji, M. & Nakahara, S. (2001). Phylogenetic relationships among fruit flies, *Bactrocera* (Diptera: Tephritidae), based on the mitochondrial rDNA sequences. *Insect Molecular Biology*, 10(6): 549-559.
- Muraji, M. & Nakahara, S. (2002). Discrimination among pest species of *Bactrocera* (Diptera: Tephritidae) based on PCR-RFLP of the mitochondrial DNA. *Applied Entomology and Zoology*, 37(3): 437-446.
- Nakahara, S., Kobashigawa, Y. & Muraji, M. (2008). Genetic variations among and within populations of the Oriental fruit fly, *Bactrocera dorsalis* (Diptera; Tephritidae), detected by PCR-RFLP of the mitochondrial control region. *Applied Entomology and Zoology*, 43(3): 457-465.
- Nakahara, S. & Muraji, M. (2008). Phylogenetic analyses of *Bactrocera* fruit flies (Diptera: Tephritidae) based on nucleotide sequences of the mitochondrial COI and COII genes. *Research Bulletin of the Plant Protection Service Japan*, 44: 1-12.
- Narayanan, E. S. & Batra, H. N. (1960). *Fruit flies and their control* (pp: 1-68). New Delhi, India: Indian Council of Agricultural Research.
- Newton, C. R. & Graham, A. (1997). *PCR* (2nd ed.). Oxford: BIOS Scientific Publishers.

- Nishikawa, K., Matsui, M., Yong, H. S., Ahmad, N., Yambun, P., Belabut, D. M., Sudin, A., Hamidy, A., Orlov, N. L., Ota, H., Yoshikawa, N., Tominaga, A. & Shimada, T. (2012). Molecular phylogeny and biogeography of caecilians from Southeast Asia (Amphibia, Gymnophiona, Ichthyophiidae), with special reference to high cryptic species diversity in Sundaland. *Molecular Phylogenetics and Evolution*, 63: 714-723.
- Norrbom, A. L., Carroll, L. E., Thompson, F. C., White, I. M. & Freidberg, A. (1998). Systematic database of names. In Thompson, F. C. (ed.), *Fruit Fly Expert Identification System and Systematic Information Database. MYIA, Volume 9* (pp: 65-251). Leiden: Bakchuys Publishers.
- Ooi, C. S. (1991). Genetic variation in populations of two sympatric taxa in the *Dacus dorsalis* complex and their relative infestation levels in various fruit hosts. In Vijaysegaran, S. & Ibrahim, A. G. (eds.), *Proceedings, First International Symposium on Fruit Flies in the Tropics, March 1988* (pp: 71-80). Kuala Lumpur: Malaysian Plant Protection Society.
- Parker, P. G., Snow, A. A., Schug, M. D. & Fuerst, P. A. (1998). What molecules can tell us about populations: choosing and using a molecular marker. *Ecology*, 79: 361-382.
- Pelt-Perkuil, van E., van Belkum, A. & Hays, J. P. (2008). *Principles and Technical Aspects of PCR Amplification*. Springer.
- Perkins, M. V., Fletcher, M. T., Kitching, W., Drew, R. A. I. & Moore, C. J. (1990). Chemical studies of rectal gland secretions of some species of *Bactrocera dorsalis* complex of fruit flies (Diptera: Tephritidae). *Journal of Chemical Ecology*, 16: 2475-2487.
- Pevsner, J. (2009). *Bioinformatics and Functional Genomics* (2nd Ed). Wiley-Blackwell.
- Piffaretti, J., Vanlerberghe-Masutti, F., Tayeh, A., Clamens, A-L., Coeur d'Acier, A. & Jouselin, E. (2012). Molecular phylogeny reveals the existence of two sibling species in the aphid pest *Brachycaudus helichrysi* (Hemiptera: Aphididae). *Zoologica Scripta*, 41(3): 266-280.
- Prabhakar, C. S., Mehta, P. K., Sood, P., Singh, S. K., Sharma, P. & Sharma, P. N. (2012). Population genetic structure of the melon fly, *Bactrocera cucurbitae* (Coquillett) (Diptera: Tephritidae) based on mitochondrial cytochrome oxidase (COI) gene sequences. *Genetica*, 140(1-3): 83-91.

- Raghu, S. (2003). The autecology of *Bactrocera cacuminata* (Hering) (Diptera: Tephritidae: Dacinae): Functional significance of resources. Doctor of philosophy thesis (pp: 240). Brisbane: Griffith University.
- Rand, D. M. & Harrison, R. G. (1986). Mitochondrial DNA transmission genetics in crickets. *Genetics*, 114: 955-970.
- Raupach, M. J., Astrin, J. J., Hannig, K., Peters, M. K. & Stoeckle, M. Y. (2010). Molecular species identification of Central European ground beetles (Coleoptera: Carabidae) using nuclear rDNA expansion segments and DNA barcodes. *Frontiers in Zoology*, 7: 26.
- Rosenberg, M. S. (2009). Sequence alignment: Concepts and history. In Rosenberg, M. S. (ed.), *Sequence alignment: methods, models, concepts, and strategies* (pp: 1-22). California: University of California Press.
- Ruiz-Garcia, M. & Pinedo-Castro, M. O. (2010). Molecular systematics and phylogeography of the genus *Lagothrix* (Atelidae, Primates) by means of the mitochondrial COII gene. *Folia Primatol*, 81: 109-128.
- Sadeghi, S., Kyndt, T. & Dumont, H. J. (2010). Genetic diversity, population structure and taxonomy of *Calopteryx splendens* (Odonata: Calopterygidae): An AFLP analysis. *European Journal of Entomology*, 107: 137-146.
- Saitou, N. & Nei, M. (1987). The neighbour-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, 4(4): 406-425.
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain terminator inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*, 74: 5463-5467.
- Sattath, S. & Tversky, A. (1977). Additive similarity trees. *Psychometrika*, 42(3): 19-345.
- Sauers-Muller, van A. E. (1991). An overview of the Carambola fruit fly *Bactrocera* species (Diptera: Tephritidae), found recently in Suriname. *Florida Entomologist*, 74: 432-440.

- Schlötterer, C. (2004). The evolution of molecular markers – just a matter of fashion? *Nature Reviews Genetics*, 5: 63-69.
- Schroeder, H., Klotzbach, H., Elias, S., Augustin, C. & Pueschel, K. (2003). Use of PCR-RFLP for differentiation of calliphorid larvae (Diptera, Calliphoridae) on human corpses. *Forensic Science International*, 132: 76-81.
- Schutze, M. K., Krosch, M. N., Armstrong, K. F., Chapman, T. A., Englezou, A., Chomič, S. L., Hailstones, D. & Clarke, A. R. (2012). Population structure of *Bactrocera dorsalis* s.s., *B. papayae*, *B. philippinensis* (Diptera: Tephritidae) in southeast Asia: evidence for a single species hypothesis using mitochondrial DNA and wing-shape data. *BMC Evolutionary Biology*, 12(130).
- Segura, M. D., Callejas, C., Fernández, M. P. & Ochando, M. D. (2006). New contributions towards the understanding of the phylogenetic relationships among economically important fruit flies (Diptera: Tephritidae). *Bulletin of Entomological Research*, 96: 279-288.
- Segura, M. D., Callejas, C. & Ochando, M. D. (2008), *Bactrocera oleae*: a single large population in Northern Mediterranean basin. *Journal of Applied Entomology*, 132: 706–713.
- Shearman, D. C. A., Gilchrist, A. S., Crisafulli, D., Graham, G., Lange, C. & Frommer, M. (2006). Microsatellite markers for the pest fruit fly, *Bactrocera papayae* (Diptera: Tephritidae) and other *Bactrocera* species. *Molecular Ecology Notes*, 6: 4–7.
- Shi, W., Kerdelhué, C. & Ye, H. (2009). Population genetic structure of the oriental fruit fly, *Bactrocera dorsalis* (Hendel) (Diptera: Tephritidae) from Yunnan province (China) and nearby sites across the border. *Genetica*, 138(3): 377-385.
- Shi, W., Kerdelhué, C. & Ye, H. (2012) Genetic Structure and Inferences on Potential Source Areas for *Bactrocera dorsalis* (Hendel) Based on Mitochondrial and Microsatellite Markers. *PLoS ONE* 7(5): e37083.
doi:10.1371/journal.pone.0037083
- Shono, H. (2000). Efficiency of the finite correction of Akaike's information criteria. *Fisheries Science*, 66: 608-610.

- Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H. & Flook, P. (1994). Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annals of the Entomological Society of America*, 87: 651-701.
- Sinclair, E. A., Pérez-Losada, M. & Crandall, K. A. (2005). Molecular phylogenetics for conservation biology. In Purvis A., Gittleman, J. L. & Brooks, T. (eds.), *Conservation Biology 8: Phylogeny and Conservation* (pp: 19-56). New York: Cambridge University Press.
- Smith, P. T., McPherson, B. A. & Kambhampati, S. (2002). Phylogenetic analysis of mitochondrial DNA supports the monophyly of Dacini fruit flies (Diptera: Tephritidae). *Annals of the Entomological Society of America*, 95(6): 658-664.
- Smith, P. T., Kambhampati, S. & Armstrong, K. A. (2003). Phylogenetic relationships among *Bactrocera* species (Diptera: Tephritidae) inferred from mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution*, 26(1): 8-17.
- Spooner, D., van Treuren, R. & de Vicente, M. C. (2005). Molecular markers for genebank management. In *IPGRI Technical Bulletin No. 10*. Rome: International Plant Genetics Resources Institute.
- Su, B., Wang, Y-W., Lan, H., Wang, W. & Zhang, Y. (1999). Phylogenetic studies of complete cytochrome b genes in musk deer (Genus *Moschus*) using museum samples. *Molecular Phylogenetics and Evolution*, 15: 115-123.
- Swofford, D. L. (2002). PAUP*: Phylogenetic analysis using parsimony (*and other methods). Version 4. Sunderland, Massachusetts: Sinauer Associates.
- Tan, K-H. (2003). Interbreeding and DNA analysis of sibling species within the *Bactrocera dorsalis* complex. In *Recent Trends Sterile Insect Tech Area-Wide Integrated Pest Management – Economic Feasibility, Control Projects, Farmer Organization and Bactrocera dorsalis Complex Control Study*. (pp: 113-122). Okinawa: Research Institute For Subtropics.
- Tanabe, A. S. (2007). Kakusan: a computer program to automate the selection of a nucleotide substitution model and the configuration of a mixed model on multilocus data. *Molecular Ecology Notes*, 7: 962-964.

- Tang, R. W. K., Yau, C. & Ng, W. C. (2010). Identification of stomatopod larvae (Crustacea: Stomatopoda) from Hong Kong waters using DNA barcodes. *Molecular Ecology Resources*, 10: 439-448.
- Templeton, A. R. (2006). *Population genetics and microevolutionary theory*. Hoboken, New Jersey: John Wiley & Sons.
- Tsuruta, K. & White, I. M. (2001). Eleven new species of the genus *Bactrocera* Macquart (Diptera: Tephritidae) from Sri Langka. *Entomological Science*, 4: 69-87.
- Vijaysegaran, S. (1983). The occurrence of oriental fruit fly on starfruit in Serdang and the status of its parasitoids. *Journal of Plant Protections in the Tropics*, 1(2): 93-98.
- Vijaysegaran, S. (1996). Fruit Fly Research and Development in Tropical Asia. In Allwood, A. J. & Drew, R. A. I. (eds.), *Management of fruit flies in the pacific. A regional symposium*, 28th–31st October 1996, Nadi, Fiji.
- Walker, K. (2005). Carambola Fruit Fly (*Bactrocera carambolae*). <http://www.padil.gov.au>. Accessed on 7th February 2013.
- Wan, X., Nardi, F., Zhang, B. & Liu, Y. (2011). The Oriental Fruit Fly, *Bactrocera dorsalis*, in China: Origin and Gradual Inland Range Expansion Associated with Population Growth. *PLoS ONE*, 6(10): e25238. doi:10.1371/journal.pone.0025238
- Wan, X., Liu, Y. & Zhang, B. (2012). Invasion history of the oriental fruit fly, *Bactrocera dorsalis*, in the Pacific-Asia region: two main invasion routes. *PLoS ONE*, 7(5): e36176. doi:10.1371/journal.pone.0036176
- Wanwisa, J., Visut, B. & Pattamaporn, K. (2003). Molecular phylogeny of tephritid fruit flies in the *Bactrocera tau* complex using the mitochondrial COI sequences. *Genome*, 46: 112-118.
- Waterhouse, D. F. (1993). The major arthropod pests and weeds of agriculture in Southeast Asia. In *ACIAR Monograph 21* (pp: 141). Canberra, Australia: Australian Center for International Agricultural Research.

- Waterhouse, D. F. (1997). The Major Invertebrate Pests and Weeds of Agriculture and Plantation Forestry in the Southern and Western Pacific. *ACIAR Monograph 44*. Canberra, Australia: Australian Center for International Agricultural Research.
- White, I. M. (1996). Fruit fly taxonomy: recent advances and new approaches. In McPherson, B. A. & Steck, G. J. (eds.), *Fruit fly pests, a world assessment of their biology and management* (pp: 253-258). Del-ray Beach, Florida: St. Lucie Press.
- White, I. M. & Elson-Harris, M. M. (1992). *Fruit flies of economic significance: their identification and bionomics*. Wallingford: CAB International.
- Wu, Y., Li, Y., Ruiz-Arce, R., McPherson, B. A., Wu, J. & Li, Z. (2011). Microsatellite Markers Reveal Population Structure and Low Gene Flow Among Collections of *Bactrocera cucurbitae* (Diptera: Tephritidae) in Asia. *Journal of Economic Entomology*, 104(3): 1065-1074.
- Yang, Z. (2006). *Computational molecular evolution*. Oxford: Oxford University Press.
- Yong, H. S. (1995). Genetic differentiation and relationships in five taxa of the *Bactrocera dorsalis* complex (Insecta: Diptera: Tephritidae). *Bulletin of Entomological Research*, 85:431-435.
- Yu, H., Frommer, M., Robson, M. K. & Sved, J. (2000). A population analysis of the Queensland fruit fly *Bactrocera tryoni* using microsatellite markers. In Tan K. H. (ed.), *Area-wide control of fruit flies and other insect pests* (pp: 497-508). Penang, Malaysia: USM Press.
- Yu, D. J., Chen, Z. L., Zhang, R. J. & Yin, W. Y. (2005). Real-time qualitative PCR for the inspection and identification of *Bactrocera philippinensis* and *Bactrocera occipitalis* (Diptera: Tephritidae) using SYBR green assay. *The Raffles Bulletin of Zoology*, 53(1): 73-78.
- Yu, D. J., Xu, L., Nardi, F., Li, J. G. & Zhang, R. J. (2007). The complete nucleotide sequence of the mitochondrial genome of the oriental fruit fly, *Bactrocera dorsalis* (Diptera: Tephritidae). *Gene*, 396(1): 66-74.

- Zahran, M. M., El-Fandary, O. O. & Mahmoud, Y. A. (2009). Assessment of genetic variability and genotyping of *Ceratitis capitata* and *Bactrosera zonata* by molecular techniques. *Archives of Phytopathology and Plant Protection*, 42(1): 92–98.
- Zaslavsky, L. & Tatusova, T. A. (2008). Accelerating the neighbour-joining algorithm using the adaptive bucket data structure. In Mandoiu, I., Sunderraman, R. & Zelikovsky, A. (eds.), *Bioinformatics research and applications: Fourth International Symposium ISBRA 2008, Atlanta, GA, USA, May 6-9, 2008, Proceedings* (pp: 122-133). Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg.
- Zhang, B., Liu, Y. H., Wu, W. X. & Wang, Z. L. (2010). Molecular phylogeny of *Bactrocera* species (Diptera: Tephritidae: Dacini) inferred from mitochondrial sequences of 16s rDNA and COI sequences. *Florida Entomologist*, 93(3): 369-377.
- Zhang, D. X. & Hewitt G. M. (1997). Insect mitochondrial control region: a review of its structure, evolution and usefulness in evolutionary studies. *Biochemical Systematics and Ecology*, 25: 99–120.
- Zhu, Z-H., Ye, H. & Zhang, Z-Y. (2005a). Molecular identification of six *Bactrocera* species (Diptera: Tephritidae) based on mtDNA. *Acta Entomologica Sinica*, 48(3): 386-390.
- Zhu, Z-H., Ye, H. & Zhang, Z-Y. (2005b). Genetic relationships among four *Bactrocera cucurbitae* geographic populations in Yunan province. *Chinese Journal of Applied Ecology*, 16: 1889–1992.