PHYLOGENETICS OF SUBFAMILY MURIN EA (RODENTIA: MURIDAE) IN MALAYSIA INFERRED FROM MITOCHONDRIAL AND NUCLEAR GENES

Nur Aida Binti Md Tamrin

Master of Science
2011
PHYLOGENETICS OF SUBFAMILY MURINEA (RODENTIA: MURIDAE) IN MALAYSIA INFERRED FROM MITOCHONDRIAL AND NUCLEAR GENES

NUR AIDA BINTI MD TAMRIN

A thesis submitted
in fulfillment of the requirement for the Degree of
Master of Science
(Molecular Ecology and Genetic)

Faculty of Resources Sciences and Technology
UNIVERSITI MALAYSIA SARAWAK
2011
DECLARATION

I hereby declare that no portion of the work referred to this thesis has been submitted in support of an application for another degree or qualification to this or any other university or institute of higher learning.

(Nur Aida binti Md Tamrin)
Acknowledgements

Bismillahhirrahmannirahim. Alhamdulillah, I am forever grateful to Allah SWT for giving me strength and determination, health and faith, granted me with great and comfort surroundings, valuable knowledge, loving family and friends during these years. Finally, I have completed my research for my Master's degree. I am indebted to my supervisor, Prof. Dr. Mohd Tajuddin Abdullah for the guidance, advices, motivations, constructive comments, concerns and supports without limits throughout this study.

This study would not be possible without various administrative and financial supports from UNIMAS. I am thankful to Vice Chancellor of UNIMAS for granting me the Vice Chancellor Scholarship (UNIMAS 2008/09) for two years of my study. I would like to thank Sarawak Forestry Corporation and Sarawak Forestry Department for granting permission to conduct wildlife research in national parks in Sarawak under the State Wild Life Protection Rules 1998; for research permit number NPW.907.4.2(II)-5 and permit to enter park. I would also like to thank Department of Wildlife and National Parks for inviting my team to various field trips in Peninsular Malaysia. All the sweet memories working with dearest colleagues in the field trips will always remain.

Special thanks go to Dr. Lim Boo Liat for assisting me in identifying Rattus species during taxidermy workshop held in UNIMAS and to my colleague, Nurhaika Atikah Azmi, who made some of the RAG 2 sequences available at the Molecular Ecology Laboratory (MEL), UNIMAS. I wish to express my sincere gratitude to the members in
Department of Zoology; All lecturers for their concerns, Mr Besar Ketol, Mr. Isa Sait, Mr. Huzal Irwan Husin, Mr. Wahap Marni and Mr. Mohd Jalani Mortada for their hard works and assistances throughout this study. To MEL seniors; Roberta Chaya Tawie Tingga, Nurhaliza Hassan, Mohd Ridwan Abd Rahman, Anang Setiawan Achmadi, Sigit Wiantoro, Mohd Fizl Sidq Ramji, Nur Salmizar Azmi and Eileen Lit, I wish to thank them for being remarkable mentors for me since 2008 until the very end of my study. Big thanks also go to MEL colleagues and dearest friends; Zahirunisa Abd Rahim, Mohd. Hanif Ridzuan Mat Daud, Wan Nurainie Wan Ismail, Mohd Shahfiz Azman, Nurul Ashikeen Abd Razak, Siti Zuriani Ismail, Irene Cristianus, Mohd Isham Azhar, Muhammad Ikhwan Idris, Madinah Adrus, Nurhafidah Afandi and Faezah Abdullah for the wonderful friendship, for the great times, shared knowledge and support. I am grateful to have Zahie and Hanif as my best friends, for being great companions, lending their sympathetic ears through my student’s life and helping each other through good and bad times.

I would like to show my biggest gratitude to my beloved parents, Md. Tamrin Aliman and Fatimah Jamal for the endless encouragement, supports and prayers while waiting patiently for me to finish my Master’s degree. I would also like to thank my siblings; Iran Irwan Md Tamrin and Fara Dilla Md Tamrin together with their families for always have faith in me.
Abstract

Traditionally, taxonomic status of Murinae was based on morphological characteristics. The classification was always outdated due to the variation of morphological traits caused by rapid adaptation towards ecological habitats and high rate of evolution in Murinae. The variation of external features sometimes does not indicate the species to be in distinct taxa, at least not in Genetic Species Concept. As closely related species in the subfamily Murinae are morphologically similar to each other, the taxonomic status of Murinae is poorly resolved up until recent. Many studies have been done using genetic data, morphology, immunology, albumin and karyotypic analyses but the information of Murinae in Malaysia is still lacking. Phylogenetic relationship of Murinae were reviewed in this study inferred by mitochondrial gene (mtDNA) of cytochrome oxidase I (COI) and cytochrome b (cyt b) as well as one nuclear gene (nucDNA) of recombination activating gene (RAG2). The evolutionary, interspecific and intraspecific relationship between selected species in the Murinae and their divergence time were investigated during this study. The phylogenetic trees were reconstructed using four methods for each gene, namely, neighbour-joining (NJ), maximum parsimony (MP), maximum-likelihood (ML) and Bayesian methods. From this study, six monophyletic lineages were observed, namely, Rattus, Berylmis, Maxomys, Sundamys, Leopoldamys and Niviventer. The interrelationship of Murinae was incongruent between the genes analyses and some of the species were unresolved in the groupings. This study also found that there were cryptic species within M. whiteheadi and M. ochraceiventer population with external
morphology similar to *M. whiteheadi* but separated with high genetic divergence whilst closely related by genetic distance to *M. ochraceiventer*. Genetically, two types (Type A and Type B) of *M. rajah*, *R. rattus* and *S. muelleri* were also found in this study by observing the phylogenetic trees. NJ and Bayesian methods gave the most resolved topologies. MP and ML methods might give better resolution if more data were added. Cyt *b* was observed to be the best gene to study the evolution of Murinae.

Keywords: phylogenetic, Murinae, COI, RAG2, cyt *b*, Bayesian, cryptic species
Filogenetik subfamili Murinae (Rodentia: Muridae) di Malaysia disimpulkan dari gen mitokondria dan nuklear

Abstrak


Kata kunci: filogenetik, Murinae, COI, RAG2, cyt b, Bayesian, spesies samar
Table of Contents

Declaration ii
Acknowledgements iii
Abstract v
Abstrak vii
Table of contents ix
List of figures xiv
List of tables xx
Abbreviations xxiii

CHAPTER ONE General Introduction

1.1 Rodents of Family Muridae (Grey, 1821) 1
1.2 Sundaland 3
1.3 Cryptic species 4
1.4 Genetic species 5
1.5 Rationale 6
1.6 Problem statement and hypothesis 7
1.7 Objectives 8
1.8 Thesis organisation 8
# CHAPTER TWO  Literature Review

## 2.1 Subfamily Murinae

- **2.1.1** Morphology of subfamily Murinae
- **2.1.2** Behavior of subfamily Murinae
- **2.1.3** Ecological roles and contribution of murine rodents
- **2.1.4** Classification and taxonomy of subfamily Murinae
- **2.1.5** Phylogenetic relationship of subfamily Murinae
- **2.1.6** Distribution and conservation status
- **2.1.7** Divergence time of Murinae

## 2.2 Study taxa

- **2.2.1** *Rattus*
- **2.2.2** *Sundamys*
- **2.2.3** *Berylmys*
- **2.2.4** *Maxomys*
- **2.2.5** *Leopoldamys*
- **2.2.6** *Niviventer*

## 2.3 Cryptic species

## 2.4 Phylogenetic inference

## 2.5 Mitochondrial DNA (mtDNA)

- **2.5.1** Cytochrome oxidase I (COI)
- **2.5.2** Cytochrome b (cyt b)
CHAPTER THREE Materials and Methods

3.1 Study Area

3.1.1 Sampling site

3.1.2 Trapping

3.1.3 Specimen identification

3.1.4 Tissue and blood sample collection

3.1.5 Sample preservation

3.2 Molecular methodology

3.2.1 Laboratory materials and solutions

3.2.2 DNA extraction

3.2.3 DNA amplification

3.2.4 DNA visualisation

3.2.5 Purification and sequencing

3.3 Molecular analyses

3.3.1 Sequence analyses

3.3.2 Phylogenetic analyses

3.4 Skull extraction and measurements
CHAPTER FOUR  Results

4.1 Taxonomic list of species captured  58
4.2 Mitochondrial gene analyses  60
  4.2.1 Cytochrome Oxidase I (COI)  60
  4.2.2 Cytochrome b (cyt b)  76
4.3 Nuclear Gene Analyses  91
  4.3.1 Recombinant Activating Gene 2 (RAG2)  91
4.4 Skull and dental measurements  105

CHAPTER FIVE  Discussion

5.1 General discussion  106
5.2 Phylogenetic relationship of Murinae using mitochondrial gene  111
  5.2.1 Cytochrome Oxidase I (COI)  111
  5.2.2 Cytochrome b (cyt b)  120
5.3 Phylogenetic relationship of Murinae using nuclear gene  126
  5.3.1 Recombinant Activating Gene 2 (RAG2)  126
5.4 Skull and dental comparison  130
CHAPTER SIX General Conclusion and Recommendation

6.1 Conclusion
6.2 Limitations during this study
6.3 Recommendation

References

Appendix A List of sample used in phylogenetic analyses.
Appendix B Pairwise distance for COI gene by species.
Appendix C Pairwise distance for cyt b gene by species.
Appendix D Pairwise distance for RAG2 gene by species.
Appendix E Summary of variable sites of COI.
Appendix F Summary of variable sites of cyt b.
Appendix G Summary of variable sites of RAG2.
Appendix H Pictures of subfamily Murinae.
List of Figures

Figure 2.1 Mitochondrial DNA map. 30

Figure 3.1 Thirteen study sites for this study. 44

Figure 4.1 Plot of transition and transversion against divergence using Kimura (1980) distance method shows there is a little saturation occurred in COI gene region. 62

Figure 4.2 Phylogenetic tree of 12 species of subfamily Murinae based on 476 bp partial COI mtDNA gene sequences. The phylogeny is a single tree recovered using NJ analysis with C. notatus as an outgroup. Values on the branches represent NJ bootstrap estimates, based on 1000 replicates. Only bootstrap values >50% are shown. 67

Figure 4.3 Rooted (with outgroup) MP tree based on nucleotide data set of 476 bp partial COI mtDNA gene (tree length=682; CI=0.3959; HI= 0.6041; RC=0.3083; RI=0.7789). Values on the branches represent MP bootstrap estimates, based on 1000 replicates. Only bootstrap values >50% are shown. 68
Figure 4.4 Rooted ML tree (-Ln likelihood=3590.56574) generated based on nucleotide data set of 476 bp partial COI mtDNA gene. Values on the branches represent ML bootstrap estimates, based on 100 replicates. Only bootstrap values >50% are shown.

Figure 4.5 A Bayesian inference with 50% majority rule consensus tree of 476 bp partial mtDNA COI gene sequences. Values of Bayesian posterior probabilities (bpp) are shown in the branch nodes.

Figure 4.6 Chronogram showing the estimated time divergence of species in subfamily Murinae inferred from partial COI gene. The divergence time were calculated following the formula from Rutschmann (2006).

Figure 4.7 Plot of transition and transversion against divergence using Kimura (1980) distance method onto the third codon position shows saturation occurrence in cyt b gene region.
Figure 4.8  Phylogentic tree of 10 species of Murinae based on 850 bp cyt b mtDNA gene sequences. The phylogeny is a single tree recovered using NJ analysis with *C. notatus* as an outgroup. Values on the branches represent NJ bootstrap estimates, based on 1000 replicates. Only bootstrap values >50% are shown.

Figure 4.9  Rooted (with outgroup) MP tree based on nucleotide data set of 850 bp cyt b mtDNA gene (tree length=959; CI=0.4828; HI=0.5172; RI=0.8517). Values on the branches represent MP bootstrap estimates, based on 1000 replicates. Only bootstrap values >50% are shown.

Figure 4.10  Rooted ML tree (-Ln likelihood=5404.12784) generated based on nucleotide data set of 850 bp cyt b mtDNA gene. Values on the branches represent ML bootstrap estimates, based on 100 replicates. Only bootstrap values >50% are shown.

Figure 4.11  A Bayesian inference with 50% majority rule consensus tree of 850 bp mtDNA cyt b gene sequences. Values of bpp are shown in the branch nodes.
Figure 4.12  Chronogram showing the estimated time divergence of species in subfamily Murinae inferred from partial COI gene. The divergence time were calculated following the formula from Rutschmann (2006).

Figure 4.13  Plot of transition and transversion against divergence using Kimura (1980) distance method shows there is no saturation occurs in RAG2 gene region.

Figure 4.14  Phylogentic tree of 10 species of subfamily Murinae based on 775 bp RAG2 nuclear gene sequences. The phylogeny is a single tree recovered using NJ analysis with C. notatus as an outgroup. Values on the branches represent NJ bootstrap estimates, based on 1000 replicates. Only bootstrap values >50% are shown.

Figure 4.15  Rooted (with outgroup) MP tree based on nucleotide data set of 775 bp RAG2 nuclear gene (tree length=220; CI=0.8; HI= 0.2; RI=0.8585). Values on the branches represent MP bootstrap estimates, based on 1000 replicates. Only bootstrap values >50% are shown.
Figure 4.16 Rooted ML tree (-Ln likelihood=2291.61359) generated based on nucleotide data set of 775 bp RAG2 nuclear gene. Values on the branches represent ML bootstrap estimates, based on 100 replicates. Only bootstrap values >50% are shown.

Figure 4.17 A Bayesian inference with 50% majority rule consensus tree of 775 bp RAG2 nuclear gene sequences. Values of Bayesian posterior probabilities (bpp) are shown in the branch nodes.

Figure 5.1 Lateral views of the skulls show variation between *M. whiteheadi* (A), intermediate form of *Maxomys* sp. (B) and *M. ochraceiventer* by observing the flatness of the braincase, the greatest skull length and the zygomatic plate (a). Skull C is flatter and have longer greatest skull length (GSL) compared to skulls B and C.

Figure 5.2 Ventral views of the skulls show variation between *M. whiteheadi* (A), intermediate form of *Maxomys* sp. (B) and *M. ochraceiventer* by observing the posterior edge of bony palate (b) in relation to M3 and the palatal foramina (c) in relation to M1.
Figure 5.3  
Dorsal views of skull show variation between *M. whiteheadi* (A), intermediate form of *Maxomys* sp. (B) and *M. ochraceiventer* by observing the broadness of the braincase and the features of the zygomatic plate (a). Skull A is the broadest followed by skull C and B. The zygomatic plate of skull A is relatively small compared to skull B and C. Skull C has longer and narrower zygomatic plate compared to skull B.
List of Tables

Table 3.1  Solutions used for preservation and molecular approaches.  48

Table 3.2  Sequence of each primer that is used in this study.  49

Table 3.3  PCR reaction mixture.  50

Table 3.4  Parameters and profile for amplification of DNA.  50

Table 4.1  Taxonomic list of species captured at 10 sampling sites with number of individuals and relative abundance in parentheses.  59

Table 4.2  Comparison (%) of nucleotide composition inferred from partial COI mtDNA gene sequence among murine specimens analysed in the phylogenetic tree with and without the outgroup.  62
Table 4.3  Average percentage of Kimura two-parameter distance values within (boldface type along diagonal) and among species in subfamily Murinae from different clades based on COI gene sequences. n = sample size of each species. NA = not available.

Table 4.4  Comparison (%) of nucleotide composition inferred from 850 bp cyt b mtDNA gene sequence among subfamily Murinae specimens analysed in the phylogenetic tree with and without the outgroup.

Table 4.5  Average percentage of Kimura two-parameter distance values within (boldface type along diagonal) and among species in subfamily Murinae from different clades based on cyt b gene sequences. n = sample size of each species. NA = not available.

Table 4.6  Comparison (%) of nucleotide composition inferred from 775 bp RAG2 nuclear gene sequence among Murinae specimens analysed in the phylogenetic tree with and without the outgroup.
Table 4.7  Average percentage of Kimura two-parameter distance values within (boldface type along diagonal) and among species in subfamily Murinae from different clades based on RAG2 gene sequences. n = sample size of each species. NA = not available.

Table 4.8  Skull and dental measurements among complex lineage in Maxomys.
# List of Abbreviations

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>°</td>
<td>degree</td>
</tr>
<tr>
<td>°C</td>
<td>degree Celcius (temperature)</td>
</tr>
<tr>
<td>°N</td>
<td>degree North</td>
</tr>
<tr>
<td>°E</td>
<td>degree East</td>
</tr>
<tr>
<td>µl</td>
<td>microliter</td>
</tr>
<tr>
<td>%</td>
<td>percentage</td>
</tr>
<tr>
<td>&gt;</td>
<td>more than</td>
</tr>
<tr>
<td>&lt;</td>
<td>less than</td>
</tr>
<tr>
<td>X</td>
<td>multiply or times</td>
</tr>
<tr>
<td>-</td>
<td>negative</td>
</tr>
<tr>
<td>/</td>
<td>per</td>
</tr>
<tr>
<td>=</td>
<td>equal</td>
</tr>
<tr>
<td>≈</td>
<td>approximately</td>
</tr>
</tbody>
</table>

**A**
- A: Adenine
- AIC: Akaike Information Criterion
- AP5: acid phosphatase V
- a.s.l: above sea level
- Asap: Sungai Asap, Belaga

**B**
- Benom: Mount Benom
- Bera: Tasik Bera or Bera Lake
- Bp: base pair
- bpp: Bayesian posterior probabilities
- Btg Ai: Batang Ai National Park

**C**
- C: Cytosine
- CBL: condylobasal length
- CI: consistency index
- cm: centimeter
- COI: cytochrome oxidase I
- CTAB: cetyltrimethylammonium bromide
- cyt b: cytochrome b

**D**
- ddH$_2$O: deionised distilled water