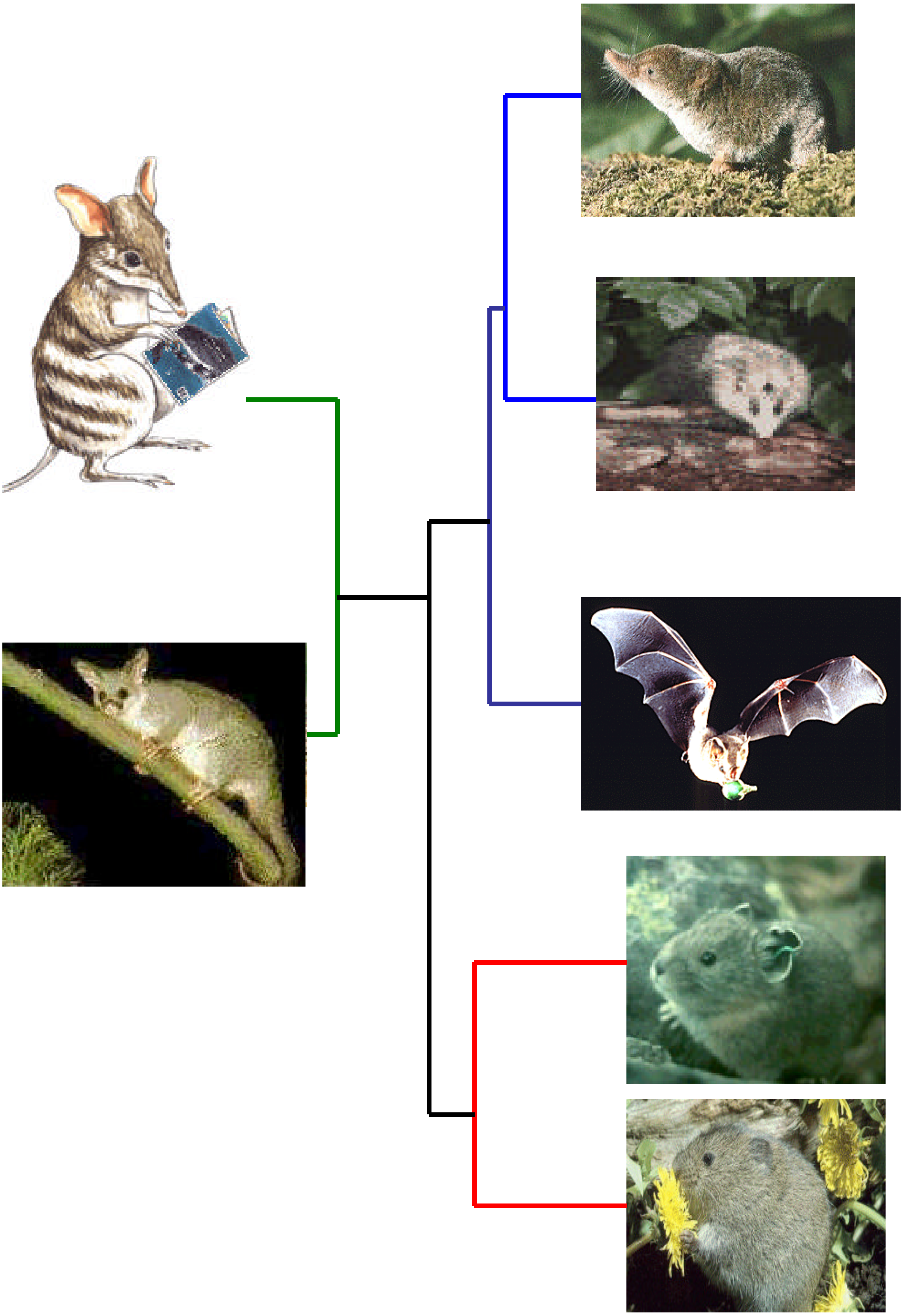
Mammalian Evolution and Phylogeny from Complete Mitochondrial Genomes

A thesis in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Molecular BioSciences
at Massey University

Yu-Hsin Lin

2001

Complete mitochondrial genomes and mammalian phylogeny



ACKNOWLEDGEMENTS

I would like to express my gratitude to my supervisor, David Penny, for his patient guidance through my Ph.D. project. He gave me far more than knowledge of research methods in biology. His wisdom, tolerance, patience and open-mindedness, all the requirements of a good scientist, were an inspiration to me.

I also thank my other supervisors, Michael Hendy and Peter Lockhart, who have taught me the mysteries of tree building in the simplest language.

Abby Harrison and Trish McLenachan are the people I need to give my acknowledge most for their help and suggestion in my lab work. I am also indebted to all the people, who helped me to arrange and provide the samples, who undertook the sequencing and who wrote the computer programs. Matt Phillip always provided good suggestions and discussions, thank you. I also thank the New Zealand Marsden Fund for their financial support.

I am also very appreciative to Trish McLenachan , Abby Harrison, and my supervisor, David Penny for their help with my manuscript.

Finally, my adorable wife and two daughters, they made me enjoy my life while I was studying.

ABSTRACT

The evolutionary tree of mammals is being resolved quickly. Complete mitochondrial DNA sequences are valuable data for deep mammalian phylogenetic relationships. From this study, the use of long-range PCR followed by short-range PCR and sequencing was proven to be a successful strategy for sequencing complete mt-genomes. This method is more efficient and cheaper than current cloning approaches. This method is also able to avoid most of the nuclear mitochondrial copies. Long branch attraction is a problem confusing the deep mammalian phylogeny. By sequencing complete mt-genomes of key taxa (2 marsupials, 2 bats, a pika, a gymnure, a rodent and a shrew) to break up long branches, this study resolve some ambiguous relationships in mammalian phylogeny.

The 8 mammalian mitochondrial DNA sequences from this study give additional support for the 4 groupings (Xenarthrans, Afrotheria, Supraprimates and Laurasiatheria) of placental mammals from current molecular studies. Some of the ambiguous relationships of higher mammalian relationships also get improved resolution. Bats are a monophyletic group but megabats may be paraphyletic. Eulipotyphla is a monophyletic group and deepest in the Laurasiatheria. Rodents are monophyletic and apart from a problem with the tree shrew, are sister to lagomorphs (Glires).

With the new gymnure complete mt-DNA available, the aberrant hedgehog mt-genome is returning to its traditional position in the placental tree and joins other Eulipotyphla (mole, shrew). This monophyletic Eulipotyphla is observed for the first time in the mammalian mitochondrial tree. The Erinaceidae (hedgehog and gymnure) and murid rodent seem to be under different processes of evolution and are attracted to the outgroups. By comparing trees without outgroups (unrooted trees) and with outgroups (rooted trees) and by constraining group(s) with unstable positions, the influence of marsupials/platypus outgroups on Erinaceidae and murid rodent can be investigated. The results from this study suggest that there is a long branch attraction problem between marsupials/platypus outgroups and murid rodent and Erinaceidae; the basal postions of Erinaceidae and murid rodent found in previous studies may be long branch attraction artifacts.

The resolved mammalian tree will be the basis for further molecular studies for estimating the time of divergence of extant mammalian orders, for the prediction of protein secondary structure, for the processes of transition of nucleotides and amino acids sequences in the tree, etc. Having a resolved mammalian tree is not the end for this research, rather a pivotal step for understanding evolution in molecular level.

TABLE OF CONTENTS

Acknowledgements..ii
Abstract...iii
Table of contents...iv

Chapter 1 Introduction

1.1 Fossil records and mammal evolution...1
1.2 Morphology versus molecules in mammal phylogeny....................................3
1.3 Theria and marsupionta...4
1.4 Systematics of placental mammals...7
  1.4.1 Xenarthra..8
  1.4.2 Afrotheria..9
  1.4.3 Supraprimates..10
  1.4.4 Laurasiatheria..11
1.5 Rooting the placental tree..15
1.6 Explosive radiation after K-T boundary VS before K-T boundary radiation..........16
1.7 Congruence of morphological and molecular characters..................................20
1.8 Mitochondrial genomes and mammalian evolution..21
  1.8.1 Origin and structure of mitochondrial DNA..21
  1.8.2 Advantages of mitochondrial genome in molecular phylogenetic studies..........24
  1.8.3 Mitochondrial genomes for deep-level mammalian phylogenetic reconstruction.24
  1.8.4 Pitfalls of using mitochondrial DNA in phylogeny...................................25
1.9 Phylogenetic inference..29
  1.9.1 The neutral theory of evolution...29
  1.9.2 Tree reconstruction..29
  1.9.3 Evaluation of tree reconstruction...33
  1.9.4 Assessing the reliability of individual branches..34
  1.9.5 Problems of inconsistency..35
  1.9.6 Rooting evolutionary tree...37
  1.9.7 The molecular clock...39
Reference list...41

Chapter 2 Material and methods

2.1 Introduction 54
2.2 Development of long-range polymerase chain reaction...…………………..………..54
2.3 Primer design...…..........................55
2.4 Sequencing complete mitochondrial genomes...................................…..................57
  2.4.1 DNA extraction...………..........57
  2.4.2 Polymerase Chain Reaction, PCR.......................................…….............…..59
  2.4.3 PCR product purification...……....…........59
  2.4.4 PCR product quantification...……..........…..60
  2.4.5 Cloning...……......…...62
  2.4.6 Sequencing..………........65
2.5 Data alignment..…......…..............…......67
2.6 Programs used for phylogenetic inference in the present study..........…............…...68
Reference list...…70

Chapter 3 Results

Background and Overview of my contribution..….72
1. Yu-Hsin Lin and David Penny (2001) Implications for bat evolution from two new complete mitochondrial genomes. Molecular Biology and Evolution 18(4): 684-688.
2. Matthew J Phillips, Yu-Hsin Lin, Gabrielle L. Harrison and David Penny (2001) Mitochondrial genomes of a bandicoot and a brushtail possum confirm the monophyly of australidelphian marsupials. Proceedings of the Royal Society of London Series B-Biological Sciences. 268: 1533-1538.
3. Yu-Hsin Lin, Peter J Waddell and David Penny (2001) Pika and vole mitochondrial genomes add support to both rodent monophyly and Glires. (submitted to Gene)….….88
4. Yu-Hsin Lin, Patrica A McLenachan, Alica R Gore, Matthew J Phillips and David Penny (2001) Four new mitochondrial genomes, and the stability of evolutionary trees of mammals. (prepare to submit to Molecular Biology and Evolution)………..……..113

Chapter 4 Discussion

4.1 Is our strategy: long PCR→short PCR→sequencing successful?......................…....138

4.2 Data alignment and database manipulation...139
4.3 General conclusions from this study..140
4.4 How reliable is the tree from mt-genomes?..142
4.5 Are the trees inferred consistent with palaeontological and biological evidence?.......144
4.6 Future perspectives...……146
4.6.1 Future progress in mammalian mitochondrial tree...146
4.6.2 Secondary structure prediction and phylogenetic inference...................................149
4.6.3 Application of the mammalian mitochondrial tree: molecular evolution, timing..150
4.6.4 Reference list..151