

University of Wollongong - Research Online

Thesis Collection

Title: The evolution of mitochondrial genomes and phylogenetic relationships in the hymenoptera

Author: Lyda Raquel Castro

Year: 2006

Repository DOI:

Copyright Warning

You may print or download ONE copy of this document for the purpose of your own research or study. The University does not authorise you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site.

You are reminded of the following: This work is copyright. Apart from any use permitted under the Copyright Act 1968, no part of this work may be reproduced by any process, nor may any other exclusive right be exercised, without the permission of the author. Copyright owners are entitled to take legal action against persons who infringe their copyright. A reproduction of material that is protected by copyright may be a copyright infringement. A court may impose penalties and award damages in relation to offences and infringements relating to copyright material.

Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.

Unless otherwise indicated, the views expressed in this thesis are those of the author and do not necessarily represent the views of the University of Wollongong.

Research Online is the open access repository for the University of Wollongong. For further information contact the UOW Library: research-pubs@uow.edu.au

2006

The evolution of mitochondrial genomes and phylogenetic relationships in the hymenoptera

Lyda Raquel Castro
University of Wollongong

Follow this and additional works at: <https://ro.uow.edu.au/theses>

University of Wollongong

Copyright Warning

You may print or download ONE copy of this document for the purpose of your own research or study. The University does not authorise you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site.

You are reminded of the following: This work is copyright. Apart from any use permitted under the Copyright Act 1968, no part of this work may be reproduced by any process, nor may any other exclusive right be exercised, without the permission of the author. Copyright owners are entitled to take legal action against persons who infringe their copyright. A reproduction of material that is protected by copyright may be a copyright infringement. A court may impose penalties and award damages in relation to offences and infringements relating to copyright material.

Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.

Unless otherwise indicated, the views expressed in this thesis are those of the author and do not necessarily represent the views of the University of Wollongong.

Recommended Citation

Castro, Lyda Raquel, The evolution of mitochondrial genomes and phylogenetic relationships in the hymenoptera, PhD thesis, School of Biological Sciences, University of Wollongong, 2006.
<http://ro.uow.edu.au/theses/660>

NOTE

This online version of the thesis may have different page formatting and pagination from the paper copy held in the University of Wollongong Library.

UNIVERSITY OF WOLLONGONG

COPYRIGHT WARNING

You may print or download ONE copy of this document for the purpose of your own research or study. The University does not authorise you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site. You are reminded of the following:

Copyright owners are entitled to take legal action against persons who infringe their copyright. A reproduction of material that is protected by copyright may be a copyright infringement. A court may impose penalties and award damages in relation to offences and infringements relating to copyright material. Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.

**The evolution of mitochondrial genomes and phylogenetic
relationships in the Hymenoptera**

A thesis submitted in fulfillment
of the requirement for the award of the degree of

Doctorate of Philosophy

from the

University of Wollongong

by

Lyda Raquel Castro BSc, MSc

School of Biological Sciences

2006

Thesis Declaration

I, Lyda Raquel Castro, declare that this thesis, submitted in partial fulfillment of the requirements for the award of Doctor of Philosophy, in the School of Biological Sciences, University of Wollongong, is wholly my own work unless otherwise referenced and acknowledged. This document has not been submitted for qualification at any other academic institution.

Lyda Raquel Castro

January 2006

Acknowledgments

I am indebted to Dr. Mark Dowton from the University of Wollongong for supervising this project. I specially want to thank him for his continual support, guidance, encouragement, and knowledge over the past years. He has been always there for me and since the beginning has made my stay in Australia the best experience I could ever have had. I feel privileged to have worked in his lab. He has been the best example, not only as a professional, but as a person. I also want to thank all the members of his lab. for their support and friendship.

I would like to thank my family, especially my mum and dad. Their continual support, guidance and confidence made me achieve this goal. I also would like to thank Carlos, for his love, support and encouragement.

Thanks to Andy Austin, Dave Smith, and John Jennings for provision of the *Perga condei*, *Vanhornia eucnemidarum* and *Primeuchroeus* sp. specimens. For the kind donation of the other specimens, I want to thank Eveline Bartowsky, Robert Belshaw, Ferdinando Bin, Paul Dangerfield, John Early, F. Felipe, Scott Field, G. Fitt, Ian Gauld, Gary Gibson, E. Grissell, Paul Hanson, B. Hatami, John Heraty, Marianne Hellers, P. Horne, G. Jackson, John Jennings, Norm Johnson, Mike Keller, D. Kent, J. King, J. Kitt, M. Kulbars, Nina Laurence, Paul Lewis, Patrick Mardulyn, Lubomir Masner, Gwen Mayo, D. Murray, Ian Naumann, J. O'Hara, Donald Quicke, Nathan Schiff, Scott Shaw, Aguiar Sharkov, D. Smith, R. Storey, Gary Taylor, N. Tonkin, G. Tribe, G. Walter, Q. Wang, A. Wells, Bob Wharton, Jim Whitfield, N. Zareh.

I would like to thank Gary Gibson who generously shared his knowledge on apocritan morphology with us. Also Andy Austin for his suggestions and comments on preliminary versions of this thesis. Thanks to Rajkumar Kunaparaju for preliminary laboratory work on the sequencing of *Perga condei*, and Kalani Ruberu also for her help in cloning and sequencing *Vanhornia eucnemidarum*.

Finally, I want to thank the University of Wollongong, the Australian Museum and the Australian Research Council. Grants from these institutions supported this work.

Abstract

This thesis studies the phylogenetic relationships in the Hymenoptera as well as the mitochondrial (mt) genomics of the group with a comparative approach. My principal purposes were to (1) reconstruct the phylogenetic relationships among the Apocrita, and (2) characterize the mt genome of the Hymenoptera and its utility as a phylogenetic marker both within the Apocrita and within the Holometabola. In order to achieve these aims: (1) 550 bp of the *18S* gene were sequenced in 87 apocritan taxa and analyzed using a Bayesian phylogenetic approach, including the sequences of two mitochondrial genes (*cox1*, *16S*) and another nuclear gene (*28S*) from Dowton and Austin (2001). Although the phylogeny of the Apocrita was not totally resolved, I was able to support some groups. In particular, the monophyly of the Proctotrupomorpha, and within this group the Chalcidoidea as sister taxon to the Diapriidae + Monomachidae + Maamingidae clade were consistently recovered and supported by high posterior probabilities. (2) Most of the mt genome of the sawfly *Perga condei* was sequenced. 12 protein coding genes, 16 *trn* genes, and the small and large rRNA genes, for a total of 13,416 bp. This mt genome has a conserved gene order, with the exception that *tnaL*^{CUN} was not found in the position considered ancestral to insects and crustaceans (Boore et al. 1998, Flook et al. 1995). Apart from this rearrangement, the organization of the genes in *Perga condei* matches perfectly with distant species such as *Drosophila melanogaster* (Lewis et al. 1994) or *Triatoma dimidiata* (Dotson and Beard 2001). The base composition, the amino acid composition, and the codon usage of the mt genome of *P. condei* were reported. Similar to other insect mt genomes, this genome is A+T rich, and there is a correlation between the base composition and amino acid occurrence, with A+T rich codons predominating. (3) Two other mt

genomes of the Hymenoptera were also sequenced. The mt genome of *Vanhornia eucnemidarum* and of *Primeuchroeus* sp., both from the suborder Apocrita. Within the Apocrita, high rates of molecular evolution, compositional bias and gene rearrangements had been reported (Dowton and Austin 1997, Dowton et al. 2003). The mt genomes of *Vanhornia* and *Primeuchroeus* are further evidence of an increased rate of gene rearrangement within the Apocrita. In particular, there is a total of six *trn* genes rearranged in *Vanhornia eucnemidarum*. Additionally, several non-coding regions were found in the mt genome of *Vanhornia eucnemidarum*. One of these non-coding regions is around 600 bp long and has a high AT content, but does not seem to correspond to the typical A+T rich region present in other insect mt genomes. There are at least nine *trn* genes rearranged in the mt genome of *Primeuchroeus* sp. Further, the large and small rRNA genes are inverted. In both species, rearrangements of *trn* genes are the most common. The gene rearrangements found in the mt genomes of the hymenopteran taxa sequenced were characterized; however no synapomorphies were detected. Since the rate of gene rearrangement appears to be increased in this group of insects, only with increased taxon sampling will phylogenetically informative rearrangements be found. (4) Finally, the mt genome sequences previously described were tested as phylogenetic markers to reconstruct relationships both within the Holometabola and within the Hymenoptera. Results indicated that phylogenetic analyses using mt genomes were susceptible to outgroup and ingroup selection as well as analytical model. Analyses excluding 3rd codon positions were found to be the best model to analyze this type of data, but an increased taxonomic sampling within the Apocrita as well as within the outgroups is required to recover appropriate phylogenetic relationships.

Table of Contents

Thesis Declaration.....	ii
Acknowledgments.....	iii
Abstract	v
List of Abbreviations	xi
List of Tables	xiv
List of Figures	xvi
CHAPTER 1 INTRODUCTION	1
1.1 MITOCHONDRIAL GENOME: EVOLUTION.....	2
1.1.1 Substitutional Rate Heterogeneity	2
1.1.2 Compositional Bias	5
1.1.3 Gene rearrangements.....	9
<i>1.1.3.1 Gene rearrangement mechanisms</i>	10
<i>1.1.3.2 Gene rearrangements as phylogenetic markers</i>	12
1.2 CONCATENATED MITOCHONDRIAL GENE SEQUENCES AND THEIR ...	14
USE FOR PHYLOGENETIC RECONSTRUCTION	14
<i>1.2.1 Cases in which concatenated mitochondrial gene sequences have been relatively</i>	
<i>successful in recovering phylogenetic relationships</i>	15
<i>1.2.2 Cases in which concatenated mitochondrial gene sequences have been</i>	
<i>problematic in recovering phylogenetic relationships</i>	16
1.3 EVOLUTION OF THE HYMENOPTERA.....	19
<i>1.3.1 Phylogenetic position of the Hymenoptera among other Holometabolan insects</i>	
.....	19
<i>1.3.2 Phylogenetic relationships within the Hymenoptera</i>	22
<i>1.3.2.1 Suborder Symphyta</i>	23
<i>1.3.2.2 Suborder Apocrita</i>	23
1.4 PARASITISM IN THE HYMENOPTERA.....	24
<i>1.4.1 Parasitoid Natural History</i>	25
<i>1.4.2 The Evolution of Parasitism</i>	28
1.5 AIMS.....	29

1.6 THESIS FORMAT	31
CHAPTER 2 The position of the Hymenoptera within the Holometabola as inferred from the mitochondrial genome of <i>Perga condei</i> (Hymenoptera:Symphyta:Pergidae)	34
2.1 Introduction.....	34
2.2 Materials and Methods.....	38
2.2.1 <i>Insects and DNA Extraction</i>	38
2.2.2 <i>PCR amplification and cloning</i>	39
2.2.3 <i>Identification of tRNAs</i>	40
2.2.4 <i>Alignments and phylogenetic analysis</i>	41
2.3 Results.....	43
2.3.1 <i>Genome structure and organization</i>	43
2.3.2 <i>Base composition, amino acid composition and codon usage</i>	45
2.3.3 <i>Phylogenetic analysis</i>	46
2.4 Discussion	48
2.4.1 <i>Perga condei mitochondrial genome structure</i>	48
2.4.2 <i>Phylogenetic analysis</i>	50
2.5 Conclusion	52
CHAPTER 3 Molecular Bayesian Analysis of the Apocrita (Insecta: Hymenoptera) suggests that the Chalcidoidea are sister to the (Monomachidae+Diapriidae+Maamingidae)	61
3.1 Introduction.....	61
3.2 Materials and Methods.....	63
3.2.1 <i>DNA Extraction and sequencing</i>	64
3.2.2 <i>Sequence alignment and phylogenetic analysis</i>	65
3.3 Results.....	67
3.3.1 <i>Molecular analyses</i>	68
3.3.1.1 <i>Unrooted analysis</i>	69
3.3.2 <i>Analysis including only the Proctotrupomorpha</i>	70
3.3.3 <i>Comparison between Bayesian runs</i>	71
3.4 Discussion	71
3.4.1 <i>Chalcidoidea+(Diapriidae+Monomachidae+Maamingidae)</i>	73
3.4.2 <i>Other hymenopteran relationships</i>	74
3.4.2.1 <i>Evaniiidae not sister to the Gasteruptiidae</i>	74
3.4.2.2 <i>Aculeata within the Evaniomorpha</i>	75

3.5 Conclusion	75
CHAPTER 4 Mitochondrial genomes of <i>Vanhornia eucnemidarum</i> and <i>Primeuchroeus</i> sp.: Evidence of rearranged mitochondrial genomes within the Apocrita (Hymenoptera)	
4.1 Introduction	84
4.2 Materials and Methods.....	86
4.2.1 <i>Insects and DNA Extraction</i>	87
4.2.2 <i>PCR amplification and cloning</i>	87
4.2.3 <i>Identification of protein-coding, rrn, trn genes and gene rearrangements</i>	89
4.3 Results and Discussion.....	90
4.3.1 <i>Genome size and composition</i>	90
4.3.1.1 <i>Vanhornia eucnemidarum</i>	90
4.3.1.2 <i>Primeuchroeus</i> sp.	92
4.3.2 <i>Protein-coding genes and codon usage</i>	94
4.3.2.1 <i>Vanhornia eucnemidarum</i>	94
4.3.2.2 <i>Primeuchroeus</i> sp.	95
4.3.3 <i>Transfer RNA genes and ribosomal RNA genes</i>	95
4.3.3.1 <i>Vanhornia eucnemidarum</i>	95
4.3.3.2 <i>Primeuchroeus</i> sp.	96
4.3.4 <i>Genome organization</i>	97
4.3.4.1 <i>Vanhornia eucnemidarum</i>	97
4.3.4.2 <i>Primeuchroeus</i> sp.	98
4.3.5 <i>The ARNSEF region of Vanhornia eucnemidarum</i>	98
4.3.6 <i>Comparison of gene rearrangements</i>	99
4.4 Conclusion	100
CHAPTER 5 Mitochondrial genomes in the Hymenoptera and their utility as phylogenetic markers.....	
5.1 Introduction	114
5.2 Materials and Methods.....	118
5.2.1 <i>Taxon Sampling</i>	118
5.2.2 <i>Alignment</i>	119
5.2.3 <i>Phylogenetic analysis</i>	119
5.3 Results and Discussion.....	120
5.3.1 <i>Model Specification and stationarity</i>	121
5.3.2 <i>Hymenopteran Relationships</i>	121
5.3.2.1 <i>All taxa included</i>	121
5.3.2.2 <i>Only Hymenoptera included</i>	122
5.3.3 <i>Holometabolan Relationships</i>	124

5.3.2.1 <i>Locusta</i> as outgroup	124
5.3.2.2 <i>Triatoma</i> as outgroup	125
5.3.2.3 <i>No outgroup</i>	125
5.4 Conclusion	126
Chapter 6 General Discussion	138
6.1 Significance of the study	138
6.2 Major results and implications	139
6.2.1 <i>An improved resolution of the Apocritan phylogeny</i>	139
6.2.2 <i>A more accurate estimate of the degree of gene rearrangements in</i> <i>hymenopteran mt genomes</i>	142
6.2.3 <i>The potential of mt genomes as phylogenetic markers</i>	144
6.3 Recommendations and future research	145
6.4 Conclusions	148
REFERENCES	151

List of Abbreviations

μg	micro-gram
μL	micro-liter
μM	micro-molar
Ala [A]	Alanine (amino acid)
Arg [R]	Arginine (amino acid)
Asn [N]	Asparagine (amino acid)
Asp [D]	Aspartic acid (amino acid)
<i>atp6</i>	gene for ATP synthase subunit 6
<i>atp8</i>	gene for ATP synthase subunit 8
<i>cob</i>	cytochrome oxydase B
<i>cox1</i>	subunit 1 of cytochrome oxydase
<i>cox2</i>	subunit 2 of cytochrome oxydase
<i>cox3</i>	subunit 3 of cytochrome oxydase
Cys [C]	Cysteine (amino acid)
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide-triphosphate
EDTA	Ethylendiaminetetraacetic acid
Gln [Q]	Glutamine (amino acid)
Glu [E]	Glutamic acid (amino acid)
Gly [G]	Glycine (amino acid)
GTR	General time-reversible model
GTR + I + Γ	General time-reversible model with rates invgamma
His [H]	Histidine (amino acid)
Ile [I]	Isoleucine (amino acid)
Kb	Kilo-base
Leu [L]	Leucine (amino acid)
Lys [K]	Lysine (amino acid)
Met [M]	Metionine (amino acid)
ML	Maximum likelihood
mM	mili-molar

MP	Maximum parsimony
Mt	Mitochondrial
<i>nd1</i>	gene for NADH dehydrogenase subunit 1
<i>nd2</i>	gene for NADH dehydrogenase subunit 2
<i>nd3</i>	gene for NADH dehydrogenase subunit 3
<i>nd4</i>	gene for NADH dehydrogenase subunit 4
<i>nd4L</i>	gene for NADH dehydrogenase subunit 4L
<i>nd5</i>	gene for NADH dehydrogenase subunit 5
<i>nd6</i>	gene for NADH dehydrogenase subunit 6
ORF	Open reading frame
PCR	Polimerase chain reaction
PEG	Polyethylene glycol
Pfu	Pyrococcus furiosus
Phe [F]	Phenylalanine (amino acid)
PP	Posterior probability
Pro [P]	Proline (amino acid)
<i>rnl</i>	gene for ribosomal RNA large-subunit
<i>rns</i>	gene for ribosomal RNA small-subunit
RGR	rate of gene rearrangement
SDS	Sodium dodecyl sulfate
Ser [S]	Serine (amino acid)
Thr [T]	Threonine (amino acid)
<i>trn</i>	transfer RNA
<i>trnA</i>	transfer RNA gene for alanine
<i>trnC</i>	transfer RNA gene for cytosine
<i>trnD</i>	transfer RNA gene for aspartic acid
<i>trnE</i>	transfer RNA gene for glutamic acid
<i>trnF</i>	transfer RNA gene for phenylalanine
<i>trnH</i>	transfer RNA gene for histidine
<i>trnK</i>	transfer RNA gene for lysine
<i>trnL</i>	transfer RNA gene for leucine

<i>trnM</i>	transfer RNA gene for methionine
<i>trnN</i>	transfer RNA gene for asparagine
<i>trnP</i>	transfer RNA gene for proline
<i>trnR</i>	transfer RNA gene for arginine
<i>trnS</i>	transfer RNA gene for serine
<i>trnT</i>	transfer RNA gene for threonine
<i>trnV</i>	transfer RNA gene for valine
<i>trnW</i>	transfer RNA gene for tryptophan
<i>trnY</i>	transfer RNA gene for tyrosine
Trp [W]	Tryptophan (amino acid)
Tyr [Y]	Tyrosine (amino acid)
U	units
Val [V]	Valine (amino acid)

List of Tables

Table 2.1 Primer pairs used in PCR amplification of the <i>Perga condei</i> mitochondrial genome.....	54
Table 2.2 Information of the taxa from which the whole mitochondrial genome have been sequenced, and were used in the phylogenetic analyses.....	55
Table 2.3 AT content in different regions of the <i>Perga condei</i> mt genome.....	56
Table 2.4 Codon usage in the <i>Perga condei</i> mt genome.....	57
Table 3.1 List of taxonomic groups included in the phylogenetic analysis.....	77
Table 3.2 Estimated model likelihood obtained in the different analyses.....	79
Table 4.1 Primer pairs used in PCR amplification of the <i>V. eucnemidarum</i> and <i>Primeuchroeus</i> sp. mt genomes.....	101
Table 4.2 Mt gene profile of <i>V. eucnemidarum</i> (Hymenoptera: Vanhnornidae).....	102
Table 4.3 AT content in different regions of the <i>V. eucnemidarum</i> mt genome.....	103
Table 4.4 Possible gene junctions in <i>Primeuchroeus</i> mt genome. Primers were carefully designed at the end of the genes and used in PCR reactions assuming all possible orientations.....	104

Table 4.5 AT content in different regions of the <i>Primeuchroeus</i> sp. mt genome...	105
Table 4.6 Comparison of the mt protein-encoding genes of <i>V. eucnemidarum</i> with those of other hymenopteran mt genomes available in GenBank.....	106
Table 4.7 Codon usage in the <i>V. eucnemidarum</i> mt genome.....	107
Table 4.8 Codon usage in the <i>Primeuchroeus</i> sp. mt genome.....	108
Table 5.1 Information on the taxa whose whole mitochondrial genome sequence was used in the phylogenetic analyses.....	128
Table 5.2 Estimated model likelihood obtained in the different analyses.....	129
Table 5.3 Within hymenopteran relationships as recovered by the different phylogenetic analyses.....	130
Table 5.4 Position of the Hymenoptera within the Holometabola as recovered by the different phylogenetic analyses.....	131
Table 6.1 Assessment of the relative rate of genome rearrangement among hymenopteran mitochondrial genomes, using the RGR test.....	150

List of Figures

Figure 1.1 Previous hypotheses for holometabolan relationships.....	32
Figure 1.2 Summary of what is known about symphytan relationships.....	33
Figure 1.3 Major apocritan relationships proposed by Rasnitsyn (1988).....	33
Figure 2.1 Holometabolan relationships obtained from the amino acid dataset.....	58
Figure 2.2 Holometabolan relationships obtained from the nucleotide datasets.....	59
Figure 3.1 Bayesian analysis including <i>28S</i> , <i>16S</i> , <i>COI-1</i> , <i>COI-2</i> (DA2001).....	80
Figure 3.2 Bayesian analysis including <i>28S</i> , <i>16S</i> , <i>COI-1</i> , and <i>COI-2</i> partitions from DA2001, and <i>18S</i> partition.	81
Figure 3.3 Bayesian analysis of the Proctotrupomorpha; including <i>28S</i> , <i>16S</i> , <i>COI-1</i> , and <i>COI-2</i> partitions from Downton and Austin (2001), and <i>18S</i> partition.....	82
Figure 3.4 Generation plots of some of the substitution rates for the <i>COI-1</i> and <i>COI-2</i> partitions after 5,000,000 generation runs. A. Generation plot of the C-G rate parameter for the <i>COI-1</i> partition (diffuse posterior distribution). B. Generation plot of the C-G rate parameter for the <i>28S</i> partition (posterior remains focused).	83
Figure 4.1 Linearized representation of the mt genome arrangement of <i>Vanhornia eucnemidarum</i> (Hymenoptera: Apocrita) and the putative ancestral arthropod organization	

represented by <i>Drosophila</i>	
<i>yacuba</i>	109

Figure 4.2 Linearized representation of the mt genome arrangement of <i>Primeuchroeus sp.</i> (Hymenoptera: Apocrita) and the putative ancestral arthropod organization represented by <i>Drosophila yacuba</i>	110
--	-----

Figure 4.3 Putative secondary structure for the 18 <i>trn</i> genes of the <i>V. eucnemidarum</i> mt genome.....	111
---	-----

Figure 4.4 Inferred secondary structure of the <i>trn</i> genes of the mt genome of <i>Primeuchroeus sp.</i>	113
---	-----

Figure 5.1 Accepted phylogenetic relationships between the five hymenopteran taxa included in this analysis.....	132
---	-----

Figure 5.2 Phylogenetic relationships of the Holometabola. Bayesian analyses.....	134
--	-----

Figure 5.3 Phylogenetic relationships of the Holometabola. Parsimony analyses..	136
--	-----

Figure 5.4 Only Hymenopteran representatives included. A. Bayesian analyses of nucleotide sequences excluding 3 rd codon positions. B. Tree topology recovered in all other analyses (parsimony and Bayesian using proteins, parsimony and Bayesian including all codon positions and parsimony excluding 3 rd codon positions).....	137
---	-----