Molecular Phylogenetics and Evolution 85 (2015) 230–237

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

A mitochondrial genome phylogeny of owlet moths (Lepidoptera: Noctuoidea), and examination of the utility of mitochondrial genomes for lepidopteran phylogenetics

Xiushuai Yang [a,1], Stephen L. Cameron [b,1], David C. Lees [c,1], Dayong Xue [a], Hongxiang Han [a,*]

[a] Key Laboratory of Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China
[b] Earth, Environmental & Biological Sciences School, Science & Engineering Faculty, Queensland University of Technology, GPO Box 2434, Brisbane, QLD 4001, Australia
[c] Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, United Kingdom

ARTICLE INFO

Article history:
Received 15 April 2014
Revised 27 January 2015
Accepted 6 February 2015
Available online 17 February 2015

Keywords:
Noctuoidea
Mitochondrial genome
Phylogeny
MAFFT
Gblocks
Gaps

ABSTRACT

A phylogenetic hypothesis for the lepidopteran superfamily Noctuoidea was inferred based on the complete mitochondrial (mt) genomes of 12 species (six newly sequenced). The monophyly of each noctuoid family in the latest classification was well supported. Novel and robust relationships were recovered at the family level, in contrast to previous analyses using nuclear genes. Erebidae was recovered as sister to (Nolidae + (Euteliidae + Noctuidae)), while Notodontidae was sister to all these taxa (the putatively basalmost lineage Oenosandridae was not included). In order to improve phylogenetic resolution using mt genomes, various analytical approaches were tested: Bayesian inference (BI) vs. maximum likelihood (ML), excluding vs. including RNA genes (rRNA or tRNA), and Gblocks treatment. The evolutionary signal within mt genomes had low sensitivity to analytical changes. Inference methods had the most significant influence. Inclusion of tRNAs positively increased the congruence of topologies, while inclusion of rRNAs resulted in a range of phylogenetic relationships varying depending on other analytical factors. The two Gblocks parameter settings had opposite effects on nodal support between the two inference methods. The relaxed parameter (GBRA) resulted in higher support values in BI analyses, while the strict parameter (GBDH) resulted in higher support values in ML analyses.

1. Introduction

The application of molecular markers in phylogenetics has made a tremendous contribution to our knowledge of the evolution of life on Earth (Delsuc et al., 2005; Wheat and Wahlberg, 2013). Inference of phylogenetic relationships within Lepidoptera (butterflies and moths), as well their macroevolutionary history (Wahlberg et al., 2013), have been greatly improved by a series of studies based on an ever-increasing numbers of molecular markers, largely nuclear protein-coding genes (PCGs) (Mutanen et al., 2010; Regier et al., 2013). For example, a comprehensive phylogenetic study of Bombycoidea (silkworms and hawkmoths) based on a combination of up to 25 nuclear genes (19.3 kb total), the largest to date for a lepidopteran superfamily, found stronger nodal support for many more subclades (families and subfamilies) than either a parallel 5-gene analyses or previous studies (Regier et al., 2008; Zwick et al., 2011). Surprisingly, the relationships of Bombycidae, Saturniidae, and Sphingidae are still contentious (e.g. Timmermans et al., 2014), although Breinholt and Kawahara (2013) found phylogenomic evidence for a sister relationship of the last two. The most recent transcriptomic analysis using 46 lepidopteran taxa recovered a novel placement of butterflies and established a preliminary framework to understand the evolution of butterflies and moths (Kawahara and Breinholt, 2014).

Noctuoidea is the biggest superfamily of Lepidoptera, numbering over 42,400 species (Nieukerken et al., 2011). Many species are pests of considerable importance in agriculture and forestry, such as armyworms (*Spodoptera* spp.), bollworms (*Helicoverpa* spp.) and cutworms (*Agrotis* spp.) (Zhang, 1994). The superfamily also includes the tiger moths (Arctiinae), one of the more attractive groups of day-flying moths. The taxonomic history of Noctuoidea is complicated, but at present six families are recognized: Oenosandridae, Notodontidae, Erebidae, Euteliidae, Nolidae and Noctuidae (Nieukerken et al., 2011; Zahiri et al., 2011). Several other well-known noctuoid groups, including the Lymantriinae and Arctiinae, previously accorded family rank are now treated as erebid subfamilies (Zahiri et al., 2012).

* Corresponding author.
E-mail address: hanhx@ioz.ac.cn (H. Han).
[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.ympev.2015.02.005
1055-7903/

Phylogenetic studies of Noctuoidea have followed the trend of increased reliance on molecular markers over morphological characters during the past few decades. The monophyly of Noctuoidea (*sensu lato*) is supported by the presence of a metathoracic tympanal organ, which is regarded as uniquely gained apomorphic character (Miller, 1991). Speidel et al. (1996) conducted an analysis of 30 subfamilies based on 10 morphological characters and confirmed the monophyly of Noctuidae, but found poor resolution among subordinate groups. Kitching et al. (1998) extended Speidel's (1996) dataset with larval characters for resolving the monophyly of the families Arctiidae, Nolidae, Lymantriidae and Pantheidae, as well as their interrelationships. Lafontaine and Fibiger found further support for the respective monophyly of the 'trifid' noctuoids (Oenosandridae and Notodontidae, with an apparently 3-branching forewing cubital vein) and 'quadrifids' (with a 4-branching forewing cubital vein) based on adult and larval characters (Fig. 1B and C) (Fibiger and Lafontaine, 2005; Lafontaine and Fibiger, 2006). In their system, all quadrifid moths, including the families Arctiidae, Lymantriidae and Aganaidae, were redefined as subfamilies of Noctuidae, and therefore Noctuidae (*sensu lato*) became the largest (~38,000 of 42,000 species) and most complex noctuoid family (48 subfamilies and 79 tribes).

Molecular studies of Noctuoidea were initially based on one or two genes, and limited taxon sampling (three families, 29–49 species) (Fang et al., 2000; Mitchell et al., 1997, 2000; Weller et al., 1994). Mitchell et al. (2000, 2006) initially conducted comprehensive systematic analysis of Noctuidae based on two nuclear genes and a wide range of taxon sampling. Besides finding strong support for the monophyly of its subfamilies, the "LAQ" clade was proposed for the first time. Lymantriidae and Arctiidae became subordinate subfamilies within the quadrifid noctuids (Fig. 1D). Mitchell's research, in concert with that of Lafontaine and Fibiger (Fibiger and Lafontaine, 2005), has greatly clarified several hitherto elusive relationships within Noctuoidea. However, the support values above the subfamily level were low, possibly resulting from a sampling biased towards 'trifine' species (with an apparently 3-branching hindwing cubital vein), with relatively few 'quadrifines' (with a 4-branching hindwing cubital vein), and/or a result of the limits of phylogenetic resolution from the two nuclear genes sequenced. Recently however, Zahiri et al. (2011) conducted the most extensive molecular phylogenetic analyses of Noctuoidea to date, based both on combined sequences of eight PCGs (nuclear and mitochondrial) and on comprehensive sampling of 152 species, which covered nearly all of the subfamilies recognized within Noctuoidea. Zahiri et al.'s (2011) results differed significantly from all previous studies, both morphological and molecular ones (Fig. 1E). The traditional group of quadrifid noctuids was broken up, and the families Nolidae and Erebidae (the last family newly defined to encompass a still vast clade of some 18 subfamilies and about 24,500 species) were reconstituted and Euteliidae was newly recognized (previously a noctuid subfamily) (Zahiri et al., 2012). In so doing, they proposed a novel six-family system that has currently been accepted within synoptic classifications of Lepidoptera (Nieukerken et al., 2011). While the datasets of Zahiri et al. (2011, 2012, 2013a,b) provide strong support for the monophyly of the six families and most of the subfamilies, nodal support for Erebidae + Nolidae and most inter-subfamily relationships was not significant. This is despite a sizeable molecular dataset (6407 bp) including most of the nuclear genes shown to be most effective at deep phylogenetic levels within Lepidoptera (Mutanen et al., 2010; Wahlberg et al., 2009a); further resolution of noctuoid relationships thus requires novel data sources. An additional outcome of Zahiri et al. (2011, 2012, 2013a,b) was that morphological and molecular studies were contradictory, with no morphological support for Erebidae. In practice, the Erebidae is impossible to recognize without comprehensive faunal knowledge (Barbut and Lalanne-Cassou, 2011; Holloway, 2011). Although some potential autapomorphies were proposed in a recent study, such as the loss of a muscle in the male genitalia, none was indisputable and they were absent in some subfamilies or some taxa (Minet et al., 2012). Despite rapid acceptance by a range of lepidopteran systematists, the higher-level phylogenetic relationships of Noctuoidea were once again potentially contentious.

Complete mitochondrial (mt) genomes have been increasingly used to address phylogenetic questions where multi-gene analyses have been either unresolved or poorly supported. They have been widely used in invertebrates in general and insects in particular (Cameron, 2014). Within Lepidoptera, multiple studies have used mt genomes to address relationships between superfamilies and within major families such as Nymphalidae (Tian et al., 2012; Yang et al., 2009). These studies have demonstrated significant and novel findings, including the first strong evidence for the non-monophyly of Macrolepidoptera, primarily due to an earlier diverging position of the butterflies (Papilionoidea) than traditionally thought (Yang et al., 2009), a position subsequently confirmed by analyses of nuclear PCGs (Mutanen et al., 2010; Regier et al., 2009). Most recently, a comprehensive phylogenetic study of Lepidoptera was conducted using mt genomes of 106 species from 24 subfamilies (Timmermans et al., 2014). The topologies are highly congruent with prior nuclear and/or morphological studies.

Various analytical approaches using insect mt genomes to resolve phylogeny have been tested (Cameron, 2014); however, optimal approaches appear to vary between insect lineages. For example, the inclusion of third codon positions positively affected phylogenetic resolution in Orthoptera, Lepidoptera and Diptera (Cameron et al., 2007; Fenn et al., 2008; Yang et al., 2013), while the exclusion of third codon positions performed better in Dictyoptera and Hymenoptera (Cameron et al., 2012; Dowton et al., 2009). So reporting sensitivity tests for partitioning is of interest for different lineages and at different taxonomic scales. In this study, we present six new noctuoid mt genomes, including representatives of two newly sequenced families (Euteliidae and Nolidae), and three additional erebid subfamilies. We conducted phylogenetic analyses of mt genomes from 57 lepidopteran taxa. Our aim was to address the infraordinal relationships within Lepidoptera, particularly the family-level relationships within Noctuoidea and subfamily relationships within Erebidae (the largest noctuoid family) using mt genome data. In addition, we employ a range of analytical approaches to explore the phylogenetic utility of mt genomes.

2. Materials and methods

2.1. Samples and sequencing

Complete mt genomes were sequenced for six noctuoid species: *Asota plana lacteata* (Erebidae, Aganainae), *Agylla virilis* (Erebidae, Arctiinae), *Catocala deuteronympha* (Erebidae, Erebinae), *Eutelia adulatricoides* (Euteliidae, Euteliinae), *Gabala argentata* (Nolidae, Chloephorinae), and *Risoba prominens* (Nolidae, Risobinae). Three legs from each specimen were removed and preserved in 100% ethanol during collection. Legs were stored at −20 °C until DNA extraction, and voucher specimens were deposited at the Museum of the Institute of Zoology, Chinese Academy of Sciences. Total genomic DNA was extracted from the leg muscle tissue with the DNeasy Blood and Tissue Kit (QIAGEN, USA). Mt genomes were amplified and sequenced as described in our previous study (Yang et al., 2013). Generally, primer walking was employed for long fragments; fragments containing the control region were cloned using TOPO® TA Cloning® Kit (Invitrogen, USA), and the resultant plasmid DNA was isolated using the TIANpure Midi Plasmid Kit (Tiangen, China). The sequence data have been deposited in