Molecular Phylogenetics and Evolution 68 (2013) 42–54

Contents lists available at SciVerse ScienceDirect

Molecular Phylogenetics and Evolution

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

Reconstructing the phylogeny of aphids (Hemiptera: Aphididae) using DNA of the obligate symbiont *Buchnera aphidicola*

Eva Nováková [a,b,*], Václav Hypša [a], Joanne Klein [b], Robert G. Foottit [c], Carol D. von Dohlen [d], Nancy A. Moran [b]

[a] Faculty of Science, University of South Bohemia, and Institute of Parasitology, Biology Centre, ASCR, v.v.i., Branisovka 31, 37005 Ceske Budejovice, Czech Republic
[b] Department of Ecology and Evolutionary Biology, Yale University, 300 Heffernan Dr., West Haven, CT 06516-4150, USA
[c] Agriculture & Agri-Food Canada, Canadian National Collection of Insects, K.W. Neatby Bldg., 960 Carling Ave. Ottawa, Ontario, Canada K1A 0C6
[d] Department of Biology, Utah State University, UMC 5305, Logan, UT 84322-5305, USA

ARTICLE INFO

Article history:
Received 21 August 2012
Revised 7 March 2013
Accepted 13 March 2013
Available online 29 March 2013

Keywords:
Aphid
Evolution
Buchnera
Phylogeny
Informative markers

ABSTRACT

Reliable phylogenetic reconstruction, as a framework for evolutionary inference, may be difficult to achieve in some groups of organisms. Particularly for lineages that experienced rapid diversification, lack of sufficient information may lead to inconsistent and unstable results and a low degree of resolution. Coincidentally, such rapidly diversifying taxa are often among the biologically most interesting groups. Aphids provide such an example. Due to rapid adaptive diversification, they feature variability in many interesting biological traits, but consequently they are also a challenging group in which to resolve phylogeny. Particularly within the family Aphididae, many interesting evolutionary questions remain unanswered due to phylogenetic uncertainties. In this study, we show that molecular data derived from the symbiotic bacteria of the genus *Buchnera* can provide a more powerful tool than the aphid-derived sequences. We analyze 255 *Buchnera* gene sequences from 70 host aphid species and compare the resulting trees to the phylogenies previously retrieved from aphid sequences, only. We find that the host and symbiont data do not conflict for any major phylogenetic conclusions. Also, we demonstrate that the symbiont-derived phylogenies support some previously questionable relationships and provide new insights into aphid phylogeny and evolution.

1. Introduction

Aphids form a distinctive insect clade that features considerable variability in interesting biological traits, such as the presence of many distinct, yet genetically identical, forms of females during the life cycle (polyphenism), alternation of sexual and asexual reproduction, sterile soldier morphs, and seasonal alternation between unrelated groups of host plants. These traits vary among species, reflecting a long evolutionary history of biogeographical expansions and contractions and codiversification with plant hosts. Understanding the evolution of these traits thus requires a reliable phylogeny as a framework for particular evolutionary hypotheses. However, except for a few generally accepted aspects, studies on aphid phylogeny have not yet produced a clear picture of many relationships within this group.

Earlier attempts to reconstruct aphid phylogeny using morphology resulted in conflicting evolutionary scenarios (Heie, 1987; Wojciechowski, 1992). While an overall picture describing three distinct groups with either viviparous (Aphididae) or oviparous (Adelgidae and Phylloxeridae) parthenogenetic females was generally accepted, the phylogeny of the most diverse group, the Aphididae, remained unclear. More recent efforts have applied DNA sequence data to the reconstruction of Aphididae phylogeny. An emergent problem with such data is a lack of sufficient phylogenetic signal, mostly ascribed to a rapid adaptive radiation within Aphididae (von Dohlen and Moran, 2000; Martínez-Torres et al., 2001). Recently, this difficulty has been partly compensated by combining data from nuclear and mitochondrial genes (Ortiz-Rivas and Martínez-Torres, 2010); however, even this extended source of information leaves many relationships within Aphididae unresolved, and several evolutionary questions unanswered. Some questions concern the number of origins of particular aspects of biology. For example, three tribes with similar life cycles that include dwarf sexual forms lacking mouthparts (Eriosomatini, Fordini, Pemphigini) have been traditionally classified together in the Eriosomatinae (formerly Pemphiginae) (Heie, 1980; Remaudière and Remaudière, 1997). Furthermore, mainly on the basis of shared

* Corresponding author at: Faculty of Science, University of South Bohemia, and Institute of Parasitology, Biology Centre, ASCR, v.v.i., Branisovka 31, 37005 Ceske Budejovice, Czech Republic. Fax: +420 385 310 388.
E-mail addresses: novaeva@paru.cas.cz (E. Nováková), vacatko@paru.cas.cz (V. Hypša), Robert.Foottit@agr.gc.ca (R.G. Foottit), carol.vondohlen@usu.edu (C.D. von Dohlen), nancy.moran@yale.edu (N.A. Moran).

1055-7903/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ympev.2013.03.016

life cycle characteristics, Eriosomatinae, Hormaphidinae, and Anoeciinae have been hypothesized to share a most recent common ancestor (Heie, 1987). Most molecular phylogenies fail to support these ideas, however, because they recover Eriosomatinae as paraphyletic, and position Eriosomatinae, Hormaphidinae, and Anoeciinae as unrelated lineages. Other biological questions that could be illuminated by improved phylogenetic reconstruction concern whether the common ancestor of extant Aphididae fed on gymnosperms or angiosperms, and whether transitions between these major plant groups were accompanied by changes in diversification rates.

Inclusion of additional markers with appropriate information capacity, as well as additional taxa to bisect long terminal branches, are useful approaches to resolving ambiguous nodes of a rapid radiation (Whitfield and Lockhart, 2007). Nuclear gene markers would be the obvious, potentially informative additions for reconstructing the late Cretacous radiation of Aphididae. However, due to the possibility of paralogy, designation of orthologous nuclear sequences across a broad selection of species could require extensive experimental work. This is particularly relevant for aphids, because a vast amount of gene duplication affecting more than 2000 gene families was confirmed within the *Acyrthosiphon pisum* genome (The International Aphid Genomics Consortium, 2010). This extensive duplication of nuclear genes has been ongoing during aphid evolution, resulting in complex gene trees in which orthologs and paralogs are difficult to distinguish and the species phylogeny is obscured (e.g., Rispe et al., 2008; Nováková and Moran, 2012).

Aphids fortunately possess an additional source of inherited genetic material apart from their own genomes, in the form of obligate, maternally transferred, and highly derived bacterial mutualists. The association between aphids and bacteria of the genus *Buchnera* is one of the most extensively studied symbiotic systems since Buchner (1953) suggested their mutualistic association. As in many other mutualistic relationships, the bacteria supply essential nutrients to their hosts, which in turn provide a stable environment for their bacterial partners. Genomic studies have revealed that particular amino acids, vitamins and sterols are supplemented by *Buchnera* (Akman Gunduz and Douglas, 2009; Moran et al., 2005; Moran and Degnan, 2006; Moya et al., 2008; Hansen and Moran, 2011). *Buchnera* is one of numerous groups of insect-associated symbionts for which evidence supports a long-term, strict cospeciation with the host (Moran et al., 1993, 1995; Clark et al., 2000). The acquisition of *Buchnera* symbionts by aphids is thought to have been a single event that took place by 150–200 MYA and was followed by cospeciation (Moran et al., 1993; Martínez-Torres et al., 2001). Due to this close evolutionary association, the symbiont and host phylogenies mirror each other for deeper evolutionary divergences. Thus, symbiont-derived data in principle can be used to reconstruct the evolutionary history of hosts.

Such an approach is appealing, because the small and simple genomes of symbionts may be an easier source of suitable sequences for phylogenetic analysis. This idea has been considered for some other insect-symbiont associations. For example, Kölsch and Pedersen (2009) suggested using endosymbionts for elucidating unresolved questions of reed beetle phylogeny. In addition to resolving host relationships, endosymbiont genes have been used for calculating divergence times in their hosts (e.g., *Cryptocercus* woodroaches (Maekawa et al., 2005)). *Buchnera* genomes consist of single-copy genes, most of which are shared across different *Buchnera* genomes (Moran et al., 2009); thus, they lack the complications presented by gene duplication and paralogy. Bacterial genomes possess several additional advantages, such as haploidy and absence of introns. Attempts to infer aphid and *Buchnera* phylogenies in a common framework have been undertaken several times (e.g., Munson et al., 1991; Moran et al., 1993; Martínez-Torres et al., 2001), mainly with the aim of testing phylogenetic correspondence of host and symbiont. Only a few studies were focused explicitly on reconstructing aphid phylogeny using *Buchnera* markers; these included analyses performed on *Uroleucon*, *Brachycaudus*, and *Mollitrichosiphum* species (Clark et al., 2000; Jousselin et al., 2009; Liu et al., 2013). All three studies ruled out the possibility of occasional horizontal transfers, even between closely related aphids sharing the same host plant. Even within species, phylogenies based on *Buchnera* genes are congruent with those based on mitochondrial sequences, confirming that this symbiont is strictly maternally inherited (Peccoud et al., 2009; Funk et al., 2000). Within an individual aphid, *Buchnera* genes also have higher copy number than aphid nuclear genes, so that *Buchnera* genes are relatively easy to amplify. DNA of *Buchnera* symbionts thus serves as a useful source of information for inferring aphid phylogeny. However, phylogenetic studies using *Buchnera* genes so far have been limited by sparse sampling across aphid taxa, and especially by taxonomic bias towards the subfamily Aphidinae.

In this study, we use *Buchnera*-derived sequences of five genes (*groEL*, *trpB*, *dnaB*, *ilvD* and 16S rDNA) from an extended taxonomic set to reconstruct phylogenetic relationships within the Aphididae. We then compare the resulting topologies to those reported from recent multilocus analyses of aphid-derived genes. We focus mainly on the most problematic questions in aphid evolution, such as monophyly of individual subfamilies/tribes and their relationships, variation in DNA substitution rates, and rooting of the Aphididae tree. We show that symbiont genes yield informative phylogenetic signal and have several methodological advantages.

2. Materials and methods

2.1. Sample collection

Our study was designed to reconstruct evolutionary relationships among a wide diversity of aphids, represented by a broad sample of most major Aphididae taxa. The collection includes 70 species from 15 of 25 subfamilies and 25 of 36 tribes recognized in the most recent, comprehensive classifications of aphids (Remaudière and Remaudière, 1997; Nieto Nafría et al., 1997) (Table S1).

2.2. DNA extraction, primer design and sequencing

For all species, several individuals were pooled and homogenized, and genomic DNA was extracted using QIAamp DNA Micro Kit (Qiagen). In efforts to compile sufficient data for a reliable phylogenetic reconstruction, we aimed to amplify five *Buchnera*-derived genes, namely 16S rDNA, *groEL*, *trpB*, *dnaB*, and *ilvD*. To confirm the aphid host species identification, the sequence of the aphid mitochondrial gene COI was obtained for each sample.

Three primer pairs allowing for nine combinations were designed for each of the targeted *Buchnera* genes, except for 16S rDNA. Pairs of highly degenerate primers, corresponding to the coding region of *groEL*, *trpB*, and *dnaB* genes, were designed according to alignments of partial sequences available in GenBank using Primaclade software (Gadberry et al., 2005). Primers for *ilvD* were designed in CLC Genomic Workbench (CLC bio A/S) based on an alignment of *Buchnera* genome sequences, with the reverse primers corresponding to a highly conserved region of the flanking tRNA. Primers for amplifying the 16S rDNA gene and aphid mitochondrial gene COI were from previous studies (Folmer et al., 1994; Hajibabaei et al., 2006; Munson et al., 1991; Mateos et al., 2006). More detailed information on primer pairs, including the sequence and the length of amplified regions, is summarized in Table S2.

To amplify PCR products from diverse aphid lineages, we used the two best-performing primer pairs for each gene (Table S2). In