Molecular Phylogenetics and Evolution 80 (2014) 1–10

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

Exploring new dating approaches for parasites: The worldwide Apodanthaceae (Cucurbitales) as an example

Sidonie Bellot *, Susanne S. Renner

Systematic Botany and Mycology, University of Munich (LMU), Menzinger Str. 67, 80638 Munich, Germany

ARTICLE INFO

Article history:
Received 7 January 2014
Revised 4 July 2014
Accepted 7 July 2014
Available online 22 July 2014

Keywords:
Local clocks
Parasitic plants
Random local clocks
Substitution rates
Uncorrelated lognormal clock
Bayes factors

ABSTRACT

Gene trees of holoparasitic plants usually show distinctly longer branch lengths than seen in photosynthetic closest relatives. Such substitution rate jumps have made it difficult to infer the absolute divergence times of parasites. An additional problem is that parasite clades often lack a fossil record. Using nuclear and mitochondrial DNA sequences of Apodanthaceae, a worldwide family of endoparasites living inside Fabaceae and Salicaceae, we compared several dating approaches: (i) an uncorrelated lognormal (UCLN) model calibrated with outgroup fossils, (ii) ages of host lineages as a maximal age in an UCLN model, (iii) user-assigned local clocks, and (iv) outgroup-fossil-calibrated random local clocks (RLC) with varying prior probabilities on the number of permitted rate changes (RLCu and RLCp models), a variable that has never been explored. The resulting dated phylogenies include all 10 species of the family, three in Australia, one in Iran, one in Africa, and the remainder in the Americas. All clock models infer a drastic rate jump between nonparasitic outgroups and Apodanthaceae, but since they distribute the rate heterogeneity differently, they result in much-different age estimates. Bayes factors using path and stepping-stone sampling indicated that the RLCp model fit poorly, while for *matR*, topologically unconstrained RLCu and UCLN models did not differ significantly and for 18S, the UCLN model was preferred. Under the equally well fitting models, the Apodanthaceae appear to be a relatively old clade, with a stem age falling between 65 and 81 my, the divergence of *Apodanthes* from *Pilostyles* between 36 and 57 my ago, and the crown age of the Australian clade 8–18 my ago. In our study system, host-age calibrations did not yield well-constrained results, but they may work better in other parasite clades. For small data sets where statistical convergence can be reached even with complex models, random local clocks should be explored as an alternative to the exclusive reliance on UCLN clocks.

© 2014 Elsevier Inc. All rights reserved.

1. Introduction

Over the past 15 years, relaxed molecular clock approaches (Sanderson, 1997, 2002; Thorne et al., 1998; Huelsenbeck et al., 2000; Drummond et al., 2006; Rannala and Yang, 2007) have become the dominant strategy for inferring absolute times from molecular data. This is because fossil-calibrated studies have often revealed punctuated rates of substitutions, which prevent the use of global clocks (Li and Tanimura, 1987; Yoder and Yang, 2000; Bell et al., 2010). To accommodate such situations, workers have assumed either that substitution rates of ancestors and descendants are correlated (Sanderson, 1997, 2002; Thorne et al., 1998; Rannala and Yang, 2007) or that rate variation is unrelated to species relationships (Drummond et al., 2006; Rannala and Yang, 2007). In the latter case, the rate for each branch is drawn from a single underlying distribution, such as a lognormal or exponential distribution of which the parameters are estimated in the analysis (Drummond et al., 2006). A problem with the uncorrelated lognormal (UCLN) model of Drummond et al. is that it may bias results in favor of rate variation, making UCLN clocks poorly suited for data sets in which a strict clock best explains the branch lengths. Rannala and Yang (2007) suggested that this bias should decrease in trees with many branches.

Another approach for dealing with abrupt rate variation is to use local clocks in which regions of a phylogeny are calibrated using paleontological time points, with each region being given its own substitution rate (Li and Tanimura, 1987; Yoder and Yang, 2000). Local molecular clocks have been used in relatively few studies (e.g., Bailey et al., 1991; Yoder and Yang, 2000; Baum et al., 2004; Aguileta et al., 2006; Nunome et al., 2007; Aris-Brosou, 2007), perhaps because they require an *ad hoc* choice of nodes in the phylogeny where a rate change is assumed to have occurred. This permits neither the modeling of uncertainty in tree topology nor in the clade most likely to have undergone the abrupt rate change. To overcome this problem, Drummond and Suchard

* Corresponding author.
E-mail address: sido.bellot@gmail.com (S. Bellot).

http://dx.doi.org/10.1016/j.ympev.2014.07.005
1055-7903/© 2014 Elsevier Inc. All rights reserved.

(2010) developed the random local clock (RLC) approach, which employs a Monte Carlo Markov chain to investigate nested local clock configurations. Their method co-estimates the phylogenetic tree and the number, magnitude and location of rate changes along it. Importantly, the RLC model includes the possibility of zero rate changes, meaning it can also serve to test whether a single global rate fits a particular data set. To compare the fit of RLC and UCLN models, authors have used Bayes factors because these models are not nested (Avaria-Llautureo et al., 2012). So far, nobody seems to have studied the impact of changing the *a priori* permitted number of rate changes in an RLC model.

Abrupt shifts in molecular rates characterize the holoparasitic lineages of flowering plants, judging from branch lengths in trees from nuclear 18S and 26S ribosomal RNA sequences, mitochondrial *matR*, *atp1*, *cox1*, and *nad1*, and the plastid 16S ribosomal RNA (Nickrent and Starr, 1994; Barkman et al., 2007; Bromham et al., 2013). Parasitism has evolved at least 13 times in the angiosperms (Barkman et al., 2007) and is associated with drastic changes in photosynthesis-related genes of the plastome, probably also in the mitochondrial and nuclear genomes (Krause, 2012; Wicke et al., 2013). In addition, many physiological, ecological, and other life history changes occur with the onset of a parasitic life style, perhaps including changes in effective population size or mating system; most parasitic lineages are long-lived and many are dioecious (Bellot and Renner, 2013). These changes likely modified selection on mutation rates or DNA repair. Regardless of their causes, the extremely long stem lineages seen in holoparasitic angiosperm clades have made it difficult to fit clock models for these groups. Nevertheless, Naumann et al. (2013) recently calculated the stem ages of all 13 parasitic lineages using an UCLN clock model and (mostly) single representatives of each parasite lineage.

Here we use the parasite family Apodanthaceae to compare local clocks, random local clocks, and clock models relying on the UCLN prior rate distribution, calibrated with either outgroup fossils or host ages. If a parasite species or clade occurs only on one host species or genus, the divergence time of the host provides a maximal age for the parasite's age, because a host-specific parasite is unlikely to be older than its host. This type of calibration has not been tried in plants, but may be useful for specialized endoparasites, such as Apodanthaceae, at least as a cross-validation of other calibration approaches. All species in this family lack any green parts and live permanently inside trees or shrubs of Fabaceae or Salicaceae from the stems of which they emerge to flower and fruit (Bellot and Renner, 2014). The family is distributed in North and South America, Africa, Iran, and Australia, and its 10 species belong to two genera, *Pilostyles* and *Apodanthes* (Bellot and Renner, 2014).

From the worldwide distribution of Apodanthaceae, it is almost certain that transoceanic dispersal must have played a role in the family reaching the different continents. Compared to the three other endoparasitic lineages of flowering plants, Cytinaceae with 10 species, Mitrastemonaceae with two, and Rafflesiaceae with 34, Apodanthaceae have by far the largest geographic and host range, which suggests that they may be older than these other lineages. With a temporal framework it would be possible to infer how long it took these parasites to evolve host specificity to either Fabaceae or Salicaceae and to acquire their extreme adaptations. Absolute ages also are needed to test hypotheses about trait correlations (Hardy and Cook, 2012) and genomic changes, for example in the plastome (Wicke et al., 2013). To achieve these ultimate goals, we here focus on three questions, (i) what is the most plausible model of substitution rate change in the Apodanthaceae; (ii) how different are the divergence times estimated with user-specified local clocks, random local clocks, or an uncorrelated lognormal clock model; and (iii) can plausible ages be inferred by using host ages as maximal constraints on parasite ages (as an alternative calibration instead of an outgroup fossil).

2. Material and methods

2.1. Plant material, DNA isolation, PCR amplifications and sequencing

DNA was extracted from 21 specimens of Apodanthaceae representing most of the named species (de Vattimo, 1971). Table S1 shows herbarium vouchers and Genbank accession numbers. Total genomic DNA was extracted from herbarium specimens or silica-dried flowers using the commercial plant DNA extraction kit NucleoSpin (Macherey–Nagel, Düren, Germany) or alternatively the NucleoSpin® Food kit of the same company. To overcome problems with fragmented DNA, we designed custom primers for *matR* and 18S using Primer3Plus v. 2.3.6 (Untergasser et al., 2012); Table S2 lists all primers, which were used for both polymerase chain reactions (PCRs) and sequencing. PCR products were purified with the ExoSAP or FastAP clean-up kits (Fermentas Life sciences, St. Leon-Rot, Germany), and sequencing relied on the Big Dye Terminator cycle sequencing kit (Applied Biosystems, Foster City, CA, USA) and an ABI 3130-4 automated capillary sequencer. In total, 35 *matR* and 18S sequences were newly obtained for this study. The 18S sequence of *P. aethiopica* was obtained by a BLAST search against Illumina data of its genomic DNA (Bellot and Renner, unpublished). Additional 18S and *matR* sequences of Malpighiales, Fagales and Cucurbitales were retrieved from GenBank, and their accession numbers are provided in Table S1.

2.2. Sequence alignment and phylogenetic analyses

All sequences were blasted against GenBank to rule out contamination. Chromatograms were checked and sequences were edited using BioEdit v. 7.2.0 (Hall, 1999) and Geneious R7 (Biomatters, available from http://www.geneious.com). Two alternative alignments were performed for each marker, using the program MAFFT v. 7 (Katoh, 2013). The 18S alignments were obtained using either the "E-INS-i" strategy with manual editing or the "Q-INS-i" strategy, which takes into account the secondary structure of the RNA, without further manual editing. The *matR* alignments were obtained using the E-INS-i strategy with some manual editing, with and without translation into amino acids. Since the alignment methods did not affect the maximum likelihood (ML) topologies obtained, we chose the Q-INS-i-based alignment of 18S and the alignment performed without considering amino acids for *matR* in final analyses. Alignments and ML trees have been deposited in TreeBASE (accession number 15658). To select the best fitting model of nucleotide substitutions, JModelTest v. 2.1 (Darriba et al., 2012) was run on the final matrices with the following parameters: 11 substitution schemes, +F, +G 4 categories, ML optimized, and Best tree search, and the best model was then chosen using the corrected Akaike Information Criterion with model averaging allowed. For both markers, GTR + G was the preferred model. Phylogenetic searches using the separate or combined DNA matrices were performed under maximum likelihood optimization as implemented in RAxML-7.2.8-ALPHA (Stamatakis, 2006), with 100 bootstrap replicates, and rooting on *Clusia rosea* (Malpighiales).

2.3. Analysis of substitution rates and molecular dating

For clock dating, we constructed DNA alignments of nuclear 18S and mitochondrial *matR* with and without outgroups of Apodanthaceae. Matrices included 8 ingroup *matR* sequences, 10 ingroup 18S sequences, and 11, 12 or no outgroup sequences (the monospecific genus *Apodanthes* was then used as the outgroup). We used the Bayesian approach implemented in the software BEAST 1.8.0 (Drummond et al., 2006, 2012). All analyses used the GTR + G (4

Download English Version:

https://daneshyari.com/en/article/2833797

Download Persian Version:

https://daneshyari.com/article/2833797

Daneshyari.com