Molecular Phylogenetics and Evolution 58 (2011) 297–303

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

Reconstructing the phylogeny of "*Buarremon*" brush-finches and near relatives (Aves, Emberizidae) from individual gene trees

Alexander Flórez-Rodríguez [a,*], Matthew D. Carling [b], Carlos Daniel Cadena [a]

[a] *Laboratorio de Biología Evolutiva de Vertebrados, Departamento de Ciencias Biológicas, Universidad de los Andes, Apartado 4976, Bogotá, Colombia*
[b] *Berry Biodiversity Conservation Center, Department of Zoology and Physiology, University of Wyoming, Laramie, WY 82071, USA*

ARTICLE INFO

Article history:
Received 26 July 2010
Revised 2 November 2010
Accepted 15 November 2010
Available online 25 November 2010

Keywords:
Anomalous genes trees
BEST
COAL
Minimizing deep coalescences
Species tree
Topological incongruence

ABSTRACT

Gene trees are often assumed to be equivalent to species trees, but processes such as incomplete lineage sorting can generate incongruence among gene topologies and analyzing multilocus data in concatenated matrices can be prone to systematic errors. Accordingly, a variety of new methods have been developed to estimate species trees using multilocus data sets. Here, we apply some of these methods to reconstruct the phylogeny of *Buarremon* and near relatives, a group in which phylogenetic analyses of mitochondrial DNA sequences produced results that were inconsistent with relationships implied by a taxonomy based on variation in external phenotype. Gene genealogies obtained for seven loci (one mitochondrial, six nuclear) were varied, with some supporting and some rejecting the monophyly of *Buarremon*. Overall, our species-tree analyses tended to support a monophyletic *Buarremon*, but due to lack of congruence between methodologies, resolution of the phylogeny of this group remains uncertain. More generally, our study indicates that the number of individuals sampled can have an important effect on phylogenetic reconstruction, that the use of seven markers does not guarantee obtaining a strongly-supported species tree, and that methods for species-tree reconstruction can produce different results using the same data; these are important considerations for researchers using these new phylogenetic approaches in other systems.
© 2010 Elsevier Inc. All rights reserved.

1. Introduction

The accurate reconstruction of phylogenetic relationships among species using molecular data can be complicated for a variety of reasons (Carstens and Knowles, 2007; Brumfield et al., 2008; Degnan and Rosenberg, 2009). One obstacle is the stochastic sorting of ancestral polymorphisms following species divergence at deep or shallow levels of the phylogeny, which can result in discordant topologies between gene and species trees (Pamilo and Nei, 1988; Takahata, 1989; Maddison and Knowles, 2006; McCormack et al., 2009). Furthermore, if reproductive isolation between taxa is not complete, then gene flow can cause incongruent topologies across genes (Meng and Kubatko, 2009). In addition, processes of gene duplication (Fitch, 1970) and horizontal transfer (Cummings, 1994) can also complicate the traditional assumption that gene trees always reflect species trees (Nichols, 2001). Some of these problems are particularly acute when phylogenies are reconstructed from single-locus datasets (e.g. mitochondrial DNA; Jennings and Edwards, 2005).

The ability to obtain sequence data from multiple loci across taxa is one of the major recent breakthroughs in molecular systematics, and has brought with it new opportunities and challenges for phylogeny reconstruction (Brito and Edwards, 2009; Knowles, 2010; Knowles and Kubatko, 2010). To date, the most common approach used in multilocus phylogenetics is concatenation of data using supermatrices, an approach assuming that all gene trees have the same topology (Rokas et al., 2003; Philippe et al., 2009). However, this approach might be positively misleading because of the existence of anomalous gene trees (i.e. gene trees that are more likely than the tree matching the species tree; Degnan and Rosenberg, 2006; Liu and Edwards, 2009). In addition, obtaining well-supported trees consistent with the true phylogeny using concatenated data might require a large number of loci in comparison to other, novel methods for species-tree reconstruction (Edwards et al., 2007; see below). Another often employed approach for phylogeny reconstruction from multilocus data is to construct consensus trees based on genealogies obtained independently for each locus (De Queiroz, 1993), but this requires a greater number of genes than concatenation to obtain a similarly supported tree (Gadagkar et al., 2005) and is also prone to be positively misleading as the number of genes increases when anomalous gene trees exist (Degnan et al., 2009). Owing to these limitations, developing alternatives to concatenation and consensus methods for the reconstruction of robust species trees has become an important priority.

Novel analytical tools have allowed a movement towards multilocus methodologies for phylogenetic reconstruction more robust

* Corresponding author. Fax: +57(1) 33934949x2817.
E-mail addresses: al-flore@uniandes.edu.co (A. Flórez-Rodríguez), mdc248@cornell.edu (M.D. Carling), ccadena@uniandes.edu.co (C.D. Cadena).

1055-7903/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2010.11.012

than concatenation or consensus methods (Jennings and Edwards, 2005; Carling and Brumfield, 2008; Degnan and Rosenberg, 2009; Knowles, 2009, 2010). For example, one can attempt to reconcile gene trees contained within species trees by minimizing the number of deep coalescences, i.e., the coalescence of two gene copies that predates a particular speciation event (Maddison, 1997; Maddison and Knowles, 2006; Leaché, 2009). Alternatively, the BEST (Bayesian Estimation of Species Trees) method estimates the joint posterior distribution of gene trees for each locus and uses the resulting joint posterior distribution of gene trees to approximate the Bayesian posterior distribution of the species tree based on coalescent theory (Liu and Pearl, 2007; Edwards et al., 2007). Yet another alternative is to calculate the probability of a genealogy given a species tree under the coalescent (Degnan and Salter, 2005). Despite these developments, as the field of multilocus phylogenetics is still maturing, studies using species-tree approaches with empirical data are scarce (Brumfield et al., 2008; Liu et al., 2008; Linnen and Farrell, 2008; Hird et al., 2010; Waters et al., 2010). In this study, we use three methods of species-tree reconstruction based on multilocus data to revisit an empirical phylogenetic question of interest in systematic ornithology. In so doing, we come across some practical issues related to the effect of taxon sampling and the variation in results among methods that should be of interest to developers and users of such methods.

The genus *Buarremon* (Aves, Emberizidae), as traditionally defined, consists of three morphologically similar species: *Buarremon torquatus*, *Buarremon brunneinucha*, and *Buarremon virenticeps*. However, a recent study rejected the monophyly of the genus. Based on analyses involving sequences of four mitochondrial genes (ND2, cyt b, ATPase 6, ATPase 8) and two nuclear introns (ACOI and MUSK), the clade formed by representatives of multiple populations of *B. torquatus* was recovered as sister to the monophyletic genus *Arremon*, whereas *B. brunneinucha* and *B. virenticeps* formed a clade that was sister to a clade formed by species in the genus *Lysurus* albeit with low support (Cadena et al., 2007; Fig. 1a). This result was unexpected considering the overall phenotypic similarity of *Buarremon* taxa, but it led to the merging of the three genera in an expanded genus *Arremon* (Remsen et al., 2010). However, the results of this study were supported mainly by mitochondrial DNA sequences and the deep internodes of the mitochondrial topologies were notably shorter than the terminal branches. The existence of short internal branches can lead to retention of ancestral polymorphisms, representing one of the most difficult scenarios for inferring phylogenies from single-locus data sets owing the high stochasticity in gene sorting. Under such scenarios, mitochondrial DNA can reveal trees with good nodal support that are incongruent with the species tree (Carling and Brumfield, 2008; Leaché, 2009, McCormack et al., 2009). In addition, the mitochondrial DNA topology was at least partly inconsistent with the topologies of two nuclear introns (Cadena et al., 2007). Thus, in this study we revisit the relationships of *Buarremon* and related genera by reconstructing the species tree based on phylogenetic analyses of sequences from multiple loci. We discuss the implications of our results in relation to challenges in sampling design and in the use of different methods, which might be common to other studies seeking to reconstruct species trees using multilocus sequence data.

2. Materials and methods

2.1. Sampling, PCR amplification and sequencing

We obtained frozen tissue samples from the collections of Instituto Alexander von Humbolt (IAvH) and the Banco de Tejidos of the Museo de Historia Natural, Universidad de los Andes (ANDES-BT) for a single individual of four focal taxa (*B. torquatus* (IAvH-1145), *B. brunneinucha* (ANDES-BT-0120), *Arremon schelegeli* (ANDES-BT-0016) and *Lysurus castaneiceps* (IAvH-CT-825)) and one outgroup (*Atlapetes latinuchus* (ANDES-BT-0130), a valid strategy considering these species are strongly supported monophyletic groups (Cadena et al., 2007). Note we did not include *B. virenticeps* in our study because this species was nested with strong support within a clade formed by populations of *B. brunneinucha* in phylogenetic analyses of mitochondrial and nuclear DNA sequences (Cadena et al., 2007). Therefore, the inclusion of this taxon (and of populations of *B. torquatus* that likely merit species status; Cadena and Cuervo, 2010) was not necessary to address the monophyly of *Buarremon*. The important question in this regard is whether the *brunneinucha–virenticeps* clade and the *torquatus* clade form a monophyletic group to the exclusion of the *Lysurus* and *Arremon* clades.

Total DNA was extracted from all samples using a DNeasy tissue kit (QIAGEN, Valencia, CA), following the manufacturer's protocol. Subsequently, we amplified six nuclear introns (four autosomal, two z-linked) and one mitochondrial protein-coding gene (Table 1) using primers published by Slade et al. (1993), Sorenson et al. (1999), and Kimball et al. (2009). The concentrations and conditions used for PCR were those described by Cadena et al. (2007). Amplicons were cleaned using Exosap IT (USB corporation, Cleveland, Ohio) and then sequenced in both directions. Resulting chromatographs were assembled in Geneious Basic 4.02. (Drummond et al., 2007). In cases where double peaks of equal height were detected in the sequence, the site was considered ambiguous (i.e. we did not attempt to phase haplotypes because sites with double peaks were scarce and because we had data for a single individual per species, which impeded haplotype estimations).

2.2. Alignment and conventional phylogenetic analyses

Sequences were aligned using the MUSCLE algorithm (Edgar, 2004) implemented in Geneious (Drummond et al., 2007) and edited manually. Intralocus recombination was tested using the program RecombiTEST (Piganeau et al., 2004). Genealogies were reconstructed individually for each locus using maximum likelihood (ML) and Bayesian inference (BI) methods, and we also conducted analyses using a concatenated matrix that included sequences of all seven genes and a partitioned matrix specifying a substitution model for each of the seven loci. For each analysis, we implemented the model of nucleotide substitution selected as the best-fit to the data (Table 1) based on the Akaike Information Criterion using ModelTest 3.7 for ML (Posada and Crandall, 1998) and MrModelTest 2.3 for BI (Nylander, 2004). Branch-and-bound searches were conducted using PAUP* 4.0b10 (Swofford, 2002) for ML analyses; nodal support was assessed with 1000 ML heuristic bootstrap replicates with tree bisection-reconnection (TBR) branch swapping. Bayesian analyses were conducted in MrBayes 3.1.2 (Ronquist and Huelsenbeck, 2003) and consisted of two runs of four MCMC chains for 15,000,000 generations sampled every 1000 generations; the first 25% of the trees sampled was discarded as burn-in.

We used the program Tracer v.1.4.1 (Rambaut and Drummond, 2007) to evaluate sampling of the tree and parameter space in Bayesian analyses. Because plots of number of generations vs. likelihood showed stabilization, effective sample sizes for all parameters was always greater than 200, and the average standard deviation of split frequencies across runs was less than 0.002 in all the analyses, chains likely sampled the posterior distributions adequately. To assess convergence of MCMC runs, we plotted posterior probabilities of clades as a function of generation number and compared results of different runs by plotting the posterior probabilities of all splits for paired runs using AWTY (Wilgenbusch et al., 2004).

Download English Version:

https://daneshyari.com/en/article/5920790

Download Persian Version:

https://daneshyari.com/article/5920790

Daneshyari.com