Molecular Phylogenetics and Evolution 61 (2011) 958–963

Contents lists available at SciVerse ScienceDirect

Molecular Phylogenetics and Evolution

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

Short Communication

Estimating phylogenies for species assemblages: A complete phylogeny for the past and present native birds of New Zealand

Robert Lanfear *, Lindell Bromham

Centre for Macroevolution and Macroecology, Evolution Ecology and Genetics, Research School of Biology, The Australian National University, Canberra ACT 0200, Australia

ARTICLE INFO

Article history:
Received 9 May 2011
Revised 19 July 2011
Accepted 23 July 2011
Available online 3 August 2011

Keywords:
Community
GeneFinder
Supermatrix
Bayesian
Extinct
Extant
Macroecology

ABSTRACT

Phylogenies of species assemblages are vital to many studies of community ecology and macroecology. However, few assemblage phylogenies are currently available, and the approaches that have been used to construct them have had a number of important limitations. Here, we estimate a Bayesian posterior sample of phylogenies for all 275 known extinct and extant native birds of New Zealand. To do this, we use an approach that uses freely-available data and software, can incorporate both extinct and extant taxa, does not rely on having DNA sequence data available for all species in the assemblage, and explicitly accounts for phylogenetic uncertainty. This approach produces a set of phylogenies that describes our knowledge and uncertainty about the relationships among the NZ birds. We also present a python script, GeneFinder, which can be used to efficiently gather publicly available sequence data in order to construct a supermatrix of DNA sequences for constructing assemblage phylogenies. The approach we describe paves the way for estimating assemblage phylogenies for any species assemblage.

1. Introduction

Phylogenetic trees are playing an increasingly important role in understanding biodiversity and macroecology (Diniz-Filho and Bini, 2008; Emerson and Gillespie, 2008). For example, phylogenies have been used to assess the degree to which species assemblages are structured by competition (e.g. Alexandrou et al., 2011), to estimate timescales for the study of diversification (e.g. Rabosky, 2009), and to correct for statistical non-independence in species data due to patterns of descent in macroecological studies (e.g. Vamosi and Vamosi, 2007). In particular, phylogenies are being increasingly used in studies of community ecology (Kembel et al., 2010; Webb and Donoghue, 2005). Studies such as these require assemblage phylogenies – that is, phylogenetic trees that show the relationships between all of the species found in a given assemblage.

Because of the difficulty in obtaining assemblage phylogenies, phylogenetically-based studies of macroecology and community ecology have typically been restricted to well studied flora and fauna, and have focused on taxa where substantial phylogenies already exist or can readily be constructed (e.g. Webb and Donoghue, 2005). For phylogenetically-based studies of community ecology to have a wider application, techniques must be developed for constructing assemblage phylogenies for a much wider taxonomic and geographic sample of assemblages.

Several approaches have been taken to producing assemblage phylogenies. One is to use taxonomy as a proxy for phylogeny, using the nested taxonomic hierarchy to represent levels of relatedness (e.g. Vamosi and Vamosi, 2007). The advantage of using taxonomy is that it is generally available for all species assemblages, however, it is limited by the lack of resolution within taxonomic categories. For example, in the absence of any other information, it must be assumed that all species in a genus are equally closely related. Furthermore, taxonomy gives no indication of evolutionary divergence, which precludes the use of metrics that require information on branch lengths, such as many tests of the phylogenetic structure of communities (e.g. Kembel et al., 2010).

An alternative approach is to use a supertree constructed from published estimates of phylogeny for different members of the assemblage, either by combining different phylogenies into one tree (e.g. Vamosi and Vamosi, 2007) or using a formal algorithm that uses a range of source trees to produce a single phylogeny of all component species (e.g. Kembel and Hubbell, 2006). The supertree approach has been extremely fruitful for comparative biology and community ecology (Bininda-Emonds, 2004; Webb and Donoghue, 2005). However, it can only be used to produce assemblage phylogenies when all members of an assemblage have been included in past phylogenetic studies. Furthermore, some supertree methods produce phylogenies of limited resolution and without meaningful branch lengths (Bininda-Emonds, 2004). These drawbacks mean that assemblage phylogenies constructed from supertrees can be of limited utility in many cases.

* Corresponding author. Fax: +61 2 6125 5573.
E-mail address: rob.lanfear@anu.edu.au (R. Lanfear).

1055-7903/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2011.07.018

Another approach to constructing assemblage phylogenies is to sample DNA from all species in an assemblage and use this data to construct the phylogeny. For example, researchers used a DNA barcoding approach to estimate an assemblage phylogeny all of the 296 woody plant species found on Barro Colorado Island in Panama (Kress et al., 2009). However, this approach will be only be possible for those research programs for which DNA samples can be obtained for all species in the assemblage, and for which time and resources are available to sequence and analyse such samples.

Ideally, we would like to have an approach to constructing assemblage phylogenies that is able to incorporate species both with and without available DNA sequences, and that takes advantage of freely-available data and methods. Furthermore, it is important to take phylogenetic uncertainty into account, as studies that use a single phylogeny as a point estimate of evolutionary history are vulnerable to error if that phylogeny is incorrect. Here, we describe an approach for constructing assemblage phylogenies that fulfils these criteria. This approach differs from approaches designed to estimate phylogenies for entire clades (Smith et al., 2009) because it makes use of the fact assemblage phylogenies rarely contain all the species within a given clade.

The approach we describe allows us to estimate a set of phylogenetic trees that represents the current state of knowledge (and uncertainty) about the phylogenetic relationships among the species of an assemblage (see Fig. 1). Our approach takes advantage of Bayesian molecular phylogenetic methods, but does not rely on having DNA sequence data available for all species in the assemblage. It enables us to include both extinct and extant taxa in the analysis, and combines publically available DNA sequence data with additional information on species relationships to produce a posterior distribution of assemblage phylogenies with meaningful branch lengths. To facilitate this approach, we have written a Python script (GeneFinder) that can be used to quickly and efficiently assemble large DNA supermatrices from GenBank (http://www.ncbi.nlm.nih.gov/GenBank/), a key step in estimating assemblage phylogenies.

We demonstrate our approach by constructing a complete phylogeny of the native birds of New Zealand (NZ). NZ birds are a unique case study in understanding diversification and extinction, because of their fascinating evolutionary history, high level of endemism, and because a large proportion of the assemblage is either extinct or currently threatened (Trewick and Gibb, 2010). We show how publicly available information can be used to produce a Bayesian posterior sample of assemblage phylogenies of NZ birds, including species for which no DNA sequence data is available and species that have gone extinct either recently or in the distant past.

2. Materials and methods

2.1. List of New Zealand bird species

Based on the Checklist of the Birds of New Zealand (OSNZ Checklist Committee, 2010), we compiled a list of all native birds of New Zealand, defined as species that are naturally present and maintain or have maintained successful breeding populations in New Zealand (step 1, Fig. 1). We did not include species that were deliberately introduced, but we did include recent natural colonists (for example, the silver eye *Zosterops lateralis* which arrived in New Zealand unassisted in 1832). We included all known species from the North and South Islands, as well as the Chatham Islands, Stuart Island, Antipodes Island, Auckland Island, Bounty Island, Campbell Island, the Snares Island group, and all offshore islands close to the main islands. Extinct species known from the fossil or subfossil record of NZ were included if they have been identified to at least genus level in the scientific literature. This resulted in a list of 275 species and subspecies, of which 188 are extant and 87 are extinct (see Supplementary information). Hereafter, we will use the term "species" to refer to each of the 275 taxa in this list, whether they have a binomial (*Genus species*) or trinomial (*Genus species subspecies*) epithet.

2.2. DNA sequence alignment

We searched for appropriate DNA sequences from GenBank for each of the 275 species in the list. Any species with sequence data in GenBank was assigned a unique GenBank TaxonID (step 2, Fig. 1), taking into account synonyms and recent taxonomic revisions. For example, *Dinornis struthoides* (slender bush moa), *Dinornis giganteus*, (giant moa), and *Dinornis novaezealandiae* (large bush moa) were synonymised to *D. novaezealandiae* when genetic evidence demonstrated they were the same species (Bunce et al., 2003). These taxa are represented by separate TaxonIDs in GenBank, but because they have been synonymised into the same species we used genetic data from all three taxa (TaxonIDs 8818, 147464, 237965) to represent *D. novaezealandiae*. A key advantage of using TaxonIDs is that the work involved in matching species lists with entries in GenBank (see Thomson and Shaffer, 2010) need

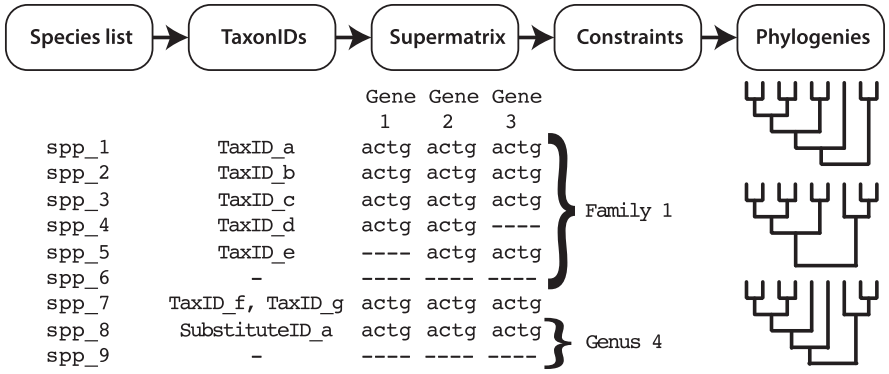

Fig. 1. Five steps to building an assemblage phylogeny. (1) Assemble a list all species in the assemblage. (2) Assign GenBank TaxonIDs to species in the list if: (i) a species has appropriate DNA sequence data available in GenBank; or (ii) when DNA from a suitable substitute taxon, such as a congeneric not found in the target assemblage, is available (in which case the TaxonID of the substitute taxon is used, see main text). In certain cases, for instance due to recent taxonomic revisions, more than one TaxonID may be assigned to a single species in the list (e.g. *spp_7*). (3) Assemble a supermatrix of DNA sequences based on the TaxonIDs. Species without any appropriate DNA sequence data are included as empty sequences (e.g. *spp_6*). (4) Identify reliable constraints from the literature (e.g. well-supported monophyly of some taxonomic groups), to provide additional information on known species relationships and to aid with the placement of species for which DNA sequence data is unavailable (e.g. *Family 1* and *Genus 4*). (5) Analyse the data in a Bayesian framework, to produce a Bayesian posterior sample of phylogenetic trees.

be performed only once, and lists of available sequences for species in the assemblage can be easily updated, which may help to reduce phylogenetic uncertainty as new DNA sequences become available on GenBank.

Once we had assigned TaxonIDs to species, we used a Python script, GeneFinder, to determine which species had DNA sequences available that could be used to estimate the assemblage phylogeny (see Supplementary information). GeneFinder searches GenBank using a list of TaxonIDs, and retrieves the longest available sequence of a given gene (e.g. CytB) for each TaxonID. GeneFinder can be used to quickly and efficiently explore and assemble the available data on GenBank for any given list of TaxonIDs, and thus provides a useful starting point for constructing DNA supermatrices for assemblage phylogenies or other purposes. Using GeneFinder, we determined that 163 of the 275 taxa in our list had appropriate DNA sequence data available on GenBank.

We had two strategies for including species that did not have any appropriate DNA sequence data in GenBank. First, we exploited the fact that DNA from any member of a lineage can be used to estimate that lineage's evolutionary history, by using sequences available for close relatives to place missing species in our phylogeny (using substitute TaxonIDs; step 2, Fig. 1). This approach is of particular utility when estimating assemblage phylogenies, as lineages from within one assemblage can often be placed using sequences from species that belong to other assemblages (see e.g. Thomas, 2008). For instance, there is currently no DNA sequence data available for the New Zealand scaup, *Aythya novaeseelandiae*. This species is the only member of the *Aythya* genus in the assemblage, so we can use DNA sequence data from any other member of this genus to estimate the placement of *Aythya* in our assemblage phylogeny. We used sequences from the Greater Scaup (*Aythya marila*) to infer the placement of *A. novaeseelandiae* by assigning the TaxonID for *A. marila* to stand in as a substitute for *A. novaeseelandiae* (step 2, Fig. 1). Using substitute TaxonIDs involves making assumptions about the monophyly of taxa (e.g. in this example we make the assumption that the *Aythya* genus is monophyletic), so we only used this approach when published molecular or morphological studies provided robust support for the monophyly of the taxa of interest, and where these studies were not controversial. 33 of the 275 taxa in our list could be represented by substitute taxa (see Supplementary information).

We used information on taxonomy to place the remaining 79 species, that had no appropriate data on GenBank, in the phylogeny. To do this, we used constraints to place each species without DNA data into the clade comprising the lowest taxonomic rank to which it could be unambiguously assigned. For example, there are no DNA sequences available for the North Island adzebill (*Aptornis otidiformis*), but there is DNA sequence data for the South Island adzebill (*Aptornis defossor*). To place the North Island adzebill in the assemblage phylogeny, we made the assumption that the two species of *Aptornis* are more closely related to each other than either is to any other species in the assemblage, by defining a constraint that forces both *Aptornis* species to be monophyletic in our phylogeny. A full list of constraints is included in the Supplementary information.

We assembled a DNA supermatrix for all 196 of the taxa for which we had representative DNA sequence data. To do this, we used GeneFinder to search GenBank for homologous sequences that were available for at least 10 species in our dataset. We aligned the sequences for each gene by eye, then concatenated the gene sequences using Geneious version 5.0.3 (Drummond et al., 2011). We removed sites in the alignments for which homology could not be confidently inferred. This resulted in a DNA supermatrix of 11 mitochondrial genes (ND3, ND4, ND5, ND6, ATP6+8, CytB, COI, COII, COIII, 12S, 16S) and 3 nuclear genes (RAG-1, c-mos, and HBA2) with a total of 16,436 base pairs representing 196 species. We then added the 79 species without sequences to the alignment as empty DNA sequences (i.e. all gaps). Accession numbers, alignments, TaxonIDs, and information on substitute taxa and empty sequences are included as Supplementary information.

2.3. Definition of constraints

We included constraints in the phylogenetic analysis to provide *a priori* information on relationships of some species. Constraints allow us to place the species for which no DNA sequence data is available (see above), and also allow us to make maximum use of well-established phylogenetic groupings of birds. For example, there is little doubt that the passerines (order Passeriformes) form a monophyletic group of species with respect to other birds, so we constrained all 57 NZ passerines in our dataset to be monophyletic. We examined the literature on avian phylogenetic relationships, and defined a total of 53 constraints (19% of the total number of nodes in the phylogenetic tree) that were well supported across the literature, including constraining 12 of the 19 orders in the assemblage to be monophyletic. A full list of constraints is included as Supplementary information, and all constrained nodes are indicated on Fig. 2.

2.4. Phylogenetic analysis

We analysed our data in Bayesian phylogenetic framework. This framework allows the flexible specification of prior expectations on the placement of species, for example in the form of topological constraints and expectations of the distribution of branch lengths. These priors are then combined with the information contained in the DNA sequences to estimate a posterior distribution of phylogenetic trees. This framework is particularly appropriate here, as it allows the position of species with little or no DNA sequence data to be estimated, while formally incorporating the uncertainty of those estimates.

We used MrBayes version 3.1.2 (Ronquist and Huelsenbeck, 2003) to analyse our DNA dataset. The publicly-released version of MrBayes version 3.1.2 limits users to defining a maximum of 30 constraints, so we edited the source code to increase this limit to 150. This code, and Macintosh executables are available from the authors. We analysed our dataset as a partitioned supermatrix, by dividing our data into six partitions: (i) RNA + Intron sequences; (ii) 1st and 2nd codon positions from nuclear protein coding genes; (iii) 3rd codon positions from nuclear protein coding genes; (iv) 1st codon positions from mitochondrial protein coding genes; (v) 2nd codon positions from mitochondrial protein coding genes; (vi) 3rd codon positions from mitochondrial protein coding genes. The best model of molecular evolution for each of the six partitions was chosen using the Akaike Information Criterion implemented in ModelGenerator (Keane et al., 2006).

Initial analyses using default priors in MrBayes failed to converge due to well-characterised problems with increasing branch lengths in the tree (Brown et al., 2010). These problems result in branch lengths that continue to increase without bounds as the analysis progresses. To counter this, we implemented the solution suggested by Brown et al. (2010), which involved altering the branch length prior from its default value (an exponential distribution with a mean value of 0.1) to a more appropriate value for our dataset (an exponential distribution with a mean value of 0.05, determined by preliminary ML analysis using RaxML: Stamatakis, 2006). We also increased the proposal rate for the Dirichlet distribution of the rate-multipler parameter from 1 to 10, and changed the parameter defining the size of proposed change in the distribution from 500 to 300. This alteration increases the rate at which the MCMC explores different ways in which rates may differ between data partitions. These modifications solved the problem of increasing branch lengths, and led to convergence of all MCMC

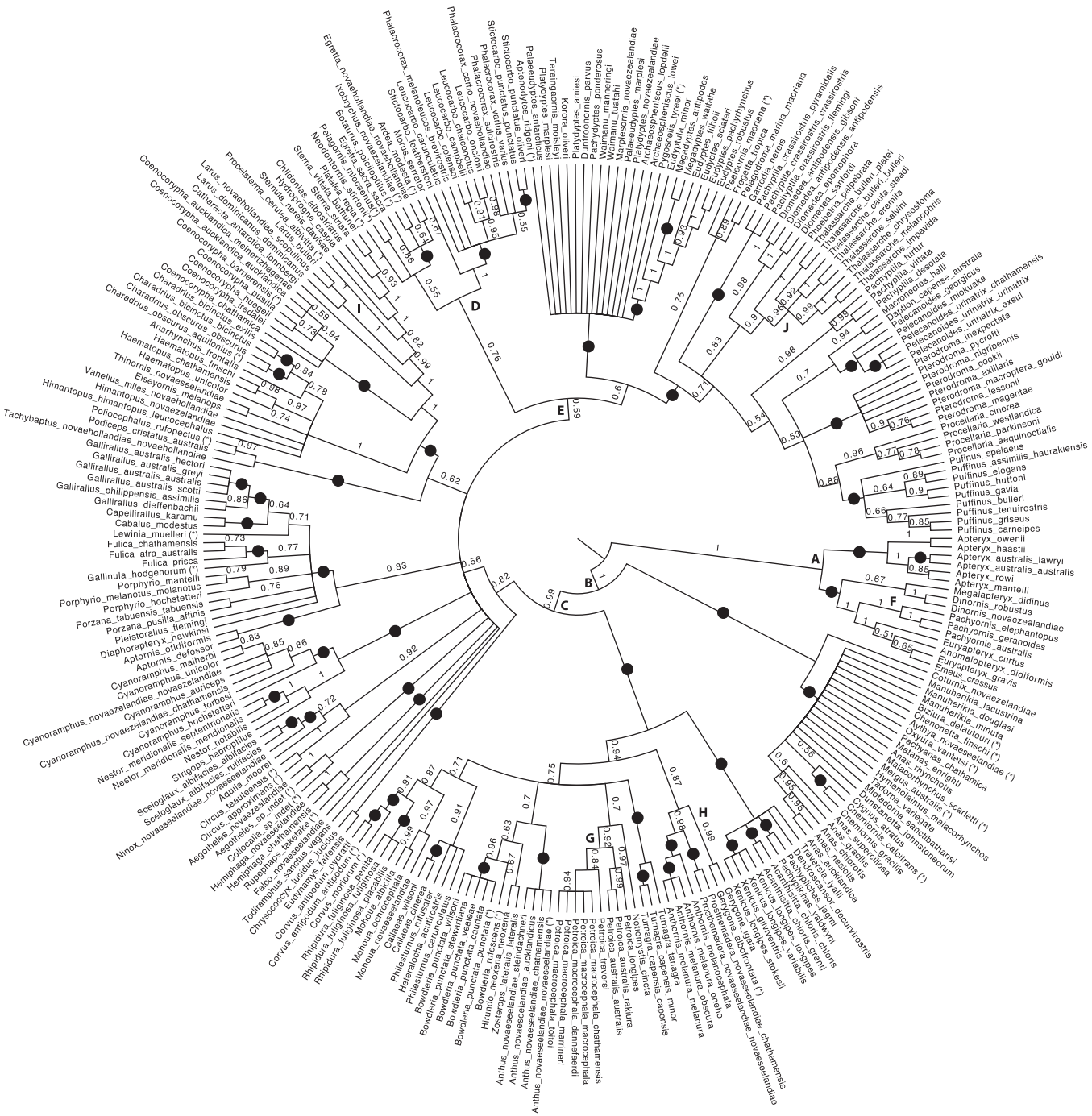

Fig. 2. A cladogram of all known extinct and extant native birds of New Zealand. This tree is a 50% majority rule tree that summarises for display purposes a Bayesian posterior distribution of 133,000 trees. Black circles on nodes represent constraints defined *a priori*. Numbers on branches show the Bayesian Posterior Probability (BPP) of a given grouping. Lettered nodes are discussed in the text. The full Bayesian posterior sample of trees, with branch lengths, is available from the authors.

runs on the same topologies, log likelihoods (ln L) and branch length distributions. The MrBayes input file is available as Supplementary information.

We analysed the data using seven independent runs of MrBayes. Each run consisted of 21,000,000 generations, with one hot and one cold chain, and samples taken from the cold chain every 1000 generations. All runs started with randomly-generated trees, and all runs had converged within 2,000,000 generations of burn-in. This resulted in a total of 133,000 samples of trees and parameters from the posterior distribution (19,000 samples per run, for 7 independent runs). All runs were checked for convergence and to determine that we had taken a large enough sample from the posterior distribution (Effective Sample Size >200) in TRACER v1.3 (Rambaut and Drummond, 2003). Topological convergence and stationarity for each run was assessed using AWTY (Nylander et al., 2008). To obtain a single tree for display purposes, the posterior distribution of trees was summarised by calculating the 50% majority rule tree from the full sample of 133,000, using the 'sumt' command in MrBayes 3.1.2.

3. Results and discussion

The 50% majority rule phylogeny of the NZ avifauna is shown in Fig. 2. This tree summarises the entire Bayesian posterior sample of

133,000 trees. Both the 50% majority rule tree and the full posterior sample of trees are available from the authors. We have placed the root of the phylogeny in Fig. 2 between Palaeognathae (represented in NZ by kiwis and moas) and the Neognathae (all other NZ birds), following the general consensus across morphological and molecular phylogenetic studies of birds (Garcia-Moreno and Mindell, 2000; Hackett et al., 2008; Mayr and Clarke, 2005; Sibley and Ahlquist, 1990; Slack et al., 2007).

To assess the accuracy of our assemblage phylogeny, we examined whether the well-supported clades in our tree are concordant with current understanding of avian systematics. Our analysis recovered extremely strong support for the monophyly of many widely-accepted major groupings of birds, including Palaeognathae (Bayesian Posterior Probability (BPP) = 1.0; Fig. 2 node A), Neognathae (BPP = 1.0; Fig. 2 node B), and Neoaves (Neognathae excluding Anseriformes and Galliformes (land- and water-fowl); BPP = 0.99; Fig. 2 node C). In concordance with other recent studies (Ericson et al., 2006; Hackett et al., 2008), we found that NZ birds traditionally assigned to Ciconiiformes (storks and allies) and Pelecaniformes (totipalmate birds) are not reciprocally monophyletic, but do form a single clade (BPP = 0.76, Fig. 2 node D). Our analysis also recovered some support, albeit weak, for a single clade of water birds (a clade comprising the NZ representatives of the Ciconiiformes, Pelecaniformes, Procellariiformes (tube-nosed seabirds) and Sphenisciformes (penguins); BPP = 0.59, Fig. 2 node E), which has been described in other recent studies (Ericson et al., 2006; Hackett et al., 2008; Van Tuinen et al., 2001). At a shallower phylogenetic depth, many of the strongly supported clades in our analysis described monophyletic genera of NZ native birds. For example we recover strong support for *Pachyornis* (an extinct genus of Moas, BPP = 1.0, Fig. 2 node F); *Petroica* (Australasian robins, BPP = 0.92, Fig. 2 node G); *Gerygone* (peep-warblers, BPP = 1.0, Fig. 2 node H); *Larus* (gulls, BPP = 1.0, Fig. 2 node I); and *Thalassarche* (mollymawks, BPP = 0.96, Fig. 2 node J), among others. That the well-supported nodes in our tree are consistent with the expected relationships at all phylogenetic depths indicates that the approach we describe here for building assemblage phylogenies is able to produce accurate phylogenetic trees.

In addition to producing accurate assemblage phylogenies, the approach we describe here formally accounts for the precision of the phylogenetic estimate. Many studies that use assemblage phylogenies effectively ignore phylogenetic uncertainty by assuming that a single tree is a perfectly precise representation of the relationships among taxa. This approach is convenient, but may give misleading results if the phylogeny is inaccurate, as will often be the case when there data are limited (Lanfear and Bromham, 2008). Assemblage phylogenies will vary in uncertainty, as the amount and quality of available data are likely to vary between taxa. For instance, the precision of the assemblage phylogeny presented here is improved by the large amount of previous taxonomic, phylogenetic, and molecular work on that provides information on the interrelationships among NZ birds (Trewick and Gibb, 2010). A Bayesian posterior sample of trees, as produced using the approach described here, provides a convenient way to account for phylogenetic uncertainty. In a Bayesian posterior sample of trees, all phylogenetic groupings are represented proportionally to their posterior probability. The trees in this sample can be used to account for phylogenetic uncertainty when testing a given hypothesis by calculating the test statistic on each tree. The result is a Bayesian posterior distribution of the test statistic, which can be used to determine the level of support for a given hypothesis while accounting for uncertainty in the underlying phylogeny.

The approach we describe here and the software we have produced can potentially be used to efficiently estimate assemblage phylogenies for any species assemblage. These phylogenetic estimates combine DNA sequence data with *a priori* knowledge on the relationships of taxa, and produce accurate phylogenies that also incorporate phylogenetic uncertainty. We hope that approaches such as these will be useful for phylogenetic community ecology.

Acknowledgments

We thank Phill Cassey and John Ewen for providing the impetus for this project, and thanks to both Phill Cassey and Gillian Gibb for valuable input throughout.

Appendix A. Supplementary material

The GeneFinder python script, list of NZ birds, and MrBayes input files associated with this article can be found, in the online version, at doi:10.1016/j.ympev.2011.07.018. The full Bayesian posterior sample of trees is available from the authors.

References

Alexandrou, M.A. et al., 2011. Competition and phylogeny determine community structure in Mullerian co-mimics. Nature 469, 84–88.
Bininda-Emonds, O.R., 2004. The evolution of supertrees. Trends Ecol. Evol. 19, 315–322.
Brown, J.M. et al., 2010. When trees grow too long: investigating the causes of highly inaccurate bayesian branch-length estimates. Syst. Biol. 59, 145–161.
Bunce, M. et al., 2003. Extreme reversed sexual size dimorphism in the extinct New Zealand moa *Dinornis*. Nature 425, 172–175.
Diniz-Filho, J.A.F., Bini, L.M., 2008. Macroecology, global change and the shadow of forgotten ancestors. Global Ecol. Biogeogr. 17, 11–17.
Drummond, A.J. et al., 2011. Geneious v5.4. <http://www.geneious.com/>.
Emerson, B.C., Gillespie, R.G., 2008. Phylogenetic analysis of community assembly and structure over space and time. Trends Ecol. Evol. 23, 619–630.
Ericson, P.G.P. et al., 2006. Diversification of Neoaves: integration of molecular sequence data and fossils. Biol. Lett. 2, 543–547.
Garcia-Moreno, J., Mindell, D.P., 2000. Rooting a phylogeny with homologous genes on opposite sex chromosomes (gametologs): a case study using avian CHD. Mol. Biol. Evol. 17, 1826–1832.
Hackett, S.J. et al., 2008. A phylogenomic study of birds reveals their evolutionary history. Science 320, 1763–1768.
Keane, T.M. et al., 2006. Assessment of methods for amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified. BMC Evol. Biol. 6, 29.
Kembel, S.W., Hubbell, S.P., 2006. The phylogenetic structure of a neotropical forest tree community. Ecology 87, 86–99.
Kembel, S.W. et al., 2010. Picante: R tools for integrating phylogenies and ecology. Bioinformatics 26, 1463–1464.
Kress, W.J. et al., 2009. Plant DNA barcodes and a community phylogeny of a tropical forest dynamics plot in Panama. Proc. Natl. Acad. Sci. USA 106, 18621–18626.
Lanfear, R., Bromham, L., 2008. Statistical tests between competing hypotheses of hox cluster evolution. Syst. Biol. 57, 708–718.
Mayr, G., Clarke, J., 2005. The deep divergences of neornithine birds: a phylogenetic analysis of morphological characters. Cladistics 21, 527–553.
Nylander, J.A. et al., 2008. AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. Bioinformatics 24, 581–583.
OSNZ Checklist Committee, 2010. Checklist of the Birds of New Zealand, Norfolk and Macquarie Islands, and the Ross Dependency, Antarctica, fourth ed. Ornithological Society of New Zealand & Te Papa Press, Wellington.
Rabosky, D.L., 2009. Ecological limits and diversification rate: alternative paradigms to explain the variation in species richness among clades and regions. Ecol. Lett. 12, 735–743.
Rambaut, A., Drummond, A.J., 2003. Tracer version 1.3. Evolutionary Biology Group, Univ. of Oxford, Oxford.
Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.
Sibley, C.G., Ahlquist, J.E., 1990. Phylogeny and Classification of Birds. Yale University Press, New Haven, CT.
Slack, K.E. et al., 2007. Resolving the root of the avian mitogenomic tree by breaking up long branches. Mol. Phylogen. Evol. 42, 1–13.
Smith, S.A. et al., 2009. Mega-phylogeny approach for comparative biology: an alternative to supertree and supermatrix approaches. BMC Evol. Biol. 9, 37.
Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22, 2688–2690.
Thomas, G.H., 2008. Phylogenetic distributions of British birds of conservation concern. Proc. Biol. Sci. 275, 2077–2083.
Thomson, R.C., Shaffer, H.B., 2010. Sparse supermatrices for phylogenetic inference: taxonomy, alignment, rogue taxa, and the phylogeny of living turtles. Syst. Biol. 59, 42–58.

Trewick, S.A., Gibb, G.C., 2010. Vicars, tramps and assembly of the New Zealand avifauna: a review of molecular phylogenetic evidence. Ibis 152, 226–253.
Vamosi, J.C., Vamosi, S.M., 2007. Body size, rarity, and phylogenetic community structure: insights from diving beetle assemblages of Alberta. Divers. Distrib. 13, 1–10.
Van Tuinen, M. et al., 2001. Convergence and divergence in the evolution of aquatic birds. Proc. Roy. Soc. London Ser. B – Biol. Sci. 268, 1345–1350.
Webb, C.O., Donoghue, M.J., 2005. Phylomatic: tree assembly for applied phylogenetics. Mol. Ecol. Notes 5, 181–183.