New Zealand Passerines Help Clarify the Diversification of Major Songbird Lineages during the Oligocene

Gillian C. Gibb1,*,†, Ryan England2,4,‡, Gerrit Hartig2,5, Patricia A. (Trish) McLenachan2, Briar L. Taylor Smith1, Bennet J. McComish2,6, Alan Cooper3, and David Penny2

1Ecology Group, Institute of Agriculture and Environment, Massey University, Palmerston North, New Zealand
2Institute of Fundamental Sciences, Massey University, Palmerston North, New Zealand
3Australian Centre for Ancient DNA, School of Earth and Environmental Sciences, University of Adelaide, South Australia, Australia
4Present address: Forensic Business Group, Institute of Environmental Science and Research (ESR Ltd.), Mt Albert Science Centre, Auckland, New Zealand
5Present address: Starlims Germany GmbH An Abbott Company, Witten, Germany
6Present address: School of Physical Sciences, University of Tasmania, Hobart, Australia
*Corresponding author: E-mail: gillian.c.gibb@gmail.com.
†These authors contributed equally to this work.
Accepted: October 7, 2015
Data deposition: This project has been deposited at GenBank under the accession numbers KC545397-KC545409, KT894672.

Abstract

Passerines are the largest avian order, and the 6,000 species comprise more than half of all extant bird species. This successful radiation probably had its origin in the Australasian region, but dating this origin has been difficult due to a scarce fossil record and poor biogeographic assumptions. Many of New Zealand’s endemic passerines fall within the deeper branches of the passerine radiation, and a well resolved phylogeny for the modern New Zealand element in the deeper branches of the oscine lineage will help us understand both oscine and passerine biogeography. To this end we present complete mitochondrial genomes representing all families of New Zealand passerines in a phylogenetic framework of over 100 passerine species. Dating analyses of this robust phylogeny suggest Passeriformes originated in the early Paleocene, with the major lineages of oscines “escaping” from Australasia about 30 Ma, and radiating throughout the world during the Oligocene. This independently derived conclusion is consistent with the passerine fossil record.

Key words: Passeriformes, mitochondrial genomes, oscine biogeography.

Introduction

Passerines (perching birds) are the most speciose avian order, and oscine songbirds (the major clade within passerines) account for almost half of all bird species (Clements 2007). The DNA-DNA hybridization studies of Sibley and Ahlquist (1990) found oscines formed two sister groups (suborders Passerida and Corvida) that appeared broadly European and Australasian in composition. Further studies have shown that Passerida is nested within the paraphyletic Corvida (Barker et al. 2002, 2004; Ericson et al. 2002). “Corvida” is now largely superseded by “core Corvoidea” and “basal oscines” (fig. 1). Sister to oscines, suboscines form two clades, broadly New World (NW) and Old World (OW) in composition (fig. 1). The species-rich New World clade originated in South America, and the relatively species-poor Old World lineages are mostly distributed through Africa, Madagascar, and Asia (Fjeldsa et al. 2003; Barker et al. 2004). Sister to all passerines is the small clade of New Zealand (NZ) wrens (Acanthisittidae). There is strong evidence that oscines arose in Australasia, with most deep lineages located in Australia, New Zealand, and Papua New Guinea (e.g., Jonsen et al. 2011). Because of Passeriformes incredibly speciose nature, both deep and shallow relationships within the passerines are still being resolved. The main hindrance to phylogenetic resolution within Passeriformes is the accessibility of DNA samples for some groups. Recent work suggests that as more groups within Passeriformes are sequenced, many nodes have high support (e.g., Aggerbeck et al. 2013).
Dating Passeriformes

Biogeographic hypotheses regarding the worldwide distribution of passerines often begin by assuming the isolation of the NZ wrens on the Zealandian landmass since it rifted from Australia approximately 82 Ma (Barker et al. 2004), because extant and recently extinct NZ wrens are either flightless or poor fliers (Heather and Robertson 2005). Vicariance and the breakup of Gondwana are then often posited as the major influence on the diversity of passerines found today (reviewed in Ericson et al. 2003), although dispersal must then be inferred for many more recent distributions. This assumption is clearly unfounded—there is no good evidence one way or the other that the NZ wren lineage originated in Zealandia, nor that they have always been poor fliers. Despite this, almost all dating of passerines has used this vicariant calibration (e.g., Ericson et al. 2002, 2014; Barker et al. 2004; Shepherd and Lambert 2007; Jönsson, Irestedt, et al. 2008; Jönsson et al. 2011; Kennedy et al. 2012; Moyle et al. 2012), even though this date (82 Ma) significantly predates many modern estimations of the radiation of the wider Neoaves group (Neoaves contains the majority of modern bird species). Both fossil and molecular evidence suggests Neoaves is more likely to have radiated in the late Cretaceous (Ho and Phillips 2009; Jetz et al. 2012; Gibb et al. 2013; Mayr 2013; Jarvis et al. 2014). Recent molecular evidence suggests it is unlikely passerines were one of the first lineages to diverge within Neoaves (Hackett et al. 2008; Suh et al. 2011; Kimball et al. 2013; Jarvis et al. 2014), and if they were then they would need to have an extended hidden fossil record (Mayr 2013).

Further geological evidence suggests the Zealandian continent was still closely associated with Australia at its northern end until more recently (~55 Ma, e.g., Schellart et al. 2006). Ericson et al. (2014) use this broader range in their analysis of passerine evolution, but this does not resolve the underlying circular argument. This geographic date variation (82–55 Ma) does not change our assertion that the NZ wrens are a poor, though tempting, choice for passerine evolution calibration. Because the NZ wrens are a single isolated lineage (both phylogenetically and geographically), it is very hard to infer anything about their distribution and composition through time (Trewick and Gibb 2010). Other calibrations must be found to date passerine evolution.

So what calibrations can be used? The sometimes-used molecular clock rate of 2% per site per million years is not recommended for use in passerines (Nabholz et al. 2009). Dating lineages using fossil calibrations from within the passerines is also a difficult challenge. There are no good deep fossils for calibration, because most passerine fossils are found in Europe, belong to crown groups, and are relatively recent (Mourer-Chauviré et al. 1989). The only deep passerines are early Eocene fossils from Australia (~54 Ma) reported by Boles (1995, 1997). Unfortunately, these Eocene fossils are likely to be stem passerines (Mayr 2009; Worthy et al. 2010) and cannot confidently be assigned to any modern lineage within passerines (such as oscines), so using this date for calibration is difficult. Nevertheless, these fossils support the biogeographic hypothesis of an Australasian origin of oscines, followed by a later radiation throughout the rest of the world. Another suggestion would be to date passerines directly using external avian fossils, but this is also difficult because passerines are a relatively fast evolving group (e.g., Nabholz et al. 2009). The best deep avian fossil calibrations (see Benton and Donoghue 2007; Parham et al. 2012; Ksepka and Clarke 2015) are in Galloanserae (Clarke et al. 2005) and penguins (Slack et al. 2006), which are both relatively slowly evolving lineages, and not closely related to passerines. Evolution during the early neoavian radiation has resulted in many short and therefore difficult-to-resolve internodes (e.g., Hackett et al. 2008; Pratt et al. 2009; Pacheco et al. 2011; Gibb et al. 2013; Kimball et al. 2013). The root can easily be drawn towards passerines, as they are likely to have the longest relative internal branch lengths (e.g., Brown et al. 2008; Pacheco et al. 2011). A data set with calibrations on a few slowly evolving outgroups can lead to incorrect tree topology, with passerines artifically drawn to the root of the tree through long branch attraction, and overestimation of passerine node ages, akin to what was seen with early studies of few taxa (e.g., Harlid and Arnason 1999). Recent studies from independent sources (nuclear and transposon, Hackett et al. 2008; Suh et al. 2011) using large numbers of taxa indicate passerines are part of the broad “higher land bird” radiation, and parrots are likely the closest extant relatives; however, these studies did not calculate a timeframe for passerine evolution. We have therefore chosen to calibrate passerine...
evolution using secondary calibrations derived from studies with broader taxon sampling using these deep fossils rather than biogeographic calibrations (Pratt et al. 2009; Schweizer et al. 2011; White et al. 2011).

Passerines have also been difficult to resolve morphologically, because convergence evolution in life history traits in this speciose group is common, and much more work is needed to define valid taxonomic constraints. Recent genetic studies have discovered convergence in a range of traits, such as the adaptation for nectivory of both New Zealand and Hawaiian honeyeaters (Ewen et al. 2006; Fleischer et al. 2008), evolution of toxicity in the paraphyletic Pitohui (Dumbacher et al. 2008; Jønsson, Bowie, et al. 2008), and behavior and plumage in birds of paradise (Cracraft and Feinstein 2000). In addition, passerines have a rather uniform skeletal morphology making the deeper nodes difficult to resolve as well (Mayr 2013). The large diversity within passerine families means molecular studies often find paraphyly and overturn previous taxonomic classifications (e.g., Moyle et al. 2012).

It is important to have agreement between independent lines of evidence such as mitochondrial, nuclear, and morphological data sets. Here, we provide a well-resolved mitochondrial phylogeny to complement existing molecular and morphological data sets. Mitochondrial genomes bridge the gap between exons (which perhaps evolve too slowly) and introns (which may not be alignable at deeper divergences). In addition it is relatively straightforward to sequence complete mitochondrial genomes from highly divergent avian species using conserved primers or high-throughput sequencing. A conservative alignment of protein and RNA genes provides over 13 kb of sequence after the removal of indels and noncoding sequence. Because the mitochondrion is a single molecule, all genes on it are considered to have the same gene order (Irestedt and Ohlson 2008). Improved taxon sampling has been shown to provide greater phylogenetic resolution (Slack et al. 2007), and we expect that to be the case here also.

Until recently, very few passerine mitochondrial genomes were completed; only ten were included by Pacheco et al. (2011). However, since 2011 there has been a dramatic increase in oscine mitochondrial genome generation. We report 14 new oscine mitochondrial genomes, including one extinct species and at least one representative from all extant New Zealand families. New Zealand has 23 extant native passerine species, and most are endemic (Gill et al. 2010). A number of families are endemic (e.g., Acanthisittidae, Callaeidae, Notiomystidae, Mohouidae) and many fall within the deeper branches of the passerine radiation. With our robust phylogeny including over 100 passerines, we increase understanding of the composition of the New Zealand avifauna, strengthen deeper oscine phylogenetic resolution, and greatly improve the timing of biogeographic hypotheses regarding the worldwide radiation of this most successful order of birds.

Materials and Methods

DNA Extraction, Polymerase Chain Reaction, and Sequencing of Modern Samples

The species sampled are given in table 1, and sample collection details are in supplementary table S1, Supplementary Material online. DNA was extracted from blood or tissue using Roche High Pure PCR template preparation kit (Roche Applied Science, Mannheim, Germany) following the manufacturers’ instructions. Two approaches were used to sequence complete mitochondrial genomes (supplementary tables S2 and S3, Supplementary Material online). The first approach was used for Rhipidura fuliginosa, Gerygone igata, Notiomystis cincta, Petroica australis, Philesturnus carunculatus, Turdus philomelos, Anthornis melanura, and Prosthemadera novaeseelandiae. The genomes were amplified in 2–3 overlapping segments using the Expand Long Template PCR System (Roche). These long-range polymerase chain reaction (PCR) products were used as templates for subsequent PCR of short overlapping fragments. This process is described in more detail in Gibb et al. (2007) and references therein. All short PCR products were amplified using primers identified from our database (available from the authors on request) employing standard protocols (e.g., Slack et al. 2007). Mitochondrial genomes were assembled using Sequencher 4.8 (Gene Codes Corp, Ann Arbour, MI) and electropherograms checked by eye.

For the second approach, short paired reads were generated from either Illumina HiSeq or MiSeq runs following standard protocols. Samples were either run in their own lane (Anthus novaeseelandiae and Petroica macroephalia), indexed (Mohoua novaeseelandiae and Zosterops lateralis) or as part of a mixture (Megalurus punctatus, mixed with a frog, fish, and mollusc). De novo assembly of contigs from the paired reads run individually or indexed used Velvet v1.1.06 (Zerbino and Birney 2008) with a K-mer sweep to identify the highest N50 for each species (typically around k = 65). De novo assembly of the mitochondrial genome from a mixture of total genomic extracts followed a modified version of the pipeline described in McComish et al. (2010). Because of the presence of nuclear DNA, Velvet assemblies, as expected, produced much larger numbers of contigs than the assembly described in McComish et al. (2010). This meant that the assembly graphs could not be used to help separate contigs from the different samples, because the graphs consisted largely of isolated short contigs of nuclear origin. However, it was possible to identify almost all of the mitochondrial contigs by simply altering the script used to separate contigs so that it called Exonerate (v. 2.2.0, Slater and Birney 2005) with a score threshold of 150. This was sufficient to filter out all of the
nuclear contigs, as these gave scores lower than 150 when aligned to the mitochondrial reference genomes. For all species, Geneious 6.1.7 (Biomatters Ltd, www.geneious.com, last accessed July 2015), was used to align the assembled contigs against reference genomes (supplementary tables S2 and S3, Supplementary Material online), typically using “map to reference” with medium-low sensitivity. Contig overlaps were identified and contigs were combined into preliminary whole mitochondrial genome assemblies. Finished assemblies were verified by mapping the original sequence reads to the assembly using Geneious 6.1.7. Any small remaining gaps were filled by sequencing shorter PCR products as described above for the first approach. DNA from *A. melanura* and *M. novaeseelandiae* was of insufficient quantity to complete these genomes, resulting in partial genomes for these two species.

**Extraction and Sequencing of Ancient DNA**

DNA extraction from a toepad of the extinct *Turnagra capensis* (piopio) was performed using strict ancient DNA procedures at the Australian Centre for Ancient DNA, University of Adelaide, South Australia. Total genomic DNA was extracted using proteinase K digestion followed by phenol chloroform extraction. Subsequently, library preparation and indexing was performed using the TrueSeq Nano DNA sample preparation kit (Illumina Inc.) in a dedicated ancient DNA facility at Massey University, New Zealand. Because ancient DNA is already degraded, the protocol was modified to omit the fragmentation step. Magnetic bead size selection criteria were relaxed to capture a wider range of DNA fragments. DNA was sequenced as part of one lane on an Illumina HiSeq machine running a 100 bp paired end protocol by New Zealand Genomics Ltd. The resulting reads were trimmed of any 3’-adapter sequence using cutadapt v1.1 (Martin 2011) and de novo assembled using ABySS v1.3.0 (Simpson et al. 2009) with a kmer sweep of 45, 55, and 65, and parameters *n* (number of pairs) = 2, *c* (coverage) = 5, and *e* (erosion) = 5. The resulting contigs were imported into Geneious 6.1.7 and mitochondrial contigs were filtered by assembly to another passerine mitochondrial genome (supplementary tables S2 and S3, Supplementary Material online). The finished assembly was checked by mapping the original sequence reads to the consensus using Geneious 6.1.7, and select short regions were independently verified by PCR. All genomes were checked to confirm protein regions translated and RNA stem loop structures matched appropriately.

In addition to the 14 new genomes reported here, a further 84 passerine genomes were included from GenBank (supplementary table S4, Supplementary Material online). Four genomes were also assembled from 454 transcriptomic data (from Nabholz et al. 2010), and rRNAs and some tRNAs were recoverable when the reads were mapped to reference genomes (see supplementary information and table S3, Supplementary Material online) using Geneious 6.1.7 as described above. Four diverse parrot species were also included as a more distant outgroup, as a number of recent studies place parrots as the closest extant outgroup to passerines, and use quite independent data sets to reach this conclusion (nuclear sequences [Hackett et al. 2008; Wang et al. 2012; Kimball et al. 2013], and retrotransposon data [Suh et al. 2011]). The details of all genomes used in these analyses can be found in supplementary table S4, Supplementary Material online.

Multiple genomes were manually aligned and checked in Geneious 6.1.7, at the amino acid level for protein-coding genes, and based on stem and loop secondary structure for RNA genes (Gutell et al. 1994; Springer and Douzery 1996), see supplementary tables S5 and S6, Supplementary Material online, for RNA stem annotations. Gaps, ambiguous sites
adjacent to gaps, the ND6 (light-strand encoded), noncoding regions, and stop codons (often incomplete in the DNA sequence), were excluded from the alignment. Conserved amino acids or stem columns were used to define the boundaries of ambiguous regions next to gaps. The third position of protein-coding genes have been coded as RY. RY recoding of third codons increases the proportion of observable changes on internal branches of the tree (treeness) and decreases the differences in nucleotide composition (relative compositional variability; e.g., Harrison et al. 2004; Gibb et al. 2007; Phillips et al. 2010). The full data set is 13,837 bp long and is separated into five partitions: three codon positions, and RNA stems and loops. The best partitioning and models were found using Bayesian information criterion model selection with a greedy search scheme and unlinked branch lengths in PartitionFinder (Lanfear et al. 2012), which recommended three final partitions: first codons plus loops, second codons plus stems, and third codons. These three partitions were used in all subsequent analyses. Alignments and trees can be downloaded from TreeBASE (TB2:15732).

Data were analyzed using maximum likelihood (ML) methods implemented in RAxML (Stamatakis et al. 2008), using a general time reversible model with gamma distribution (GTR + G). Bayesian analyses were performed using MrBayes (Huelsenbeck and Ronquist 2001) using GTR + I + G. Data sets were run for 10 million generations, and sampled every 2,500 generations after a burnin of 1,000,000 generations. Independent runs were checked for convergence, and trace files analyzed using Tracer (Rambaut and Drummond 2003) to ensure effective sample size (ESS) values greater than 200. Trees were viewed using Figtree v1.4 (http://tree.bio.ed.ac.uk/software/figtree/ [accessed June 2015]).

The data were also analyzed using the CAT–GTR mixture model of PhyloBayes-MPI 1.5a (Lartillot et al. 2013) with no RY coding, and constant sites removed. Two independent chains were run for 26,700 iterations and checked for convergence in likelihood and model parameters (tracecomp subprogram) and clade posterior probability (bpcomp subprogram). The first 10% of trees were discarded as burnin and a 50% majority rule Bayesian consensus tree with associated posterior probabilities was computed from the remaining trees using bpcomp.

**Results**

**Mitogenomic Features**

Mitochondrial gene rearrangements are very common within birds (see Gibb et al. 2007), and passerines are no exception. The majority of the genomes reported here have the standard gene order found in the chicken (Desjardins and Morais 1990). Two species (*Z. lateralis*, silvereye and *Me. punctatus*, fernbird) have the duplicate control region (CR) gene order (a copy of the CR is located between tRNA Thr and tRNA Pro, respectively). These calibrations are based on the results of Schweizer et al. (2011) and are very similar to other recent publications (White et al. 2011; Jetz et al. 2012). Schweizer et al. (2011) used the penguin fossil Waimanu as a calibration for the split between Sphenisciformes and other seabirds (Slack et al. 2007). This calibration used a mean of 66 Ma (SD 3.06) with normal distribution. The lower bound takes into account potential dating errors of the fossil, and the upper bound allows for putative members of Gaviiformes (loons) from the late Cretaceous. In addition they tested uniform and lognormal priors, and alternate sister groups for penguins (Procellariformes, or a group consisting of Gaviiformes, Procellariformes, Pelecaniformes, and Suliformes). They also used two additional fossil calibrations. First, a stem flamingo (Mayr 2005) as minimum age of the Podicipediformes–Phoenicopteriformes (grebe-flamingo) split with uniform prior and conservative upper bound (30–124 Ma). Second, a uniform distribution for the split between Galloanserae and Neoaves with the same 124 Ma upper bound and lower bound of 66 Ma based on the Vegavis fossil (Clarke et al. 2005). The results of Schweizer et al. (2011) were very similar with all calibration options.

Runs totaling 40,000,000 Markov chain Monte Carlo generations ensured ESS values greater than 200 (as estimated in Tracer v 1.5 Rambaut and Drummond 2003). Chains were sampled every 5000th generation after removing a burnin of 10%. Trees were viewed using Figtree v1.4 (http://tree.bio.ed.ac.uk/software/figtree/ [accessed June 2015]).

To compare alternative calibration schemes, the dating analysis was repeated using slightly younger calibrations based on the results of Pratt et al (2009). Four calibration priors were used, all with a normal distribution. These were: root prior, mean 61 Ma, SD 5 Ma, all passerines excluding *Ac. chloris*, mean 57.3 Ma, SD 5 Ma, all oscines 32.6 Ma, SD 6 Ma, all suboscines 46.3 Ma, SD 5.3 Ma. All other parameters were unaltered.

**Mitochondrial gene** rearrangements are very common within birds (see Gibb et al. 2007), and passerines are no exception. The majority of the genomes reported here have the standard gene order found in the chicken (Desjardins and Morais 1990). Two species (*Z. lateralis*, silvereye and *Me. punctatus*, fernbird) have the duplicate control region (CR) gene order (a copy of the CR is located between tRNA Thr and tRNA Pro, respectively). These calibrations are based on the results of Schweizer et al. (2011) and are very similar to other recent publications (White et al. 2011; Jetz et al. 2012). Schweizer et al. (2011) used the penguin fossil Waimanu as a calibration for the split between Sphenisciformes and other seabirds (Slack et al. 2007). This calibration used a mean of 66 Ma (SD 3.06) with normal distribution. The lower bound takes into account potential dating errors of the fossil, and the upper bound allows for putative members of Gaviiformes (loons) from the late Cretaceous. In addition they tested uniform and lognormal priors, and alternate sister groups for penguins (Procellariformes, or a group consisting of Gaviiformes, Procellariformes, Pelecaniformes, and Suliformes). They also used two additional fossil calibrations. First, a stem flamingo (Mayr 2005) as minimum age of the Podicipediformes–Phoenicopteriformes (grebe-flamingo) split with uniform prior and conservative upper bound (30–124 Ma). Second, a uniform distribution for the split between Galloanserae and Neoaves with the same 124 Ma upper bound and lower bound of 66 Ma based on the Vegavis fossil (Clarke et al. 2005). The results of Schweizer et al. (2011) were very similar with all calibration options.

Runs totaling 40,000,000 Markov chain Monte Carlo generations ensured ESS values greater than 200 (as estimated in Tracer v 1.5 Rambaut and Drummond 2003). Chains were sampled every 5000th generation after removing a burnin of 10%. Trees were viewed using Figtree v1.4 (http://tree.bio.ed.ac.uk/software/figtree/ [accessed June 2015]).

To compare alternative calibration schemes, the dating analysis was repeated using slightly younger calibrations based on the results of Pratt et al (2009). Four calibration priors were used, all with a normal distribution. These were: root prior, mean 61 Ma, SD 5 Ma, all passerines excluding *Ac. chloris*, mean 57.3 Ma, SD 5 Ma, all oscines 32.6 Ma, SD 6 Ma, all suboscines 46.3 Ma, SD 5.3 Ma. All other parameters were unaltered.

**Results**

**Mitogenomic Features**

Mitochondrial gene rearrangements are very common within birds (see Gibb et al. 2007), and passerines are no exception. The majority of the genomes reported here have the standard gene order found in the chicken (Desjardins and Morais 1990). Two species (*Z. lateralis*, silvereye and *Me. punctatus*, fernbird) have the duplicate control region (CR) gene order (a copy of the CR is located between tRNA Thr and tRNA Pro, respectively). These calibrations are based on the results of Schweizer et al. (2011) and are very similar to other recent publications (White et al. 2011; Jetz et al. 2012). Schweizer et al. (2011) used the penguin fossil Waimanu as a calibration for the split between Sphenisciformes and other seabirds (Slack et al. 2007). This calibration used a mean of 66 Ma (SD 3.06) with normal distribution. The lower bound takes into account potential dating errors of the fossil, and the upper bound allows for putative members of Gaviiformes (loons) from the late Cretaceous. In addition they tested uniform and lognormal priors, and alternate sister groups for penguins (Procellariformes, or a group consisting of Gaviiformes, Procellariformes, Pelecaniformes, and Suliformes). They also used two additional fossil calibrations. First, a stem flamingo (Mayr 2005) as minimum age of the Podicipediformes–Phoenicopteriformes (grebe-flamingo) split with uniform prior and conservative upper bound (30–124 Ma). Second, a uniform distribution for the split between Galloanserae and Neoaves with the same 124 Ma upper bound and lower bound of 66 Ma based on the Vegavis fossil (Clarke et al. 2005). The results of Schweizer et al. (2011) were very similar with all calibration options.

Runs totaling 40,000,000 Markov chain Monte Carlo generations ensured ESS values greater than 200 (as estimated in Tracer v 1.5 Rambaut and Drummond 2003). Chains were sampled every 5000th generation after removing a burnin of 10%. Trees were viewed using Figtree v1.4 (http://tree.bio.ed.ac.uk/software/figtree/ [accessed June 2015]).

To compare alternative calibration schemes, the dating analysis was repeated using slightly younger calibrations based on the results of Pratt et al (2009). Four calibration priors were used, all with a normal distribution. These were: root prior, mean 61 Ma, SD 5 Ma, all passerines excluding *Ac. chloris*, mean 57.3 Ma, SD 5 Ma, all oscines 32.6 Ma, SD 6 Ma, all suboscines 46.3 Ma, SD 5.3 Ma. All other parameters were unaltered.
see figure 1 in Gibb et al. 2007) also found in many other Sylvioida. The two Petroica have the remnant CR 2 gene order as do other published Petroica (Cooke et al. 2012). The N. cincta (hihi) and T. philomelos (song thrush) genomes have independently evolved a novel gene order involving duplication of tRNA-Pro-CR. This is contra to a previous report showing a standard gene order for these two genomes (Barker 2014). This gene rearrangement differs from previously recorded avian mitochondrial duplications, as tRNA-Thr is not part of the duplicated segment (see fig. 1 in Gibb et al. 2007). As has been seen in other CR gene duplications, the duplicated segments are nearly identical (e.g., Gibb et al. 2007, 2013). It is surprising to note that all published genomes from Passeroidea and Muscicapoidea, except our thrush (T. philomelos), have only one CR. Given the regularity and repeated occurrence of gene rearrangements in passerines, it is likely researchers will soon uncover more species with duplicated regions in Passeroidea and Muscicapoidea. It is also very likely some of the genomes published by other groups actually have a hidden duplication not detected due to genome misassembly. This highlights the ongoing need for researchers to be aware of the possibility of hidden duplicate CRs when sequencing mitochondrial genomes (Gibb et al. 2007). Researchers need to actively check (as we do), using PCR or high throughput data, for “out of order” gene arrangements (such as the CR before tRNA-Thr or tRNA-Pro) and signs such as increased read coverage depth between cyt b and 12S.

The majority of taxa are grouped as expected based on previous studies and taxonomic classification. We were, however, surprised by the placement of Pseudopodoces humilis (Tibetan ground-jay). Further analysis of this genome against a wider range of shorter mitochondrial genes indicates this genome may have been mislabeled and is likely to be Phoenicurus ochruros (black redstart, supplementary fig. S1, Supplementary Material online).

The complete genome of Turn. capensis (piopio) shows significant differences to a number of other published Turn. capensis mitochondrial gene sequences, and we suspect these other published sequences are incorrect. Portions of some sequences are more divergent than would be expected between sequences from the same order, let alone the same species. To elaborate, cyt b (GenBank U51734) is 16% divergent, the second half of cyt b (JN614879) is 10% divergent and ND3 (JN614657) is 20% divergent. Compare this with the correct ND2 (JN614707) where the sequences are less than 1% divergent. This highlights the challenge of accurately verifying novel sequences amplified from short PCRs of degraded ancient DNA. Generated by high throughput sequencing, our Turnagra mitochondrial genome has an average coverage of 171× (SD of 36×) with no ambiguous regions, giving us high confidence in the accuracy of the full sequence (supplementary table S2, Supplementary Material online). We find Turnagra to fall within the core Corvoidea, diverging from other orioles about 22 Ma. This is contra to the earlier cyt b study of Christidis et al. (1996) which found Turnagra to be a basal bowerbird. Its position with orioles in the core Corvoidea agrees with our finding from a few short mtDNA regions (Gibb 2010) and recent nuclear studies (Johansson et al. 2011; Zuccon and Ericson 2012).

Phylogenetic Analyses

All ML and Bayesian analyses resulted in congruent phylogenies with the majority of branches having high bootstrap and Bayesian posterior probability support (fig. 2 and supplementary figs. S2–S5, Supplementary Material online). Those few branches that do differ between analyses have low support. As expected, passerines form three groups: oscines, suboscines, and the Ac. chloris (rifleman). Within oscines, the traditional Passerida forms three broad lineages: Passeroidea, Sylvioida, and Muscicapoidea. We find Sylvioida to diverge first, in agreement with some nuclear and mitochondrial studies (e.g., Barker et al. 2004; Treplin et al. 2008; Barker 2014), although alternative arrangements are also seen (e.g., Hackett et al. 2008; Johansson et al. 2008; Price et al. 2014). Support for the arrangement is still weak and the internodes are short, however, it is “locally stable” (Cooper and Penny 1997) in that the alternate arrangements are only one step away. The deepest split within Muscicapoidea has sometimes defined the separation of Muscicapoidea from Certhioidea (labeled in fig. 2, though it is a similar depth to the radiation of Sylvioida). The core corvoids are monophyletic, and the majority of the New Zealand lineages fall within the basal oscine, core Corvoidea, and basal Passerida lineages. Menura novae-hollandiae (lyrebird) is sister to all other oscines.

Notiomystis cincta and Ph. carunculatus (saddleback) are deeply diverging sister taxa, as predicted from nuclear studies (Ewen et al. 2006; Driskell et al. 2007). Interestingly, they are sister to the rest of Passerida, near the Australasian robins (here, Petroica + Eopsaltria, fig. 2). This is contra to the results of studies including the Rag-1 gene (e.g., Barker et al. 2004) where Callaeidae (represented in our study by Ph. carunculatus) is sister to the core Corvoidea. That Callaeidae is sister to Passerida was also found by Irestedt and Ohlson (2008), who analyzed short introns and exons from a combination of Callaeas cinerea (kokako) and Ph. carunculatus sequences (both Callaeidae). The different results seen with some intron, exon, and mitochondrial data highlights the danger of simply combining different data sets without first checking for congruence between data sets. In the case of the Callaeidae, it appears Rag-1 overpowered the signal for an alternative topology now found independently by intron, exon, and mitochondrial data (fig. 1 and Irestedt and Ohlson 2008). This is akin to the case of β-fibrinogen intron 7 (Morgan-Richards et al. 2008) that appeared to provide signal supporting a Metaves/Coronaves split (Fain and Houde 2004) even in the presence of additional genes with alternate signal (Ericson et al. 2006; Kimball et al. 2013).
FIG. 2.—Beast chronogram of 102 passerine species derived from complete mitochondrial genomes. The 14 New Zealand native species are in bold. Branch thickness indicates ML bootstrap support, and the timescale is in millions of years. The insert ML tree shows relative branch length, and a full size version is shown in supplementary figure S2, Supplementary Material online. Colors are the same as in figure 1 for reference. 

1Pseudopodoces humilis is likely Phoenicurus ochruros, see Results section and supplementary figure S1, Supplementary Material online.
However, given the agreement here from intron, exon, and mitochondrial data, we suggest these “basal passerids” should be included within the Passerida; this would simplify the formal classification of oscines.

**Dating and Diversification**

The two dating analyses resulted in trees with the same topology, as expected (fig. 2; supplementary figs. S6 and S7, Supplementary Material online). The calibrations based on Schweizer et al. (2011) were slightly older than those of Pratt et al. (2009), and unsurprisingly, the resultant tree also has a deeper root node. While there was approximately 10 Ma difference in root calibration ages, the resultant trees only differ by 5 Ma at the deepest nodes (table 2), and differences between node ages in the two trees decrease as the nodes get younger. All node dates compared between the two trees fall within the 95% posterior probability intervals of the two trees, and we find these results very comparable. The broader, more conservative date ranges based on Schweizer et al. (2011) (fig. 2; supplementary fig. S6, Supplementary Material online) are referenced in the following paragraphs.

We find the NZ wrens diverge from all other passerines around 62 Ma (with a range of 48.5–76.3 Ma, fig. 3), with oscines and suboscines diverging only a (geologically) short time later (fig. 2; supplementary fig. S3, Supplementary Material online). The major “backbone” branches of the oscines diverged from each other between 30 and 44 Ma. The deepest of these branches are all Australian-centric in origin, and it is around 30 Ma that the three speciose Passerida clades diverge. The subsequent extant radiation of all major oscine lineages, both Australasian corvids and worldwide passerids, appears to have happened over a similar period of time (22–28 Ma), with the Australasian elements radiating slightly earlier than the Passerida lineages (figs. 2 and 3).

We do not have enough information to comment on the timing of the radiation within the suboscines, with only one Old World suboscine representative. The extant New World suboscines appear to have begun radiating earlier than the oscines, as more than 30 Ma separates the two sampled NW suboscines, and the Furnarii (ovenbirds, antbirds and allies) probably diverged earlier (e.g., Barker et al. 2004). Greater sampling within suboscines will be needed to explore this further.

**Discussion**

It does appear that the addition of New Zealand lineages has helped clarify oscine radiation timing, for example, the deep split of *G. igata* in Meliphagoidea (28 Ma) and the separation of *N. cincta* and *P. carunculatus* around 30 Ma. Our finding of approximately 15 Ma for the split of *Epithanura* and *Prosthemadera* is a similar timeframe to Joseph et al. (2014) who used very different calibration methods for honeyeaters (a modified molecular rate estimation for ND2). Similar ages from different methods help increase our confidence in our findings. The extant core Corvoidea also began radiating around 28 Ma. BEAST and PhyloBayes analyses show *Turnagra* plus *Oriolus* as the deepest branch (fig. 2; supplementary figs. S4, S6, and S7, Supplementary Material online); however, RAxML and MrBayes analyses slightly favored *M. novaeseelandiae* (brown creeper) as the deepest lineage in the core Corvoidea (supplementary figs. S2 and S3, Supplementary Material online), in agreement with other recent work (Aggerbeck et al. 2013; Aidala et al. 2013). ML support for the two configurations is split around 50% for each topology (fig. 2 and supplementary fig. S2, Supplementary Material online), but they are locally stable.

Within the basal oscines there are a number of Australasian lineages that will be needed to further resolve our understanding of the basal oscine radiation, including catbirds, logrunners, babbler, berrycreepers and treecreepers, and others (work is currently underway on this project). These taxa will indicate whether more Australian lineages were diversifying before 30 Ma, or if all modern diversification of oscine lineages predominantly occurred 20–30 Ma.

**Dating the Modern New Zealand Avifauna**

Many of the New Zealand endemic lineages are quite deeply diverged and species poor (*Acanthisitta, Turnagra, Mohoua, Notiomystis, Philesturnus*). What we do know about the modern diversity within lineages is that it is very shallow—for example, the NZ Meliphagidae, *Pr. novaeseelandiae* (tui) and *A. melanura* (bellbird), diverged about 3.8 Ma, and *P. australis* and *P. macrocephala* (NZ robin and tomtit) are separate invasions of New Zealand approximately 5–6 Ma (figs. 2 and 3). It is tempting to infer the deeper lineages, (e.g., *Mohoua*) have been isolated in New Zealand as they diverged; however, the fossil record is able to tell us little about the geographic location of these lineages through time. From our phylogeny, *Notiomystis* and *Philesturnus* are the only hint of a deep split within Zealandia. Most, if not all, of Zealandia was submerged during the late Oligocene (22–25 Ma, Landis et al. 2008), which certainly caused great turnover in the New Zealand fauna. Many factors are involved in assembling a biota over time (Trewick and Gibb 2010; Bromham et al. 2012). The early Miocene fossil deposits at St Bathans (fig. 3, 16–19 Ma) show a great diversity of avian species (e.g., *Worthy et al. 2009, 2007, 2011*), some of which are still present today. These fossils include passerines from Cisticolidae (currawongs and Australian magpies, not currently native NZ taxa, *Worthy et al. 2007*) and Acanthisittidae (*Worthy et al. 2010*), indicating at least the NZ wrens were present by this time. Further undescribed passerine fossils from St Bathans will improve our understanding of lineage assembly over time (Trevor Worthy, personal communication).

We do not expect this study to be the final word on the timing of passerine evolution. As our understanding of deep
<table>
<thead>
<tr>
<th>Reference</th>
<th>Passerine Split Date</th>
<th>95% Range</th>
<th>Suboscines and Oscines Split Date</th>
<th>95% Range</th>
<th>Includes Waimanu and Vegavis Fossils in Calibrations?</th>
<th>Includes NZ Wrens as Biogeographic Calibration?</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pratt et al. (2009)</td>
<td>61.05</td>
<td>50–70</td>
<td>57.26</td>
<td>48–67</td>
<td>Y</td>
<td>N</td>
<td>Focused on broad range of birds including passerines</td>
</tr>
<tr>
<td>Schweizer et al. (2011)</td>
<td>70.36</td>
<td>52.6–88.13</td>
<td>62.4</td>
<td>45.5–79.7</td>
<td>Y</td>
<td>N</td>
<td>Focused on parrots but uses broad range of birds for dating including passerines.</td>
</tr>
<tr>
<td>White et al. (2011)</td>
<td>68.03</td>
<td>55–82</td>
<td>63.95</td>
<td>51–78</td>
<td>Y</td>
<td>N</td>
<td>Focused on parrots but uses broad range of birds for dating including passerines.</td>
</tr>
<tr>
<td>Pacheco et al. (2011)</td>
<td>78.5</td>
<td>71–85</td>
<td>74.9</td>
<td>—</td>
<td>Y</td>
<td>Y</td>
<td>Includes broad range of birds. Dating includes NZ wrens calibration</td>
</tr>
<tr>
<td>Jarvis et al. (2014)</td>
<td>39a</td>
<td>32–43</td>
<td>32*</td>
<td>27–36</td>
<td>Y</td>
<td>N</td>
<td>Primarily uses conservative (i.e. young) minimum bounds with no maximum. The surprisingly young passerine split requires further investigation.</td>
</tr>
<tr>
<td>Kennedy et al. (2012)</td>
<td>44*</td>
<td>—</td>
<td>42*</td>
<td>—</td>
<td>NA</td>
<td>Y</td>
<td>Root calibration has uniform prior of 85–0 Ma. All other calibrations are young (&lt; 20 Ma)</td>
</tr>
<tr>
<td>Ericson et al. (2014)</td>
<td>74*</td>
<td>59–85</td>
<td>73*</td>
<td>58–84</td>
<td>NA</td>
<td>Y</td>
<td>Uses uniform root prior of 85–52 Ma.</td>
</tr>
<tr>
<td>Price et al. (2014)</td>
<td>—</td>
<td>—</td>
<td>51</td>
<td>—</td>
<td>Waimanu</td>
<td>Y</td>
<td>Many calibrations have broad uniform distribution with 80 Ma lower bound</td>
</tr>
<tr>
<td>This study</td>
<td>62.5</td>
<td>48.5–76.3</td>
<td>59.7</td>
<td>40.5–73.2</td>
<td>Y (indirectly)</td>
<td>N</td>
<td>Calibrations based on Schweizer et al. (2011)</td>
</tr>
<tr>
<td>This study</td>
<td>57.61</td>
<td>51.0–64.4</td>
<td>54.75</td>
<td>48.9–60.8</td>
<td>Y (indirectly)</td>
<td>N</td>
<td>Calibrations based on Pratt et al. (2009)</td>
</tr>
</tbody>
</table>

*Approximate values inferred from figures.
avian evolution improves, so will our calibrations for the timing of passerine evolution. The basic structure of the passerine tree appears to be consistent across many studies, and as different calibrations are applied to the tree, the branch lengths will shrink or expand proportionately. The recent work of Jarvis et al. (2014) used 1,156 clocklike exons to examine avian evolution, and found a much more recent timing for the radiation of Passerines (~39 Ma, table 2). Jarvis et al. (2014) found oscines and suboscines diverged approximately 32 Ma, with the core Corvoidea and Passeriodea diverging approximately 20 Ma. If this is indeed the case, then late Oligocene fossil oscines from Europe (fig. 3 and Mayr 2009) would likely fall outside the modern oscine radiation. These surprisingly young divergence times for passerines will require further investigation.

In our study, two sets of secondary calibrations separated by approximately 10 Ma have resulted in trees approximately 5 Ma divergent at the root. The root divergence of around 57–62 Ma (with 95% posterior probabilities around 48–76 Ma) is much younger than the divergence of Zealandia and Australia around 82 Ma. It is certainly possible the ancestors of the modern NZ wrens were present on Zealandia around 55 Ma when the landmasses’ separation completed (Schellart et al. 2006). However, there is no physical evidence to indicate their location between the 57 and 62 Ma divergence inferred here and the Acanthisittidae fossils found in New Zealand during the Miocene 16–19 Ma (Worthy et al. 2010). Therefore we reiterate the use of NZ wrens as a vicariant biogeographic calibration should stop.
Many studies of passerines have been calibrated using the rifting of New Zealand and Australia around 82 Ma for the separation of Acanthisittidae from other passerines (table 2). We find no evidence to support this suggestion. Our study, which calibrates passerine evolution using secondary calibrations of the basal nodes from previous more widely sampled studies that used well-defined fossil calibrations, avoids this problem and allows us to more realistically begin to interpret passerine evolution and biogeography. We find crown group Passeriformes began diverging in the early Paleocene, with major expansion of the speciose oscine lineages during the Oligocene.

Our timescale for passerine evolution is concordant with the scant fossil record available for deep oscines (fig. 3 and Mayr 2013). The oldest “passerine” fossils are found in the early Eocene of Australia, 45–55 Ma (Boles 1997). From our results these fossils are older than the modern oscine radiation, which would be concordant with the stem passerine status suggested for them (Mayr 2009; Worthy et al. 2010). Early Oligocene fossils from Europe are not crown oscines and could be subsocines or near passerines, but oscine passerines are known from the late Oligocene in Europe (fig. 3 and Mayr 2009). This is in agreement with our timescale of a radiation of the modern passerid lineages beginning in the late Oligocene (22–28 Ma). The timing of this “escape” of passerines from the Australasian region into the rest of the world is coincident with the collision of the Australian and South East Asian plates (22–28 Ma). The timing of this “escape” of passerines from the Australasian region into the rest of the world is coincident with the collision of the Australian and South East Asian plates and the emergence of Wallacea 20–30 Ma (fig. 3, Hall 2009). How much influence this increased SE Asian land area had on lineages that were already volant is still to be determined. Much denser sampling of these incredibly speciose lineages will be needed to fully understand the worldwide radiation of Passerida.

Supplementary Material

Supplementary figures S1–S7 and tables S1–S6 are available at Genome Biology and Evolution online (http://www.gbe.oxfordjournals.org).

Acknowledgments

The authors thank many people for providing tissue and bone samples for this study including John Ewen and Phil Cassey (samples from Tiritiri Matangi), Dick Gill (Waikanae Department of Conservation), Sara Shepherd and Gillian Stone (Museum of New Zealand Te Papa Tongarewa), and Sandra Burles (Wairarapa DoC). Thanks to Philippa Horton and Trevor Worthy for access to the piopio specimen at the South Australian Museum, and to Trevor Worthy, Dorothee Huchon and anonymous referees for helpful comments on the manuscript. This work was supported by the Allan Wilson Centre for Molecular Ecology and Evolution, the New Zealand Marsden Fund, and Massey University.

Literature Cited


Associate editor: Dorothee Huchon

**Associate editor:** Dorothee Huchon

---