Supporting Information

Section 1: Genotype-by-sequencing

Whole genomic DNA was quantified with picogreen. Two single-end genotype-by-sequencing libraries (*see published protocol in* 1) were constructed using the EcoT221 restriction enzyme at the Cornell Institute of Genomic Diversity, each with 95 samples uniquely barcoded and multiplexed (in addition to the 95 individuals analyzed in the present study, these libraries included kiwi samples that failed during sequencing as well as a number of samples unrelated to the present study). One library was sequenced on a single, and the other on two Illumina HiSeq 2000 lanes (the latter because sufficient coverage was not obtained from a single lane) using 100 base pair, single end sequencing. We obtained 290,409,851 reads for the first library, and 257,188,302 reads for the two runs of the second library. Raw sequences were trimmed to 90 base pairs, quality filtered, and demultiplexed using the process_radtags script (-c, -q, -r options invoked, all other options set to default), of the Stacks 1.28 pipeline (2). Any portion of the 3' adaptors that occurred within the 90bp fragment were filtered using the fastx_clipper from the FASTX-Toolkit v 0.0.14 (4). Fragments were aligned to the kiwi (*A. mantelli*) reference genome (3) using default settings in Bowtie2-2.2.6 (4), and genotypes were called using the ref_map option of the Stacks v1.35 pipeline (2). All settings were kept at their default values except: 1) the minimum depth of coverage to report a stack was set to 3 and 2) the bounded-error SNP calling model was used which estimates the sequencing error rate at each nucleotide position, but does not allow the rate to exceed 0.05. Following others (e.g. 3), we retained loci that occurred > 10,000 bp from coding regions and that were > 50,000 bp from each other, or

25,000 bp from the ends of the kiwi reference genome contigs. Using custom scripts, we further filtered for a minimum depth of coverage of at least 10, a maximum SNP heterozygosity of 0.75 and locus coverage of at least 75%, and an individual coverage of at least 50% (which resulted in retention of 96 kiwi samples). This filtering scheme resulted in a baseline dataset used for all analyses involving all individuals (6332 loci with one or more variable sites). For analyses involving a subset of individuals, further filtering was applied (see below for each analysis).

Section 2: Lineage delimitation

We used several methods to determine if the 11 extant mtDNA lineages (see Fig. 1) also represented genetically differentiated lineages in the nuclear genome. First, we used principal coordinate analyses (PCoA) and analyses of population structure to assess the number of genetically distinct extant populations within *A. mantelli* and *A. australis*. For each variable locus, a single SNP was extracted (the first SNP) for a total of 6332 SNPs. For PCoA, genotypes were coded as 0 and 2 for homozygotes and 1 for heterozygotes and analyses were performed in the program PAST (6). We ran analyses of population structure using the program STRUCTURE 2.3 (7). Sixteen separate runs (each with different random starting parameters) were performed with an admixture model for 100,000 generations, following a 10,000 generation burn-in. The number of distinct ancestral populations (k) was set between 1 and 8. The Evanno method was used to assess the best value of k for each dataset using the program Structure Harvester (8). For *A. mantelli*, k = 4 was best supported. For *A. australis*, k = 3 was best supported. We calculated genome-wide pairwise Hudson's F_{st} (9) and JC69 genetic distances between each kiwi lineage (Table S1). F_{st} and its 95% CI (obtained using 500 bootstrap

replicates) were calculated using custom scripts for 6332 SNPs. Average JC69 genetic distances between and within lineages were calculated using the R package *Ape* (10) using the sequence dataset in Section 5 below .

We used Bayesian species delimitation (11) under a coalescent, population genetic framework to quantify the number of distinct lineages in *A. australis* and *A. mantelli* using nuclear genomic data. All analyses were performed using the program bpp 2.3 (11). Six high coverage individuals (12 gene copies) were chosen at random for each of the four populations of *A. mantelli* and *A. australis*, and loci 90bp in length with no missing data (complete coverage across individuals and loci) and with a minimum of 3 SNPs were extracted (*A. mantelli* = 131 loci and *A. australis* = 114 loci). Our filtering strategy of not including genes with missing data, while including at least 1 male, meant that female specific (e.g. haploid) W-linked loci were filtered from the dataset. An approach that simultaneously delimitated lineages and the species tree was fit (species delimitation algorithm 0, with fine-tune parameter set to 20). The model has two key population genetic parameters Θ and τ, with priors for each set with a gamma distribution with parameters α and β and mean value = α/β. $\Theta = 4N_e\mu$, where N_e is the effective population size and μ is the mutation rate. Θ can be approximated by the average pairwise divergence between individuals within a population. We therefore calculated JC69 genetic distances (the distance used by bpp) and estimated Θ to be 0.00544 and 0.00556 for *A. mantelli* and *A. australis* respectively. We set the mean (α/β) of the gamma prior to these values. We used a rather diffuse gamma prior with $\alpha = 2$ and set β accordingly. $\tau = \mu T$ where T is the time of population divergence. We approximated τ for the basal split within *A. mantelli* and *A. australis* to be the average genetic distance at the root minus ½ the average pairwise divergence

between individuals within extant populations. This gave a value of 0.00397 and 0.00859 respectively for *A. mantelli* and *A. australis*. We set the mean (α/β) of our gamma priors to these values and set $\alpha = 2$.

We used a conservative burnin of 300,000 generations, and then sampled every ten generations for the next 100,000 generations for a total of 10,000 samples. Bayesian species delimitation in bpp can get stuck, especially when the starting model is set to number of species = 1, and many loci and populations are involved (11). We therefore ran 20 independent runs with the number of starting species set from 1 to 4. For *A. australis*, all 20 runs appeared to mix well and converged on a model with four species regardless of the starting number of species (posterior probability for a model with four species = 1.0 across all runs). For *A. mantelli*, 19 of 20 runs converged on a model of four species (each with a posterior probably of 1.0), and one run failed to reach convergence in the burnin period and was discarded.

Section 3: Phylogenetic Analyses

A Bayesian phylogeny was generated in Mr. Bayes 3.2.6 (12) for 203 individuals using the entire mtDNA dataset (Figs. 1, S1; Database S1). The mtDNA dataset was partitioned according to each codon position for cytochrome *b* and a single partition created for the non-coding control region. Models were selected for each partition using the software MrModeltest v.2.3 (13). The GTR-y-I model was selected for each mtDNA partition. We performed two separate runs, for a

total of 10 million generations each. We used a burnin of 2 million generations and sampled post burnin samples every 1000 generations. A majority-rule consensus tree was generated.

We generated a species tree using SNAPP 1.0 (14) from nuclear SNPs found in two or more individuals. To reduce computation time, we used a random sample of 1000 SNPs. We included three individuals for each of the 11 extant lineages, and included only loci with no missing data across individuals. For each locus, homozygotes for the major allele (defined as the allele with a frequency > 0.5 across all samples used in the SNAPP analyses) were coded as 2, heterozygotes as 1 and homozygotes for the minor allele as 0. Forward and backward mutation rates were estimated by SNAPP from the data and all other priors were kept at default values. We ran five independent runs, each of 1.1 million generations, with the first 0.1 million generations discarded as the burnin. Post-burnin trees were sampled every 1000 generations and were combined across the five runs to make a maximum clade credibility tree in TreeAnnotator (15).

The species tree generated in SNAPP (which doesn't include extinct lineages) and the mtDNA gene tree (which does include extinct lineages) disagreed on the topology within *A. australis*. In order to better determine how fossil and extant lineages of *A. australis* relate we generated a phylogeny in MrBayes using mtDNA and nuclear DNA combined. Only a single individual per lineage was used, with a sample of *A. haastii* and *A. rowi* used as outgroups, and the analysis rooted with the former. mtDNA COI, mtDNA cyt B and the GBS generated nDNA (the entire sequence, not just the SNPs were used) were each assigned a different partition, and parameters of the GTR- y model were estimated separately for each partition. The nDNA was assigned a single partition because each GBS fragment (at most 90bp in length) included

only a small number of SNPs and are thus unlikely to exhibit rate variation that cannot be adequately handled by the γ correction for rate variation used in the partition. Run conditions and generation of a majority-rule consensus tree were the same as for the MrBayes analysis of the full mtDNA dataset above. Though posterior probabilities were low across nodes, the consensus topology (Fig. S2) was consistent with the SNAPP species topology for extant lineages and placed the extinct Mt. Cookson lineage as basal to the rest of the complex, and the extinct Mt. Somers lineage as sister to a clade that included the two Fiordland and Haast lineages. We used this topology in our multigene coalescent analysis of divergence times in Section 5 below.

Section 4: Calibration of kiwi crown group

We used 16 nuclear genes to generate a dated phylogeny of birds (Fig. S3; Table S5) which included a representative from the five kiwi species, other ratite groups, and lineages from a variety of other avian orders for which well-supported external calibration points exists. We vetted a pool of available genes from previously published sources (16–18) for mutation saturation, and retained only those genes or codon positions that showed minimal signs of saturation over the crown group of birds. We included the ten protein-coding genes used by Haddrath & Baker (18) and excluded third codon positions for four genes (CMOS, NT3, PTPN12 and TRAF6) in which there was a strong departure from a linear relationship between genetic distances estimated for fast evolving third codon positions versus slower 1st and 2nd codon positions (an indication of saturation; Fig. S6). We also considered 17 largely intronic sequences

(16, 17), but included only six of these in our analyses for which a nearly linear relationship occurred between the proportion of transitions versus transversions in pairwise comparisons (as estimated using DAMBE5, (19, 20)) across the dataset (a non-linear relationship suggests saturation; Fig. S7). This latter dataset already included one species of kiwi and we sequence these six genes for the remaining four species (Table S5). In order to maximize our inclusion of high quality calibration points and to break up long branches we extracted these same genes from recently published genomes (Jarvis et al. 2014), for a total of 45 species (Table S5).

Seventeen well supported avian fossils (Haddrath & Baker 2012; Jarvis *et al*. 2014) were used as calibrations (Table S2 and Fig. S3), with the age of the fossil being treated as a minimum age calibration using uniform priors that ranged from 133 Mya (18) to the age of the fossil for all but two calibrated nodes. The exceptions were the node defining Galloanserae whose fossil calibration (*Vegavis iaai*) was given an upper limit of 20 million years older than the fossil, following Mitchell *et al.* 2014 and the Spheniscidae stem which was also given an upper limit 20 million years older than the fossil. We also used the date of splitting of New Zealand from Australia (between 85 and 65 Mya) as a maximum age calibration of when the Strigopidae (an endemic family of New Zealand parrots which represent the basal split within Psittaciformes; (21–23) and the Acanthisitti (the endemic New Zealand wrens which represent the basal split within Passeriformes; (24, 25)) could have split from their sister groups. These calibrations were implemented with a uniform distribution spanning 85 to 0 Mya. The molecular dating was performed using BEAST2.3.1 (26). The 16 gene (6,190 bp) data set was divided into nine partitions, three for each of the codon positions of the 10 exclusively protein-coding genes (as in (18)) and a separate partition for each of the six largely intronic gene sequences. The

branching order of the tree was constrained in the BEAST2 analysis to match the tree topologies of (27) and (18). Three independent BEAST2 MCMC chains were run for 200 million generations, each sampled every 20,000. After completion, the first 10% of the trees from each run were discarded as burnin and the results combined in TreeAnnotator. Tracer version 1.6 (28) was used to assess convergence of parameter values. We used the median node age of the basal split within kiwi (dated to 5.96 Ma; 95% confidence region, 2.90-10.80 Ma) as a calibration point for our dated analyses (see next section).

Section 5: Phylogenetic dating

We generated time-calibrated phylogenies of kiwi lineages using two methods. First, we generated a relaxed-clock phylogeny (with rate variation following a log-normal distribution and with Yule speciation prior) in BEAST2 (26) for mtDNA using a single individual per each of the 17 lineages (Fig. S4). The basal kiwi node was set to 5.96 Ma and the ages of all other nodes were estimated. A GTR-y model (with rate variation following a y distribution with 5 categories and with shape parameter estimated from the data) was used. Separate partitions were set for cytochrome B and the control region. Clock rates for each partition were estimated from the data. We used lognormal priors for transition rates between nucleotide categories (with default settings for those priors), but otherwise used default priors for other parameters. The analysis was run for 10 million generations, with the first two million generations set as the burnin (as assessed by Tracer) and samples obtained every 1000 generations following the burnin. A maximum clade credibility tree with median node ages was calculated in TreeAnnotator (15) (Fig S4).

The BEAST node ages estimate mtDNA coalescent times which are expected to predate population or lineage divergence times. We therefore used the population genetic modelling framework in the program G-PHOCS (5) to obtain dates of population splitting. G-PHOCS does not estimate the species tree topology, but rather estimates dates of population splitting and effective population sizes along a supplied species tree topology. Two datasets were used. The first included only nuclear genes and included 4231 nuclear loci (each 90 bp in length and with no missing data, no W-linked loci, but including loci with no variable sites) for three individuals (each with two gene copies for a total of six copies of each locus) for each of the 11 extant lineages. This dataset used the well supported species tree topology generated in SNAPP (Fig. 1C). The second dataset was the same as the first, but also included the two mitochondrial genes for the sample individuals (treated as a single non-recombining locus) as well as for the extinct lineages. This dataset also used the SNAPP species tree topology, with extinct lineages added according to their placement in mtDNA phylogeny shown in Fig 1A for *A. rowi* and *A. owenii*, and according to the topology of Fig S2 for the *A. australis* clade.

Like bpp used above (Section 2), G-PHOCS has two key population genetic parameters, Θ and τ, and allows priors to be set for these at each node in the phylogeny (bpp allows priors for these to be set only for the basal node). We used similar methods as for the bpp analysis to set these priors. JC69 genetic distances were calculated for the nuclear dataset using the R ape package (10). We then calculate the average pairwise genetic distance within lineages as rough estimate of Θ. The average of these across the 11 extant lineages was 0.000857 and we used this value as the mean (α/β) of our gamma prior for Θ, and set $\alpha = 2$ (resulting in a rather diffuse prior). For each node we estimated the mean value for the gamma prior of τ to be equal

to the sequence divergence between the daughter clades descending from a node minus ½ the average pairwise distances within the extant populations in each of those clades (see Table S3). α was likewise set to 2 for each of these priors.

G-PHOCS was run either with or without gene flow between geographically adjacent extant lineages within *A. australis* and *A. mantelli*. For the model with gene flow, we allowed bi-directional migration between the following pairs of populations of 1) *A. australis*: S. Fiordland/N. Fiordland, Haast/N. Fiordland, Stewart Is./S.; and 2) *A. mantelli*: Coromandel/Northland, Coromandel/Taranaki, Taranaki/Eastern. Each of the two models (with or without gene flow) for each dataset (with or without mtDNA included) were run 96 times (for a total of 48 hours, the maximum time allowed on our supercomputing cluster) from different random starting parameters for a total of 27,000 (with mtDNA) or 31,000 (no mtDNA) generations each. For each run the first 10,000 generations were discarded as burnin. This burnin was sufficient for most runs, but a few runs did not converge. Such runs typically have elevated variance in posterior probabilities across generations. To detect and delete these runs, we used the following approach. First, we calculated the standard deviation in posterior probabilities ("Full.ld.ln") and data likelihood ("Data.ld.ln") provided by G-PHOCS for each run, and from these calculated the 3rd quartile and interquartile range in standard deviations across runs. We excluded runs whose standard deviations in either measure exceeded the third quartile by 1.5 times the interquartile range. The remaining runs were combined, visually inspected for individual runs that may have converged on a different region of likelihood space (indicating a local optima) (none were detected), and median values of Θ and τ were obtained for each branch in our analyses.

To calibrate τ values into actual time, we used the following approach adapted from (5) and illustrated in Fig. S5. The τ value of the basal kiwi split, τ_{root}, represents the date of population divergence, while our 5.96 Ma calibration point for the genomic divergence (based on 16 nuclear loci) at the base of the kiwi phylogeny represents an average date of gene coalescence, τ_{coal} which is expected to predate τ_{root} by ½ Θ_{root}. The τ value for each node *i* can be calibrated into actual time, T, as follows:

$$T_i = \tau_i / (\tau_{coal} / 5.96 \text{ Ma}),$$

where $\tau_{coal} = \tau_{root} + \frac{1}{2}\Theta_{root}$. Equation 1

Table S3 lists values of T and τ for each node.

Section 6: Ancestor State Reconstructions

We performed parsimony based ancestor state reconstruction of kiwi distribution in the North and South Islands to determine where kiwi originated, the number of in situ diversification events within each island and the number of inter-island diversification events. Reconstructions were performed in Mesquite 3.04 (29) along the topology used in the G-PHOCS analysis (see Fig. S5). We added the extinct fossil lineage *Proapteryx* (30) from the early Miocene (16 to19 Ma) of the South Island to the topology (Fig. 4). This is the only pre-Holocene fossil lineage of kiwi, with all Holocene fossil lineages that we are aware of already included on the phylogeny. We assume that lineages can occur in only one character state. Lineages dispersing across the landbridge connecting the North and South Island during glacial cycles

may have briefly occurred in both character states, but would have been genetically isolated shortly after (during the next interglacial period of low high sea level), thus supporting our use of discrete character states. Two equally parsimonious reconstructions were obtained and are shown in Fig. 4).

Section 7: Effective Population Size

We used the approximate Bayesian computation (ABC) pipeline in PopSizeABC (31) to estimate effective population size (N_e) through time over the course of the last glacial cycle from 130 Ka (the end of the interglacial preceding the last glacial cycle) to 2.7 Ka. PopSizeABC estimates empirical summary statistics and matches these to summary statistics simulated under a range of N_e at each time point in order to determine how N_e changes through time.

N_e was estimated in 24 discrete time bins, the size of which increased exponentially from the recent to the past (parameter α = 0.0005 in PopSizeABC determines the length of each time bin) with the breakpoints between time bins as follows: 2.74 Ka, 5.63Ka, 8.68 Ka, 11.90 Ka, 15.29 Ka, 18.88 Ka, 22.65 Ka, 26.63 Ka, 30.83 Ka, 35.26 Ka, 39.93 Ka, 44.87 Ka, 50.07 Ka, 55.56 Ka, 61.34 Ka, 67.45 Ka, 73.88 Ka, 80.68 Ka, 87.84 Ka, 95.40 Ka, 103.38 Ka, 111.78 Ka, 120.64 Ka, 130.00 Ka. The model estimates N_e for each of the 24 time bins given the per generation per site mutation rate μ and recombination rate r.

Kiwi have a life expectancy of 40 to 50 years with high adult survivorship (32). Based on these values and our extensive experience working with kiwi populations, we estimate that the

average age of breeding females in a population to be about 25 years and use this as our estimate of the generation time. We calculate a genome-wide mutation rate for kiwi using the result from our G-PHOCS analysis where τ_{coal} equaled 37.52. Given $\tau = \mu T$, where T is the number of generations, 238,400 generations would have occurred during the past 5.96 Ma, resulting in $\mu = 1.57e^{-8}$ (or a per year rate of $3.93e^{-7}$). For recombination rate we used a uniform prior from $2e^{-8}$ to $4e^{-8}$ reflecting the rate estimated for the chicken (33). We used genomic data for five diploid individuals for each kiwi lineage, except for *A. owenii* for which only three were available. The mpileup function of samtools 1.3 (34)(options: -q 20 -Q 20 -C 50 -u) and the call function of bcftools 1.3 (34)(options: -c -f GQ -V indels -v) were used to generate VCF files for each kiwi lineage from our BAM alignment files to the kiwi reference genome. The VCF files were used by PopSizeABC to calculate a number of summary statistics on our empirical datasets and to compare these to summary statistics obtained from genomic data simulated under the priors of our model. We used the same genome-wide allele frequency spectrum and time bin specific linkage disequilibrium summary statistics as (31). These statistics were calculated only for SNPs above a minor allele count of two. PopSizeABC cut our empirical datasets into chromosome sizes of 200,000 bp or less (as appropriate over the timescales we tested given our rate of mutation and recombination) and then obtained 400,000 simulated datasets, each with 100 200,000 bp segments. All other PopSizeABC simulation settings were kept at their null values.

We used an almost identical routine as (31) to obtain the final ABC estimates of N_e. We used a tolerance rate of 0.005 to accept simulated samples and used 1000 (rather than only

100) neural network regressions to link parameters and summary statistics. For each analysis 95% CI were calculated. Results are summarized in Fig. 4 and shown in full in Fig S8.

References

1. Elshire RJ, et al. (2011) A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. *PLoS ONE* 6(5):e19379.

2. Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) Stacks: an analysis tool set for population genomics. *Mol Ecol* 22(11):3124–3140.

3. Le Duc D, et al. (2015) Kiwi genome provides insights into evolution of a nocturnal lifestyle. *Genome Biol* 16(1):1–15.

4. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9(4):357–359.

5. Gronau I, Hubisz MJ, Gulko B, Danko CG, Siepel A (2011) Bayesian inference of ancient human demography from individual genome sequences. *Nat Genet* 43(10):1031–1034.

6. Hammer Ø, Harper DAT, Ryan PD (2001) *PAST: Paleontological Statistics Software Package for education and data analysis. Palaeontolia Electronica 4*.

7. Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155(2):945–959.

8. Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14(8):2611–2620.

9. Bhatia G, Patterson N, Sankararaman S, Price AL (2013) Estimating and interpreting FST: The impact of rare variants. *Genome Res* 23(9):1514–1521.

10. Paradis E, Claude J, Strimmer K (2004) APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20(2):289–290.

11. Yang Z, Rannala B (2010) Bayesian species delimitation using multilocus sequence data. *Proc Natl Acad Sci* 107(20):9264–9269.

12. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19(12):1572–1574.

13. Nylander JAA (2004) MrModeltest v2. Program distributed by the author. *Evol Biol Cent Upps Univ* 2.

14. Bryant D, Bouckaert R, Felsenstein J, Rosenberg NA, RoyChoudhury A (2012) Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. *Mol Biol Evol*:mss086.

15. Rambaut A, Drummond AJ (2013) *TreeAnnotator v1. 7.5*.

16. Hackett SJ, et al. (2008) A phylogenomic study of birds reveals their evolutionary history. *science* 320(5884):1763–1768.

17. Harshman J, et al. (2008) Phylogenomic evidence for multiple losses of flight in ratite birds. *Proc Natl Acad Sci* 105(36):13462–13467.

18. Haddrath O, Baker AJ (2012) Multiple nuclear genes and retroposons support vicariance and dispersal of the palaeognaths, and an Early Cretaceous origin of modern birds. *Proc R Soc Lond B Biol Sci*:rspb20121630.

19. Xia X, Xie Z (2001) DAMBE: software package for data analysis in molecular biology and evolution. *J Hered* 92(4):371–373.

20. Xia X, Xie Z, Salemi M, Chen L, Wang Y (2003) An index of substitution saturation and its application. *Mol Phylogenet Evol* 26(1):1–7.

21. de Kloet RS, de Kloet SR (2005) The evolution of the spindlin gene in birds: sequence analysis of an intron of the spindlin W and Z gene reveals four major divisions of the Psittaciformes. *Mol Phylogenet Evol* 36(3):706–721.

22. Astuti D, Azuma N, Suzuki H, Higashi S (2006) Phylogenetic relationships within parrots (Psittacidae) inferred from mitochondrial cytochrome-b gene sequences. *Zoolog Sci* 23(2):191–198.

23. Wright TF, et al. (2008) A multilocus molecular phylogeny of the parrots (Psittaciformes): support for a Gondwanan origin during the Cretaceous. *Mol Biol Evol* 25(10):2141–2156.

24. Ericson PG, et al. (2002) A Gondwanan origin of passerine birds supported by DNA sequences of the endemic New Zealand wrens. *Proc R Soc Lond B Biol Sci* 269(1488):235–241.

25. Barker FK, Cibois A, Schikler P, Feinstein J, Cracraft J (2004) Phylogeny and diversification of the largest avian radiation. *Proc Natl Acad Sci U S A* 101(30):11040–11045.

26. Bouckaert R, et al. (2014) BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Comput Biol* 10(4):e1003537.

27. Jarvis ED, et al. (2014) Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science* 346(6215):1320–1331.

28. Rambaut A, Suchard MA, Xie D, Drummond AJ (2014) *Tracer v1. 6*.

29. Maddison WP, Maddison DR (2011) *Mesquite: a modular system for evolutionary analysis. Version 2.75.* Available at: http://mesquiteproject.org.

30. Worthy TH, et al. (2013) Miocene fossils show that kiwi (Apteryx, Apterygidae) are probably not phyletic dwarves. *Paleornithological Research 2013—Proceedings of the 8th International Meeting of the Society of Avian Paleontology and Evolution. Vienna: Natural History Museum Vienna*, pp 63–80.

31. Boitard S, Rodríguez W, Jay F, Mona S, Austerlitz F (2016) Inferring Population Size History from Large Samples of Genome-Wide Molecular Data - An Approximate Bayesian Computation Approach. *PLOS Genet* 12(3):e1005877.

32. Robertson HA, de MONCHY PJM (2012) Varied success from the landscape-scale management of kiwi Apteryx spp. in five sanctuaries in New Zealand. *Bird Conserv Int* 22(04):429–444.

33. Groenen MAM, et al. (2008) A high-density SNP-based linkage map of the chicken genome reveals sequence features correlated with recombination rate. *Genome Res* 19(3):510–519.

34. Li H, et al. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25(16):2078–2079.

SI Figure legends

Figure S1

Majority-rule consensus phylogeny generated in MrBayes using mtDNA for 203 individuals. Each tip represents a haplotype and is often comprised of several individuals as listed in the tip label names. Posterior probabilities are shown at all nodes.

Figure S2

Majority-rule consensus phylogeny generated in MrBayes for mtDNA and nuclear genomic data for the *A. australis* species complex. Posterior probabilities are shown at nodes. The analysis is rooted with *A. haastii.*

Figure S3

A) Fossil calibrations and B) the resulting time-calibrated phylogeny for the crown group of birds.

Figure S4

A time-calibrated maximum clade credibility phylogeny (with median node ages) for mtDNA using a relaxed-clock model in BEAST with a 5.96 Ma calibration point set for the basal kiwi node. Node bars indicate 95% credible intervals for each node age. Posterior probabilities at nodes are shown.

Figure S5

Schematic of the population genetic model fit using G-PHOCS to kiwi nuclear and mtDNA data. The model estimates splitting times (e.g. each scaled by the mutation rate) for the 16 nodes (τ_{root}, τ_2 to τ_{16}). The value τ_{coal} (not a parameter of the model) was calculated from the parameters τ_{root} (the basal kiwi splitting event) and θ_{root} (effective population size scaled by the mutation rate for the ancestral kiwi population), and was calibrated with a date of 5.96 Ma (the mean estimated date of gene coalescence for the basal split). τ_2 to τ_{16} were then calibrated using Equation 1 (see calibrated values in Table S3).

Figure S6

Assessment of saturation at 3rd codon positions for 10 protein-coding nuclear genes. Genetic distances (GTR-gamma) are plotted on the x-axis for 3rd codon positions and on the y-axis for 1st and 2nd codon positions. If the faster evolving third codon positions exhibit saturation, then a strong non-linear relationship should be evident. We used linear regression (lm function in R) to obtain the best fit linear

model and a quadratic polynomial model (red), both fit through the origin. Red stars indicate genes for which we excluded 3rd codon positions.

Figure S7

Assessment of saturation for 17 nuclear introns. Genetic distances (GTR-gamma) are plotted on the x-axis for transversions and on the y-axis for transitions. Saturation is indicated if distances for transversions and transitions are not linear. We used linear regression (lm function in R) to obtain the best fit linear model (blue) and a quadratic polynomial model (red), both fit through the origin. Blue stars indicate genes whose relationship was linear and which we included in our analysis.

Figure S8

Effective population size (N_e) through time over the last glacial cycle for the eleven extant lineages of kiwi. N_e is plotted on a log10 scale. Solid lines show median estimates, and dotted lines show 95% credible intervals.

Figure S1

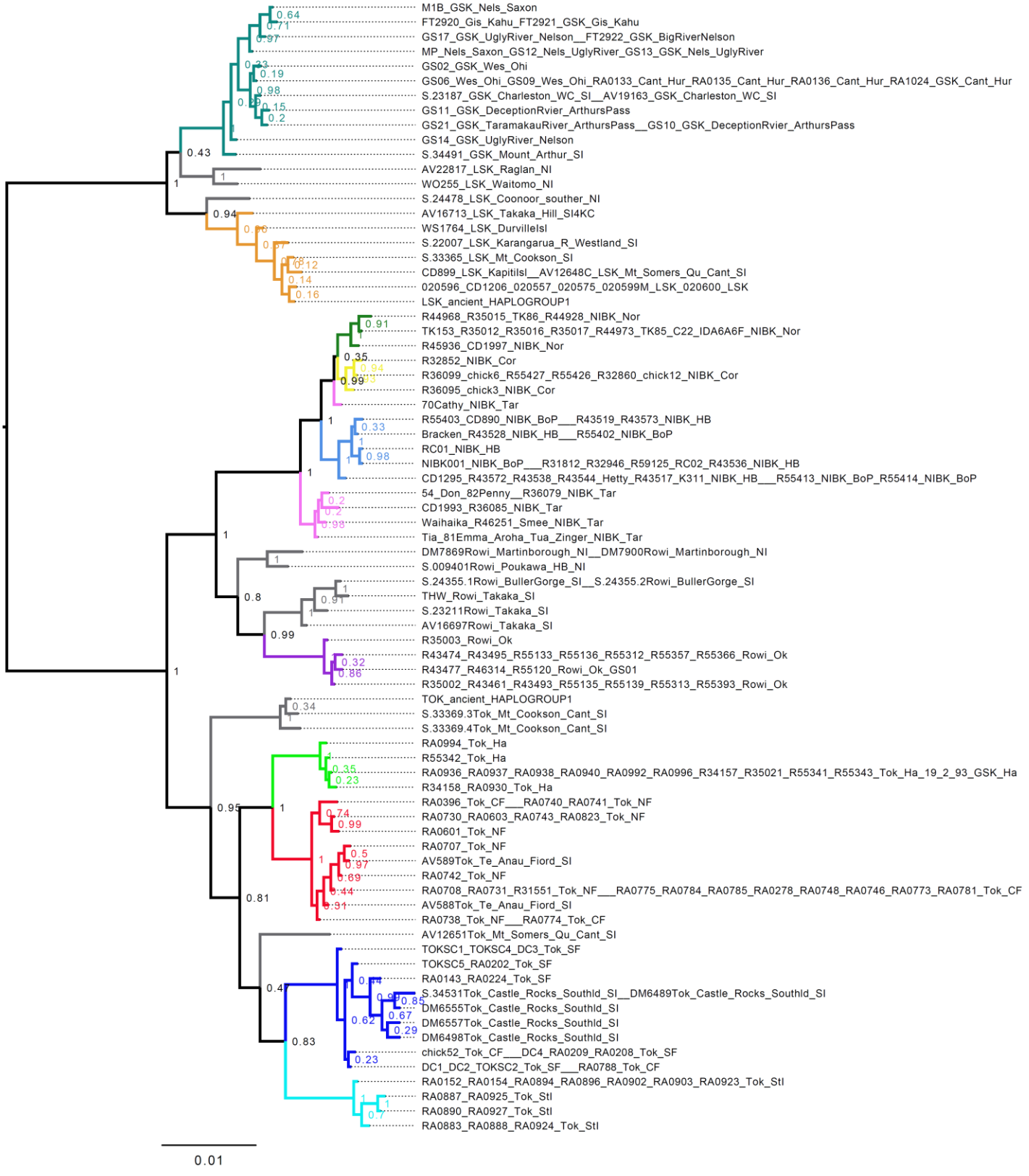

M1B_GSK_Nels_Saxon
FT2920_Gis_Kahu_FT2921_GSK_Gis_Kahu
GS17_GSK_UglyRiver_Nelson__FT2922_GSK_BigRiverNelson
MP_Nels_Saxon_GS12_Nels_UglyRiver_GS13_GSK_Nels_UglyRiver
GS02_GSK_Wes_Ohi
GS06_Wes_Ohi_GS09_Wes_Ohi_RA0133_Cant_Hur_RA0135_Cant_Hur_RA0136_Cant_Hur_RA1024_GSK_Cant_Hur
S.23187_GSK_Charleston_WC_SI__AV19163_GSK_Charleston_WC_SI
GS11_GSK_DeceptionRvier_ArthursPass
GS21_GSK_TaramakauRiver_ArthursPass__GS10_GSK_DeceptionRvier_ArthursPass
GS14_GSK_UglyRiver_Nelson
S.34491_GSK_Mount_Arthur_SI
AV22817_LSK_Raglan_NI
WO255_LSK_Waitomo_NI
S.24478_LSK_Coonoor_souther_NI
AV16713_LSK_Takaka_Hill_SI4KC
WS1764_LSK_DurvilleIsl
S.22007_LSK_Karangarua_R_Westland_SI
S.33365_LSK_Mt_Cookson_SI
CD899_LSK_KapitiIsl__AV12648C_LSK_Mt_Somers_Qu_Cant_SI
020596_CD1206_020557_020575_020599M_LSK_020600_LSK
LSK_ancient_HAPLOGROUP1
R44968_R35015_TK86_R44928_NIBK_Nor
TK153_R35012_R35016_R35017_R44973_TK85_C22_IDA6A6F_NIBK_Nor
R45936_CD1997_NIBK_Nor
R32852_NIBK_Cor
R36099_chick6_R55427_R55426_R32860_chick12_NIBK_Cor
R36095_chick3_NIBK_Cor
70Cathy_NIBK_Tar
R55403_CD890_NIBK_BoP___R43519_R43573_NIBK_HB
Bracken_R43528_NIBK_HB___R55402_NIBK_BoP
RC01_NIBK_HB
NIBK001_NIBK_BoP___R31812_R32946_R59125_RC02_R43536_NIBK_HB
CD1295_R43572_R43538_R43544_Hetty_R43517_K311_NIBK_HB___R55413_NIBK_BoP_R55414_NIBK_BoP
54_Don_82Penny__R36079_NIBK_Tar
CD1993_R36085_NIBK_Tar
Waihaika_R46251_Smee_NIBK_Tar
Tia_81Emma_Aroha_Tua_Zinger_NIBK_Tar
DM7869Rowi_Martinborough_NI__DM7900Rowi_Martinborough_NI
S.009401Rowi_Poukawa_HB_NI
S.24355.1Rowi_BullerGorge_SI__S.24355.2Rowi_BullerGorge_SI
THW_Rowi_Takaka_SI
S.23211Rowi_Takaka_SI
AV16697Rowi_Takaka_SI
R35003_Rowi_Ok
R43474_R43495_R55133_R55136_R55312_R55357_R55366_Rowi_Ok
R43477_R46314_R55120_Rowi_Ok_GS01
R35002_R43461_R43493_R55135_R55139_R55313_R55393_Rowi_Ok
TOK_ancient_HAPLOGROUP1
S.33369.3Tok_Mt_Cookson_Cant_SI
S.33369.4Tok_Mt_Cookson_Cant_SI
RA0994_Tok_Ha
R55342_Tok_Ha
RA0936_RA0937_RA0938_RA0940_RA0992_RA0996_R34157_R35021_R55341_R55343_Tok_Ha_19_2_93_GSK_Ha
R34158_RA0930_Tok_Ha
RA0396_Tok_CF___RA0740_RA0741_Tok_NF
RA0730_RA0603_RA0743_RA0823_Tok_NF
RA0601_Tok_NF
RA0707_Tok_NF
AV589Tok_Te_Anau_Fiord_SI
RA0742_Tok_NF
RA0708_RA0731_R31551_Tok_NF___RA0775_RA0784_RA0785_RA0278_RA0748_RA0746_RA0773_RA0781_Tok_CF
AV588Tok_Te_Anau_Fiord_SI
RA0738_Tok_NF___RA0774_Tok_CF
AV12651Tok_Mt_Somers_Qu_Cant_SI
TOKSC1_TOKSC4_DC3_Tok_SF
TOKSC5_RA0202_Tok_SF
RA0143_RA0224_Tok_SF
S.34531Tok_Castle_Rocks_Southld_SI__DM6489Tok_Castle_Rocks_Southld_SI
DM6555Tok_Castle_Rocks_Southld_SI
DM6557Tok_Castle_Rocks_Southld_SI
DM6498Tok_Castle_Rocks_Southld_SI
chick52_Tok_CF___DC4_RA0209_RA0208_Tok_SF
DC1_DC2_TOKSC2_Tok_SF___RA0788_Tok_CF
RA0152_RA0154_RA0894_RA0896_RA0902_RA0903_RA0923_Tok_StI
RA0887_RA0925_Tok_StI
RA0890_RA0927_Tok_StI
RA0883_RA0888_RA0924_Tok_StI
0.01

Figure S2

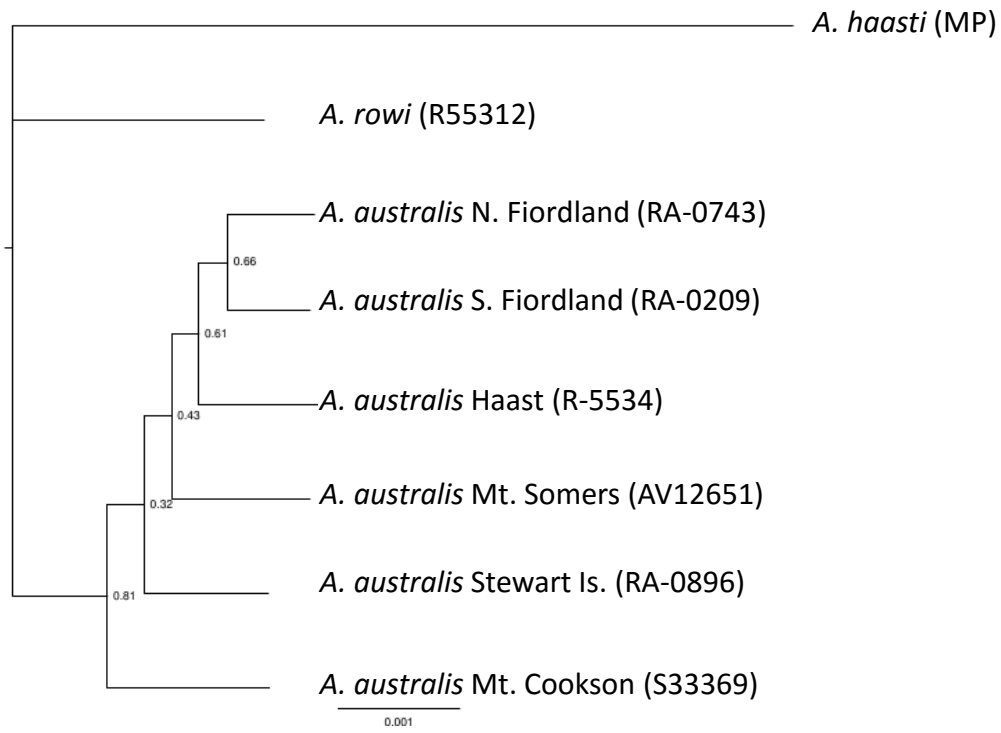
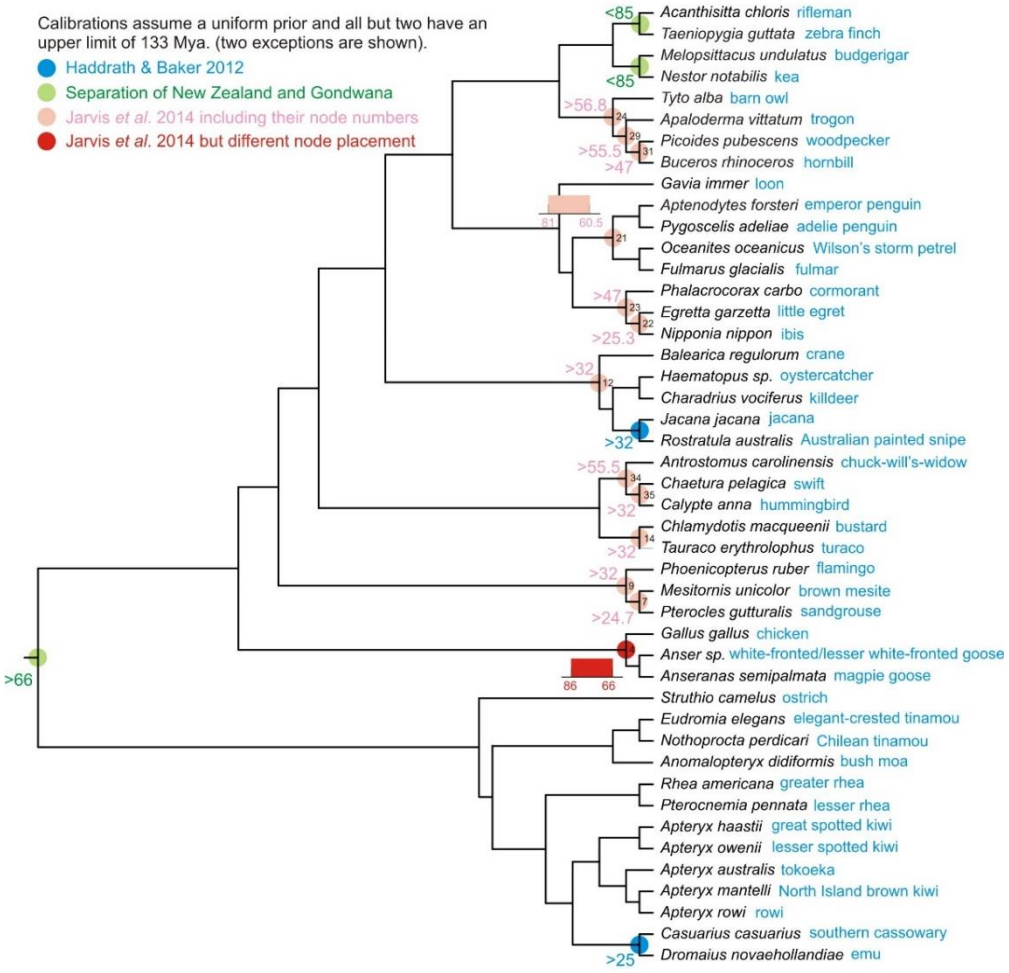

A. haasti (MP)
A. rowi (R55312)
A. australis N. Fiordland (RA-0743)
A. australis S. Fiordland (RA-0209)
A. australis Haast (R-5534)
A. australis Mt. Somers (AV12651)
A. australis Stewart Is. (RA-0896)
A. australis Mt. Cookson (S33369)
0.66
0.61
0.43
0.32
0.81
0.001

Figure S3

A) Fossil calibrations



B) Time calibrated phylogeny of birds

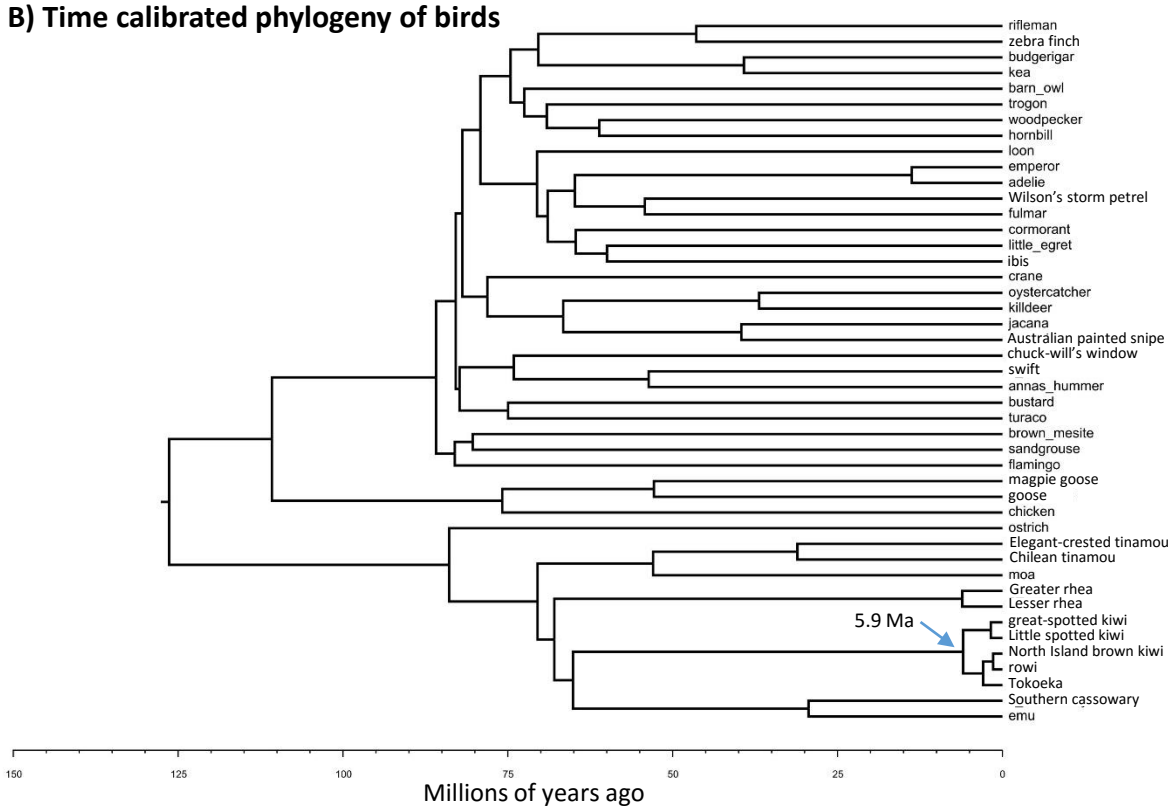


Figure S4

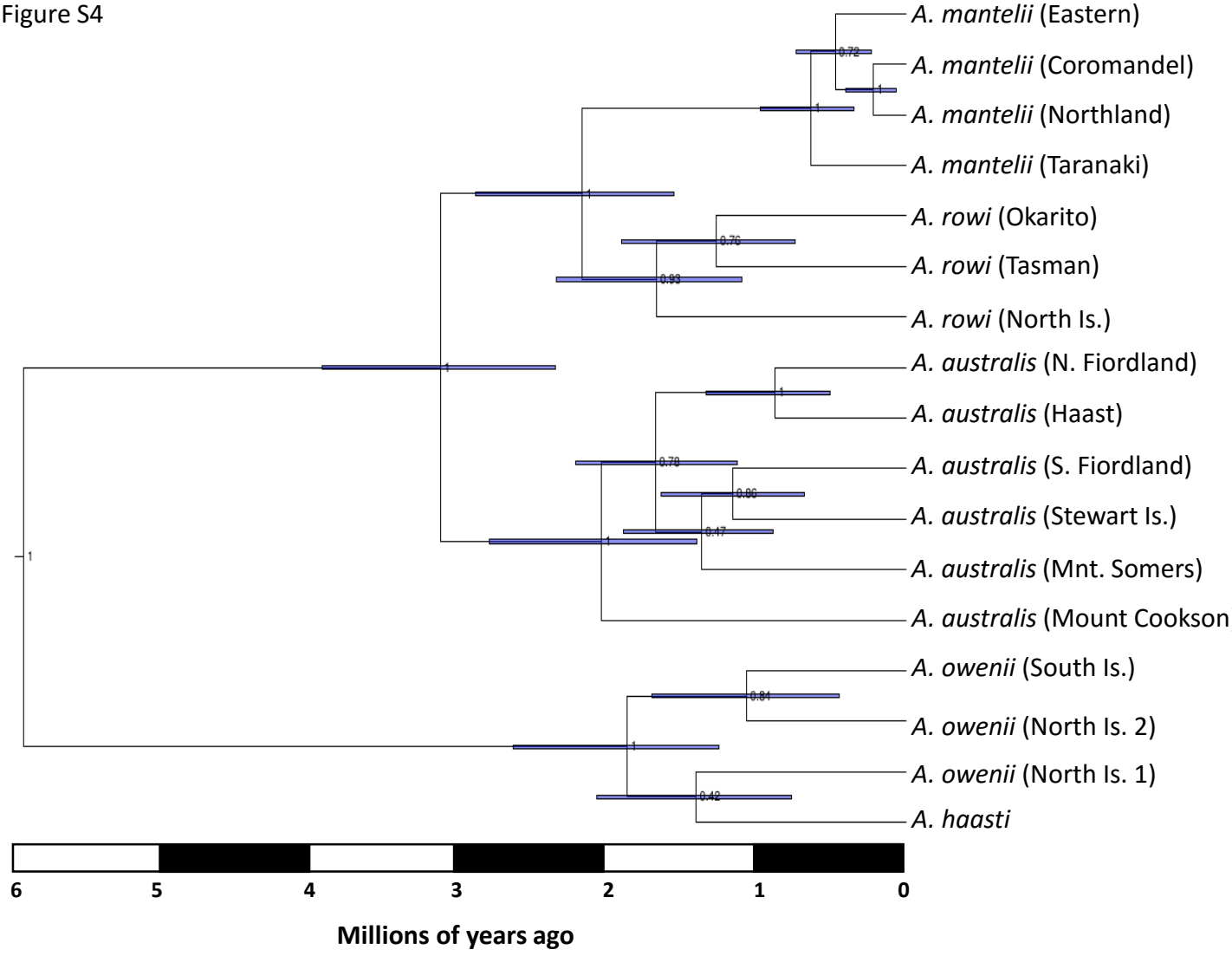

A. mantelii (Eastern)
A. mantelii (Coromandel)
A. mantelii (Northland)
A. mantelii (Taranaki)
A. rowi (Okarito)
A. rowi (Tasman)
A. rowi (North Is.)
A. australis (N. Fiordland)
A. australis (Haast)
A. australis (S. Fiordland)
A. australis (Stewart Is.)
A. australis (Mnt. Somers)
A. australis (Mount Cookson
A. owenii (South Is.)
A. owenii (North Is. 2)
A. owenii (North Is. 1)
A. haasti
0.72
1
1
1
1
0.76
0.93
1
1
0.78
0.86
0.47
1
1
0.84
1
0.48
6
5
4
3
2
1
0
Millions of years ago

Figure S5

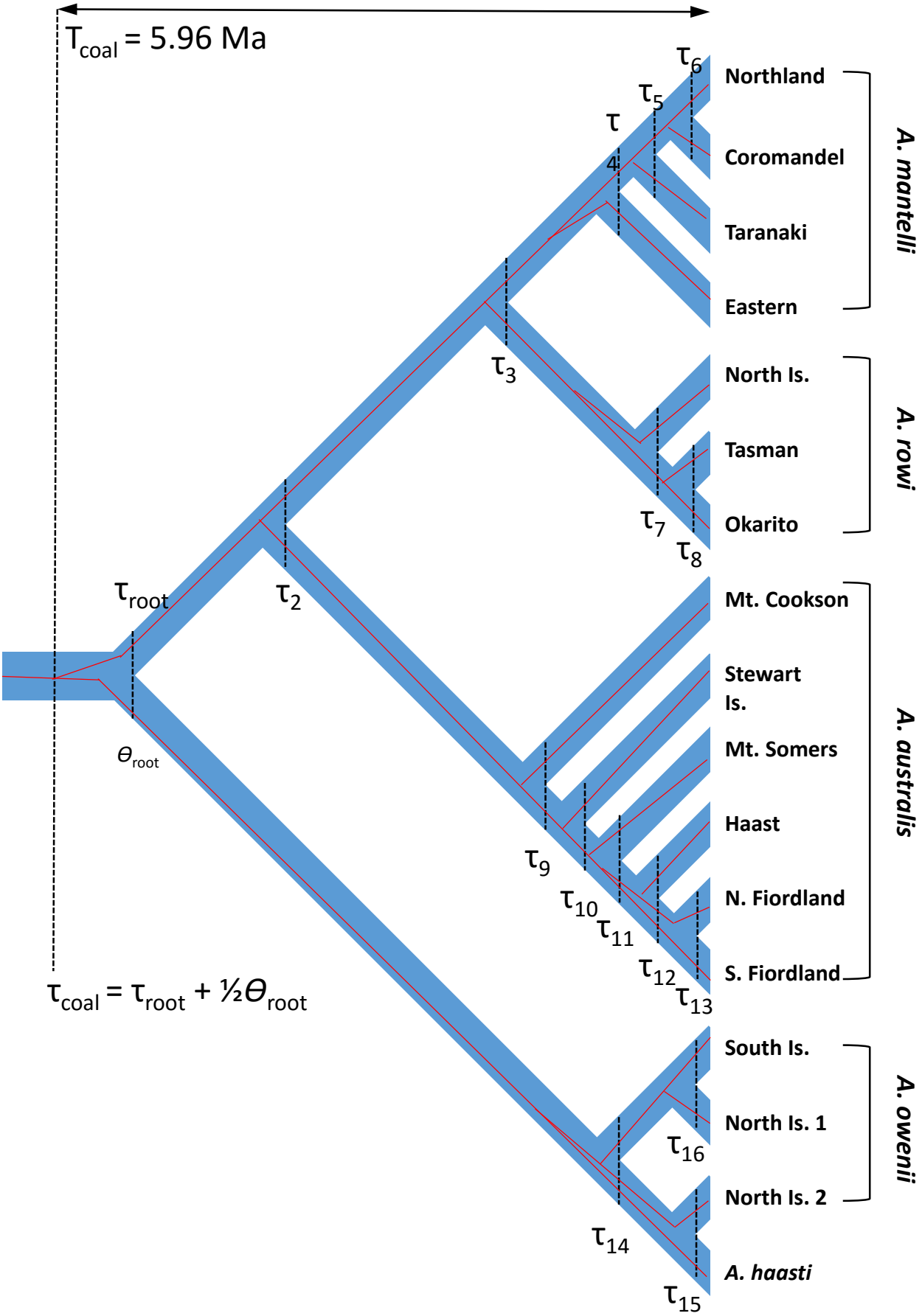

T_{coal} = 5.96 Ma
$\tau_{coal} = \tau_{root} + ½\theta_{root}$
τ_{root}
τ_2
τ_3
τ_4
τ_5
τ_6
τ_7
τ_8
τ_9
τ_{10}
τ_{11}
τ_{12}
τ_{13}
τ_{14}
τ_{15}
τ_{16}
θ_{root}
Northland
Coromandel
Taranaki
Eastern
North Is.
Tasman
Okarito
Mt. Cookson
Stewart Is.
Mt. Somers
Haast
N. Fiordland
S. Fiordland
South Is.
North Is. 1
North Is. 2
A. haasti
A. mantelli
A. rowi
A. australis
A. owenii

Figure S6

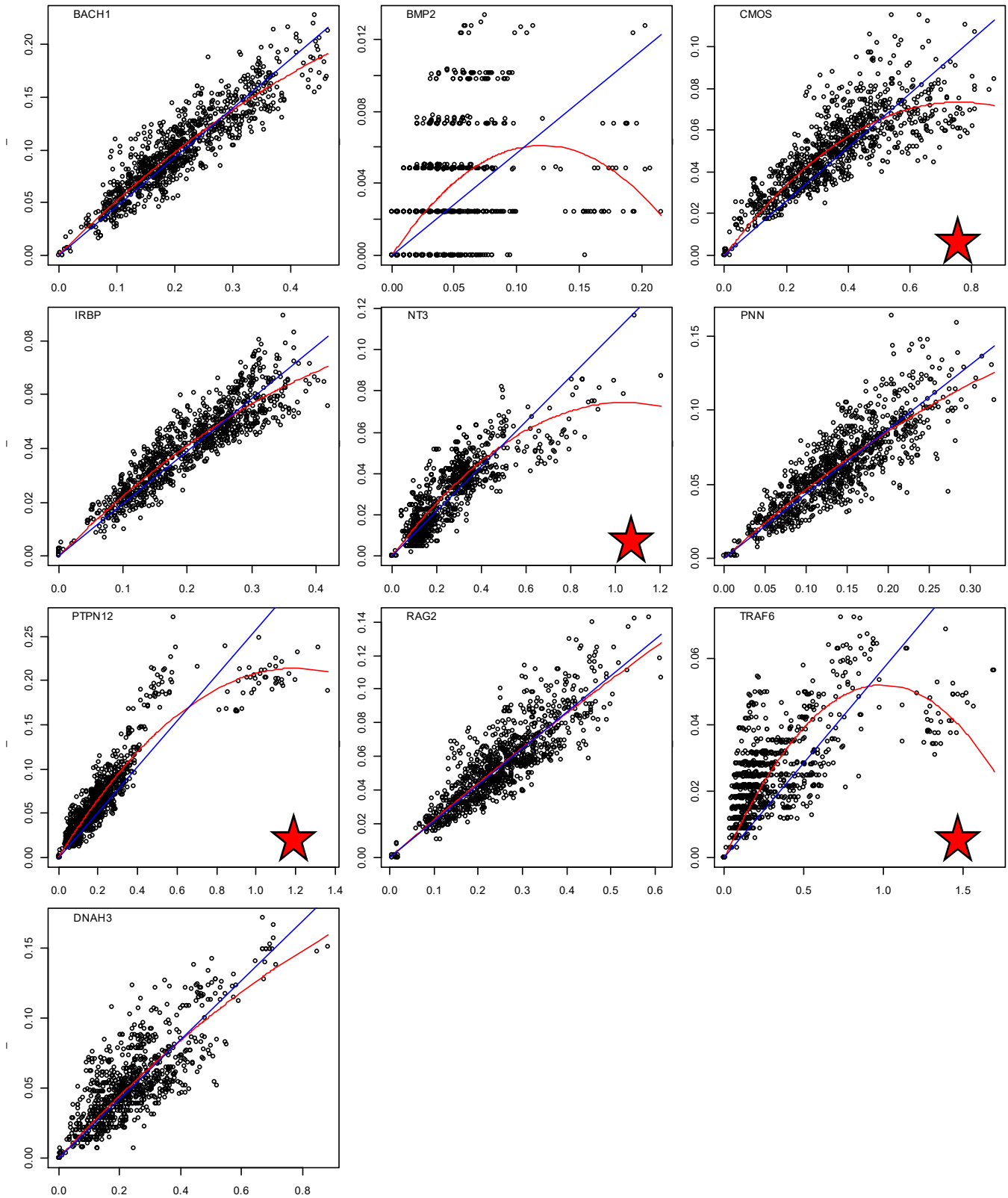
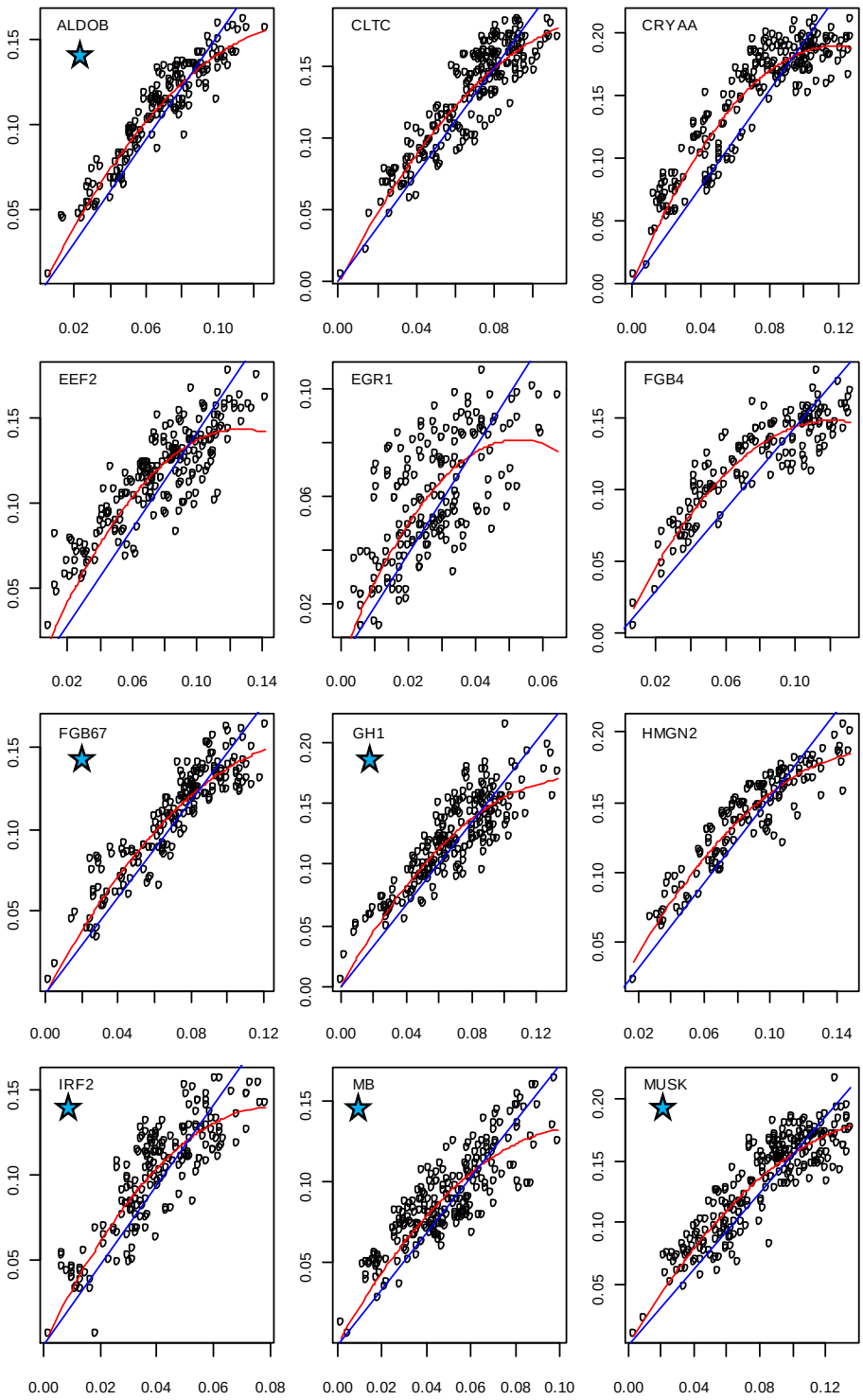


Figure S7



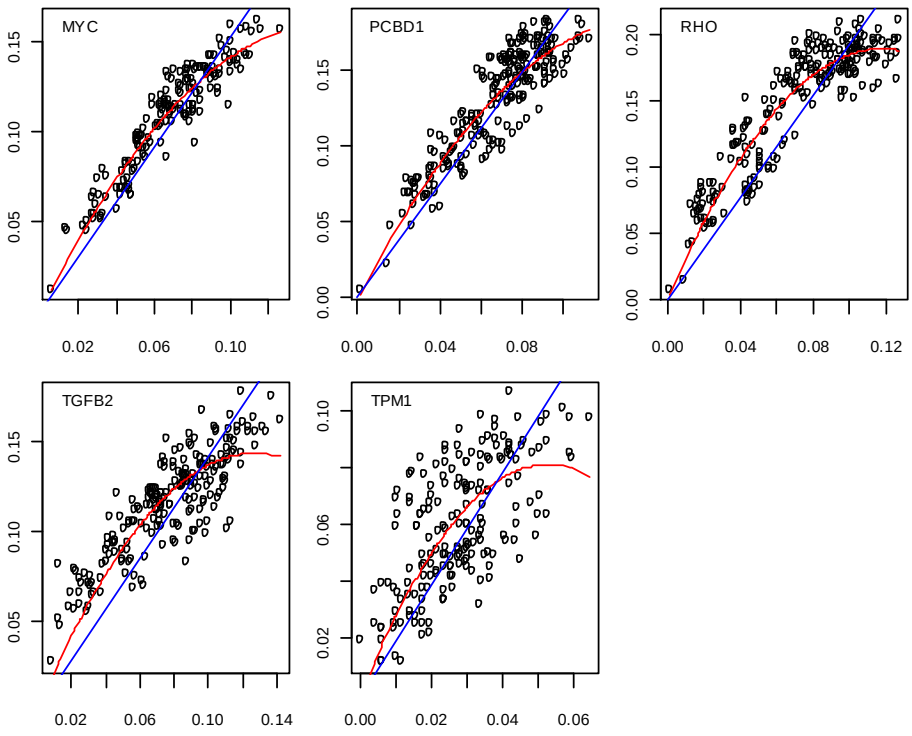
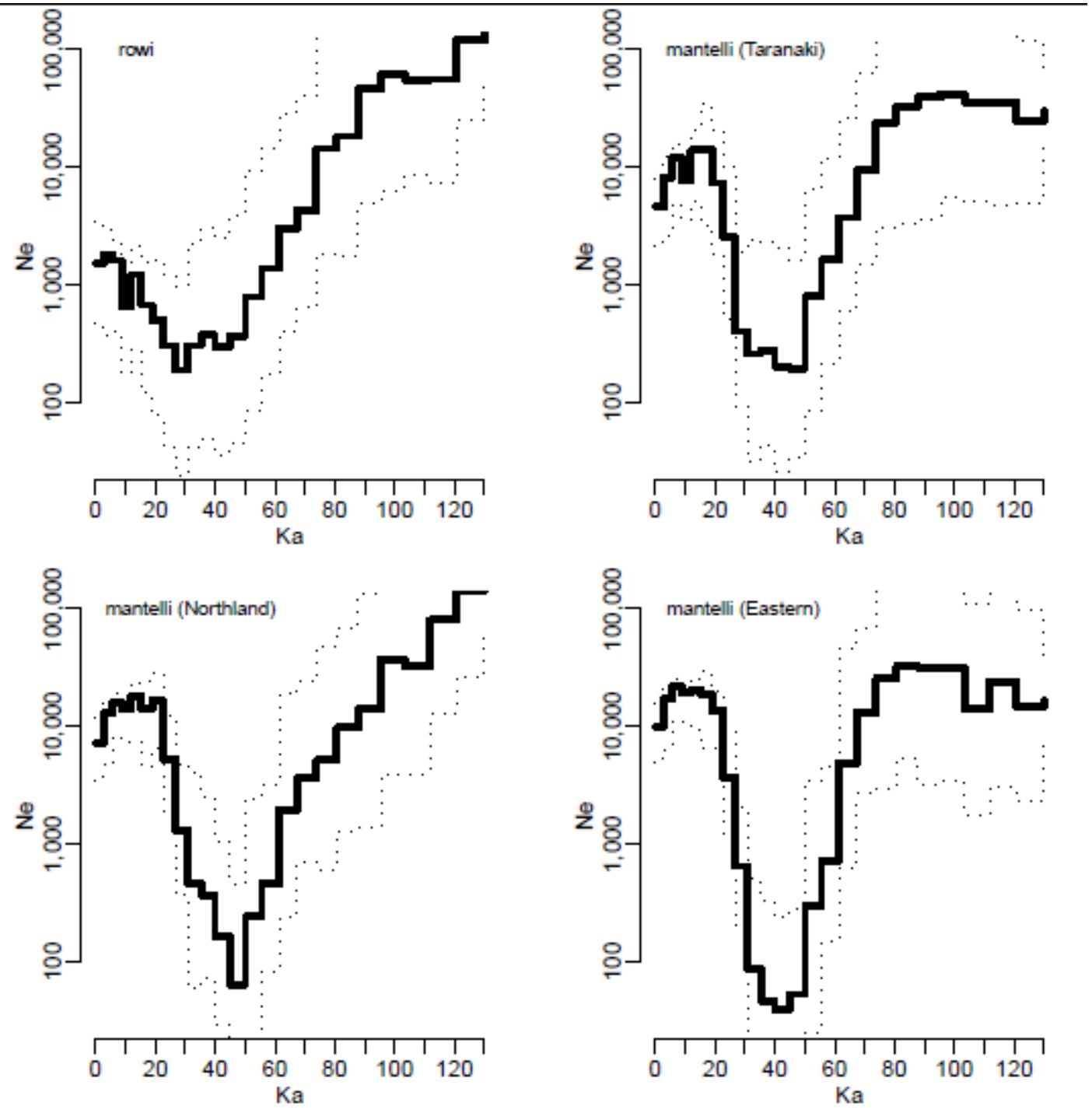
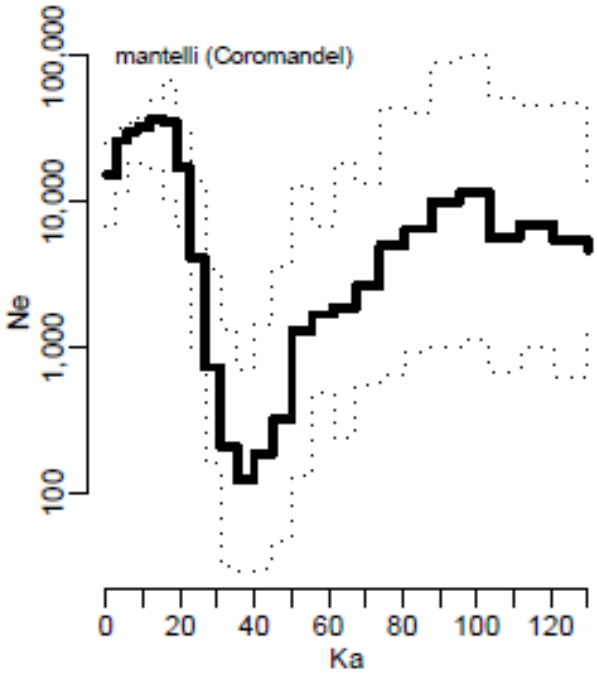
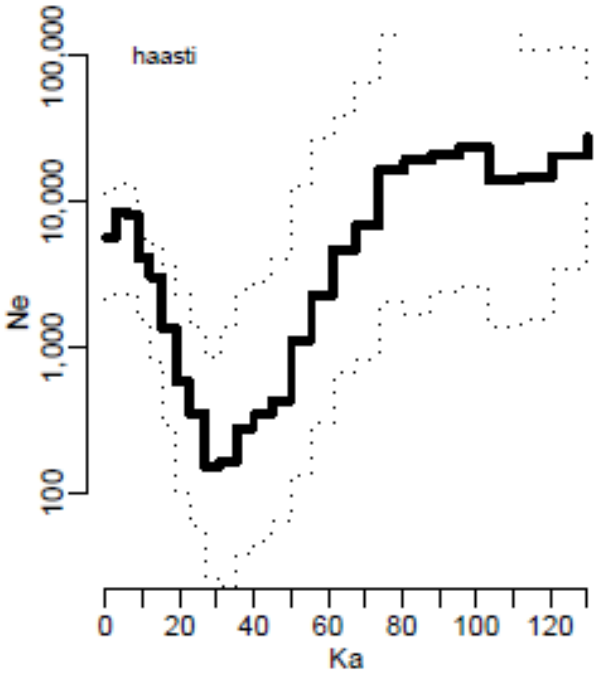
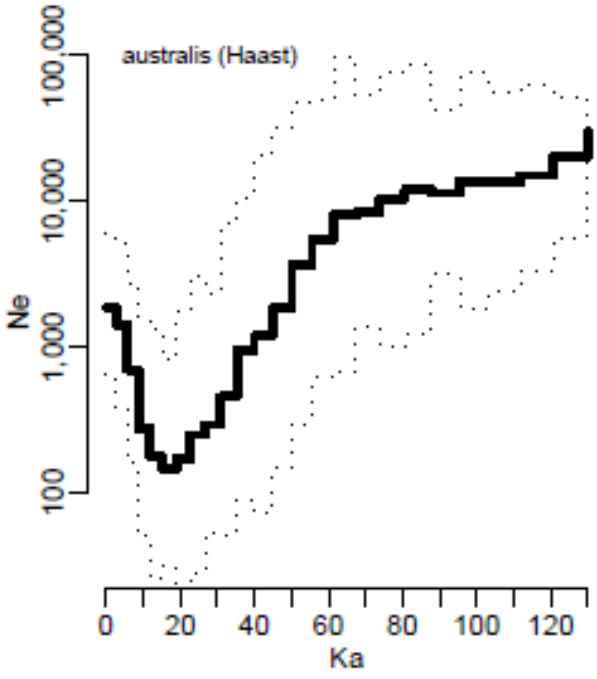

MYC
PCBD1
RHO
TGFB2
TPM1
0.05
0.10
0.15
0.00
0.20
0.02
0.06
0.04
0.08
0.12
0.14

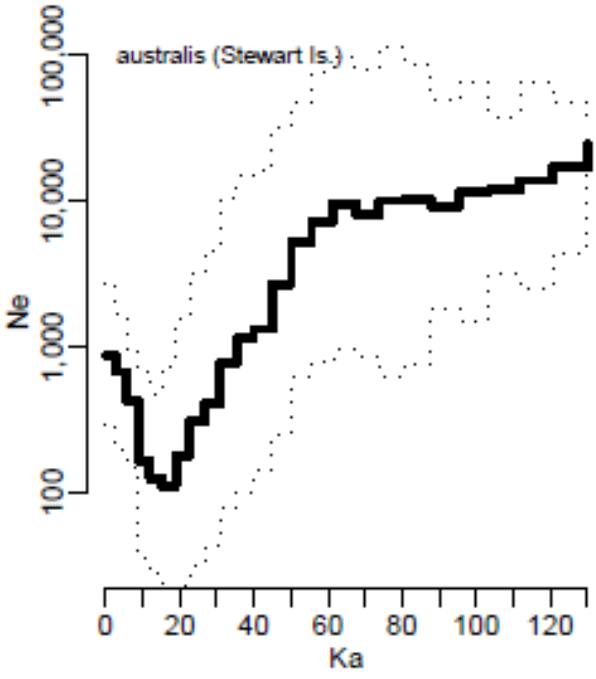
Figure S8

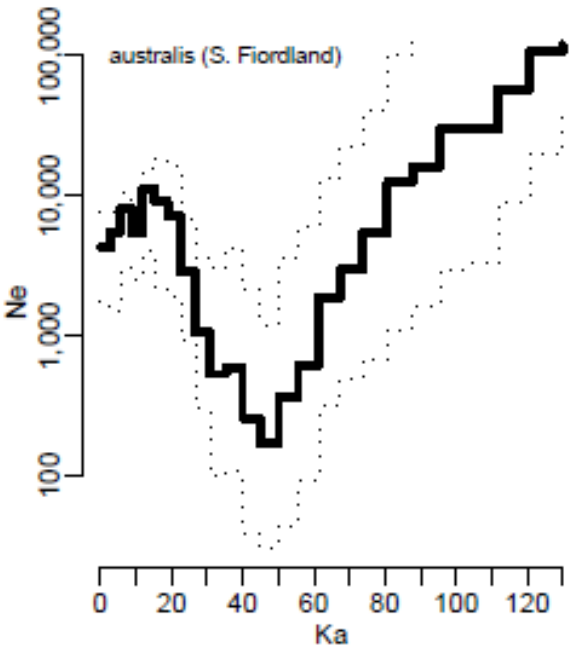

rowi
mantelli (Taranaki)
mantelli (Northland)
mantelli (Eastern)
Ne
100
1,000
10,000
100,000
0
20
40
60
80
100
120
Ka

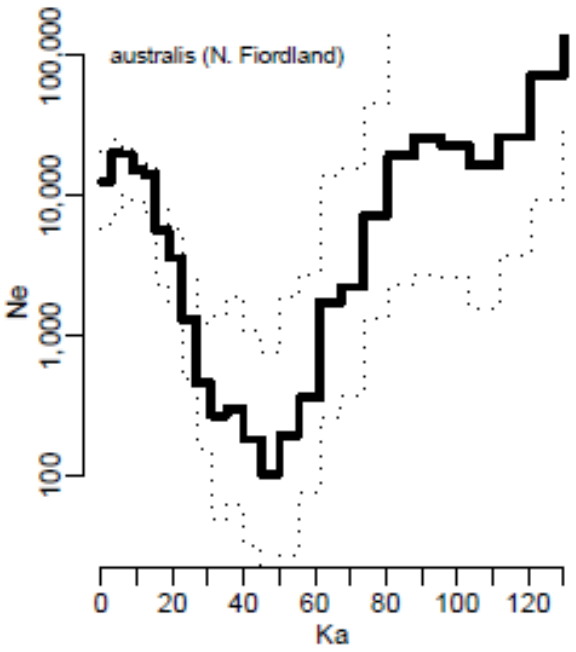

mantelli (Coromandel)
Ne
100
1,000
10,000
100,000
0
20
40
60
80
100
120
Ka


haasti
Ne
100
1,000
10,000
100,000
0
20
40
60
80
100
120
Ka


australis (Haast)
Ne
100
1,000
10,000
100,000
0
20
40
60
80
100
120
Ka


australis (Stewart Is.)
Ne
100
1,000
10,000
100,000
0
20
40
60
80
100
120
Ka


australis (S. Fiordland)
Ne
100
1,000
10,000
100,000
0
20
40
60
80
100
120
Ka


australis (N. Fiordland)
Ne
100
1,000
10,000
100,000
0
20
40
60
80
100
120
Ka

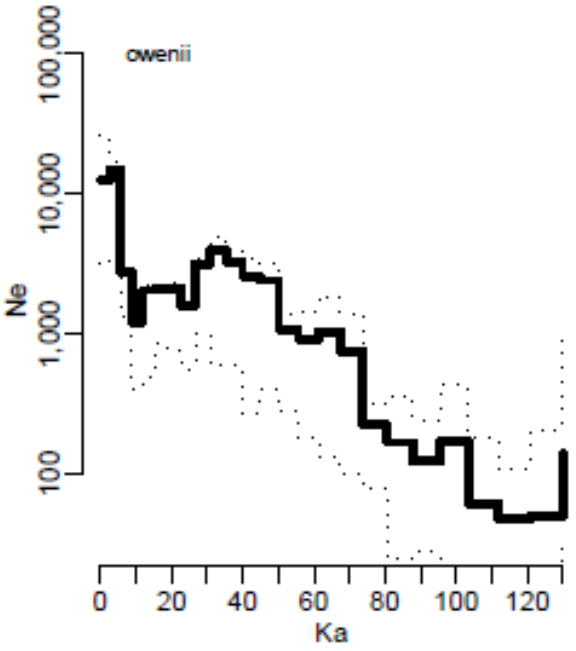

owenii
Ne
100
1,000
10,000
100,000
0
20
40
60
80
100
120
Ka

Table S1. Pairwise Hudson's Fst (below diagonal with 95% CI from 500 bootstrap replicates), JC69 genetic distances (above diagonal) and average JC69 distances within lineages (diagonal).

	A. australis **(N. Fiordland)**	***A. australis*** **(S. Fiordland)**	***A. australis*** **(Haast)**	***A. australis*** **(Stewart Is.)**	***A. mantelli*** **(Eastern)**	***A. mantelli*** **(Northland)**	***A. mantelli*** **(Taranaki)**	***A. mantelli*** **(Cormandel)**	***A. rowi***	***A. owenii***	***A. haastii***
A. australis **(N. Fiordland)**	0.000782	0.001271	0.001359	0.001593	0.003469	0.003459	0.003477	0.003422	0.003432	0.007298	0.007294
A. australis **(S. Fiordland)**	0.213 (0.192 - 0.236)	0.001107	0.001525	0.001533	0.003429	0.003414	0.003426	0.003371	0.003407	0.007275	0.007289
A. australis **(Haast)**	0.472 (0.444 - 0.501)	0.372 (0.343 - 0.403)	0.000695	0.001729	0.003519	0.003198	0.003526	0.003476	0.003483	0.007388	0.0074
A. australis **(Stewart Is.)**	0.500 (0.465 - 0.531)	0.490 (0.458 - 0.519)	0.627 (0.599 - 0.651)	0.000798	0.003422	0.003418	0.003416	0.003368	0.003386	0.007277	0.007286
A. mantelli **(Eastern)**	0.712 (0.694 - 0.728)	0.681 (0.660 - 0.701)	0.719 (0.701 - 0.737)	0.769 (0.752 - 0.785)	0.00116	0.001278	0.001303	0.001255	0.00271	0.007586	0.007601
A. mantelli **(Northland)**	0.733 (0.712 - 0.750)	0.704 (0.685 - 0.722)	0.741 (0.724 - 0.760)	0.788 (0.773 - 0.803)	0.159 (0.138 - 0.180)	0.001047	0.00124	0.001137	0.002724	0.007589	0.007607
A. mantelli **(Taranaki)**	0.711 (0.691 - 0.728)	0.681 (0.663 - 0.700)	0.717 (0.699 - 0.737)	0.766 (0.750 - 0.783)	0.118 (0.103 - 0.136)	0.139 (0.120 - 0.158)	0.001135	0.001188	0.002722	0.007584	0.0076
A. mantelli **(Cormandel)**	0.735 (0.715 - 0.752)	0.705 (0.685 - 0.721)	0.742 (0.726 - 0.759)	0.789 (0.773 - 0.805)	0.197 (0.177 - 0.220)	0.163 (0.143 - 0.187)	0.161 (0.141 - 0.183)	0.000966	0.002703	0.007549	0.00757
A. rowi	0.770 (0.753 - 0.786)	0.738 (0.718 - 0.756)	0.777 (0.759 - 0.796)	0.826 (0.808 - 0.841)	0.667 (0.644 - 0.688)	0.697 (0.674 - 0.718)	0.669 (0.646 - 0.689)	0.701 (0.679 - 0.721)	0.000627	0.007593	0.007603
A. owenii	0.913 (0.904 - 0.921)	0.897 (0.887 - 0.905)	0.915 (0.907 - 0.923)	0.940 (0.932 - 0.946)	0.901 (0.893 - 0.910)	0.908 (0.899 - 0.916)	0.896 (0.887 - 0.904)	0.909 (0.899 - 0.917)	0.934 (0.927 - 0.941)	0.000272	0.002211
A. haastii	0.879 (0.870 - 0.887)	0.863 (0.853 - 0.872)	0.881 (0.871 - 0.889)	0.906 (0.897 - 0.914)	0.866 (0.856 - 0.875)	0.875 (0.867 - 0.885)	0.863 (0.852 - 0.872)	0.875 (0.864 - 0.884)	0.899 (0.890 - 0.907)	0.750 (0.723 - 0.773)	0.00084

Table S2 Fossil calibration used in dating the avian tree. All calibrations use a uniform prior and all but two use an upper bound of 133 Mya (upper limit set by oldest estimated date in Haddrath & Baker 2012)

Node	Calibration	Constraint	Published Source
Acanthisitta chloris/Taeniopygia guttata	Middle Cretaceous, greater than 85 Mya	Geological constraint, beginning of separation between New Zealand and Gondwana	this paper
Melopsittacus undulatus/Nestor notabilis	Middle Cretaceous, greater than 85 Mya	Geological constraint, beginning of separation between New Zealand and Gondwana	this paper
Tyto alba/(Apaloderma vittatum, Picoides pubescens, Buceros rhinoceros silvestris)	Late Paleocene, greater than 56.8 Mya	Fossil constraint, terminal branch *Tyto*, see Jarvis *et al.* 2014 and references therein	Jarvis *et al*. 2014; node 24
Apaloderma vittatum/(Picoides pubescens, Buceros rhinoceros silvestris)	Late Paleocene, greater than 55.5Mya	Fossil constraint, terminal branch *Apaloderma*, see Jarvis *et al.* 2014 and references therein	Jarvis *et al*. 2014; node 29
Picoides pubescens/Buceros rhinoceros silvestris	Lower Middle Eocene, greater than 47 Mya	Fossil constraint, terminal branch *Buceros*, see Jarvis *et al.* 2014 and references therein	Jarvis *et al*. 2014; node 31
Aptenodytes forsteri/Pygoscelis adeliae	Late Early Paleocene, greater than 60.5 Mya but less than 81 Mya	Fossil constraint, internal branch *Pygoscelis+Aptenodytes*, see Jarvis *et al.* 2014 and references therein	Jarvis *et al*. 2014; node 21
Phalacrocorax carbo/(Egretta garzetta, Nipponia nippon)	Lower Middle Eocene, greater than 47 Mya	Fossil constraint, terminal branch *Plegadis*+sister, see Jarvis *et al*. 2014 and references therein	Jarvis *et al*. 2014; node 23
Egretta garzetta/Nipponia nippon	Late Oligocene, greater than 25.3 Mya	Fossil constraint, internal branch *Egretta*+sister, see jarvis et al. 2014 and references therein	Jarvis *et al*. 2014; node 22
Jacana jacana/Rostratula australis	Earliest Oligocene, greater than 32 Mya	Fossil constraint, oldest fossil identified within the Jacanidae, see Haddrath & Baker 2012 and references therein	Haddrath & Baker 2012
Antrostomus carolinensis/(Chaetura pelagica, Calypte anna)	Earliest Eocene, greater than 55.5 Mya	Fossil constraint, internal branch *Chaetura+Calypte*, see Jarvis et al. 2014 and references therein	Jarvis *et al*. 2014; node 34
Chaetura pelagica/Calypte anna	Early Oligocene, greater than 32 Mya	Fossil constraint, terminal branch *Calypte*, see Jarvis et al. 2014 and references therein	Jarvis *et al*. 2014; node 35
Chlamydotis macqueenii/Tauraco erythrolophus	Early Oligocene, greater than 32 Mya	Fossil constraint, terminal branch *Turaco*, see Jarvis et al. 2014 and references therein	Jarvis *et al*. 2014; node 14

Phoenicopterus ruber ruber/(Mesitornis unicolor, Pterocles gutturalis)	Early Oligocene, greater than 32 Mya	Fossil constraint, internal branch *Phoenicopterus+Podiceps*, see Jarvis et al. 2014 and references therein	Jarvis *et al*. 2014; node 9
Mesitornis unicolor/Pterocles gutturalis	Late Oligocene, greater than 24.7 Mya	Fossil constraint, terminal branch *Pterocles*, see Jarvis et al. 2014 and references therein	Jarvis *et al*. 2014; node 7
Gallus gallus/(Anser sp., Anseranas semipalmata)	Late Creatceous, greater than 66 Mya but less than 86 Mya	Fossil constraint, stem Anseriformes, see Jarvis et al. 2014 and references therein	Jarvis *et al*. 2014; node 4
Dromaius novaehollandiae/Casuarius casuarius	Late Oligocene, greater than 25 Mya	Fossil constraint, terminal branch *Dromaius*, see Haddrath & Baker 2012 and references therein	Haddrath & Baker 2012
root	Late Creatceous, greater than 66 Mya but less than 133 Mya	Fossil constraint, internal branch Neognathae, see Jarvis et al. 2014 and references therein	Jarvis *et al*. 2014; node 4

Table S3 Lineage divergence times estimated from two models (with and without migration) fit to each of two datasets (with and without mtDNA) using G-PHOCS. Each line shows the uncalibrated parameter value (scaled by the mutation rate x 10^4) and its posterior 95% credible interval, followed by the time calibrated value (in units of millions of years ago) based on a date of 5.96 Ma for τ_{coal}.

		SNPs + mtDNA		SNPs only	
parameter	**mean prior**[1]	**migration**	**no migration**	**migration**	**no migration**
τ_{coal}[2]	NA	37.52 5.96	37.52 5.96	38.08 5.96	38.12 5.96
τ_{root}	70.37	24.22 (22.46 – 26.22) 3.85 (3.58 – 4.17)	24.34 (22.57 – 26.27) 3.86 (3.59 – 4.18)	23.84 (21.94 – 25.93) 3.74 (3.49 – 4.12)	23.95 (22.08 – 26.09) 3.75 (3.51 – 4.15)
τ_2	29.80	9.84 (9.12 – 10.54) 1.56 (1.45-1.68)	9.84 (9.14 – 10.55) 1.56 (1.46 – 1.68)	9.96 (9.31 – 10.63) 1.56 (1.48 – 1.69)	10.03 (9.39 – 10.70) 1.56 (1.49 – 1.70)
τ_3	22.21	7.02 (6.42 – 7.62) 1.12 (1.02 – 1.21)	7.19 (6.61 – 7.79) 1.14 (1.05 – 1.23)	6.12 (5.45 – 6.87) 0.99 (0.86 – 1.07)	6.38 (5.89 – 6.93) 1.0 (0.94 – 1.11)
τ_4	7.40	1.63 (1.42 – 1.95) 0.26 (0.22 – 0.31)	1.56 (1.38 – 1.74) 0.24 (0.22 – 0.28)	1.78 (1.36 – 2.34) 0.28 (0.21 – 0.37)	1.54 (1.35 – 1.71) 0.24 (0.21 – 0.28)
τ_5	7.19	1.05 (0.74 – 1.41) 0.17 (0.12 – 0.22)	1.05 (0.80 – 1.31) 0.17 (0.13 – 0.21)	0.87 (0.60 – 1.23) 0.14 (0.10 – 0.19)	0.95 (0.70 – 1.35) 0.15 (0.11 – 0.21)
τ_6	6.33	0.68 (0.46 – 0.94) 0.11	0.75 (0.52 – 0.99) 0.12	0.58 (0.39 – 0.83) 0.10	0.62 (0.39 – 0.89) 0.10

		(0.07 – 0.15)	(0.09 – 0.16)	(0.06 – 0.13)	(0.06 – 0.14)
τ_7	21.21*	4.54 (2.46 – 6.73) 0.72 (0.39 – 1.07)	5.07 (2.68 – 7.13) 0.81 (0.43 – 1.14)		
τ_8	20.21*	2.39 (1.38 – 3.48) 0.38 (0.22 – 0.55)	2.39 (1.29 – 3.63) 0.38 (0.20 – 0.57)		
τ_9	12.96*	3.43 (1.68 – 8.10) 0.54 (0.27 – 1.29)	3.79 (1.83 – 7.42) 0.61 (0.29 – 1.18)		
τ_{10}	11.96	1.56 (1.29 – 1.99) 0.24 (0.20 – 0.32)	1.48 (1.12 – 1.92) 0.23 (0.18 – 0.31)	1.55 (1.27 – 1.89) 0.24 (0.20 – 0.30)	1.50 (1.22 – 1.76) 0.23 (0.19 – 0.28)
τ_{11}	10.96*	1.43 (1.18 – 1.89) 0.22 (0.19 – 0.30)	1.34 (0.91 – 1.85) 0.21 (0.15 – 0.30)		
τ_{12}	10.12	1.28 (1.07 – 1.54) 0.20 (0.17 – 0.24)	0.72 (0.52 – 0.95) 0.12 (0.09 – 0.15)	1.14 (0.89 – 1.43) 0.18 (0.14 – 0.22)	0.75 (0.54 – 1.08) 0.12 (0.09 – 0.17)
τ_{13}	7.99	1.22 (0.99 – 1.50) 0.19 (0.16 – 0.23)	0.37 (0.22 – 0.56) 0.06 (0.03 – 0.09)	1.04 (0.76 – 1.36) 0.16 (0.12 – 0.21)	0.36 (0.19 – 0.54) 0.05 (0.03 – 0.09)
τ_{14}	19.33	3.55 (3.03 – 4.60)	3.57 (3.09 – 4.70)	3.48 (3.26 – 4.19)	3.50 (3.25 – 4.21)

		0.56 (0.46 – 0.69)	0.56 (0.46 – 0.70)	0.54 (0.49 – 0.63)	0.54 (0.49 – 0.63)
τ_{15}	18.33*	2.51 (1.30 – 3.82)	2.54 (1.22 – 3.96)		
		0.40 (0.20 – 0.61)	0.40 (0.19 – 0.63)		
τ_{16}	18.33*	1.92 (0.94 – 2.89)	2.00 (1.16 – 2.97)		
		0.31 (0.15 – 0.46)	0.32 (0.18– 0.47)		

*nodes subtended by a fossil lineages (only mtDNA) for which mean prior was set manually as follows: $\tau_{16} = \tau_{14} - 1$, $\tau_{15} = \tau_{14} - 1$, $\tau_{11} = \tau_{1} = - 1$, $\tau_{9} = \tau_{10} + 1$, $\tau_{5} = \tau_{3} - 2$, $\tau_{4} = \tau_{3} - 1$.

[1] The mean of the gamma distribution for each prior

[2] τ_{coal} is not a parameter of the model but is calculated as $\tau_{root} + ½\ \Theta_{root}$.

Table S4 Net diversification rates calculated for different periods. Calculation of net rates and their 95% confidence intervals (CI) follow methods in (26), with lower CI set to 0 for negative values.

Time period	Time bin	Net Rate	Low 95%CI	Upper 95%CI
Pliocene	2.6 to 5.3	0.25	0	0.75
Early Pleistocene	0.78 to 2.6	0.32	0.01	0.62
Middle and Late Pleistocene	0 to 0.78	1.75	1.48	2.01

Table S5

Taxon	Sequence Source for 10 protein-coding genes[1]	Sequence Source for 6 largely intronic gene sequences[2]
Acanthisitta chloris	Haddrath & Baker 2012	Jarvis et al. 2014
Taeniopygia guttata	Haddrath & Baker 2012	Haddrath & Baker 2012
Gavia immer	Haddrath & Baker 2012	Harshman et al. 2008
Aptenodytes forsteri	Haddrath & Baker 2012	Jarvis et al. 2014
Pygoscelis adeliae	Jarvis et al. 2014	Jarvis et al. 2014
Oceanites oceanicus	Haddrath & Baker 2012	Hackett et al. 2008
Haematopus ater / H. ostralegus	Haddrath & Baker 2012	Hackett et al. 2008
Jacana jacana	Haddrath & Baker 2012	Hackett et al. 2008
Rostratula australis	Haddrath & Baker 2012	Hackett et al. 2008
Chaetura pelagica	Haddrath & Baker 2012	Jarvis et al. 2014
Anser albifrons / A. erythropus	Haddrath & Baker 2012	Hackett et al. 2008
Anseranas semipalmata	Haddrath & Baker 2012	Hackett et al. 2008
Gallus gallus	Haddrath & Baker 2012	Harshman et al. 2008
Eudromia elegans	Haddrath & Baker 2012	Harshman et al. 2008
Nothoprocta perdicari	Haddrath & Baker 2012	Harshman et al. 2008
Struthio camelus	Haddrath & Baker 2012	Harshman et al. 2008
Rhea americana	Haddrath & Baker 2012	Harshman et al. 2008
Pterocnemia pennata	Haddrath & Baker 2012	Harshman et al. 2008
Apteryx haastii	Haddrath & Baker 2012	KX668224,KX668229,KX668234, KX668239,KX668244 & KX668249
Apteryx owenii	Haddrath & Baker 2012	KX668225,KX668230,KX668235 KX668240,KX668245 & KX668250
Apteryx mantelli	Haddrath & Baker 2012	KX668226,KX668231,KX668236, KX668241,KX668246 & KX668251
Apteryx rowi	Haddrath & Baker 2012	KX668227,KX668232,KX668237, KX668242,KX668247 & KX668252
Apteryx australis	Haddrath & Baker 2012	KX668228,KX668233,KX668238, KX668243,KX668248 & KX668253
Casuarius casuarius	Haddrath & Baker 2012	Harshman et al. 2008
Dromaius novaehollandiae	Haddrath & Baker 2012	Harshman et al. 2008
Anomalopteryx didiformis	Haddrath & Baker 2012	Haddrath & Baker 2012
Melopsittacus undulatus	Jarvis et al. 2014	Jarvis et al. 2014
Nestor notabilis	Jarvis et al. 2014	Jarvis et al. 2014
Picoides pubescens	Jarvis et al. 2014	Jarvis et al. 2014
Buceros rhinoceros silvestris	Jarvis et al. 2014	Jarvis et al. 2014
Apaloderma vittatum	Jarvis et al. 2014	Jarvis et al. 2014
Tyto alba	Jarvis et al. 2014	Jarvis et al. 2014
Egretta garzetta	Jarvis et al. 2014	Jarvis et al. 2014
Nipponia nippon	Jarvis et al. 2014	Jarvis et al. 2014
Phalacrocorax carbo	Jarvis et al. 2014	Jarvis et al. 2014

Fulmarus glacialis	Jarvis et al. 2014	Jarvis et al. 2014
Charadrius vociferus	Jarvis et al. 2014	Jarvis et al. 2014
Balearica regulorum gibbericeps	Jarvis et al. 2014	Jarvis et al. 2014
Calypte anna	Jarvis et al. 2014	Jarvis et al. 2014
Antrostomus carolinensis	Jarvis et al. 2014	Jarvis et al. 2014
Chlamydotis macqueenii	Jarvis et al. 2014	Jarvis et al. 2014
Tauraco erythrolophus	Jarvis et al. 2014	Jarvis et al. 2014
Mesitornis unicolor	Jarvis et al. 2014	Jarvis et al. 2014
Pterocles gutturalis	Jarvis et al. 2014	Jarvis et al. 2014
Phoenicopterus ruber ruber	Jarvis et al. 2014	Jarvis et al. 2014

1. 10 protein-coding gene sequences: BACH1, BMP2, CMOS, DNAH3, IRBP, NT3, PNN, PTPN12, RAG2, TRAF6
2. Six largely intronic gene sequences: ALDOB, FGB67, GH1, IRF2, MB, MUSK