ELECTRONIC SUPPLEMENTARY MATERIAL

Convergent evolution of "creepers" in the Hawaiian honeycreeper radiation

Dawn M. Reding, Jeffrey T. Foster, Helen F. James, H. Douglas Pratt, and Robert C. Fleischer

1) Methods

(a) Samples

We analyzed 19 individuals belonging to 11 relevant honeycreeper taxa and one outgroup. Common and scientific names of taxa sampled and respective sample sizes (*n*) are as follows: Hawaii creeper, *Oreomystis mana* (3); Kauai creeper, *O. bairdi* (2); Maui alauahio, *Paroreomyza montana* (2); akepa, *Loxops coccineus* (3); akekee, *L. caeruleirostris* (2); Hawaii amakihi, *Hemignathus virens* (1); Kauai amakihi, *H. kauaiensis* (1); Laysan finch, *Telespiza cantans* (1); palila, *Loxioides bailleui* (1); apapane, *Himatione sanguinea* (1); iiwi, *Vestiaria coccinea* (1); and house finch, *Carpodacus mexicanus* (1). See appendix S1 for information on provenance of samples. Samples consisted of muscle tissue or blood, and genomic DNA was extracted with DNeasy (Qiagen, Valencia, CA) tissue kits following manufacturer's protocols.

(b) Laboratory analysis

A total of approximately 2500 bp of sequence data were generated from four nuclear introns and one exon. Intron XI from the glyceraldehydes-3-phosphate dehydrogenase (Gapd) gene, intron III from the lamin (Lam) gene, an intron of the lactate dehydrogenase (Lac) gene, and an intron of ribosomal protein 40 (RP40) were amplified and sequenced using primers described in Friesen *et al.* (1997, 1999). The intron dataset consisted of a total of 1496 bp: 326 bp Gapd, 605 bp Lac, 209 bp Lam, and 348 bp RP40. A portion of the large RAG-1 gene was amplified using two non-

overlapping primer sets: RAG1-3L (GAGAAAGAAGAGGGCGGTGA) and RAG1-7R (TGTGAAAGAAAAGCGAACAGC); and RAG1-6L (GCTGAAAATCTGGAGCGATA) and RAG1-1R (CATGAGGATCGCCACACTG). The RAG-1 dataset consisted of a total of 977 bp.

Products were amplified using polymerase chain reaction (PCR) carried out in 25 µl reactions containing 1X buffer (Applied Biosystems, Foster City, CA), 2 mM $MgCl_2$, 200 µM each dNTP, 0.5 µM each primer, 0.8 mg/ml BSA, 1 U Amplitaq Gold DNA polymerase (Applied Biosystems), and 25-100 ng DNA. Amplifications were carried out with 35 cycles and an annealing temperature of 50°C for RAG, 52°C for Gapd, 55°C for Lam and Lac, and 64°C for RP40. PCR products were cleaned using 96-well Qiaquick (Qiagen) kits following manufacturer's protocols. Sequencing reactions were performed using BigDye 3.1 terminator chemistry (Applied Biosystems) and were run on an ABI 3100 capillary-based DNA sequencer. Both directions were sequenced using the same primers as used for PCR.

(c) Phylogenetic analyses

Sequences were edited, assembled, and aligned using program Sequencher 4.6 (Gene Codes Corporation, Ann Arbor, MI) and deposited in GenBank under accession numbers FJ266094-FJ266110 (Gapd), FJ266111-FJ266128 (Lam), FJ266129-FJ266166 (RAG-1), FJ266167-FJ266185 (RP40), and FJ266278-FJ266296 (Lac). The RAG-1 and intron datasets were analyzed separately as well as concatenated. A homogeneity partition test (ILD, Farris *et al.* 1995) with heuristic search, as implemented in program PAUP* 4.0b10 (Swofford 2002), was also conducted to evaluate the congruence of phylogenetic signal between the two sets of sequences.

Phylogenies were reconstructed using three approaches: maximum parsimony (MP), maximum likelihood (ML), and Bayesian inference. For MP, we conducted an unweighted analysis with a heuristic search, TBR branch-swapping, 10 random-addition repetitions, and gaps treated as missing data using program PAUP*. To indicate nodal support, we conducted a subsequent bootstrap analysis with 1000 replications. For the ML approach, we first used Akaike's information criterion (AIC) in program MODELTEST (Posada & Crandall 1998) to select the most appropriate model of sequence evolution. The K81uf model was selected for the RAG-1 dataset, the HKY+I model was selected for the introns, and the TrN+I model was selected for the combined dataset. The ML analysis involved a heuristic search as described above and a bootstrap of 1000 heuristic repetitions in PAUP*. Bayesian analyses were run in MrBayes 3.1.2 (Ronquist & Huelsenbeck 2003). Since MrBayes does not implement all models available in MODELTEST, we used program MrModelTest 2.3 (Nylander 2004) and AIC to select the best model (HKY+I) for the entire dataset. We also partitioned the combined dataset into two sections, allowing for the RAG-1 and intron data to assume different models of evolution: GTR and HKY+I, respectively. Bayesian analyses consisted of 2 runs of 4 chains of 1,000,000 generations MCMC. Sampling frequency was set at 100 generations, and the first 2500 samples (25%) were excluded as burn-in. All trees were rooted using the House Finch as the outgroup.

(d) Hypothesis testing

To evaluate the strength of the evidence for placement of the Hawaii creeper with akepa/amakihi vs. Kauai creeper, constraints were subsequently applied in the MP and ML methods to force the alternate topology, and analyses were repeated as described in the previous section. The scores

of the best ML trees obtained from the unconstrained vs. constrained analyses were then compared with a one-tailed Shimodaira-Hasegawa (1999) test using RELL bootstrapping with 1000 replicates, as implemented in PAUP*.

(e) mtDNA analysis

The mtDNA dataset consisted of a total of 1254 bp from 5 different regions: 480 bp cytb in two non-overlapping pieces (190 bp and 290 bp), 345 bp control region, 173 bp ATP8, 132 bp ATP6, and 124 bp ND2. The same samples were used as in the nucDNA analysis, except sequence from a different iiwi (*Vestiaria coccinea*) sample was used for ATP6, ATP8, and cytb. This is unlikely to affect the results, since Iiwi have very low levels of mtDNA variation (Jarvi et al. 2004; Foster et al. 2007).

The smaller cytb fragment was amplified using primers cytb-wow and cytb-2rc, the larger cytb fragment using primers cytb-1 and cytb-2 (Fleischer et al. 2006), control region fragment using primers LGL2 and H417 (Tarr 1995), ATP8 using primers t-lys and A6MNH, ATP6 using primers ATP6.L and ATP6.R, and ND2 using primers L5419-H5578 (Fleischer et al. 2006). Products were amplified using polymerase chain reaction (PCR) carried out in 25 µl reactions containing 1X buffer (Applied Biosystems, Foster City, CA), 2 mM $MgCl_2$, 200 µM each dNTP, 0.5 µM each primer, 1.2 mg/ml BSA, 1 U Amplitaq Gold DNA polymerase (Applied Biosystems), and 25-100 ng DNA. Amplifications were carried out with 40 cycles and an annealing temperature of 50°C for all loci. PCR products were cleaned using 96-well Qiaquick (Qiagen) kits following manufacturer's protocols. Sequencing reactions were performed using

BigDye 3.1 terminator chemistry (Applied Biosystems) and the same primers as used for PCR, and the products were run on an ABI 3100 capillary-based DNA sequencer.

Sequences were edited, assembled, and aligned using program Sequencher 4.6 and deposited in GenBank under accession numbers FJ266186-FJ266204 (ATP6), FJ266205-FJ266221 (ATP8), FJ266222-FJ266240 (control region), FJ266241-FJ266277 (cytb), and FJ266297-FJ266312 (ND2). Phylogenetic and hypothesis testing methods used were as described for the nucDNA analysis, except the HKY+I+G model was used in ML and Bayesian methods.

2) Results

(a) Separate analysis of introns and RAG-1

In the introns dataset, there were 57 parsimony-uninformative and 27 parsimony-informative characters total. Maximum parsimony (MP) analysis of the dataset resulted in 63 equally most parsimonious trees of length 92, and maximum likelihood (ML) analysis resulted in 2 equally most likely trees with a –ln likelihood score of 2770.0. Based on strict consensus, all of these trees placed the Hawaii creeper in a clade with the akepas and amakihis (Figure S1). Constraining the topology such that the Hawaii and Kauai creepers formed a monophyletic group increased the length of the MP tree by 6 steps and lowered the score of the ML tree by 20.7 ($P = 0.028$). Bayesian analysis also showed strong support for placement of Hawaii creeper with the akepas and amakihis, as noted by a high posterior probability value of 0.94 for this clade. MP yielded a 68% bootstrap value for this Hawaii creeper/akepa/amakihi clade, and ML yielded 60%. We also analyzed each intron separately. Individually, they provided little resolution

except for LAC, which showed moderate support for the Hawaii creeper/akepa/amakihi clade (data not shown).

The RAG-1 dataset contained 14 parsimony-uninformative and 5 parsimony-informative characters. Analysis resulted in 24 MP trees 19 steps in length and 2 ML trees of score 1474.2, none of which grouped Hawaii creeper in a monophyletic clade with Kauai creeper (Figure S2). Again, constraints increased tree length by 3 steps and lowered the ML score by 20.2 (P = 0.035). The Bayesian, ML, and MP methods provided little resolution among the different topologies, as seen by overall low posterior probability and bootstrap values. This is likely due to the small number of parsimony-informative sites.

(b) mtDNA analysis

In the mtDNA dataset, there were 83 parsimony-uninformative and 212 parsimony-informative characters. Maximum parsimony (MP) analysis of the dataset resulted in a single most parsimonious tree of length 535, and maximum likelihood (ML) analysis resulted in a single most likely tree with a –ln likelihood score of 4329.1. Both of these trees placed the Hawaii creeper in a clade with the akepas and amakihis (Figure S3). MP yielded a 88% bootstrap value for this Hawaii creeper/akepa/amakihi clade, and ML yielded 92%. Bayesian analysis also showed strong support for placement of the Hawaii creeper with the akepas and amakihis, as noted by a high posterior probability value of 1.0 for this clade. All 3 methods produced a tree of the same topology (Figure S3). Constraining the topology such that the Hawaii and Kauai creepers formed a monophyletic group increased the length of the MP tree by 24 steps and lowered the score of the ML tree by 31.0 (P = 0.003).

3) References

Farris, J. S., Kallersjo, M., Kluge, A. G. & Bult, C. 1995 Testing significance of incongruence. *Cladistics* **10**, 315-319.

Fleischer, R. C., Kirchman, J. J., Dumbacher, J. P., Bevier, L., Dove, C., Rotzel, N. C., Edwards, S. V., Lammertink, M., Miglia, K., & Moore, W. S. 2006 Mid-Pleistocene divergence of Cuban and North American ivory-billed woodpeckers. *Biol. Lett.* **2**, 466-469.

Foster, J. T., Woodworth, B. L., Eggert, L. E., Hart, P. J., Palmer, D., Duffy, D. C., & Fleischer, R. C. 2007. Genetic structure and evolved malaria resistance in Hawaiian honeycreepers. *Mol. Ecol.* **16**, 4738-4746.

Friesen, V. L., Congdon, B. C., Kidd, M. G. & Birt, T. P. 1999 Polymerase chain reaction (PCR) primers for the amplification of five nuclear introns in vertebrates. *Mol. Ecol.* **8**, 2141-2152.

Friesen, V. L., Congdon, B. C., Walsh, H. E. & Birt, T. P. 1997 Intron variation in marbled murrelets detected using analyses of single-stranded conformational polymorphisms. *Mol. Ecol.* **6**, 1047-1058.

Jarvi, S., Tarr, C. McIntosh, C., Atkinson, C., & Fleischer, R. 2004 Natural selection of the major histocompatibility complex (*Mhc*) in Hawaiian Honeycreepers (Drepanidinae). *Mol. Ecol.* **13**, 2157-2168.

Nylander, J. A. A. 2004 MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.

Posada, D. & Crandall, K. A. 1998 Modeltest: testing the model of DNA substitution. *Bioinformatics* **14**, 817-818.

Ronquist, F. & Huelsenbeck, J. P. 2003 MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572-1574.

Shimodaira, H. & Hasegawa, M. 1999 Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol. Biol. Evol*. **16**, 1114-1116.

Swofford, D. L. 2002 PAUP*. Phylogenetic analysis using parsimony (*and other methods), version 4.0. Sunderland, MA: Sinauer Associates.

Tarr, C. L. 1995 Primers for amplification and determination of mitochondrial control-region sequences in oscine passerines. *Mol. Ecol.* **4**, 527–529.

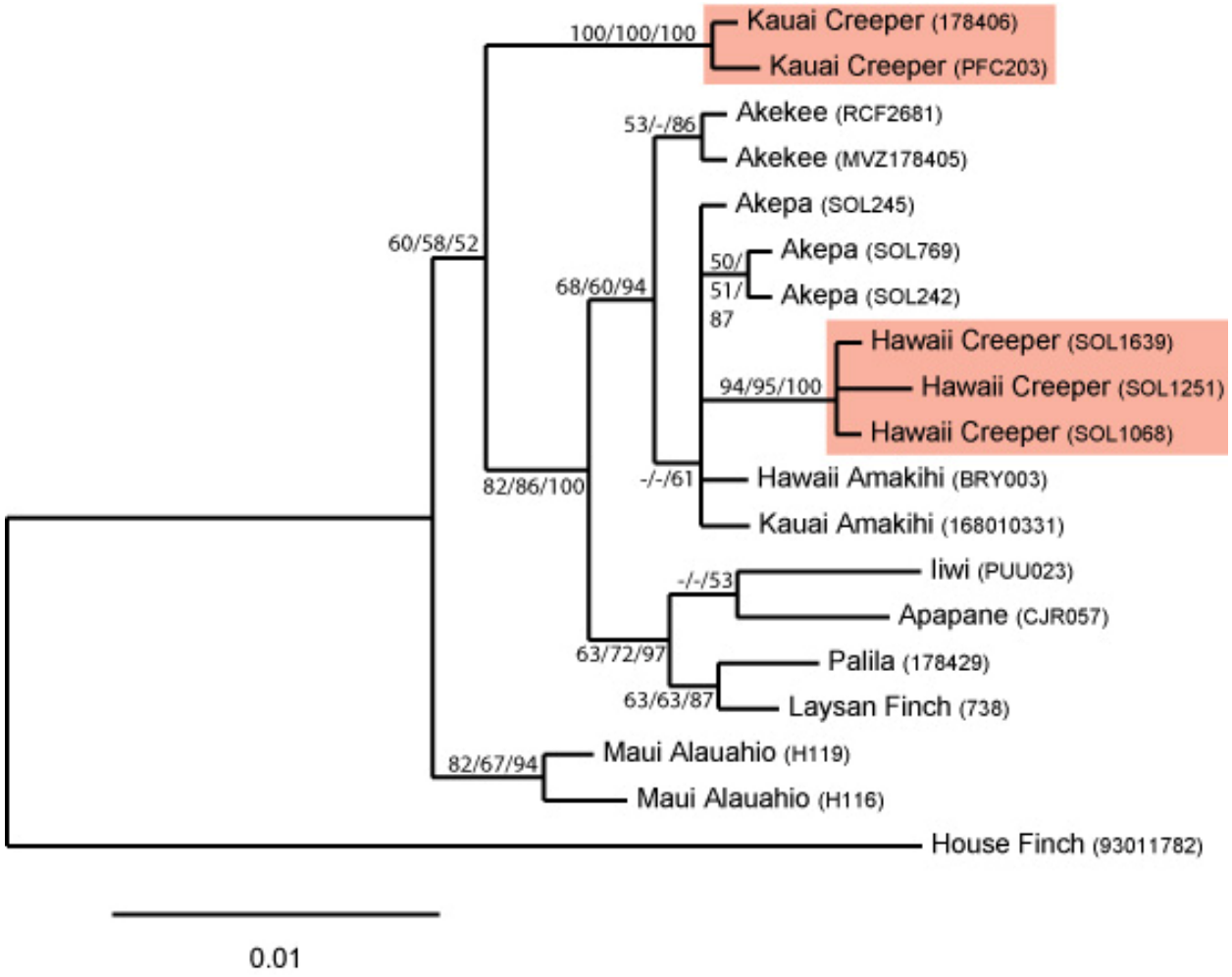

Figure S1. Bayesian consensus topology obtained from analysis of sequences from 4 introns. Bootstrap values from MP and ML analysis (1000 replicates each), and posterior probability values (x100) from Bayesian analysis (15000 trees sampled) are provided at each node. A dash (-) indicates <50% support. Collapsed nodes had low bootstrap and posterior probability values (i.e., <50%). The scale at the bottom indicates 1% divergence along a branch.

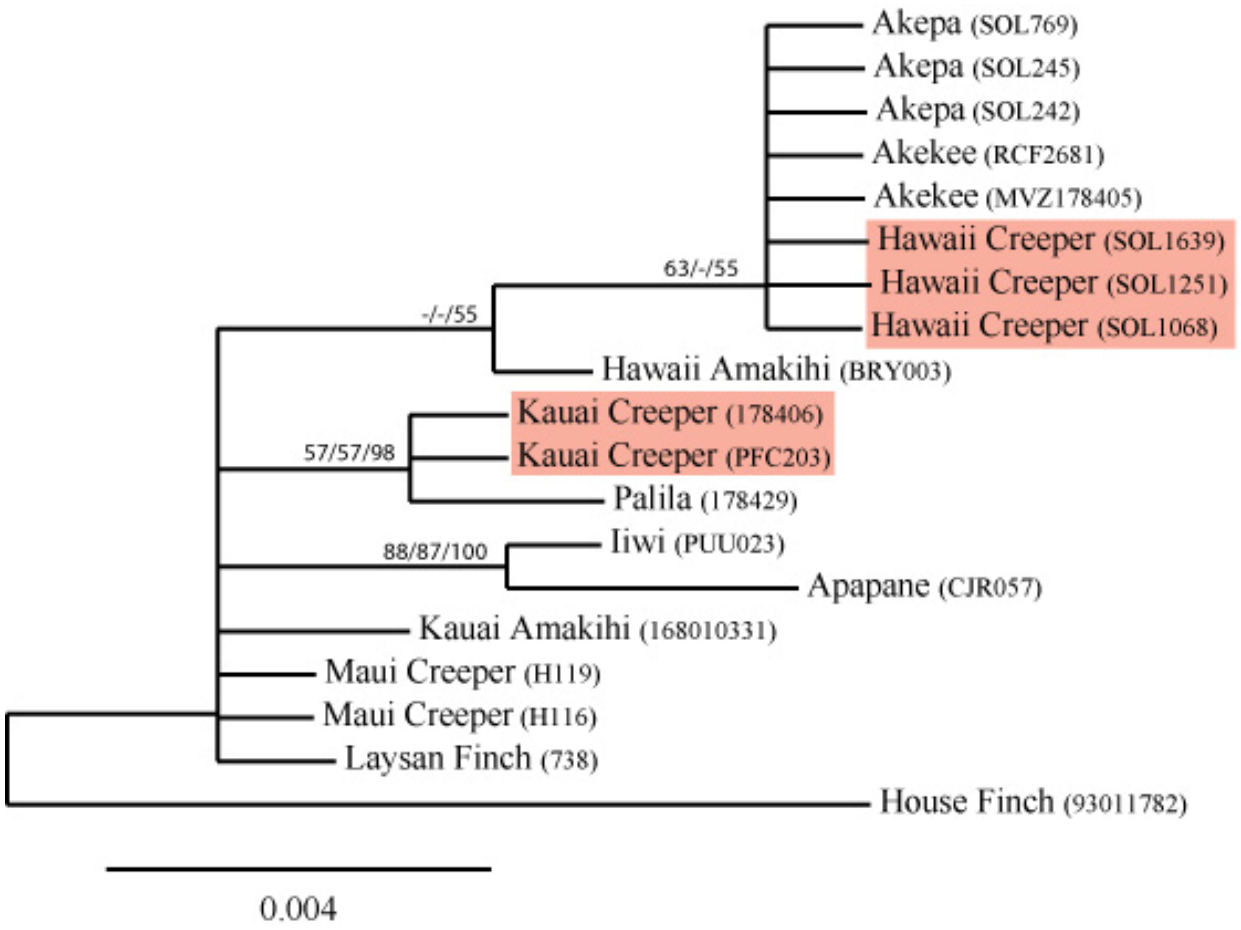

Figure S2. Bayesian consensus topology obtained from analysis of sequences from the RAG exon. Bootstrap values from MP and ML analysis (1000 replicates each), and posterior probability values (x100) from Bayesian analysis (15000 trees sampled) are provided at each node. A dash (-) indicates <50% support. Collapsed nodes had low bootstrap and posterior probability values (i.e., <50%). The scale at the bottom indicates 0.4% divergence along a branch.

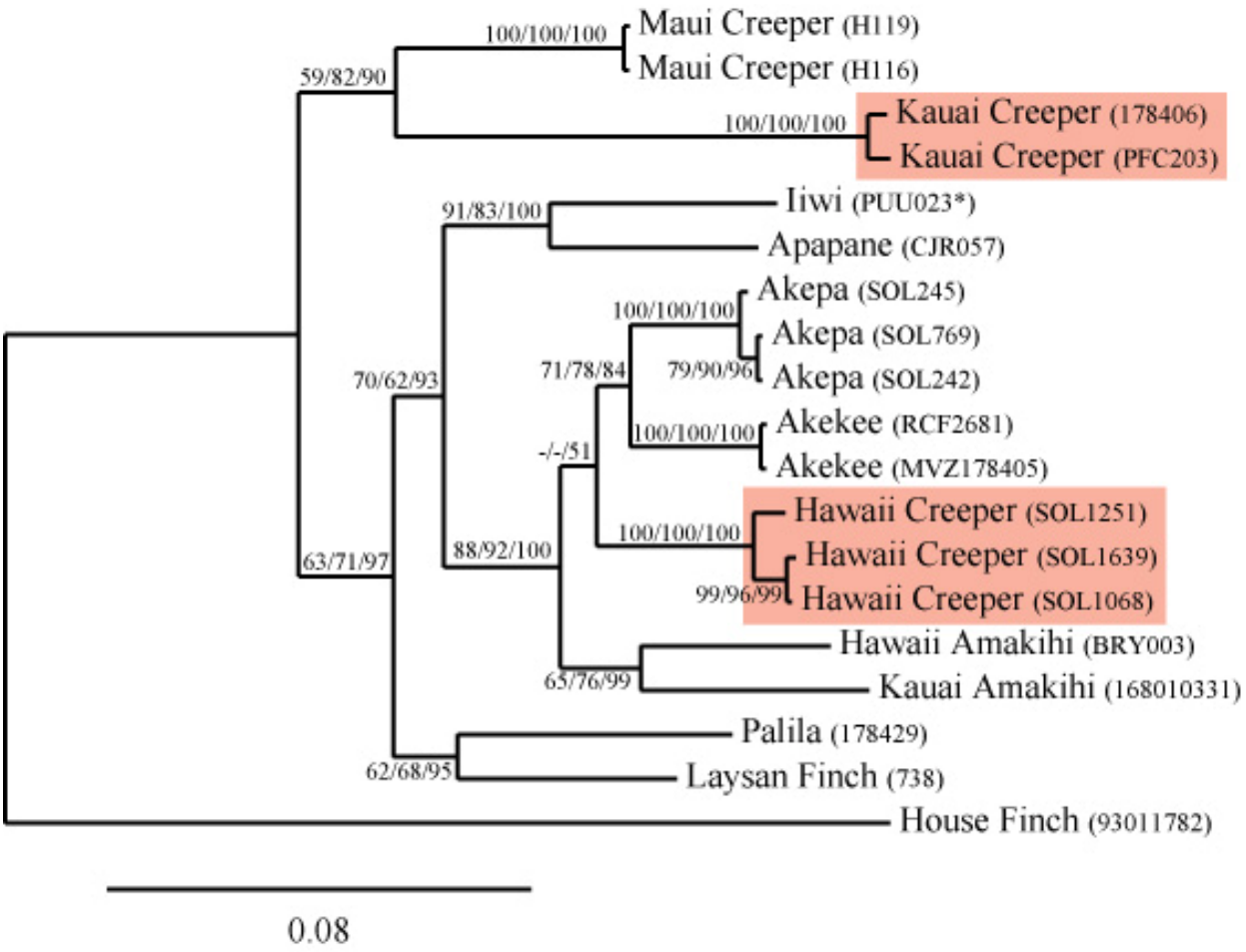

Figure S3. Bayesian consensus topology obtained from analysis of concatenated sequences from cytb, control region, ATP-8, ATP-6, and ND2. The same topology was obtained as the single best MP and ML tree. Bootstrap values from MP and ML analysis (1000 replicates each), and posterior probability values (x100) from Bayesian analysis (15000 trees sampled) are provided at each node. A dash (-) indicates <50% support. The scale at the bottom indicates 8% divergence along a branch. *Also contained sequence from a different iiwi sample for ATP6, ATP8, and cytb.

Appendix S1: List of specimens used in the analyses.

Common Name	Species	Sample ID#	Collection Location
Kauai creeper	*Oreomystis bairdi*	PFC203	Alakai Swamp, Kauai
Kauai creeper	*Oreomystis bairdi*	178406	Alakai Swamp, Kauai
Hawaii creeper	*Oreomystis mana*	SOL1251	Solomons Waterhole, Hawaii
Hawaii creeper	*Oreomystis mana*	SOL1639	Solomons Waterhole, Hawaii
Hawaii creeper	*Oreomystis mana*	SOL1068	Solomons Waterhole, Hawaii
Maui alauahio	*Paroreomyza montana*	H116	Haleakala, Maui
Maui alauahio	*Paroreomyza montana*	H119	Haleakala, Maui
palila	*Loxioides bailleui*	178429	Pu'u La'au, Hawaii
Laysan finch	*Telespiza cantans*	738	Laysan Island
Kauai amakihi	*Hemignathus kauaiensis*	1680-10331	Alakai Swamp, Kauai
Hawaii amakihi	*Hemignathus virens*	BRY003	Bryson's Cinder Cone, Hawaii
akekee	*Loxops caeruleirostris*	RCF 2681	Alakai Swamp, Kauai
akekee	*Loxops caeruleirostris*	MVZ 178405	Kokee State Park, Kauai
akepa	*Loxops coccineus*	SOL245	Solomons Waterhole, Hawaii
akepa	*Loxops coccineus*	SOL769	Solomons Waterhole, Hawaii
akepa	*Loxops coccineus*	SOL242	Solomons Waterhole, Hawaii
apapane	*Himatione sanguinea*	CJR057	C.J. Ralph Site, Hawaii
iiwi	*Vestiaria coccinea*	PUU023	Puu Unit-Olaa Tract, Hawaii
house finch (outgroup)	*Carpodacus mexicanus*	930-11782	Manuka, Hawaii