THE ROYAL SOCIETY

A phylogeny of Darwin's finches based on microsatellite DNA length variation

Kenneth Petren*, B. Rosemary Grant and Peter R. Grant

Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544-1003, USA

Allele length variation at 16 microsatellite loci was used to estimate the phylogeny of 13 out of the 14 species of Darwin's finches. The resulting topology was similar to previous phylogenies based on morphological and allozyme variation. An unexpected result was that genetic divergence among Galápagos Island populations of the warbler finch (*Certhidea olivacea*) predates the radiation of all other Darwin's finches. This deep split is surprising in view of the relatively weak morphological differentiation among *Certhidea* populations and supports the hypothesis that the ancestor of all Darwin's finches was phenotypically similar to *Certhidea*. The results also resolve a biogeographical problem: the Cocos Island finch evolved after the Galápagos finch radiation was under way, supporting the hypothesis that this distant island was colonized from the Galápagos Islands. Monophyletic relationships are supported for both major groups, the ground finches (*Geospiza*) and the tree finches (*Camarhynchus* and *Cactospiza*), although the vegetarian finch (*Platyspiza crassirostris*) appears to have diverged prior to the separation of ground and tree finches. These results demonstrate the use of microsatellites for reconstructing phylogenies of closely related species and interpreting their evolutionary and biogeographic histories.

Keywords: phylogenetic; biogeography; simple sequence repeats; dinucleotide; Cocos; Galápagos

1. INTRODUCTION

Adaptive radiations are a major source of information about the evolutionary origins of biological diversity (Givnish & Sytsma 1997; Grant 1998). Darwin's finches are one of a few classical examples of such radiations (Lack 1947; Grant 1986; Givnish & Sytsma 1997). Species in this group show adaptive variation in beak size, beak shape and body size that is more typical of differences among families of birds (Sushkin 1929), yet the entire radiation is believed to have occurred in less than three million years (Grant 1994). While much has been learned about adaptation and speciation in the group, their phylogenetic relationships remain poorly known. Lack (1947) offered a phylogenetic reconstruction for the group based on a non-quantitative comparison of morphological characteristics (plumage, size and shape, see figure 1*a*; see also Schluter 1984). Yang & Patton (1981) produced a phylogeny from allozyme variation among 11 out of the 14 currently recognized species, but support for the tree finch was limited and the results differed according to the methodology used for analysis (Stern & Grant 1996; figure 1*b*). Variation in mitochondrial (mt) and nuclear DNA sequences appears to be insufficient for resolving relationships among the more closely related members of the group (Freeland 1997; Sato *et al*. 1999*a*).

We have estimated the evolutionary history of Darwin's finches using microsatellite DNA length variation. Microsatellites are multilocus genetic markers with high mutation rates that have been used frequently to test parentage, assess population differentiation and detect hybridization (MacDonald & Potts 1997). Although it has been suggested that allele length polymorphism at these loci may be useful for resolving phylogenetic relationships (Takezaki & Nei 1996; MacDonald & Potts 1997), few interspecific microsatellite phylogenies have been reconstructed to date (Pollock *et al*. 1998; Primmer & Ellegren 1998; but see Roy *et al*. 1994).

We analysed microsatellite length variation among 13 out of the 14 currently recognized species of Darwin's finches including the Cocos Island finch (*Pinaroloxias*). The missing species from the analysis is the rarest, *Cactospiza heliobates* (mangrove finch), but it is extremely similar morphologically to its congener *Cactospiza pallida* (woodpecker finch) (Lack 1947; Grant 1986). Two continental species, *Tiaris olivacea* (yellow-faced grassquit) and *Sporophila aurita* (variable seedeater), are included in the analysis. *Tiaris* is among a small group of emberizines (seedeaters and tanagers) which are believed to be the closest mainland relatives of Darwin's finches (Lack 1947; Baptista & Trail 1988; Sato *et al*. 1999*b*).

2. MATERIALS AND METHODS

(a) *Sampling*

Blood was collected (Petren *et al*. 1999) and standard laboratory protocols were used for DNA extraction, genomic library screening and genotype determination (Sambrook *et al*. 1989; Primmer *et al*. 1995). Specific protocols and primer sequences are available elsewhere (Petren 1998). All loci contained a pure $(CA)_{>13}$ core motif except for two that contained a $(GA)_{>13}$ motif. Eight of these loci have been used to test parentage in 159 *Geospiza scandens* offspring (Petren *et al*. 1999) without detection of a single 'null' allele (Callen *et al*. 1993). Alleles at

*Author for correspondence (petren@princeton.edu).

Received 28 September 1998 *Accepted* 9 November 1998

one locus were inherited in a sex-linked (Z-linked) fashion (Petren *et al*. 1999).

To estimate genetic distances among taxa, we use Nei's (1972) unbiased genetic distance (G_{ST}). We also present results using $(\delta\mu)^2$, a distance measure recently developed for microsatellites (Goldstein *et al*. 1995). G_{ST} makes no assumption about the mechanism of mutation whereas $(\delta\mu)^2$ assumes stepwise mutation. We expect G_{ST} to be better than $(\delta\mu)^2$ at shorter time intervals, while $(\delta\mu)^2$ should perform better than G_{ST} at longer time-scales because it is expected to remain more linear (Goldstein *et al*. 1995).

Populations were grouped into species according to current taxonomic classification, which is based primarily on morphology (Lack 1947; Grant 1986). *Certhidea olivacea* occurs on all of the 17 major islands of the Galápagos (Grant 1986). The six populations analysed here were divided into two groups, *C. olivacea* and *Certhidea fusca*, because the mean G_{ST} between these groups (2.06) was larger than any other distance among Darwin's finch species. The mean G_{ST} distances among *C. olivacea* (0.62) and *C. fusca* (0.52) are comparable to distances among populations of other taxa (below). Following Swarth (1931), we use the name *C. fusca* to refer to the populations from the outer islands, while retaining *C. olivacea* for the central-island populations (Santa Cruz and Santiago). Differentiation among populations is the focus of a separate study (Grant *et al*. 1999). Excluding individual populations from the analysis resulted in only minor rearrangements that do not affect any of our conclusions.

Sample sizes and locations for each species are as follows (abbreviated name, number of populations and mean G_{ST} among populations): *Geospiza fuliginosa* (*G. fu.*, 6, 0.14), Dm-18, Sc-14, So-9, Ra-10, Es-10, Pi-10; *Geospiza fortis* (*G. fo.*, 4, 0.30), Dm-36, Sc-24, Ra-3, So-2; *Geospiza magnirostris* (*G. ma.*, 3, 0.22), Dm-14, Sc-4, Ge-19; *G. scandens* (*G. sc.*, 4, 0.15), Dm-68, Sc-15, Ra-5, So-4; *Geospiza conirostris* (*G. co.*, 2, 0.62), Ge-49, Es-23; *Geospiza difficilis* (*G. di.*, 6, 0.69), So-14, Ge-30, Wo-10, Da-12, Pi-23, Fe-8; *C. pallida* (*C. pl.*, 1), Sc-16; *Platyspiza crassirostris* (*P. cr.*, 4, 0.15), Sc-23, So-3, Ma-7, Pi-20; *Camarhynchus parvulus* (*C. pv.*, 2, 0.03), Sc-11, Fl-22; *Camarhynchus psittacula* (*C. ps.*, 3, 0.10), Sc-5, Ma-3, Pi-8; *Camarhynchus pauper* (*C. pp.*, 1), Fl-19; *Pinaroloxias inornata* (*P. in.*, 1), Co-30; *C. olivacea* (*C. ol.*, 2, 0.62), Sc-13, So-20; *C. fusca* (*C. fu.*, 4, 0.52), Ge-13, Es-20, Ma-8, Pi-10. (Locations: Sc, Santa Cruz; So, Santiago; Dm, Daphne Major; Ge, Genovesa; Pi, Pinta; Fl, Floreana; Wo, Wolf; Co, Cocos; Da, Darwin; Fe, Fernandina; Ra, Rábida; Es, Española; Ma, Marchena.) *Tiaris olivacea* and *S. aurita* were collected in Panama. Genetic distances among species are given in table 1.

(b) *Phylogenetic analysis*

There is no obvious single method of analysis because microsatellite loci have higher mutation rates and a fundamentally different mechanism of mutation than other phylogenetic markers such as nucleotide sequences or allozymes (Goldstein & Pollock 1997). Given uncertainty regarding the mechanisms of mutation, we present phylogenetic reconstructions based on four methods that make different assumptions: UPGMA (Sokal & Sneath 1963), the Fitch–Margoliash least-squares method (Fitch & Margoliash 1967), the Fitch–Margoliash method with contemporaneous taxa (KITCH; Felsenstein 1984, 1993) and maximum likelihood (CONTML; Edwards & Cavalli-Sforza 1964; Felsenstein 1981, 1993).

The FITCH and KITCH methods (with power set to 2.0; Felsenstein 1993) minimize the weighted squared distances along branches of the tree. The UPGMA and KITCH methods assume a molecular clock, so the tips of the tree are constrained to be contemporaneous (Felsenstein 1984, 1993). CONTML is based on a model of Brownian motion diffusion which is a questionable assumption because large mutational jumps are not uncommon in microsatellites (Primmer & Ellegren 1998). Therefore we expect CONTML to be less reliable for comparisons across longer time-scales. Analyses and bootstraps were performed with PHYLIP (Felsenstein 1993). The hypotheses we address do not depend on uncovering a single exact tree topology. Therefore, we discuss the implications of the consensus topology among the different methods and then consider the bootstrap support for alternative topologies that would lead to different conclusions.

(c) *Microsatellite variation*

We examined microsatellite length variation for indications of development bias and homoplasy. Development bias is evident when microsatellite primers developed in one species fail to produce a product or show little or no variation in other species (Ellegren *et al*. 1995). A decrease in allele size is generally accompanied by reduced polymorphism (Primmer *et al*. 1995): smaller repeat regions are known to have substantially reduced mutation rates (Weber & Wong 1993). Reduced mutation rates at longer genetic distances will lead to non-linear distance estimates for both G_{ST} and $(\delta\mu)^2$. Among Darwin's finches, polymorphism remains relatively high and allele size declines only slightly as genetic distance increases from *G. fortis*, the species used for microsatellite development (table 2). Allele size and variation decline more abruptly in the mainland taxa, suggesting that at this time-scale genetic distances may be non-linear.

Homoplasy obscures phylogenetic signal. It occurs when characters scored as the same are identical by convergence and not by common descent. In microsatellites, homoplasy can be caused by mutational length changes in regions flanking the microsatellite repeat (Orti *et al*. 1997). The low frequency of odd-sized alleles (table 2) implies minimal homoplasy due to insertions and deletions in regions flanking the repeat region because insertions and deletions should be equally likely to involve odd and even numbers of bases. In addition, the core repeat region of alleles sequenced in both *G. fortis* and *C. olivacea* were the expected length for all 11 loci tested to date (K. Petren, unpublished data).

Homoplasy may also arise because microsatellite repeat regions are bounded in size (Garza *et al*. 1995; Goldstein & Pollock 1997). These size constraints allow only a finite number of character states and, at longer time intervals, homoplasy is expected to rise and the phylogenetic signal will be obscured as the signal becomes saturated (Takezaki & Nei 1996; Pollock *et al*. 1998). In Darwin's finches, the mean allele size range for each species considered separately (across all loci) is 22.2 bp, while the mean range for all species combined is over twice this figure (46.4 bp). The large difference in allele size range suggests saturation of phylogenetic signal through this type of homoplasy has not been extensive. The wide range of genetic distances (table 1) suggests that genetic distances are not likely to be saturated, at least among the more closely related taxa.

3. RESULTS

(a) *Tree comparisons*

The phylogenetic reconstructions (figure 2) share a number of common elements. Every method supports

Table 1. *Genetic distances among Darwin's finches and two mainland relatives based on microsatellite length variation*
(Below diagonal G_{ST} (Nei 1972; Felsenstein 1993) and above diagonal $(\delta\mu)^2$ (Goldstein *et al.* 1995).)

	G. fu.	*G. fo.*	*G. ma*	*G. sc.*	*G. co.*	*G. di.*	*C. pl.*	*P. cr.*
G. fuliginosa	—	1.4	3.6	4.8	4.5	3.9	10.1	22.7
G. fortis	0.11	—	5.5	3.8	6.9	6.5	13.9	29.4
G. magnirostris	0.25	0.19	—	8.7	5.1	7.4	13.7	23.2
G. scandens	0.35	0.35	0.47	—	12.1	11.7	22.3	33.2
G. conirostris	0.39	0.36	0.35	0.37	—	5.8	5.1	14.6
G. difficilis	0.24	0.38	0.54	0.56	0.56	—	6.7	13.7
C. pallida	0.57	0.60	0.90	0.56	0.61	0.66	—	12.4
P. crassirostris	0.82	1.02	1.26	1.05	1.06	0.91	0.71	—
C. parvulus	0.50	0.56	0.84	0.65	0.60	0.69	0.18	0.83
C. psittacula	0.59	0.69	1.00	0.71	0.62	0.68	0.22	0.85
C. pauper	0.53	0.56	0.88	0.66	0.59	0.66	0.12	0.80
P. inornata	1.08	1.31	1.64	1.16	1.22	0.88	0.99	1.25
C. fusca	0.96	1.02	1.32	1.14	0.90	0.96	1.08	1.32
C. olivacea	1.32	1.35	1.41	1.50	1.52	1.41	2.25	1.73
T. olivacea	2.47	2.42	3.35	2.79	2.48	2.56	2.45	2.26
S. aurita	1.55	1.61	1.74	1.66	1.74	1.73	1.89	2.78

	C. pv.	*C. ps.*	*C. pp.*	*P. in.*	*C. fu.*	*C. ol.*	*T. ol.*	*S. au.*
G. fuliginosa	18.0	23.3	14.1	10.3	10.6	17.5	85.4	131.0
G. fortis	21.8	27.3	16.8	13.1	11.5	20.0	102.3	150.6
G. magnirostris	16.7	23.4	14.0	12.0	15.4	16.8	82.3	131.9
G. scandens	31.7	37.2	26.8	10.4	14.2	31.1	97.7	145.2
G. conirostris	10.5	14.5	7.4	14.4	15.1	12.9	81.6	127.3
G. difficilis	11.4	13.2	9.0	10.4	8.5	16.1	80.3	122.6
C. pallida	7.4	9.2	4.1	16.2	12.0	14.6	72.8	122.1
P. crassirostris	9.9	8.4	11.9	18.9	24.2	21.4	60.1	90.9
C. parvulus	—	1.6	1.4	18.3	15.0	9.0	69.4	100.8
C. psittacula	0.04	—	3.0	22.4	18.1	12.3	74.4	103.5
C. pauper	0.05	0.09	—	17.6	13.1	10.3	75.9	116.1
P. inornata	1.18	1.17	1.05	—	8.3	25.3	67.6	113.2
C. fusca	0.98	1.13	0.94	1.07	—	21.7	84.3	128.7
C. olivacea	2.17	2.32	2.31	1.80	1.81	—	71.5	110.4
T. olivacea	2.55	2.58	2.37	2.14	2.39	1.81	—	44.1
S. aurita	1.85	2.00	1.91	1.78	3.19	3.00	1.69	—

(i) the monophyly of Darwin's finches; (ii) the basal placement of *C. olivacea*; (iii) the non-sister relationship between the two *Certhidea*; (iv) the derivation of the Cocos Island finch (*Pinaroloxias*) from the Galápagos finches; (v) monophyly of the ground finches (*Geospiza*); and (vi) monophyly of the combined tree finch genera *Camarhynchus* and *Cactospiza*. The FITCH tree topology is largely congruent with the KITCH topology. This shows that the conclusions are not dependent upon assuming a molecular clock. The main difference between these two methods lies in the placement of *Platyspiza*, which is basal to the ground finch–tree finch split in all but the FITCH tree. We refer primarily to the UPGMA tree (figure 2), because it captures the elements most commonly observed among the other methodologies and because this method consistently revealed the same topology when subsets of populations and species were analysed.

The microsatellite tree (figure 3) is in general agreement with the morphological tree (figure 1). The sections of the microsatellite tree that disagree most with Lack's (1947) tree occur towards the tips of the branches among the ground finches (*Geospiza*) and the tree finches (*Camarhynchus*) and in the placement of *Platyspiza*. The microsatellite tree also generally agrees with the allozyme tree and the G_{ST} distance matrices are significantly similar by the Mantel (1967) test ($R_M = 0.66$ and $p < 0.005$). However, the microsatellite tree provides higher resolution and greater concordance among different methods of analysis than the allozyme tree (figure 1*b*; Stern & Grant 1996).

(b) *Monophyly of Darwin's finches*

The microsatellite data support monophyly of Darwin's finches not only by showing agreement across all methods of analysis, but also because the patterns of allelic variation (table 2) show a marked change in allele size, variation and amplification ability in *Tiaris* and *Sporophila*. G_{ST} does not incorporate information about allele size or non-amplifying loci which clearly set the outgroup taxa apart. The $(\delta\mu)^2$ distance provides high bootstrap support for monophyly (figure 2*e*). Similarly, $(\delta\mu)^2$ better reflects the pattern of relationships evident in

Table 2. *Variation at 16 microsatellite loci among Darwin's finches and two mainland species*

species	*n*	distance (G_{ST}) to *G. fortis*	mean values for 16 loci: expected heterozygosity	number of alleles	allele size range[a] (bp)
G. fortis	65	—	0.74	10.9	27
G. fuliginosa	71	0.11	0.78	12.9	33
G. magnirostris	37	0.19	0.65	7.4	25
G. scandens	92	0.35	0.68	9.9	26
G. conirostris	72	0.36	0.68	8.1	24
G. difficilis	97	0.38	0.70	10.2	27
C. parvulus	33	0.56	0.52	6.8	22
C. pauper	19	0.56	0.50	5.2	18
C. pallida	16	0.60	0.41	4.2	15
C. psittacula	16	0.69	0.45	4.4	15
P. crassirostris	53	1.02	0.52	6.7	17
C. fusca	51	1.02	0.52	5.7	21
P. inornata	30	1.31	0.47	4.6	14
C. olivacea	33	1.35	0.69	9.8	27
T. olivacea	17	2.42	0.44	4.1	13
S. aurita	8	1.61	0.34	2.7	8

species	*n*	mean values for 16 loci: allele size s.d.	allele size deviation[b]	variable loci (%)	odd-sized alleles[c] (%)
G. fortis	65	7.7	0.7	100	< 0.1
G. fuliginosa	71	8.3	−0.6	100	0.0
G. magnirostris	37	6.5	−0.3	100	0.0
G. scandens	92	7.2	0.3	100	0.0
G. conirostris	72	6.5	−0.2	100	0.3
G. difficilis	97	6.8	−2.0	100	< 0.1
C. parvulus	33	5.6	−3.0	94	0.2
C. pauper	19	5.1	−1.8	94	0.2
C. pallida	16	3.8	−2.3	94	3.3
C. psittacula	16	4.2	−3.2	94	0.6
P. crassirostris	53	3.5	−3.0	94	0.0
C. fusca	51	4.1	−3.0	100	0.1
P. inornata	30	3.7	−2.5	88	0.0
C. olivacea	33	6.7	−2.6	100	0.5
T. olivacea	17	4.0	−11.8	75	0.0
S. aurita	8	2.4	−17.4	63	1.2

[a] Allele size range is calculated as the difference between the maximum and minimum allele sizes for each locus.
[b] The deviation of a species' mean allele size from the mean allele size of all alleles across all species.
[c] Alleles with lengths not multiples of two bp different from most other alleles at the same locus.

table 2, as *Sporophila* appears to be more distantly related to Darwin's finches than *Tiaris*. Alternative topologies that would contradict the monophyly hypothesis were not observed (< 0.1%) in bootstrap re-samplings when $(\delta\mu)^2$ was used with KITCH.

(c) *Divergence of* Certhidea

An unexpected result is that populations of *Certhidea*, currently classified as a single species, *C. olivacea*, diverged well before any other species of Darwin's finch arose (figure 1*c*). Regardless of the exact phylogenetic topology, this deep split among *Certhidea* populations on central and peripheral islands, coupled with relatively little divergence in morphology (Lack 1947) and song (Bowman 1983), stands in stark contrast to the great diversity that evolved among other descendent lineages (figure 3). A second result is that *Certhidea* are not depicted as sister taxa in any reconstruction. Topologies depicting *Certhidea* species as sister taxa were rarely observed in bootstrap replicates (UPGMA < 5%, KITCH < 3%, FITCH < 1%, CONTML < 14% and $(\delta\mu)^2$/KITCH < 1%).

(d) *The origin of* Pinaroloxias

The microsatellite phylogeny resolves the biogeographic problem of determining the sequence of colonization of the Galápagos and Cocos Islands (Grant 1986). Cocos Island is approximately midway between mainland Costa Rica (500 km) and the Galápagos Islands (630 km), while the Galápagos Islands are over 900 km from the coast of Ecuador. There are three main

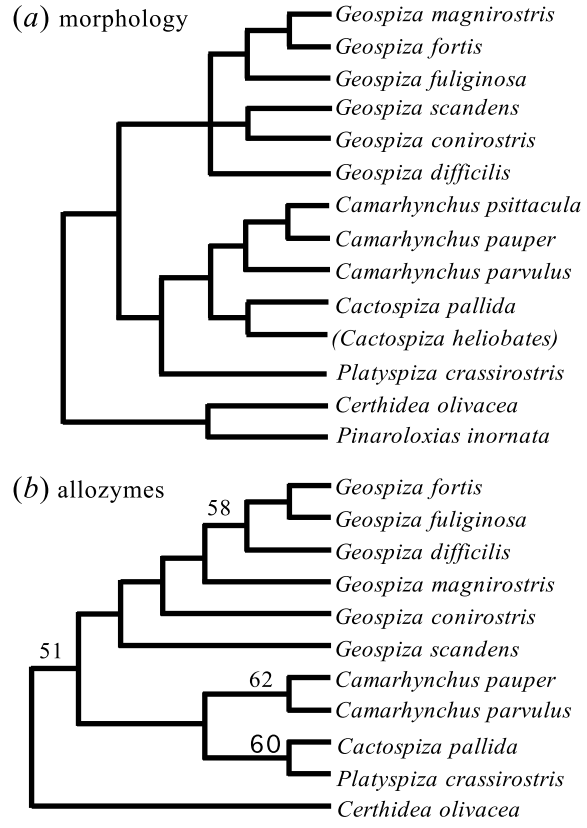

Figure 1. Phylogenetic hypotheses for Darwin's finches. (*a*) The phylogenetic topology proposed by Lack (1947) based on morphological characteristics. (*b*) The phylogenetic topology based on allozyme variation (Yang & Patton 1980; Stern & Grant 1996) using Nei's (1972) distance (G_{ST}) and UPGMA (Sokal & Sneath 1963). Numbers indicate the percentage of bootstrap replicates ($> 50\%$) that supported the node.

hypotheses. Either the Galápagos Islands were colonized from the mainland, followed by colonization of Cocos Island from the Galápagos Islands (Snodgrass 1903), Cocos Island was colonized first, followed by colonization of the Galápagos Islands by emigrants from Cocos Island (Harris 1973), or Cocos Island and the Galápagos Islands were colonized independently from the mainland (Steadman 1982).

All five methods place *Pinaroloxias* within the Darwin's finch clade, which is consistent with the first hypothesis and inconsistent with the other two. The Cocos Island finch (*Pinaroloxias*) was derived from the Galápagos Islands' lineage after radiation was underway (figure 1*c*). This fits with the geological evidence. Cocos Island appears to be much younger (*ca.* 2 Myr ago; Castillo *et al.* 1988) than the Galápagos Islands (> 10 Myr ago; Christie *et al.* 1992) and when formed it was closer to the Galápagos Islands. The third hypothesis is highly unlikely because, given the phylogenetic topology, a complex colonization history would be required. Topologies depicting *Pinaroloxias* basal to all finch lineages, which would contradict this interpretation, were not commonly observed in bootstrap re-samplings (UPGMA $< 3\%$, KITCH $< 8\%$, FITCH $< 7\%$, CONTML $< 28\%$ and $(\delta\mu)^2$/KITCH $< 2\%$).

(e) *Patterns of divergence within the tree and ground finches*

All methods support the monophyletic relationship of the ground finches (*Geospiza*) and the *Camarhynchus* and *Cactospiza* tree finches (figure 2). All methods except CONTML support the monophyletic arrangement of *G. magnirostris*, *G. fortis* and *G. fuliginosa*. Four out of five methods denote *Camarhynchus* as monophyletic with bootstrap values in the range of 90–100% for three methods. However, the vegetarian tree finch (*Platyspiza*) is placed outside this clade (basal to the tree and ground finches) in all but the FITCH reconstruction, which, however, has low bootstrap support for this alternative arrangement. Bootstrap support for the alternate grouping of *Platyspiza* with the tree finches ranged from 17–27% among the other four methods.

4. DISCUSSION

(a) *Phylogenetic reconstruction*

For the six main conclusions of the analysis, there was complete agreement across five different methods of phylogenetic reconstruction despite the fact that they differ in their underlying assumptions. The tree-building methods which assume a molecular clock (UPGMA and KITCH) revealed much higher bootstrap support (reflecting concordance among loci) than the FITCH method, which makes no such assumption. A recent simulation study revealed that UPGMA performs better when evolutionary rates are high (Huelsenbeck & Kirkpatrick 1996) and evolutionary rates are expected to be high with microsatellites. The UPGMA method gives greater weight to the genetic relationships among more closely related taxa (Sokal & Sneath 1963). This may be appropriate for microsatellite analysis because G_{ST} is expected to vary more linearly with shorter time-scales (Goldstein *et al.* 1995; table 2). As expected, $(\delta\mu)^2$ provided strong support for conclusions involving deeper nodes of the tree. Although $(\delta\mu)^2$ may be better with these relatively longer time-scales, the use of this measure may be limited because of nonlinearities associated with changes in the mutation rate. The CONTML method showed less support for the deeper nodes; however, this maximum-likelihood method unrealistically assumes that mutation is absent (Felsenstein 1981, 1993).

Hybridization will hinder the recovery of true phylogenetic relationships if introgression occurs frequently among taxa (Grant & Grant 1992; Avise 1994). This is a problem that is not unique to microsatellites. We expect that hybridizing species will tend to cluster together more closely than if they did not hybridize (Grant 1986). Some evidence of rare hybridization has been recorded among most species pairs within the ground finch clade, within the tree finch clade and between the warbler and tree finches (Grant 1986). However, in the best-studied case of introgression in Darwin's finches, between *G. fortis* and *G. scandens* on the island of Daphne Major (Grant 1993; Grant & Grant 1994), the genetic affinity of these populations is still much closer to other populations of the same species (K. Petren, B. R. Grant and P. R. Grant, unpublished data) and these taxa do not cluster together phylogenetically (figure 3). *Pinaroloxias* on the isolated

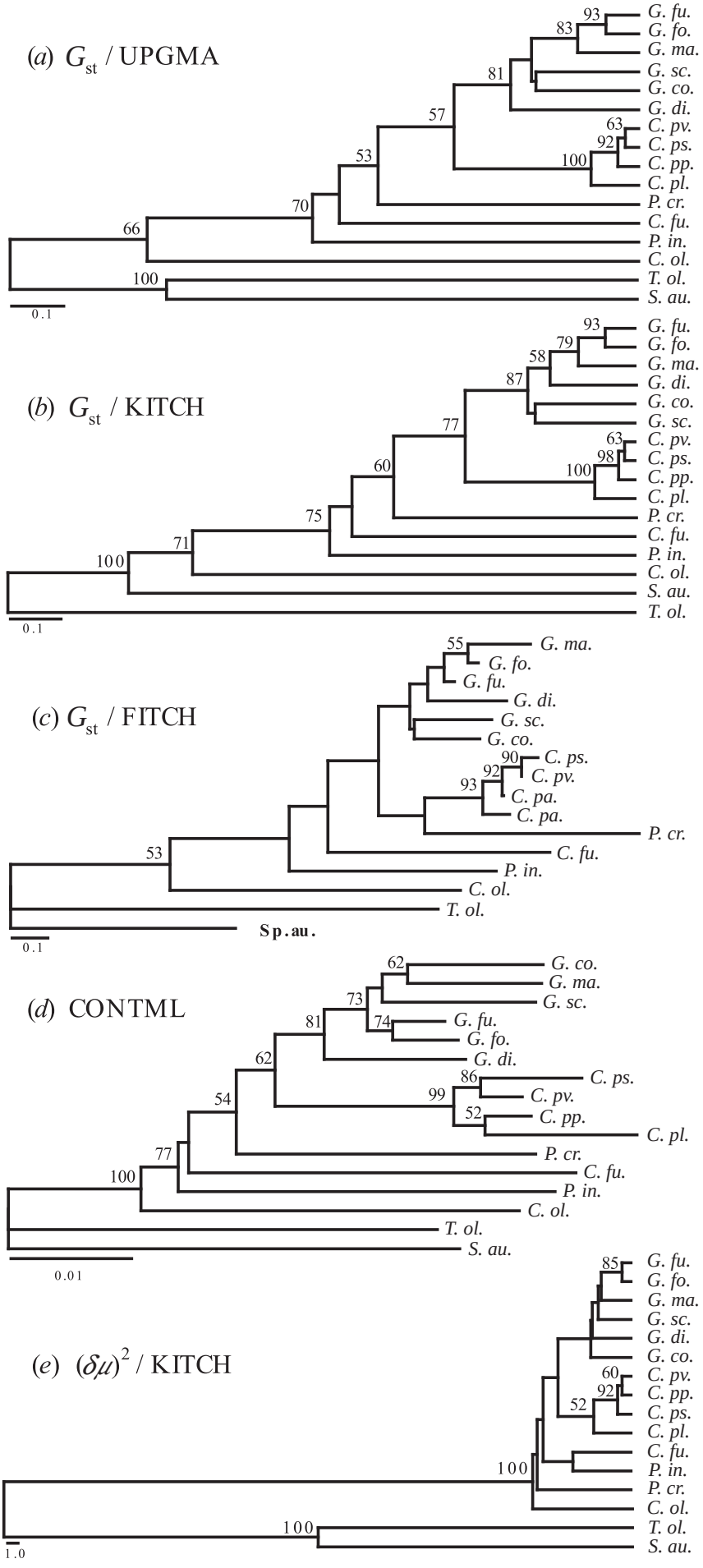

Figure 2. (*a–e*) Phylogenetic reconstructions using microsatellite length variation (see § 2 for descriptions of techniques). Primer pairs developed in *G. fortis* ($n = 16$; Petren 1998; Petren *et al*. 1999) were used to obtain more than 11 000 genotypes from 710 birds. Branch lengths are proportional to genetic distance as indicated by the scales. Numbers indicate the percentage of 1000 bootstrap replicates (> 50%) that supported the node.

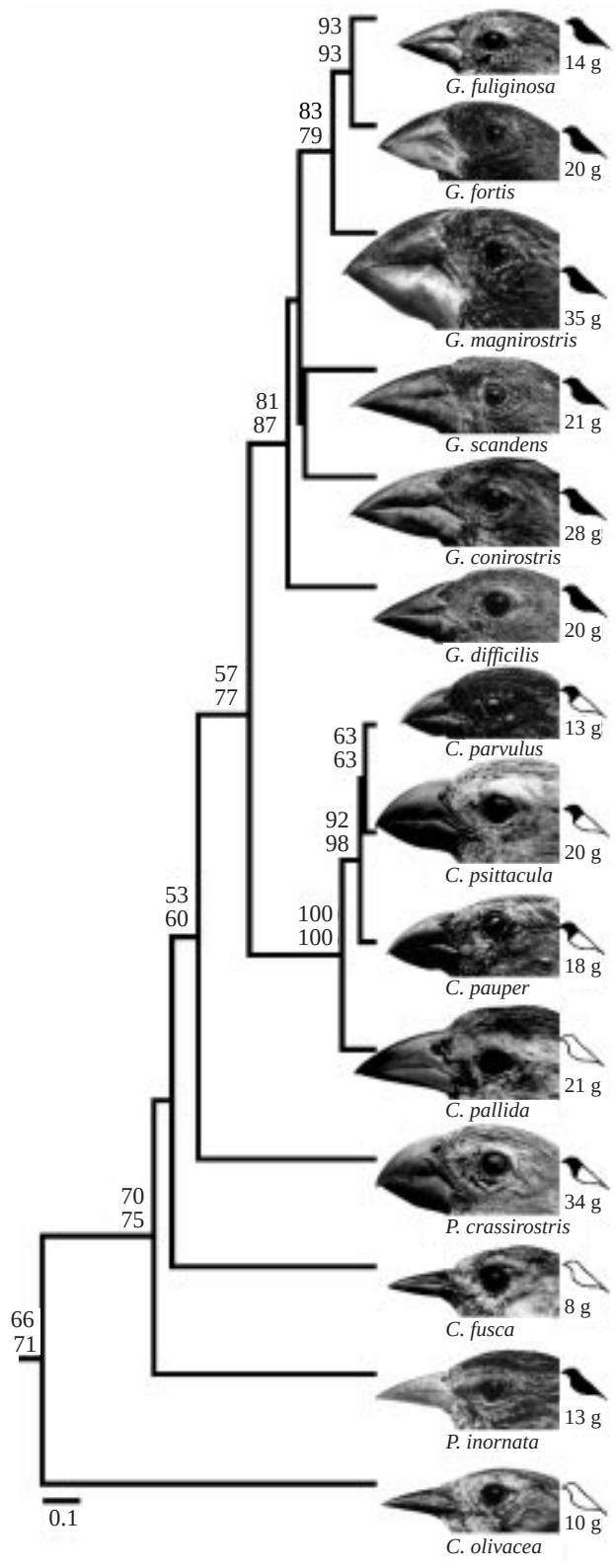

Figure 3. A phylogram of Darwin's finches based on microsatellite length variation constructed using G_{ST} and UPGMA. Photographs of birds are proportional to actual size. The maximum amount of black colouring in male plumage and the mean body mass among populations is indicated for each species (Grant 1986). Horizontal branch lengths are proportional to units of genetic distance (G_{ST}) as indicated by the scale. Numbers indicate the percentage of 1000 bootstrap replicates ($>$ 50%) that supported the node (UPGMA method above and KITCH method below). Names of genera are given in full in table 1. On the basis of these results, reclassification is justifiable but has not been done.

Cocos Island is immune from problems arising from hybridization.

(b) *Evolution in Darwin's finches*

This is the first molecular (or biochemical) study to support the monophyletic classification of Darwin's finches, as well as the placement of *Pinaroloxias* within the group. These results agree with other studies that clearly place Darwin's finches in a monophyletic group based on morphology, plumage, song and other characteristics (Bowman 1961, 1983; Grant 1986). Molecular sequence and microsatellite analysis of other mainland taxa not included here are consistent with these results (Freeland 1997; Sato *et al*. 1999*a*; K. Petren, unpublished data).

There are indications of the deep split among *Certhidea* populations from allozyme data (Polans 1983) as well as recent mtDNA sequence analysis (Freeland 1997). However, until now, the division of *Certhidea* has not been placed in a phylogenetic context. Lack (1947, 1961) argued for a *Geospiza*-like ancestor of all Darwin's finches (but see Swarth 1931). If the non-sister relationship among *Certhidea* is correct and these lineages represent two independent branchings from the main lineage, the argument for a more *Certhidea*-like ancestor to all of Darwin's finches is strengthened. We cannot reject Lack's (1947, 1961) view that the ancestor of all Darwin's finches possessed *Geospiza*-like traits such as black plumage and a blunt beak. Yet if this was the case then the two *Certhidea* lineages would represent a remarkable case of convergence in morphology, behaviour, plumage and song.

The *Certhidea* results suggest that, in some instances, morphology may be a poor guide to the genetic distinctness of populations. This is of particular relevance to management strategies. Similar findings have emerged from molecular studies of birds (Avise & Nelson 1989) and reptiles (Daugherty *et al*. 1990), but, to our knowledge, discovery of an unsuspected divergence occurring at the base of an adaptive radiation has not been previously reported. There are no *Certhidea* populations currently in danger of extinction, but, if they become threatened in the future, more than one will deserve protection by virtue of their genetic distinctiveness. This emphasizes the importance of verifying the genetic ancestry of not only threatened species, but also threatened populations.

Many phenotypic traits of Darwin's finches, such as beak size and shape, body size and plumage coloration, have been studied extensively (Lack 1947; Bowman 1961, 1983; Grant 1986). Reconstructing their evolution is not straightforward and is not attempted here because many of these traits (particularly beak shape) are subject to strong ecological and genetic constraints (Grant & Grant 1999). Furthermore, non-parsimonious evolutionary reconstructions are biologically plausible since evolution can proceed very rapidly in this system (Grant & Grant 1995). However, given the improved phylogenetic resolution provided by microsatellites (figure 3) we offer two relevant observations.

First, with the exception of *Platyspiza*, we note that species that root basally on the tree and the basal members of the tree and ground finches have relatively long, pointed beaks. This beak form is generally associated with an insectivorous diet (Bowman 1961). Two novel blunt-beaked forms evolved later in the ground

finches (*G. magnirostris*, *G. fortis* and *G. fuliginosa*), and in the tree finches (*Camarhynchus*). The ground finch beak is efficient for crushing seeds at the base, while the tree finch beak permits greater biting strength at the tip for tearing vegetation (Bowman 1961). Once a novel beak shape evolved in these two groups, beak shape remained relatively constant while body size and beak size changed. This supports the view that evolution most easily proceeds along lines of allometry (Grant & Grant 1995): once a novel shape is formed away from the line of allometry, a rapid divergence in size along a new line of allometry is possible.

Second, the microsatellite phylogeny generally supports Lack's (1947) view that plumage is a relatively conserved trait: species with similar plumage patterns generally cluster together (figure 3). Yet the tree also suggests that multiple evolutionary transitions have taken place in plumage coloration, as *Geospiza* and *Pinaroloxias* males share black plumage, while *Certhidea* and *Cactospiza* males share similar green–grey plumage. It is likely that one of these instances of a shared trait is the result of convergence. This is perhaps not surprising because melanism in other birds is controlled by a small number of loci (Buckley 1987).

In conclusion, we suggest that microsatellites may be useful for reconstructing the evolutionary history of other groups of organisms which, like Darwin's finches, have radiated relatively recently and rapidly. Microsatellite studies could complement DNA sequence and allozyme studies by providing resolution at shorter time-scales. Microsatellite variation is expected to be most useful when divergence times are short and when populations are small (Nauta & Weissing 1996; Takezaki & Nei 1996), as is the case with Darwin's finches. All 16 loci used in this study were simple (pure) dinucleotide repeats, which show greater promise for phylogenetic reconstruction than other microsatellite motifs (Primmer & Ellegren 1998). Nevertheless, in view of continuing uncertainty over mutation mechanisms, homoplasy and the best methods of analysis for microsatellite data, there is a need for further theoretical and empirical investigation of their use in estimating phylogenies.

We thank the Galápagos and Costa Rica National Parks Services and the Charles Darwin Research Station (Galápagos Islands), the staff on Cocos Island and the STRI for administrative and logistical support, the National Science Foundation (USA) and the NSF–Sloan Foundation for financial support, P. T. Boag, B. Fessl, G. Seutin, S. Tebbich, and many assistants for help in the Galápagos Islands, M. Wikelski and M. Hau for mainland samples, E. K. Monson, V. Schulz and V. A. Zakian for laboratory advice, and L. F. Keller, T. D. Price, J. T. Bonner, T. J. Case and five anonymous reviewers for comments on earlier drafts.

REFERENCES

Avise, J. 1994 *Molecular markers, natural history and evolution*. New York: Chapman & Hall.

Avise, J. C. & Nelson, W. S. 1989 Molecular genetic relationships of the extinct dusky seaside sparrow. *Science* **243**, 646–648.

Baptista, L. F. & Trail, P. W. 1988 On the origin of Darwin's finches. *Auk* **105**, 663–671.

Bowman, R. I. 1961 Morphological differentiation and adaptation in the Galápagos finches. *Univ. Calif. Publ. Zool.* **58**, 1–302.

Bowman, R. I. 1983 The evolution of song in Darwin's finches. In *Patterns of evolution in Galápagos organisms* (ed. R. I. Bowman, M. Berson & A. E. Leviton), pp. 237–537. San Francisco: American Association for the Advancement of Sciences.

Buckley, P. A. (ed.) 1987 Mendelian genes. In *Avian genetics* (ed. F. Cooke & P. A. Buckley), pp. 1–44. London: Academic Press.

Callen, D. F., Thompson, A. D. & Shen, Y. 1993 Incidence and origin of 'null' alleles in the $(AC)_n$ microsatellite markers. *Am. J. Hum. Genet.* **52**, 922–927.

Castillo, P., Batiza, R., Vanko, D., Malvassi, E., Basqueso, J. & Fernandez, E. 1988 Anomalously young volcanoes on old hot-spot traces. I. Geology and petrology of Cocos Island. *Geol. Soc. Am. Bull.* **100**, 1400–1414.

Christie, D. M., Duncan, R. A., McBirney, A. R., Richards, M. A., White, W. M., Harpp, K. S. & Fosc, C. G. 1992 Drowned islands downstream from the Galápagos hotspot imply extended speciation times. *Nature* **355**, 246–248.

Daugherty, C. H., Cree, A. H., Hay, J. M. & Thompson, M. B. 1990 Neglected taxonomy and continuing extinctions of *Tuatara sphenodon*. *Nature* **347**, 177–179.

Edwards, A. F. W. & Cavalli-Sforza, L. L. 1964 Reconstruction of evolutionary trees. In *Phenetic and phylogenetic classification* (ed. V. H. Heywood & J. McNeill), pp. 67–76. London: Systematics Association.

Ellegren, H., Primmer, C. R. & Sheldon, B. C. 1995 Microsatellite 'evolution': directionality or bias? *Nat. Genet.* **11**, 360–362.

Felsenstein, J. 1981 Evolutionary trees from gene frequencies and quantitative characters: finding maximum likelihood estimates. *Evolution* **35**, 1229–1242.

Felsenstein, J. 1984 Distance methods for inferring phylogenies: a justification. *Evolution* **38**, 16–24.

Felsenstein, J. 1993 *PHYLIP* (*phylogeny inference package*), *version 3.5c*. Zoology Department, University of Washington, Seattle.

Fitch, W. M. & Margoliash, E. 1967 Construction of phylogenetic trees. *Science* **155**, 279–284.

Freeland, J. R. 1997 The genetic evolutionary history of the Darwin's finches (Aves: Geospizinae). PhD thesis, Queen's University, Kingston.

Garza, J. C., Slatkin, M. & Freimer, N. B. 1995 Microsatellite allele frequencies in humans and chimpanzees, with implications for constraints on allele size. *Mol. Biol. Evol.* **12**, 594–603.

Givnish, T. J. & Sytsma, K. J. (ed.) 1997 *Molecular evolution and adaptive radiation*. Cambridge University Press.

Goldstein, D. B. & Pollock, D. D. 1997 Launching microsatellites: a review of mutation processes and methods of phylogenetic inference. *J. Hered.* **88**, 335–342.

Goldstein, D. B., Linares, A. R., Cavalli-Sforza, L. L. & Feldman, M. W. 1995 An evaluation of genetic distances for use with microsatellite loci. *Genetics* **139**, 463–471.

Grant, P. R. 1986 *Ecology and evolution of Darwin's finches*. Princeton University Press.

Grant, P. R. 1993 Hybridization of Darwin's finches on Isla Daphne Major, Galápagos. *Phil. Trans. R. Soc. Lond.* B **340**, 127–139.

Grant, P. R. 1994 Population variation and hybridization: comparison of finches from two archipelagos. *Evol. Ecol.* **8**, 598–617.

Grant, P. R. (ed.) 1998 *Evolution on islands*. Oxford University Press.

Grant, P. R. & Grant, B. R. 1992 Hybridization of bird species. *Science* **256**, 193–197.

Grant, P. R. & Grant, B. R. 1994 Phenotypic and genetic effects of hybridization in Darwin's finches. *Evolution* **48**, 297–316.

Grant, P. R. & Grant, B. R. 1995 Predicting microevolutionary responses to directional selection on heritable variation. *Evolution* **49**, 241–251.

Grant, P. R. & Grant, B. R. 1999 Quantitative genetic variation in populations of Darwin's finches. In *Adaptive variation in the wild* (ed. T. A. Mousseau, B. L. Sinervo & J. Endler). Oxford University Press. (In the press.)

Grant, P. R., Grant, B. R. & Petren, K. 1999 The allopatric phase of speciation: the sharp-beaked ground finch (*Geospiza difficilis*) on the Galápagos Islands. (In preparation.)

Harris, M. P. 1973 *Coereba flaveola* and the Geospizinae. *Bull. Br. Ornithol. Club* **92**, 164–168.

Huelsenbeck, J. P. & Kirkpatrick, M. 1996 Do phylogenetic methods produce trees with biased shapes? *Evolution* **50**, 1418–1424.

Lack, D. 1947 *Darwin's finches*. Cambridge University Press.

Lack, D. 1961 *Darwin's finches*, 2nd edn. Cambridge University Press.

McDonald, D. B. & Potts, W. K. 1997 DNA microsatellites as genetic markers at several scales. In *Avian molecular evolution and systematics* (ed. D. P. Mindell), pp. 29–49. New York: Academic Press.

Mantel, N. 1967 The detection of disease clustering and a generalized regression approach. *Cancer Res.* **27**, 209–220.

Minch, E., Ruiz-Linares, A., Goldstein, D. B., Feldman, M. W. & Cavalli-Sforza, L. L. 1996 *Microsat program, version 1.5e.* Biology Department, University of Stanford.

Nauta, M. J. & Weissing, F. J. 1996 Constraints on allele size at microsatellite loci: implications for genetic differentiation. *Genetics* **143**, 1021–1032.

Nei, M. 1972 Genetic distance between populations. *Am. Nat.* **106**, 283–292.

Orti, G., Pearse, D. E. & Avise, J. 1997 Phylogenetic assessment of length variation at a microsatellite locus. *Proc. Natl Acad. Sci. USA* **94**, 10 745–10 749.

Petren, K. 1998 Microsatellite primers from *Geospiza fortis* and cross-species amplification in Darwin's finches. *Mol. Ecol.* **7**, 1782–1784.

Petren, K., Grant, B. R. & Grant, P. R. 1999 Low extra-pair paternity in the cactus finch (*Geospiza scandens*). *Auk* **116**, 252–256.

Polans, N. O. 1983 Enzyme polymorphisms in Galápagos finches. In *Patterns of evolution in Galapágos organisms* (ed. R. I. Bowman, M. Berson & A. E. Leviton), pp. 219–236. San Francisco: American Association for the Advancement of Sciences.

Pollock, D. D., Bergman, A., Feldman, M. W. & Goldstein, D. B. 1998 Microsatellite behavior with range constraints: parameter estimation and improved distances for use in phylogenetic reconstruction. *Theor. Popul. Biol.* **53**, 256–271.

Primmer, C. R. & Ellegren, H. 1998 Patterns of molecular evolution in avian microsatellites. *Mol. Biol. Evol.* **15**, 997–1008.

Primmer, R., Møller, A. P. & Ellegren, H. 1995 Resolving genetic relationships with microsatellite markers: a parentage testing system for the swallow *Hirundo rustica*. *Mol. Ecol.* **4**, 493–498.

Roy, M. S., Geffen, E., Smith, D., Ostrander, E. A. & Wayne, R. K. 1994 Patterns of differentiation and hybridization in North American wolflike canids, revealed by analysis of microsatellite loci. *Mol. Biol. Evol.* **11**, 553–570.

Sambrook, J., Fritsch, E. F. & Maniatis, T. 1989 *Molecular cloning: a laboratory manual*, 2nd edn. New York: Cold Spring Harbor Laboratory Press.

Sato, A., O'Huigin, C., Figueroa, F., Grant, P. R., Grant, B. R., Tichy, H. & Klein, J. 1999*a* Phylogeny of Darwin's finches as revealed by mitochondrial DNA sequences. (In preparation.)

Sato, A., Tichy, H., O'Huigin, C., Grant, P. R., Grant, B. R. & Klein, J. 1999*b* On the origin of Darwin's finches. *Proc. Natl Acad. Sci. USA*. (Submitted.)

Schluter, D. 1984 Morphological and phylogenetic relations among the Darwin's finches. *Evolution* **38**, 921–930.

Snodgrass, R. E. 1903 Notes on the anatomy of *Geospiza*, *Cocornis*, and *Certhidea*. *Auk* **20**, 402–417.

Sokal, R. R. & Sneath, P. H. 1963 *Principles of numerical taxonomy*. San Francisco: Freeman.

Steadman, D. 1982 The origin of Darwin's finches (Fringillidae, Passeriformes). *Trans. San Diego Soc. Nat. Hist.* **19**, 279–296.

Stern, D. L. & Grant, P. R. 1996 A phylogenetic reanalysis of allozyme variation among populations of Galápagos finches. *Zool. J. Linn. Soc.* **118**, 119–134.

Sushkin, P. P. 1929 On some peculiarities of adaptive radiation presented by insular faunae. In *Proceedings of the Sixth International Ornith. Congress*, Kopenhagen. pp. 375–378.

Swarth, H. S. 1931 The avifauna of the Galápagos Islands. *Occ. Pap. Calif. Acad. Sci.* **18**, 1–299.

Takezaki, N. & Nei, M. 1996 Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics* **144**, 389–399.

Weber, J. & Wong, C. 1993 Mutation of human short tandem repeats. *Hum. Mol. Genet.* **2**, 1123–1128.

Yang, S. Y. & Patton, J. L. 1981 Genetic variability and differentiation of Galápagos finches. *Auk* **98**, 230–242.

As this paper exceeds the maximum length normally permitted, the authors have agreed to contribute to production costs.