

# A Test of Founder Effect Speciation Using Multiple Loci in the Auklets (*Aethia* spp.)

H. E. Walsh,<sup>\*,1</sup> I. L. Jones<sup>†</sup> and V. L. Friesen<sup>\*</sup>

<sup>\*</sup>Department of Biology, Queen's University, Kingston, Ontario K7L 3N6, Canada and <sup>†</sup>Department of Biology, Memorial University of Newfoundland, St. John's, Newfoundland A1B 3X9, Canada

Manuscript received March 16, 2005  
Accepted for publication August 22, 2005

## ABSTRACT

Whether speciation results more frequently from the genetic consequences of founder events or from gradual genetic divergence of large populations is a matter of debate. In this study, multiple analyses were applied to data from three loci (cytochrome *b*,  $\alpha$ -enolase intron VIII, and MHC class II B) to test for founder effects associated with speciation in *Aethia* (Aves: Alcidae), a genus of seabirds thought to have undergone a rapid founder-induced radiation. Effective population sizes ( $N_e$ ) were derived from estimators of  $\theta$  based on allelic diversity and the coalescent and from data on trans-species polymorphism. Results indicated that  $N_e$  has been on the order of  $10^5$ – $10^6$  individuals throughout the evolutionary histories of least and crested auklets (*A. pusilla* and *A. cristatella*, respectively) and that  $N_e$  of the ancestral species was at least 16,000 individuals. Computer simulations of MHC evolution indicated that a single-generation bottleneck at speciation could not have involved <85 individuals for each species. More moderate simulation scenarios indicated that population size could not have dropped below 2000 individuals at the time of species founding. Demographic history appears to have been stable for the auklets throughout the past several million years, and a founder effect associated with their speciation is unlikely.

**F**OUNDER-INDUCED speciation is arguably one of the most controversial of all potential mechanisms of speciation (MAYR 1942; BARTON and CHARLESWORTH 1984; CARSON and TEMPLETON 1984; TEMPLETON 1986; PROVINE 1989; GAVRILETS and HASTINGS 1996; HOLLOCHER 1996; CLEGG *et al.* 2002; RUNDLE 2003; COYNE and ORR 2004). As a result of a founder event (*i.e.*, population founding by a small number of individuals), gene frequencies may change rapidly and particular alleles may be lost or fixed due to sampling error alone. Founder events also may induce linkage disequilibrium, and new or rare mutations and/or gene combinations (including those that are slightly deleterious) have a higher probability of drifting to fixation in small populations (WRIGHT 1969; GAVRILETS and BOAKE 1998). Such founder effects, in concert with environmental differences experienced by founder populations, have the potential to precipitate speciation (but see RUNDLE 2003). For example, by accelerating divergence in small populations, founder effects may facilitate rapid and/or adaptive radiations of taxa. Founder effect speciation is therefore among the speciation hypotheses considered most commonly for taxa whose histories appear to have contained radiations (CARSON 1983; D. KLEIN *et al.* 1993; HOLLOCHER 1996; VINCEK *et al.* 1997; CLEGG

*et al.* 2002). Founder effect speciation is consistent with both vicariant and dispersal scenarios for population isolation.

Although the incidence of a founder event alone is not sufficient proof of the occurrence of speciation by founder effect, a first assessment of the relevance of a founder effect model to a given case of speciation involves evaluation of whether or not a founder event (such as a bottleneck in population size) plausibly occurred in the history of the examined taxa. To date, however, no empirical data have demonstrated a convincing association between bottlenecks in population size and speciation in vertebrates. Among animals, documented examples of what may be founder effect speciation exist, controversially, for *Drosophila* species (CARSON 1983; HOLLOCHER 1996; TEMPLETON 1996), and cases of founder-induced speciation may have occurred in seabirds such as the shy albatross (*Thalassarche cauta*) (ABBOTT and DOUBLE 2003) and Armenian gull (*Larus armenicus*) (LIEBERS *et al.* 2001). The genus *Drosophila*, for example, has radiated throughout the Hawaiian islands to comprise almost 1000 species in the few million years since formation of the archipelago, and a founder-flush mechanism of speciation has been proposed for these flies (CARSON 1983; HOLLOCHER 1996).

Founder effect has been viewed as a theoretically attractive mechanism for vertebrate speciation (TEMPLETON 1986; GAVRILETS and HASTINGS 1996; GAVRILETS and BOAKE 1998; RUNDLE *et al.* 1998; TURELLI *et al.* 2001), but

<sup>1</sup>Corresponding author: Department of Biology, University of Washington, Box 351800, Seattle, WA 98195-1800.  
E-mail: hwalsh@u.washington.edu

its occurrence may be rare (GAVRILETS 2004) and it is not supported in the few cases in which it has been tested empirically. For example, hundreds of species of cichlid fishes originated in the African Rift Lakes in <1 million years (MEYER *et al.* 1990; JOHNSON *et al.* 1996), but extensive intraspecific major histocompatibility complex (MHC) variability suggests that their speciation was not due to a founder effect (D. KLEIN *et al.* 1993). Similarly, despite evidence for rapid radiation, no evidence indicates that speciation was associated with founder effects in either threespine sticklebacks (*Gasterosteus aculeatus*) (McKINNON and RUNDLE 2002) or island-colonizing birds in the silvereye species complex (*Zosterops lateralis*) (CLEGG *et al.* 2002). On the basis of human leukocyte antigen (HLA) polymorphism and evidence from nuclear introns, several researchers concluded that founder effect was not an important component of human evolution (TAKAHATA 1991, 1993; TAKAHATA *et al.* 1995; AYALA and ESCALANTE 1996). Likewise, founder effect has been discounted as the primary mode of speciation among Darwin's finches (*Geospiza* spp.) since the discovery of extensive MHC variability in these species (VINCEK *et al.* 1997).

Methods have been developed for inferring both long-term and ancestral effective population sizes ( $N_e$  and  $N_a$ , respectively) from DNA polymorphism in extant populations, as well as from variation shared by closely related species (*e.g.*, KIMURA 1969; WATTERSON 1975; KINGMAN 1982a,b; TAJIMA 1983; ROGERS and HARPENDING 1992; TAKAHATA *et al.* 1995; YANG 1997, 2002; RANNALA and YANG 2003; WALL 2003). These methods allow tests of explicit hypotheses on the origin and radiation of taxa and have been particularly informative in the controversy surrounding the origin of modern humans (TAKAHATA 1993; ROGERS 1995; TAKAHATA *et al.* 1995; AYALA and ESCALANTE 1996). In this study, we applied several of these methods to test the hypothesis that founder effects were involved in speciation in *Aethia* auklets, a group of North Pacific seabirds.

*Aethia* (Charadriiformes: Alcidae) comprises four species (*Aethia cristatella*, *A. psittacula*, *A. pusilla*, and *A. pygmaea*) endemic to the North Pacific, Bering Sea, and Sea of Okhotsk. As wing-propelled divers, auklets spend most of their lives at sea, returning to islands and continental coasts to nest in sympatry in the spring. All four species are sexually monomorphic and display facial plumes and bill ornamentation during the breeding season; they appear to be monogamous and can produce only a single chick per breeding pair each breeding season (GASTON and JONES 1998; JONES 1999). Phylogenetic relationships within the genus are unresolved, and power analysis suggested that internodes between divergence events in this group were extremely short (relative to population sizes) (WALSH *et al.* 1999). The above results have led to the hypothesis that speciation in the genus was facilitated by population bottlenecks and/or founder events, precipitated by one

or more vicariant events (JONES 1999; WALSH *et al.* 1999). Contemporary census population sizes are on the orders of millions of individuals for both least and crested auklets (*A. pusilla* and *A. cristatella*, respectively) (STEPHENSON and IRONS 2003) and, on the basis of life history tables, generation times are 6.7 and 8.3 years, respectively (WALSH 1999; WALSH and FRIESEN 2003). DNA sequences of the cytochrome *b* gene indicate that lineages leading to least and crested auklets diverged  $\sim 2.8$  MYA (FRIESEN *et al.* 1996; WALSH and FRIESEN 2003). Given this time frame, divergence of the auklets may have followed isolation within glacial refugia of the Plio-Pleistocene, consistent with the glacier-induced speciation models proposed for many groups of species native to the Americas (*e.g.*, O'REILLY *et al.* 1993; HEWITT 1996; AVISE and WALKER 1998; HOLDER *et al.* 1999; but see KLICKA and ZINK 1997, 1999). Auklets exhibit both strong mutual sexual selection and behavioral isolation among species (JONES and HUNTER 1993; GASTON and JONES 1998), both of which are among the proposed consequences of founder effect speciation (POWELL 1989).

Recent methods for estimating ancestral effective population sizes ( $N_a$ ) rely on identifying conflicts between gene trees and the presumed known species tree (YANG 2002; RANNALA and YANG 2003) or on maximizing the likelihood of a given number of inferred mutations along branches in a known phylogeny (WALL 2003). For auklets, however, the species tree is unknown despite analyses of a large amount of sequence data (MOUM *et al.* 1994; FRIESEN *et al.* 1996; WALSH *et al.* 1999). Without a known species tree, some methods for estimating  $N_a$  cannot be applied; however, data from loci that exhibit trans-species polymorphism can be used to evaluate  $N_a$  for species with unknown species trees. Using data on genetic diversity and allelic genealogy at three loci, we estimated  $N_e$  at several depths in the evolutionary histories of least and crested auklets to test the hypothesis that their speciation was associated with a founder effect.

## THEORY AND METHODS

Because genetic diversity is a function of long-term  $N_e$ , estimates of  $N_e$  can be derived from the neutral mutation parameter,  $\theta$  ( $= 4N_e\mu$  for nuclear DNA and  $2N_f\mu$  for mitochondrial DNA, where  $\mu$  is the neutral mutation rate per sequence per generation, and  $N_f$  is female effective population size). (In what follows, we adopt the convention of Felsenstein to use  $\Theta$  to denote  $4N_e\mu$  per site and  $\theta$  to denote  $4N_e\mu$  per locus.)

Several methods have been advanced to estimate  $\theta$  from DNA sequences (KIMURA 1969; WATTERSON 1975; NEI 1987; FELSENSTEIN 1992; GRIFFITHS and TAVARÉ 1993; FU 1994a,b; KUHNER *et al.* 1995; TAKAHATA *et al.* 1995; TAJIMA 1996), and it can be scaled for balancing

selection by applying TAKAHATA'S (1990) correction factor. In this study,  $\theta$  was estimated using data on nucleotide polymorphism previously collected for least and crested auklets (WALSH and FRIESEN 2003). These data include 306 bp of sequence from the 5' end of the mitochondrial cytochrome *b* gene, intron VIII (248 bp) from the nuclear gene for  $\alpha$ -enolase, and a 154-bp fragment of the second exon of an MHC class II B gene. Analyses of 89 least auklets and 81 crested auklets sampled from throughout their breeding ranges revealed 5 and 8 alleles for cytochrome *b*, and 19 and 33 alleles for  $\alpha$ -enolase, respectively. At least 40 alleles at each of two MHC II B loci were observed for each species (WALSH and FRIESEN 2003).

**Substitution rates ( $\mu$ ):** Given the 2%/MY rate of cytochrome *b* divergence reported by AVISE and WALKER (1998), the substitution rate for cytochrome *b* is  $2.1 \times 10^{-5}$  and  $2.5 \times 10^{-5}$  substitutions per sequence per generation for least and crested auklets, respectively (given generation times of 6.7 and 8.3 years, respectively). The substitution rate for  $\alpha$ -enolase intron VIII was estimated at  $3.7 \times 10^{-6}$  and  $4.5 \times 10^{-6}$  substitutions per sequence per generation for least and crested auklets, respectively, by comparison of auklet sequences with marbled murrelet sequences (FRIESEN *et al.* 1997) and assuming a divergence time of 12 MYA (FRIESEN *et al.* 1996). The substitution rate for the MHC II B fragment was estimated on the basis of the rate of  $3.5 \times 10^{-9}$  substitutions per site per year of VINCEK *et al.* (1997), adjusted for generation times of the auklets to yield  $\mu = 3.6 \times 10^{-6}$  and  $4.5 \times 10^{-6}$  substitutions per sequence per generation, respectively.

**Estimating  $N_e$ :** For cytochrome *b* and  $\alpha$ -enolase, both  $\Pi$  (average pairwise differences) (TAJIMA 1983) and  $\theta_W$  (Watterson's  $\theta$ ) (WATTERSON 1975) were calculated.  $N_e$  was derived from  $\Pi$  and  $\theta_W$ , assuming substitution rates given above. We used the algorithm FLUCTUATE (KUHNER *et al.* 1995) to estimate contemporary  $\Theta$  for least and crested auklets, using data on intraspecific allelic sequences for cytochrome *b* and  $\alpha$ -enolase intron VIII.

We used the maximum-likelihood methods of TAKAHATA *et al.* (1995) to estimate  $N_e$  for each species and their associated  $N_a$ , given substitution rates for each locus.  $N_e$  for each species was estimated by evaluating TAKAHATA *et al.*'s (1995) Equation 2 at its maximum ( $\Theta_{ML}$ ), given the data for cytochrome *b* and  $\alpha$ -enolase. The maximum-likelihood value for  $N_a$  (and associated 95% confidence interval) was estimated numerically from TAKAHATA *et al.*'s (1995) Equation 3, with divergence time fixed at 2.8 MYA and a mean of four substitutions between two random sequences (WALSH and FRIESEN 2003).

Six polymorphisms are shared between least and crested auklets in intron VIII of the gene for  $\alpha$ -enolase (*i.e.*,  $S_S = 6$ ), and no fixed differences were observed between the two species for this locus. We thus used an

abbreviated form of the method developed by WAKELEY and HEY (1997) to estimate  $\theta_A (= 4N_a\mu)$  from the data for  $\alpha$ -enolase. We equated the number of polymorphisms shared between least and crested auklets with  $E(S_S)$ , the expectation for this quantity [given by WAKELEY and HEY'S (1997) Equation 12], substituting estimated  $\theta$ -values from  $\alpha$ -enolase intron VIII for  $\theta_1$  and  $\theta_2$ . This equation was solved for  $\theta_A$  and the corresponding value for  $N_a$  was derived, assuming that the substitution rate for this locus in auklets is  $4 \times 10^{-6}$  substitutions per sequence per generation (the average for the two species, above).

*$N_e$  from trans-species polymorphism:* If trans-species polymorphism exists in a sample of nucleotide sequences (and is sufficiently common that it cannot reasonably be attributed to convergence), then coalescence of alleles within species predates speciation (TAKAHATA 1991). Measured genetic distances between alleles can be superimposed on a genealogy to estimate  $k$ , the number of lineages extant today that also were present prior to a speciation event (TAKAHATA 1991; J. KLEIN *et al.* 1993).

If time is measured in units of  $2N_e$  generations [that is,  $t' = t/(2N_e)$ , where  $t$  is the known number of generations since a speciation event], the probability  $g_{nk}(t')$  that at  $t'$  (in the past) there were  $k$  alleles in a population, given that the population now carries  $n$  alleles, is given by

$$g_{nk}(t') = \sum_{j=k}^n \frac{(2j-1)(-1)^{j-k} k_{(j-1)} n_{[j]} e^{-(j(j-1)t'/2)}}{k!(j-k)!n_{(j)}} \quad (1)$$

when  $2 \leq k \leq n$  (TAVARÉ 1984), where  $a_{(j)} = a(a+1) \dots (a+j-1)$  and  $a_{[j]} = a(a-1) \dots (a-j+1)$  [ $a$  represents variables with subscripts in (1)].  $k$  was determined as described below.

Relationships among least and crested auklet alleles for  $\alpha$ -enolase intron VIII were estimated by the neighbor-joining method using Jukes-Cantor distances in MEGA v. 1.02 (KUMAR *et al.* 1993). The expected Jukes-Cantor corrected distance (JUKES and CANTOR 1969) between sequences that diverged 2.8 MYA (*i.e.*, about the time of divergence of least and crested auklets) is 0.0125 substitutions per site. For each species, the minimum number of allelic lineages present prior to the speciation event,  $k$ , was estimated as the number of lineages whose coalescence on the tree occurs prior to 0.0125 distance units ago (J. KLEIN *et al.* 1993). For each species, the maximum of Equation 1 was estimated numerically and  $N_e$  at this point was calculated from its relation with  $t'$ .

The timescales over which our estimates of long-term  $N_e$  apply are given by an exponentially distributed time parameter with mean  $4N_e[1 - (1/i)]$ , where  $i$  is the number of sampled alleles ( $4N_e$  is replaced with  $2N_f$  for mitochondrial DNA) (KINGMAN 1982a,b).

*$N_e$  by computer simulation:* Extant lineages whose coalescence predates speciation are lineages that persisted

through any putative bottleneck at the time of species founding. Thus, if the number of extant allelic lineages whose coalescence occurred at or prior to species divergence is known, then the minimum size of the founding population can be inferred from results of computer simulations. Allelic lineages for which coalescence predates speciation generally are observed at loci under balancing selection, which acts to maintain variation through time. Trans-species polymorphism also may be evident at neutral loci when  $\theta$  is large with respect to time since divergence, as noted by TAKAHATA (1986, 1989) and observed in our study (see RESULTS, Figure 1). VINCEK *et al.* (1997) used a computer simulation to determine the number of alleles retained in a population through bottlenecks of various sizes, given data on polymorphism for the second exon of an MHC II B gene. This locus is known to be under balancing selection of an intensity on the order of  $s = 0.01$  in humans (SATTA *et al.* 1994).

1

In this study we used the simulation algorithm of VINCEK *et al.* (1997) to determine the minimum population size of the immediate ancestor of least and crested auklets. We used available data on allelic polymorphism in a 154-bp fragment of the second exon of an MHC class II B gene for the auklets, at which a minimum of 40 alleles have been documented in each of least and crested auklets (WALSH and FRIESEN 2003). The number of allelic lineages older than the speciation event (2.8 MYA) was estimated by two methods. First, we estimated the average time to coalescence of two alleles under balancing selection,  $2N_e f_s$  generations [where  $f_s$  is given by TAKAHATA's (1990) Equation 15], using a selection intensity of 0.01 (SATTA *et al.* 1994).  $N_e$  was taken as 253,000 individuals for least auklets and 375,000 individuals for crested auklets (the average of four estimates of long-term  $N_e$  based on neutral markers, for each species; see RESULTS). Thus, the average time to coalescence of two alleles for exon 2 of the class II MHC B gene is  $1.9 \times 10^6$  generations ( $12.6 \times 10^6$  years) for least auklets and  $2.4 \times 10^6$  generations ( $20.3 \times 10^6$  years) for crested auklets. This result is consistent with the observation that some human allelic lineages at this locus are >30 million years old (O'HUIGIN *et al.* 1991; J. KLEIN *et al.* 1993). Given this result, most extant auklet alleles at this locus predate the divergence of least and crested auklets.

2

To obtain a more conservative estimate of the number of alleles at this locus that predate the speciation event, we considered the observation that for some species that diverged within the last 5 MY [Darwin's finches (VINCEK *et al.* 1997) and several extant primate species (J. KLEIN *et al.* 1993)], no less than 62% of the allelic lineages present today (for MHC II B exon 2) also were present prior to speciation. Given this generalization, we estimated that of the (minimum) 40 alleles currently present in each of least and crested auklets, 25 (~62%) were present prior to speciation.

We ran the computer simulation to determine the minimum number of founder individuals that would allow persistence of 40 and 25 alleles for the second exon of an MHC II B locus. The ultimate parental population was assumed to have the same number and frequency of alleles as a known human (Caucasoid) population (as in VINCEK *et al.* 1997). Simulations were run under conditions of immediate expansion following an initial bottleneck of size  $N_b$  and expansion following a bottleneck 10 generations in duration. As in VINCEK *et al.* (1997), two growth rates were applied for each of the expansion scenarios (5 and 50% per generation), and an additional case, that of no expansion (growth = 0%), also was considered.

*Historical demographic events inferred from mismatch distributions:* ROGERS and HARPENDING (1992) argued that the distribution of pairwise nucleotide differences between individuals in a population sample (the mismatch distribution) carries information on the demographic history of the population (but see DI RIENZO and WILSON 1991; LUNDSTROM *et al.* 1992). We plotted the mismatch distributions for cytochrome *b* and  $\alpha$ -enolase intron VIII for our sample of least and crested auklets to judge whether evidence exists for a population expansion in either species. Comparisons between all pairs of individuals were conducted using Arlequin ver. 2.000 (SCHNEIDER *et al.* 2000). Ninety-five percent confidence intervals on the raggedness statistics ( $r$ ) (HARPENDING 1994) were generated in DnaSP v 4.0 (ROZAS *et al.* 2003) by 5000 replicates in coalescent simulations based on the observed genetic diversities and samples sizes, under no recombination and under intermediate recombination ( $R = 10$ ).

To address some of the assumptions of the above methods, tests of intraspecies population differentiation and of deviations from Hardy-Weinberg equilibrium were conducted using GENEPOP (v. 3.1d) (RAYMOND and ROUSSET 1995) and Arlequin (data given in WALSH and FRIESEN 2003).

## RESULTS

Estimates of  $\theta$  and  $\Theta$  and corresponding values for  $N_e$  or  $N_f$  derived from variation in cytochrome *b* and the  $\alpha$ -enolase intron indicated that population sizes for the auklets have been large throughout the course of their evolution (minimum 201,000 and 220,000 individuals for least and crested auklets, respectively; Tables 1 and 2). Timescales over which estimates of  $\theta$  (or  $\Theta$ ) and  $N$  apply varied from  $5.1 \times 10^5$  to  $1.1 \times 10^7$  years (Tables 1 and 2). Fixing divergence time at 2.8 MYA gave a maximum-likelihood value of  $N_a = 3.4 \times 10^4$  individuals (95% C.I. ~0.0– $1.9 \times 10^6$ ; maximum  $L = -1.69$ ) by the method of TAKAHATA *et al.* (1995).

Equating  $S_S$  with WAKELEY and HEY's (1997) Equation 12 and substituting estimated  $\theta_W$  values for  $\theta_1$  and  $\theta_2$  gave a value of 0.256 for  $\theta_A$  at the  $\alpha$ -enolase locus. The

**TABLE 1**  
**Estimates of  $\theta$  and  $\Theta$  and corresponding values for  $N_e$  or  $N_f$  for least auklets, based on data for cytochrome *b* and  $\alpha$ -enolase intron VIII**

Estimator	$\hat{\theta}$ or $\hat{\Theta}$	$N_f$	Generations (yr <sup>a</sup> ) before present throughout which estimates apply
Cytochrome <i>b</i>			
$\Pi$ (TAJIMA 1983)	2.00	~48,000	$\sim 7.6 \times 10^4$ ( $5.1 \times 10^5$ )
$\theta_W$ (WATTERSON 1975)	2.40	~57,000	$\sim 9.1 \times 10^4$ ( $6.1 \times 10^5$ )
$\Theta_P$ (KUHNER <i>et al.</i> 1995)	0.012	~87,000	Contemporary $N_f$
$\Theta_{ML}$ (TAKAHATA <i>et al.</i> 1995)	0.0065	~47,000	$\sim 7.6 \times 10^4$ ( $5.1 \times 10^5$ )
Estimator	$\hat{\theta}$ or $\hat{\Theta}$	$N_e$	Generations (yr <sup>a</sup> ) before present throughout which estimates apply
$\alpha$ -Enolase			
$\Pi$ (TAJIMA 1983)	3.08	~208,000	$\sim 7.9 \times 10^5$ ( $5.3 \times 10^6$ )
$\theta_W$ (WATTERSON 1975)	4.29	~290,000	$\sim 1.1 \times 10^6$ ( $7.4 \times 10^6$ )
$\Theta_P$ (KUHNER <i>et al.</i> 1995)	0.043	~720,000	Contemporary $N_e$
$\Theta_{ML}$ (TAKAHATA <i>et al.</i> 1995)	0.012	~201,000	$\sim 7.6 \times 10^5$ ( $5.1 \times 10^6$ )
Trans-species polymorphism	NA <sup>b</sup>	313,000	$4.2 \times 10^5$ ( $2.8 \times 10^6$ )

<sup>a</sup> Given a generation time of 6.7 years for least auklets (WALSH and FRIESEN 2003).

<sup>b</sup> Not applicable.

corresponding value for the effective size of the ancestral population was 16,000 individuals. Substituting values obtained from other estimators of  $\theta$  (Tables 1 and 2) for  $\theta_1$  and  $\theta_2$  produced higher values for  $N_a$  ( $\sim 117,000$  using  $\Pi$  and  $\sim 123,000$  using  $\theta_{ML}$ ).

According to the phylogeny of least and crested auklet alleles for  $\alpha$ -enolase intron VIII (Figure 1), seven allelic lineages predate the divergence of the two species, three of which have descendants in our sample of modern least auklets and all seven of which have descendants in our sample of modern crested auklets. The

maximum of Equation 1 occurred at  $t' = 0.673$  for least auklets ( $k = 3$ ,  $n = 19$ ) and at  $t' = 0.243$  for crested auklets ( $k = 7$ ,  $n = 33$ ), giving  $N_e$ 's of  $3.13 \times 10^5$  and  $6.99 \times 10^5$  individuals, respectively. These are long-term  $N_e$ 's for the last 2.8 MY (*i.e.*, since divergence of the lineages leading to the two species; Tables 1 and 2).

Results of the computer simulations (Figure 2) showed that the minimum population size required to maintain 25 MHC class II B allelic lineages through a founder event was 85 individuals (under simulation of immediate expansion from bottleneck size, with a population

**TABLE 2**  
**Estimates of  $\theta$  and  $\Theta$  and corresponding values for  $N_e$  or  $N_f$  for crested auklets, based on data for cytochrome *b* and  $\alpha$ -enolase intron VIII**

Estimator	$\hat{\theta}$ or $\hat{\Theta}$	$N_f$	Generations (yr <sup>a</sup> ) before present throughout which estimates apply
Cytochrome <i>b</i>			
$\Pi$ (TAJIMA 1983)	2.61	~52,000	$\sim 9.1 \times 10^4$ ( $7.6 \times 10^5$ )
$\theta_W$ (WATTERSON 1975)	3.47	~69,000	$\sim 1.2 \times 10^5$ ( $1.0 \times 10^6$ )
$\Theta_P$ (KUHNER <i>et al.</i> 1995)	0.017	~104,000	Contemporary $N_f$
$\Theta_{ML}$ (TAKAHATA <i>et al.</i> 1995)	0.0098	~60,000	$\sim 1.0 \times 10^5$ ( $8.7 \times 10^5$ )
Estimator	$\hat{\theta}$ or $\hat{\Theta}$	$N_e$	Generations (yr <sup>a</sup> ) before present throughout which estimates apply
$\alpha$ -Enolase			
$\Pi$ (TAJIMA 1983)	3.96	~220,000	$\sim 8.5 \times 10^5$ ( $7.1 \times 10^6$ )
$\theta_W$ (WATTERSON 1975)	6.41	~360,000	$\sim 1.4 \times 10^6$ ( $1.1 \times 10^7$ )
$\Theta_P$ (KUHNER <i>et al.</i> 1995)	0.092	~1,300,000	Contemporary $N_e$
$\Theta_{ML}$ (TAKAHATA <i>et al.</i> 1995)	0.016	~220,000	$\sim 8.6 \times 10^5$ ( $7.1 \times 10^6$ )
Trans-species polymorphism	NA <sup>a</sup>	699,000	$3.4 \times 10^5$ ( $2.8 \times 10^6$ )

<sup>a</sup> Given a generation time of 8.3 years for crested auklets (WALSH and FRIESEN 2003).

<sup>b</sup> Not applicable.

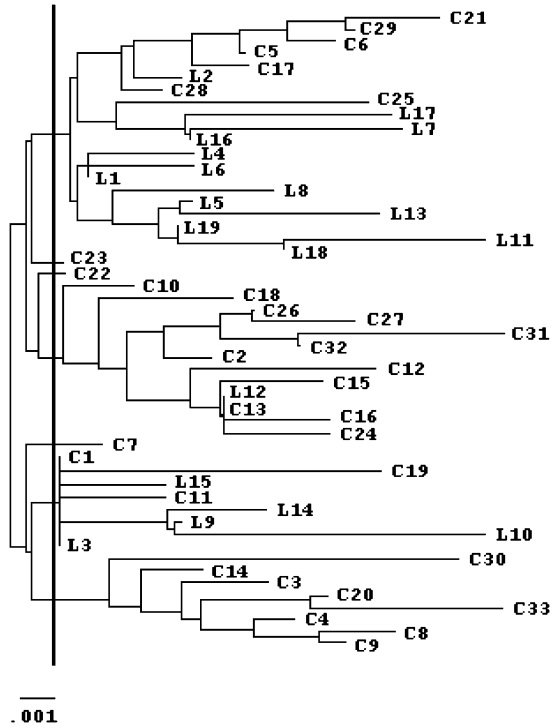


FIGURE 1.—Neighbor-joining tree of  $\alpha$ -enolase alleles for least (L) and crested (C) auklets, based on Jukes-Cantor distances. Allelic lineages present prior to the speciation event ( $n = 7$ ) are those that cross the vertical line indicated in bold-face type (which marks the estimated divergence time of lineages leading to modern least and crested auklets). L12 and C13 are identical, as are C1 and L3.

growth rate of 50% per generation). The largest estimate was 650 founder individuals (under simulation of expansion with a growth rate of 5% following a 10-generation bottleneck). For retention of all or almost all (90% of) 40 alleles through an initial founder event, 740 individuals were required under immediate expansion with growth rate 50% and >2000 individuals were required under expansion with growth rate 5% after

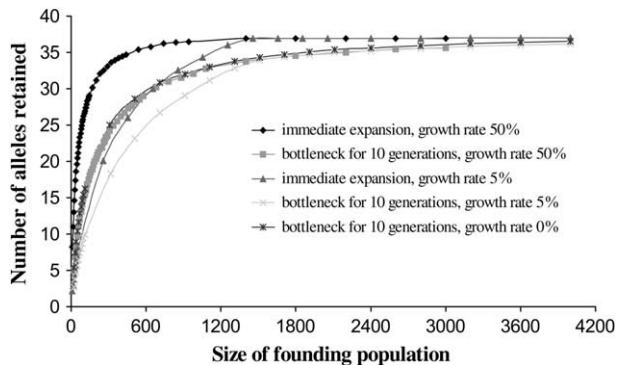


FIGURE 2.—The number of alleles retained in a population as a function of founder population size, after a 0- or 10-generation bottleneck followed by 0, 5, or 50% per generation population growth (from computer simulations). Bottleneck duration and subsequent population growth rate are indicated for each curve.

10 bottlenecked generations. Under a scenario of no expansion, 310 and >2000 individuals were required for retention of 25 and 40 alleles, respectively. In all cases, allelic loss was minimal once the size of the population was  $> \sim 2000$  individuals.

None of the frequency distributions of pairwise nucleotide differences for cytochrome *b* and  $\alpha$ -enolase intron VIII in least and crested auklets (Figure 3) displayed a unimodal shape characteristic of a population expansion following an extended bottleneck, nor were many peaks observed at high numbers of differences, as expected for a brief bottleneck (ROGERS and HARPENDING 1992). Raggedness statistics ( $r$ ; Figure 3) all were within 95% confidence intervals for  $r$  under coalescent simulations given observed sample sizes and nucleotide diversities, under conditions of both no recombination and intermediate recombination.

Exact tests of population differentiation showed no significant heterogeneity in haplotype or genotype distributions among populations for either cytochrome *b* or  $\alpha$ -enolase ( $P > 0.10$ ). Similarly, pairwise  $F_{ST}$  values detected no significant genetic structure among populations (all  $P > 0.05$ ), and no significant deviations from Hardy-Weinberg equilibrium were observed (all  $P > 0.21$ ).

## DISCUSSION

Population sizes unequivocally have been large for least and crested auklets throughout the course of their evolution. This conclusion is consistent across loci and estimation methods (Tables 1 and 2). The absence of detectable population bottlenecks during the evolutionary history of the auklet species implies that founder effect could not have been a causal factor in their speciation. Large values for estimates of  $N_e$  at different depths in their history indicate that the observed genetic variation has been maintained over the long term and is not a product of a recent population expansion. Absence of detectable genetic structure among populations indicates that estimated  $N_e$  values are unlikely to have been inflated due to population subdivision.

Due to its haploid state and primarily maternal transmission,  $N_e$  for mtDNA (*i.e.*,  $N_f$ ) is expected to be approximately one-quarter that for nuclear DNA (TAJIMA 1983). Effective population sizes estimated from mtDNA in our study are consistent with this expectation (Tables 1 and 2).

TAJIMA (1996) showed that, given sufficient substitution rate variation across sites, estimates of  $\theta$  based on the infinite-sites model will be substantial underestimates. Although site-specific substitution rate variation was not addressed in our work, such rate variation is observed frequently for samples of DNA sequences (*e.g.*, WAKELEY 1993; VOELKER and EDWARDS 1998), and TAJIMA's (1996) conclusion makes our estimates of  $N_e$  more conservative.

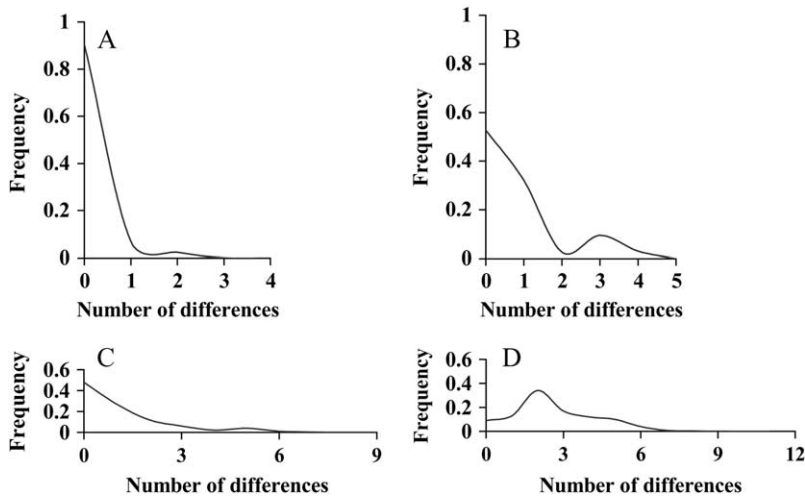


FIGURE 3.—Frequency distributions of nucleotide differences among pairs of individuals in population samples of least and crested auklets. (A) Cytochrome *b* alleles among least auklets [ $r = 0.704$ ; raggedness statistic (HARPENDING 1994)]; (B) cytochrome *b* alleles among crested auklets ( $r = 0.137$ ); (C)  $\alpha$ -enolase alleles among least auklets ( $r = 0.0926$ ); (D)  $\alpha$ -enolase alleles among crested auklets ( $r = 0.0472$ ).

The estimates of  $N_e$  in Tables 1 and 2 assume a single panmictic population in mutation-drift equilibrium. To address these assumptions, pairwise  $F_{ST}$  values and raggedness statistics on mismatch distributions were computed.  $F_{ST}$  values did not reveal any significant genetic structure among populations, suggesting panmixia, and  $r$ -values of all mismatch distributions fell within 95% confidence intervals of  $r$  estimated by coalescent simulations (Figure 3), suggesting equilibrium between mutation and genetic drift.

Our estimate of  $N_a$  derived by the method of TAKAHATA *et al.* (1995) may be biased in that it is based on data for a single locus, using a method that originally was developed to treat data for multiple loci (specifically, the variation of sequence divergences among loci). This may explain the width of the confidence interval on our estimate of  $N_a$  by this method (YANG 1997). However, we do not expect among-locus variation in sequence divergences for these species to be very high because their recent divergence (with respect to effective population sizes, so much so that trans-species polymorphism is observed) predicts low divergence at all loci. Therefore, per locus sequence divergence for least and crested auklets may not be as biased by factors such as substitution rate variation among loci (YANG 1997), and single-locus estimates of ancestral parameters may be more robust for these species than they would be for species whose divergence times are older.

Bias is introduced into the estimation of  $\theta_A$  when using empirically derived values for  $\theta_1$  and  $\theta_2$  in the evaluation of WAKELEY and HEY's (1997) Equation 12. Watterson's  $\theta$ , for example, assumes constant population size through time, a hypothesis that has not been rejected for these birds, but that remains an assumption by virtue of it being a null hypothesis. The use of different  $\theta$ -estimators as variables in this equation suggests, however, that  $\theta_W$ -values are underestimates (see RESULTS). Thus, our estimates of  $\theta_A$  are conservatively low within the framework of testing the plausibility of founder effect as a mechanism for speciation of the auklets.

If fewer than seven allelic lineages for  $\alpha$ -enolase intron VIII actually predate the speciation event (as may occur either if divergence of the two species is older than estimated or due to error associated with phylogeny estimation), the resulting  $N_e$  is reduced, although still sufficiently large to reject the hypothesis of founder effect speciation in the auklets. For example, if  $k = 3$  for crested auklets, the maximum of Equation 1 occurs approximately at  $t' = 0.8$ , and if  $k = 2$  (the minimum value that  $k$  can take because trans-species polymorphism is observed), the maximum of (1) occurs approximately at  $t' = 1.2$ . Given these values, estimates of  $N_e$  are  $\sim 212,000$  and  $\sim 142,000$  individuals, respectively.

The computer simulations in this study were based on conservative scenarios for the duration of population bottlenecks and the rate of expansion following the bottleneck phase and consider the minimum amount of variability at an MHC II B locus for least and crested auklets. Because single-generation bottlenecks allow retention of a greater number of alleles than do multi-generation bottlenecks, the estimates derived from computer simulation represent the minimum number of individuals required in a bottlenecked founder population. Given population bottlenecks of  $>10$  generations, even larger population sizes would be required for retention of the observed number of ancient alleles.

The assumption of the computer model that the ancestral auklet population contained the same number and frequencies of alleles as in the current Caucoid population is probably unrealistic, although less so than in the original simulations of VINCEK *et al.* (1997) due to greater similarity of population size estimates for auklets and humans than for Darwin's finches and humans (VINCEK *et al.* 1997). The model also assumes that mating is random and that rates of population growth are constant through time, and thus it ignores fluctuations due to sudden population explosions and/or catastrophes in addition to the original bottleneck; however, these assumptions are typical of similar models (*e.g.*, KUO and JANZEN 2004) and are necessary to assure

feasibility of the computer analysis. The intensity of selection at MHC II B loci in least and crested auklets is unknown and the value of  $s$  (0.01) used in the computer model may not be accurate. However, empirical data for humans (SATA *et al.* 1994) indicate that  $s$  for exon 2 of MHC II B loci is generally less than or equal to this value. Because balancing selection acts to maintain more variation than would be maintained by drift alone, a smaller value for  $s$  (*i.e.*, reduced selection intensity) would necessitate larger population sizes through bottlenecks to allow persistence of the same amount of variation.

Our simulations suggest that throughout the evolutionary history of least and crested auklets and their immediate ancestor, no population size reductions to fewer than several thousand individuals occurred. This result is similar to that obtained for human populations over the past several million years (AYALA and ESCALANTE 1996). In addition, mismatch distributions for the auklets do not fit models of population expansion subsequent to either long or brief bottlenecks. The above results suggest that the demographic history of the auklets has been relatively stable throughout the past several million years. For effective population sizes to be of the orders of magnitude reported in Tables 1 and 2, no dramatic bottlenecks, even for a single generation, could have occurred in the evolutionary histories of least and crested auklets. To have maintained observed levels of variability, even population reductions of moderate size lasting many generations are unlikely.

Methods for estimating ancestral population parameters are in an early stage of development. As such, estimates presented herein incorporate the biases of their respective models (*e.g.*, model assumptions described above). Future development of more realistic models and their application to data for a large number of unlinked loci will allow more accurate estimation of ancestral population sizes (RANNALA and YANG 2003). Despite uncertainties in the estimation methods applied herein, results across multiple estimation methods and three types of molecular markers provide unambiguous evidence that auklets did not speciate as a consequence of founder effect. Alternative explanations for both the relatively short internodes that separated divergence events among species in this group and their mode of speciation should focus on the demonstrated strong and mutual sexual selection exhibited by these species (JONES and HUNTER 1993).

Though hypothesized to be a plausible and even widespread mechanism of speciation, to date available data do not support the hypothesis that speciation by founder effect has occurred commonly in vertebrates. Whether this is due to a deficit of relevant data or to lack of the effect itself is difficult to determine. A systematic evaluation across many vertebrate taxa, especially those expected to have experienced a history of reduced population size (*e.g.*, island species, refugial populations, species in areas of wide ecological gradients), is required

before generalizations can be established regarding the overall importance of this potential mechanism of speciation.

We are grateful to C. O'hUigin for providing the original source code for the simulation and to S. Uellenberg for assistance in compiling and running many versions of the algorithm. We thank J. Felsenstein, P. Beerli, and the members of the Friesen, Rohwer, and Swanson lab groups, who participated in helpful discussions and commented on the manuscript. We are indebted to Y. Artukhin, G. V. Byrd, L. Climo, S. Ebbert, J. Piatt, L. Scharf, D. Siegel-Causey, A. Sowls, S. Stephensen, J. Wachtel, J. Williams, and the crew of R/V Tiglax for their generous assistance with tissue collections and to F. Tajima and the anonymous reviewers who provided valuable comments on the manuscript. Financial and/or logistic support for this study was provided by the Natural Sciences and Engineering Research Council of Canada (NSERC discovery grant held by V.L.F. and postgraduate scholarships to H.E.W.); the Northern Sciences Training Program; the U.S. Fish and Wildlife Service, Alaska Maritime National Wildlife Refuge, Bering Sea and Aleutian Islands units; and the Leaders Five endowed fellowship of the University of Washington Burke Museum. 3

#### LITERATURE CITED

- ABBOTT, C. L., and M. C. DOUBLE, 2003 Genetic structure, conservation genetics and evidence of speciation by range expansion in shy and white-capped albatrosses. *Mol. Ecol.* **12**: 2953–2962.
- AVISE, J. C., and D. WALKER, 1998 Pleistocene phylogeographic effects on avian populations and the speciation process. *Proc. R. Soc. Lond. Ser. B* **265**: 457–463.
- AYALA, F. J., and A. A. ESCALANTE, 1996 The evolution of human populations: a molecular perspective. *Mol. Phylogenet. Evol.* **5**: 188–201.
- BARTON, N. H., and B. CHARLESWORTH, 1984 Genetic revolutions, founder effects, and speciation. *Annu. Rev. Ecol. Syst.* **15**: 133–164.
- CARSON, H. L., 1983 Chromosomal sequences and interisland colonizations in Hawaiian *Drosophila*. *Genetics* **103**: 465–482.
- CARSON, H. L., and A. R. TEMPLETON, 1984 Genetic revolutions in relation to speciation phenomena: the founding of new populations. *Annu. Rev. Ecol. Syst.* **15**: 97–131.
- CLEGG, S. M., S. M. DEGNAN, J. KIKKAWA, C. MORITZ, A. ESTOUP *et al.*, 2002 Genetic consequences of sequential founder events by an island-colonizing bird. *Proc. Natl. Acad. Sci. USA* **99**: 8127–8132.
- COYNE, J. A., and H. A. ORR, 2004 *Speciation*. Sinauer Associates, Sunderland, MA.
- DI RIENZO, A., and A. C. WILSON, 1991 Branching pattern in the evolutionary tree for human mitochondrial DNA. *Proc. Natl. Acad. Sci. USA* **88**: 1597–1601.
- FELSENSTEIN, J., 1992 Estimating effective population size from samples of sequences: a bootstrap Monte Carlo integration method. *Genet. Res.* **60**: 209–220.
- FRIESEN, V. L., A. J. BAKER and J. F. PIATT, 1996 Phylogenetic relationships within the Alcidae (Charadriiformes: Aves) inferred using total molecular evidence. *Mol. Biol. Evol.* **13**: 359–367.
- FRIESEN, V. L., B. C. CONGDON, H. E. WALSH and T. P. BIRT, 1997 Intron variation in marbled murrelets detected using analyses of single-stranded conformational polymorphisms. *Mol. Ecol.* **6**: 1047–1058.
- FU, Y.-X., 1994a A phylogenetic estimator of effective population size or mutation rate. *Genetics* **136**: 685–692.
- FU, Y.-X., 1994b Estimating effective population size or mutation rate using the frequencies of mutations of various classes in a sample of DNA sequences. *Genetics* **138**: 1375–1386.
- GASTON, A. J., and I. L. JONES, 1998 *The Auks: Alcidae*. Oxford University Press, Oxford.
- GAVRILETS, S., 2004 *Fitness Landscapes and the Origin of Species*. Princeton University Press, Princeton, NJ.
- GAVRILETS, S., and C. R. B. BOAKE, 1998 On the evolution of pre-mating isolation after a founder event. *Am. Nat.* **152**: 706–716.

- GAVRILETS, S., and A. HASTINGS, 1996 Founder effect speciation: a theoretical reassessment. *Am. Nat.* **147**: 466–491.
- GRIFFITHS, R. C., and S. TAVARÉ, 1993 Sampling theory for neutral alleles in a varying environment. *Philos. Trans. R. Soc. Lond. Ser. B* **344**: 403–410.
- HARPENDING, H. C., 1994 Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Hum. Biol.* **66**: 591–600.
- HEWITT, G., 1996 Some genetic consequences of ice ages, and their role in divergence and speciation. *Biol. J. Linn. Soc.* **58**: 247–276.
- HOLDER, K., R. MONTGOMERIE and V. L. FRIESEN, 1999 A test of the glacial refugium hypothesis using patterns of mitochondrial and nuclear DNA sequence variation in rock ptarmigan (*Lagopus mutus*). *Evolution* **53**: 1936–1950.
- HOLLOCHER, H., 1996 Island hopping in *Drosophila*: patterns and processes. *Philos. Trans. R. Soc. Lond. B* **351**: 735–743.
- JOHNSON, T. C., C. A. SCHOLZ, M. R. TALBOT, K. KELTS, R. D. RICKETTS *et al.*, 1996 Late Pleistocene desiccation of Lake Victoria and rapid evolution of cichlid fishes. *Science* **273**: 1091–1093.
- JONES, I. L., 1999 Assessing the role of sexual selection in adaptive radiation of the auklets (Alcidae, Aethiini). Proceedings of the 22nd International Ornithological Congress, edited by N. ADAMS and R. SLOTOW. University of Natal, Durban, South Africa, pp. ■■■.
- JONES, I. L., and F. M. HUNTER, 1993 Mutual sexual selection in a monogamous seabird. *Nature* **362**: 238–239.
- JUKES, T. H., and C. R. CANTOR, 1969 Evolution of protein molecules, pp. 21–132 in *Mammalian Protein Metabolism*, edited by H. N. MUNRO. Academic Press, New York.
- KIMURA, M., 1969 The number of heterozygous nucleotide sites maintained in a finite population due to steady flux of mutations. *Genetics* **61**: 893–903.
- KINGMAN, J. F. C., 1982a The coalescent. *Stoch. Proc. Appl.* **13**: 235–248.
- KINGMAN, J. F. C., 1982b On the genealogy of large populations. *J. Appl. Probab.* **19A**: 27–43.
- KLEIN, D., H. ONO, C. O'HUIGIN, V. VINCEK, T. GOLDSCHMIDT *et al.*, 1993 Extensive *Mhc* variability in cichlid fishes of Lake Malawi. *Nature* **364**: 330–334.
- KLEIN, J., Y. SATTI, N. TAKAHATA and C. O'HUIGIN, 1993 Trans-specific *Mhc* polymorphism and the origin of species in primates. *J. Med. Primatol.* **22**: 57–64.
- KLICKA, J., and R. M. ZINK, 1997 The importance of recent ice ages in speciation: a failed paradigm. *Science* **277**: 1666–1669.
- KLICKA, J., and R. M. ZINK, 1999 Pleistocene effects on North American songbird evolution. *Proc. R. Soc. Lond. Ser. B* **266**: 695–700.
- KUHNER, M. K., J. YAMATO and J. FELSENSTEIN, 1995 Estimating effective population size and mutation rate from sequence data using Metropolis-Hastings sampling. *Genetics* **140**: 1421–1430.
- KUMAR, S., K. TAMURA and M. NEI, 1993 *MEGA: Molecular Evolutionary Genetics Analysis (v. 1.0)*. Pennsylvania State University, University Park, PA.
- KUO, C.-H., and F. J. JANZEN, 2004 Genetic effects of a persistent bottleneck on a natural population of ornate box turtles (*Terrapene ornata*). *Conserv. Genet.* **5**: 425–437.
- LIEBERS, D., A. J. HELBIG and P. DE KNIJFF, 2001 Genetic differentiation and phylogeography of gulls in the *Larus cachinnans-fuscus* group (Aves: Charadriiformes). *Mol. Ecol.* **10**: 2447–2462.
- LUNDSTROM, R., S. TAVARÉ and R. H. WARD, 1992 Modeling evolution of the human mitochondrial genome. *Math. Biosci.* **112**: 319–335.
- MAYR, E., 1942 *Systematics and the Origin of Species*. Columbia University Press, New York.
- McKINNON, J. S., and H. D. RUNDLE, 2002 Speciation in nature: the threespine stickleback model systems. *Trends Ecol. Evol.* **17**: 480–488.
- MEYER, A., T. D. KOCHER, P. BASASIBWAKI and A. C. WILSON, 1990 Monophyletic origin of Lake Victoria cichlid fishes suggested by mitochondrial DNA sequences. *Nature* **347**: 550–553.
- MOUM, T., S. JOHANSEN, K. E. ERIKSTAD and J. F. PIATT, 1994 Phylogeny and evolution of the auks (subfamily Alcinae) based on mitochondrial DNA sequences. *Proc. Natl. Acad. Sci. USA* **91**: 7912–7916.
- NEI, M., 1987 *Molecular Evolutionary Genetics*. Columbia University Press, New York.
- O'REILLY, P., T. E. REIMCHEN, R. BEECH and C. STROBECK, 1993 Mitochondrial DNA in *Gasterosteus* and Pleistocene glacial refugium on the Queen Charlotte Islands, British Columbia. *Evolution* **47**: 679–684.
- POWELL, J. R., 1989 The effects of founder-flush cycles on ethological isolation in laboratory populations of *Drosophila*, pp. 239–251 in *Genetics, Speciation and the Founder Principle*, edited by L. V. GIDDINGS, K. Y. KANESHIRO and W. W. ANDERSON. Oxford University Press, New York.
- PROVINE, W. B., 1989 Founder effects and genetic revolutions in microevolution and speciation: an historical perspective, pp. 43–76 in *Genetics, Speciation and the Founder Principle*, edited by L. V. GIDDINGS, K. Y. KANESHIRO and W. W. ANDERSON. Oxford University Press, New York.
- RANNALA, B., and Z. H. YANG, 2003 Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. *Genetics* **164**: 1645–1656.
- RAYMOND, M., and F. ROUSSET, 1995 GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *J. Hered.* **86**: 248–249.
- ROGERS, A. R., 1995 Genetic evidence for a Pleistocene population explosion. *Evolution* **49**: 608–615.
- ROGERS, A. R., and H. C. HARPENDING, 1992 Population growth makes waves in the distribution of pairwise genetic differences. *Mol. Biol. Evol.* **9**: 552–569.
- ROZAS, J., J. C. SÁNCHEZ-DELBARRIO, X. MESSEGUER and R. ROZAS, 2003 DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**: 2496–2497.
- RUNDLE, H. D., 2003 Divergent environments and population bottlenecks fail to generate premating isolation in *Drosophila pseudoobscura*. *Evolution* **57**: 2557–2565.
- RUNDLE, H. D., A. O. MOOERS and M. C. WHITLOCK, 1998 Single founder-flush events and the evolution of reproductive isolation. *Evolution* **52**: 1850–1855.
- SATTA, Y., C. O'HUIGIN, N. TAKAHATA and J. KLEIN, 1994 Intensity of natural selection at the major histocompatibility complex loci. *Proc. Natl. Acad. Sci. USA* **91**: 7184–7188.
- SCHNEIDER, S., D. ROESSLI and L. EXCOFFIER, 2000 *Arlequin: A Software for Population Genetic Data Analysis*. Genetics and Biometry Laboratory, University of Geneva, Geneva.
- STEPHENSEN, S. W., and D. B. IRONS, 2003 Comparison of colonial breeding seabirds in the eastern Bering Sea and Gulf of Alaska. *Mar. Ornithol.* **31**: 167–173.
- TAJIMA, F., 1983 Evolutionary relationship of DNA sequences in finite populations. *Genetics* **105**: 437–460.
- TAJIMA, F., 1996 The amount of DNA polymorphism maintained in a finite population when the neutral mutation rate varies among sites. *Genetics* **143**: 1457–1465.
- TAKAHATA, N., 1989 Gene genealogy in three related populations: consistency probability between gene and population trees. *Genetics* **122**: 957–966.
- TAKAHATA, N., 1990 A simple genealogical structure of strongly balanced allelic lines and trans-species evolution of polymorphism. *Proc. Natl. Acad. Sci. USA* **87**: 2419–2423.
- TAKAHATA, N., 1991 Trans-species polymorphism of HLA molecules, founder principle, and human evolution, pp. 29–49 in *Molecular Evolution of the Major Histocompatibility Complex*, edited by J. KLEIN and D. KLEIN. Springer-Verlag, New York.
- TAKAHATA, N., 1993 Allelic genealogy and human evolution. *Mol. Biol. Evol.* **10**: 2–22.
- TAKAHATA, N., Y. SATTI and J. KLEIN, 1995 Divergence time and population size in the lineage leading to modern humans. *Theor. Popul. Biol.* **48**: 198–221.
- TAVARÉ, S., 1984 Lines-of-descent and genealogical processes, and their applications in population genetics models. *Theor. Popul. Biol.* **26**: 119–164.
- TEMPLETON, A. R., 1986 The relation between speciation mechanisms and macroevolutionary patterns, pp. 497–512 in *Evolutionary Processes and Theory*, edited by S. KARLIN and E. NEVO. Academic Press, New York.
- TEMPLETON, A. R., 1996 Experimental evidence for the genetic-transience model of speciation. *Evolution* **50**: 909–915.
- TURELLI, M., N. H. BARTON and J. A. COYNE, 2001 Theory and speciation. *Trends Ecol. Evol.* **16**: 330–343.
- VINCEK, V., C. O'HUIGIN, Y. SATTI, N. TAKAHATA, P. T. BOAG, *et al.*, 1997 How large was the founding population of Darwin's finches? *Proc. R. Soc. Lond. Ser. B* **264**: 111–118.

- VOELKER, G., and S. V. EDWARDS, 1998 Can weighting improve bushy trees? Models of cytochrome *b* evolution and the molecular systematics of pipits and wagtails (Aves: Motacillidae). *Syst. Biol.* **47**: 589–603.
- WAKELEY, J., 1993 Substitution rate variation among sites in hyper-variable region 1 of human mitochondrial DNA. *J. Mol. Evol.* **37**: 613–623.
- WAKELEY, J., and J. HEY, 1997 Estimating ancestral population parameters. *Genetics* **145**: 847–855.
- WALL, J. D., 2003 Estimating ancestral population sizes and divergence times. *Genetics* **163**: 395–404.
- WALSH, H. E., 1999 Polymorphism at neutral and non-neutral loci in least and crested auklets (*Aethia* spp.) and implications for speciation. M.Sc. Thesis, Queen's University, Kingston, ON, Canada.
- WALSH, H. E., and V. L. FRIESEN, 2003 A comparison of intraspecific patterns of DNA sequence variation in mitochondrial DNA,  $\alpha$ -enolase, and MHC Class II B loci in auklets (Charadriiformes: Alcidae). *J. Mol. Evol.* **57**: 681–693.
- WALSH, H. E., M. G. KIDD, T. MOUM and V. L. FRIESEN, 1999 Polytomies and the power of phylogenetic inference. *Evolution* **53**: 932–937.
- WATTERSON, G. A., 1975 On the number of segregating sites in genetical models without recombination. *Theor. Popul. Biol.* **7**: 256–276.
- WRIGHT, S., 1969 *Evolution and the Genetics of Populations, Vol. 2: The Theory of Gene Frequencies*. University of Chicago Press, Chicago.
- YANG, Z., 1997 On the estimation of ancestral population sizes of modern humans. *Genet. Res.* **69**: 111–116.
- YANG, Z., 2002 Likelihood and Bayes estimation of ancestral population sizes in hominoids using data from multiple loci. *Genetics* **162**: 1811–1823.

Communicating editor: J. B. WALSH

**Genetics December 2005**  
**Author query sheet Walsh (43380)**

- 1 Please add Takahata (1986) to Literature Cited or delete from text.
- 2 Please add O'hUigin *et al.* 1991 to Literature Cited or delete from text.
- 3 Definition of AMNWR as meant?
- 4 Please provide page range for Jones 1999.