MOLECULAR ECOLOGY

Molecular Ecology (2013) **22**, 5922–5935 doi: 10.1111/mec.12525

# Heteropatric speciation in a duck, *Anas crecca*

KEVIN WINKER,*§ KEVIN G. MCCRACKEN,*† DANIEL D. GIBSON* and JEFFREY L. PETERS†‡§
*Department of Biology & Wildlife, University of Alaska Museum, 907 Yukon Drive, Fairbanks, AK 99775, USA, †Institute of Arctic Biology, University of Alaska, 902 N. Koyukuk Drive, Fairbanks, AK 99775, USA, ‡Department of Biological Sciences, Wright State University, 3640 Colonel Glenn Hwy, Dayton, OH 45435, USA

## Abstract

**Heteropatric differentiation is a mode of speciation with gene flow in which divergence occurs between lineages that are in sympatry and allopatry at different times during cyclic spatial movements. Empirical evidence suggests that heteropatric differentiation may prove to be common among seasonally migratory organisms. We examined genetic differentiation between the sedentary Aleutian Islands population of green-winged teal (*Anas crecca-nimia*) and its close migratory relative, the Eurasian, or Old World (OW), *Anas c. crecca* population, a portion of which passes through the range of *nimia* during its seasonal migrations. We also examined its relationship with the parapatric North American, New World (NW), *A. c. carolinensis* population. Sequence data from eight nuclear introns and the mtDNA control region showed that the *nimia-crecca* divergence occurred much more recently than the deeper *crecca-carolinensis* split (~83 000 years vs. ~1.1 Myr). Despite considerable spatial overlap between *crecca* and *nimia* during seasonal migration, three key predictions of heteropatric differentiation are supported: significant genetic divergence (overall mean $\Phi_{ST}$ = 0.07), low gene flow ($2N_em \sim 1.8$), and an effective population size in *nimia* that is not especially low ($N_e \sim 80\,000$ individuals). Similar levels of gene flow have come into *nimia* from *carolinensis*, but no detectable nuclear gene flow has gone out of *nimia* into either OW (*crecca*) or NW (*carolinensis*) populations. We infer that adaptations of these populations to local optima in different places (e.g. each matching their reproductive effort to different resource blooms) promote genetic isolation and divergence despite periods of sympatry between them, as the heteropatric model predicts.**

*Keywords*: birds, divergence, ecological speciation, leapfrog migration, population genetics, seasonal migration, speciation with gene flow

*Received 29 June 2013; revision received 31 August 2013; accepted 11 September 2013*

## Introduction

How speciation can occur with gene flow is an important research area, and it is clear that understanding attributes of the divergence process prior to the completion of speciation is critical (Rundle & Nosil 2005; Via & West 2008; Rundell & Price 2009; Pyron & Burbrink 2010; Nosil 2012). Heteropatric speciation is a mode of speciation with gene flow in which divergence occurs between lineages that are in sympatry and allopatry at different times during cyclic movements associated with seasonal migrations. Phenotypic evidence suggests that heteropatric differentiation may be common among seasonally migratory organisms, and although few relevant studies have yet focused on migratory lineages, these organisms offer diverse opportunities to better understand speciation with gene flow (Winker 2010). Seasonal migration increases the annual movements of animals, and in birds this behaviour has long been considered to inhibit divergence between populations (Montgomery 1896). Indeed, the increased movements inherent in a migratory life history strategy are associated with greater dispersal distances among migratory species relative to sedentary ones, and this does decrease the occurrence of population divergence (Paradis *et al.* 1998; Belliure *et al.* 2000). However, increased movements can

Correspondence: Kevin Winker, Fax: (907) 474 5469;
E-mail: kevin.winker@alaska.edu
§These authors contributed equally to this work.

make new environments accessible, and although these movements increase the likelihood of gene flow and decrease opportunities for strict allopatry between diverging lineages, there are nevertheless many examples of divergence occurring within migratory lineages (table 1 in Winker 2010). The driver of this heteropatric differentiation is thought to be divergent selection operating on ecological adaptations to exploit cyclically available resources that are heterogeneously distributed in space and time (Winker 2010).

A common pattern of differentiation among seasonal migrants is that of a migratory population crossing the range of another population in its seasonal movements. This often occurs when a population from higher latitudes migrates to wintering grounds at lower latitudes than another population. The geographic range of the less mobile population is transected or passed through twice each year during the seasonal migrations of the more mobile population. This phenomenon is called 'leapfrog migration' and is usually recognized through movements of populations that are morphologically distinguishable as distinct subspecies, but it also occurs among full species (Swarth 1920; Salomonsen 1955; Baker 1978; Lundberg & Alerstam 1986; Boland 1990; Kondo *et al*. 2008; Winker 2010). This leapfrog pattern is common among avian migrants, demonstrating that differentiation can progress directly in the path of high rates of movement, and the heteropatric model offers a framework for how this differentiation can accrue and persist.

The three hypotheses within the heteropatric model that are probably most relevant for leapfrog patterns of divergence occurring within species are (Winker 2010): (i) There should be genetic differentiation. (ii) Gene flow should be low. (iii) The effective population size ($N_e$) of the population being leaped will not be especially low; that is, allopatry is not effectively achieved through rarity. We test these hypotheses in the green-winged teal (*Anas crecca* sensu lato), in which a sedentary population (*A. c. nimia*), resident in the Aleutian Islands of western Alaska (Friedmann 1948), is passed through twice annually by individuals of the Eurasian nominate subspecies, *A. c. crecca* (Fig. 1). Nominate *crecca* occur in sometimes considerable numbers during

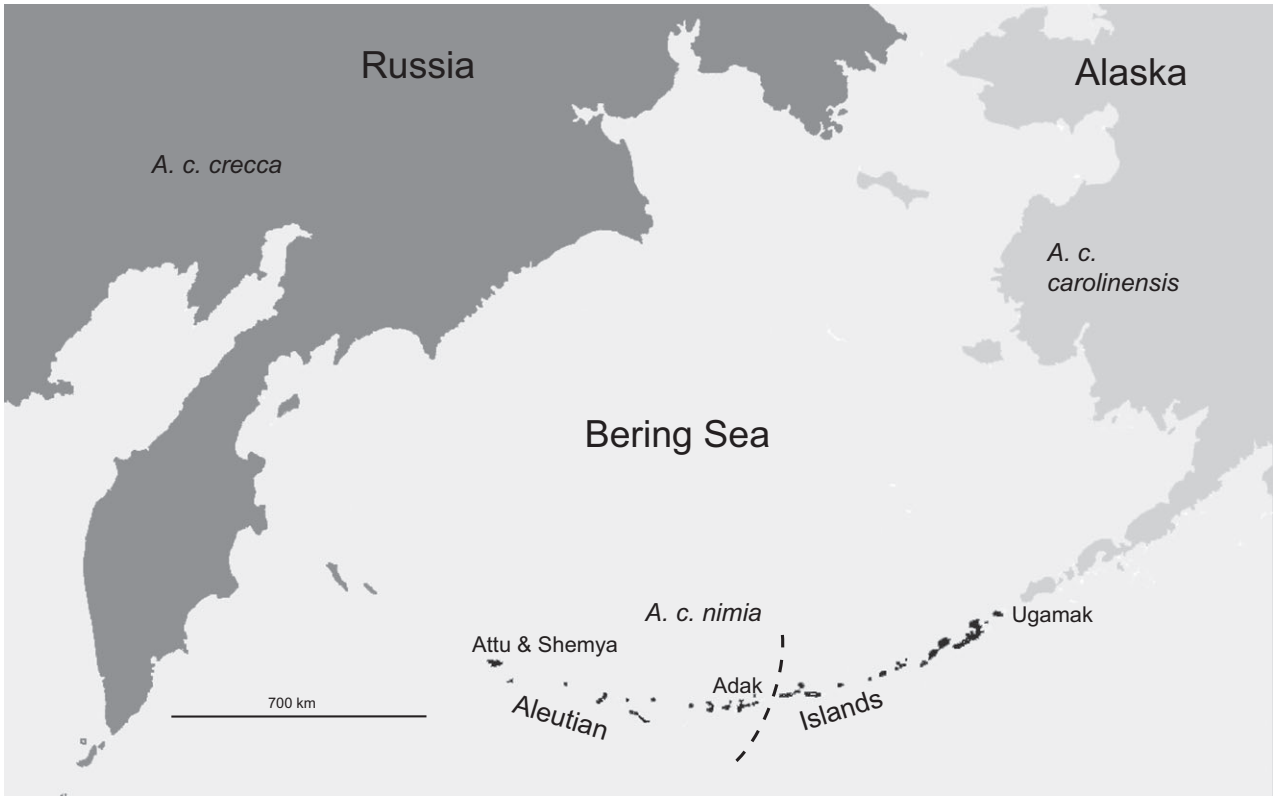


**Fig. 1** Bering Sea (North Pacific) region breeding ranges of green-winged teal (*Anas crecca* sensu lato), with subspecies *A. c. crecca* occurring mainly in Asia (dark grey; east in seasonal migration at least to the dashed line in the Aleutian Islands), subspecies *A. c. carolinensis* occurring in North America (pale grey) and subspecies *A. c. nimia* occurring as a year-round resident in the Aleutian Islands (black). Shemya Island is a small island (15.3 $km^2$) just 44 km from Attu Island. The dotted line in the Aleutians represents the easternmost routine occurrence of substantial numbers of *crecca* during seasonal migration.

spring and autumn migration in the western Aleutian Islands as these birds alternately go to seasonally occupied breeding grounds in northeast Asia and then return to wintering grounds elsewhere in eastern Asia (Gibson & Byrd 2007; Brazil 2009). In addition, the range of *nimia* abuts that of the migratory mainland North American *A. c. carolinensis* population, which is rare in the eastern Aleutians and intermittent in the western Aleutians—although intergrades are well known (i.e. males with plumage characteristics of both *carolinensis* and *crecca*/*nimia* or males without the identifying plumage characteristics of either), especially from the eastern Aleutians (Gibson & Byrd 2007). Thus, there is phenotypic evidence for gene flow into the Aleutian population from the North American population. Morphologically, adult-male breeding plumage suggests that the Aleutian form *nimia* is more closely related to the Eurasian form *crecca* than to the North American *carolinensis*. Another species, *Anas flavirostris*, which may be more closely related to *carolinensis* than the latter is to *crecca* (the nuclear and mitochondrial data on this relationship are in conflict; Peters *et al.* 2012a), occurs in South America entirely in allopatry; we do not include this species in this study.

North and east of the island range of *nimia*, the Beringia region (including the Bering land bridge and great expanses to the east and west) was unglaciated during the Quaternary (Kaufman & Manley 2004). Fossil and palynological evidence shows that rich vertebrate and plant communities existed across Beringia even at the height of the Wisconsinan glacial episode (110 000–10 000 years bp; Elias 2001; Harrington 2003; Abbott *et al.* 2010). The famous Pleistocene megafauna (mammoths, steppe bison, etc.) occupied Beringia at this time, and these larger vertebrates would have required free water to drink. It is likely that continental populations of *Anas crecca* seasonally migrated to and successfully bred in those same wetlands. Fossil evidence shows that Beringia has hosted a rich assemblage of migratory waterfowl since the early Quaternary (early Pleistocene), including *Anas crecca* specifically at >54 000 years and at mixed times during the Quaternary (Fitzgerald 1991). Here, we examine the relationships among all three subspecies (the sedentary *nimia* and the seasonally migratory *crecca* and *carolinensis*), testing the three hypotheses above regarding whether or not *nimia* is undergoing genomic differentiation as predicted under the heteropatric model.

## Methods

We collected 55 birds from the following Aleutian Islands (Fig. 1; Table S1, Supporting Information): Attu ($n$ = 24; 52°50′50″N 173°12′E), Shemya ($n$ = 11; 52°43′25″N 174°6′40′E), Adak ($n$ = 19; 51°52′50″N 176°39′W), and Ugamak ($n$ = 1; 54°11′59″N 164°48′25″W). Our sampling on Attu and Shemya islands was purposely conducted during migration periods, when we expected to obtain at least some continental *crecca* migrants, especially at Shemya, where numbers in migration can be an order of magnitude larger than the estimated number of breeding pairs (Schwitters 2008). Morphological and anatomical data were taken when these birds were prepared as museum specimens. In addition, we used the corresponding nuclear data sets for 25 individuals each of Eurasian *crecca* (Old World, or OW) and North American *carolinensis* (New World, or NW), spanning the continental distributions of each, reported in Peters *et al.* (2012a). The corresponding mtDNA data set included haplotypes from 58 *crecca* and 86 *carolinensis*.

We sequenced 979–982 bp of the mitochondrial (mt) DNA control region and the flanking tRNA-Phe. We also sequenced seven autosomal introns, including CRYAB (334 bp), GRIN1 (330–331 bp), ENO1 (309–313 bp), ODC1 (345–349 bp), PCK1 (343 bp), FGB (448–459 bp) and LDHB (531–534 bp), and one intron from the Z-chromosome, CHD1Z (308–310 bp). Primer sequences and PCR and sequencing protocols are described by Peters *et al.* (2012b). For the nuclear introns, we resolved the gametic phases of alleles using three methods. First, for sequences that were heterozygous for indels, we compared the ambiguous 3′-end (downstream of the indel) with the unambiguous 5′-end from the complementary sequence (in both forward and reverse directions) to determine the size and placement of indels. Because indels cause a shift in the peaks of chromatograms, we were able to determine the linkage of polymorphisms throughout the sequence using this method (Peters *et al.* 2007). Second, we used the program PHASE v. 2.1.1 (Stephens *et al.* 2001) to algorithmically determine the most likely allelic phases of sequences containing ≥2 polymorphic positions. For this analysis, we defined alleles resolved on the basis of indels as known alleles. We also included homologous sequences from 25 *crecca* and 25 *carolinensis* (Peters *et al.* 2012a). Input files were created using the program SEQPHASE (Flot 2010). Finally, for all sequences resolved with <0.95 probability, we used allele-specific priming to preferentially amplify and sequence one of the alleles (Bottema *et al.* 1993), and we subtracted this allele from the heterozygous sequence to determine the phase of the other allele.

### *Genetic differentiation*

We calculated $\Phi_{ST}$, the proportion of genetic variation partitioned between populations, using ARLEQUIN ver. 3.1 (Excoffier *et al.* 2005). We defined five populations for this analysis: three populations of *nimia* (Adak, Shemya,

and Attu), OW *crecca*, and NW *carolinensis*. Significance was tested with 10 000 permutations, and a Bonferroni correction was applied to each series of pairwise locus tests such that alpha was 0.005 (0.05/10 tests). We also used the clustering algorithm in STRUCTURE ver. 2.2.3 to test for population differences using multilocus nuclear genotypes (Pritchard *et al.* 2000). STRUCTURE's likelihood approach tests for deviations from Hardy–Weinberg equilibrium and linkage disequilibrium, calculates the likelihood of a user-defined number of populations, and assigns individuals to those populations. We numbered the alleles for each locus from 1 to $n$ ($n$ being the locus-specific number of different alleles observed). We used a no-admixture model assuming independent allele frequencies and tested models of varying numbers of populations ($K$ = 1–5), reporting results from the model with the highest likelihood (there were no appreciable differences when using the admixture model). Only nuclear loci were included, and STRUCTURE was run for 100 000 generations of burn-in and 500 000 generations of sampling. Each analysis was replicated 10 times. We chose the best value of $K$ using the method of Evanno *et al.* (2005) as implemented in STRUCTURE-HARVESTER (Earl & von Holdt 2012).

We calculated allelic richness for each locus using rarefaction to standardize the number of alleles for each population to the smallest sample size (Shemya; $n$ = 22 alleles for autosomal DNA and 18 for CHD1Z). Allelic richness was calculated using the program RAREFACTION CALCULATOR (University of Alberta; www.biology.ualberta.ca/jbrzusto/rarefact.php). We tested for significant differences in allelic richness among the five populations using a repeated measures ANOVA (with loci being the repeated variable) and a Tukey's post hoc test.

### *Population demography and gene flow*

We fit the sequence data to a three-population isolation-with-migration model in the program IMa2 (Hey 2010), using an input tree of ((*crecca*, *nimia*), *carolinensis*) based on morphology and mtDNA evidence. For *nimia*, based on STRUCTURE results and our intent to sample migratory *crecca* in the Aleutians, we excluded samples from Shemya and we lumped samples from Adak and Attu. We used IMa2 to estimate 15 parameters scaled to the per-locus mutation rate ($u$), including $\theta_{crecca}$, $\theta_{carolinensis}$, $\theta_{nimia}$, $\theta_{A1}$ and $\theta_{A2}$ (a measure of neutral genetic variation for *crecca*, *carolinensis*, *nimia*, the common *crecca-nimia* ancestor and the *crecca/nimia-carolinensis* ancestor, respectively, where $\theta = 4N_e u$ and $N_e$ is the effective population size), $t_{crecca\text{-}nimia}$ and $t_{crecca/nimia\text{-}carolinensis}$ (where $t = Tu$, and $T$ is the number of years since divergence), and $M_{crecca \leftarrow carolinensis}$, $M_{crecca \leftarrow nimia}$, $M_{carolinensis \leftarrow crecca}$, $M_{carolinensis \leftarrow nimia}$, $M_{nimia \leftarrow crecca}$, $M_{nimia \leftarrow carolinensis}$, $M_{carolinensis \leftarrow ancestor1}$, and $M_{ancestor1 \leftarrow carolinensis}$ (where $M = m/u$, and $m/u$ is the ratio by which new alleles immigrate into population 1 from population 2 relative to the mutation rate). We treated mtDNA and nuDNA in separate analyses because we suspected sex-biased dispersal (Peters *et al.* 2012a) and defined inheritance scalars as 0.25 for mtDNA, 0.75 for CHD1Z, and 1.0 for autosomal DNA to reflect modes of inheritance.

To meet IMa2's assumption of no intralocus recombination, we tested for recombination within nuDNA using the four-gamete test (Hudson & Kaplan 1985). For loci that showed evidence of recombination, we used the program IMgc to subsample our data (Woerner *et al.* 2007); IMgc optimizes the removal of copies of loci and/or base pairs to retain the maximum number of polymorphic sites that are consistent with no recombination. We ran the program iteratively, changing the weighted preference for removing chromosomes so that a maximum of 5% of copies were removed for any one population at each locus. By default, IMgc recodes nucleotide positions segregating for three or more states so the sequences are consistent with an infinite-sites model of evolution; however, we only used IMgc as a guide for truncating sequences and retained those sites for IM analyses.

We initially ran IMa2 using results from Peters *et al.* (2012a) as guide for setting priors. On the basis of these preliminary runs, we set uniform priors (assumed to be uninformative) containing the entirety of the posterior distributions for most parameters. However, for both mtDNA and nuDNA, the posterior distributions for $\theta_{crecca}$ rose sharply and plateaued, even at large values, and the posteriors for three of the eight migration parameters were flat over a large range of values (note that high upper priors on migration caused the program to crash). In addition, $\theta_{A2}$ was flat in the mtDNA analysis. Because we used the same prior range for all θs and for all migration parameters, we used the parameters that had finite posterior distributions to set the priors. We also recorded $2Nm$ (the number of effective migrants) during the Markov chains, because these values are directly comparable between nuDNA and mtDNA; all posteriors for $2Nm$ had finite posterior distributions. For nuclear DNA, we ran a cold chain and 39 heated chains with a geometric heating scheme for a burn-in of 500 000 steps followed by 10 000 000 steps with a sampling interval of 100 steps. For mtDNA, we ran a cold chain and 19 heated chains for the same number of steps. Each run was replicated with a different random number seed to check for consistent results.

### *Coalescent simulations of genetic differentiation*

To assess assumptions of neutrality (Kuhner 2009) and genomic patterns of differentiation, we conducted

coalescent simulations to examine the expected levels of neutral genetic differentiation among the three teal taxa given the population history inferred from IMa2. For each nuclear locus and each pair of taxa, we simulated 1000 replicates of genetic diversity in the program ms (Hudson 2002). Following Peters *et al.* (2012b), we incorporated three sources of uncertainty into these simulations. (i) We randomly sampled 1000 values from the posterior distributions of each parameter estimated in IMa2, and these values were randomly assigned to one of the 1000 replicates. (ii) We incorporated relative substitution rates ($\mu_R$) among the eight loci; estimates of $\mu_R$ were obtained from Peters *et al.* (2012b), which were based on evolutionary rates inferred from eight lineages deep in the anseriform tree. To account for uncertainty in these rates, we randomly sampled 1000 values from the posterior distributions of $\mu_R$ for each locus. (iii) We sampled 1000 values from the posterior distributions of recombination rates for each locus that were estimated in the program LAMARC (Kuhner 2006); these values were obtained from Peters *et al.* (2012a) and were estimated using DNA sequences from *crecca*. We scaled all parameters in the simulations to $\theta_{carolinensis}$, which was calculated as $\theta$ (from IMa2) $\times$ $\mu_R$ $\times$ $l_R$, where $l_R$ is the locus-specific length relative to the geometric mean of all fragment lengths used in the IMa2 analysis (see Hudson 2002 & Peters *et al.* 2012b for more details on converting parameters). $\theta_{carolinensis}$ for CHD1Z was multiplied by a factor of 0.75 to account for the ¾ effective size of Z-linked loci. In summary, we simulated genetic diversity for eight loci over 1000 population histories while incorporating locus-specific substitution rates, locus-specific recombination rates, and uncertainty in those rates for a total of 8000 simulated data sets.

For each simulated data set, we calculated pairwise $\Phi_{ST}$ among the three teal taxa in the program MS.OUTPUT v.4.1 (Peters *et al.* 2012b). For each locus, we generated posterior predictive distributions (Meng 1994) of $\Phi_{ST}$ for each pairwise comparison and performed goodness-of-fit tests as described by Becquet & Przeworski (2007). We considered a locus as deviating significantly from the neutral population history if the empirical value of $\Phi_{ST}$ fell within the 2.5% tail of the posterior predictive distribution for that locus. We also calculated mean expected $\Phi_{ST}$ among all eight loci and compared that to the empirical mean value to test the overall fit of the data to the model.

Lastly, we estimated the demographic history of all *nimia* individuals from Attu, Adak, and Ugamak islands ($n$ = 44) using mtDNA data and Bayesian skyline plots (Drummond *et al.* 2005) as implemented in the software BEAST (Drummond *et al.* 2012). Repeated runs of 50 million generations were performed, using a mutation rate of $4.8 \times 10^{-8}$ (Peters *et al.* 2005), the HKY model of nucleotide substitution, empirical base frequencies, no site heterogeneity, a strict clock, a uniform prior and an upper population limit of 200 000 and a burnin of 5 million.

## Results

### *Genetic differentiation*

Comparison of our population samples from three Aleutian islands (Attu, Shemya and Adak) with those of the OW and NW populations revealed a complex relationship of differentiation, with substantial variation among loci and among locations (Table 1). The Shemya population was exceptional in that it did not differ significantly from OW *crecca* in any of the nine loci; all other pairwise population comparisons showed from one to as many as nine significant differences (Table 1). Inasmuch as we expected to obtain OW *crecca* migrants in our Shemya sample, we set that population aside for the moment and consider the other populations first.

**Table 1** Genetic differentiation ($\Phi_{ST}$) at eight nuclear loci and the mtDNA control region among five populations of green-winged teal

| | CRYAB | FGB | ENO1 | LDHB | GRIN1 | ODC1 | PCK1 | CHD1Z | nuDNA (avg.) | SD | mtDNA |
|---|---|---|---|---|---|---|---|---|---|---|---|
| NW-Adak | **0.322** | **0.033** | **0.059** | **0.319** | **0.176** | **0.102** | **0.094** | **0.094** | 0.150 | 0.113 | **0.949** |
| NW-Attu | **0.412** | **0.037** | **0.040** | **0.321** | **0.113** | **0.195** | **0.055** | **0.066** | 0.155 | 0.142 | **0.954** |
| NW-Shemya | **0.270** | 0.014 | 0.011 | **0.287** | 0.015 | 0.005 | 0.028 | **0.052** | 0.082 | 0.123 | **0.875** |
| NW-OW | **0.369** | 0.004 | 0.001 | **0.350** | 0.007 | 0.005 | **0.064** | **0.035** | 0.103 | 0.160 | **0.885** |
| Adak-Attu | 0.019 | 0.004 | **0.038** | 0.014 | 0.011 | 0.001 | 0.045 | 0.035 | 0.016 | 0.021 | **0.201** |
| Adak-Shemya | -0.019 | 0.011 | **0.039** | 0.019 | **0.101** | 0.006 | **0.115** | **0.063** | 0.037 | 0.052 | **0.162** |
| Adak-OW | -0.006 | 0.013 | **0.078** | 0.025 | **0.119** | **0.094** | **0.129** | **0.077** | 0.066 | 0.050 | **0.182** |
| Attu-Shemya | 0.011 | 0.012 | 0.024 | 0.008 | **0.071** | **0.079** | 0.005 | 0.042 | 0.029 | 0.032 | **0.148** |
| Attu-OW | 0.009 | **0.028** | **0.036** | 0.020 | **0.068** | **0.194** | 0.018 | 0.026 | 0.050 | 0.061 | **0.158** |
| Shemya-OW | -0.001 | 0.025 | 0.003 | 0.017 | 0.003 | 0.009 | 0.007 | 0.005 | 0.002 | 0.012 | -0.037 |

Bold indicates $P < 0.05$; underline indicates significant after Bonferroni correction.

The mtDNA control region showed a relationship corroborating that of male breeding plumage, indicating that Aleutian birds are more closely related to OW *crecca* ($\Phi_{ST}$ ~ 0.16–0.18) than to NW *carolinensis* ($\Phi_{ST}$ ~ 0.88–0.95; Table 1, Figs 2 and S1, Supporting Information). This closer relationship between *crecca* and *nimia* was also corroborated with nuDNA, which showed a higher average $\Phi_{ST}$ between Aleutian and NW birds (~0.10–0.15) than between Aleutian and OW birds (~0.05–0.07; Table 1).

The Aleutian populations from Attu and Adak were more strongly differentiated from NW teal than OW teal were from NW teal, showing significant differences at more loci and a higher level of differentiation in both the nuclear and mitochondrial genomes (Table 1). The Attu and Adak populations were significantly differentiated from the OW population in four to six of the nine loci examined (Table 1), and when considered together as a single population, the Aleutian *nimia* differed from OW *crecca* at three loci (GRIN1, ODC1 and mtDNA after Bonferroni correction) and showed an average $\Phi_{ST} = 0.07$ (Tables 1 and S2, Supporting Information). These two island populations of *nimia* differed significantly from each other in mtDNA and ENO1. Genetic diversity, measured as allelic richness, was normally distributed after a log transformation (Shapiro–Wilk $W$ Test, $W = 0.9723$, $P = 0.54$). Allelic richness differed significantly among populations (repeated measures ANOVA, $F_{4,28} = 13.3$, $P < 0.0001$). OW, NW and Shemya each had significantly higher allelic richness than Attu and Adak (Fig. 3; Tukey post-test; $q_{28,5} > 5.0$, $P < 0.05$). There were no significant differences among the remaining populations. Including all birds in STRUCTURE for analysis of nuDNA genotypes, with no a priori information on individual origin, resulted in the most likely number of populations ($K$) being two (Supporting Information, Table S1, Fig. S2). All birds from Attu, Adak, and Ugamak were assigned to population 1 (mean $Q_1 = 1.0 \pm 0.01$ SD), whereas most of the remainder (42/61 or 69%) were assigned to population 2. Two NW birds (8%) were assigned to population 1 ($Q_1$ of 0.59 and 0.67), as were five from Shemya (45%, mean $Q_1 = 1.0 \pm 0.004$ SD) and 12 from the OW (48%, mean $Q_1 = 0.90 \pm 0.11$ SD; Table S1, Supporting Information).

Because weaker population structure in a data set (e.g. Table 1) may not be detected by STRUCTURE (Hubisz *et al.* 2009), and because we had determined previously that OW and NW populations are genomically different at these nuclear loci (Peters *et al.* 2012a), we did a second analysis removing NW birds, considering that the parameter space dominated by the large differences between Aleutian-NW divergence might preclude detection of a third population (i.e. the model's pooling of 69% of the birds from two different continents into one population is *contra* the findings of Peters *et al.* 2012a). In this analysis of OW and Aleutian birds, with no a priori information on individual origin, the most probable number of populations ($K$) was again 2. Among OW birds, 16/25 (64%) were assigned to population 1 (mean $Q_1 = 0.90 \pm 0.13$ SD; Table S1, Supporting Information). Six birds from Shemya (55%) were

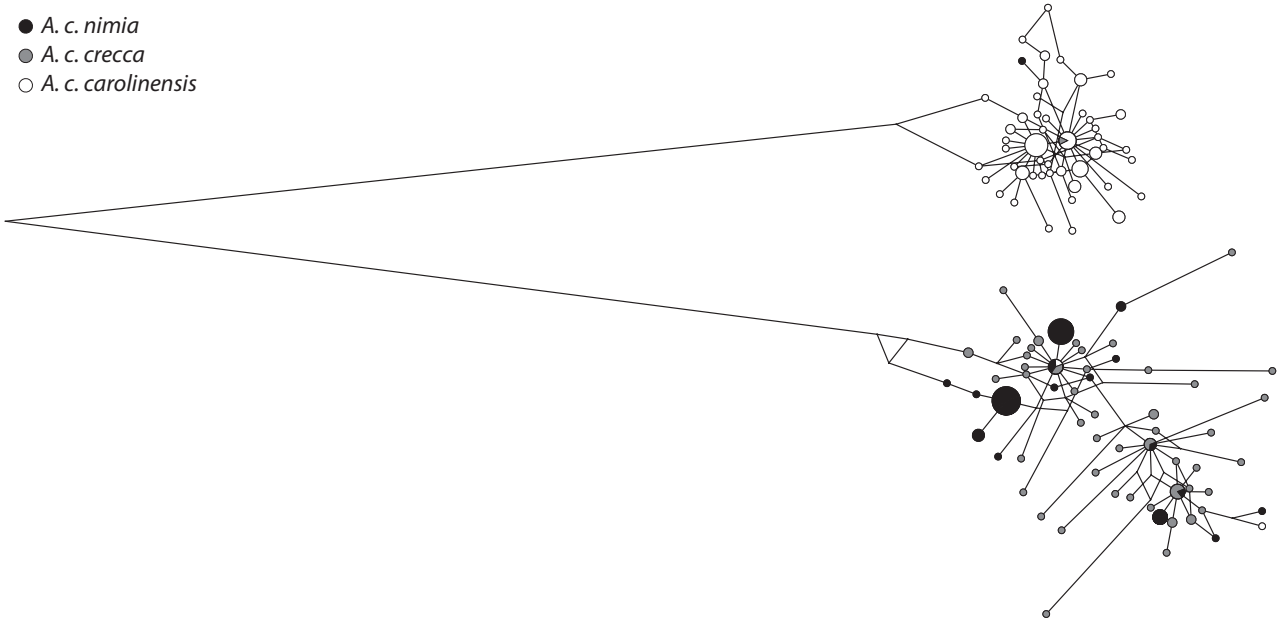


**Fig. 2** Haplotype network of the mtDNA control region.

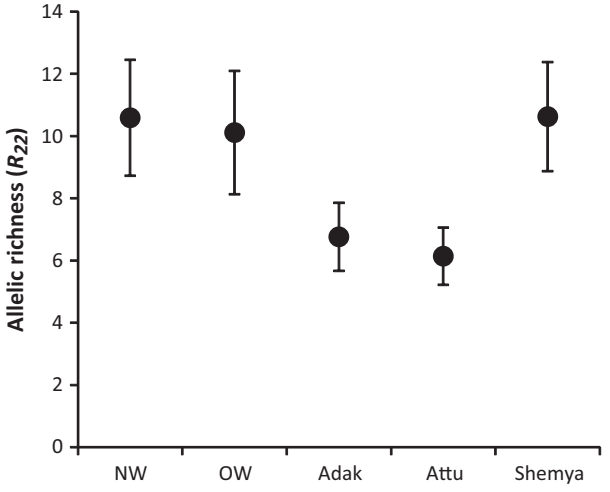


**Fig. 3** Allelic richness of the five sampled populations (and standard error bars), from Eurasia (Old World) through the Aleutian Islands, to North America (New World).

also assigned to population 1 (mean $Q_1 = 0.96 \pm 0.04$ SD), whereas all birds from Attu, Adak and Ugamak were assigned to population 2 (mean $Q_2 = 0.98 \pm 0.06$ SD; Fig. 4; Table S1, Supporting Information; population numbers in these analyses are based solely on STRUCTURE output and by chance were reversed in this analysis from the one above).

### *Gene flow and population demography*

Excluding the Shemya population and pooling the remaining Aleutian birds, our three-population analyses of nuDNA using IMa2 showed significant, nonzero gene flow in three of the six possible directions tested: from *crecca* and *carolinensis* into *nimia* (Fig. 5) and from *crecca* into *carolinensis* as documented earlier by Peters *et al.* (2012a). The estimated number of effective immigrants coming into *nimia* from *crecca* and *carolinensis* per generation ($2N_em$) peaked at 1.8 (95% HPD of 0.28–3.46) and 1.3 (95% HPD of 0.52–2.51), respectively (Fig. 5). In contrast, the estimate for the number of effective immigrants from *crecca* into *carolinensis* peaked at 28.0 individuals per generation (95% HPD of 4.40–72.06), and gene flow estimates from *nimia* into *crecca*, from *nimia* into *carolinensis*, and from *carolinensis* into *crecca* were not distinguishable from zero (Fig. S3, Supporting Information).

Coalescent analysis of mtDNA in IMa2 (as with the nuDNA analysis, excluding the Shemya population) yielded different results, with significant nonzero gene flow occurring into *nimia* from *crecca* only (peak $2N_em = 1.5$, 95% HPD 0.2–4.2), and out of *nimia* into both *carolinensis* (peak $2N_em = 1.6$, 95% HPD 0.3–8.0) and *crecca* (peak $2N_em = 12.9$, 95% HPD 1.9–28.1; Fig. S3, Supporting Information). Other estimates of gene flow were not significantly different from zero. In contrasting nuDNA and mtDNA gene flow results (Fig. S3, Supporting Information), the low levels of nuclear gene flow into *nimia* from both continental populations are corroborated by mtDNA (low from *crecca*, zero from *carolinensis*). However, while there was no detectable nuclear gene flow out of *nimia*, mtDNA suggested possible low levels of gene flow coming out of

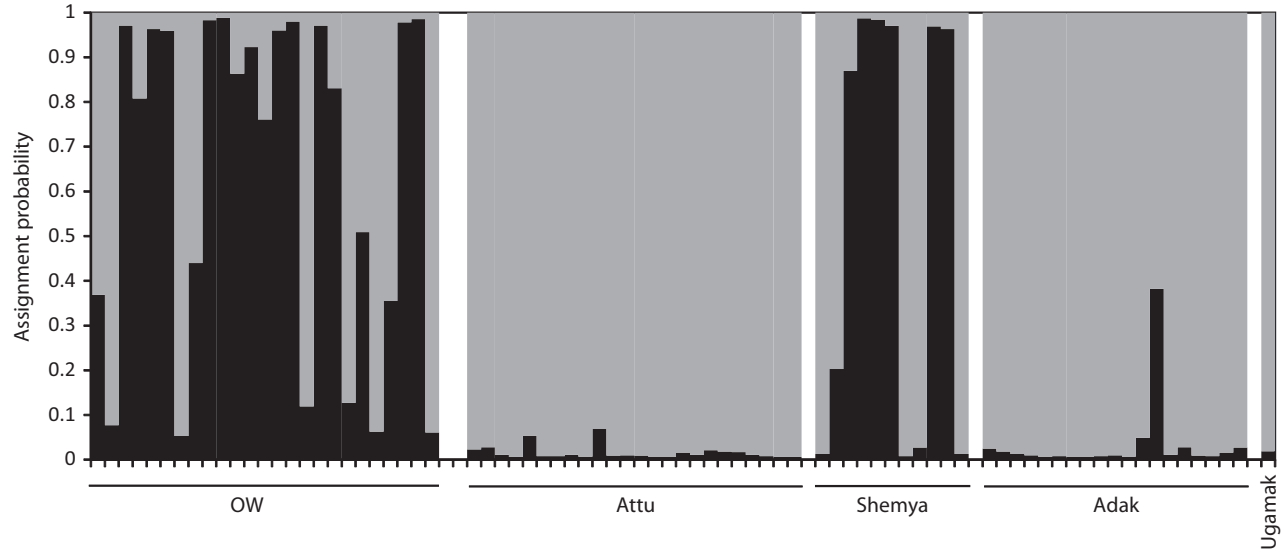


**Fig. 4** Assignment probabilities of 55 individual teal to one of two populations based on eight nuclear introns; population 1 = dark slices; population 2 = light slices (using STRUCTURE, ver. 2.2.3). Individuals are given from westernmost Eurasia (Italy) to Ugamak Island in the Aleutian Islands; no New World *carolinensis* are included in this Figure. White columns separate clusters from the following locations: Old World, Attu, Shemya, Adak, Ugamak ($N = 1$).

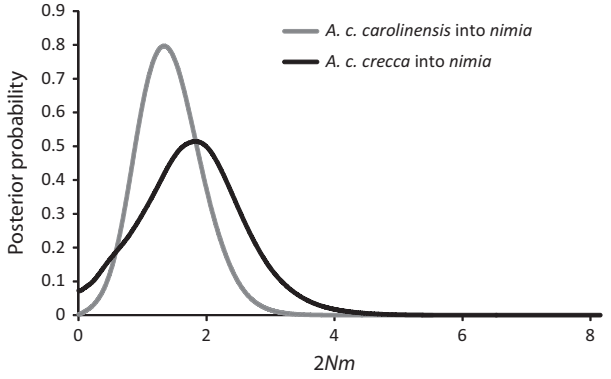


**Fig. 5** The number of effective immigrants ($2N_em = \theta M/2$) into Aleutian green-winged teal (*nimia*) from the New World (*carolinensis*) and Old World (*crecca*) continental populations, based on nuDNA.

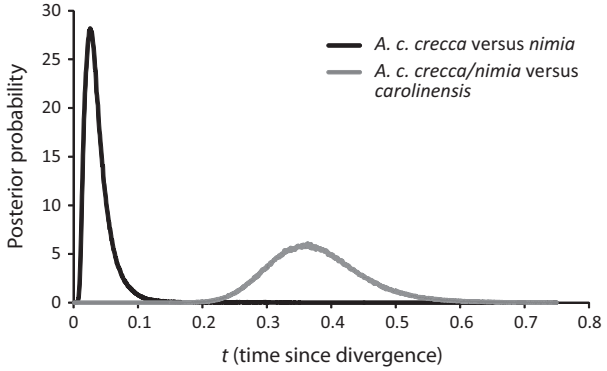


**Fig. 6** Estimates of divergence ($t$) for the timing of splits between *Anas. c. crecca* and *nimia* (~83 000 years) and between the two major clades *crecca+nimia* and *carolinensis* (~1.1 Myr).

the sedentary population (Fig. S3, Supporting Information). Unlike the results from Peters *et al.* (2012a), these analyses did not show mtDNA gene flow from *crecca* into *carolinensis*, despite inclusion of the same data, emphasizing the importance of caution when interpreting the results of the complex IMa2 models when using the limited data in a single locus.

Estimates of θ ($4N_eu$) from nuDNA (again, excluding the Shemya individuals) indicated that the effective population size of *nimia* is significantly smaller than both of the continental populations, as expected given differences in breeding range size (Fig. S4, Supporting Information). Assuming a mean substitution rate of $1.2 \times 10^{-9}$ substitutions per site per year for nuDNA and a generation time of approximately 3 years (Peters *et al.* 2008), we estimated that the effective population size ($N_e$) of *nimia* is ~80 000 individuals. For comparison, the effective size ($N_e$) of *carolinensis* was ~1.6 million and that of *crecca* was ~5.9 million [note that the curve peaked within the upper bin (not shown), and setting a wider prior was unrealistic for this population]. Using the same substitution rate, nuDNA indicated that the time of divergence between *nimia* and *crecca* is substantially more recent than that between the continental populations (~83 000 years vs. ~1.1 Myr; Fig. 6). For mtDNA, using a substitution rate of $\sim 4.8 \times 10^{-8}$ substitutions per site per year (Peters *et al.* 2005), the *crecca-nimia* divergence was estimated to be ~23 000 years bp, and the *crecca/nimia-carolinensis* divergence was ~600 000 years bp (however, the curve was flat and thus does not provide a very reliable estimate for the deeper divergence).

Reconstruction of the mtDNA demographic history of Aleutian birds (minus individuals from Shemya) showed a relatively stable population from ~28 000 years bp to near the present, when there may have been a small decline (Fig. S5, Supporting Information).

### *Heterogeneity among loci*

Simulating genetic diversity under neutrality revealed that empirical values of $\Phi_{ST}$ for the loci CRYAB and LDHB were significantly higher than expected based on the inferred model ($P < 0.001$; Fig. 7), but only in tests involving *carolinensis* (see also Peters *et al.* 2012a). These loci also contributed to higher than expected mean $\Phi_{ST}$ for the empirical data ($P = 0.005$ for *carolinensis* vs. *crecca*; $P < 0.001$ for *carolinensis* vs. *nimia*). For comparisons between *crecca* and *nimia*, values of $\Phi_{ST}$ fit model expectations fairly well; there were no significant deviations (Fig. 7).

### *Shemya Island*

The Shemya population did not differ significantly at any locus from OW *crecca* (Table 1), and STRUCTURE showed mixed and strongly divergent (bimodal) assignment probabilities for the birds from Shemya, with six assigning with high probability to the OW (mean $Q = 96 \pm 4\%$ SD) and five with high probability to the Aleutian population (mean $Q = 95 \pm 8\%$ SD; Table S1, Fig. S1, Supporting Information). This suggests that migrant OW *crecca* were obtained there. Considering the birds from Shemya as representing two groups, we were unable to find diagnostic phenotypic differences in wing length, condition (mass/wing length), or plumage to separate them. The volume of the left testis of the males (calculated after Hoyt 1979) tended to be larger in Aleutian individuals ($\sim = 3774$ mm$^3$ $\pm$ 1046 SD; $n = 3$) than in those assigned to the OW population ($\sim = 2382$ mm$^3$ $\pm$ 1611 SD; $n = 4$), but this was not significant. In addition, the fat levels (determined during specimen preparation) of birds assigned to the OW population tended to be higher than those assigned to the Aleutian population ($\sim = 3.5$ vs. 2.9), but this, too, was not significant. Both of these trends are in the

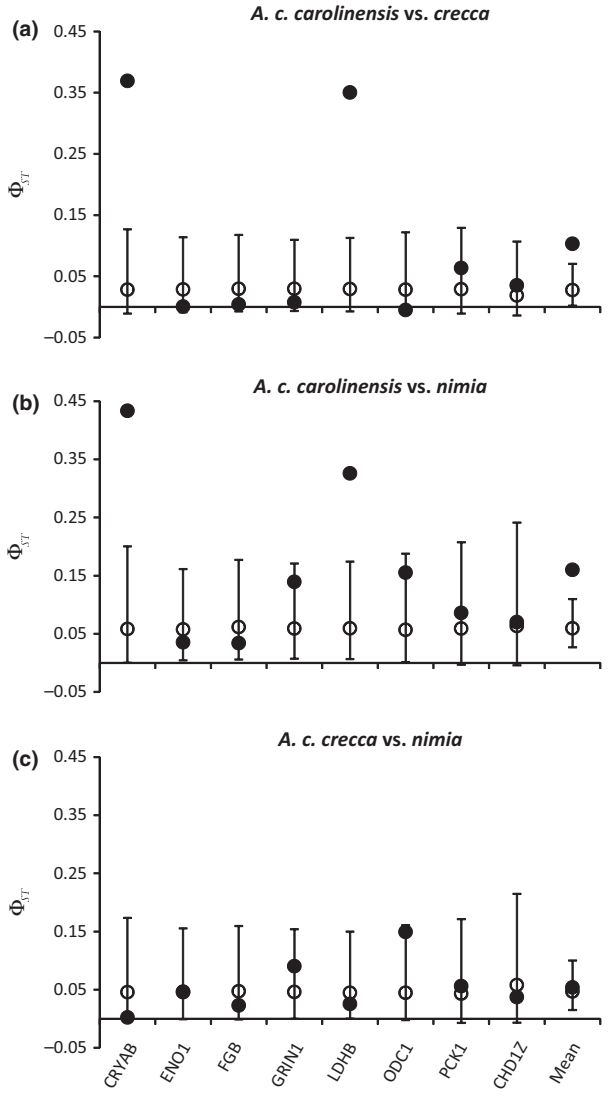


**Fig. 7** Simulating genetic diversity under neutrality (open circles and 95% confidence intervals) relative to empirical values of $\Phi_{ST}$ (black circles) for the nuDNA loci in this study, between (a) *carolinensis* and *crecca*, (b) *carolinensis* and *nimia*, and (c) *crecca* and *nimia*.

direction predicted by migrant vs. sedentary individuals. Also, behavioural observations prior to collection indicated that four of the six individuals assigned to the OW population were mated pairs (31 May 2000 and 8 June 2001, Table S1, Supporting Information).

## Discussion

Three key predictions of the heteropatric speciation model were borne out when examining the possible genetic divergence of the Aleutian population of green-winged teal: (i) significant divergence has occurred (Tables 1, S2; Figs 3 and S3, Supporting Information), with (ii) rather low gene flow between a sedentary population (*nimia*) and the migratory population (*crecca*) that passes through it twice annually (Fig. 5), and (iii) the sedentary population is of substantial size (Fig. S4, Supporting Information). Thus, divergence with gene flow appears to be occurring, commensurate with the heteropatric model of differentiation. This model predicts that this divergence should occur due to population-specific adaptations to cyclically available resources that differ in space and time. This causes divergent selection between different populations focusing (in an evolutionary sense) on local optima in different places, despite the fact that these populations occur in sympatry during portions of the annual cycle. In this case, we also see that a parapatric relationship between *nimia* and the more distantly related NW *carolinensis* involves a similarly low level of gene flow into *nimia*. The case of parapatric speciation between OW *crecca* and NW *carolinensis* was examined in depth by Peters *et al*. (2012a).

What specific factors might be preventing more interbreeding between *nimia* and *crecca* despite their occurrence in sympatry? The timing and location of pair bonding is crucial; migratory waterfowl form pairs on their wintering grounds, and the male migrates with the female to her breeding grounds (Palmer 1976). Thus, in this case, pairing occurs in the migratory population on different wintering grounds than the sedentary population, and, with such low evidence for gene flow, our data suggest a rather high degree of nonoverlap in wintering populations (i.e. apparently substantial populational site fidelity to wintering regions). Given the flexibility that individual waterfowl often exhibit in wintering areas, with ranges shifting to take advantage of open waters in winter (which occurs in both of the continental subspecies *crecca* and *carolinensis*; Cramp *et al*. 1977; Johnson 1995a), the inferred degree of separation in this case is somewhat surprising. Indeed, the levels of gene flow that we found were to us surprisingly low for three reasons: (i) semiannual sympatry; (ii) the clear suitability of island habitats for wintering birds and evidence in waterfowl for wintering site flexibility; and, finally, c) because of the notorious behaviour of waterfowl to crossbreed (McCarthy 2006). Further, unlike our previous study of *crecca* and *carolinensis* (Peters *et al*. 2012a), we did not find a striking difference in gene flow between mtDNA and nuDNA (Fig. S3, Supporting Information). In our previous study, male-biased gene flow dominated, being an order of magnitude higher from *crecca* into *carolinensis* than gene flow estimates from mtDNA. Here, we found a lack of male-biased gene flow into *nimia*. Instead, immigration into the Aleutian Islands seems to be low for both sexes, which is consistent with the heteropatric model.

The degree of genomic identity and the low levels of gene flow present suggest that *nimia* is well along in the speciation process, despite being morphologically nearly indistinguishable from *crecca* (Gibson & Kessel 1997). While this divergence is less than that found between the subspecies *crecca* and *carolinensis*, the levels of gene flow between the latter seem to have them evolutionarily tied together in a classic dumbbell model of parapatry (Peters *et al.* 2012a), whereas while it represents a much younger split, the divergence between *crecca* and *nimia* exhibits a much cleaner break with respect to gene flow.

Our results emphasize another aspect of the heteropatric speciation model, in that divergence is probably occurring along ecological dimensions rather than through sexual selection, because the most pronounced phenotypic differences appear in the presence and absence of migration and in where and when the populations breed and winter rather than in morphology. Another noteworthy aspect is that waterfowl differ from most birds in having male-biased dispersal and, in migratory lineages, forming pair bonds on their nonbreeding grounds. Our results support predictions of the heteropatric model that divergent selection operates in such systems on individuals and populations adapting to different local optima, mostly in resource peaks that enable successful reproduction and that occur in different places and at different times. Different wintering areas (wintering allopatry, or allohiemy; Lack 1944, 1968; Salomonsen 1955) were also considered to be potential contributors to divergence (Winker 2010), but the data presented here suggest that this can be critical, at least in species that pair before they arrive on their breeding grounds. What might promote wintering allopatry? Intraspecific competition (competitive exclusion) is the most likely explanation (Lack 1968; Gauthreaux 1982). In this case, Aleutian populations are at their annual peak in autumn when resources are diminishing and more northern birds are migrating south to suitable wintering areas. Bird populations generally experience food-based limitation during the nonbreeding season (Gill 2007), and, therefore, there is probably strong competition for resources when *crecca* and *nimia* are sympatric. We thus infer that incoming autumn migrants from other breeding populations do not fare well if they try to winter in the Aleutians.

The divergence between *nimia* and *crecca* does not seem likely to be a case of peripatric speciation, in which the Aleutian Islands population arose from Eurasian colonists and was isolated in allopatry for an extended period before Eurasian birds began a modern period of semiannual sympatry. Evidence suggests that it is unlikely that *nimia* has ever undergone a substantial period of isolation. Northeastern Asia did not undergo the extensive glaciations that occurred widely across North America (Zamoruyev 2004; Melles *et al.* 2012), and during the last glacial maximum (LGM), the Beringian breeding grounds hosted a diverse mammalian and avian fauna (Hopkins 1967; Fitzgerald 1991). Thus, environmental conditions were probably suitable for continual interactions among all three of these teal taxa during the LGM and afterwards.

Might the lack of significant differentiation in the Shemya population be due to hybridization rather than to the presence of OW individuals in the sample? Several lines of evidence suggest to us that migratory OW birds, and not hybrids, were responsible for this result. First, the assignment probabilities were high and strongly bi-modal. Hybrids would be expected to have intermediate assignment probability values. Second, these birds exhibited phenotypic trends concordant with being migrants (smaller testes in males and higher levels of fat); though not significant (samples sizes were small), both trends were in the predicted direction. Third, four of the six Shemya birds assigned to OW genotypes were collected as two mated pairs, the only such taken in this study. Thus, behaviourally, two-thirds of the Shemya birds with OW genotypes stuck out as being different. We thus consider that these were indeed migrants from the OW population, as we had expected to obtain in our sample given that migrants can be an order of magnitude more common at Shemya than the estimated number of breeding pairs (Schwitters 2008). Although nearby Attu Island birds were also sampled heavily during migration (though more in autumn than in spring), no birds in that population were assigned to the OW population (Table S1, Supporting Information). This could be due to differences between seasons, between years, or to the much larger breeding population on Attu.

Some OW individuals were assigned to *nimia*, despite being sampled more than 8000 km away in some cases (e.g. Italy). Rather than being migrants from the Aleutians, we suspect these assignments resulted from the stochastic sampling of alleles in the large populations. More specifically, individual Eurasian *crecca* that happened to carry alleles that were rare on the mainland (see Fig. S6), but common on the Aleutians, might have been assigned to *nimia*. Indeed, these individuals generally received lower assignment scores than individuals that were sampled in the Aleutians (Fig. 4). Individual *crecca* that carried alleles that are common in the *nimia* population could thus be assigned to either population without causing large deviations from Hardy–Weinberg equilibrium, especially if they carried common alleles at two or more loci. Genetic drift has thus decreased genetic diversity of *nimia*, whereas time since divergence has not been sufficiently long to allow allele sorting in

the large continental population because the effects of genetic drift are weaker (see also McCracken *et al.* 2013). Finally, our coalescent analyses indicated that the genetic patterns were consistent with no detectable gene flow into the continental populations from the Aleutians, further supporting our interpretation that these continental '*nimia*' individuals were simply misassigned. Sampling more loci with allele frequency differences between populations will be necessary to better understand the influence of stochastic sampling vs. gene flow.

Our simulations demonstrated that the nuclear loci we used met the assumptions of neutrality in our coalescent analyses between *crecca* and *nimia* (Fig. 7). This assumption should be tested more often when performing coalescent analyses to estimate population parameters, as we have done here, because departures may prove to be common (Sella *et al.* 2009; Peters *et al.* 2012b). There were, however, notable differences among loci between the two divergence events of *crecca-nimia* and *crecca-carolinensis*, with two loci in the latter exhibiting significant departures from neutrality (see also Peters *et al.* 2012a). Different genomic patterns of divergence are expected to occur between allopatric speciation vs. speciation with gene flow (Osada & Wu 2005; Via 2009). Here, parapatric and heteropatric speciation are both types of speciation with gene flow, yet the divergence processes occurring are clearly affecting these homologous loci differently. We cannot determine whether these differences might be due to different divergent selection regimes or to stochastic processes operating between population pairs with quite different effective sizes over very different timescales. However, the *nimia* population is likely to be orders-of-magnitude smaller than the two continental populations, and therefore, selection is probably less effective and drift is more effective in driving differences across the genome in the sedentary island population.

Genetic drift has thus probably had a prominent role in our results. While there is no evidence for bottlenecks, and modern census sizes assure *nimia* populations in the thousands (Gibson & Byrd 2007), these populations are nonetheless a small fraction (Fig. S4, Supporting Information) of the global populations of *crecca* and *carolinensis* (~3.7–4.7 and 2.9 million individuals, respectively; Delany & Scott 2002). Genetic drift in *nimia* is likely to occur more rapidly than in either of the continental forms. The signals of drift rather than selection in causing the genomic identity of the sedentary population (*nimia*) to be so strong (Figs 3 and S2, Supporting Information) appear in (i) assignment of some OW birds (as far away as Italy) to the Aleutians population (Fig. 4; see above), and (ii) the absence of loci exhibiting more structure than expected by chance (Fig. 7C). Drift is also evident in higher levels of mtDNA than nuDNA divergence between nearly all pairwise population comparisons, as expected due the smaller effective population size of mtDNA (Table 1). In contrast, selection has probably had a prominent influence in the divergence between *crecca* and *carolinensis* at some loci, resulting in heterogeneous patterns of divergence in these large population sizes (Peters *et al.* 2012a). Drift is an important factor affecting biodiversity (Lynch 2007), but we doubt that it is driving the divergence between these two lineages. While strong selection can override drift in small populations, we consider that our small sample of the genome in the case of *crecca* vs. *nimia* has a predominant signal of drift and that a broader sampling of the genome would be required to detect direct evidence of the divergent selection that we infer is occurring.

Among a number of avian taxa, the mapping of migration onto phylogenetic trees shows that some differentiation or speciation occurs when a subgroup of the ancestral lineage drops out of migration to establish resident (or less migratory) breeding populations, usually in tropical or subtropical regions (Klein & Brown 1994; Johnson 1995b; Burns 1998; Cicero & Johnson 1998; Kondo *et al.* 2008). In the theory of heteropatric speciation, the sequence of origin of the sedentary population being passed through or leapfrogged by the migratory form is not particularly important; the ancestral population of the diverging lineages could be migratory or sedentary. In this case, the *Anas* phylogeny does not clearly indicate the ancestral migratory status of the [(*crecca*, *nimia*) *carolinensis*] lineage (Johnson & Sorenson 1999).

Mallet (2008) pointed out that in speciation research we are mainly concerned with how diverging lineages evolve to occur in sympatry. Heteropatric speciation considers how divergence develops when lineages may never escape from some degree of sympatry. Green-winged teal and other cases of leapfrog migration may provide key insights into speciation with gene flow.

## Acknowledgements

This study was supported by the University of Alaska Museum, Alaska EPSCoR (NSF EPS-0346770), the National Science Foundation (DEB-9981915, DEB-0444748, DEB-0746365, DEB-0926162), and the United States Department of Agriculture (SCA 58-6612-2-217). We thank K Millam and KA Bolender for assistance with data collection, S Houston and UAF Life Science Informatics for computer resources, and C Bickford, K Campbell, K Everson, J Harley, M Smith, L Stephens, and two anonymous reviewers for comments on earlier drafts.

## References

Abbott MB, Edwards ME, Finney BP (2010) A 40,000-yr record of environmental change from Burial Lake in northwest Alaska. *Quaternary Research*, **74**, 156–165.

Baker RR (1978) *The Evolutionary Ecology of Animal Migration*. Holmes & Meier Publishers, Inc., New York, New York.

Becquet C, Przeworski M (2007) A new approach to estimate parameters of speciation models with application to apes. *Genome Research*, **17**, 1505–1519.

Belliure J, Sorci G, Møller AP, Clobert J (2000) Dispersal distances predict subspecies richness in birds. *Journal of Evolutionary Biology*, **13**, 480–487.

Boland JM (1990) Leapfrog migration in North American shorebirds: intra- and interspecific examples. *Condor*, **92**, 284–290.

Bottema CDK, Sarkar G, Cassady JD *et al.* (1993) Polymerase chain reaction amplification of specific alleles: a general method of detection of mutations, polymorphisms, and haplotypes. *Methods in Enzymology*, **218**, 388–402.

Brazil M (2009) *Birds of East Asia*. Princeton University Press, Princeton, New Jersey.

Burns K (1998) Molecular phylogenetics of the genus *Piranga*: implications for biogeography and the evolution of morphology and behavior. *Auk*, **115**, 621–634.

Cicero C, Johnson NK (1998) Molecular phylogeny and ecological diversification in a clade of New World songbirds (genus *Vireo*). *Molecular Ecology*, **7**, 1359–1370.

Cramp S, Simmons KEL, Ferguson IJ *et al.* (eds) (1977) *Handbook of the Birds of Europe the Middle East and North Africa. Volume I. Ostriches to Ducks*. Oxford University Press, Oxford.

Delany S, Scott D (2002) *Waterbird Population Estimates*, 3rd edn. Wetlands International, Wageningen, the Netherlands.

Drummond AJ, Rambaut A, Shapiro B, Pybus OG (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology & Evolution*, **22**, 1185–1192.

Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, **29**, 1969–1973.

Earl DA, von Holdt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, **4**, 359–361.

Elias SA (2001) Beringian paleoecology: results from the 1997 workshop. *Quaternary Science Reviews*, **20**, 7–13.

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

Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47–50.

Fitzgerald GR (1991) Pleistocene ducks of the Old Crow Basin, Yukon Territory, Canada. *Canadian Journal of Earth Sciences*, **28**, 1561–1571.

Flot J (2010) Seqphase: a web tool for interconverting phase input / output files and fasta sequence alignments. *Molecular Ecology Resources*, **10**, 162–166.

Friedmann H (1948) The green-winged teal of the Aleutian Islands. *Proceedings of the Biological Society of Washington*, **61**, 157–158.

Gauthreaux Jr SA (1982) The ecology and evolution of avian migration systems. *Avian Biology*, **6**, 93–168.

Gibson DD, Byrd GV (2007) *The Birds of the Aleutian Islands, Alaska*. Nuttall Ornithological Club, Cambridge, Massachusetts.

Gibson DD, Kessel B (1997) Inventory of the species and subspecies of Alaska birds. *Western Birds*, **28**, 45–95.

Gill F (2007) *Ornithology*, 3rd edn. W. H. Freeman and Company, New York, New York.

Harrington CR (ed.) (2003) *Annotated Bibliography of Quaternary Vertebrates of Northern North America*. University of Toronto Press, Toronto, Ontario.

Hey J (2010) Isolation with migration models for more than two populations. *Molecular Biology and Evolution*, **27**, 905–920.

Hopkins DM (1967) *The Bering Land Bridge*. Stanford University Press, Stanford, California.

Hoyt DF (1979) Practical methods of estimating volume and fresh weight of bird eggs. *Auk*, **96**, 73–77.

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, **9**, 1322–1332.

Hudson RR (2002) Generating samples under a Wright-Fisher neutral model of genetic variation. *Bioinformatics*, **18**, 337–338.

Hudson RR, Kaplan NL (1985) Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics*, **111**, 147–164.

Johnson K (1995a) Green-winged teal (*Anas crecca*). In: *The Birds of North America Online* (ed. Poole A). Cornell Lab of Ornithology, Ithaca, New York. Available from: http://bna.birds.cornell.edu/bna/species/193.

Johnson NK (1995b) Speciation in vireos. I. Macrogeographic patterns of allozymic variation in the Vireo solitarius complex in the contiguous United States. *Condor*, **97**, 903–919.

Johnson K, Sorenson MD (1999) Phylogeny and biogeography of dabbling ducks (genus *Anas*): a comparison of molecular and morphological evidence. *Auk*, **116**, 195–208.

Kaufman DC, Manley WF (2004) Pleistocene maximum and Late Wisconsinan glacier extents across Alaska, U.S.A. In: *Quaternary Glaciations—Extent and Chronology, Part II* (eds Ehlers J, Gibbard PL), pp. 9–27. Elsevier, Amsterdam.

Klein NK, Brown WM (1994) Intraspecific molecular phylogeny in the Yellow Warbler (*Dendroica petechia*) and its implications for avian biogeography in the West Indies. *Evolution*, **48**, 1914–1932.

Kondo B, Peters JL, Rosensteel BB, Omland KE (2008) Coalescent analyses of multiple loci support a new route to speciation in birds. *Evolution*, **62**, 1182–1191.

Kuhner MK (2006) LAMARC 2.0: maximum likelihood and Bayesian estimation of population parameters. *Bioinformatics*, **22**, 768–770.

Kuhner MK (2009) Coalescent genealogy samplers: windows into population history. *Trends in Ecology and Evolution*, **24**, 86–93.

Lack D (1944) Ecological aspects of species-formation in passerine birds. *Ibis*, **1944**, 260–286.

Lack D (1968) Bird migration and natural selection. *Oikos*, **19**, 1–9.

Lundberg S, Alerstam T (1986) Bird migration patterns: conditions for stable geographical population segregation. *Journal of Theoretical Biology*, **123**, 403–414.

Lynch M (2007) The frailty of adaptive hypotheses for the origins of organismal complexity. *Proceedings of the National Academy of Sciences USA*, **104**(Suppl. 1), 8597–8604.

Mallet J (2008) Hybridization, ecological races and the nature of species: empirical evidence for the ease of speciation. *Philosophical Transactions of the Royal Society B*, **363**, 2971–2986.

McCarthy EM (2006) *Handbook of Avian Hybrids of the World*. Oxford University Press, Oxford.

McCracken KG, Wilson RE, Peters JL, Winker K, Martin AR (2013) Late Pleistocene colonization of South Georgia by yellow-billed pintails predates the Last Glacial Maximum. *Journal of Biogeography*, **40**. doi: 10.1111/jbi.12162 (in press).

Melles M, Brigham-Grette J, Minyuk PS *et al.* (2012) 2.8 million years of arctic climate change from Lake El'gygytgyn, NE Russia. *Science*, **337**, 315–320.

Meng X (1994) Posterior predictive *P*-values. *The Annals of Statistics*, **22**, 1142–1160.

Montgomery Jr TH (1896) Extensive migration in birds as a check upon the production of geographical varieties. *American Naturalist*, **1896**, 458–464.

Nosil P (2012) *Ecological Speciation*. Oxford University Press, Oxford, UK.

Osada N, Wu C-I (2005) Inferring mode of speciation from genomic data: a study of the great apes. *Genetics*, **169**, 259–264.

Palmer RS (1976) *Handbook of North American Birds*, vol. 2. Yale University Press, New Haven, Connecticut.

Paradis E, Baillie SR, Sutherland WJ, Gregory RD (1998) Patterns of natal and breeding dispersal in birds. *Journal of Animal Ecology*, **67**, 518–536.

Peters JL, Gretes W, Omland KE (2005) Late Pleistocene divergence between eastern and western populations of wood ducks (*Aix sponsa*) inferred by the 'isolation with migration' coalescent method. *Molecular Ecology*, **14**, 3407–3418.

Peters JL, Zhuravlev Y, Fefelov I *et al.* (2007) Nuclear loci and coalescent methods support ancient hybridization as cause of mitochondrial paraphyly between gadwall and falcated duck (*Anas* spp.). *Evolution*, **61**, 1992–2006.

Peters JL, Zhuravlev YN, Fefelov I *et al.* (2008) Multilocus phylogeography of a Holarctic duck: colonization of North America from Eurasia by gadwall (*Anas strepera*). *Evolution*, **62**, 1469–1483.

Peters JL, McCracken KG, Pruett CL *et al.* (2012a) A parapatric propensity for breeding in teal (*Anas crecca*, sensu lato) precludes the completion of speciation. *Molecular Ecology*, **21**, 4563–4577.

Peters JL, Roberts TE, Winker K, McCracken KG (2012b) Heterogeneity in genetic diversity among non-coding loci fails to fit neutral coalescent models of population history. *PLoS One*, **7**, e31972.

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

Pyron RA, Burbrink FT (2010) Hard and soft allopatry: physicallyand ecologically mediated modes of geographic speciation. *Journal of Biogeography*, **37**, 2005–2015.

Rundell RJ, Price TD (2009) Adaptive radiation, nonadaptive radiation, ecological speciation and nonecological speciation. *Trends in Ecology and Evolution*, **24**, 394–399.

Rundle HD, Nosil P (2005) Ecological speciation. *Ecology Letters*, **8**, 336–352.

Salomonsen F (1955) The evolutionary significance of bird-migration. *Danske Biologiske Meddelelser*, **22**, 1–62.

Schwitters MT (2008) *Bird species found at Shemya Island, Alaska, 1999–2007*. US Fish & Wildlife Service Report AMNWR 08/15. Homer, Alaska.

Sella G, Petrov DA, Przeworski M, Andolfatto P (2009) Pervasive natural selection in the *Drosophila* genome. *PLoS Genetics*, **5**, e1000495.

Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics*, **68**, 978–989.

Swarth HS (1920) Revision of the avian genus *Passerella*, with special reference to the distribution and migration of the races in California. *University of California Publications in Zoology*, **21**, 75–224.

Via S (2009) Natural selection in action during speciation. *Proceedings of the National Academy of Sciences USA*, **106**, 9939–9946.

Via S, West J (2008) The genetic mosaic suggests a new role for hitchhiking in ecological speciation. *Molecular Ecology*, **17**, 4334–4345.

Winker K (2010) On the origin of species through heteropatric differentiation: a review and a model of speciation in migratory animals. *Ornithological Monographs*, **69**, 1–30.

Woerner AE, Cox MP, Hammer MF (2007) Recombination-filtered genomic datasets by information maximization. *Bioinformatics*, **23**, 1851–1853.

Zamoruyev V (2004) Quaternary glaciation of north-eastern Asia. In: *Quaternary Glaciations—Extent and Chronology, Part III* (eds Ehlers J, Gibbard PL), pp. 321–323. Elsevier, Amsterdam.

---

Designed research: K.W., J.L.P., K.G.M.; performed research, analysed data, and wrote paper: K.W., J.L.P., K.G.M., D.D.G.

---

## Data accessibility

GenBank accessions are numbers KF588710-KF589200. Table S1 contains sampling locations and specimen details. Input files are available on Dryad: doi:10.5061/dryad.76d82.

## Supporting information

Additional supporting information may be found in the online version of this article.

**Table S1** Specimen data and results of assignment tests from STRUCTURE analyses.

**Table S2** Genetic differentiation (ΦST) at eight nuclear loci and the mtDNA control region among three populations of green-winged teal (Aleutian population excludes Shemya Island birds).

**Fig. S1** MtDNA haplotype networks (top Old World clade; bottom New World).

**Fig. S2** Individual assignment probabilities to one of two populations using STRUCTURE (ver. 2.2.3).

**Fig. S3** Estimates of effective immigrants ($2N_em = \theta M/2$) from nuDNA and mtDNA.

**Fig. S4** Estimate of θ ($4N_eu$) from nuDNA for *Anas crecca nimia*.

**Fig. S5** Bayesian skyline plot for mtDNA data for all *nimia* individuals from Attu, Adak, and Ugamak islands, reconstructing historical demographics of this population ($n = 44$).

**Fig. S6** NuDNA haplotype networks.