Reproductive isolation and cryptic introgression in a sky island enclave of Appalachian birds

Brian S. Davidson1,2, Gene D. Sattler3, Sara Via2 & Michael J. Braun1,2

1Department of Vertebrate Zoology, National Museum of Natural History, MRC 534, Smithsonian Institution, 4210 Silver Hill Road, Suitland, Maryland 20746
2Program in Behavior, Ecology, Evolution, and Systematics, University of Maryland, College Park, Maryland 20742
3Department of Biology and Chemistry, Liberty University, 1971 University Boulevard, Lynchburg, Virginia 24502

Keywords
Chickadees, hybridization, introgression, metapopulation dynamics, Poecile, reproductive isolating mechanism, speciation.

Abstract
Reproductive isolation is central to the speciation process, and cases where the strength of reproductive isolation varies geographically can inform our understanding of speciation mechanisms. Although generally treated as separate species, Black-capped chickadees (Poecile atricapillus) and Carolina chickadees (P. carolinensis) hybridize and undergo genetic introgression in many areas where they come into contact across the eastern United States and in the northern Appalachian Mountains. The Great Smoky Mountains harbor the last large breeding population of atricapillus in the southern Appalachians, isolated from the species’ main range by nearly 200 km. This population is believed to be reproductively isolated from local carolinensis due to an unusual, behaviorally mediated elevational range gap, which forms during the breeding season and may function as an incipient reproductive isolating mechanism. We examined the effectiveness of this putative isolating mechanism by looking for genetic introgression from carolinensis in Great Smoky Mountain atricapillus. We characterized this population and parental controls genetically using hundreds of amplified fragment length polymorphism (AFLP) loci as well as mitochondrial DNA (mtDNA) sequence data from cytochrome-b. Great Smoky Mountain atricapillus have experienced nuclear genetic introgression from carolinensis, but at much lower levels than other populations near the hybrid zone to the north. No mitochondrial introgression was detected, in contrast to northern contact areas. Thus, the seasonal elevational range gap appears to have been effective in reducing gene flow between these closely related taxa.

Introduction
Population isolation results from changes in the geographic range of a species. Range expansion may result in colonization of previously unoccupied habitat patches, while range contraction can isolate peripheral populations into “islands” separate from the main range. Peripheral isolates can act as natural laboratories for evolutionary processes because they may experience different ecological and evolutionary pressures than populations in the species’ main range. While individual populations within a larger metapopulation may differ mildly from one another, geographically isolated populations are released from the homogenizing effects of gene flow and may take unique evolutionary trajectories. Geographical isolates may preserve and accumulate these differences over time, resulting in replicate natural experiments on speciation (Key 1968; Themudo and Arntzen 2007).

The evolution and ecology of peripherally isolated populations may also be influenced by interactions with parapatric (or sympatric) relatives along contact zones, where hybridization or ecological competition may occur (Barton and Hewitt 1985; Sætre and Sæther 2010). A peripherally isolated population completely surrounded by populations of a close relative is known as an enclave (Arntzen 1978). The formation of enclaves may be facilitated by the presence of habitat mosaics, or differential rates of hybrid zone movement. Moving contact zones, in which one species expands its range at the expense of another (see review in Buggs 2007), may result in either replacement or assimilation of the species whose range retracts, depending on the frequency of hybridization and

© 2013 The Authors. Ecology and Evolution published by John Wiley & Sons Ltd.
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
genetic introgression (Rhymer and Simberloff 1996; Mallet 2005). These represent special cases of range boundary dynamics.

Comparisons between peripheral and main range populations of a species can help elucidate the ecological, environmental, and population genetic processes that shape an organism’s responses to life at its range boundary. Here, we make such a comparison between enclave and main range populations of Black-capped Chickadee (*Poecile atricapillus*) in the Appalachian region. We assess genetic variation in these populations and those of a congener, Carolina Chickadee (*P. carolinensis*), with which *atricapillus* is known to hybridize. Using a multilocus molecular survey, we searched for genetic evidence of hybridization and introgression in an *atricapillus* enclave in the Great Smoky Mountains (GSM), the highest range in the southern Appalachians. We particularly wished to investigate whether geographically separate contact zones between the same taxa can result in fundamentally different levels of hybridization, based on local ecological differences, because this enclave appears to have evolved a unique reproductive isolating mechanism based on elevational movement (Tanner 1952; Tove 1980). This study also has significant conservation implications, as human impact on Appalachian ecosystems has been severe, in the form of habitat destruction, invasive species (Tingley et al. 2002; Simons et al. 2002), and climate change (Thomas and Lennon 1999; Inouye et al. 2000; Crick 2004).

*Poecile atricapillus* are small songbirds that are found throughout northern North America (Foote et al. 2010). Like many northern taxa, the range of *atricapillus* includes a southern salient through the Appalachian Mountains (Fig. 1). Other northern birds with Appalachian range extensions include Common Raven (*Corvus corax*), Red-breasted Nuthatch (*Sitta canadensis*), Black-throated Green Warbler (*Setophaga virens*), and Dark-eyed Junco (*Junco hyemalis*) to name just a few (Price et al. 1995).

In this region, the species’ continuous range extends as far south as southern West Virginia (WV) and southwestern Virginia (VA) in upper elevation forest characterized by northern tree species, while in the southern Appalachians *atricapillus* are only found in high elevation habitats that support Red Spruce (*Picea rubens*)/Fraser Fir (*Abies fraseri*)/Yellow Birch (*Betula alleghaniensis*) ecosystems. These “sky island” communities are widely scattered in the Blue Ridge province and have long been subject to anthropogenic habitat disturbance. With one exception, all *atricapillus* populations in this region have gone extinct within the last century (Lee 1999). The only remaining large population of *atricapillus* in the southern Appalachians is in the GSM National Park, on the border of Tennessee (TN) and North Carolina (NC).

Today, this sky island population is separated from the species’ contiguous range by nearly 200 km of marginal or unsuitable habitat, making it probable that the level of genetic exchange between it and main range populations is low. Due to its small geographic range, and restrictive habitat requirements, *atricapillus* is considered a species of concern in both TN and NC. Although protected by the national park, *atricapillus* has been identified as the southern Appalachian bird species most likely to become extirpated due to habitat destruction and the least likely to become reestablished in suitable but unoccupied habitat (Hunter et al. 1999).

Extensive hybridization between *atricapillus* and *carolinensis* occurs along the main range interface from New Jersey to Kansas (Brewer 1963; Rising 1968; Merritt 1978; Robbins et al. 1986; Bronson et al. 2003a; Curry 2005; Reudink et al. 2007; Olson et al. 2010) and in the northern Appalachians (Johnston 1971; Sattler and Braun 2000; Sattler et al. 2007). Southern Appalachian *atricapillus* are under threat of ecological replacement or genetic assimilation by the more southerly distributed, morphologically similar *P. carolinensis*. Each Appalachian sky island left vacant by extirpated *atricapillus* populations in the past 100 years has been colonized by *P. carolinensis* (Tanner 1952; Lee 1999).

In contrast to the main range contact zone between the species, hybridization between *atricapillus* and *carolinensis* has not been observed in the GSM (Tanner 1952; Tove 1980). Although the two forms occur together in winter
flocks in this area, a well-documented gap in their elevational distributions forms before the breeding season (Tanner 1952; Tove 1980), and has been implicated as a reproductive isolating mechanism. This gap in breeding distribution develops during early April, when carolinensis begin nesting below 900 m and atricapillus move upslope to the remaining spruce/fir forests above 1150 m. This distance is equivalent to at least 1.6 km horizontally, depending on slope (Tanner 1952). After the breeding season, the gap disappears as atricapillus move back downslope. The ultimate reason this elevational gap occurs is unknown, but carolinensis can be found breeding at elevations over 1800 m on nearby mountains where atricapillus are absent (Tanner 1952; Simpson 1992), suggesting that the gap is mediated by interspecific interactions rather than a difference in breeding habitat preferences.

Tanner (1952) and Tove (1980) concluded that the GSM population of atricapillus did not hybridize with local carolinensis based on the lack of morphological and vocal admixture, respectively. However, due to their similar morphological and the fact that their vocalizations are learned, molecular methods are more sensitive for differentiating these species and identifying hybrids (Sattler and Braun 2000; Bronson et al. 2003; Sattler et al. 2007). In fact, although hybridization was not always suspected in advance, all populations studied near the main range contact zone of atricapillus with carolinensis have been found to be heavily introgressed at the molecular level (Robbins et al. 1986; Sawaya 1990; Sattler and Braun 2000; Bronson et al. 2003a; Curry 2005; Reudink et al. 2007; Sattler et al. 2007; Olson et al. 2010). Thus, if the conclusions of Tanner (1952) and Tove (1980) regarding absence of hybridization are correct, the GSM population of atricapillus is unique in its purity.

The main goal of the present study was thus to assess the efficacy of the elevational gap as a reproductive isolating mechanism by determining whether cryptic hybridization or introgression between GSM P. atricapillus and local P. carolinensis is evident in the multilocus genotype of the GSM population. Previous studies of hybridization and introgression between these species have used relatively low numbers of highly differentiated molecular markers (Braun and Robbins 1986; Sattler and Braun 2000). Such diagnostic markers facilitate identification of hybrids, but may often underestimate the extent of genome-wide introgression because they are under selection opposing it (Sattler and Braun 2000). In fact, the degree of differentiation may vary dramatically among loci and genomic regions, probably as a result of the homogenizing effects of gene flow on some regions, while genetic incompatibilities or adaptive processes promote divergence in others (Harr 2006; Via and West 2008; Yuri et al. 2009). In order to gain a genome-level perspective on hybridization and introgression, we surveyed both mitochondrial cytochrome-b sequences and a relatively large number of amplified fragment length polymorphism (AFLP) loci from the nuclear genome. AFLP loci have several advantages over other marker types: They are largely neutral, being randomly generated from the whole genome, they require no prior sequence knowledge, they have high reproducibility and hundreds of loci can readily be studied in order to provide an approximation of genome-wide variation at low cost (Bensch and Åkesson 2005).

**Methods**

**Sampling design**

The sampling design for this study comprises 171 chickadees from seven populations (Fig. 1, Table 1). All sampling was performed with protocols approved by Smithsonian and/or University of Maryland Animal Care and Use Committees, under permits issued by state and national wildlife authorities. The focal population of P. atricapillus was sampled in the GSM National Park. All these samples were taken from areas >1500 m in elevation during the breeding season (29 May–27 June 2009) to minimize the possibility of sampling transient P. carolinensis (Tanner 1952; Simpson 1992). Due to conservation concerns, individuals in this population were mist-netted using song playback, then weighed, measured, banded, photographed, bled, and released. Blood was obtained by brachial vein puncture with a 26-gauge needle and 50–100 µL was preserved in the field using lysis buffer (Longmire et al. 1997). Two additional atricapillus population samples were used to represent main range parental: a sample from the northern Appalachian peninsular range of the species in WV, and a sample from Pennsylvania (PA), distantly allopatric from the hybrid zone with carolinensis. Both were previously described by Sattler and Braun (2000).

Four populations of P. carolinensis were sampled: Ohio (OH) and VA previously collected by Sattler and Braun (2000), and newly collected populations from NC and TN (Fig. 1, Table 1). Locations for sampled carolinensis populations were chosen to represent potential sources for introgression into Appalachian atricapillus populations. Samples NC and TN were collected by shotgun and tissue samples frozen in the field. Specimens were measured in the field and will be prepared as study skins for deposit at the Smithsonian Institution’s National Museum of Natural History (NMNH). An additional three carolinensis individuals from Louisiana (LA; Braun and Robbins 1986) known to represent the western carolinensis mitochondrial haplotype (Sawaya 1990; Gill et al. 1999) were also sequenced as controls to help characterize carolinensis haplotypes as eastern or western.
Table 1. Population sampling.

<table>
<thead>
<tr>
<th>Population</th>
<th>Sample size</th>
<th>Date</th>
<th>Coordinates</th>
<th>Tissue numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poecile atricapillus</td>
<td>Pennsylvania (PA)</td>
<td>20</td>
<td>1991</td>
<td>81.52 N, 77.65 W</td>
</tr>
<tr>
<td></td>
<td>West Virginia (WV)</td>
<td>20</td>
<td>1990</td>
<td>81.90 N, 75.25 W</td>
</tr>
<tr>
<td></td>
<td>Great Smoky Mts (GSM)</td>
<td>30</td>
<td>2009</td>
<td>35.60 N, 83.45 W</td>
</tr>
<tr>
<td>Poecile carolinensis</td>
<td>Ohio (OH)</td>
<td>20</td>
<td>1991</td>
<td>38.72 N, 82.57 W</td>
</tr>
<tr>
<td></td>
<td>Virginia (VA)</td>
<td>21</td>
<td>1991</td>
<td>37.33 N, 77.85 W</td>
</tr>
<tr>
<td></td>
<td>Tennessee (TN)</td>
<td>30</td>
<td>2009</td>
<td>35.70 N, 85.25 W</td>
</tr>
<tr>
<td></td>
<td>North Carolina (NC)</td>
<td>30</td>
<td>2009</td>
<td>35.00 N, 79.50 W</td>
</tr>
<tr>
<td></td>
<td>Louisiana (LA)</td>
<td>3</td>
<td>1979</td>
<td>30.05 N, 90.39 W</td>
</tr>
</tbody>
</table>

West Virginia, Virginia, Ohio, and Pennsylvania were collected by G. Sattler (Sattler and Braun 2000). Louisiana samples were collected by M. Braun (Braun and Robbins 1986). All specimen numbers are NMNH (USNM) tissue numbers.

Molecular methods

DNA was extracted from blood and breast muscle tissue using a standard proteinase K/phenol-chloroform protocol (mouse tail) on an AutoGenprep 965 extraction system (Autogen Inc., Holliston, MA). DNA concentration and purity were assessed using a NanoDrop ND-1000 spectrophotometer (Nanodrop Products, Wilmington, DE). All GSM, NC, and TN individuals were sexed with the PCR-based assay of Griffiths et al. (1998), except two GSM individuals where PCR amplification failed (PA, OH, WV, and VA birds were previously sexed by gonadal inspection; Sattler et al. 2007). The mtDNA cytochrome-b gene was amplified and sequenced using primers L14841 (Kocher et al. 1989) and H15498 (Mariaux and Braun 1996), yielding a 656 bp fragment. The amplification PCR included 1 μl GoTaQ PCR buffer (Promega, Madison, WI), 0.2 mmol/L of each dNTP, 0.75 μmol/L of each primer, 1.5 mmol/L MgCl2, 0.625 U Taq polymerase (Promega GoTaq), and 5 ng of whole genomic DNA in a 25 μL reaction vessel. The cycling profile consisted of 35 repetitions of 95°C for 30 sec, 50°C for 30 sec, and 72°C for 60 sec, with a final 10 min hold at 72°C for fragment extension. PCR products were cleaned using 3 μL of Exosap-IT (United States Biochemical, Cleveland, OH). The sequencing reactions included 80 mmol/L Tris pH 9.0, 2 mmol/L MgCl2, 1 μmol/L primer, 2 μL amplification product, and 0.75 μL BigDye (Life Technologies, Grand island, NY) in a 10 μL reaction. The cycle-sequencing profile consisted of 30 cycles of 94°C for 30 sec, 50°C for 15 sec, and 60°C for 4 min, followed by a 10°C hold. Sequencing products were cleaned with Sephadex G-50 columns (GE Healthcare Biosciences, Pittsburgh, PA). Sequencing was performed on an ABI 3730xl DNA Analyzer (Life Technologies). Chromatograms were examined and sequences trimmed, assembled, and edited using Sequencher 4.10.1 (Gene Codes Corp., Ann Arbor, MI). Consensus sequences for all individuals were aligned with Sequence Alignment Editor v2.0a11 (http://iubio.bio.indiana.edu/soft/iubionew/molbio/dna/analysis/Pist/main.html). Haplotypes were identified using MacClade 4.08 (Sinauer Associates Inc., Sunderland, MA), and a cytochrome-b haplotype network constructed using the median-joining method in the program Network 4.5 (fluxus-engineering.com, Bandelt et al. 1999).

AFLPs were generated following the protocol of Vos et al. (1995) as modified for vertebrates by Kingston and Rosel (2004). Ten selective PCR primer pairs were used, consisting of all combinations of two fluorescently labeled EcoR1 + ANN primers and five Taq1 + ANN primers (Davidson 2011). Fragment profiles were generated by capillary electrophoresis on an ABI PRISM 3130xl Genetic Analyzer (Life Technologies).

Fragments were scored using GeneMapper 4.0 software (Life Technologies). All samples were scored concurrently and blindly for each selective primer pair. AFLP fragments were grouped in 1 bp bins ranging in size from 90 to 350 bp. The scoring protocol developed by Kingston and Rosel (2004) was used to minimize potential noise associated with underamplification of large fragments or uneven amplification among samples. A threshold of 100 fluorescence units was set as the minimum amplitude for fluorescence peaks to qualify as potential marker loci, while baseline fluorescence was generally below 50 units. The presence of false-negative peaks (<100 fluorescence units) in a bin with scorable peaks from other individuals resulted in the rejection of the marker. For each primer pair, the fragment length of the largest monomorphic marker was taken as the upper size limit for scorable loci to prevent scoring problems resulting from PCR drop off with fragment length. One sample from each population was reprocessed for all primer pairs and scored anonymously and concurrently with all other samples to verify reproducibility of AFLP marker generation.

Data analyses

Molecular diversity indices for mtDNA and the partitioning of variation in mtDNA and AFLP within and among populations and species were assessed with the Analysis of
Molecular Variance (AMOVA) routine in Arlequin 3.5 (Excoffier et al. 2005). We also used Arlequin 3.5 to quantify population differentiation in AFLP loci among populations and calculate pairwise $F_{ST}$ values. AFLP profiles were coded as binary haplotype restriction fragment length polymorphism (RFLP) data, and significance was calculated from 100,172 permutations. These $F_{ST}$ values provide useful measures of genetic differentiation between pairs of populations, but they are not calculated from allele frequencies and therefore cannot be directly compared to $F_{ST}$ values determined using codominant markers (Excoffier et al. 2005).

Using multilocus AFLP scores, we constructed three-dimensional clouds of all sampled individuals by Nonmetric Multidimensional Scaling (NMDS) ordination using NTSYSpc (Rohlf 2000). A Jaccard similarity matrix was calculated from AFLP data for all pairs of individuals using the equation $J_{ij} = n_{11} / (n_{11} + n_{10} + n_{01})$, where $n_{ij}$ is the number of polymorphic markers for which the character states (1 or 0) are found for a pair of samples $i$ and $j$. This approach is appropriate for determining similarity between AFLP profiles because it is conservative in that it does not assume that the band-absent phenotypes are homologous. The Jaccard matrix was used to generate a Principal Coordinates Analysis, which served as the input for NMDS. A stress value set from 0.0 to 1.0 was used to measure goodness of fit, where zero indicated perfect fit between the NMDS coordinates and the Jaccard matrix, and one indicates no relationship between the two. Each individual was plotted in ordinal space using its three-dimensional coordinates.

We used the Bayesian population genetic clustering algorithm in the software STRUCTURE 2.3 (Pritchard et al. 2000) adjusted for dominant markers (Falush et al. 2007) to cluster individual AFLP profiles. Three analyses were performed: one using both species and all seven sampled populations; and one for each of the two individual species to detect intraspecific population clustering. To determine the most appropriate model we tested three analyses performed: one using both species and all seven sampled populations; and one for each of the two individual species to detect intraspecific population clustering. To determine the most appropriate model we tested several combinations of input parameters at all levels of $K$ clusters. We ultimately used a model that assumed admixture, alleles correlated between populations, and used sampling locations as priors (Hubisz et al. 2009). We then determined the Q-values of all individuals by running the model for each value of $K$ (from $K = 1$–14) replicated ten times with burn-in of 100,000 steps followed by Markov chain Monte Carlo sampling of 1,000,000 steps. After completing STRUCTURE runs, the results of the unsupervised STRUCTURE replicates were aligned using CLUMPP 1.1.2 (Jakobsson and Rosenberg 2007), and graphics generated using DISTRACT 1.1 (Rosenberg 2004). We chose to present all informative values of $K$ generated by the unsupervised models (following Rosenberg et al. 2002) instead of choosing a specific $K$ value for two reasons: (1) there may be more than one biologically informative value of $K$ (e.g., see Wang et al. 2007); and (2) the established criteria for choosing an optimal $K$ value rely on ad hoc methods (Pritchard et al. 2000; Evanno et al. 2005; Pritchard et al. 2010).

We augmented the STRUCTURE results with a simple test for introgression, taking advantage of the fact that GSM atricapillus were divergent in allele frequency from their parental populations at a number of AFLP loci. If this divergence was due to introgression from carolinensis, it should be possible to predict the direction of GSM divergence for each locus from the carolinensis frequency for that locus. If divergence was due to local differentiation of the GSM population (through drift, adaptation, or isolation by distance), its directionality should be random with respect to carolinensis. To determine whether these loci showed an overall pattern of introgression, a subset of markers were chosen where the frequency difference of positive AFLP phenotypes between parental atricapillus and the focal population was above 7.5% (corresponding to 2/28 GSM birds or 3/40 parentals). The frequency of the band present phenotype in the GSM population also had to be free to vary in either direction around parental atricapillus frequency, necessitating an upper boundary of 90% and a lower boundary of 10% in parentals. These restrictions left 17 loci that met the criteria. A sign test was applied to determine whether, for these 17 loci, divergence of the GSM population varied randomly in direction from parental atricapillus or was biased toward carolinensis (indicative of introgression).

**Results**

**mtDNA**

Over a 656 bp cytochrome- $b$ amplicon, there were 25 fixed differences between atricapillus and eastern carolinensis haplotype groups (3.8%), and 30 fixed differences between atricapillus and western carolinensis haplotype groups (4.6%). There were 17 fixed differences between eastern and western carolinensis haplotype groups (2.6%). All carolinensis in our four study populations flanking the Appalachians exhibited the eastern carolinensis haplotype group (Table 2).

To insure haplotype accuracy, further comparisons were focused on a 535 bp region of the cytochrome- $b$ amplicon for which double-stranded sequence was obtained for all individuals. There were no shared haplotypes between atricapillus and carolinensis, and no admixture of atricapillus and carolinensis haplotype groups in any population (Table 2, Fig. 1). Each haplotype group was characterized by a single abundant haplotype that
predominated in all populations of that species and a series of closely related haplotypes differing by one or two substitutions (Table 2, Fig. 2). No individuals were identified as migrants, and there was no evidence of carolinensis mtDNA introgression into the GSM atricapillus population (Fig. 1). AMOVA indicated that 97.57% of mtDNA variation was between species, the rest within. Cytochrome-b molecular diversity was lower for GSM than for all other populations, with only one variant haplotype among 30 individuals (Table 2, Fig. 2).

### Table 2. Diversity indices for each population calculated from cytochrome-b mtDNA sequence data.

<table>
<thead>
<tr>
<th>Population</th>
<th>Sample size</th>
<th>No. haplotypes (No. unique haplotypes)</th>
<th>( \Theta_p )</th>
<th>Average pairwise seq. divergence</th>
<th>( \Theta_S ) Nucleotide polymorphism</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSM</td>
<td>30</td>
<td>2(1)</td>
<td>0.06667</td>
<td>0.25242</td>
<td></td>
</tr>
<tr>
<td>WV</td>
<td>20</td>
<td>3(2)</td>
<td>0.40000</td>
<td>1.12748</td>
<td></td>
</tr>
<tr>
<td>PA</td>
<td>20</td>
<td>6(5)</td>
<td>0.61988</td>
<td>1.43057</td>
<td></td>
</tr>
<tr>
<td>VA</td>
<td>21</td>
<td>4(2)</td>
<td>0.55238</td>
<td>1.11181</td>
<td></td>
</tr>
<tr>
<td>OH</td>
<td>20</td>
<td>5(2)</td>
<td>0.40000</td>
<td>1.12748</td>
<td></td>
</tr>
<tr>
<td>NC</td>
<td>30</td>
<td>4(2)</td>
<td>0.49195</td>
<td>0.75726</td>
<td></td>
</tr>
<tr>
<td>TN</td>
<td>30</td>
<td>7(4)</td>
<td>0.50000</td>
<td>1.79881</td>
<td></td>
</tr>
<tr>
<td>LA</td>
<td>3</td>
<td>1(1)</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

Population codes – atricapillus: PA, Pennsylvania; WV, West Virginia; GS, Great Smoky Mountains; carolinensis: OH, Ohio; VA, Virginia; NC, North Carolina; TN, Tennessee.

**AFLP analysis**

We scored 276 AFLP loci from 10 primer pairs. Of these, 11 were monomorphic and 265 were polymorphic. Two loci had fixed differences between parental atricapillus and carolinensis and three more had frequency differences >0.9, but the vast majority of loci (~90%) had frequency differences <0.1 (Fig. 3). When diagnostic AFLP locus scores were compared with cytochrome-b haplotypes, no individuals exhibited cytonuclear mismatch and no migrants were detected.

NMDS of multilocus AFLP profiles revealed two clouds of individuals, corresponding to the two forms (Fig. 4). Population centroids of carolinensis were tightly clustered, while those of atricapillus, especially GSM, were more dispersed (Fig. 4). There was no evidence of F1 or early backcross hybrids, which would appear spatially intermediate between the two forms in NMDS, nor was there obvious intermediacy in the GSM population according to this analysis.

Results of AMOVA on the AFLP data indicated that genetic variation is partitioned 22.27% by species, 1.63% among populations within species, and 76.10% within populations. Pairwise \( F_{ST} \) showed that GSM atricapillus were less distant from carolinensis populations (mean \( F_{ST} = 0.215 \)) than were the parental atricapillus populations (mean \( F_{ST} = 0.251 \); Table 3). \( F_{ST} \) values for GSM compared to the parental atricapillus populations were higher than for PA and WV compared to each other, but only the GSM to PA comparison was significant (Table 3). One intraspecific \( F_{ST} \) in carolinensis was also significant (TN to VA).

In STRUCTURE analyses, posterior probabilities were highest for models including admixture, population priors, and correlated alleles. For all seven populations, models were run for values of \( K \) from 1 to 7. The results of the \( K = 2 \) model suggested that introgression of carolinensis alleles into the GSM atricapillus comprises ~5% of

---

2490

© 2013 The Authors. *Ecology and Evolution* published by John Wiley & Sons Ltd.
their genome (Fig. 5). A small amount of introgression of atricapillus alleles into OH was also evident with $K = 2$, consistent with introgression observed by Sattler and Braun (2000) in a RFLP marker. At $K = 3$, a low-level cluster appeared in all populations, apparently reflecting individual variation. In the $K = 4$ model, there was a signal of local differentiation in the GSM atricapillus population (5–10% of the genomic signal), some of which was shared at a lower frequency by NC carolinensis (Fig. 5). Models with $K > 4$ were all very similar to results with $K = 4$, differing only in finer and finer subdivision of the low-level signal of individual variation apparent at $K = 3$ (not shown). STRUCTURE analyses restricted to atricapillus populations only yielded no appreciable population structure (Fig. 6), as did runs restricted to carolinensis only (not shown).

Seventeen AFLP loci met the criteria for inclusion in the sign test for introgression. The GSM population frequency was shifted toward carolinensis for 14 of these 17 loci. This ratio is significantly different from the expectation of 0.5 under the null hypothesis of random variation in direction from the parental atricapillus frequencies ($P = 0.013$).

### Discussion

**Cryptic introgression in GSM**

Based on morphological markers, Tanner (1952) suggested that hybridization between *P. atricapillus* and *P. carolinensis* in the GSM was prevented by a behaviorally mediated elevational gap in the species’ breeding ranges, a conclusion substantiated by Tove (1980) based on vocal studies. We were also unable to detect substantial signals of hybridization or introgression in mtDNA, with NMDS ordination of AFLP phenotypes, or through
genetic clustering of AFLP data from *P. atricapillus* alone using STRUCTURE. However, the consistently lower $F_{ST}$ values of GSM relative to other *atricapillus* populations in comparison to *carolinensis*, suggested the possibility of nuclear introgression into GSM, which was then confirmed by the significant sign test for the directionality of GSM’s divergence from parental AFLP frequencies. The STRUCTURE analyses of AFLP data from both species revealed a small but consistent signal of *carolinensis* introgression in all members of the GSM population, amounting to ~5% of the nuclear genomic signal. This signal indicates that a low level of hybridization actually is occurring now or has occurred in the past.

Given the low levels of introgression observed here, it is not surprising that hybridization between chickadee species in the GSM was previously hard to detect. Our earlier studies of the northern Appalachian hybrid zone demonstrated that genetic techniques would provide more sensitive measures of hybridization and introgression (Sattler and Braun 2000; Sattler et al. 2007), and application of genetic data now demonstrates that nuclear introgression has occurred in the GSM enclave, albeit at much lower levels than in the northern Appalachian contact zone. Introgression was not detected in mtDNA, a popular marker (Zink and Barrowclough 2008), but mtDNA may not introgress freely between these species; it has been shown to act in a nonneutral fashion in earlier studies of chickadees (Sattler and Braun 2000) and other systems (Ballard and Whitlock 2004; Bazin et al. 2006).

Ancestral polymorphism is often mentioned as an alternative explanation for the presence of alleles characteristic of one species in another. However, previous studies of the hybrid zone between these species show clear patterns of clinal variation in allele frequency for diagnostic markers (e.g., Table 2 in Sattler and Braun 2000). Such clinal variation is a hallmark of introgression, but is not expected with ancestral polymorphism. This argument can be extended to the AFLP data presented here, as the GSM population is in contact with *carolinensis* during much of the year and only narrowly separated during the breeding season, while WV and PA *atricapillus* are allopatric. Thus, we would expect to see introgression affect GSM first, while ancestral polymorphism should be equally apparent in all three *atricapillus* populations.

Interestingly, *atricapillus*-only STRUCTURE analyses were unable to detect any well-defined divergence of GSM from parental *atricapillus* populations, despite the clear signal of introgression when *carolinensis* was included. This highlights the general difficulty in
classifying low levels of divergence in multilocus data (Evanno et al. 2005; Pritchard et al. 2010). It seems that STRUCTURE was unable to distinguish between low-level introgression and individual variation in the atricapillus-only runs, perhaps due to the lack of genetic clusters from pure carolinensis, which may provide the algorithm with a basis for identifying the carolinensis signal in GSM. Similarly, a signal suggesting local differentiation of the GSM population from main range atricapillus emerged at values of K = 4 and higher in the STRUCTURE analyses of all populations, but not atricapillus only. We speculate that this signal is also reinforced when all populations are included, because it also appears in geographically adjacent carolinensis from NC.

It is unclear whether hybridization is currently occurring in the GSM population. The low level of introgression observed and our failure to detect immigrant carolinensis or early generation hybrids suggests that levels of ongoing hybridization may be quite low. An alternative explanation is that the observed introgression may be due to former contact between these populations. However, an inherent bias in our sampling strategy should be recognized. Our samples were heavily skewed toward males due to the use of song playback for collecting birds (n = 3 for females, n = 25 for males in GSM, two birds were unsexed). In many songbirds, including atricapillus, females disperse farther than males (Weise and Meyer 1979; Greenwood 1980). Thus, the sex most likely to immigrate was underrepresented, limiting to some degree our ability to detect migrants. For this reason, we cannot exclude the possibility of ongoing immigration and hybridization in GSM.

Life history considerations suggest that low-level ongoing hybridization would be hard to detect. GSM atricapillus move downslope to more sheltered woodlands in the nonbreeding season and flock with carolinensis there (Tanner 1952; Simpson 1992). It is during this time that chickadees form pair bonds in winter flocks (Smith 1991). Females of both species prefer larger, more dominant males (Bronson et al. 2003b). Unlike other areas of contact between these two species, in the GSM atricapillus are substantially larger than carolinensis (Tanner 1952; Brewer 1963; Merritt 1981), which may make the males more attractive as mates. This suggests a mechanism by which hybridization might occur in this region: carolinensis females pairing with atricapillus males on the wintering grounds, then moving upslope with them during the breeding season. This phenomenon would be cryptic, as the species are very similar morphologically, and females rarely sing and are less likely to be attracted to song playback (Smith 1991).

Breeding season range gaps between atricapillus and carolinensis, with apparently suitable habitat in the unoccupied zones, have been reported in Illinois and Indiana (Brewer 1963; Merritt 1981). Like Tanner (1952), Merritt (1981) observed movement of atricapillus away from the range interface immediately preceding the advent of the nesting season in Indiana, albeit the movement was latitudinal and not elevational. However, the difference between suitable nonbreeding habitat and optimal breeding season habitat may have been overlooked in the Midwestern studies (see discussion in Robbins et al. 1986; Grubb et al. 1994). Whether these gaps are transient or a permanent feature of the contact zone in these areas deserves further study, as does their impact on genetic admixture between the two forms.

A novel reproductive isolating mechanism

Unlike other areas of contact between atricapillus and carolinensis, hybridization in the GSM seems to be rare. We did not detect immigrant carolinensis or F1 individuals in GSM, and the impact of introgression is low. The most likely explanation is that the elevational gap between the species in the GSM actually does act to retard hybridization and introgression. These results suggest that a unique reproductive barrier is at work at this enclave contact zone.

Breeding carolinensis populations exist at all elevations in all nearby southern Appalachian sky islands (Tanner 1952; Simpson 1992; Lee 1999) demonstrating that the upper reaches of the GSM should be suitable habitat for both chickadees. However, only atricapillus currently occupy this territory. Tanner (1952) suggested that the continued existence of GSM atricapillus was due not only to the breeding season gap, but also to the relatively large and dense population. The high population density of atricapillus in GSM relative to carolinensis, especially when compared to other contact zone studies (e.g., Brewer 1963) may act to insulate the population from genetic erosion. Taken together, these considerations suggest that the behaviorally mediated reproductive isolation in this case may be more effective at preventing hybridization than other areas of microallopatry between these species (Brewer 1963; Merritt 1981).

Marker choice in studies of hybridization and phylogeography

Mitochondrial DNA introgression was not detected in any of the GSM sample, and the haplotype groups of atricapillus and carolinensis are substantially divergent, making mtDNA among the most structured markers between carolinensis and atricapillus in this study. Often diagnostic markers such as mtDNA have been used in hybridization research because intermediacy in a diagnostic marker is
an unambiguous signal of hybridization, and because mtDNA has often been considered a leading indicator of differentiation (Zink and Barrowclough 2008). However, markers chosen for high levels of differentiation are more likely to be under purifying selection, which limits the ability to detect and measure introgression (Yuri et al. 2009). In this study, AFLP profiles were more informative than mtDNA haplotypes in that they were able to detect low-level introgression in GSM. This may be due to the larger number of AFLP loci and lower average levels of purifying selection on them, allowing more prevalent introgression than mtDNA. This finding is consistent with the now frequent observation that mtDNA can be a useful marker of population structuring, but that multiple independent loci are needed to insure an accurate portrayal of phylogeography, especially when gene flow and introgression are likely (e.g., Carling and Brumfield 2008).

Acknowledgments

We thank Paul Super at the Great Smoky Mountains National Park, as well as the National Parks Service; Daron Barnes and Christine Kelly at the NC Wildlife Resources Commission; Richard Kirk at the TN Wildlife Resources Agency; the U.S. Fish and Wildlife Service; and the U.S. Geological Survey. The Smithsonian Institution granted Animal Care and Use permission for collecting samples. The Smithsonian Institution and the University of Maryland provide financial support and laboratory access. We would also like to thank Sarah Kingston, Chris Huddleston, and Jeff Hunt at NMNH, Charlie Fenster at the University of Maryland, and Ellen Paul at the Ornithological Council for their help with this project.

Author Contributions

B. S. D. designed and performed research, collected samples, analyzed the data and wrote the manuscript; M. J. B. designed the research and wrote the manuscript; G. D. S. collected samples and edited the manuscript; S. V. helped design the research and edited the manuscript.

Data Accessibility

DNA sequences: GenBank accessions KF134214-KF134386. AFLP data deposited in the Dryad Repository: http://dx.doi.org/10.5061.dryad.3nr84. Tissue samples (see Table 1) are archived at the US National Museum of Natural History (USNM).

Conflict of Interest

None declared.

References


and Carolina chickadees at a contact zone in the Appalachian Mountains. Auk 117:427–444.