MOLECULAR ECOLOGY

Molecular Ecology (2011) **20**, 3102–3115 doi: 10.1111/j.1365-294X.2011.05159.x

Genetic variation and seasonal migratory connectivity in Wilson's warblers (*Wilsonia pusilla*): species-level differences in nuclear DNA between western and eastern populations

DARREN E. IRWIN,*† JESSICA H. IRWIN*† and THOMAS B. SMITH*‡

*Center for Tropical Research, Institute of the Environment, University of California Los Angeles, 619 Charles E. Young Dr, Los Angeles, CA 90095, USA, †Department of Zoology and Biodiversity Research Centre, University of British Columbia, 6270 University Blvd., Vancouver, BC, V6T 1Z4, Canada, ‡Department of Ecology and Evolutionary Biology, University of California Los Angeles, Los Angeles, CA 90095, USA

Abstract

There is growing interest in understanding patterns of seasonal migratory connectivity between breeding and wintering sites, both because differences in migratory behaviour can be associated with population differentiation and because knowledge of migratory connectivity is essential for understanding the ecology, evolution and conservation of migratory species. We present the first broad survey of geographic variation in the nuclear genome of breeding and wintering Wilson's warblers (*Wilsonia pusilla*), which have previously served as a research system for the study of whether genetic markers and isotopes can reveal patterns of migratory connectivity. Using 153 samples surveyed at up to 257 variable amplified fragment length polymorphism markers, we show that Wilson's warblers consist of highly distinct western and eastern breeding groups, with all winter samples grouping with the western breeding group. Within the west, there is weak geographic differentiation, at a level insufficient for use in the assignment of wintering samples to specific areas. The distinctiveness of western and eastern genetic groups, with no known intermediates, strongly suggests that these two groups are cryptic species. Analysis of mitochondrial cytochrome *b* sequence variation shows that the estimated coalescence time between western and eastern clades is approximately 2.3 Ma, a surprisingly old time of divergence that is more typical of distinct species than of subspecies. Given their morphological similarity but strong genetic differences, western and eastern Wilson's warblers present a likely case of association between divergence in migratory behaviour and the process of speciation.

Keywords: AFLP, Aves, cryptic species, migratory connectivity, Parulidae, speciation

Received 22 February 2011; revision received 6 May 2011; accepted 11 May 2011

Introduction

How seasonal migratory behaviour relates to population differentiation is of growing interest to ecologists, conservationists and evolutionary biologists. A full understanding of the year-round ecology of migratory organisms depends critically on the knowledge of patterns of migratory connectivity between specific breeding and wintering populations (Webster *et al.* 2002; Smith *et al.* 2005; Carlisle *et al.* 2009). Effective conservation planning also depends on such knowledge, because threats to population persistence can occur in the breeding and wintering ranges as well as the migratory routes in between and because migratory differences can point to genetically divergent populations (Bensch *et al.* 1999; Ruegg & Smith 2002) that are important to conserve. In an evolutionary context,

Correspondence: Darren Irwin, Fax: 604-822-2416;
E-mail: irwin@zoology.ubc.ca

© 2011 Blackwell Publishing Ltd

differences in migratory pathways can reveal populations that show some degree of reproductive isolation (Ruegg 2007; Bensch *et al.* 2009), and it is even possible that the migratory differences themselves may lead to reproductive isolation (Helbig 1991; Irwin & Irwin 2005; Irwin 2009; Rohwer & Irwin 2011), thereby contributing to speciation.

Given the importance of migratory connectivity to ecology, conservation and evolution, there is intense interest in developing better methods for revealing patterns of connectivity (Webster *et al.* 2002). While mark–recapture techniques have been somewhat useful in revealing broad patterns for some species (e.g. Swainson's thrush; Ruegg & Smith 2002), for most species the recovery rate is extremely low, requiring immense effort for very low resolution of patterns of connectivity (Webster *et al.* 2002). Thus, new approaches using telemetry, stable isotopes and/or genetic signatures are receiving much attention (Webster *et al.* 2002; Kelly *et al.* 2005; Boulet & Norris 2006). While all of these approaches show potential for uncovering patterns of connectivity, particularly when used in combination, genetic analysis is particularly informative for characterizing evolutionary divergence between populations.

The Wilson's warbler (*Wilsonia pusilla* Wilson 1811) has emerged as a focal species for the study of how stable isotopes and genetic markers can be used to describe patterns of migratory connectivity (Kelly *et al.* 2002; Kimura *et al.* 2002; Clegg *et al.* 2003; Smith *et al.* 2005). With a breeding range extending through much of Canada and the United States and a wintering range through much of Mexico and Central America (Fig. 1; Dunn & Garrett 1997; Ammon & Gilbert 1999), this species offers considerable potential for revealing detailed patterns of connectivity. As a colourful and widespread Nearctic–Neotropical migrant that has shown large population declines (e.g. Breeding Bird Survey survey-wide growth rate of −2.3% per year between 1996 and 2007; Sauer *et al.* 2008), it is also of considerable conservation interest. Being too small for even the smallest available telemetry devices that could track movements over large geographic areas, these birds require alternative methods, such as feather isotopes and genetic signatures, to determine their patterns of connectivity. Feather isotopes have been most useful as an indicator of latitude (Kelly *et al.* 2002), enabling, for example, the elucidation of a pattern of 'leapfrog migration' in which birds that breed furthest north spend the winter furthest south (Kelly *et al.* 2002; Clegg *et al.* 2003; Paxton *et al.* 2007).

In contrast, molecular analyses of Wilson's warblers across the breeding range have been most informative in revealing longitudinal rather than latitudinal variation, showing large differences between western and

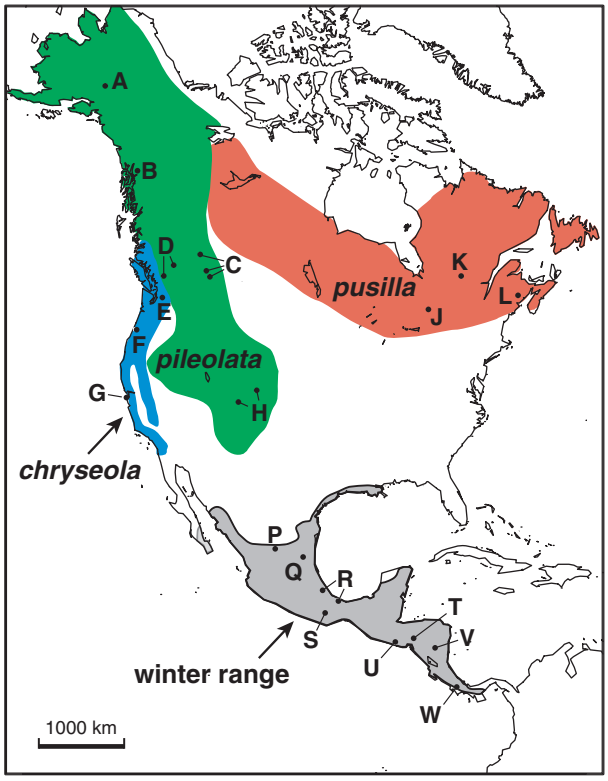

Fig. 1 Map of North and Central America showing breeding ranges of three subspecies of Wilson's warbler (blue, *Wilsonia pusilla chryseola*; green, *W. p. pileolata*; red, *W. p. pusilla*) and their combined winter range (grey), along with sampling sites (designated with letters, according to Table 1). Range information was obtained from Dunn & Garrett (1997) and Ammon & Gilbert (1999). The boundaries between subspecies have not been well studied; hence, the map shows their approximate positions.

eastern regions in mitochondrial DNA (Kimura *et al.* 2002) and microsatellites (Clegg *et al.* 2003) compared with the variation within each of those regions. These differences parallel previously recognized but subtle differences in coloration of western and eastern populations, which have been used as the basis for designating three named subspecies (Dunn & Garrett 1997). These include two subspecies breeding in the west (*W. p. pileolata*, from Alaska south through interior British Columbia and the Rocky Mountains to northern New Mexico; *W. p. chryseola*, from southwestern British Columbia through southern California; Fig. 1) and one in the east (*W. p. pusilla*, from Alberta and the Northwest Territories to the east coast of Canada; Dunn & Garrett 1997). However, previous genetic studies were unable to resolve patterns of migratory connectivity within the western and eastern groups (Kimura *et al.* 2002; Clegg *et al.* 2003). Given that previous research surveyed only small numbers of genetic loci (i.e.

© 2011 Blackwell Publishing Ltd

mtDNA, which is inherited as a single unit, and eight microsatellite loci), we decided to survey of a much larger number of genetic markers distributed throughout the genome, with the goal of answering several questions. First, do these markers enable assignment of wintering and migratory individuals to specific breeding regions? Second, how closely do patterns of genomic variation correspond to the three named subspecies? Third, are the strong east–west differences observed in mitochondrial DNA representative of overall genomic divergence such that western and eastern forms of Wilson's warbler might actually be reproductively isolated species?

Motivated by these questions, we conducted a survey of geographic variation in a large number (257) of genetic markers spread throughout the Wilson's warbler genome, using the amplified fragment length polymorphism (AFLP) method (Vos *et al.* 1995; Mueller & Wolfenbarger 1999; Bensch & Åkesson 2005; Bonin *et al.* 2007). Compared with previously used methods (mtDNA sequencing and microsatellites), AFLP has the advantage that it surveys a large number of markers and as a consequence is better representative of overall genomic divergence and therefore has more potential to capture geographic variation. Given that AFLP has been highly successful in resolving within-species population structure in other study systems (e.g. Bensch *et al.* 2002; Campbell & Bernatchez 2004; Irwin *et al.* 2005, 2009b; Milá *et al.* 2007, 2009; Freedman *et al.* 2010; Brelsford *et al.* 2011) and in discovering cryptic species (Irwin *et al.* 2005; Toews & Irwin 2008), we use it in Wilson's warblers to determine (i) whether there is sufficient population structure within western and eastern groups of Wilson's warblers to allow the use of these markers to infer breeding locations of wintering birds and (ii) how divergent the three named subspecies are in their nuclear genomes. We also estimate the mitochondrial divergence time between western and eastern groups of Wilson's warblers using a novel analysis of cytochrome *b* sequences available from GenBank.

Methods

Sampling

Blood and/or feather samples of Wilson's warblers were obtained throughout much of the breeding and wintering ranges (Table 1, Fig. 1) as part of a long-term effort to examine patterns of connectivity in this species. Most of the samples used in this study were also used by Kimura *et al.* (2002) and Clegg *et al.* (2003); see those studies for detailed descriptions of sampling and DNA extraction methods. Table 1 describes 24 sampling locations included in the present study as well as their sample sizes. Nearby locations were grouped into 'sites' that are designated with letters in Fig. 1 according to Table 1.

AFLP

Genomic DNA was extracted according to the method of Kimura *et al.* (2002). To survey variation in AFLP markers, we used a protocol designed by Vos *et al.* (1995), with slight modifications by Bensch *et al.* (2002). We adapted this protocol for the visualization of bands on an ABI 3700 genetic analyzer at the UCLA Core DNA Sequencing Facility; later, the identical protocol was used by Milá *et al.* (2007, 2009), where it is described more clearly. The average per locus methodological error rate, as estimated for this protocol by Milá *et al.* (2007), is 1.8%. We also calculated a rough estimate of the error rate for our study, by running and blindly scoring two samples at two primer combinations multiple times (one-three times, the other twice) on different days and in different sets of other samples. This resulted in an error rate of 2.0%, small compared with the differences between individuals within populations (ranging from 12.7% to 25.3% on average, see Results) and between individuals of different populations (ranging from 15.3% to 33.6% on average).

We generated AFLP signatures for 237 samples at 2–6 primer combinations. After the inspection of the trace files, we rejected 84 of these samples because they generated trace files that were visibly irregular compared with those produced by the bulk of the samples. Irregularities included too many bands, too few bands or unusual distributions of band sizes (e.g. too many short bands, too few long bands). Many of these DNA samples had been extracted from feathers, which often produce less DNA than blood, or had been extracted many years earlier, possibly causing degraded DNA or other problems. The decision of whether to include or reject samples was made before subsequent analysis and without considering individual identity or origin. Of the remaining 153 samples, 108 were successfully profiled using six primer combinations (Table 2), and the rest were successfully profiled at two primer combinations. (Logistical and time constraints prevented those from being profiled at the other four primer combinations.) We are confident that these samples produced reliable AFLP signatures because there were no noticeable outliers, every sample falling into one of two distinct clusters (see Results).

We examined trace files and scored AFLP bands using the program Genographer 1.6 (Benham 2001). We scored all variable AFLP bands, except those that appeared in three or fewer individuals. Variation in the

© 2011 Blackwell Publishing Ltd

Table 1 Sampling locations and numbers of samples from each used in the two amplified fragment length polymorphism analyses (six primer combinations and two primer combinations)

Location	Grouped site	Lat.	Long.	*N* (6combo)	*N* (2combo)
Breeding (west)					
1. Denali, AK	A	63.4	−150.4	—	4
2. Juneau, AK	B	58.3	−134.4	8	8
3. Hinton, AB	C	53.5	−117.5	10	10
4. Ram Falls, AB	C	52.0	−115.8	—	4
5. Benjamin Creek, AB	C	51.5	−115.0	—	2
6. 100 Mile House, BC	D	51.7	−121.3	10	10
7. Cayoosh Mountain, BC	D	50.3	−122.3	—	1
8. Mt. Baker Nat. Forest, WA	E	48.1	−121.5	—	8
9. Siuslaw Nat. Forest, OR	F	44.3	−123.9	—	8
10. Half Moon Bay, CA	G	37.5	−122.5	—	8
11. Rocky Mountain Nat. Park, CO	H	40.4	−105.6	2	5
12. Grand Mesa, CO	H	39.0	−107.9	1	6
Breeding (east)					
13. New Liskeard, ON	J	47.5	−79.7	5	5
14. Camp Myrica, QC	K	49.7	−73.3	13	13
15. Fredericton, NB	L	45.8	−66.7	4	4
Wintering					
16. Concordia, Sinaloa, Mex.	P	23.8	−102.3	3	3
17. El Cielo Biosphere Reserve, Tamaulipas, Mex.	Q	23.0	−99.1	7	7
18. Coatepec, Veracruz, Mex.	R	19.5	−97.0	9	9
19. Tuxtlas, Veracruz, Mex.	R	18.4	−95.2	2	2
20. Oaxaca. Mex.	S	17.1	−96.8	9	9
21. Tegucigalpa, Honduras	T	14.1	−87.2	7	7
22. San Salvador, El Salvador	U	13.7	−89.2	9	9
23. Nicaragua	V	13.0	−85.0	—	2
24. Las Alturas Biological Station, San Vito, Costa Rica	W	8.8	−83.0	9	9
Totals				108	153

Table 2 Primer combinations used in the amplified fragment length polymorphism analysis, the dye used to visualize bands, the number of samples genotyped at each combination and the number of variable bands. For the lower two combinations, the number of variable bands is shown for the 108-sample data set (six primer combinations) followed by the number of variable bands in the 153-sample data set (two primer combinations)

EcoRI primer*	M-primer†	Dye	Samples	Variable bands
TCT	CGA	VIC	108	38
TCT	CGT	VIC	108	54
TGA	CGT	VIC	108	48
TGC	CGA	FAM	108	44
TAG	CTA	FAM	153	42/45
TGC	CTT	FAM	153	31/32

*EcoRI primer: 5′-GACTGCGTACCAATTCNNN-3′.
†MseI primer: 5′-GATGAGTCCTGAGTAANNN-3′.

resulting presence/absence matrix was summarized using principal components analysis (PCA) applied to the covariance matrix, using R (R Development Core Team 2008). We used Structure 2.2 (Pritchard *et al.* 2000; Falush *et al.* 2007) using models of both admixture and no admixture and correlated allele frequencies to estimate the optimal number of genetic clusters and to determine the probabilities that each individual belongs to each cluster. As recommended by Pritchard *et al.* (2010) for initial studies of geographic structure, sampling locations were not used as priors. Each Structure run consisted of 50 000 burn-in steps, followed by 50 000 steps for parameter estimation, with K (the number of clusters) ranging from 1 to 15. In the admixture model, α (the parameter for degree of admixture) was inferred from the data. Three replicate runs were performed for each K.

To describe the amount of genetic differentiation between populations, we used two forms of F_{ST} statistic (Bonin *et al.* 2007; Toews & Irwin 2008). The first estimates the fraction of variance in allele frequencies that is explained by differences between populations. The second is similar but based on AFLP band frequencies rather than allele frequencies. The allele-based F_{ST} has more relevance to population genetic theory and can be compared more directly with F_{ST} calculated from

© 2011 Blackwell Publishing Ltd

codominant markers such as microsatellites, whereas the band-based F_{ST} is more straightforward to estimate, requiring fewer assumptions. Both have been used in previous studies of AFLP variation in birds; hence, we calculated both for purposes of comparison with other studies. Allele-based F_{ST} was estimated using AFLP-SURV (Vekemans 2002), which used a Bayesian approach (with a nonuniform prior distribution of allele frequencies) assuming Hardy–Weinberg equilibrium within each species. We estimated 95% confidence intervals as follows: we used AFLP-SURV to calculate F_{ST} from each of 99 bootstrapped data sets by resampling with replacement among AFLP markers and then took the third smallest value and the third highest value as the 95% confidence interval. In addition to the overall F_{ST} calculated by AFLP-SURV based on all loci, we calculated F_{ST} for each locus, according to the method used by Toews & Irwin (2008). The program Arlequin 3.11 (Excoffier *et al.* 2005) was used to calculate F_{ST} based on band frequencies, for purposes of comparing with other studies that have used the same method (summarized by Irwin *et al.* 2009b).

Arlequin was also used to perform analysis of molecular variance (AMOVA), which calculates the percentage of genetic variation that is explained by the difference between groups of interest, among sampling sites within groups and within sampling sites. Using the 2-primer-combination data set, because it had greater geographic sampling, we performed two AMOVA analyses: the first calculated the percentage of AFLP variation in all breeding samples that was explained by the difference between western (sites A–H, subspecies *chryseola* and *pileolata*; Table 1, Fig. 1) and eastern (sites J–L, subspecies *pusilla*) breeding populations, whereas the second calculated the percentage of variation within the western breeding group that was explained by the difference between the two western subspecies (*chryseola* and *pileolata*). Statistical significance of the variance components was tested using 1023 permutations in Arlequin.

Mitochondrial DNA divergence time

While it has been previously established that western and eastern breeding groups of Wilson's warbler correspond to divergent mitochondrial DNA clades (Kimura *et al.* 2002), no previous study has applied a reliable molecular clock to estimate the time of that split. Kimura *et al.* (2002) presented an analysis of variation in the control region, but the rate of evolution in that region is not well established and is highly variable. In contrast, cytochrome *b* has been intensively studied and has a fairly reliable molecular clock of 2% sequence divergence per million years, at least in birds (Weir & Schluter 2008). Kimura *et al.* (2002) identified 33 unique cytochrome *b* haplotypes within Wilson's warblers and used variation within them to design a PCR–RFLP assay to quickly assign samples to divergent western or eastern clades; they did not present a phylogenetic analysis of these sequences. We now present the first phylogenetic analysis of cytochrome *b* variation within Wilson's warblers, comparing them with related species.

We downloaded the 33 sequences, each of 316 bp, contributed by Kimura *et al.* (2002), from Wilson's warblers (GenBank accession numbers AF499563–AF499595). We used Arlequin 3.11 (Excoffier *et al.* 2005) to calculate average pairwise differences between and within western and eastern groups and to calculate F_{ST} between these haplotype clusters. Note that this F_{ST} cannot be taken as a precise population-based F_{ST} because it is not based on actual haplotype frequencies. (Each haplotype was entered once into the analysis, because the frequency data from Kimura *et al.* 2002 are presently unavailable; M. Kimura, pers. comm.) To compare variation in Wilson's warblers with molecular distance between other related and distinct species, we downloaded sequences for the four species that are apparently most closely related to Wilson's warblers (Lovette *et al.* 2010): *Ergaticus ruber*, *Ergaticus versicolor*, *Wilsonia canadensis* (GU932413–GU932415; Lovette *et al.* 2010) and *Cardellina rubifrons* (AF383026; Lovette & Bermingham 2002). We then estimated a phylogeny using the program Tree-Puzzle (Schmidt *et al.* 2002), using a maximum-likelihood approach. We specified exact parameter estimation using a neighbour-joining tree and an HKY + Γ model of sequence evolution (Price 2008), with eight gamma-distributed rate categories. A likelihood ratio test using Tree-Puzzle revealed that the assumption of a molecular clock did not significantly decrease the likelihood of the phylogeny. We therefore use the clock-like tree for illustration, assuming 2% sequence divergence per million years (Weir & Schluter 2008). This rate of divergence is an expectation given much time and a large sequence, whereas in any particular short span of time and short sequence, there will be stochastic variation in the number of mutations that actually occur. To estimate a 95% confidence interval on the estimated coalescence time owing to this source of variation, we determined the lower 2.5% and highest 2.5% values of a Poisson distribution with a mean of 12.6, which is the average basepair difference between western and eastern haplotypes (see Results).

Results

A total of 257 variable AFLP markers were identified using six primer combinations (Table 2) applied to a sample of 108 Wilson's warblers (Table 1). An additional 45 samples were successfully profiled at two of

© 2011 Blackwell Publishing Ltd

Table 3 Pairwise genetic differentiation between major breeding sites of Wilson's warblers, estimated using 77 amplified fragment length polymorphism (AFLP) markers. Sites are indicated by letters according to Fig. 1 and Table 1, with A–H being western sites (*chryseola* and *pileolata*) and J–K being eastern (*pusilla*). On the diagonal (shaded) are average AFLP distances within sites (i.e. the number of bands that differ between individuals in presence/absence). Above the diagonal are average AFLP distances between sites. Below the diagonal are band-based F_{ST} values between sites, calculated using Arlequin (Excoffier *et al*. 2005). Italicized text indicates F_{ST} values that are significantly different from zero at the $P < 0.05$ level, and bold italics indicates significance at the $P < 0.005$ level

	A	B	C	D	E	F	G	H	J	K	L
A	19.0	17.9	18.8	17.3	17.0	20.4	18.9	18.1	23.8	25.2	23.0
B	***0.18***	11.9	12.6	12.2	13.1	16.3	15.3	13.4	18.9	20.2	18.4
C	***0.20***	0.01	13.0	12.9	13.7	16.9	15.3	13.8	18.5	19.5	18.0
D	***0.12***	−0.01	0.01	12.7	12.8	16.7	15.4	13.0	19.4	20.7	18.8
E	*0.07*	0.04	*0.04*	−0.01	13.4	16.6	15.6	13.8	19.4	20.5	18.6
F	*0.08*	*0.06*	***0.08***	*0.07*	0.03	18.7	17.8	16.9	21.9	22.4	20.4
G	*0.11*	***0.12***	***0.09***	***0.10***	*0.09*	0.05	15.1	15.0	22.7	22.9	21.8
H	***0.16***	***0.07***	***0.06***	0.02	*0.05*	***0.08***	*0.08*	12.8	19.9	20.8	18.8
J	*0.38*	***0.39***	***0.33***	***0.38***	***0.36***	***0.30***	***0.41***	***0.39***	11.2	12.8	11.5
K	***0.40***	***0.37***	***0.32***	***0.37***	***0.35***	***0.30***	***0.39***	***0.37***	0.03	13.3	11.6
L	*0.38*	***0.40***	***0.34***	***0.38***	***0.36***	***0.26***	***0.40***	***0.38***	0.09	−0.01	9.5

these primer combinations accounting for 77 variable AFLP markers (Tables 1 and 2). Among these 77 markers, individuals within breeding sites differed on average by 9.5–19.0 markers (12.3–24.7%), and individuals from different breeding sites differed on average by 11.5–25.2 markers (14.9–32.7%), providing abundant genetic variation for the study of population differentiation (Table 3).

A principal component analysis on the 257-marker data set reveals that the major axis of variation (PC1) separates two highly distinct genetic clusters (Fig. 2), explaining a large amount (17.0%) of the variation while subsequent PCs explained much less (e.g. PC2: 3.1%). The clusters separated along PC1 correspond perfectly to (i) all western breeding samples and all wintering samples (the left cluster in Fig. 2) and (ii) all eastern breeding samples (the right cluster). Analysis of the 77-marker data set produces similar results (Fig. S1, Supporting information), the notable difference being that there is a bit more scatter in the clusters compared with the distance between the clusters, presumably due to the lower signal-to-noise ratio owing to less information (77 vs. 257 markers).

While there is clear separation between western and eastern clusters, principal component analysis reveals relatively weak geographic variation within each cluster. Among western breeding individuals (Fig. 3), there is some differentiation between breeding sites, but a fair amount of overlap as well (see below for F_{ST} analyses and tests of differentiation). For example, individuals from the eastern slopes of the Rocky Mountains in Alberta (site C in Fig. 3) tend to fall on the lower left side of the graph, individuals from Denali, Alaska (site A) in the upper right and individuals from California (site G) in the lower right. Among our three eastern breeding sites, there is little if any noticeable differentiation revealed by principal component analysis (not shown).

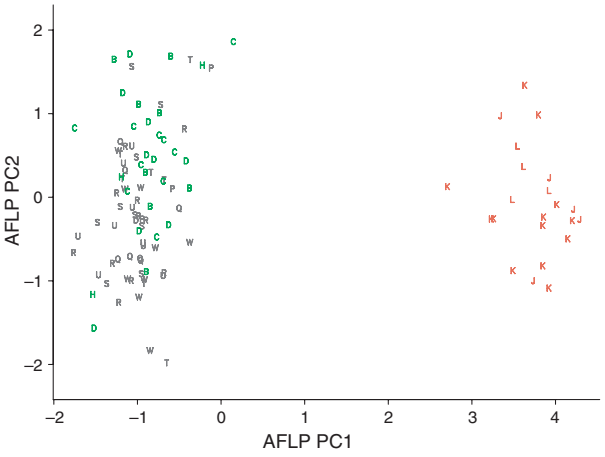

Fig. 2 Amplified fragment length polymorphism fingerprinting using 257 variable markers from six primer combinations (Table 2) followed by principal component analysis reveals two highly distinct genetic clusters of Wilson's warblers. Each letter (corresponding to grouped sampling sites; Table 1, Fig. 1) represents a single individual (108 total, sampled across the breeding and wintering ranges), plotted according to its position on the first two principal components. Colours indicate sampling region (see Fig. 1). The two major clusters correspond perfectly to western (sites A–H, green) and eastern (sites J–L, red) breeding groups, with all winter samples in the study (sites P–W, grey) grouping with the western breeding cluster.

The above patterns are reflected in our analyses of population differentiation using *F*-statistics. In the

© 2011 Blackwell Publishing Ltd

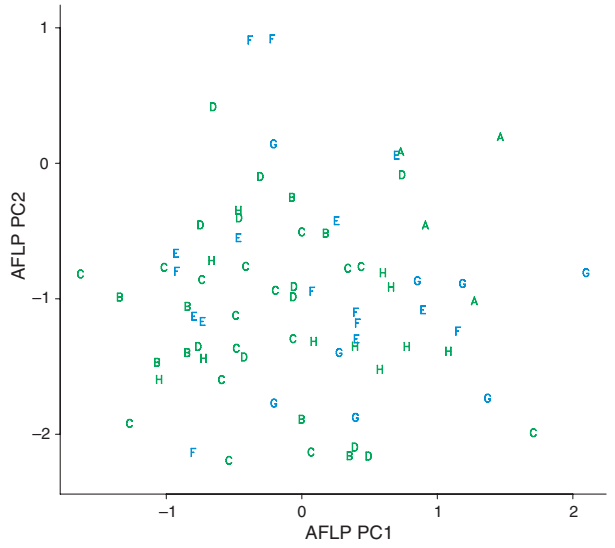

Fig. 3 A principal component analysis performed only on those breeding samples in the western cluster using the 77-marker data set (also see Fig. S1, Supporting information) suggests a small but detectable amount of geographic variation. PC1 explains 8.2% of the variation, and PC2 explains 6.5%. To maximize the number of samples shown, the 77-marker data set was used for this figure; the 257-marker data set shows a similarly weak level of differentiation between western breeding sites.

breeding range, band-based F_{ST} is quite high between western (sites A–H) and eastern (sites J–L) regions, ranging from 0.26 and 0.41 in pairwise comparisons (Table 3). When sites within each region (western and eastern) are lumped, band-based F_{ST} between west and east is 0.33 (significantly different from zero at the $P < 0.00001$ level) for the 77-marker data set and 0.34 ($P < 0.00001$) for the 257-marker data set. Allele-based F_{ST} between west and east breeding regions is 0.23 (95% C.I.: 0.12–0.30) for the 77-marker data set and 0.23 (0.18–0.29) for the 257-marker data set. Consistent F_{ST} values for each data set, which differ markedly in numbers of markers and samples, suggest that the AFLP analysis produces highly reliable results. While the signal of divergence between west and east is strong, the signal comes from a relatively small number of markers: the distribution of F_{ST} values among markers is highly skewed, with only 29 markers having $F_{ST} > 0.2$ and only five markers having $F_{ST} > 0.6$ (Fig. S2, Supporting information). In contrast to these high overall values of F_{ST} between western and eastern breeding regions, F_{ST} values are generally much lower between pairs of sites within each region (Table 3). However, among western breeding sites, F_{ST} values are mostly positive (26 of 28 cases) and many are highly significant, indicating that there is genetic differentiation between many western breeding areas. In particular, our sample of four individuals from Denali (site A) differed fairly strongly from other sites (band-based F_{ST} ranging from 0.07 to 0.20, $P < 0.05$ in all cases; Table 3).

Analysis of molecular variance (AMOVA) allows the quantification of the amount of variation explained by differences between previously designated subspecies. The difference between the eastern subspecies (*pusilla*) and the western pair of subspecies (*chryseola* and *pileolata*) explains 31.6% of the AFLP variation, whereas differences between sampling sites within those regions explain only 4.0%, leaving 64.4% of the variation explained by differences between individuals within sites. In stark contrast, when only the sites within the west are analysed, only 1.2% of the variation is partitioned between *chryseola* and *pileolata*, with 5.8% explaining differences between sites within those subspecies and 93.0% being within-site variation. The difference between *pusilla* and the two western subspecies is highly statistically significant ($P < 0.005$), whereas the difference between *chryseola* and *pileolata* is not ($P = 0.14$).

In contrast to the breeding range, our sample of warblers across the wintering range shows virtually no geographic structure in AFLP signatures (sites P–W; Table 4). Band-based F_{ST} between pairs of wintering populations are mostly small, with only 2 of 28 being significantly different than zero at the $P < 0.05$ level. (Hence, these could be false-positives.) Each of the wintering sites is strongly divergent from the eastern breeding region (sites J–L, lumped as 'East' in Table 4), with band-based F_{ST} ranging from 0.33 to 0.43, and highly similar to the western breeding region (sites A–H, lumped as 'West'), with estimates of band-based F_{ST} ranging from −0.02 to 0.05. These results are consistent with the principal component analysis in showing that each of our wintering individuals has an AFLP profile that is similar to those of the western breeding cluster.

Analysis using the program Structure (Pritchard *et al.* 2000) also showed that samples in our study fit into two distinct genetic clusters. Running Structure with all samples (both breeding and wintering) in the 257-marker data set under the no-admixture model showed that the likelihood of the data [i.e. Ln P(D)], given a number of populations, K, rises dramatically from $K = 1$ to $K = 2$, where it reaches a plateau before declining at higher values of K (Fig. S3a). This suggests that $K = 2$ is the best description of the number of genetic clusters (Evanno *et al.* 2005). While the likelihood is sometimes higher when $K > 2$, the difference is small compared with the likelihood when $K = 2$, and the individuals are still assigned with high probability to only two clusters. Under this model ($K = 2$), each of the breeding individuals from western areas (sites A–H) and each of the wintering individuals were assigned

© 2011 Blackwell Publishing Ltd

Table 4 Pairwise genetic differentiation between major wintering sites of Wilson's warblers, estimated using 77 amplified fragment length polymorphism (AFLP) markers, along with comparisons to western (sites A–H) and eastern (sites J–L) breeding clusters. Sites are indicated by letters according to Fig. 1 and Table 1. On the diagonal (shaded) are average AFLP distances within sites. Above the diagonal are average AFLP distances between sites. Below the diagonal are band-based F_{ST} values between sites, calculated using Arlequin (Excoffier *et al.* 2005). Italicized text indicates F_{ST} values that are significantly different from zero at the $P < 0.05$ level, and bold italics indicates significance at the $P < 0.005$ level

	P	Q	R	S	T	U	V	W	West	East
P	14.0	12.5	12.1	12.1	12.4	12.1	14.3	12.0	14.2	20.3
Q	0.02	11.1	11.1	11.5	11.9	11.1	13.9	10.5	13.3	19.7
R	−0.02	−0.02	11.3	11.5	12.0	11.0	13.1	10.8	13.5	20.3
S	0.03	0.06	*0.06*	10.4	11.3	10.7	12.3	11.0	13.2	20.2
T	−0.05	0.01	0.01	−0.01	12.5	11.5	13.9	11.5	13.6	19.8
U	0.01	0.01	−0.01	0.01	−0.01	10.8	12.9	10.3	13.3	20.6
V	0.08	0.18	0.12	0.12	0.11	0.14	12.0	13.1	14.8	19.0
W	0.03	−0.02	−0.01	0.05	0.01	−0.03	*0.18*	10.3	13.2	20.2
West	−0.02	0.01	*0.02*	*0.03*	−0.01	0.02	0.05	*0.03*	14.8	20.5
East	***0.37***	***0.39***	***0.41***	***0.42***	***0.37***	***0.43***	***0.35***	***0.42***	***0.33***	12.4

with 100% confidence to one cluster, and each of the breeding individuals from eastern areas (sites J–L) was assigned with 100% confidence to the other cluster. Similar results were obtained with the admixture model: when $K = 2$, each individual was assigned strongly to one of the two clusters (Fig. 4; Fig. S3b, Supporting information). When $K = 3$, no individual had a strong assignment probability to the third cluster, and all winter individuals were still assigned to the same cluster as all western breeding individuals; hence, we interpret $K = 2$ as being the most accurate description of the number of clusters.

Given the strong differentiation between west and east and given the assignment of all winter samples to the western breeding group, we again ran Structure without the eastern breeding samples. In both the no-admixture (Fig. S4a, Supporting information) and admixture (Fig. S4b, Supporting information) models, the likelihood is highest and approximately equal for $K = 1$ and $K = 2$. When $K > 1$, all individuals are assigned either entirely to one cluster (in the no-admixture model) or in approximately equal probabilities to the K clusters (in the admixture model). This showed that the optimal K is one; all wintering samples and western breeding samples form a single genetic cluster.

The strong differences in nuclear DNA (i.e. AFLPs) between western and eastern breeding groups are reflected in our mitochondrial DNA phylogeny based on 316 bp of the cytochrome *b* gene (Fig. 5). Average pairwise difference between western and eastern haplotypes is 12.6 bp (4.0%), much larger than average pairwise differences in the western group (3.4 bp; 1.1%) and in the eastern group (2.1 bp; 0.7%). F_{ST} between western and eastern clades is 0.76, a very high value that reflects the large distance between groups compared to the variation within.

The maximum-likelihood phylogeny that includes these samples as well as four outgroup species (Fig. 5) allows us to estimate the timing of the split between western and eastern mitochondrial clades. Applying the well-accepted molecular clock for cytochrome *b* of 2% sequence divergence per million years (Weir & Schluter

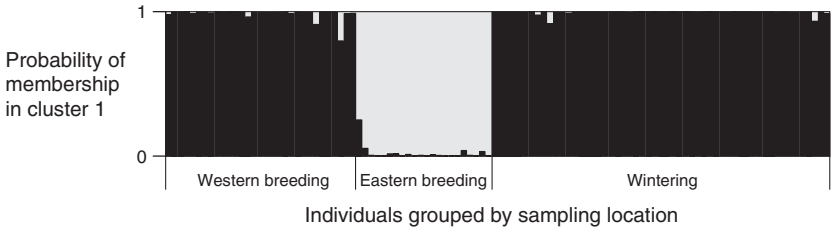

Fig. 4 Analysis using Structure (Pritchard *et al.* 2000) shows that each individual Wilson's warbler can be assigned with high confidence to one of two genetic clusters, corresponding to western ($n = 31$) and eastern ($n = 22$) breeding regions. Each column represents a single individual, with black and grey representing the probabilities of membership in two genetic clusters. All wintering individuals ($n = 55$) are clearly assigned to the western breeding region. This figure shows results based on the admixture model, the 257-marker data set and $K = 2$.

© 2011 Blackwell Publishing Ltd

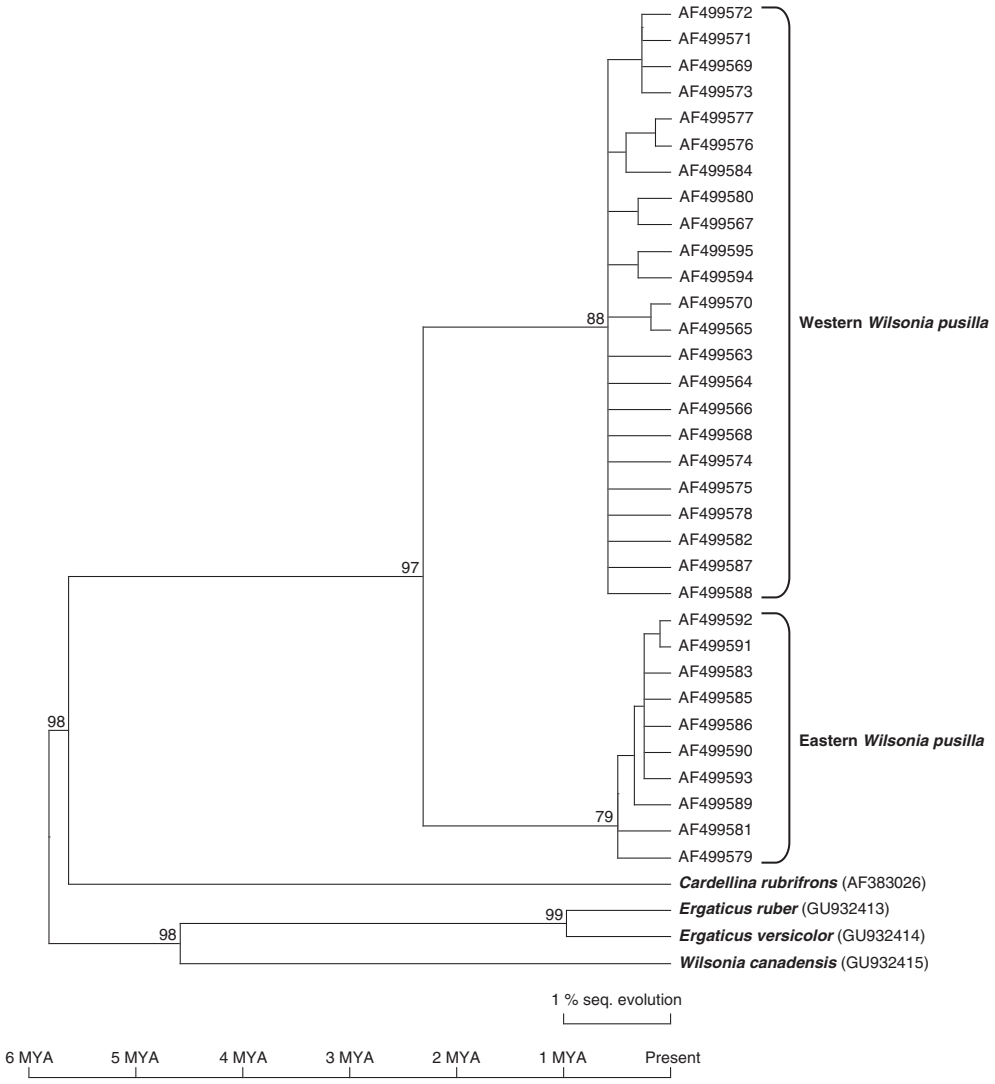

Fig. 5 Mitochondrial DNA phylogeny showing highly divergent western and eastern groups of Wilson's warbler (*Wilsonia pusilla*) as well as their relationships to the four species most related to them (Lovette *et al.* 2010). The phylogeny is based on 316 bp of cytochrome *b* sequence, using sequences downloaded from GenBank (accession numbers are shown) and the program Tree-Puzzle (Schmidt *et al.* 2002). Wilson's warbler sequences were generated by Kimura *et al.* (2002); only unique haplotypes are shown. The scale bar shows the branch length equivalent to 1% sequence evolution within a lineage, as well as estimated coalescence times in millions of years before present, based on a rate of 2% sequence divergence per million years (Weir & Schluter 2008). Numbers at major nodes are quartet puzzling support values (Schmidt *et al.* 2002).

2008), we estimate that the two mitochondrial clades split approximately 2.3 Ma (Fig. 5), with a 95% confidence interval of 1.1–3.7 Ma (based on stochastic variation in the number of mutations that are expected to occur given a 2% molecular clock). The point estimate of the coalescence time is almost half the span of time back to the splits between the Wilson's warbler lineage and lineages leading to the four most closely related species, about 5.6 and 5.8 Ma, and more than twice as deep as the split between the two species *Ergaticus ruber* and *Ergaticus versicolor*, which is dated about 1.0 Ma (Fig. 5).

Discussion

Our survey of variation in a large number (257) of nuclear genetic markers in Wilson's warblers revealed a remarkable amount of differentiation between western (subspecies *chryseola* and *pileolata*) and eastern (*pusilla*) breeding groups and comparatively very little variation within each. There is no indication of gene flow between these clusters, and no individuals were found with intermediate genetic signatures. While our sampling in the breeding season had a large gap that includes the area of putative contact between western

© 2011 Blackwell Publishing Ltd

and eastern forms, our winter sampling to some extent mitigates this gap in sampling, as the wintering birds must have come from many breeding areas that we did not sample directly. The lack of any intermediate genetic signatures in the winter suggests that genetic intermediates may not actually exist in any sizeable numbers.

These patterns raise an intriguing question: Might western and eastern Wilson's warblers in fact be reproductively isolated cryptic species? To our knowledge, this possibility has not been suggested previously, although strong differentiation has been seen in both mitochondrial DNA (Kimura *et al.* 2002) and microsatellites (Clegg *et al.* 2003), presumably because of the high phenotypic similarity of the western and eastern forms. The level of differentiation in AFLP markers is at a level that is often seen between species; in other studies of AFLP differentiation between taxa at or near the species level, band-based F_{ST} has ranged from 0.14 to 0.42 between taxa generally considered separate species (Irwin *et al.* 2009b); the observed F_{ST} of 0.34 in Wilson's warblers is towards the upper end of this range. Mitochondrial differentiation is also very large between western and eastern clades compared with the variation within each, and the estimated date of the split between these groups (2.3 Ma) is more typical of separate species than of subspecies within species (Weir & Schluter 2004; Lovette 2005; Price 2008). Even with the uncertainty in the exact timing of the split (95% CI: 1.1–3.7 Ma), the reciprocal monophyly and high sequence divergence support our conclusion from AFLP divergence that the western and eastern populations of Wilson's warblers have experienced substantial genetic isolation over a long period of time. While species designations should not be based solely on such an analysis of genetic patterns, the observed level of differentiation, without any evidence for intermediates, raises the likelihood that the two forms are reproductively isolated.

A full test of this possibility would benefit from intensive research in a zone of contact (e.g. Toews & Irwin 2008) between western and eastern forms as well as analysis of variation in other traits such as vocalizations and response to vocalizations. One of the authors (DI) has spent much time attempting to find a zone of contact in west-central Alberta (between Hinton and Lesser Slave Lake), without success; there appears to be a gap in distribution between the western and eastern forms that spans much of central Alberta. It is possible, however, that the two forms come into contact in northern Alberta and the Northwest Territories; future attempts to find contact zones should examine those areas. Based on the present results, we foresee four possible alternatives: (i) no breeding range contact between western and eastern populations, (ii) contact between distinct species that are fully reproductively isolated (as recently documented in the same general region for Pacific and winter wrens; Toews & Irwin 2008), (iii) a narrow hybrid zone between western and eastern forms (as recently documented in northeastern British Columbia between MacGillivray's and mourning warblers and between Townsend's and black-throated green warblers; Irwin *et al.* 2009a; Toews *et al.*, in press) or (iv) a more complex pattern in which there are a series of transitional zones between western and eastern forms, with clines in different traits and genetic markers being discordant in their locations and widths. Given the high amount of genetic divergence between the western and eastern groups that we have studied, we feel the latter alternative is less likely than the first three.

As a tool for elucidating patterns of migratory connectivity, our AFLP analysis was partially successful. The analysis revealed strong differences between western and eastern breeding regions, and these can be used to assign wintering samples to one of those regions with very high confidence. The fact that each of the wintering samples was assigned clearly to a breeding cluster shows that the AFLP method in principle can work well for studying migratory connectivity. However, in this case, Wilson's warblers show little geographic differentiation within each of the western and eastern breeding regions. While different western sites are genetically divergent, as shown by the F_{ST} analysis, they are not different enough to allow assignment of wintering birds to areas within the western region using the markers that we surveyed. This is also true for the eastern region, although our sampling there was confined to only three sites, limiting the scope of our conclusions. While our AFLP survey was insufficient for revealing detailed patterns of connectivity within each of the western and eastern regions, the significant and often fairly high values of F_{ST} between various western sites suggests that a larger AFLP analysis using many more primer combinations could reveal a small number of markers that differ strongly between western breeding regions. Such an approach was used successfully with willow warblers (*Phylloscopus trochilus*), revealing two AFLP markers that were useful for studying migratory connectivity (Bensch *et al.* 2002, 2009).

Despite the above limitations of our ability to assign birds to specific areas within west and east, our survey of wintering individuals revealed some interesting patterns in terms of migratory connectivity. First, all of our 57 wintering samples were assigned clearly to the western breeding region. We can conclude from this that the large majority of wintering Wilson's warblers at the winter sites included in this study are from the western breeding group. But this raises an important question:

© 2011 Blackwell Publishing Ltd

why is it so difficult to find wintering birds from the eastern breeding group? We designed our winter sampling as a broad survey of wintering areas for Wilson's warblers in general (not targeting the western group); hence, it was a surprise to not find any eastern birds among our winter samples. We suggest several possible reasons: First, perhaps the most likely explanation is that there may be some geographic localization of the eastern group in the winter, and our winter sampling could have unintentionally favoured the western group. While our sampling was distributed widely over the winter range, it did not cover some regions that may be used by wintering birds from the eastern breeding region. In particular, the Yucatan peninsula and Panama were not sampled in this study; using mtDNA, Kimura *et al.* (2002) identified approximately 10 wintering eastern birds (and approximately 176 wintering western birds), and these were from sites near these areas (specifically, from southeast Veracruz, Oaxaca, Chiapas, Belize and Costa Rica). It is also possible that eastern Wilson's warblers differ behaviourally or ecologically in ways that make them less likely to be sampled compared with western Wilson's warblers. But this seems unlikely as a sole explanation given that our sampling was from a broad range of habitats and used methods (e.g. mist netting) that are generally considered good at sampling Wilson's warblers. Such ecological or behavioural differences in the winter would provide further evidence for the distinctiveness of the western and eastern forms. A final possibility as to why eastern breeders were not sampled in the winter is that they might be very low in numbers compared with western breeders. This is consistent with a breeding bird survey (BBS) map for Wilson's warblers, which shows a very low density across their range from central Alberta to Nova Scotia, whereas densities are higher across most of the western range (Sauer *et al.* 2008). A problem with the BBS data is that much of the breeding range extends far north of the northern boundary of BBS data, making population size very difficult to estimate accurately. Our results point to the possibility of more accurately monitoring relative population size of western and eastern Wilson's warblers by repeated sampling across the wintering range (rather than the breeding range, which is large and difficult to survey broadly) and then using genetic methods to assign samples to west or east.

The second pattern with respect to migratory connectivity revealed by our study is that there is evidence for greater geographic genetic structure in the breeding region than in the wintering region. Support for this conclusion comes mainly from the F_{ST} analysis: generally, genetic differentiation as measured by F_{ST} is usually moderate and statistically significant between western breeding sites, whereas it is lower and not statistically significant between winter sites (where all samples were from the western group). This suggests that winter sites contain a mix of individuals from different western breeding areas.

This study shows that AFLP can be a valuable tool in revealing patterns of genetic differentiation and migratory connectivity. By detecting patterns of covariation in a large number of markers spread throughout the genome, AFLP can show both large differences between populations (e.g. between west and east) and subtle patterns of differentiation that might not be detectable from examinations of a small number of neutral loci (e.g. within the west). While the differentiation seen among western samples is insufficient to allow the use of these markers in assigning individual migrants to specific western breeding regions, the analysis nonetheless uncovered some important patterns, including the very strong differentiation between west and east, the widespread wintering range of western birds, the rarity and apparent geographic confinement of eastern birds in the winter and the apparent mixing of wintering birds from different western breeding regions. While similar patterns were also seen in analyses of mitochondrial DNA (Kimura *et al.* 2002) and microsatellites (Clegg *et al.* 2003), the wide genomic coverage of AFLP gives higher confidence that the patterns are accurately representative of genomic relationships rather than peculiar patterns of one or a few loci (Irwin 2002; Irwin *et al.* 2005; Brelsford *et al.* 2011). These findings suggest that AFLP is a useful tool to apply to studies of migratory connectivity in other species. Furthermore, patterns of subtle variation revealed by an AFLP survey such as that done here can be further investigated by examining a much larger number of markers, with the goal of finding those few parts of the genome that differ strongly between populations, presumably due to selection (Bensch *et al.* 2002, 2009).

In conclusion, our broad survey of breeding and wintering samples shows that what is currently known as the Wilson's warbler in fact consists of two highly divergent genetic groups with no known intermediates. Of the three previously named subspecies, one (*pusilla*) is highly divergent genetically, whereas the other two (*chryseola* and *pileolata*) are extremely similar genetically compared to the amount of variation within each; in fact, the largest between-site differences within the western group are between sites within *pileolata* (specifically, between Denali, AK and more southern sites). These results are a good demonstration that named subspecies can sometimes correspond to highly distinct genetic groups and other times can be highly genetically similar. High levels of divergence between western and eastern breeding groups have been observed

© 2011 Blackwell Publishing Ltd

in a survey of a large number of nuclear markers (257 AFLP markers in this study and eight microsatellite loci in Clegg *et al.* 2003) and in mtDNA (Kimura *et al.* 2002). These two genetic groups appear to have differing winter distributions, although our information on the eastern group is sparse owing to its apparently low population size and to its apparently more limited winter distribution compared with the western group. The patterns also strongly suggest that these groups have been reproductively isolated for a long period of time (the cytochrome *b* divergence time is estimated at 2.3 Ma) and therefore may be best described as separate species, although some researchers may feel that more research in potential areas of contact (e.g. Toews & Irwin 2008) is needed before making a taxonomic change. Whether they are designated as distinct species or simply as highly divergent subspecies, it is clear that both western and eastern forms should be treated as distinct conservation units and as divergent evolutionary and ecological entities. Given their visual similarity, these two forms of Wilson's warblers thus provide an especially striking case of cryptic biological diversity and remind us that many more such cases probably await discovery.

Acknowledgements

For contributing genetic samples, we thank D. DeSante and The Institute for Bird Populations, the Monitoring Avian Productivity and Survivorship (MAPS) Program banders, J. Booker, R. Brito, R. Colwell, T. Doyle, P. Marra, J. Steele; the Colorado Bird Observatory; D. Dittmann and F. Sheldon at the Museum of Natural Science, Louisiana State University; the Point Reyes Bird Observatory; and T. Gaston. For sharing their study sites and/or facilitating fieldwork, we thank Y. Aubry, P. Bichier, R. Carlson, A. Cruz, O. Cruz, R. Dickson, O. Figueroa, T. Gavin, K. Holl, L. Imbeau, O. Komar, B. Murphy, E. Rodriguez, E. Ruelas, E. Santana, W. Schaldach, C. Spytz, P. Thorn and the Belize Audubon Society. For research permits, we thank Environment Canada's Canadian Wildlife Service, Provinces of Quebec, New Brunswick, Ontario, Alberta and British Columbia; US Fish & Wildlife Service; National Park Service; F. Villaseñor for help obtaining permits in Mexico; Instituto Nacional de Ecologia, Secretaria de Medio Ambiente y Recursos Naturales in Mexico; Forestry Department of the Ministry of Natural Resources in Belize; Organization for Tropical Studies and Ministerio del Ambiente y Energia in Costa Rica; and the Government of El Salvador. For collecting samples and extracting DNA, we especially thank M. Kimura and B. Milá. For assistance in the field and/or laboratory, we also thank S. Clegg, A. Alvarado, T. Erwin, H. Gates, J. Gearlds, L. Gentle, J. Larson, M. Macabeo-Ong, L. Nguyen, K. Ruegg, C. Saux and B. Wang. Helpful comments on the manuscript were provided by S. Clegg, J. Irwin, M. Kimura, J. Lee-Yaw, B. Milá, S. Seneviratne and D. Toews, in addition to Dany Garant and two anonymous reviewers. This work was supported by grants from San Francisco State University, NIH Office of Research on Minority Health (grant #5P20RR11805), The Turner Foundation, the Environmental Protection Agency (R827109-01-0) and National Science Foundation (DEB-9726425 and IRCEB-9977072) to TBS, and Discovery Grants (311931-2005 and 311931-2010) to DEI from the Natural Sciences and Engineering Research Council of Canada during the analysis and writing phase.

References

Ammon EM, Gilbert WM (1999) Wilson's Warbler (*Wilsonia pusilla)*. In: *The Birds of North America Online* (ed. Poole A). Cornell Lab of Ornithology, Ithaca, New York. Retrieved from the Birds of North America Online: http://bna.birds.cornell.edu/bna/species/478. Accessed on 12th November 2010.

Benham J (2001) Genographer. Available from: http://hordeum.msu.montana.edu/genographer/. Accessed on January 2009.

Bensch S, Åkesson M (2005) Ten years of AFLP in ecology and evolution: why so few animals? *Molecular Ecology*, **14**, 2899–2914.

Bensch S, Andersson T, Åkesson S (1999) Morphological and molecular variation across a migratory divide in willow warblers, *Phylloscopus trochilus*. *Evolution*, **53**, 1925–1935.

Bensch S, Åkesson S, Irwin DE (2002) The use of AFLP to find an informative SNP: genetic differences across a migratory divide in willow warblers. *Molecular Ecology*, **11**, 2359–2366.

Bensch S, Grahn M, Müller N, Gay L, Åkesson S (2009) Genetic, morphological, and feather isotope variation of migratory willow warblers show gradual divergence in a ring. *Molecular Ecology*, **18**, 3087–3096.

Bonin A, Ehrich D, Manel S (2007) Statistical analysis of amplified fragment length polymorphism data: a toolbox for molecular ecologists and evolutionists. *Molecular Ecology*, **16**, 3737–3758.

Boulet M, Norris DR (2006) The past and present of migratory connectivity. *Ornithological Monographs*, **61**, 1–13.

Brelsford A, Milá B, Irwin DE (2011) Hybrid origin of Audubon's warbler. *Molecular Ecology*, **20**, 2380–2389.

Campbell D, Bernatchez L (2004) Generic scan using AFLP markers as a means to assess the role of directional selection in the divergence of sympatric whitefish ecotypes. *Molecular Biology and Evolution*, **21**, 945–956.

Carlisle JD, Skagen SK, Kus BE, van Riper III C, Paxton KL, Kelly JF (2009) Landbird migration in the American West: recent progress and future research directions. *Condor*, **111**, 211–225.

Clegg S, Kelley JF, Kimura M, Smith TB (2003) Combining genetic markers and stable isotopes to reveal population connectivity and migration patterns in a Neotropical migrant, Wilson's warbler (*Wilsonia pusilla*). *Molecular Ecology*, **12**, 819–830.

Dunn J, Garrett K (1997) *Warblers*. Houghton Mifflin Company, New York.

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 ver. 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47–50.

© 2011 Blackwell Publishing Ltd

Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes*, **7**, 574–578.

Freedman AH, Thomassen HA, Buermann W, Smith TB (2010) Genomic signals of diversification along ecological gradients in a tropical lizard. *Molecular Ecology*, **19**, 3773–3788.

Helbig AJ (1991) Inheritance of migratory direction in a bird species: a cross-breeding experiment with SE- and SW-migrating blackcaps (*Sylvia atricapilla*). *Behavioral Ecology and Sociobiology*, **28**, 9–12.

Irwin DE (2002) Phylogeographic breaks without geographic barriers to gene flow. *Evolution*, **56**, 2383–2394.

Irwin DE (2009) Incipient ring speciation revealed by a migratory divide. *Molecular Ecology*, **18**, 2923–2925.

Irwin DE, Irwin JH (2005) Siberian migratory divides: the role of seasonal migration in speciation. In:*Birds of Two Worlds: The Ecology and Evolution of Migration* (eds Greenberg R, Marra PP). pp. 27–40, Johns Hopkins University Press, Baltimore, Maryland.

Irwin DE, Bensch S, Irwin JH, Price TD (2005) Speciation by distance in a ring species. *Science*, **307**, 414–416.

Irwin DE, Brelsford A, Toews DPL, MacDonald C, Phinney M (2009a) Extensive hybridization in a contact zone between MacGillivray's and mourning warblers (*Oporornis tolmiei* and *O. philadelphia*) detected using molecular and morphometric analyses. *Journal of Avian Biology*, **40**, 539–552.

Irwin DE, Rubtsov AS, Panov EN (2009b) Mitochondrial introgression and replacement between yellowhammers (*Emberiza citrinella*) and pine buntings (*E. leucocephalos*; Aves, Passeriformes). *Biological Journal of the Linnean Society*, **98**, 422–438.

Kelly JF, Atudorei V, Sharp ZD, Finch DM (2002) Insights into Wilson's warbler migration from analyses of hydrogen stable-isotope ratios. *Oecologia*, **130**, 216–221.

Kelly JF, Ruegg KC, Smith TB (2005) Combining isotopic and genetic markers to identify breeding origins of migrant birds. *Ecological Applications*, **15**, 1487–1494.

Kimura M, Clegg SM, Lovette IJ *et al.* (2002) Phylogeographical approaches to assessing demographic connectivity between breeding and overwintering regions in a Nearctic-Neotropical warbler (*Wilsonia pusilla*). *Molecular Ecology*, **11**, 1605–1616.

Lovette IJ (2005) Glacial cycles and the tempo of avian speciation. *Trends in Ecology and Evolution*, **20**, 57–59.

Lovette IJ, Bermingham E (2002) What is a wood-warbler? Molecular characterization of a monophyletic Parulidae. *Auk*, **119**, 695–714.

Lovette IJ, Pérez-Emán JL, Sullivan JP *et al.* (2010) A comprehensive multilocus phylogeny for the wood-warblers and a revised classification of the Parulidae (Aves). *Molecular Phylogenetics and Evolution*, **57**, 753–770.

Milá B, McCormack JE, Castañeda G, Wayne RK, Smith TB (2007) Recent postglacial range expansion drives the rapid diversification of a songbird lineage in the genus *Junco*. *Proceedings of the Royal Society of London B*, **274**, 2653–2660.

Milá B, Wayne RK, Fitze P, Smith TB (2009) Divergence with gene flow and fine-scale phylogeographical structure in the wedge-billed woodcreeper, *Glyphorynchus spirurus*, a Neotropical rainforest bird. *Molecular Ecology*, **18**, 2979–2995.

Mueller UG, Wolfenbarger LL (1999) AFLP genotyping and fingerprinting. *Trends in Ecology and Evolution*, **14**, 389–394.

Paxton KL, Van Riper III C, Theimer TC, Paxton EH (2007) Spatial and temporal migration patterns of Wilson' Warbler (*Wilsonia pusilla*) in the southwest as revealed by stable isotopes. *Auk*, **124**, 162–175.

Price T (2008) *Speciation in Birds*. Roberts and Company, Greenwood Village, Colorado.

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

Pritchard JK, Wen X, Falush D (2010) Documentation for structure software: version 2.3. Available from http://pritch.bsd.uchicago.edu/structure.html. Accessed on 12th November 2010.

R Development Core Team (2008) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. Available from http://www.R-project.org. Accessed on 12th November 2010.

Rohwer S, Irwin DE (2011) Molt, orientation, and avian speciation. *Auk*, **128**, 419–425.

Ruegg K (2007) Genetic, morphological, and ecological characterization of a hybrid zone that spans a migratory divide. *Evolution*, **62**, 452–466.

Ruegg KC, Smith TB (2002) Not as the crow flies: a historical explanation for circuitous migration in Swainson's thrush (*Catharus ustulatus*). *Proceedings of the Royal Society of London B*, **269**, 1375–1381.

Sauer JR, Hines JE, Fallon J (2008) *The North American Breeding Bird Survey, Results and Analysis 1966–2007. Version 5.15.2008*. USGS Patuxent Wildlife Research Center, Laurel, MD. Available from http://www.mbr-pwrc.usgs.gov/bbs/bbs.html (Accessed on 12th November, 2010).

Schmidt HA, Strimmer K, Vingron M, von Haeseler A (2002) TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics*, **18**, 502–504.

Smith TB, Clegg SM, Kimura M, Ruegg K, Milá B, Lovette I (2005) Molecular genetic approaches to linking breeding and overwintering areas in five Neotropical migrant passerines. In: *Birds of Two Worlds: The Ecology and Evolution of Migration* (eds Greenberg R, Marra PP). pp. 222–234, Johns Hopkins University Press, Baltimore, Maryland.

Toews DPL, Irwin DE (2008) Cryptic speciation in a Holarctic passerine revealed by genetic and bioacoustic analyses. *Molecular Ecology*, **17**, 2691–2705.

Toews DPL, Brelsford A, Irwin DE (in press) Hybridization between Townsend's and black-throated green warblers in an avian suture zone. *Journal of Avian Biology*, in press.

Vekemans X (2002) *AFLP-SURV version 1.0*. Distributed by the author. Laboratoire de Génétique et Ecologie Végétale, Université Libre de Bruxelles, Belgium.

Vos P, Hogers R, Bleeker M *et al.* (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research*, **23**, 4405–4414.

Webster MS, Marra PP, Haig SM, Bensch S, Holmes RT (2002) Links between worlds: unraveling migratory connectivity. *Trends in Ecology and Evolution*, **17**, 76–83.

Weir JT, Schluter D (2004) Ice sheets promote speciation in boreal birds. *Proceedings of the Royal Society of London B*, **271**, 1881–1887.

© 2011 Blackwell Publishing Ltd

Weir JT, Schluter D (2008) Calibrating the avian molecular clock. *Molecular Ecology*, **17**, 2321–2328.

D.E.I., an Associate Professor of Zoology at the University of British Columbia, is interested in the roles of geography, behaviour, genetics, morphology, and ecology in the diversification of vertebrates in the boreal forests of North America and Eurasia. J.H.I. shares those interests, but has recently been focused on maximizing their fitness while minimizing parent-offspring conflict. T.B.S., Director of the Center for Tropical Research and Professor of Ecology and Evolutionary Biology at the University of California, Los Angeles, is interested in evolutionary genetics, speciation and conservation biology of vertebrates.

Data accessibility

AFLP data tables for the 2-combination and 6-combination data sets: DRYAD entry: doi:10.5061/dryad.vd8b0.

Cytochrome *b* data matrix and phylogeny in Fig. 5: TreeBASE Study Accession URL: http://purl.org/phylo/treebase/phylows/study/TB2:S11522.

Supporting information

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

Fig. S1 Genetic variation in 153 Wilson's warblers sampled in across the breeding and wintering ranges, as illustrated using principal components analysis of 77 AFLP markers from two primer combinations (Table 2).

Fig. S2 Distribution of allele-based F_{ST} values between western and eastern breeding populations of Wilson's Warblers, among 257 AFLP markers.

Fig. S3 Results of Structure analysis of AFLP data from all samples in the 257-marker dataset, showing the likelihood of the data [Ln P(D)] given different values of the number of genetic clusters (*K*). Three replicates are shown (mean ± standard deviation) for each *K*, under the no admixture model (a) and the admixture model (b).

Fig. S4 Results of Structure analysis of AFLP data within the western breeding region (and including the winter samples, which are all assigned to the west; Fig. 4), showing the likelihood of the data [i.e. Ln P(D)] given different values of the number of genetic clusters (*K*). Three replicates are shown (mean ± standard deviation) for each *K*, under the no admixture model (a) and the admixture model (b).

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

© 2011 Blackwell Publishing Ltd