

# HALDANE'S RULE IN AN AVIAN SYSTEM: USING CLINE THEORY AND DIVERGENCE POPULATION GENETICS TO TEST FOR DIFFERENTIAL INTROGRESSION OF MITOCHONDRIAL, AUTOSOMAL, AND SEX-LINKED LOCI ACROSS THE *PASSERINA* BUNTING HYBRID ZONE

Matthew D. Carling<sup>1,2,3</sup> and Robb T. Brumfield<sup>1,2</sup>

<sup>1</sup>Museum of Natural Science, Louisiana State University, Baton Rouge, Louisiana 70803

<sup>2</sup>Department of Biological Sciences, Louisiana State University, Baton Rouge, Louisiana 70803

<sup>3</sup>E-mail: mcarli1@lsu.edu

Received May 1, 2008

Accepted July 10, 2008

Using cline fitting and divergence population genetics, we tested a prediction of Haldane's rule: autosomal alleles should introgress more than z-linked alleles or mitochondrial haplotypes across the *Passerina amoena/Passerina cyanea* (Aves: Cardinalidae) hybrid zone. We screened 222 individuals collected along a transect in the Great Plains of North America that spans the contact zone for mitochondrial (two genes), autosomal (four loci) and z-linked (two loci) markers. Maximum-likelihood cline widths estimated from the mitochondrial (223 km) and z-linked (309 km) datasets were significantly narrower on average than the autosomal cline widths (466 km). We also found that mean coalescent-based estimates of introgression were larger for the autosomal loci (0.63 genes/generation, scaled to the mutation rate  $\mu$ ) than for both the mitochondrial (0.27) and z-linked loci (0.59). These patterns are consistent with Haldane's rule, but the among-locus variation also suggests many independently segregating loci are required to investigate introgression patterns across the genome. These results provide the first comprehensive comparison of mitochondrial, sex-linked, and autosomal loci across an avian hybrid zone and add to the body of evidence suggesting that sex chromosomes play an important role in the formation and maintenance of reproductive isolation between closely related species.

**KEY WORDS:** Gene flow, Indigo Bunting, Lazuli Bunting, reproductive isolation, sex chromosomes, speciation.

Hybrid zones, formed after the secondary contact of partially reproductively isolated taxa (Hewitt 1988; Harrison 1990; Arnold 1997; Mallet 2005), are often used to investigate Haldane's rule (Haldane 1922) in nonmodel organisms for which laboratory crosses are impractical (Hagen and Scriber 1989; Brumfield et al. 2001; Saetre et al. 2003). The dominance theory of Haldane's

rule allows for the development of testable hypotheses related to the differential introgression of different loci across these hybrid zones. Based on the Dobzhansky-Muller (D-M) incompatibility model (Dobzhansky 1937; Muller 1940, 1942), the dominance theory states that inviability or sterility arises from the interaction between two genes that evolved incompatible alleles in allopatry.

If the alleles causing hybrid incompatibility are recessive, their impact will be much larger if one or both of the genes are located on the sex chromosomes.

Support for the dominance theory comes from a variety of empirical sources, including direct tests of a simple prediction: if genes causing hybrid sterility or inviability are recessive (thus being masked in  $F_1$  hybrids), it should be possible to find the causative alleles if they can be made homozygous or hemizygous. Such alleles have been found, mainly in *Drosophila* (Sawamura et al. 2000; Presgraves 2003; Tao and Hartl 2003). The dominance theory also is supported by comparative contrasts; researchers have found that Haldane's rule evolved more quickly in *Drosophila* with larger X chromosomes than in those with smaller X chromosomes (Turelli and Begun 1997). This "large-X effect" likely plays a role in non-*Drosophila* systems as well. It predicts that if the sex chromosomes play a large role in reproductive isolation, sex-linked alleles will show less introgression than autosomal alleles across hybrid zones. Such reductions in sex-linked introgression have been found in studies of natural hybrid zones between species of butterflies, mice, and birds (Hagen and Scriber 1989; Tucker et al. 1992; Saetre et al. 2003).

Mechanisms other than the dominance theory have been proposed to explain Haldane's rule (Coyne and Orr 2004), but they either do not apply to organisms with ZW sex determination (faster-male theory; Wu and Davis 1993; Wu et al. 1996), rely partially on the dominance theory (faster-x theory; Charlesworth et al. 1987), or are difficult to test in natural systems (meiotic drive theory; Frank 1991; Tao et al. 2003). Thus, our work focuses on testing introgression predictions based on the dominance theory.

In birds, females are the heterogametic sex (ZW) and thus, female hybrids are expected to show greater fitness reductions than male hybrids. Haldane's rule predicts that mitochondrial haplotypes, because they are maternally inherited, would show patterns of reduced introgression compared to autosomal alleles. Therefore, in taxa with heterogametic females, Haldane's rule predicts autosomal alleles will show greater levels of introgression than either mitochondrial or sex-linked alleles (z-linked hereafter).

Tests of Haldane's rule predictions in hybrid zones often involve investigating the relative selective forces acting on different loci by employing cline analyses to explain the changes in allele frequencies occurring along a geographic transect (Porter et al. 1997; Brumfield et al. 2001; Blum 2002; Payseur et al. 2004; Macholan et al. 2007). Cline theory predicts a relatively straightforward relationship between natural selection against hybrids and cline width—stronger selection against introgression results in a narrower cline (Slatkin 1973; Slatkin and Maruyama 1975; Endler 1977). Accordingly, loci (or linked loci) having alleles with narrower clines are thought to be more important in maintaining reproductive isolation between the hybridizing taxa than loci with wider clines. It is expected then, that mitochondrial haplotypes

(in taxa with heterogametic females) and z-linked alleles should show narrower clines than autosomal alleles.

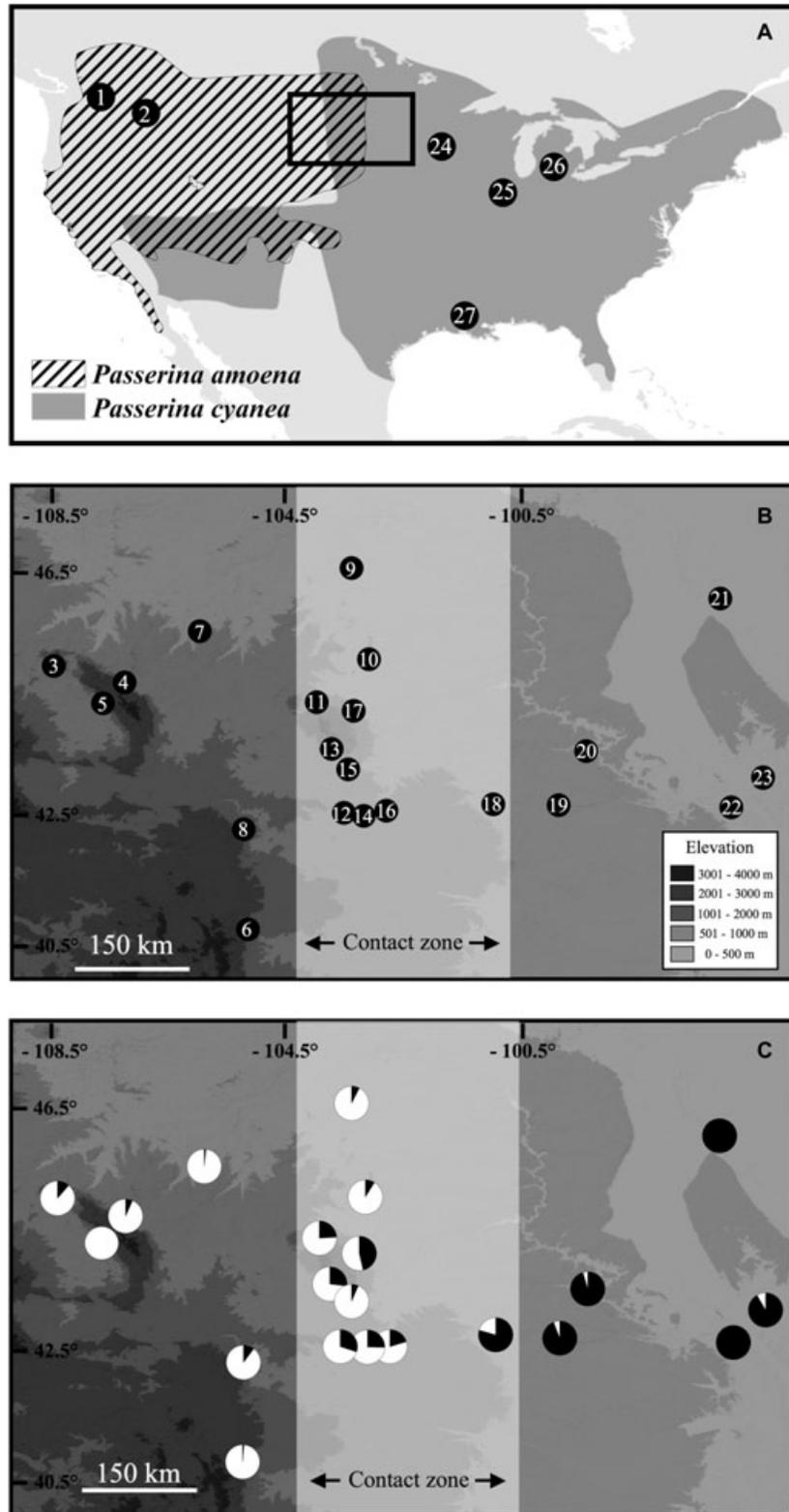
More recently, the increased sophistication of coalescent-based analytical techniques (Wakeley and Hey 1997; Hey 2005; Putnam et al. 2007; Rosenblum et al. 2007) has provided a new manner in which to investigate the selective forces acting on the movement of alleles across hybrid zones. Such methods use information contained in the genealogy of population samples to estimate a variety of population genetic parameters, including effective population size, introgression rates and divergence time, that can be used to explore the evolutionary history of populations. For example, these techniques have been used to reject strict allopatric speciation models in a number of taxa (Machado et al. 2002; Osada and Wu 2005). Additionally, by comparing gene flow parameters estimated from different loci, such coalescent-based methods can be used to investigate patterns of differential introgression. Loci with reduced rates of introgression may contribute more to the maintenance of reproductive isolation than loci with greater rates. Again, according to Haldane's rule, mitochondrial and z-linked alleles are predicted to show decreased rates of introgression compared to autosomal alleles.

In this study we had three principal objectives. The first was to provide the first genetic characterization of the hybrid zone between *Passerina amoena* and *Passerina cyanea*. Our second objective was to test the prediction that, according to Haldane's rule, both mitochondrial haplotypes and z-linked alleles should introgress less, on average, than autosomal alleles. Because two-thirds of z-linked alleles are carried by males, mitochondrial haplotypes should also show a reduction in introgression when compared with z-linked alleles. Our third objective was to investigate the utility of integrating cline-based and coalescent-based analyses of multilocus datasets to test the above predictions.

#### STUDY SYSTEM: PASSERINA BUNTINGS

*Passerina amoena* (Lazuli Bunting) and *P. cyanea* (Indigo Bunting) are closely related species (Carling and Brumfield 2008) that form a hybrid zone in which their breeding ranges overlap. The *Passerina* bunting hybrid zone (Fig. 1) is one of many hybrid zones clustered in a "suture zone" (Remington 1968) located near the junction between the western edge of the Great Plains and the eastern foothills of the Rocky Mountains (Swenson and Howard 2004, 2005). Although a diversity of taxa have contact zones in the same geographical location, the clustering of avian hybrid zones in the Northern Rocky Mountain zone is particularly striking (Rising 1983). Besides *Passerina* buntings, the distributions of 13 pairs of avian taxa meet in the Great Plains and form 10 additional hybrid zones (Rising 1983).

Previous morphological studies of the *Passerina* hybrid zone reported differences in the extent of introgression between *P. amoena* and *P. cyanea*. Sibley and Short (1959) reported



**Figure 1.** (A) Breeding distributions and allopatric sampling localities for *Passerina amoena* and *P. cyanea*. Region outlined by black box magnified in B. Digital maps (Ridgely et al. 2003) for each species were downloaded from NatureServe (2006) and modified. (B) More detailed view of sampling localities in and near the contact zone. (C) Allele frequencies (averaged across all seven loci) at each sampling locality in and near the contact zone. White proportion of pie charts refers to *P. amoena* alleles, black proportion refers to *P. cyanea* alleles. White box in (B) and (C) indicates location of contact zone. Background shading in (B) and (C) indicates elevation, darker shading = increasing elevation. Numbers of sampling localities refer to more detailed location information in Table 1.

widespread hybridization—nearly 70% of specimens collected from South Dakota, Wyoming, Nebraska, and Colorado were classified as hybrids. In contrast, two later studies reported much lower levels. Of 153 buntings collected across North Dakota and eastern Montana, only 35 were identified as hybrids (Kroodsma 1975), and ~5% of the 53 buntings collected in Nebraska and Wyoming were identified as hybrids (Emlen et al. 1975). These three studies differed in the location of the transect sampled as well as in the methods used to classify individuals as *P. amoena*, *P. cyanea*, or hybrids. Although the discrepancies in the reported frequency of hybridization may reflect different hybrid zone dynamics among the three transects, at least a portion of the difference is due to methodological differences. Both Kroodsma (1975) and Emlen et al. (1975) used a more conservative hybrid index score than Sibley and Short (1959), and a reassessment of the specimens collected by Sibley and Short suggested only ~7% of their specimens were hybrids (Emlen et al. 1975).

Behavioral studies of the *Passerina* bunting hybrid zone provide some support for assortative mating and interspecific male territoriality. Using captive individuals, Baker (1991) showed that *P. amoena* and *P. cyanea* males from sympatric populations respond antagonistically to each other's songs and that sympatric females give more copulation solicitation displays when exposed to conspecific male traits than when exposed to heterospecific male traits (Baker and Baker 1990; Baker 1996). Subsequent field-based research suggested male plumage may be a better predictor of conspecific versus heterospecific status when females are choosing mates (Baker and Boylan 1999). Because male buntings are able to alter their songs phrases from one breeding season to the next (Payne 1981), females may opt to rely on male plumage patterns, because they are likely less plastic than song characteristics. Field research also provided some support for selection against hybrids. Pairings that involved at least one hybrid individual (either the male or female, there were no instances in which both sexes were hybrids) resulted in fewer nestlings and fledglings when compared to pairings involving nonhybrid individuals (Baker and Boylan 1999).

## Methods

### SAMPLING, AMPLIFICATION AND SEQUENCING

During the summers of 2004–2007, we collected vouchered population samples of *P. amoena*, *P. cyanea*, and *P. amoena* × *P. cyanea* hybrids from 21 localities spanning the contact zone (Fig. 1, Table 1 and Supporting Information Table S1). For reference parental samples, we also acquired tissues from allopatric populations of *P. amoena* and *P. cyanea* west and east of the contact zone, respectively (Fig. 1, Table 1 and Supporting Information Table S1), resulting in 248 individuals collected from 27 focal populations. Four additional localities represented by a sin-

gle individual (Supporting Information Table S1) were included in the coalescent-based analyses, but not in the cline-based analyses. All skin and tissue specimens collected for this project were deposited in the Louisiana State University Museum of Natural Science.

We extracted genomic DNA from ~25 mg of pectoral muscle using either standard phenol/chloroform methods or a DNeasy Tissue Kit (QIAGEN Inc., Valencia, CA). Each individual was amplified at a suite of eight loci (two mitochondrial, four autosomal, and two z-linked markers; Table 2) using the following PCR conditions in a 25  $\mu$ l reaction: ~40 ng template DNA (2  $\mu$ l of DNA extracts), 1  $\mu$ l of 10 mM dNTPs (2.5 mM each dATP, dTTP, dCTP, dGTP), 1  $\mu$ l of each primer (10 mM, Table 2), 2.5  $\mu$ l 10 × Buffer with MgCl<sub>2</sub> (15 mM), 0.1  $\mu$ l Taq (5 U/ $\mu$ l of either AmpliTaq DNA Polymerase, Applied Biosystems Inc., Foster City, CA or Taq DNA Polymerase, New England Biolabs, Ipswich, MA), and 17.4  $\mu$ l sterile diH<sub>2</sub>O. The thermocycling profile was as follows: an initial 95°C denaturation for 2 min followed by 35 cycles consisting of a 30 sec, 95°C denaturation step, a 30 sec, locus-specific temperature primer annealing step (Table 2), and a 2 min, 72°C extension step, and a final extension of 5 min at 72°C. To check for amplification we electrophoresed 2.5  $\mu$ l of each PCR product on a 1% agarose gel.

We cycle-sequenced both strands of all PEG-purified PCR amplicons in a 7  $\mu$ l reaction using 1.5  $\mu$ l of 5 X sequencing buffer (Applied Biosystems), 1  $\mu$ l of 10 mM primer, 2.5  $\mu$ l of template, 0.15–0.25  $\mu$ l Big Dye Terminator Cycle-Sequencing Kit v 3.1 (A), and 1.85–1.95  $\mu$ l sterile diH<sub>2</sub>O. We cleaned cycle-sequencing products on Sephadex (G-50 fine) columns and electrophoresed the cleaned products on a 3100 Genetic Analyzer (Applied Biosystems). All sequences were edited and assembled using Sequencher ver. 4.7 (GeneCodes Corp., Ann Arbor, MI). When direct sequencing of purified PCR amplicons of autosomal and z-linked loci revealed more than one heterozygous site within a sequence, we resolved haplotypes probabilistically using PHASE (Stephens et al. 2001; Stephens and Donnelly 2003).

To identify the largest independently segregating block of sequence data for all nuclear loci (autosomal and z-linked), we tested for intralocus recombination using the four-gamete test as implemented in DnaSP ver. 4.10 (Rozas et al. 2003). We also used DnaSP to calculate haplotype diversity ( $H_d$ ; Nei 1987) and nucleotide diversity ( $\pi$ ; Tajima 1983) within each species. Species identification was based on plumage pattern (see below). To assign haplotypes as belonging to either *P. amoena* or *P. cyanea*, we used TCS ver. 1.21 (Clement et al. 2000) to build parsimony-based haplotype networks that included all phased haplotypes. Individual haplotypes with frequencies  $\geq 0.80$  in the “allopatric” *P. amoena* population (WA) were classified as *P. amoena* haplotypes. Alternately, haplotypes with frequencies  $\geq 0.80$  in the “allopatric” *P. cyanea* population (LA) were designated as *P. cyanea*

**Table 1.** Sampling localities, sample sizes, and distance along sampling transect of individuals analyzed in this study.

Code	Locality	Latitude	Longitude	<i>N</i>	Distance from locality 1 (km)
1	WA	47.8	-118.9	29	0
2	ID	44.1	-116.2	3	232
3	MT: Custer National Forest #1	45.1	-108.5	4	842
4	WY: Bighorn National Forest #1	44.8	-107.3	8	943
5	WY: Bighorn National Forest #2	44.6	-107.7	2	944
6	CO: Roosevelt National Forest	40.7	-105.2	7	946
7	MT: Custer National Forest #2	45.7	-106	5	1003
8	WY: Medicine Bow National Forest	42.4	-105.3	3	1072
9	ND: Little Missouri National Grassland	46.8	-103.5	13	1083
10	SD: Custer National Forest	45.3	-103.2	10	1191
11	WY: Sand Creek <sup>1</sup>	44.5	-104.1	14	1198
12	NE: White River <sup>1</sup>	42.6	-103.5	8	1203
13	SD: Black Hills National Forest	43.7	-103.8	5	1225
14	NE: Ponderosa State Wildlife Management Area	42.6	-103.3	7	1227
15	SD: The Nature Conservancy Whitney Preserve	43.3	-103.6	3	1243
16	NE: Nebraska National Forest	42.8	-102.9	8	1245
17	SD: Ft. Meade National Recreation Area	44.4	-103.5	4	1250
18	NE: Nenzel <sup>1</sup>	42.8	-101.1	4	1395
19	NE: The Nature Conservancy Niobrara Valley Preserve	42.8	-100	18	1479
20	SD: Carpenter Game Production Area	43.7	-99.5	13	1492
21	ND: The Nature Conservancy Pigeon Point Preserve	46.3	-97.3	4	1557
22	NE: Wiseman State Wildlife Area	42.8	-97.1	10	1670
23	SD: Newton Hills Game Production Area	43.2	-96.6	13	1721
24	MN	44.9	-93.7	10	1861
25	IL	40.5	-88.9	10	2322
26	MI	42.3	-83.7	7	2714
27	LA <sup>2</sup>			26	-
	CA <sup>2</sup>			1	
	MO <sup>2</sup>			1	
	SC <sup>2</sup>			1	
	TX <sup>2</sup>			1	

<sup>1</sup>Access to these localities generously provided by private landowners.

<sup>2</sup>Individuals from these localities were excluded from cline analyses; see Supporting Information Table S1 for exact sampling localities.

haplotypes. The most common MYO2 haplotype could not be assigned as either *P. amoena* or *P. cyanea* based on the above criterion. For the cline analyses (see below) we estimated cline shape parameters for the most common MYO2 haplotype.

### CLINE ANALYSES

For the cline analyses we excluded those individuals for which PHASE was unable to assign haplotypes at a probability greater than 0.75 (range: 4–21 individuals per locus), and the 26 individuals from Louisiana, because they did not fall on a natural linear transect from west to east. With the remaining individuals (Table 3), we used a procedure similar to Macholan et al. (2007) to generate a linear transect spanning the hybrid zone. First, for

localities 3–23 we averaged the frequency of *P. cyanea* alleles across all seven loci (Fig. 1C). In the case of MYO2, we only included those alleles that could be unambiguously assigned to either *P. amoena* or *P. cyanea*, thus the most common allele was not included in this calculation. Next, using the Contour Plot option in JMP 5.1, we plotted the 0.5-isocline using *x* and *y* coordinates and average allele frequencies of localities 3–23, which are those populations within the previously described area of overlap between *P. amoena* and *P. cyanea*. Treating this 0.5-isocline as the center of the zone, we measured the shortest straight-line distance between all sampling localities and the 0.5-isocline. Because all individuals from Washington (locality 1), Minnesota (24), Illinois (25), and Michigan (26) were not from the same sampling locality

**Table 2.** PCR annealing temperatures, sequence lengths, primer sequences, diversity measures, and GenBank accession numbers for mitochondrial, autosomal, and z-linked loci.

Locus	Annealing Temp	Length (bp) <sup>1</sup>	PCR/sequencing primers <sup>2</sup>	Internal sequencing primers <sup>3</sup>	<i>P. cyanea</i>		<i>P. amoena</i>		Access. Nos.
					$H_d$	$\pi$	$H_d$	$\pi$	
ND3 <sup>4</sup>	46	356	L10755 H11151	NA NA	0.96	0.008	0.92	0.019	EU859404–652
Cont. Region <sup>4</sup>	55	168	F: TAC CTA GGA GGT GGG CGA AT R: CCC AAA CAT TAT CTC CAA AA	NA NA					EU858464–710
ALDOB3 <sup>5</sup>	61	111	F: GGCAGGAACAAATGGAGAACT R: GCCAGAACCTGAAAACAGGAG	F: CACTTGGCAAAGTCAGCACCATCT NA	0.6	0.008	0.41	0.005	EU858008–237
BRM15 <sup>5</sup>	60	300	F: AGC ACC TTT GAA CAG TGG TT R: TAC TTT ATG GAG ACG ACG GA	NA NA	0.64	0.004	0.55	0.003	EU858238–463
GADPH <sup>6</sup>	62	101	F: ACCTTTCATGCGGGTGCTGGCATTGC R: CATCAAGTCCACAACACGGTTGCTGTA	F: CAGAGGGTAGATGGGA NA	0.55	0.007	0.71	0.01	EU858711–950
MC1R <sup>6</sup>	62.5	202	F: CTGGCTCCGGAAGGCRTAGAT R: AYGCCAGYGAGGGCAACCA	F: SGCRTAGAAGATGGTGATGTAGC R: GTBGACCGCTACATCACCAT	0.38	0.002	0.33	0.002	EU858951–9172
MYO2 <sup>6</sup>	58	113	F: CTGGCTCCGGAAGGCRTAGAT R: AYGCCAGYGAGGGCAACCA	F: GCATGCCTGGGAAACTGTAT R: GGTTGCAGAGCCTGGAAATA	0.57	0.006	0.51	0.005	EU859173–403
RHO1 <sup>6</sup>	60	114	F: TGCTACATCGAGGGCTTCTT R: GCAGTGACCAGAGAGCGATT	F: CCTGGCAGAAGCCTAATCTCTGA R: CATGTCAGAGTGTGACCACAGGC	0.12	0.001	0.27	0.004	EU859653–897

<sup>1</sup>Length of longest independently segregating block.

<sup>2</sup>PCR and external sequencing primer references: ND3—Chesser 1999; Cont. Region—this study; GADPH—Primmer et al. 2002; MC1R—Cheviron et al. 2006; MYO2—Heslewood et al. 1998 (F), Slade et al. 1993 (R); RHO1—Primmer et al. 2002; ALDOB3—G. P. Saetre, pers comm.; BRM15—Borge et al. 2005.

<sup>3</sup>Internal sequencing primer references: GADPH—this study; MC1R—Cheviron et al. 2006; RHO1—Carling and Brumfield 2008, ALDOB3—this study.

<sup>4</sup>ND3 and Cont. Region were analyzed together as a single mtDNA dataset.

<sup>5</sup>z-linked locus.

<sup>6</sup>Autosomal locus.

**Table 3. Cline shape parameters (two-unit likelihood support limits) for locus-specific and combined loci datasets.**

Dataset	$N^1$	$c$ (km from locality 1)	$w$ (km)	$\theta_L$	$\theta_R$	$z_L$	$z_R$
mtDNA <sup>2</sup>	222	1262 (1232–1297)	223 (158–305)	0.242 (0.016–0.816)	0.410 (0.044–0.819)	945 (81–994)	604 (80–997)
<i>z</i> -linked							
ALDOB3	388	1353 (1319–1394)	268 (185–346)	0.442 (0.050–0.858)	0.035 (0.015–0.180)	595 (163–994)	692 (44–965)
BRM15	379	1328 (1294–1360)	245 (159–310)	0.172 (0.046–0.780)	0.018 (0.009–0.045)	791 (80–1000)	1000 (570–1000)
<i>z</i> -linked <sup>3</sup>	–	1361 (1334–1385)	309 (241–375)	0.152 (0.064–0.837)	0.014 (0.009–0.040)	511 (123–967)	999 (539–1000)
autosomal							
GADPH	428	1259 (1225–1350)	175 (110–236)	0.002 (0.001–0.004)	0.020 (0.010–0.739)	1000 (780–1000)	1000 (108–1000)
MC1R	386	1312 (1258–1365)	461 (295–648)	0.031 (0.013–0.101)	0.036 (0.018–0.146)	1000 (494–1000)	1000 (418–1000)
MYO2	402	1703 (1427–1915)	523 (399–632)	0.008 (0.006–0.019)	0.008 (0.007–0.017)	999 (722–1000)	999 (711–1000)
RHO1	434	1309 (1282–1362)	243 (159–313)	0.016 (0.008–0.073)	0.022 (0.009–0.048)	999 (275–1000)	1000 (640–1000)
autosomal <sup>4</sup>	–	1333 (1294–1354)	466 (344–545)	0.024 (0.014–0.038)	0.026 (0.015–0.035)	1000 (815–1000)	999 (816–1000)

<sup>1</sup>Number of individuals (mtDNA) or number of chromosomes (all other loci) sampled for cline analyses.<sup>2</sup>Combined dataset of ND3 and Cont. Region haplotypes.<sup>3</sup>Combined dataset of ALDOB3 and BRM15 alleles.<sup>4</sup>Combined dataset of GADPH, MC1R and RHO1 alleles.

within the respective states, we plotted all sampling localities in Google Earth and then picked a point at the estimated center of the resultant centroid. In this way, the samples from Washington, Minnesota, Illinois, and Michigan were collapsed into a single locality for each state (Table 1). The farthest sampling locality (WA) was then set to 0 km and the location of all other localities was recalculated accordingly, resulting in a linear transect from WA to MI (Table 1).

Along this transect, we investigated introgression patterns using the methods developed by Szymura and Barton (1986, 1991) and implemented in the program ClineFit (Porter et al. 1997). These methods estimate cline shape parameters using three equations that model the relationship between allele frequency data within a sampling locality and the geographic location of those localities. The first equation (1) describes a symmetrical, sinusoidal shape in the center of the cline and the other two equations (2 and 3) describe the exponential change in allele frequencies on the left and right sides of the cline:

$$p = \frac{1}{2} \left[ 1 + \tanh \left( \frac{2[x - c]}{w} \right) \right], \quad (1)$$

$$p = \exp \left[ \frac{4(x - [c + z_L])\sqrt{\theta_L}}{w} \right], \quad (2)$$

$$p = 1 - \exp \left[ \frac{-4(x - [c - z_R])\sqrt{\theta_R}}{w} \right]. \quad (3)$$

In these equations,  $c$  represents the location of the center of the cline (measured in kilometers from the WA locality),  $w$  is the cline width (1/max slope), and  $x$  is the geographic location along the sampling transect (kilometers from WA locality). In equations (2 and 3),  $z_L$  and  $z_R$  describe the distance from the center ( $c$ ) of a vertical asymptote for the exponential decay of allele frequencies on the left and right side of the zone, respectively. The parameters,  $\theta_L$  and  $\theta_R$  are the exponential decay values relative to the shape of the central cline (eq. 1) on the left and right sides. The three equations are related in that as the parameters  $z$  and  $\theta$  approach zero and 1, equation (2) and (3) approach the shape of equation (1) on their respective sides of the zone. The center of the zone ( $c$ ) is the point along the transect in which allele frequencies change most rapidly and the width of the zone ( $w$ ) provides an estimate of the geographic distance over which that rapid change in allele frequencies occurs. The parameters  $\theta_L$  and  $\theta_R$  describe the exponential rate of change in allele frequencies in the western and eastern tails of the cline and  $z_L$  and  $z_R$  provide information on the geographic distance over which the exponential decay in the tails occurs.

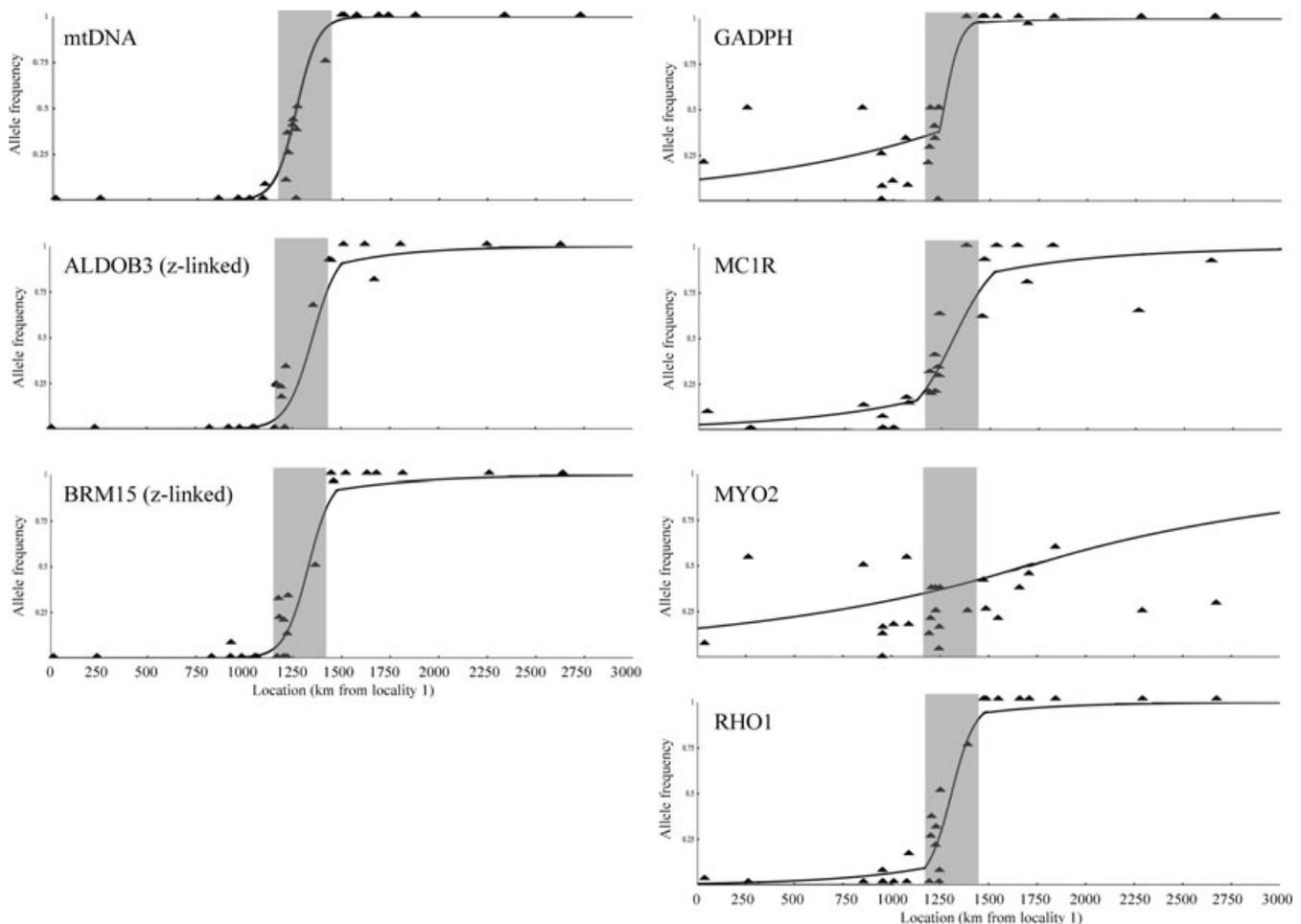
For each locus, as well as combined autosomal loci and combined *z*-linked loci datasets ( $n = 9$  datasets), we tested the fit of two different models using likelihood-ratio tests: the

two-parameter model, which includes only the center and width parameters, and the six-parameter model, which estimates the six shape parameters described (Porter et al. 1997; Brumfield et al. 2001). In the majority of tests (seven of nine), the six-parameter model was a significantly better fit to the data (likelihood-ratio tests,  $P < 0.05$ ). Thus, to facilitate comparisons between the different datasets, we used the six-parameter model to describe the cline shape for every dataset. The following search parameters were used in each analysis: burn-in: parameter tries per step—300; sampling for support: replicates saved  $\geq 2000$ , and 30 replicates between saves. Differences in parameter estimates between different loci were assessed using the two-unit likelihood support limits, which are analogous to 95% confidence limits (Edwards 1992), and by comparing the log likelihood score of the cline parameter estimates from an unconstrained model to the log likelihood score when the width was constrained to equal the width estimated from an unconstrained search of another dataset. Significance was assessed using a likelihood-ratio test. Because an initial inspection suggested the change in MYO2 allele fre-

quency data across the transect was not clinal (Fig. 2), it was excluded from the estimates of combined autosomal cline shape parameters.

### COALESCENT ANALYSES

If introgression of autosomal alleles is greater than that of mitochondrial and z-linked markers, autosomal loci should show increased levels of interspecific gene flow. We assessed this using the coalescent-based nonequilibrium Isolation with Migration model implemented in the software program IM (Hey and Nielsen 2004; Hey 2005). By combining coalescent theory with Bayesian methodologies, IM simultaneously estimates multiple population genetic parameters for two populations. These parameters, which are scaled to the neutral mutation rate,  $\mu$ , include:  $\theta_1 = 4N_1\mu$ ,  $\theta_2 = 4N_2\mu$ ,  $\theta_A = 4N_A\mu$ ,  $t = t\mu$ ,  $m_1 = m_1/\mu$ , and  $m_2 = m_2/\mu$ . The two migration rates,  $m_1$  and  $m_2$ , allow for inferences of asymmetric introgression, and can be used to compare gene flow rates between different marker classes (e.g., mitochondrial vs. autosomal, and z-linked vs. autosomal).



**Figure 2.** Maximum-likelihood cline shapes and allele frequencies for each locus-specific cline, as well as for the combined autosomal and z-linked loci datasets. Gray shading indicates approximate location of contact zone. Cline shape parameters are provided in Table 3.

Because IM assumes no intralocus recombination and requires fully resolved haplotypes (i.e., no ambiguous sites) with no gaps, we had to pare down the nuclear DNA sequence datasets. First, we removed all gaps in the alignments. Second, we included only those individuals for which PHASE was able to assign haplotypes with a probability of greater than 0.75. We tested for recombination as described above. Lastly, we used the dominant plumage motif (e.g., presence of white wing bars and rusty breast band are characteristic of *P. amoena* males) of each specimen to assign individual birds to either the *P. amoena* or *P. cyanea* population. Because females of the two species are difficult to distinguish morphologically, we excluded eight females collected in the contact zone. We also excluded two males (Supporting Information Table S1) for which definitive population assignment was difficult (Table 4). Because all individual classifications were the same in each dataset, these exclusions are unlikely to systematically bias the results.

To facilitate comparisons of introgression estimates between different loci, we combined the sequence data for all loci into a single dataset. From this dataset we estimated one effective population size (we forced  $\theta_1 = \theta_2 = \theta_A$ ), divergence time ( $t$ ), and asymmetric introgression rates ( $m_1$  and  $m_2$ ) for each locus. Additionally, we included inheritance scalars in the input file (mitochondrial—0.25, autosomal—1.0, z-linked—0.75) to account for the difference in effective population sizes due to different ploidy and inheritance modes of the three marker types. We initially ran IM using the HKY finite substitution model with wide, uninformative priors for > 2 million steps and used these initial runs to identify more appropriate priors (Won and Hey 2005). These adjusted priors ( $\theta_1 = 10$ ,  $m_1 = 10$ ,  $m_2 = 10$ ,  $t = 10$ ) were then used in two replicate “final” analyses that differed only in starting random number seed. All “final” analyses were run with a burn-in of 200,000 steps and were allowed to continue

until the Effective Sample Size (ESS) values for each parameter were > 100 (Hey 2005). Convergence was also assessed by inspecting the plots of parameter trend lines and by comparing the results of the two replicate runs. Because all parameter estimates for the two replicate runs were qualitatively similar, we only present the estimates from the longest “final” run of each dataset (number of steps greater than  $5 \times 10^7$ ). The 14 directional introgression rate estimates (two mitochondrial rates—from *P. cyanea* into *P. amoena* and vice-versa, eight autosomal rates (two for each of four loci), and four z-linked rates (two for each of two loci)) from the longest “final” run were used in comparisons of differential introgression.

The introgression estimates are scaled to the mutation rate, which is approximately 10 times faster for mitochondrial than for nuclear loci (Graur and Li 2000). We assumed the following mutation rates:  $3.6 \times 10^{-9}$  substitutions/site/year for autosomal and  $3.9 \times 10^{-9}$  substitutions/site/year for z-linked loci (Axelsson et al. 2004). We divided the introgression estimates for the mitochondrial data by 10 to make the estimates more readily comparable across loci.

## Results

No heterozygous sites in the two presumed z-linked loci (ALDOB3 and BRM15) were found in any females. For four of seven loci, both haplotype diversity ( $H_d$ ) and nucleotide diversity ( $\pi$ ) were larger in *P. cyanea* than in *P. amoena* (Table 2). In two datasets, RHO1 and GADPH, both diversity measures were larger in *P. amoena* and in the mitochondrial dataset, haplotype diversity was larger in *P. cyanea*, but nucleotide diversity was higher in *P. amoena*.

Males that were defined as *P. amoena* based on plumage patterns (see section “Coalescent Analyses”) were found as far

**Table 4.** Smoothed highpoint coalescent-based estimates (90% Highest Posterior Densities) of per generation asymmetric introgression rates between *P. amoena* and *P. cyanea*.

Locus	$N^1$ <i>P. amoena</i>	$N^1$ <i>P. cyanea</i>	Introgression from <i>P. amoena</i> into <i>P. cyanea</i>	Introgression from <i>P. cyanea</i> into <i>P. amoena</i>
mtDNA	121	117	0.065 (0.005–0.286)	0.473 (0.252–0.882)
z-linked				
ALDOB3	214	222	1.225 (0.385–3.055)	0.305 (0.005–1.115)
BRM15	223	193	0.375 (0.075–1.105)	0.455 (0.085–1.295)
autosomal				
GADPH	234	242	0.355 (0.045–1.155)	1.505 (0.525–3.225)
MC1R	236	200	0.015 (0.005–2.435)	0.405 (0.005–2.355)
MYO2	222	222	1.675 (0.005–5.685)	0.685 (0.005–5.325)
RHO1	244	242	0.175 (0.015–0.625)	0.225 (0.005–0.925)

<sup>1</sup>Sample sizes ( $N$ ) are the number of individuals (mtDNA) or the number of chromosomes (all other loci) designated as belonging to either *P. amoena* or *P. cyanea*.

east as population 18 and *P. cyanea* males were found as far west as population 11 (Fig. 1). We refer to this region, which is approximately 250-km wide, as the “contact zone.” No genetically pure individuals of the wrong type (e.g., an individual with all *P. amoena* alleles but identified as *P. cyanea* based on plumage) were found outside of the contact zone. After excluding the MYO2 data, there were no individuals collected on the east side of the contact zone with a frequency of *P. amoena* alleles greater than 0.273. A similar pattern was found on the west side: no single individual had a frequency of *P. cyanea* alleles greater than 0.364.

### CLINE SHAPE ESTIMATES

The centers of all locus-specific clines, with the exception of MYO2, were coincident and fell within a 64 km range (extremes: GADPH—1259 km east of locality 1 and ALDOB3—1353 east of locality 1; Fig. 2 and Table 3). This 64 km range begins just east of locality 17 and encompasses the midpoint between localities 11 and 18, which form the boundaries of the contact zone (see above). Similarly, the cline widths of all locus-specific clines were concordant (Table 3). Widths ranged from 175 km (GADPH) to 461 km (MC1R). Cline widths of the other four loci (mtDNA, ALDOB3, BRM15, and RHO1) were all between 223 and 268 km, approximately the same width as the contact zone (~250 km). Overall, the estimates of cline center were less variable (coefficient of variation = 0.028) than the estimates of cline width (c.v. = 0.368). There was no association between the locus-specific estimates of cline center and width (ANOVA  $P > 0.4$ ).

The cline centers of the combined autosomal (1333 km east of locality 1; minus MYO2) and z-linked (1361 km east of locality 1) datasets were coincident (Table 3), but shifted east from the mitochondrial cline center (1262 km east of locality 1; Fig. 2 and Table 3). The autosomal cline width was 466 km (two-unit likelihood support limits: 344–545 km), significantly greater than the mitochondrial width (223 km; two-unit likelihood support limits: 158–305 km), and greater than the z-linked cline width (309 km). The difference in cline widths between the autosomal and z-linked clines was almost significant using the nonoverlapping support limits criterion (autosomal two-unit likelihood support limits: 344–545 km; z-linked: 241–375 km). Forcing the z-linked width to equal the autosomal width (466 km) resulted in a significantly worse fit of the data to the model ( $\chi^2 = 32.1$ ,  $df = 1$ ,  $P < 0.0001$ ). The difference between the mitochondrial and z-linked cline widths was not significant by either method (two-unit likelihood support limits; Table 3), although forcing the mitochondrial width to be 309 km (the z-linked width) resulted in a worse fit that was close to significant ( $\chi^2 = 2.85$ ,  $df = 1$ ,  $P = 0.09$ ). As with the locus-specific clines, the cline center estimates for the combined datasets were less variable (c.v. = 0.015) than the cline width estimates (c.v. = 0.286).

### INTROGRESSION ESTIMATES

With the exception of divergence time ( $t$ ) the distributions of the posterior probabilities of all parameter estimates showed a single peak. The smoothed highpoint estimate of  $\theta_1$  was 2.285 (90% HPD: 1.763–2.895). Smoothed high point estimates along with the 90% highest posterior density intervals (90% HPD) are given for all introgression parameters in Table 4 and Supporting Information Figure S1. Although none of the introgression parameter estimates were significantly different by this criterion, as in the cline shape estimates, the variation in locus-specific introgression rates was considerable (Table 4 and Supporting Information Fig. S1). Introgression from *P. amoena* into *P. cyanea* estimated from the MYO2 data was more than 100-fold greater than mitochondrial introgression in the same direction (Table 4 and Supporting Information Fig. S1). No pairwise comparisons of locus-specific introgression estimates from *P. cyanea* into *P. amoena* differed by more than a factor of seven (Table 4).

For all appropriate pairwise comparisons, IM tallies which parameter is larger during each recorded step (every 10 steps in the chain), and this tally can also be used to assess differences in parameter estimates. For any given locus the proportion of steps in which the estimate of introgression from *P. cyanea* into *P. amoena* is greater than in the opposite direction is indicative of asymmetric introgression at that locus; the opposite pattern is also possible. Using this criterion, mitochondrial introgression from *P. cyanea* into *P. amoena* is significantly greater than mitochondrial introgression from *P. amoena* into *P. cyanea* (proportion of recorded steps in which the introgression estimate into *P. amoena* was larger than in the opposite direction: 0.979). Although not significant at an  $\alpha = 0.05$ , introgression estimates at two other loci were asymmetric. Like the estimates for the mitochondrial data, introgression estimates for GADPH were greater from *P. cyanea* into *P. amoena* (proportion: 0.946). Introgression of ALDOB3 alleles was also asymmetric, but in the opposite direction. In most steps the estimate from *P. amoena* into *P. cyanea* was larger than from *P. cyanea* into *P. amoena* (proportion: 0.902). In all other locus-specific comparisons of asymmetric introgression estimates, the proportion of steps in which introgression in one direction was larger than in the other direction was never greater than 0.643.

To compare the coalescent-based introgression estimates with the average mitochondrial, autosomal, and z-linked cline widths, we calculated a mean introgression rate for each class of genetic marker (mitochondrial, autosomal, z-linked) from the IM parameter estimates. These means were calculated as the sum of all smoothed highpoint introgression parameter estimates within a marker class, divided by the number of parameters. The average autosomal introgression estimate was the largest (0.63 genes per generation, scaled to the mutation rate  $\mu$ ), followed by the z-linked estimate (0.59), and the mitochondrial estimate (0.27).

## Discussion

Using DNA sequence data from a suite of mitochondrial, autosomal, and z-linked loci, we conducted the first detailed genetic characterization of the *Passerina* bunting hybrid zone. Our results demonstrate that, on average, mitochondrial and z-linked markers exhibit reduced introgression relative to autosomal markers (Fig. 2 and Table 3). This finding is consistent with theoretical predictions of Haldane's rule, stemming from the dominance theory, as well as with empirical results in other systems, but is the first of its kind in an avian system. Within each marker class (i.e., mitochondrial, autosomal, and z-linked), we found considerable among-locus variance in introgression estimates (Fig. 2 and Table 3). Because patterns of differential introgression are due to the amount of genetic drift and the strength of selection acting on each locus, our results suggest that many more independently segregating markers than were used here are required (1) to adequately quantify mean levels of introgression across the genome and (2) to help disentangle the relative roles of genetic drift and natural selection on each locus.

We also found some support for the hypothesis that mitochondrial haplotypes should introgress less than z-linked alleles. The coalescent-based estimate of mitochondrial introgression was lower (0.27 genes per generation, scaled to the mutation rate  $\mu$ ) than for z-linked loci (0.59). The estimates of cline width suggest that mitochondrial haplotypes (223 km) introgressed less than z-linked alleles (309 km), but the difference was not significant. To our knowledge, there has been only one other attempt to systematically compare introgression patterns of mitochondrial, autosomal, and z-linked loci between taxa with a ZW sex-determination system. Cianchi et al. (2003) found less introgression of mitochondrial and sex-linked loci than of autosomal loci across a hybrid zone between the swallowtail butterflies *Papilio machaon* and *P. hospiton*, but they did not statistically test the significance of the reductions.

Other hybrid zone studies between taxa with heterogametic females, particularly between Lepidopteran species, have found reductions in mitochondrial introgression relative to autosomal alleles (Hagen and Scriber 1989; Jiggins et al. 1997; Dasmahapatra et al. 2002; Kronforst et al. 2006), and a few have reported lower introgression of sex-linked alleles than autosomal alleles (Hagen 1990; Putnam et al. 2007). Although these studies were not designed to explicitly test the different hypotheses for the genetic mechanism of Haldane's rule, they do, along with our study, strongly support the predictions of the dominance theory and are also consistent with the dominance requiring faster-x theory (Charlesworth et al. 1987; Orr 1997). As noted above, the faster-male theory (Wu and Davis 1993; Wu et al. 1996) cannot explain Haldane's rule in Lepidopteran and avian taxa. Thus it is clear that the dominance theory of Haldane's rule plays an important

role in causing the fitness reductions seen in hybrids between taxa with heterogametic females.

### MAINTENANCE OF *PASSERINA* HYBRID ZONE

Multiple lines of evidence suggest that the narrow clines (cline analysis) and low migration rates (coalescent analysis) observed across most markers sampled in this study cannot be explained solely by neutral diffusion following secondary contact. In the absence of selection, the neutral diffusion model (Endler 1977; Barton and Gale 1993) can be used to approximate the expected width of a cline using estimates of root-mean-square (RMS) dispersal distance ( $\sigma$ ) and time ( $t$ ), in generations, since contact ( $w = 2.51\sigma\sqrt{t}$ ; Barton and Gale 1993). In *Passerina* buntings reliable estimates of both RMS and time since contact are unavailable, but we can place limits on these variables to explore how the hybrid zone might appear under the neutral diffusion hypothesis.

Representative estimates of RMS in other North American passerines range from  $\sim 30$  km (*Dendroica occidentalis* and *D. townsendi*; Rohwer and Wood 1998) to  $\sim 150$  km (*Catharus ustulatus*; Rugg 2008). Although based on only two individuals, recapture data of *P. cyanea* banded as nestlings suggest that the dispersal distance from natal sites can be quite large (52 and 350 km; Payne 1992). Given the lack of any RMS estimates in *Passerina* buntings, we use the estimates from other species as a rough guide. Individuals of *P. amoena* were reported to be breeding in eastern South Dakota (near population 22) in 1888 (Sibley and Short 1959), which indicates that the two species have likely been in contact for a minimum of 120 generations (assuming generation time = 1 year). Using a conservative estimate of 30 km for RMS dispersal and 120 years as limits on time since contact, the neutral diffusion model predicts the *Passerina* hybrid zone should be  $\sim 830$  wide. The widths of all clines were narrower (range: 165–523 km) than even the most conservative estimate provided by the neutral diffusion model ( $\sim 830$  km), thus selection appears to be maintaining the width of the hybrid zone, at least at some loci.

Because neutral diffusion is an unlikely explanation for the introgression patterns, alternative explanations for the maintenance of the *Passerina* hybrid zone are required. The general geographic location of the hybrid zone coincides with the Great Plains–Rocky Mountain ecotone (Swenson and Howard 2004, 2005), raising the possibility that the bounded-superiority hypothesis (Moore 1977), wherein hybrid individuals have increased fitness over parental types in the intermediate environment occurring at the transition between two ecosystems' may explain the *Passerina* hybrid zone maintenance. However, ecological niche modeling suggests that the fundamental niche of the *Passerina* hybrid zone is much wider ( $\sim 750$  km from eastern Wyoming to western Iowa; Swenson 2006) than estimated in this study. If

hybrids are indeed more fit than parentals in transitional habitats, the hybrid zone should cover the entire area of transitional habitat, but that is not the case. Additionally, numerous “pure” parentals of both species are found within the hybrid zone (Baker and Boylan 1999), which is in the area of transitional habitat (Swenson 2006), and field studies conducted in the hybrid zone found that pairings involving at least one hybrid individual fledge fewer young than pairings between nonhybrid individuals (Baker and Boylan 1999).

The available data, from our genetic work, field studies of reproductive success in hybrids, and ecological niche modeling, all point to selection against hybridization playing a prominent role in the maintenance of the narrow *Passerina* hybrid zone. What the selection pressure is remains unknown, but previous studies suggest a number of possibilities. One interesting postmating hypothesis implicates differences in the molt schedules of *P. amoena* and *P. cyanea* (Rohwer and Johnson 1992; Voelker and Rohwer 1998). *Passerina amoena* individuals begin their prebasic molt before leaving breeding areas in August and migrate to a molting “hot spot” in southern Arizona and New Mexico or in southern Baja California (Greene et al. 1996). They generally stay at these molting “hot spots” for one month while they complete the prebasic molt before continuing on to their wintering grounds in western Mexico. In contrast, *P. cyanea* individuals leave the breeding areas approximately one month later, having already undergone their prebasic molt, and migrate directly to their wintering grounds in Central and northern South America (Payne 1992). It is possible then, that hybrid individuals could undergo two prebasic molts (instead of the usual single prebasic molt), one on the breeding grounds, as seen in *P. cyanea*, and another sometime later, as in *P. amoena*. Alternately, there may be delays in either the timing of molt or migration that jeopardize hybrid survival or future reproductive success (Helbig 1991; Ruegg 2008). Although our data cannot address the role the differences in the timing of molt and migration may play in reducing hybrid fitness, these potential problems seem unlikely to be the sole selective agent. At least some hybrid individuals do return to the breeding grounds (Sibley and Short 1959; Emlen et al. 1975; Kroodsmma 1975; Baker and Boylan 1999; M. D. Carling, pers. obs.).

A premating hypothesis is that the intermediate plumages and songs of hybrid males limit their ability to acquire territories and mates. Females of both species routinely give more copulation-solicitation displays to conspecific males than to heterospecific males in experiments with captive birds (Baker 1996). Interestingly, females appear to take more cues from plumage than song, perhaps because male *Passerina* buntings are capable of altering their song from one year to the next (Payne 1981), which subsequently reduces the reliability of song as an accurate indicator of a male’s genetic makeup. In sympatry, pure plumaged parentals mate assortatively but hybrid females mate randomly,

which also implies sexual selection against hybrid males (Baker and Boylan 1999). Careful field studies that directly examine pre and postmating reproductive isolating mechanisms in this system are needed.

### COMPARING CLINEFIT AND IM RESULTS

In general, the cline-based and coalescent-based analytical methods produced the same results—introgression of mitochondrial and z-linked markers was less, on average, than that of autosomal markers. The differences seen can likely be attributed to the differences in what the two techniques are estimating. ClineFit describes the sigmoidal change in allele frequencies across the geographic transect. Alleles found on the “wrong” side of the hybrid zone (e.g., a *P. cyanea* GADPH allele found in a *P. amoena* individual in population 7) extend the width of the hybrid zone. As such, there is a chance that if alternate alleles are not fixed at each end of the transect, cline widths may be overestimated because the presence of the “wrong” allele is assumed to be introgression when it may in fact be retained ancestral polymorphism. This may explain the difference between the wide MC1R cline (Table 3) and the fairly limited rates of introgression of MC1R estimated using IM (Table 4 and Supporting Information Fig. S1). The presence of *P. cyanea* alleles on the west end of the transect resulted in a wide cline, but IM likely attributed the presence of those alleles to the retention of ancestral polymorphisms that have not yet completely sorted.

The geographic location of the samples can also influence cline-based and coalescent-based inferences of introgression differently. No geographic information is provided as input for IM analyses, with individual DNA sequences simply classified as belonging to one of two populations. Therefore, differences like those seen for GADPH, which had the narrowest cline of any single locus dataset (Table 3), but the second highest coalescent-based estimate of introgression (Table 4), might be best explained by geography. There were a relatively large number of individuals with *P. amoena* plumage that possessed a *P. cyanea* GADPH allele, and vice-versa, leading to the high IM estimates of introgression, but most of those individuals were clustered in the contact zone, leading to the narrow cline width. For some loci, the two methods are quite consistent. For example, the mitochondrial cline was relatively narrow and IM estimates of introgression were low, because few individuals possessed the “wrong” mitochondrial haplotype for their plumage type, but those that did were contained within a narrow geographic region.

Despite the potential for inconsistencies, combining cline-based and coalescent-based methods offers a great deal of promise for investigating the role of post-divergence introgression in the formation and maintenance of species. Certainly, cline-based analyses are more limited in their ability to separate introgression from the retention of ancestral polymorphisms and cannot be used

to estimate divergence times of loci with interesting introgression patterns. Although coalescent-based methods are continually improving (Putnam et al. 2007; Rosenblum et al. 2007), they are not precisely designed to explore the impact of introgression in a geographical setting. That said, it seems possible that researchers could set up a geographically structured experimental design to take advantage of the benefits of coalescent-based methods. By hierarchically analyzing two populations at increasing distances from each other, levels of introgression across a landscape may be estimated using current coalescent-based methods.

### GENETICS OF REPRODUCTIVE ISOLATION

Our results support the findings of both theoretical (Coyne and Orr 1989; Kirkpatrick and Ravigne 2002; Coyne and Orr 2004; Kirkpatrick and Hall 2004) and empirical studies in other systems (True et al. 1996; Prowell 1998; Jiggins et al. 2001; Tao et al. 2003; Counterman et al. 2004; Payseur et al. 2004) that the sex chromosomes likely play a relatively important role in maintaining reproductive isolation between closely related taxa. Cline-based analyses such as this have identified X-linked candidate regions for reproductive isolation as well as candidate regions for adaptive introgression across the *Mus musculus*/*M. domesticus* hybrid zone (Payseur et al. 2004). Given the reduction in z-linked introgression, relative to autosomal loci, seen in this study, further investigations of the introgression patterns of additional loci distributed along the z-chromosome are warranted.

In *Ficedula* flycatchers, phylogenetic analyses of mitochondrial, autosomal, and z-linked loci suggest sex chromosomes play an important role in reproductive isolation between *Ficedula albicollis* and *F. hypoleuca* (Saetre et al. 2003), two species that hybridize where their breeding ranges overlap in Europe (Saetre et al. 1997, 1999, 2003, 2001; Veen et al. 2001). Although introgression patterns of z-linked loci along a geographic sampling transect have yet to be explored in the *Ficedula* hybrid zone, their research is the only other study, to our knowledge, to have used z-linked loci to investigate the genetics of reproductive isolation between hybridizing avian species. As the availability of genetic and genomic resources continues to grow (Backstrom et al. 2008), it is likely that more researchers will investigate introgression patterns of z-linked loci across avian hybrid zones. Such studies will be critical to comparative investigations of avian speciation, and will help address whether genes of the same functional class are routinely involved in reproductive isolation.

### ACKNOWLEDGMENTS

We thank S. Birks (University of Washington Burke Museum of Natural History and Culture), A. Capparella (Illinois State University), D. Dittmann (Louisiana State University Museum of Natural Science), J. Hinshaw (University of Michigan Museum of Zoology), and M. Westberg (University of Minnesota Bell Museum) for generously providing tissue samples for this project. S. Carling and J. Maley were outstanding

field assistants. To C. Burney, Z. Cheviron, A. Cuervo, E. Derryberry, R. Eytan, J. Klicka, I. Lovette, J. Maley, and M. Webster all provided helpful comments on the manuscript. J. Hey and A. Porter assisted greatly with their software packages IM and ClineFit, respectively. Scientific collecting permits were granted by the Colorado Division of Wildlife, Montana Fish, Wildlife and Parks, Nebraska Game and Parks, North Dakota Game and Fish, South Dakota Game, Fish and Parks, Wyoming Game and Fish, and the United States Fish and Wildlife Service. Louisiana State University Institutional Animal Care and Use Committee Protocol: 08–025. The map data were provided by NatureServe in collaboration with R. Ridgely, J. Zook, The Nature Conservancy—Migratory Bird Program, Conservation International—CABS, World Wildlife Fund—US, and Environment Canada—WILDSPACE. This work was supported by grants from NSF (DEB-0543562, DEB-0808464, and DBI-0400797) to RTB, as well as by grants from the AMNH Chapman Fund, AOU, Explorer's Club, LSUMNS Birdathon and Prepathon Funds, LSU Department of Biological Sciences, and Sigma-Xi.

### LITERATURE CITED

- Arnold, M. L. 1997. Natural hybridization and evolution. Oxford Univ. Press, New York.
- Axelsson, E., N. G. C. Smith, H. Sundstrom, S. Berlin, and H. Ellegren. 2004. Male-biased mutation rate and divergence in autosomal, Z-linked and W-linked introns of chicken and turkey. *Mol. Biol. Evol.* 21:1538–1547.
- Backstrom, N., S. Fagerberg, and H. Ellegren. 2008. A gene-based set of reference markers evenly spread across the avian genome. *Mol. Ecol.* 17:964–980.
- Baker, M. C. 1991. Response of male Indigo and Lazuli buntings and their hybrids to song playback in allopatric and sympatric populations. *Behaviour* 119:225–242.
- . 1996. Female buntings from hybridizing populations prefer conspecific males. *Wilson Bull.* 108:771–775.
- Baker, M. C., and A. E. M. Baker. 1990. Reproductive behavior of female buntings—isolating mechanisms in a hybridizing pair of species. *Evolution* 44:332–338.
- Baker, M. C., and J. T. Boylan. 1999. Singing behavior, mating associations and reproductive success in a population of hybridizing Lazuli and Indigo Buntings. *Condor* 101:493–504.
- Barton, N., and K. S. Gale. 1993. Genetic analysis of hybrid zones. Pp. 13–45 in R. Harrison, ed. *Hybrid zones and the evolutionary process*. Oxford Univ. Press, New York.
- Blum, M. J. 2002. Rapid movement of a *Heliconius* hybrid zone: evidence for phase III of Wright's shifting balance theory? *Evolution* 56:1992–1998.
- Borge T., M. T. Webster, G. Andersson, and G. P. Saetre. 2005. Contrasting patterns of polymorphism and divergence on the Z chromosome and autosomes in two *Ficedula* flycatcher species. *Genetics* 171:1861–1873.
- Brumfield, R. T., R. W. Jernigan, D. B. McDonald, and M. J. Braun. 2001. Evolutionary implications of divergent clines in an avian (*Manacus*: Aves) hybrid zone. *Evolution* 55:2070–2087.
- Carling, M. D., and R. T. Brumfield. 2008. Integrating phylogenetic and population genetic analyses of multiple loci to test species divergence hypotheses in *Passerina* buntings. *Genetics* 178:363–377.
- Charlesworth, B., J. A. Coyne, and N. Barton. 1987. The relative rates of evolution of sex chromosomes and autosomes. *Am. Nat.* 130: 113–146.
- Chesser, R. T. 1999. Molecular systematics of the rhynchoptid genus *Pteroptochos*. *Condor* 101:439–446.
- Cheviron, Z. A., S. J. Hackett, and R. T. Brumfield. 2006. Sequence variation in the coding region of the melanocortin-1 receptor gene (MC1R) is not associated with plumage variation in the Blue-crowned Manakin (*Lepidothrix coronata*). *Proc. R. Soc. Lond. B* 273:1613–1618.

- Cianchi, R., A. Ungaro, M. Marini, and L. Bullini. 2003. Differential patterns of hybridization and introgression between the swallowtails *Papilio machaon* and *P. hospiton* from Sardinia and Corsica islands (Lepidoptera, Papilionidae). *Mol. Ecol.* 12:1461–1471.
- Clement, M., D. Posada, and K. A. Crandall. 2000. TCS: a computer program to estimate gene genealogies. *Mol. Ecol.* 9:1657–1660.
- Counterman, B. A., D. Ortiz-Barrientos, and M. A. F. Noor. 2004. Using comparative genomic data to test for fast-x evolution. *Evolution* 58:656–660.
- Coyne, J. A., and H. A. Orr. 1989. Two rules of speciation. Pp. 180–207 in D. Otte, and J. A. Endler, eds. *Speciation and its consequences*. Sinauer Associates, Sunderland, MA.
- . 2004. *Speciation*. Sinauer Associates, Inc., Sunderland, MA.
- Dasmahapatra, K. K., M. J. Blum, A. Aiello, S. Hackwell, N. Davies, E. Bermingham, and J. Mallet. 2002. Inferences from a rapidly moving hybrid zone. *Evolution* 56:741–753.
- Dobzhansky, T. 1937. *Genetics and the origin of species*. Columbia Univ. Press, New York.
- Edwards, A. 1992. *Likelihood*. Johns Hopkins Univ. Press, Baltimore, MD.
- Emlen, S. T., J. D. Rising, and W. L. Thompson. 1975. Behavioral and morphological study of sympatry in Indigo and Lazuli buntings of Great-Plains. *Wilson Bull.* 87:145–179.
- Endler, J. A. 1977. *Geographic variation, speciation, and clines*. Princeton Univ. Press, Princeton, NJ.
- Frank, S. H. 1991. Divergence of meiotic drive-suppressors as an explanation for sex-biased hybrid sterility and inviability. *Evolution* 45:262–267.
- Graur, D., and W. H. Li. 2000. *Fundamentals of molecular evolution*. Sinauer Associates, Inc., Sunderland, MA.
- Greene, E., V. R. Muehter, and W. Davison. 1996. Lazuli bunting (*Passerina amoena*). In A. Poole, and F. Gill, eds. *The birds of North America*. The Academy of Natural Sciences and The American Ornithologists' Union, Philadelphia, PA and Washington DC.
- Hagen, R. H. 1990. Population structure and host use in hybridizing subspecies of *Papilio glaucus*. *Evolution* 44:1914–1930.
- Hagen, R. H., and M. Scriber. 1989. Sex-linked diapause, color, and allozyme loci in *Papilio glaucus*: linkage analysis and significance in a hybrid zone. *J. Hered.* 80:179–185.
- Haldane, J. B. S. 1922. Sex ratio and unisexual sterility in animal hybrids. *J. Genet.* 12:101–109.
- Harrison, R. 1990. Hybrid zones: windows on the evolutionary process. Pp. 69–128 in D. J. Futuyma, and J. Antonovics, eds. *Oxford surveys in evolutionary biology*. Oxford Univ. Press, Oxford.
- Helbig, A. J. 1991. SE- and SW- migrating Blackcap (*Sylvia atricapilla*) populations in Central Europe: orientation of birds in the hybrid zone. *J. Evol. Biol.* 4:657–670.
- Heslewood, M. M., M. S. Elphinstone, S. C. Tidemann, and P. R. Baverstock. 1998. Myoglobin intron variation in the Gouldian Finch *Erythrura gouldiae* assessed by temperature gradient gel electrophoreses. *Electrophoresis* 19:142–151.
- Hewitt, G. M. 1988. Hybrid zones—natural laboratories for evolutionary studies. *Trends Ecol. Evol.* 3:158–167.
- Hey, J. 2005. On the number of New World founders: a population genetic portrait of the peopling of the Americas. *Plos Biol.* 3:965–975.
- Hey, J., and R. Nielsen. 2004. Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics* 167:747–760.
- Jiggins, C. D., W. O. McMillan, P. King, and J. Mallet. 1997. The maintenance of species differences across a *Heliconius* hybrid zone. *Heredity* 79:495–505.
- Jiggins, C. D., R. E. Linares, R. E. Naisbit, C. Salazar, Z. H. Yang, and J. Mallet. 2001. Sex-linked hybrid sterility in a butterfly. *Evolution* 55:1631–1638.
- Kirkpatrick, M., and D. W. Hall. 2004. Sexual selection and sex linkage. *Evolution* 58:683–691.
- Kirkpatrick, M., and V. Ravigne. 2002. Speciation by natural and sexual selection: models and experiments. *Am. Nat.* 159:S22–S35.
- Kronforst, M. R., L. G. Young, L. M. Blume, and L. E. Gilbert. 2006. Multilocus analyses of admixture and introgression among hybridizing *Heliconius* butterflies. *Evolution* 60:1254–1268.
- Kroodsma, R. L. 1975. Hybridization in Buntings (*Passerina*) in North-Dakota and Eastern Montana. *Auk* 92:66–80.
- Machado, C. A., R. M. Kliman, J. A. Markert, and J. Hey. 2002. Inferring the history of speciation from multilocus DNA sequence data: the case of *Drosophila pseudoobscura* and close relatives. *Mol. Biol. Evol.* 19:472–488.
- Macholan, M., P. Munclinger, M. Sugerkova, P. Dufkova, B. Bimova, E. Bozikova, J. Zima, and J. Pialek. 2007. Genetic analysis of autosomal and x-linked markers across a mouse hybrid zone. *Evolution* 61:746–771.
- Mallet, J. 2005. Hybridization as an invasion of the genome. *Trends Ecol. Evol.* 20:229–237.
- Moore, W. S. 1977. Evaluation of narrow hybrid zones in vertebrates. *Q. Rev. Biol.* 52:263–277.
- Muller, H. J. 1940. Bearing of the *Drosophila* work on systematics. Pp. 185–268 in J. B. S. Haldane, ed. *The new systematics*. Clarendon Press, Oxford.
- . 1942. Isolating mechanisms, evolution, and temperature. *Biol. Symp.* 6:71–125.
- NatureServe. 2006. NatureServe explorer: an online encyclopedia of life, version 6.1. <http://www.natureserve.org/explorer>.
- Nei, M. 1987. *Molecular evolutionary genetics*. Columbia Univ. Press, New York City.
- Orr, H. A. 1997. Haldane's rule. *Ann. Rev. Ecol. Syst.* 28:195–218.
- Osada, N., and C. I. Wu. 2005. Testing the mode of speciation with genomic data—examples from the great apes. *Genetics* 169:259–264.
- Payne, R. B. 1981. Song learning and social interaction in Indigo Buntings. *Anim. Behav.* 29:688–697.
- . 1992. Indigo bunting. In A. Poole, P. Stettenheim, and F. Gill, eds. *The birds of North America*. The Academy of Natural Sciences, Washington, DC; The American Ornithologists' Union, Philadelphia.
- Payseur, B. A., J. G. Krenz, and M. W. Nachman. 2004. Differential patterns of introgression across the X chromosome in a hybrid zone between two species of house mice. *Evolution* 58:2064–2078.
- Porter, A. H., R. Wenger, H. Geiger, A. Scholl, and A. M. Shapiro. 1997. The *Pontia daplidice-edusa* hybrid zone in northwestern Italy. *Evolution* 51:1561–1573.
- Presgraves, D. C. 2003. A fine-scale genetic analysis of hybrid incompatibilities in *Drosophila*. *Genetics* 163:955–972.
- Primmer, C. R., T. Borge, J. Lindell, and G. P. Saetre. 2002. Single-nucleotide polymorphism characterization in species with limited available sequence information: high nucleotide diversity revealed in the avian genome. *Mol. Ecol.* 11:603–612.
- Prowell, D. P. 1998. Sex linkage and speciation in Lepidoptera. Pp. 309–319 in D. J. Howard, and S. H. Berlocher, eds. *Endless forms: species and speciation*. Oxford Univ. Press, New York.
- Putnam, A. S., J. M. Scriber, and P. Andolfatto. 2007. Discordant divergence times among z-chromosome regions between two ecologically distinct swallowtail butterfly species. *Evolution* 61:912–927.
- Remington, C. L. 1968. Suture-zones of hybrid interaction between recently joined biotas. Pp. 321–428 in T. Dobzhansky, M. K. Hecht, and W. C.

- Steere, eds. Evolutionary biology, Volume 2. Appleton Century Crofts, New York.
- Ridgely, R. S., T. F. Allnutt, T. Brooks, D. K. McNicol, D. W. Mehlman, B. E. Young, and J. R. Zook. 2003. Digital distribution maps of the birds of the Western hemisphere, version 1.0. NatureServe, Arlington, VA.
- Rising, J. D. 1983. The Great Plains hybrid zones. Pp. 131–157 in R. F. Johnston, ed. Current ornithology. Plenum, New York.
- Rohwer, S., and M. S. Johnson. 1992. Scheduling differences of molt and migration for Baltimore and Bullocks orioles persist in a common environment. *Condor* 94:992–994.
- Rohwer, S., and C. Wood. 1998. Three hybrid zones between Hermit and Townsend's Warblers in Washington and Oregon. *Auk* 115:284–310.
- Rosenblum, E. B., M. J. Hickerson, and C. Moritz. 2007. A multilocus perspective on colonization accompanied by selection and gene flow. *Evolution* 61:2971–2985.
- Rozas, J., J. C. Sanchez-DelBarrio, X. Messeguer, and R. Rozas. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19:2496–2497.
- Ruegg, K. 2008. Genetic, morphological, and ecological characterization of a hybrid zone that spans a migratory divide. *Evolution* 62:452–466.
- Saetre, G. P., M. Kral, and S. Bures. 1997. Differential species recognition abilities of males and females in a flycatcher hybrid zone. *J. Avian Biol.* 28:259–263.
- Saetre, G. P., M. Kral, S. Bures, and R. A. Ims. 1999. Dynamics of a clinal hybrid zone and a comparison with island hybrid zones of flycatchers (*Ficedula hypoleuca* and *F. albicollis*). *J. Zool.* 247:53–64.
- Saetre, G. P., T. Borge, J. Lindell, T. Moum, C. R. Primmer, B. C. Sheldon, J. Haavie, A. Johnsen, and H. Ellegren. 2001. Speciation, introgressive hybridization and nonlinear rate of molecular evolution in flycatchers. *Mol. Ecol.* 10:737–749.
- Saetre, G. P., T. Borge, K. Lindroos, J. Haavie, B. C. Sheldon, C. Primmer, and A. C. Syvanen. 2003. Sex chromosome evolution and speciation in *Ficedula* flycatchers. *Proc. R. Soc. Lond. B* 270:53–59.
- Sawamura, K., A. W. Davis, and C. I. Wu. 2000. Genetic analysis of speciation by means of introgression into *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* 97:2652–2655.
- Sibley, C. G., and L. L. J. Short. 1959. Hybridization in the buntings (*Passerina*) of the Great Plains. *Auk* 76:443–463.
- Slade, R. W., C. Moritz, A. Heideman, and P. T. Hale. 1993. Rapid assessment of single-copy nuclear DNA variation in diverse species. *Mol. Ecol.* 2:359–373.
- Slatkin, M. 1973. Gene flow and selection in a cline. *Genetics* 75:733–756.
- Slatkin, M., and T. Maruyama. 1975. Genetic drift in a cline. *Genetics* 81:209–222.
- Stephens, M., and P. Donnelly. 2003. A comparison of Bayesian methods for haplotype reconstruction from population genotype data. *Am. J. Hum. Genet.* 73:1162–1169.
- Stephens, M., N. Smith, and P. Donnelly. 2001. A new statistical method for haplotype reconstruction from population data. *Am. J. Hum. Genet.* 68:978–989.
- Swenson, N. G. 2006. GIS-based niche models reveal unifying climatic mechanisms that maintain the location of avian hybrid zones in a North American suture zone. *J. Evol. Biol.* 19:717–725.
- Swenson, N. G., and D. J. Howard. 2004. Do suture zones exist? *Evolution* 58:2391–2397.
- . 2005. Clustering of contact zones, hybrid zones, and phylogeographic breaks in North America. *Am. Nat.* 166:581–591.
- Szymura, J. M., and N. Barton. 1986. Genetic analysis of a hybrid zone between the fire-bellied toads, *Bombina orientalis* and *Bombina orientalis*, near Cracow in southern Poland. *Evolution* 40:1141–1159.
- . 1991. The genetic structure of the hybrid zone between the fire-bellied toads *Bombina orientalis* and *B. orientalis*—Comparisons between transects and between loci. *Evolution* 45:237–261.
- Tajima, F. 1983. Evolutionary relationships of DNA sequences in finite populations. *Genetics* 105:437–460.
- Tao, Y., and D. L. Hartl. 2003. Genetic dissection of hybrid incompatibilities between *Drosophila simulans* and *D. mauritiana*. III. Heterogeneous accumulation of hybrid incompatibilities, degree of dominance, and implications for Haldane's rule. *Evolution* 57:2580–2598.
- Tao, Y., S. Chen, D. L. Hartl, and C. C. Laurie. 2003. Genetic dissection of hybrid incompatibilities between *Drosophila simulans* and *D. mauritiana*. I. Differential accumulation of hybrid male sterility factors on the X and autosomes. *Genetics* 164:1383–1397.
- True, J., B. S. Weir, and C. C. Laurie. 1996. A genome-wide survey of hybrid incompatibility factors by the introgression of marked segments of *Drosophila mauritiana* chromosomes into *Drosophila simulans*. *Genetics* 142:819–837.
- Tucker, P. K., R. D. Sage, J. Warner, A. C. Wilson, and E. M. Eicher. 1992. Abrupt cline for sex-chromosomes in a hybrid zone between two species of mice. *Evolution* 46:1146–1163.
- Turelli, M., and D. J. Begun. 1997. Haldane's rule and x-chromosome size in *Drosophila*. *Genetics* 147:1799–1815.
- Veen, T., T. Borge, S. C. Griffith, G. P. Saetre, S. Bures, L. Gustafsson, and B. C. Sheldon. 2001. Hybridization and adaptive mate choice in flycatchers. *Nature* 411:45–50.
- Voelker, G., and S. Rohwer. 1998. Contrasts in scheduling of molt and migration in eastern and western Warbling Vireos. *Auk* 115:142–155.
- Wakeley, J., and J. Hey. 1997. Estimating ancestral population parameters. *Genetics* 145:847–855.
- Won, Y. J., and J. Hey. 2005. Divergence population genetics of chimpanzees. *Mol. Biol. Evol.* 22:297–307.
- Wu, C.-I., and A. W. Davis. 1993. Evolution of post-mating reproductive isolation—The composite nature of Haldane's rule and its genetic bases. *Am. Nat.* 142:187–212.
- Wu, C.-I., N. A. Johnson, and M. F. Palopoli. 1996. Haldane's rule and its legacy: why are there so many sterile males? *Trends Ecol. Evol.* 11:281–284.

Associate Editor: M. Webster

## *Supporting Information*

The following supporting information is available for this article:

**Figure S1.** Posterior probability distributions of migration rate parameters between *P. amoena* and *P. cyanea* for locus-specific datasets.

**Table S1.** Voucher numbers and sampling localities of all individuals.

This material is available as part of the online article from:

<http://www.blackwell-synergy.com/doi/abs/10.1111/j.1558-5646.2008.00477.x>

(This link will take you to the article abstract).

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