Molecular Ecology (2009) **18**, 834–847 doi: 10.1111/j.1365-294X.2008.04038.x

Speciation in *Passerina* buntings: introgression patterns of sex-linked loci identify a candidate gene region for reproductive isolation

MATTHEW D. CARLING* and ROBB T. BRUMFIELD
Museum of Natural Science, 119 Foster Hall, Louisiana State University, Baton Rouge, LA, 70803, USA; Department of Biological Sciences, Louisiana State University, Baton Rouge, LA, 70803, USA

Abstract

Sex-chromosomes are thought to play an important role in speciation, but few studies of non-model organisms have investigated the relative influence of multiple sex-linked markers on reproductive isolation. We collected 222 individuals along a geographical transect spanning the hybrid zone between *Passerina amoena* and *P. cyanea* (Aves: Cardinalidae). Using maximum-likelihood cline fitting methods, we estimated locus-specific introgression rates for 10 z-linked markers. Although the cline width estimates ranged from 2.8 to 584 km, eight of 10 loci had cline widths between 224 and 271 km. We also used coalescent-based estimates of locus-specific divergence times between *P. amoena* and *P. cyanea* to test a recently proposed hypothesis of an inverse relationship between divergence time and cline width but did not find a significant association. The narrow width (2.8 km) of the cline estimated from the *VLDLR9* locus indicates strong selection retarding introgression of alleles at this locus across the hybrid zone. Interestingly, a mutation in the very low density lipoprotein receptor (*VLDLR*) gene, in which *VLDLR9* is an intron, is known to reduce the egg-laying ability of some chickens, suggesting a possible link between this gene region and reproductive isolation between *P. amoena* and *P. cyanea*. These results underscore the importance of sampling multiple loci to investigate introgression patterns across a chromosome or genome and support previous findings of the importance of sex-linked genes in speciation.

Keywords: cline theory, divergence population genetics, Indigo Bunting, Lazuli Bunting

Received 2 September 2008; revised 7 November 2008; accepted 10 November 2008

Introduction

Theoretical and empirical studies support a view in which the sex-chromosomes are important in speciation (Charlesworth *et al*. 1987; Coyne & Orr 1989; Prowell 1998; Jiggins *et al*. 2001; Sætre *et al*. 2003; Tao *et al*. 2003; Payseur *et al*. 2004). The theoretical arguments are derived primarily from Haldane's rule (Haldane 1922), which states that in the formation of a hybrid between differentiated taxa it is usually the heterogametic sex that is inviable or infertile. The dominance theory of Haldane's rule, based on the Dobzhansky–Muller (D–M) incompatibility model (Dobzhansky 1937; Muller 1940, 1942), states that inviability or sterility arise from the interaction between two genes that evolved incompatible alleles in allopatry. If the alleles causing hybrid incompatibility are recessive, genes located on the sex-chromosomes will have a much larger impact than genes located on the autosomes.

Much of the empirical support for Haldane's rule comes from the study of hybrid zones; geographical regions where closely related taxa that are partially reproductively isolated meet and interbreed (Hewitt 1988; Harrison 1990; Arnold 1997; Mallet 2005). Through the use of cline-based analyses (Haldane 1948; Slatkin 1973; Slatkin & Maruyama 1975; Endler 1977; Barton 1979), hybrid-zone research has made important contributions to speciation genetics (Szymura & Barton 1986, 1991; Porter *et al*. 1997; Rieseberg *et al*. 1999; Price & Bouvier 2002; Barbash *et al*. 2003; Presgraves *et al*. 2003; Fitzpatrick 2004; Brumfield 2005). Many hybrid zones are thought to be maintained by a balance between

Correspondence: Matthew D Carling, Fax: 607–254–2486; E-mail: mdc248@cornell.edu
*Present address: Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, NY, 14853, USA; Cornell Laboratory of Ornithology, 159 Sapsucker Woods Road, Ithaca, NY, 14850, USA

natural selection against hybrids and the dispersal of parental types into the zone (Barton & Hewitt 1985, 1989). Under this tension-zone model, the amount of genetic introgression across a zone can be inferred from the estimated cline widths, which are related to both dispersal rates and the strength of selection against hybrids. For a given dispersal distance, narrower clines indicate stronger selective pressure. Narrow clines may point to chromosomal regions involved in genetic incompatibilities between divergent taxa. In support of the predictions of Haldane's rule, many studies have found that sex-linked loci show patterns of reduced introgression relative to autosomal loci (Hagen & Scriber 1989; Tucker *et al*. 1992; Sætre *et al*. 2003; Carling & Brumfield 2008a).

While hybrid-zone analyses have been successful in identifying genetic regions with reduced introgression (Hagen & Scriber 1989; Tucker *et al*. 1992; Payseur *et al*. 2004), one potential complication of using cline-based analyses in the search for candidate regions of reproductive isolation is the inability of such analyses to distinguish between those loci that may have been important in the initial stages of population divergence and those loci that diverged more recently (Coyne & Orr 2004). For example, Presgraves (2003) estimated that approximately 190 genes are involved in causing hybrid inviability between *Drosophila melanogaster* and *D. simulans*. It is less clear how many of these genes were important in maintaining reproductive isolation during the early stages of divergence in allopatry vs. how many evolved incompatibilities after reproductive isolation was complete. Because later substitutions may be incompatible with earlier substitutions, genetic incompatibilities between species can accumulate quickly once reproductive isolation has been completed (Orr 1995). This supports the hypothesis that the number of genes initially involved in reproductive isolation might be lower than what can be identified using hybrid crosses and illustrates the difficulty in trying to determine the timing of the evolution of reproductive isolation.

Methods rooted in coalescent theory can be combined with traditional cline-based analyses to provide new ways of investigating the genetics of reproductive isolation. Coalescent-based analyses are capable of jointly estimating a variety of population genetic parameters, such as effective population size, introgression rates and divergence time between differentiated populations, by using information contained in the genealogies of population samples (Wakeley & Hey 1997; Hey 2005; Hey & Nielsen 2007; Putnam *et al*. 2007; Rosenblum *et al*. 2007). Thus, it is possible to investigate differences in divergence time among different loci thought to be candidates for reproductive isolation. Loci important in the initial stages of divergence between the focal taxa should have divergence times that are older than divergence-time estimates for loci that contributed less to the early divergence (Putnam *et al*. 2007). If loci with older divergence times make a relatively large contribution to the evolution of reproductive isolation, they may also show relatively narrow cline widths that are consistent with a pattern of strong selection opposing introgression. As such, there may be a negative relationship between cline width and divergence time (Putnam *et al*. 2007), a hypothesis we test here.

In this paper, we combine cline-based and coalescent-based methods to investigate the relative contribution of a suite of z-linked loci to reproductive isolation between *Passerina amoena* and *P. cyanea* (Aves: Cardinalidae).

Passerina *bunting hybrid zone*

Passerina cyanea (Indigo Bunting) and *P. amoena* (Lazuli Bunting) are closely related oscine passerines (Carling & Brumfield 2008b) that hybridize where their breeding ranges overlap in the western Great Plains and eastern Rocky Mountains of North America (Fig. 1; Sibley & Short 1959; Emlen *et al*. 1975; Kroodsma 1975; Baker & Baker 1990; Baker 1991, 1996; Baker & Johnson 1998). Analyses of mitochondrial DNA (mtDNA), autosomal and z-linked loci from individuals collected along a geographical transect spanning the contact zone demonstrated that the genomes of these two species are semi-permeable to introgression (Carling & Brumfield 2008a). In accordance with the predictions of the dominance theory of Haldane's rule, autosomal alleles showed greater average levels of introgression than either mtDNA haplotypes or z-linked alleles. Interestingly, introgression patterns of two z-linked loci (*BRM15* and *ALDOB3*) were asymmetric, with greater introgression from *P. amoena* into *P. cyanea* than in the opposite direction.

These results suggest that the z-chromosome may be fertile ground for investigations of the genetic basis of reproductive isolation in avian systems. Additionally, the overall pattern of asymmetric introgression suggests adaptive introgression may also be occurring. Here, using the *Passerina* bunting hybrid zone as a model, we characterize patterns of introgression among 10 z-linked loci to explore the role of the sex-chromosomes in avian speciation. We also investigate the interplay between locus-specific divergence time and cline width to test the hypothesis that loci with older divergence times between *P. cyanea* and *P. amoena* show reduced cline widths when compared to loci with more recent divergence times.

Materials and methods

Sampling, amplification and sequencing

We collected population samples of *Passerina cyanea*, *P. amoena* and *P. cyanea* × *P. amoena* hybrids, during May, June and July of 2004–2007 from 21 localities spanning the contact zone (Fig. 1, Table 1 and Supplemental Material). We also

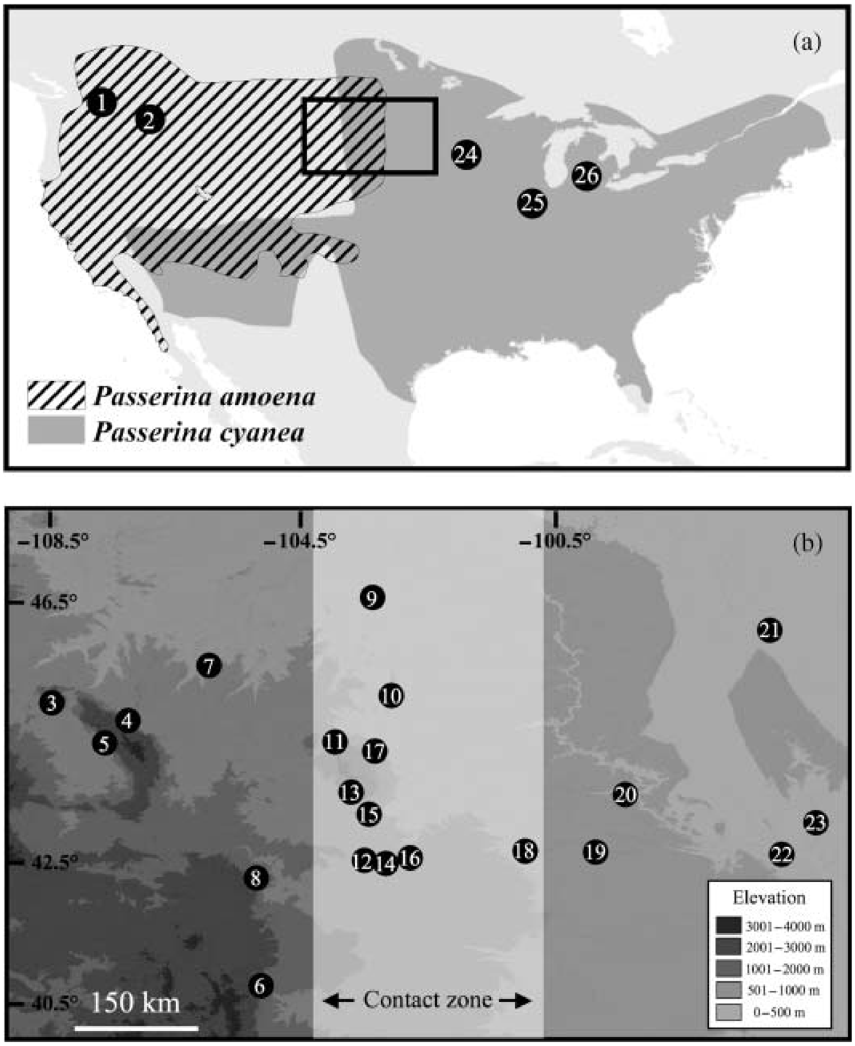

Fig. 1 (a) Principal breeding distributions and allopatric sampling localities for *Passerina amoena* and *P. cyanea*. (b) Sampling localities in and near the contact zone. White box indicates approximate location of the contact zone and dark background shading represents elevation in the region. Digital maps (Ridgely *et al*. 2003) for each species were downloaded from NatureServe (2006) and modified. Numbers of sampling localities refer to more detailed location information in Table 1.

acquired samples from allopatric populations of *P. cyanea* and *P. amoena* east and west of the contact zone, respectively (Fig. 1, Table 1 and Table S1, Supporting information), for inclusion as reference parental populations, resulting in 222 individuals collected from 26 populations. All of these individuals were included in previous research on the *Passerina* hybrid zone (Carling & Brumfield 2008a). Voucher specimens for these individuals are deposited in the Louisiana State University Museum of Natural Science.

We extracted genomic DNA from ~25 mg of pectoral muscle from all sampled individuals using either standard phenol/chloroform methods or a DNeasy Tissue Kit (QIAGEN Inc., Valencia, CA), following the manufacturer's recommend protocol. Each individual was amplified at 10 z-linked loci (Table 2), two of which (*ALDOB3* and *BRM15*) were examined in a previous study of the hybrid zone (Carling & Brumfield 2008a). We found no heterozygous sites in any females. We also checked the genomic location of all loci by BLASTing *Passerina* sequence data against the zebra finch (*Taeniopygia guttata*) genome (Assembly 3.2.4; http://www.genome.wustl.edu) and the chicken (*Gallus gallus*) genome (build 2.1; International Chicken Genome Sequencing Consortium 2004).

To amplify loci, we used the following polymerase chain reaction (PCR) conditions in a 25-μL reaction: ~40 ng template DNA (2 μL of DNA extracts), 1 μL of 10 mM dNTPs (2.5 mM each of dATP, dTTP, dCTP and dGTP), 1 μL of each primer (10 mM, Table 2), 2.5 μL 10 X Buffer with MgCl$_2$ (15 mM), 0.1 μL Taq (5 U/μL of either AmpliTaq DNA Polymerase, Applied Biosystems Inc., Foster City, CA or Taq DNA Polymerase, New England Biolabs, Ipswich, MA), and 17.4 μL sterile dH$_2$O. The thermocycling profile was as follows: an initial 95° denaturation for 2 mins followed by 35 cycles consisting of a 30-s, 95° denaturation step, a 30-s, locus specific temperature primer annealing step (Table 2), a 2-min, 72° extension step, and a final extension of 5 mins at 72°. To check for amplification we electrophoresed 2.5 μL of each PCR product on a 1% Agarose gel.

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

Code	Locality	Latitude	Longitude	*N*	Distance from Locality 1 (km)
1	WA	47.77	–118.92	29	0
2	ID	44.1	–116.15	3	232
3	MT: Custer National Forest #1	45.08	–108.5	4	842
4	WY: Bighorn National Forest #1	44.83	–107.3	8	943
5	WY: Bighorn National Forest #2	44.57	–107.67	2	944
6	CO: Roosevelt National Forest	40.67	–105.23	7	946
7	MT: Custer National Forest #2	45.68	–106.03	5	1003
8	WY: Medicine Bow National Forest	42.38	–105.32	3	1072
9	ND: Little Missouri National Grassland	46.77	–103.5	13	1083
10	SD: Custer National Forest	45.3	–103.22	10	1191
11	WY: Sand Creek*	44.53	–104.08	14	1198
12	NE: White River*	42.62	–103.53	8	1203
13	SD: Black Hills National Forest	43.72	–103.82	5	1225
14	NE: Ponderosa State Wildlife Management Area	42.62	–103.32	7	1227
15	SD: The Nature Conservancy Whitney Preserve	43.33	–103.55	3	1243
16	NE: Nebraska National Forest	42.75	–102.92	8	1245
17	SD: Ft. Meade National Recreation Area	44.38	–103.45	4	1250
18	NE: Nenzel*	42.8	–101.1	4	1395
19	NE: The Nature Conservancy Niobrara Valley Preserve	42.8	–100.02	18	1479
20	SD: Carpenter Game Production Area	43.74	–99.53	13	1492
21	ND: The Nature Conservancy Pigeon Point Preserve	46.25	–97.3	4	1557
22	NE: Wiseman State Wildlife Area	42.75	–97.12	10	1670
23	SD: Newton Hills Game Production Area	43.24	–96.57	13	1721
24	MN	44.87	–93.65	10	1861
25	IL	40.48	–88.92	10	2322
26	MI	42.25	–83.72	7	2714

*Access to these localities generously provided by private landowners.

Both strands of all PEG-purified PCR amplicons were cycle-sequenced in a 7-μL reaction using 1.5 μL of 5 × sequencing buffer (ABI), 1 μL of 10 mm primer (Table 2), 2.0 μL of template, 0.15–0.25 μL Big Dye Terminator Cycle-Sequencing Kit version 3.1 (ABI), and 2.25–2.35 μL sterile dH_2O. We cleaned cycle-sequencing products on Sephadex (G-50 fine) columns and electrophoresed the cleaned products on a 3100 Genetic Analyzer (ABI). All sequences were edited and assembled using Sequencher version 4.7 (GeneCodes Corp., Ann Arbor, MI). When direct sequencing of purified PCR amplicons revealed more than one heterozygous site within a sequence, we resolved haplotypes probabilistically using PHASE (Stephens *et al*. 2001; Stephens & Donnelly 2003).

To identify the largest independently segregating block of sequence data for each locus, we used a panel of individuals from outside the contact zone (populations 1–2 and 24–26; Fig. 1) to test for intralocus recombination using the four-gamete test (Hudson & Kaplan 1985) as implemented in DnaSP version 4.10 (Rozas *et al*. 2003). To assign within locus haplotypes as belonging to either *P. cyanea* or *P. amoena*, we used TCS version 1.21 (Clement *et al*. 2000) to build a parsimony-based haplotype network. Individual haplotypes with frequencies greater than 0.80 in the 'allopatric' *P. amoena* population (WA) were classified as *P. amoena* haplotypes. Alternately, haplotypes with frequencies greater than 0.80 in the 'allopatric' *P. cyanea* population were designated as *P. cyanea* haplotypes. For two loci, *GPBP1* and *RIOK2*, phased haplotypes could not be assigned to *P. cyanea* or *P. amoena* using the above criterion; thus, in the cline analyses (see below), we estimated shape parameters for the most common haplotype only. All statistical analyses were performed using JMP version 5.0.1.2 (SAS Institute Inc., Cary, NC).

One general caveat to this study stems from the lack of a linkage map for *Passerina* buntings. In all analyses for which we investigated the relationship between introgression and genomic location of each locus, we replicated the analyses assuming the genomic location in *Passerina* buntings is the location in the zebra-finch genome (Assembly 3.2.4; http://www.genome.wustl.edu) and the location in the chicken genome (build 2.1; International Chicken Genome Sequencing Consortium 2004). While few avian linkage maps are available for comparison, z-chromosome linkage maps have been recently published for *Ficedula* flycatchers (Backstrom *et al*. 2006) and zebra finches (Itoh *et al*. 2006).

Table 2 PCR conditions, sequence lengths, primer sequences, and GenBank Accession Nos for z-linked loci

Locus	Annealing temp	Length (bp)*	PCR/sequencing primers†	Internal sequencing primers‡	GenBank Accession Nos.
24555	54	197	F: `CCT CCA GAT ATT TCA TTC CC` R: `AAT GGA AAT GGC TGA ACT TG`	NA R: `ATC ACA CGT AAT CTT GCC CAG GAC`	FJ539556–759
PTCH6	49	73	F: `CCA TTT TCT TCC AAG CAA TA` R: `TTT CTT GAC AGT CCA TAG CA`	F: `CCA ACA TGG TCT TTA ACA TCC AT` NA	FJ540719–902
ALDOB3	61	111	F: `GGC AGG AAC AAA TGG AGA AAC T` R: `GCC AGA ACC TGA AAA CAG GAG`	F: `CAC TTG GCA AAG TCA GCA CCA TCT` NA	EU858008–237
RIOK2	49	69	F: `ATG GGT GTT GGC AAA GAA TC` R: `GCT CCT CTT CRT TWG CAA CAA T`	F: `TGG AAC GCT ATC TTG GTT CYT GCT` R: `TCC TTC CTG AGC AAT RGA CAG ACA`	FJ540368–520
GPBP1	54	110	F: `CTT TTG TGG ACG GAG AAT CG` R: `ATT TCT GCC TTG TGA ACG CC`	NA NA	FJ539760–961
PPWD1	56	85	F: `AAC TGT GGA AAA CTT CTG TG` R: `TCA TCT TCA AAT TCT CCT CC`	NA NA	FJ540160–367
IQGAP2	58	70	F: `TCG ATT CAG AGT TTG CGA GC` R: `TCA TGG AGT GAG CTC TTC AG`	F: `GGA GTG AGC TCT TCA GAA CTT TGG CT` R: `TGG CAA ACC TCT CCC TGT ACT TGA`	FJ539962–40159
VLDLR9	60	240	F: `AAG TGT GAA TGT AGC GCT GG` R: `TCG GTT GGT GAA AAT CAG AC`	NA NA	FJ540521–718
BRM15	60	300	F: `AGC ACC TTT GAA CAG TGG TT` R: `TAC TTT ATG GAG ACG ACG GA`	NA NA	EU858238–463
24105	58	199	F: `CCY AAA GGA GCT GTA CCA GC` R: `CTC TTT CCT GTA CGA ATC AC`	F: `ACC GCA TAT GCT CAC ATT GTC C` R: `AAA GGA GCT GTA CCA GCA TAC CTG C`	FJ539348–555

*Length of longest independently segregating block.

†PCR and external sequencing primer references: *GPBP1*, *PPWD1*, *24105*, *IQGAP2*, *24555* — Backstrom *et al.* 2006; *BRM15*, *VLDLR9*, *PTCH6* — Borge *et al.* 2005, *RIOK2* — A. Brelsford personal communication; *ALDOB3* — G. P. Saetre personal communication.

‡Internal sequencing primer references: *24105*, *PTCH6*, *24555*, *RIOK2*, *IQGAP2* — this study; *ALDOB3* — Carling & Brumfield 2008a.

Both, *Ficedula* flycatchers and zebra finches, are members of the avian order Passeriformes, which also includes *Passerina* buntings, and comparisons between these linkage maps and the chicken genome suggest a high degree of gene-content conservation but less conservation of the physical location of z-linked genes. A comparison of the location of the loci, on either the chicken or zebra finch z-chromosome, sampled in this study supported earlier findings that there have been substantial rearrangements of sex-linked loci across bird species (Table 3). While we are confident these loci are located on the z-chromosome in *Passerina* buntings, their exact location remains uncertain.

Cline analyses

Before performing the cline analyses, we excluded those individuals for which PHASE was unable to assign haplotypes to a posterior probability greater than 0.75 (sample sizes given in Table 3). Cline shape parameters were estimated using the same linear transect as described in Carling & Brumfield (2008a). To generate the transect, we plotted the 0.5-isocline using x and y coordinates and average allele frequencies of sampling localities 3–23, which are those populations within the previously described area of overlap between *P. amoena* and *P. cyanea*. This 0.5-isocline was considered the centre of the zone and the shortest straight-line distance between all sampling localities and the 0.5-isocline was measured. 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 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).

The program ClineFit (Porter *et al.* 1997) implements the methods developed by Szymura & Barton (1986, 1991) to investigate introgression patterns along the sampling transect. These methods, which do not consider the species identity of sampled individuals, estimate cline shape parameters using three equations that explore the relationship between the geographical location of the sampling localities and allele frequency data within each locality. The first, eqn 1, describes a symmetrical, sinusoidal cline in the centre of the cline and the other two, eqns 2 and 3, describe the exponential change in allele frequencies on the left and right sides of the centre of the cline:

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

Table 3 Maximum-likelihood cline shape parameters with 2-unit support limits for z-linked loci analysed in this study

Dataset	N*	Loc. (Mb)†	Loc. (Mb)‡	c (km from pop1)	w (km)	θ_L	θ_R	z_L	z_R	p_L	p_R
24555	416	6.7	43.3	1299 (1265–1338)	230 (165–319)	—	—	—	—	—	—
PTCH6	367	9.7	41.2	1293 (1256–1367)	258 (178–441)	0.088 (0.031–0.845)	0.012 (0.007–0.098)	529 (70–998)	998 (269–998)	—	—
ALDOB3	388	11.1	63.9	1353 (1319–1394)	268 (185–346)	0.442 (0.050–0.858)	0.035 (0.015–0.180)	595 (163–994)	692 (44–965)	—	—
RIOK2	163	24.9	50.4	1404 (1284–1822)	584 (418–919)	0.024 (0.010–0.103)	0.010 (0.006–0.032)	1000 (476–1000)	1000 (718–1000)	—	—
GPBP1	411	48.0	17.0	1273 (1241–1355)	260 (157–318)	0.033 (0.017–0.081)	0.352 (0.033–0.790)	871 (394–999)	850 (745–1000)	0.108 (0.085–0.123)	0.488 (0.421–0.562)
PPWD1	391	50.9	19.9	1346 (1290–1420)	224 (140–400)	0.012 (0.004–0.053)	0.005 (0.002–0.017)	999 (437–999)	999 (558–1000)	—	—
IQGAP2	401	57.0	59.6	1307 (1263–1422)	271 (147–436)	0.009 (0.003–0.043)	0.344 (0.030–0.828)	988 (256–998)	881 (91–995)	—	—
VLDLR9	301	64.7	27.3	1227 (1225–1228)	2.8 (1.45–17.5)	0.469 (0.003–0.997)	0.000 (0.000–0.000)	779 (7.4–999)	999 (745–1000)	—	0.71 (0.639–0.772)
BRM15	379	64.9	27.0	1328 (1294–1360)	245 (159–310)	0.172 (0.046–0.780)	0.018 (0.009–0.045)	791 (80–1000)	1000 (570–1000)	—	—
24105	416	67.1	23.6	1306 (1272–1347)	227 (165–314)	—	—	—	—	—	—

*Number of sampled chromosomes in each dataset.
†Location along the zebra finch z-chromosome.
‡Location along the chicken z-chromosome.

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

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

In each equation, c represents the location of the centre of the cline (measured in km from the WA locality), w is the cline width (1/max slope) and x is the geographical location along the sampling transect (km from WA locality). In eqns 2 and 3, parameters z_L and z_R describe the distance from the centre (c) of a vertical asymptote for the exponential decay of allele frequencies on the left and right side of the zone, respectively. The parameters θ_L and θ_R represent the exponential decay values relative to the shape of the central cline (eqn 1) on the left and right sides. The three equations are related in that, as the parameters z and θ approach 0 and 1, eqns 2 and 3 approach the shape of eqn 1 on their respective sides of the zone.

Together, these six parameters can be used to explore introgression patterns of different loci across a hybrid zone. The centre of the zone (c) is the point along the transect at which allele frequencies change most rapidly and the width of the zone (w) provides an estimate of the geographical distance over which that rapid change in allele frequencies occurs. The parameters θ_L and θ_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 geographical distance over which the exponential decay in the tails occurs.

For each of the 10 loci, we tested the fit of three different cline models using likelihood ratio tests: (i) the two-parameter model, which includes only the centre and width parameters; (ii) the six-parameter model, which includes c, w, θ_L, θ_R, z_L and z_R; and (iii) an eight-parameter model that allows for allele frequencies that do not reach fixation (0.0 and 1.0) for alternate alleles on the right and left sides of the cline, respectively (Porter *et al.* 1997; Brumfield *et al.* 2001). For each locus, we used the most appropriate model to estimate the final cline shape parameters (Table 3). The following search parameters were used in each analysis: burn-in, parameter tries per step ≥ 50X the number of parameters being estimated (2, 6 or 8); sampling for support, replicates saved ≥ 2000, and 50 replicates between saves (Porter *et al.* 1997).

Differences in parameter estimates between different loci were assessed using the two-unit support limits, which are analogous to 95% confidence limits (Edwards 1992), and which provide a very conservative test of statistically significant differences. In this test, if the two-unit support limits for the width estimates from two different loci do not overlap, the difference in the widths is statistically significant (Porter *et al.* 1997). The two-unit support limits account for sampling error within each population sample and stochastic fluctuations in allele frequencies along the sampling transect (A. Porter, personal communication), both of which contribute to differences in parameter estimates among loci.

Coalescent analyses

We estimated divergence times for each locus using the coalescent-based isolation-with-migration model as implemented in the software program IM (Hey & Nielsen 2004; Hey 2005). Through the combination of coalescent theory and Bayesian methodologies, IM simultaneously estimates multiple population genetic parameters for two diverging populations. These parameters (scaled to the neutral mutation rate, μ) 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 estimates of different, asymmetric introgression. Because IM assumes no intralocus recombination and requires fully resolved haplotypes (i.e. no ambiguous sites) with no gaps, we culled the datasets. First, we removed all gaps in the alignments. Second, as in the cline analyses, we included only those individuals for which PHASE (Stephens *et al.* 2001; Stephens & Donnelly 2003) was able to assign haplotypes with a probability of greater than 0.75 (Table 4). Last, we only included sequence data from the largest independently segregating block after testing for evidence of recombination as described above.

To estimate divergence times between *P. cyanea* and *P. amoena* for each locus, we analysed datasets containing 21–30 sequences (Table 4) of *P. cyanea* and *P. amoena* from allopatric populations (IL and MI for *P. cyanea*, WA for *P. amoena*). Initially, we estimated one standardized mutation rate (we forced $\theta_1 = \theta_2 = \theta_A$), divergence time ($t$) and asymmetric introgression rates (m_1 and m_2) for each locus. We first ran IM using the Hasegawa Kishino Yano (HKY) finite-sites-substitution model (Hasegawa *et al.* 1985) with wide, uninformative priors and an inheritance scalar of 0.75 (since all loci are sex-linked) for more than 50 million steps. We used these initial runs to identify more appropriate priors (Won & Hey 2005). The adjusted priors were then used in two replicate 'final' analyses that differed only in starting random-number seed. For those datasets in which both introgression parameters (m_1 and m_2) peaked at zero, we only estimated standardized mutation rate (θ) and divergence time (t) in subsequent runs (Table 4). All 'final' analyses were run with a burn-in of 100 000 steps and were allowed to continue until the effective sample size (ESS) values for each parameter were greater than 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.

We also estimated divergence times for a subset of loci using the same procedure as outlined above with the

Table 4 Maximum-likelihood estimates (90% highest posterior density (HPD) intervals) of locus-specific divergence times, θ, and introgression rates

Locus	Loc. (Mb)†	Loc. (Mb)‡	Samples sizes*		Divergence time ($t = t\mu$)§	$\theta = 4N\mu$	$m_1 = m_1/\mu$¶	$m_2 = m_2/\mu$¶
			Passerina cyanea	*P. amoena*				
24555	6.7	43.3	25	21	1.055 (0.365–2.595)	1.488 (0.651–2.953)	—	—
PTCH6	9.7	41.2	24	26	0.595 (0.305–9.995?)	0.2813 (0.067–0.8797)	0.945 (0.015–6.335)	0.005 (0.005–4.155)
ALDOB3	11.1	63.9	27	30	1.025 (0.365–2.595)	1.323 (0.603–2.659)	—	—
RIOK2	24.9	50.4	25	27	0.215 (0.075–9.995?)	0.349 (0.096–1.019)	7.35 (1.95–19.05)	0.005 (0.005–15.19?)
GPBP1	48.0	17.0	30	30	0.29 (0.090–9.995?)	0.820 (0.299–2.017)	6.225 (0.025–37.675?)	0.005 (0.005–7.265)
PPWD1	50.9	19.9	30	27	0.655 (0.385–9.995?)	0.167 (0.036–0.9755)	4.325 (1.695–9.865)	0.005 (0.005–7.265)
IQGAP2	57.0	59.6	27	30	0.775 (0.355–9.995?)	0.963 (0.356–1.969)	0.005 (0.005–1.595)	0.415 (0.005–2.755)
VLDLR9	64.7	27.3	30	30	0.265 (0.075–9.995?)	1.238 (0.428–2.214)	0.015 (0.005–8.345?)	0.005 (0.005–7.065)
BRM15	64.9	27.0	23	29	1.305 (0.485–3.015)	1.759 (0.852–3.331)	—	—
24105	67.1	23.6	28	23	1.615 (0.615–3.435)	3.625 (2.052–6.124)	—	—

*Number of chromosomes sampled from allopatric populations (IL, MI for *P. cyanea*, WA for *P. amoena*).

†Location along the zebra finch z-chromosome.

‡Location along the chicken z-chromosome.

§'?' indicates a plateau in the posterior probability distribution.

¶m_1 — introgression from *P. amoena* into *P. cyanea*; m_2 — introgression from *P. cyanea* into *P. amoena*; estimates of 0.005 are effectively 0.

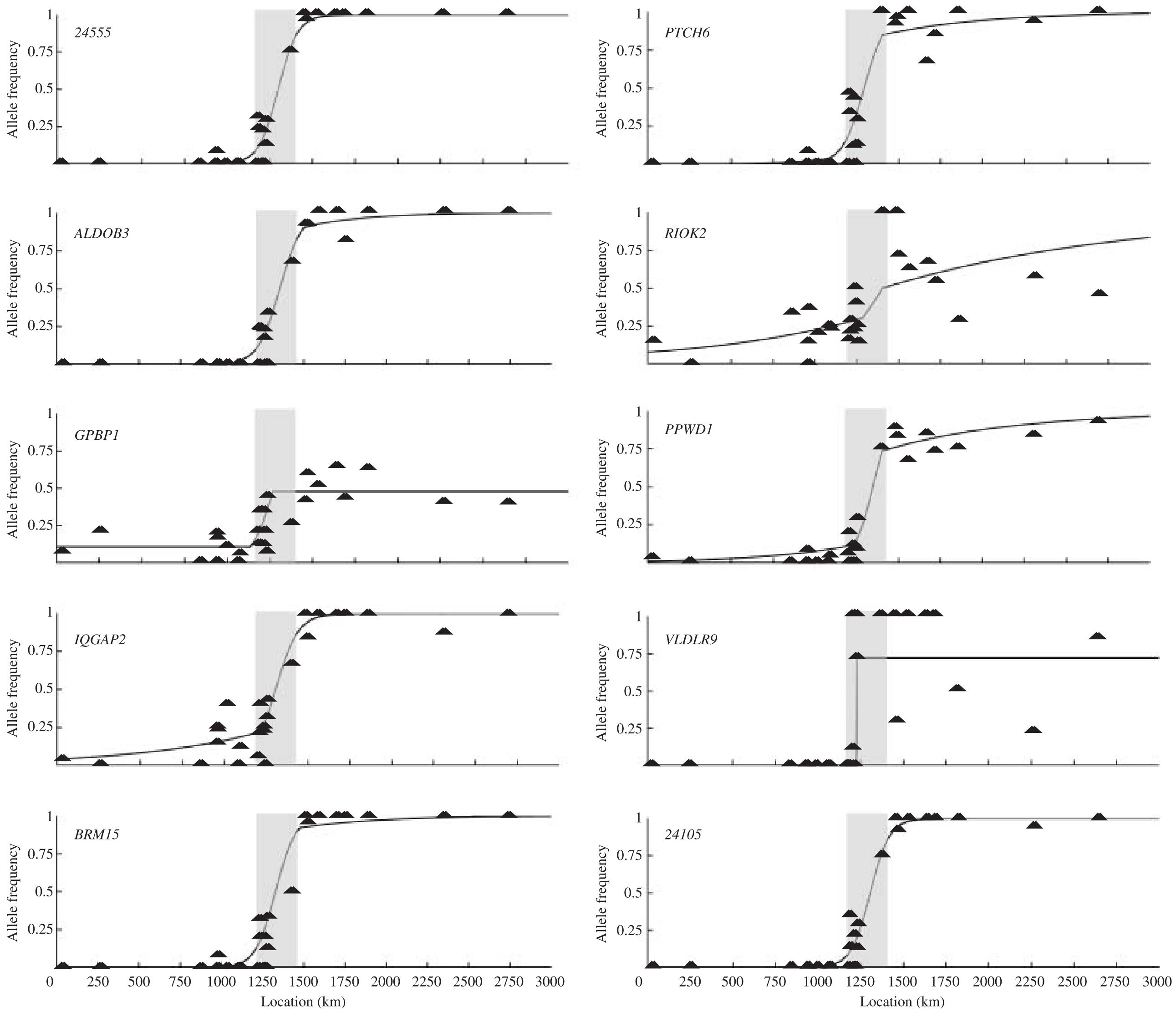

Fig. 2 Locus-specific maximum-likelihood cline shapes (line) and allele frequencies (triangles) plotted against geographical location of sampling locality along transect. Loci presented in order of their location on zebra finch z-chromosome. Grey box indicates the approximate location of the contact zone. Cline shape parameter estimates are provided in Table 3.

exception of analyzing the data using the infinite-sites substitution model (Kimura 1969) in place of the HKY finite-sites model.

Results

The contact zone between these species, defined as the geographical area bounded by the western most locality in which *Passerina cyanea* was collected and the eastern most locality in which *P. amoena* was collected (both identified using external plumage patterns), is approximately 250 km wide and extends from Population 11 to Population 18 (Fig. 1; Carling & Brumfield 2008a). We found no genetically pure individuals of the 'wrong' type (e.g. an individual with all *P. amoena* alleles but possessing a predominantly *P. cyanea* plumage) outside of the contact zone.

Cline shape estimates

Eight of the ten locus-specific clines were concordant, with widths between 224 and 271 km (Fig. 2 and Table 3). Of the other two loci, one, *RIOK2*, had a much greater width (584 km; 2-unit support limits: 418–919 km), although the difference was not significant (Table 3). The other, *VLDLR9*, had a significantly narrower width (2.8 km; 2-unit support limits: 1.45–17.5 km). There was no relationship between cline width and the genomic location of the loci on either the chicken- or zebra-finch z-chromosome (both ANOVAS

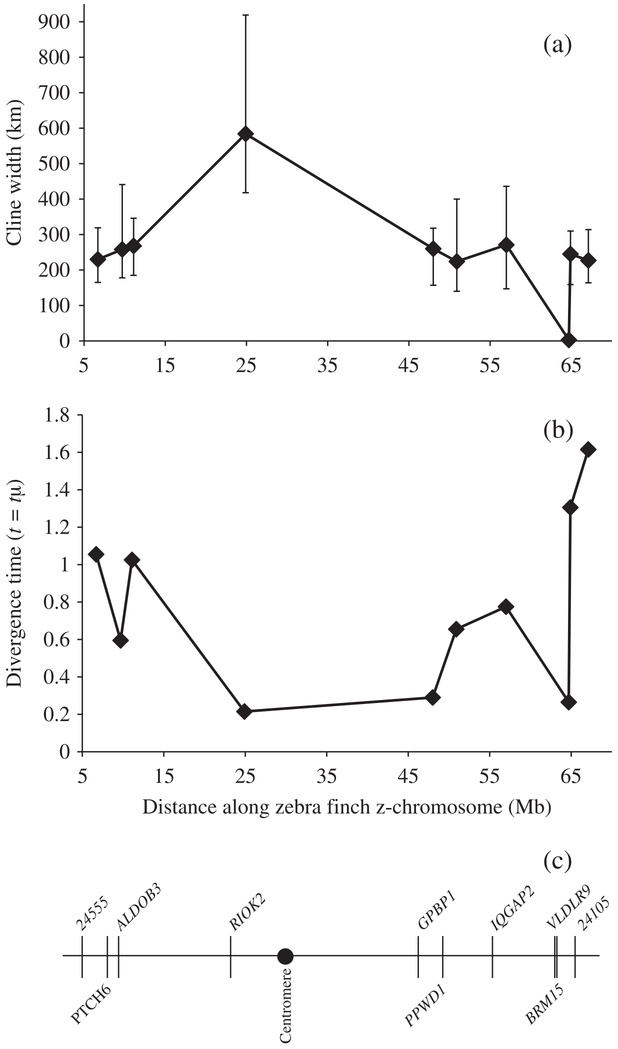

Fig. 3 (a) Scatterplot of cline widths (km) vs. genomic location of loci (Mb) on the zebra finch z-chromosome. Error bars indicate 2-unit likelihood support limits, (b) Divergence time ($t = t\mu$) vs. genomic location. Note that confidence intervals of all estimates overlapped. Genomic location did not correlate with cline width or divergence time. (c) Schematic of genomic location of loci and the centromere along the zebra finch z-chromosome.

$P > 0.1$; Fig. 3a). The coefficient of variation of the width estimates was 0.58.

The centres of all clines, with the exception of *VLDLR9*, were coincident and fell within a region 131 km wide (Fig. 2 and Table 3). The wide support limits of the *RIOK2* centre estimate (1284–1822 km east of Population 1) were likely the result of reduced sample size of that locus compared to the others (Table 3). The centre of the *VLDLR9* cline was shifted significantly to the west (1227 km east of Population 1; 2-unit support limits: 1225–1228). This may be because of substantial variation in allele frequencies among localities on the *P. cyanea* side of the zone (Fig. 2). Overall, the large degree of overlap among the 2-unit support limits of cline centre indicates the cline centres are generally coincident across the different loci. There was a significant positive relationship between cline centre and width ($y = -2677 + 2.22x$, ANOVA $P = 0.012$) although, when either *VLDLR9* (centre located farthest west) or *RIOK2* (farthest east) was excluded, the relationship lost significance. The estimates of the cline centres were less variable than the estimates of cline width (c.v. = 0.03).

Estimates of the rate of exponential decay of allele frequencies in the eastern and western tails of the clines (θ_L and θ_R) provided further insight into introgression patterns along the geographical transect. The estimate of θ_L was greater than the estimate of θ_R in six of the eight datasets (Table 3), suggesting that introgression from *P. amoena* into *P. cyanea* is greater, on average, than in the opposite direction (Sign test, $n_{obs} = 8$, $P = 0.29$). Support for unidirectional introgression was also found in the coalescent-based estimates (see below).

Because samples sizes at some localities were small (Table 1), we also estimated cline shape parameters for a subset of loci, including *VLDLR9*, after excluding localities containing fewer than eight sampled alleles. There were no qualitative differences in the parameter estimates between the full and reduced datasets.

Divergence time estimates

Two independent IM analyses produced similar results, so we present only those results from the longest run (number of steps $> 3 \times 10^8$). There was considerable inter-locus variation in the estimates of divergence times and the 90% highest posterior densities for all estimates overlapped (Table 4). The most recent divergence time was estimated for *RIOK2* (0.215) and the oldest divergence time for *24105* (1.615). When the divergence times were compared with the genomic location of the loci, multiple peaks and valleys were found (Fig. 3b), with no significant relationship between divergence time and chromosomal location in either the zebra-finch or chicken genome (both ANOVAS $P > 0.1$).

In the initial IM analyses for four loci (*24105*, *24555*, *ALDOB3*, *BRM15*), both introgression estimates peaked at zero, so the two migration parameters (m_1 and m_2) were removed from the model for those loci (Table 4). In five of the remaining six loci, introgression from *P. amoena* into *P. cyanea* (m_1 in Table 4) was higher than in the opposite direction (m_2). Although this difference was not significant (Sign test, $n_{obs} = 6$, $P > 0.2$), it may suggest a pattern of greater introgression from *P. amoena* into *P. cyanea* than in the other direction. This was consistent with the asymmetric introgression pattern found in the cline analyses (Table 4).

If loci with older divergence times have narrower cline widths, there should be an inverse relationship between

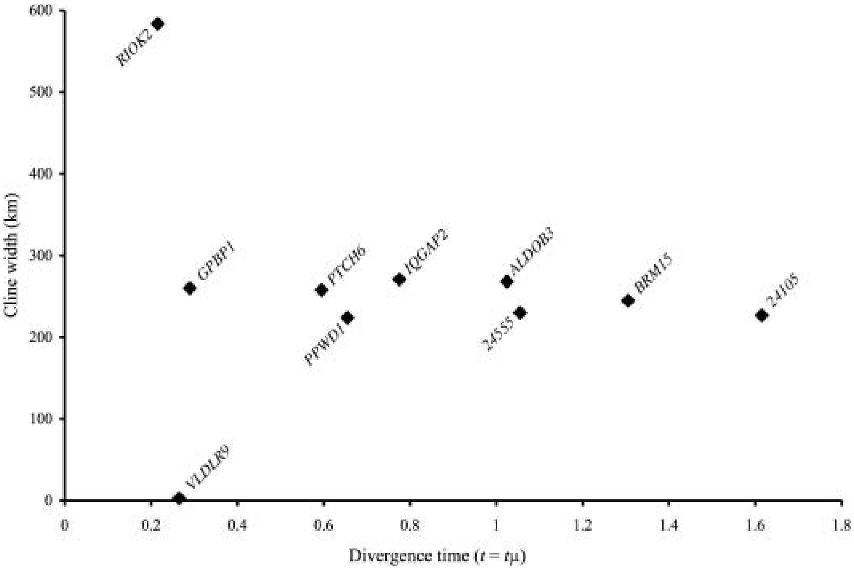

Fig. 4 Scatterplot of divergence time, estimated using IM, and cline width, estimated using ClineFit.

divergence time and width. We found no evidence for such a relationship (ANOVA $P > 0.8$; Fig. 4).

Divergence time estimates made using the HKY and infinite-sites substitution models were statistically correlated ($r^2 = 0.965$, $P = 0.0028$).

Discussion

This first detailed investigation of sex-linked introgression patterns across an avian hybrid zone identified a candidate region on the z-chromosome that may be involved in the reproductive isolation of *Passerina cyanea* and *P. amoena*. One locus in particular, *VLDLR9*, had an extremely narrow cline width (Table 3), suggestive of strong selective pressures that could be impeding the introgression of alleles at this locus.

We also tested the hypothesis proposed by Putnam *et al.* (2007) that loci with older divergence times should introgress less than loci with more recent divergence times. The confidence limits on the divergence estimates were large, but we found no evidence suggesting a negative relationship between divergence time and cline width (Fig. 4). Assuming divergence began in allopatry, the predominant mode of speciation in avian systems (Mayr 1942, 1963; Coyne & Price 2000; Price 2008), early genetic differences between *P. cyanea* and *P. amoena* were likely the result of genetic drift in isolation and these changes may result in older divergence times. However, if these early changes were largely neutral, there is no a priori reason to expect narrow clines of loci with older divergence times upon secondary contact. In contrast, adaptive evolution has been shown to be important in reproductive isolation in some systems (Ting *et al.* 1998; Barbash *et al.* 2003; Presgraves *et al.* 2003). If loci involved in reproductive isolation are the targets of diversifying selection between species, they may have relatively young divergence times but narrow cline widths. A strong selective sweep could result in a recent coalescent time for a tightly linked neutral allele as well. Such diversifying selection would be unlikely to produce a negative relationship between divergence time and cline width.

To our knowledge, this is only the second study that has explored the relationship between cline width and divergence time, and neither study found a negative association between divergence time and cline width (Putnam *et al.* 2007). Until there are more studies with which to compare our results, it is difficult to know whether there is a general correlation between timing of divergence and introgression rates.

Candidate region for reproductive isolation

Given the theoretical relationship between cline width and selection (Endler 1977; Barton & Gale 1993), estimates of cline width are the most direct assessment of the level of introgression between *P. cyanea* and *P. amoena* and offer the greatest ability to detect loci contributing to reproductive isolation. If alleles at a particular locus are involved in D–M incompatibilities between hybridizing taxa, selection acting to limit introgression of those alleles will result in a narrow cline at that locus. In contrast, selection preventing introgression of alleles at loci not involved in genetic incompatibilities will be less and alleles at those loci will show broader cline widths.

The locus with the narrowest cline width (*VLDLR9*) is located approximately 64 Mb from the end of the zebra finch z-chromosome (Fig. 3c). It is certainly possible that *VLDLR9* is not directly involved in reproductive isolation but is linked, physically or epistatically, to other speciation gene(s). There are ~100 genes known to occur within a 10-Mb region centred around *VLDLR* in the chicken genome (the current draft version of zebra finch genome is not annotated) (International Chicken Genome Sequencing Consortium 2004), which represents ~14% of the chicken z-chromosome. Unfortunately, the functions of many of these genes are unknown, limiting our ability to speculate which genes offer the most potential for future research, but there is one interesting candidate.

The *VLDLR9* locus we examined is an intron in the very low-density lipoprotein receptor (*VLDLR*) gene. In chickens, the *VLDLR* gene is expressed almost exclusively in oocytes, where it functions in the uptake of circulating yolk precursor macromolecules (Bujo *et al*. 1994; Ocon-Grove *et al*. 2007). In certain females of the white leghorn chicken strain, a single nucleotide mutation in *VLDLR* causes these females to produce a less than completely functional protein on the oocyte membrane, which in turn generally results in the failure to lay eggs (Ocon-Grove *et al*. 2007). There is the potential that the divergent *P. cyanea* and *P. amoena VLDLR* alleles do not function as well in a heterospecific genomic background, which could cause hybrid females to lay fewer eggs. Although there is no direct evidence that hybrid females produce fewer eggs, field studies suggest the possibility. In their field study of mating associations between *P. cyanea* and *P. amoena*, Baker & Boylan (1999) tracked the reproductive success of mated pairs in the contact zone over the course of four years. They found that when the female in a pair was a hybrid, the pair had a lower chance of producing eggs than if the female was not a hybrid. The difference was not significant and the sample sizes were small (only six pairings involved a hybrid female), but the results point to hybrid female buntings having greater difficulty producing eggs than pure *P. cyanea* or *P. amoena* females.

Interestingly, in a comparison of polymorphism and divergence between autosomal and sex-linked loci between *Ficedula albicollis* and *F. hypoleuca*, two flycatchers that hybridize in Europe, results from the *VLDLR* gene are consistent with a selective sweep at that locus in *F. hypoleuca* (Borge *et al*. 2005). A selective sweep at *VLDLR* in *Passerina* buntings could help explain the observation that allopatric *P. cyanea* populations harboured *P. amoena* alleles (see below and Fig. 2).

We caution that, while the combination of data from chickens and *Ficedula* flycatchers with the reduced introgression of *VLDLR9* in our study suggest a link between *VLDLR* and reproductive isolation between *P. cyanea* and *P. amoena*, our data are insufficient to adequately address that possibility directly. Further exploration of the gene region surrounding *VLDLR* is required.

Interestingly, allopatric populations of *P. cyanea* and *P. amoena* were not fixed for alternate alleles at the *VLDLR9* locus (Fig. 2), which may indicate a recent origin of the *P. cyanea* allele. In this case, the derived *P. cyanea* allele could be incompatible with a *P. amoena* allele at an unidentified gene. Under this scenario, individuals with a *P. cyanea VLDLR9* allele in a *P. amoena* genetic background would suffer reduced fitness. In contrast, the general patterns of asymmetric introgression suggest that fitness may be more influenced by whether individuals possess *P. amoena* alleles in a *P. cyanea* genetic background (see below). Alternately, since the *VLDLR9* cline was estimated from intronic sequence data, any alleles in the coding regions of the gene may not show the exact same pattern. Presently, our data cannot address these different possibilities.

Asymmetric introgression

Our cline-based analyses indicated that six of eight loci had introgression tails that extended further into *P. cyanea* than into *P. amoena* (Table 3), suggestive of asymmetric introgression. This finding was also supported by the coalescent-based analyses of divergence time and introgression. In IM analyses of the locus-specific datasets for which we included introgression parameters (Table 4), five of six estimates of introgression into *P. cyanea* from *P. amoena* were greater than into *P. amoena* from *P. cyanea*, although the confidence intervals of all estimates overlapped. Furthermore, there was a significant positive correlation between the location of the cline centre and width. This general pattern of asymmetric introgression of multiple sex-linked loci is similar to the pattern seen across the *Mus domesticus*/*Mus musculus* hybrid zone in central Europe (Payseur *et al*. 2004).

Such asymmetry indicates that the patterns of introgression across the *Passerina* hybrid zone may be more strongly influenced by selection against *P. cyanea* alleles in a *P. amoena* genomic background than by selection against *P. amoena* alleles in a *P. cyanea* genomic background. In this scenario, individuals carrying some z-linked *P. cyanea* alleles in a primarily *P. amoena* genome suffer greater reductions in fitness. Interestingly, this pattern was also observed in a previous study. Carling & Brumfield (2008a) showed that introgression of both mtDNA haplotypes and autosomal alleles was more symmetrical than z-linked alleles across this hybrid zone.

There are several mechanisms that could produce asymmetric introgression patterns at the z-linked loci. First, parental *P. amoena* males could be dispersing east into the breeding range of *P. cyanea* and mating with parental *P. cyanea* females. If the F_1 males return in the next breeding season and backcross with pure *P. cyanea* females, *P. amoena* alleles would be incorporated into the *P. cyanea* genetic background without the incorporation of *P. amoena* mtDNA haplotypes. This seems unlikely because *P. amoena* males are rarely seen east of the contact zone; we did not detect any *P. amoena*-plumaged males east of Population 18. Furthermore, in birds, the general pattern of natal dispersal is that females disperse much greater distances than males

(Greenwood 1980; Greenwood & Harvey 1982). This raises the question that if *P. amoena* females are moving east, why are there no *P. amoena* mtDNA haplotypes found east of the contact zone?

A second hypothesis that would address this question is that the entire hybrid zone is actually moving west, and the *P. amoena* z-linked alleles found on the eastern side of the zone have been left behind in the 'wake' of the moving hybrid zone (Rohwer *et al*. 2001; Secondi *et al*. 2006). The long tail of *P. amoena* autosomal alleles extending east is also consistent with the moving-hybrid-zone hypothesis (Carling & Brumfield 2008a). That said, a long tail of *P. cyanea* autosomal alleles extending west was also found (Fig. 2 in Carling & Brumfield 2008a), such that *P. cyanea* alleles are found much further west than the advancing front if the hybrid zone were moving west.

While we cannot completely reject either the long-distance dispersal of *P. amoena* males or the moving hybrid zone hypotheses, the available data are more consistent with a hypothesis wherein introgression is primarily determined by the reduced fitness of individuals possessing *P. amoena* z-linked alleles in a *P. cyanea* genetic background. As more studies investigating differential introgression of multiple sex-linked loci are completed, it may be found that asymmetry is fairly common as it has now been documented in *Passerina* buntings (this study), *Mus* (Payseur *et al*. 2004) and European rabbits (*Oryctolagus cuniculus*; Geraldes *et al*. 2006).

Conclusions

In our analyses of the introgression of z-linked loci across the *Passerina* hybrid zone, we found evidence of a severe reduction in introgression of one locus, highlighting the potential of z-chromosome as a source of genes involved in the maintenance of reproductive isolation. The discovery of a candidate region for reproductive isolation contributes to a growing body of literature on the role the sex-chromosomes play in speciation.

Whether the patterns we found are part of a larger pattern in avian speciation genetics remains to be seen, but our results suggest that many of the evolutionary forces shaping the speciation genetics of model organisms may also be important in the formation and maintenance of reproductive isolation between closely related bird species.

Acknowledgements

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. S. Bensch, C. Burney, Z. Cheviron, A. Cuervo, E. Derryberry, J. Maley and three anonymous reviewers 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. All individuals were collected in accordance with Louisiana State University Institutional Animal Care and Use Committee Protocol 08-025. The map data were provided by NatureServe in collaboration with Robert Ridgely, James Zook, The Nature Conservancy — Migratory Bird Program, Conservation International — CABS, World Wildlife Fund—US, and Environment Canada —WILDSPACE. Zebra finch sequence data used as reference for genomic locations of all loci were produced by the Genome Sequencing Center at Washington University School of Medicine in St. Louis and can be obtained from ftp://genome.wustl.edu/pub/organism/Other_Vertebrates/Taeniopygia_guttata/assembly/Taeniopygia_guttata-3.2.4/. This work was supported by grants from NSF to RTB (DEB-0543562 and DBI-0400797) and MDC (DEB-0808464), as well as by grants from the AMNH Chapman Fund, AOU, Explorer's Club, LSUMNS Birdathon and Prepathon Funds, LSU Dept. of Biological Sciences, and Sigma-Xi.

References

Arnold ML (1997) *Natural Hybridization and Evolution*. Oxford University Press, New York.

Backstrom N, Brandstrom M, Gustafsson L *et al*. (2006) Genetic mapping in a natural population of collared flycatchers (*Ficedula albicolis*): conserved synteny but gene order rearrangements on the avian Z-chromosome. *Genetics*, **174**, 377–386.

Baker MC (1991) Response of male Indigo and Lazuli buntings and their hybrids to song playback in allopatric and sympatric populations. *Behaviour*, **119**, 225–242.

Baker MC (1996) Female buntings from hybridizing populations prefer conspecific males. *Wilson Bulletin*, **108**, 771–775.

Baker MC, Baker AEM (1990) Reproductive-behavior of female buntings — isolating mechanisms in a hybridizing pair of species. *Evolution*, **44**, 332–338.

Baker MC, Boylan JT (1999) Singing behavior, mating associations and reproductive success in a population of hybridizing Lazuli and Indigo buntings. *Condor*, **101**, 493–504.

Baker MC, Johnson MS (1998) Allozymic and morphometric comparisons among Indigo and Lazuli buntings and their hybrids. *Auk*, **115**, 537–542.

Barbash DA, Siino DF, Tarone M, Roote J (2003) A rapidly evolving MYB-related protein causes species isolation in Drosophila. *Proceedings of the National Academy of Sciences, USA*, **100**, 5302–5307.

Barton NH (1979) Gene flow past a cline. *Heredity*, **43**, 333–339.

Barton NH, Gale KS (1993) Genetic analysis of hybrid zones. In: *Hybrid Zones and the Evolutionary Process* (ed. Harrison R). Oxford University Press, New York.

Barton NH, Hewitt GM (1985) Analysis of hybrid zones. *Annual Review of Ecology and Systematics*, **16**, 113–148.

Barton NH, Hewitt GM (1989) Adaptation, speciation and hybrid zones. *Nature*, **341**, 497–503.

Borge T, Webster MT, Andersson G, Sætre GP (2005) Contrasting patterns of polymorphism and divergence on the Z chromosome and autosomes in two *Ficedula* flycatcher species. *Genetics*, **171**, 1861–1873.

Brumfield RT (2005) Mitochondrial variation in Bolivian populations of the Variable Antshrike (*Thamnophilus caerulescens*). *Auk*, **122**, 414–432.

Brumfield RT, Jernigan RW, McDonald DB, Braun MJ (2001) Evolutionary implications of divergent clines in an avian (*Manacus*: Aves) hybrid zone. *Evolution*, **55**, 2070–2087.

Bujo H, Hermann M, Kaderli MO *et al*. (1994) Chicken oocyte growth is mediated by an eight ligand binding repeat member of the LDL receptor family. *The EMBO Journal*, **13**, 5165–5175.

Carling MD, Brumfield RT (2008a) 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. *Evolution*, **62**, 2600–2615.

Carling MD, Brumfield RT (2008b) Integrating phylogenetic and population genetic analyses of multiple loci to test species divergence hypotheses in *Passerina* buntings. *Genetics*, **178**, 363–377.

Charlesworth B, Coyne JA, Barton N (1987) The relative rates of evolution of sex chromosomes and autosomes. *American Naturalist*, **130**, 113–146.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1660.

Coyne JA, Orr HA (1989) Two rules of speciation. In: *Speciation and Its Consequences* (eds Otte D, Endler JA), pp. 180–207. Sinauer Associates, Sunderland, MA.

Coyne JA, Orr HA (2004) *Speciation*. Sinauer Associates, Inc., Sunderland, MA.

Coyne JA, Price TD (2000) Little evidence for sympatric speciation in island birds. *Evolution*, **54**, 2166–2171.

Dobzhansky T (1937) *Genetics and the Origin of Species*. Columbia University Press, New York.

Edwards A (1992) *Likelihood*. Johns Hopkins Univ. Press, Baltimore, MD.

Emlen ST, Rising JD, Thompson WL (1975) Behavioral and morphological study of sympatry in Indigo and Lazuli buntings of Great-Plains. *Wilson Bulletin*, **87**, 145–179.

Endler JA (1977) *Geographic Variation, Speciation, and Clines*. Princeton University Press, Princeton, NJ.

Fitzpatrick BM (2004) Rates of evolution of hybrid inviability in birds and mammals. *Evolution*, **58**, 1865–1870.

Geraldes A, Ferrand N, Nachman MW (2006) Contrasting patterns of introgression at X-linked loci across the hybrid zone between subspecies of the European rabbit (*Oryctolagus cuniculus*). *Genetics*, **173**, 919–933.

Greenwood PJ (1980) Mating systems, philopatry and dispersal in birds and mammals. *Animal Behaviour*, **28**, 1140–1162.

Greenwood PJ, Harvey PH (1982) The natal and breeding dispersal of birds. *Annual Review of Ecology and Systematics*, **13**, 1–21.

Hagen RH, Scriber M (1989) Sex-linked diapause, color, and allozyme loci in *Papilo glaucus*: linkage analysis and significance in a hybrid zone. *Journal of Heredity*, **80**, 179–185.

Haldane JBS (1922) Sex ratio and unisexual sterility in animal hybrids. *Journal of Genetics*, **12**, 101–109.

Haldane JBS (1948) The theory of a cline. *Journal of Genetics*, **48**, 277–284.

Harrison R (1990) Hybrid zones: windows on the evolutionary process. In: *Oxford Surveys in Evolutionary Biology* (eds Futuyma DJ, Antonovics J), pp. 69–128. Oxford University Press, Oxford.

Hasegawa M, Kishino H, Yano TA (1985) Dating of the human ape splitting by a molecular clock of mitochondrial-DNA. *Journal of Molecular Evolution*, **22**, 160–174.

Hewitt GM (1988) Hybrid zones — natural laboratories for evolutionary studies. *Trends in Ecology and Evolution*, **3**, 158–167.

Hey J (2005) On the number of New World founders: a population genetic portrait of the peopling of the Americas. *Plos Biology*, **3**, 965–975.

Hey J, Nielsen R (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.

Hey J, Nielsen R (2007) Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proceedings of the National Academy of Sciences, USA*, **104**, 2785–2790.

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

International Chicken Genome Sequencing Consortium (2004) Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature*, **432**, 695–716.

Itoh Y, Kampf K, Arnold AP (2006) Comparison of the chicken and zebra finch Z chromosomes shows evolutionary rearrangements. *Chromosome Research*, **14**, 805–815.

Jiggins CD, Linares RE, Naisbit RE *et al*. (2001) Sex-linked hybrid sterility in a butterfly. *Evolution*, **55**, 1631–1638.

Kimura M (1969) The number of heterozygous nucleotide sites maintained in a finite population due to steady flux of mutations. *Genetics*, **61**, 893–903.

Kroodsma RL (1975) Hybridization in buntings (*Passerina*) in North Dakota and eastern Montana. *Auk*, **92**, 66–80.

Mallet J (2005) Hybridization as an invasion of the genome. *Trends in Ecology and Evolution*, **20**, 229–237.

Mayr E (1942) *Systematics and the Origin of Species*. Columbia University Press, New York.

Mayr E (1963) *Animal Species and Evolution*. Belknap Press, Cambridge.

Muller HJ (1940) Bearing of the Drosophila work on systematics. In: *The New Systematics* (ed. Haldane JBS), pp. 185–268. Clarendon Press, Oxford.

Muller HJ (1942) Isolating mechanisms, evolution, and temperature. *Biological Symposia*, **6**, 71–125.

NatureServe (2006) *NatureServe Explorer: an Online Encyclopedia of Life*, Version 6.1. http://www.natureserve.org/explorer.

Ocon-Grove OM, Maddineni S, Hendricks III GL *et al*. (2007) Pituitary progesterone receptor expression and plasma gonadotropic concentrations in the reproductively dysfunctional mutant restricted ovulator chicken. *Domestic Animal Endocrinology*, **32**, 201–215.

Orr HA (1995) The population genetics of speciation: the evolution of hybrid incompatibilities. *Genetics*, **139**, 1805–1813.

Payseur BA, Krenz JG, Nachman MW (2004) Differential patterns of introgression across the X chromosome in a hybrid zone between two species of house mice. *Evolution*, **58**, 2064–2078.

Porter AH, Wenger R, Geiger H, Scholl A, Shapiro AM (1997) The *Pontia daplidice-edusa* hybrid zone in northwestern Italy. *Evolution*, **51**, 1561–1573.

Presgraves DC (2003) A fine-scale genetic analysis of hybrid incompatibilities in Drosophila. *Genetics*, **163**, 955–972.

Presgraves DC, Balagopalan L, Abymayr SM, Orr HA (2003) Adaptive evolution drives divergence of a hybrid inviability gene between two species of Drosophila. *Nature*, **423**, 715–719.

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

Price TD, Bouvier MM (2002) The evolution of F-1 postzygotic incompatibilities in birds. *Evolution*, **56**, 2083–2089.

Prowell DP (1998) Sex linkage and speciation in Lepidoptera. In: *Endless Forms: Species and Speciation* (eds Howard DJ, Berlocher SH), pp. 309–319. Oxford University Press, New York.

Putnam AS, Scriber JM, Andolfatto P (2007) Discordant divergence times among z-chromosome regions between two ecologically distinct swallowtail butterfly species. *Evolution*, **61**, 912–927.

Ridgely RS, Allnutt TF, Brooks T *et al*. (2003) *Digital Distribution Maps of the Birds of the Western Hemisphere*, Version 1.0. NatureServe, Arlington, VA.

Rieseberg LH, Whitton J, Gardner K (1999) Hybrid zones and the genetic architecture of a barrier to gene flow between two sunflower species. *Genetics*, **152**, 713–727.

Rohwer S, Bermingham E, Wood C (2001) Plumage and mitochondrial DNA haplotype variation across a moving hybrid zone. *Evolution*, **55**, 405–422.

Rosenblum EB, Hickerson MJ, Moritz C (2007) A multilocus perspective on colonization accompanied by selection and gene flow. *Evolution*, **61**, 2971–2985.

Rozas J, Sanchez-De JC, Barrio I, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, **19**, 2496–2497.

Sætre GP, Borge T, Lindroos K *et al*. (2003) Sex chromosome evolution and speciation in *Ficedula* flycatchers. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, **270**, 53–59.

Secondi J, Faivre B, Bensch S (2006) Spreading introgression in the wake of a moving contact zone. *Molecular Ecology*, **15**, 2463–2475.

Sibley CG, Short LLJ (1959) Hybridization in the buntings (*Passerina*) of the Great Plains. *Auk*, **76**, 443–463.

Slatkin M (1973) Gene flow and selection in a cline. *Genetics*, **75**, 733–756.

Slatkin M, Maruyama T (1975) Genetic drift in a cline. *Genetics*, **81**, 209–222.

Stephens M, Donnelly P (2003) A comparison of Bayesian methods for haplotype reconstruction from population genotype data. *American Journal of Human Genetics*, **73**, 1162–1169.

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

Szymura JM, Barton N (1986) Genetic analysis of a hybrid zone between the fire-bellied toads, *Bombina bombina* and *Bombina variegata*, near Cracow in southern Poland. *Evolution*, **40**, 1141–1159.

Szymura JM, Barton NH (1991) The genetic structure of the hybrid zone between the fire-bellied toads *Bombina bombina* and *B. variegata* — comparisons between transects and between loci. *Evolution*, **45**, 237–261.

Tao Y, Chen S, Hartl DL, Laurie CC (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.

Ting CT, Tsaur SC, Wu ML, Wu CI (1998) A rapidly evolving homeobox at the site of a hybrid sterility gene. *Science*, **282**, 1501–1504.

Tucker PK, Sage RD, Warner J, Wilson AC, Eicher EM (1992) Abrupt cline for sex-chromosomes in a hybrid zone between two species of mice. *Evolution*, **46**, 1146–1163.

Wakeley J, Hey J (1997) Estimating ancestral population parameters. *Genetics*, **145**, 847–855.

Won YJ, Hey J (2005) Divergence population genetics of chimpanzees. *Molecular Biology and Evolution*, **22**, 297–307.

Matthew Carling completed his PhD in the Museum of Natural Science and Department of Biological Sciences at Louisiana State University. This work is a portion of his dissertation, which focussed on speciation genetics in birds. Robb Brumfield is an Associate Professor in the Department of Biological Sciences at Louisiana State University with research interests in avian population genetics, speciation, biogeography, systematics, genomics and behaviour.

Supporting information

Additional Supporting information may be found in the online version of this article:

Table S1 Voucher numbers and sampling localities of all individuals