MOLECULAR ECOLOGY

Molecular Ecology (2010) **19**, 3171–3192 doi: 10.1111/j.1365-294X.2010.04727.x

Secondary contact between *Lycaeides idas* and *L. melissa* in the Rocky Mountains: extensive admixture and a patchy hybrid zone

ZACHARIAH GOMPERT,* LAUREN K. LUCAS,† JAMES A. FORDYCE,‡ MATTHEW L. FORISTER§ and CHRIS C. NICE¶

*Department of Botany, Program in Ecology, University of Wyoming, Laramie, WY 82071, USA, †Department of Secondary Education, University of Wyoming, Laramie, WY 82071, USA, ‡Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, TN 37996, USA, §Department of Biology/MS 314, University of Nevada, Reno, NV 89557, USA, ¶Department of Biology, Population and Conservation Biology Program, Texas State University, San Marcos, TX 78666, USA

Abstract

Studies of hybridization have increased our understanding of the nature of species boundaries, the process of speciation, and the effects of hybridization on the evolution of populations and species. In the present study we use genetic and morphological data to determine the outcome and consequences of secondary contact and hybridization between the butterfly species *Lycaeides idas* and *L. melissa* in the Rocky Mountains. Admixture proportions estimated from *structure* and geographical cline analysis indicate *L. idas* and *L. melissa* have hybridized extensively in the Rocky Mountains and that reproductive isolation was insufficient to prevent introgression for much of the genome. Geographical patterns of admixture suggest that hybridization between *L. idas* and *L. melissa* has led to the formation of a hybrid zone. The hybrid zone is relatively wide, given estimates of dispersal for *Lycaeides* butterflies, and does not show strong evidence of cline concordance among characters. We believe the structure of the *Lycaeides* hybrid zone might be best explained by the patchy distribution of *Lycaeides*, local extinction and colonization of habitat patches, environmental variation and weak overall selection against hybrids. We found no evidence that hybridization in the Rocky Mountains has resulted in the formation of independent hybrid species, in contrast to the outcome of hybridization between *L. idas* and *L. melissa* in the Sierra Nevada. Finally, our results suggest that differences in male morphology between *L. idas* and *L. melissa* might contribute to isolation, or perhaps even that selection has favoured the spread of *L. melissa* male genitalia alleles.

Keywords: admixture, hybrid zone, Lepidoptera, reproductive isolation, speciation

Received 2 April 2010; revision accepted 18 May 2010

Introduction

The study of natural hybridization provides insights into the nature of species boundaries and the process of speciation (Barton & Hewitt 1985; Hewitt 1988; Harrison 1990; Jiggins & Mallet 2000; Mallet *et al.* 2007; Buerkle & Lexer 2008; Gompert & Buerkle 2009). Although hybrids are uncommon at an individual level, many nominal species hybridize with at least one other species (e.g. 5.8–25% of vascular plants, Ellstrand *et al.* 1996; Rieseberg 1997; Mallet 2005; 9.2% of birds, Grant & Grant 1992; 35% of *Heliconius* butterflies, Mallet *et al.* 2007; 6% of European mammals, Mallet 2005). These studies suggest that many of the important phenotypic and genetic discontinuities that delineate lineages we recognize as species persist despite hybridization. Moreover, studies of natural hybrid zones have demonstrated substantial heterogeneity in the extent of introgression among different genetic regions and

Correspondence: Zachariah Gompert, Fax: 307 766 2851; E-mail: zgompert@uwyo.edu

phenotypic traits (e.g. Parsons *et al.* 1993; Rieseberg *et al.* 1999; Stinchcombe & Hoekstra 2007; Carling & Brumfield 2008; Fitzpatrick *et al.* 2008; Nolte *et al.* 2009; Teeter *et al.* 2010). These findings are consistent with the view that reproductive isolation and adaptive differentiation are due to differences at individual genetic regions as suggested by Wu (2001), not the genome as a whole. This heterogeneity arises because recombination in hybrid or admixed individuals breaks up parental genotypes creating independence among genetic regions (Barton & Hewitt, 1985; Buerkle & Lexer 2008). Thus, patterns of introgression for individual genetic regions are affected by the fitness effects of allelic variants at these regions. Consequently, genetic regions associated with reproductive isolation or adaptive differentiation often experience little introgression, while the remainder of the genome experiences substantial introgression and is homogenized (Barton 1979, 1983; Rieseberg *et al.* 1999; Mallet 2005; Turner *et al.* 2005; Nolte *et al.* 2009; Teeter *et al.* 2010).

Hybridization and hybrid zone formation can affect evolutionary processes, including speciation. Incomplete reproductive isolation between species with overlapping geographical distributions can result in the formation of a hybrid zone, which can be defined as a region where genetically distinct lineages meet and produce hybrid offspring. Hybrid zones can persist for relatively long stretches of time (Barton & Hewitt 1985; Moore & Buchanan 1985) and numerous models have been developed to explain the structure and maintenance of hybrid zones. These include dispersal-independent models where hybrids are more fit than non-hybrid individuals in narrow ecotones (Moore 1977) and dispersal-dependent models where hybrid zones are maintained by a balance of dispersal into the zone and intrinsic (i.e. tension zones; Barton & Hewitt 1981, 1985) or environment-dependent (Endler 1977) selection against hybrids. The above models are generally concerned with hybrid zones in continuous habitats, although models of mosaic hybrid zones, where discrete patches of habitat preferred by or more suitable for each of two lineages are interdigitated forming a mosaic, have also been developed (Harrison & Rand 1989). Natural hybrid zones that have structures generally consistent with each of these models have been identified and studied (e.g. Grant 1971; Endler 1977; Moore 1977; Harrison 1986; Howard 1986; Szymura & Barton 1986, 1991; Harrison & Rand 1989; Vines *et al.* 2003; Yanchukov *et al.* 2006).

Hybrid zones can function as conduits for the spread of adaptive genetic variation among differentiated lineages, while limiting gene flow for other regions of the genome, particularly those regions associated with reproductive isolation (Ellstrand 2003; Grant *et al.* 2004; Martin *et al.* 2006; Whitney *et al.* 2006; Nolte *et al.* 2009). Hybrid zones can also be much more ephemeral. For example, when reproductive isolation is weak or species abundances differ markedly, hybrid zones may expand and cause the fusion of differentiated lineages into a single lineage (Levin *et al.* 1996; Avise *et al.* 1997; Fitzpatrick & Shaffer 2007; Morgan-Richards *et al.* 2009). Alternatively, if hybrids generally have reduced fitness relative to non-hybrid individuals, selection against hybrid formation can reinforce pre-zygotic isolation driving the speciation process towards completion (Lukhtanov *et al.* 2005; Nosil & Yukilevich 2008; Ortiz-Barrientos *et al.* 2009). However, reinforcement depends on limited gene flow from allopatric populations or specific genetic architectures of pre-zygotic isolation and might be uncommon in nature (Felsenstein 1981; Butlin 1987; Sætre *et al.*, 1997; Servedio & Noor 2003; Coyne & Orr 2004; Ortiz-Barrientos *et al.* 2009). Finally, hybridization can facilitate the origin of new evolutionary lineages via homoploid or polyploid hybrid speciation (Arnold 1997; Rieseberg 1997; Buerkle *et al.* 2000; Otto & Whitton 2000; Gompert *et al.* 2006a; Mavárez *et al.* 2006; Mallet 2007; Duenez-Guzman *et al.* 2009). Theoretical and empirical studies suggest that hybridization is more likely to result in hybrid speciation when admixed individuals are able to occupy a niche not used by their parental species (Buerkle *et al.* 2000; Rieseberg *et al.* 2003; Gompert *et al.* 2006a; Rieseberg 2006). Clearly, the outcomes and evolutionary consequences of hybridization are variable and depend on numerous factors including: the fitness of hybrid offspring, the strength and nature of selection, patterns of dispersal and the ecological context in which hybridization occurs (Barton 1979, 2001; Barton & Hewitt 1985; Buerkle *et al.* 2000; Jiggins & Mallet 2000; Duenez-Guzman *et al.* 2009; Lepais *et al.* 2009).

In the current study we investigate hybridization dynamics between populations of two nominal butterfly species that occur in the Rocky Mountains, *Lycaeides idas* and *L. melissa*. The goals of this study are to determine the extent, outcome and consequences of hybridization between these lineages and to explore potential phenotypic traits contributing to their isolation. *Lycaeides* is a holartic genus of small blue butterflies in the family Lycaenidae. Between two and five *Lycaeides* species are recognized in North America, including *L. idas*, which is found in Eurasia, Alaska, much of Canada and mountainous regions of the western United States, and *L. melissa*, which occupies the western and north-eastern United States as well as portions of southern Canada (Fig. 1; Scott 1986; Guppy & Shepard 2001; Brock & Kaufman 2003; Gompert *et al.* 2006a). *Lycaeides idas* populations generally occur in more mesic environments, whereas *L. melissa* populations occupy drier grassland or agricultural areas (Scott 1986; Nice &

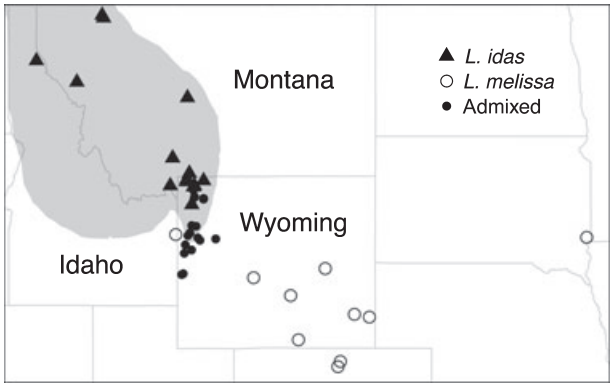

Fig. 1 Map depicting sampling localities for *Lycaeides* butterflies in the north-western United States. The approximate range of *L. idas* is shown in grey (Scott 1986; Brock & Kaufman 2003). *Lycaeides* melissa occupies most of the area depicted in the map. Designations of *L. idas*, *L. melissa* and admixed populations are based on morphological, ecological, genetic and distribution data.

Shapiro 1999). *Lycaeides* butterflies tend to have low dispersal distances (generally 500 m or less) and patchy distributions governed by the availability of suitable habitat and host plants (Scott 1986; Knutson *et al.* 1999; U.S. Fish and Wildlife Service 2003). Some *Lycaeides* populations display metapopulation dynamics with local extinction and recolonization of at least a subset of habitat patches occurring at an appreciable rate (U.S. Fish and Wildlife Service 2003).

Numerous regions of hybridization have been documented in North American *Lycaeides* butterflies with various outcomes (Gompert *et al.* 2006a, b, 2008b, 2010; Lucas *et al.* 2008). These cases of hybridization provide an interesting contrast to our present study. Our previous work suggests that hybridization between *L. idas* (recently designated *L. anna* by Guppy & Shepard 2001) *L. melissa* populations in the vicinity of the Sierra Nevada gave rise to an alpine-associated homoploid hybrid species (Gompert *et al.* 2006a; Lucas *et al.* 2008). In fact, this hybrid lineage might be one of four or more independent hybrid lineages occupying the mountainous regions of the western United States (additional putative hybrid lineages occur in the Siskyou Mountains, Warner Mountains and White Mountains of California, Oregon and Nevada; Gompert Z, *et al.* 2008b, 2010; Nice C, *et al.*, unpublished). Conversely, hybridization between populations of *L. melissa* and the Karner blue butterfly resulted in the replacement of Karner blue mitochondrial haplotypes with a common *L. melissa* haplotype in Karner blue populations west of Lake Michigan (Nice *et al.* 2005; Gompert *et al.* 2006b, 2008a, b). This mitochondrial introgression was associated with, and likely facilitated by, the spread of the endoparasitic bacterium *Wolbachia* from *L. melissa* to Karner blue butterfly populations (Gompert *et al.* 2008b; Nice *et al.* 2009). Despite this extensive mitochondrial introgression, little evidence for nuclear introgression has been detected between *L. melissa* and the Karner blue butterfly. Mitochondrial introgression has also been detected between *L. melissa* and putative high-elevation hybrid lineages in the White Mountains and Siskyou Mountains in western North America (Gompert *et al.* 2008a, b).

Lycaeides idas and *L. melissa* are broadly sympatric in the central and northern Rocky Mountains (i.e. the mountains of north-western Wyoming, Idaho and Montana), which is the focus of our current manuscript. Specifically, nominal *L. idas* and *L. melissa* populations can be found in close proximity to each other in the vicinity of the Jackson Hole valley and the Grand Teton mountain range of north-western Wyoming. This distribution likely constitutes relatively recent secondary contact following post-Pleistocene range shifts, as much of the Jackson Hole valley and surrounding area was glaciated until approximately 14 000 years BP (Wisconsin glaciation; Harris *et al.* 1997). Nabokov (1949, 1952) identified individual *Lycaeides* from several populations in this region with male genitalia intermediate between *L. idas* and *L. melissa*, which he believed was the result of hybridization between these species. *Lycaeides idas* and *L. melissa* populations in this region have identical mitochondrial haplotypes (Gompert *et al.* 2008a), which provides additional evidence of hybridization.

Near this region of secondary contact *L. idas* and *L. melissa* populations differ ecologically and morphologically. For example, *L. idas* populations use several native species of *Astragalus* (e.g. *A. miser* and *A. alpinus*) as larval host plants, whereas most *L. melissa* populations feed on cultivated or feral alfalfa (*Medicago sativa*; personal observation, Gompert Z). *Lycaeides idas* populations have a single brood per year with adults active from early July into early August (personal observation, Gompert Z). Alternatively, *L. melissa* populations have two or more broods with the adults flying from June to the middle of July and again in August. Compared with *L. melissa*, *L. idas* individuals have smaller scleritized male genitalic structures and somewhat reduced wing pattern elements (i.e. black spots and orange iridescent aurorae; Nabokov 1949; Scott 1986; Lucas *et al.* 2008).

These morphological differences are the basis for much of the current taxonomic classification of *Lycaeides* butterflies (Nabokov 1949; Scott 1986), and might contribute to reproductive isolation. Variation in female wing pattern coupled with variation in male mate preference is known to cause pre-zygotic isolation between butterfly species (Jiggins *et al.* 2001; Kronforst *et al.* 2006; Chamberlain *et al.* 2009), including some *Lycaeides* populations (Fordyce *et al.* 2002). Differences in insect genitalia have often been used to delineate insect

species, because of a proposed link to mechanical isolation (Nabokov 1949; Shapiro & Porter 1989). Although little support exists for the once popular lock-and-key hypothesis that selection for pre-insemination avoidance between different species causes insect genitalia to diverge (e.g. Eberhard 1985; Shapiro & Porter 1989; Arnqvist *et al.* 1997; Ohno *et al.* 2003; but see, Polak & Rashed 2010), these differences might nevertheless be driven by selection and contribute to reproductive isolation (Eberhard 1985, 1993; Eberhard *et al.* 1998; Takami 2003; Nagata *et al.* 2007). We do not know whether differences in male genitalia cause mechanical isolation between *L. idas* and *L. melissa*.

Herein we characterize patterns of hybridization and admixture among Rocky Mountain *Lycaeides* populations. The motivation for this study was to determine whether Rocky Mountain *L. idas* and *L. melissa* hybridize and explore the dynamics and consequences of hybridization in this region. We addressed the following specific questions: (i) Has admixture occurred between Rocky Mountain *L. idas* and *L. melissa*? (ii) Has hybridization led to the formation of a hybrid zone or the establishment of an isolated hybrid lineage? (iii) What is the geographical distribution, genetic and morphological composition, and ecological context of admixture in Rocky Mountain *Lycaeides*? and (iv) Do morphological differences in wing pattern or male genitalic morphology contribute to reproductive isolation between *L. idas* and *L. melissa* populations in the Rocky Mountains? We contrast the patterns and outcomes of this potential case of hybridization with other cases of hybridization in *Lycaeides* butterflies and other hybrid zones.

Methods

Sampling and molecular markers

We collected adult *Lycaeides* from 45 localities in western North America (Table 1; Fig. 1). Samples were collected from as far south as Colorado to northern Canada as the geographical extent of hybridization was initially unknown. Collections within U.S. National Parks (NP) were made in accordance with National Park Service permits: Yellowstone NP (YELL-2008-SCI-5682), Grand Teton NP (GRTE-2008-SCI-0024) and Glacier NP (GLAC-2009-SCI-0140). Collected butterflies were stored at −20 °C prior to processing for molecular and morphological data. We isolated and purified DNA from approximately 10 mg of thoracic tissue using Qiagen's DNeasy 96 Blood and Tissue Kit (Cat. No. 69581; Qiagen Inc., Valencia, CA, USA) in accordance with the manufacturer's recommended protocol.

We genotyped each of 360 DNA samples for two single nucleotide polymorphism (SNP) markers, four microsatellites, and fragment presence/absence polymorphism with five AFLP selective primer pair combinations. We identified SNPs in two single copy nuclear genes: glyceraldehyde 3 phosphate (gapdh; forward primer: 5′-CACAACCATTGACAAAGCWTC-3′, reverse primer: 5′-TTGGTCTTCRGTGTATCCAAG-3′) and invected (inv; forward primer: 5′-TCGTTTTACAAGGCCCAAAG-3′, reverse primer: 5′-CGTCAGGTACCGGTTCTCC-3′), by sequencing four *L. idas* and four *L. melissa*. The SNPs identified were variable within and between *Lycaeides* species. Subsequently, we scored these SNPs by visualizing PCR products on 2% agarose gels following digestion with *Taq*α1 and *Hin*dIII for gapdh and inv respectively. The four microsatellite markers we amplified were Msat4, Msat6, MsatZ12 and Msat201, which were designed from a *L. melissa* genomic DNA library (Anthony *et al.* 2001) and were known to be polymorphic within and among *L. melissa* and *L. idas* populations (Anthony *et al.* 2001; Gompert *et al.* 2006a). Fluorescently labelled microsatellite products were separated and visualized on an ABI 3100 DNA analyser (Applied Biosystems Inc., Carlsbad, CA, USA) and scored using GeneMapper v4.0. We generated AFLP profiles for each of 360 butterflies using the following five selective primer pair combinations: *Eco*RI-CACT and *MSE*I-CTA, *Eco*RI-CACT and *Mse*I-CAT, *Eco*RI-CAAT and *Mse*I-CAT, *Eco*RI-CACT and *Mse*I-CTT, and *Eco*RI-CACT and *Mse*I-CAA. Fluorescently labelled amplicons were separated and visualized on an ABI 3730 DNA analyser (Applied Biosystems Inc.) at the University of Nevada Genomics Center. We used the software GeneMapper v4.0 to evaluate, bin and call AFLP markers. A set of eight individuals was re-run on every 96-well plate to verify repeatability of the AFLP markers and assess interplate variability in amplification and sizing. Aberrant AFLP markers (i.e. markers that did not consistently amplify in replicate runs or markers that could not be reliably scored) were removed from the data set. We retained 360 polymorphic AFLP markers for analysis. We used principal components analysis (PCA) as an initial means to explore variation in the AFLP data for the sampled populations. PCA was performed using singular value decomposition on scaled and centred data with the PRCOMP function in R (R Development Core Team 2009).

Morphological characters

We measured three components of the scleritized portion of the male genitalic anatomy: falx length, humerulus length and uncus length (Fig. 2A). These scleritized structures interact with female reproductive morphology during mating. This structure has been used extensively in *Lycaeides* systematics and is believed to be important for copulation (Nabokov 1949; Nice &

Table 1 *Lycaeides* sample localities

Designation	Locality	Taxon	Latitude (°)	Longitude (°)	Elevation (m)
GVL	Gardnerville, NV	*L. melissa*	38.9608	119.7788	1435
BSD	Brandon, SD	*L. melissa*	43.5946	96.5723	412
BHC	Buckhorn Canyon, CO	*L. melissa*	40.5763	105.3777	2077
GRR	Greyrock, CO	*L. melissa*	40.6963	105.2917	1747
GOS	Goshen, WY	*L. melissa*	41.7346	104.2645	1364
CGW	Chugwater, WY	*L. melissa*	41.7993	104.7940	1577
IBT	IndianBathtubs, WY	*L. melissa*	41.2040	106.7714	2214
GLR	Glenrock, WY	*L. melissa*	42.8662	105.8285	1518
BMT	Bear Mountain, WY	*L. melissa*	42.2378	107.0492	1982
LAN	Lander, WY	*L. melissa*	42.6533	108.3551	1787
VIC	Victor, ID	*L. melissa*	43.6590	111.1114	1850
DBS	Dubois, WY	*L. melissa*	43.5622	109.6991	2153
TSS†	Teton Science School, WY	*L. melissa*	43.4233	110.7757	1918
PSP†	Periodic Spring, WY	*L. melissa*	42.7468	110.8493	2113
SWC†	Swift Creek, WY	*L. melissa*	42.7251	110.9066	1949
BCR†	Bull Creek, WY	*L. idas*	43.3007	110.5530	2195
SDL†	Soda Lake, WY	*L. idas*	43.5283	110.2573	2359
ETL†	East Table, WY	*L. idas*	43.2249	110.8117	1865
USL†	Upper Slide Lake, WY	*L. idas*	43.5829	110.3328	2246
BTB†	Blacktail Butte, WY	*L. idas*	43.6382	110.6820	2220
SHA†	Shadow Mountain, WY	*L. idas*	43.6974	110.6102	2180
MRF†	Mt. Randolf, WY	*L. idas*	43.8547	110.3918	2221
LZH†	Lozier Hill, WY	*L. idas*	43.8729	110.5497	2122
AVP	Avalanche Peak, WY	*L. idas*	44.4860	110.1307	2998
RDL	Riddle Lake, WY	*L. idas*	44.3617	110.5465	2395
NBR	Natural Bridge, WY	*L. idas*	44.5278	110.4479	2373
TRL	Trout Lake, WY	*L. idas*	44.9019	110.1291	2124
HNV	Hayden Valley, WY	*L. idas*	44.6823	110.4945	2344
MWB	Mt. Washburn, WY	*L. idas*	44.7832	110.4494	2740
CDL	Cascade Lake, WY	*L. idas*	44.7550	110.4927	2424
ICR	Indian Creek, WY	*L. idas*	44.8787	110.7387	2211
BNP	Bunsen Peak, WY	*L. idas*	44.9337	110.7212	2260
JAR	Jardine, MT	*L. idas*	45.0747	110.6335	1985
WTC	Watkins Creek, MT	*L. idas*	44.7849	111.3088	2150
GNP	Garnet Peak, MT	*L. idas*	45.4323	111.2245	1910
KHL	King's Hill, MT	*L. idas*	46.8407	110.6990	2239
SDC	Soldier Creek, MT	*L. idas*	47.2062	114.6118	1148
SYC	Siyeh Creek, MT	*L. idas*	48.7026	113.6681	1764
GNG	Grinell Glacier, MT	*L. idas*	48.7673	113.7192	1856
BRC	Brown Creek, ID	*L. idas*	47.7044	116.0437	857
HSB	Hailstone Butte, AB	*L. idas*	50.2010	114.4460	2033
PCT	Prospect Creek, AB	*L. idas*	52.9670	117.3840	1767
EBR	Brülé, AB	*L. idas*	53.2831	117.8683	1002
MMR	Montana Mountain, YT	*L. idas*	60.2269	134.6804	914
BEL	Bennet Lake, YT	*L. idas*	60.1760	134.7240	664
WHR	Whitehorse, YT	*L. idas*	60.7227	135.1752	819

†Core hybrid zone populations.

Shapiro 1999; Lucas *et al*. 2008). Soft tissue surrounding the scleritized portion of the male genitalia was dissolved by submersing the posterior-most abdominal segments of each *Lycaeides* butterfly in 100 °C 5 M KOH. We then dissected and measured the genitalia under a dissecting scope using an optical micrometer. We measured the genitalia of 360 male *Lycaeides* butterflies.

We measured 23 wing pattern characters from 243 male and 104 female *Lycaeides* (Fig. 2B). First, we removed the wings from each butterfly and photographed them using a digital SLR Canon Rebel camera (Cannon Inc. USA, Lake Success, NY, USA). We designated 147 wing points at wing veins and around the edges of wing pattern elements (Fig. 2B). We then

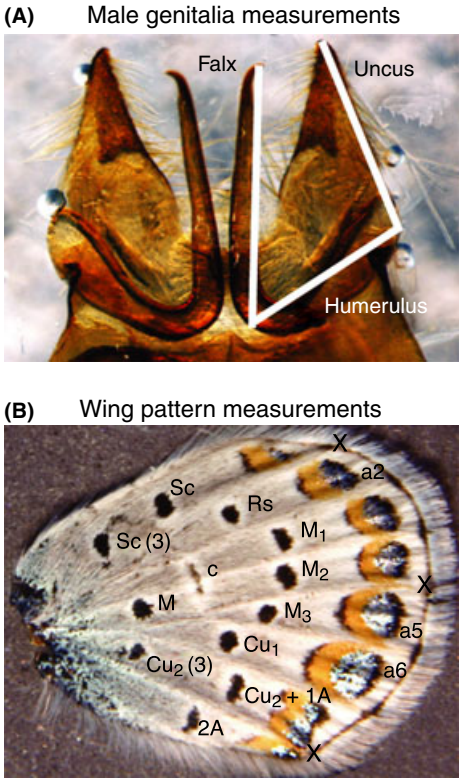

Fig. 2 Photographs depicting male genitalia (A) and wing pattern (B) characters measured in this study. Wing pattern vein margins used to calculate wing area or perimeter are denoted by 'X'. Wing aurorae include the orange, black and iridescent blue pattern elements near the wing margins (i.e. elements a2, a5 and a6). The orange portion of each aurorae is referred to as the b2 aurorae region.

obtained coordinates for these points using the IMAGEJ software. We then used these points to calculate the area of the following 18 wing pattern elements (black wing pattern spots and aurorae): a2, b2 (the orange portion) of a2, a5, b2 of a5, a6, b2 of a6, Cu_2(3), M, Sc(3), Sc, Rs, M_1, M_2, M_3, Cu_1, Cu_2 + 1A, 2A and c, as well as straight line distances between the centroids of c and M, c and M_2, Rs and 2A, c and a6, and Sc(3) and m (Fig. 2B). All measurements were standardized by wing size by dividing them by wing area (for area characters) or wing perimeter (for linear distance characters; calculated from wing edge landmarks at veins Rs, M_3 and 2A). Area and distance measurements were calculated from point coordinates using the R statistical computing environment (R Development Core Team 2009) with custom scripts and functions from the *splancs* R package (Rowlingson & Diggle 1993). We used PCA to explore variation in the morphological data. We conducted separate PCAs for the genitalic measurements and wing pattern characters as described above for the AFLP data.

Analysis of molecular data

Our first objective was to determine whether admixture occurred between Rocky Mountain *L. idas* and *L. melissa* populations and what evolutionary consequences hybridization has had. Thus, we used the Bayesian model-based analysis implemented in *structure* 2.2 (Pritchard *et al.* 2000; Falush *et al.* 2007) to determine whether our sample of *Lycaeides* included individuals with admixed ancestry, and whether admixed individuals represented a cohesive genetically distinct lineage. Although we were primarily interested in a model with two clusters for assessing evidence of admixture, we were also interested in whether models with more than two clusters were supported by the data and whether admixed individuals formed their own clusters. Support for the latter would constitute evidence that admixed populations have achieved a degree of evolutionary independence from parental populations, as would be expected if they constituted a stable hybrid species. Therefore, we conducted analyses with the number of clusters (k) ranging from one to ten to determine whether a model with two clusters explained the genetic data well and whether models with additional clusters were supported. These analyses were based on a total of 366 molecular markers (360 AFLP amplicons, four microsatellites and two SNPs). We used the admixture model with correlated allele frequencies, as this model is most appropriate for individuals with admixed ancestry (Pritchard *et al.* 2000). Parameter estimation was based on 150 000 Markov chain Monte Carlo iterations following a 50 000 iteration burnin. Ten replicate analyses for each number of clusters were conducted. We then determined support for specific numbers of clusters by plotting the marginal likelihood of each model as a function of k and using the *ad hoc* Δk statistic proposed by Evanno *et al.* (2005).

We used Hartigan's dip test of unimodality (Hartigan 1985; Hartigan & Hartigan 1985) to characterize the distribution of admixture proportions for samples from the core of the hybrid zone (localities: TSS, PSP, SWC, BCR, SDL, ETL, USL, BTB, SHA, MRF, LZH; see Results). This test determines whether an empirical distribution departs from a unimodal distribution, and thus can be used to estimate the extent of reproductive isolation between hybridizing lineages. Bimodal distributions of admixture proportions and cluster membership would be expected if substantial pre-zygotic isolation exists between *L. idas* and *L. melissa* (Jiggins & Mallet 2000). To obtain a null distribution of dip test statistics for significance testing, we generated 100 000 samples of $n = 82$ from a uniform distribution bounded by zero and one, where n was equal to our

sample size for admixture proportions from the core hybrid zone populations, and calculated a dip test statistic for each sampled replicate. This analysis was conducted in the R statistical computing environment (R Development Core Team 2009) using the *diptest* package (Maechler & Ringach 2009) and additional code written by Z.G.

Analysis of morphological data

Prior to testing for morphological evidence of admixture, we conducted several tests to determine if genetically differentiated allopatric *L. idas* and *L. melissa* populations (based on Bayesian clustering; see Results) were morphologically distinct. These analyses of allopatric populations provide a context for the morphological characterization of individuals from admixed populations. The allopatric populations used for this analysis were BNP, BRC, CDL, GNG, GNP, ICR, JAR, KHL, MWB, RDL, SDC, SYC, TRL and WTC for *L. idas* and BMT, BSD, GOS, GRR, GVL, IBT and VIC for *L. melissa* (Table 1). Designated allopatric *L. melissa* populations were outside the known range of *L. idas*, whereas designated allopatric *L. idas* populations were at least 30 km from known *L. melissa* populations. We contrasted the first principal component for male genitalia (male genitalia size, see Results) between these *L. idas* and *L. melissa* populations using a single factor analysis of variance (ANOVA). Similarly we contrasted the first five principal components for wing pattern between *L. idas* and *L. melissa* using multivariate ANOVA (MANOVA; see Results). For the latter, sex was also included as a factor. In addition, we used linear discriminant function analysis to determine whether male genitalia or wing pattern could be used to distinguish between Rocky Mountain *L. idas* and *L. melissa* populations. Linear discriminant function analysis was performed separately for male genitalia, male wing pattern, and female wing pattern. ANOVA, MANOVA and linear discriminant function analysis were performed in R.

We tested for the presence of morphological clusters and morphologically admixed individuals based on male genitalic morphology and wing pattern characters. Specifically, we used a fuzzy c-means clustering algorithm to identify clusters and estimate cluster membership based on these morphological characters. Unlike classic k-means clustering, fuzzy c-means clustering assumes that each observation has a degree of belonging to each cluster (Bezdek 1981), and thus is appropriate for instances of expected admixture. We used the objective function from Kaufamn & Rosseeuw (1990) for fuzzy c-means clustering, where the aim of the clustering analysis is to minimize

$$\Sigma_{v=1}^{k} \frac{\Sigma_{i,j} u(i,v)^r u(j,v)^r d(i,j)}{2 \times \Sigma_j u(j,v)^r}.$$

In this function, k is the number of clusters, v is a specific cluster, $u(i,v)$ and $u(j,v)$ denote the membership of observation i and j in cluster v respectively, $d(i,j)$ is the dissimilarity between i and j, and $r \in [1,\infty]$ is the membership exponent (Kaufamn & Rosseeuw 1990). We measured dissimilarity $d(i,j)$ as sum-of-squares Euclidean distances. The membership exponent (r) determines the degrees of 'fuzziness' for fuzzy c-means clustering. As r approaches 1 each observation is assigned wholly to a cluster and fuzzy c-means clustering is equivalent to classic k-means clustering, whereas as r approaches ∞ clustering becomes more fuzzy until each observation is distributed equally among the k clusters (Bezdek 1981; Pal & Bezdek 1995). The choice of r is expected to affect the proportion of individuals classified as morphological hybrids at any specific threshold. Unfortunately, the choice of r is somewhat subjective and dependent on the data and nature of the clustering problem (Bezdek 1981; Pal & Bezdek 1995; Okeke & Karnieli 2006), although values between 1.5 and 2.5 have generally been suggested. Preliminary analyses suggested that values of $r > 2$ tended towards complete fuzziness. Therefore we used three values of r: 1.5, 1.75 and 2.0. We also used $r = 1.25$ for wing pattern data, because of difficulties clustering with higher values of r. We conducted fuzzy clustering with $k = 2$, as the genetic data were most consistent with two clusters (see Results). Separate fuzzy c-means clustering analyses were performed for male genitalic morphology, male wing pattern and female wing pattern. These analyses were performed using the FANNY function from the R package *cluster* (Struyf *et al.* 1997).

We used Hartigan's dip test to determine whether fuzzy cluster membership for male genitalia and male wing pattern for samples from the core of the hybrid zone followed unimodal or multimodal distributions (Hartigan 1985; Hartigan & Hartigan 1985). Female wing pattern was excluded from this analysis, as clusters were difficult to identify (see Results). Bimodal cluster membership could result from nearly complete isolation or if the specific morphological characters measured have experienced divergent selection or were associated with reproductive isolation. Null distributions of dip test statistics were obtained as described for admixture proportions but with $n = 102$ (male genitalia) and 57 (male wing pattern).

Cline analysis and geographical patterns of admixture

Geographical patterns of admixture (i.e. geographical clines) provide insight into the processes affecting

hybridization dynamics (Endler 1977; Barton & Hewitt 1981; Harrison & Rand 1989; Barton 1993; Gay *et al.* 2008). We investigated geographical patterns of genetic and morphological admixture between *L. idas* and *L. melissa* to determine whether these patterns were consistent with the presence of a hybrid zone and to increase our understanding of the phenotypic basis and strength of reproductive isolation. First, we fit unimodal, bimodal and trimodal character cline models separately for genomic composition (AFLP PC1), male genitalia size (male genitalia PC1), wing pattern aurorae and spot size (wing pattern PC1) and relative wing pattern spot size (wing pattern PC2; see Results for a description of these principal components) using *Cfit7*, which implements the cline models developed by Gay *et al.* (2008). Our analyses treat genomic composition, which is a continuous variable with positive and negative values, as a quantitative character. For wing pattern models we constrained males and females to share a cline centre and slope, but allowed them to have different trait value ranges and variances. We used Akaike Information Criterion (AIC) to contrast unimodal, bimodal and trimodal models for each character. We then tested for coincidence (common centre) and concordance (common width) of geographical clines in genomic composition and each morphological character. We tested for cline concordance and coincidence by constraining each morphological character to have the same slope or centre as the cline in genomic composition. We used the unimodal character distribution model for these geographical cline comparisons (see Results). We then used AIC to compare the constrained and unconstrained models. For each model, we ran the simulated annealing algorithm implemented in *Cfit7* from multiple independent starting positions to determine whether the algorithm was exploring parameter space adequately.

Prior to fitting the cline models above, we converted the latitude and longitude coordinates of our sampling localities into locations along a single axis. First, we calculated pair-wise geographical distances among the sampling localities i and j as the great circle distance. We then used classical multidimensional scaling to ordinate these pair-wise distances onto a single dimension that retained the overall pair-wise distance structure. This ordinated dimension was equated with geographical location along a linear transect in kilometres, with km 0 at the approximate centre of the sampled sites. Initial analyses suggested that two *L. melissa* localities (GVL and VIC) were west of the main axis of sampling. Therefore, we excluded these localities from geographical analyses.

We also used a non-spatial analysis to compare patterns of genetic and morphological variation in Rocky Mountain *Lycaeides* populations and gain insights into the potential role of male genitalia and wing pattern variation in reproductive isolation. If patterns of variation for genomic composition and each morphological character are similar, a simple linear model should describe the relationship between these characters. However, if cline widths for these characters differ, a cubic regression model should describe their relationship better (Szymura & Barton 1991; Bridle & Butlin 2002; Brennan *et al.* 2009). Moreover, the strength of the relationship between genomic composition and each morphological character serves as a general measure of the extent that these characters display similar patterns of variation. We tested for a relationship between genomic composition and three morphological characters: male genitalia size, wing aurorae and spot size, and relative wing spot size. We fit linear and cubic models for each of these characters using the LM function in R. Wing pattern models included sex as a covariate. We determined whether a cubic model better fit the data for each character using a likelihood ratio test.

Ecological context and hybrid zone structure

Frequent dispersal is an important component of many hybrid zone models (e.g. Endler 1977; Barton & Hewitt 1985). Dispersal from parental populations and ongoing hybridization resulting in F_2 and back-cross individuals is expected to increase within-population linkage disequilibrium (LD) and variance in morphological characters (Barton & Hewitt 1985; Szymura & Barton 1986; Lynch & Walsh 1998; Gay *et al.* 2007). Therefore, we tested for increased linkage disequilibrium and increased variability in male genitalia size, wing aurorae and spot size, relative wing spot size, and genomic composition in the centre of the hybrid zone relative to populations farther from the hybrid zone. For each pair of variable codominant loci, we calculated the correlation-based composite linkage disequilibrium coefficient, R^2, described by Zaykin *et al.* (2008) using the program *RxC*. As a measure of character variability, we calculated the absolute deviation of each individual's character value from the locality median (this is equivalent to the response variable used for the Brown–Forsythe test for equality of variances; Brown & Forsythe 1974). Linkage disequilibrium and character deviations were calculated for sampling localities with five or more individuals. Only males were used for estimating character deviations. Females were excluded because of sexual dimorphism coupled with differences in the relative number of females and males sampled from each locality. We fit quadratic regressions of mean linkage disequilibrium or character deviations as a function of sampling location using the LM function in R.

A significant overall model fit coupled with a quadratic term significantly different from zero and yielding a higher expected value in the centre of the hybrid zone was taken as evidence for increased linkage disequilibrium or character variability associated with the hybrid zone centre.

Environmental variation might affect the structure of the Rocky Mountain *L. idas* × *L. melissa* hybrid zone, as suggested by mosaic and environment-dependent hybrid zone models (e.g. Endler 1977; Moore 1977; Harrison & Rand 1989). To explore this possibility we first asked whether admixed populations occupying similar environments had similar admixture proportions after controlling for geographical proximity. Populations with mean admixture proportions between 0.1 and 0.9 were included in this analysis (13 populations). Environmental data were obtained from WestMap, which uses the Parameter-elevation Relationships on Independent Slopes Model (PRISM) to interpolate climate data for ~800 square metre grids (http://www.cefa.dri.edu/Westmap/Westmap_home.php; Daly *et al.* 2008). For each location we obtained interpolated data for four climatic variables: mean July precipitation, mean daily mean temperature for July, mean daily maximum temperature for July and mean daily minimum temperature for January. These variables were chosen to provide a concise summary of aspects of climate that likely affect *Lycaeides* butterflies. We then used a partial Mantel test to determine whether there was a correlation between pair-wise differences in mean admixture proportions and pair-wise climatic Euclidean distances while controlling for pair-wise geographical distances among populations. This analysis was conducted in R using the PARTIAL.MANTEL function. We performed 10 000 permutations to determine whether the observed correlation was significantly greater than zero.

We then asked whether populations more genetically similar to *L. idas* or *L. melissa* occupied an environment more similar to the environment occupied by *L. idas* or *L. melissa* respectively. We addressed this question by first testing for a correlation between the mean admixture proportion for each of the admixed populations and each climate variable and then asking whether this correlation was in the same direction as the environmental differences between locations occupied by *L. idas* and *L. melissa*.

Results

Molecular data

The first principal component for the AFLP data accounted for 4.2% of the variation in the AFLP data and included many markers with positive and negative loadings. *Lycaeides idas* and *L. melissa* differed significantly for AFLP PC1 ($F_{1,154} = 738.82$, $P < 0.0001$), suggesting that PC1 corresponds to the axis of species differentiation. Hereafter, we refer to this principal component as 'genomic composition'. AFLP PC2 accounted for 2.4% of the AFLP variation, but AFLP PC2 scores did not differ between *L. idas* and *L. melissa* ($F_{1,154} = 0.126$, $P = 0.7236$). Additional AFLP principal components accounted for little variation and were not examined in detail.

Initial analysis of all 45 sampled populations using *structure* suggested that five *L. idas* populations sampled from Canada (PCT, HSB, EBR, MMR, WHR) were genetically distinct from *L. idas* populations farther south in Montana and Wyoming, including putatively admixed populations (results not shown). *structure* results including the Canadian populations and assuming three gene pools were similar to results excluding these populations with two gene pools (i.e. the Canadian populations simply made up the third gene pool; results not shown). As results excluding the Canadian populations were easier to interpret in the context of two hybridizing lineages, these five Canadian *L. idas* populations were excluded from all further analyses. A plot of the marginal likelihood of each *structure* model vs. the number of clusters assumed indicated that the likelihood increased substantially with an increase in k from one to two and then increased to a lesser extent or decreased with additional increases in k (Fig. S1A, Supporting Information). Similarly, Δk was roughly six times greater for $k = 2$ than all other values of k assessed (Fig. S1B, Supporting Information). Taken together, these results suggest that the sampled *Lycaeides* individuals come from two distinct gene pools, but do not completely exclude the possibility that three gene pools exist.

Bayesian admixture proportions estimated from *structure* for $k = 2$ indicated that most individuals were derived from a single gene pool (cluster) (Fig. 3A). Specifically 25.88% of the sampled individuals were classified as *L. melissa* (cluster 1) and 55.27% of individuals were classified as *L. idas* (cluster 2) assuming the following admixture proportion bins: 0.0–0.1 (*L. idas*), 0.1–0.9 (admixed) and 0.9–1.0 (*L. melissa*). These thresholds were shown to be efficient for identifying admixed individuals in a simulation study by Vaha & Primmer (2006) and have been used in other hybrid zone analyses (e.g. Beaumont *et al.* 2001; Lancaster *et al.* 2006). Although many non-admixed individuals were identified, 18.85% of the sampled *Lycaeides* individuals could not be assigned to the *L. idas* or *L. melissa* gene pools, but instead appeared to have admixed ancestry. The majority of admixed individuals were sampled from a relatively small geographical region near the Jackson Hole valley (localities: TSS, PSP, SWC, BCR, SDL, ETL,

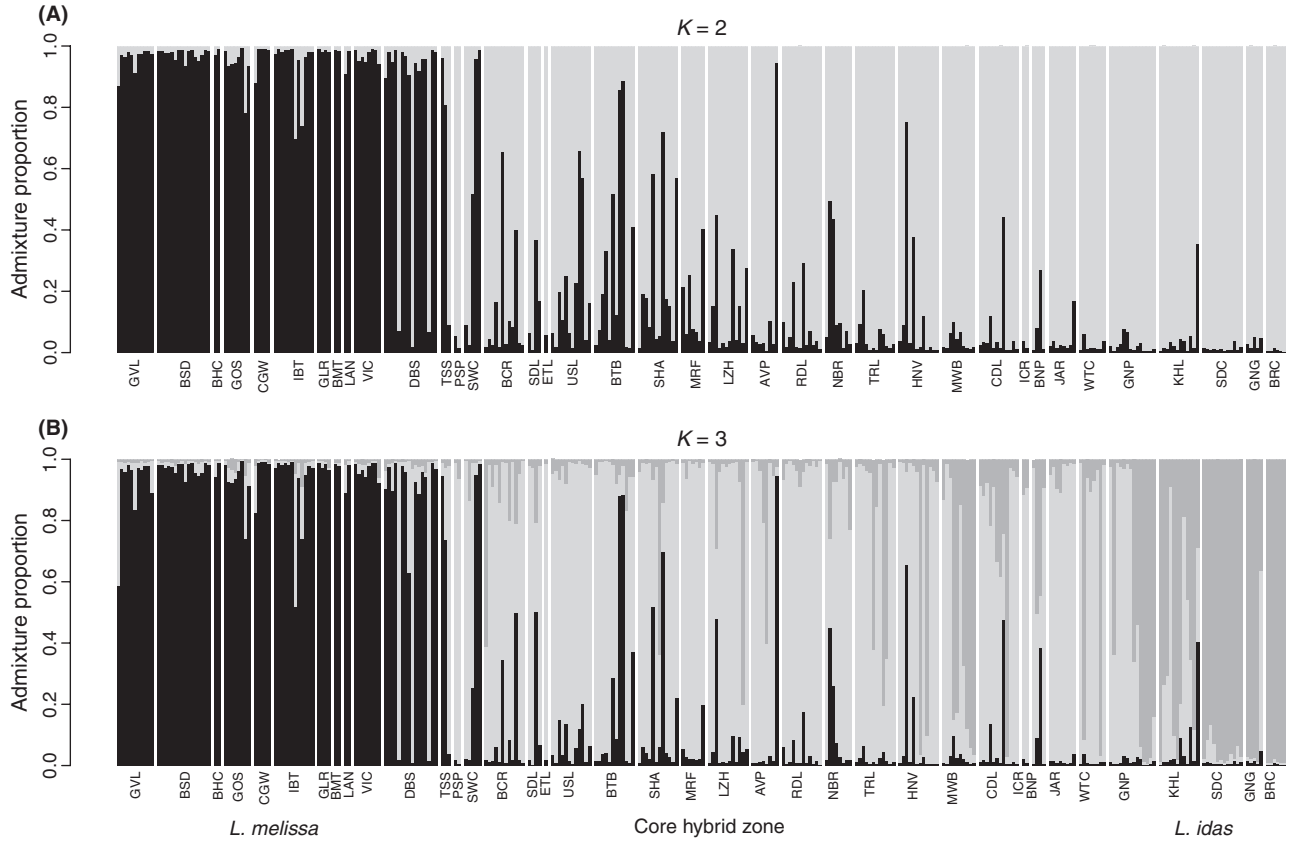

Fig. 3 Barplots depicting Bayesian admixture proportions estimated using *structure* with $k = 2$ (A) and $k = 3$ (B). Each bar depicts an individual and is coloured to denote the proportion of an individual's genome inherited from cluster 1 (black, *Lycaeides melissa*), cluster 2 (light grey, *L. idas* for $k = 2$ or southern *L. idas* for $k = 3$) and cluster 3 (dark grey, northern *L. idas* for $k = 3$). White bars separate sampling localities, which are labelled using abbreviations from Table 1.

USL, BTB, SHA, MRF, LZH), which appears to represent the core of the hybrid zone (45.12% of individuals from these localities had mixed ancestry). However, admixed individuals were also sampled outside this core hybrid zone suggesting that introgression is not limited to this core region. For example, several admixed individuals were sampled from NBR and HNV on the Yellowstone plateau (Fig. 3A). Bayesian admixture proportions when three gene pools were assumed were similar to those for two gene pools, but divided the *L. idas* gene pool into a more southern (cluster 2) and more northern (cluster 3) gene pool (although the boundary between these was not distinct, Fig. 3B). Assuming three gene pools had little effect on the number of individuals that were classified as *L. melissa* (24.28% of individuals were assigned to *L. melissa*, cluster 1), and only a moderate effect on the number of individuals classified as admixed between *L. melissa* and either *L. idas* cluster (12.46% of individuals). Visual inspection of admixture proportions suggested that individuals identified as admixed given two gene pools were also generally identified as admixed given three gene pools and did not constitute a third unique gene pool. The slight reduction in the proportion of admixed individuals with three gene pools might reflect a greater similarity between the southern *L. idas* and *L. melissa* gene pools with $k = 3$ relative to the complete *L. idas* and *L. melissa* gene pools with $k = 2$. Based on Hartigan's dip test of unimodality, we failed to reject the null hypothesis that the distribution of admixture proportions for individuals sampled from the core hybrid zone was unimodal (dip = 0.02913, $P = 0.9261$; Fig. S3A, Supporting Information). The distribution of admixture proportions was right skewed with a median of 0.0915 (first quantile = 0.0383, third quantile = 0.320), but included a wide range of admixture proportions.

Morphological data

Most variation in male genitalia (85.7%) was explained by the first principal component. All three measurements of male genitalia had positive loadings on PC1 suggesting that PC1 corresponded to genitalia size, thus

we refer to this principal component as 'male genitalia size' (previous morphometric analysis of other *Lycaeides* populations found very similar results; Nice & Shapiro 1999; Lucas *et al*. 2008). *Lycaeides idas* and *L. melissa* differed significantly in male genitalia size, with the latter having larger genitalia than the former ($F_{1,220} = 817.79$, $P < 0.0001$). The first principal component for wing pattern accounted for 38.3% of the variation in wing pattern and had moderate positive loadings for all area measurements but not for distance measurements. We refer to this principal component as 'wing pattern aurorae and spot size'. Wing pattern PC2 had positive loadings for wing aurorae, but negative loadings for most black spots and accounted for 12.5% of wing pattern variation. We refer to this principal component as 'relative wing pattern spot size'. Wing pattern PC1–PC5 each accounted for more than 5% of the total variation in wing pattern (69.1% total). Based on the first five principal components, we detected significant differences in wing pattern between *L. idas* and *L. melissa* ($Pillali_{1,184} = 0.4149$, $P < 0.0001$) and between males and females ($Pillali_{1,184} = 0.3948$, $P < 0.0001$), as well as a significant interaction between these two factors ($Pillali_{1,184} = 0.0628$, $P < 0.0381$). Inspection of treatment means for wing pattern principal components indicated that this interaction occurred because the difference between *L. idas* and *L. melissa* wing patterns was greater in males than in females (Fig. S2, Supporting Information).

Linear discriminant function analysis based on male genitalic morphology correctly classified all but a single individual from allopatric populations as *L. idas* or *L. melissa* (99.5% classified correctly; Table S1, Supporting Information). Many individuals were also correctly classified as *L. idas* or *L. melissa* based on wing pattern (males: 88.5% correctly classified; females: 73.7% correctly classified). Coefficients of linear discriminants indicate that falx length was most important for classifying individuals based on male genitalia (standardized coefficients: falx = 2.116, uncus = 0.350, humerulus = −0.002). A number of characters were important for classifying individuals based on wing pattern and interestingly these characters differed between males and females. Specifically, the area of Rs and b2 of a2 were most important for distinguishing between *L. idas* and *L. melissa* males, whereas the area of M, M_1 and $Cu_1 + 1A$ were most important for distinguishing between *L. idas* and *L. melissa* females (Table S2, Supporting Information).

Based on analyses of male genitalia measurements using fuzzy c-means clustering, many individuals had membership primarily in one of two morphological clusters corresponding to *L. idas* and *L. melissa* (i.e. memberships >0.9; Table 2). However, regardless of the value used for *r*, many morphologically intermediate individuals were identified as well (19.7–49.3% of individuals). Results were qualitatively similar for fuzzy c-means clustering based on male wing pattern, although more individuals with membership spread among both clusters were identified, and with $r \geq 1.75$ no individuals had a membership of >0.9 in a single cluster (Table 2). Even fewer individuals had membership primarily in one cluster based on female wing patterns. Contrary to results based on male genitalic morphology and male wing pattern, the clusters defined based on female wing patterns did not clearly correspond to *L. idas* and *L. melissa* (Table 2).

Table 2 Proportion of *Lycaeides* individuals classified as *L. idas*, *L. melissa* and admixed based on c-means fuzzy clustering with $k = 2$

Character	Cluster	$r = 1.25$	$r = 1.5$	$r = 1.75$	$r = 2.0$
Male genitalia	*L. idas*	NA	0.479	0.449	0.365
	L. melissa	NA	0.323	0.245	0.142
	Admixed	NA	0.198	0.306	0.493
Male wing pattern	*L. idas*	0.401	0.179	0.000	0.000
	L. melissa	0.297	0.052	0.000	0.000
	Admixed	0.302	0.769	1.000	1.000
Female wing pattern	Non-admixed	0.650	0.217	0.000	0.000
	Admixed	0.351	0.784	1.000	1.000

As with our analysis of admixture proportions, we failed to reject the null hypothesis of a unimodal distribution for fuzzy cluster membership of core hybrid zone *Lycaeides* based on male genitalic morphology ($r = 1.5$: dip = 0.03406, $P = 0.5607$; $r = 1.75$: dip = 0.03303, $P = 0.6134$; $r = 2$: dip = 0.03323, $P = 0.6044$). Unlike the admixture proportion data, male genitalia in the core hybrid zone were more *L. melissa*-like with a median membership in the *L. melissa* cluster of 0.7443–0.8978 (depending on *r*; Fig. S3B, Supporting Information). Finally, we found no evidence that fuzzy cluster membership based on male wing patterns followed a multimodal distribution ($r = 1.25$: dip = 0.04300, $P = 0.6260$; $r = 1.5$: dip = 0.04091, $P = 0.7132$; $r = 1.75$: dip = 0.03655, $P = 0.8715$; $r = 2$: dip = 0.02800, $P = 0.9960$). The median membership in the *L. melissa* fuzzy cluster varied substantially as a function of *r* (0.0285–0.3738), but generally suggested that the core hybrid zone was dominated by *L. idas*-like male wing patterns (Fig. S3C, Supporting Information).

Cline analysis and geographical patterns of admixture

The geographical clines in genomic composition and male genitalia size were best described by unimodal character distribution models (Table 3). Given the unimodal character distribution model, unconstrained

Table 3 Geographical cline model comparison for each quantitative character

Character	Model	Parameters	ln likelihood	AIC_C	Δ_{AIC}	AIC weights
Genomic composition	Bimodal	6	−253.7738	519.9050	65.9176	∼0
	Trimodal	10	−234.9276	490.8076	36.8201	∼0
	Unimodal	7	−219.7544	453.9874	0	∼1
Male genitalia size	Bimodal	6	−454.3228	920.8979	169.1827	∼0
	Trimodal	10	−399.9137	820.4961	68.780	∼0
	Unimodal	7	−368.6889	751.7151	0	∼1
Aurorae and spot size	Bimodal	10	−674.0406	1368.8990	0	0.7697
	Trimodal	18	−667.5497	1373.7201	4.8210	0.0691
	Unimodal	12	−673.4283	1372.0252	3.1260	0.1612
Relative spot size	Bimodal	10	−537.6326	1096.0830	1.5339	0.3169
	Trimodal	18	−534.6530	1107.9266	13.3775	0.0008
	Unimodal	12	−534.6903	1094.5490	0	0.6823

AIC, Akaike Information Criterion.

maximum likelihood estimates for the centre (defined as the inverse of the slope) and width of the cline in genomic composition were km 9.77 and 256.13 km respectively (Fig. 4A). Maximum likelihood estimates for the centre and width of the geographical cline in male genitalia size were km 38.21 and 227.39 km respectively (Fig. 4B). Clines in aurorae and spot size and in relative spot size were consistent with both uni-

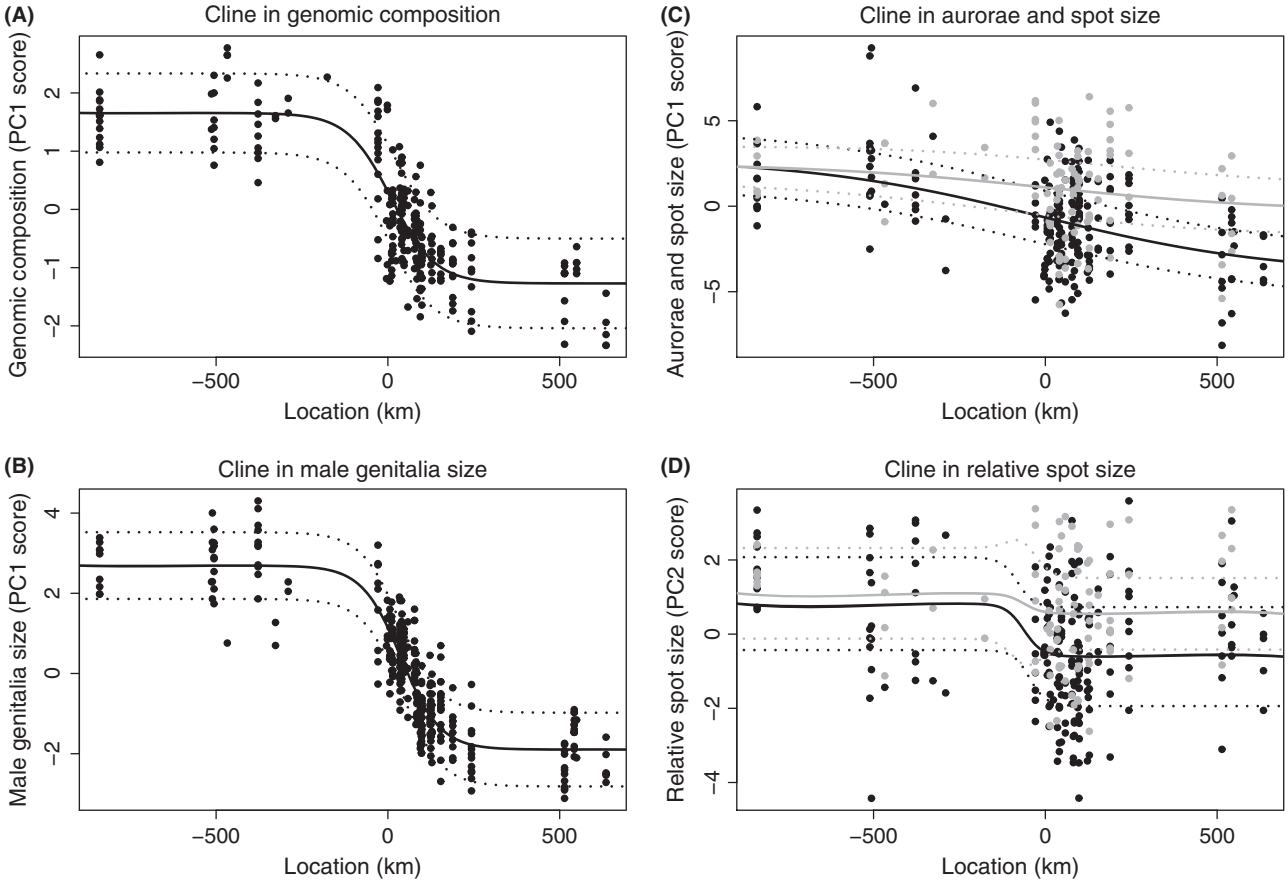

Fig. 4 Geographical cline in genomic composition (A), male genitalia size (B), wing aurorae and spot size (C) and relative spot size (D). *Lycaeides melissa* populations are in the south-east (negative location values) and *L. idas* populations are in the north-west (positive location values). Lines denote mean expected character values (solid) and standard deviations in character values (dotted) along the geographical cline based on the unimodal cline model for males (black) and females (grey). Solid circles depict individual character values.

modal and bimodal character distribution models (i.e. $\Delta_{AIC}\sim3$ or less and moderate AIC weights; Table 3). In both cases the likelihood of the unimodal model was greater than the likelihood of the bimodal model, although AIC suggested that the latter should be preferred slightly for aurorae and spot size. The likelihood surfaces for the slope and centre of the cline in aurorae and spot size under the bimodal model were relatively flat, although the likelihood increased slightly as the slope approached ∞ (i.e. as the cline width approached 0 km). As the data were described well by the unimodal model and because of difficulties obtaining reasonable parameter estimates under alternative models, we report the results from the unimodal model. The unimodal model was also used for the constrained analysis. Given the unimodal model, maximum likelihood estimates of cline centre and width were km −13.81 and 1392.57 km for aurorae and spot size and km −62.25 and 106.73 km for relative spot size (Fig. 4C,D).

Morphological clines were not generally coincident or concordant with the cline in genomic composition. Geographical clines in genomic composition and male genitalia size were better fit with models allowing for different cline centres or different slopes and cline centres than for models forcing these clines to have the same centre (Table 4). We obtained similar results in our comparison of clines in genomic composition and relative wing pattern spot size (Table 4). A model forcing genomic composition and wing aurorae and spot size to have the same centre and slope was less likely than alternative models allowing for different slopes or centres. However, we could not discriminate among these alternative models (Table 4).

The relationship between genomic composition and male genitalia size was better described by a cubic function than a linear function ($D = 7.4880$, d.f. $= 2$, $P = 0.0237$). However, wing aurorae and spot size ($D = 2.6674$, d.f. $= 2$, $P = 0.2635$) and relative spot size ($D = 3.2376$, d.f. $= 2$, $P = 0.1981$) were better fit by linear functions. Genomic composition and the cube of genomic composition were significant predictors of male genitalia size ($t = 10.934$, $P < 0.0001$ and $t = -2.731$, $P = 0.0070$ respectively). Moreover, genomic composition explained 65% of the observed variation in male genitalia size (Fig. 5A). Sex and genomic composition were significant predictors of wing aurorae and spot size ($t = -3.587$, $P = 0.0004$ and $t = 4.896$, $P < 0.0001$ respectively). However, this model explained only 18% of the variation in aurorae and spot size (Fig. 5B). A similar pattern was detected for relative wing spot size (sex: $t = -3.443$, $P = 0.0007$; genomic composition: $t = 3.200$, $P = 0.0016$; $r^2 = 0.12$; Fig. 5C).

Ecological context and hybrid zone structure

Based on the permutation test implemented in *RxC*, we detected significant LD for a single locus pair in each of seven populations. However, none of these estimates of LD remained significant following false discovery rate correction (Benjamini & Hochberg 1995). The mean LD for a sampling locality was not dependent on its geographical location (overall model: $F_{2,23} = 0.9355$, $P = 0.4068$; quadratic term: $t = 1.237$, $P = 0.2280$; Fig. S4, Supporting Information). The absolute deviation of an individual's genomic composition from the sampling locality median had a marginally significant relationship with geographical location ($F_{2,234} = 2.686$, $P = 0.0703$, $r^2 = 0.0224$; Fig. 6A) and the quadratic term for this model was significantly different from zero ($t = -2.213$, $P = 0.0279$). The absolute deviation of male genitalia size from each sampling locality median was also dependent on geographical location ($F_{2,325} = 2.868$, $P = 0.0583$, $r^2 = 0.0173$; Fig. 6B), but not on the quadratic term in the model ($t = -1.206$, $P = 0.2287$). The absolute deviations

Table 4 Model comparison for coincidence and concordance of clines

Character	Model	Parameters	ln likelihood	AIC_C	Δ_{AIC}	AIC weights
Male genitalia size	Slopes and centres equal	12	−592.8902	1210.3287	6.3770	0.0276
	Centres equal	13	−592.5683	1211.7774	7.8258	0.0134
	Slopes equal	13	−588.6554	1203.9520	0	0.6694
	Unconstrained	14	−588.4434	1204.8867	1.6758	0.2896
Aurorae and spot size	Slopes and centres equal	17	−899.2350	1833.5552	6.0312	0.0214
	Centres equal	18	−895.8909	1828.9967	1.4727	0.2088
	Slopes equal	18	−895.4219	1828.0587	0.5347	0.3338
	Unconstrained	19	−894.0858	1827.5240	0	0.4361
Relative spot size	Slopes and centres equal	17	−760.6814	1556.5770	8.1737	0.0088
	Centres equal	18	−757.5909	1552.5416	4.1382	0.0659
	Slopes equal	18	−755.7772	1548.9142	0.5109	0.4039
	Unconstrained	19	−754.4447	1546.8890	0	0.5215

AIC, Akaike Information Criterion.

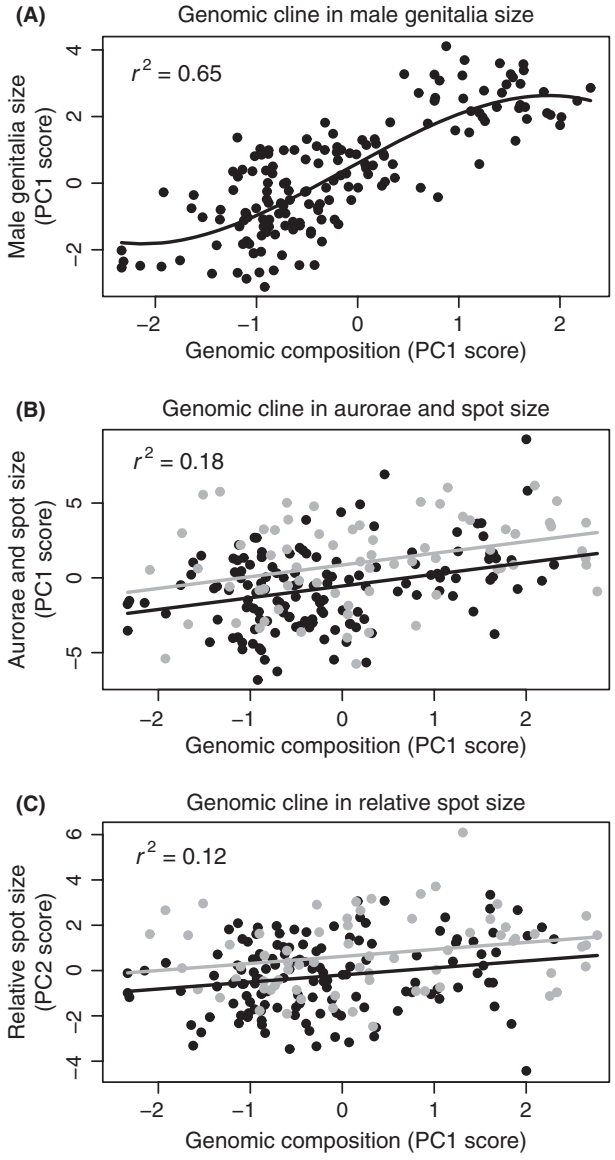

Fig. 5 Genomic cline in male genitalia size (A), wing aurorae and spot size (B), and relative spot size (C). Principal component scores are polarized so that *Lycaeides melissa* individuals have positive values. Lines denote mean expected character values and solid circles depict individual character values for males (black) and females (grey).

of wing aurorae and spot size and relative wing spot size from sampling localities were not dependent on geographical location (aurorae and spot size: $F_{2,181} = 0.2219$, $P = 0.8012$; relative spot size: $F_{2,181} = 0.6277$, $P = 0.5350$; Fig. 6C,D). Similar results were obtained when LD and character deviations were modelled as a function of admixture proportions (results not shown).

We detected a significant positive correlation between pair-wise differences in mean population admixture proportion and Euclidean distance in climate while controlling for pair-wise geographical distance ($r = 0.461$, $P = 0.0024$). However, this pattern was driven primarily by the presence of two groups of populations: (i) three populations at the south-eastern edge of the hybrid zone (DBS, TSS and SWC) with *L. melissa*-like climate and genomic compositions and (ii) 10 populations covering the remainder of the hybrid zone with *L. idas*-like climate and genomic compositions. When we analysed only the second group of populations, we no longer detected a correlation between admixture proportions and climate ($r = 0.229$, $P = 0.1268$). We did not detect differences in precipitation among localities occupied by *L. idas*, *L. melissa*, and admixed populations ($F_{2,35} = 3.141$, $P = 0.0556$). However, we found that *L. melissa* populations occurred in sites with significantly warmer mean daily maximum and mean daily mean July temperatures than *L. idas* or admixed populations (maximum July temperature: $F_{2,35} = 30.936$, $P < 0.0001$; mean July temperature: $F_{2,35} = 48.39$, $P < 0.0001$). Localities occupied by *L. idas* and admixed populations did not differ for either of these characters based on Tukey's HSD. Mean daily minimum January temperature differed among *L. idas*, *L. melissa*, and admixed population sites, with *L. idas* sites having the coolest temperatures and *L. melissa* sites having the warmest temperatures ($F_{2,35} = 12.719$, $P < 0.0001$). Consistent with these differences, for admixed populations we found a positive correlation between mean population admixture proportions and each of the temperature variables, indicating that warmer sites within the hybrid zone had more *L. melissa*-like individuals (maximum July temperature: $r = 0.579$, $P = 0.0383$; mean July temperature: $r = 0.604$, $P = 0.0287$; minimum January temperature: $r = 0.705$, $P = 00071$).

Discussion

Our results indicate *L. idas* and *L. melissa* have hybridized in the Rocky Mountains and that reproductive isolation is insufficient to prevent introgression for much of the genome. Admixture has been particularly prevalent in the vicinity of the Jackson Hole valley and Yellowstone plateau. Hybridization between *L. idas* and *L. melissa* has led to the formation of a hybrid zone. The hybrid zone is relatively wide given the dispersal abilities of *Lycaeides* butterflies and does not show strong evidence of cline concordance. We believe the structure of the *Lycaeides* hybrid zone might be best explained by the patchy distribution of *Lycaeides*, local extinction and recolonization of habitat patches, environmental variation and weak overall selection against hybrids. Unlike hybridization between *L. idas* and *L. melissa* in the Sierra Nevada, Siskyou Mountains, Warner Mountains and White Mountains (Gompert *et al.* 2006a, 2008b, 2010, Nice *et al.* unpublished), hybridization between Rocky Mountain *L. idas* and *L. melissa* has not resulted

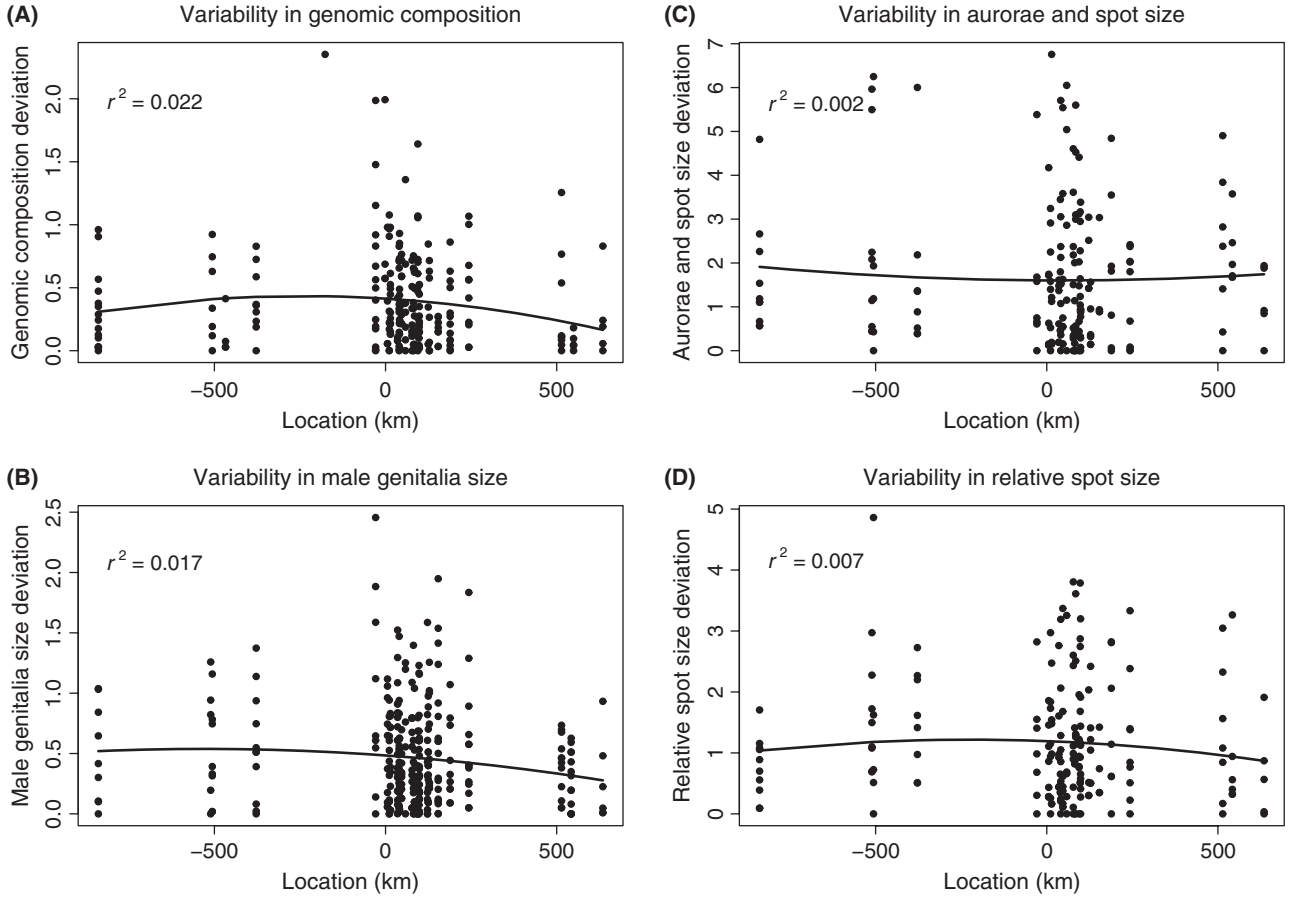

Fig. 6 Scatterplot depicting the relationship between geographical location and individual deviations from median genomic composition (A), male genitalia size (B), wing aurorae and spot size (C) and relative spot size (D). Solid lines denote the expected values from quadratic models.

in the establishment of an isolated hybrid lineage or species. Finally, we found little evidence that wing pattern variation contributes to reproductive isolation. However, our results are consistent with the hypothesis that differences in male genitalia size have contributed to isolation or that *L. melissa*-like genitalia tend to increase fitness relative to *L. idas*-like genitalia, but other valid explanations for these patterns exist as well. In the following paragraphs we provide a detailed discussion of our results and their implications.

Outcome and consequences of hybridization

Bayesian admixture proportions estimated from our molecular data suggest that admixed individuals are relatively common, particularly in and around the Jackson Hole valley and to a lesser extent on the Yellowstone plateau (Fig. 3A). These findings are supported by the presence of many morphologically intermediate individuals (inferred from fuzzy c-means clustering) in the same region (Fig. S3, Supporting Information; Table 2). The admixed individuals range from *L. idas*-like to *L. melissa*-like, with many individuals containing roughly equal genetic contributions from both. The distribution of admixture proportions and morphological cluster memberships within the *L. idas* × *L. melissa* hybrid zone was not different from a unimodal distribution. This result was generally supported by model comparison of geographical clines, although bimodal models could not be rejected in some instances (Table 3). Unimodal admixture and character distributions are expected when reproductive isolation, particularly pre-zygotic isolation, between hybridizing populations is weak (Jiggins & Mallet 2000; Gay *et al.* 2008). Therefore, our results suggest relatively weak reproductive isolation between *L. idas* and *L. melissa*. This conclusion is further supported by weak evidence for concordance among geographical clines, which would not be expected if there was a strong barrier to gene flow (Barton 1983). It is still possible that specific hybrid genotypes or phenotypes have experienced relatively strong selection in this hybrid zone (e.g. all admixed populations we have examined appear to be univoltine like *L. idas*, not bivoltine like *L. melissa*;

personal observation). Our results simply suggest that most of the genome is not prevented from introgressing by strong selection. Despite the weak isolation inferred between Rocky Mountain *L. idas* and *L. melissa* populations, many of these populations remain genetically and morphologically distinct. The hypothesis that divergence and differentiation are possible despite substantial gene flow has received considerable recent support in the literature (e.g. Emelianov *et al.* 2004; Fitzpatrick *et al.*, 2008; Niemiller *et al.* 2008; Nadachowska & Babik 2009). Although our results are consistent with this hypothesis (i.e. most *L. idas* and *L. melissa* populations remain distinct despite considerable gene flow), additional time and dispersal of admixed individuals might further erode the geographical region where *L. idas* and *L. melissa* remain differentiated.

Our data indicate that secondary contact between *L. idas* and *L. melissa* in the Rocky Mountains has resulted in the formation of a hybrid zone. Specifically, genomic composition and morphological characters vary from *L. melissa*-like to *L. idas*-like along a geographical transect and the transition for genomic composition and male genitalia size has a clear sigmoidal form that is often associated with hybrid zones (Fig. 4; e.g. Endler 1977; Barton & Hewitt 1985; Mallet *et al.* 1990; Ruegg 2008; Brennan *et al.* 2009; Teeter *et al.* 2010). The structure of this *Lycaeides* hybrid zone differs in subtle ways from expectations from classic models for the structure and maintenance of hybrid zones. Unlike hybrid zones studied in many taxonomic groups (e.g. *Bombina*, Szymura & Barton 1986; *Mus*, Macholan *et al.* 2007; *Larus*, Gay *et al.* 2008; *Senecio*, Brennan *et al.* 2009; *Microcebus*, Gligor *et al.* 2009), we found little evidence of elevated LD or character variability in the centre of the hybrid zone, which is contrary to expectations from dispersal-dependent hybrid zone models (Endler 1977; Barton & Hewitt 1985, Fig. 6; Fig. S4, Supporting Information). Interestingly, the hybrid zone centre does not correspond to an area of low population density (in fact, admixed populations have some of the highest densities of individuals we have observed; personal observation, Gompert Z), contrary to predictions from tension zone models and empirical patterns from other well-studied hybrid zones (Bazykin 1969; Barton 1979; Barton & Hewitt 1985; Nichols & Hewitt 1986; Hewitt 1989; Ruegg 2008). Our results suggest that admixed individuals generally occupy sites with similar habitat to *L. idas* populations and feed on *L. idas* larval host plants (e.g. *Astragalus* sp.). This distribution is inconsistent with the dispersal-independent model of bounded hybrid superiority within narrow ecotones proposed by Moore (1977) and suggested for the Great Plains hybrid zone between Red- and Yellow-shafted flickers (Moore & Buchanan 1985; Moore & Koenig 1986). Our results indicate that the Rocky Mountain hybrid zone should not be regarded as a mosaic hybrid zone despite the patchy distribution of *Lycaeides* and a correlation between local climate and the genetic makeup of individual populations. This is because the hybrid zone is not composed of interdigitated patches of more *L. idas*-like and more *L. melissa*-like habitat with corresponding admixed populations having more *L. idas* and more *L. melissa* ancestry, but instead has a south-eastern region with more *L. melissa*-like habitat and ancestry with the remainder of the zone behind composed of more *L. idas*-like individuals occupying more *L. idas*-like habitat.

The patchy distribution of *Lycaeides* populations coupled with local extinction and colonization of habitat patches (i.e. metapopulation dynamics) might explain the structure of the Rocky Mountain *Lycaeides* hybrid zone. Geographical clines in genomic composition, male genitalia size, aurorae and spot size, and relative spot size are relatively wide. Following Endler's equation for the width of a neutral character cline, which assumes a continuously distributed population (Endler 1977), even with a very generous mean dispersal distance of 500 m per generation, 20 882 generations of admixture would be required for the 256.13 km cline in genomic composition. Assuming one generation per year (the generation time of *L. idas* and admixed populations), this estimate amounts to approximately 6000 years more than the maximum age for the hybrid zone based on patterns of glaciation (Harris *et al.* 1997). Moreover, genomic composition is unlikely to be completely neutral and secondary contact may have occurred much more recently than 14 000 years BP. The historical range of *L. melissa* might have been substantially smaller than the present range. Range expansion of *L. melissa* may have been facilitated by the colonization of, or host switch to, alfalfa, *M. sativa*, which was introduced to North America approximately 200 years ago and has since become widespread (Michaud *et al.* 1988). Utilization of alfalfa might have been the mechanism generating secondary contact between *L. idas* and *L. melissa* in the Rocky Mountains relatively recently. Although it is possible that mean dispersal for *Lycaeides* is much greater than 500 m, this seems unlikely based on the extensive field work conducted on the U.S. endangered Karner blue butterfly (U.S. Fish and Wildlife Service 2003). However, as pointed out by Hewitt (1989), local extinction and recolonization of habitat patches of a patchily distributed species in a hybrid zone could result in much wider clines than would be predicted for a continuously distributed species. This is because colonizing individuals have a much greater effect on local allele frequencies than individuals dispersing across habitat with higher population densities. A patchy distribution coupled with local extinction and recolonization likely account

for the wide hybrid zone between the flightless grasshoppers *Chorthippus erythropus* and *C. paralellus* in the Pyrenees range (Hewitt 1989). This might explain the wide geographical clines in the *Lycaeides* hybrid zone and might also account for our lack of evidence for increased LD and character variability in the centre of the *Lycaeides* hybrid zone. Rare colonization events coupled with overall low dispersal among populations would provide sufficient time for recombination to erode LD and genetic drift to reduce admixture-induced character variability. Thus, we propose the hypothesis that the structure of the Rocky Mountain *Lycaeides* hybrid zone represents the outcome of local extinction-recolonization dynamics (i.e. dispersal in a metapopulation context) coupled with weak selection against some hybrid genotypes. Selection is likely both environment-dependent and environment-independent.

The Rocky Mountain *Lycaeides* hybrid zone has likely persisted for some time (given the width of the zone), and has not yet resulted in the formation of an independent hybrid lineage. Our molecular data were better explained by two gene pools or populations than three, and even when the existence of three gene pools was assumed the third gene pool did not correspond to populations with admixed individuals (Fig. 3B; Fig. S1, Supporting Information). This conclusion is further supported by the results from geographical cline analyses, which indicated our molecular and morphological data were better described by unimodal than trimodal distributions (Table 3). Thus, the results of this study differ substantially from outcomes of hybridization between *Lycaeides* populations elsewhere in North America. Hybridization between *L. i. anna* (also known as *L. anna*) and *L. melissa* in the Sierra Nevada Mountains resulted in the evolution of an isolated ecologically distinct hybrid lineage associated with alpine-habitat (Gompert *et al.* 2006a). A similar, perhaps independent hybrid lineage exists above tree-line in the White Mountains and two other, likely independent ecologically unique hybrid lineages exists in the Warner Mountains and Siskyou Mountains (Gompert *et al.* 2008b, 2010; Nice C, Gompert Z, Fordyce J and Forister M, unpublished data). Why have these hybridization events resulted in the establishment of differentiated hybrid lineages, whereas hybridization between *L. idas* and *L. melissa* in the Rocky Mountains has resulted in a series of admixed populations along a patchy hybrid zone? This difference might be due to the fact that the western hybrid lineages generally occupy habitat or use host plants not utilized by other nearby *Lycaeides* populations (Gompert *et al.* 2006a). For example, many of the hybrid lineages in the western United States occupy higher elevations than other *L. idas* and *L. melissa* populations, which is not necessarily the case for the hybrid populations in the Rocky Mountains (Table 1). In some cases these western hybrid lineages posses unique traits that are adaptive in these novel habitats (e.g. a lack of egg adhesion in the Sierra Nevada and White Mountains, Fordyce & Nice 2003; Gompert *et al.* 2006a). Conversely, the admixed Rocky Mountain populations occupy similar habitats and use the same host plants as nearby *L. idas* populations. The use of similar habitats and the same host plants by admixed and *L. idas* populations might have contributed to gene flow and prevented divergence among these populations. Differences in population connectivity in the Sierra Nevada and Rocky Mountains might have contributed to these different outcomes as well. For example, the likelihood of hybrid speciation might have been higher in the Sierra Nevada relative to the Rocky Mountains if habitat patches of the former experienced less connectivity (Buerkle *et al.* 2000). However, these are just a few of many synergistic explanations for these different consequences of hybridization. Some alternatives include differences in: (i) the genetic architecture of isolation among different *Lycaeides* populations, (ii) the time since initial admixture and (iii) various stochastic and historical factors (i.e. deterministic forces might not be responsible for these different outcomes). The outcome of hybridization in the Rocky Mountains also differs from the pattern of recent mitochondrial introgression in the absence of substantial nuclear introgression observed between *L. melissa* and the Karner blue in the eastern United States and between *L. melissa* and several hybrid lineages in the western United States (e.g. the White and Warner Mountains populations; Gompert *et al.* 2006b, 2008b). This difference might indicate that more of the genome is involved in isolation in cases where introgression is limited to mitochondrial DNA than in the Rocky Mountains where nuclear introgression is also prevalent. This possibility is supported by the relatively low level of genetic differentiation between Rocky Mountain *L. idas* and *L. melissa* (Gompert *et al.*, 2010). However, it is not clear whether the relatively low genetic differentiation between these Rocky Mountain *Lycaeides* is the cause or effect of nuclear introgression.

Phenotypic basis of isolation and adaptive introgression

Geographical cline analyses suggest that the geographical region where the transition occurs from *L. idas* to *L. melissa* male genitalia size is shifted approximately 28 km to the north-west relative to the transition in genomic composition (Fig. 4; Table 4). These analyses also provide some support the possibility that this geographical transition in male genitalia size occurs over a more narrow region than the transition in genomic

composition (227.39 km as opposed to 256.13 km). This lack of concordance is further supported by the better fit of a cubic model than a simple linear model to the relationship between genomic composition and male genitalia size (i.e. male genitalia size remains relatively constant over a range of high and low genomic composition scores). The north-western shift in this geographical transition might indicate that selection has favoured the geographical spread of *L. melissa*-like male genitalia. However, this possibility is inconsistent with general expectations that male insect genitalia are under balancing selection (e.g. Eberhard *et al*. 1998; but see Bertin & Fairbairn 2007). An alternative explanation for this shift is that some or all *L. melissa* alleles affecting male genitalia size are dominant to *L. idas* alleles such that heterozygotes display *L. melissa*-like male genitalia. This possibility cannot be evaluated at present, as the genetic basis of male genitalia size in *Lycaeides* is unknown. Geographical cline theory suggests that, in the absence of a strong barrier to gene flow, clines for characters involved in reproductive isolation should be narrower than clines for neutral characters (Barton 1983; Barton & Hewitt 1985). Thus, the more rapid geographical transition in male genitalia size than genomic composition might indicate that differences in male genitalia size contribute to isolation between *L. idas* and *L. melissa*. Alternatively, this difference might reflect the genetic architecture of male genitalia size or phenotypic plasticity. Although the current data and results do not allow for definitive conclusions regarding male genitalia and selection in *Lycaeides*, these results suggest that this character warrants further study.

The geographical cline in wing aurorae and spot size was very wide (1392.57 km) and the relationship between genomic composition and wing aurorae and spot size was linear (Figs 4 and 5). Taken together, these results do not suggest that differences in wing aurorae and spot size contribute to isolation between Rocky Mountain populations of *L. idas* and *L. melissa*. The cline centre for aurorae and spot size was shifted about 24 km to the south-east relative to the cline for genomic composition (Table 4). Although this shift might reflect adaptive introgression of *L. idas*-like smaller aurorae and spots, it could also be explained by plasticity or dominance of *L. idas* alleles and seems somewhat trivial relative to the great geographical distance required for this transition. Finally, unlike male genitalia size, aurorae and spot size varies much more within species and less between species, which would not be expected if these characters were involved in reproductive isolation (Fig. 4). The cline in relative wing spot size was shifted 72 km to the southeast of the cline in genomic composition. However, the overall difference in relative wing pattern size between *L. idas* and *L. melissa* was minimal, particularly for females. Thus, we do not believe that this apparent geographical shift should be interpreted as evidence that selection has favoured *L. idas*-like relative spot size. Finally, although geographical cline analysis indicated that the cline in relative spot size might be significantly more narrow than the cline in genomic composition, this was not supported by the relationship between genomic composition and relative spot size, which was linear (Figs 4 and 5; Table 4). These results, in conjunction with the limited difference in relative spot size between *L. idas* and *L. melissa*, suggest that relative spot size has little or no contribution to reproductive isolation. However, these results do not necessarily mean that wing pattern differences do not contribute to isolation between these species. Wing pattern elements might simply be highly plastic, and thus wing pattern clines could reflect low heritability of wing pattern elements (relative to male genitalia size). Specific wing pattern elements, such as the areas of spots M and M_1, are highly differentiated between female *L. idas* and *L. melissa* and might be important for mate recognition and isolation even if the composite characters we analysed (i.e. wing aurorae and spot size and relative spot size) are not involved in isolation. We have observed subtle differences in the colour of wing pattern elements between *L. idas* and *L. melissa* that might contribute to isolation as well (personal observation, Gompert Z). Finally, preliminary experimental data involving paper wing models suggest that males from some admixed and *L. idas* populations in the Rocky Mountains approach *L. idas*-like female wing patterns more frequently than *L. melissa*-like wing patterns (unpublished). Thus, wing pattern variation might contribute to isolation despite the results of this study.

Conclusions

This study provides another dimension to the varied outcomes of hybridization previously observed in North American *Lycaeides* butterflies. Whereas hybridization in other geographical regions led to the establishment of novel evolutionary lineages (Gompert *et al*. 2006a, 2010; unpublished) or resulted in mitochondrial introgression with little to no nuclear gene flow (Gompert *et al*. 2006b, 2008a, b), hybridization between *L. idas* and *L. melissa* in the Rocky Mountains has led to substantial admixture and nuclear introgression over a relatively large geographical area without the establishment of isolated hybrid lineages. The structure of the hybrid zone formed between *L. idas* and *L. melissa* in the Rocky Mountains is interesting and likely reflects the patchy distribution of *Lycaeides* butterflies. The extent of admixture observed in the Rocky Mountains suggests much of the *Lycaeides* genome is not prevented from

introgressing between *L. idas* and *L. melissa* populations by reproductive isolation. This does not mean that these lineages are not isolated to some degree, but simply that the overall strength of isolation is weak or involves few characters and genetic regions. Finally, we found tentative evidence that male genitalia size might be one of potentially many characters contributing to reproductive isolation between *L. idas* and *L. melissa*.

Acknowledgements

This manuscript was improved by comments from two anonymous reviewers. We are indebted to the following for help collecting *Lycaeides* or donating collected specimens: R. Lund, M. Moore, P. Opler and C. Schmidt. We would also like to thank the Nevada Genomics Center at the University of Nevada and C. A. Buerkle for use of laboratory facilities and valuable discussion. Z.G. was funded by a NSF graduate research fellowship, a NSF EPSCoR WySTEP summer fellowship, a NSF EPSCoR Ecology Project research grant and a graduate student research award from the Society of Systematic Biologists. L.K.L. was funded by a NSF EPSCoR WySTEP summer fellowship. C.C.N. was funded by Texas State University, J.A.F. was funded by the University of Tennessee and M.L.F. was funded by the University of Nevada.

References

Anthony N, Gelembiuk G, Raterman D, Nice C, ffrench Constant R (2001) Isolation and characterization of microsatellite markers from the endangered Karner blue butterfly *Lycaeides melissa samuelis* (Lepidoptera). *Hereditas*, **134**, 271–273.

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

Arnqvist G, Thornhill R, Rowe L (1997) Evolution of animal genitalia: morphological correlates of fitness components in a water strider. *Journal of Evolutionary Biology*, **10**, 613–640.

Avise JC, Pierce PC, VandenAvyle MJ, Smith MH, Nelson WS, Asmussen MA (1997) Cytonuclear introgressive swamping and species turnover of bass after an introduction. *Journal of Heredity*, **88**, 14–20.

Barton NH (1979) The dynamics of hybrid zones. *Heredity*, **43**, 341–359.

Barton NH (1983) Multilocus clines. *Evolution*, **37**, 454–471.

Barton NH (1993) The probability of fixation of a favoured allele in a subdivided population. *Genetical Research, Cambridge*, **62**, 149–157.

Barton NH (2001) The role of hybridization in evolution. *Molecular Ecology*, **10**, 551–568.

Barton NH, Hewitt GM (1981) A chromosomal cline in the grasshopper *Podisma pedestris*. *Evolution*, **35**, 1008–1018.

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

Bazykin AD (1969) Hypothetical mechanism of speciation. *Evolution*, **23**, 685–687.

Beaumont M, Barratt E, Gottelli D *et al.* (2001) Genetic diversity and introgression in the Scottish wildcat. *Molecular Ecology*, **10**, 319–336.

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B-Methodological*, **57**, 289–300.

Bertin A, Fairbairn DJ (2007) The form of sexual selection on male genitalia cannot be inferred from within-population variance and allometry – a case study in *Aquarius remigis*. *Evolution*, **61**, 825–837.

Bezdek JC (1981) *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press, Norwell, MA, USA.

Brennan AC, Bridle JR, Wang AL, Hiscock SJ, Abbott RJ (2009) Adaptation and selection in the *Senecio* (Asteraceae) hybrid zone on Mount Etna, Sicily. *New Phytologist*, **183**, 702–717.

Bridle JR, Butlin RK (2002) Mating signal variation and bimodality in a mosaic hybrid zone between *Chorthippus* grasshopper species. *Evolution*, **56**, 1184–1198.

Brock JP, Kaufman K (2003) *Butterflies of North America*. Hillstar Editions L. C., New York, NY, USA.

Brown MB, Forsythe AB (1974) Robust tests for equality of variances. *Journal of the American Statistical Association*, **69**, 364–367.

Buerkle CA, Lexer C (2008) Admixture as the basis for genetic mapping. *Trends in Ecology & Evolution*, **23**, 686–694.

Buerkle CA, Morris RJ, Asmussen MA, Rieseberg LH (2000) The likelihood of homoploid hybrid speciation. *Heredity*, **84**, 441–451.

Butlin R (1987) Speciation by reinforcement. *Trends in Ecology & Evolution*, **2**, 8–13.

Carling MD, Brumfield RT (2008) 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.

Chamberlain NL, Hill RI, Kapan DD, Gilbert LE, Kronforst MR (2009) Polymorphic butterfly reveals the missing link in ecological speciation. *Science*, **326**, 847–850.

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

Daly C, Halbleib M, Smith JI *et al.* (2008) Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *International Journal of Climatology*, **28**, 2031–2064.

Duenez-Guzman EA, Mavarez J, Vose MD, Gavrilets S (2009) Case studies and mathematical models of ecological speciation. 4. Hybrid speciation in butterflies in a jungle. *Evolution*, **63**, 2611–2626.

Eberhard WG (1985) *Sexual Selection and the Evolution of Animal Genitalia*. Harvard University Press, Cambridge, MA, USA.

Eberhard W (1993) Evaluating models of sexual selection – genitalia as a test-case. *The American Naturalist*, **142**, 564–571.

Eberhard WG, Huber BA, Rodriguez RL, Briceno RD, Salas I, Rodriquez V (1998) One size fits all? Relationships between the size and degree of variation in genitalia and other body parts in twenty species of insects and spiders. *Evolution*, **52**, 415–431.

Ellstrand NC (2003) Current knowledge of gene flow in plants: implications for transgene flow. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **358**, 1163–1170.

Ellstrand NC, Whitkus R, Rieseberg LH (1996) Distribution of spontaneous plant hybrids. *Proceedings of the National*

Academy of Sciences of the United States of America, **93**, 5090–5093.

Emelianov I, Marec F, Mallet J (2004) Genomic evidence for divergence with gene flow in host races of the larch budmoth. *Proceedings of the Royal Society B-Biological Sciences*, **271**, 97–105.

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

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

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

Felsenstein J (1981) Skepticism towards Santa Rosalia, or why are there so few kinds of animals? *Evolution*, **35,** 124–138.

Fitzpatrick BM, Shaffer HB (2007) Hybrid vigor between native and introduced salamanders raises new challenges for conservation. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 15793–15798.

Fitzpatrick BM, Placyk JS, Niemiller ML, Casper GS, Burghardt GM (2008) Distinctiveness in the face of gene flow: hybridization between specialist and generalist gartersnakes. *Molecular Ecology*, **17**, 4107–4117.

Fordyce JA, Nice CC (2003) Variation in butterfly egg adhesion: adaptation 968 to local host plant senescence characteristics? *Ecology Letters*, **6**, 23–27.

Fordyce JA, Nice CC, Forister ML, Shapiro AM (2002) The significance of wing pattern diversity in the Lycaenidae: mate discrimination by two recently diverged species. *Journal of Evolutionary Biology*, **15**, 871–879.

Gay L, Neubauer G, Zagalska-Neubauer M *et al.* (2007) Molecular and morphological patterns of introgression between two large white-headed gull species in a zone of recent secondary contact. *Molecular Ecology*, **16**, 3215–3227.

Gay L, Crochet PA, Bell DA, Lenormand T (2008) Comparing clines on molecular and phenotypic traits in hybrid zones: a window on tension zone models. *Evolution*, **62**, 2789–2806.

Gligor M, Ganzhorn JU, Rakotondravony D *et al.* (2009) Hybridization between mouse lemurs in an ecological transition zone in southern Madagascar. *Molecular Ecology*, **18**, 520–533.

Gompert Z, Buerkle CA (2009) A powerful regression-based method for admixture mapping of isolation across the genome of hybrids. *Molecular Ecology*, **18**, 1207–1224.

Gompert Z, Fordyce JA, Forister ML, Shapiro AM, Nice CC (2006a) Homoploid hybrid speciation in an extreme habitat. *Science*, **314**, 1923–1925.

Gompert Z, Nice CC, Fordyce JA, Forister ML, Shapiro AM (2006b) Identifying units for conservation using molecular systematics: the cautionary tale of the Karner blue butterfly. *Molecular Ecology*, **15**, 1759–1768.

Gompert Z, Fordyce JA, Forister ML, Nice CC (2008a) Recent colonization and radiation of North American *Lycaeides* (*Plebejus*) inferred from mtDNA. *Molecular Phylogenetics and Evolution*, **48**, 481–490.

Gompert Z, Forister ML, Fordyce JA, Nice CC (2008b) Widespread mito-nuclear discordance with evidence for introgressive hybridization and selective sweeps in *Lycaeides*. *Molecular Ecology*, **17**, 5231–5244.

Gompert Z, Forister ML, Fordyce JA, Nice CC, Williamson R, Buerkle CA (2010) Bayesian analysis of molecular variance in pyrosequences quantifies population genetic structure across the genome of Lycaeides butterflies. *Molecular Ecology*, **19**, 2455-2473.

Grant V (1971) *Plant Speciation*. Columbia University Press, New York.

Grant PR, Grant BR (1992) Hybridization of bird species. *Science*, **256**, 193–197.

Grant PR, Grant BR, Markert JA, Keller LF, Petren K (2004) Convergent evolution of Darwin's finches caused by introgressive hybridization and selection. *Evolution*, **58**, 1588–1599.

Guppy C, Shepard J (2001) *Butterflies of British Columbia*. UBC Press, Vancouver, BC, Canada.

Harris AG, Tuttle E, Tuffle SD (1997) *Geology of National Parks*, 5th edn. Kendall/Hunt Publishing, Dubuque, IA, USA.

Harrison RG (1986) Pattern and process in a narrow hybrid zone. *Heredity*, **56**, 337–349.

Harrison RG (1990) Hybrid zones: windows on evolutionary process. *Oxford Surveys in Evolutionary Biology*, **7**, 69–128.

Harrison RG, Rand DM (1989) Mosaic hybrid zones and the nature of species boundaries. In: *Speciation and its Consequences* (eds Otte D, Endler J), pp. 110–133. Sinauer Associates, Sunderland, Massachusetts.

Hartigan P (1985) Computation of the dip statistic to test for unimodality. *Applied Statistics-Journal of the Royal Statistical Society Series C*, **34**, 320–325.

Hartigan J, Hartigan P (1985) The dip test of unimodality. *Annals of Statistics*, **13**, 70–84.

Hewitt GM (1988) Hybrid zones – natural laboratories for evolution studies. *Trends in Ecology & Evolution*, **3**, 158–166.

Hewitt GM (1989) The subdivision of species by hybrid zones. In: *Speciation and its consequences* (eds Otte D, Endler J), pp. 85-110. Sinauer Associates, Sunderland, Massachusetts.

Howard DJ (1986) A zone of overlap and hybridization between two ground crickets. *Evolution*, **40**, 34–43.

Jiggins CD, Mallet J (2000) Bimodal hybrid zones and speciation. *Trends in Ecology & Evolution*, **15**, 250–255.

Jiggins C, Naisbit R, Coe R, Mallet J (2001) Reproductive isolation caused by colour pattern mimicry. *Nature*, **411**, 302–305.

Kaufamn L, Rosseeuw PJ (1990) *Finding Groups in Data: an Introduction to Cluster Analysis*. Wiley, Hoboken, NJ, USA.

Knutson RL, Kwilosz JR, Grundel R (1999) Movement patterns and population characteristics of the Karner blue butterfly (*Lycaeides melissa samuelis*) at Indiana Dunes National Lakeshore. *Natural Areas Journal*, **19**, 109–120.

Kronforst M, Young L, Kapan D, McNeely C, O'Neill R, Gilbert L (2006) Linkage of butterfly mate preference and wing color preference cue at the genomic location of wingless. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 6575–6580.

Lancaster ML, Gemmell NJ, Negro S, Goldsworthy S, Sunnucks P (2006) Menage a trois on Macquarie Island: hybridization among three species of fur seal (*Arctocephalus* spp.) following historical population extinction. *Molecular Ecology*, **15**, 3681–3692.

Lepais O, Petit R, Guichoux E *et al.* (2009) Species relative abundance and direction of introgression in oaks. *Molecular Ecology*, **18**, 2228–2242.

Levin DA, Francisco-Ortega JK, Jansen RK (1996) Hybridization and the extinction of rare plant species. *Conservation Biology*, **10**, 10–16.

Lucas LK, Fordyce JA, Nice CC (2008) Patterns of genitalic morphology around suture zones in North American *Lycaeides* (Lepidoptera : Lycaenidae): implications for

taxonomy and historical biogeography. *Annals of the Entomological Society of America*, **101**, 172–180.

Lukhtanov VA, Kandul NP, Plotkin JB, Dantchenko AV, Haig D, Pierce NE (2005) Reinforcement of pre-zygotic isolation and karyotype evolution in *Agrodiaetus* butterflies. *Nature*, **436**, 385–389.

Lynch M, Walsh B (1998) *Genetics and Analysis of Quantitative Traits*. Sinauer Associates, Sunderland, Massachusetts.

Macholan M, Munclinger P, Sugerkova M *et al.* (2007) Genetic analysis of autosomal and X-linked markers across a mouse hybrid zone. *Evolution*, **61**, 746–771.

Maechler M, Ringach D (2009) *diptest: Hartigan's Dip Test Statistic for Unimodality-Corrected Code*. R package version 0.25-2. http://CRAN.R-project.org/package=diptest/.

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

Mallet J (2007) Hybrid speciation. *Nature*, **445**, 279–283.

Mallet J, Barton N, Lamas G, Santisteban J, Muedas M, Eeley H (1990) Estimates of selection and gene flow from measures of cline width and linkage disequilibrium in *Heliconius* hybrid zones. *Genetics*, **124**, 921–936.

Mallet J, Beltran M, Neukirchen W, Linares M (2007) Natural hybridization in Heliconiine butterflies: the species boundary as a continuum. *BMC Evolutionary Biology*, **7**, 28.

Martin NH, Bouck AC, Arnold ML (2006) Detecting adaptive trait introgression between *Iris fulva* and *I. brevicaulis* in highly selective field conditions. *Genetics*, **172**, 2481–2489.

Mavárez J, Salazar CA, Bermingham E, Salcedo C, Jiggins CD, Linares M (2006) Speciation by hybridization in *Heliconius* butterflies. *Nature*, **441**, 868–871.

Michaud R, Lehman WF, Rumbaugh MD (1988) World distribution and historical developments. In: *Alfalfa and Alfalfa Improvement* (eds Hanson AA, Barnes DK, Hill RR), Vol. 29. pp. 25–91, Madison, Wisconsin.

Moore WS (1977) An evaluation of narrow hybrid zones in vertebrates. *Quarterly Review of Biology*, **52**, 263–267.

Moore WS, Buchanan DB (1985) Stability of the Northern Flicker hybrid zone in historical times – implications for adaptive speciation theory. *Evolution*, **39**, 135–151.

Moore WS, Koenig WD (1986) Comparative reproductive success 1071 of Yellow-shafted, Red-shafted, and hybrid flickers across a hybrid zone. *AUK*, **103**, 42–51.

Morgan-Richards M, Smissen RD, Shepherd LD *et al.* (2009) A review of genetic analyses of hybridisation in New Zealand. *Journal of the Royal Society of New Zealand*, **39**, 15–34.

Nabokov V (1949) The nearctic members of *Lycaeides* Hübner (Lycaenidae, Lepidoptera). *Bulletin of the Museum of Comparative Zoology*, **101**, 479–541.

Nabokov V (1952) Butterfly collecting in Wyoming, 1952. *The Lepidopterists' News*, **6**, 41.

Nadachowska K, Babik W (2009) Divergence in the face of gene flow: the case of two newts (Amphibia: Salamandridae). *Molecular Biology and Evolution*, **26**, 829–841.

Nagata N, Kubota K, Yahiro K, Sota T (2007) Mechanical barriers to introgressive hybridization revealed by mitochondrial introgression patterns in *Ohomopterus* ground beetle assemblages. *Molecular Ecology*, **16**, 4822–4836.

Nice CC, Shapiro AM (1999) Molecular and morphological divergence in the butterfly genus *Lycaeides* (Lepidoptera: Lycaenidae) in North America: evidence of recent speciation. *Journal of Evolutionary Biology*, **12**, 936–950.

Nice CC, Anthony N, Gelembiuk G, Raterman D, ffrench Constant R (2005) The history and geography of diversification within the butterfly genus *Lycaeides* in North America. *Molecular Ecology*, **14**, 1741–1754.

Nice CC, Gompert Z, Forister ML, Fordyce JA (2009) An unseen foe in arthropod conservation efforts: the case of *Wolbachia* infections in the Karner Blue butterfly. *Biological Conservation*, **14**, 3137–3146.

Nichols RA, Hewitt GM (1986) Population-structure and the shape of a chromosomal cline between two races of *Podisma pedestris* (Orthoptera, Acrididae). *Biological Journal of the Linnean Society*, **29**, 301–316.

Niemiller ML, Fitzpatrick BM, Miller BT (2008) Recent divergence with gene flow in Tennessee cave salamanders (Plethodontidae : Gyrinophilus) inferred from gene genealogies. *Molecular Ecology*, **17**, 2258–2275.

Nolte AW, Gompert Z, Buerkle CA (2009) Variable patterns of introgression in two sculpin hybrid zones suggest that genomic isolation differs among populations. *Molecular Ecology*, **18**, 2615–2627.

Nosil P, Yukilevich R (2008) Mechanisms of reinforcement in natural and simulated polymorphic populations. *Biological Journal of the Linnean Society*, **95**, 305–319.

Ohno S, Hoshizaki S, Ishikawa Y, Tatsuki S, Akimoto S (2003) Allometry of male genitalia in a lepidopteran species, *Ostrinia latipennis* (Lepidoptera: Crambidae). *Applied Entomology and Zoology*, **38**, 313–319.

Okeke F, Karnieli A (2006) Linear mixture model approach for selecting fuzzy exponent value in fuzzy c-means algorithm. *Ecological Informatics*, **1**, 117–124.

Ortiz-Barrientos D, Grealy A, Nosil P (2009) The genetics and ecology of reinforcement implications for the evolution of prezygotic isolation in sympatry and beyond. In: *Year in Evolutionary Biology 2009*. *Annals of the New York Academy of Sciences*, **1168**, 156–182.

Otto SP, Whitton J (2000) Polyploid incidence and evolution. *Annual Review of Genetics*, **34**, 401–437.

Pal NR, Bezdek JC (1995) On cluster validity for the fuzzy c-means model. *IEEE Transactions on Fuzzy Systems*, **3**, 370–379.

Parsons TJ, Olson SL, Braun MJ (1993) Unidirectional spread of secondary sexual plumage traits across an avian hybrid zone. *Science*, **260**, 1643–1646.

Polak M, Rashed A (2010) Microscale laser surgery reveals adaptive function of male intromittent genitalia. *Proceedings of the Royal Society B-Biological Sciences*, **277**, 1371-1376.

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

R Development Core Team (2009) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Rieseberg LH (1997) Hybrid origins of plant species. *Annual Review of Ecology and Systematics*, **28**, 359–389.

Rieseberg LH (2006) Hybrid speciation in wild sunflowers. *Annals of the Missouri Botanical Garden*, **93**, 34–48.

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.

Rieseberg LH, Raymond O, Rosenthal DM *et al.* (2003) Major ecological transitions in wild sunflowers facilitated by hybridization. *Science*, **301**, 1211–1216.

Rowlingson BS, Diggle PJ (1993) SPLANCS – spatial point pattern-analysis code in S-plus. *Computers and Geosciences*, **19**, 627–655.

Ruegg K (2008) Genetic, morphological, and ecological characterization of a hybrid zone that spans a migratory divide. *Evolution*, **62**, 452–466.

Sætre GP, Moum T, Bureš S, Král M, Adamjan M, Moreno J (1997) A sexually selected character displacement in flycatchers reinforces premating isolation. *Nature*, **387**, 589–592.

Scott J (1986) *The Butterflies of North America: A Natural History and Field Guide*. Stanford University Press, Palo Alto, CA, USA.

Servedio MR, Noor MAF (2003) The role of reinforcement in speciation: theory and data. *Annual Review of Ecology, Evolution, and Systematics*, **34**, 339–364.

Shapiro AM, Porter AH (1989) The lock-and-key hypothesis – evolutionary and biosystematic interpretation of insect genitalia. *Annual Review of Entomology*, **34**, 231–245.

Stinchcombe JR, Hoekstra HE (2007) Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. *Heredity*, **100**, 158–170.

Struyf A, Hubert M, Rousseeuw PJ (1997) Integrating robust clustering techniques in S-PLUS. *Computational Statistics and Data Analysis*, **26**, 17–37.

Szymura JM, Barton NH (1986) Genetic analysis of a hybrid zone between the fire-bellied toads, *Bombina bombina* and *B. 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.

Takami Y (2003) Experimental analysis of the effect of genital morphology on insemination success in the ground beetle *Carabus insulicola* (Coleoptera Carabidae). *Ethology, Ecology and Evolution*, **15**, 51–61.

Teeter KC, Thibodeau LM, Gompert Z, Buerkle CA, Nachman MW, Tucker PK (2010) The variable genomic architecture of isolation between hybridizing species of house mouse. *Evolution*, **64**, 472–485.

Turner T, Hahn M, Nuzhdin S (2005) Genomic islands of speciation in *Anopheles gambiae*. *PLoS Biology*, **3**, 1572–1578.

U.S. Fish and Wildlife Service (2003) *Karner Blue Butterfly (Lycaeides melissa samuelis)*. Recovery Plan. Tech. rep., Region 3, U.S. Fish and Wildlife Service, Fort Snelling, Minnesota.

Vaha J, Primmer C (2006) Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. *Molecular Ecology*, **15**, 63–72.

Vines TH, Kohler SC, Thiel A *et al.* (2003) The maintenance of reproductive isolation in a mosaic hybrid zone between the fire-bellied toads *Bombina bombina* and *B. variegata*. *Evolution*, **57**, 1876–1888.

Whitney KD, Randell RA, Rieseberg LH (2006) Adaptive introgression 1176 of herbivore resistance traits in the weedy sunflower *Helianthus annuus*. *The American Naturalist*, **167**, 794–807.

Wu CI (2001) The genic view of the process of speciation. *Journal of Evolutionary Biology*, **14**, 851–865.

Yanchukov A, Hofman S, Szymura JM, Mezhzherin SV (2006) Hybridization of *Bombina bombina* and *B. variegata* (Anura, Discoglossidae) at a sharp ecotone in western Ukraine: Comparisons across transects and over time. *Evolution*, **60**, 583–600.

Zaykin DV, Pudovkin A, Weir BS (2008) Correlation-based inference for linkage disequilibrium with multiple alleles. *Genetics*, **180**, 533–545.

Zach Gompert is a PhD student in the ecology program at the University of Wyoming. His research interests include speciation genetics, hybridization, and statistical methods in population genetics. Lauren Lucas has research interest in evolutionary biology, morphometrics, ecological genetics, and science education. Matt Forister, assistant professor at the University of Nevada, Reno, has research interests that include diet specialization and the evolutionary ecology of plant-insect interactions. James Fordyce is an associate professor at the University of Tennessee with research interests in ecological factors that promote population differentiation and maintain variation. Chris Nice is an associate professor at Texas State University with interests in evolutionary ecology and genetics.

Supporting Information

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

Table S1 Classification from linear discriminant function analysis. Rows correspond to taxonomic designation based on genetic data and locality, whereas columns correspond to predicted classification from the linear discriminant function analysis

Table S2 Coefficients of linear discriminants and standardized coefficients of linear discriminants for wing pattern

Fig. S1 Plots show the marginal likelihood of the model (A) and ΔK (B) for different assumed numbers of clusters. Mean values plus and minus one standard error are depicted for the former. Calculations of ΔK follow Evanno *et al.* (2005).

Fig. S2 Boxplots depicting variation in wing pattern PC1 and PC2 scores for *Lycaeides* by species and sex. Wing pattern differences between species are greater for males than females.

Fig. S3 Distribution of admixture proportions (A), fuzzy cluster membership based on male genitalia (B), and fuzzy cluster membership based on male wing pattern (C) for individuals from the core of the Rocky Mountain hybrid zone. Admixture proportions were estimated using *structure* with $k = 2$. The distributions of admixture proportion and fuzzy cluster memberships were not different from unimodal. See the main text for more details.

Fig. S4 Scatterplot of geographical location by LD. Points show LD estimates for individual pairs of markers.

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