GENETIC LEAKAGE AFTER ADAPTIVE AND NONADAPTIVE DIVERGENCE IN THE ENSATINA ESCHSCHOLTZII RING SPECIES

Ricardo J. Pereira¹,²,³,⁴ and David B. Wake¹

¹Museum of Vertebrate Zoology and Department of Integrative Biology, 3101 Valley Life Sciences Building, University of California, Berkeley, California 94720
²E-mail: ricardo@berkeley.edu
³CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Universidade do Porto, Campus Agrário de Vairão, 4485-661 Vairão, Portugal
⁴Departamento de Zoologia e Antropologia, Faculdade de Ciências da Universidade do Porto, 4099-002 Porto, Portugal

Received October 24, 2008
Accepted April 9, 2009

The salamander Ensatina eschscholtzii is an example of a ring species in which extant intermediate stages of terminal forms have a nearly continuous range, offering replicated interactions at several stages of divergence. We employ a greatly expanded allozyme database and individual-based analyses to separate the effects of divergence time and gene flow to evaluate how gradual divergence of populations around the ring contributes to the development of reproductive isolation. Despite the high degree of genetic (D ≤ 0.39) and ecomorphological divergence observed in secondary contacts around the ring, reproductive isolation or rare hybridization is observed only at the terminus of the ring. Instead, in the secondary contacts sampled around the ring, hybrids are common and reproductively successful, enabling genetic leakage between parental genomes and the potential for genetic merger. Nevertheless, genetic admixture is geographically broad (<100 km) only in contacts between ecomorphologically similar populations (within subspecies). When divergence is accompanied by alternative patterns of adaptive divergence (between subspecies), zones of intergradation are narrower and affect populations only locally (>8 km). Diversification and consequent genetic interactions in Ensatina reveal a continuum between populations, ecological races, and species, where polytypic traits and high genetic differentiation are maintained without reproductive isolation.

KEY WORDS: Adaptive divergence, gene flow, hybridization, polytypic species, secondary contact, speciation.
finches, Grant and Grant 2002) and potentially to the complete merger of the once divergent taxa (e.g., three-spined stickleback, Taylor et al. 2006). In contrast, reproductive isolation via nonadaptive divergence is usually a slower but more irreversible process due to the evolution of intrinsic isolating mechanisms (e.g., Dobzhansky-Muller incompatibilities, Coyne and Orr 1989).

Microevolutionary processes often fail to result in full reproductive isolation and genetic exchange follows at some level, as a consequence of interaction between parental populations. Despite the recognized continuity of the divergence process (see de Queiroz 1998), the genetic exchange upon secondary contact has been interpreted traditionally to fall into two main categories, with distinct evolutionary outcomes. When admixture is episodic and does not have major consequences for the genetic “integrity” of the parental groups (Mayr 1963), it is classified as hybridization and species-level divergence between parental forms is recognized. In contrast, when genetic exchange is unrestricted, leading to introgression and further merging of once differentiated gene pools, it is identified as secondary intergradation between populations within the same species. Ring species are particularly interesting to study the development of reproductive isolation, because gradual population level divergence of the persistent intermediate forms results in species-level divergence between the terminal overlapping forms. Therefore, both kinds of genetic interactions are expected to occur, with free gene flow connecting contiguous populations in the terminal contact across the ring (Mayr 1963; Dobzhansky 1958).

The *Ensatina eschscholtzii* complex is a textbook example of a ring species that illustrates the role of geography in species formation (Dobzhansky 1958; Futuyma 1998; Dawkins 2004). The ancestor is hypothesized to have expanded from northern California around the inhospitable habitat that currently constitutes the Central Valley, gradually diverging along each arm of expansion (Stebbins 1949; Wake and Yanev 1986). When populations met at the terminus, in southern California, species-level divergence has occurred and they overlap with full reproductive isolation or rare hybridization occurring in the terminal contact across the ring (Mayr 1963; Dobzhansky 1958).

The *Ensatina eschscholtzii* complex is a textbook example of a ring species that illustrates the role of geography in species formation (Dobzhansky 1958; Futuyma 1998; Dawkins 2004). The ancestor is hypothesized to have expanded from northern California around the inhospitable habitat that currently constitutes the Central Valley, gradually diverging along each arm of expansion (Stebbins 1949; Wake and Yanev 1986). When populations met at the terminus, in southern California, species-level divergence has occurred and they overlap with full reproductive isolation or rare hybridization occurring in the terminal contact across the ring (Mayr 1963; Dobzhansky 1958).

In this salamander ring species, population-level divergence occurred most probably during periods of geographic isolation, resulting in deep phylogenetic breaks (Moritz et al. 1992; Kuchta et al. 2009a). These periods were more or less prolonged, resulting in varying levels of genetic differentiation between the presently contiguous populations around the ring (Jackman and Wake 1994; Wake 1997). In some instances, this divergence process is presumed to have occurred in similar habitats, resulting in populations that are only genetically distinct (nonadaptive divergence), while in others in ecologically distinct habitats, promoting alternative divergence in ecologically relevant traits (adaptive divergence). The color pattern in *Ensatina* is believed to represent one such trait that may have evolved via alternative predator avoidance strategies (Stebbins 1949; Brown 1974; Wake 2006). As a result, a remarkably high phenotypic variability of the salamander is strongly regionalized in areas with ecologically similar habitats. Geographically concordant regionalization also occurs in plants (Hickman 1993) and herpetological communities (Feldman and Spicer 2006; Rissler et al. 2006), suggesting that the physiogeographic provinces of California share features that promote adaptive divergence across taxa. Experimental studies in *Ensatina* (Kuchta et al. 2008) support the adaptive value of color pattern by showing that a predator, following presentation with the presumed highly toxic model (*Taricha tarosa*), was more hesitant to contact the presumed mimic (*E. e. xanthoptica*) than a control subspecies lacking the postulated aposematic colors (*E. e. oregonensis*). The regionalization of *Ensatina* ecomorphotypes is reflected taxonomically by the recognition of seven subspecies (Fig. 1): the putative ancestral form *picta* in northern California; the cryptic forms *platensis*, *croceater*, and *klauberi* in the interior mountain ranges, and the variously aposematic and mimetic forms *oregonensis*, *xanthoptica*, and *eschscholtzii* along the Pacific coast from central British Columbia to northern Baja California (Stebbins 1949; Wake 2006). More recently, during the cooler and/or wetter periods of the Holocene, a presumed temporary corridor across the Central Valley allowed the coastal *xanthoptica* to colonize the foothills of the Sierra Nevada, providing a second closure of the ring in the mid-Sierra Nevada (see Alexandrino et al. 2005).

The use of molecular markers (Wake and Yanev 1986, Moritz et al. 1992; Jackman and Wake 1994; Wake 1997), besides confirming the high genetic differentiation expected between different ecomorphotypes, showed that populations with the same color pattern (within subspecies) can be genetically more differentiated than are populations with divergent color patterns (between subspecies). Recent wide range studies using mitochondrial DNA (Kuchta et al. 2009a), revealed high subdivision with fractal spatial structure, beyond that earlier reported, suggesting that long-term isolation might be more frequent than previously considered and might account for most of the genetic differentiation observed. Empirical work using model organisms (e.g., *Drosophila*, Coyne and Orr 1989) predicts that the extent of reproductive isolation should increase with divergence time and ecomorphological divergence (Coyne and Orr 1989, 2004). Therefore, we may expect that in *Ensatina* reproductive isolation may have evolved also between populations around the ring as a function of divergence time (genetic differentiation), particularly when coupled with ecomorphological divergence (differentiation in color pattern). If reproductive isolation has arisen between populations around the ring as a product of extended geographic isolation, in regions of secondary contact we can expect to observe only pure individuals or accidental hybrids that will not enable gene flow between contiguous groups (Mayr 1963). However, if high
genetic differentiation does not result in reproductive barriers, we can expect genetically admixed individuals to be abundant and potentially lead to the genetic merger of adjacent populations.

In fact, fine-scale studies that approach the cruising distance of these salamanders (Kuchta et al. 2009b) manifest sharp parapatric boundaries in mitochondrial DNA, even between ecomorphologically similar populations (e.g., oregonensis lineages in San Francisco Bay Area), suggesting that gene flow might be compromised in several contacts around the ring.

Several authors (see Avise 2000) strongly advocate against the use of single gene trees, particularly from uniparentally inherited markers, to infer processes that occur at the species or population level, such as generation and maintenance of genetic boundaries. Theoretical (Irwin 2002; Kuo and Avise 2005) and empirical (Irwin et al. 2001) studies suggest that genetic discontinuities may readily arise in single gene trees without historical isolation, due to stochastic processes that strongly affect nonrecombining stretches of DNA, especially if mean dispersal distances of individuals and/or population sizes are low. In contrast, spatial concordance between independent markers is expected in a historical isolation scenario (Kuo and Avise 2005). The extremely low dispersal of Ensatina (mean final distance for individual dispersal of 22 m, Staub et al. 1995) makes it a valuable model to test whether deep and parapatric mitochondrial breaks represent lineage boundaries or if they are due to stochastic processes.

In this study, we evaluate the evolutionary processes that shaped the diversification of E. eschscholtzii and how they contribute to reproductive and/or genetic isolation of the differentiated forms. We greatly expand previous works to consider 1132 individuals collected throughout the entire ring and genotyped for 22–27 nuclear markers (allozymes). We use individual-based (multilocus genotypes) methods to separate the effects of divergence time (increased genetic differentiation) from those associated with gene flow (low genetic differentiation), ignoring a priori assumptions on the divergence level of the units studied here (i.e., populations, species). We first evaluate how often historical isolation affected populations through the ring distribution by identifying genetic breaks in the nuclear DNA, contrasting this analysis to those previously recovered using mtDNA. We then examine how different periods of geographic isolation may have contributed to the development of reproductive isolation by analyzing the frequency of hybrids at zones of secondary contact between populations that diverged only genetically (within subspecies) and populations that also diverged in the presumed adaptive color patterns (between subspecies). Finally, we ask if populations around and across the ring are at a reversible stage of divergence by determining if hybrids are occasional and parental.

Figure 1. Distribution of the ring species Ensatina eschscholtzii in California. The range of each subspecies (morphologically divergent units) is represented by different colors, and sampling localities are signed with white circles. Inset refers to a detailed view of the sampling scheme applied to the contact between oregonensis and xanthoptica (included in study area C, Fig. 2).
populations are evolving independently, or if reproduction can occur over several generations of backcrossing, thus enabling gene flow and potential merger.

**Methods**

**SPECIMEN AND GENETIC SAMPLING**

In this work, we sampled the entire species range around the Central Valley of California, the main geographic barrier that resulted in the ring-like distribution of the complex. Our sampling includes 1130 individuals from 128 localities (Fig. 1, Table S1) representing all morphologically distinct populations of *Ensatina* subspecies, as well as all strongly supported mitochondrial clades nested within these (Kuchta et al. 2009a). Our sampling is particularly dense around the areas of morphological and genetic transitions, which enables us to detect potential genetic transitions in nDNA and regions of secondary contact where reproductive isolation can be tested.

Starch-gel electrophoresis was used to examine protein variation in the samples for 27 allozyme loci, following the methods of Wake and Yanev (1986). Due to the high genetic variability of this system, by adjusting the electrophoretic conditions we captured the maximum number of alleles per region. Hence, the samples were partitioned into five spatially overlapping study areas (Fig. 2A–E) that were analyzed independently. We integrate allozymic data for three previously unpublished studies (Fig. 2B–D), and two other published studies (Jackman and Wake 1994 in Fig. 2A, Wake et al. 1989 in Fig. 2E), using novel individual-based (multilocus genotypes) analyses that are appropriate at the interface between population- and species-level divergence.

Study area A (Fig. 2A) covers the population from the north coast of California and along the Sierra Nevada and includes the subspecies *picta, oregonensis*, and *platensis* (261 individuals, 33 localities, 23 variable loci; data from Jackman and Wake 1994). Study area B (Fig. 2B) concerns coastal populations north of San Francisco Bay and includes *oregonensis* and the transition to *xanthoptica*, where a “population based” sampling design was applied, with many individuals collected from few localities (274 individuals, 20 localities, 25 variable loci). Study area C (Fig. 2C), around San Francisco Bay, focuses on the same transition between phenotypes, but the sampling design of this contact zone is “individual based,” prioritizing the number of geographic points sampled over the number of individuals per locality (247 individuals, 39 localities, 22 variable loci). Study area D (Fig. 2D) is concentrated on the San Francisco Peninsula, a geologically dynamic region marked by several tectonic faults, which delimit the distribution of well-supported mitochondrial lineages (Kuchta et al. 2009b). In this study, we include samples from all the tectonic plates and from the three distinct phenotypes: *xanthoptica, oregonensis*, and *eschscholtzii* (141 individuals, 19 localities, and 27 variable loci). And finally, we focus on the closure of the ring in study area E, in southern California, which includes populations morphologically assigned to *platensis, croceater*, and *klauberi*, which meets the coastal *eschscholtzii* (207 individuals, 17 localities, 23 variable loci; combined dataset from Jackman and Wake 1994 and Wake et al. 1989).

**POPULATION STRUCTURE AND HISTORICAL ISOLATION**

To assure that the genetic markers used meet all the assumptions of the methods applied (neutral markers without null alleles), we tested for deviations from Hardy–Weinberg (HW) equilibrium across all sampled localities. We assessed their ability to reflect potential population structure while testing the partition of genetic variability with an analysis of molecular variance, using ARLEQUIN (Schneider et al. 2000).

Using the individual, and its multilocus genotype, as the operational unit of our analysis, we were able to ignore subspecific taxonomy, mitochondrial clade, or geographic location of the samples, and thus avoid both a priori definitions of “population” and issues arising from uneven sampling (we used from 1 to 51 individuals per locality). We diagnosed distinct population groupings of randomly mating individuals with differentiated allele frequencies that could be distinguished with high statistical support (posterior probability, *pp* ≥ 0.95). To simultaneously assign individuals to populations and infer their allele frequencies, we applied the Bayesian approach implemented in the software STRUCTURE (Pritchard et al. 2000). Populations are identified as groups of randomly mating individuals, with minimum HW and linkage disequilibrium. We chose to run the “independent allele frequency” model because we expect allele frequencies in different ancestral populations to diverge by drift during nonadaptive divergence events, together with the “admixture” model to accommodate potential gene flow. We ran five pseudo-replicates with 10⁶ Markov chain Monte Carlo (MCMC) iterations, after a burn-in of 10⁵ steps. We estimated the smallest number of parental populations (*K*) that captures the most population structure in the data for each study by allowing *K* to increase up to 10 hypothetical populations while monitoring their posterior probabilities (ln Pr(X|*K*)) and variance (Pritchard et al. 2000).

If geographic isolation was a frequent process in the diversification of *E. eschscholtzii*, we expect to identify several genetically distinct populations that share the same color pattern (within subspecies). If diversification was mainly adaptive and the mitochondrial breaks result from stochastic effects, we do not predict being able to observe nuclear genetic discontinuities that are spatially concordant with those reported for mtDNA (Kuchta et al. 2009a).
Figure 2. Population structure and gene flow in the ring species *Ensatina eschscholtzii*. The results refer to five separate studies from spatially overlapping areas of the ring in California: A, Sierra Nevada; B, North of San Francisco Bay (SFB); C, SFB; D, South of SFB; and E, Southern California. For each study, the genetic discontinuities in nuclear DNA were identified from individual multilocus genotypes using the program STRUCTURE (Pritchard et al. 2000). The assignment of individuals is shown on the color graphic at the side of each map, where ancestral populations are represented by different colors. Populations that diverged both in neutral genetic markers.
DIVERSITY PROCESS AND REPRODUCTIVE ISOLATION

If the assignment method described above allows us to unambiguously distinguish pure individuals, one may expect that samples from geographic isolates or which are distant to secondary contacts will be assigned to a single cluster \((Q \geq 0.95)\). In contrast, in zones of secondary contacts one can expect to find individuals in sympatry with ancestry in different parental populations or their descendents \((Q < 0.95)\). Due to the almost contiguous distribution of the *Ensatina* complex, we had the opportunity to test for reproductive isolation between every pair of contiguous populations that meets around the ring (with or without divergence in color pattern), and also across the ring (between the most divergent forms).

In the first step, we calculated genetic differentiation between populations as defined by STRUCTURE as a surrogate for neutral divergence between ancestral populations. To avoid any bias caused by gene flow after secondary contact and benefit from large sample sizes, we combined all the “pure” individuals assigned to the same population \((Q \geq 0.95)\) and calculated Nei’s \(D\) (1972) using GENETIX (Belkhir et al. 2001). As a result, the magnitude of genetic distances calculated here is higher than those calculated by locality-pooling methods previously used (as in Wake and Yanev 1986; Jackman and Wake 1994; Highton 1998; Wake and Schneider 1998).

In the second step, we assessed the frequency of hybrid individuals in each secondary contact, by assigning individuals to categories of parentals or hybrids, using the program NEW HYBRIDS (Anderson and Thompson 2002) as described below. For the purpose of this work, all the individuals that are genetic mosaics of more than one parental genome are called “hybrids,” referring to their genetic structure and not to the divergence level of the parental populations. The frequency of hybrids in all contacts was plotted against the genetic distance between parental populations.

If reproductive isolation increases with divergence time (Mayr 1963), we would expect a fast decay of the frequency of hybrids with increasing degrees of genetic differentiation. This decay would be more extreme if genetic differentiation is accompanied by ecomorphological divergence (Coyne and Orr 2004).

GENETIC ISOLATION

Despite the frequency of hybridization events, the presence of postzygotic isolating mechanisms (e.g., natural or sexual selection) acting on early-generational hybrids may prevent effective gene flow between parental populations. We evaluated the reproductive success of the initial hybrids in each contact by distinguishing recent from older hybrids (originating from two or more generations of backcrosses). We used the assignment test implemented in the program NEW HYBRIDS (Anderson and Thompson 2002) to classify individuals into categories of parentals, F1s, F2s, first generation backcrosses, or later-generation hybrids.

Because of the broad geographic extent of genetic admixture observed in some cases (see Results, Fig. 2), we suspected that F1 hybrids might have high reproductive success, backcrossing over several generations. To test our power of distinguishing old generations of backcrosses from pure parentals and recent generations of hybrids, we simulated multilocus genotypes for increasing generations of hybrids using HYBRIDLAB (Nielsen et al. 2006). We simulated genotypes for several generations of hybrids, starting from two equally abundant parental populations \((N = 200)\) genotyped for 25 codominant diagnostic loci. In the first scenario, we simulated genotypes for all possible classes of individuals after two generations since secondary contact: 28 genotypes from each parental, 12 F1s, 12 F2, and 12 first generations of backcrosses with each parental population. Second, we allowed the first-generation backcrosses to reproduce with all the other categories; simulating 18 genotypes that are not expected to fall under the classes readily determined using genetic methods (F1, F2, or first generation backcrosses). In a third scenario, we simulated genotypes produced in a “hybrid swarm” for three following generations of random mating (total of six generations of admixture since secondary contact). The individuals simulated for the

---

*and morphology (adaptive divergence) are taxonomically recognized at the subspecies level and are here represented in different main colors, according to the legend. Within subspecies, populations that diverged only in genetic markers (nonadaptive divergence) are in different shades of the same main color. Each line corresponds to one individual with colors representing the percentage of its genome inherited from each parental population; hybrid individuals will have proportions of different colors. Assignments for the detailed studies of the four contacts at the terminus of the ring (between *eschscholtzii* and *klauberi*, study area E; Wake et al. 1989) are laid out in the lower graphs; from North to South: S. Sawmill Canyon, J. San Jacinto, P. Palomar, and C. Cuyamaca Mts. Bidirectional arrows (→) indicate genetic breaks that lack reproductive isolation (presence of hybrids), (♯) indicate breaks without evidence of secondary contact, and crosses (×) sign areas of secondary contact with full reproductive isolation. The distribution of the major phylogeographic breaks in mitochondrial DNA is represented for each study by colored bars at the side of each graph, following Kuchta et al. (2009a). Sympathy of mitochondrial lineages is here represented by parallel bars. The maps represent the distribution of the differentiated populations, as reflected by the graph and considering the geographic coordinates of sampling. The presence of hybrid individuals or sympathy of differentiated individuals is translated into overlapping ranges of populations. Genetic differentiation (Nei’s \(D\)) between contiguous ancestral populations is represented by the lines, continuous when there is evidence of connecting gene flow and dashed when there is no such evidence.*
three scenarios were analyzed with NEW HYBRIDS (Anderson and Thompson 2002) using the same methods and thresholds defined for the collected data. We included 10 individuals from each parental group as references for the assignment test.

For our collected data, we conducted independent analyses for all pairs of contiguous populations within each study, because the algorithm assumes a two-population model. The assignment was based solely on the multilocus genotypes without any prior information on the ancestry or admixture of each individual. We pseudo-replicated the MCMC from different starting points and ran the analysis long after reaching stability to assure convergence to the same result. We summarized for each secondary contact the proportion of each category of hybrids, using a threshold of 0.9 posterior probability. We present detailed results for all contacts between morphologically similar groups (within platensis at Wagner Ridge, within oregonensis at Longvale, and within xanithoptica at Mt. Hamilton), morphologically divergent groups (between croceater and platensis at Breckenridge, picta and oregonensis at Salyer, oregonensis and xanithoptica at Pressley Road), and also results for all four contacts after the more extreme genetic and morphological divergence observed in the closure of the ring (klauberi and eschscholtzii from North to South: Sawmill Canyon, San Jacinto, Palomar and Cuyamaca Mts.).

If gene flow is enabled between parental populations in secondary contacts, we expect to find older hybrids that result from several generations of backcrosses. On the other hand, if initial hybrids can occur but cannot reproduce successfully, we only expect to find recent generations of hybrids.

Results

POPULATION STRUCTURE AND HISTORICAL ISOLATION

We do not detect significant deviation from HW equilibrium in any marker across all sampling localities, except for localities in southern California where the two most divergent subspecies (eschscholtzii and klauberi, Fig. 2E) are sympatric, suggesting Wahlund effect (Hartl and Clark 2007). The analysis of molecular variance shows that the allozyme markers are highly informative for population differentiation, with most of the genetic variability attributed to differences among sampling localities (52%, 40%, 57%, 60%, and 53%, for studies A–E, respectively; \( P \gg 0.01 \)).

Cluster analysis of individuals reveals strong population structure. The posterior probability of the data (\( \ln \Pr(X|K) \)) increases exponentially with \( K \), as we identify nuclear genetic discontinuities at increasingly finer scales, from geographic areas characterized by transition in morphology (reflected by subspecies), to areas of transition in mitochondrial lineage. Pure individuals (\( Q \geq 0.95 \)) are mainly found in isolated localities or far from contact zones, whereas admixed individuals are found close to secondary contacts, reflecting the high power to accurately distinguish populations. Hybrid individuals (\( Q < 0.95 \)) have appropriately mixed ancestry in adjacent ancestral populations. Their geographic setting enables estimation of the extent of genetic admixture. Figure 2 summarizes the results for the five studies around the ring species complex, for a conservative number of ancestral populations, i.e., before \( \ln \Pr(X|K) \) reaches a plateau. Subsequent increases in \( K \) uncover further substructure at finer scales, but significant increase of variance of the estimated parameters would lower our confidence in distinguishing pure individuals.

In all studies, we not only confirm the major genetic breaks suggested by mtDNA, but in addition always identify finer substructure in the nuclear genome, which typically agrees with the geographic limits of well-supported mitochondrial lineages within major clades (see Kuchta et al., 2009a). Overlapping areas of different studies generally show concordant patterns. The occasional mismatches refer to areas in which we uncover a finer scale of population structure with higher confidence (i.e., light-green oregonensis in Fig. 2B is further substructured in light and dark groups in Fig. 2C), probably as a result of different sampling strategy and adjustments of the electrophoretic conditions.

In study area A (Fig. 2A), the individuals collected along the Sierra Nevada show high spatial structure, representing at least seven genetically distinct populations. Groups with divergent color pattern (different subspecies) are always genetically differentiated (in Fig. 2A: platensis in blue, oregonensis in green, and picta in maroon). But we often identified genetic differentiation between groups with similar color pattern, revealing fractal genetic differentiation within subspecies (in Fig. 2A: at least four genetically distinct populations within platensis in shades of blue, and three within oregonensis in shades of green). Genetic discontinuities detected by allozymes are geographically concordant with the five major mitochondrial clades, but they reveal finer substructure within platensis and oregonensis, which is reflected in well-supported branches in mitochondrial DNA phylogeny (see Kuchta et al. 2009a). Genetic admixture between morphologically divergent populations is generally broad (around 40 km between picta and oregonensis), and is more spread out between morphologically similar groups of platensis (spanning from about 20 to 100 km). An exception is the break between oregonensis and platensis (Lassen Gap) where individuals from each side of the genetic discontinuity are assigned to a separate cluster (green and blue, Fig. 2). Despite sampling through this geographical break in the otherwise continuous range (minimum distance between parental, \( md = 15 \) km), we are unable to detect sympathy between these two populations. However, this is a region with poor habitat and low population density, and sampling cannot exclude current geographic isolation (see Jackman and Wake 1994).

In study area B (Fig. 2B) populations are also structured in a strong linear pattern along coastal northern California, with at
least five differentiated clusters. Within *oregonensis*, we uncover finer population structure than suggested by mtDNA (shades of green in Fig. 2B), with four differentiated groups (two within each mitochondrial clade) showing genetic admixture in secondary contacts. In the transition between the morphologically divergent *oregonensis* and *xanthoptica* (green to orange in Fig. 2B), neither sympatry of pure parents nor hybrid individuals were observed (reducing distance between groups to 18 km).

In study C (Fig. 2C), by collecting individuals spread out through the same contact zone (*md* = 8.8 km), we detect several hybrid individuals (*N* = 6) along a restricted area of approximately 3 km (Pressley Road in Fig. 2C). Individuals collected 2 km East and 7 km North are assigned as pure *xanthoptica* and *oregonensis*, respectively (*Q* ≥ 0.95), showing that genetic admixture spans less than 9 km. Here, in addition to the genetic discontinuities concordant with the mitochondrial breaks, we find that *oregonensis* is subdivided into several genetically differentiated populations, which admix in intervening localities (Fig. 2C). Our sampling detects no hybrids or sympathy (dashed lines) between *xanthoptica* and a group of coastal *oregonensis* (*md* = 20 km) or between *xanthoptica* and *eschscholtzii* (*md* = 16 km).

In study D, we identify at least five clusters, two of which are within *xanthoptica* (orange and red in Fig. 2D), another two within *oregonensis* (shades of green in Fig. 2D), and another cluster encompassing all *eschscholtzii* individuals collected from Pajaro Gap to the San Jacinto Mountains in southern California. All groups are spatially concordant with tectonic faults (San Andreas and San Gregorio faults) and well-supported mitochondrial lineages (Kuchta et al. 2009b). All secondary contacts show hybrid individuals, but sympathy is not detected between some populations that are either morphologically similar (e.g., within *oregonensis*, in green; *md* = 22 km) or distinct (e.g., *xanthoptica* and *eschscholtzii* in red and brown; *md* = 32 km).

In southern California (study E, Fig. 2E), we can distinguish all regions with different phenotypes and distinct mitochondrial clades (e.g., *platensis* and *croceater*), and furthermore, at least two populational units within *klauberi* that are isolated on different mountain masses (shades of purple in Fig. 2E). We find individuals with shared ancestry in contiguous populations among the inland groups. We do not find evidence of recent hybridization between *croceater* and the northernmost population of *klauberi* (*md* = 160 km), which are currently separated by an area known to be largely inhospitable or unoccupied by salamanders (Bob’s Gap, Jackman and Wake 1994). We identify four contact zones between the terminal tips across the ring, involving *klauberi* and *eschscholtzii* (in Sawmill Canyon, San Jacinto, Palomar and Cuyamaca, Fig. 2E). We observed sympathy of pure individuals from both parental populations, in agreement with the overlap of the mitochondrial lineages observed in the three areas.

**DIVERGENCE PROCESS AND REPRODUCTIVE ISOLATION**

Throughout the ring-like distribution of *Ensatina*, higher levels of genetic differentiation were observed between populations that also diverged morphologically (i.e., *klauberi* vs. *croceater*: *D* = 0.36, *platensis* vs. *oregonensis*: *D* = 0.35, *xanthoptica* vs. *eschscholtzii*: *D* = 0.24, *picta* vs. *oregonensis*: *0.39 ≤ D ≤ 0.29*, *oregonensis* vs. *xanthoptica*: *0.31 ≤ D ≤ 0.26*). The first three comparisons occur between presently disjunct populations of *Ensatina* (Bob’s, Lassen, and Pajaro gaps, Fig. 2), whereas the last two establish secondary contact. An exception to this trend is the subspecies *croceater*, which is morphologically divergent from its contiguous population of *platensis*, but presents one of the lowest levels of genetic differentiation observed throughout the ring (*D* = 0.11).

Genetic differentiation of the same degree was also observed between morphologically similar groups, as identified by the clustering analysis (e.g., within *platensis*: *D* = 0.25, within *oregonensis*: *D* = 0.20 and *D* = 0.25). However, other groups within subspecies frequently exhibit lower degrees of genetic differentiation (e.g., within *platensis*: *D* = 0.06, within *oregonensis* *D* = 0.07, and within *xanthoptica D* = 0.12).

Without respect to the level of genetic differentiation found among populations with or without divergent color patterns, we always observe that hybrids are abundant in regions of secondary contact (Fig. 3). In contacts between morphologically similar groups the frequency of hybrids (*f*) does not rapidly decay with increasing genetic divergence (0.75 ≤ *f* ≤ 1). Even where genetic differentiation is highest (*D* = 0.25 within *platensis*), all of the individuals collected in the secondary contact were hybrids. Furthermore, when genetic differentiation is accompanied by morphological divergence, we observed that hybrids dominate the contact zones (0.75 ≤ *f* ≤ 1), despite increased levels of genetic differentiation (e.g., between *oregonensis* vs. *xanthoptica*: *D* = 0.31, *f* = 1.0). It is only at the terminus of the ring, where genetic and morphological divergence is highest (0.56 ≤ *D* ≤ 0.60), that we observe low rates of hybridization in the three northernmost contacts (0.07 ≤ *f* ≤ 0.20), or full reproductive isolation in the southernmost contact.

**GENETIC ISOLATION**

Our power-analysis simulations show that we can accurately distinguish pure parents from hybrid individuals, but the confidence of assignment within hybrid classes depends on the number of generations after secondary contact (Fig. 4). If hybridization extends only for two generations, we can accurately distinguish all possible categories of hybrids: F1, F2, and first-generation backcross with either parental. If hybrids reproduce over more generations, we can still distinguish F1 and first-generation backcross hybrids, but older generations of hybrids are frequently assigned
to the F2 category with high posterior probability ($pp \geq 0.90$). This effect is extreme for recent hybrid swarms, where almost all individuals in the hybrid zone are assigned to the F2 category. This simulation shows that even with 25 diagnostic loci, in a secondary contact where hybrids can reproduce for more than one generation, we cannot distinguish between the expected multilocus genotypes for an F2 from older generations of backcrosses. In our analysis, all the individuals that were not assigned to parental, F1, or first-generation backcross classes ($pp \geq 0.90$) are considered older hybrids.

Figure 5 reports detailed assignments for secondary contacts around the ring between populations that are only genetically divergent (Fig. 5A), that are both genetically and morphologically divergent (Fig. 5B), and also for contacts across the ring (Fig. 5C). Contacts around the ring show very similar profiles, either when populations are morphologically divergent or not (Fig. 5A,B). Pure individuals are rare or absent ($f \leq 0.43$) and usually represent only one parental population. Moreover, hybrids always result from several generations of backcrossing (cannot be assigned to F1 or first-generation backcross categories by NEWHYBRIDS).

Figure 4. Power analysis of assignment to hybrid categories in three simulated hybrid zones for a varying number of generations (G) after secondary contact. The expected results (exp) refer to the chosen percentage of individuals simulated for each category and the observed results (obs) refer to the assignment of the simulated multilocus genotypes by the program NEW HYBRIDS (Anderson and Thompson 2002).
The high frequency of old generations of hybrids strongly suggests that hybridization is not only common but unrestricted, and with enough time can result in complete genetic merger, as observed in four of these contacts (Wagner Ridge, Longvale, Orr Springs, and Pressley Rd.). In contrast, with higher levels of genetic and morphologic divergence, the four contacts across the ring sampled at the ring terminus present a different dynamic: pure individuals largely dominate the area (≥80%), and most of the hybrids are F1s or first-generation backcrosses (Fig. 5C).

**Discussion**

**POPULATION STRUCTURE AND HISTORICAL ISOLATION**

Single locus phylogeographic breaks, often deep, can arise in continuously distributed species if mean dispersal distances of individuals and/or population sizes are low, as is often found in genealogies based on nonrecombining genetic units such as mitochondrial and chloroplast DNA (Irwin 2002). Such breaks will be idiosyncratic, i.e., each locus will break in a different place. In contrast, breaks caused by longstanding historical barriers to gene flow (Kuo and Avise 2005) or in regions of low environmental suitability (Endler 1977; Barton and Hewitt 1989) will be spatially concordant across independent neutral loci. In a well-studied case of ring speciation, the greenish warbler, spatial discordance between mitochondrial lineages and other traits (i.e., song, morphology and nuclear genes) suggested that divergence occurred without geographic barriers to gene flow (Irwin et al. 2001; Irwin 2002) but as a stochastic product due to low dispersal and population size.

The analysis of 22–27 variable allozymes loci of *Ensatinia* revealed fractal geographic structure of neutral genetic diversity (Fig. 2), strongly suggesting that geographical isolation was frequent throughout the entire ring distribution. Transitions in color patterns (subspecies borders) were always concordant with nuclear genetic discontinuities, often with high levels of genetic differentiation. This supports the hypothesis that divergence in the putative adaptive color patterns was probably facilitated by periods of historical isolation (Wake 1997). Although purely adaptive divergence (without interruption of gene flow) may also result in high genetic differentiation given enough time (Endler 1973), contiguous populations would be expected to be the most phylogenetically related, which is not observed in *Ensatinia* (Moritz et al. 1992; Kuchta et al. 2009a).

In addition to the adaptive divergence reflected by the color patterns, episodes of historical isolation have been frequent between populations that are ecomorphologically similar. This is more evident within subspecies with wider ranges, such as *platensis* that extends through the Sierra Nevada, and *oregonensis* that extends from the Pacific Northwest to south of San Francisco Bay.
The dynamic geological history of California, with intermittent glacial episodes in the northern California mountains and the Sierra Nevada and tectonic movements involving major geological plates along the coast, coupled with the high habitat dependency and low dispersal rate of salamanders, probably resulted in periods of cessation of gene flow between contiguous populations. This geographic isolation in similar habitats resulted in a fractal genetic structure of phenotypically cryptic units that may be as or more distinct then groups with divergent color patterns. This process can be observed currently in southern California, where montane isolates of the morphologically similar *klauberi* are distinct in allozymes (Fig. 2E).

The few potential secondary contacts that were not detected in our sampling (Fig. 2, #) are typically between populations that diverged both in neutral genetic markers and in morphology: Bob’s Gap (between *klauberi* and *croceater*, Fig. 2E), Lassen Gap (between *platensis* and *oregonensis*, Fig. 2A), and Pajaro Gap (between *xanthoptica* and *eschscholtzii*, Fig. 2C). Some of such cases represent current geographic isolation (e.g., Bob’s Gap is characterized by 160 km of greatly fragmented and generally unsuitable habitat), but it is unclear if other gaps reflect genetic transitions that are narrower than our sampling (Lassen Gap: 15 km; Pajaro Gap: 32 km). We think that some of these gaps occur in areas of strong ecological gradients, although it is not clear if these transitions resulted in noncontiguous populations (micro-vicariance) or very narrow parapatric boundaries (Endler 1977). In comparison with our fine-scale study of the contact between *oregonensis* and *xanthoptica* (study area C, Fig. 2C), a sampling design scaled both to the rate of dispersal of the organism and the degree of environmental change needs to be applied to these potential secondary contacts, to test these hypotheses. However, extensive collecting effort in these areas failed in recovering specimens or finding suitable habitat between localities included in this study (D. B. Wake, pers. obs), suggesting that these are gaps in distribution of the complex that may have been maintained by a combination of the geological history of the terrain (i.e., glaciation, volcanism and recurrent seaways) and human-induced habitat disturbance.

Our results show that nonadaptive divergence was, and probably still is, an important evolutionary process operating in the Ensatina ring species complex, and can result in equivalent levels of genetic differentiation to the ones reflected by adaptive divergence.

Although expectations from theory (Irwin 2002; Kuo and Avise 2005) would predict high frequency of stochastic mitochondrial breaks in low dispersal ring species, which Ensatina exemplifies (maximum of 20 m per generation; Staub et al. 1995), we can refute this hypothesis. The spatial concordance of mitochondrial breaks with nuclear discontinuities (Fig. 2) strongly suggests that the formation of lineages boundaries is reflective of periods of historical isolation, followed by secondary contact.

### Divergence Process and Reproductive Isolation

Although our results confirm that the genetic discontinuities previously reported for mtDNA (Kuchta et al. 2009a) reflect historical interruptions of gene flow, the lack of sympathy of the mitochondrial lineages at a fine scale (Kuchta et al. 2009b) and high genetic differentiation (Fig. 2) opens the possibility that adaptive and nonadaptive divergence around the ring resulted in full reproductive isolation.

Around the ring, with few exceptions, we found regions of secondary contact after adaptive and nonadaptive divergence where reproductive isolation could be tested. We always found hybrid individuals revealing the lack of reproductive isolation between contiguous differentiated populations (Fig. 2, ↔), even when genetic differentiation is high ($D = 0.39$, between *picta* and *oregonensis*). Moreover, the frequency of hybrids is always high ($f > 50\%$) and often reaches 100%, whether or not genetic differentiation is accompanied by morphological divergence (Fig. 3). Population differentiation around the ring, regardless of its extent, was not accompanied by reproductive isolation; genetic admixture in high frequency or complete merger follows secondary contacts.

A contrasting scenario is found in southern California, in the four localities where the two subspecies *klauberi* and *eschscholtzii* are sympatric (Figs. 2E and 3) and where the only deviations from HW equilibrium were detected. Although these populations are at the extremes of the morphological and genetic gradients, hybridization between differentiated populations in three contact zones is rare but possible ($f \leq 20\%$). It is only in the southernmost contact (Cuyamaca Mts) that all individuals were unambiguously assigned to one or the other parental, indicating that reproductive isolation has been achieved (Wake et al. 1989). Although in this contact parents are found in uneven ratios (Fig. 2C), our sample is sufficiently large ($N = 51$) to capture hybridization at lower rates than in the other contacts analyzed.

Contrary to what was expected by several reviews of the genetic differentiation in this ring species (Coyne and Orr 2004; Highton 1998), our analysis shows that genetic and ecomorphological divergence at the levels observed around the ring did not result in reproductive isolation (Fig. 3). Evidence of barriers to reproduction, such as full reproductive isolation or rare hybridization, is only observed at the terminus of the ring, where the most extreme levels of phenotypic and genetic divergence are found.

Moreover, although the mitochondrial evidence reflects historical subdivision of the ancestral populations, the parapatric boundaries of mitochondrial lineages do not reflect the current degree of reproductive isolation. This discordant pattern probably results from the intrinsic characteristics of mitochondrial markers (see Avise 2000). Furthermore, if migration is mainly led by males, as suggested by field estimates in Ensatina (Staub et al. 1995), the admixture rate would also be higher in markers carried...
by both sexes, which could further enhance the differential effect of genetic drift. Therefore, these results call for caution in the interpretation of species boundaries based solely on mitochondrial markers, which could be misleading especially in cases of organisms that are philopatric and that have sex-biased dispersal and/or low vagility.

**GENETIC ISOLATION**

Secondary contacts challenge the genetic integrity of the diverging genomes (Mayr 1963). In the absence of reproductive isolation, it is the reproductive success of the hybrids that will dictate the reversibility of the divergence process. If hybrids can backcross with the parentals, gene flow can effectively link the divergent gene pools and completely reverse the divergence process (e.g., three-spined sticklebacks, Taylor et al. 2006). On the other hand, if hybrids are not reproductively successful due to intrinsic or extrinsic selection, a stable hybrid zone genetically isolates the gene pools and completely reverse the divergence process (e.g., green-eyed tree frog, Hoskin et al. 2005).

Both kinds of interactions are hypothesized to occur in the *Ensatina* ring species. The gradual divergence around the Central Valley was thought to result in secondary intergradation, with gene flow effectively linking contiguous populations around the ring (Dobzhansky 1958). Across the ring, secondary contact would result in hybridization, which could not reverse the species-level divergence achieved by the parental taxa (Stebbins 1949). The dynamic of hybrid zones can inform us about the degree and reversibility of the divergence process. Using methodologies that take advantage of the individual multilocus genotypes, we have analyzed regions of secondary contact throughout the ring of *Ensatina*, without a priori assumptions of the divergence level of the parental populations, and objectively tested the hypothesis initially posed by Stebbins (1949).

Our data show that morphologically similar populations are often characterized by broad genetic transitions, reflected by a high frequency of admixed individuals that occur through several localities neighboring the secondary contact. Good examples are the genetically distinct groups within *platensis* (Fig. 2A) and *oregonensis* (Fig. 2B), which, in addition to the geographically fine population structure and high genetic differentiation ($D \leq 0.25$), show extensive admixture of the parental genomes (spanning about 8 km within *oregonensis* to more than 100 km within *platensis*). These contacts are dominated by hybrids (70–100%) and lack one or both parental populations (Fig. 5A). Moreover, these hybrids likely result from several generations of backcrossing, suggesting that the initial F1s are successfully reproducing over many generations, which affects the genetic integrity of the parental populations over a broad geographic scale. This fits the classical view of a zone of secondary intergradation between populations (Mayr 1963), which, with enough time, may lead to the complete merging of divergent gene pools.

A contrasting dynamic is observed in the secondary contacts across the ring, between *eschscholtzii* and *klauberi* in southern California, even in those without evidence of full reproductive isolation (Fig. 5C). These contacts are dominated by parental individuals (above 80%), which very occasionally produce hybrids that do not reproduce over more than a few generations (mostly F1 or first generational backcross with parentals, Fig. 5C). These observations suggest that initial hybrids are not part of a functioning population and do not affect the genetic integrity of parental populations. This fits the criteria of an interspecific hybrid zone, as is expected at the terminus of a ring species (Mayr 1963).

However, as anticipated by Wake (1997), our results also show that the genetic interactions around the ring between morphologically divergent populations (e.g., subspecies) do not fall in the two discrete categories of secondary intergradation versus hybridization. Between the adaptive groups in *Ensatina*, genetic admixture is not restricted to recent generations of hybrids (Fig. 5B), suggesting reproductively successful hybrids that allow ongoing gene flow between parental populations, which results in relatively broad genetic transitions (i.e., about 30 km between *picata* and *oregonensis*, about 10 km between *croceater* and *platensis*). This scenario is closer to the secondary intergradation around the ring hypothesized by Stebbins (1949), although genetic admixture is never as extensive as between some nonadaptive groups (e.g., within *platensis*, Fig. 2A).

Yet, it is only between adaptive groups where we observe that genetic admixture can be spatially confined, even if not restricted to recent generations of hybrids. Such is the case of the contact between *oregonensis* and *xanthoptica*, in the San Francisco Bay Area, an apparently parapatric boundary that we analyzed at two spatial scales. In the population-based sampling no evidence was found of sympathy between parental populations or introgression in neighboring localities (Fig. 5B). However, by sampling individuals at a much finer scale throughout the contact we identified a region, about 3 km in extent (Pressley Road, Fig. 2C), where all the individuals are hybrids that resulted from several generations of backcrossing (Fig. 5B). The dense spatial sampling around this contact revealed that individuals from neighboring localities (2 km East and 7 km North) are genetically pure for either parental population (Fig. 2C). Together, these observations indicate that, despite the lack of reproductive isolation and the high reproductive success of the hybrids, the narrowness of the genetic transition suggests that gene flow is geographically restricted and this contact zone between adaptively divergent groups might be acting to sharply restrict gene flow. An alternative hypothesis is that secondary contact may be too recent to allow extensive gene diffusion. Although further studies would be necessary to distinguish between these two hypotheses, HW equilibrium and the
lack of pure parental individuals in Pressley Road indicate that the secondary contact is not recent. This interaction does not fit readily in the secondary intergradation scenario, because gene flow is spatially restricted and does not affect adjacent population. Nor does the interaction fit in the hybridization scenario, because hybrids are part of a functioning population and are, at least locally, as or more successful than parentals, producing a hybrid swarm.

Although populations at the extremes of the ring have diverged to the point of full reproductive isolation, or apparent low reproductive success of hybrids, at the lower levels of divergence observed around the ring the differentiated genomes are still porous to gene flow, even when historical isolation was coupled with morphological differentiation. We have shown that in the interactions between nonadaptive groups, where genetic admixture is geographically broad, parental populations seem to be merging back to a single gene pool. When genetic differentiation is coupled with morphological divergence (adaptive groups) this potential to merge seems to be restricted to defined transition zones, and may not affect neighboring populations (e.g., between oregonensis and xanthoptica), compromising the idea of species-wide gene flow, as envisioned by Dobzhansky (1958). Throughout its range, Ensatina has suffered several periods of historical isolation, with varying duration, which resulted in high genetic differentiation that is not necessarily reflected in ecologically relevant traits such as color pattern (e.g., oregonensis). Furthermore, short periods of geographical isolation seem to have allowed morphological differentiation, despite lower levels of genetic differentiation (e.g., croceater). When the genetic integrity of these taxa is challenged in secondary contacts, we observed that the outcome of the divergence process does not fall into discrete categories (secondary intergradation vs. hybridization; Mayr 1963). Instead these interactions reflect a large continuum of possibilities, from (1) unrestricted gene flow between differentiated populations, both spatially and in generations of genetic admixture; (2) to spatially restricted gene flow with high reproductive success of the hybrids; (3) hybridization limited to few generations; and (4) ultimately to full reproductive isolation. We conclude that species formation in Ensatina is proceeding in a manner that might be termed Darwinian (as explicated by Mallet 2008). Adaptive divergence in different directions, extinction, and vicariant processes leading to temporary separation, and diverse outcomes of interactions following recontact have resulted in a vivid example of stages in the establishment of lineage independence, the ultimate criterion of species establishment (de Queiroz 2007). Although prized as an example of evolutionary clarity, Ensatina presents a pattern of taxonomic irresolution (different systematists might recognize from one to many species, depending on what criteria they choose) that one expects from Darwinian species formation.

ACKNOWLEDGMENTS
We thank C. Brown, K. Yanev, M. Frelow, T. Jackman, and D. Nguyen for their valuable assistance in field and laboratory work. We also thank T. Jackman and S. Kuchta for access to published data, and C. Moritz, E. Anderson, and S. Kuchta for helpful discussion. We thank M. Fujita, J. Patton, J. Martinez-Solano, D. Buckley, S. Singhal, T. Devitt, A. Ribeiro, S. Schoville, S. Rovito, and A. Carnaval for their comments on early manuscripts. In addition, the manuscript benefited from comments of two anonymous reviewers and the Associate Editor S. Steppan. This work was supported by National Science Foundation grant EF 0334939 and DEB 0641078. Ricardo Pereira was supported by “Fundaçao para a Ciência e Tecnologia” grant SFRH/BD/18073/2004.

LITERATURE CITED


Supporting Information

The following supporting information is available for this article:

**Table S1.** Collecting localities, sample sizes, and geographic location.

Supporting Information may be found in the online version of this article.

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

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.

Additional results and discussion can be found in a document at http://www.repository.naturalis.nl/record/289893.