JOURNAL OF Evolutionary Biology

doi: 10.1111/j.1420-9101.2011.02418.x

Regional divergence and mosaic spatial distribution of two closely related damselfly species (*Enallagma hageni* and *Enallagma ebrium*)

A. BOURRET*, M. A. McPEEK† & J. TURGEON*

*Département de biologie, Université Laval, Quebec City, QC, Canada
†Department of Biological Sciences, Dartmouth College, Hanover, NH, USA

Keywords:

amplified fragment length polymorphism;
distribution;
local divergence;
Odonata;
patch;
radiation;
speciation.

Abstract

North American *Enallagma* damselflies radiated during the Pleistocene, and species differ mainly by reproductive structures. Although morphologically very different, *Enallagma hageni* and *Enallagma ebrium* are genetically very similar. Partitioning of genetic variation (AFLP), isolation by distance and clustering analyses indicate that these morphospecies are locally differentiated genetically. Spatial analyses show that they are rarely sympatric at local sites, and their distributions form a mosaic of patches where one is clearly dominant over hundreds of square kilometers. However, these morphospecies are also not genetically more similar when they are sympatric, indicating that hybridization is probably not occurring. Given that these morphospecies are ecologically equivalent, strong assortative mating, reproductive interference and fast post-glacial recolonization may explain the origin and maintenance of these distributional patches across eastern North America. By limiting opportunities for gene flow, reproductive interference may play an unsuspected role in accelerating genetic differentiation in the early phases of nonecological speciation.

Introduction

Current research on speciation often aims at identifying and understanding the mechanisms driving the divergence process (Via, 2001; Kirkpatrick & Ravigné, 2002; Butlin *et al.*, 2008; Fitzpatrick *et al.*, 2008, 2009). Ecological speciation, whereby barriers to gene flow result from ecologically based divergent selection, has retained much attention in the recent past (Schluter, 2000; Rundle & Nosil, 2005). However, limits to gene flow arising from stochastic processes (e.g. speciation by polyploidization) or sexual selection are also well documented (Gavrilets, 2000; Panhuis *et al.*, 2001).

This recent focus on mechanisms contrasts with the traditional classification of speciation modes centered on the geographical setting reflecting the initial level of gene flow among diverging populations: null in allopatry to panmixia in sympatry and all intermediate levels of isolation in parapatry, first proposed by Mayr (1942). This classification has been criticized because it dissects a continuum into artificially discrete categories (Rice & Hostert, 1993; Schluter, 2001; Gavrilets, 2003). After extended debates on the possibility and frequency of allopatric vs. sympatric speciation, it is now apparent that speciation generally occurs in parapatry (Gavrilets, 2003; Mallet *et al.*, 2009). Another critique is that current species distributions might not reflect historical geographical ranges, which can shift substantially after speciation (Losos & Glor, 2003). Nevertheless, contemporary species distributions can inform studies of speciation. The spatial context observed at each stage of the speciation process determines the extent of gene flow between diverging populations (Butlin *et al.*, 2008), which in turn affects the probability of speciation (Mallet *et al.*, 2009). Overall, the joint analysis of the mechanisms driving divergence and the spatial distributions of diverging populations may contribute to a better understanding

Correspondence: Julie Turgeon, Département de biologie, Université Laval, 1045 avenue de la Médecine (Vachon 3048), Quebec City, QC G1V 0A6, Canada.
Tel.: 418 656 3135; fax: 418 656 2043;
e-mail: julie.turgeon@bio.ulaval.ca

of the speciation process. For instance, ecological speciation is often first hypothesized on the basis of the distribution of distinct phenotypes or lineages in different environments (Schluter, 2000). In contrast, sexual selection does not necessarily lead to species having distribution tightly linked to any ecological gradients (e.g. Siepielski *et al.*, 2010).

With 38 species, *Enallagma* is the most diverse genus of damselflies in North America (Westfall & May, 2006). The vast majority of these species are morphologically very distinct, and are best discriminated on the basis of adult male and female reproductive structures (Westfall & May, 2006; McPeek *et al.*, 2008, 2009). These reproductive structures constitute a lock and key mechanism, which acts as a prezygotic reproductive barrier among closely related species (Paulson, 1974; Robertson & Paterson, 1982). *Enallagma* species form two old monophyletic clades marked by distinct speciation rates (Brown *et al.*, 2000; Turgeon *et al.*, 2005). One of these major clades is comprised of two groups of very recently radiated species, whose histories have been influenced by cycles of glaciation during the Pleistocene (McPeek & Brown, 2000; Turgeon & McPeek, 2002; Turgeon *et al.*, 2005). Within each radiation, most of these young species are genetically distinct and appear reciprocally monophyletic, even though they share extensive mtDNA polymorphisms (Turgeon *et al.*, 2005).

In North America, the distributions of recently radiated *Enallagma* species are influenced by ecological and nonecological factors. Four species are exclusively found in habitats where dragonfly larvae are the main predators, and these four species were the result of three independent habitat shifts that involved adaptation to dragonfly predation via the evolution of morphological, physiological and behavioural traits (McPeek, 1990a, 1995, 1999, 2000; Stoks *et al.*, 2003). However, many species inhabiting the ancestral habitat where fish are the top predators are ecological equivalents (Siepielski *et al.*, 2010). Local species diversity is often very high, with up to 12 species co-occurring in a lake (e.g. Johnson & Crowley, 1980; McPeek, 1998), and local species assemblages conform to random expectations (Siepielski *et al.*, 2010).

Enallagma hageni and *Enallagma ebrium* are both members of one of these recent radiations (Turgeon *et al.*, 2005). They are morphologically distinct, with adult individuals being easily distinguished based on the shapes of their reproductive structures (i.e. cerci for males and mesostigmal plates for females) (Westfall & May, 2006; McPeek *et al.*, 2008). The shapes of these structures are distinctive and nonoverlapping between species (McPeek *et al.*, 2008, 2011), and individuals are assigned to morphospecies based on the shapes of these structures (we use the term morphospecies throughout this article to emphasize that individuals are assigned to 'species' based on the morphologies of these reproductive structures). Like other sister species derived from the same radiation, they share extensive mtDNA polymorphisms, with the exception of a subgroup of haplotypes specific to *E. hageni* found in New England (the 'Atlantic' clade of Turgeon & McPeek, 2002). Unlike most other species in this radiation, a previous amplified fragment length polymorphism (AFLP) based phylogeny suggested that *E. hageni* and *E. ebrium* are polyphyletic (Turgeon *et al.*, 2005). Although the two morphospecies have nearly identical ranges that extend across North America from the Atlantic Ocean to eastern British Columbia (Westfall & May, 2006), field observations suggest that *E. hageni* and *E. ebrium* are rarely sympatric at local sites, with one clearly predominating over large areas covering tens to hundreds of square kilometers (M. A. McPeek & J. Turgeon, pers. obs.). Moreover, we have found no ecological differences between these abundant and very common morphospecies or among the lakes in which they are found (McPeek, 1989, 1990b; Siepielski *et al.*, 2010).

The goal of this study was to disentangle the evolutionary history of these two seemingly very young morphospecies. To do so, we first establish the relative spatial distributions of these morphospecies over a large area of the eastern portion of their range to gain insight on the potential for reproductive interactions. Then, we analyse patterns of genetic variation within and across regions within this large area. We explore the full range of evolutionary hypotheses to explain the unusually high genetic similarity between *E. hageni* and *E. ebrium* (Turgeon *et al.*, 2005).

First, *E. hageni* and *E. ebrium* may correspond to two young but good species (e.g. as defined by the unified species concept, de Queiroz, 2007), as morphologically based taxonomy readily suggests (H1 – 'two species'). If true, the morphospecies should define distinct and reciprocally monophyletic lineages for sequence data, and independent gene pools at other types of loci (e.g. microsatellites and AFLPs) at all spatial scales.

Second, occasional hybridization between largely independent lineages corresponding to morphospecies may blur genetic distinctiveness (H2 – 'hybridization'). This is suspected because of the rare observations of intermediate morphologies in reproductive structures (Catling, 2001; M. A. McPeek & J. Turgeon, pers. obs.). However, this seems improbable, given that morphospecies rarely co-occur, and that attempts at experimentally crossing these morphospecies have always failed (Fincke *et al.*, 2007; M. A. McPeek & J. Turgeon, unpubl. results). Evidence for this hypothesis would consist of detecting two lineages corresponding to morphospecies, with morphospecies being less genetically differentiated in sympatry than in allopatry.

Third, these morphospecies may be reproductively isolated, but still be polyphyletic and sorting, regionally or locally (H3 – 'multiple lineages'). This hypothesis predicts that interspecific genetic differentiation will occur only below a certain spatial scale, whereas morphospecies will exhibit common genetic patterns at larger spatial scales.

Finally, morphospecies may represent a single biological species with a polymorphic phenotype (H4 – 'one species'). In this case a single, global pattern of distribution of genetic variation within a single lineage is expected, and differentiation between morphs should not exceed that within morph. Note that these hypotheses are not mutually exclusive; for example, hybridization could be detected under H1 (between two main lineages) and H3 (between pairs of regional lineages).

Materials and methods

Morphospecies distributions

The degree of effective sympatry or allopatry between *E. hageni* and *E. ebrium* can help in interpreting genetic signals, suggesting interbreeding as well as weighing evidence for the geographic context of speciation. Herein, our main purpose was to ascertain whether *E. hageni* and *E. ebrium* are randomly distributed relative to one another. To do so, we analysed the distribution of morphospecies frequency at local sites, within regions, and across a wide zone in eastern North America.

Field sampling and museum data were used to establish the relative distributions of the two morphospecies. In 2008 and 2009, we sampled 96 sites in western Quebec and eastern Ontario (Canada), 17 sites in New England (Vermont, New-Hampshire and Maine in the USA) and 25 sites in Prince Edward Island (Canada). At each site, 30–45 individuals were collected whenever possible and morphospecies were identified as above. In all, 4188 individuals (including 135 females) were collected. In addition to field sampling, data from museum and government collections were compiled (hereafter 'museum data'). The complete database included 31 393 individual observations (coded as 0 or 1, with 51.5% *E. hageni*) from 2128 sites distributed across Prince Edward Island, Nova Scotia, New Brunswick, southern Quebec and eastern Ontario in Canada, as well as Maine, New Hampshire and Vermont in the USA (Fig. 1a). Details on data sources and treatment are provided in Appendix S1.

To characterize morphospecies frequency at the scale of the local site, we used sampling data for sites with a minimal sample size of 30 individuals, and we examined the distribution of morphospecies frequency. We also tested the null hypothesis that morphospecies were homogeneously distributed across sites by means of permutations (PERM 1.0, Duchesne *et al.*, 2006).

To estimate the distribution of morphospecies frequency within each region and over the entire study area, we performed ordinary kriging interpolation, a minimum mean squared error method for spatial prediction (Cressie, 1993; Croucher *et al.*, 2007). Using SAS 9.2 (SAS Institute, 2008), we first established an empirical semivariogram (VARIOGRAM procedure). We then used the NLIN procedure to test and choose the best fit function (spherical, exponential, or Gaussian) to estimate the range, the nugget and the sill of the empirical semivariogram. The nugget/sill ratio was used as an indicator of the spatial autocorrelation strength over short distances (Webster & Oliver, 1990). ARCGIS 9.2 (Environmental Systems Research Institute) was used for kriging as well as for mapping.

To assess how the distribution of sites affected morphospecies frequencies, we compared kriging interpolation maps obtained with empirical vs. randomized datasets. The empirical data comprises museum and sampling datasets for the entire zone considered. To randomize morphospecies distributions, we built 100 independent datasets, where morphospecies identity of each individual observation was randomly assigned while respecting the overall empirical morphospecies proportion and site locations. The percentage of territory (divided in 10 km^2 cells) where one morphospecies represented more than 75% of the individuals was calculated and compared between kriging maps based on empirical vs. random datasets. This proportion is somewhat arbitrary, but was suggested by the results for the random dataset where either morphospecies rarely represented ≥75% of the individuals.

Genetic variation

Biological material and genetic characterization

In 2008 and 2009, we sampled adult *E. hageni* and *E. ebrium* from lakes according to a hierarchical design in three geographic regions: Prince Edward Island (PE, Canada), Quebec and Ontario (QO, Canada) and New England (NE, USA) (Fig. 1a). Within each region, we sampled allopatric and sympatric lakes. A lake was considered allopatric for one morphospecies when one morphospecies represented >90% of the individuals (among ≥30 individuals) collected at that lake. At each lake, we aimed to collect 24 adults of each morphospecies if present. Males were directly identified in the field on the basis of cerci morphology (Westfall & May, 2006; McPeek *et al.*, 2008). Females were kept only if caught in tandem or copulating with a male, and morphospecies was confirmed in the laboratory using a dissecting microscope (Westfall & May, 2006; McPeek *et al.*, 2009). Specimens were dried in the field and stored at −80 °C until DNA analyses were performed.

Within each region, we used AFLP to characterize the genotypes of 6–24 individuals of each morphospecies from each of two sympatric and four to six allopatric lakes (Fig. 2; Table S1 for more details on sampling locations). In total, 264 *E. hageni* and 277 *E. ebrium* from 20 lakes were analysed. We extracted DNA following Aljanabi & Martinez (1997), and DNA quality was verified on 2% agarose gels. We quantified DNA using spectrophotometry and diluted samples to 100 ng μL^{-1}. We generated AFLP fragments using the restriction enzymes *Eco*RI and *Mse*I (New England Biolabs, Ipswich,

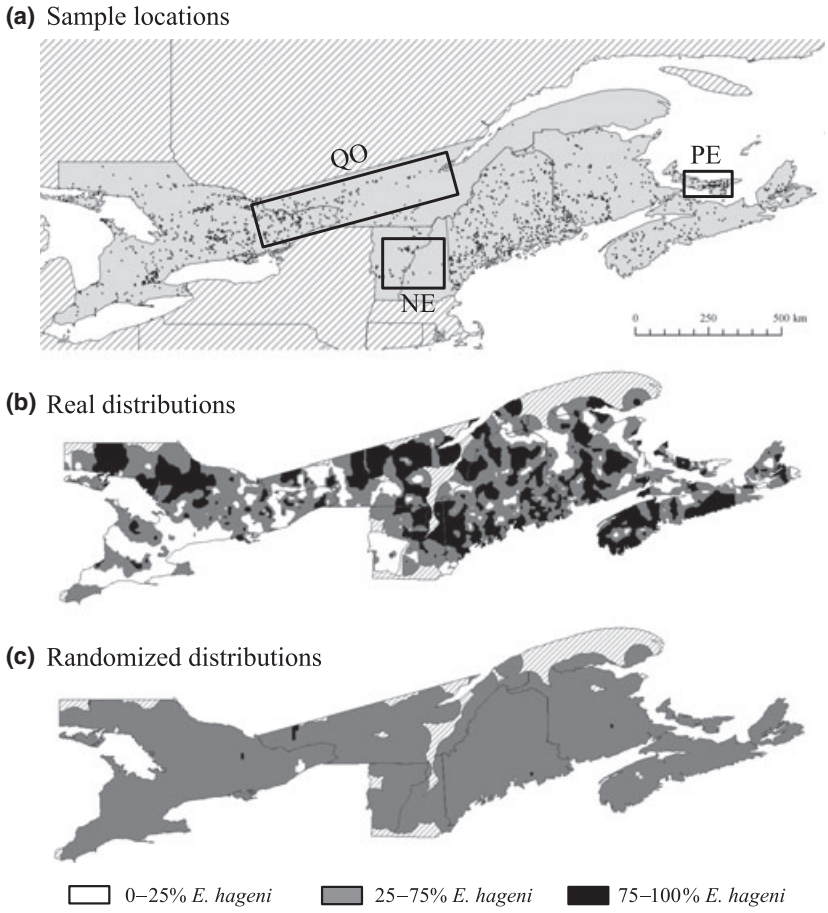

Fig. 1 Sampling regions for genetic analyses and location of sites used in global kriging interpolation of morphospecies frequency distribution. (a) Distribution of sites used in spatial analysis (dots), with boxes delimiting the three regions (QO, NE and PE) where samples were obtained for genetic analyses (see also Fig. 2), (b) kriging results for empirical dataset showing areas where *Enallagma hageni* or *Enallagma ebrium* represent more than 75% of the individuals and (c) a representative distribution map obtained with randomized datasets. Hatched areas could not be used for interpolation because sites were too sparse.

MA, USA) following AFLP® Plant Mapping protocol of Applied Biosystems (2007–2010) with slight modifications. Three *Eco*RI/*Mse*I primer pairs were used in selective PCRs: AGG/CACG, ACC/CACA and ACA/CACA (note that the *Mse*I primer has four selective nucleotides). Selective PCRs included a denaturation step of 20 s at 94 °C, nine 'step-down' cycles with 30 s annealing step beginning at 69 °C and ending at 61 °C, 20 cycles with 30 s annealing step at 52 °C and a final 2 min extension step at 72 °C. We ran PCRs products on an ABI 3100 capillary sequencer with LIZ size standard (Applied Biosystems, Carlsbad, CA, USA). AFLP profiles were checked and scored manually using GeneMapper 3.7 analysis software (Applied Biosystems) with a minimum relative fluorescence set at 100 units.

As a complement to the nuclear data, we sequenced an 884 bp mtDNA fragment (COI and COII genes and the intervening leucine tRNA) following Turgeon & McPeek (2002) for five individuals by morphospecies for three lakes in PE and QO and for all sites in NE. We used a previously published haplotype network to establish the mutational relationships among the observed haplotypes and to interpret associations with morphospecies and/or putative refugial lineages (Turgeon & McPeek, 2002; Turgeon *et al.*, 2005).

Genetic analyses

We used several complementary approaches to ascertain whether genetic polymorphisms conformed to the predictions of our four hypotheses. To assess H1, H2 and H4, we contrasted patterns of genetic variation partitioning, isolation by distance (IBD) and genetic clustering when comparing morphospecies within and among regions. To test H2 more specifically, we compared results for sympatric vs. allopatric (or random, see below) morphospecies within region.

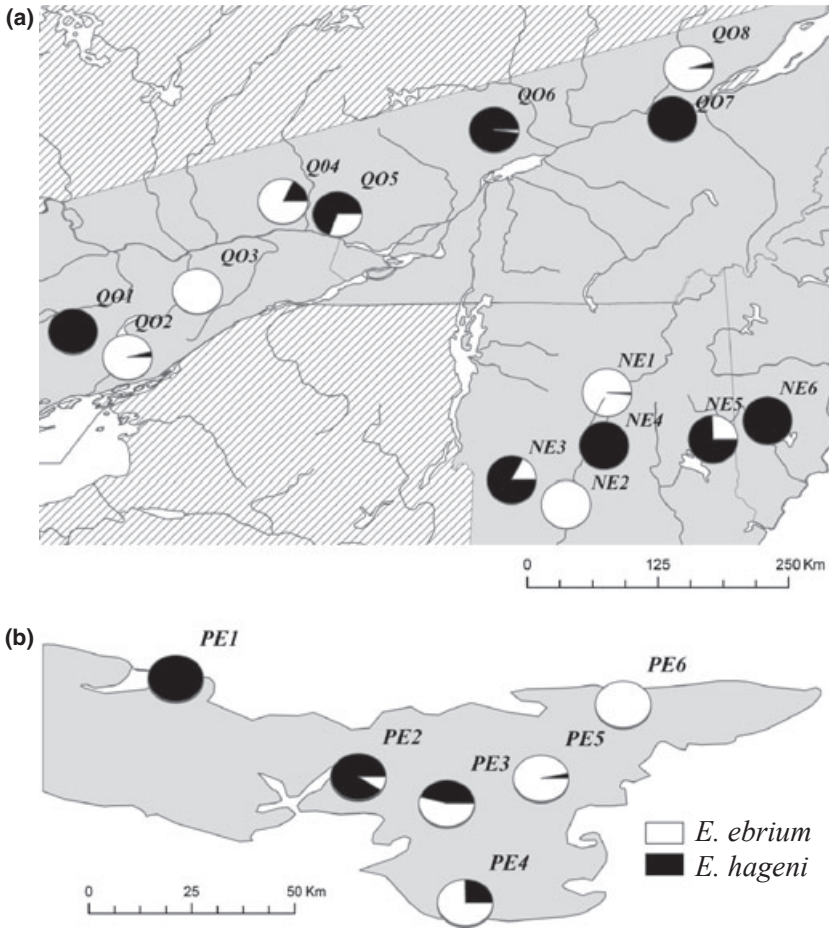

Fig. 2 Morphospecies frequency in sampling sites used in genetics analysis in three regions: (a) Quebec-Ontario (QO) and New England (NE), and (b) Prince Edward Island (PE). See Table S1 for more information on sites and sample sizes.

First, we partitioned genetic variation along contrasting hierarchical grouping models using morphological species, regional morphospecies (i.e. morphospecies within each region) or regions as the higher grouping factor. H1, H3 and H4 would be supported by the prime importance of each of these factors respectively. We performed AMOVAS using ARLEQUIN 3.5 (Excoffier & Lischer, 2010) and compared the explanatory power of models using the corrected Akaike Information Criterion (AIC_c) following Halverson *et al.* (2008). A first analysis was performed with all sites. We then used only those sites where morphospecies were sympatric. In doing so, the influence of distances between sites on interspecific comparisons was mitigated, thereby improving the comparison between expectations associated with H1 and H3.

Second, we examined patterns of genetic differentiation in relation to geographic distances (IBD). H1 predicts that differentiation between morphospecies will always exceed within-morphospecies differentiation. Under H3, higher interspecific differentiation will be restricted to small spatial scales (e.g. a morphospecies signal at the regional scale only). Finally, similar levels of differentiation within and between morphospecies at any spatial scale would support H4. Pairwise F_{ST} between samples (morphospecies at each site) were estimated using Bayesian allele frequency estimation with nonuniform prior distribution with AFLP-SURV 1.0 (Vekemans *et al.*, 2002). Evidence for IBD was assessed by relating $F_{ST}/(1-F_{ST})$ to Euclidian geographic distance between sites for intra- and interspecific comparisons using IBDWS (Jensen *et al.*, 2005).

Third, we identified genetic clusters and assessed their correspondence with morphospecies and regions. H1 predicts that two clusters should be detected, with each cluster being a morphospecies across all regions. H2 largely predicts the same as H1, with some mixing at sympatric lakes. In contrast, H3 predicts that clusters corresponding to morphospecies will be detected only at reduced spatial scale (e.g. within region) and that a cluster associated with a morphospecies in one region will not necessarily be associated with the same morphospecies in another region. Finally, H4 predicts that

clustering will mostly reflect geographical location and proximity, and clusters should not correspond to morphospecies. If hybridization occurs, there should be a higher incidence of mixed ancestry (intermediate *q* values) in clusters associated with each morphospecies in sympatric vs. allopatric sites.

We used STRUCTURE 2.3.1 (Pritchard *et al.*, 2000; Falush *et al.*, 2007; Hubisz *et al.*, 2009) and performed the analyses with and without information about samples (morphospecies by site as loc-prior). We performed analyses with the entire dataset as well as for each region independently. Given the low genetic differentiation within regions (see Results), the model considering sample information (with loc-prior option) was preferred (Hubisz *et al.*, 2009). We set burn-in to 50 000 iterations and subsequent run lengths to 200 000 iterations using the Admix model (using the model with No Admixture provided qualitatively very similar results). We did 10 runs for each K value tested (K = 1–8 clusters for PE and NE, K = 1–10 for QO and the entire dataset). We used Ln $P(X|K)$ (Pritchard *et al.*, 2000) and ΔK (Evanno *et al.*, 2005) as criteria to infer the number of clusters (K). Figures were made with DISTRUCT (Rosenberg, 2004).

Finally, we tested the hybridization hypothesis (H2) using the population-level randomization procedure of Mims *et al.* (2010). If morphospecies exchange genes when co-occurring, then pairs of sympatric morphospecies should be less differentiated than random population pairs of each morphospecies. Otherwise, genetic similarity more probably reflects shared ancestral polymorphisms in these recently radiated morphospecies. These analyses were performed within each region because F_{ST} for random allopatric pairs from different regions would unduly boost F_{ST} values such that sympatric pairs (necessarily within the same region) could fall in the (potentially significant) low range of values solely because of their geographic proximity. The mean F_{ST} observed for sympatric pairs was compared with the null distribution of mean F_{ST} values between two random pairs of samples involving different morphospecies (10 000 randomizations). The proportion of randomized F_{ST} values that are smaller or equal to the observed value provides a P-value for the one-tailed test of no-association between differentiation and co-occurrence. Following the same logic, if hybridization occurs, sympatric morphospecies should contain a greater proportion of intermediate, hybrid-like genotypes. We used AFLPOP (Duchesne & Bernatchez, 2002) to estimate the proportion of genotypes belonging to each parental (morphospecies) and hybrid (F_1, F_2, backcrosses) classes for all pairs of morphospecies. We then used the above randomization procedure to test whether observed sympatric pairs comprised, on average, an equal or higher proportion of hybrid-like genotypes than randomly chosen pairs. Again, analyses were performed independently within regions.

Results

Spatial distribution

One morphospecies was generally dominant at any given lake, with one morphospecies accounting for more than 90% of the individuals in 62% of the sites (Fig. 3). Across lakes, the distributions of morphospecies were clearly heterogeneous (PERM membership homogeneity test, $P < 0.001$).

At the regional scale, moderate spatial autocorrelation was evident over small distances, as estimated by the nugget/sill ratio (range: 0.24–0.74, Table 1). The kriging map based on empirical data revealed a mosaic of patches where one or the other morphospecies clearly dominated (Fig. 1b). Overall, one morphospecies represented 75% or more of the individuals over 53% of the studied area

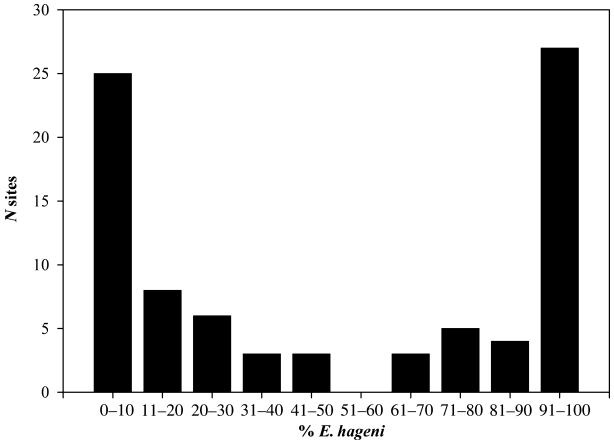

Fig. 3 Frequency of *Enallagma hageni* among 84 sites sampled in PE, NE and QO.

Table 1 Results of kriging interpolation analyses of morphospecies distribution. The proportion of the territory where *one* morphospecies was dominant is indicated.

			% Where dominant (≥75%)		
State/province	*N* entries	Nugget/sill ratio	*Enallagma hageni*	*Enallagma ebrium*	Total
ON	9391	0.27	21	30	51
QC	2278	0.47	42	13	55
VT/NH	1156	0.37	25	31	56
ME	8217	0.63	43	11	54
NB	4503	0.74	27	24	51
NS	2574	0.24	44	13	57
PE	3274	0.32	33	35	68
Global	31393	–	33	21	53

ON, Ontario; QC, Quebec; NB, New Brunswick; NS, Nova Scotia; PE, Prince Edward Island in Canada; VT, Vermont; NH, New Hampshire; ME, Maine in the USA (see also Fig. 1).

(Table 1). This pattern was also apparent within each region (range: 51–68%, Table 1). This is in sharp contrast with the kriging interpolation using randomized datasets where, on average, only 0.7% (range: 0.1–2.2%) of the territory was similarly dominated by either morphospecies (Fig. 1c).

Genetic variation

We amplified a total of 347 AFLP loci, of which 120 were polymorphic using a 5% criterion with all individuals. We replicated 45 genotypes (7.6%) from the restriction step, yielding a low genotyping error rate of 1.2% (Bonin *et al.*, 2004).

Partitioning of genetic variance

AMOVA analyses revealed that genetic variation was best partitioned by regional morphospecies (Table 2), as predicted by H3. Using all sites, the regional morphospecies grouping model (AIC_c = 1493) was preferred (Table 2a). No significant variation was explained by morphospecies (H1; P = 0.185). Partitioning variation by regions (H4) explained slightly less variance and was associated with higher AIC_c values than the preferred model (4.00% vs. 4.34%, AIC_c = 1503 vs. 1493, Table 2a). Analyses of only sympatric sites provided very similar support for H3 (Table 2b); regional morphospecies grouping (AIC_c = 634) better explained genetic variation partitioning than morphospecies (AIC_c = 647) or regional grouping (AIC_c = 640), and again, no variation was explained by morphospecies (P = 0.432).

Isolation by distance

Isolation by distance was apparent when all sites and both morphospecies were considered ($P < 0.001$, Fig. 4a). Clearly, interspecific differentiation was not always higher than intraspecific differentiation, offering no support for H1. Likewise, the different IBD patterns for each morphospecies ($P < 0.001$ for both *E. ebrium* and *E. hageni*), with a steeper regression slope for *E. hageni* than *E. ebrium*, is not compatible with H4.

In contrast, the pattern predicted under H3 was apparent. Interspecific differentiation seemed to exceed intraspecific differentiation only at a relatively small spatial scale (<*ca.* 200 km), and it was intermediate between intraspecific-levels at large distances (e.g. >600 km; Fig. 4a). To test whether F_{ST} values were significantly larger than expected at small distances only, we used an approach combining stepwise regression and permutations of F_{ST} residuals (Fig. 4b). Under the hypothesis that differentiation is solely related to

Table 2 Partitioning of genetic variance for grouping models considering morphospecies (H1), regional morphospecies (H3) or region (H4) as the main grouping factor using (a) all sites and (b) only sympatric sites. Corrected Akaike Information Criteria (AIC_c) were calculated following Halverson *et al.* (2008).

Grouping model	d.f.	SS	% Variation	Fixation index	*P*
(a)					
By morphospecies – H1 (AIC_c = 1519)					
Between morphospecies	1	51	0.26	0.003	0.185
Among sites within morphospecies	24	920	6.72	0.067	< 0.001
Within sites	515	7906	93.02	0.070	< 0.001
By regional morphospecies – H3 (AIC_c = 1493)					
Among regional morphospecies	5	456	4.34	0.043	< 0.001
Among sites within regional morphospecies	20	515	3.06	0.032	< 0.001
Within sites	515	7906	92.60	0.074	< 0.001
By regions – H4 (AIC_c = 1503)					
Among regions	2	300	4.00	0.040	< 0.001
Among sites within regions	23	671	4.00	0.042	< 0.001
Within sites	515	7906	92.00	0.080	< 0.001
(b)					
By morphospecies – H1 (AIC_c = 647)					
Between morphospecies	1	38	0.00	0.000	0.431
Among sites within morphospecies	10	356	6.54	0.065	< 0.001
Within sites	218	3348	93.46	0.065	< 0.001
By regional morphospecies – H3 (AIC_c = 634)					
Among regional morphospecies	5	252	4.06	0.041	< 0.001
Between sites within regional morphospecies	6	142	2.82	0.029	< 0.001
Within sites	218	3348	93.12	0.068	< 0.001
By regions – H4 (AIC_c = 641)					
Among regions	2	146	3.49	0.035	< 0.001
Among sites within regions	9	248	3.95	0.041	< 0.001
Within sites	218	3348	92.56	0.074	< 0.001

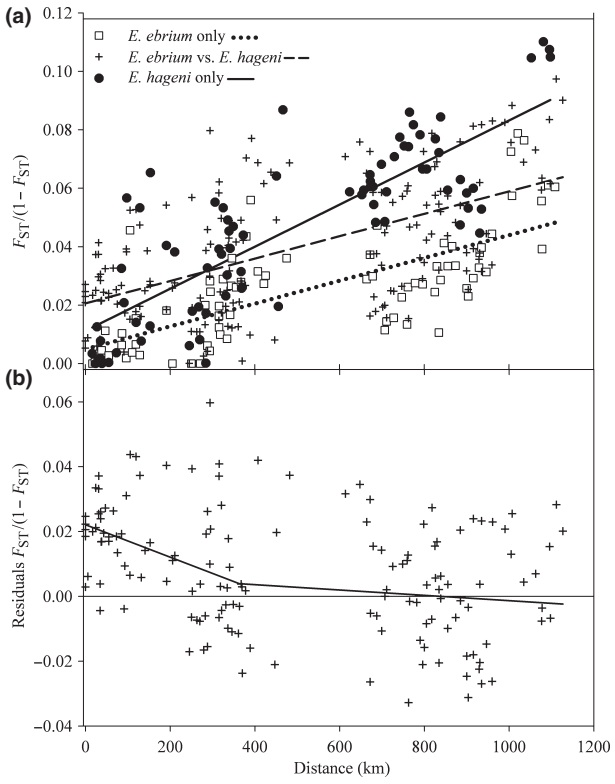

Fig. 4 Genetic differentiation in relation with distance: (a) Isolation by distance pattern for intraspecific and interspecific comparisons. (b) Interspecific residual differentiation in relation with distance. A piecewise regression slope is fitted (solid line), indicating a break at 363 km. See Results for rationale for this analysis.

distance, residuals of F_{ST} regressed on distance should be distributed around zero with no systematic bias related to distance. The global IBD regression slope for long distance comparisons (>600 km) was used as a prediction for the expected interspecific IBD pattern at all spatial scales. Indeed, F_{ST} between regions (>600 km) should not be influenced by the local divergence processes underlying H3. In contrast, higher F_{ST} residuals at small distances would support H3. We fitted a piecewise regression and found that slopes differed below and above 363 km ($P < 0.001$, Fig. 4b). Residuals seemed larger below this threshold distance. To assess the statistical significance of this apparent trend, residuals were randomly permuted across distances. This was done 10 000 times and the sum of squared residuals originally located within the short distance group (<363 km) were generally larger than the sum generated by an equal number of residual values randomly chosen from the set of all residuals ($P = 0.018$). Note that for this method, we excluded comparisons involving *E. hageni* sampled in two sympatric lakes: NE5 and NE6. As is shown below, these two sites proved to harbor representatives of a different mitochondrial lineage (Atlantic clade, Turgeon & McPeek, 2002) that were also strongly differentiated at AFLP loci, distorting the general IBD pattern (see Appendix S2 for F_{ST} values).

Clustering analyses

When cluster analyses using STRUCTURE were applied to the entire dataset, the preferred value for K was $K = 6$ (see Appendix S3). The two clusters at $K = 2$ did not correspond to morphospecies, contrary to the prediction of H1. At $K = 3$, each region was slightly dominated by one cluster, and one of the two clusters present in PE was associated with *E. ebrium*, a pattern that is not compatible with H4.

Regional clustering analyses supported H3 by establishing a general correspondence between clusters and morphospecies within each region (Fig. 5). In PE, clustering by morphospecies was unambiguous, with $K = 2$ clusters using both Pritchard's and Evanno's criteria (Fig. 5a). In NE, both criteria supported the existence of

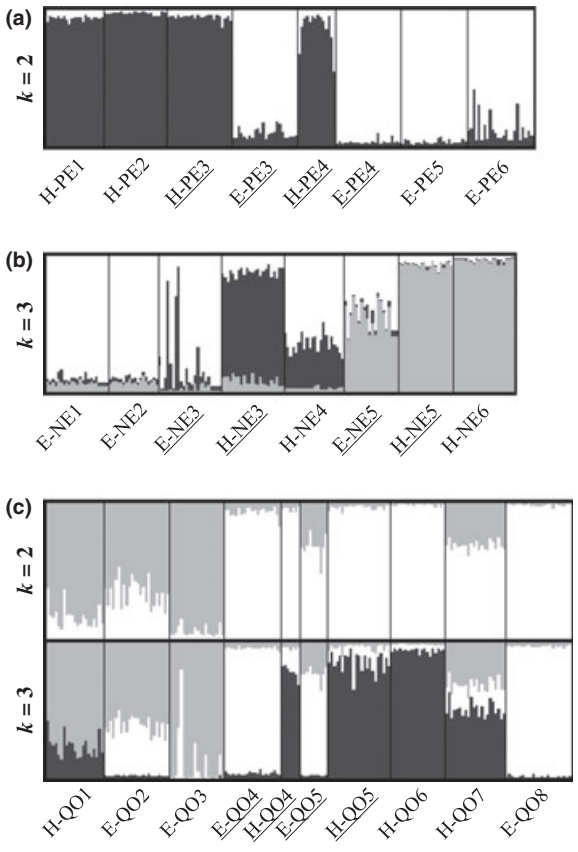

Fig. 5 Regional clustering of amplified fragment length polymorphism genotypes for the most likely number(s) of clusters (K) in (a) PE, (b) NE and (c) QO. Within each region, sites are presented following a west to east gradient. Sites where both morphospecies were present are underlined, with the first letter specifying morphospecies (H: *Enallagma hageni*, E: *Enallagma ebrium*).

three clusters (Fig. 5b). Three samples of *E. ebrium* from westernmost sites clearly formed the bulk of the first cluster (*E. ebrium* from NE1, NE2, NE3), and the other *E. ebrium* sample (from site NE5, but see below) also had ancestry in this cluster. The other two clusters were more strongly associated with *E. hageni*, one with the two easternmost sites (NE5 and NE6) and the other with western sites (NE3 and NE4). In QO, the most likely number of clusters depended on the criterion used and reflected geography ($K = 2$, Evanno's criterion) or morphospecies ($K = 3$ Pritchard's criterion) (Fig. 5c). Indeed, at $K = 2$, the genetic clusters correspond approximately to Ontario (QO1–QO3) vs. Quebec (QO4–QO8) sites, whereas at $K = 3$, the third cluster clearly corresponded to *E. hageni*, especially in Quebec. Note that results from analyses within each region largely correspond to those obtained with the entire dataset for $K = 6$ (see Appendix S3).

Comparisons of morphospecies in sympatry vs. allopatry

We found no evidence for hybridization (H2). When in sympatry, morphospecies were not less differentiated than random pairs of morphospecies. In all three regions, average F_{ST} values were not lower for sympatric when compared with random pairs of samples (all P-values >0.23, see Appendix S4). Similarly, we found no evidence that *E. ebrium* or *E. hageni* sampled in sympatry comprised a higher proportion of genotypes resembling hybrids (all P-values >0.246, see Appendix S4). It is worth noting that detecting genotypes that suggest hybrids does not mean that true hybrids are present. Instead, these statistical figures suggest that morphospecies are only slightly differentiated and share polymorphisms.

Moreover, when in sympatry, morphospecies belonged to distinct genetic clusters (Fig. 5). In PE, this clustering pattern was very clear (Fig. 5a), as well as in Quebec for $K = 3$ (QO4–QO8, Fig. 5c). In both regions, each morphospecies had much stronger ancestry in the cluster typical of the morphospecies in allopatry, and individuals were no more of mixed ancestry in sympatry than when morphospecies were allopatric. In NE, sympatric morphospecies were generally very distinct at site NE3, but *E. ebrium* from site NE5 had mixed ancestry in clusters associated with each morphospecies. In a few cases, allopatric morphospecies also displayed mixed ancestry, such as *E. hageni* in NE4 and QO7.

Distribution of mtDNA haplotypes

MtDNA polymorphisms revealed different patterns of genetic similarities between morphospecies within each region (Fig. 6). Within each of PE and QO, haplotypes were extensively shared between morphospecies. Each region comprised distinct sets of haplotypes, but all belonged to what we have previously called the 'Continental' group of haplotypes documented in both morphospecies (Turgeon & McPeek, 2002). In contrast, morphospecies in NE were associated with distinct sets of haplotypes. Haplotypes previously labelled as those of the Atlantic *E. hageni* lineage were found not only in all *E. hageni* belonging to the AFLP cluster characterizing the easternmost NE sites but also in *E. hageni* from sites located more to the west. *Enallagma hageni* also possessed haplotypes typical of the Continental lineage (NE4). Most *E. ebrium* from NE carried one of two haplotypes (H-008, H-009, Fig. 6) very commonly documented in *E. ebrium* from an extensive area in NE (see also Turgeon & McPeek, 2002).

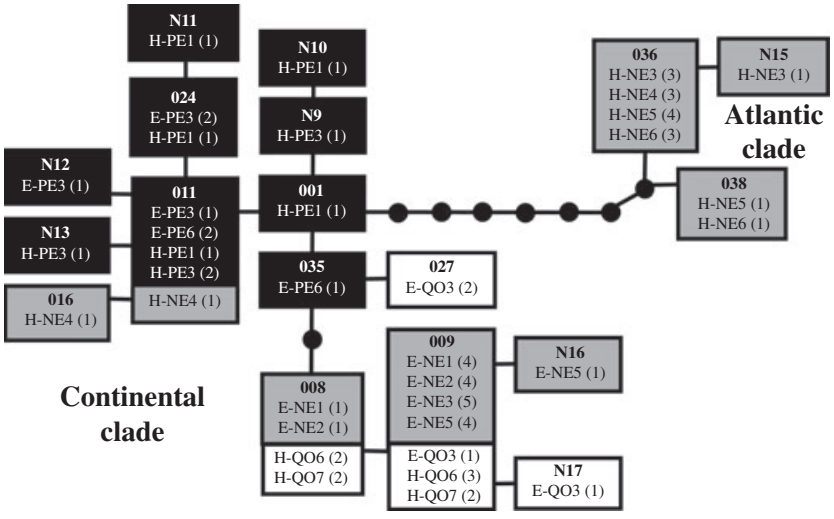

Fig. 6 MtDNA haplotype network relating sequences of *Enallagma hageni* (H) and *Enallagma ebrium* (E) documented in this study. Haplotypes are numbered as per Turgeon *et al.* (2005), with new haplotypes starting with '*N*'. Numbers in parentheses indicate the observed occurrence of each haplotype. Haplotypes collected in each region are shown by different fill colour: white = QO, gray = NE, and black = PE.

Discussion

Ongoing regional differentiation of *E. hageni* and *E. ebrium*

Genetic analyses concur in supporting the predictions associated with ongoing regional differentiation between *E. hageni* and *E. ebrium* (H3). We found evidence for genetic differentiation at a small spatial scale only. Indeed, genetic variation between morphospecies at a regional scale best explained overall variation partitioning, and genetic differentiation between morphospecies was most pronounced at this regional scale.

Likewise, morphospecies from different regions did not form clusters; rather, morphospecies could only be identified as clusters within each region. This was most clearly revealed in the smallest and insular region of PE, where genetic clusters were sharply defined and these genetic clusters perfectly matched morphospecies identities. In NE, *E. ebrium* was associated with one cluster, whereas *E. hageni* was split into two clusters. One *E. hageni* cluster probably corresponds to an Atlantic refugial lineage already documented in other studies, thus reflecting the historical, probably more ancient split within this morphospecies (Brown *et al.*, 2000; Turgeon & McPeek, 2002; Turgeon *et al.*, 2005). Indeed, this cluster included all *E. hageni* individuals from the easternmost sites (NE5 and NE6) and all these possessed mitochondrial haplotypes typical of the Atlantic clade. This may explain why *E. hageni* from these sites are more differentiated from other *E. hageni* samples than is expected from IBD alone (see Appendix S2 for F_{ST} values). In QO, the morphospecies signal was confounded by a geographic signal. This region is much larger than PE and NE, and so interspecific comparisons are being made between very distant sites (nearly 500 km). These comparisons are not likely to reveal strong differentiation between morphospecies, as our extended IBD analysis suggested. In fact, IBD within this region probably blurs the clustering pattern (Guillot *et al.*, 2009). Our sampling design could not *a priori* match the scale of the divergence processes under investigation, resulting in the imperfect clustering of morphospecies within this region. Nevertheless, morphospecies formed clusters over smaller distances, for instance in eastern QO. In addition, distinct collections of mtDNA haplotypes existed within each region, further supporting the local genetic divergence of these morphospecies. Similar geographical clustering of haplotypes shared among species has been documented for this clade (Turgeon & McPeek, 2002; Turgeon *et al.*, 2005).

Ongoing regional differentiation (H3) is also much better supported than the alternative hypotheses. We found no convincing evidence that morphospecies form two globally distinct lineages (i.e. reject H1). The morphospecies were clearly not genetically partitioned based on the AMOVA analysis, and interspecific differentiation was not generally substantially larger than intraspecific levels. Each morphospecies displayed distinct IBD patterns, suggesting independent gene pools. However, the higher dispersal propensity of *E. ebrium* (McPeek, 1989), characterized by a shallower IBD slope, is probably a better explanation than strictly distinct species. Moreover, as mentioned above, morphospecies from different regions did not form clusters. Also, the hypothesis of a single lineage comprising two alternative morphotypes (H4) is easily refuted. Within region, morphospecies most generally belonged to distinct genetic clusters, particularly when in sympatry.

Finally, we found no evidence that contemporary hybridization is commonly occurring (H2). F_{ST} values and genetic clustering patterns provided no evidence that morphospecies are less different when sympatric, or that they comprise a higher proportion of genotypes that could have originated from hybridization within the last few generations. Sharing of ancestral polymorphisms is a better explanation than hybridization for the genetic similarity of these young morphospecies. Moreover, despite a large sampling effort, only eight individuals (i.e. 0.2%), all from different sites, possessed unusual cerci morphologies. The unusual morphology may well be the result of developmental malformation rather than the consequence of hybridization. It is important to note, however, that hybridization may have been more common in the past. For example, clustering analysis suggest that *E. ebrium* individuals at site NE5 in New England (Fig. 5b) are of mixed ancestry between the Atlantic *E. hageni* lineage and *E. ebrium*. This hybridization may have been asymmetrical (female *E. ebrium* × male *E. hageni*) given that none of these *E. ebrium* individuals possessed *E. hageni* Atlantic mtDNA haplotypes. Localized, past asymmetrical hybridization has also been documented between other pairs of species within this recently radiated *Enallagma* clade (Turgeon *et al.*, 2005). The historical character of hybridization is also revealed by the fact that allopatric *E. hageni* individuals (NE4) had mixed ancestry in clusters associated with both morphospecies.

Enallagma hageni and *E. ebrium* mosaic distribution

The full ranges of *E. hageni* and *E. ebrium* distribution ranges are nearly coincident across northern North America, but these morphospecies generally do not co-occur at local and regional scales within these ranges. Locally, one morphospecies is usually very dominant, and morphospecies are rarely equally frequent when sympatric. Across a large portion of these morphospecies ranges, lakes with the same dominant morphospecies are aggregated, creating a mosaic of patches alternating in morphospecies dominance (i.e. Fig. 1b). Our previous experiences in other parts of their ranges also suggest that this to be true across their entire ranges (M.A. McPeek, pers. obs.) and confirm the field experiences

reported by many other workers (Walker, 1953; P.M. Catling, M.R.L. Forbes and P.M. Brunelle, pers. comms.).

Habitat heterogeneity is the common explanation for such mosaic distributions. Competitive interaction for resources and habitat preference (e.g. insect host plant) can create patchiness in the distributions of closely related species (Miller, 1963; Howard & Harrison, 1984; Bridle *et al.*, 2001). However, no ecological differences are known between *E. hageni* and *E. ebrium*, despite extensive work to identify them (e.g. McPeek, 1989, 1990b; Siepielski *et al.*, 2010). Moreover, the scale and distribution of these patches do not, to our knowledge, correspond to those of any biotic or abiotic environmental factors. Up to 12 *Enallagma* species can be found together at lakes containing fish across eastern North America (Johnson & Crowley, 1980; McPeek, 1990a, 1998), and most *Enallagma* species have very broad and overlapping ranges. Almost all species can be found at every lake containing fish. Only *E. hageni* and *E. ebrium* show such a mosaic pattern of distributions relative to one another (e.g. McPeek, 1989, 1990b, 1998; Siepielski *et al.*, 2010). We have identified no environmental factor that can account for this segregation after many years of searching. To our knowledge, very few other examples of closely related species with similar relative distributions have been identified (but see below). Such mosaic patterns certainly go undetected when entire ranges overlap, suggesting that species should be sympatric. Likewise, studies considering only presence vs. absence data cannot detect mosaic patterns, because rare and dominant species are given the same importance.

We hypothesize that reproductive interference between *E. hageni* and *E. ebrium* is more likely to explain local allopatry and the maintenance of patches where one morphospecies is clearly dominant. Reproductive interference comprises any interspecific sexual interaction that negatively affects the fitness of at least one of the species involved (Gröning & Hochkirch, 2008; Burdfield-Steel & Shuker, 2011). Reproductive interference can take many forms (e.g. misdirected courtship, heterospecific mating, hybridization) and may lead to different issues (e.g. sexual exclusion, spatial segregation, reproductive character displacement). In a fashion similar to the ecological competitive exclusion process, reproductive interference between closely related species may lead to local exclusion (Kuno, 1992). For example, the mosaic distribution pattern observed between two ground-hopper species (*Tetrix ceperoi* and *T. subalata*) could be a consequence of reproductive interference (Gröning *et al.*, 2007; Hochkirch *et al.*, 2007). These closely related species broadly overlap in their range and general ecological requirements, but rarely co-occur at local scale (Gröning & Kocum, 2005 in Gröning *et al.*, 2007). The incomplete mate recognition systems may be more relevant than habitat partitioning to explain such distributional pattern (Gröning *et al.*, 2007).

In *Enallagma*, males are highly promiscuous and attempt to mate with all *Enallagma* females they encounter, regardless of species (Paulson, 1974; Fincke *et al.*, 2007). Males initiate mating by grasping females with their cerci, but males cannot force females to mate. Females identify males to species based on the tactile cues she receives as the male's cerci grasp her thoracic plates (Paulson, 1974; Robertson & Paterson, 1982), which are the same structures used by taxonomists to identify species (Westfall & May, 2006). Females signal rejection to heterospecific males by refusing to mate and such interactions can last up to 2 min before the female is released by a male she has rejected as a mate (Fincke *et al.*, 2007; McPeek & Turgeon, unpubl. data). Thus, when either *E. hageni* or *E. ebrium* is relatively rare at a site, as is commonly observed, heterospecific mating attempts may lead to a greatly reduced mating success of the rare morph, which may in turn lead to its gradual exclusion locally.

Post-glacial recolonization processes may also have played a significant role in first establishing the contemporary mosaic distribution of *E. hageni* and *E. ebrium*. Quaternary climatic oscillations shaped the genetic and spatial structure of many species (Hewitt, 2004), and *E. hageni* and *E. ebrium* experienced past range expansions following this period (Turgeon *et al.*, 2005). The colonization of open habitat after the last ice age may have favoured long distance dispersal, a process that can create patchy population structure (Nichols & Hewitt, 1994; Ibrahim *et al.*, 1996; Bialozyt *et al.*, 2006; Ray & Excoffier, 2010). Moreover, simulations have shown that long distance dispersal, when coupled with assortative mating, can lead to both the formation and maintenance of mosaic distributions (M'Gonigle & FitzJohn, 2010). Female control over mating result in strong assortative mating in *Enallagma* (see above), and reproductive interference, if occurring, probably strengthens assortative mating between these sexually interacting species. Thus, both phenomena may have helped establishing and maintaining the mosaic distribution pattern by countering the immigration of morphospecies into a patch dominated by the other morphospecies.

Evolutionary history of *E. hageni* and *E. ebrium*

Enallagma hageni and *E. ebrium* are very young species that are part of a recent radiation linked to the last glaciation (Turgeon *et al.*, 2005). Like most *Enallagma* species, *E. ebrium* and *E. hageni* represent classic morphospecies long recognized in taxonomy (Westfall & May, 2006). These are principally discriminated on the basis of the male caudal cerci and female thoracic plates (Paulson, 1974; Robertson & Paterson, 1982; Westfall & May, 2006). Our genetic results confirm that *E. hageni* and *E. ebrium* do not regularly interbreed, and, in addition, that their current spatial distribution offers little opportunity to do so. However, unlike other species

from this radiation, these two morphospecies are not yet clearly genetically differentiated from one another. These morphospecies are most likely undergoing local or regional differentiation, and thus are still largely polyphyletic when the entire species are considered.

The regional pattern of genetic differentiation also suggests that one morphospecies [most likely *E. ebrium* (Turgeon *et al.*, 2005)], and the associated typical caudal cerci morphology, may have appeared in parallel more than once. Distinct regional pools of mtDNA haplotypes shared by both morphospecies offer support for this hypothesis. Moreover, multiple appearances of the same reproductive structure have already been observed in *Enallagma*. For example, Palaearctic *E. cyathigerum* and Neartic *E. annexum* were previously regarded as one species on the basis of highly similar cerci (Westfall & May, 1996); however, phylogenetic relationships clearly show that they are highly divergent species and that the same cerci morphology very likely evolved twice (Stoks *et al.*, 2005; Turgeon *et al.*, 2005; Westfall & May, 2006). Such repeated evolution of the same cerci type could result from developmental or genetic constraints. However, the wide variety of caudal cerci morphology observed among the 17 species that recently radiated rather suggests that the diversification of cerci morphology is not severely constrained (McPeek *et al.*, 2008). Hence, if our results are representative of the spatial scale at which speciation proceeds, this could potentially imply an unusually high number of instances of parallel evolution of identical reproductive structures over the very large distributional ranges of these morphospecies.

Alternatively, the new morphospecies may have appeared only once, early in the radiation, such that both morphospecies could have participated in establishing the mosaic of allopatric patches during the postglacial recolonization process. In this case, patterns of different regional groups of haplotypes shared between morphospecies would result from past hybridization after colonization. We have found no clear evidence for contemporary hybridization in this study, but shared cluster membership in some sampling sites may be indicative of past reproductive contacts between morphospecies following the establishment of their morphological differentiation. Also, hybridization is known to have occurred between very distant species upon colonization of recently deglaciated areas (Turgeon *et al.*, 2005). Notwithstanding that a unique appearance seems more parsimonious than multiple, parallel evolution events, these alternative scenarios are still currently very speculative. Our data are based on a large number of genetic markers, but these are probably neutral markers mostly reflecting the history of migration and drift within morphospecies. To discern between single and multiple origins of traits discriminating these morphospecies, sequence information of functional genes controlling cerci shape would be much more instructive.

In conclusion, the joint analysis of genetic variation and relative distribution of *E. hageni* and *E. ebrium* revealed that these young morphospecies are still polyphyletic and locally diverging. Although ecologically equivalent and with similar distribution ranges, they rarely co-occur locally and their distributions form an unusual mosaic of patches. Assortative mating, and possibly reproductive interference, coupled with postglacial recolonization may have played a role in generating and maintaining this peculiar distributional pattern, which is conducive to local divergence. Our results call for more attention on the action of frequency dependent selection in the study of recent speciation events involving species that are ecologically similar or derived from sexual selection. Although reinforcement helps consolidate reproductive isolation between diverged lineages, reproductive interference may play a crucial role by limiting interactions early in the divergence process.

Acknowledgments

This work is the central part of A. Bourret M.Sc. thesis under the supervision of J. Turgeon. We thank P. Favriou and C. Lehoux for field and laboratory assistantship and P. Duchesne for statistical help. We also thank N. Donnelly and B. Mauffray (International Odonata Research Institute), D. Doucet, S. H. Gerriets and R. Meherzad (Atlantic Canada Conservation Data Center), B. Henson and A. Lapenna (Ontario Natural Heritage Information Center), J. Louton (National Museum of Natural History), D. McAlpine (New Brunswick Museum), G. Pelletier (Laurentian Forestry Centre of Canadian Forest Service), R. Pupedis and M. Thomas (Yale Peabody Museum) for sharing database from entomological collection, as well as P. M. Brunelle, P. M. Catling, R. Curley, D. Doucet, M. R. L. Forbes, G. Lemelin, M. Ludvik and J. M. Perron for discussion and hints about species distribution, and anonymous reviewers for constructive comments. This work was supported by a FQRNT scholarship to A. Bourret, an NSERC (Canada) Discovery Grant to J. Turgeon, and National Science Foundation (USA) grant DEB-0516104 to M. A. McPeek.

References

Aljanabi, S.M. & Martinez, I. 1997. Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Res.* **25**: 4692–4693.

Applied Biosystems. 2007–2010. *AFLP® Plant Mapping protocol.*

Bialozyt, R., Ziegenhagen, B. & Petit, R.J. 2006. Contrasting effects of long distance seed dispersal on genetic diversity during range expansion. *J. Evol. Biol.* **19**: 12–20.

Bonin, A., Bellemain, E., Eidesen, P.B., Pompanon, F., Brochmann, C. & Taberlet, P. 2004. How to track and assess genotyping errors in population genetics studies. *Mol. Ecol.* **13**: 3261–3273.

Bridle, J.R., Baird, S.J.E. & Butlin, R.K. 2001. Spatial structure and habitat variation in a grasshopper hybrid zone. *Evolution* **55**: 1832–1843.

Brown, J.M., McPeek, M.A. & May, M.L. 2000. A phylogenetic perspective on habitat shifts and diversity in the north American *Enallagma* damselflies. *Syst. Biol.* **49**: 697–712.

Burdfield-Steel, E. & Shuker, D.M. 2011. Reproductive interference. *Curr. Biol.* **21**: R450–R451.

Butlin, R.K., Galindo, J. & Grahame, J.W. 2008. Sympatric, parapatric or allopatric: the most important way to classify speciation? *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **363**: 2997–3007.

Catling, P.M. 2001. Morphological evidence for the hybrid *Enallagma ebrium* × *hageni* (Zygoptera: Coenagrionidae) from Ontario. *Proc. Entomol. Soc. Ontario* **132**: 99–100.

Cressie, N.A.C. 1993. *Statistics for Spatial Data*. J. Wiley & Sons, New York, NY.

Croucher, P.J.P., Jones, R.M., Searle, J.B. & Oxford, G.S. 2007. Contrasting patterns of hybridization in large house spiders (*Tegenaria atrica* group, Agelenidae). *Evolution* **61**: 1622–1640.

de Queiroz, K. 2007. Species concepts and species delimitation. *Syst. Biol.* **56**: 879–886.

Duchesne, P. & Bernatchez, L. 2002. AFLPOP: a computer program for simulated and real population allocation, based on AFLP data. *Mol. Ecol. Notes* **2**: 380–383.

Duchesne, P., Etienne, C. & Bernatchez, L. 2006. PERM: a computer program to detect structuring factors in social units. *Mol. Ecol. Notes* **6**: 965–967.

Evanno, G., Regnaut, S. & Goudet, J. 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol. Ecol.* **14**: 2611–2620.

Excoffier, L. & Lischer, H.E.L. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* **10**: 564–567.

Falush, D., Stephens, M. & Pritchard, J.K. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Mol. Ecol. Notes* **7**: 574–578.

Fincke, O.M., Fargevieille, A. & Schultz, T.D. 2007. Lack of innate preference for morph and species identity in mate-searching *Enallagma* damselflies. *Behav. Ecol. Sociobiol.* **61**: 1121–1131.

Fitzpatrick, B.M., Fordyce, J.A. & Gavrilets, S. 2008. What, if anything, is sympatric speciation? *J. Evol. Biol.* **21**: 1452–1459.

Fitzpatrick, B.M., Fordyce, J.A. & Gavrilets, S. 2009. Pattern, process and geographic modes of speciation. *J. Evol. Biol.* **22**: 2342–2347.

Gavrilets, S. 2000. Rapid evolution of reproductive barriers driven by sexual conflict. *Nature* **403**: 886–889.

Gavrilets, S. 2003. Perspective: models of speciation: what have we learned in 40 years? *Evolution* **57**: 2197–2215.

Gröning, J. & Hochkirch, A. 2008. Reproductive interference between animal species. *Q. Rev. Biol.* **83**: 257–282.

Gröning, J. & Kocum, A. 2005. Dornschrecken (Orthoptera, Tetrigidae) auf den Ostfriesischen Inseln-Verbreitung, Koexistenz und Okologie. *Entomol. heute* **17**: 47–63.

Gröning, J., Lücke, N., Finger, A. & Hochkirch, A. 2007. Reproductive interference in two ground-hopper species: testing hypotheses of coexistence in the field. *Oikos* **116**: 1449–1460.

Guillot, G., Leblois, R., Coulon, A. & Frantz, A.C. 2009. Statistical methods in spatial genetics. *Mol. Ecol.* **18**: 4734–4756.

Halverson, K., Heard, S.B., Nason, J.D. & Stireman, J.O. 2008. Origins, distribution, and local co-occurrence of polyploid cytotypes in *Solidago altissima* (Asteraceae). *Am. J. Bot.* **95**: 50–58.

Hewitt, G.M. 2004. Genetic consequences of climatic oscillations in the Quaternary. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **359**: 183–195.

Hochkirch, A., Gröning, J. & Bücker, A. 2007. Sympatry with the devil: reproductive interference could hamper species coexistence. *J. Anim. Ecol.* **76**: 633–642.

Howard, D.J. & Harrison, R.G. 1984. Habitat segregation in ground crickets: the role of interspecific competition and habitat selection. *Ecology* **65**: 69–76.

Hubisz, M.J., Falush, D., Stephens, M. & Pritchard, J.K. 2009. Inferring weak population structure with the assistance of sample group information. *Mol. Ecol. Notes* **9**: 1322–1332.

Ibrahim, K.M., Nichols, R.A. & Hewitt, G.M. 1996. Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. *Heredity* **77**: 282.

Jensen, J.L., Bohonak, A.J. & Kelley, S.T. 2005. Isolation by distance, web service. *BMC Genet.* **6**: 13.

Johnson, D.M. & Crowley, P.H. 1980. Habitat and seasonal segregation among coexisting odonate larvae. *Odonatologica* **9**: 297–308.

Kirkpatrick, M. & Ravigné, V. 2002. Speciation by natural and sexual selection: models and experiments. *Am. Nat.* **159**: S22–S35.

Kuno, E. 1992. Competitive-exclusion through reproductive interference. *Res. Popul. Ecol.* **34**: 275–284.

Losos, J.B. & Glor, R.E. 2003. Phylogenetic comparative methods and the geography of speciation. *Trends Ecol. Evol.* **18**: 220–227.

Mallet, J., Meyer, A., Nosil, P. & Feder, J.L. 2009. Space, sympatry and speciation. *J. Evol. Biol.* **22**: 2332–2341.

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

McPeek, M.A. 1989. Differential dispersal tendencies among *Enallagma* damselflies (Odonata) inhabiting different habitats. *Oikos* **56**: 187–195.

McPeek, M.A. 1990a. Behavioral-differences between *Enallagma* species (Odonata) influencing differential vulnerability to predators. *Ecology* **71**: 1714–1726.

McPeek, M.A. 1990b. Determination of species composition in the *Enallagma* damselfly assemblages of permanent lakes. *Ecology* **71**: 83–98.

McPeek, M.A. 1995. Morphological evolution mediated by behavior in the damselflies of 2 communities. *Evolution* **49**: 749–769.

McPeek, M.A. 1998. The consequences of changing the top predator in a food web: a comparative experimental approach. *Ecol. Monogr.* **68**: 1–23.

McPeek, M.A. 1999. Biochemical evolution associated with antipredator adaptation in damselflies. *Evolution* **53**: 1835–1845.

McPeek, M.A. 2000. Predisposed to adapt? Clade-level differences in characters affecting swimming performance in damselflies. *Evolution* **54**: 2072–2080.

McPeek, M.A. & Brown, J.M. 2000. Building a regional species pool: diversification of the *Enallagma* damselflies in eastern North America. *Ecology* **81**: 904–920.

McPeek, M.A., Shen, L., Torrey, J.Z. & Farid, H. 2008. The tempo and mode of three-dimensional morphological evolution in male reproductive structures. *Am. Nat.* **171**: E158–E178.

McPeek, M.A., Shen, L. & Farid, H. 2009. The correlated evolution of three-dimensional reproductive structures between male and female damselflies. *Evolution* **63**: 73–83.
McPeek, M.A., Symes, L.B., Zong, D.M. & McPeek, C.L. 2011. Species recognition and patterns of population variation in the reproductive structures of a damselfly genus. *Evolution* **65**: 419–428.
M'Gonigle, L.K. & FitzJohn, R.G. 2010. Assortative mating and spatial structure in hybrid zones. *Evolution* **64**: 444–455.
Miller, R.S. 1963. Pattern and process in competition. In: *Advances in Ecological Researches* (J.B. Cragg, ed.), pp. 1–74. Academic Press, New York, NY.
Mims, M.C., Hulsey, C.D., Fitzpatrick, B.M. & Streelman, J.T. 2010. Geography disentangles introgression from ancestral polymorphism in Lake Malawi cichlids. *Mol. Ecol.* **19**: 940–951.
Nichols, R.A. & Hewitt, G.M. 1994. The genetic consequences of long-distance dispersal during colonization. *Heredity* **72**: 312–317.
Panhuis, T.M., Butlin, R., Zuk, M. & Tregenza, T. 2001. Sexual selection and speciation. *Trends Ecol. Evol.* **16**: 364–371.
Paulson, D.R. 1974. Reproductive isolation in damselflies. *Syst. Zool.* **23**: 40–49.
Pritchard, J.K., Stephens, M. & Donnelly, P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.
Ray, N. & Excoffier, L. 2010. Infering processes during introduction and range expansion: a first step towards inferring levels of long-distance dispersal during past expansions. *Mol. Ecol. Resour.* **10**: 902–914.
Rice, W.R. & Hostert, E.E. 1993. Laboratory experiments on speciation: what have we learned in 40 years? *Evolution* **47**: 1637–1653.
Robertson, H.M. & Paterson, H.E.H. 1982. Mate recognition and mechanical isolation in *Enallagma*-damselflies (Odonata, Coenagrionidae). *Evolution* **36**: 243–250.
Rosenberg, N.A. 2004. DISTRUCT: a program for the graphical display of population structure. *Mol. Ecol. Notes* **4**: 137–138.
Rundle, H.D. & Nosil, P. 2005. Ecological speciation. *Ecol. Lett.* **8**: 336–352.
SAS Institute. 2008. *SAS/STAT® 9.2 User's Guide*. SAS Institute Inc., Cary, NC.
Schluter, D. 2000. *The Ecology of Adaptative Radiation: Oxford Series in Ecology and Evolution*. Oxford University Press, New York, NY.
Schluter, D. 2001. Ecology and the origin of species. *Trends Ecol. Evol.* **16**: 372–380.
Siepielski, A.M., Hung, K.L., Bein, E.E.B. & McPeek, M.A. 2010. Experimental evidence for neutral community dynamics governing an insect assemblage. *Ecology* **91**: 847–857.
Stoks, R., McPeek, M.A. & Mitchell, J.L. 2003. Evolution of prey behavior in response to changes in predation regime: Damselflies in fish and dragonfly lakes. *Evolution* **57**: 574–585.
Stoks, R., Nystrom, J.L., May, M.L. & McPeek, M.A. 2005. Parallel evolution in ecological and reproductive traits to produce cryptic damselfly species across the holarctic. *Evolution* **59**: 1976–1988.
Turgeon, J. & McPeek, M.A. 2002. Phylogeographic analysis of a recent radiation of *Enallagma* damselflies (Odonata: Coenagrionidae). *Mol. Ecol.* **11**: 1989–2001.
Turgeon, J., Stoks, R., Thum, R.A., Brown, J.M. & McPeek, M.A. 2005. Simultaneous Quaternary radiations of three damselfly clades across the Holarctic. *Am. Nat.* **165**: E78–E107.
Vekemans, X., Beauwens, T., Lemaire, M. & Roldán-Ruiz, I. 2002. Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. *Mol. Ecol.* **11**: 139–151.
Via, S. 2001. Sympatric speciation in animals: the ugly duckling grows up. *Trends Ecol. Evol.* **16**: 381–390.
Walker, E.M. 1953. *The Odonata of Canada and Alaska*. University of Toronto Press, Toronto.
Webster, R. & Oliver, M.A. 1990. *Statistical Methods in Soil and Land Resource Survey*. Oxford University Press, Oxford.
Westfall, M.J. & May, M.L. 1996. *Damselflies of North America*. Scientific publishers, Gainesville, FL.
Westfall, M.J. & May, M.L. 2006. *Damselflies of North America*, revised edition. Scientific publishers, Gainesville, FL.

Supporting information

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

Appendix S1 Spatial analysis of the relative distribution of *Enallagma hageni* and *E. ebrium* in eastern North America.

Appendix S2 Pairwise regional F_{ST} between samples of *Enallagma hageni* and *E. ebrium* sampled in northeastern North America.

Appendix S3 Results of clustering analyses with the entire dataset (both morphospecies from all three regions) with STRUCTURE 2.3.1.

Appendix S4 Comparisons of *Enallagma hageni* and *E. ebrium* morphospecies in sympatry vs. allopatry.

Table S1 Description of sites sampled for *Enallagma hageni* (H) and *E. ebrium* (E) in three geographical regions indicating morphospecies frequency, as well as sample sizes for genetic analyses using AFLP and mtDNA.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

Data deposited at Dryad: doi: 10.5061/dryad.s26qt314

Received 9 September 2011; revised 17 October 2011; accepted 19 October 2011