

AUDREY BOURRET

**DIVERGENCE RÉGIONALE ET RÉPARTITION  
SPATIALE EN MOSAÏQUE CHEZ LES DEMOISELLES  
*ENALLAGMA HAGENI* ET *E. EBRIUM***

Mémoire présenté  
à la Faculté des études supérieures de l'Université Laval  
dans le cadre du programme de maîtrise en Biologie  
pour l'obtention du grade de Maître ès sciences (M. Sc.)

DÉPARTEMENT DE BIOLOGIE  
FACULTÉ DE SCIENCES ET GÉNIE  
UNIVERSITÉ LAVAL  
QUÉBEC

2011

## Résumé

Les demoiselles *Enallagma* ont radié en Amérique du Nord durant le Pléistocène et les espèces sont principalement différenciées par des structures sexuelles secondaires. Bien qu'elles soient morphologiquement très différentes, *Enallagma hageni* et *E. ebrium* sont génétiquement très similaires. Leur répartition spatiale est généralement chevauchante, mais il semble que ces espèces co-occurrent rarement. Dans cette étude, nous investiguons l'histoire évolutive de ces deux jeunes espèces morphologiques par l'analyse conjointe de données génétiques et spatiales. Les analyses de génétique des populations basées sur le polymorphisme AFLP indiquent que ces espèces morphologiques divergent régionalement. Ces deux espèces morphologiques sont rarement sympatriques à l'échelle locale et leur distribution forme une mosaïque d'îlots où une espèce domine. L'interférence reproductive couplée à la recolonisation suite à la dernière glaciation pourrait expliquer ce patron. Cette répartition spatiale limite probablement les opportunités contemporaines de flux génique, et accélère par le fait même la différenciation régionale.

## Abstract

North American *Enallagma* damselflies radiated during the Pleistocene and species differ mainly by secondary sexual structures. Although morphologically very different, *Enallagma hageni* and *E. ebrium* are genetically very similar. Their ranges broadly overlap but these ecologically equivalent species rarely co-occur. Here, we investigated the evolutionary history of these two very young morphospecies by coupling population genetics and spatial analyses. Based on AFLP polymorphisms, population genetics analyses indicate that these morphospecies are locally diverging. They are rarely sympatric at local sites and their distribution form a mosaic of patches where one morphospecies is dominant. Reproductive interference coupled with postglacial recolonization likely explains this pattern. By limiting contemporary opportunities for gene flow, it probably also contributes to accelerating regional differentiation.

## Avant-Propos

Le présent mémoire est composé d'une introduction et d'une conclusion en français, et d'un chapitre en anglais. Ce chapitre prend la forme d'un article scientifique et est rédigé en anglais en vue d'être soumis dans une revue spécialisée en biologie évolutive. Les auteurs de cet article sont Julie Turgeon, ma directrice de maîtrise, Mark A. McPeck, professeur à Dartmouth College et collaborateur à ce projet de recherche, ainsi que moi-même. Toutes les étapes sous-jacentes à la réalisation de ce projet, de l'échantillonnage à la rédaction de l'article, en passant par les manipulations en laboratoire et les diverses analyses, ont été principalement à ma charge.

C'est une grande opportunité que m'a offerte Julie Turgeon en me proposant de travailler sur le système des demoiselles *Enallagma*. Elle a cru que je pouvais relever ce défi, et avec son aide, je crois bien que j'y suis arrivée. Je tiens à la remercier sincèrement pour la confiance qu'elle m'a accordée tout au long du projet. C'est une chercheuse toujours disponible pour ses étudiants, rigoureuse et passionnée. Outre l'encadrement en lien direct avec ce projet de maîtrise, elle m'a appris à développer un fort esprit critique, une qualité essentielle pour une future carrière scientifique. Encore une fois, merci!

Le petit laboratoire Turgeon regorge de personnes exceptionnelles que je tiens à remercier pour l'aide et le soutien qu'ils m'ont apportés tout au cours de ma maîtrise. Pour leur aide au niveau technique, je tiens à remercier Catherine Potvin qui m'a donné un grand coup de mains avec les AFLP, Pierre Duchesne pour ses statistiques et ses permutations, ainsi que Caroline Lehoux pour son aide avec les bases de données. J'ai eu aussi la chance d'avoir un assistant de terrain à l'été 2009, Pierre Favriou, avec qui les heures sous la pluie à chercher désespérément quelques demoiselles ont été moins longues. Je tiens aussi à remercier tous les membres passés et présents du laboratoire Turgeon, Marie-Claude Gagnon, Geneviève Parent, Gabriel Colbert, Charlotte Roby et David Mérette, qui sont devenus plus que des collègues de travail.

De nombreuses ressources professionnelles ont contribué à ce projet. Merci aux nombreux musées et organismes gouvernementaux qui nous ont envoyé des données pour les analyses spatiales. Merci aussi à P. Brunelle, P. Catling, R. Curley, D. Doucet, M. R. L. Forbes, G. Lemelin, J. M. Perron pour les discussions sur la probable répartition de ces deux espèces. Finalement un merci tout spécial à M. A. McPeck pour avoir collaboré avec nous sur ce projet touchant le système *Enallagma* qu'il affectionne particulièrement, pour nous avoir fourni informations et échantillons et pour les discussions sur ce système complexe, mais si intéressant.

Finalement, je voudrais remercier ma famille et mes amis pour le soutien qu'ils m'ont apporté tout au long de ma maîtrise.

Ce projet de recherche n'aurait pu exister sans le soutien financier du Fonds québécois de recherche sur la nature et les nouvelles technologies (FQRNT) pour ma bourse d'études et du Conseil de recherches en sciences naturelles et en génie (CRSNG) pour les fonds de recherche de J. Turgeon.

*Biology is the science. Evolution is the  
concept that makes biology unique*  
Jared Diamond

# Table des matières

Résumé .....	ii
Abstract .....	iii
Avant-Propos .....	iv
Table des matières .....	vii
Liste des tableaux.....	ix
Liste des figures.....	x
Introduction Générale .....	1
L'étude de la spéciation .....	1
Définir ce qu'est une espèce .....	1
Le contexte spatial de la spéciation .....	3
Les mécanismes menant à la spéciation .....	4
<i>Enallagma</i> spp. ....	7
Cycle vital.....	7
Répartition écologique.....	8
Histoire évolutive du genre <i>Enallagma</i> .....	9
Différenciation entre <i>Enallagma hageni</i> et <i>E. ebrium</i> .....	10
Chapitre 1: Regional divergence and mosaic spatial distribution of two closely related damselfly species ( <i>Enallagma hageni</i> and <i>E. ebrium</i> ).....	13
Abstract.....	14
Résumé.....	15
Introduction.....	16
Materials and methods .....	19
Patterns of genetic variation .....	19
Species distribution.....	22

Results.....	24
Genetic analysis.....	24
Spatial distribution.....	26
Discussion.....	27
Ongoing regional differentiation of <i>E. hageni</i> and <i>E. ebrium</i> .....	27
<i>E. hageni</i> and <i>E. ebrium</i> mosaic distribution.....	29
Evolutionary history of <i>E. hageni</i> and <i>E. ebrium</i> .....	31
Conclusion.....	33
Appendix I: Sites description.....	41
Appendix II: Spatial analysis.....	42
Section 1. Museum data source.....	42
Section 2. Data treatment and kriging specification.....	43
Section 3. Data validation.....	44
Appendix III : Regional Pairwise $F_{ST}$ .....	46
Appendix IV : MtDNA haplotype network.....	48
Conclusion.....	49
Divergence locale et répartition spatiale en mosaïque.....	49
L'importance du contexte spatial.....	50
Perspectives.....	52
Bibliographie.....	54



## Liste des tableaux

<b>Table 1</b>	Partitioning of genetic variance for grouping models considering morphospecies, region or regional morphospecies as the main grouping factor using a) all sites and b) only sympatric sites. Corrected Akaike Information Criteria (AIC <sub>c</sub> ) were calculated following Halverson <i>et al.</i> (2008).....	34
<b>Table 2</b>	Species distribution kriging interpolation results (see also Figure 3) for regional and global analyses. ....	35
<b>Table A1</b>	Description of sites sampled for <i>Enallagma hageni</i> (H) and <i>E. ebrium</i> (E) in three geographical regions indicating species frequency, as well as sample sizes for genetic analyses using AFLP and mtDNA.....	41
<b>Table A2</b>	F <sub>ST</sub> between pairs of sampling sites with <i>E. hageni</i> (H) and <i>E. ebrium</i> (E) for three regions: a) PE, b) NE and c) QO. F <sub>ST</sub> value associated P-values < 0.05 are in italic; bold F <sub>ST</sub> value indicate Bonneferroni corrected comparisons for $\alpha=0.05$ . ....	46

## Liste des figures

- Figure 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 used in genetic analyses (see also Figure 2), b) a representative distribution map obtained with randomized datasets and c) kriging results for empirical dataset showing areas where *E. hageni* or *E. ebrium* represent more than 75% of the individuals. Hatched areas could not be used for interpolation because sites were too sparse. ....36
- Figure 2** Sampling sites used in genetics analysis and morphospecies relative distribution in the three study region: a) QO, NE and c) PE (see Table A1 for more information on sample location and size). .....37
- Figure 3** Genetic differentiation in relation with distance. a) IBD 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. ....38
- Figure 4** Regional clustering of AFLP 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 species were present are underlined. ....39
- Figure 5** Frequency of *E. hageni* among 84 sites sampled in PE, NE and QO. ....40

**Figure A1** MtDNA haplotype network relating sequences of *E. hageni* (H) and *E. 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.....48

# Introduction Générale

## L'étude de la spéciation

La spéciation est un processus central en biologie évolutive dont l'étude suscite un intérêt toujours aussi important (Sobel et coll. 2010). C'est un phénomène complexe et difficile à bien cerner puisqu'il se déroule à une échelle temporelle souvent trop longue pour l'observer directement. De plus, ce phénomène est en interaction avec plusieurs autres facteurs évolutifs (p. ex. la structure spatiale) (Kirkpatrick et Ravigné 2002). La compréhension qu'en ont les biologistes évolue continuellement, ce qui se traduit par la manière dont ont été abordées les questions de spéciation au fil du temps. Par exemple, la récente réévaluation critique de la définition d'espèce semble avoir des répercussions directes sur son étude (Hart 2010). La prévalence des différents mécanismes menant à la spéciation, tels les agents sélectifs écologiques et non-écologiques, et l'importance du contexte spatial observé tout au cours de la divergence des populations sont des questions actuelles qui font couler beaucoup d'encre.

### *Définir ce qu'est une espèce*

Bien qu'il soit généralement accepté que l'espèce est une entité réelle, sa définition reste un sujet controversé. Depuis la première catégorisation des différents concepts d'espèce par Mayr (1942) et sa proposition du concept d'espèce biologique, plus d'une vingtaine de concepts alternatifs ont été proposés (Mayden 1997; Coyne et Orr 2004). On retrouve parmi ceux-ci les concepts d'espèce morphologique, écologique, phénétique et différentes versions du concept d'espèce phylogénétique (revu dans Coyne et Orr 2004). Ces multiples définitions d'espèce s'appuient souvent sur des propriétés biologiques différentes. Par exemple, le concept d'espèce biologique repose sur l'isolement reproducteur, l'écologique sur l'occupation de différentes niches et le phylogénétique sur la monophylie ou la diagnose phylogénétique. Les nombreux sous-groupes de biologistes utilisent le concept d'espèce le plus approprié pour leur domaine d'étude, et ils se concentrent les propriétés

biologiques qu'ils croient les plus importantes. Force est d'admettre que tous ces différents concepts d'espèce, bien qu'ils ne fassent pas l'unanimité, s'appuient sur d'importantes propriétés biologiques (de Queiroz 2005).

Les nombreux concepts d'espèces proposés dans la littérature sont incompatibles entre eux puisqu'ils mènent à la reconnaissance d'un nombre différent d'espèces selon celui préconisé (de Queiroz 2005). Par exemple, un nombre plus grand d'espèces peut être reconnu avec le concept d'espèce phylogénétique dans sa version diagnostique qu'avec celui biologique (Martin 1996; Zink 1996). Selon de Queiroz (2005), le problème commun de ces nombreux concepts d'espèce est qu'ils traitent les propriétés biologiques acquises par les populations divergentes comme des éléments essentiels à leur définition alors qu'elles ne le sont pas. Par contre, tous ces concepts semblent se rejoindre sur une même base qui constitue le cœur du concept d'espèce unifié de de Queiroz (2007), soit que les espèces sont des lignées de métapopulations qui évoluent indépendamment. Les autres propriétés biologiques sur lesquelles s'appuient les différents concepts d'espèce sont des propriétés importantes, mais elles ne sont pas nécessaires dans la définition même; elles deviennent des propriétés contingentes (de Queiroz 2005). Ces propriétés contingentes sont, dans le concept d'espèce unifié, des lignes d'évidences pour la présence ou non d'espèces distinctes. Durant le processus de spéciation, les métapopulations divergentes vont acquérir ou non ces différentes propriétés biologiques. Au cours du temps, ces métapopulations vont devenir distinctes en termes de traits quantitatifs et dans leur écologie. Elles vont atteindre l'isolement reproducteur. Elles vont aussi passer par des phases polyphylétiques, paraphylétiques et monophylétiques en regard de leurs gènes. Tous ces changements n'arrivent pas au même moment au cours de la spéciation, pas plus qu'ils arrivent à chaque fois dans le même ordre (de Queiroz 1998). Choisir l'une ou l'autre de ces propriétés biologiques comme un élément central dans une définition d'espèce revient à placer la limite de l'espèce plus ou moins tôt dans le processus de la spéciation.

Le débat historique sur la définition de l'espèce semble prendre une nouvelle direction qui influence directement l'étude de la spéciation. Avec son concept d'espèce unifié, de Queiroz (2007) a fait passer le problème de l'espèce du niveau conceptuel à celui

méthodologique, c'est-à-dire non plus à savoir quelle est la meilleure propriété biologique pour définir l'espèce, mais plutôt comment en déterminer les limites. Cette tendance se reflète dans les études actuelles de spéciation où la définition même de l'espèce semble peu importante par rapport aux méthodes pour les délimiter (Hart 2010). En ce sens, on semble d'éloigner graduellement du critère unique pour délimiter les espèces (Hey 2006).

### *Le contexte spatial de la spéciation*

Traditionnellement, la spéciation a été classée et définie en fonction du contexte spatial des populations divergentes (suivant les idées de Mayr 1942). Cette classification permet la distinction entre les différents niveaux de flux géniques initiaux possibles, c'est-à-dire d'une totale isolation en allopatrie à la panmixie en sympatrie, en passant par tous les niveaux intermédiaires en parapatrie. Depuis une vingtaine d'années, la publication d'études portant sur la spéciation sympatrique, un sujet soulevant parfois de vives oppositions, a augmenté exponentiellement (Bolnick et Fitzpatrick 2007). La spéciation sympatrique a été largement explorée au niveau théorique et elle serait possible dans certains contextes. Par exemple, une étude de simulations de Dieckmann et Doebeli (1999) a montré que ce mode de spéciation sympatrique est probable lors de situation de prédation ou de mutualisme par la sélection contre les phénotypes intermédiaires ou dans des cas de compétition pour les ressources. Plusieurs événements probables de spéciation sympatrique sont maintenant largement acceptés, p. ex. des cas de déplacement d'espèces hôtes comme chez *Rhagoletis pomonella* ou par spéciation instantanée par la polyploïdisation (revu dans Bolnick et Fitzpatrick 2007). La variation génétique nécessaire à la spéciation sympatrique peut être acquise lors de phases allopatriques (Bolnick et Fitzpatrick 2007). Le nombre restreint d'exemples empiriques de spéciation sympatrique suggère que ce mode de spéciation reste peu fréquent (Bolnick et Fitzpatrick 2007), et il est généralement accepté que la majorité des événements de spéciation se déroule dans un contexte parapatrique, c'est-à-dire à des niveaux intermédiaires de flux géniques (Gavrilets 2003; Mallet et coll. 2009).

Plusieurs auteurs remettent maintenant en doute la pertinence d'une classification géographique de la spéciation (Via 2001; Kirkpatrick et Ravigné 2002; Butlin et coll. 2008; Fitzpatrick et coll. 2008; 2009; mais voir Mallet et coll. 2009 pour des arguments en faveur d'une classification géographique). La division en grande classes géographiques briserait le continuum des possibilités de flux géniques initiaux en catégories discrètes et artificielles (Rice et Hostert 1993; Schluter 2001; Gavrilets 2003). De plus, l'inférence du contexte spatial observé au temps de la spéciation nécessaire pour ce type de classification n'est pas une tâche facile. La répartition contemporaine des espèces ne représente pas nécessairement celle au temps de la spéciation, laquelle peut changer considérablement au fil du temps (Losos et Glor 2003). Par exemple, les oscillations climatiques comme ceux de la dernière glaciation, en plus de favoriser la spéciation, ont modifié grandement la répartition spatiale de nombreuses espèces (Hewitt 2004). Selon de nombreux auteurs, une classification basée sur les mécanismes menant à la spéciation serait plus appropriée (Via 2001; Butlin et coll. 2008; Fitzpatrick et coll. 2008; 2009).

Malgré le désaccord sur la manière la plus adéquate de classer les différents modes de spéciation, la répartition contemporaine des espèces peut jouer un rôle majeur dans son étude. Le contexte spatial observé à chaque étape de la spéciation détermine l'importance du flux génique entre les populations divergentes, qui en retour va affecter la probabilité de spéciation (Butlin et coll. 2008; Mallet et coll. 2009). En ce sens, le contexte géographique de la spéciation peut aussi être vu comme une forme d'homogamie (*assortative mating*, Kirkpatrick et Ravigné 2002). La répartition spatiale contemporaine ne doit donc pas être totalement exclue des études de spéciations.

### *Les mécanismes menant à la spéciation*

Les études actuelles sur la spéciation cherchent à identifier et comprendre les mécanismes menant à la spéciation, une tendance qui se reflète dans la classification des modes de spéciation. Parmi les récentes classifications de ces mécanismes proposées dans la littérature, ceux contrastant les processus écologiques avec ceux non-écologiques sont

particulièrement communs (p. ex. Schluter 2001; Via 2001; Rundle et Nosil 2005; Rundle et Price 2009; Schluter 2009).

La spéciation écologique est le processus par lequel les barrières au flux génique évoluent en raison d'un agent sélectif écologique (Schluter 2000; Schluter 2001; Rundle et Nosil 2005). Dans la littérature, le rôle de l'environnement dans le processus de la spéciation est nettement plus étudié que celui des mécanismes non-écologiques (Hendry 2009). Dans un cas classique de spéciation écologique, l'isolement reproducteur arrive progressivement entre des populations qui accumulent des adaptations particulières liées à un contraste environnemental. Les différences dans l'environnement peuvent être au niveau des éléments biotiques et abiotiques de l'habitat (p. ex. les ressources, la température) et dans les interactions avec les autres espèces (p. ex. la compétition pour les ressources, la prédation, le mutualisme) (revu dans Schluter 2001). L'évolution des barrières au flux génique est souvent un produit dérivé de la sélection naturelle sur des traits morphologiques, physiologiques ou comportementaux liés à l'environnement, mais peut aussi être sous l'influence de sélection directe. En effet, la spéciation écologique pourrait aussi être impliquée dans des cas où la sélection sexuelle est induite par une différence dans l'environnement, p. ex. des cas de variations spatiales des traits sexuels secondaires ou de différences entre les modes de transmission des signaux sexuels selon l'environnement (revu dans Schluter 2001). C'est le cas des isopodes d'eau douce *Asellus aquaticus* chez qui une différence dans l'environnement induit une réponse différente à la garde précopulatoire (Karlsson et coll. 2010). L'étude de la spéciation écologique demande une rigueur particulière puisqu'un contraste écologique observé entre deux espèces n'est pas garant d'un tel mécanisme de spéciation; cette différence doit absolument être la cause de la divergence adaptative pour que l'on parle de spéciation écologique (Rasanen et Hendry 2008). Plusieurs études affirmant confirmer des cas de spéciation écologique n'auraient pas réussi à démontrer une signature simple et prononcée d'un tel phénomène pour ces raisons (Hendry 2009).

La spéciation parallèle est un cas particulier de spéciation écologique où, sous l'influence de conditions environnementales similaires, les traits permettant l'isolement reproducteur



évoluent en parallèle dans différentes populations (Schluter et Nagel 1995; Rundle et coll. 2000). Les deux écotypes observés chez les corégones (*Coregonus clupeaformis*) seraient apparus en parallèle dans différents lacs suivant la recolonisation postglaciaire (Bernatchez et coll. 2010). Les formes naines et normales diffèrent morphologiquement au niveau de traits liés à la spécialisation trophique, respectivement planctonique et épibenthique. La spéciation parallèle n'est pas nécessairement facilement identifiable en nature puisqu'elle requiert à la fois des analyses écologiques, morphologiques et moléculaires (Johannesson 2001). De plus, le signal de spéciation parallèle historique peut facilement être perdu en raison de flux génique contemporain entre les populations d'une même forme (Johannesson 2001).

Les processus non-écologiques pouvant mener à la spéciation semblent être moins représentés dans la littérature. On retrouve deux grandes classes de processus non-écologiques, soit ceux totalement stochastiques et ceux dont l'agent sélectif est non-écologique (Rundle et Nosil 2005). Les mécanismes stochastiques menant à la spéciation les plus souvent invoqués sont ceux par dérive génétique et ceux par polyploïdisation. Du côté des agents sélectifs non-écologiques, la sélection sexuelle a un grand potentiel de créer des divergences entre les populations et de mener à la spéciation sans contrastes environnementaux (revu dans Panhuis et coll. 2001). Par exemple, la divergence entre les intérêts reproducteurs des mâles et des femelles peut causer de la sélection sexuelle et éventuellement mener à la spéciation (Parker et Partridge 1998; Gavrilets 2000). L'évolution rapide de barrières reproductives en raison de conflits entre les sexes pourrait expliquer le grand taux de spéciation observé lors de la colonisation de nouveaux habitats (Gavrilets 2000). Le choix des femelles, comme dans le cas des cichlidés du lac Malawi, peut aussi créer la diversité nécessaire à la spéciation (Albertson et coll. 1999). Il est par contre difficile de démontrer sans l'ombre d'un doute qu'aucun aspect écologique n'a contribué à ces mécanismes de spéciation dits non-écologiques (Sobel et coll. 2010).

## *Enallagma spp.*

Les demoiselles du genre *Enallagma* (Odonate : Zygoptère) forment le groupe d'Odonates le plus diversifié au monde avec environ 70 espèces définies selon la morphologie (Westfall et May 2006). De ce nombre, trente-huit se retrouvent en Amérique du Nord, majoritairement dans l'est du continent (Westfall et May 2006).

### *Cycle vital*

Le cycle vital du genre *Enallagma* s'étend sur une année et est majoritairement aquatique. Il débute avec la ponte des œufs dans les tissus de plantes aquatiques. Ces œufs se développent généralement entre deux et quatre semaines jusqu'à l'apparition du premier stade larvaire (Westfall et May 2006). Environ douze stades larvaires se succèdent chez les *Enallagma spp.*, mais ce nombre varie en fonction de l'espèce et des conditions abiotiques, p. ex. la température et la photopériode (Corbet 1999). Pour passer l'hiver, les larves entrent en diapause à un stade larvaire plus ou moins avancé. Au printemps, les larves à leur dernier stade émergent et grimpent sur des tiges de végétation aquatique pour devenir des demoiselles sexuellement immatures (stade ténéral). À l'émergence, le rapport des sexes est généralement de 1 : 1 (Fincke 1982). Les dates de ponte et d'émergence varient considérablement selon l'espèce et la latitude (Corbet 1999).

La courte phase ailée des agrions *Enallagma* permet leur reproduction. Après avoir passé les quelques jours suivant leur émergence à se nourrir dans les boisés avoisinants, les mâles devenus sexuellement matures reviennent au lieu de reproduction à la recherche constante de femelles. Les femelles s'y rendent uniquement pour la reproduction, biaisant le rapport des sexes à cet endroit (3 : 1 mâle : femelle; Fincke 1982). Pour qu'il y ait copulation, un mâle doit attraper une femelle par le thorax à l'aide de ses pattes; il forme alors un tandem avec celle-ci en utilisant son appareil anal pour s'accrocher aux plaques thoraciques mésostigiales de cette dernière. Alors que les mâles ne reconnaissent pas visuellement les femelles de leur espèce et qu'ils attrapent tous ce qui pourrait y ressembler (Fincke et coll.

2007), les femelles usent de récepteurs tactiles, stimulés par le contact des cercoïdes des mâles, pour reconnaître leur conspécifiques (Robertson et Paterson 1982). Si une femelle reconnaît un mâle conspécifique et l'accepte, elle formera alors un cœur copulatoire avec ce dernier en recourbant son abdomen, permettant ainsi l'insémination. Les appendices anaux des mâles et les plaques thoraciques mésostigiales des femelles sont en étroite association, réduisant ainsi la reproduction interspécifique (Paulson 1974; Robertson et Paterson 1982). D'ailleurs, plusieurs espèces apparentées ne sont reconnues qu'à la morphologie de ces structures (Westfall et May 2006).

### *Répartition écologique*

En Amérique du Nord, les différentes espèces du genre *Enallagma* sont réparties dans les milieux lenticules selon le régime de prédation qui y prévaut. On note en effet une nette distinction entre les espèces qui co-existent avec les poissons ou les larves de libellules comme prédateur principal. Alors que la majorité des espèces sont retrouvées dans des lacs dits « à poissons », quatre espèces (*E. annexum*, *E. aspersum*, *E. boreale* et *E. doubledayi*) possèdent des adaptations physiologiques, morphologiques et comportementales les rendant moins vulnérables à la prédation par les larves de libellules (McPeck 1990a; 1995; 1999; 2000; Stoks et coll. 2003).

Outre cette différence écologique majeure, les espèces « à poissons » sont écologiquement très similaires (McPeck 1990a; McPeck et coll. 1996; McPeck 1999; Stoks et coll. 2003). Ainsi, jusqu'à douze espèces du genre *Enallagma* peuvent être retrouvées dans un même lac « à poissons » ainsi que cinq à sept espèces au stade larvaire dans un même micro-habitat. De plus, l'abondance relative des espèces n'est corrélée avec aucun gradient environnemental majeur, et des évidences expérimentales indiquent que la mortalité *per capita* et le taux de croissance sont influencés uniquement par l'abondance totale des espèces et non par leur abondance relative comme on aurait pu s'y attendre dans le cas de coexistence en raison de niches différentes (Siepielski et coll. 2010). Ainsi, plusieurs espèces d'*Enallagma* sont considérées comme des équivalents écologiques (Siepielski et coll. 2010).

## *Histoire évolutive du genre Enallagma*

Le genre *Enallagma* forme un groupe monophylétique d'origine ancienne au sein duquel se sont succédé des épisodes de divergence à vitesse variable (McPeck et Brown 2000; Turgeon et coll. 2005). En Amérique du Nord, deux clades primaires seraient apparus il y a environ 15 millions d'années (clade Sud et clade Nord, Brown et coll. 2000; McPeck et Brown 2000; Turgeon et coll. 2005). Ces deux clades sont caractérisés par des centres de diversité et des rythmes de spéciation très distincts. À l'échelle régionale, la richesse spécifique en espèces du clade Sud est nettement plus grande dans le sud-est des États-Unis alors que la diversité en espèces du clade Nord est plus élevée dans le nord-est américain (McPeck et Brown 2000). De plus, les divergences s'étalent régulièrement dans le temps au sein du clade Sud alors que le clade Nord est marqué par deux radiations récentes (McPeck et Brown 2000).

Les deux radiations du clade Nord se seraient produites pendant le Pleistocène (Turgeon et McPeck 2002). Les radiations des clades *hageni* et *carunculatum*, nommés selon l'espèce possédant la plus vaste répartition au sein du clade, comptent respectivement dix et sept espèces (Brown et coll. 2000; Turgeon et coll. 2005). Vu leur origine récente, les espèces issues de ces deux radiations sont le plus souvent marquées par un polymorphisme génétique ancestral sous forme de partage d'haplotypes mitochondriaux (Turgeon et coll. 2005). Néanmoins, la plupart de ces espèces sont génétiquement distinctes lorsqu'elles sont considérées un grand nombre de locus nucléaires (Turgeon et coll. 2005). En effet, une hypothèse phylogénétique basée sur les AFLP (*Amplified Fragment Length Polymorphism*) révèle que toutes les espèces du clade *carunculatum* sont différenciées génétiquement, formant chacune un groupe monophylétique correspondant à une espèce morphologique reconnue (Turgeon et coll. 2005).

Les dix espèces du clade *hageni* présentent quant à elles des niveaux de différenciation différents (Turgeon et coll. 2005). Cinq espèces suivent un patron de différenciation similaire à celui observé au sein du clade *carunculatum*; elles sont nettement distinctes sur la base de l'ADN nucléaire. Ces espèces ont des répartitions géographiques restreintes sur

la côte est des États-Unis (*E. laterale*, *E. recurvatum*, *E. minusculum*, *E. davisii*) ou dans la région des Grands Lacs (*E. clausum*). Les cinq autres espèces, chacune à vaste distribution, forment deux clades. Le premier clade comprend les espèces *E. annexum*, *E. boreale* et *E. vernalis* et le second les espèces *E. hageni* et *E. ebrium* (Turgeon et coll. 2005). Les espèces *E. annexum* et *E. boreale* sont des espèces « à libellules » et elles sont souvent sympatriques, alors que *E. hageni* et *E. ebrium* sont des espèces « à poissons » qui sont le plus souvent retrouvées en allopatrie (McPeck 1989; 1990b).

Autre particularité intéressante observée chez ce clade, *E. hageni* est composé de deux groupes génétiques distincts séparés par sept mutations à un locus mitochondrial (COI, COII et une leucine-ARNt interposée; race Continentale et Atlantique, Turgeon et McPeck 2002). Ces deux groupes ne se chevauchent pas géographiquement, les races continentale et atlantique étant respectivement présentes à l'ouest et à l'est de la limite des bassins versants des rivières Connecticut et Hudson. La race atlantique a une répartition très limitée (New Hampshire, Vermont, Maine) et une diversité haplotypique faible, alors que la race continentale a une vaste répartition et est très diversifiée. Cette différenciation entre les deux races au niveau mitochondrial ne semble pas se refléter au niveau nucléaire en raison du jeu de données limitées (Turgeon et coll. 2005).

### **Différenciation entre *Enallagma hageni* et *E. ebrium***

La similarité génétique entre *E. hageni* et *E. ebrium* sera au coeur de ce projet de recherche. Il est tout d'abord utile de rappeler ici que ces deux espèces sont très différentes au niveau morphologique. Les mâles des espèces *E. hageni* et *E. ebrium* ont des appendices anaux qui sont sans doute les plus distincts parmi ceux des espèces du clade Nord et des différences au niveau des plaques mésostigiales des femelles peuvent être notées sous le binoculaire. Par contre, cette différenciation observée au niveau morphologique ne semble pas se refléter au niveau génétique. Au niveau mitochondrial, comme toutes les autres espèces issues de la même radiation, *E. hageni* et *E. ebrium* partagent du polymorphisme ancestral, à l'exception des *E. hageni* de la race atlantique qui possèdent un sous-ensemble d'haplotypes distincts. Par contre, contrairement à la majorité des espèces du clade Nord,

ces deux espèces morphologiques ne semblent pas différenciées au niveau nucléaire et elles forment presque des regroupements géographiques (Turgeon et coll. 2005). L'objectif général de ce projet de recherche est donc de comprendre pourquoi ces espèces morphologiquement distinctes sont si similaires génétiquement.

Certains aspects des interactions reproductives entre *E. hageni* et *E. ebrium* sont intrigants sachant la grande similarité génétique observée entre eux. Tout d'abord, la reproduction interspécifique n'a pas été observée lors de tests de croisements (M.A. McPeck et J. Turgeon, données non publiées). En effet, les femelles *E. hageni* et *E. ebrium* refusent les mâles hétérospecifics. Une étude menée par Fincke et coll. (2007) indique tout de même que les mâles *E. ebrium* attrapent les femelles *E. hageni* autant que celles de leur espèce en vue de la formation d'un tandem précopulatoire. Cependant, seulement 15 % des tentatives de tandem entre espèces fonctionnent, suggérant fortement la présence d'une barrière reproductrice prézygotique entre ces espèces. Par contre, Catling (2001) rapporte l'observation en Ontario d'un individu possiblement d'origine hybride puisqu'il possède des cercoïdes morphologiquement intermédiaires entre *E. hageni* et *E. ebrium*. L'hybridation est une hypothèse possible en raison d'évidences d'hybridation récente impliquant *E. hageni*. En effet, sur la côte est américaine, des mâles *E. hageni* possèdent des haplotypes mitochondriaux des espèces *E. aspersum* et *E. geminatum* (Turgeon et coll. 2005). Ces appendices anaux anormaux pourraient toutefois être simplement le résultat de problèmes de développement.

Autre fait particulier, bien que ces deux espèces morphologiques aient des répartitions géographiques globales presque entièrement chevauchantes, s'étendant d'est en ouest de l'Amérique du Nord et du centre des États-Unis jusqu'à la limite de la zone bioclimatique subarctique (Pilon et Lagacé 1998; Westfall et May 2006), elles semblent former une mosaïque d'allopatrie à l'échelle régionale. En effet, dans une région donnée, il semble beaucoup plus fréquent de trouver une espèce qui domine, et au niveau de chaque site, la grande majorité des individus sont généralement d'une seule espèce. Les sites où les deux espèces sont en sympatrie existent, mais sont très rares. Ces distributions relatives inusitées ont été notées dans la littérature (Walker 1953) et semblent faire consensus auprès des

spécialistes consultés (P.M. Catling, Agriculture Canada, M.R.L. Forbes, Carleton University, M.A. McPeck, Darmouth College et J. Turgeon, communication personnelle). Notons enfin qu'on ne connaît pas de différences écologiques liées à l'habitat pouvant expliquer ce patron de répartition spatiale entre *E. hageni* et *E. ebrium* (Siepielski et coll. 2010).

L'objectif général de ce projet de recherche est donc de discerner les raisons de la similarité génétiques entre *E. hageni* et *E. ebrium*, des espèces pourtant morphologiquement distinctes. Dans un premier temps, une approche classique de génétique des populations nous permettra de mieux comprendre certains aspects de l'histoire évolutive de ces deux espèces morphologiques. Le niveau de différenciation génétique observé entre les morphes sera contrasté à la fois entre les sites sympatriques et allopatriques et à différentes échelles spatiales. Cette comparaison permettra une fine résolution entre les différentes explications de similarité génétique possibles exprimées en termes de lignées évolutives. Par la suite, une approche spatiale sera utilisée pour déterminer la répartition relative de ces deux espèces morphologiques, toujours à différentes échelles spatiales. Ces considérations génétiques et spatiales nous aideront à mieux comprendre l'étroite relation entre *E. hageni* et *E. ebrium*.

**Chapitre 1:**  
**Regional divergence and mosaic spatial distribution of  
two closely related damselfly species (*Enallagma hageni*  
and *E. ebrium*)**

Bourret, A., McPeck, M. A., Turgeon, J.



## Abstract

North American *Enallagma* damselflies radiated during the Pleistocene and species differ mainly by secondary sexual structures involved in mating pair formation. Although morphologically very different, *Enallagma hageni* and *E. ebrium* are genetically very similar. Their ranges broadly overlap but these ecologically equivalent species rarely co-occur. Here, we investigated the evolutionary history of these two very young morphospecies by coupling population genetics and spatial analyses. Based on AFLP polymorphisms, partitioning of genetic variation, isolation by distance and clustering analyses indicate that these morphospecies are locally diverging. These two morphospecies are rarely sympatric at local sites and their distribution form a mosaic of patches where one morphospecies is dominant. Reproductive interference coupled with postglacial recolonization likely explains the origin and maintenance of patches. By limiting contemporary opportunities for gene flow, it probably also contributes to accelerating regional differentiation. The unsuspected effects of reproductive exclusion on the spatial and genetic structure of these closely related morphospecies suggest that this process may similarly affect other species.

## Résumé

Les demoiselles *Enallagma* ont radié en Amérique du Nord durant le Pléistocène et les espèces sont principalement différenciées par des structures sexuelles secondaires impliquées dans la formation des paires copulatoires. Bien qu'elles soient morphologiquement très différentes, *Enallagma hageni* et *E. ebrium* sont génétiquement très similaires. Leur répartition spatiale est généralement chevauchante, mais il semble que ces espèces écologiquement équivalentes co-occurrent rarement. Dans cette étude, nous investiguons l'histoire évolutive de ces deux jeunes espèces morphologiques en combinant des analyses de génétique de populations et des analyses spatiales. La répartition de la variation génétique, l'isolement par distance et les analyses de groupements à partir de polymorphismes d'AFLP indiquent que ces espèces morphologiques divergent régionalement. Elles sont rarement sympatriques à l'échelle locale et leur distribution forme une mosaïque d'îlots où l'une domine. L'interférence reproductive couplée à la recolonisation suivant la dernière glaciation pourrait expliquer l'origine et le maintien des îlots. Cette répartition spatiale relative limite probablement les opportunités contemporaines de flux génique, et accélère par le fait même la différenciation régionale. L'exclusion reproductive entre ces espèces morphologiques apparentées a eu des effets insoupçonnés sur leur structure génétique et spatiale et pourrait avoir un rôle à jouer dans l'histoire évolutive d'autres espèces.

## Introduction

Current research on speciation often aims at identifying and understanding the mechanisms driving the divergence process (Via 2001; Kirkpatrick and Ravigné 2002; Butlin *et al.* 2008; Fitzpatrick *et al.* 2008; 2009). Among the recent classifications proposed in the literature, those contrasting ecological versus non-ecological processes are particularly common (*e.g.* Schluter 2001; Via 2001; Rundle and Nosil 2005; Rundell and Price 2009; Schluter 2009). Ecological speciation, the process by which barriers to gene flow evolve as a result of ecologically-based divergent selection, has retained much attention in the recent past. Most case studies involve adaptive divergence along contrasting values of an environmental factor, but some special cases of sexual selection and other ecological interactions can also promote ecological speciation (reviewed in Rundle and Nosil 2005). Non-ecological speciation processes have also been proposed. These comprise cases where stochasticity plays a central role (*e.g.* speciation by polyploidization) and where the selection agent is non-ecological (*e.g.* speciation by sexual conflict).

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 in 1942). This classification is criticized because it breaks up a continuum into artificially discrete categories (Rice and Hostert 1993; Schluter 2001; Gavrilets 2003). After long lasting 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 distribution might not reflect historical geographical range, which can shift substantially after speciation (Losos and Glor 2003). Nevertheless, contemporary species distribution can help understanding speciation. The spatial context observed at each stage of speciation 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 distribution of diverging populations may contribute to a better understanding of the speciation process.

*Enallagma* is the most diverse group of damselflies, with 38 species, in North America (Westfall and May 2006). The vast majority of these species are morphologically very distinct, and are best discriminated on the basis of adult male caudal appendages. These reproductive structures are used in a lock and key mechanism to grasp females during mating, acting as a pre-zygotic barrier between closely related species (Paulson 1974; Robertson and Paterson 1982). *Enallagma spp.* form two old monophyletic clades (Brown *et al.* 2000) marked by distinct speciation rates (Turgeon *et al.* 2005) and one of these major clades comprises two groups of recently radiated species whose history has been influenced recent cycles of glaciation (Brown *et al.* 2000; Turgeon and McPeck 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 distribution of recently radiated *Enallagma* species is influenced by ecological and non-ecological factors. Four species are exclusively found in habitats where dragonfly larvae are the main predators, and there is ample evidence these habitat shifts were possible owing to adaptive evolution of morphological, physiological and behavioural traits (McPeck 1990a; 1995; 1999; 2000; Stoks *et al.* 2003). However, the many species inhabiting in the ancestral habitat where fish are the main predators can be considered as 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 and Crowley 1980; McPeck 1998), and species assemblage conform to random expectations (Siepielski *et al.* 2010). Moreover, relative species abundance is uncorrelated with any major environmental gradient, and experimental evidence indicates that *per capita* mortality and growth rates respond to total but not to relative species abundance (Siepielski *et al.* 2010).

*Enallagma hageni* and *E. ebrium*, which are both members of one recent radiation (Turgeon *et al.* 2005), are morphologically distinct. Individuals are easily distinguished based on the shapes of the reproductive structures (*i.e.* cerci) that males use to clasp females while mating (Westfall and May 2006, McPeck *et al.* 2008). The shapes of these

structures are distinctive and non-overlapping between species (McPeck *et al.* 2008), and so we assign species to morphospecies based on the shapes of these structures. Like other sister species issued from the same radiation, they share extensive mtDNA polymorphism, with the exception of a subgroup of haplotypes specific to *E. hageni* found in New England (the “Atlantic” clade of Turgeon and McPeck 2002). However, unlike most other species, an AFLP-based phylogeny suggested that *E. hageni* and *E. ebrium* are polyphyletic. There are no known ecological differences between these abundant and very common morphospecies (McPeck 1989; 1990b; Siepielski *et al.* 2010). Both species are widely distributed across North America, with their ranges being nearly identical (Westfall and May 2006). However, field observations suggest that both morphospecies are rarely sympatric at local sites with one morphospecies clearly predominating over large areas covering tens of square kilometers (M.A. McPeck and J. Turgeon pers. observ.).

The goal of this study was to disentangle the evolutionary history of these two very young morphospecies. To do so, we analyze patterns of genetic differentiation between morphospecies as well as their relative spatial distribution at local, regional and global (*e.i.* at all regions studied) scales. We first explore the full range of evolutionary hypotheses, or scenario, possibly explaining 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 good young species (*e.g.* as defined by the unified species concept, de Queiroz 2007), as morphology-based taxonomy readily suggests (H1 – ‘two species’). Each should then define a distinct lineage and each should display independent spatial distributions of genetic polymorphisms at all spatial scales. Second, occasional hybridization between lineages corresponding to morphospecies may blur genetic distinctiveness (H2 – ‘hybridization’). This is suspected because of the occasional observation of intermediate morphologies in reproductive structures (Catling 2001, M.A. McPeck and J. Turgeon pers. observ.); however, this seems improbable given the general local allopatry of species (see above). Evidence for this hypothesis would consist in detecting two lineages with species being less genetically differentiated in sympatry than in allopatry. Third, these morphospecies may still be polyphyletic and sorting, regionally or locally (H3 – ‘multiple lineages’). This hypothesis predicts that interspecific differentiation will occur only below a

certain spatial scale, while morphospecies will exhibit common patterns at largest scale. 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. Second, we establish the relative spatial distribution of these morphospecies. The degree of effective sympatry or allopatry between *E. hageni* and *E. ebrium* will help interpreting any genetic signals suggesting interbreeding as well as weighting evidence for the geographic context of speciation.

## Materials and methods

### *Patterns of 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) (Figure 1a). Within each region, we sampled allopatric (only one species present) and sympatric (both species present) sites. A site was considered allopatric for one morphospecies when it represented > 90% of individuals (among a minimum of 30 individuals). At each site, 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 and May 2006, McPeck *et al.* 2008). Females were kept only if caught while in tandem or copulating with a male and species identification was confirmed in the lab for every individual, using a dissecting microscope to examine mesostigmal plates in females (Westfall and May 2006, McPeck *et al.* 2009). Specimens were preserved by drying in the field and transferred to -80 °C until DNA analyses were performed.

Within each region, we characterized the AFLP genotypes of 6 to 24 individuals of each morphospecies from each of two sympatric and four to six allopatric sites (Figure 2; see

Appendix I for more details on sampling locations). In total, 264 *E. hageni* and 277 *E. ebrium* from 20 sites were analysed. We also analysed 10 individuals of *E. minusculum* (five from PE and five from NE) to ascertain that the strong geographical signal in genetic differentiation that we detected was not an artefact (see Results). We extracted DNA from non-diagnostic anatomical parts, *i.e.* thorax muscle tissue for males and abdomen for females, following Aljanabi and Martinez (1997); DNA quality was verified on 2% agarose gel. We quantified DNA with spectrophotometry and diluted samples to 100 ng/ul. We generated AFLP fragments using the restriction enzymes *EcoRI* and *MseI* (New England Biolabs, Ipswich, MA) following AFLP® Plant Mapping protocol of Applied Biosystems (2007-2010) with slight modifications. Three *EcoRI/MseI* primer pairs were used in selective PCRs: AGG/CACG, ACC/CACA and ACA/CACA (note that the *MseI* primer has four selective nucleotides). Selective PCRs included a denaturation step of 20 s at 94°C, 9 ‘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 PCR products on ABI 3100 capillary sequencer with LIZ size standard (Applied Biosystems). AFLP profiles were checked and scored manually using GeneMapper 3.7 analysis software (Applied Biosystems) with a minimum relative fluorescence units set at 100.

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 and McPeck (2002) for five individuals by morphospecies for three sites in PE and QO and 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 species and/or putative refugial lineages (Turgeon and McPeck 2002; Turgeon *et al.* 2005).

### Genetic analyses

We used three complementary approaches to ascertain whether genetic polymorphisms conformed to the predictions of our four hypotheses. Specifically, we contrasted patterns of genetic variation partitioning, isolation by distance, and genetic clustering when comparing

morphospecies, sympatric and allopatric morphospecies populations at global and regional spatial scales.

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 with Arlequin 3.5 (Excoffier and Lischer 2010) and compared the explanatory power of our different grouping 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 levelled off, thereby strengthening the comparison between expectations associated with H1 and H3.

Second, we examined patterns of genetic differentiation in relation to geographic distances (isolation by distance, IBD). Higher or similar interspecific vs. intraspecific differentiation at all spatial scales would support H1 and H4, respectively. Higher interspecific differentiation restricted to small spatial scale (*i.e.* a morphospecies signal at small scale only) would support H3. Pairwise  $F_{ST}$  between samples (morphospecies at each site) were estimated using Bayesian allele frequency estimation with non-uniform prior distribution with AFLP-SURV 1.0 (Vekemans *et al.* 2002).  $F_{ST}/(1-F_{ST})$  was related to Euclidian geographic distance between sites for intra- and interspecific comparisons, and significance was evaluated with Mantel tests performed with IBDWS (Jensen *et al.* 2005).

Finally, we identified genetic clusters and assessed their correspondence with morphospecies, sampling regions and sites. Genetic clusters corresponding to morphospecies are expected at any spatial scales under H1 but only at reduced spatial scale under H3. Clustering primarily by regions is expected under H4 while higher incidence of mixed ancestry (intermediate  $q$  values) in sympatric than allopatric sites would offer support for H2. We used STRUCTURE 2.3.1 (Pritchard *et al.* 2000; Falush *et coll.* 2007, Hubisz *et al.* 2009); this bayesian clustering analysis attempts to identify the number of



genetic clusters ( $K$ ) without *a priori* information on population membership and then estimates a coefficient of ancestry ( $q$ ) in each group. We performed the analyses with and without information about samples (morphospecies by site as loc-prior), and 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 and subsequent run lengths to 200 000 iterations. We did 10 runs for each  $K$  value tested ( $K= 1-8$  clusters for PE and NE,  $K= 1-10$  for QO and over all regions). We used  $\ln P(X|K)$  (Pritchard *et al.* 2000) and  $\Delta K$  (Evanno *et al.* 2005) as criteria to infer the number of clusters ( $K$ ). Figures were made with DISTRUCT (Rosenberg 2004).

### *Species distribution*

Our main purpose was to ascertain whether *E. hageni* and *E. ebrium* are randomly distributed relative to one another, or if, as our field observations suggest, these species are rarely sympatric. To do so, we analyzed the distribution of species frequency at local, regional and global scales 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 (VT, NH and ME, USA), and 25 sites in Prince Edward Island, (PE, Canada). At each site, 30 to 45 individuals were collected whenever possible and species 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'). Details on museum data source and treatment are given in Appendix II. The complete database included 31393 individual observations (51.5% *E. hageni*) each of which was given a discrete value (0 for *E. hageni* and 1 for *E. ebrium*). These covered 2128 sites distributed across PE, NS, NB, southern Quebec, and eastern Ontario in Canada, as well as ME, NH, and VT in the USA (Figure 1a).

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

To estimate the distribution of species frequency at the regional and global scales, we performed ordinary kriging interpolation, a minimum mean squared error method for spatial prediction (Cressie 1993). Using SAS version 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 and Oliver 1990). ARCGIS version 9.2 (Environmental Systems Research Institute) was used for kriging *per se* as well as for mapping.

In order to assess how the distribution of sites affected that of species frequency, we compared kriging interpolation maps obtained with empirical vs. randomized datasets. The empirical data comprises museum and sampling datasets for the entire zone considered. Validation of museum data and rationale for conducting the analyses at the overall scale are detailed in Appendix II. To randomize species distribution, we built 100 independent datasets where species identity (0 or 1) of each individual observation was randomly assigned while respecting the empirical proportion of each species as well as empirical site locations. The percentage of territory (divided in 10 km<sup>2</sup> cells) where one species 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 species rarely represented  $\geq 75\%$  of the individuals.

## Results

### *Genetic analysis*

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

AMOVA analyses revealed that genetic variation was best partitioned by regional morphospecies (Table 1). Using all sites, the regional morphospecies grouping model ( $AIC_c = 1493.1$ ) explained a larger component of total genetic variation than models considering either only morphospecies ( $AIC_c = 1518.5$ ) or regions ( $AIC_c = 1503.0$ ) (Table 1a). Clearly, no significant variation was explained by morphospecies ( $P = 0.185$ ). Analyses of only sympatric sites provided very similar results (Table 1b): regional morphospecies grouping ( $AIC_c = 633.6$ ) better explained genetic variation partitioning than morphospecies ( $AIC_c = 647.3$ ) or regional grouping ( $AIC_c = 640.6$ ), and again, no variation was explained by morphospecies ( $P = 0.432$ ).

Genetic differentiation was related to geographic distance between sites and a global IBD pattern was detected using all sites and both morphospecies (Figure 3a,  $P < 0.001$ ). Differentiation within each morphospecies also yielded an IBD signal ( $P < 0.001$  for both *E. ebrium* and *E. hageni*), but with a steeper regression slope for *E. hageni* than *E. ebrium*. At large distances (*e.g.*  $> 600$  km), interspecific differentiation was intermediate between intraspecific levels (Figure 3a). However, interspecific differentiation seemed to exceed intraspecific differentiation at small spatial scale (*ca.* 200 km). This variation in levels of inter- vs. intraspecific differentiation with geographic distances was not an artifact. Indeed, *E. minisculum* intraspecific differentiation between PE and NE ( $F_{ST} = 0.120$ ) was always smaller than interspecific differentiation with *E. hageni* or *E. ebrium* from the same region ( $F_{ST}$ : 0.521 to 0.636).

To test whether  $F_{ST}$  values were significantly larger than expected at small distances only, we used an approach combining stepwise regression and permutation procedures. First, we calculated the global IBD regression slope for long distance comparisons ( $> 600$  km). This regression line was used as a prediction for the expected interspecific IBD pattern at all spatial scales. Then, we plotted interspecific  $F_{ST}$  residual values against all distances. Under the hypothesis that differentiation is solely related to distance, residuals should be distributed around zero without any systematic distance effect. We fitted a piecewise regression and found that regressions differed below and above 363 km ( $P < 0.001$ , Figure 3b). 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 ( $P = 0.0181$ ). Note that for this method, we excluded comparisons involving *E. hageni* sampled in ME1 and NH4. As is shown below, these two sites proved to harbor representatives of a different mitochondrial lineage (Atlantic clade, Turgeon and McPeck 2002) that were also strongly differentiated at AFLP loci, distorting the general IBD pattern (see Appendix III for  $F_{ST}$  values).

Clustering analyses using the entire dataset provided no statistical support for two or three clusters ( $K=2$  and  $K=3$ ). When considered, there was clearly no correspondence between these genetic clusters and morphospecies or regions (results not shown). By contrast, there was evidence for the presence of two or three clusters within each region, and these generally corresponded to morphospecies (Figure 4). In PE, clustering by morphospecies was unambiguous, with  $K=2$  clusters using both Pritchard's and Evanno's criteria. In NE, both criteria supported the existence of three clusters. Three samples of *E. ebrium* from westernmost sites clearly formed the bulk of the first cluster (*E. ebrium* from VT1, VT10, VT3), and the other *E. ebrium* sample (from site NH4) also had ancestry in this cluster. The other two clusters were more strongly associated with *E. hageni*, one with the two easternmost sites (ME1 and NH4), and the other with western sites (VT3 and VT11). 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). Indeed, at K=2, the genetic clusters correspond approximately to Quebec vs. Ontario sites, while at K=3, the third cluster clearly corresponded to *E. hageni*, especially in Quebec.

When in sympatry, morphospecies belonged to distinct genetic clusters. In PE, this pattern was very clear, as well as in Quebec for K=3. In both regions, each morphospecies had much stronger ancestry in the cluster typical of that morphospecies in allopatry, and individuals were no more of mixed ancestry than when species were allopatric. In NE, sympatric morphospecies were generally very distinct at site VT3 but *E. ebrium* from site NH4 had mixed ancestry in clusters associated with each morphospecies.

MtDNA polymorphisms revealed different patterns of genetic similarities between species within each region (see Appendix IV for haplotype network). 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 that has been previously documented in both species (Turgeon and McPeck 2002). By contrast, in NE, morphospecies were associated with distinct sets of haplotypes. Haplotypes previously labelled as those of an Atlantic *E. hageni* lineage were found 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. *E. hageni* also possessed haplotypes typical of the Continental lineage (VT11). Most *E. ebrium* from NE carried one of two haplotypes (H-008, H-009, see Appendix IV) very commonly documented in *E. ebrium* from an extensive area in NE (Turgeon and McPeck 2002).

### *Spatial distribution*

The distribution of morphospecies frequency revealed that one morphospecies was generally dominant at any given site (Figure 5). Indeed, one morphospecies accounted for more 90% of the individuals in 62% of the sites. At that local scale, the distribution of morphospecies was clearly heterogeneous (PERM membership homogeneity test,  $P < 0.001$ ).

At the regional scale, there was moderate spatial autocorrelation over small distances, as estimated by the nugget/sill ratio (range: 0.24 - 0.74, Table 2). The kriging map based on empirical data revealed a mosaic of patches where one or the other morphospecies clearly dominates (Figure 1c). Overall, one species represented 75% or more of the individuals over 53% of the territory (Table 2). This pattern was also apparent within each region (range: 51 - 68%, Table 2). This is in sharp contrast with the kriging interpolation using randomized datasets where, in average, only 0.7% (range: 0.1 - 2.2%) of the territory was similarly dominated by either morphospecies (Figure 1b).

## 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 reproductive independence at small spatial scale only. Indeed, genetic variation between regional morphospecies best explained global variation partitioning, and interspecific genetic differentiation between morphospecies was more pronounced at small geographic scale (< 363 km). 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, while *E. hageni* was split into two clusters. One *E. hageni* cluster likely corresponds to an Atlantic refugial lineage already documented in other studies, thus reflecting the historical, likely more ancient split within this morphospecies (Brown *et al.* 2000; Turgeon and McPeck 2002; Turgeon *et al.* 2005). Indeed, this cluster included all *E. hageni* individuals from the easternmost sites (ME1 and NH4) and these all possessed mitochondrial haplotypes typical the Atlantic clade. This may explain why *E. hageni* from these sites are more differentiated from other *E. hageni* samples than is expected from global IBD alone (see Appendix III for  $F_{ST}$  values). In QO, the morphospecies signal was confounded by a geographic signal. This region is wider than

PE and NE, such that it involves interspecific comparisons between very distant sites (nearly 500 km). These comparisons are not likely to reveal strong differentiation, as our extended IBD analysis suggests. Nevertheless, morphospecies formed clusters over smaller distances, for instance within Quebec. In addition, there were distinct collections of mtDNA haplotypes within each region, further supporting the local divergence of these morphospecies.

Ongoing regional differentiation is also much better supported than the alternative hypotheses. First, there was no convincing evidence that morphospecies form two globally distinct lineages, which rejects our H1 that these are two reproductively isolated and phylogenetically independent species. There was clearly no genetic variation associated with morphospecies partitioning in 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* (McPeck 1989), characterized by a shallower IBD slope, is probably a better explanation than globally distinct species. Moreover, as mentioned above, morphospecies from different regions did not form clusters. Second, the unlikely possibility of a single lineage comprising two alternative morphotypes (H4) is easily refuted. Within region, morphospecies most generally belong to distinct genetic clusters, particularly when in sympatry. Finally, we found little evidence that contemporary hybridization is commonly occurring (H3).  $F_{ST}$  values and genetic clustering patterns provided no evidence that species are less different when sympatric. Moreover, despite a large sampling effort, only 8 individuals (*i.e.* 0.2%), all from different sites, possessed unusual cerci morphology. The unusual morphology may well be the result of developmental malformation rather than the consequence of hybridization. It is worth noting, however, that hybridization may have been more common in the past. For example, clustering analysis suggest that *E. ebrium* individuals at site NH4 in New England (Figure 4b) are of mixed ancestry between the Atlantic *E. hageni* lineage and *E. ebrium*. Moreover, this hybridization may have been asymmetrical (female *E. ebrium* x male *E. hageni*) given that none of these *E. ebrium* individuals possessed *E. hageni* Atlantic haplotypes.

### *E. 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 species generally do not co-occur at local and regional scales within these ranges. Locally, one morphospecies is usually very dominant, and species are rarely equally frequent when sympatric. Regionally and globally, lakes with the same dominant species are aggregated, creating a mosaic of patches alternating in species dominance. Our previous experiences in other parts of their ranges also suggest this to be true across their entire ranges (M.A. McPeck, pers. observ.). This first formal description of these species relative spatial distributions confirms the field experience of many odonatologists (Walker 1953; P.M. Catling, M.R.L. Forbes and P.M. Brunelle pers. comm.).

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 and Harrison 1984; Bridle *et al.* 2001). However, no ecological differences are known between *E. hageni* and *E. ebrium* (McPeck 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 and Crowley 1980, McPeck 1990a, 1998), and all these 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.* McPeck 1989, 1990b, 1998, Siepielski *et al.* 2010). We have identified no environmental factor that account for this segregation after many years of searching.

Reproductive interference between *E. hageni* and *E. ebrium* is more likely to explain local allopatry and the maintenance of patches where one species is clearly dominant. Reproductive interference comprises any interspecific sexual interaction which negatively affects the fitness of at least one of the species involved (reviewed in Gröning and



Hochkirch 2008). Reproductive interference can take many forms (*e.g.* misdirect courtship, heterospecific mating, hybridization) and may lead to different issues (*e.g.* sexual exclusion, spatial segregation, reproductive character displacement). In a fashion similar to Lotka-Volterra's competitive exclusion process, reproductive interference between closely related species may lead to local exclusion (Kuno 1992). The strength and symmetry of interference modulate the issue of sexual interactions and the initial density is determinant; in many cases, the rarer species has a higher probability of being driven to exclusion. 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 and Kocum 2005 *in* Gröning *et al.* 2008). The incomplete mate recognition systems may be more relevant than habitat partitioning to explain the distributional pattern (Gröning *et al.* 2007).

Reproductive interference is very likely happening in *Enallagma*. Males are highly promiscuous, trying to mate with all *Enallagma* females they encounter, regardless of species (Paulson 1974, Fincke *et al.* 2007). Males try to grasp females with their caudal cerci but female control the final stage of reproduction as they must bend their abdomen to be inseminated. Overall, female preferences exert control over reproduction (Robertson and Paterson 1982), and copulation does not take place between heterospecific couples (Fincke *et al.* 2007, McPeck and Turgeon, unpubl. data). Thus, when either *E. hageni* or *E. ebrium* is less frequent at a site, as is commonly observed, heterospecific mating attempts by the rare morph may lead to a reduced mating success, causing its gradual exclusion. Moreover, reproductive interference is highly compatible with the maintenance of the mosaic distribution pattern by countering the migration of morphospecies in a patch dominated by the other morphospecies (see below).

In addition to reproductive interference, postglacial recolonization processes may also have played a significant role in first establishing the basis for 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 expansion 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 and Hewitt 1994; Ibrahim *et al.* 1996; Bialozyt *et al.* 2006; Ray and 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 and FitzJohn 2010). Given that reproductive interference strengthens assortative mating between sexually interacting species, postglacial colonization and sexual interaction between *E. hageni* and *E. ebrium* may allowed the establishment of the observed mosaic distribution.

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

*E. 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* define classic morphospecies long recognized in taxonomy (Westfall and May 2006). These are principally discriminated on the basis of male caudal cerci, a structure used by male to grasp female prior to copulation (Paulson 1974; Robertson and Paterson 1982; Westfall and May 2006). This lock and key mechanism is an effective pre-zygotic isolation mechanism (Fincke *et al.* 2007) that also helps defining good biological species. Our results confirm that *E. hageni* and *E. ebrium* do not regularly interbreed, and, additionally, that their current spatial distribution offers little opportunity to do so. Unlike other species from this radiation, however, these two species are not clearly genetically differentiated from one another. Here, we show that these species are most likely undergoing local or regional differentiation, and thus, still largely polyphyletic.

The peculiar distribution of these species over a large part of their common range defines a spatial context that is favourable both for local divergence and persisting polyphyletic relationships. Locally, and within large patches of territory, species are essentially allopatric. This setting limits reproductive interference and facilitates the action of genetic drift within patches of metapopulations. Local sorting may even be accelerated by this

spatial context. This is suggested by the much clearer interspecific differentiation on Prince Edward Island (PE). *E. hageni* and *E. ebrium* form very sharp patches on the western end of the island (Figure 1c, J. Turgeon pers. observ.), these being surrounded by water. Lower migration pressure from surrounding patches may have helped the formation of these patches and easing their differentiation. Globally, however, the mosaic of allopatric patches probably contributes to preserving multiple and increasingly differentiated sub-lineages of *E. hageni* and *E. ebrium*. Reproductive interference would then restrict the successful establishment of individuals of one morphospecies in patches dominated by the other species. These processes may contribute to the relative spatial and temporal stability of patches, and consequently, to maintaining differences among patches occupied by the same species. However, if patches are temporally dynamic, migration and recolonization may eventually homogenize the genetic composition of each species.

The regional pattern of genetic differentiation suggests that one morphospecies, and the associated typical caudal cerci morphology, may have appeared in parallel more than once. Distinct regional pools of mtDNA haplotypes shared by both species could support this hypothesis. Moreover, multiple apparition of the same reproductive structure has already been observed in *Enallagma*. For example, Palaearctic *E. cyathigerum* and Nearctic *E. annexum* were previously regarded as one species on the basis of highly similar cerci (Westfall and May 1996); however, phylogenetic relationships clearly show that they are highly divergent species and that the same cerci morphology very likely evolved twice (Turgeon *et al.* 2005, Westfall and 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 (Shen *et al.* 2009) rather suggests that the diversification of cerci morphology is not severely constrained. 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 distribution range of these species.

Alternatively, the new morphospecies may have appeared only once, early in the radiation, such that both species 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 species 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. 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 apparition seems more parsimonious than multiple parallel evolution, these alternative scenarios are still speculative at this point in time. Our data is based on a large number of genetic markers, but these are likely neutral markers mostly reflecting the history of migration and drift in these morphospecies. To discern between single and multiple origins of traits discriminating these species, sequence information of functional genes would be much more instructive.

## Conclusion

In conclusion, this study revealed the polyphyletic relationships between *E. hageni* and *E. ebrium*, and the joint analysis of genetic variation and relative species distribution allowed disentangling recent, complex evolutionary histories marked by the unexpected influence of reproductive interference on speciation. This phenomenon probably allowed the establishment and maintenance of a mosaic species distribution, which in return permitted the evolution of distinct lineages in each patch. Our results call for more attention on the action of such frequency dependent selection in the study of recent speciation events involving ecologically equivalent species in the context of rapid colonization.

**Table 1** Partitioning of genetic variance for grouping models considering morphospecies, region or regional morphospecies 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).

a)

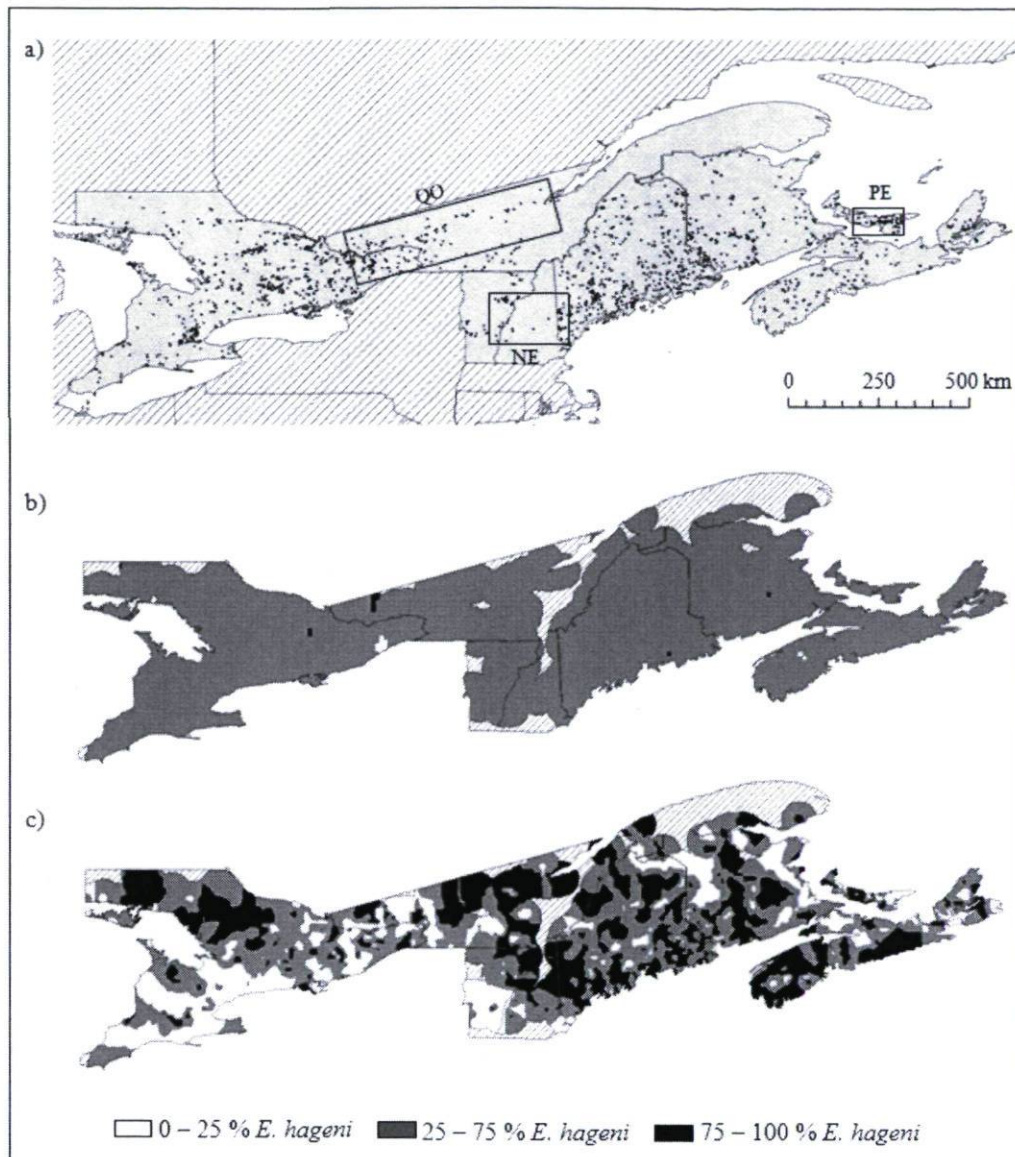
Grouping model	df	SS	% Variation	Fixation index	P
By morphospecies ( $AIC_c = 1518.5$ )					
Between morphospecies	1	51.0	0.26	0.003	0.185
Among sites within morphospecies	24	920.0	6.72	0.067	< 0.001
Within sites	515	7905.5	93.02	0.070	< 0.001
By regions ( $AIC_c = 1503.0$ )					
Among regions	2	300.3	4.00	0.040	<0.001
Among sites within regions	23	670.6	4.00	0.042	< 0.001
Within sites	515	7905.5	92.00	0.080	< 0.001
By regional morphospecies ( $AIC_c = 1493.1$ )					
Among regional morphospecies	5	455.6	4.34	0.043	< 0.001
Among sites within regional morphospecies	20	515.3	3.06	0.032	< 0.001
Within sites	515	7905.5	92.60	0.074	< 0.001

b)

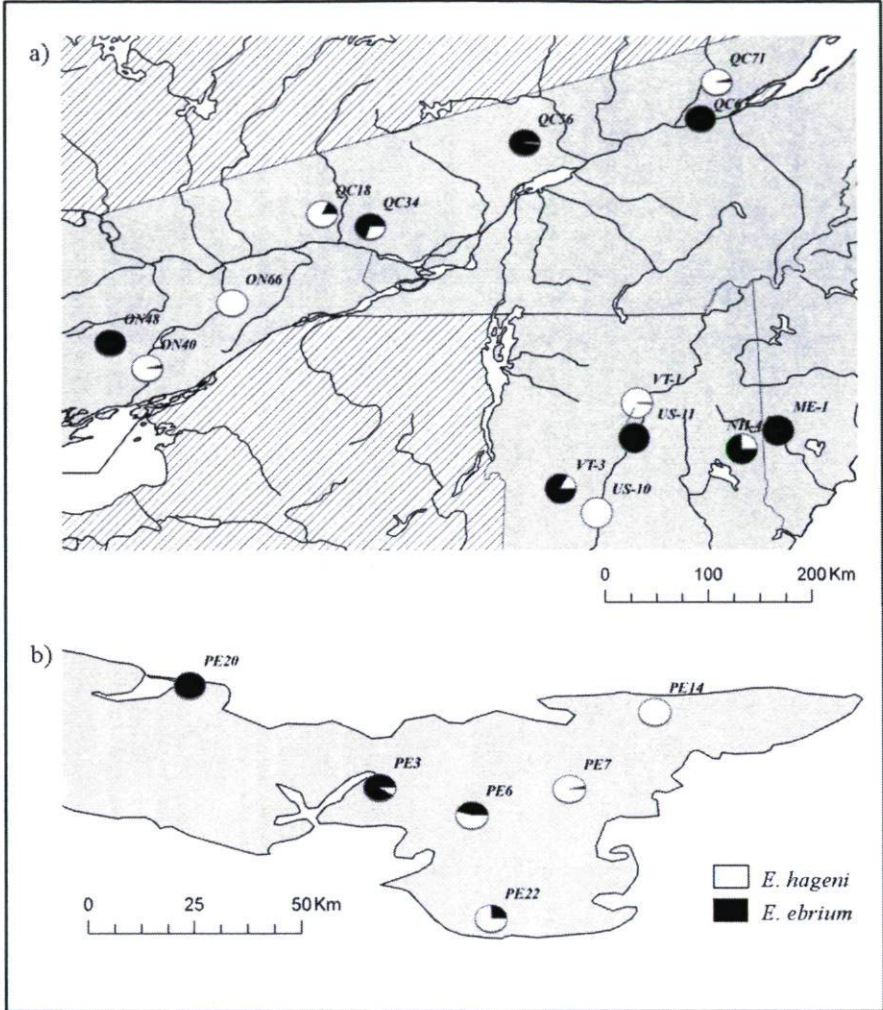
Grouping model	df	SS	% Variation	Fixation index	P
By morphospecies ( $AIC_c = 647.3$ )					
Between morphospecies	1	37.8	0.00	0.000	0.431
Among sites within morphospecies	10	355.9	6.54	0.065	< 0.001
Within sites	218	3347.7	93.46	0.065	< 0.001
By regions ( $AIC_c = 640.6$ )					
Among regions	2	145.5	3.49	0.035	< 0.001
Among sites within regions	9	248.2	3.95	0.041	< 0.001
Within sites	218	3347.7	92.56	0.074	< 0.001
By regional morphospecies ( $AIC_c = 633.6$ )					
Among regional morphospecies	5	252.1	4.06	0.041	< 0.001
Between sites within regional morphospecies	6	141.6	2.82	0.029	< 0.001
Within sites	218	3347.7	93.12	0.068	< 0.001

**Table 2** Species distribution kriging interpolation results (see also Figure 3) for regional and global analyses.

State / Province	N entries	Nugget/sill ratio	% territory where one morphospecies is dominant ( $\geq 75\%$ )		
			<i>E. hageni</i>	<i>E. ebrium</i>	Total
ON	9391	0.27	21	30	51
QC	2278	0.47	42	13	55
NE	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

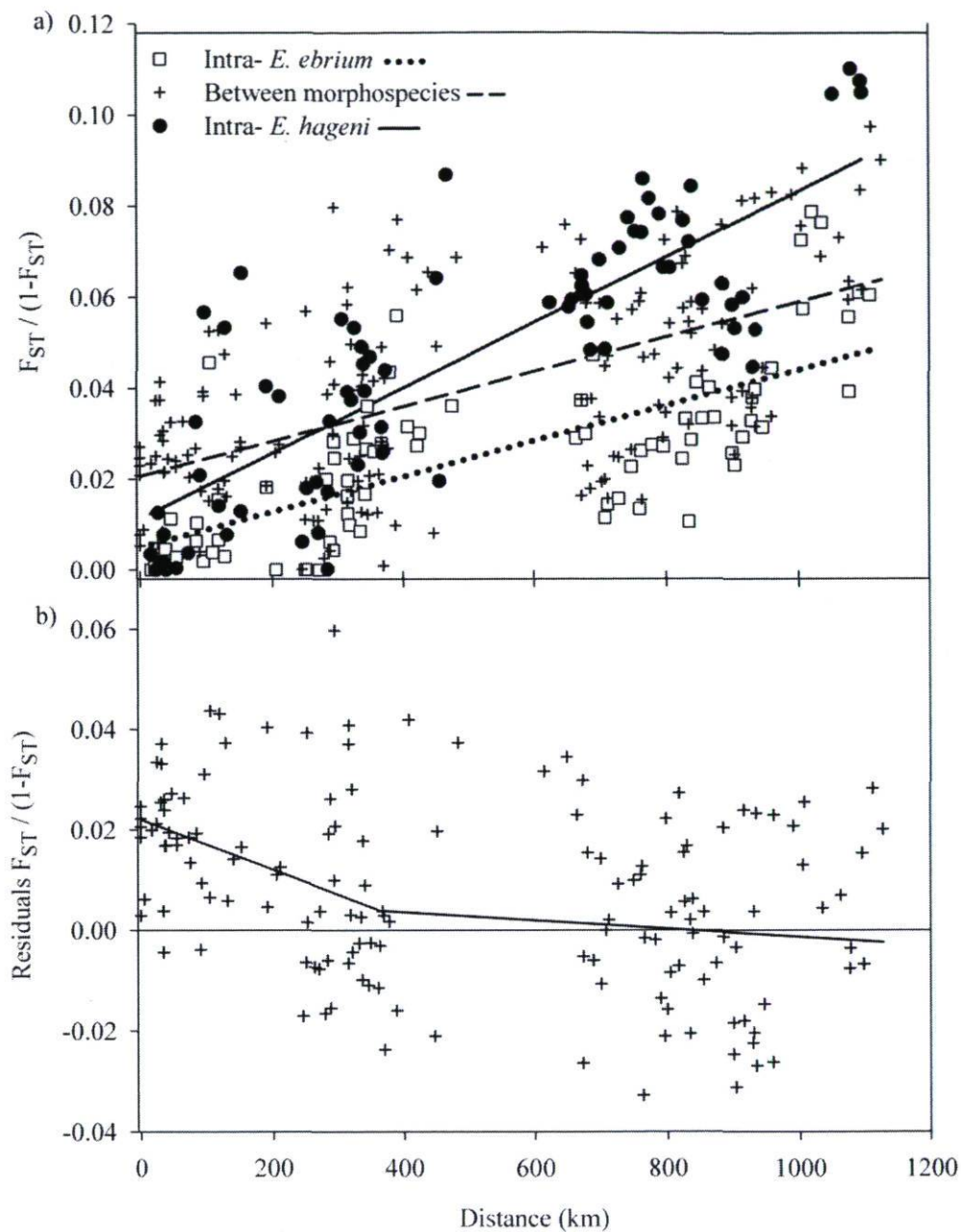


**Figure 1** Sampling regions for genetic analyses and location of sites used in global kriging interpolation of morphospecies frequency distribution. a) Distribution of sites used in spatial analysis (dots), with boxes delimiting the three regions (QO, NE and PE) where samples were used in genetic analyses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

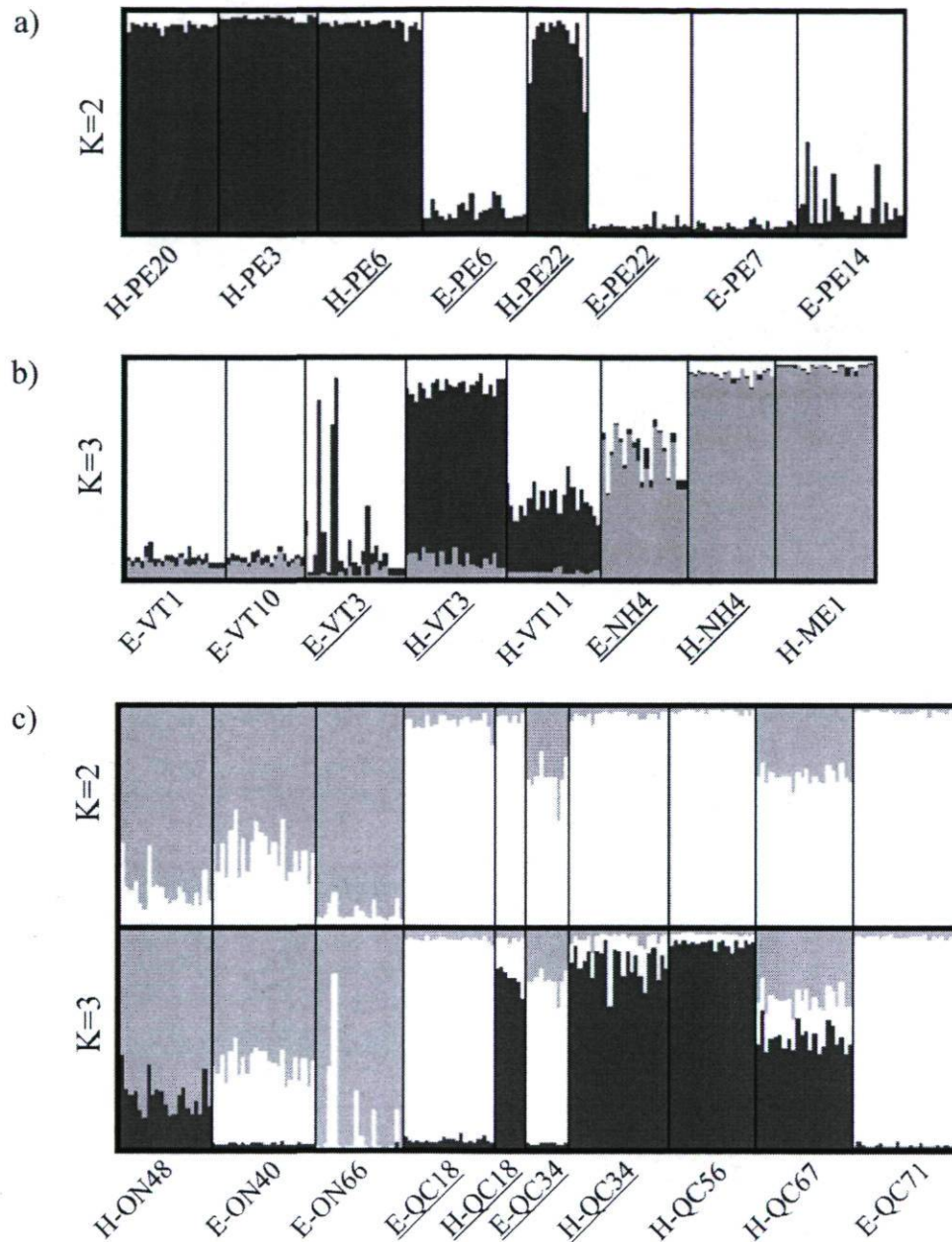


**Figure 2** Morphospecies relative distribution in sampling sites used in genetics analysis in the three study regions: a) QO, NE and c) PE (see Table A1 for more information on sites and sample sizes).

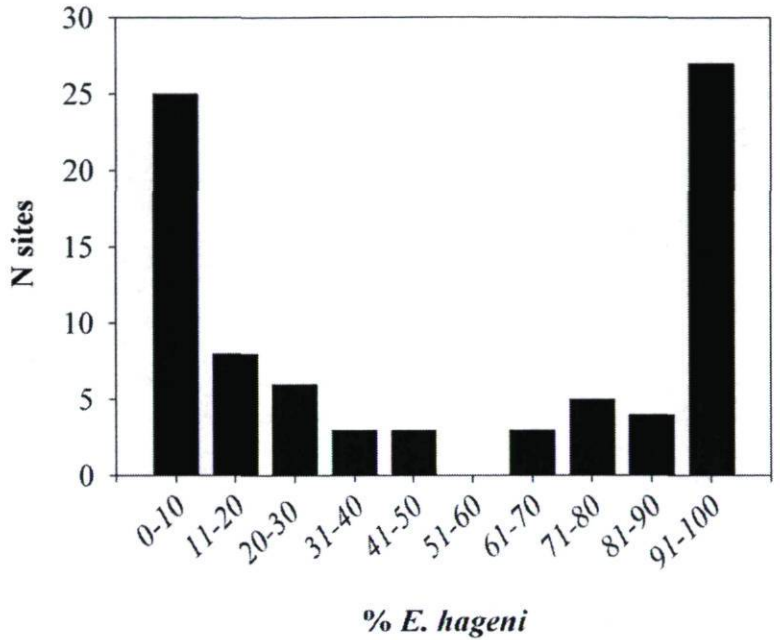




**Figure 3** Genetic differentiation in relation with distance. a) IBD pattern for intraspecific and interspecific comparisons. b) Interspecific residual differentiation in relation with distance. A piecewise regression slope is fitted (solid line), indicating a break at 363 km. See Results for rationale for this analysis.



**Figure 4** Regional clustering of AFLP 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 species were present are underlined.



**Figure 5** Frequency of *E. hageni* among 84 sites sampled in PE, NE and QO.

## Appendix I: Sites description

**Table A3** Description of sites sampled for *Enallagma hageni* (H) and *E. ebrium* (E) in three geographical regions indicating species frequency, as well as sample sizes for genetic analyses using AFLP and mtDNA.

Region / Site	Latitude °N	Longitude °W	Species observed	% H	AFLP		mtDNA	
					H	E	H	E
<i>Prince Edward Island (PE)</i>								
PE20	46.49	-63.40	H	100.0	21		5	
PE3	46.27	-63.00	H	91.2	23			
PE6	46.21	-62.81	H/E	46.0	24	24	5	5
PE22	46.00	-62.77	H/E	25.4	14	24		
PE7	46.27	-62.60	E	3.2		24		
PE14	46.43	-62.42	E	0.0		24		5
<i>New England (NE)</i>								
ME1	44.00	-70.93	H	100.0	23		5	
VT11	44.19	-72.12	H	100.0	22		5	
NH4	43.84	-71.19	H/E	74.4	23	18	5	5
VT3	43.49	-72.71	H/E	82.8	22	20	5	5
VT10	43.28	-72.45	E	0.0		23		5
VT1	44.24	-72.10	E	1.6		20		5
<i>Quebec/Ontario (QO)</i>								
ON48	44.77	-76.69	H	100.0	21			
ON40	44.55	-76.37	E	3.5		24		
ON66	45.11	-75.62	E	0.0		20		5
QC18	45.89	-74.84	H/E	17.5	6	22		
QC34	45.78	-74.42	H/E	71.4	23	10		
QC56	46.51	-73.07	H	97.9	20		5	
QC67	46.72	-71.54	H	100.0	22		5	
QC71	47.04	-71.40	E	4.0		24		

## Appendix II: Spatial analysis

### *Section 1. Museum data source*

**Museum collections:** National Museum of Natural History (Washington, DC, USA), New Brunswick Museum (Saint John, NB, CAN), Peabody Museum (Yale University, CT, USA)

**Province, state and country collections:** Atlantic Canada Conservation Data Center (Sackville, NB, CAN), Laurentian Forestry Centre (Canadian Forest Service, Quebec, QC, CAN), International Odonata Research Institute (Gainesville, FL, USA), Natural Heritage Information Centre (Peterborough, ON, CAN)

**Private collections:** P.M. Brunelle, M.A. McPeck (Dartmouth College, NH, USA)

### **Data from publications:**

Harper, P. P., and L. Cloutier. 1986. Spatial structure of the insect community of a small dimictic lake in the Laurentians (Quebec). *International Revue der Gesamten Hydrobiologie* 71:655-685.

Hellebuyck, V. 1993. Confirmation of *Enallagma civile* (Hagen) (Odonata :Zygoptera) in southern Quebec. *Argia* 5:9.

Hutchison, R. 1978. Récoltes d'odonates au Québec et dans les états de New York, Vermont et New Hampshire de 1972 à 1978. *Cordulia* 4:109-137.

Jobin, L.-J., and J.-M. Perron. 2004. Odonatofaune du parc écologique du mont Shefford, division de recensement de Shefford, Québec. *Le naturaliste canadien* 128:27-30.

Larochelle, A. 1979. Observations sur l'accouplement et la ponte de 22 espèces d'odonates du Québec. *Cordulia* 5:34-37.

Larochelle, A. 1979. Capture d'odonates au Québec en 1979. *Cordulia* 5:80-81.

Lebuis, M.-A., and J.G. Pilon. 1976. Analyse comparative de la faune odonatologique de quatre milieux de la région de Saint-Hyppolyte, comté de Prévost, Québec. *Annales de la Société entomologique du Québec* 21:3-25

- Legault, J. 1979. Liste préliminaire des odonates de l'étang Streit à Philipsburg, comté de Missisquoi, sud du Québec. *Cordulia* 5:39.
- Legault, J. 1979. Liste des odonates capturés à Frelighsburg (Missisquoi), Québec. *Cordulia* 5:40.
- Legault, J. 1979. Capture d'odonates à l'île-aux-Noix (Saint-Jean), Québec. *Cordulia* 5:40.
- Ménard, B. 1991. Capture d'odonates dans la vallée de l'Outaouais et dans la Haute-Gatineau. *Fabriques* 16:69-77.
- Ménard, B. 1996. Liste annotée des odonates de la vallée de l'Outaouais. *Fabriques* 21:29-61.
- Pilon, J.-G., and D. Lagacé. 1993. Additions à la faune odonatologique du lac Martin et de son émissaire. *Fabriques* 18:33-34.
- Pilon, J.-G., L. Pilon, and D. Lagacé. 1993. Composition et caractérisation du peuplement odonatologique d'un marais blainvillois situé dans la zone bioclimatique de l'érablière à caryers du Québec, Canada. *Notulae odonatologicae* 4:23-29.
- Roch, J.-F. 1989. Liste des odonates récoltés à Granby, division de recensement de Shefford, Québec. *Fabrique* 14:44-45.
- Samson, P. 1978. Premier aperçu sur les odonates de la région de Saint-Joseph-de-la-Rive, comté de Charlevoix, Québec. *Cordulia* 4:35-38.

## *Section 2. Data treatment and kriging specification*

Prior to analysis, museum database entries were checked to keep only informative and precise data. We excluded entries specifying locations that likely covered several specific water bodies (*e.g.* 'Toronto' or 'Ottawa River' with imprecise coordinates), as well as those dating from before 1990. In Quebec, sampling has been less extensive; we kept observations from 1970 and on. In both Quebec and Ontario, northern areas were less well-covered. We retained sites from the southern parts of these two provinces (Québec: south of *ca.* 200 km north of Ottawa and St-Lawrence Rivers; Ontario: west 47.0 N, 83.0 W).

For kriging interpolation, we created an individual database where each individual observation had a unique geographic coordinate (kriging interpolation cannot analyze more

than one observation by location). When there was more than one observation at a given site, individual coordinates were shifted by 1m relative to the initial coordinate. In order to limit the influence of sites with large sampling size, we also retained a maximum of 100 individuals per site. *E. hageni* and *E. ebrium* were given a discrete value (respectively 0 and 1), and our objectives was to estimate species proportion in each cell. We specified a search radius of 30 observations around the predicted cell. For mapping, we excluded cells located farther than 50 km from an individual observation. It is worth noting that the final database easily allowed for kriging interpolation given that these two morphospecies were very common and present over all studied territory.

### *Section 3. Data validation*

#### 3.1 Sampling vs. Museum data

Museum databases are an important source of information for mapping species distribution. However, they may comprise many errors (reviewed in Suarez and Tsutsui 2004) and they are rarely validated. To verify that museum data reflected field observation, we compared kriging results obtained with museum vs. sampling data for Prince Edward Island (PE). First, we performed independent kriging interpolation with the best fit function for each semivariogram (respectively exponential and Gaussian for museum and sampling data) using 5 km<sup>2</sup> cells. For each data type, the kriging interpolation was mapped with three categories of relative morphospecies proportion ( $\geq 75\%$  *E. hageni*,  $\leq 25\%$  *E. hageni*). We performed cell to cell comparison of kriging interpolation maps obtained with each data type and counted cells belonging to different frequency categories. We found that 54% of cells remained in the same category, 38% of cells corresponded to a change by one category (more than 75% of a morphospecies to intermediate value, or the contrary), and only 8% of cell comparisons corresponded to a change within the dominant morphospecies. Museum data, although probably less precise than those from sampling, were therefore valuable to interpolate relative morphospecies distribution. Thus, we pooled field sampling and museum data for subsequent kriging interpolation reported in the main text.

### 3.2 Province/State vs. Global kriging

We compared regional semivariograms to ascertain the influence of variable site density and data richness on the estimated distribution of relative species frequency. We calculated the semivariograms and fitted functions for each province or state with 10 km<sup>2</sup> cell size. Results were very similar across regions, with the exponential model best fitting in all regions except in Quebec where the Gaussian function fitted better. Moreover, estimated function parameters such nugget and sill had similar values. Thus, we performed the analysis with the entire data set for the entire zone. The global mosaic pattern was highly similar to that found within each region, and was smoother near province/state boundaries.



### Appendix III : Regional Pairwise $F_{ST}$

Pairwise  $F_{ST}$  between samples revealed that morphospecies are differentiated within each region (Table A1). In PE, intraspecific comparisons show no significant differentiation, but all interspecific comparisons are significantly different from zero. Moreover, sympatric and allopatric comparisons are alike. In NE, genetic differentiation seemingly reflects the mtDNA clade dominating at each site.  $F_{ST}$  comparisons involving *E. hageni* from NH4 and ME1 (Atlantic clade) and other sample sites (Continental clade) have higher values than expected under isolation by distance alone. The only exception is the small differentiation observed with *E. ebrium* from NH4, a result congruent with clustering analysis. In QO, the largest part of genetic differentiation was explained by distance (between Quebec and Ontario comparisons). Within Quebec, we found a large genetic differentiation between morphospecies, but not within Ontario.

**Table A2**  $F_{ST}$  between pairs of sampling sites with *E. hageni* (H) and *E. ebrium* (E) for three regions: a) PE, b) NE and c) QO.  $F_{ST}$  value associated P-values < 0.05 are in italic; bold  $F_{ST}$  value indicate Bonneferroni corrected comparisons for  $\alpha=0.05$ .

a)

	H-PE20	H-PE3	H-PE6	E-PE6	H-PE22	E-PE22	E-PE7	E-PE14
H-PE20	0							
H-PE3	0	0						
H-PE6	0.0004	0.0034	0					
E-PE6	<b><i>0.0234</i></b>	<b><i>0.0276</i></b>	<b><i>0.024</i></b>	0				
H-PE22	0.0037	<i>0.0077</i>	0	<b><i>0.0244</i></b>	0			
E-PE22	<b><i>0.0246</i></b>	<b><i>0.0288</i></b>	<b><i>0.0359</i></b>	0.0033	<b><i>0.0263</i></b>	0		
E-PE7	<b><i>0.0316</i></b>	<b><i>0.036</i></b>	<b><i>0.0228</i></b>	0	<b><i>0.0314</i></b>	<i>0.0111</i>	0	
E-PE14	<b><i>0.0199</i></b>	<b><i>0.0238</i></b>	<b><i>0.0211</i></b>	0.0045	<b><i>0.0221</i></b>	0.0028	0.0005	0

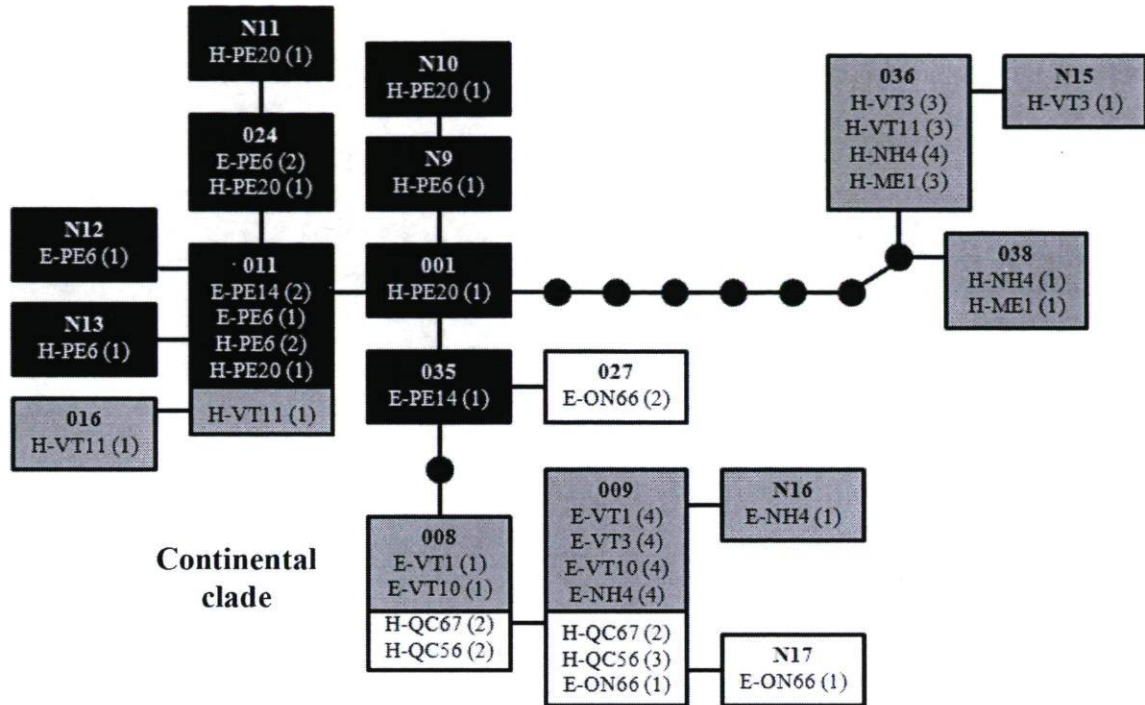
b)

	H-ME1	H-US11	H-NH4	E-NH4	E-VT3	H-VT3	E-US10	E-VT1
H-ME1	0							
H-US11	<b>0.0536</b>	0						
H-NH4	<b>0.0124</b>	<b>0.0315</b>	0					
E-NH4	<b>0.0059</b>	<b>0.026</b>	<b>0.0076</b>	0				
E-VT3	<b>0.0262</b>	0.0039	<b>0.0192</b>	0.0029	0			
H-VT3	<b>0.0613</b>	<b>0.0204</b>	<b>0.0506</b>	<b>0.0452</b>	<b>0.0224</b>	0		
E-US10	<b>0.0371</b>	<b>0.0149</b>	<b>0.0174</b>	0.0065	0.0003	<b>0.0397</b>	0	
E-VT1	<b>0.0368</b>	<b>0.0087</b>	<b>0.0204</b>	<b>0.0063</b>	0.0019	<b>0.0377</b>	0.0038	0

c)

	H-ON48	E-ON40	E-ON66	H-QC18	E-QC18	H-QC34	E-QC34	H-QC56	H-QC67	E-QC71
H-ON48	0									
E-ON40	<b>0.0296</b>	0								
E-ON66	<b>0.017</b>	<b>0.0102</b>	0							
H-QC18	<b>0.0388</b>	<b>0.0181</b>	<b>0.0498</b>	0						
E-QC18	<b>0.0514</b>	<b>0.0178</b>	<b>0.0436</b>	<b>0.0204</b>	0					
H-QC34	<b>0.0368</b>	<b>0.0251</b>	<b>0.05</b>	0.0018	<b>0.0082</b>	0				
E-QC34	<b>0.0268</b>	0	<b>0.0152</b>	0	0.0012	0.0052	0			
H-QC56	<b>0.0379</b>	<b>0.0388</b>	<b>0.0537</b>	<b>0.0127</b>	<b>0.0273</b>	<b>0.0076</b>	<b>0.0158</b>	0		
H-QC67	<b>0.0192</b>	<b>0.008</b>	<b>0.0206</b>	0.0081	<b>0.0107</b>	<b>0.0061</b>	0	<b>0.0139</b>	0	
E-Q571	<b>0.0642</b>	<b>0.0348</b>	<b>0.053</b>	<b>0.029</b>	0.0043	<b>0.0219</b>	0	<b>0.0243</b>	<b>0.0209</b>	0

## Appendix IV : MtDNA haplotype network



**Figure A1** MtDNA haplotype network relating sequences of *E. hageni* (H) and *E. ebrivum* (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.

## Conclusion

### *Divergence locale et répartition spatiale en mosaïque*

La similarité génétique observée précédemment entre les espèces morphologiques *Enallagma hageni* et *E. ebrium* était au cœur de cette étude. Il y a quelques années, Turgeon et coll. (2005) ont mis en lumière plusieurs aspects intrigants de l'histoire évolutive récente des *Enallagma spp.* Il était déjà connu que les nombreuses espèces issues des deux radiations partageaient du polymorphisme ancestral à un locus mitochondrial, mais leur nouvelle phylogénie à partir de marqueurs nucléaires révélait maintenant une différenciation pour la plupart des espèces. Étrangement, certaines espèces morphologiquement très distinctes, dont *Enallagma hageni* et *E. ebrium*, se sont révélées très similaires génétiquement. Afin de mieux comprendre les raisons d'une telle ressemblance génétique, une investigation en profondeur s'imposait, et ce fut le point de départ du présent projet de recherche.

Afin de discerner les raisons de la similarité génétique entre *E. hageni* et *E. ebrium*, nous avons opté pour une approche combinant des analyses génétiques et spatiales. D'un côté, les analyses des patrons de partition de la variation génétique, d'isolement par distance et de groupements génétiques ont révélé une différenciation génétique associée aux espèces morphologiques restreinte à l'échelle régionale puisque ce signal génétique ne s'est pas reflété à l'échelle globale. Parmi les différentes hypothèses que nous avons proposées, celle de la divergence locale semble la plus probable. D'un autre côté, les analyses spatiales des répartitions relatives des deux espèces morphologiques ont été très informatives. Ces analyses ont montré la prédominance de lacs allopatriques au niveau local, et, à plus grande échelle, une mosaïque d'ilots où l'une ou l'autre des espèces morphologiques domine. L'interférence reproductive entre *E. hageni* et *E. ebrium* couplée à la recolonisation lors de la dernière glaciation pourrait expliquer l'établissement et le maintien de cette répartition spatiale en mosaïque. Les résultats des analyses génétiques et spatiales, lorsque considérés ensemble, nous permettent d'entrevoir certains aspects de l'histoire évolutive de *E. hageni* et *E. ebrium*. La divergence locale récente observée entre ces deux espèces morphologiques est

probablement maintenant possible en raison de leur répartition spatiale en mosaïque, un contexte spatial qui doit aussi ralentir l'atteinte d'une relation monophylétique.

### *L'importance du contexte spatial*

La variation génétique observée entre différentes populations est influencée par des processus écologiques et évolutifs ayant une composante spatiale; ainsi, l'analyse conjointe de données génétiques et spatiales permet une meilleure compréhension de ces processus (Guillot et coll. 2009). Dans la présente étude, les analyses génétiques utilisées comportaient une composante spatiale en soi. En effet, la structure hiérarchique des différents sites d'échantillonnage (plusieurs sites dans trois régions) nous a permis de comparer la variation génétique à plusieurs échelles spatiales et ainsi de mettre en lumière la relation polyphylétique entre *E. hageni* et *E. ebrium*. Cette observation n'aurait pas été possible avec une analyse restreinte à une région unique et rappelle l'importance de vérifier la variation intraspécifique à différentes échelles spatiales. Il est aussi à noter que nous n'avons pas présenté de phylogénie dans cette étude afin de supporter la relation polyphylétique inférée. La solidité des quelques arbres phylogénétiques réalisés, bien qu'ils permettaient généralement de grouper adéquatement les populations en régions puis en espèces morphologiques, n'était pas suffisante (faible valeur de bootstrap). Ceci pourrait être expliqué par le faible niveau de différenciation génétique observé entre *E. hageni* et *E. ebrium* au sein d'une région et l'échelle spatiale différente intra- et inter-région.

La comparaison du patron d'isolement par distance (IBD) intra- et interspécifique à différentes échelles spatiales est une approche à notre connaissance rarement utilisée. Les patrons d'IBD sont observés dans plusieurs populations naturelles puisque la distance de migration d'un individu est beaucoup plus petite que l'aire géographique de l'espèce, et que la reproduction est souvent plus fréquente entre les individus plus près géographiquement. Classiquement, la caractérisation des patrons d'IBD a été utilisée au sein d'une même espèce uniquement. Dans notre étude, l'analyse des comparaisons intra- et interspécifique s'est révélée très efficace pour faire ressortir la différenciation génétique entre *E. hageni* et *E. ebrium* à petite échelle par leur contraste à différentes échelles spatiales. Ainsi, nous

avons quantifié la distance à laquelle la différenciation interspécifique ne répondait pas uniquement à un patron d'IBD. Ce résultat est d'autant plus intéressant maintenant que l'on sait que les *E. hageni* et *E. ebrium* forment une mosaïque d'ilots et que la distance trouvée pourrait être associée à la taille des régions de différenciation. Cette caractérisation du patron d'IBD intra- et interspécifique par la vérification du comportement des résidus à différentes échelles spatiales pourrait être utilisée dans l'étude de différents contextes complexes évolutifs comme les cas de divergence de jeunes espèces ou de spéciation en parallèle.

Les analyses spatiales effectuées dans cette étude ont permis d'aller au-delà des connaissances générales sur les répartitions spatiales globalement chevauchantes de *E. hageni* et *E. ebrium* et de mieux comprendre leurs interactions interspécifiques. L'étude de la répartition relative de ces espèces morphologiques à l'échelle locale et régionale a confirmé ce que plusieurs chercheurs semblaient observer sur le terrain : *E. hageni* et *E. ebrium* se retrouvent rarement en sympatrie et ils forment une mosaïque d'ilots où l'une ou l'autre domine. La répartition spatiale relative d'espèces apparentées est couramment étudiée dans les zones hybrides afin d'établir la répartition des espèces parentales. La méthode d'interpolation par kriging utilisée dans notre étude est d'ailleurs grandement inspirée d'une analyse semblable pour la zone hybride des araignées du groupe *Tegenaria atrica* au Royaume-Uni (Croucher et coll. 2007). Dans ce système, l'analyse de leur répartition spatiale relative a permis de calculer la largeur de la zone de tension. Ce type d'analyse aurait pu être utilisé dans notre système pour calculer la superficie des ilots, mais nous ne pensions pas avoir des données adéquates pour le faire. En effet, bien que nous ayons beaucoup de données en provenance des musées, l'abondance et la répartition de ces données n'étaient pas égales sur le territoire, et, de ce fait, la délimitation des ilots semblait être influencée par l'effort d'échantillonnage. Par contre, les deux étapes de validation que nous avons utilisées confortaient les résultats trouvés. Nous avons tout d'abord comparé les résultats obtenus avec les données d'échantillonnage et celle de musée pour vérifier la qualité de ces dernières. Puis, nous avons fait la même chose avec notre base de données finale et une répartition aléatoire de *E. hageni* et *E. ebrium* pour s'assurer de la validité de la répartition interpolée. Cette validation nous a donc donné confiance dans le patron

observé puisque nous savons qu'il n'est pas un artéfact du type de données ou de leur répartition. Ces étapes de validation sont critiques pour la bonne interprétation des résultats, et sont malheureusement absentes des études précédentes. Les musées regorgent de données qui pourraient être utilisées dans d'autres études similaires, à condition qu'il y ait des étapes de validation.

### *Perspectives*

Les interactions reproductives entre *E. hageni* et *E. ebrium* semblent avoir joué un rôle central dans leur histoire évolutive. La reproduction interspécifique n'avait pas été observée lors de tests de croisements précédents (Fincke 2007; M.A. McPeck et J. Turgeon, données non publiées), un phénomène qui semble se refléter dans nos résultats génétiques en raison d'un niveau de différenciation similaire dans les sites allopatriques et sympatriques. Nous avons soulevé dans cette étude un autre aspect reproductif intéressant, soit l'interférence reproductrice qui pourrait avoir eu un grand rôle à jouer dans l'établissement du patron de répartition spatiale en mosaïque. Les conséquences de ce type d'interaction reproductrice sont peu connues et gagnent à être étudiées. Notre système, comme celui des crickets *Tetrix ceperoi* et *T. subalata* (Gröning et coll. 2007; Hochkirch et coll. 2007), est bon un système pour l'étudier. Il serait intéressant de vérifier la symétrie et la force de cette interaction reproductrice.

La question du nombre d'apparitions du second morphe a été soulevée dans la dernière section de la discussion du Chapitre 1. Bien que nous n'ayions pas testé directement pour cette question, nous avons discuté de deux hypothèses possibles, soit une apparition unique ou multiple. Dans le futur, il serait intéressant d'aller chercher réponse à cette question par l'analyse de marqueurs génétiques non neutres. Par exemple, en trouvant de marqueurs associés aux gènes associés à la morphologie des appendices annaux, il pourrait être possible d'inférer l'espèce ancestrale par la comparaison de ces marqueurs entre les espèces et entre les régions. Ainsi, nous pourrions savoir si la morphologie des appendices annaux de *E. hageni* et *E. ebrium* est associée à différents génotypes dans différentes régions. Ce

sont des avenues qui sont regardées présentement dans le laboratoire de M.A. McPeck à Dartmouth College par l'étude de transcriptomes.

Les thèmes de la spéciation et de la délimitation des espèces étaient sous-entendus dans la question de distinction génétique entre *E. hageni* et *E. ebrium*. Ces deux espèces morphologiques semblent maintenant avoir acquis plusieurs propriétés biologiques importantes les élevant au rang de bon candidat au statut d'espèces distinctes : elles sont morphologiques différentes, elles s'excluent reproductivement en raison d'un fort isolement prézygotique et elles se retrouvent rarement ensemble sur le territoire. Par contre, il semble qu'elles n'ont pas encore atteint la monophylie. À long terme, tout dépendant de la migration entre les îlots dominés par la même espèce morphologique, le signal polyphylétique pourrait devenir monophylétique. Malgré ces incertitudes, *E. hageni* et *E. ebrium* semblent représenter des métapopulations évoluant indépendamment, le critère au cœur de la définition d'espèce de de Queiroz (2007).

En conclusion, cette étude a permis de mieux comprendre l'histoire évolutive des deux espèces morphologiques *E. hageni* et *E. ebrium*. Les deux approches utilisées dans cette étude, soit génétique et spatiale, ont permis de mettre en lumière l'importance insoupçonnée de l'interférence reproductive. Tout porte à croire que l'interférence reproductive est en cause dans l'établissement et le maintien de leur répartition spatiale en mosaïque. Ainsi géographiquement séparées dans leurs îlots respectifs, *E. hageni* et *E. ebrium* ont pu évoluer en lignées indépendantes. L'ajout d'une vision spatiale à la structure génétique polyphylétique de ces deux espèces morphologiques, en plus de l'intégration de connaissances antérieures sur leurs interactions reproductives, nous a permis de mieux cerner leur histoire évolutive unique.



## Bibliographie

- Albertson, R. C., J. A. Markert, P. D. Danley et T. D. Kocher. 1999. Phylogeny of a rapidly evolving clade: The cichlid fishes of Lake Malawi, East Africa. *Proceedings of the National Academy of Sciences of the United States of America* 96: 5107-5110.
- Aljanabi, S. M. et I. Martinez. 1997. Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Research* 25: 4692-4693.
- Applied Biosystems. 2007-2010. AFLP® Plant Mapping protocol.
- Bernatchez, L., S. Renaut, A. R. Whiteley, N. Derome, J. Jeukens, L. Landry, G. Q. Lu et coll. 2010. On the origin of species: insights from the ecological genomics of lake whitefish. *Philosophical Transactions of the Royal Society B-Biological Sciences* 365: 1783-1800.
- Bialozyt, R., B. Ziegenhagen et R. J. Petit. 2006. Contrasting effects of long distance seed dispersal on genetic diversity during range expansion. *Journal of Evolutionary Biology* 19: 12-20.
- Bolnick, D. I. et B. M. Fitzpatrick. 2007. Sympatric speciation: Models and empirical evidence. *Annual Review of Ecology Evolution and Systematics* 38: 459-487.
- Bonin, A., E. Bellemain, P. B. Eidesen, F. Pompanon, C. Brochmann et P. Taberlet. 2004. How to track and assess genotyping errors in population genetics studies. *Molecular Ecology* 13: 3261-3273.
- Bridle, J. R., S. J. E. Baird et R. K. Butlin. 2001. Spatial structure and habitat variation in a grasshopper hybrid zone. *Evolution* 55: 1832-1843.
- Brown, J. M., M. A. McPeck et M. L. May. 2000. A phylogenetic perspective on habitat shifts and diversity in the north American *Enallagma* damselflies. *Systematic Biology* 49: 697-712.
- Butlin, R. K., J. Galindo et J. W. Grahame. 2008. Sympatric, parapatric or allopatric: the most important way to classify speciation? *Philosophical Transactions of the Royal Society B: Biological Sciences* 363: 2997-3007.
- Catling, P. M. 2001. Morphological evidence for the hybrid *Enallagma ebrium* x *hageni* (Zygoptera: Coenagrionidae) from Ontario. *Proceedings of the Entomological Society of Ontario* 132: 99-100.
- Corbet, P. S. 1999, Dragonflies : behavior and ecology of Odonata. New-York, *Comstock Pub. Associates*.

- Coyne, J. A. et H. A. Orr. 2004. Speciation. Sunderland, Massachusetts, *Sinauer Associates, Inc.*
- Cressie, N. A. C. 1993. Statistics for spatial data. New York, *J. Wiley & Sons.*
- Croucher, P. J. P., R. M. Jones, J. B. Searle et G. S. Oxford. 2007. Contrasting patterns of hybridization in large house spiders (*Tegenaria atrica* group, Agelenidae). *Evolution* 61: 1622-1640.
- de Queiroz, K. 1998. The general lineage concept of species, species criteria, and the process of speciation: A conceptual unification and terminological recommendations, Pages 57-77 in D. J. Howard, and S. H. Berlocher, eds. *Endless forms: species and speciation*. New York, Oxford University Press.
- de Queiroz, K. 2005. Ernst Mayr and the modern concept of species. *Proceedings of the National Academy of Sciences of the United States of America* 102: 6600-6607.
- de Queiroz, K. 2007. Species concepts and species delimitation. *Systematic Biology* 56: 879-886.
- Dieckmann, U. et M. Doebeli. 1999. On the origin of species by sympatric speciation. *Nature* 400: 354-357.
- Duchesne, P., C. Etienne et L. Bernatchez. 2006. PERM: a computer program to detect structuring factors in social units. *Molecular Ecology Notes* 6: 965-967.
- Evanno, G., S. Regnaut et J. Goudet. 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology* 14: 2611-2620.
- Excoffier, L. et H. E. L. Lischer. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* 10: 564-567.
- Falush, D., M. Stephens et J. K. Pritchard. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes* 7: 574-578.
- Fincke, O. M. 1982. Lifetime mating success in a natural-population of the damselfly, *Enallagma-hageni* (Walsh) (Odonata, Coenagrionidae). *Behavioral Ecology and Sociobiology* 10: 293-302.
- Fincke, O. M., A. Fargevieille et T. D. Schultz. 2007. Lack of innate preference for morph and species identity in mate-searching *Enallagma* damselflies. *Behavioral Ecology and Sociobiology* 61: 1121-1131.
- Fitzpatrick, B. M., J. A. Fordyce et S. Gavrilets. 2008. What, if anything, is sympatric speciation? *Journal of Evolutionary Biology* 21: 1452-1459.

- Fitzpatrick, B. M., J. A. Fordyce et S. Gavrilets. 2009. Pattern, process and geographic modes of speciation. *Journal of Evolutionary Biology* 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. et A. Hochkirch. 2008. Reproductive interference between animal species. *Quarterly Review of Biology* 83: 257-282.
- Gröning, J. et A. Kocum. 2005. Dornschröcken (Orthoptera, Tetrigidae) auf den Ostfriesischen Inseln—Verbreitung, Koexistenz und Ökologie. *Entomologie heute* 17: 47-63.
- Gröning, J., N. Lücke, A. Finger et A. Hochkirch. 2007. Reproductive interference in two ground-hopper species: testing hypotheses of coexistence in the field. *Oikos* 116: 1449-1460.
- Guillot, G., R. Leblois, A. Coulon et A. C. Frantz. 2009. Statistical methods in spatial genetics. *Molecular Ecology* 18: 4734-4756.
- Halverson, K., S. B. Heard, J. D. Nason et J. O. Stireman, III. 2008. Origins, distribution, and local co-occurrence of polyploid cytotypes in *Solidago altissima* (Asteraceae). *American Journal of Botany* 95: 50-58.
- Hart, M. W. 2010. The species concept as an emergent property of population biology. *Evolution*: accepté.
- Hendry, A. P. 2009. Ecological speciation! Or the lack thereof? *Canadian Journal of Fisheries and Aquatic Sciences* 66: 1383-1398.
- Hewitt, G. M. 2004. Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society B: Biological Sciences* 359: 183-195.
- Hey, J. 2006. On the failure of modern species concepts. *Trends in Ecology & Evolution* 21: 447-450.
- Hochkirch, A., J. Gröning et A. Bückler. 2007. Sympatry with the devil: reproductive interference could hamper species coexistence. *Journal of Animal Ecology* 76: 633-642.
- Howard, D. J. et R. G. Harrison. 1984. Habitat segregation in ground crickets: The role of interspecific competition and habitat selection. *Ecology* 65: 69-76.

- Hubisz, M. J., D. Falush, M. Stephens et J. K. Pritchard. 2009. Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources* 9: 1322-1332.
- Ibrahim, K. M., R. A. Nichols et G. M. Hewitt. 1996. Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. *Heredity* 77: 282.
- Jensen, J. L., A. J. Bohonak et S. T. Kelley. 2005. Isolation by distance, web service. *Bmc Genetics* 6.
- Johannesson, K. 2001. Parallel speciation: a key to sympatric divergence. *Trends in Ecology & Evolution* 16: 148-153.
- Johnson, D. M. et P. H. Crowley. 1980. Habitat and seasonal segregation among coexisting odonate larvae. *Odonatologica* 9: 297-308.
- Karlsson, K., F. Eroukhmanoff, R. Härdling et E. I. Svensson. 2010. Parallel divergence in mate guarding behaviour following colonization of a novel habitat. *Journal of Evolutionary Biology* 23: 2540-2549.
- Kirkpatrick, M. et V. Ravigné. 2002. Speciation by Natural and Sexual Selection: Models and Experiments. *The American Naturalist* 159: S22-S35.
- Kuno, E. 1992. Competitive-exclusion through reproductive interference. *Researches on Population Ecology* 34: 275-284.
- Losos, J. B. et R. E. Glor. 2003. Phylogenetic comparative methods and the geography of speciation. *Trends in Ecology & Evolution* 18: 220-227.
- M'Gonigle, L. K. et R. G. FitzJohn. 2010. Assortative mating and spatial structure in hybrid zones. *Evolution* 64: 444-455.
- Mallet, J., A. Meyer, P. Nosil et J. L. Feder. 2009. Space, sympatry and speciation. *Journal of Evolutionary Biology* 22: 2332-2341.
- Martin, G. 1996. Taxonomy - Birds in double trouble. *Nature* 380: 666-667.
- Mayden, R. L. 1997. Hierarchy of species concepts: the denouement in the saga of the species problem, Pages 381-384 in M. F. Claridge, H. A. Dawah, and M. R. Wilson, eds. *Species: The units of biodiversity*. London, Chapman and Hall.
- Mayr, E. 1942, Systematics and the origin of species. New York, *Columbia University Press*.
- McPeck, M. A. 1989. Differential dispersal tendencies among *Enallagma* damselflies (Odonata) inhabiting different habitats. *Oikos* 56: 187-195.
- McPeck, M. A. 1990a. Behavioral-differences between *Enallagma* species (Odonata) influencing differential vulnerability to predators. *Ecology* 71: 1714-1726.

- McPeck, M. A. 1990b. Determination of species composition in the *Enallagma* damselfly assemblages of permanent lakes. *Ecology* 71: 83-98.
- McPeck, M. A. 1995. Morphological evolution mediated by behavior in the damselflies of two communities. *Evolution* 49: 749-769.
- McPeck, M. A. 1998. The consequences of changing the top predator in a food web: A comparative experimental approach. *Ecological Monographs* 68: 1-23.
- McPeck, M. A. 1999. Biochemical evolution associated with antipredator adaptation in damselflies. *Evolution* 53: 1835-1845.
- McPeck, M. A. 2000. Predisposed to adapt? Clade-level differences in characters affecting swimming performance in damselflies. *Evolution* 54: 2072-2080.
- McPeck, M. A. et J. M. Brown. 2000. Building a regional species pool: Diversification of the *Enallagma* damselflies in eastern North America. *Ecology* 81: 904-920.
- McPeck, M. A., A. K. Schrot et J. M. Brown. 1996. Adaptation to predators in a new community: Swimming performance and predator avoidance in damselflies. *Ecology* 77: 617-629.
- McPeck, M. A., L. Shen, J. Z. Torrey et H. Farid. 2008. The tempo and mode of three-dimensional morphological evolution in male reproductive structures. *American Naturalist* 171: E158-E178.
- McPeck, M. A., L. Shen et H. Farid. 2009. The correlated evolution of three-dimensional reproductive structures between male and female damselflies. *Evolution* 63: 73-83.
- Miller, R. S. 1963. Pattern and process in competition, Pages 1-74 in J. B. Cragg, ed. New York, Advances in ecological researches, Academic Press.
- Nichols, R. A. et G. M. Hewitt. 1994. The genetic consequences of long-distance dispersal during colonization. *Heredity* 72: 312-317.
- Panhuis, T. M., R. Butlin, M. Zuk et T. Tregenza. 2001. Sexual selection and speciation. *Trends in Ecology & Evolution* 16: 364-371.
- Parker, G. A. et L. Partridge. 1998. Sexual conflict and speciation. *Philosophical Transactions of the Royal Society B-Biological Sciences* 353: 261-274.
- Paulson, D. R. 1974. Reproductive isolation in damselflies. *Systematic Zoology* 23: 40-49.
- Pilon, J.-G. et D. Lagacé. 1998. Les odonates du Québec. Chicoutimi, *Entomofaune du Québec*.
- Pritchard, J. K., M. Stephens et P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959.

- Rasanen, K. et A. P. Hendry. 2008. Disentangling interactions between adaptive divergence and gene flow when ecology drives diversification. *Ecology Letters* 11: 624-636.
- Ray, N. et L. Excoffier. 2010. Inferring processes during introduction and range expansion: A first step towards inferring levels of long-distance dispersal during past expansions. *Molecular Ecology Resources* 10: 902-914.
- Rice, W. R. et E. E. Hostert. 1993. Laboratory experiments on speciation: What have we learned in 40 years? *Evolution* 47: 1637-1653.
- Robertson, H. M. et H. E. H. Paterson. 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. *Molecular Ecology Notes* 4: 137-138.
- Rundell, R. J. et T. D. Price. 2009. Adaptive radiation, nonadaptive radiation, ecological speciation and nonecological speciation. *Trends in Ecology & Evolution* 24: 394-399.
- Rundle, H. D., L. Nagel, J. W. Boughman et D. Schluter. 2000. Natural Selection and Parallel Speciation in Sympatric Sticklebacks. *Science* 287: 306-308.
- Rundle, H. D. et P. Nosil. 2005. Ecological speciation. *Ecology Letters* 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. New York, *Oxford University Press*.
- Schluter, D. 2001. Ecology and the origin of species. *Trends in Ecology & Evolution* 16: 372-380.
- Schluter, D. 2009. Evidence for ecological speciation and its alternative. *Science* 323: 737-741.
- Schluter, D. et L. M. Nagel. 1995. Parallel speciation by natural selection. *The American Naturalist* 146: 292-301.
- Shen, L., H. Farid et M. A. McPeck. 2009. Modeling three-dimensional morphological structures using sperical harmonics. *Evolution* 63: 1003-1016.
- Siepielski, A. M., K. L. Hung, E. E. B. Bein et M. A. McPeck. 2010. Experimental evidence for neutral community dynamics governing an insect assemblage. *Ecology* 91: 847-857.
- Sobel, J. M., G. F. Chen, L. R. Watt et D. W. Schemske. 2010. The biology of speciation. *Evolution* 64: 295-315.

- Stoks, R., M. A. McPeck et J. L. Mitchell. 2003. Evolution of prey behavior in response to changes in predation regime: Damselflies in fish and dragonfly lakes. *Evolution* 57: 574-585.
- Suarez, A. V. et N. D. Tsutsui. 2004. The value of museum collections for research and society. *Bioscience* 54: 66-74.
- Turgeon, J. et M. A. McPeck. 2002. Phylogeographic analysis of a recent radiation of *Enallagma* damselflies (Odonata : Coenagrionidae). *Molecular Ecology* 11: 1989-2001.
- Turgeon, J., R. Stoks, R. A. Thum, J. M. Brown et M. A. McPeck. 2005. Simultaneous Quaternary radiations of three damselfly clades across the Holarctic. *American Naturalist* 165: E78-E107.
- Vekemans, X., T. Beauwens, M. Lemaire et I. Roldán-Ruiz. 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. *Molecular Ecology* 11: 139-151.
- Via, S. 2001. Sympatric speciation in animals: the ugly duckling grows up. *Trends in Ecology & Evolution* 16: 381-390.
- Walker, E. M. 1953. The Odonata of Canada and Alaska, *University of Toronto Press*.
- Webster, R. et M. A. Oliver. 1990. Statistical methods in soil and land resource survey. Oxford, *Oxford University Press*.
- Westfall, M. J. et M. L. May. 1996. Damselflies of North America, *Scientific publishers*.
- Westfall, M. J. et M. L. May. 2006. Damselflies of North America, revised edition, *Scientific publishers*.
- Zink, R. M. 1996. Bird species diversity. *Nature* 381: 566-566.