Thèse de doctorat de l'Université de Lille

Structure génétique spatiale, flux de gènes et tailles efficaces des populations chez l'Agrion de Mercure (*Coenagrion mercuriale, Zygoptera*)

Préparée à l'UMR 8198, unité Évolution, Écologie et Paléontologie (Evo-Eco-Paléo) et à Office Génie Écologique (OGE)

Ecole doctorale n°104, Science de la Matière, du Rayonnement et de l'Environnement

Thèse soutenue à Villeneuve d'Ascq, le 2 Juillet 2024, par

Agathe Lévêque

Pour obtenir le grade de Docteur en
Biologie de l'environnement, des organismes, des populations, écologie

Membres du jury :

Pierre Saumitou-Laprade, Directeur de recherche CNRS, Université de Lille	Président du jury
Thomas Broquet, Chargé de recherche CNRS, Station Biologique de Roscoff	Rapporteur
Christoph Haag, Directeur de recherche CNRS, Centre d'Écologie Fonctionnelle et Évolutive	Rapporteur
Frédéric Austerlitz, Directeur de recherche CNRS, Muséum national d'Histoire naturelle	Examinateur
Fabienne Van Rossum, Research scientist, Jardin Botanique de Meise	Examinatrice
Jean-François Arnaud, Professeur, Université de Lille	Directeur de thèse
Anne Duputié, Maître de conférences, Université de Lille	Co-encadrante de thèse
Vincent Vignon, Directeur de projet, Alkios	Invité Co-encadrant de thèse

Remerciements

Je tiens à remercier toutes les personnes qui m'ont entouré et ont contribué à l'aboutissement de ce projet de thèse. Cette aventure de trois ans (et quelques mois) n'aurait pas été possible sans l'enthousiasme et la bienveillance des gens que j'ai pu rencontrer avec qui j'ai pu travailler de près ou de loin au cours de ce voyage de thèse, le long des cours d'eau avec mes petits agrions de mercure. Même s'il est difficile de trouver les mots justes, ce sont des sentiments bien personnels qui me motivent à écrire ces lignes pour vous exprimer toute ma gratitude et vous remercier pour tous ces bons moments partagés ensemble.

Tout d'abord, je tiens à remercier grandement les sociétés SOCOS et ARCOS d'avoir financé ma thèse à travers une collaboration CIFRE entre le bureau d'étude Office de Génie Écologique (OGE) et le laboratoire Évolution Écologie et Paléontologie (EEP), ce qui m'a permis travailler durant ces années de thèse dans d'excellentes conditions et de manière indépendante.

Je remercie Christophe Haag et Thomas Broquet d'avoir accepté de rapporter cette thèse, Fabienne Van Rossum et Frédéric Austerlitz d'avoir accepté de faire partie du jury, ainsi que Pierre Saumitou-Laprade d'avoir accepté de présider ce jury. Je remercie également les membres de mes comités de thèse, Éric Petit, Frédéric Austerlitz et Xavier Vekemans pour leurs questions et leurs conseils avisés, qui m'ont permis de faire avancer ma réflexion au cours de cette thèse.

Un immense merci à mes encadrants, Jean-François, Anne et Vincent. Je ne les remercierai jamais assez pour la confiance qu'ils m'ont accordée tout au long de cette aventure, parfois sinueuse, qu'est la thèse, ainsi que pour leur accompagnement sans faille pendant ces années, qui m'ont permis de grandir tant scientifiquement qu'humainement. Merci à vous trois pour votre bonne humeur, vos conseils à chaque étape, que ce soit pour préparer et réaliser les campagnes de terrain, échanger sur la progression du projet et discuter des idées, mais aussi pour les nombreuses relectures…

Jean-François, durant ces années de thèse, j'ai pu compter sur ta disponibilité, ta pédagogie et ta patience indéfectibles, je te remercie donc autant pour le partage de tes connaissances et ton expérience que pour le cadre de confiance et de sérénité instauré autour de ma thèse. J'ai énormément apprécié ta bonne humeur constante et ton soutien lors des passages plus périlleux de la thèse (les pol'potes, les bières et le chocolat n'y sont pas pour rien dans la finalisation de ce manuscrit). *Du coup, techniquement*, c'est aussi grâce à ton humour et tes plaisanteries que j'ai pu avancer dans ce projet de thèse passionnant ;)

Anne, je te remercie pour ton implication, tes conseils, ton aide, pour ton dynamisme et ta curiosité sans limites tout au long de ces années de thèse. Tu as accordé une place importante à la qualité des relations humaines durant toute cette aventure scientifique et je te remercie de m'avoir donné l'opportunité de travailler à tes côtés et de poursuivre durant ces derniers mois de thèse pour finir sur ce beau projet.

Vincent, cette thèse n'aurait pas été possible sans ta participation, je tiens à te remercier de m'avoir accueilli au sein d'OGE ainsi que pour ta bienveillance et ta gentillesse. Même si tu étais géographiquement plus éloigné, je te remercie pour la richesse de nos échanges, pour tes remarques pertinentes et tes conseils, ce qui m'a permis de ne pas me focaliser sur les problématiques techniques du projet, mais de toujours revenir au sens « biologique » de ce que nous faisions.

À la source du projet, un grand merci à Cédric Vanappelghem, le « pape de l'agrion » pour ta bonne humeur et tes précieux conseils et pour avoir partagé avec moi tes connaissances sur notre petit agrion de mercure.

Je remercie aussi l'ensemble des membres du CEN qui ont participé aux échantillonnages de pattes d'agrions réalisés avant le début de la thèse dans les Hauts-de-France, m'assurant de belles analyses dans la région et plus de sérénité dans mon échantillonnage. Je remercie également l'association IMAGO pour les échanges que nous avons eus et l'accès aux données d'occurrence de l'agrion de mercure en Alsace. Enfin je tiens à remercier Sylvain Santoni pour ses précieux conseils (notamment le fameux « That's life » qui m'a marquée) concernant les manips génomiques.

Merci à l'Université de Lille ainsi que l'école doctorale SMRE pour le support et l'accompagnement des démarches administratives des doctorants et plus particulièrement Christophe Van Brussel pour ta bienveillance et ta patience.

Merci à toute l'équipe OGE de m'avoir accueillie pendant ces trois ans, où j'ai pu découvrir la vie dans ce bureau d'étude en écologie auprès de naturalistes passionnés. Cela a été une expérience professionnelle très enrichissante, que ce soit lors des réunions à distance, lors de mes rares déplacements à Paris ou lors des moments partagés avec mes collègues sur les terrains en Alsace à la recherche de la fameuse « marmotte de Strasbourg » lors des campagnes de suivi de Hamster géant *Cricetus*. Toute ma gratitude va à Olivier et Caroline pour leur support administratif.

Le laboratoire EEP a constitué un cadre de travail extrêmement agréable, me permettant d'évoluer dans d'excellentes conditions tout au long de ces années de thèse. Je voudrais remercier notamment Cécile et Clément qui ont contribué plus particulièrement à ce projet de thèse et avec qui j'ai eu énormément de plaisir à travailler.

Cécile, merci beaucoup pour ta bonne humeur, ta pédagogie et ton investissement. Grâce à toi, les manips de biologie moléculaire se sont déroulées de manière efficace et dans les meilleures conditions possibles, bien que la biomol reste toujours pour moi un domaine à mi-chemin entre la grande cuisine et la magie. Merci aux filles du plateau biomol Laurence, pour ses blagues et sa bonne humeur, mais aussi pour son soutien technique lors des manips, Christelle pour sa gentillesse et ses précieux conseils de génomiques et Anne-Cat' pour son énergie et sa spontanéité.

Merci beaucoup Clément, pour ta patience, ta bienveillance et pour ton aide précieuse pour la réalisation des analyses bio-informatiques, sans ton aide et au soutien technique du plateau bio-informatique que vous formez avec Mathieu, je n'y serai pas arrivé ! Même s'il me reste encore pas mal de marge de progression dans les méandres des codes sur le cluster, j'ai beaucoup appris à vos côtés (dire qu'au début je voulais ouvrir un vcf de données génomiques sous Excel, un comble !).

À mes stagiaires, tout d'abord Fabien qui m'a accompagné pour deux stages, un grand merci pour ton aide précieuse, ton investissement, ton énergie et ta bonne humeur que ce soit sur le terrain en équipe avec Louise (et malgré les allergies) ou en manip au labo. Merci beaucoup à Anaëlle, Clément et Hana, la dream team terrain, pour votre curiosité, votre sérieux et votre efficacité qui ont fait de 2022 une super expérience de terrain.

Je remercie aussi l'ensemble de l'équipe du labo, sans pouvoir citer toutes les belles personnes que j'ai rencontrées, pour leur accueil chaleureux digne du NORD et leur bienveillance (pour m'avoir supporté) durant ces trois années.

Merci à Isabelle, Eléonore, Anne, Zoé, François, Fabien et Guillaume pour les services d'enseignement réalisés à leurs côtés.

Merci à toi Sandrine qui m'a permis de réaliser toutes mes missions dans les meilleures conditions possibles face l'administration parfois capricieuse et à Xavier pour le suivi et pour ses conseils sur les points clés à valider pour le doctorat.

Aux « grands » du labo, je sais la chance que j'ai eu d'intégrer votre équipe dans un environnement composé de gens passionnés et bienveillants qui font à travers les discussions et les moments partager ensemble, de ce labo une maison où il fait bon vivre. Je remercie tous ceux avec qui j'ai bien rigolé et discuté en salle de pause. Mes remerciements vont aussi particulièrement Pierre (notre phare dans la nuit, je ne sais pas ce qu'on ferait sans lui) et Philippe pour leur présence et leurs discussions passionnées sur les plantes qui ont rendu mon début de thèse en fin de déconfinement, beaucoup moins solitaire.

Aux « jeunes EE », je remercie l'ensemble de la team des doctorants pour ces supers moment passés à vos côtés, pour les bières et jeux de société de la MDE, le mystère des poèmes de la salle de pause, les blagues et pour toutes les activités réalisées … difficile de citer tous ces bons moments

partagés ensembles. Plus que de simples collègues de travail j'ai eu la chance de passer ces années de thèse avec des amis.

Aux plus anciens des jeunes : Audrey, tu étais un peu la maman des doctorants merci pour tous tes super conseils, Zoé merci pour ces bons moments, notamment lors des TP ACP on a bien géré quand même.

A ma génération : Émilie merci pour ta gentillesse et ta bonne humeur (eh oui c'est le début du bonheur), François notre King des doctorants, merci pour ton humour et le fait qu'on ait rarement été d'accord ce qui a donné lieu à de nombreux débats, Flavia merci pour ta bonne humeur permanente, tous tes délires et tes blagues comme tu l'as si bien dit « on aura bien rigolé ! ».

Aux jeunes : un très grand merci à Justine, merci d'avoir partagé avec moi ta passion des bourdons, merci pour ton amitié, ton sens de l'humour et toutes nos discussions, tu resteras ma co-bureau ! et si tu vois une abeille mellifère ..… tu comprendras ;) Merci à Guillaume, pour toutes les petites attentions et la gentillesse (la douceur du panda n'est pas accessible à qui veut) que tu as pour nous tous, tes amis doctorants, Claire merci pour ton esprit joueur et pétillant et pour ton humour noir. Aux p'tits nouveaux, avec qui j'ai pu boire des bières à la MDE, merci beaucoup, Camille, Achille, Fabien et Muskaan pour ces derniers mois de thèse partagés avec vous. Et merci à tout ceux que je n'ai pas pu citer ici par manque de place.

À l'ensemble de mes proches qui m'ont accompagné et qui m'ont aidé à « déconnecter » un peu, merci. Merci à mes parents Antoine et Maryline, mes premiers supporters, pour votre amour qui m'a porté jusque-là et à mon frère Matthieu et à Eva avec qui j'ai passé de super moments et de super vacances pendant ces trois années. Merci à ma famille du Nord, Nicolas, Corinne, Félix, Édouard et Florentine pour avoir marqué avec moi les grands évènements pendants ces trois années, merci aussi à nos deux poilus Nounours et Mimichat qui ont rendu ces années plus douces. Un grand merci à mon super cousin Thomas, qui par son sens de l'humour, sa répartie et son esprit vif m'a permis de me changer les idées à chaque fois que je rentrais à la maison.

Je voudrais terminer ces remerciements en dédiant cette thèse à ma grand-mère Anita qui a amplement participé (parfois malgré elle) à ce travail de thèse, merci pour ton soutien indéfectible durant toutes ces années, merci d'avoir pris soin de moi avec tes bons petits plats qui m'ont permis de prendre des forces (et un peu de poids), pour ton écoute (surtout quand je râlais un peu) et tes précieux conseils, sans toi jamais ces années de thèse n'auraient pu être aussi douces et agréables, MERCI MAMIE !

Table des matières

Introduction générale

I. Organisation spatiale de la diversité génétique

1) La diversité génétique comme composante de la biodiversité

La diversité du vivant peut être exprimée à différentes échelles, toutes interconnectées. On distingue trois niveaux auxquels cette diversité peut être observée et étudiée : les niveaux écosystémiques, spécifiques et génétiques, reconnus par la Convention sur la diversité biologique de 1992 (Dirzo & Raven, 2003; Figure 1). Au niveau le plus large, les écosystèmes se caractérisent par des assemblages d'espèces retrouvés de manière récurrente dans l'espace, avec par exemple les écosystèmes tempérés forestiers ou de prairies. Toutefois, au-delà de la simple diversité des communautés d'espèces retrouvée à cette échelle, le terme d'écosystème intègre aussi de manière plus large les processus et les relations écologiques qu'entretiennent les espèces entre elles et avec leur environnement. Ainsi, le terme de biodiversité inclut de nombreuses définitions (Jax, 2007). Les espèces retrouvées dans ces écosystèmes correspondent, quant à elles, à des lignées généalogiques séparées par des barrières à la reproduction. Ces dernières sont distinguées et classifiées à partir de données taxonomiques fondées sur des critères morphologiques, écologiques et/ou génétiques (Padial *et al.*, 2010; Puillandre *et al.*, 2012). Toutefois, au sein d'une même espèce, l'ensemble des individus ne sont pas identiques, mais sont chacun porteurs de combinaisons génétiques et phénotypiques uniques.

Ces variations génétiques, quant à elles, constituent le niveau le plus fin de diversité du monde vivant, et sont à l'origine de l'ensemble de la biodiversité que l'on peut rencontrer. Elles expliquent non seulement les différences observées entre individus d'une même espèce, mais aussi entre individus appartenant à des populations distinctes. L'expression de cette variabilité génétique, combinée à des facteurs environnementaux, s'exprime à travers le phénotype des individus, pouvant impliquer par exemple la couleur, la morphologie générale, la phénologie, ou encore la résistance immunitaire. Contraction de « diversité biologique » proposée par Walter Rosen en 1986, le terme de biodiversité englobe ainsi l'ensemble des variations du vivant. Or, depuis le sommet de la Terre à Rio de Janeiro en 1992, la préservation et la conservation de la biodiversité sont devenues une préoccupation majeure (Wilson, 1988; Dirzo & Raven, 2003).

La biodiversité est actuellement menacée par des facteurs à la fois déterministes et stochastiques (Caughley, 1994). En effet, cinq grandes menaces liées aux activités humaines impactant la biodiversité ont été identifiées (Figure 1) : la destruction et la fragmentation des habitats (Fahrig, 2003; Mimura *et al.*, 2017), la pollution des environnements, comprise au sens large (Peñuelas *et al.*, 2013), le changement climatique (Kannan & James, 2009; Prakash, 2021), l'introduction d'espèces exotiques envahissantes (Vitousek *et al.*, 1996; Vuillaume *et al.*, 2015) et la surexploitation (Pinsky & Palumbi, 2014). Ces menaces anthropiques altèrent les interactions entre espèces et avec leurs environnements. Ces menaces entraînent aussi un déclin des effectifs des populations et espèces, bouleversent les patrons

de diversité génétique des populations et modifient l'aire de distribution géographique des espèces (Figure 1).

Toutefois, à ces impacts déterministes s'ajoutent des facteurs stochastiques affectant déjà les populations, avec des variations aléatoires, démographiques, génétiques, mais aussi environnementales (Figure 1).

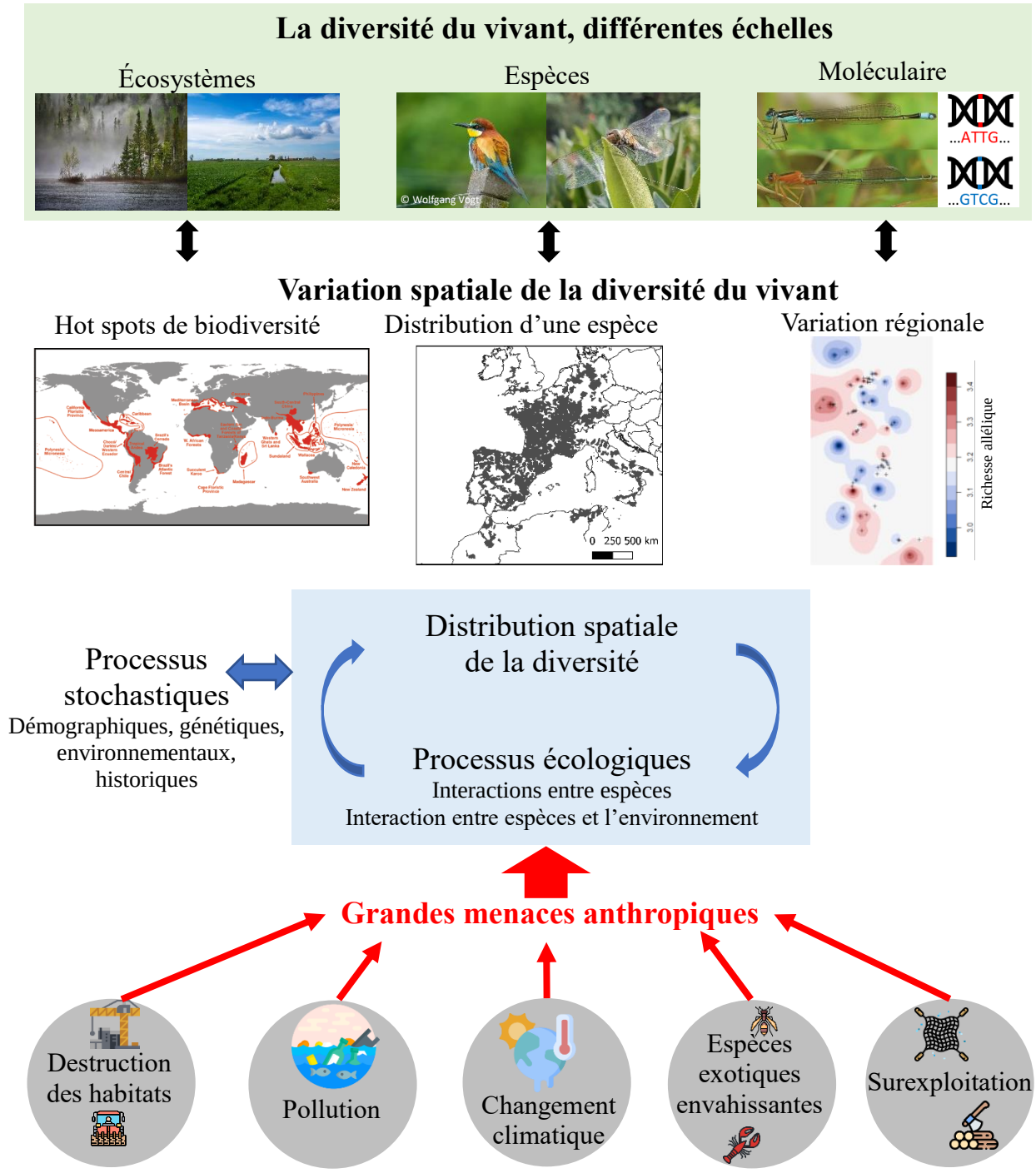


Figure 1. *Schéma simplificateur des niveaux d'organisation du vivant, variation spatiale de la diversité du vivant et présentation des cinq menaces pesant sur la biodiversité. Carte des Hot spots de la biodiversité (Myers et al., 2000).*

2) Comment préserver et surveiller les niveaux de biodiversité ?

L'échelle à laquelle les priorités de conservation doivent être menées fait l'objet de nombreux débats (Bowen, 1999). Jusqu'à présent, les efforts de conservation ont généralement mis l'accent sur la protection des espèces (Bowen, 1999; Armsworth *et al.*, 2007). Or, la conservation d'une espèce nécessite une approche équilibrée tenant compte non seulement de l'histoire et de la viabilité des espèces, mais aussi d'une protection de leur biotope et biocénose. En effet, il est illusoire d'envisager la protection d'une espèce sans protéger les processus écologiques ainsi que l'ensemble des facteurs biotiques et abiotiques nécessaire à sa persistance, aussi bien que de conserver des écosystèmes sans prendre en compte les espèces qui les composent. De plus, la considération d'écosystèmes et d'assemblages d'espèces permet d'identifier des points chauds ou « hot spots » de biodiversité, lieux de radiations évolutives d'où la biodiversité du futur est susceptible de provenir. Or, ces « hot spots » de biodiversité ne sont pas répartis de manière aléatoire à l'échelle de la planète (Myers *et al.*, 2000; Miraldo *et al.*, 2016; Figure 1).

En outre, l'identification et la délimitation des espèces peuvent être controversées. En effet, à travers l'aire de distribution géographique d'une espèce, les populations peuvent être soumises à des contraintes environnementales différentes entraînant des évolutions adaptatives uniques (Funk *et al.*, 2016). De plus, au sein d'une même espèce, l'on peut retrouver une très grande variabilité en termes de diversité génétique et de probabilité d'extinction locale de populations. Certains modèles et observations suggèrent d'ailleurs que la position centrale ou périphérique de groupes d'individus à travers l'aire de distribution influence leurs niveaux de diversité génétique et de connectivité, et détermine ainsi leur sensibilité à des phénomènes d'extinction et de recolonisation (Eckert *et al.*, 2008).

Néanmoins, la variation génétique au sein des espèces n'est souvent pas considérée dans les plans de gestion des applications de conservation (Laikre, 2010; Garner *et al.*, 2020). Or, afin de protéger la biodiversité, celle-ci doit être prise en compte à travers l'ensemble de ses niveaux de complexité (Allendorf *et al*., 2022). Il a donc été suggéré que les efforts de conservation devraient se concentrer davantage sur l'extinction de populations génétiquement distinctes, et moins sur l'extinction d'espèces (Hughes *et al.*, 1997; Hobbs & Mooney, 1998). Ainsi, un quatrième niveau de biodiversité composé de populations locales génétiquement distinctes serait nécessaire pour assurer la survie des espèces à long terme et la stabilité des écosystèmes. Ce dernier niveau permet alors l'identification d'unités de conservation pertinentes d'un point de vue génétique et adaptatif. En outre, la protection de populations génétiquement distinctes est essentielle pour maximiser un certain potentiel évolutif et minimiser ainsi le risque d'extinction locale (Hughes *et al.*, 1997; Luck *et al.*, 2003; Funk *et al.*, 2012). Enfin, la perte ou le rétablissement de variabilité génétique dans les populations influencent la diversité des communautés et les fonctions des écosystèmes ainsi que leur résilience, élevant alors l'entité populationnelle au statut d'unité fonctionnelle de de conservation (Whitham *et al.*, 2008; Des Roches *et al.*, 2021).

3) La répartition de la diversité génétique des populations, le résultat de 4 forces micro-évolutives majeures

La diversité du vivant ne se répartit pas aléatoirement à la surface de la planète, mais implique des points chauds de biodiversité, des distributions géographiques non aléatoires de chaque espèce, ainsi que des variations notables sur une échelle régionale (Figure 1). Des processus stochastiques ainsi que les processus écologiques entre les espèces et avec leur environnement, engendrés par des patrons distincts de biodiversité, influencent également la distribution spatiale de la diversité du vivant. De manière plus fine, l'organisation géographique de la diversité génétique des populations est gouvernée par l'action de quatre forces micro-évolutives majeures : la mutation, la sélection naturelle, la migration et la dérive génétique (Bergstrom & Dugatkin, 2016; Figure 2).

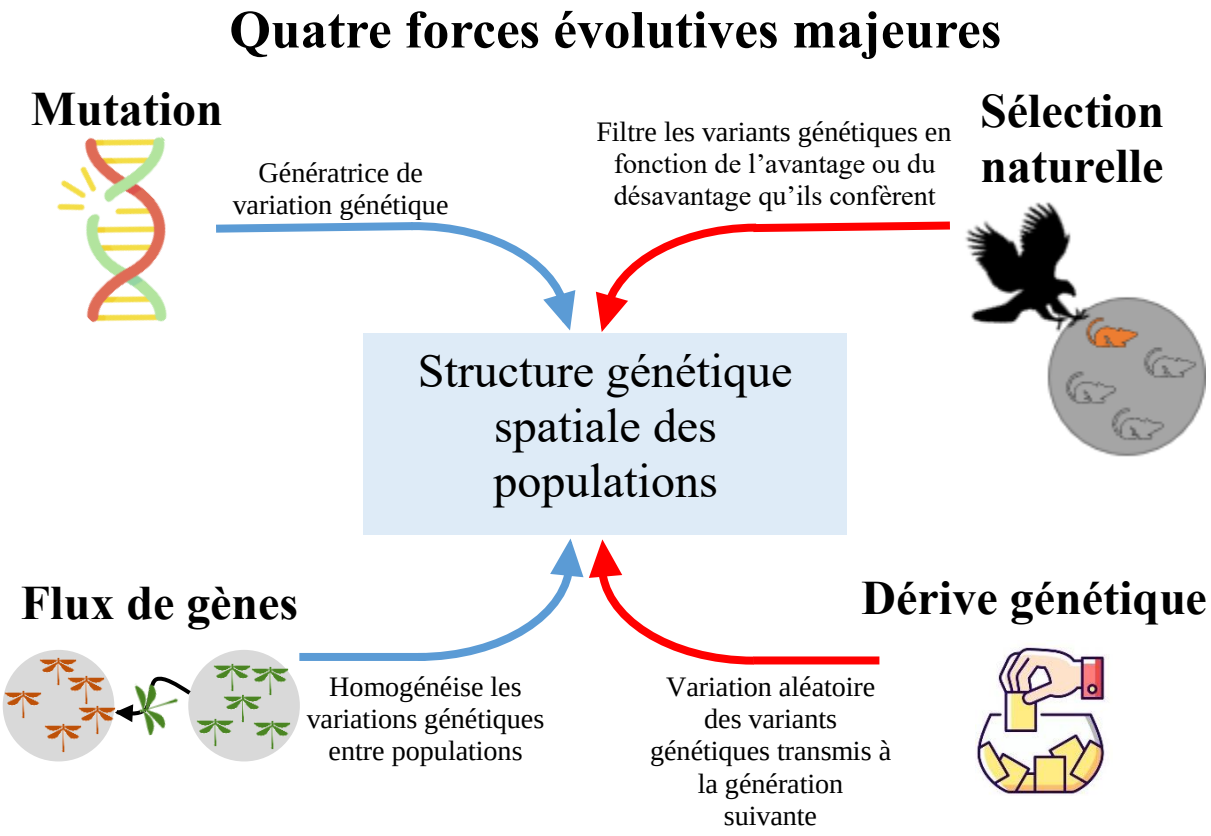


Figure 2. *Schéma général des quatre forces micro-évolutives à l'origine des niveaux de diversité génétique neutre et adaptative et de la structure génétique spatiale des populations. Les forces évolutives créant de la diversité génétique intra-population sont indiquées par les flèches bleues tandis que les forces évolutives provoquant une érosion de cette diversité génétique portent des flèches rouges. Les icônes de cette figure sont issues du site internet suivant : https://www.flaticon.com/fr/*

La mutation est induite par des modifications du matériel génétique, impliquant par exemple les changements, la suppression ou l'insertion de nucléotides, ainsi que les recombinaisons et les inversions de séquences d'ADN. Cette force évolutive est génératrice de diversité à travers l'apparition aléatoire de

nouveaux variants génétiques (Figure 2). Cette force évolutive crée aléatoirement le substrat sur lequel la sélection naturelle pourra éventuellement agir (Futuyma & Kirkpatrick, 2018). Ainsi, la plupart des variants génétiques (ou allèles) aléatoirement générés par la mutation n'apportent pas d'avantage ou de désavantage à la survie et au succès reproducteur, le produit de ces deux traits déterminant la valeur sélective des individus. On parle alors de variation génétique neutre. Toutefois, certains variants génétiques et associations génotypiques peuvent modifier la valeur sélective des individus sous certaines conditions environnementales, on parle alors de variation génétique adaptative, pouvant donner lieu à une évolution également d'ordre adaptatif.

La sélection naturelle, seul processus micro-évolutif déterministe et incluant également la notion de sélection sexuelle, va alors agir comme un filtre lié au succès différentiel des génotypes dans leur contribution à la génération suivante (Bergstrom & Dugatkin, 2016; Futuyma & Kirkpatrick, 2018). Ceci va, de génération en génération, modifier la fréquence des variants observés, avec une augmentation de la fréquence de ceux conférant un avantage sélectif, et au contraire une réduction de la fréquence des variants désavantageux à un temps donné dans un environnement donné. Par ailleurs, la différenciation génétique entre populations adaptées localement peut être importante pour la persistance de ces populations : en effet, la migration d'individus non porteurs de variations adaptatives relatives à un environnement local peut présenter un effet négatif sur la survie des populations, on parle alors de dépression d'allofécondation ou d'« outbreeding depression » (Frankham *et al.*, 2017; Figure 2).

La dérive génétique, quant à elle, correspond à des fluctuations fortuites de la fréquence d'allèles entre générations. Ces fluctuations aléatoires résultent du simple fait que, dans une population de taille finie, l'ensemble des individus constituant une population ne participera pas d'égale manière à la génération suivante, indépendamment de la sélection naturelle. La taille de population, et plus particulièrement la taille efficace de population, constitue l'un des paramètres démographiques majeurs déterminant l'intensité de cette dérive génétique. Contrairement à la sélection qui agit comme un filtre déterministe, la dérive génétique représente le produit du tirage aléatoire des gamètes transmis d'une génération à une autre. Cette force évolutive peut ainsi à terme entraîner la fixation ou la disparition d'allèles au sein des populations, augmentant alors les niveaux de différenciation génétiques observés entre populations naturelles, tout en réduisant le niveau de diversité de chacune de ces populations. Cette force évolutive non déterministe aura d'autant plus d'impact que les populations concernées seront isolées géographiquement et de petites tailles (Figure 2).

Enfin, les flux de gènes entre populations sont un processus fondamental qui, sous réserve qu'ils soient suffisamment importants, vont homogénéiser les fréquences de gènes entre populations. Ces flux génétiques sont la conséquence de la dispersion des gamètes et des individus dans l'espace (Broquet & Petit, 2009). Leur intensité est donc intimement liée aux caractéristiques biologiques des espèces, tels que la capacité de dispersion des individus ou le système de reproduction et le régime d'appariement. Par ailleurs, les individus migrants évoluent aussi au sein d'une matrice environnementale qui peut faciliter ou au contraire restreindre ces événements de migration (Storfer *et al.*, 2010; Grafius *et al.*,

2017). Enfin, l'effet ultime des flux géniques permet de limiter ainsi l'isolement et la différenciation génétique entre populations conduisant à une certaine cohésion génétique des espèces (Slatkin, 1985; Frankham, 2015; Bergstrom & Dugatkin, 2016).

Dans le cadre de cette thèse, nous nous focaliserons sur l'influence de la dérive génétique et des flux de gènes sur la structure génétique des populations. L'étude des patrons spatiaux de variabilité génétique peut en effet permettre de comprendre les phénomènes qui déterminent la distribution de la diversité intra et inter-population d'une espèce biologique. Cette distribution peut être purement fortuite, liée à l'histoire évolutive des populations étudiées, mais elle peut aussi être associée à des différences d'ordre adaptatif qui n'ont pas été effacées par des facteurs homogénéisants tels que les flux de gènes (Slatkin, 1987; Marko & Hart, 2011; Ellegren & Galtier, 2016). Ces patrons peuvent aussi révéler des systèmes complexes de métapopulations[1], mais aussi des processus d'extinction et de recolonisation non aléatoires, dont les goulots d'étranglement résultants peuvent accentuer la différenciation génétique entre populations (Whitlock & McCauley, 1990; Haag *et al.*, 2005). Ainsi, la structure spatiale de la variation génétique des populations est un phénomène complexe, dont la mise en place dépendra aussi bien de pressions sélectives, de facteurs historiques et géographiques, que des traits d'histoire de vie de l'espèce considérée.

II. Fragmentation du paysage et isolement géographique : impact sur la structure génétique des populations

1) Fragmentation du paysage

Les changements d'occupation des sols liés aux activités anthropiques entraînent une fragmentation des habitats et constituent une menace majeure pour les populations et plus largement pour la biodiversité à tous ses niveaux (Haddad *et al.*, 2015; Wilson *et al.*, 2016). La fragmentation des habitats correspond au processus par lequel des continuums d'habitats favorables sont transformés en un morcellement d'habitats de surface réduite et isolés les uns des autres par des espaces non favorables à la survie, à la dispersion et/ou à la reproduction des individus (Fahrig, 2003). Cette fragmentation des habitats se traduit donc par deux conséquences dont il est souvent difficile de distinguer les effets respectifs : une réduction des surfaces d'habitat favorables, conduisant à la perte *sensu stricto* d'habitat, ainsi qu'un isolement des habitats restants, correspondant à la fragmentation au sens strict du terme (Fahrig, 2003; Ewers & Didham, 2006; Didham *et al.*, 2012, Figure 3).

À l'échelle des populations, la fragmentation des habitats entraîne une réduction des tailles de populations ainsi qu'une diminution des flux de gènes entre populations occupant les parcelles d'habitat restants. Cet isolement de populations, entraînant une réduction des flux de gènes entre populations,

[1] Métapopulations : ensemble de populations d'individus d'une même espèce séparées spatialement ou temporellement, mais interconnectées par la dispersion

peut conduire à une augmentation des niveaux de différenciation génétique entre populations. Les populations restantes, plus petites et isolées, vont alors faire face à des conséquences à la fois démographiques et génétiques, avec un risque de consanguinité accrue, une plus grande force de la dérive génétique ainsi qu'une perte de diversité génétique (Figure 3). Toutes ces conséquences peuvent conduire les populations à entrer dans ce qui est dénommé un « vortex d'extinction » (Frankham, 2015; Haddad *et al.*, 2015 ; Figure 4).

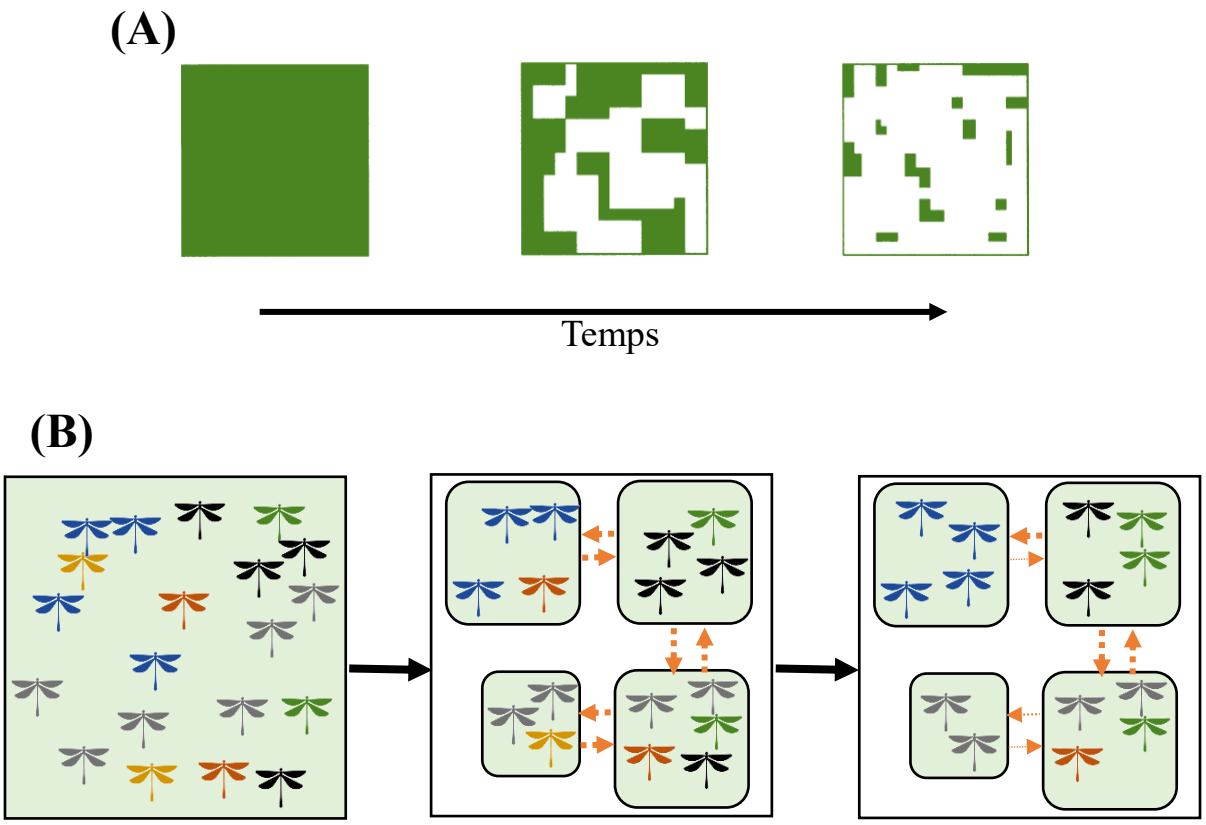


Figure 3. *Schématisation de l'impact de la fragmentation des habitats sur la diversité génétique et les flux de gènes entre populations. (A) Processus de fragmentation d'habitats, modifié à partir de Fahrig (2003). Les zones vertes représentent les habitats favorables et les zones blanches représentent la matrice d'espaces non favorables à la survie ou la dispersion des individus. On observe ici les effets combinés de la réduction de la taille des habitats, et de la perte de leur connectivité. (B) Schéma illustrant les effets de la fragmentation du paysage sur une population de libellules. La couleur des individus reflète les variants génétiques retrouvés au sein des populations. Les flèches orange représentent les niveaux de flux génique s'opérant entre populations.*

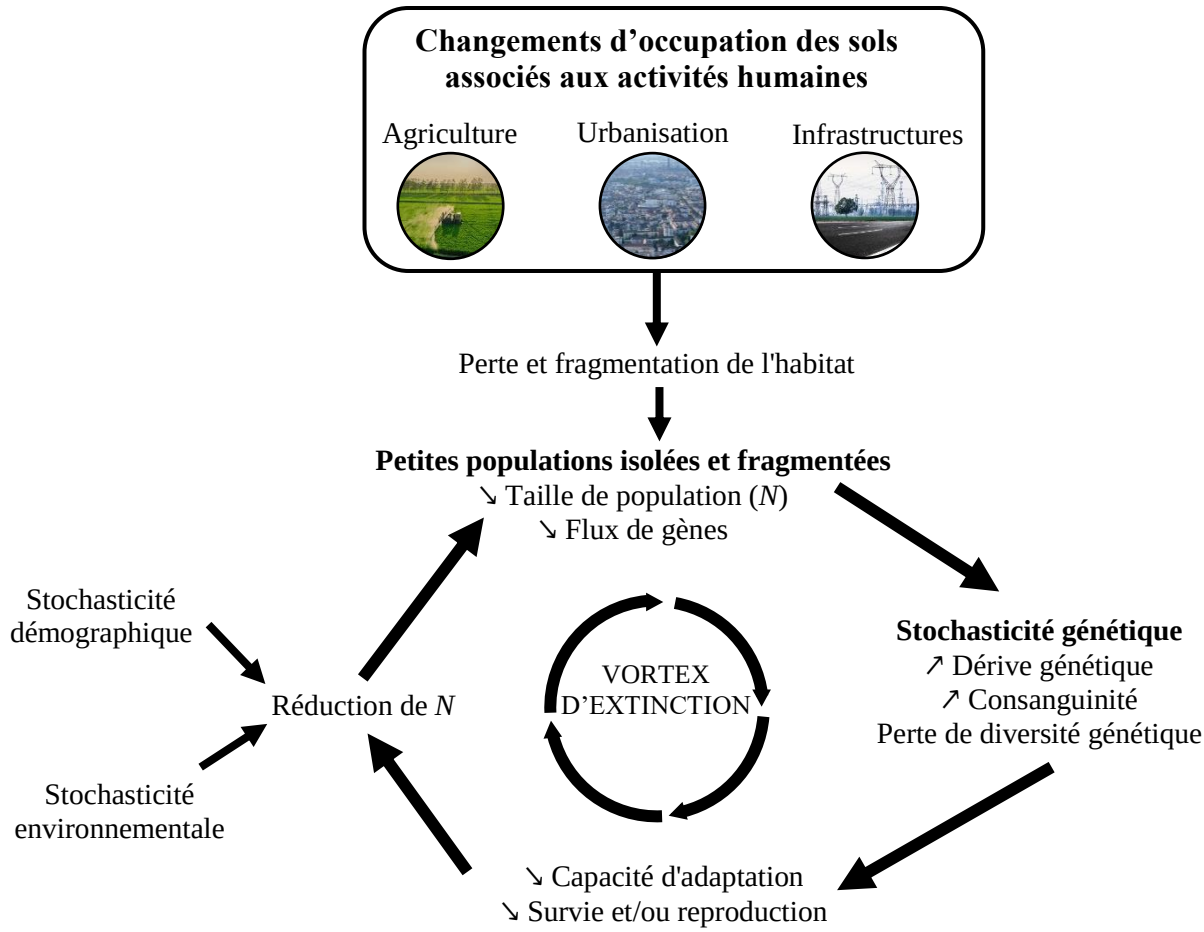


Figure 4. *Schéma illustrant le vortex d'extinction auquel les petites populations fragmentées et isolées peuvent être soumises, modifié à partir de Frankham et al., (2013).*

Ce type d'impact négatif a notamment été étudié dans le cadre de l'urbanisation des espaces naturels, la dégradation des habitats par la construction d'espaces urbains entraînant en effet une amplification de la dérive génétique et diminuant de fait les événements de flux de gènes entre les populations (voir la synthèse dans Johnson & Munshi-South, 2017a; Miles *et al.*, 2019). Toutefois, les conséquences des modifications de paysage peuvent être très différentes en fonction des traits d'histoire de vie des espèces, des types d'habitats rencontrés ainsi que de l'interaction entre les espèces considérées et ces nouveaux environnements (Miles *et al.*, 2019). En effet, les espèces à faible capacité de dispersion seront plus sensibles aux phénomènes de fragmentation urbaine, tandis que les espèces à forte capacité de dispersion pourront traverser plus librement les matrices paysagères défavorables, pouvant conduire à un impact neutre ou même facilitateur des changements paysagers sur l'espèce (voir pour exemple Richardson *et al.*, 2021). De plus, les modifications humaines du paysage peuvent aussi conduire à une augmentation indirecte des flux de gènes entre populations en créant de nouveaux corridors favorables à certaines espèces ou en impliquant le transport involontaire de gamètes ou d'individus (par exemple Miles *et al.*, 2018). Enfin, pour un même type de perturbation, les conséquences des changements d'occupation des sols peuvent varier d'une région à l'autre (Santangelo *et al.*, 2018; Miles *et al.*, 2019;

Diamond & Martin, 2021). Ainsi, il convient d'évaluer les caractéristiques génétiques des populations au cas par cas afin d'évaluer l'impact des modifications anthropiques d'occupation des sols sur une espèce donnée. De surcroît, certains paramètres de génétique des populations, tels que les niveaux de différenciation génétique et les niveaux de diversité génétique, peuvent répondre de manière très différente et éventuellement différée à une même contrainte paysagère (e.g. Trumbo *et al.*, 2019; Fusco *et al.*, 2021).

2) Diversité génétique intra-population

De nombreux estimateurs permettent de quantifier la variation génétique présente au sein des populations. Cette diversité génétique intra-populationnelle peut être estimée afin de comparer les niveaux de cette diversité entre différentes populations, mais aussi de suivre sur plusieurs pas de temps l'évolution d'une même population, ce qui permet de fournir des informations sur les goulots d'étranglement passés ou l'expansion démographique des populations, ou encore d'identifier une perte de diversité génétique accompagnée d'une augmentation de la consanguinité locale due à une fragmentation et/ou une perte d'habitat. Il s'agit ici d'un paramètre clé de persistance à long terme des populations. En effet, la diversité génétique intra-population est intimement liée au « potentiel évolutif » d'une population, c'est-à-dire sa capacité à éventuellement s'adapter face à des changements environnementaux (Frankham *et al.*, 2013; Ellegren & Galtier, 2016). Ci-dessous, nous allons brièvement décrire deux types d'estimateurs de diversité génétique couramment utilisés en génétique des populations.

La diversité génétique d'une population peut ainsi se mesurer au moyen de l'estimation des niveaux d'hétérozygotie. Pour une espèce diploïde, l'hétérozygotie observée permet de mesurer la proportion d'individus porteurs de deux allèles (ou variants génétiques) différents à un locus (localisation génétique) donné. Cette mesure peut également être moyennée sur l'ensemble des locus utilisés pour génotyper les individus. Ainsi, l'hétérozygotie observée (H_o) correspond à la fréquence réelle des individus hétérozygotes au sein d'une population. L'hétérozygotie attendue (H_e), aussi appelée diversité génétique (Nei, 1973), correspond quant à elle à la proportion d'hétérozygotes que l'on attendrait au sein d'une population panmictique, impliquant l'hypothèse d'une population de taille infinie avec un accouplement aléatoire des individus et sans intervention des quatre processus micro-évolutifs majeurs. Il s'agit ici de l'équilibre de Hardy-Weinberg. Cette mesure présente l'avantage d'être généralement peu sensible à la taille d'échantillon utilisée pour son estimation. En effet, quelques individus suffisent pour estimer ce paramètre H_e, si toutefois un grand nombre de locus sont examinés (Gorman & Renzi, 1979).

La richesse allélique (A_r) est une mesure complémentaire à l'hétérozygotie attendue pour l'estimation de la diversité génétique. En effet, cette dernière est plus sensible que l'hétérozygotie attendue à la perte de variation génétique due à une réduction de la taille d'une population (Allendorf *et al.*, 2022). Il s'agit ici d'une mesure du nombre d'allèles par locus qui utilise une méthode de raréfaction

pour minimiser les biais associés à la taille inégale des échantillons (El Mousadik & Petit, 1996; Petit *et al.*, 1998).

3) Niveau de consanguinité intra-population

Outre la diversité génétique, les niveaux de consanguinité des populations sont également des paramètres importants à estimer, car ils peuvent être directement liés aux régimes de reproduction des individus. Ces niveaux locaux de consanguinité peuvent être le reflet d'un isolement géographique prononcé et d'une faible taille de populations. La consanguinité est le fruit d'accouplements entre individus génétiquement apparentés. Ce type d'unions entre individus ayant une co-ascendance commune génère une augmentation de la charge d'allèles identiques par descendance hérités d'un ancêtre commun ; les niveaux d'homozygotie sont donc de fait augmentés, au détriment des niveaux d'hétérozygotie. Cette consanguinité peut être liée aux traits d'histoire de vie d'une espèce, soit par un phénomène d'homogamie dû au fait que les individus génotypiquement ou phénotypiquement semblables ont plus tendance à s'accoupler entre eux, soit du fait d'une compétition sexuelle intense où seuls quelques individus s'accaparent les évènements de reproduction, soit du fait d'un régime de reproduction permettant l'autofécondation (Hall *et al.*, 1994). Même lorsque la panmixie est vérifiée, une consanguinité locale importante peut se produire dans les petites populations au sein desquelles les individus sont génétiquement apparentés. Or, dans les populations naturelles, la dépression de consanguinité, c'est-à-dire la diminution de la valeur sélective des individus issus de croisements consanguins, peut contribuer à moyen terme à l'extinction locale de populations (Keller & Waller, 2002; Charlesworth & Willis, 2009; Frankham *et al.*, 2013; Spurgin & Gage, 2019; Figure 4).

4) Taille efficace des populations (N_e)

Si les estimations de niveaux de diversité génétique et de consanguinité locale permettent d'avoir un instantané de la structure génétique d'une population, la taille efficace d'une population (N_e) permet de mesurer le taux de perte de variation génétique de la population en raison de l'influence accrue de la dérive génétique et de la consanguinité (Nei & Tajima, 1981; Waples, 2005; Hare *et al.*, 2011; Wang *et al.*, 2016). Ce concept est lié à la relation directe entre le taux de changements génétiques et la taille d'une population tenant compte des facteurs pouvant influencer le nombre d'individus contribuant réellement à la génération suivante (Figure 5). Dans une population dite "idéale" sous un modèle de Wright-Fisher qui satisfait aux hypothèses d'accouplements aléatoires, d'une taille de population constante et de générations non chevauchantes (Fisher, 1930; Wright, 1931), le nombre d'individus de la population correspond à sa taille efficace. Cependant, les populations naturelles se conforment rarement à ce modèle car divers facteurs, par exemple des sex-ratios déséquilibrés ou des contributions inégales à la génération suivante, peuvent réduire le nombre d'individus participant à la formation de la génération suivante. Ainsi la taille efficace d'une population reste très souvent beaucoup plus faible que

la taille démographique, correspondant au nombre d'individus recensés de cette dernière (Waples, 2016).

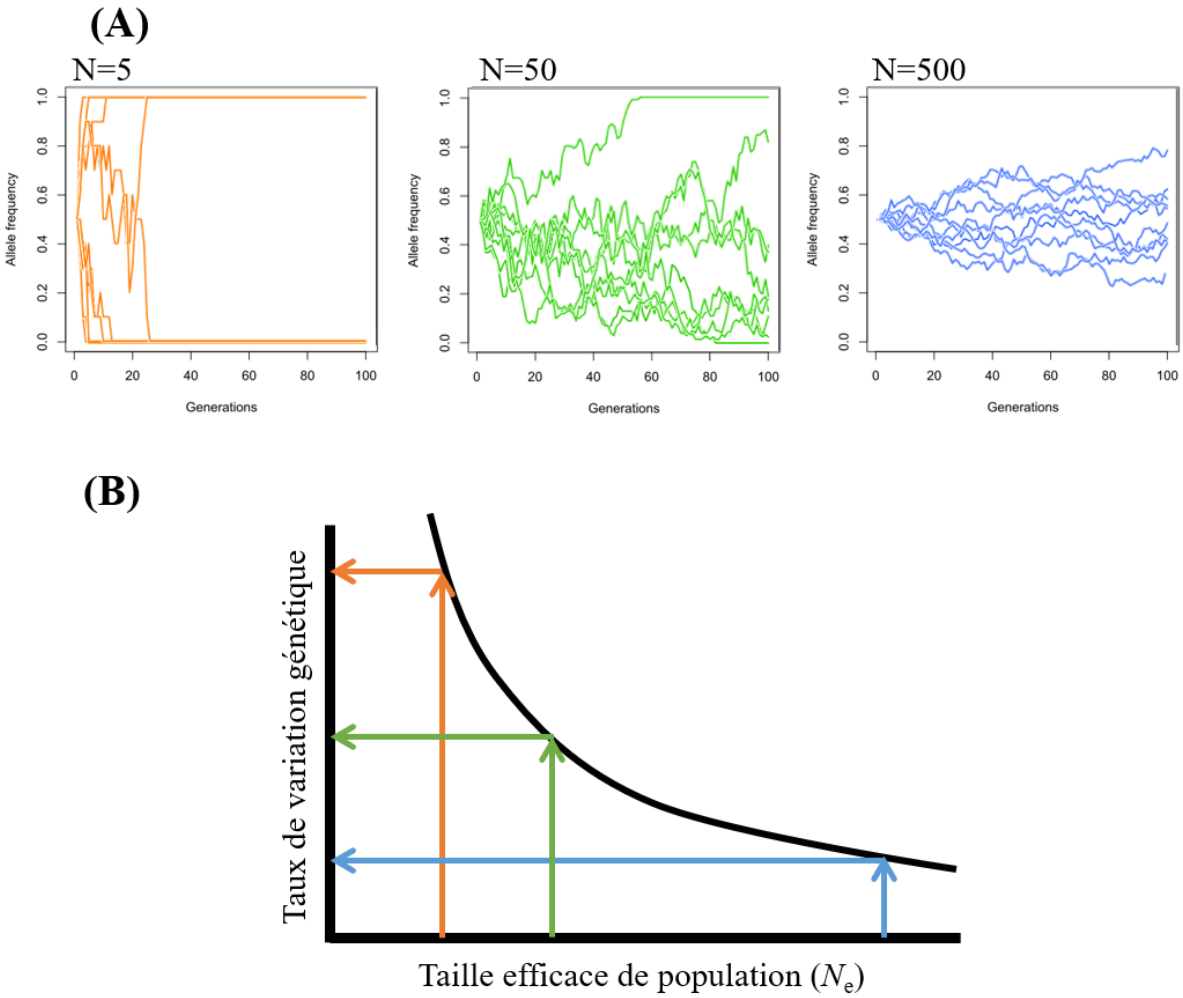


Figure 5. *Illustration des effets de la dérive génétique sur la taille efficace des populations (N_e). (A) Illustration du phénomène de dérive génétique impliquant des simulations avec 10 réplicas de la variation aléatoire des fréquences alléliques, avec une fréquence initiale de deux allèles de 0.5 et une taille de population de 5, 50 ou 500 individus. Avec une petite taille de population égale à 5, l'ensemble des simulations entraînent la perte d'un des deux allèles (fréquence = 0) et la fixation de l'autre (fréquence =1) à très court terme (moins de 30 générations). Cette issue est atteinte plus lentement à mesure que N augmente. (B) Figure représentant la relation entre le taux de variation génétique au sein des populations et leurs tailles efficaces. Cette figure est inspirée de l'exposé de Robin Waples, Effective population size (09/03/2015) National Marine Fisheries Service, NOAA. ConGen 2015.*

5) Estimation des flux de gènes s'opérant entre populations structurées

La mesure des flux de gènes contemporains entre populations peut permettre d'estimer les voies de migration et les niveaux de connectivité à long terme au sein d'un système en métapopulations structurées géographiquement. Ceci peut permettre d'aider à identifier les éléments facilitants ou restreignant la dispersion des individus. En effet, la dispersion et la reproduction d'un individu migrant d'une population à une autre vont favoriser l'homogénéisation de la diversité génétique présente au sein

de ces deux populations. Or dans les populations naturelles, les niveaux de différenciation génétique entre populations représentent un continuum de scénarii de migration. À un extrême, impliquant des accouplements complètement aléatoires, chaque individu a une probabilité égale de s'accoupler avec un autre individu, cette situation représente donc classiquement la panmixie (Figure 6A). L'autre extrémité du continuum de différenciation des populations est caractérisée par un isolement complet au sein de groupes d'individus distincts où la panmixie est également supposée (Waples *et al.*, 2018; Figure 6D). Les populations naturelles sont, quant à elles, caractérisées par un gradient de niveaux de différenciation intermédiaires, représentés par les scénarii B et C (Figure 6).

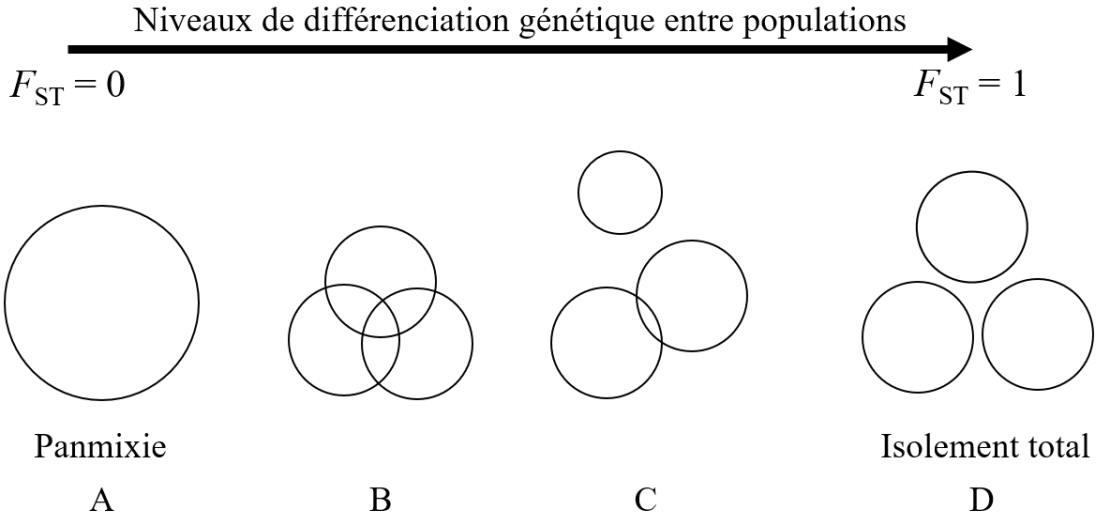


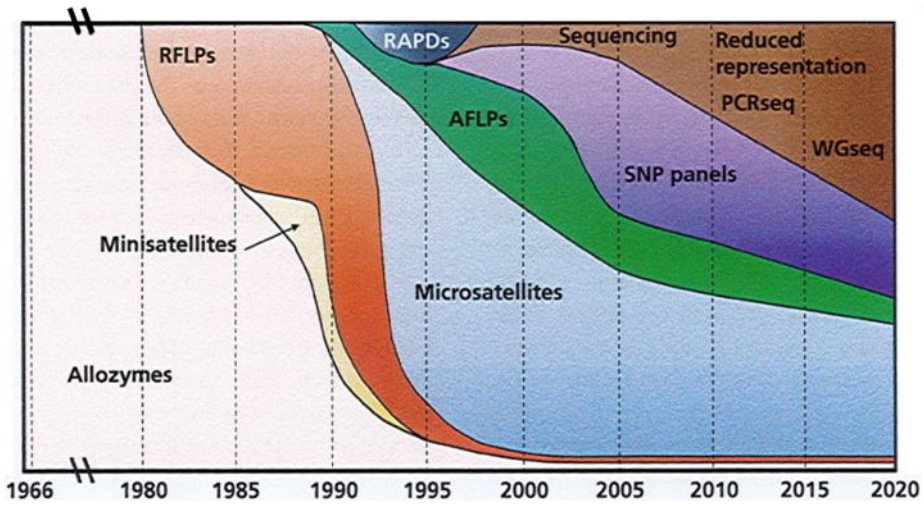
Figure 6. *Illustration du continuum de différenciation génétique des populations pouvant être observés en populations naturelles. Chaque cercle représente un groupe d'individus qui peuvent ou non constituer un groupe ou une population distincte. Quatre scénarii, avec différents degrés de connectivité impliquant un chevauchement géographique et/ou de la migration, sont identifiés sur le continuum : (A) Panmixie, (B) Connectivité substantielle, (C) Connectivité réduite, (D) Isolement complet. Cette figure est inspirée de Waples & Gaggiotti (2006).*

6) Bref historique du développement et de l'utilisation des marqueurs génétiques en biologie évolutive

En complément des approches dites « directes » permettant un suivi démographique des populations, impliquant pour exemple des suivis d'effectifs de populations (par exemple Couvreur *et al.*, 2008; Beaune & Sellier, 2021) ou des études de capture-marquage-recapture ou de radiométrie dédiées au suivi de la dispersion des individus (par exemple Purse *et al.*, 2003; Naveda-Rodríguez *et al.*, 2022), les approches génétiques représentent des outils de choix pour définir des mesures de conservation appropriées. En effet, les outils et concepts de génétique et génomique des populations permettent de répondre à de nombreuses questions centrales en biologie de la conservation, telles que, pour exemples, l'identification d'unités de conservation pertinentes, l'évaluation de la taille efficace de

population, la mesure de la connectivité réelle entre populations, ou encore l'évaluation de la diversité génétique locale des populations et de leur potentiel évolutif (Allendorf *et al.*, 2010; Funk *et al.*, 2012; Holderegger *et al.*, 2019; Hohenlohe *et al.*, 2021; Allendorf *et al*., 2022). La génétique et la génomique de la conservation permettent donc de fournir des informations sur l'histoire évolutive des populations, des instantanés de la structure génétique actuelle au sein et entre populations, ainsi qu'un aperçu du potentiel des populations à persister et à s'adapter aux changements environnementaux. Ceci fournit ainsi des informations essentielles à la conception de stratégies efficaces de gestion des espèces et populations (Allendorf *et al.*, 2010; Hohenlohe *et al.*, 2021). Dans ce contexte de génétique de la conservation, de nombreux marqueurs génétiques ont été développés au fil du temps afin de caractériser au mieux les populations naturelles (revues dans Davey *et al.*, 2011; Allendorf, 2017; Charlesworth & Charlesworth, 2017; Figure 7). Or, chaque type de marqueur génétique présente ses propres avantages et limites, associés entre autres à leurs coûts économiques, à la quantité d'ADN de départ nécessaire, aux équipements disponibles en laboratoire, à la complexité associée à leur développement et à leur utilisation, à leurs niveaux de reproductibilité, à la dominance ou codominance des marqueurs, à leur niveau de variabilité, ou encore au nombre de marqueurs assurant la précision des estimations réalisées (Schlötterer, 2004; Davey *et al.*, 2011; Allendorf, 2017). La Figure 7 présente une synthèse globale du type de marqueurs moléculaires utilisés depuis la découverte des allozymes dans les années 1960.

(A)



(B)

	Approche empirique	Avancée génétique en matière de conservation
1966	Allozymes	Description de l'ampleur de la variation génétique au sein des populations et entre elles Détection d'espèces apparentées
1979	ADN Mitochondrial	Phylogéographie Capacité à distinguer les flux de gènes via la voie mâle ou femelle
1990	Microsatellites	Plus de puissance pour décrire la structure des populations Détection de structuration génétique des populations Détection des goulets d'étranglement ou de l'expansion des populations Estimation de la taille efficace des populations Mesure de la perte de diversité génétique Association entre patrons génétiques et paysage
2000	SNPs	Plus grande capacité à décrire la structure de la population. Capacité à détecter les loci qui affectent la valeur sélective des individus
2010	Génomique des populations	Détection des régions génomiques sous sélection Détection d'adaptation locale des populations Estimation de la proportion du génome identique par descendance Détection et compréhension de la dépression de consanguinité Mesure de la perte de variation adaptative

Figure 7. *Illustration de l'évolution des types de marqueurs moléculaires utilisés en génétique de la conservation. (A) Évolution de l'importance relative au fil du temps des marqueurs moléculaires en génétique puis génomique de la conservation. RFLP, restriction fragment length polymorphism; AFLP, amplified fragment length polymorphism; RAPD, randomly amplified polymorphic DNA; SNP, single nucleotide polymorphism. Le panel SNP (violet) comprend le génotypage fondé sur du séquençage de type Sanger. Le séquençage (marron) est subdivisé en trois approches : PCRseq (séquençage Sanger d'un seul ou de quelques locus), séquençage par représentation réduite du génome (Reduced representation, principalement RADseq, mais aussi capture ciblée d'ADN et séquençage d'amplicons) ou reséquençage du génome entier (WGseq). Cette figure est issue de Allendorf et al. (2022), elle-même pour partie modifiée de Schlötterer (2004). (B) Chronologie des principales avancées dans l'utilisation des marqueurs génétiques neutres au sein de populations naturelles en biologie de la conservation, figure modifiée à partir d'Allendorf et al. (2010; 2017).*

Après les marqueurs allozymiques initialement développés dans les années 1960 et permettant une première description des variations génétiques observées entre et au sein des populations, les marqueurs mitochondriaux ont ensuite permis de mener des études phylogéographiques ainsi que

d’estimer les flux de gènes différentiels entre voies mâle et femelle (Prakash *et al.*, 1969; Avise *et al.*, 1979, 2016). Depuis les années 1990 et encore aujourd’hui, les marqueurs microsatellites, de par leur polymorphisme, codominance, reproductibilité et faible coût économique, permettent de décrire la variation génétique nucléaire des populations ainsi que les goulots d'étranglement passés et d'estimer la taille efficace actuelle des populations, ce qui explique leur large application en génétique des populations et en biologie de la conservation (Goldstein & Pollock, 1997; Frankham *et al.*, 2013; Allendorf *et al.*, 2022; Figure 7). Enfin, le développement des nouvelles technologies de séquençage, associé à l’essor des marqueurs de type SNP pour « single nucleotide polymorphism », permet de générer une quantité de données génomiques considérable à partir de pratiquement n'importe quel organisme (Morin *et al.*, 2004, 2012; Harrisson *et al.*, 2014; Allendorf, 2017). Ces marqueurs génomiques de type SNP s’ajoutent donc aux marqueurs neutres traditionnels de génétique de la conservation, leur abondance et leur large distribution à travers le génome offrant la promesse d'une plus grande précision et d'une plus grande puissance statistique dans l'estimation des paramètres de génétique des populations (Allendorf *et al.*, 2010; Narum *et al.*, 2013; Harrisson *et al.*, 2014). De plus, les marqueurs SNPs permettent de détecter les régions du génome qui sont affectées par la sélection naturelle et ainsi de répondre à de nouvelles questions concernant les mécanismes sous-jacents de l’évolution des populations et de leurs réponses adaptatives (Allendorf *et al.*, 2010; Hohenlohe *et al.*, 2021).

III. L’agrion de Mercure : un modèle d’étude particulièrement intéressant en biologie de la conservation

1) Description et caractéristiques de l’agrion de Mercure

L’agrion de Mercure (*Coenagrion mercuriale*) est un arthropode, de la classe des insectes et de l’ordre des odonates (Figure 8). Au sein des Odonates l’on peut distinguer deux sous-ordres : les Anisoptères (communément nommés Libellules) et les Zygoptères (communément nommés Demoiselles). À l’état adulte, les zygoptères sont caractérisés par leur aspect plus petit et fin que les anisoptères, leurs ailes antérieures et postérieures sont identiques, permettant aux individus de maintenir leurs ailes fermées au repos. De plus, leurs yeux sont distants et ne se touchent pas, contrairement aux anisoptères.

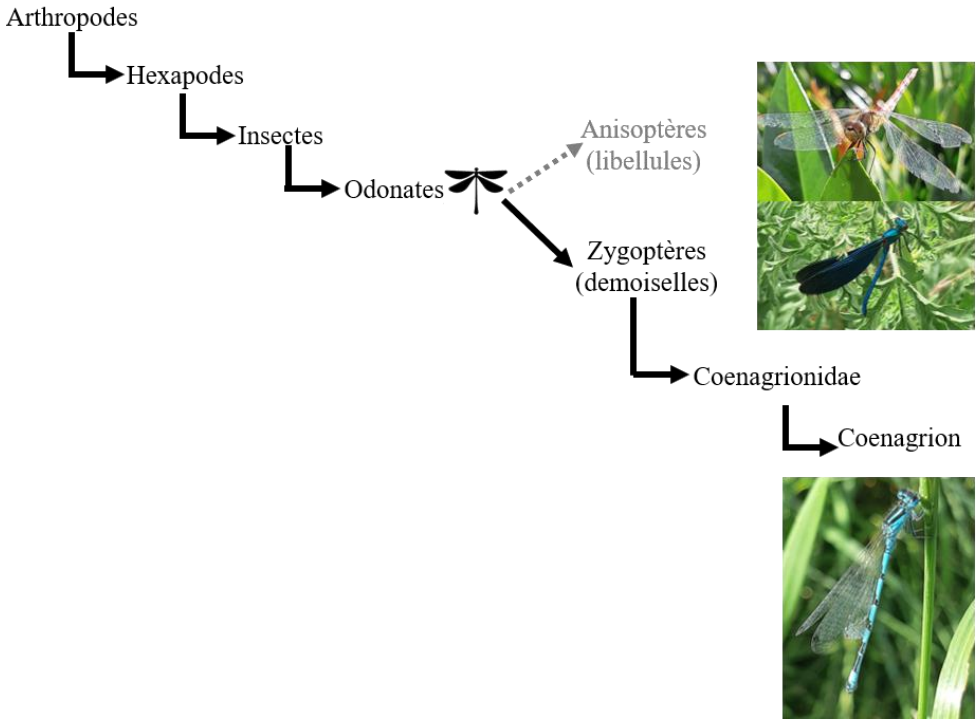


***Figure 8**. Phylogénie simplifiée de l'Agrion de Mercure (Coenagrion mercuriale).*

D'une longueur d'environ 3 cm, l'agrion de Mercure est la plus petite des quatre espèces du genre *Coenagrion* que l'on retrouve actuellement en France, à savoir l'Agrion jouvencelle (*Coenagrion puella*), l'agrion joli (*Coenagrion pulchellum*), l'agrion de Mercure (*Coenagrion mercuriale*) et l'agrion mignon (*Coenagrion scitulum*). Les mâles adultes de l'espèce, de couleur dominante bleue, sont particulièrement reconnaissables au dessin noir qu'ils portent sur leur deuxième segment abdominal. Ce dessin abdominal est relativement proche du symbole de Mercure, symbole ressemblant à une tête avec un casque à cornes (Figure 9). Compte tenu de la variation morphologique de cette marque, d'autres critères fondés sur l'examen des marques noires sur les autres segments abdominaux ainsi que les appendices anaux permettent de confirmer l'identification de l'espèce (Dijkstra *et al.*, 2015). Les femelles sont, quant à elles, beaucoup plus sombres, vert et noir, et sont plus difficiles à identifier. Dans ce cas précis, l'ornementation de leur pronotum reste le meilleur critère d'identification (Askew, 1988; Dijkstra *et al.*, 2015).

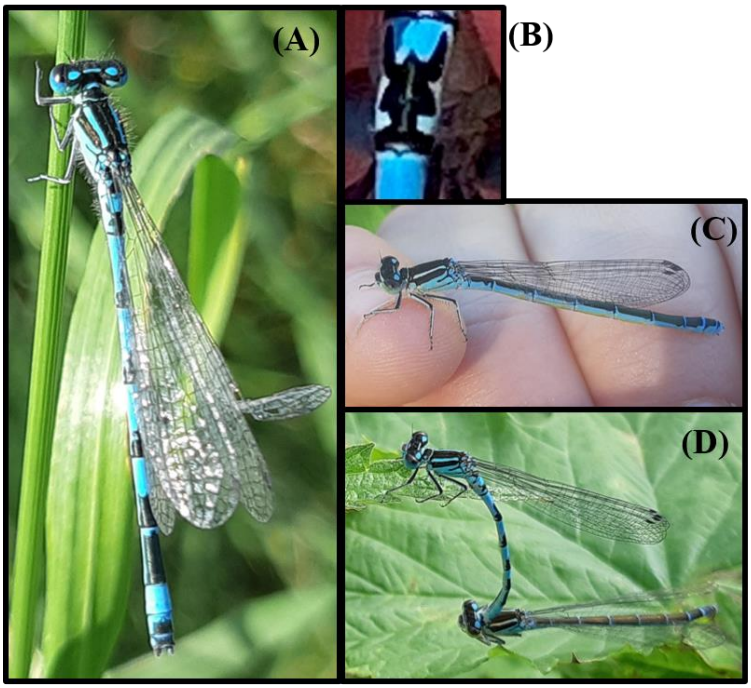


Figure 9. *Photographies de description de l'Agrion de Mercure (Coenagrion mercuriale). (A) Adulte mâle portant le symbole de Mercure sur son deuxième segment abdominal (B), (C) femelle agrion de Mercure et (D) photographie d'un mâle se saisissant d'une femelle au niveau du thorax grâce à ses appendices anaux afin de former un tandem copulateur.*

L'agrion de Mercure (*Coenagrion mercuriale*, Odonata, Charpentier, 1840) est une espèce protégée particulièrement intéressante pour des questions à la fois de biologie de la conservation, de génétique des populations et de biologie évolutive au regard de son écologie. En effet, cette espèce présente un cycle de vie et de reproduction, un type d'habitat, des modalités de dispersion, ainsi qu'une sensibilité aux pressions anthropiques particulièrement intéressantes. De fait, cette espèce a particulièrement été étudiée car elle :

- Revêt des enjeux de conservation liés au déclin de ses populations en limite d'aire de répartition géographique, déclin également associé aux activités humaines ;
- Dépend d'environnements à la fois terrestres et aquatiques avec des habitats spécifiques de reproduction ;
- Présente un gradient latitudinal de son voltinisme et un patron d'émergence asynchrone ;
- Présente une faible capacité de dispersion des individus, conduisant à des patrons de structuration génétique à faible échelle ;
- Montre une forte sensibilité aux perturbations anthropiques qui impactent la présence de l'espèce, mais impactent aussi la dispersion effective des individus.

La Figure 10 ci-dessous synthétise l'ensemble des études ayant été menées sur cette espèce, ainsi que les lieux et thématiques de recherche conduites.

Écologie & habitat
1-Purse, 2001
2-Thompson, Rouquette, *et al.*, 2003
3-Rouquette and Thompson, 2005
4-Rouquette and Thompson, 2007b
5-Couvreur *et al.*, 2008
6-Ferreira *et al.*, 2015
7-Mahdjoub, 2017

Reproduction & développement
8-Martens, 2000
1-Purse, 2001
9-Purse and Thompson, 2002
10-Purse and Thompson, 2003
11-Purse and Thompson, 2003b
12-Purse and Thompson, 2005
13-Purse and Thompson, 2009
14-Watts and Thompson, 2012
15-Mahdjoub *et al.*, 2014
16-Mahdjoub *et al.*, 2015
17-Mahdjoub, 2017
18-Khelifa *et al.*, 2019

Développement et mise au point de marqueurs génétiques
19-Watts *et al.*, 2004a
20-Watts *et al.*, 2004b
21- Watts *et al.*, 2007

Diversité génétique, dispersion & impact du paysage
22-Hunger and Röske, 2001
23-Purse *et al.*, 2003
24-Watts *et al.*, 2004c
25-Watts *et al.*, 2005
26-Thompson and Watts, 2006
27-Watts *et al.*, 2006
28-Watts, Rousset, *et al.*, 2007
29-Rouquette and Thompson, 2007a
30-Keller *et al.*, 2012
31-Van Strien *et al.*, 2012
32-Hassall and Thompson, 2012
33-Keller and Holderegger, 2013
34-Johansson *et al.*, 2013
35-Lorenzo-Carballa *et al.*, 2015
36-La Porta and Goretti, 2020

Suivi et abondance
37-Thompson, Purse, *et al.*, 2003
38-Couvreur *et al.*, 2008
39-Allen and Thompson, 2014
6-Ferreira *et al.*, 2015
17-Mahdjoub, 2017
40-La Porta and Goretti, 2019
36-La Porta and Goretti, 2020
41-Beaune and Sellier, 2021

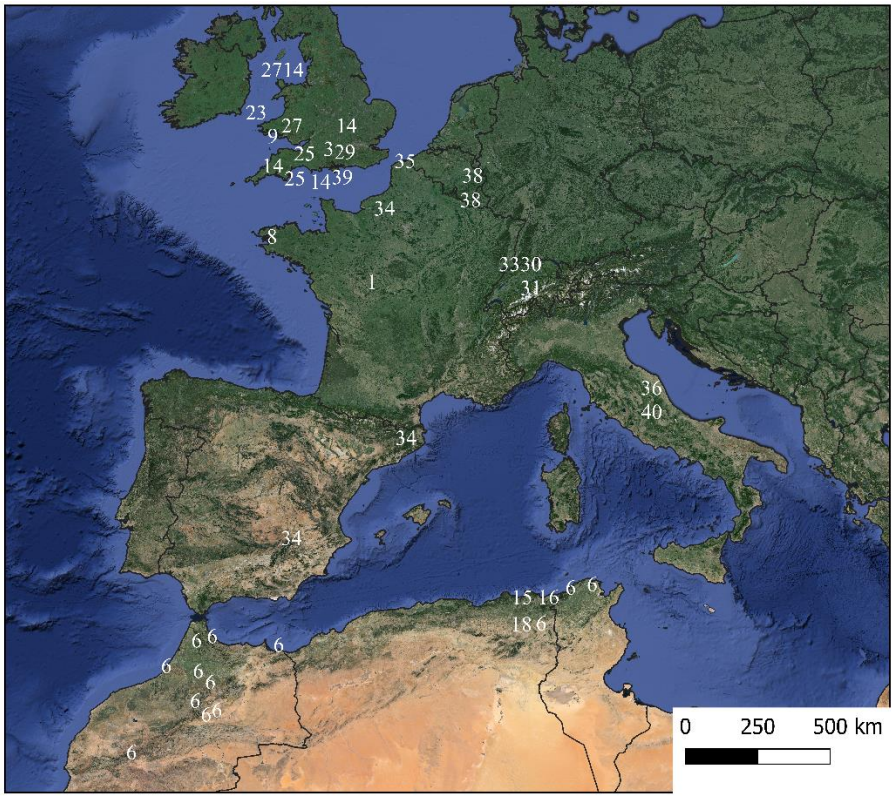


Figure 10. *Résumé des articles scientifiques réalisé sur l'agrion de Mercure (Coenagrion mercuriale). Les études sont rangées par classe thématique de recherche et la carte montre la répartition géographique de ces études menées sur cette espèce (les chiffres blancs correspondant aux études citées).*

2) Distribution et statuts de conservation

L'aire de distribution géographique de l'agrion de Mercure est limitée au Sud-Ouest de l'Europe et au nord du Maghreb. L'espèce est commune en France, mais plus dispersée en Espagne et au Portugal (Figure 11). L'agrion de Mercure est retrouvé plus rarement au Royaume-Uni en limite de répartition septentrionale. L'espèce trouve sa limite de répartition orientale en Allemagne. Enfin l'on rencontre des populations en limite méridionale de répartition au nord du Maroc, de l'Algérie et de la Tunisie (Ferreira *et al.*, 2015; Boudot, 2020). La distribution spatiale connue de l'espèce a toutefois évolué entre 2006 et 2023 avec l'actualisation des données de suivi et des statuts de conservation de l'espèce. S'ajoutent notamment des données de présence de l'espèce en Afrique du Nord et le retrait de l'Italie comme zone de présence en 2023. Ces évolutions récentes sont illustrées en Annexe 1 (Thompson & Watts, 2006; Lorenzo-Carballa *et al.*, 2015; Boudot, 2020; De Knijf *et al.*, 2023). Les effets du changement climatique sur les populations de l'espèce, avec des périodes estivales d'assèchement des cours d'eau de plus en plus fréquentes, risquent d'entraîner un déclin significatif de l'espèce au cours des prochaines décennies. Ainsi, la distribution géographique de l'agrion de Mercure sera très certainement amenée à évoluer très rapidement dans les années futures (De Knijf *et al.*, 2023).

En France, l'agrion de Mercure est largement distribué et localement abondant, à l'exception des régions les plus septentrionales et de l'Île-de-France. L'espèce est également absente des régions montagneuses à haute altitude (Houard, 2020; Fierimonte & Vanappelghem, 2021).

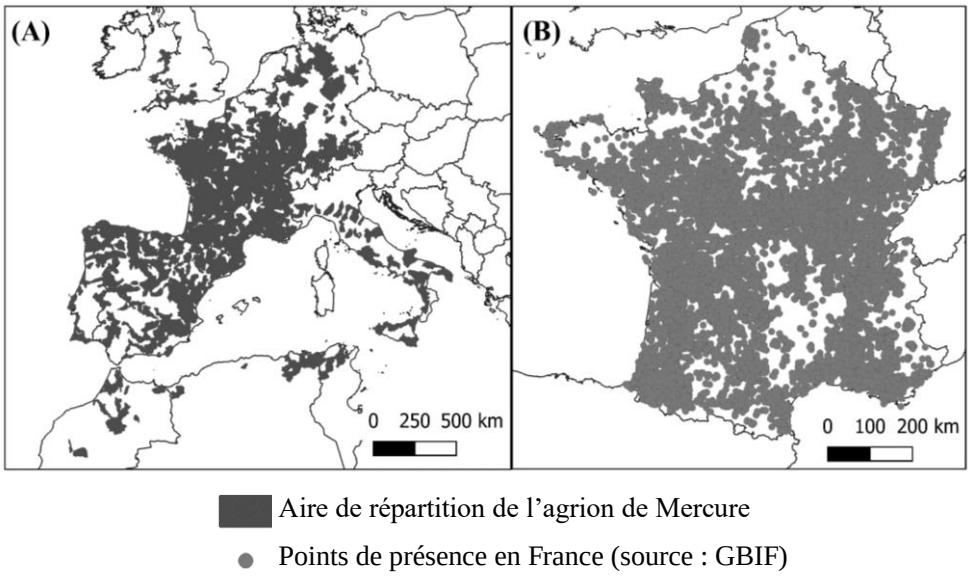


Figure 11. *Distribution spatiale de l'agrion de Mercure (Coenagrion mercuriale) en Europe et en France. (A) Carte montrant la répartition actuelle connue (zones gris foncé) de l'agrion de Mercure. Les données ont été obtenues auprès de l'IUCN SSC Odonata Specialist Group 2019. The IUCN Red List of Threatened Species. Version 2022-2. https://www.iucnredlist.org/ Downloaded on 28 July 2023. (B) Carte montrant les occurrences de l'agrion de Mercure enregistrées en France (points gris). Les données ont été obtenues sur GBIF.org (5 mars 2024) GBIF Occurrence Download https://doi.org/10.15468/dl.5z9gue*

L'agrion de Mercure fait l'objet de mesures de protection internationales et européennes liées à son statut de conservation. A l'échelle internationale, l'espèce est considérée comme vulnérable (IUCN Red List of Threatened Species ; De Knijf *et al.*, [2023]) alors qu'elle était évaluée comme quasi-menacée en 2020 (NT Near Threatened ; Boudot, 2020 ; Figure 12). Historiquement, le statut de l'espèce s'explique par le déclin de ses populations, notamment en périphérie de son aire de répartition (Grand, 1996; Ferreira *et al.*, 2015). Toutefois, une nouvelle analyse réalisée pour l'évaluation régionale européenne semble aussi indiquer un déclin présumé des populations en France et surtout en Espagne et au Portugal. Ce déclin des populations d'agrion de Mercure est directement lié à la destruction et à la fragmentation des habitats de l'espèce avec le développement de pratiques agricoles intensives modernes, le drainage et l'assèchement des cours d'eau, ainsi que la fermeture de ces derniers par des forêts ou embroussaillements incompatibles avec la survie ou la reproduction de l'espèce (Purse, 2001; Thompson *et al.*, 2003; Boudot, 2020).

Statut de conservation Liste rouges

Statut		Liste	Niveau
Vunérable	(VU)	Liste rouge mondiale des espèces menacées	International
Préoccupation mineur	(LC)	Liste rouge des odonates de France métropolitaine (2016)	France
Vunérable	(VU)	Liste rouge régionale de la faune menacée en Picardie	Picardie
En danger	(EN)	Liste rouge régionale des Odonates du Nord - Pas-de-Calais	Nord - Pas-de-Calais
Quasi menacé	(NT)	Liste rouge des Odonates du Grand Est	Grand Est

Protection et législation

Liste de la Convention de Bonn pour la conservation des espèces migratrices International d'animaux sauvages	International
Convention relative à la conservation de la vie sauvage et du milieu naturel de l'Europe (Convention de Berne): Annexe II	International
Directive 92/43/CEE (Directive européenne dite Directive Habitats-Faune-Flore): Annexe II	Europe
Liste des insectes protégés sur l'ensemble du territoire et les modalités de leur protection: Article 3	France

Figure 12. *Statut de conservation et mesures de protection de l'Agrion de Mercure (Coenagrion mercuriale) de l'échelle internationale à l'échelle régionale. Les données ont été adaptées d'après Thompson et al., 2003, et le site internet suivant : https://inpn.mnhn.fr/).*

3) Cycle de vie et habitats

Les odonates sont des insectes thermophiles, la température influence la distribution des espèces de ce taxon mais également le développement des individus (Purse, 2001). L'agrion de Mercure est une espèce semi-voltine au Nord et à l'Est de son aire de distribution géographique, il faut donc à cette

espèce deux ans pour réaliser une génération (Corbet, 1957; Sternberg *et al.*, 1999; Purse & Thompson, 2002). Cependant, le voltinisme de l'agrion de Mercure peut être considéré comme un trait plastique le long d'un gradient latitudinal, les populations localisées au Sud de son aire de répartition présentant un développement uni- et bi-voltine (Mahdjoub *et al.*, 2014, 2015). Le temps de développement des individus peut en outre être diminué d'une année par des facteurs environnementaux tels que les eaux de refroidissement industriel, comme cela a été montré dans une population en Allemagne (Thelen, 1992). Pour la suite, nous nous focaliserons sur l'écologie des populations présentant un développement typiquement caractéristique d'une espèce semi-voltine.

Les espèces d'odonates sont caractérisées par un cycle de vie dépendant de deux environnements différents, avec un développement larvaire au sein de milieux aquatiques et des adultes volants dépendant au contraire de milieux terrestres. Les larves d'agrion de Mercure évoluent dans des milieux aquatiques d'eau courante et se retrouvent principalement au niveau de cours d'eau de faibles calibres à courant lent et avec des températures favorables à l'espèce (Figure 13). Ces larves vivent dans la végétation aquatique immergée et les sédiments (Fierimonte & Vanappelghem, 2021). Le développement larvaire dure deux ans, avec la réalisation d'une dizaine de stades de développement impliquant des mues successives (Purse & Thompson, 2002).

Les adultes émergent de leur stade larvaire final entre la mi-mai et la fin-juin en France, mais ces périodes d'émergence peuvent varier en fonction de la latitude et de l'altitude. Les larves de deuxième année stoppent leur développement en hiver à différents stades de développement, l'espèce montre ainsi un patron d'émergence long et asynchrone, on parle alors de « Summer species » (Purse & Thompson, 2002; Thompson *et al.*, 2003). Après émergence, les individus suivent ensuite un stade de maturation durant lequel ils quittent les environs immédiats des cours d'eau et se déplacent vers les sites d'alimentation, les mâles acquièrent alors leur coloration mature (Thompson *et al.*, 2003; Purse & Thompson, 2003b). Après la maturation, on retrouve les mâles au niveau des sites de reproduction qu'ils survolent de manière erratique ainsi qu'au niveau de sites de perchage situés à proximité (Thompson *et al.*, 2003; Rouquette & Thompson, 2007b). Les femelles, quant à elles, sont retrouvées moins fréquemment sur les sites de reproduction où elles se rendent une fois qu'elles ont une réserve d'œufs à pondre (Purse & Thompson, 2003a). La copulation des odonates est caractérisée par la formation dite du cœur copulatoire (Figure 13). La femelle réalise ensuite une ponte endophyte en déposant ses œufs dans les tiges creuses de végétaux hélophytes (Martens, 2000; Rouquette & Thompson, 2005; Purse & Thompson, 2005, 2009).

Au final, les individus adultes d'Agrion de Mercure dépendent ainsi de deux types d'habitats principaux, avec des sites de perchage (« roosting sites ») localisés au niveau des cours d'eau ainsi qu'au niveau des bandes enherbées des berges. Ceci permet aux individus de se protéger notamment lors d'intempéries et de se reposer. De plus, l'espèce dépend d'habitats de reproduction spécifique, liés au type de ponte réalisé par les femelles. En effet, la reproduction nécessite des habitats lotiques avec de petits cours d'eau ou des fossés à écoulement lent, ainsi que des habitats ouverts avec peu d'ombre

permettant la présence d'espèces végétales hélophytes indispensables à la ponte de l'espèce (Thompson *et al.*, 2003; Rouquette & Thompson, 2005; Purse & Thompson, 2009, Figure 13).

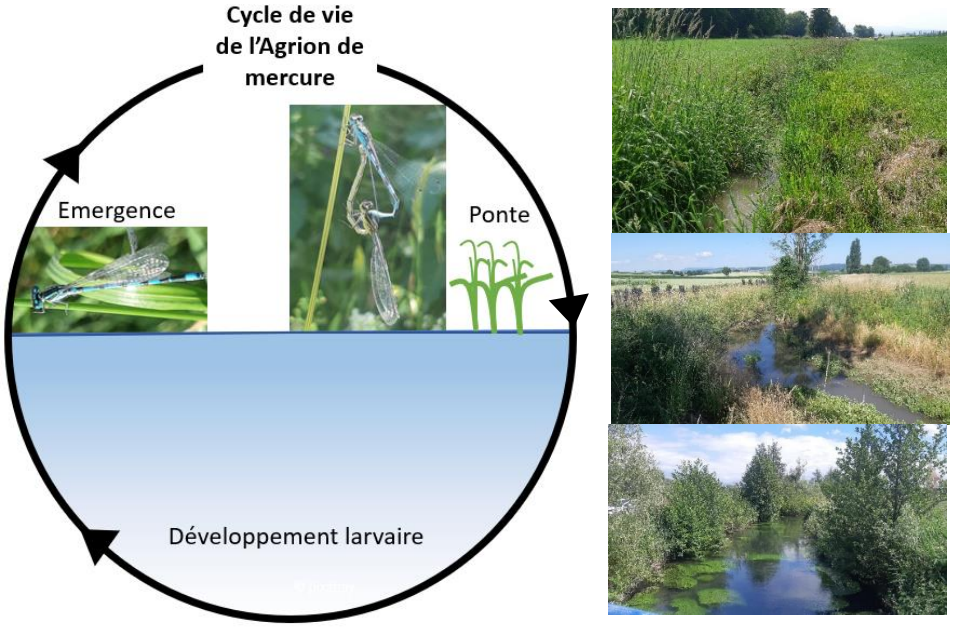


Figure 13. *Cycle de développement de l'agrion de Mercure (Coenagrion mercuriale) et photographies d'habitats de reproduction.*

4) **Dispersion des individus et flux de gènes**

L'agrion de Mercure a la réputation d'être une espèce sédentaire avec une faible capacité de dispersion. En effet, les études de capture-marquage-recapture, menées au Royaume-Uni et permettant une estimation directe des capacités de dispersion de l'espèce, ont montré que la grande majorité des mouvements des adultes volants étaient réalisés à l'échelle d'une centaine de mètres (impliquant de 75% à 85% des individus suivis) avec quelques rares observations de dispersion à plus grandes distances et supérieures à 1 km (Purse *et al.*, 2003; Thompson & Watts, 2006; Watts *et al.*, 2007, 2004c, Figure 14A). La majorité des mouvements d'individus était ainsi réalisée au sein de patches d'habitats relativement restreints, les mouvements entre ces patches locaux d'habitats étant plus rares (Purse *et al.*, 2003; Thompson &Watts, 2006, 2006). Au final, la distance maximale de déplacement observée pour l'ensemble des individus marqués et recapturés a été de 1790 m pour un mâle (Rouquette & Thompson, 2007a; Watts *et al.*, 2004c).

Ces observations directes de déplacements individuels ont ensuite été confortées par le patron de structuration génétique observé à fine échelle ainsi que les niveaux de différenciation génétique relativement élevés (F_{ST} variant de 0.0030 à 0.2556 avec des marqueurs microsatellites) entre populations (Watts *et al.*, 2005, 2007, 2004c). Toutefois, ces études génétiques ont aussi montré des flux

de gènes jusqu'à une distance de 2 km et des taux de migration relativement élevés par rapport aux observations directes de déplacement des individus (Watts *et al.*, 2006). Ces estimations indirectes d'évènements de migration à longue distance restent concordantes avec une observation ponctuelle de dispersion d'un individu d'Agrion de Mercure jusqu'à 3 km de l'habitat de reproduction le plus proche (Hunger & Röske, 2001).

De manière générale, des études de génétique des populations ont montré à plusieurs reprises une structuration génétique spatiale s'accordant avec un patron d'isolement par la distance, impliquant une augmentation de la différenciation génétique entre paires de populations avec la distance géographique séparant ces mêmes paires de populations (Watts *et al.*, 2006, 2007, 2004c; Lorenzo-Carballa *et al.*, 2015, Figure 14B).

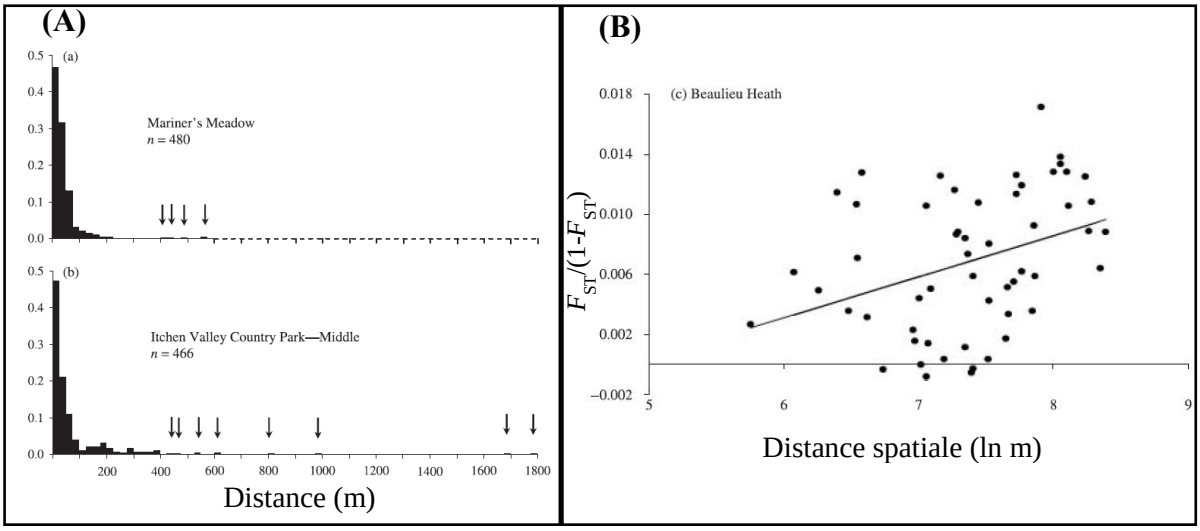


Figure 14. *Capacité de dispersion de l'agrion de Mercure (Coenagrion mercuriale). (A) Distribution des mesures directes des mouvements individuels dans deux régions localisées au sud du Royaume-Uni, Mariner's Meadow et Itchen Valley Country Park. n: nombre d'individus recapturés. Les flèches soulignent les mouvements peu fréquents impliquant une longue distance de dispersion. (Figure adaptée de Watts* et al., *2004c). (B) Patron d'isolement par la distance illustré par la relation entre la différenciation génétique par paire de populations (définie par $F_{ST}/[1 - F_{ST}]$) et leur distance géographique sur le site d'étude de Beaulieu Heath (UK) (Figure adaptée de Watts* et al., *2007).*

De plus, plusieurs études ont montré des voies de dispersion variables à courte et longue distance géographique, ainsi qu'un lien entre le paysage ou les activités anthropiques et les voies de dispersion et l'intensité des flux de gènes entre populations d'agrion de Mercure. Ainsi, sur de courtes distances (< 3 km), les déplacements des individus se limitent à l'habitat de reproduction localisé le long des cours d'eau, tandis qu'à plus grande distance, la dispersion des individus semble mieux expliquée par les distances euclidiennes entre populations (Hunger & Röske, 2001; Rouquette & Thompson, 2007a; Keller & Holderegger, 2013). Par ailleurs, plusieurs éléments du paysage ont été identifiés comme ayant probablement un impact sur la dispersion des individus. La fermeture des cours d'eau par les forêts et

l'embroussaillement ainsi que les reliefs élevés représenteraient ainsi des barrières à la dispersion de l'espèce (Purse, 2001; Purse *et al.*, 2003; Watts *et al.*, 2006; Keller *et al.*, 2012; Lorenzo-Carballa *et al.*, 2015). Au contraire, les zones aquatiques, les prairies et les landes représenteraient des espaces facilitant les flux de gènes s'opérant entre populations (Watts *et al.*, 2006; Lorenzo-Carballa *et al.*, 2015). De manière intéressante, les infrastructures de transport comme les routes et chemins de terre ne semblent pas impacter les flux de gènes entre populations de part et d'autre de ces voies de communication (Thompson & Watts, 2006; Keller *et al.*, 2012; Watts *et al.*, 2004c). Toutefois, l'impact des espaces agricoles et urbains semble contrasté. En effet, si certaines études montrent un effet facilitateur de ces environnements pour la dispersion de l'espèce, d'autres études montrent en revanche un impact négatif contraignant les flux de gènes (Watts *et al.*, 2006, 2004c; Keller *et al.*, 2012; Lorenzo-Carballa *et al.*, 2015). Ces impacts contrastés s'expliqueraient peut-être par les types différents de pressions environnementales présentes au sein de ces différents sites d'étude et par le degré variable d'anthropisation de ces espaces.

Enfin, même si la dérive larvaire liée au courant le long des cours d'eau concerne de nombreux invertébrés d'eau douce (Elliott, 2003), ce type de dispersion est considéré comme négligeable pour l'agrion de Mercure car l'espèce est retrouvée au niveau de cours d'eau peu profonds, à faible débit et à végétation abondante (Watts *et al.*, 2007, 2004c; Cédric Vanappelghem, communication personnelle).

IV. Contexte et objectifs de la thèse

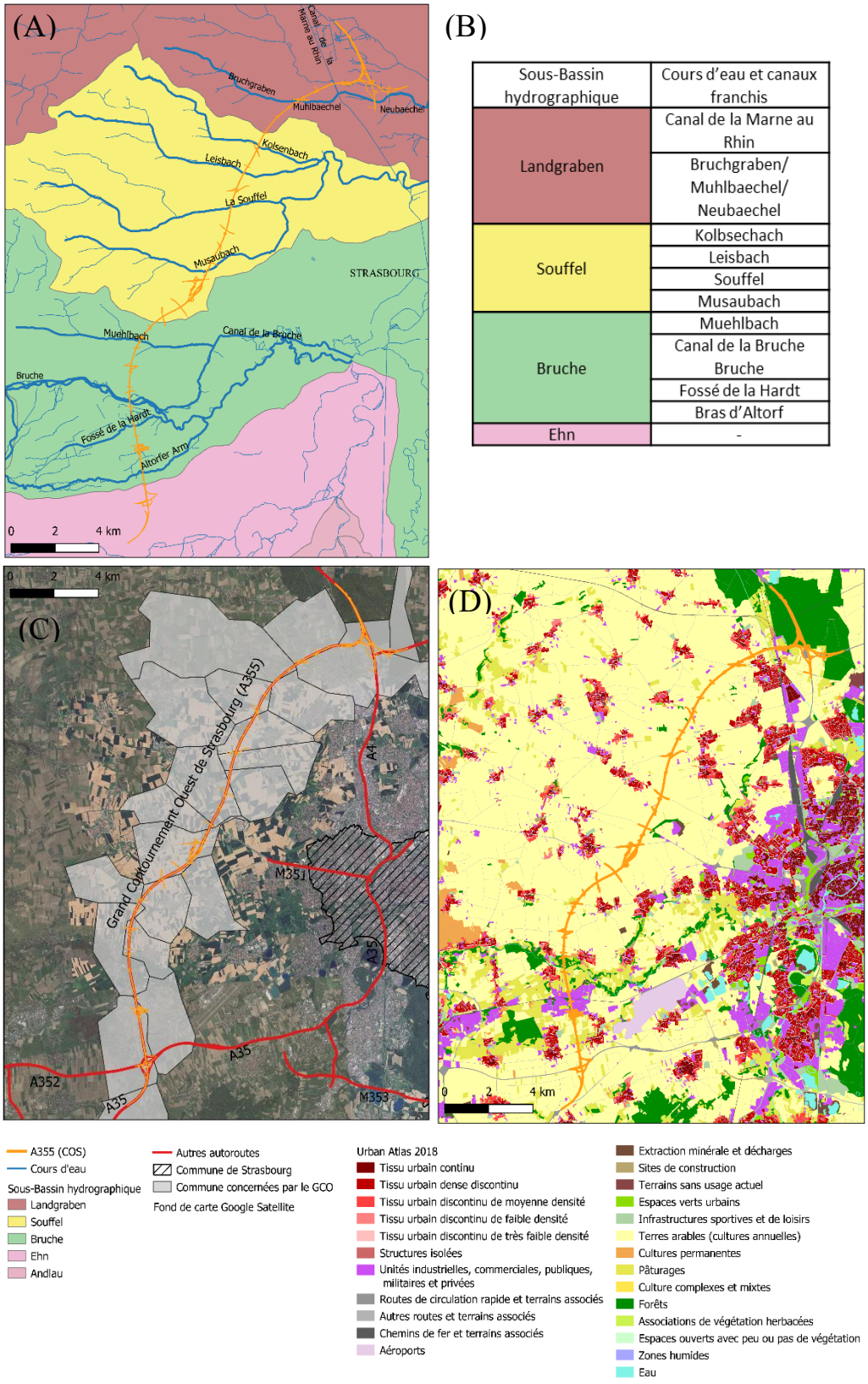
1) Contexte général de la thèse

Cette thèse de doctorat bénéficie d'un financement de type CIFRE qui s'inscrit dans un projet collaboratif entre le bureau d'étude « Office de Génie Écologique » et l'Unité Évolution, Écologie et Paléontologie (UMR 8198 CNRS - Université de Lille). Dans le contexte des changements environnementaux actuels, l'étude de l'évolution de la structure génétique d'espèces soumises à de fortes pressions anthropiques, *i.e*. urbanisation, agriculture intensive ou fragmentation par des infrastructures, est en plein essor pour des raisons à la fois appliquées de biologie de la conservation, mais aussi fondamentales. En effet, ces zones soumises à de fortes pressions sont le théâtre d'évolution neutre et adaptative très rapide (Donihue & Lambert, 2015; Munshi-South *et al.*, 2016; Johnson & Munshi-South, 2017; Santangelo *et al.*, 2018; Miles *et al.*, 2019; Halfwerk *et al.*, 2019). Cette thèse se place ainsi dans le contexte directement appliqué de l'étude des effets des modifications des cours d'eau et des recréations d'habitats dans le cadre de la construction de l'autoroute de contournement Ouest de Strasbourg. L'Agrion de Mercure est, en effet, historiquement présent au niveau de plusieurs cours d'eau traversés par cette autoroute, dont certains ont été réaménagés lors de sa construction.

Éléments de connaissance concernant l'autoroute de Contournement Ouest de Strasbourg, l'A355

L'autoroute A355 de Contournement Ouest de Strasbourg (COS), également appelée GCO (Grand Contournement Ouest), est une infrastructure autoroutière dont la vocation est de constituer un itinéraire nord-sud dans le département du Bas-Rhin visant à délester par l'ouest l'A35, autoroute traversant l'agglomération Strasbourgeoise (Figure 15C,D). Les origines de ce projet remontent à 1973 dans le cadre du Schéma directeur d'aménagement et d'urbanisme de l'agglomération de Strasbourg. Le projet a toutefois fait l'objet de vives oppositions, avec la création en 2003 d'un collectif contre la construction du GCO. En 2008, le projet est déclaré d'utilité publique par décret en conseil d'État et en 2013, le Conseil Général de l'Environnement et du Développement Durable confirme la nécessité de l'infrastructure. Suite à un nouvel appel d'offres de l'état en 2015, la conception de cet ouvrage repose sur un partenariat public-privé avec un contrat de concession entre l'état et la société ARCOS (filiale du groupe VINCI, société concessionnaire du contournement ouest de Strasbourg) qui a confié la construction de l'ouvrage à un groupement d'entreprises, SOCOS. Les travaux de préparation et de construction de l'infrastructure ont commencé en 2016 et l'ouvrage a été inauguré en 2021. Dans le cadre de mesures compensatoires liées à la construction de l'autoroute A355, SOCOS a réalisé un partenariat avec OGE en finançant trois thèses co-encadrées par l'entreprise.

L'autoroute A355, d'une longueur d'environ 24 kilomètres, concerne 22 communes et s'insère dans un contexte rural au sein d'une plaine agricole avec principalement des parcelles cultivées (maïs, blé, maraîchage), des vignes et vergers, des prairies fauchées, mais aussi des milieux boisés (Figure 15D). L'autoroute franchit 11 cours d'eau dont deux canaux, avec une emprise s'étendant sur quatre sous-bassins hydrographiques (Figure 15A,B). Les cours d'eau rencontrés sont du Nord au Sud : le canal de la Marne au Rhin, le Muhlbaechel, le Kolbsenbach, le Leisbach, la Souffel, le Musaubach, le Muehlbach, la Bruche et le canal de la Bruche, le fossé de la Hardt et le bras d'Altdorf. Dans ce contexte, plusieurs aménagements hydrauliques ont été réalisés au niveau de cours d'eau présents sur la zone de travaux, ces aménagements impliquant par exemple des busages, des fascinages et des dérivations provisoires ou définitives des cours d'eau (Figure 16). Des travaux de restauration et de reméandrage ont aussi été réalisés au niveau de plusieurs sites aux abords directs de l'autoroute (Figure 16).



Sous-Bassin hydrographique	Cours d'eau et canaux franchis
Landgraben	Canal de la Marne au Rhin
	Bruchgraben/ Muhlbaechel/ Neubaechel
Souffel	Kolbsechach
	Leisbach
	Souffel
	Musaubach
Bruche	Muehlbach
	Canal de la Bruche Bruche
	Fossé de la Hardt
	Bras d'Altorf
Ehn	-

Figure 15. *Contexte entourant le Contournement Ouest de Strasbourg (A355). (A, B) Cours d'eau et sous-bassins hydrographiques traversés par l'autoroute. (C) Réseaux autoroutiers, les communes concernées par le GCO sont illustrées en gris, l'emprise de l'agglomération strasbourgeoise est hachurée. (D) Occupation des sols autour de l'autoroute, couche vecteur Urban Atlas 2018.*

Figure 16. *Photographie d'un site échantillonné en 2021 (A, B) et 2022 (C, D). Des busages du cours d'eau étaient visibles en 2021 (A). Des travaux de reméandrage ont été réalisés (B, C, D).*

Des inventaires écologiques ont été réalisés entre 2015 et 2021 au sein d'une aire d'étude comprenant l'emprise des travaux et ont permis de recenser la présence d'espèces d'intérêt patrimonial, notamment le Grand Hamster d'Alsace (*Cricetus cricetus*), la Grenouille verte (*Pelophylax kl. esculentus*) ainsi que l'Agrion de Mercure (*Coenagrion mercuriale*), une espèce à enjeux de conservation[2]. Dans ce cadre, des enjeux écologiques ont été identifiés au niveau des cours d'eau constituant des zones de reproduction pour l'Agrion de Mercure, il s'agit du fossé de la Hardt, du Musaubach, de la Souffel, du Leisbach, du Kolbsenbach, et du Muhlbaechel (Figure 15A). Afin de limiter l'impact sur cette espèce, des opérations de déplacement d'une partie des populations ont été effectuées pendant l'hiver 2018-2019 sur certains cours d'eau (communication interne OGE). Les fonds du cours d'eau ont été dragués au niveau de la zone de travaux et les herbiers ainsi qu'une partie du substrat ont été prélevés puis redéposés en aval avec pour objectif de déplacer une partie de la population d'Agrion de Mercure présente sous forme de larves dans le cours d'eau.

Éléments de description concernant la région strasbourgeoise

De manière plus large, le COS s'inscrit en Alsace dans la région entourant l'Eurométropole de Strasbourg. Cette aire d'étude d'une superficie d'environ 4 000 km² présente des espaces variés dans des milieux à la fois semi-naturels ou plus anthropisés (Figure 15D).

En effet, on y retrouve l'agglomération de Strasbourg, qui s'étend sur 78 km² et compte environ 290 000 habitants, avec une densité de population humaine de plus de 3 700 personnes/km². Elle est traversée par l'Ill et le canal du Bruche ainsi qu'un ensemble de canaux aménagés, avec une zone portuaire située le long du Rhin. La ville de Strasbourg est elle-même située au sein de l'aire urbaine de

[2] https://www.bas-rhin.gouv.fr/contenu/telechargement/27310/188753/file/DAU-SOCOS-EIA_1.pdf

l'Eurométropole de Strasbourg, qui compte environ 500 000 habitants pour 330 km². Cet espace forme une vaste nappe urbaine en patchwork dense et arboré (Figure 17 et Annexe 2A).

Autour de l'Eurométropole de Strasbourg, on retrouve deux espaces principalement agricoles, le Kochersberg et les Rieds. Ces espaces sont caractérisés par des étendues agricoles avec des grandes cultures (blé, maïs, orge) ponctuées de villages répartis régulièrement. On y retrouve de multiples réseaux hydrographiques avec la présence de cours d'eau ouverts de faibles calibres et favorables à l'Agrion de Mercure (Figure 17, Figure 18, et Annexe 2BC).

À l'Ouest, on retrouve ensuite le piedmont viticole, qui forme un coteau viticole nord/sud de plus de 100 kilomètres de long, adossé au contrefort vosgien avec une plus faible densité de populations d'agrion de Mercure recensées (Figure 17, Figure 18 et Annexe 2E). Enfin, à l'Est le long du Rhin, on retrouve la bande Rhéanae et le Ried Nord, deux zones mosaïques présentant des forêts alluviales, avec des espaces cultivés parsemés de villages et de zones industrielles. On retrouve de nombreuses populations d'agrion de Mercure localisées le long des contre-canaux du Rhin ponctués d'installation hydroélectriques (Strasbourg, Gerstheim, Rhinau ; Figure 17 et Annexe 2D).

L'agrion de Mercure est classé comme vulnérable par le Plan Régional d'Action (PRA) en Alsace. L'espèce est historiquement présente et disséminée dans la région (Figure 18, Annexe 3). Au niveau de la bande Rhénane, on trouve des populations montrant des densités localement importantes. L'espèce est toutefois plus rare à proximité des reliefs vosgiens ainsi que dans les massifs forestiers. L'Agrion de Mercure est présent sur différents types d'habitats lotiques, allant de cours d'eau de faibles calibres (<1 m, fossés alimentés par la nappe ou petits réseaux hydrographiques de plaines) jusqu'au contre-canal du Rhin (largeur de plus de 15 m) dès lors que l'on retrouve une eau faiblement courante, peu profonde, avec des taches d'hélophytes, et un ensoleillement suffisant (Figure 17, Figure 18).

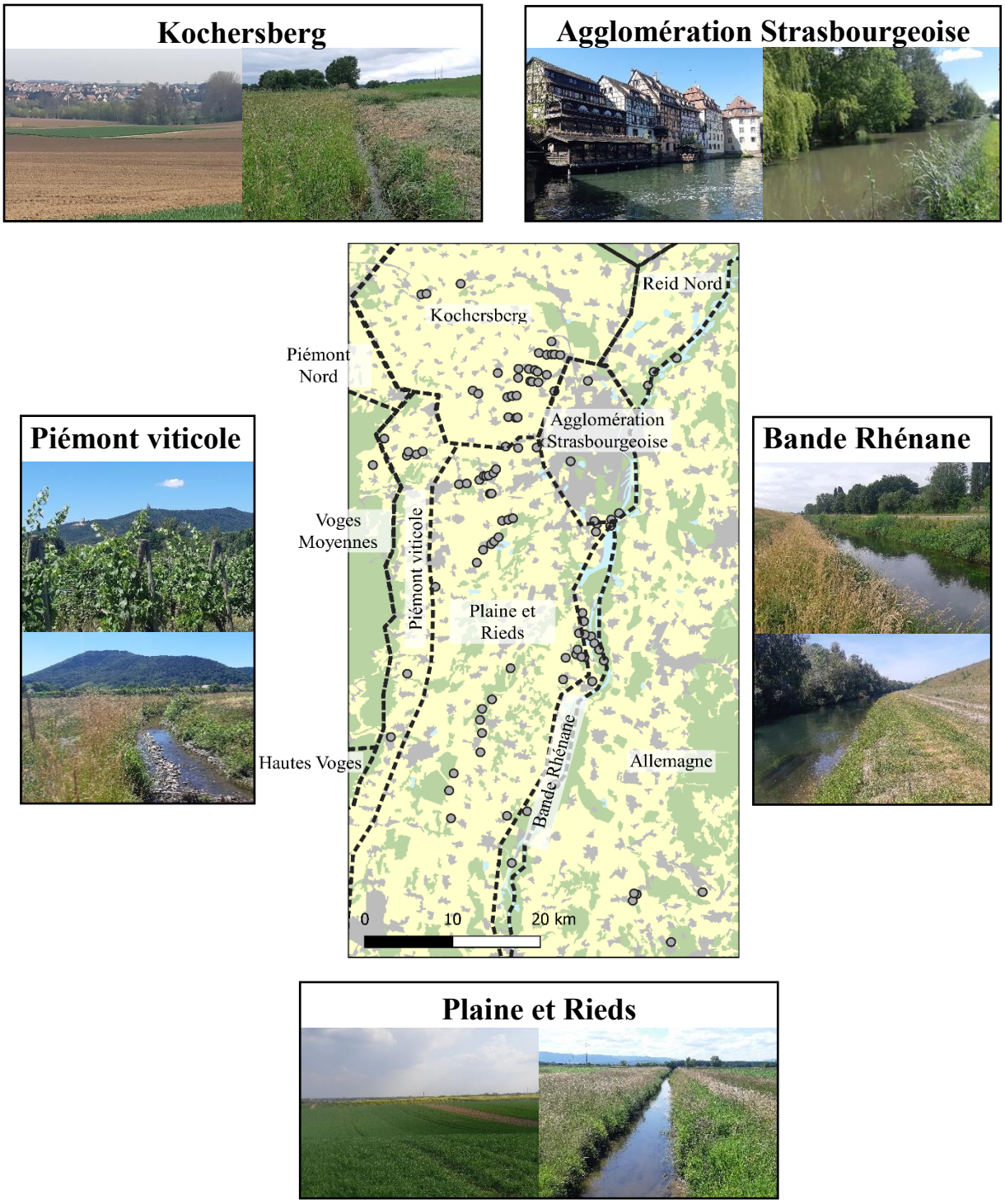


Figure 17 : *Grands types de paysage intégrés dans la zone d'étude. Les espaces ont été délimités à l'aide de l'Atlas des paysages d'Alsace*[3]*. L'occupation du sol a été simplifiée à partir de Corine Land Cover Edition 2018. Les sites échantillonnés dans le cadre de la thèse sont indiqués par les points gris. Pour chaque zone, des photographies prises lors de l'échantillonnage de populations ont été ajoutées. L'espèce a été retrouvée sur l'ensemble des photographies représentant un cours d'eau.*

[3] http://www.paysages.alsace.developpement-durable.gouv.fr/spip.php?rubrique32 © ministère de l'Écologie, du Développement durable et de l'Énergie - DREAL Alsace

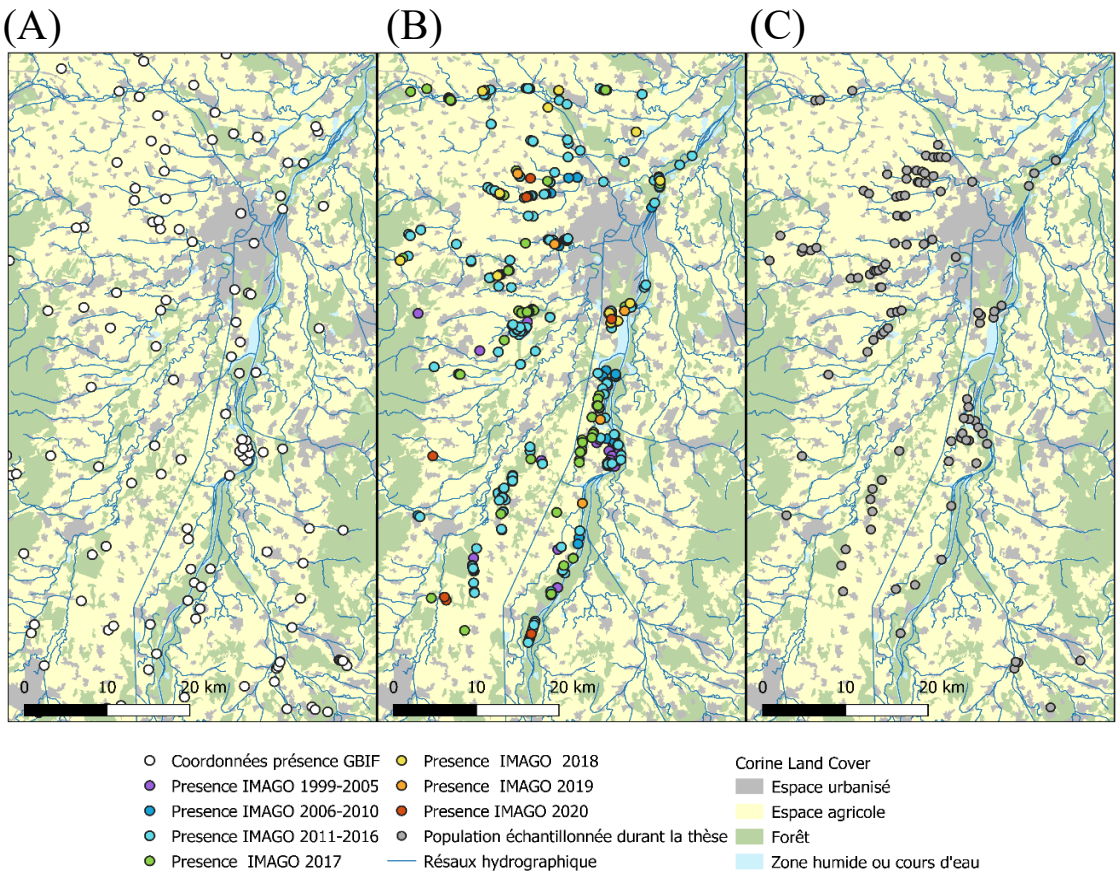


Figure 18 : *Données de présence de l'agrion de Mercure dans la région d'étude. (A) Données de présence obtenue via GBIF.org (5 mars 2024)[4]. (B) Données de présence obtenue de l'association IMAGO[5]; (C) Sites échantillonnés dans le cadre de la thèse. Les sites échantillonnés, mais où aucun agrion de Mercure n'a été trouvé malgré des conditions favorables, sont indiqués en Annexe 3.*

Description d'une zone géographique complémentaire à l'agglomération Strasbourgeoise : le Nord de la France

Cette région correspond à une zone d'étude secondaire permettant de généraliser les patrons de structure génétique au-delà de ce qui est observé au niveau régional en Alsace. Cette zone contraste avec la région Strasbourgeoise, car les populations d'agrion de Mercure y sont plus rares et plus isolées et ne se rencontrent que dans des habitats semi-naturels, avec des efforts de conservation pour l'espèce.

Localisées en limite Nord de répartition de l'espèce en France, les populations d'agrion de Mercure se raréfient à mesure que l'on remonte vers le nord : Haute-Normandie, Picardie et Nord-Pas-de-Calais (Vanappelghem & Hubert, 2010). En effet, l'espèce est considérée comme en danger dans le Nord-Pas-de-Calais et vulnérable pour la liste rouge de Picardie (Fierimonte & Vanappelghem, 2021). Si on retrouve l'espèce dans l'ensemble de la région des Hauts-de-France, sa présence dans chaque département est variable. En effet, on distingue des populations très localisées dans l'Aisne ainsi que

[4] GBIF Occurrence Download https://doi.org/10.15468/dl.5z9gue. https://doi.org/10.15468/dl.9h9maw
[5] http://association.imago.free.fr/

dans le Nord, où l'on ne retrouve l'espèce qu'au niveau de deux sites (Wallers et Cuyet à Saint-Amand-Les-Eaux) autour du Parc Naturel Régional Scarpe-Escaut (Figure 19). Quelques populations, découvertes à partir de 1996, sont retrouvées au sud-ouest du Pas-de-Calais, notamment dans la vallée de la Course, la Canche ainsi que de petits cours d'eau côtiers et à proximité du PNR Caps et Marais d'Opale (Figure 19). Toutefois, ces populations restent isolées des autres stations connues plus au sud, en Picardie, ou plus à l'est dans le Nord et en Belgique. Ces populations ont d'ailleurs fait l'objet d'une première étude de génétique des populations visant à établir leurs statuts de conservation ainsi que la connectivité entre ces populations (Lorenzo-Carballa *et al.*, 2015). Dans la Somme, on rencontre des populations localisées dans les vallées de la Somme et de la Bresle à proximité du PNR Baie de Somme Picardie Maritime (Figure 19). Enfin, dans l'Oise, des groupements de populations à l'ouest et au sud du département existent au niveau des PNR du Vexin français et Oise-Pays de France (Figure 19).

Préalablement à ce projet de thèse, plusieurs populations avaient été échantillonnées entre 2017 et 2020 par des membres du Conservatoire des Espaces Naturels dans le Nord, le Pas-de-Calais, la Somme et l'Oise, ainsi que sur deux sites en situés Wallonie ou l'espèce est considérée comme vulnérable (Motte *et al.*, 2021).

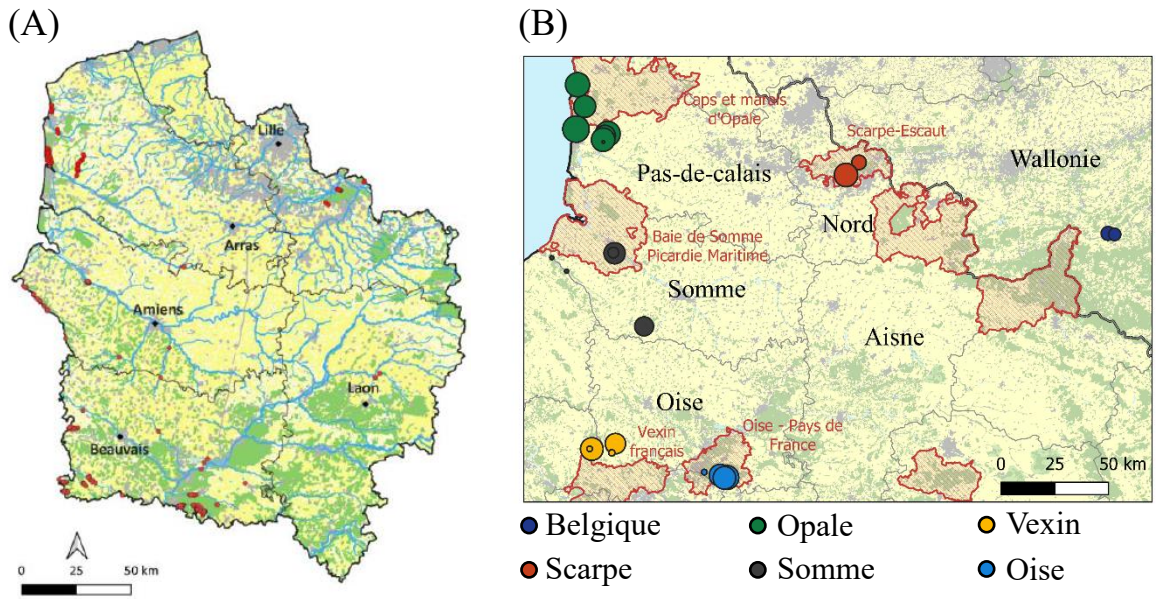


Figure 19 *(A) Carte de la répartition des populations d'Agrion de Mercure (Coenagrion mercuriale) dans les Hauts-de-France (Fierimonte & Vanappelghem, 2021). (B) Carte des populations échantillonnées et étudiées dans le cadre de la thèse. Les différents Parcs Naturels Régionaux cités dans le texte sont en rouge hachuré.*

2) Objectif général de la thèse

L'objet de ce travail de thèse est d'étudier à fines et à larges échelles spatiales la structure génétique et génomique de populations d'Agrion de Mercure (*Coenagrion mercuriale*), une espèce protégée d'intérêt patrimonial et tributaire des cours d'eau. Ces travaux de thèses répondent à quatre

objectifs en lien avec des questions de génétique de la conservation : (i) inférer l'intensité des flux de gènes s'opérant entre populations et déterminer les voies de dispersion préférentielles des individus ; (ii) évaluer les niveaux de diversité génétique, de consanguinité, et de tailles efficaces de populations, paramètres clés de persistance à long terme des populations ; (iii) quantifier l'influence de perturbations anthropiques du paysage sur ces paramètres ainsi que les patrons de différenciation génétique entre populations ; (iv) étudier les effets des modifications des cours d'eau et des recréations d'habitats dans le cadre du projet autoroutier de construction de l'A355 (Figure 20).

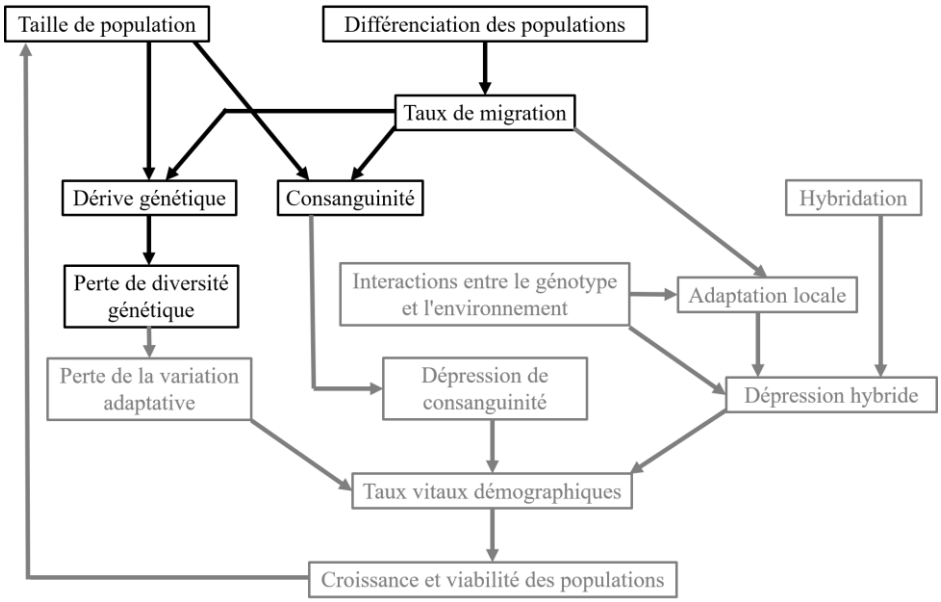


Figure 20. *Schéma des facteurs d'interaction dans la conservation des populations naturelles. Les estimations de ces facteurs d'interaction, étudiés à l'aide de locus microsatellites et de SNPs sont présentées en noir, alors que les facteurs non étudiés dans le cadre de cette thèse sont représentés en gris. Figure modifiée à partir de Allendorf et al., 2010.*

Pour répondre à ces objectifs, différents volets de recherche sont abordés à différentes échelles géographiques dans le cadre de cette thèse (Figure 21) :

1) Une étude de la structure génétique régionale des populations d'agrion de Mercure menée sur l'ensemble de l'agglomération strasbourgeoise pour avoir une vision plus large de l'effet d'un habitat anthropisé sur l'intensité des flux de gènes entre populations, ainsi que sur leur niveau de taille efficace. Toutefois, afin de généraliser ces résultats au-delà d'une image instantanée de la structure génétique spatiale régionale observée uniquement en Alsace, et de mieux connaître la biologie de l'espèce au sein de son aire de distribution géographique, il est nécessaire d'avoir un comparatif avec d'autres régions moins soumises à un impact anthropique. Dans cette optique, la structure génétique et les patrons de flux de gènes mis en évidence en Alsace ont été comparés avec des populations d'Agrion de Mercure déjà échantillonnées dans la région des Hauts-De-France et présentes dans des milieux *a priori* moins anthropisés mais plus isolés géographiquement.

2) Une étude fine de la structure génétique spatiale des populations dans la région strasbourgeoise. Il s'agissait de définir s'il existait des unités génétiques distinctes dues à l'effet barrière de certains éléments du paysage ou au contraire l'existence d'un simple patron d'isolement par la distance géographique euclidienne. L'évolution de l'apparentement génétique des individus a ainsi été étudiée à l'échelle des sous-bassins versants, des cours d'eau et de leurs tributaires.

3) Une étude de cas, à fine échelle, de la structure génétique spatiale des populations d'agrion de Mercure ainsi que des processus de recolonisation de cours d'eau suite à des travaux de restaurations et de reméandrages réalisés le long de l'axe autoroutier du contournement Ouest de Strasbourg nouvellement construit. Ce volet de recherche vise à définir si l'autoroute A355 représente une barrière aux flux de gènes entre populations localisées de part et d'autre de cette infrastructure. De plus, l'objectif est de détecter si des systèmes sources-puits sont à l'œuvre ou au contraire si les colonisateurs proviennent de sources différentes « tous azimuts », ceci afin d'établir le type de recolonisations s'effectuant de part et d'autre du tracé de l'autoroute.

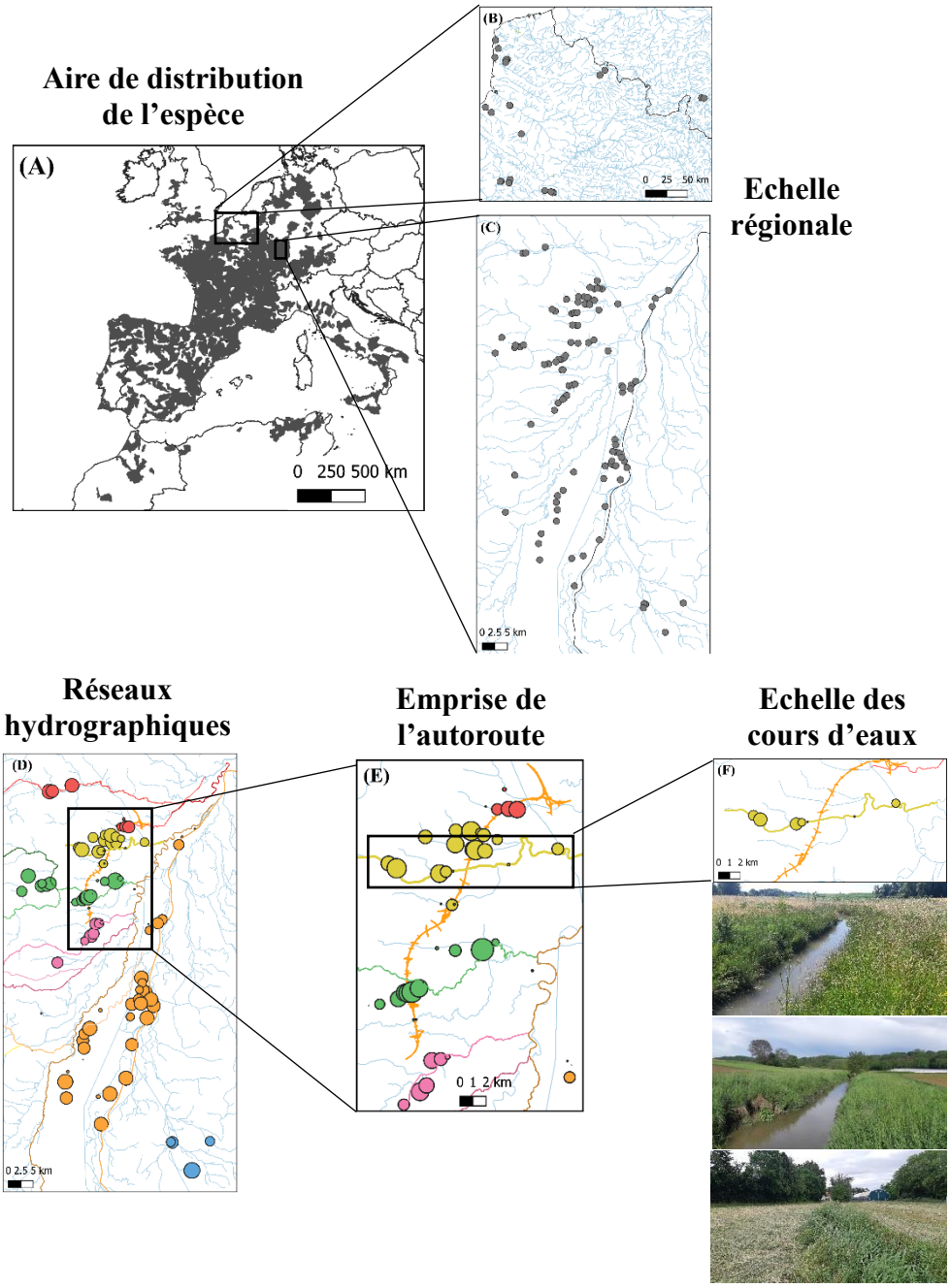


Figure 21. *Figure représentant les différentes échelles spatiales abordées dans le cadre de cette thèse.*

Les données employées dans ce projet ont été intégralement récoltées dans le cadre de cette thèse de doctorat, à l'exception de 589 échantillons récoltés dans les Hauts de France avant 2021, ces derniers m'ayant été fournis par C. Vanappelghem et récoltés par des membres du CEN. Pour cela, deux campagnes d'échantillonnages ont été réalisées à la fin du printemps 2021 et 2022 (mi-mai à fin juin) principalement en Alsace, mais aussi pour quelques populations localisées au sein des Hauts-De-France. Deux autorisations de capture ont été délivrées par la DREAL du Grand Est et la DREAL des Hauts-de-France après accord du CSRPN Grand Est (Arrêté N°2021-DREAL-EBP-0104) et du CSRPN Hauts-de-France (Arrêté N°2021-ESP-24). Ces campagnes ont permis d'échantillonner un total de 3207 individus génotypés à l'aide de marqueurs microsatellites déjà développés par Watts (Watts *et al.*,

2004ab). Une partie de ces échantillons ont ensuite été génotypés à l'aide de marqueurs SNPs mis au point durant cette thèse.

3) Organisation du manuscrit

Le chapitre I vise à étudier à fine et à large échelle les niveaux de diversité génétique et les niveaux de différenciation génétique entre populations d'agrion de Mercure retrouvées au sein de deux espaces contrastés de son aire de répartition : (i) l'un dans des habitats semi-naturels localisés dans le Nord de la France à la périphérie de l'aire de répartition géographique de l'espèce ; (ii) l'autre plus central dans l'aire de répartition de l'espèce, localisé en Alsace (Est de la France), où l'espèce est trouvée dans des sites plus impactés par les activités anthropiques (Figure 22). En se focalisant sur cette deuxième région, cette étude pilote menée à l'aide de marqueurs microsatellites a aussi permis d'étudier si les voies de dispersion de l'espèce s'effectuaient à travers le milieu terrestre ou suivaient strictement le milieu aquatique qui est l'habitat de reproduction de l'espèce. L'éventuel impact négatif de l'agglomération Strasbourgeoise, espace densément urbanisé, sur les niveaux de diversité génétique intrapopulations et de différenciation génétique interpopulations a également été étudié.

Le Chapitre II se focalise, quant à lui, sur le développement de nouveaux marqueurs de type SNP dans le but de réaliser un génotypage massif chez l'agrion de Mercure. L'utilisation des marqueurs SNPs devrait permettre une meilleure couverture du génome et une meilleure puissance statistique afin de mieux décrire l'histoire évolutive récente des populations d'agrion de Mercure, avec des estimations plus fines des tailles efficaces et des niveaux de consanguinité rencontrés en populations naturelles. Ce chapitre essentiellement méthodologique détaille une stratégie de développement de SNPs en deux étapes pour (i) assembler des contigs génomiques de référence et identifier *de novo* des milliers de marqueurs SNPs à partir d'une librairie ddRADseq, puis (ii) d'utiliser ces connaissances génomiques pour appliquer la méthode dite de « Allegro Targeted Genotyping » permettant de génotyper de manière ciblée des milliers de SNPs chez un très grand nombre d'individus. Les problèmes techniques associés à l'application d'outils génomiques d'enrichissement ciblé de SNPs chez une espèce non modèle sont ensuite discutés. Un dernier volet vise à valider l'utilisation de ces nouveaux marqueurs SNPs. Pour cela, des analyses visant à évaluer les niveaux de diversité génétique, de différenciation génétique, et à identifier l'affiliation génétique des individus, ont été menées dans un sous-ensemble de cinq populations localisées dans le nord de la France et précédemment étudiées à l'aide de marqueurs microsatellites. Ces populations sont localisées sur la côte d'Opale et sont cartographiées sur la Figure 19B.

Le Chapitre III est dédié à l'estimation, chez l'agrion de Mercure, d'un paramètre clé de conservation des populations, la taille efficace des populations (N_e). Au sein de la région Strasbourgeoise, une première analyse a été réalisée pour confirmer l'absence de cohortes génétiques

distinctes entre générations émergentes chez cette espèce semi-voltine. Les tailles efficaces de populations ont ensuite été estimées en appliquant de multiples méthodes, à la fois fondées sur un seul échantillonnage et sur des rééchantillonnages temporels, afin de comparer les estimations de N_e en utilisant des locus microsatellites déjà disponibles et un ensemble de SNPs décrits précédemment dans le chapitre II. La distribution spatiale des tailles efficaces de populations a ensuite été étudiée afin de définir des zones géographiques à faibles ou fortes tailles efficaces pouvant potentiellement être mises en lien avec des pressions anthropiques ou un système démographique de type source-puits. Enfin, des analyses ont été réalisées afin de définir l'impact de l'occupation des sols entourant les habitats de reproduction des populations échantillonnées sur les tailles efficaces estimées (Figure 22).

Le Chapitre IV est une étude de cas de l'impact à fine échelle de la construction d'une infrastructure autoroutière, l'autoroute A355, sur les populations d'agrion de Mercure (Figure 22). Cette étude vise à déterminer le potentiel effet barrière d'une telle infrastructure sur les flux de gènes entre populations, ainsi que les dynamiques d'affiliations d'individus de part et d'autre de l'autoroute dans le cadre de processus de recolonisation de cours d'eau restaurés et reméandrés lors de la construction de cette autoroute.

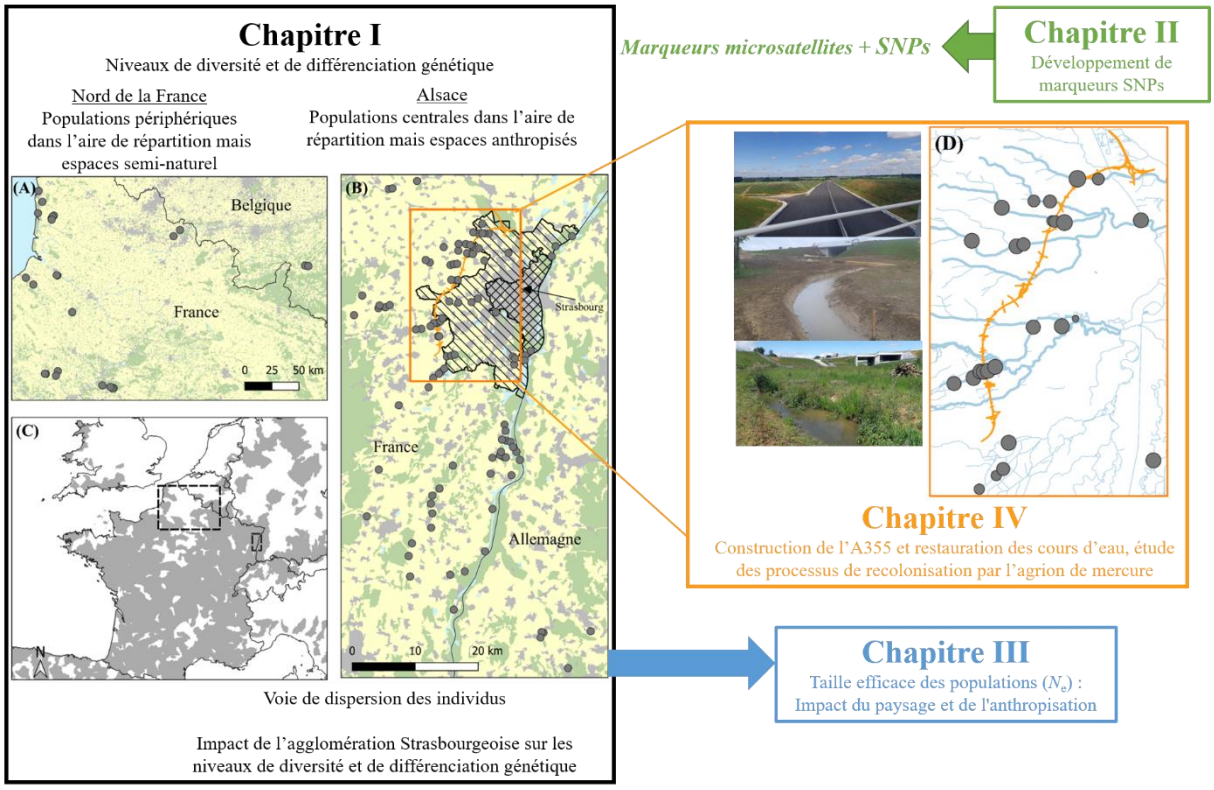


Figure 22. *Organisation du manuscrit en quatre chapitres et questions scientifiques abordées dans chacun d'entre eux. (A,B) Cartes des populations échantillonnées dans les Hauts-de-France et analysées dans le chapitre I, et dans le chapitre III pour celles situées en Alsace. (C) Carte des zones étudiées dans le chapitre I par rapport à la distribution de l'espèce en France et dans les pays frontaliers. (D) Carte focalisée sur les populations étudiées dans le Chapitre IV.*

Références

Allendorf FW (2017). Genetics and the conservation of natural populations: allozymes to genomes. *Molecular Ecology* **26**: 420–430.

Allendorf FW, Funk WC, Aitken SN, Byrne M, Luikart G (2022). *Conservation and the Genomics of Populations*, Third Edition. Oxford University Press: Oxford, UK.

Allendorf FW, Hohenlohe PA, Luikart G (2010). Genomics and the future of conservation genetics. *Nature Reviews Genetics* **11**: 697–709.

Armsworth PR, Chan KMA, Daily GC, Ehrlich PR, Kremen C, Ricketts TH, *et al.* (2007). Ecosystem-service science and the way forward for conservation. *Conservation Biology* **21**: 1383–1384.

Askew RR (1988). The Dragonflies of Europe. Harley Books Colch : Colchester, UK

Avise JC, Bowen BW, Ayala FJ (2016). In the light of evolution X: comparative phylogeography. *Proceedings of the National Academy of Sciences of the USA* **113**: 7957–7961.

Avise JC, Lansman RA, Shade RO (1979). The use of restriction endonucleases to measure mitochondrial DNA sequence relatedness in natural populations. I.Population structure and evolution in the genus Peromyscus. *Genetics* **92**: 279–295.

Beaune D, Sellier Y (2021). Stream restorations with meanders increase dragonfly and damselfly diversity and abundance, including an endangered species. *Journal for Nature Conservation* **60**: 125950.

Bergstrom CT, Dugatkin LA (2016). *Evolution*, Second Edition. W. W. Norton & Compagny Inc: New York, USA.

Boudot J-P (2006). *Coenagrion mercuriale. IUCN Red List of threatened species. Version 20141 International Union for Conservation of Nature Available wwwiucnredlistorg*.

Boudot J-P (2020). *Coenagrion mercuriale*. The IUCN Red List of Threatened Species.

Bowen BW (1999). Preserving genes, species, or ecosystems? Healing the fractured foundations of conservation policy. *Molecular Ecology* **8**: S5–S10.

Broquet T, Petit EJ (2009). Molecular estimation of dispersal for ecology and population genetics. *Annual Review of Ecology, Evolution, and Systematics* **40**: 193–216.

Caughley G (1994). Directions in conservation biology. *Journal of Animal Ecology* **63**: 215–244.

Charlesworth B, Charlesworth D (2017). Population genetics from 1966 to 2016. *Heredity* **118**: 2–9.

Charlesworth D, Willis JH (2009). The genetics of inbreeding depression. *Nature Reviews Genetics* **10**: 783–796.

Corbet PS (1957). The life-histories of two summer species of dragonfly (Odonata: Coenagriidae). *Proceedings of the Zoological Society of London* **128**: 403–418.

Couvreur JM, Dufrêne M, Goffart P, Vandevyvre X, Etienne F, Testart D (2008). Nouvelles estimations des effectifs de l'Agrion de Mercure (*Coenagrion mercuriale*) dans la plaine du Biran (Commune de Beauraing, Belgique) avec une analyse des principaux facteurs écologiques expliquant son abondance. *Bulletin de la Société Royale Belge d'Entomologie* **144**: 101–115.

Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, Blaxter ML (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics* **12**: 499–510.

De Knijf G, Billqvist M, van Grunsven RHA, Prunier F, Vinko D, Boudot JP (2023). *Coenagrion mercuriale*. The IUCN Red List of Threatened Species 2023. *IUCN Red List Threat Species 2023*: e.T203582437A220527463.

Des Roches S, Pendleton LH, Shapiro B, Palkovacs EP (2021). Conserving intraspecific variation for nature's contributions to people. *Nature Ecology & Evolution* **5**: 574–582.

Diamond SE, Martin RA (2021). Evolution in cities. *Annual Review of Ecology, Evolution, and Systematicst* **52**: 519–540.

Didham RK, Kapos V, Ewers RM (2012). Rethinking the conceptual foundations of habitat fragmentation research. *Oikos* **121**: 161–170.

Dijkstra K, Lewington R, Jourde P (2015). *Guide des libellules de France et d'Europe*, Delachaux et Niestlé : Neuchâtel Switerland

Dirzo R, Raven PH (2003). Global state of biodiversity and loss. *Annual Review of Environment and Resources* **28**: 137–167.

Donihue CM, Lambert MR (2015). Adaptive evolution in urban ecosystems. *Ambio* **44**: 194–203.

Eckert CG, Samis KE, Lougheed SC (2008). Genetic variation across species' geographical ranges: the central–marginal hypothesis and beyond. *Molecular Ecology* **17**: 1170–1188.

El Mousadik A, Petit RJ (1996). High level of genetic differentiation for allelic richness among populations of the argan tree *Argania spinosa (L.)* Skeels endemic to Morocco. *Theoretical and Applied Genetics* **92**: 832–839.

Ellegren H, Galtier N (2016). Determinants of genetic diversity. *Nature Reviews Genetics* **17**: 422–433.

Elliott JM (2003). A comparative study of the dispersal of 10 species of stream invertebrates. *Freshwater Biology* **48**: 1652–1668.

Ewers RM, Didham RK (2006). Confounding factors in the detection of species responses to habitat fragmentation. *Biological Reviews* **81**: 117–142.

Fahrig L (2003). Effects of habitat fragmentation on biodiversity. *Annual Review of Ecology, Evolution, and Systematics* **34**: 487–515.

Ferreira S, Martínez-Freiría F, Boudot J-P, El Haissoufi M, Bennas N, Alves PC, *et al.* (2015). Local extinctions and range contraction of the endangered *Coenagrion mercuriale* in North Africa. *International Journal of Odonatology* **18**: 137–152.

Fierimonte B, Vanappelghem C (2021). Plan régional d'actions Hauts-de-France en faveur des libellules menacées 2022-2031 - Agir pour la préservation des odonates menacés et de leurs habitats. Conservatoire d'espaces naturels Hauts-de-France et DREAL Hauts-de-France.

Fisher R A (1930). *The genetical theory of natural selection*, Clarendon Press. Oxford, UK.

Frankham R (2015). Genetic rescue of small inbred populations: meta-analysis reveals large and consistent benefits of gene flow. *Molecular Ecology* **24**: 2610–2618.

Frankham R, Ballou JD, Ralls K, Eldridge MDB, Dudash MR, Fenster CB, *et al.* (2017). *Genetic management of fragmented animal and plant populations*, Oxford University Press. Oxford, UK.

Frankham R, Briscoe DA, Ballou JD (2013). *Introduction to conservation genetics*, Second Edition. Cambridge University Press : Cambridge, UK

Funk WC, Lovich RE, Hohenlohe PA, Hofman CA, Morrison SA, Sillett TS, *et al.* (2016). Adaptive divergence despite strong genetic drift: genomic analysis of the evolutionary mechanisms causing genetic differentiation in the island fox (*Urocyon littoralis*). *Molecular Ecology* **25**: 2176–2194.

Funk WC, McKay JK, Hohenlohe PA, Allendorf FW (2012). Harnessing genomics for delineating conservation units. *Trends in Ecology & Evolution* **27**: 489–496.

Fusco NA, Pehek E, Munshi-South J (2021). Urbanization reduces gene flow but not genetic diversity of stream salamander populations in the New York City metropolitan area. *Evolutionary Applications* **14**: 99–116.

Futuyma D, Kirkpatrick M (2018). *Evolution*, Fourth Edition. Oxford University Press: Oxford, UK.

Garner BA, Hoban S, Luikart G (2020). IUCN Red List and the value of integrating genetics. *Conservation Genetics* **21**: 795–801.

Goldstein DB, Pollock DD (1997). Launching microsatellites: a review of mutation processes and methods of phylogenetic inference. *Journal of Heredity* **88**: 335-342.

Gorman GC, Renzi J (1979). Genetic distance and heterozygosity estimates in electrophoretic studies: effects of sample size. *Copeia* 242–249.

Grafius DR, Corstanje R, Siriwardena GM, Plummer KE, Harris JA (2017). A bird's eye view: using circuit theory to study urban landscape connectivity for birds. *Landscape Ecology* **32**: 1771–1787.

Grand D (1996). *Coenagrion mercuriale (Charpentier, 1840). Background information on invertebrates of the Habitats Directive and the Bern Convention, Part II: Mantodea, Odonata, Orthoptera, and Arachnida.*, Council of Europe Publishing (P van Helsdingen, L Willemse, and M Speight, Eds). France (Strasbourg, France).

Haag C, Riek M, Hottinger JW, Pajunen V, Ebert D (2005). Genetic diversity and genetic differentiation in Daphnia metapopulations with subpopulations of known age. *Genetics* **170**: 1809–1820.

Haddad NM, Brudvig LA, Clobert J, Davies KF, Gonzalez A, Holt RD, *et al.* (2015). Habitat fragmentation and its lasting impact on Earth's ecosystems. *Science Advances* **1**: e1500052.

Halfwerk W, Blaas M, Kramer L, Hijner N, Trillo PA, Bernal XE, *et al.* (2019). Adaptive changes in sexual signalling in response to urbanization. *Nature Ecology & Evolution* **3**: 374–380.

Hall P, Orrell LC, Bawa KS (1994). Genetic diversity and mating system in a tropical tree, *Carapa guianensis* (Meliaceae). *American Journal of Botany* **81**: 1104–1111.

Hare MP, Nunney L, Schwartz MK, Ruzzante DE, Burford M, Waples RS, *et al.* (2011). Understanding and estimating effective population size for practical application in marine species management. *Conservation Biology* **25**: 438–449.

Harrisson KA, Pavlova A, Telonis-Scott M, Sunnucks P (2014). Using genomics to characterize evolutionary potential for conservation of wild populations. *Evolutionary Applications* **7**: 1008–1025.

Hobbs RJ, Mooney HA (1998). Broadening the extinction debate: population deletions and additions in California and Western Australia. *Conservation Biology* **12**: 271–283.

Hohenlohe PA, Funk WC, Rajora OP (2021). Population genomics for wildlife conservation and management. *Molecular Ecology* **30**: 62–82.

Holderegger R, Balkenhol N, Bolliger J, Engler JO, Gugerli F, Hochkirch A, *et al.* (2019). Conservation genetics: linking science with practice. *Molecular Ecology* **28**: 3848–3856.

Houard X (2020). Plan national d'actions en faveur des « libellules » - Agir pour la préservation des odonates menacés et de leurs habitats 2020-2030. *Office pour les insectes et leur environnement – DREAL Hauts-de-France - Ministère de la transition écologique*

Hughes JB, Daily GC, Ehrlich PR (1997). Population diversity: its extent and extinction. *Science* **278**: 689–692.

Hunger H, Röske W (2001). Short-range dispersal of the southern damselfly (*Coenagrion mercuriale*: Odonata) defined experimentally using UV fluorescent ink. *Journal for Nature Conservation* **9**: 181–187.

Jax K (2007). Can we define ecosystems? On the confusion between definition and description of ecological concepts. *Acta Biotheoretica* **55**: 341–355.

Johnson MTJ, Munshi-South J (2017). Evolution of life in urban environments. *Science* **358**: eaam8327.

Kannan R, James D (2009). Effects of climate change on global biodiversity: a review of key literature. *Tropical Ecology* **50**: 31-39.

Keller D, Holderegger R (2013). Damselflies use different movement strategies for short- and long-distance dispersal. *Insect Conservation and Diversity* **6**: 590–597.

Keller D, Van Strien MJ, Holderegger R (2012). Do landscape barriers affect functional connectivity of populations of an endangered damselfly? *Freshwater Biology* **57**: 1373–1384.

Keller LF, Waller DM (2002). Inbreeding effects in wild populations. *Trends in Ecology & Evolution* **17**: 230–241.

Laikre L (2010). Genetic diversity is overlooked in international conservation policy implementation. *Conservation Genetics* **11**: 349–354.

Lorenzo-Carballa MO, Ferreira S, Sims AM, Thompson DJ, Watts PC, Cher Y, *et al.* (2015). Impact of landscape on spatial genetic structure and diversity of *Coenagrion mercuriale* (Zygoptera:Coenagrionidae) in northern France. *Freshwater Science* **34**: 1065–1078.

Luck GW, Daily GC, Ehrlich PR (2003). Population diversity and ecosystem services. *Trends in Ecology & Evolution* **18**: 331–336.

Mahdjoub H, Khelifa R, Zebsa R, Bouslama Z, Houhamdi M (2015). Bivoltinism in *Coenagrion mercuriale* (Zygoptera: Odonata) in the southern margin of its distribution range: emergence pattern and larval growth. *African Entomology* **23**: 59–67.

Mahdjoub H, Khelifa R, Zebsa R, Mellal MK, Bouslama Z, Houhamdi M (2014). Aspects of reproductive biology and ecology of *Coenagrion mercuriale* at its southern range margin. *International Journal of Odonatology* **17**: 173–180.

Marko PB, Hart MW (2011). The complex analytical landscape of gene flow inference. *Trends in Ecology & Evolution* **26**: 448–456.

Martens A (2000). Group oviposition in *Coenagrion mercuriale* (Charpentier) (Zygoptera:Coenagrionidae). *Odonatologica* **29**: 329–332.

Miles LS, Dyer RJ, Verrelli BC (2018). Urban hubs of connectivity: contrasting patterns of gene flow within and among cities in the western black widow spider. *Proceedings of the Royal Society* **285**: 20181224.

Miles LS, Rivkin LR, Johnson MTJ, Munshi-South J, Verrelli BC (2019). Gene flow and genetic drift in urban environments. *Molecular Ecology* **28**: 4138–4151.

Mimura M, Yahara T, Faith DP, Vázquez-Domínguez E, Colautti RI, Araki H, *et al.* (2017). Understanding and monitoring the consequences of human impacts on intraspecific variation. *Evolutionary Applications* **10**: 121–139.

Miraldo A, Li S, Borregaard MK, Flórez-Rodríguez A, Gopalakrishnan S, Rizvanovic M, *et al.* (2016). An Anthropocene map of genetic diversity. *Science* **353**: 1532–1535.

Morin PA, Archer FI, Pease VL, Hancock-Hanser BL, Robertson KM, Huebinger RM, *et al.* (2012). An empirical comparison of SNPs and microsatellites for population structure, assignment, and demographic analyses of bowhead whale populations. *Endangered Species Research* **19**: 1–27.

Morin PA, Luikart G, Wayne RK, Group the S workshop (2004). SNPs in ecology, evolution and conservation. *Trends in Ecology & Evolution* **19**: 208–216.

Motte G, Dufrêne M, Mayon N, Goffart P, Barbier Y, Cors R, *et al.* (2021). Liste rouge 2021 des Libellules de Wallonie. *Naturalistes Belges* **102**: 1–21.

Munshi-South J, Zolnik CP, Harris SE (2016). Population genomics of the Anthropocene: urbanization is negatively associated with genome-wide variation in white-footed mouse populations. *Evolutionary Applications* **9**: 546–564.

Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J (2000). Biodiversity hotspots for conservation priorities. *Nature* **403**: 853–858.

Narum S, Buerkle A, Davey JW, Miller MR, Hohenlohe PA (2013). Genotyping-by-sequencing in ecological and conservation genomics. *Molecular Ecology* **22**: 2841–2847.

Naveda-Rodríguez A, Campbell-Thompson E, Watson RT, McCabe J, Vargas FH (2022). Dispersal and space use of captive-reared and wild-rehabilitated harpy eagles released in central American landscapes: implications for reintroduction and reinforcement management. *Diversity* **14**: 886.

Nei M (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of the USA* **70**: 3321–3323.

Nei M, Tajima F (1981). Genetic drift and estimation of effective population size. *Genetics* **98**: 625–640.

Padial JM, Miralles A, De la Riva I, Vences M (2010). The integrative future of taxonomy. *Frontiers in Zoology* **7**: 1-14.

Peñuelas J, Poulter B, Sardans J, Ciais P, van der Velde M, Bopp L, *et al.* (2013). Human-induced nitrogen–phosphorus imbalances alter natural and managed ecosystems across the globe. *Nature Communications* **4**: 2934.

Petit RJ, El Mousadik A, Pons O (1998). Identifying populations for conservation on the basis of genetic markers. *Conservation Biology* **12**: 844–855.

Pinsky ML, Palumbi SR (2014). Meta-analysis reveals lower genetic diversity in overfished populations. *Molecular Ecology* **23**: 29–39.

Prakash S (2021). Impact of climate change on aquatic ecosystem and its biodiversity: an overview. *Journal of Biological Innovations* **3**: 312-317

Prakash S, Lewontin RC, Hubby JL (1969). A molecular approach to the study of genic heterozygosity in natural populations. IV. Patterns of genic variation in central, marginal and isolated populations of *Drosophila pseudoobscura*. *Genetics* **61**: 841–858.

Puillandre N, Modica MV, Zhang Y, Sirovich L, Boisselier M-C, Cruaud C, *et al.* (2012). Large-scale species delimitation method for hyperdiverse groups. *Molecular Ecology* **21**: 2671–2691.

Purse BV (2001). The ecology and conservation of the Southern Damselfly (*Coenagrion mercuriale*). *Thèse de doctorat, Université de Liverpool.*

Purse BV, Hopkins GW, Day KJ, Thompson DJ (2003). Dispersal characteristics and management of a rare damselfly. *Journal of Applied Ecology* **40**: 716–728.

Purse B, Thompson D (2002). Voltinism and larval growth pattern in *Coenagrion mercuriale* (Odonata: Coenagrionidae) at its northern range margin. *European Journal of Entomology* **99**: 11–18.

Purse BV, Thompson DJ (2003a). Reproductive morphology and behaviour in *Coenagrion Mercuriale* (Charpentier) (Zygoptera: Coenagrionidae). *Odonatologica* **32**: 29–37.

Purse BV, Thompson DJ (2003b). Emergence of the damselflies, *Coenagrion mercuriale* and *Ceriagrion tenellum* (Odonata: Coenagrionidae), at their northern range margins, in Britain. *European Journal of Entomology* **100**: 93–99.

Purse BV, Thompson DJ (2005). Lifetime mating success in a marginal population of a damselfly, *Coenagrion mercuriale*. *Animal Behaviour* **69**: 1303–1315.

Purse BV, Thompson DJ (2009). Oviposition site selection by *Coenagrion mercuriale* (Odonata: Coenagrionidae). *International Journal of Odonatology* **12**: 257–273.

Richardson JL, Michaelides S, Combs M, Djan M, Bisch L, Barrett K, *et al.* (2021). Dispersal ability predicts spatial genetic structure in native mammals persisting across an urbanization gradient. *Evolutionary Applications* **14**: 163–177.

Rouquette JR, Thompson DJ (2005). Habitat associations of the endangered damselfly, *Coenagrion mercuriale*, in a water meadow ditch system in southern England. *Biological Conservation* **123**: 225–235.

Rouquette JR, Thompson DJ (2007a). Patterns of movement and dispersal in an endangered damselfly and the consequences for its management: movement and dispersal in a damselfly. *Journal of Applied Ecology* **44**: 692–701.

Rouquette JR, Thompson DJ (2007b). Roosting site selection in the endangered damselfly, *Coenagrion mercuriale*, and implications for habitat design. *Journal of Insect Conservation* **11**: 187–193.

Santangelo JS, Rivkin LR, Johnson MTJ (2018). The evolution of city life. *Proceedings of the Royal Society* **285**: 20181529.

Schlötterer C (2004). The evolution of molecular markers - just a matter of fashion? *Nature Reviews Genetics* **5**: 63–69.

Slatkin M (1985). Gene flow in natural populations. *Annual Review of Ecology, Evolution, and Systematics* **16**: 393–430.

Slatkin M (1987). Gene flow and the geographic structure of natural populations. *Science* **236**: 787–792.

Spurgin LG, Gage MJG (2019). Conservation: the costs of inbreeding and of being inbred. *Current Biology* **29**: R796–R798.

Sternberg K, Buchwald R, Röske W (1999). *Coenagrion mercuriale* (Charpentier, 1840)-Helm Azurjungfer. *Libellen Baden-Württ*: 255–270.

Storfer A, Murphy MA, Spear SF, Holderegger R, Waits LP (2010). Landscape genetics: where are we now? *Molecular Ecology* **19**: 3496–3514.

Thelen C (1992). Untersuchung zum Larvenhabitat und zum Entwicklungszyklus der Helmazurjungfer (*Coenagrion mercuriale*, Zygoptera: Odonata) an zwei verchiedenen Gewassern der Freiburger Bucht. Institut de limnologie de Constance : Université de Fribourg.

Thompson DJ, Rouquette JR, Purse BV (2003). Ecology of the southern damselfly. *Conserving Natura 2000 Rivers Ecology Series*

Thompson DJ, Watts PC (2006). The structure of the *Coenagrion mercuriale* populations in the New Forest, southern England. *Forests and Dragonflies*, Pensoft Publishers, Sofia, Bulgaria; Moscow, Russia. 239–258.

Trumbo DR, Salerno PE, Logan KA, Alldredge MW, Gagne RB, Kozakiewicz CP, *et al.* (2019). Urbanization impacts apex predator gene flow but not genetic diversity across an urban-rural divide. *Molecular Ecology* **28**: 4926–4940.

Vanappelghem C, Hubert B (2010). Suivi de la population de *Coenagrion mercuriale* (Charpentier, 1840) dans la Réserve naturelle régionale des dunes et hauts de Dannes-Camiers (Pas-de-Calais). *Martinia* **3–4**: 131–137.

Vitousek PM, d'Antonio CM, Loope LL, Westbrooks R (1996). Biological invasions as global environmental change. *American scientist* **84**: 468–478.

Vuillaume B, Valette V, Lepais O, Grandjean F, Breuil M (2015). Genetic evidence of hybridization between the endangered native species *Iguana delicatissima* and the invasive *Iguana iguana* (Reptilia, Iguanidae) in the Lesser Antilles: management implications. *PloS One* **10**: e0127575.

Wang J, Santiago E, Caballero A (2016). Prediction and estimation of effective population size. *Heredity* **117**: 193–206.

Waples RS (2005). Genetic estimates of contemporary effective population size: to what time periods do the estimates apply? *Molecular Ecology* **14**: 3335–3352.

Waples RS (2016). Tiny estimates of the *Ne/N* ratio in marine fishes: are they real? *Journal of Fish Biology* **89**: 2479–2504.

Waples RS, Gaggiotti O (2006). What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology* **15**: 1419–1439.

Waples RS, Rus Hoelzel A, Gaggiotti O, Tiedemann R, Palsbøll PJ, Cipriano F, *et al.* (2018). Guidelines for genetic data analysis. *Journal of Cetacean Research and Management* **18**: 33-80.

Watts PC, Kemp SJ, Saccheri IJ, Thompson DJ (2005). Conservation implications of genetic variation between spatially and temporally distinct colonies of the endangered damselfly *Coenagrion mercuriale*. *Ecological Entomology* **30**: 541–547.

Watts PC, Rouquette JR, Saccheri IJ, Kemp SJ, Thompson DJ (2004c). Molecular and ecological evidence for small-scale isolation by distance in an endangered damselfly, *Coenagrion mercuriale*. *Mol Ecol* **13**: 2931–2945.

Watts PC, Rousset F, Saccheri IJ, Leblois R, Kemp SJ, Thompson DJ (2007). Compatible genetic and ecological estimates of dispersal rates in insect (*Coenagrion mercuriale*: Odonata: Zygoptera) populations: analysis of 'neighbourhood size' using a more precise estimator. *Molecular Ecology* **16**: 737–751.

Watts PC, Saccheri IJ, Kemp SJ, Thompson DJ (2006). Population structure and the impact of regional and local habitat isolation upon levels of genetic diversity of the endangered damselfly *Coenagrion mercuriale* (Odonata: Zygoptera). *Freshwater Biology* **51**: 193–205.

Watts PC, Thompson DJ, Kemp SJ (2004b). Cross-species amplification of microsatellite loci in some European zygopteran species (Odonata: Coenagrionidae). *International Journal of Odonatology* **7**: 87–96.

Watts PC, Wu JH, Westgarth C, Thompson DJ, Kemp SJ (2004a). A panel of microsatellite loci for the southern damselfly, *Coenagrion mercuriale* (Odonata: Coenagrionidae). Conservation Genetics **5**: 117–119.

Whitham TG, DiFazio SP, Schweitzer JA, Shuster SM, Allan GJ, Bailey JK, *et al.* (2008). Extending genomics to natural communities and ecosystems. *Science* **320**: 492–495.

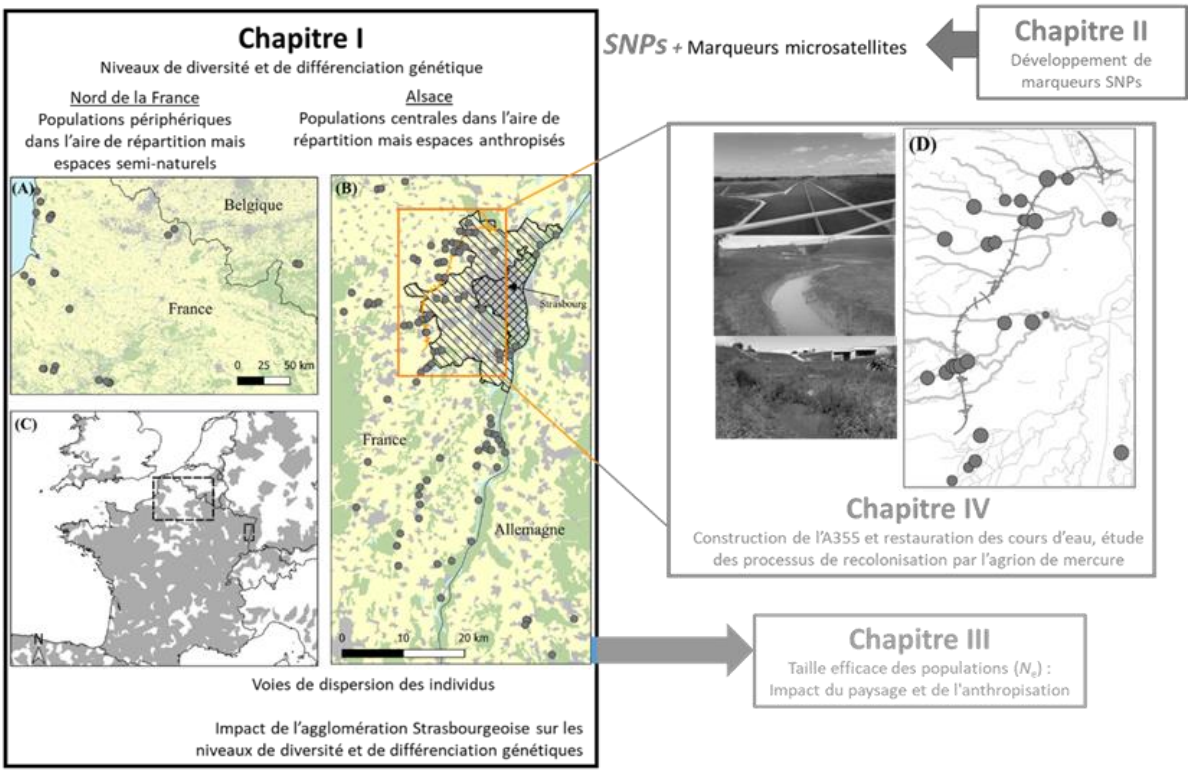
Whitlock MC, McCauley DE (1990). Some population genetic consequences of colony formation and extinction: genetic correlations within founding groups. *Evolution* **44**: 1717–1724.

Wilson EO (1988). Biodiversity. National Academy Press. Washington, DC, USA.

Wilson MC, Chen X-Y, Corlett RT, Didham RK, Ding P, Holt RD, *et al.* (2016). Habitat fragmentation and biodiversity conservation: key findings and future challenges. *Landscape Ecology* **31**: 219–227.

Wright S (1931). Evolution in Mendelian populations. *Genetics* **16**: 97–159.

Chapitre 1. Structure génétique des populations d'agrion de Mercure dans deux régions très contrastées



L'agrion de Mercure est présent sur presque tout le territoire métropolitain, mais ses populations se raréfient à mesure que l'on remonte vers le Nord. Des attendus classiques sont donc une diminution de la diversité génétique et une augmentation de la différenciation génétique entre populations à mesure que l'on s'éloigne du centre de l'aire de répartition géographique de l'espèce. En outre, une augmentation de la consanguinité locale est également prédite en limite d'aire de répartition géographique du fait d'un isolement géographique accru des populations.

À l'aide de marqueurs microsatellites initialement développés par Watts *et al.* (2004ab), nous avons donc cherché à estimer les niveaux de diversité génétique, au travers de l'estimation de l'hétérozygotie attendue ou de la richesse allélique, ainsi que les niveaux de différenciation génétique entre populations de cette espèce en comparant deux régions : l'Alsace où l'espèce est commune, et le nord de la France, où elle est beaucoup plus rarement rencontrée. Nous avons également cherché à déterminer si des groupes génétiques distincts existaient au sein de chacune des deux régions.

A l'échelle plus fine de l'Alsace, nous avons également essayé de déterminer si les patrons de flux de gènes entre populations s'opéraient uniquement le long des cours d'eau, ou également à travers terre. Pour ce faire, nous avons comparé les patrons d'isolement par la distance en utilisant deux

métriques : une distance géographique mesurée au travers des réseaux hydrographiques, et une simple distance euclidienne entre chaque paire de populations. Enfin, nous avons étudié la variation spatiale des taux de migration entre populations à l'échelle de cette région, notamment pour déterminer si la présence de l'agglomération de Strasbourg réduisait les niveaux de diversité génétique ainsi que les flux de gènes entre populations.

Ce chapitre fait l'objet d'un article, qui a été soumis à *Diversity & Distributions* en novembre 2023, puis révisé et resoumis début mars 2024.

Contrasting patterns of spatial genetic structure in endangered southern damselfly (*Coenagrion mercuriale*) populations facing habitat fragmentation and urbanisation.

Agathe Lévêque[1-2], Anne Duputié[1], Vincent Vignon[3], Fabien Duez[1], Cécile Godé[1], Cédric Vanappelghem[4], Jean-François Arnaud[1]

Affiliation :

1. Univ. Lille, CNRS, UMR 8198 - Evo-Eco-Paleo, F-59000 Lille, France

2. Office de Génie Ecologique (O.G.E.), F-67200 Strasbourg, France

3. Office de Génie Ecologique (O.G.E.), F-94100 Saint-Maur-des-Fossés, France

4. Conservatoire d'espaces naturels des Hauts-de-France, F-80480 Dury, France

Short running title: Southern damselfly spatial genetic structure

Abstract

Aim

Human-induced environmental changes result in habitat loss and fragmentation, impacting wildlife population genetic structure and evolution. Urbanised and geographically peripheral areas often represent unfavourable environments, reducing connectivity among populations and causing higher population genetic differentiation and lower intra-population genetic diversity. We examined how geographic peripherality and anthropogenic pressures affect genetic diversity and genetic differentiation in the protected southern damselfly (*Coenagrion mercuriale*, Odonata), which has low dispersal capabilities and specific habitat requirements and whose populations are declining.

Location

We studied two areas: one in semi-natural habitats at the periphery of the species geographic range (northern France) and the other more central to the species' range, in an urbanised area surrounding the city of Strasbourg (Alsace, eastern France).

Methods

We genotyped 2743 individuals from 128 populations using eleven microsatellite loci. We analysed the spatial distribution of neutral genetic diversity (allelic richness, heterozygosity, levels of inbreeding, genetic relatedness), the extent of genetic differentiation, and population affiliations (sPCA analyses) within the two areas. We also examined fine-scale patterns of gene flow in the urbanised area of Alsace by investigating patterns of isolation by distance and estimating effective migration surfaces (EEMS method).

Results

Northern peripheral populations showed lower levels of genetic diversity and higher levels of genetic differentiation than central Alsacian populations. Although located in anthropised habitats, geographically central Alsacian populations showed high levels of gene flow, with dispersal events mainly occurring overland and not restricted to watercourses. However, the highly urbanised city of Strasbourg negatively impacted nearby populations by reducing levels of genetic diversity and increasing population genetic differentiation.

Main conclusions

These results showed the need for management action by restoring breeding sites and creating migratory corridors for peripheral southern damselfly populations. However, our results also highlighted the resilience of southern damselfly in central range populations facing strong urbanisation pressures.

Key Words: central-marginal hypothesis, conservation genetics, central and peripheral populations, genetic diversity, isolation-by-distance, population genetic structure, urbanisation effects.

Introduction

In the context of major environmental changes caused by human activities, such as urbanisation, intensive agriculture or land-use changes, habitat loss and fragmentation are major threats to biodiversity (Fahrig, 2003; Wilson *et al.*, 2016). Fragmented populations face a reduction both in the amount of favourable habitats and in gene flow among populations inhabiting the remaining habitat patches (Fahrig, 2003). This may generate both demographic and genetic effects on populations, with an increased risk of inbreeding, a loss of genetic diversity, and an increase in genetic differentiation among populations, which can ultimately lead to a vortex of extinction in the long term (Frankham *et al.*, 2013; Ellegren and Galtier, 2016).

Urbanisation also impacts the fate of wildlife populations, as described by the "urban fragmentation" model in which habitat fragmentation and degradation magnify the extent of genetic drift and decrease gene flow events among populations (Johnson and Munshi-South, 2017; Miles *et al.*, 2019). This process was indeed documented in numerous species of different taxa (Desender *et al.*, 2005; Vandergast *et al.*, 2006; Munshi-South *et al.*, 2016; Miles *et al.*, 2019; Trumbo *et al.*, 2019; Fusco *et al.*, 2021). Nonetheless, the impact of urbanisation varies greatly, depending on the species' life-history traits but also on the landscape matrix (Miles *et al.*, 2019; Kimmig *et al.*, 2020; Richardson *et al.*, 2021). Urbanisation can not only reduce connectivity between populations, but also create barriers or corridors that influence species' dispersal behaviours and pathways (Kimmig *et al.*, 2020). While highly mobile species are more likely to maintain some connectivity with exurban populations, avoiding inbreeding depression and preserving high levels of genetic diversity, this might not be the case for low vagility species depending on specific habitats and facing dispersal barriers (Fusco *et al*., 2021; Johnson & Munshi-South, 2017; Richardson *et al*., 2021).

Yet, to obtain a global vision of the conservation management of a species, it is important to understand patterns of spatial genetic structure both at local scales and at the larger scale of the species' geographical range (Guo, 2012; DeWoody *et al.*, 2021; Willi *et al.*, 2022). The "abundant centre" model posits that populations near the centre of a species' geographic distribution are the most abundant, and become smaller and scarcer toward the periphery of the range (Brown, 1984). Two main predictions arise from this model: (i) peripheral populations should exhibit lower levels of genetic diversity

compared to central ones, as a result of genetic drift associated with small population size; (ii) peripheral populations are likely to exhibit higher levels of genetic differentiation than central populations, because of genetic drift and of restricted gene flow due to geographic isolation (Vucetich and Waite, 2003; Eckert *et al.*, 2008). Nonetheless, this pattern is not systematically observed, as the geometry of the species' range, historical and current ecological factors can lead to deviations from this model (Eckert *et al.*, 2008; Guo, 2012; Johansson *et al.*, 2013; Pironon *et al.*, 2017).

All these evolutionary considerations are particularly relevant to the Odonata taxa, since they are bioindicator species for freshwater ecosystems and have a biphasic life cycle, with aquatic larval and terrestrial adult stages, making them dependent on both environments for their development. While Odonata species are considered to be efficient fliers, potentially able to disperse over large spatial distances, some Zygoptera species, such as the southern damselfly (*Coenagrion mercuriale*, Charpentier, 1840), have low dispersal capability. Mark-recapture studies have documented rare long-distance movements of southern damselflies over more than one kilometre (Purse *et al.*, 2003; Rouquette and Thompson, 2007; Watts *et al.*, 2004c). The species is geographically restricted to western Europe, with northern range limits in the UK, France, and Belgium, eastern limit in Germany and southern limit in northern Africa (Grand, 1996; see Figure S1). Southern damselflies are under high conservation priority because the species has become almost extinct in seven European countries on the northern and eastern boundaries of its distribution, where many populations are threatened by the deterioration and loss of its habitat associated with changes in agricultural practices (Grand, 1996). The southern damselfly is found in lotic habitats, requiring slow-flowing small streams or ditches, open habitats with little shade, and the presence of helophytes (Rouquette and Thompson, 2005; Purse and Thompson, 2009).

In this study, using a set of microsatellite loci, we aimed to examine the spatial patterns of genetic diversity and genetic differentiation within and among southern damselfly populations in two contrasting areas: (i) an urbanised area located around the city of Strasbourg in Alsace (eastern France), towards the centre of the species' geographic distribution and characterised by the occurrence of numerous populations; (ii) a second area located in northern France and Belgium, towards the northern

limit of the species' distribution, where populations are scarcer and only occur in semi-natural habitats. We asked the following questions:

(1) Do central urbanised and peripheral semi-natural populations of southern damselfly exhibit the same levels of genetic diversity and genetic differentiation? We expected populations in northern France and Belgium to show higher levels of genetic differentiation and lower levels of genetic diversity than populations in Alsace, due to their geographical isolation close to the limit of the species' distribution. However, these differences could be mitigated by a negative effect of urbanisation on Alsacian populations, even though they are more central to the species' range.

(2) When focussing on the Alsace region, where populations are densely distributed, do migration pathways occur overland or through the hydrographic network? We compared the patterns of isolation-by-distance (IBD) using ecological distances measured along watercourses to those found using simple straight-line Euclidean geographic distances. We only conducted this analysis in Alsace, because of a larger occurrence of the species and, consequently, because of a larger number of sampled populations.

(3) Does the urban pressure of the city of Strasbourg impact population genetic features of neighbouring southern damselfly populations? The main expectation was that the city represents a substantial barrier to the dispersal of southern damselflies. This could lead to a decrease in levels of intra-population genetic diversity for populations close to the city, according to the urban fragmentation model (Miles *et al.*, 2019).

Methods

Study sites and sample collection

We sampled two areas: (i) Northern France and Southern Belgium (Figures 1A-C); (ii) the Alsace region, located in eastern France (Figures 1B-C).

In Northern France and southern Belgium, southern damselfly populations are scarce (Figure S1). We thus considered this region to be at the northern limit of the species' geographic distribution. This region is very urbanised with intensive agriculture, and is characterised by a low occurrence of natural areas. In this region, southern damselflies are only found in semi-natural sites (Fierimonte and

Vanappelchem, 2021). The “Conservatoire des Espaces Naturels Hauts de France” (CEN) is in charge of managing biodiversity in more than 540 semi-natural sites in northern France, and surveys additional sites when required. Because southern damselfly is an endangered species in this region, its occurrence is well recorded by biodiversity managers (see Fig. S2A). Over this wide region (~60 000 km²), we exhaustively sampled all 24 semi-natural areas where southern damselfly was documented by the CEN (Fierimonte and Vanappelchem, 2021). These sites are located in six subregions labelled N1 to N6 (Figure 1A; N1: Belgium; N2: Oise; N3: Opal coast; N4: Scarpe; N5: Somme; N6: Vexin).

In Alsace region (and in neighbouring areas), southern damselfly populations are more densely distributed (Figures S1, S2B). We thus considered the Alsace region to be more central to the species’ geographic distribution. In this region, southern damselflies can occur in natural or in more anthropised environments. We thus surveyed an area encompassing the city of Strasbourg. This city spans 78 km² and counts about 290 000 inhabitants, with a human population density of more than 3700 people/km². Strasbourg city is located within the urban area named “Eurométropole de Strasbourg” with about 500 000 inhabitants for 330 km². A few kilometres west are the Vosges mountains, where southern damselflies do not occur. In this region, the landscape is characterised by intensive and less intensive agriculture (mostly cereals and grapevine on the slopes). For this region of about 4 000 km², we accessed to historical records of southern damselfly occurrence (http://association.imago.free.fr/; https://www.gbif.org/fr/species/1422012). We therefore surveyed all known sites where this species was reported, but we also investigated additional sites that appeared to present favourable conditions for southern damselflies, by selecting them on Google Earth, along small streams outside forest areas. In total, we surveyed 206 sites in Alsace (see Fig. S2B), and were able to find southern damselfly populations in 104 of them. Populations were then spatially grouped into six main subwatersheds named A1 to A6 (Figure 1B; A1: Bruche/Mossig; A2: Ehn/Andlau; A3: Dreisan/Elz; A4: Ill/Rhine; A5: Souffel; A6: Zorn/Landgraben).

We collected adult individuals of southern damselfly (*Coenagrion mercuriale*) using an insect net in spring (May to July) from 2017 to 2022 for Northern France and from 2021 to 2022 for the Alsace sampling area. Given the limited dispersal capability of the southern damselfly (Watts *et al*., 2004c;

2006), we designed sampling sites following 200 metres sections located along rivers and streams, separated by at least 500 metres to avoid repeated sampling of the same population (Figure S2B).

We sampled a total of 2743 adult specimens (545 in Northern France, 2198 in Alsace; 1-39 per population; mean 21.4 ± 11.3; Figure 1 and Table S1), with a sex ratio strongly biased towards males (86.73%) due to their higher-flying activities and conspicuousness. Table S1 lists the geographic coordinates and sampling sizes of each population. We collected two kinds of samples: whole individual body, or only the right middle leg of each individual, the latter method being non-lethal and does not impact the damselfly's survival (Fincke and Hadrys, 2001). All samples were stored in 100% ethanol until DNA extraction.

Genotyping

For whole individuals, we removed abdomens from the rest of the body to avoid sampling intestinal microbiota, crushed the samples using five MN Beads Type D (Macherey-Nagel). We crushed leg samples with only three beads. We extracted total genomic DNA from each whole individual sample using NucleoMag® Tissue Kit (Macherey-Nagel) according to the manufacturer's recommendations. For leg samples, we purified DNA with 12µL NucleoMag B-beads and 12µL of pure water, and eluted in 50 µL.

We genotyped all samples using eleven unlinked nuclear microsatellite loci named LIST002, LIST023, LIST034, LIST035, LIST037, LIST042, LIST062, LIST024, LIST060, LIST063, and LIST066, isolated and described in Watts *et al* (2004a,b). Locus LIST060 was excluded from all subsequent analyses because it showed evidences for null alleles occurrence (F_{IS} = 0.377, $P < 10^{-3}$; Table S2). PCR reactions were performed in two multiplexes described in Table S2. PCR reactions were conducted in a volume of 10 µL, using 3 µL of DNA (0.5-5 ng/mL), 5 µL Multiplex PCR Master Mix (QIAGEN) and a primer mix (each primer at a final concentration of 0.2 µM). PCR amplifications were as follows: (i) 15 min at 95°C, (ii) 30 cycles for whole individuals, and 32 cycles for legs of 30 s denaturation at 94°C, 90 s annealing 55°C, and 60 s elongation at 72°C, (iii) 30 min at 60°C. 1.5 µL of PCR products (1/10 diluted for whole individual samples) were pooled with 0.25 µL of RadiantDy™ 632 500 MOB size standard (Eurogentec, Seraing, Belgium) and 9.75 µL of formamide (Applied

Biosystems, Foster City, CA), electrophoresed and sized with an ABI PRISM 3130XL sequencer (Applied Biosystems) and the software GeneMapper 5.0 software (Applied Biosystems). Amplification failure was scarce (0% for Alsace and 0.14% for Northern France).

Genetic diversity

Standard population genetic statistics were calculated using FSTAT 2.9.4 (Goudet, 2003) and R (version 4.2.1; R Core Team, 2022) libraries 'adegenet' 2.1.9 (Jombart, 2008), 'hierfstat' 0.5-11 (Goudet, 2005), 'pegas' 1.1 (Paradis, 2010), and 'demerelate' 0.9 (Kraemer and Gerlach, 2017). All statistics were weighted by population sizes (Weir and Cockerham, 1984; Nei, 1987), allowing for non-biased comparisons across populations and across regions. Moreover, to ensure statistically robust results, we estimated these standard statistics of genetic diversity for only populations with a minimal sampling size of at least eight individuals, allowing to compare relevant genetic diversity estimates between representative samples of populations (El Mousadik and Petit, 1996; Hedrick, 2011; Frankham *et al.*, 2013). These basic statistics included the total number of alleles sampled (A_T), the allelic richness (A_r), the observed (H_o) and the expected (H_e) heterozygosity, and the mean individual kinship coefficient within populations (F_{ij}; Loiselle *et al.*, 1995). Statistical differences between Alsace and Northern France in terms of mean levels of genetic diversity and of mean levels of intra-population kinship were tested by permutation tests using the R library 'perm' 1.0-0.4 (Fay and Shaw, 2010).

To investigate population genetic structure over the whole dataset and among the two study areas and the different subwatersheds and subregions, we computed estimates of *F*-statistics following Weir and Cockerham (1984) and tested their statistical significance using permutation tests (10000 permutations) implemented in FSTAT (Goudet, 2003). As for the calculation of basic genetic diversity estimates, *F*-statistics estimates were only used for populations with at least eight individuals.

To study the potential effect of Strasbourg city on southern damselfly populations, we considered a specific radius of 20 km around the city centre and we performed a linear regression between the distance of each population from the centre of Strasbourg city and i) the genetic diversity estimates (A_r, H_e), ii) the mean individual kinship coefficient within populations (F_{ij}) and iii) the mean pairwise levels of F_{ST} calculated according to Weir and Cockerham (1984) between each focal

population and all other populations within this radius. Because other ecological factors, such as land-use, forest occurrence or the presence of the Vosges mountains are expected to shape the genetic structure at larger scales of investigation, we restricted this analysis to a 20 km radius around the centre of Strasbourg to avoid confounding effects (Figure S3).

Delimitation of population boundaries and genetic discontinuities

Genetic discontinuities and grouping of genetically related populations were assessed using spatial Principal Component Analysis (sPCA). This spatially explicit multivariate method, implemented in the 'adegenet' R library (Jombart, 2008), reveals spatial genetic patterns and detects cryptic geographical variation in allele frequencies allowing to depict population boundaries, without making assumptions with regard to linkage disequilibrium or departures from Hardy-Weinberg equilibrium (Jombart *et al.*, 2008). We used the Delaunay triangulation graph to create the spatial network underlying the sPCA. To draw a comprehensive synthetic representation, we represented population coordinates along the first three principal components on the Red, Green, and Blue colour channels as in Menozzi *et al* (1978).

Analysis of spatial genetic structure

We did not investigate the patterns of isolation-by-distance (IBD) in northern France, both because of the smaller number of sampled populations in this area and because of the difference in geographic scale of sampling. Indeed, because southern damselflies have limited dispersal capabilities, it is not expected to find a migration/drift equilibrium among distant and isolated subregions. Therefore, we studied the fine-scaled spatial genetic structure by only focusing on the whole Alsace region and in the three major subwatersheds, where substantial numbers of populations and individuals ensure biologically relevant results. To search for an IBD pattern expected under migration/drift equilibrium, we considered two different estimations of geographical distances: (i) log-transformed Euclidean distances (d_{Eucli}); (ii) log-transformed shortest path distances along watercourses (d_{Stream}) because southern damselflies only reproduce along watercourses. We computed Euclidean geographical

distances using the R function *dist*, and geographical distances along waterways using the R package 'riverdist' 0.15.5 (Tyers, 2022). To calculate the latter distance, we used the BD TOPAGE® 2019 linear hydrology shapefile. We then assessed the spatial genetic structure using two complementary approaches.

First, to test for IBD among populations, we regressed pairwise F_{ST} estimates against geographical distances and tested the significance of the Mantel statistic *rz* (Smouse *et al.*, 1986) using the *mantel.rtest* function of the R package 'ade4' with 10000 permutations (Dray and Dufour, 2007).

Secondly, to get insight into the spatial genetic structure over the Alsace area, we analysed the relationship between pairwise individual kinship coefficients F_{ij} (Loiselle *et al.*, 1995) and geographical distance using SPAGEDI 1.5 (Hardy and Vekemans, 2002). To allow comparable and relevant statistical results in terms of mean genetic relatedness among individuals, nine classes of increasing geographical distances were defined, which included nearly identical number (mean of 268278 per distance class) of pairwise individual comparisons. In each distance class, we assessed 95% upper and lower confidence intervals of F_{ij} using 10000 permutations of individual locations. Finally, to compare the strength of spatial genetic structure between the different geographical distances and subwatersheds in Alsace, we used the *Sp* statistic (Vekemans and Hardy, 2004).

Estimated Effective Migration Surface (EEMS)

Finally, to identify geographical regions where migration rate was higher or lower than expected under an isolation-by-distance null model, we used the Estimated Effective Migration Surfaces method (EEMS*;* Petkova *et al.*, 2016) in the whole Alsace region. We ran the model for a range of deme values (400, 1000, 1500, and 2000) and parameters were optimised until the proposals were accepted about ~20-30% of the time, as recommended by Petkova *et al.* (2016). Convergence of MCMC chains, checked using the 'reemsplots2' R library (Petkova, 2022), was reached with the MCMC parameters by default (numMCMCIter = 2000000, numBurnIter = 1000000, numThinIter = 9999). This analysis was again restricted to the Alsace region, where populations were more densely distributed.

Results

Genetic diversity

Over all populations, single-locus F_{IS} values did not significantly differ from zero, except for locus LIST037 (F_{IS} = 0.035, $P < 0.05$; Table S2). Single-locus F_{ST} all significantly differed from zero, with a mean multilocus estimate of 0.063 ($P < 0.05$; Table S2). The number of alleles per locus ranged from two to 24, the mean H_e was 0.568, and the mean H_o was 0.507 (Table S2). Within populations, we detected no significant departures from Hardy-Weinberg equilibrium (Table S1).

In Alsace, levels of genetic diversity were high, with a mean allelic richness of 3.19 ± 0.13, a mean H_e of 0.529 ± 0.021 and a mean H_o of 0.528 ± 0.03 (Table S1, Figure 2A, Figure S4). The geographical patterns of these estimates of genetic diversity are displayed in Figure S5. The mean intra-population kinship F_{ij} was 0.022 (Table S1, Figure 2B). Five populations showed higher mean levels of genetic relatedness between individuals than other Alsacian population (populations Gox1, S17, S21, Rlei45, Rsc1; Table S1, locations visible on Figure S3). Gene diversity (H_e) and allelic richness (A_r) increased with increasing distance of populations from the centre of Strasbourg within a radius of 20 km around the city (Figure 3A, Figure S6A). On the contrary, mean pairwise F_{ST} levels and mean intra-population kinship F_{ij} decreased with increasing distance of populations from the centre of Strasbourg (Figures 3B, S6B). This significant relationship observed between genetic diversity or genetic differentiation and the geographical distance from the city of Strasbourg disappear beyond this scale of 20 km, which suggested other ecological factors than strict urbanisation level impacting population genetic features at larger scale of observation (data not shown).

Populations located in Northern France exhibited significantly lower levels of genetic diversity compared to those in Alsace, with a mean allelic richness of 2.91 ± 0.43, a mean H_e of 0.43 ± 0.07, and a mean H_o of 0.48 ± 0.07 (permutation tests, all at P <0.001; Table S1, Figure 2A, Figure S4). The mean intra-population kinship level in Northern France was 0.136 (Table S1, Figure 2B). Within Northern France, levels of genetic diversity and intra-population relatedness considerably varied among the different subregions (Figure S5 and Figure S7), with subregions N1, N2, and N6 showing levels similar

to those observed in the Alsace subwatersheds, while the more isolated subregions N3 to N5 showed lower levels of genetic diversity (permutation tests, $P < 10^{-3}$).

Delimitation of population boundaries and genetic discontinuities

The sPCA showed contrasting geographic partitions of populations for the two studied areas: while a clear spatial structure appeared in Northern France, with genetic grouping corresponding to the major geographical regions, Alsacian populations were rather characterised by a gradient of population delimitation (Figure 4).

In Northern France, the first three axes of the sPCA discriminated four distinct spatial groups: (i) populations located in subregion N3 in the north-western part, (ii) the populations located in subregion N5 in the south-western part, (iii) populations located in subregions N2 and N6 in the southern part of the sampling area, (iv) and finally populations located in subregion N1. The populations in subregion N4 showed a less clear assignment (Figure 4A).

In Alsace, five spatially structured groups could be distinguished by the first three axes of the sPCA (Figure 4B). These groups gather: (i) populations north of Strasbourg, mostly in the A6 subwatershed, (ii) populations west of Strasbourg, within subwatersheds A1, A2 and A5, (iii) populations to the north of subwatershed A4, just south of Strasbourg, (iv) populations of the remainder of subwatershed A4, further south from Strasbourg, (v) and populations located in subwatershed A3 and to the extreme south of subwatershed A4 (Figure 4B). As the boundaries between the different genetic groups did not show clear breaks, these groups reflected a gradient of genetic dissimilarity (Figure 4B).

Levels of genetic differentiation and spatial genetic structure

Populations were highly genetically differentiated in northern France, with a mean multilocus F_{ST} of 0.135 (95% confidence interval [0.113, 0.158]; Table S3, Figure S8A). Significant levels of population genetic differentiation were observed within most subregions (Table S3, Figure S8B). Accordingly, 91.9% of the 210 pairwise population F_{ST} significantly differed from 0 after Bonferroni

correction, most of the non-significant values of F_{ST} being observed among populations belonging to the same geographical regions (Figure S9).

On the contrary, Alsacian populations were weakly, yet significantly genetically differentiated, with a mean multilocus F_{ST} of 0.022 (95% confidence interval [0.020, 0.024]; Table S3, Figure S8A). Significant population differentiation was also observed within each subwatershed, except for very close populations in subwatershed A3 (Table S3, Figure S8B). Only 20.1% of the 3081 pairwise population F_{ST} significantly differed from 0 after Bonferroni correction (Figure S8). Four populations (Gox1, Rlei45, S17, and S21, mapped on Figure S3) accounted for most of the significant pairwise F_{ST} estimates (Figure S10).

When focusing on spatial patterns of genetic differentiation in Alsace, we observed a weak, yet statistically significant, pattern of isolation-by-distance (IBD), as shown by the positive relationship between pairwise population $F_{ST}/1\text{-}F_{ST}$ and log-transformed Euclidean geographical distances ($rz = 0.238$, $P < 0.01$, Figure5A). However, this pattern goes along with an increase in pairwise F_{ST} variance beyond a scale of ~8 km, which suggested a predominant effect of drift beyond this scale of observation. Consistently, genetic relatedness among individuals decreased with increasing geographical distances, with significantly positive kinship coefficients occurring up to 11 km (Figure 5C). Interestingly, when distances among populations were calculated along watercourses (d_{Stream}), this pattern of IBD disappeared ($rz = 0.045$, $P = 0.136$, Figure 5B). However, positive F_{ij} values still occurred for up to 20 km along streams (Figure 5D).

Within subwatersheds, IBD was stronger using Euclidean distances compared to aquatic watercourse (d_{Stream}) distances in subwatersheds A1, A4 and A5 (respectively $rz = 0.374$, 0.414 and 0.472; $P < 0.0001$, $P < 0.05$ and $P < 0.0001$; Figure S11). Within subwatersheds, there was a general short-distance spatial autocorrelation pattern, respectively up to 2.2 km, 4.6 km and <1 km in watersheds A1, A4 and A5 (Figure S11). A significant correlation between pairwise population genetic differentiation estimates and d_{Stream} was only found in subwatersheds A1 and A4 (Figure S11D-F). Yet, pairwise individual kinship coefficients decreased with aquatic distances (d_{Stream}), and were significantly positive for up to < 500 m, 4 km and about 700 m respectively in subwatersheds A1, A4 and A5 (Figure S12).

Estimated Effective Migration Surface (EEMS)

The results with 1500 demes represented the best balance between precision and calculation time. The EEMS plot of effective migration rates identified an area of reduced migration rates covering most of the city of Strasbourg, and extending to the South (Figure 6). This notably encompasses populations located in the north of Strasbourg city (populations named S17, S21, CCR4, CCR2). Three additional areas showed lower migration rates than expected under an IBD at equilibrium. The first one was located downstream along subwatershed A5. The second was found in the midstream of subwatershed A1. The last area of restricted migration was centred on population Gox1, the population most upstream of subwatershed A2 (Figure 6).

Discussion

Patterns of genetic diversity and population genetic differentiation between central and peripheral populations

Consistent with the first prediction of the "abundant centre" hypothesis (Eckert *et al.*, 2008; Guo, 2012), peripheral populations in northern France showed significantly lower levels of genetic diversity as compared to central Alsacian populations. In addition, the two areas displayed different spatial distributions of genetic diversity: while indices of intra-population genetic diversity were constantly high over all populations located in Alsace, they widely differed between the surveyed subregions in Northern France. Indeed, at a finer scale in Northern France, populations within the most peripheral, isolated, and small habitat patches of the species distribution (N3, N4, N5) showed lower levels of genetic diversity than those from larger and more central patches (N1, N2, N6; Figure 1).

Also consistent with the second prediction of the "abundant centre" (Eckert *et al.*, 2008; Pironon *et al.*, 2017), the mean genetic differentiation in Alsace was lower than in Northern France. This pattern observed for marginal populations may be due to larger geographical scale of population separation in Northern France. Yet, similarly high levels of population differentiation were previously observed in other peripheral regions, even at small geographical scales (see Watts *et al.*, 2004c, 2005; Lorenzo-

Carballa *et al.*, 2015). Indeed, indirect genetic approaches indicated dispersal up to 2 km and small-scale population genetic structuring (Watt *et al.*, 2004c; 2006). Besides, the sPCA analysis could not reveal any major disjunction across the Alsace area. Both the low levels of population differentiation and low genetic relatedness within Alsacian populations support the idea that there is a substantial amount of gene flow among Alsacian populations (Ellegren and Galtier, 2016). By contrast, in Northern France, the sPCA analysis revealed striking genetic discontinuities between four large spatial groups of populations and suggested restricted gene flow among populations even at small scales, as exemplified by high levels of genetic differentiation observed within isolated sets of populations located in subregions N3, N4, N5.

In Northern France, gene flow may be reduced over short distances, as illustrated by population Bai3 in subregion N3, which showed low levels of genetic diversity and high levels of genetic differentiation with close neighbouring populations (< 5 km). Our findings thus reiterate the claim of Lorenzo-Carballa *et al.* (2015), namely that habitat management should improve connectivity with geographically close populations to prevent further loss of genetic diversity. Interestingly, no sign of inbreeding was detected, even in isolated populations or in populations characterised by low levels of genetic diversity, a result consistent with previous reports on southern damselflies (Watts *et al.*, 2004c; Keller *et al.*, 2012).

Altogether, the spatial distribution of genetic diversity and the patterns of genetic differentiation observed in the two regions are thus in agreement with the "abundant centre" hypothesis (Pironon *et al.*, 2017). This pattern, observed towards the northern limit, may exist in other cardinal directions as well, because peripheral populations of the southern damselfly are declining in their northern, eastern (Grand, 1996) and southernmost limits in North Africa (Ferreira *et al.*, 2015). It may also result from colonisation history (see Johansson *et al.*, 2013), and more extensive sampling would certainly provide a more global view of the evolution of genetic structuring in populations of the southern damselfly.

Spatial genetic structure and dispersal pathways

Although the level of genetic differentiation among southern damselfly populations was weak across the Alsace area, we found an isolation-by-distance (IBD) pattern. This pattern was reflected both

by an increase in genetic differentiation among populations and a decrease of pairwise individual kinship with increasing Euclidean geographical distances. This result is consistent with previous studies that observed an IBD pattern in southern damselfly populations both at a fine scale (Watts *et al.*, 2004c; Watts and Thompson, 2012; Lorenzo-Carballa *et al.*, 2015) and at a broader scale of investigation (Watts *et al.*, 2006; Watts and Thompson, 2012).

Population genetic structure also depends on the biological attributes of a species, in particular its dispersal capabilities and habitat requirements (Hutchison and Templeton, 1999; Phillipsen *et al.*, 2015; Ellegren and Galtier, 2016). The relationship between genetic and geographic distances across the Alsace area suggested that gene flow events among southern damselfly populations may occur in a stepping-stone pattern among patches of suitable habitats (Kimura, 1953; Kimura and Weiss, 1964; Hutchison and Templeton, 1999), consistent with the patchy distribution of habitats of this species (Conrad *et al.*, 1999; Rouquette and Thompson, 2005; Purse and Thompson, 2009). Adults can disperse among suitable habitat patches, but since they are poor fliers (Watts *et al.*, 2004c), dispersal is likely to occur mostly at short distances between neighbouring populations.

The IBD patterns were stronger at the smaller scale of subwatersheds compared to the one observed across the whole Alsace region. In addition, variance in pairwise genetic differentiation in Alsace strongly increased beyond geographical distances larger than 8 km. These results suggest that at large geographical distances, gene flow among populations are influenced by ecological factors additional to the simple Euclidean distances, suggesting potential barriers to gene flow occurring over large scales over 8 km (e.g. forested or elevated areas; Hutchison & Templeton, 1999; Phillipsen *et al.*, 2015).

In this respect, the geographical distances strictly calculated along watercourses only partly explained the levels of genetic differentiation observed among populations. This suggests that dispersal events can occur overland and is not restricted to stream networks. This overland dispersal between different streams is further supported by the sPCA analysis that suggested no clear genetic discontinuity between the different river networks. These results are in line with the results of Keller and Holderegger (2013), where different movement strategies were described in southern damselflies for short- and long-distance dispersal: dispersal along watercourses appeared to be favoured over short distances (< 3 km),

whereas dispersal along straight lines was best supported for long distances (> 3 km). Altogether, a simple IBD is not likely to explain the totality of the spatial genetic structure we observed in Alsace, pinpointing the possible role of landscape features in shaping dispersal among populations (Storfer *et al.*, 2007).

Urbanisation effects on spatial genetic structure

The city of Strasbourg likely acts as a barrier to gene flow between southern damselfly populations in Alsace. Indeed, the EEMS analysis clearly showed a reduced migration rate in the vicinity of the city. This result was further supported by increasing levels of population genetic differentiation and intra-population kinship coefficients in the vicinity of the city of Strasbourg. A typical "urban fragmentation" pattern may thus be observed; urbanised areas appeared to fragment habitats, reducing gene flow and increasing genetic differentiation among populations (Lorenzo-Carballa *et al.*, 2015; Johnson and Munshi-South, 2017; Lourenço *et al.*, 2017; Miles *et al.*, 2019). However, while the city of Strasbourg represented a clearly restricted migration zone, peripheral small towns and villages did not suggest any major barriers to gene flow. Therefore, the impact of urbanisation on patterns of gene flow likely differs depending on the size and level of urbanisation (e.g. Trumbo *et al*., 2019). The fragmentation impact of the city of Strasbourg can extend to a large distance outside the city centre, because the urban area extends over more than 300 km², up to 12 km away from the city centre. Further away from the city, restricted migration among populations may be due to other landscape ecology features, such as forest patches or high-elevation areas. As a matter of fact, the significant relationship observed between levels of genetic diversity and genetic differentiation and the geographical distance from the city disappeared beyond a scale of 20 km. This finding goes along with a general increase in pairwise F_{ST} variance with increasing geographical distance, which suggested additional landscape effects other than urban areas, such as topography and land-use (Hutchison and Templeton, 1999; Storfer *et al.*, 2007).

Moreover, the “urban fragmentation” model further postulates that fragmentation and isolation of populations by unfavourable urbanised areas could increase the strength of genetic drift, which erodes neutral genetic diversity within populations (Miles *et al.*, 2019). Levels of allelic richness and gene

diversity indeed decreased in the direct neighbourhood of Strasbourg city, thus suggesting a negative relationship between urbanisation and genetic diversity, as it was described in several species (Noël *et al.*, 2007; Munshi-South *et al.*, 2016; Johnson and Munshi-South, 2017; Miles *et al.*, 2019; Kimmig *et al.*, 2020). Nonetheless, a reduction in gene flow linked to urbanisation is not systematically associated with a reduction in genetic diversity (Trumbo *et al.*, 2019; Fusco *et al.*, 2021): indeed, population S21 surprisingly showed high levels of genetic diversity despite a substantial genetic differentiation with close populations and its proximity to Strasbourg city. This could be explained by incoming migrant individuals from diverse sources outside our sampling area (Whitlock and McCauley, 1990).

Our findings on city impacts on population genetic features again require caution before being generalised, and would need to be repeated in different urban areas in order to confirm that populations evolve in the same way in response to urbanisation (see Santangelo *et al.*, 2018; Miles *et al.*, 2019; Diamond and Martin, 2021). Incorporating landscape components among populations, along with the use of more resolutive genome-wide molecular markers, should be pursued to better understand the impact of urbanisation and landscape features on southern damselfly populations at a fine scale. This may also help to explain the occurrence of zones of restricted migration observed in rural areas. Furthermore, cities can exhibit considerable spatial and temporal heterogeneity at fine scale, which could influence population adaptation (Rivkin *et al.*, 2019). It would therefore be useful to study the adaptive variation that can influence the evolution of species in urban environments (Johnson and Munshi-South, 2017; Rivkin *et al.*, 2019; Diamond and Martin, 2021; Babik *et al.*, 2023). Indeed, in another species of damselfly, *Coenagrion puella*, urbanisation and associated heat islands seem to select for slower development and higher survival (Tüzün *et al.*, 2017). Likewise, in southern damselflies, development time is also affected by water temperature associated with the release of industrial cooling waters (Thelen, 1992).

Conclusion

Central populations of southern damselflies located in urbanised environments in Alsace exhibited higher levels of genetic diversity and lower levels of genetic differentiation compared to

northern peripheral populations, indicating that the species might follow a classical "abundant centre" model in terms of population genetic structure (Guo, 2012; Pironon *et al.*, 2017). Populations found in isolated suitable habitats located at high latitudes exhibited particularly low levels of genetic diversity, which is alarming with regard to their evolutionary potential and may threaten their long-term persistence. In contrast, central Alsacian populations displayed a metapopulation structure with high intra-population levels of genetic diversity and a high amount of gene flow among populations. These findings contrast with previous studies on southern damselflies that suggested very low dispersal capabilities and high levels of genetic differentiation at fine scales of investigation (Watts *et al.*, 2004c; Watts and Thompson, 2012; Lorenzo-Carballa *et al.*, 2015). The highly urbanised city of Strasbourg likely represents a substantial barrier to gene flow among southern damselfly populations. Although this barrier effect of urban areas on southern damselfly dispersal has already been suggested by previous work conducted at a smaller scale (Lorenzo-Carballa *et al.*, 2015; Watts *et al.*, 2004c), our study is the first to clearly show a progressive decrease in genetic diversity in populations near a major urban area. The impact of urbanisation may also vary according to the type and size of urban areas encountered, as smaller Alsacian villages did not seem to impede gene flow among populations. Finally, Alsacian populations matched a classical pattern of isolation-by-distance population genetic structure. This suggests that individuals do not disperse exclusively along watercourses, their specific breeding habitat, but are also able to disperse over land. However, Euclidean distance alone did not fully explain the levels of genetic differentiation and the impact of the different landscape features occurring in the Alsace region remains to be determined. A genomic approach using SNP markers coupled with a landscape analysis would provide a more detailed and comprehensive vision of the evolutionary processes influencing patterns of gene flow and levels of genetic diversity occurring in this area of the southern damselfly geographical distribution.

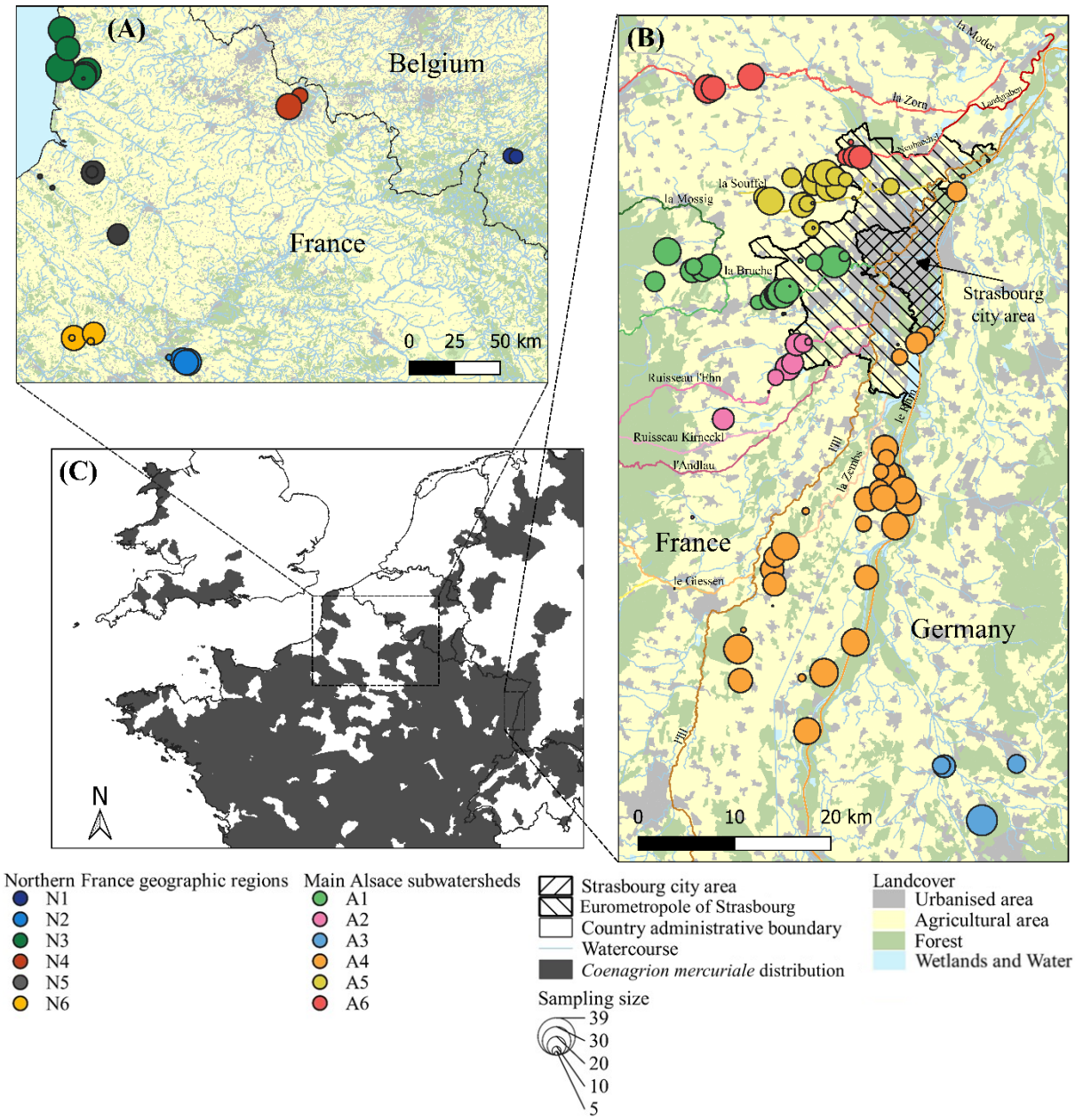


Figure 1: Location of 128 populations of southern damselfly (*Coenagrion mercuriale*) within two study areas: Northern France **(A)**, and Alsace **(B)**. Each circle represents one population, with circle size being proportional to the number of individuals sampled. Strasbourg city and the eurometropole of Strasbourg are represented by the striped areas. Around Strasbourg, main watercourses are represented and coloured depending on the hydrographic network they contribute to. Circle colours correspond to the main subwatershed or geographical region each population belongs to. Land cover was simplified from Corine Land Cover Edition 2018. **(C)** The grey zones represent the current geographic distribution of *Coenagrion mercuriale* (data obtained from IUCN SSC Odonata Specialist Group 2019. The IUCN Red List of Threatened Species. Version 2022-2).

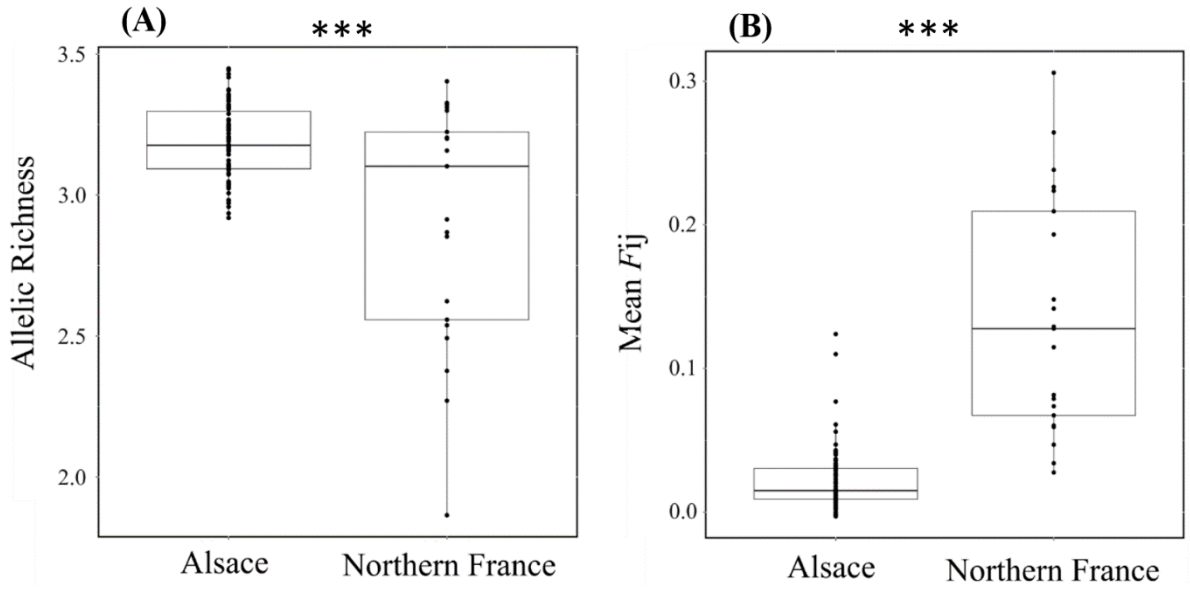


Figure 2: Distribution of allelic richness A_r **(A)** and mean kinship coefficient F_{ij} **(B)** in the southern damselfly within the Alsace and Northern France areas. ***: $P < 0.001$.

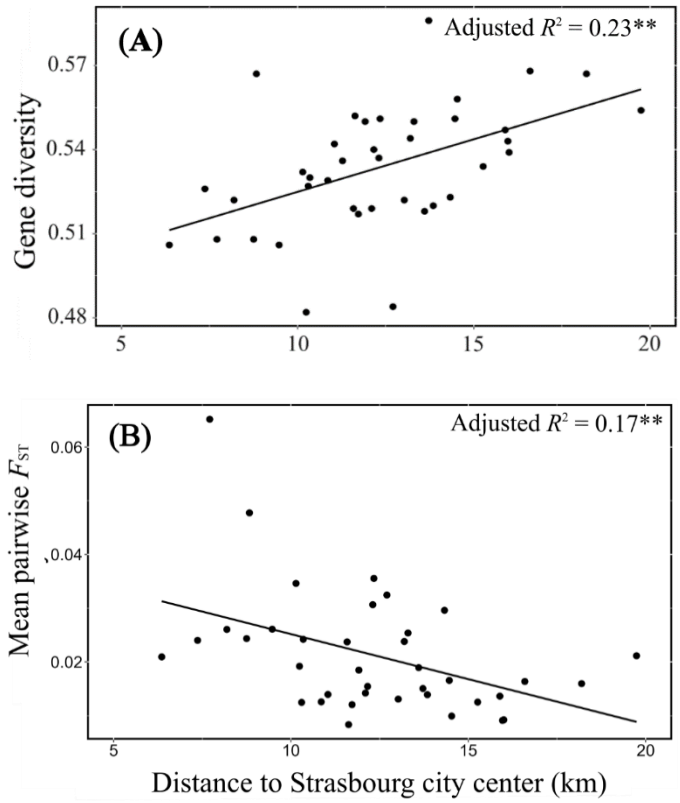


Figure 3: Relationship between the distance of populations to the centre of Strasbourg within a radius of 20 km buffer (see Figure S3) and the genetic diversity (H_e) **(A)** and the mean pairwise F_{ST} calculated according to Weir and Cockerham (1984) between each population and other populations within this radius **(B)**. Lines show the linear model. **: $P < 0.01$

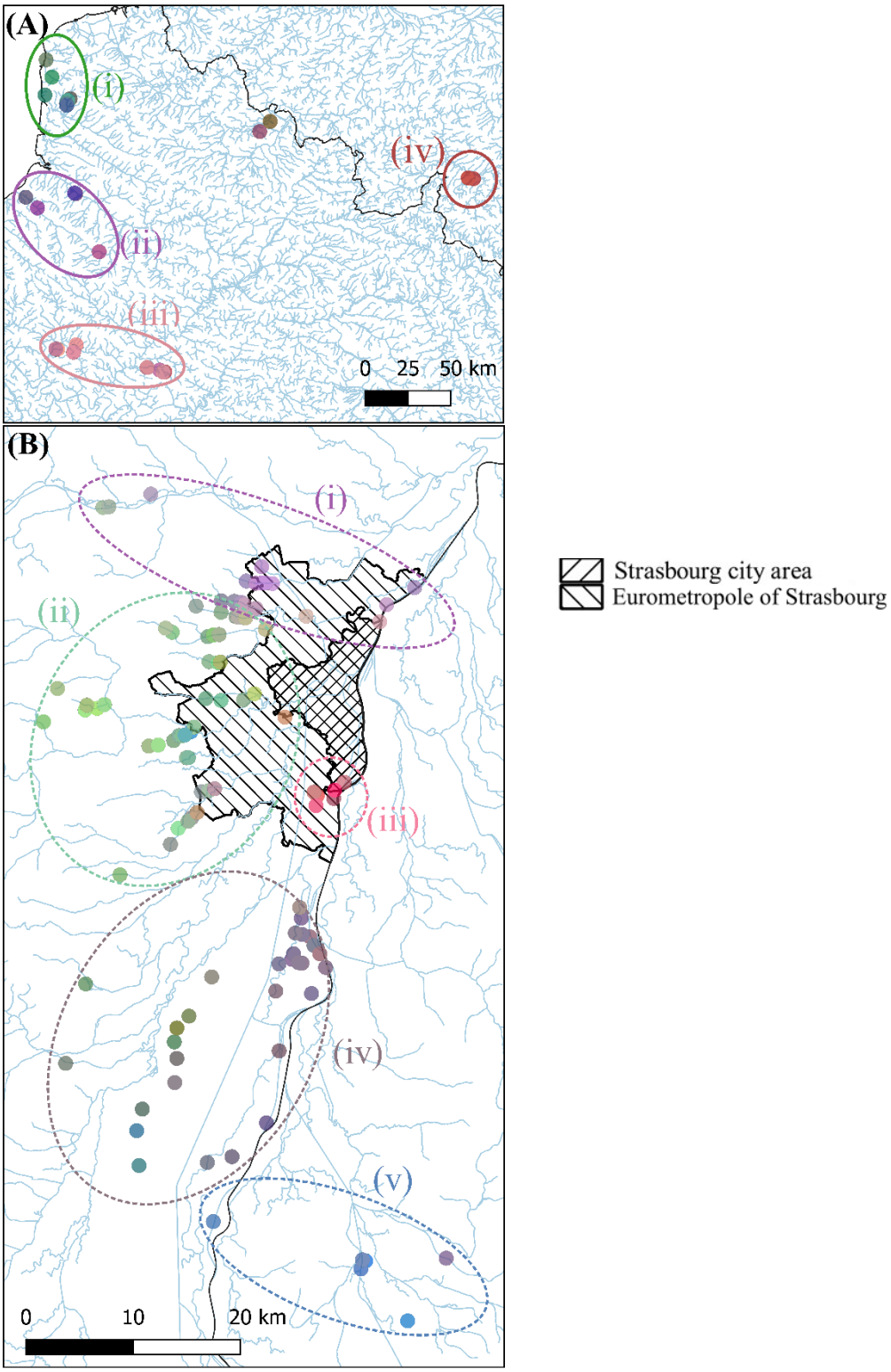


Figure 4: sPCA analyses depicting population genetic affiliation in Northern France **(A)** and Alsace **(B)**. Population colours reflect their coordinates on the first three axes of the sPCA, with coordinates along these three axes assigned to the Red, Green, and Blue channels respectively. Broad groupings of populations can be visualised by dashed labelled ellipses. Strasbourg city and the

eurometropole of Strasbourg are represented by the striped areas. Main watercourses are also represented. Watercourses from COPERNICUS Land Monitoring Service, 2019: EU-Hydro.

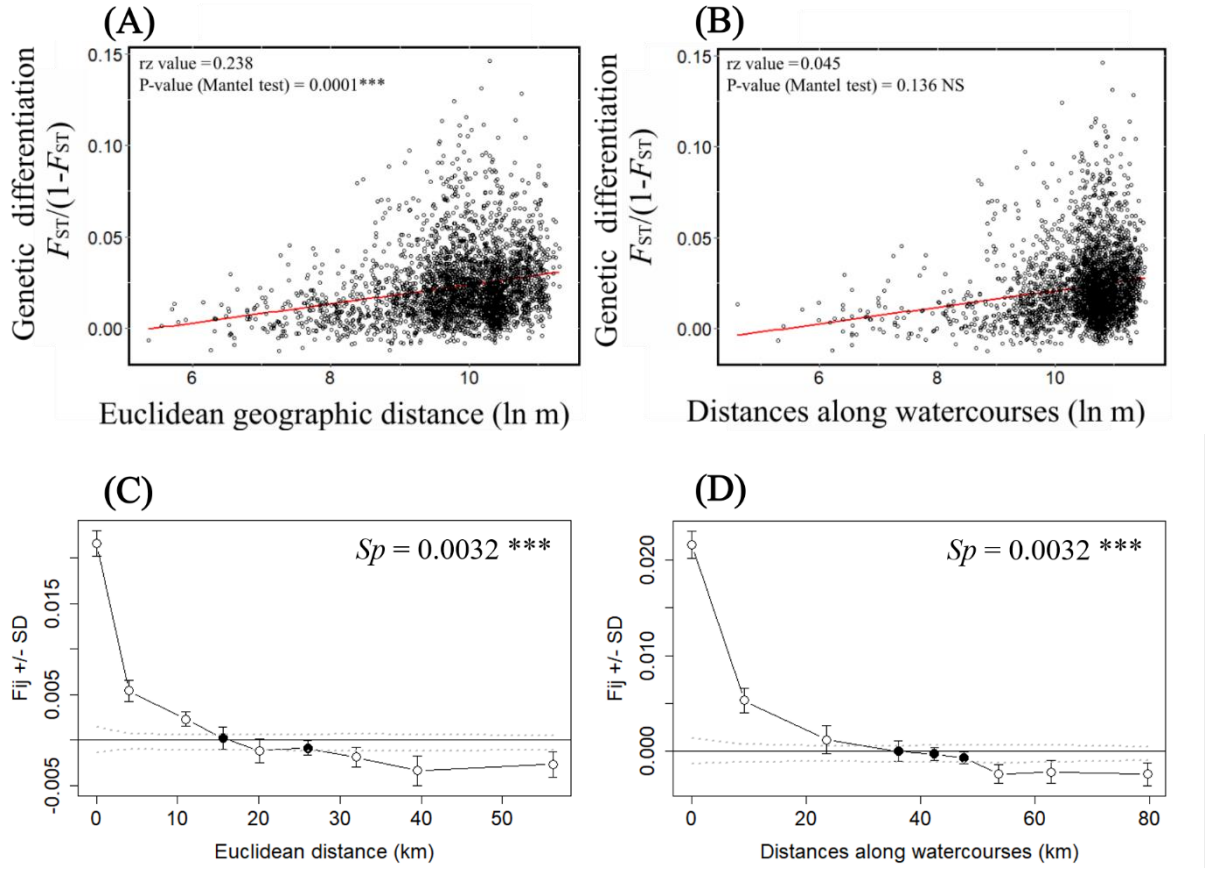


Figure 5: Spatial genetic structure in southern damselfly populations across the whole Alsace area. Relationship between levels of genetic differentiation ($F_{ST}/(1-F_{ST})$) and log-transformed Euclidean distance (d_{Eucli}) **(A)** or watercourse geographical distance (d_{Stream}) **(B)**. Average pairwise kinship coefficients (F_{ij}, Loiselle *et al.*, 1995) between individuals are plotted against increasing classes of geographical Euclidean distances **(C)** or of waterways distance **(D)**. Standard errors for F_{ij} for each distance class were obtained by jackknifing over loci. Dashed lines indicate the upper and lower 95% confidence intervals of non-significant spatial genetic structure, with white dots indicating significant estimates outside this confidence interval. *Sp* statistics based on the regression with ln(distance) and their statistical significance are also indicated.

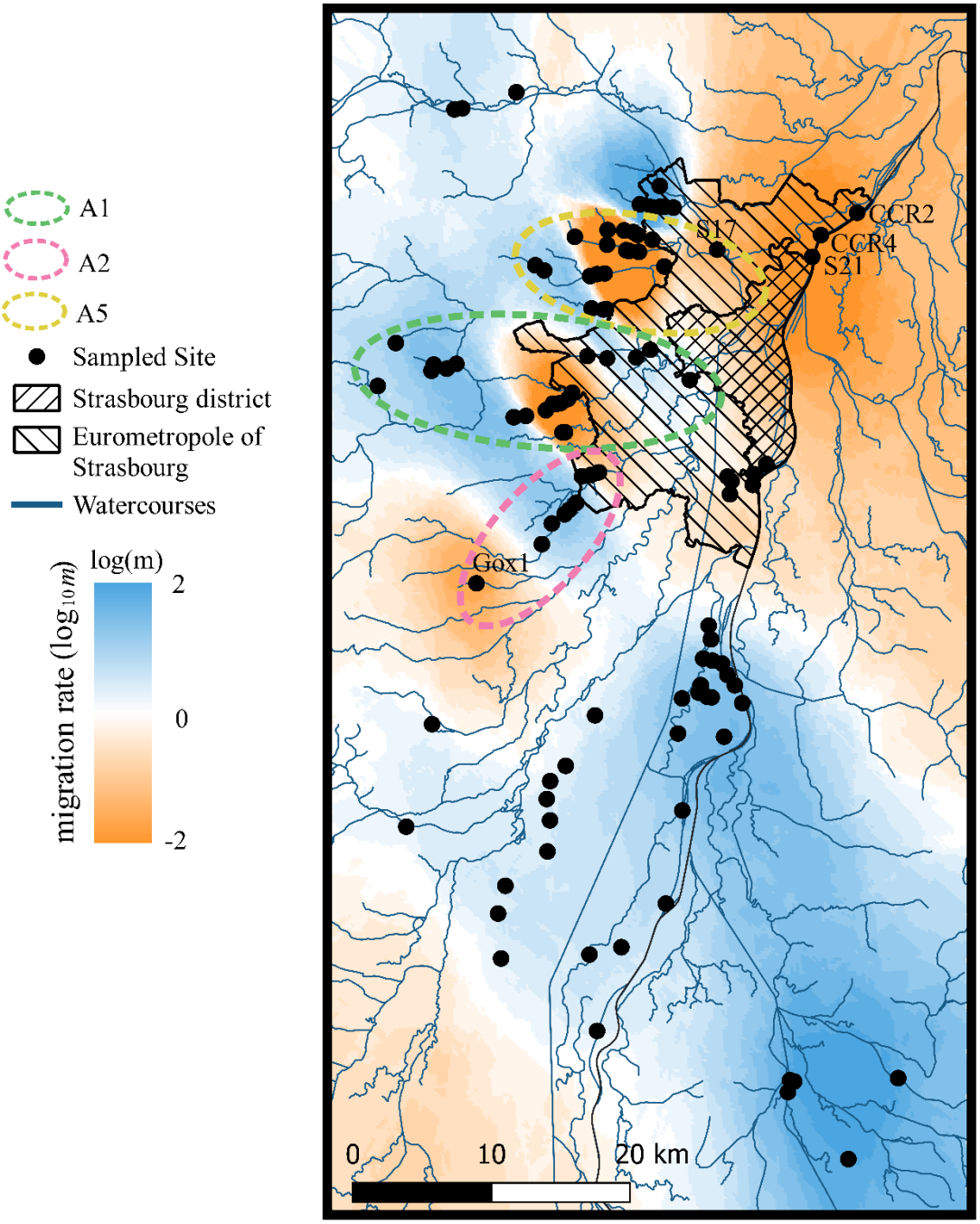


Figure 6: Estimated effective migration rates in southern damselfly in the Alsace region, as visualised by interpolated surface of the posterior mean migration rates *m* (on a log_{10} scale), where darker shades of blue indicate greater than expected migration (0-2), darker shades of orange indicate lower than expected migration (-2 to 0), and white indicates the null hypothesis of isolation-by-distance (0). The black circles indicate population locations. Main watercourses around Strasbourg City (blue lines) are also represented. The Strasbourg district area and the eurometropole of Strasbourg are represented by the striped areas. The locations and names of specific populations cited in the text (Gox1, S17, S21,

CCR4, CCR2 were added to the map. Contours of the A1, A2, and A5 sub-catchments are indicated by dotted lines.

References

Babik, W., Dudek, K., Marszałek, M., Palomar, G., Antunes, B., & Sniegula, S. (2023). The genomic response to urbanization in the damselfly *Ischnura elegans*. *Evolutionary Applications*, *16* (in press).

Brown, J. H. (1984). On the relationship between abundance and distribution of species. *The American Naturalist*, *124*, 255-279.

Conrad, K. F., Willson, K. H., Harvey, I. F., Thomas, C. J., & Sherratt, T. N. (1999). Dispersal characteristics of seven odonate species in an agricultural landscape. *Ecography*, *22*, 524-531.

Desender, K., Small, E., Gaublomme, E., & Verdyck, P. (2005). Rural-urban gradients and the population genetic structure of woodland ground beetles. *Conservation Genetics*, *6*, 51-62.

DeWoody, J. A., Harder, A. M., Mathur, S., & Willoughby, J. R. (2021). The long-standing significance of genetic diversity in conservation. *Molecular Ecology*, *30*, 4147-4154.

Diamond, S. E., & Martin, R. A. (2021). Evolution in cities. *Annual Review of Ecology, Evolution, and Systematics*, *52*, 519-540.

Dray, S., & Dufour, A.-B. (2007). The ade4 Package : Implementing the Duality Diagram for Ecologists. *Journal of Statistical Software*, *22*, 1-20.

Eckert, C. G., Samis, K. E., & Lougheed, S. C. (2008). Genetic variation across species' geographical ranges : The central–marginal hypothesis and beyond. *Molecular Ecology*, *17*, 1170-1188.

El Mousadik, A., & Petit, R. J. (1996). High level of genetic differentiation for allelic richness among populations of the argan tree *Argania spinosa (L.)* Skeels endemic to Morocco. *Theoretical and Applied Genetics*, *92*, 832-839.

Ellegren, H., & Galtier, N. (2016). Determinants of genetic diversity. *Nature Reviews Genetics*, *17*, 422-433.

Fahrig, L. (2003). Effects of habitat fragmentation on biodiversity. *Annual Review of Ecology, Evolution, and Systematics*, *34*, 487-515.

Fay, M., & Shaw, P. (2010). Exact and asymptotic weighted logrank tests for interval censored data: The interval R Package. *Journal of Statistical Software*, *36*, 1-34.

Ferreira, S., Martínez-Freiría, F., Boudot, J.-P., El Haissoufi, M., Bennas, N., Alves, P. C., Watts, P. C., Thompson, D. J., & Brito, J. C. (2015). Local extinctions and range contraction of the endangered *Coenagrion mercuriale* in North Africa. *International Journal of Odonatology*, *18*, 137-152.

Fierimonte B, Vanappelchem C (2021). Plan régional d'actions Hauts-de-France en faveur des libellules menacées 2022-2031 - Agir pour la préservation des odonates menacés et de leurs habitats. Conservatoire d'espaces naturels Hauts-de-France et DREAL Hauts-de-France : 80 p.

Fincke, O. M., & Hadrys, H. (2001). Unpredictable offspring survivorship in the damselfly, Megaloprepus *coerulatus*, shapes parental behavior, constrains sexual selection, and challenges traditional fitness estimates. *Evolution*, *55*, 762-772.

Fincke OM, Hadrys H (2001). Unpredictable offspring survivorship in the damselfly, *Megaloprepus coerulatus*, shapes parental behavior, constrains sexual selection, and challenges traditional fitness estimates. *Evolution* **55**: 762–772.

Frankham, R., Briscoe, D. A., & Ballou, J. D. (2013). *Introduction to conservation genetics* (2d edition, Cambridge University Press).

Fusco, N. A., Pehek, E., & Munshi-South, J. (2021). Urbanization reduces gene flow but not genetic diversity of stream salamander populations in the New York City metropolitan area. *Evolutionary Applications*, *14*, 99-116.

Goudet, J. (2003). *FSTAT (ver. 2.9.4), a program to estimate and test population genetics parameters.*

Goudet, J. (2005). Hierfstat, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Notes*, *5*, 184-186.

Grand, D. (1996). *Coenagrion mercuriale (Charpentier, 1840). Background information on invertebrates of the Habitats Directive and the Bern Convention, Part II: Mantodea, Odonata, Orthoptera, and Arachnida*. (P. van Helsdingen, L. Willemse, & M. Speight, Eds.; Council of Europe Publishing, Vol. 80).

Guo, Q. (2012). Incorporating latitudinal and central-marginal trends in assessing genetic variation across species ranges. *Molecular Ecology*, *21*, 5396-5403.

Hardy, O. J., & Vekemans, X. (2002). SPAGeDi : A versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes*, *2*, 618-620.

Hedrick PW (2011). *Genetics of Populations* (4th edition, Sudbury, MA: Jones and Bartlett Publishers, Inc.)

Hutchison, D. W., & Templeton, A. R. (1999). Correlation of pairwise genetic and geographic distance measures : Inferring the relative influences of gene flow and drift on the distribution of genetic variability. *Evolution*, *53*, 1898-1914.

Johansson, H., Stoks, R., Nilsson-Örtman, V., Ingvarsson, P. K., & Johansson, F. (2013). Large-scale patterns in genetic variation, gene flow and differentiation in five species of European Coenagrionid damselfly provide mixed support for the central-marginal hypothesis. *Ecography*, *36*, 744-755.

Johnson, M. T. J., & Munshi-South, J. (2017). Evolution of life in urban environments. *Science*, *358*.

Jombart, T. (2008). adegenet : A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*, 1403-1405.

Jombart, T., Devillard, S., Dufour, A.-B., & Pontier, D. (2008). Revealing cryptic spatial patterns in genetic variability by a new multivariate method. *Heredity*, *101*, 92-103.

Keller, D., & Holderegger, R. (2013). Damselflies use different movement strategies for short- and long-distance dispersal. *Insect Conservation and Diversity*, *6*, 590-597.

Keller, D., Van Strien, M. J., & Holderegger, R. (2012). Do landscape barriers affect functional connectivity of populations of an endangered damselfly? *Freshwater Biology*, *57*, 1373-1384.

Kimmig, S. E., Beninde, J., Brandt, M., Schleimer, A., Kramer-Schadt, S., Hofer, H., Börner, K., Schulze, C., Wittstatt, U., Heddergott, M., Halczok, T., Staubach, C., & Frantz, A. C. (2020). Beyond the landscape : Resistance modelling infers physical and behavioural gene flow barriers to a mobile carnivore across a metropolitan area. *Molecular Ecology*, *29*, 466-484.

Kimura, M. (1953). Stepping stone' model of population. *Annual report of the National Institute of Genetics Japan*, *3*, 62-63.

Kimura, M., & Weiss, G. H. (1964). The stepping stone model of population structure and the decrease

Kraemer, P., & Gerlach, G. (2017). Demerelate : Calculating interindividual relatedness for kinship analysis based on codominant diploid genetic markers using R. *Molecular Ecology Resources*, *17*, 1371-1377.

Loiselle, B. A., Sork, V. L., Nason, J., & Graham, C. (1995). Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *American Journal of Botany*, *82*, 1420-1425.

Lorenzo-Carballa, M. O., Ferreira, S., Sims, A. M., Thompson, D. J., Watts, P. C., Cher, Y., Damoy, V., Evrard, A., Gelez, W., & Vanappelghem, C. (2015). Impact of landscape on spatial genetic structure and diversity of *Coenagrion mercuriale* (Zygoptera:Coenagrionidae) in northern France. *Freshwater Science*, *34*, 1065-1078.

Lourenço, A., Álvarez, D., Wang, I. J., & Velo-Antón, G. (2017). Trapped within the city : Integrating demography, time since isolation and population-specific traits to assess the genetic effects of urbanization. *Molecular Ecology*, *26*, 1498-1514.

Menozzi, P., Piazza, A., & Cavalli-Sforza, L. (1978). Synthetic maps of human gene frequencies in Europeans : These maps indicate that early farmers of the Near East spread to all of Europe in the Neolithic. *Science*, *201*, 786-792.

Miles, L. S., Rivkin, L. R., Johnson, M. T. J., Munshi-South, J., & Verrelli, B. C. (2019). Gene flow and genetic drift in urban environments. *Molecular Ecology*, *28*, 4138-4151.

Munshi-South, J., Zolnik, C. P., & Harris, S. E. (2016). Population genomics of the Anthropocene : Urbanization is negatively associated with genome-wide variation in white-footed mouse populations. *Evolutionary Applications*, *9*, 546-564.

Nei M (1987). Molecular Evolutionary Genetics (Columbia Univ Press).

Noël, S., Ouellet, M., Galois, P., & Lapointe, F.-J. (2007). Impact of urban fragmentation on the genetic structure of the eastern red-backed salamander. *Conservation Genetics*, *8*, 599-606.

Paradis, E. (2010). pegas : An R package for population genetics with an integrated–modular approach. *Bioinformatics*, *26*, 419-420.

Petkova, D. (2022). *reemsplots2 : Generate plots to inspect and visualize the results of EEMS. R package version 0.1.0*.

Petkova, D., Novembre, J., & Stephens, M. (2016). Visualizing spatial population structure with estimated effective migration surfaces. *Nature Genetics*, *48*, 94-100.

Phillipsen, I. C., Kirk, E. H., Bogan, M. T., Mims, M. C., Olden, J. D., & Lytle, D. A. (2015). Dispersal ability and habitat requirements determine landscape-level genetic patterns in desert aquatic insects. *Molecular Ecology*, *24*, 54-69.

Pironon, S., Papuga, G., Villellas, J., Angert, A. L., García, M. B., & Thompson, J. D. (2017). Geographic variation in genetic and demographic performance: New insights from an old biogeographical paradigm. *Biological Reviews*, *92*, 1877-1909.

Purse, B. V., Hopkins, G. W., Day, K. J., & Thompson, D. J. (2003). Dispersal characteristics and management of a rare amselfly. *Journal of Applied Ecology*, *40*, 716-728.

Purse, B. V., & Thompson, D. J. (2009). Oviposition site selection by *Coenagrion mercuriale* (Odonata : Coenagrionidae). *International Journal of Odonatology*, *12*, 257-273.

R Core Team. (2022). *R: A language and environment for statistical computing. R Foundation for Statistical Computing*.

Richardson, J. L., Michaelides, S., Combs, M., Djan, M., Bisch, L., Barrett, K., Silveira, G., Butler, J., Aye, T. T., Munshi-South, J., DiMatteo, M., Brown, C., & McGreevy, T. J. (2021). Dispersal ability predicts spatial genetic structure in native mammals persisting across an urbanization gradient. *Evolutionary Applications*, *14*, 163-177.

Rivkin, L. R., Santangelo, J. S., Alberti, M., Aronson, M. F. J., De Keyzer, C. W., Diamond, S. E., Fortin, M., Frazee, L. J., Gorton, A. J., Hendry, A. P., Liu, Y., Losos, J. B., MacIvor, J. S., Martin, R. A., McDonnell, M. J., Miles, L. S., Munshi-South, J., Ness, R. W., Newman, A. E. M., … Johnson, M. T. J. (2019). A roadmap for urban evolutionary ecology. *Evolutionary Applications*, *12*, 384-398.

Rouquette, J. R., & Thompson, D. J. (2005). Habitat associations of the endangered damselfly, *Coenagrion mercuriale*, in a water meadow ditch system in southern England. *Biological Conservation*, *123*, 225-235.

Rouquette, J. R., & Thompson, D. J. (2007). Patterns of movement and dispersal in an endangered damselfly and the consequences for its management: Movement and dispersal in a damselfly. *Journal of Applied Ecology*, *44*, 692-701.

Santangelo, J. S., Rivkin, L. R., & Johnson, M. T. J. (2018). The evolution of city life. *Proceedings of the Royal Society*, *285*, 20181529.

Smouse, P., Long, J., & Sokal, R. (1986). Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Systematic Zoology*, *35*, 627-632.

Storfer, A., Murphy, M. A., Evans, J. S., Goldberg, C. S., Robinson, S., Spear, S. F., Dezzani, R., Delmelle, E., Vierling, L., & Waits, L. P. (2007). Putting the 'landscape' in landscape genetics. *Heredity*, *98*, 128-142.

Thelen, C. (1992). Untersuchung zum Larvenhabitat und zum Entwicklungszyklus der Helmazurjungfer (*Coenagrion mercuriale*, Zygoptera : Odonata) an zwei verchiedenen Gewassern der Freiburger Bucht. *Limnologisches Institut Konstanz: Universitat Freiburg*.

Trumbo, D. R., Salerno, P. E., Logan, K. A., Alldredge, M. W., Gagne, R. B., Kozakiewicz, C. P., Kraberger, S., Fountain-Jones, N. M., Craft, M. E., Carver, S., Ernest, H. B., Crooks, K. R., VandeWoude, S., & Funk, W. C. (2019). Urbanization impacts apex predator gene flow but not genetic diversity across an urban-rural divide. *Molecular Ecology*, *28*, 4926-4940.

Tüzün, N., Op De Beeck, L., Brans, K. I., Janssens, L., & Stoks, R. (2017). Microgeographic differentiation in thermal performance curves between rural and urban populations of an aquatic insect. *Evolutionary Applications*, *10*, 1067-1075.

Tyers, M. (2022). *Riverdist : River networkd distance computation and applications. R package version 0.15.5*.

Vandergast, A. G., Bohonak, A. J., Weissman, D. B., & Fisher, R. N. (2006). Understanding the genetic effects of recent habitat fragmentation in the context of evolutionary history : Phylogeography and landscape genetics of a southern California endemic Jerusalem cricket (Orthoptera: Stenopelmatidae: Stenopelmatus). *Molecular Ecology*, *16*, 977-992.

Vekemans, X., & Hardy, O. J. (2004). New insights from fine-scale spatial genetic structure analyses in plant populations. *Molecular Ecology*, *13*, 921-935.

Vucetich, J. A., & Waite, T. A. (2003). Spatial patterns of demography and genetic processes across the species' range : Null hypotheses for landscape conservation genetics. *Conservation Genetics*, *4*, 639-645.

Watts, P. C., Kemp, S. J., Saccheri, I. J., & Thompson, D. J. (2005). Conservation implications of genetic variation between spatially and temporally distinct colonies of the endangered damselfly *Coenagrion mercuriale*. *Ecological Entomology*, *30*, 541-547.

Watts, P. C., Rouquette, J. R., Saccheri, I. J., Kemp, S. J., & Thompson, D. J. (2004c). Molecular and ecological evidence for small-scale isolation by distance in an endangered damselfly, *Coenagrion mercuriale*. *Molecular Ecology*, *13*, 2931-2945.

Watts, P. C., Saccheri, I. J., Kemp, S. J., & Thompson, D. J. (2006). Population structure and the impact of regional and local habitat isolation upon levels of genetic diversity of the endangered damselfly *Coenagrion mercuriale* (Odonata : Zygoptera). *Freshwater Biology*, *51*, 193-205.

Watts, P. C., & Thompson, D. J. (2012). Developmental plasticity as a cohesive evolutionary process between sympatric alternate-year insect cohorts. *Heredity*, *108*, 236-241.

Watts, P. C., Thompson, D. J., & Kemp, S. J. (2004b). Cross-species amplification of microsatellite loci in some European zygopteran species (Odonata : Coenagrionidae). *International Journal of Odonatology*, *7*, 87-96.

Watts, P. C., Wu, J. H., Westgarth, C., Thompson, D. J., & Kemp, S. J. (2004a). A panel of microsatellite loci for the Southern Damselfly, *Coenagrion mercuriale* (Odonata : Coenagrionidae). *Conservation Genetics*, *5*, 117-119.

Weir, B. S., & Cockerham, C. C. (1984). Estimating *F*-Statistics for the analysis of population structure. *Evolution*, *38*, 1358-1370.

Whitlock, M. C., & McCauley, D. E. (1990). Some Population Genetic Consequences of Colony Formation and Extinction : Genetic Correlations Within Founding Groups. *Evolution*, *44*, 1717-1724.

Willi, Y., Kristensen, T. N., Sgrò, C. M., Weeks, A. R., Ørsted, M., & Hoffmann, A. A. (2022). Conservation genetics as a management tool : The five best-supported paradigms to assist the management of threatened species. *Proceedings of the National Academy of Sciences*, *119*.

Wilson, M. C., Chen, X.-Y., Corlett, R. T., Didham, R. K., Ding, P., Holt, R. D., Holyoak, M., Hu, G., Hughes, A. C., Jiang, L., Laurance, W. F., Liu, J., Pimm, S. L., Robinson, S. K., Russo, S. E., Si, X., Wilcove, D. S., Wu, J., & Yu, M. (2016). Habitat fragmentation and biodiversity conservation : Key findings and future challenges. *Landscape Ecology*, *31*, 219-227.

Supporting information

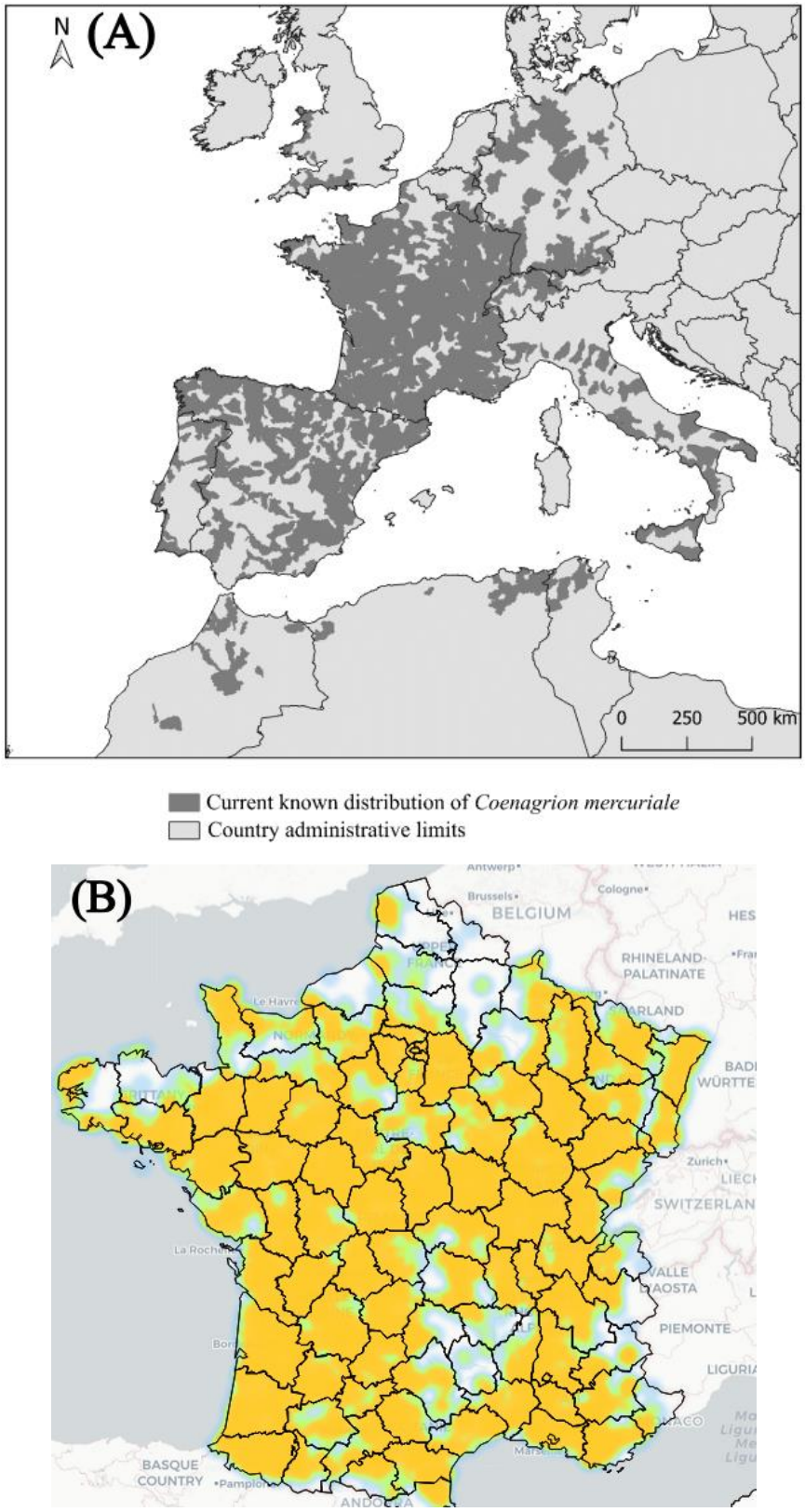


Figure S1: (A) Map showing the current known distribution (dark grey zones) of the southern damselfly (*Coenagrion mercuriale*). Data obtained from IUCN SSC Odonata Specialist Group 2019. The IUCN Red List of Threatened Species. Version 2022-2. https://www.iucnredlist.org/ Downloaded on 28 July 2023. Black lines indicate the administrative limits of countries. **(B)** Interpolation of the density of

observations of southern damselflies in France. Source: https://atlas-odonates.insectes.org/odonates-de-france/coenagrion-mercuriale.

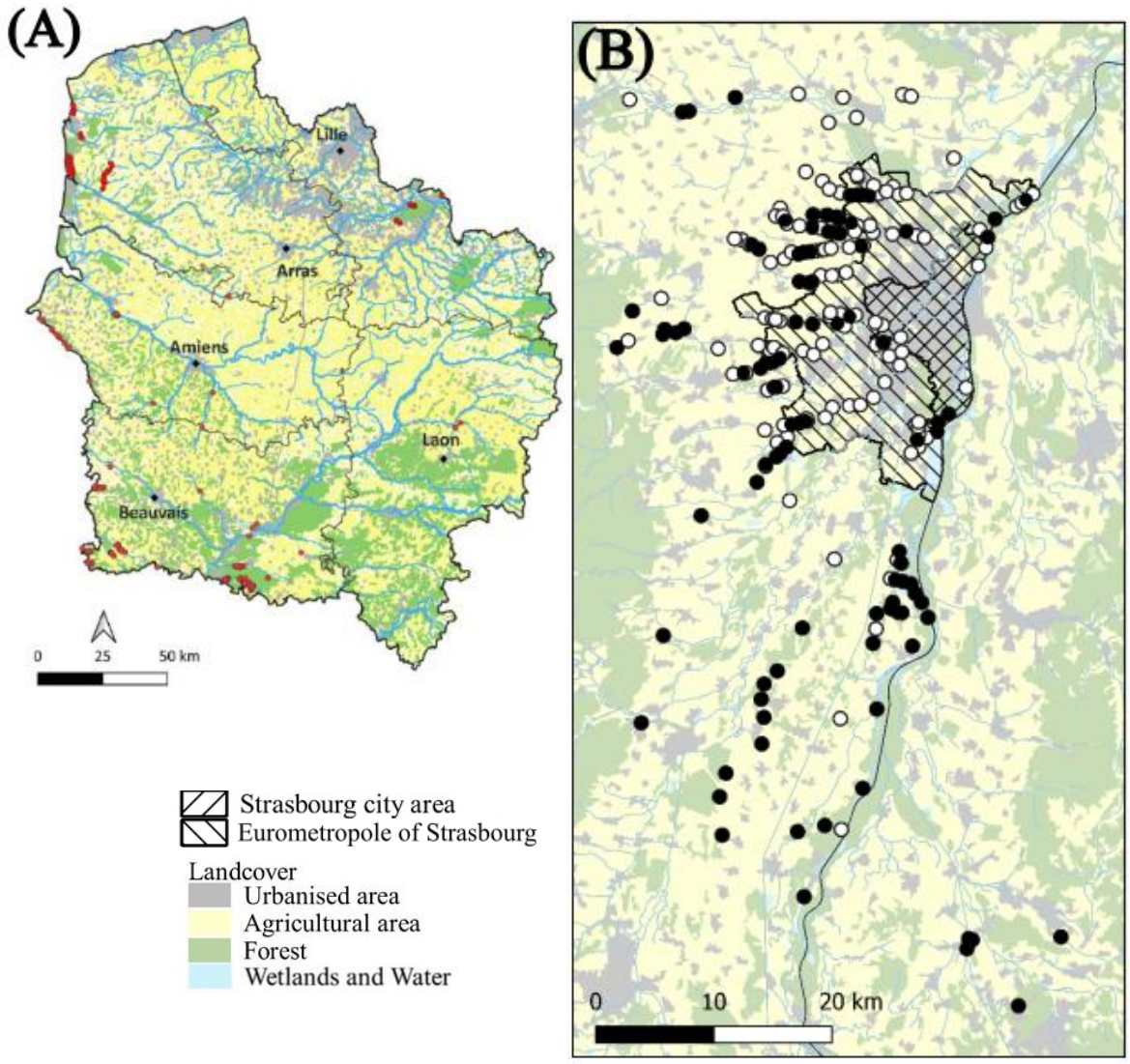


Figure S2: (A) Map showing the geographic distribution of southern damselfly in the region Hauts de France. The sites where the species is present appear in red. Blue lines; rivers; light blue regions; wetlands; yellow; agricultural zones; green: forest; grey: urban areas (https://irpn.drealnpdc.fr/wp-content/uploads/2023/04/PRA_20230420_DOC_PRA_Libellules_HdF.pdf.). **(B)** Geographical location of southern damselfly sampling sites around Strasbourg City from 2021 to 2022. Black dots represent sites successfully sampled, and white dots represent sites surveyed but where no southern damselflies were found. Main watercourses were taken from COPERNICUS Land Monitoring Service (2019: EU-Hydro). Strasbourg city and the eurometropole of Strasbourg are represented by the striped areas. Land cover was simplified from Corine Land Cover Edition 2018. Note the difference in spatial scales among the regions.

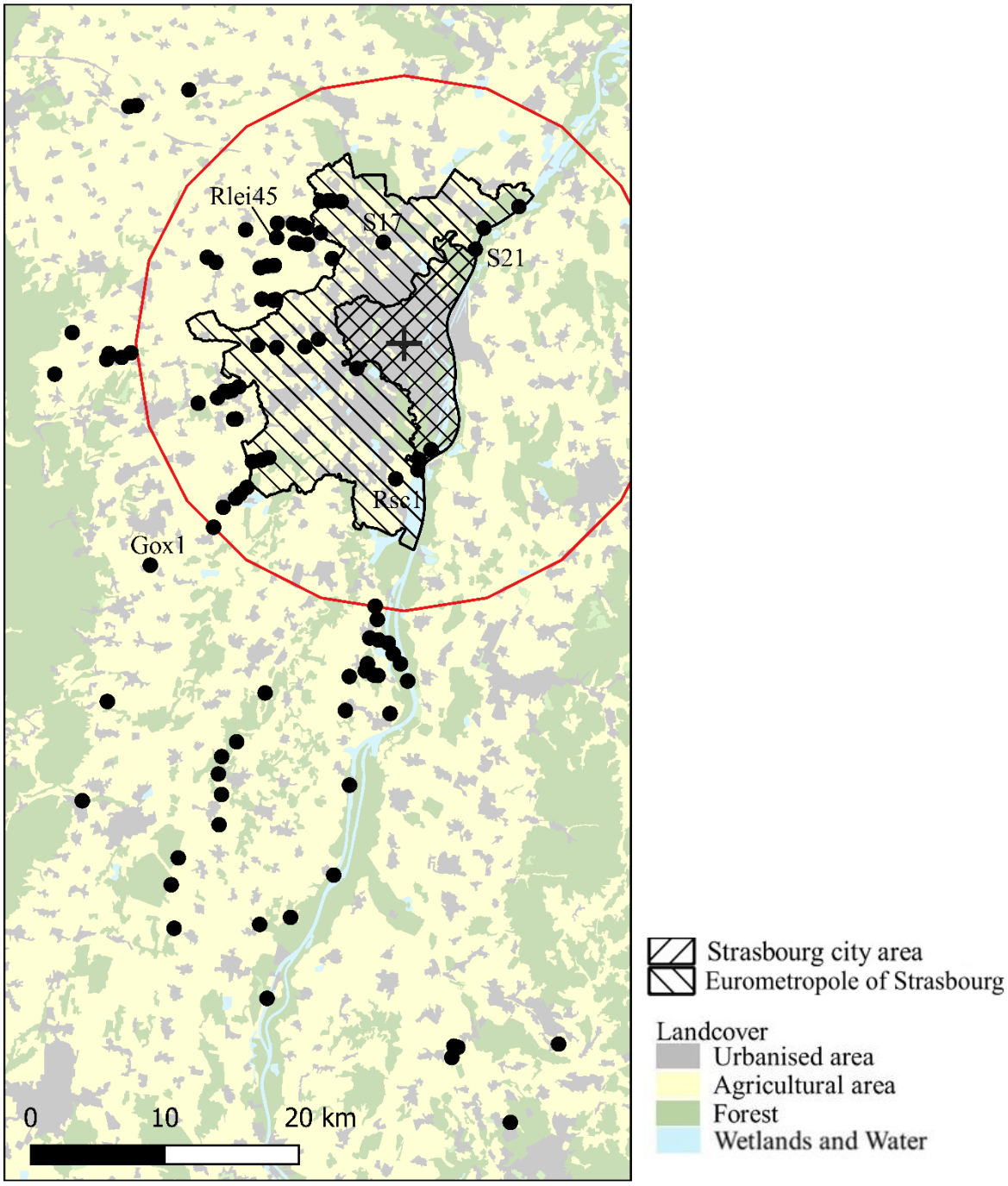


Figure S3: Map showing the 20 km radius (red line) around Strasbourg centre (black cross) within which the potential effect of the city of Strasbourg on southern damselfly populations was measured in terms of levels of genetic diversity (A_r, H_e), the mean levels of individual kinship coefficient (F_{ij}) and the mean pairwise levels of F_{ST} calculated according to Weir and Cockerham (1984). Black dots represent sites sampled in Alsace with at least eight individuals. Land cover was simplified from Corine Land Cover Edition 2018 France Métropolitaine. Locations and names of specific populations cited in the text (Gox1, S17, S21, Rlei45, and Rsc1) were added to the map. Strasbourg city and the eurometropole of Strasbourg are represented by the striped areas.

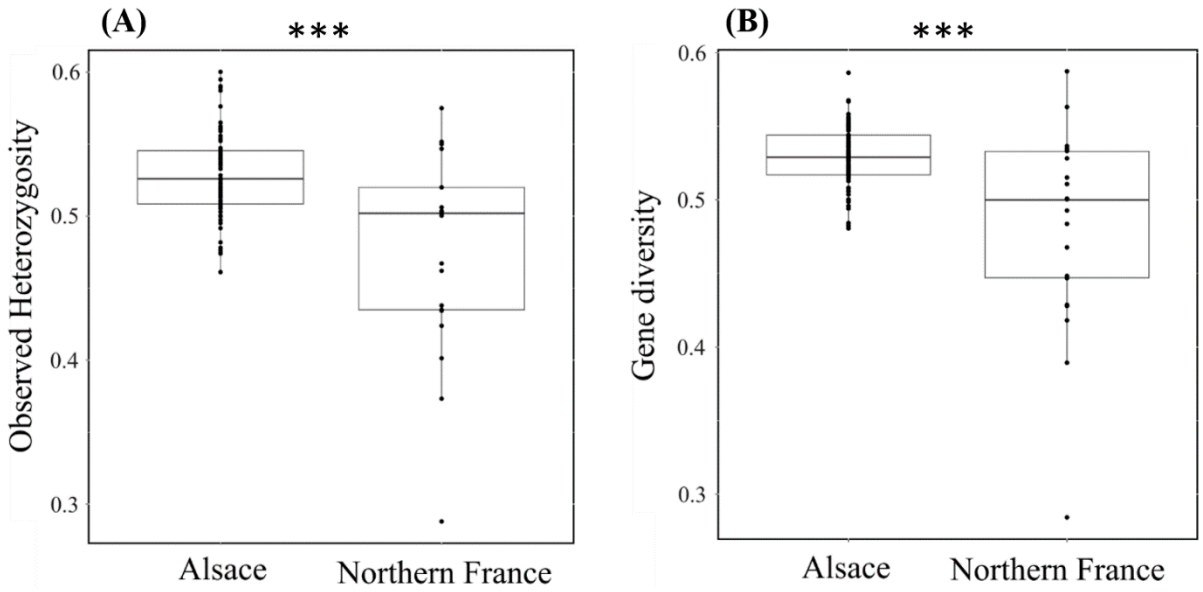


Figure S4: Distribution of observed heterozygosity H_o **(A)** and gene diversity H_e **(B)** within Alsace and northern France areas. *P*-value: *** $P < 0.001$.

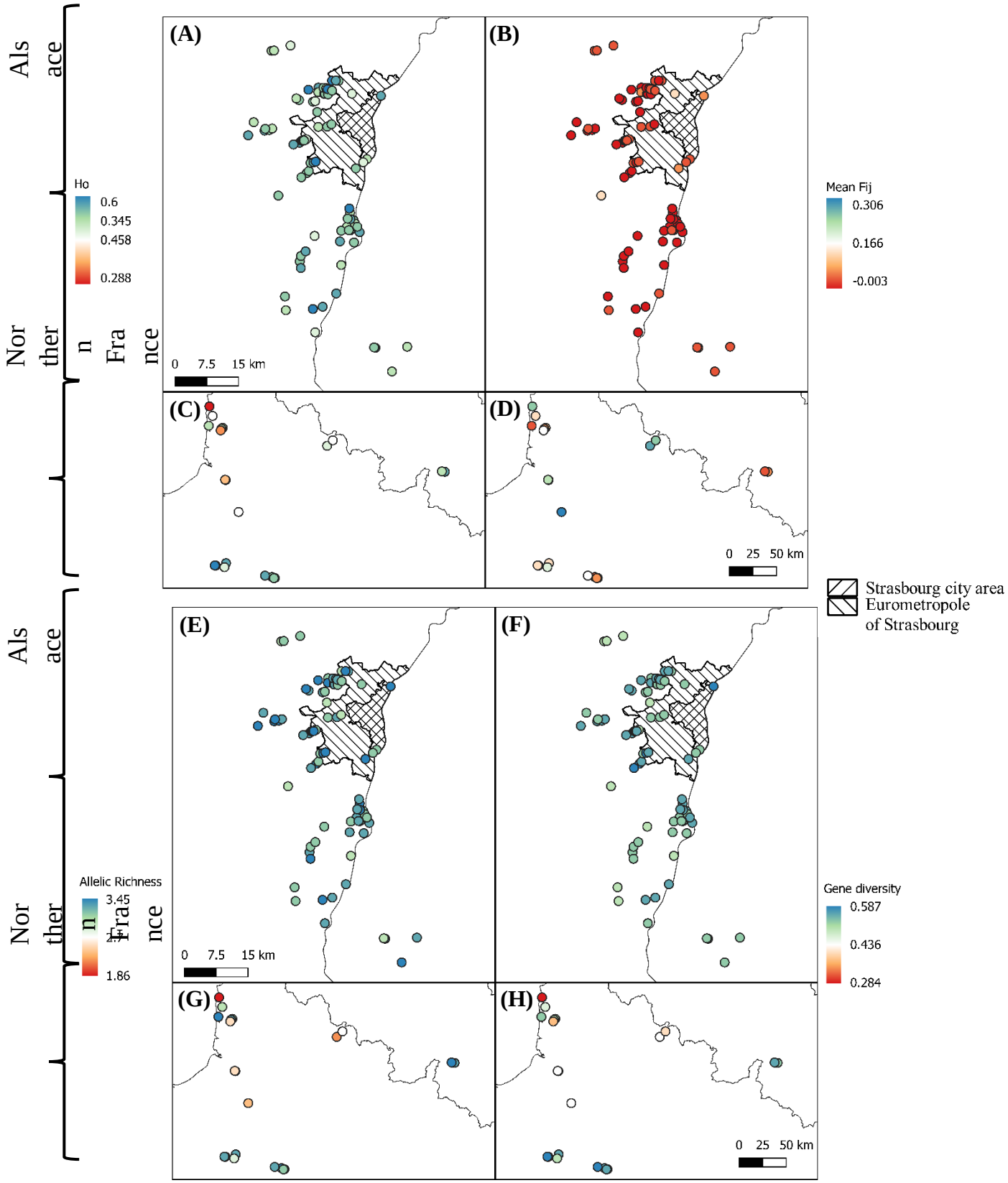


Figure S5: Geographical distribution of levels of observed heterozygosity (H_o) **(A, C)**, levels of intra-population kinship coefficient F_{ij} **(B, D)**, levels of allelic richness (A_r) **(E, G)**, and levels of expected heterozygosity (H_e) **(F, H)** in southern damselfly populations genotyped for at least eight individuals in Alsace **(A, B, E, F)** and Northern France **(C, D, G, H)**. Strasbourg city and the eurometropole of Strasbourg are represented by the striped areas.

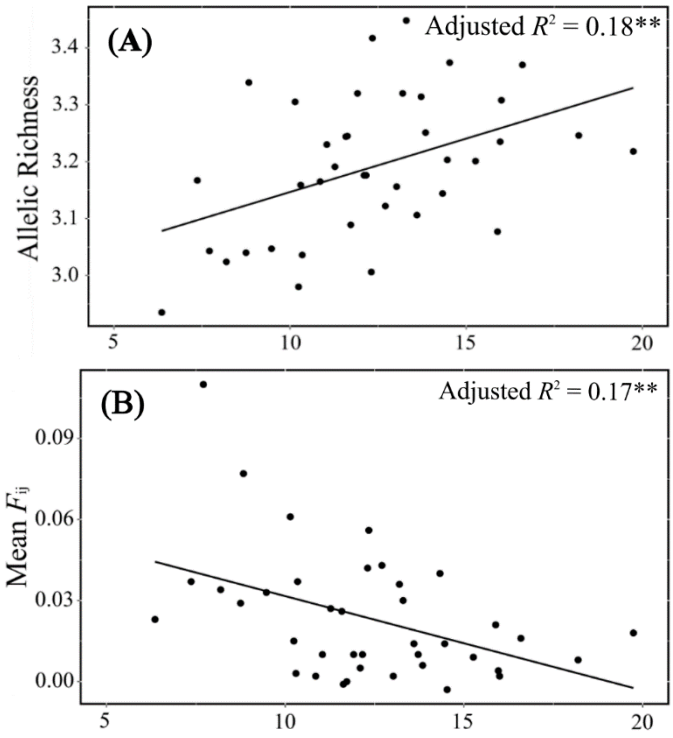


Figure S6: Relationship between allelic richness (A_r) **(A)** and mean intra-population kinship coefficient (F_{ij}) **(B)**, and the distance of the population to the centre of Strasbourg within a 20 km buffer (see Figure S3). Lines show the linear model. ***: P < 0.01*

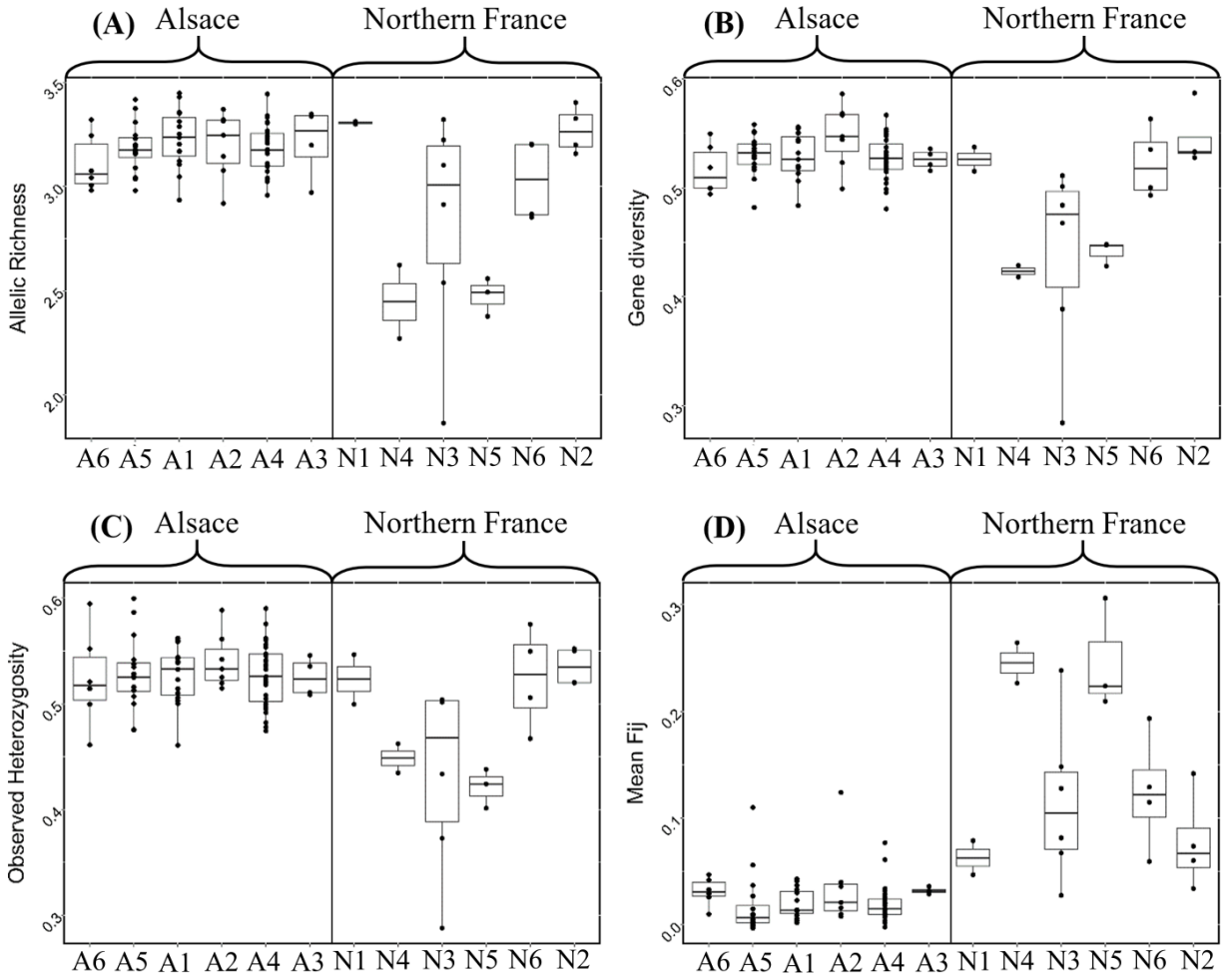


Figure S7: Distribution of allelic richness A_r **(A)**, expected heterozygosity H_e **(B)**, observed heterozygosity H_o **(C)**, and intra-population kinship coefficient F_{ij} **(D)** within each hydrographic unit or spatial regions, for populations with at least eight sampled individuals.

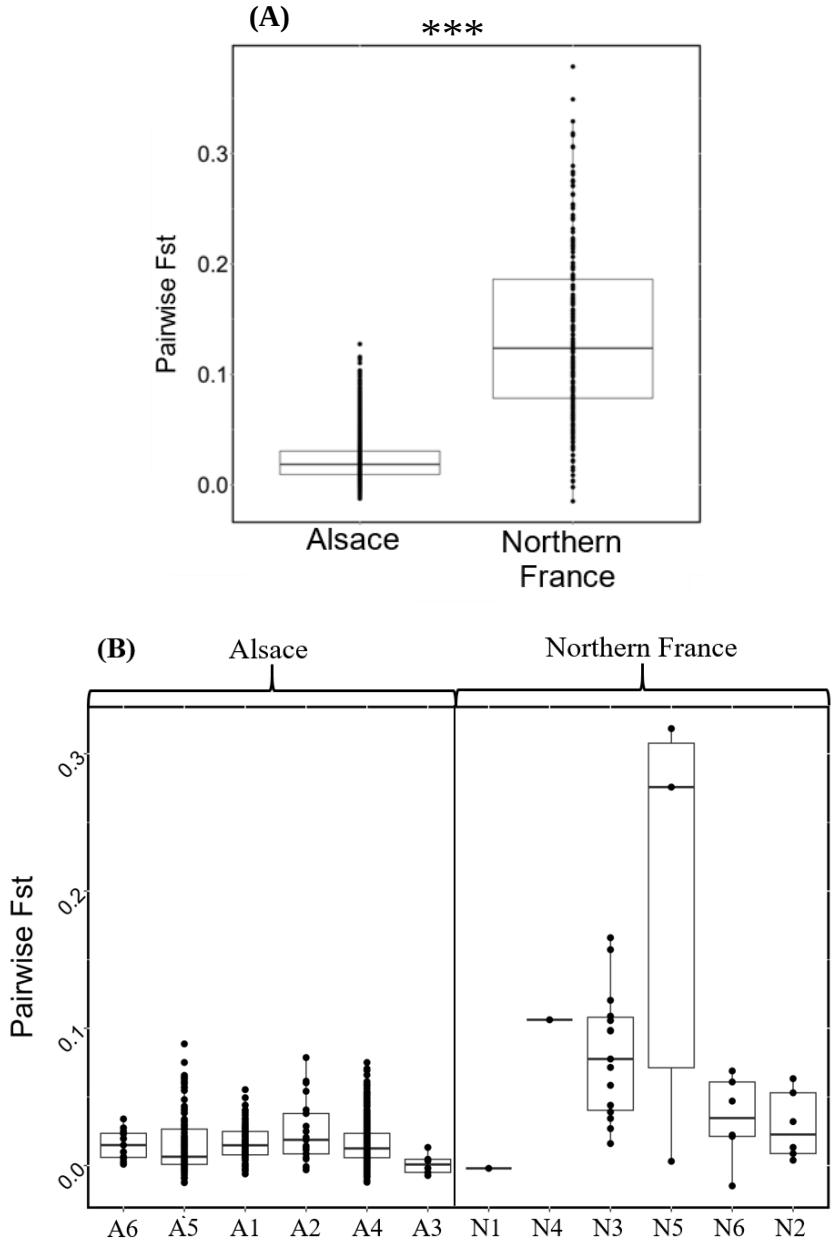


Figure S8: Distribution of pairwise F_{ST} estimates within the two main surveyed areas, Alsace and Northern France **(A)** and within hydrographic units or spatial regions within both areas **(B)**, for populations with eight or more sampled individuals. *P*-value: *** $P < 0.001$.

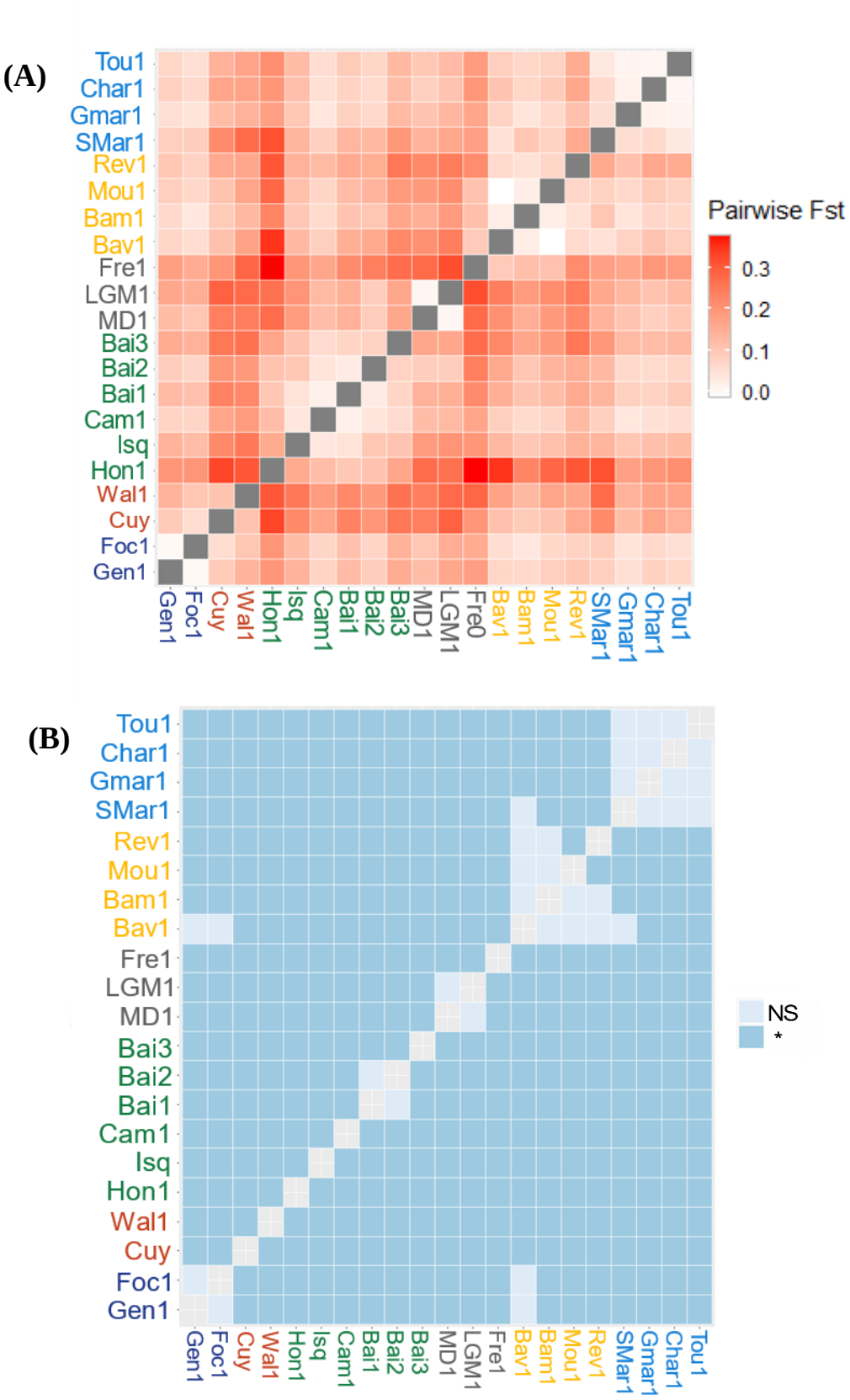


Figure S9: Matrices of pairwise F_{ST} estimates in Northern France **(A)** and their associated statistical significance assessed using 10000 permutations of multilocus genotypes among populations after Bonferroni correction **(B)**. Populations are listed and coloured by geographical regions as in Figure 1.

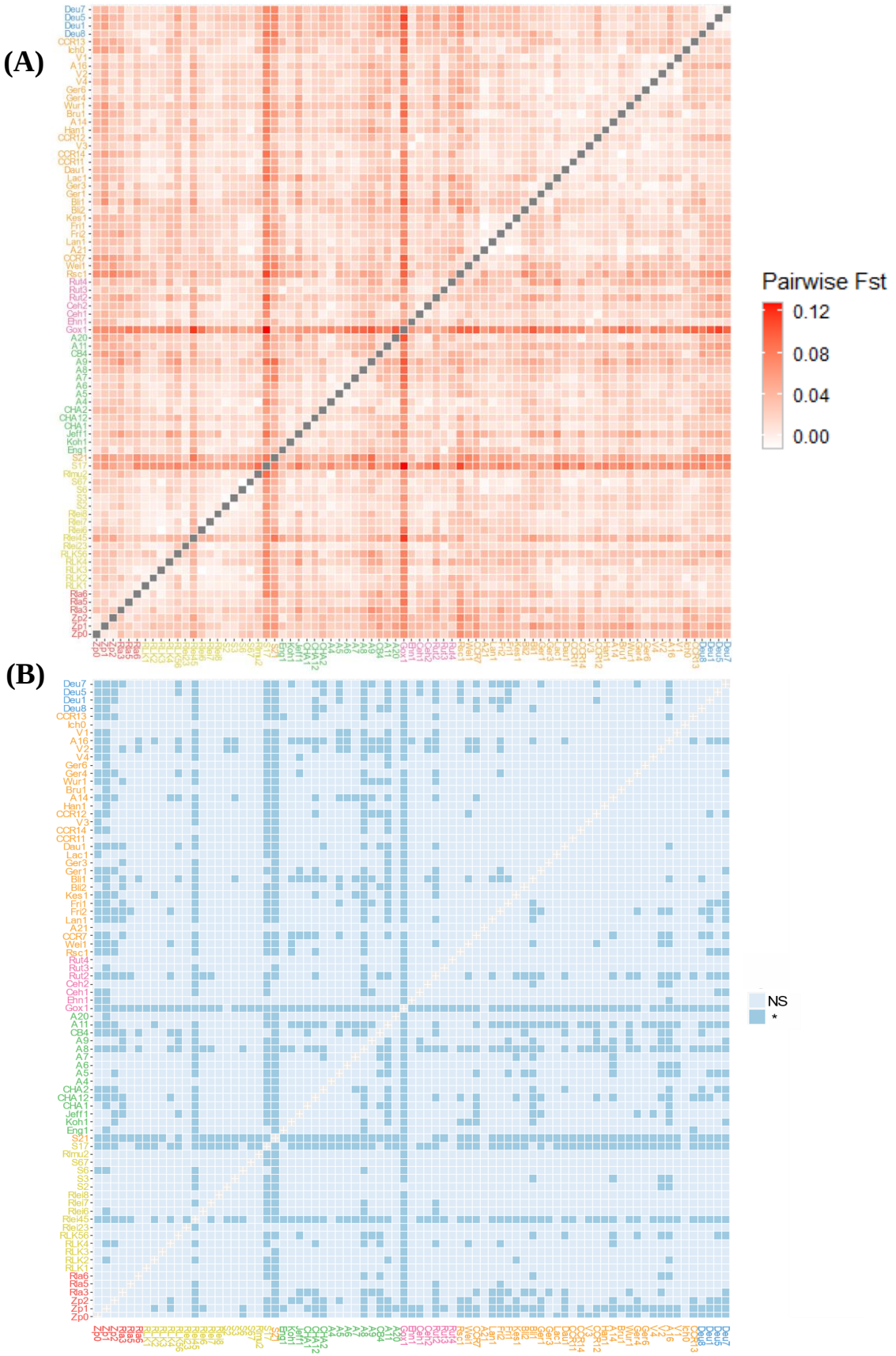


Figure S10: Matrices of pairwise F_{ST} estimates in Alsace **(A)** and their associated statistical significance assessed using 10000 permutations of multilocus genotypes among populations after Bonferroni correction (**B**). Populations are listed and coloured by subwatershed as in Figure 1.

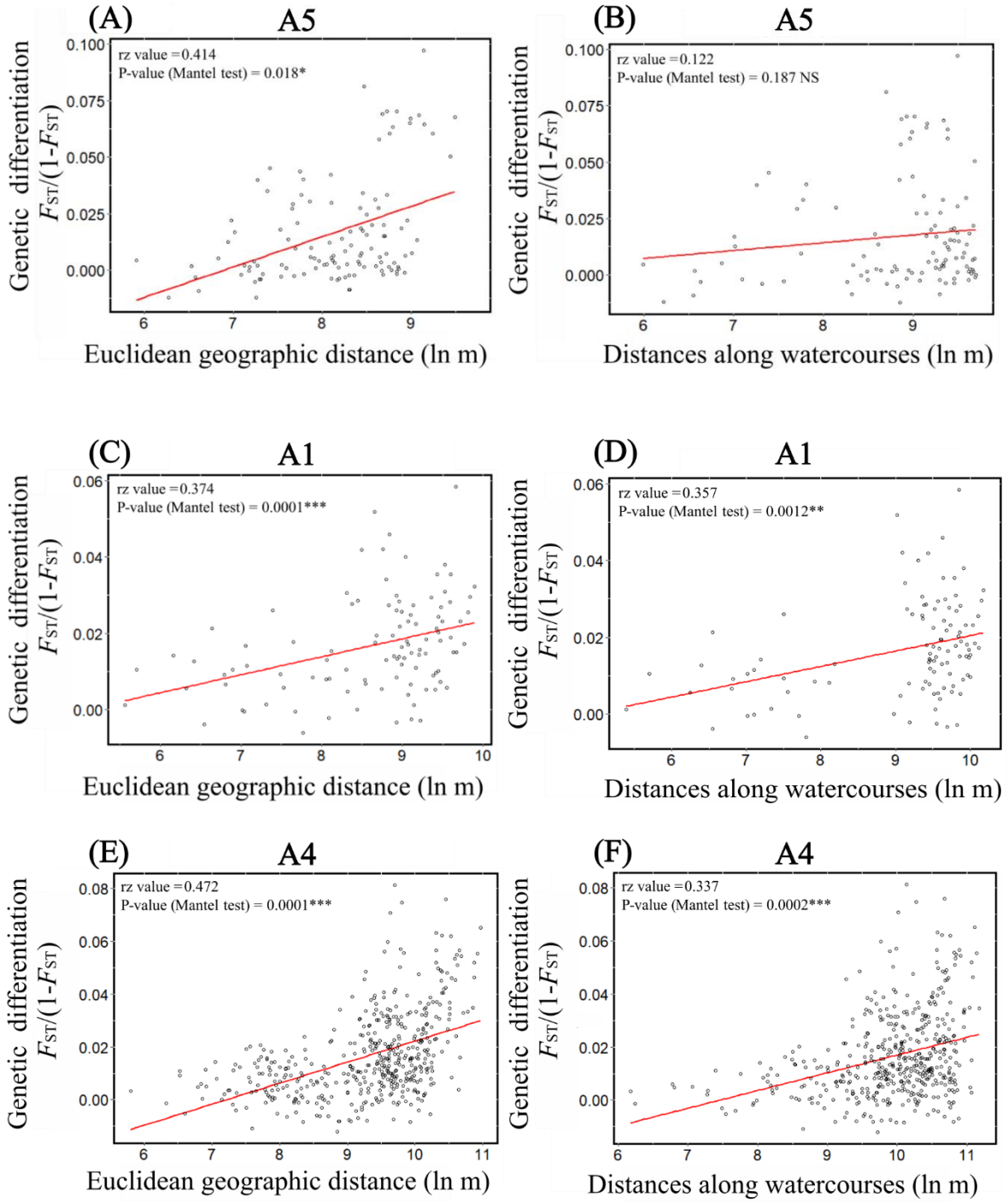


Figure S11: Relationship between levels of genetic differentiation ($F_{ST}/(1-F_{ST})$) and log-transformed Euclidean distance (d_{Eucli}) **(A, C, E)** or watercourse geographical distance (d_{Stream}) **(B, D, F)** among populations of *Coenagrion mercuriale* within three main subwatersheds: A5 **(A, B)**, A1 **(C, D)**, A4 **(E, F)**.

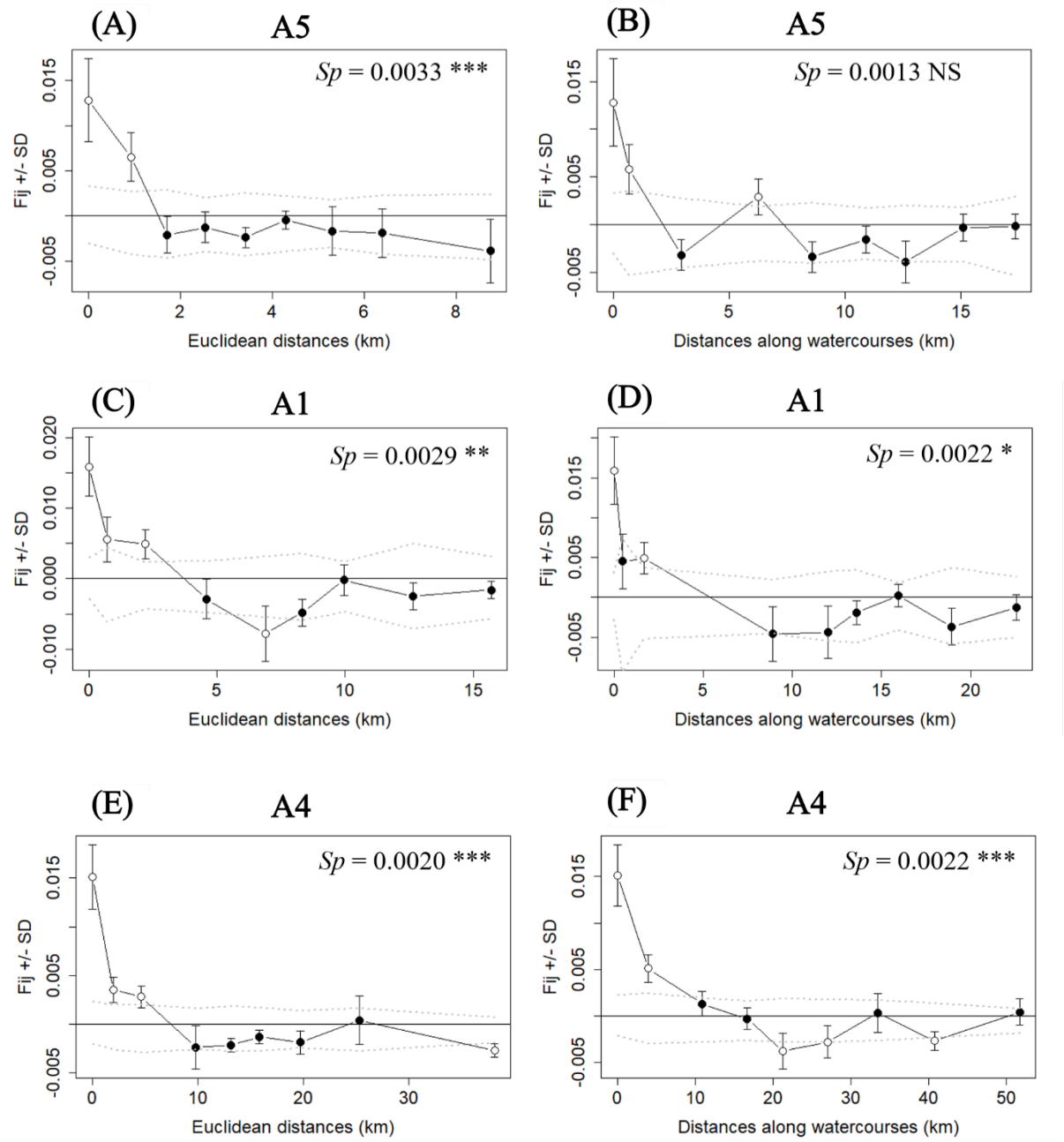


Figure S12: Broad spatial genetic structure in southern damselfly populations in three different main subwatersheds: A5 **(A, B)**, A1 **(C, D)**, A4 **(E, F)**. Average pairwise kinship coefficients (F_{ij}, Loiselle *et al.*, 1995) between individuals are plotted against increasing classes of geographical Euclidean distances **(A, C, E)** or of waterways distance **(B, D, F)**. Standard errors for F_{ij} for each distance class were obtained by jackknifing over loci. Dashed lines indicate the upper and lower 95% confidence intervals of non-significant spatial genetic structure, with white dots indicating significant estimates outside this confidence interval. *Sp* statistics based on the regression with ln(distance) and their statistical significance are also indicated, allowing for comparisons of the strength of spatial genetic structure.

Table S1: Sampling locations and basic genetic diversity estimates for each studied population. ID: population name, Sampling Year: the year of sampling, Area: sampling region and hydrographic units they belong to, geographic coordinates of the sampling site (WGS84, EPSG4326), N: sample size. Summary of population structure for nuclear polymorphism at 10 microsatellite loci for each population can also be found: the total number of alleles (A_T), the allelic richness (A_r) rarefied on eight individuals, the observed heterozygosity (H_o), the expected heterozygosity (H_e), the intrapopulation fixation index F_{IS}, and the mean intrapopulation kinship coefficient F_{ij}. The significance of F_{IS} was tested with 10000 permutations of alleles among individuals within populations (*: $P < 0.05$, **: $P < 0.01$, ***: $P < 0.001$), note that none was significant. -: no estimates because of insufficient population size.

ID	Sampling Year	Area	Hydrographic units/ Spatial Region	Longitude	Latitude	N	A_T	A_r	*Ho*	*He*	F_{IS}	F_{ij}
A11	2022	Alsace	A1	7.655745	48.574643	39	41	3.167	0.533	0.526	-0.013	0.037
A14	2021	Alsace	A4	7.697191	48.354712	27	39	3.073	0.526	0.528	0.004	0.008
A16	2022	Alsace	A4	7.653206	48.218936	34	40	3.178	0.547	0.552	0.009	0.026
A20	2022	Alsace	A1	7.669929	48.579161	13	32	2.935	0.523	0.506	-0.034	0.023
A21	2022	Alsace	A4	7.595143	48.343877	9	31	3.029	0.478	0.481	0.006	0.011
A3	2021	Alsace	A1	7.532710	48.540547	1	17	-	-	-	-	-
A4	2022	Alsace	A1	7.544702	48.540918	18	38	3.235	0.544	0.543	-0.002	0.004
A5	2021	Alsace	A1	7.564511	48.543905	26	41	3.203	0.562	0.551	-0.020	0.014
A6	2021	Alsace	A1	7.571759	48.547882	30	43	3.251	0.503	0.520	0.033	0.006
A7	2021	Alsace	A1	7.575255	48.547574	33	44	3.106	0.506	0.518	0.023	0.014
A8	2022	Alsace	A1	7.579292	48.548017	36	46	3.448	0.539	0.550	0.020	0.030
A9	2021	Alsace	A1	7.586779	48.550395	31	44	3.122	0.461	0.484	0.048	0.043
BA3	2022	Alsace	A1	7.579429	48.528832	4	26	-	-	-	-	-
BA4	2022	Alsace	A1	7.581946	48.528855	3	25	-	-	-	-	-
Bli1	2022	Alsace	A4	7.490202	48.189452	30	41	3.035	0.497	0.499	0.004	0.032
Bli2	2022	Alsace	A4	7.489811	48.218759	36	44	3.102	0.519	0.496	-0.046	0.020
Bli3	2022	Alsace	A4	7.498421	48.236530	7	29	-	-	-	-	-
Bru1	2022	Alsace	A4	7.709551	48.351107	19	37	3.110	0.495	0.517	0.043	0.011
CB2	2021	Alsace	A1	7.591170	48.554269	2	19	-	-	-	-	-
CB4	2022	Alsace	A1	7.627223	48.575197	21	36	3.047	0.514	0.506	-0.016	0.033
CCR11	2022	Alsace	A4	7.721753	48.372592	27	40	3.104	0.533	0.552	0.034	-0.002
CCR12	2022	Alsace	A4	7.739052	48.346225	32	40	3.207	0.547	0.549	0.004	0.015
CCR13	2022	Alsace	A4	7.579254	48.138813	33	45	3.267	0.482	0.508	0.051	0.025
CCR14	2022	Alsace	A4	7.726147	48.365158	21	37	3.096	0.505	0.517	0.023	0.013
CCR2	2022	Alsace	A4	7.880296	48.660017	4	28	-	-	-	-	-
CCR4	2022	Alsace	A4	7.843604	48.646993	6	29	-	-	-	-	-
CCR7	2022	Alsace	A4	7.776558	48.500519	24	37	3.024	0.508	0.522	0.027	0.034
CCR8	2022	Alsace	A4	7.762018	48.487490	3	24	-	-	-	-	-
Ceh0	2022	Alsace	A2	7.576856	48.475591	4	28	-	-	-	-	-
Ceh1	2021	Alsace	A2	7.581641	48.478382	31	43	3.370	0.542	0.568	0.046	0.016
Ceh2	2022	Alsace	A2	7.588845	48.482536	27	37	3.077	0.515	0.547	0.059	0.021
CHA1	2022	Alsace	A1	7.454961	48.573777	29	44	3.351	0.559	0.544	-0.028	0.014
CHA12	2022	Alsace	A1	7.470093	48.574551	30	45	3.429	0.540	0.556	0.029	0.002
CHA2	2022	Alsace	A1	7.479667	48.577296	30	41	3.170	0.510	0.532	0.041	0.013
Dau1	2022	Alsace	A4	7.712416	48.374920	31	43	3.305	0.561	0.550	-0.020	0.008
Deu1	2022	Alsace	A3	7.767138	48.098375	28	43	3.337	0.546	0.536	-0.019	0.029

Deu2	2022	Alsace	A3	7.760566	48.091790	2	20	-	-	-	-	-
Deu5	2022	Alsace	A3	7.868078	48.096462	23	40	3.198	0.509	0.521	0.023	0.036
Deu7	2022	Alsace	A3	7.815028	48.045885	38	50	3.347	0.511	0.531	0.038	0.032
Deu8	2022	Alsace	A3	7.763198	48.099368	22	37	2.971	0.536	0.516	-0.039	0.031
Ehn1	2021	Alsace	A2	7.563860	48.470101	20	38	3.246	0.520	0.567	0.083	0.008
Ehn9	2022	Alsace	A2	7.553003	48.457075	3	31	-	-	-	-	-
Eng1	2022	Alsace	A1	7.401690	48.565821	25	44	3.312	0.560	0.555	-0.009	0.014
Far1	2022	Alsace	A4	7.405380	48.278538	2	20	-	-	-	-	-
Fri1	2022	Alsace	A4	7.543433	48.291327	29	42	3.231	0.521	0.523	0.004	0.015
Fri2	2022	Alsace	A4	7.547780	48.302849	26	40	3.158	0.523	0.515	-0.016	0.020
Ger1	2022	Alsace	A4	7.710864	48.397795	33	40	3.218	0.576	0.554	-0.040	0.018
Ger3	2021	Alsace	A4	7.712247	48.388783	21	41	3.332	0.500	0.533	0.062	0.007
Ger4	2022	Alsace	A4	7.680718	48.351555	30	39	3.103	0.540	0.520	-0.038	0.014
Ger6	2021	Alsace	A4	7.674729	48.328929	20	39	3.174	0.560	0.526	-0.065	0.009
Gox1	2022	Alsace	A2	7.487110	48.434042	27	34	2.919	0.533	0.499	-0.068	0.124
Han1	2021	Alsace	A4	7.699740	48.359411	25	40	3.238	0.492	0.524	0.061	0.020
Ich0	2022	Alsace	A4	7.576146	48.188722	10	35	3.309	0.590	0.544	-0.085	0.021
Jeff1	2022	Alsace	A1	7.457562	48.577673	21	42	3.356	0.543	0.513	-0.058	0.041
Kes1	2022	Alsace	A4	7.545415	48.277385	29	45	3.443	0.545	0.530	-0.028	0.018
Koh1	2022	Alsace	A1	7.421644	48.593061	34	46	3.288	0.500	0.520	0.038	0.009
Lac1	2022	Alsace	A4	7.703560	48.376597	19	41	3.174	0.537	0.533	-0.008	0.023
Lan1	2022	Alsace	A4	7.563761	48.312189	34	39	3.034	0.562	0.518	-0.085	0.010
Mul5	2022	Alsace	A1	7.608110	48.577487	6	27	-	-	-	-	-
Ost1	2022	Alsace	A4	7.706748	48.558034	4	23	-	-	-	-	-
Rla23	2021	Alsace	A6	7.668104	48.673981	3	26	-	-	-	-	-
Rla3	2021	Alsace	A6	7.680047	48.671614	20	35	3.006	0.595	0.537	-0.108	0.042
Rla4	2022	Alsace	A6	7.687878	48.671665	3	23	-	-	-	-	-
Rla5	2022	Alsace	A6	7.692171	48.671474	27	45	3.320	0.552	0.550	-0.004	0.010
Rla6	2021	Alsace	A6	7.700949	48.670594	29	47	3.244	0.521	0.519	-0.004	0.026
Rlei23	2021	Alsace	A5	7.602713	48.655405	24	47	3.374	0.600	0.558	-0.075	-0.003
Rlei45	2022	Alsace	A5	7.633774	48.649009	30	45	3.417	0.513	0.551	0.069	0.056
Rlei6	2022	Alsace	A5	7.652179	48.644755	29	42	3.230	0.507	0.542	0.065	0.010
Rlei7	2022	Alsace	A5	7.654957	48.644172	34	44	3.165	0.535	0.529	-0.011	0.002
Rlei8	2022	Alsace	A5	7.664481	48.643140	26	45	3.159	0.538	0.527	-0.021	0.003
RLK1	2021	Alsace	A5	7.635142	48.658617	23	41	3.156	0.565	0.522	-0.082	0.002
RLK2	2022	Alsace	A5	7.651268	48.657640	33	43	3.176	0.542	0.540	-0.004	0.010
RLK3	2022	Alsace	A5	7.660315	48.656235	15	38	3.245	0.513	0.552	0.071	-0.001
RLK4	2022	Alsace	A5	7.664291	48.654240	23	40	3.191	0.587	0.536	-0.095	0.027
RLK56	2022	Alsace	A5	7.678233	48.650551	18	36	3.036	0.528	0.530	0.004	0.037
Rlmu1	2022	Alsace	A5	7.614947	48.608272	3	28	-	-	-	-	-
Rlmu2	2022	Alsace	A5	7.626599	48.607025	20	37	2.980	0.525	0.482	-0.089	0.015
Rlmu3	2022	Alsace	A5	7.629157	48.607368	4	26	-	-	-	-	-
Rmh1	2021	Alsace	A6	7.688742	48.685199	7	32	-	-	-	-	-
Rsc1	2022	Alsace	A4	7.739234	48.482301	20	40	3.305	0.520	0.532	0.023	0.061
Rsc2	2021	Alsace	A4	7.741570	48.490704	1	19	-	-	-	-	-
Rsc3	2021	Alsace	A4	7.737795	48.493810	3	28	-	-	-	-	-

Rut2	2022	Alsace	A2	7.595870	48.499645	28	38	3.144	0.525	0.523	-0.004	0.040
Rut3	2022	Alsace	A2	7.605300	48.500283	23	39	3.314	0.561	0.586	0.043	0.010
Rut4	2022	Alsace	A2	7.612676	48.501593	9	34	3.320	0.589	0.544	-0.083	0.036
S12	2022	Alsace	A5	7.688485	48.632729	5	26	-	-	-	-	-
S17	2021	Alsace	A5	7.741153	48.641593	21	38	3.043	0.476	0.508	0.063	0.110
S21	2021	Alsace	A4	7.833577	48.633176	25	39	3.339	0.556	0.567	0.019	0.077
S2	2022	Alsace	A5	7.562404	48.638451	26	43	3.308	0.500	0.539	0.072	0.002
S3	2021	Alsace	A5	7.570723	48.634705	34	43	3.201	0.526	0.534	0.015	0.009
S6	2021	Alsace	A5	7.615283	48.629329	31	42	3.176	0.516	0.519	0.006	0.005
S67	2022	Alsace	A5	7.622262	48.630331	20	38	3.089	0.475	0.517	0.081	0.000
S677	2022	Alsace	A5	7.629693	48.630449	4	27	-	-	-	-	-
Sci1	2022	Alsace	A4	7.541203	48.257234	2	21	-	-	-	-	-
Ser1	2022	Alsace	A2	7.436113	48.344174	4	25	-	-	-	-	-
V1	2021	Alsace	A4	7.607497	48.192280	35	45	3.263	0.554	0.537	-0.032	0.004
V2	2022	Alsace	A4	7.674399	48.278712	30	36	2.958	0.500	0.504	0.008	0.012
V3	2022	Alsace	A4	7.732873	48.358055	33	40	3.158	0.536	0.548	0.022	0.002
V4	2022	Alsace	A4	7.719338	48.325092	34	45	3.161	0.518	0.527	0.017	0.010
Wei1	2022	Alsace	A4	7.763958	48.494377	27	37	3.040	0.474	0.508	0.067	0.029
Wur1	2022	Alsace	A4	7.705175	48.351641	32	44	3.245	0.541	0.530	-0.021	0.029
Zp0	2022	Alsace	A6	7.491934	48.742628	35	40	2.982	0.514	0.494	-0.040	0.033
Zp1	2022	Alsace	A6	7.499699	48.742984	30	40	3.075	0.500	0.500	0.000	0.047
Zp2	2022	Alsace	A6	7.553629	48.751473	33	41	3.042	0.461	0.500	0.078	0.029
Bai1	2017	Nord	N3	1.814731	50.556417	34	38	3.102	0.502	0.484	-0.037	0.082
Bai2	2017	Nord	N3	1.795901	50.547208	28	40	3.223	0.504	0.501	-0.006	0.067
Bai3	2017	Nord	N3	1.784821	50.531367	32	32	2.538	0.373	0.389	0.041	0.148
Bai4	2017	Nord	N3	1.786619	50.521859	4	23	-	-	-	-	-
Bam1	2017	Nord	N6	1.743292	49.241668	31	34	2.868	0.506	0.493	-0.026	0.059
Bav1	2017	Nord	N6	1.729459	49.242210	8	32	3.200	0.575	0.563	-0.021	0.129
Cam1	2017	Nord	N3	1.606164	50.573788	36	41	3.321	0.503	0.511	0.016	0.028
Char1	2017	Nord	N2	2.605009	49.128569	33	43	3.403	0.552	0.528	-0.045	0.060
Cuy	2022	Nord	N4	3.458695	50.442091	20	31	2.623	0.435	0.418	-0.041	0.226
Foc1	2018	Nord	N1	5.072654	50.129394	19	39	3.311	0.500	0.537	0.069	0.047
Fre1	2019	Nord	N5	2.067305	49.755982	26	26	2.377	0.438	0.428	-0.023	0.306
Gam1	2017	Nord	N5	1.561604	49.981019	6	28	-	-	-	-	-
Gen1	2017	Nord	N1	5.115321	50.124713	17	39	3.299	0.547	0.515	-0.062	0.079
Gmar1	2017	Nord	N2	2.562805	49.137762	30	40	3.326	0.520	0.533	0.024	0.034
Hon1	2017	Nord	N3	1.611199	50.758635	33	21	1.865	0.288	0.284	-0.014	0.238
Isq	2017	Nord	N3	1.660956	50.668245	30	35	2.913	0.434	0.468	0.073	0.128
LGM1	2017	Nord	N5	1.870529	50.059129	29	26	2.493	0.424	0.447	0.051	0.224
MD1	2017	Nord	N5	1.862351	50.061888	15	27	2.558	0.401	0.448	0.105	0.209
Mou1	2017	Nord	N6	1.893308	49.265562	28	40	3.204	0.550	0.535	-0.028	0.115
OM1	2017	Nord	N5	1.464563	50.036032	6	30	-	-	-	-	-
Rev1	2017	Nord	N6	1.871270	49.226970	9	29	2.853	0.467	0.500	0.066	0.193
SMar1	2017	Nord	N2	2.461970	49.150454	8	32	3.200	0.550	0.587	0.063	0.142
Tou1	2017	Nord	N2	2.590224	49.127125	31	40	3.157	0.520	0.533	0.024	0.074
Wal1	2017	Nord	N4	3.373399	50.389700	32	24	2.271	0.462	0.429	-0.077	0.264

Table S2: Estimates of genetic diversity for each microsatellite locus. Multiplex: multiplex in which the microsatellite locus was amplified, Fluo: fluorophore used for genotyping (FAM: blue, NED: yellow, PET: red, VIC: green), A_N: number of alleles, H_o: observed heterozygosity, H_e: expected heterozygosity, F_{IS}: intrapopulation fixation index, F_{ST}: mean genetic differentiation across sampled populations. *: $P < 0.05$, ***: $P < 0.001$.

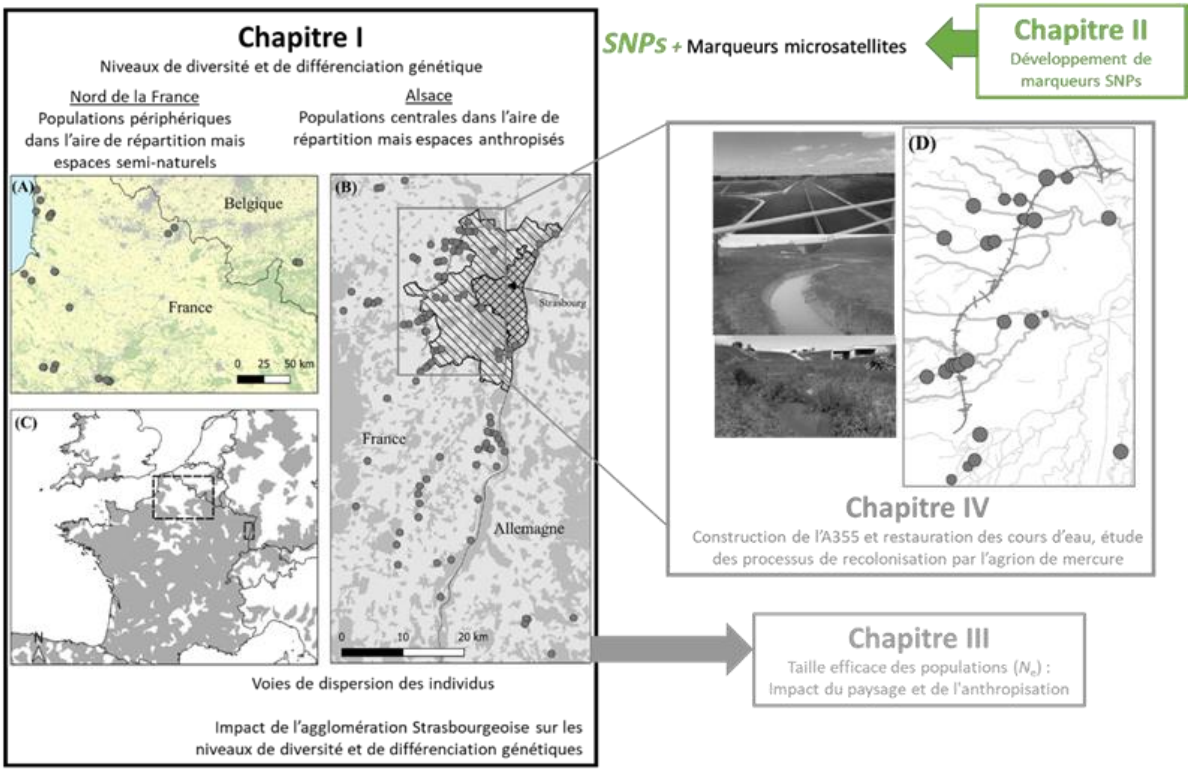
Locus	Multiplex	Fluo	A_N	H_o	H_e	F_{ST}	F_{IS}
LIST024	PLEX1	VIC	5	0.643	0.666	0.035 ***	-0.001
LIST034	PLEX1	PET	2	0.471	0.500	0.073 ***	-0.008
LIST035	PLEX1	VIC	24	0.652	0.706	0.070 ***	0.006
LIST063	PLEX1	FAM	12	0.629	0.673	0.061 ***	-0.001
LIST066	PLEX1	NED	14	0.544	0.583	0.069 ***	-0.004
LIST002	PLEX2	VIC	2	0.367	0.401	0.074 ***	-0.012
LIST023	PLEX2	VIC	11	0.639	0.687	0.054 ***	0.004
LIST037	PLEX2	FAM	11	0.320	0.347	0.062 ***	0.035 *
LIST042	PLEX2	VIC	5	0.591	0.655	0.081 ***	0.010
LIST060	PLEX2	NED	6	0.411	0.689	0.053 ***	0.377 ***
LIST062	PLEX2	PET	4	0.311	0.335	0.088 ***	-0.005
					Mean over all loci	0.063 *	0.045
					Mean over all loci without locus LIST060	0.065 *	0.002

Table S3: Weir and Cockerham's (1984) estimates of *F*-statistics for Alsace and northern France region and subregions or subwatersheds. In the two subregions where only two populations were sampled, only the pairwise F_{ST} is reported. *: $P < 0.05$.

	Mean distance between sites (km)	F_{IT}	F_{ST}	F_{IS}
Alsace	24.65 ± 15.96	0.025 ± 0.003 *	0.022 ± 0.001 *	0.003 ± 0.003
Northern France	118.24 ± 76.45	0.136 ± 0.016 *	0.135 ± 0.012 *	0.001 ± 0.01
Subwatersheds in Alsace				
A6	8.41 ± 7.32	0.007 ± 0.025	0.015 ± 0.005 *	-0.008 ± 0.022
A5	3.93 ± 2.69	0.019 ± 0.010 *	0.015 ± 0.005 *	0.004 ± 0.010
A1	7.26 ± 5.13	0.023 ± 0.011	0.016 ± 0.004 *	0.006 ± 0.012
A2	3.91 ± 3.77	0.044 ± 0.018 *	0.024 ± 0.005 *	0.020 ± 0.018
A4	14.64 ± 10.92	0.015 ± 0.009	0.016 ± 0.003 *	-0.001 ± 0.008
A3	4.55 ± 3.58	0.006 ± 0.027	0.000 ± 0.004	0.006 ± 0.026
Subregions in Northern France				
N3	12.47 ± 9.40	0.089 ± 0.012 *	0.078 ± 0.006 *	0.012 ± 0.014
N6	6.12 ± 5.12	0.013 ± 0.040	0.029 ± 0.019	-0.017 ± 0.031
N2	4.32 ± 4.21	0.023 ± 0.020	0.017 ± 0.006 *	0.006 ± 0.020
N5	16.52 ± 19.28	0.267 ± 0.045 *	0.241 ± 0.045 *	0.034 ± 0.030

	population pairs	Distance between sites (km)	Pairwise F_{ST}
N1	Foc1/Gen1	3.1	-0.002
N4	Cuy/Wal1	8.42	0.106 *

Chapitre 2. Développement de marqueurs SNP pour l'agrion de Mercure, une espèce sans génome de référence, et leur application pour caractériser la structure génétique à fine échelle de populations du Nord de la France



Les marqueurs génétiques moléculaires sont une source d'information cruciale dans le domaine de la biologie de la conservation et de la biologie évolutive. En effet, ils permettent d'étudier la distribution spatiale de la diversité génétique, d'évaluer les niveaux locaux de consanguinité, d'estimer des tailles efficaces de population, et de délimiter finement des unités de conservation. Les marqueurs microsatellites utilisés dans le chapitre précédent donnent une image partielle de la structure génétique des populations. Récemment, les marqueurs de type SNP ont gagné en popularité, parce qu'ils sont abondants et largement distribués à travers le génome, permettant ainsi une plus grande précision dans l'estimation des paramètres classiques de génétique des populations.

L'un des principaux obstacles à leur application en génétique de la conservation est la nécessité de génotyper un grand nombre d'individus avec des milliers de marqueurs, ceci chez des espèces non modèles pour lesquelles les données génomiques sont souvent absentes. Les approches de représentation réduite du génome permettent d'identifier et de séquencer de nombreux marqueurs de type SNP chez

des espèces non modèles. En complément, les méthodes d'enrichissement ciblé permettent d'obtenir ces données sur un grand nombre d'individus.

Dans ce contexte, nous avons appliqué une méthode hybride combinant une approche de représentation réduite du génome (dénommée ddRADseq) pour identifier des SNPs putatifs, que nous avons ensuite ciblés en utilisant la méthode d'enrichissement dite « Allegro Targeted Genotyping », pour génotyper 1920 agrions de Mercure sur 6000 SNPs. Nous avons pu identifier *de novo* des centaines de milliers de marqueurs SNP, mais l'enrichissement ciblé n'a permis d'obtenir des génotypes fiables que pour 2 100 SNPs et 1080 individus. En effet, nous avons rencontré de nombreux problèmes : faibles rendements en ADN après enrichissement ciblé, sondes non retrouvées, différences de profondeur de couverture après séquençage, ou encore des estimations de valeurs de F_{IS} aberrantes d'un point de vue biologique. Nous formulons donc des recommandations pour identifier et, idéalement, prévenir de tels problèmes lors de l'application d'une telle méthode à d'autres espèces sans génome de référence.

Au-delà de ces problèmes techniques, le filtrage des locus a toutefois permis d'obtenir 2 092 locus de qualité, que nous avons donc utilisés pour caractériser la structure génétique fine de cinq populations d'agrion de Mercure localisées sur la Côte d'Opale dans le nord de la France. Ces marqueurs ont notamment permis d'identifier deux individus hybrides putatifs de première génération entre deux populations adjacentes, ce que ne permettait pas la résolution des marqueurs microsatellite précédemment utilisés.

Cet article a été soumis à *Journal of Heredity* début mai 2024.

Pitfalls and recommendations for large-scale SNP genotyping in a non-model endangered species: the southern damselfly (*Coenagrion mercuriale*) as a case study.

Agathe Lévêque [1,2], Jean-François Arnaud [1], Vincent Vignon [2,3], Clément Mazoyer [1], Cécile Godé [1], Anne Duputié [1]

1. Univ. Lille, CNRS, UMR 8198 - Evo-Eco-Paleo, F-59000 Lille, France

2. Office de Génie Écologique (O.G.E.), F-67200 Strasbourg, France

3. ALKIOS, F-84911 Avignon, France

.

Abstract

Genomic markers are essential tools for studying species of conservation concern, yet non-model species often lack a genome reference. Here we describe a methodology for identifying and genotyping thousands of SNP loci in the southern damselfly (*Coenagrion mercuriale*), a bioindicator of freshwater stream quality classified as near-threatened, with locally declining populations. We used a hybrid approach combining reduced representation sequencing and target enrichment. First, we identified putative SNP loci using ddRADseq and *de novo* assembly. Then, single primer enrichment technology targeted 6,000 of these SNPs across 1,920 individuals. Challenges encountered included sequence recapture failure, coverage depth discrepancies, and aberrant F_{IS} values. We provide recommendations to address such issues. After multiple filtering, we retained 2,092 SNPs. We used them to characterise rear-edge populations of the southern damselfly in Northern France, a region where populations are sparsely distributed. Previous surveys utilising microsatellite markers allowed comparison of genetic diversity and differentiation estimates. Consistent with prior findings, genetic diversity estimates were similar across the studied populations that showed no sign of inbreeding. SNP markers exhibited greater resolution in detecting fine-scaled genetic structure, identifying two putative hybrids in adjacent populations, a feat unattainable with microsatellite loci. Altogether, this study highlighted the ongoing challenge of large-scale SNP genotyping using target sequencing techniques in non-model species to set conservation guidelines. Nonetheless, these new markers showed greater statistical power in identifying conservation units and offered the promise of greater precision in the identification of admixture events or the estimation of key population parameters such as effective population size.

Keywords: Allegro target genotyping, conservation genomics, ddRADseq, non-model species, single nucleotide polymorphisms

Introduction

Neutral molecular genetic markers are a crucial source of information in the fields of population genetics, evolutionary biology, and species conservation. They can be used for a wide range of applications, such as measuring the strength of micro-evolutionary processes of gene flow, genetic drift, mutation and selection, studying the spatial distribution of genetic diversity, estimating effective population sizes or delimiting conservation units. Therefore, it makes them essential tools in the establishment of conservation efforts, notably for endangered species (Allendorf *et al.*, 2010; Holderegger *et al.*, 2019; Hohenlohe *et al.*, 2021; Allendorf *et al.*, 2022). Many different approaches have been developed to genotype sets of molecular markers in a context of conservation biology, each with its own benefits and limitations (Davey *et al.*, 2011; Allendorf, 2017). Because of their polymorphism, reproducibility, ease of development and genotyping, microsatellite markers were widely used to study natural populations (Frankham *et al.*, 2013; Allendorf, Funk, Aitken, Byrne, Luikart, *et al.*, 2022). However, in the last decade, these markers have rapidly been replaced by single nucleotide polymorphism (SNP) markers, which are more abundant and widely distributed across the genome, offering the promise of greater precision and statistical power in the estimation of population genetics parameters (Allendorf *et al.*, 2010; Harrisson *et al.*, 2014). The success of this new type of molecular marker is intimately linked to the development of next-generation sequencing (NGS) technologies, which has led to an increasingly large amount of genomic data being generated for wild populations (Harrisson *et al.*, 2014; Allendorf, 2017; Hohenlohe *et al.*, 2021). However, integrating these new genomic tools in conservation practices remains challenging (Allendorf *et al.*, 2010; Shafer *et al.*, 2015; Allendorf, 2017). Studying the spatial genetic structure of wild populations across space requires sampling hundreds or even thousands of individuals in multiple populations (Narum *et al.*, 2013). Although it is now possible to sequence whole genomes for a subset of individuals, reduced genomic data can provide sufficient and valuable information to answer many scientific questions in conservation biology. Moreover, obtaining complete genome sequences for many individuals considerably increases the cost and complexity of bioinformatics analyses, particularly when the genome sequence of the focal species is unknown (Narum *et al.*, 2013). Consequently, one of the main impediments to the application of genomic tools to the conservation of wild populations is the need to

genotype thousands of individuals with thousands of genetic markers in species for which often, little or no reference genomic data is available (Seeb *et al.*, 2011; Narum *et al.*, 2013).

Various strategies for sequencing and genotyping a sub-part of the genome emerged as powerful alternatives to whole genome sequencing (Davey *et al.*, 2011; Jones and Good, 2016). Among these approaches, reduced representation sequencing (RRS, Narum *et al.*, 2013; Andrews *et al.*, 2016; Campbell *et al.*, 2018) and target enrichment methods (Mamanova *et al.*, 2010) appeared to be complementary in terms of genomic knowledge required for their application, the effort and time required for their development, and their ability to genotype a large number of samples.

RRS approaches such as RAD-seq (Baird *et al.*, 2008), genotyping-by-sequencing (GBS, Elshire *et al.*, 2011) and double-digest RAD sequencing (ddRAD; Peterson *et al.*, 2012) make it possible to identify and sequence many SNP markers in non-model species (Narum *et al.*, 2013). These methods rely on the reduction of genomic complexity by using restriction enzymes to produce short-sequenced fragments that provide the frame for the discovery of SNPs widely spread across the genome (Van Tassell *et al.*, 2008). The major advantage of these techniques is that they do not require any prior genomic information, which makes them particularly applicable in the conservation of non-model species for which there is often no available reference genome (Peterson *et al*., 2012). In addition, the development of ddRAD sequencing and associated computer tools (e.g. STACKS; Catchen *et al.*, 2013) has made it possible not only to discover polymorphic sites but also to construct *de novo* consensus sequences (RADloci). However, these methods have some limitations, especially because the random distribution of enzymatic cut sites impedes targeting specific loci, which reduces the reproducibility between studies and spreads the sequencing effort over a very large number of sites (Andrews *et al.*, 2016; Barchi *et al.*, 2019; Scaglione *et al.*, 2019).

In parallel with the development of RRS techniques, target enrichment methods arose as viable alternatives for obtaining cost-effective genome-wide genotypic data with higher reproducibility (Meek and Larson, 2019). These methods are based on the design of DNA capture beads or PCR amplification probes that target specific genomic regions, focusing sequencing efforts on the targeted regions, reducing the cost per sample, and thereby increasing the number of samples that can be genotyped (Mamanova *et al.*, 2010; Jones and Good, 2016; Meek and Larson, 2019). Tecan Genomics (Redwood

City, CA) developed the Allegro Targeted Genotyping system (ATG), a ready-to-use technique for custom targeting of previously identified SNPs (Scaglione *et al.*, 2019). This technique presents the advantages of moderate cost and subcontracted probe development, whilst having a flexible design and ready-to-use library preparation kits. In addition, the use of individual combinations of barcodes allows pooling several thousand samples, and interrogating thousands of SNPs in a single sequencing run. Based on a single primer enrichment technology (SPET), ATG relies on the design of ~40 bp long probes based on reference genomes or transcriptomes, enabling to target with high reproducibility specific genomic regions carrying genetic variants of interest. The method has so far mainly been used in human health (Scolnick *et al.*, 2015; Nairismägi *et al.*, 2016; Saber *et al.*, 2017) or plants of agronomic interest (Barchi *et al.*, 2019; Scaglione *et al.*, 2019; Baccichet *et al.*, 2022; Tripodi *et al.*, 2023) and its use on wildlife species for conservation purposes is only starting to develop (Gramazio *et al.*, 2020; Gavriliuc *et al.*, 2022).

The main limit to the application of target enrichment methods is the need to have prior knowledge of the genome and the genetic variants to be targeted. To overcome this issue, several hybrid approaches were developed in recent years to genotype genomic variants in a very large number of samples at a reduced cost and without any prior genomic knowledge (Campbell *et al.*, 2015; Ali *et al.*, 2016; Meek and Larson, 2019). These methods combine a RRS approach to identify putative SNPs, which are then targeted for enrichment. Here, we applied such a hybrid method to develop and genotype thousands of SNPs in numerous samples of a non-model species with high conservation concerns, the southern damselfly *Coenagrion mercuriale* (Odonata, Zygoptera).

The southern damselfly is a particularly interesting study model to integrate evolutionary genomics with ecology, as it exhibits a complex biphasic (aquatic and terrestrial) life-cycle and represents a bioindicator of freshwater stream quality (Bybee *et al.*, 2016; Watts *et al.*, 2004c). The species is restricted to western and southern Europe, and became almost extinct in seven European countries on the eastern edge of its geographic distribution, where many populations are threatened by habitat loss or deterioration (Grand, 1996). Consequently, the southern damselfly is regarded as Near Threatened by the IUCN (Boudot, 2020), listed on Appendix II of the Bern Convention and on Annex II of the EC Habitat directive, in addition to a national protection in France where the southern damselfly

is widely distributed and locally abundant, except in the northernmost regions (Houard, 2020; Fierimonte and Vanappelghem, 2021). In these northernmost regions, southern damselfly populations are stable but sparse, mainly due to modern intensive agricultural practices with the shift from grasslands to irrigated ploughed fields, in addition to the closure of waterways by scrubs and forests (Fierimonte and Vanappelghem, 2021). In northern France, southern damselfly populations are monitored by biodiversity managers, and major demographic declines were recorded with only a few individuals found in some years (Vanappelghem and Hubert, 2010). In this context, the conservation status of several populations in the region was previously described in the light of levels of genetic diversity and estimates of gene flow among populations (Lorenzo-Carballa *et al.*, 2015; Lévêque *et al.*, 2024).

In this study, we used a two-step approach for the massive genotyping of thousands of SNPs in numerous southern damselfly populations. First, we used ddRADseq library preparation and a *de novo* assembly analysis to construct reference contigs and to identify SNPs markers using paired-end sequencing (Peterson *et al.*, 2012; Rochette *et al.*, 2019). Secondly, we used a single primer enrichment technology (Allegro Targeted Genotyping) to target six thousand SNPs and massively genotype thousands of southern damselfly individuals from various populations. We discuss the technical issues associated with the application of targeted enrichment genomic techniques on a non-model species. Finally, we carried out population genetic analyses to assess the levels of genetic diversity and of genetic differentiation, and to identify population genetic affiliation in a set of five threatened southern damselfly populations located in northern France. These populations were previously studied using microsatellite markers (Lorenzo-Carballa *et al.*, 2015; Lévêque *et al.*, 2024), which enables us to compare our results with previous findings and to suggest monitoring actions of these populations.

Materials and methods

1/ Reduced-representation genome sequencing and de novo SNP discovery

Sampling and DNA extraction

To build reference contig sequences and to identify SNPs present in diverse populations, we collected twenty males of southern damselflies from various locations in France in spring 2020 and 2021

(see Table S1). Whole individual samples were stored in 100% ethanol prior to DNA extraction. Abdomens were removed to avoid sampling intestinal microbiota. We then crushed the remaining tissues using five MN Beads Type D (Macherey-Nagel, Duren, Germany). We extracted total genomic DNA from each sample using NucleoMag® Tissue Kit (Macherey-Nagel) according to the manufacturer's recommendations. We then quantified DNA using a Qubit flex fluorimeter (Thermo Fisher Scientific, Waltham, Massachusetts, USA).

ddRAD library preparation and sequencing

To reduce the genome complexity, we prepared a double-digest restriction site-associated DNA sequencing (ddRADseq) library following the standard protocol described in Peterson *et al* (2012). In brief, 250 ng of DNA was digested with the restriction enzymes PstI (recognising the rare CTGCA/G motif) and MspI (recognising the frequent motif C/CGG; both enzymes from New England Biolabs, Ipswich, MA, USA). After the ligation of two adapters P1 and P2, each containing a unique barcode sequence, DNA was cleaned with 1X AMPure XP magnetic beads (Beckman-Coulter, Brea, CA, USA). Then, ligated fragments were amplified for 16 cycles using KAPA Hifi Hot Start PCR kit (Roche, Basel, Switzerland), and purified again with 1X AMPure XP beads. During the PCR, Illumina flow cell annealing sequences and a second index barcode were appended to the sequences, giving each sample a unique combination of barcodes for downstream identification. Next, after a size selection by AMPure XP double SPRI (0.5X and 1X), libraries were pooled in equimolar concentrations. A final control step was performed with Agilent 2100 BioAnalyzer, and the pooled libraries were then sequenced for 250 bp paired-end on a NovaSEQ6000 instrument (LIGAN-PM sequencing platform at EGID, Lille, France).

Building reference sequences and identifying SNPs

We used STACKS software pipeline v2.60 (Catchen *et al.*, 2013; Rochette *et al.*, 2019) following the *de novo* analysis protocol described by Rivera-Colón and Catchen (2022) to assemble RADloci from restriction-digested short-read sequences and to identify SNPs (Figure 1). First, raw sequence reads were cleaned and demultiplexed using the *process_radtags* module with the options to

remove reads with any uncalled base (-c), discard reads with low-quality scores (-q) and rescue barcodes (-r) applied with default values. Sequence quality assessment was then performed using the software FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). All cleaned sequences were then trimmed to 200 bp using Trimmomatic v1.3.1 (Bolger *et al.*, 2014). The *denovo_map.pl* pipeline (*ustacks*, *cstacks*, *sstacks*, *tsv2bam*, *gstacks* and *populations* modules) was then used to call SNPs and build the RADtag catalog, with the number of mismatches allowed to merge into one locus as M=9 (*ustacks*) and the number of mismatches allowed among individuals when building catalogs to n=9 (*cstacks*). These parameters were chosen after extensive exploration of the parameter space and choice of parameters appropriate for this dataset, following the R80 methods proposed by Paris *et al.* (2017). Finally, loci were filtered using the STACKS *populations* module to retain only loci shared by 80% of individuals (r = 0.8; Figure 1).

2/ Massive SNP genotyping with Single Primer Enrichment Technology

Designing probes for target SNP loci

The Allegro Target Genotyping (Tecan Genomics Inc., Redwood City, CA) used in our assay implements Single Primer Enrichment Technology (SPET), which involves hybridisation of custom-designed probes near target SNPs, followed by probe extension, the addition of sequencing adapters, and high-throughput Illumina sequencing. We first selected SNPs out of the 80 bp of the start or end of the RADcontigs sequence and removed RADcontigs with uncalled bases (N), leaving 22,224 eligible SNPs distributed over 10,419 RADlocus. Probes for the target SNPs were about 40 bp long and were custom-designed by Tecan using the previously identified RADloci. Two probes were designed per target SNP, one per DNA strand. From the first panel of probes build by TECAN, we selected probes with a unique alignment, we kept only one pair per RADlocus and we then randomly sub-selected 6,000 couples of probes (Figure S1). Our final panel thus contained 12,000 probes targeting 6,000 SNPs occurring each on a different RADcontig (Table S2).

Sampling and DNA Extraction

We sampled 1,920 southern damselflies from various locations in France, with 1 to 32 individuals sampled per population (mean 16.55 ± 10.07). We collected them with an insect net in late spring (May to July) from 2020 to 2022. Whole individuals were stored in 100% ethanol, and DNA was extracted following the procedure described above. In the present paper, we only focus on 50 individuals sampled in Northern France to analyse the levels of genetic diversity and the extent of genetic differentiation. Indeed, these populations are all located at the edge of the geographic distribution of the species (Figure 2A) and were previously characterised using a set of ten microsatellite markers. This allows us to compare the genetic structure observed using classical microsatellite with that observed with the newly developed SNP markers (Lévêque *et al.*, 2024).

SPET library preparation

We followed the guidelines of the manufacturer available at https://www.tecan.com/hubfs/HubDB/Te-DocDB/pdf/UG_Allegro_Targeted_Genotyping_V2.pdf, to prepare the SPET libraries. We used the Allegro Metaplex Module, an optional module provided by TECAN to increase sample multiplexing capacity during library preparation. By default, Allegro Targeted Genotyping V2 includes 192 unique 1/i7 barcode for tagging samples. The Metaplex kit allows adding a second barcode (2/i5 barcodes) during the library amplification step. Sixteen 2/i5 barcodes are available from TECAN genomics: MP2-BC01 to MP2-BC16. We used the first ten 2/i5 barcodes, enabling all 1,920 samples to be identified in subsequent analyses by a unique combination of barcodes (Table S3). After qualitative and quantitative analysis of each pool of 48 individuals with a Qubit flex fluorimeter (Thermo Fisher Scientific, Waltham, Massachusetts, USA) and a 5200 Fragment Analyzer System (Agilent Technologies, Santa Clara, CA, USA), we pooled them equimolarly into a single tube containing all 1,920 samples. After a final quantity and quality control step was carried out, the pooled genomic library was sequenced on a MiSeq System (1x150 bp, Illumina Inc., San Diego, CA, USA sequenced at GenoScreen sequencing platform, Lille, France).

Sequence filtering and SNP genotyping

First, we removed low-quality reads and trimmed the Illumina adapters of the raw demultiplexed reads using *fastp* v0.23.4 (Chen *et al.*, 2018) with the parameters -y, -5, -3 (enabling the low-complexity filter and dropping low-quality bases from both ends of the reads; see Figure 1). Then, we removed the 768 individuals that were tagged with i5 barcodes MP2-BC07 to MP2-BC10, because they had about 9 times fewer reads than those labelled with the six remaining i5 barcodes (Table S3, Table S4). We also removed all reads not starting exactly with a 40-base sequence corresponding to one of the probes (Table S4).

The remaining reads were then aligned to the previously built RADloci using *Bowtie2* v2.4.1 (Langmead and Salzberg, 2012), with local alignment and parameters -L 32 -D 20 -R 3 -N 0 -i S,1,3 (Figure 1). SNP calling was performed using *GATK*-4.0 v4.2.2.0 (DePristo *et al.*, 2011) following the software best practices workflow for germline short variant discovery. Analyses included the following steps: (i) per-sample variant calling on target regions using *HaplotypeCaller* with default parameters to create a GVCF file for each sample; (ii) the GVCF files were then consolidated using GenomicsDBImport in order to improve scalability and speed the next step, (iii) joint genotyping using GenotypeGCVFs with default parameters to produce a set of joint-called variants; (iv) selection of SNPs using *SelectVariants* and quality filtering of SNPs using *VariantFiltration* (filter expression used : QD < 2, QUAL < 30, MQRankSum < -12.5, MQ < 30; Figure 1).

An extra filtration of the genotype matrix was performed using *VCFtools* v0.1.16 (Danecek *et al.*, 2011) and *BCFtools* v1.12 (Danecek *et al.*, 2021). We tested a range of filter thresholds at each filtering step: loci depth of coverage, amount of missing data per loci, minor allele frequency (MAF) and the number of missing data allowed per individual (Table S5). After evaluating the effect of the stringency of each of these filters, we chose to code genotypes with fewer than 8 reads in depth of coverage as missing, and to retain only biallelic SNPs present in >50% of individuals and with minor allele frequency >0.01. After that, we selected a single SNP per RADlocus (the one with the least missing data), thus reducing linkage disequilibrium issues in subsequent population genetic analyses (Waples *et al.*, 2022). We then filtered out individuals with >40% missing SNPs. After that, we removed loci showing more than twice the mean read depth, because variants that have a high mean depth across all samples can be indicative of either repetitive regions or paralogs (Li, 2014; O'Leary *et al.*, 2018;).

Finally, we only kept SNPs showing F_{IS} values within a range of [-0.2 ;0.2]. This criterion aimed to exclude loci with excessive bias in homozygotes or heterozygotes, suggestive of biologically irrelevant genotypic structures (Figure 3). Indeed, surveyed populations were known to be under classical Hardy-Weinberg equilibrium using microsatellite data (Lorenzo et al. 2015; Lévêque *et al.*, 2024*)*.

3/ Population genetic analyses

Genome scans for selection

We used the R v4.2.1 (R Core Team, 2022) package *pcadapt* v4.3.5 (Luu *et al.*, 2017; Privé *et al.*, 2020) to identify outlier loci potentially under selection, applying the protocol found at https://bcm-uga.github.io/pcadapt/articles/pcadapt.html, with the default Mahalanobis distance method. The number of principal components was selected using the scree plot method, which is based on Cattell's rule stating to keep principal components up to the first inflexion point of the eigenvalues (Cattell, 1966). All SNPs showing *P*-values < 0.05 after adjustment for multiple testing through false discovery rate (Benjamini and Hochberg, 1995) were excluded from further analyses.

Genetic diversity, genetic differentiation, and population affiliation

We selected 50 individuals from five populations located in Northern France, towards the edge of the geographic distribution of the species. These populations were named Bai1, Bai2, Bai3, Cam1 and Isq and mapped on Figure 2A. The first three populations Bai1, Bai2 and Bai3 were located on the same watercourse and less than 2 km apart; populations Cam and Isq were more isolated and about 10-15 km away from populations Bai1-3.

We used the R packages *vcfR* v1.14.0 (Knaus and Grünwald, 2017), *dartR* v2.7.2 (Gruber *et al.*, 2018; Mijangos *et al.*, 2022), *poppr* v2.9.3 (Kamvar *et al.*, 2015) to navigate between the file formats of the genotype matrix (vcf, genlight, genind). We estimated standard population genetic statistics (observed and expected heterozygosity, allelic richness) using the R library *hierfstat* v0.5-11 (Goudet, 2005). Allelic richness (A_r) was rarefied on a minimal sample size of six individuals using the procedure of El Mousadik and Petit (1996).

To assess the levels of genetic differentiation between pairwise populations, we calculated pairwise F_{ST} estimates according to Weir and Cockerham (1984) using *StAMPP* v1.6.3 R package (Pembleton *et al.*, 2013). Confidence intervals and significance of F_{ST} values were obtained using 1000 bootstraps over loci. Global levels of F_{IS} and F_{ST} were assessed using the R package *hierfstat* v0.5-11 (Goudet, 2005).

We then visualised the pattern of population affiliation using a principal component analysis (PCA) and a discriminant analysis of principal components (DAPC; Jombart *et al.*, 2010), both implemented in the R package *adegenet* v2.1.9 (Jombart, 2008; Jombart and Ahmed, 2011). We used the function *find.cluster* for a number of clusters *K* varying from 1 to 5 to identify the optimal *K* (i.e. the one minimising the BIC). In addition, we also used the sampling populations as prior groups. Then, for groups built for *K*=2 to 5, we used a cross-validation procedure (*xvalDapc* with 1,000 replicates) to identify the optimal number of principal components retained to perform the DAPC, and subsequently applied it to assign individuals to a cluster and calculate their population membership probabilities.

Results

Reduced-representation genome sequencing and de novo SNP discovery

The Illumina NovaSeq 6000 SP 2x250bp sequencing produced a total of 1.28 billion reads, of which 1.12 billion (87.2%) passed initial quality filters and demultiplexing. Reads were filtered out due to ambiguous barcode sequences (6.56%), ambiguous RAD tags (6.18%) and low-quality scores (0.06%). After trimming reads to 200 bp and removing reads shorter than 200 bp, 199 million reads remained. The catalog generated by STACKS included 320,556 RADloci with a mean per-sample depth of coverage of 43.6X (s.d. 15.3X, from 19.6X to 82.3X). This number was reduced to 135,468 RADloci (mean length of 330.72 bp, s.e. 0.20) after requiring them to be present in at least 80% of the 20 genotyped individuals, resulting in a matrix of 758,786 SNPs.

Massive SNP genotyping with Allegro Target Genotyping

The Illumina MiSeq SE 1x150bp sequencing produced a total of 1.784 billion single-end raw reads corresponding to an average of 929,223 reads per sample (from 54,628 to 3,926,551; Table S3). After

adapter trimming and quality filtering, 99.98% of the initial reads remained. When removing the 768 individuals labelled with metaplexes MP2-BC07 to MP2-BC10 (which yielded too few reads per sample to be analysed; Table S3), and when removing sequences that did not begin exactly by a designed probe, the overall alignment rate on the RADloci was on average 99%. After SNP calling and the first quality filter, we identified 154,324 bi-allelic-SNPs distributed across 5,556 RADloci. Extra filtration of the SNPs produced a matrix of 2,541 SNPs containing 1,080 individuals genotyped for at least sixty percent of the loci (Table S5). After filtering out loci showing aberrant F_{IS} values, we obtained 2,100 high confidence SNPs (Figure 3). Eight SNPs were detected as outliers by *pcadapt* (keeping 2 PCs). These SNPs were therefore excluded to obtain a final dataset of 2,092 SNPs for population genetic analyses.

Population genetic analyses

Using these newly developed SNPs, we observed similar levels of genetic diversity (H_o, H_e, A_r) among the five surveyed populations located in Northern France (Table 1), with observed and expected heterozygosity ranging from 0.195 (populations Bai3 and Isq) to 0.207 (Cam1) and from 0.185 (Isq) to 0.205 (Cam1) respectively. Rarefied allelic richness ranged from 1.568 (Bai1) to 1.667 (Bai2) and F_{IS} levels ranged from -0.054 (population Isq) to -0.010 (for populations Bai2, Bai3, Cam1) (Table 1). As observed using these SNPs, microsatellite markers showed very similar patterns of genetic diversity among populations, with no evidence of departures from Hardy-Weinberg expectations (Table S6).

In terms of genetic differentiation, all populations showed significant levels of pairwise genetic differentiation, with F_{ST} values ranging from 0.014 to 0.108 (Table 2). Although geographically very close, the three populations located on the same watercourse (Bai1, Bai2, Bai3) showed significant pairwise F_{ST} ranging from 0.030 to 0.037 (Table 2). Compared with microsatellite data (mean multilocus F_{ST} = 0.068 and mean multilocus F_{IS}=-0.089), the overall F_{ST} and F_{IS} in the studied region were of 0.055 and -0.022 respectively for the SNP data (Table S7).

With regards to population genetic affiliation, the PCA revealed a clear geographical and genetic distinctiveness between two sets of populations: (i) populations Bai1, Bai2, and Bai3, and (ii) populations Cam1 and Isq located further away (Figure 2B). For the DAPC, the optimal number of clusters was K=2, distinguishing the same two groups of populations (Figure 2D, Figure S3A).

Visualising the DAPC results for *K*=3 clusters further allowed to separate specific individuals of populations Bai3 from those of populations Bai1 and Bai2 (Figure 2D, Figure S1B). With increasing numbers of clusters (*K*=4 and *K*=5), differences arose in individual assignment within populations Bai1, Bai2 and Bai3 (Figure 2D, Figure S1C,D). Finally, using sampled populations as a prior of assignation, most individuals were assigned to their own population, except two individuals who were assigned to two geographically close populations, suggesting admixture events between populations along the Bai river (Figure 2C,D).

Discussion

1/ Pitfalls and recommendations when producing a SPET dataset for a non-model organism.

Here we summarise and discuss the problems we faced, the solutions we explored, and the recommendations that emerged when genotyping thousands of individuals at thousands of SNPs loci using Allegro Targeted Genotyping (ATG) method, with contigs constructed from a ddRAD library for a non-model species with no genomic data available.

Differences in ATG library DNA yields and their consequences on sequence data available

We used ten of the sixteen supplied Allegro i5 indexes (Table S3) to multiplex 1,920 samples using dual indexing of 48-sample pools. However, although the laboratory preparation conditions were identical for all pools, we obtained very different final DNA yields depending on the i5 indexes (Table S4). Indeed, those labelled MP2-BC07 to MP2-BC10 showed 2-3x lower DNA yields than the other six indexes. Even though the pools were brought back to equimolar proportions to form the final pooled library, we obtained around 9 times fewer reads for individuals labelled with these four indexes (Table S3).

As a consequence, all 768 individuals labelled with these four indexes exhibited low numbers of sequences resulting in a low SNP depth of coverage and subsequently a very low confidence in genotypes. Thus, we excluded all individuals marked with these four indexes from the rest of the analyses (resulting in only 1152 individuals being usable). Therefore, we recommend paying particular attention to the differences in the yield of Allegro libraries prepared with different indexes, because it

may have a major impact on the resulting data, even if the pools are brought to equimolar concentrations prior to sequencing.

Differential probe efficiency affects the number of suitable SNP

In Allegro Targeted Genotyping, the first 40 bp of each read should correspond to the sequences of a designed probe. We carried out preliminary analyses before the filtering, alignment and variant calling steps to check the sequences captured and the efficiency of the probes.

First, we examined whether all the sequences obtained began exactly with the 40 bp of the designed probes. 85.5% of the sequences we obtained did not exactly match any probe sequences, owing to insertions, deletions, or substitutions. This could indicate that during capture, probes may have hybridised with other non-targeted but genetically close portions of the genome. This may indicate possible paralogues or copy number variant in the southern damselfly genome (Verdu *et al.*, 2016; Karunarathne *et al.*, 2023). Consequently, we discarded sequences resulting from these likely parasitic hybridisations and whose first 40 base pairs did not exactly correspond to probe sequences.

Secondly, we checked which of the 12,000 designed probes were present in the first 40 base pairs of our sequences (0, 1 or 2 probes of the pair found, Table S4). Over the 6,000 initially designed pairs of probes, 704 were not found in any sequence. Consequently, the SNPs targeted by these probes cannot be found in our final dataset. Furthermore, on average over the 1,152 individuals remaining after the previous filtering step, only about half of the pairs of probes were completely recovered. Additionally, for about a third of the targeted SNPs, only one of the probes of the pair was found (Table S4). This drives variability in the sequencing depth between sites covered by two probes or by just a single probe (Scaglione *et al.*, 2019; Gavriliuc *et al.*, 2022).

Aberrant departures from Hardy-Weinberg expectations

F_{IS} levels per locus were slightly negative on average, with many loci showing values close to zero, indicating panmixia. However, some loci showed extreme F_{IS} values, with F_{IS} = -1 (all individuals were heterozygous and no homozygote was found among all individuals) or F_{IS} = 1 (all individuals were homozygous but for different states, with no heterozygote occurrence; see Figure 3). Nonetheless, the

same populations studied using microsatellite markers followed a Hardy-Weinberg equilibrium with non-significant F_{IS} levels (Lévêque *et al.*, 2024).

We hypothesised that extreme values in F_{IS} were due to differences in depth of coverage: strong excess in heterozygotes could be correlated with excessively high depths, indicating possible duplication of the target genome region, whereas biases toward excess in homozygotes could be due to too low sequencing depth, making it impossible to detect alternative variants (Song *et al.*, 2016; Lou and Therkildsen, 2022; Karunarathne *et al.*, 2023). However, there was no clear relationship between loci sequencing depth and the levels of F_{IS} (Figure 3), although loci with a high homozygote excess showed a slight tendency to exhibit low coverage.

We decided to filter our loci using F_{IS} bounds set to -0.2 and 0.2, to remove loci showing aberrant departures from Hardy-Weinberg equilibrium, given what is known on patterns of inbreeding in that species (Keller *et al.*, 2012; Lévêque *et al.*, 2024; Watts *et al.*, 2004a,b, Figure 2). From a strict population genetic affiliation point of view, these extreme F_{IS} values did not impact the delimitation of population boundaries, as the PCA and DAPC analyses carried out before and after the application of this filter on F_{IS} led to the same geographical patterns and similar conclusions in terms of genetic structuring whatever the threshold chosen (data not shown). Yet, such extreme values in F_{IS} can lead to highly false estimates of inbreeding levels and of effective population size (Hedrick, 2011; Allendorf, et al. 2022). We therefore strongly recommend that future users of SPET method check whether the F_{IS} values associated with each locus are biologically relevant, and verify any biases before carrying out any analyses devoted to describe intra-population genotypic structure before drawing any conclusions on mating system and on local level of inbreeding.

2/Population genetics of southern damselflies in Northern France

The southern damselfly populations located in northern France showed relatively similar levels of genetic diversity and genotypic structure (H_o, H_e, A_r, F_{IS}, Table 1) as previously described using microsatellite markers (Lorenzo-Carballa *et al.*, 2015). This pattern was still true in 2021 (Lévêque *et al.*, 2024, Table S6). Interestingly, all populations showed slightly negative F_{IS} values, indicating a slight

excess of heterozygotes compared with what is expected according to the Hardy-Weinberg equilibrium, an imbalance that was already markedly observed with microsatellite markers (Table S6).

In contrast to what can be observed with microsatellite data, significant levels of genetic differentiation were observed between all pairs of populations in Northern France, with an overall F_{ST} value of 0.055, indicating that gene flow among populations may be spatially constrained. This was true even at short geographical distances below 2 km, among the three populations located along the Bai watercourse. Nonetheless, pairwise levels of genetic differentiation estimated with SNP markers were lower than those observed with microsatellite markers in the same region (see Table S7 and Lorenzo-Carballa *et al.*, [2015]) and in more fragmented southern damselfly populations in the UK (Watts *et al.*, 2005, 2004c).

Finally, the study of population genetic structure using PCA and DAPC assignments showed a clear distinctiveness between two groups (Bai river and more coastal populations), which makes sense from a geographical point of view. Even at a finer spatial scale of less than 4km, population Bai3 appeared to be distinct from populations Bai1 and Bai2. This fine-scale population genetic distinctiveness is consistent with individual assignments obtained using a Bayesian clustering approach conducted on a similar set of populations using microsatellite markers (Lorenzo-Carballa *et al.*, 2015).

Thanks to these patterns of fine-scale spatial structure and genetic differentiation identified using SNP markers, the DAPC analysis identified two individuals as of mixed origin between populations Bai1 and Bai2, which may be considered as putative first-generation hybrid in each of these two populations. This is consistent with previous studies reporting that southern damselflies can migrate over short distances of about 2 km (Rouquette and Thompson, 2007; Lévêque *et al.*, 2024; Watts et al., 2004c). From a conservation point of view, this suggested that the connectivity of the population Bai3 belonging to the same watercourse population should be improved by habitat management.

More generally, the ecological methods and genetic tools used to study Odonata populations already provided relevant information on dispersal pathways and landscape barriers among populations (Watts *et al.*, 2006, 2007; Keller *et al.*, 2010; Van Strien *et al.*, 2012). These studies also highlighted a loss of genetic diversity in geographically isolated populations. Yet, little information is available on

the evolutionary consequences of genetic erosion in Odonates, including the adaptative processes involved. Genomic techniques offer the potential to address these questions (Bybee *et al.*, 2016).

Conclusion

We showed that the ddRADseq RRS approach allows to identify *de novo* hundreds of thousands of SNP markers in a non-model species of conservation concerns (Peterson *et al.*, 2012; Narum *et al.*, 2013). However, the target enrichment method, initially designed to target 6,000 SNPs in 1,920 individuals, only managed to recover 2,100 SNPs for 1,080 individuals. This low recapture rate sharply contrasted with rates achieved in species where reference genomic sequences were available for developing Allegro probes (Barchi *et al.*, 2019; Scaglione *et al.*, 2019; Tripodi *et al.*, 2023). This highlighted the ongoing challenge of large-scale SNP genotyping using targeted sequencing techniques in a non-model species. Future work using a combination of genomic and transcriptomic data targeting gene duplication, DNA loss, intron size variation or transposable elements may explain the difficulties we encountered in this study, but also may elucidate the putative mechanisms responsible for genome size variation in odonate species (Bybee *et al.*, 2016). However, these transcriptomic and genomic studies would benefit from the availability of transcriptomes and reference genomes that need to be developed in Odonates (Bybee *et al.*, 2016; Ioannidis *et al.*, 2017).

However, the set of SNPs markers we developed in this study provided valuable information on the levels of genetic diversity, genetic differentiation and the spatial genetic structure of populations in local populations initially thought to be genetically depauperate. Notably, these markers allowed the identification of fine-scale genetic structure and patterns of gene flow that were undetectable using microsatellite loci. These newly developed SNPs thus emerge as valuable resources for future conservation genetics studies designed to infer crucial population genetic parameters, such as effective population sizes and migration pathways in southern damselfly populations.

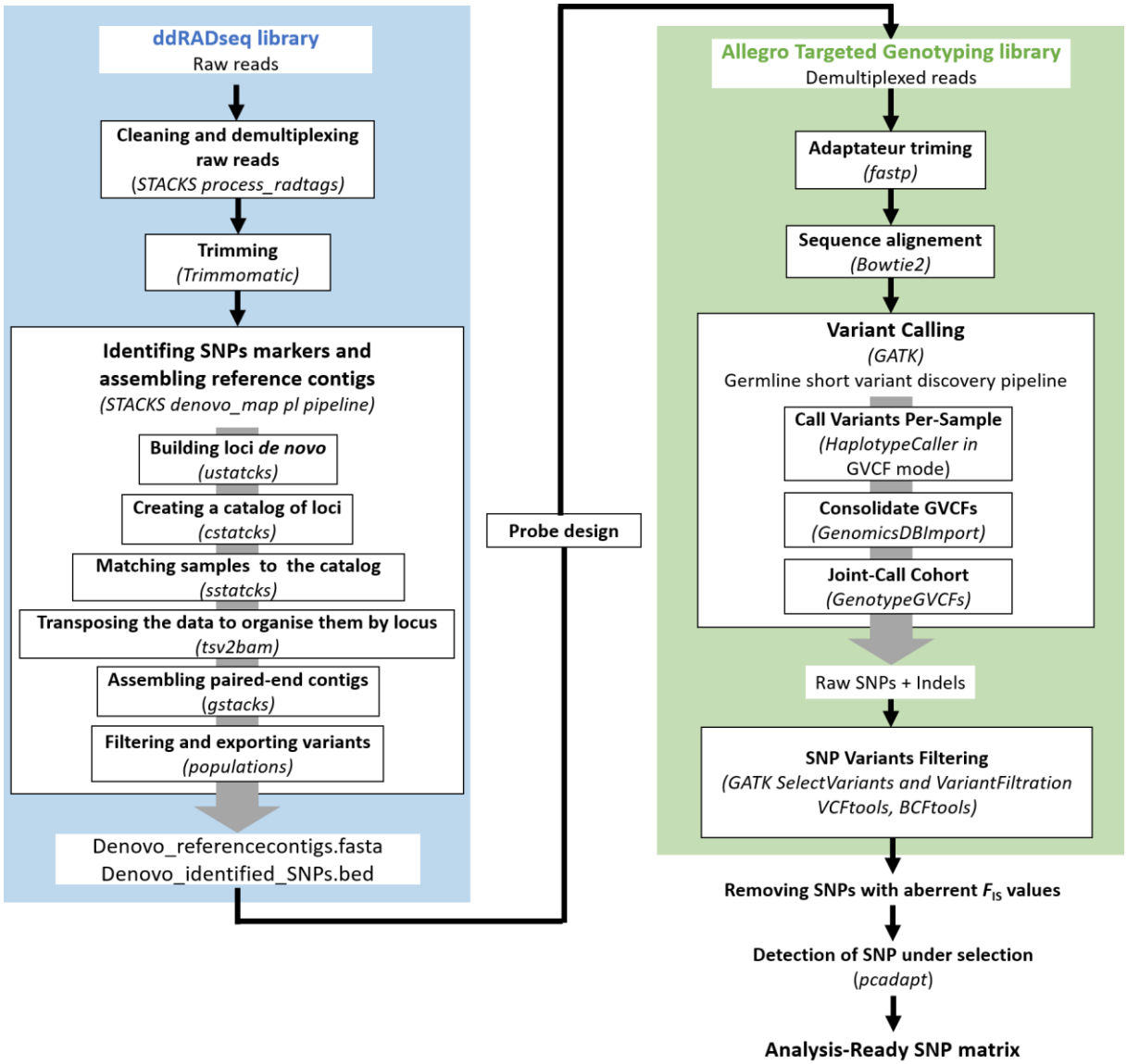


Figure 1: Bioinformatic pipeline for a two-step approach to massively genotype thousands of SNP markers in a non-model species, the southern damselfly (*Coenagrion mercuriale*). First, we used the STACKS software to identify *de novo* SNP variants and construct RADloci as reference sequences from a ddRADseq library (blue zone). Then, we massively genotyped the targeted SNPs through Allegro Targeted Genotyping developed by Tecan Genomics, and filtered the data to obtain high-quality genotypes (green zone).

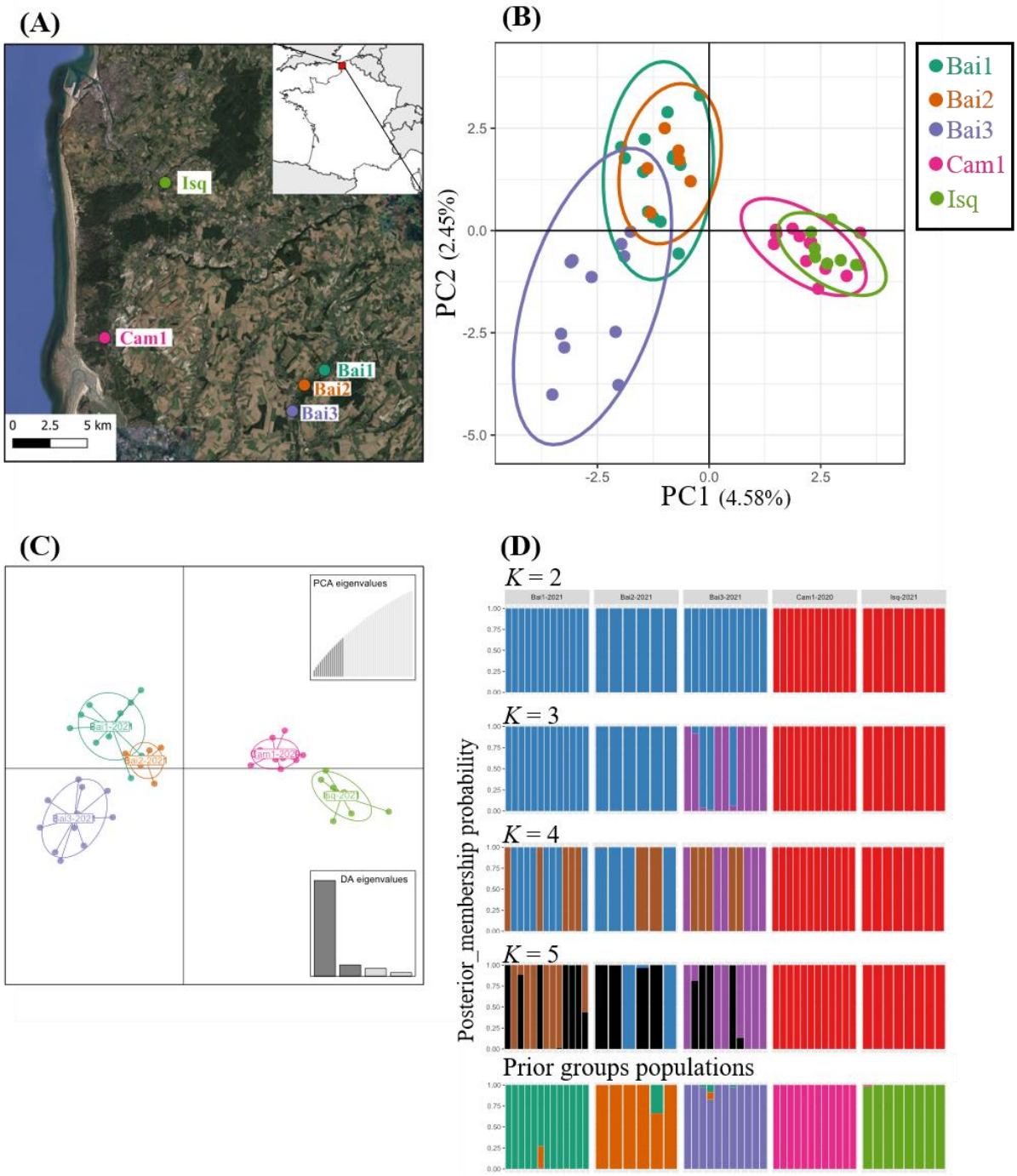


Figure 2: Patterns of genetic structure in five southern damselfly (*Coenagrion mercuriale*) populations located at range margin in northern France. **(A)** Map showing their geographical locations. **(B)** Principal Component Analysis (PCA), **(C)** Discriminant Analysis of Principal Components (DAPC) with sampling groups as priors, **(D)** assignment plots from DAPC for a number of clusters *K* varying from two to five with sampling groups as priors. Colors in panels B and C followed those in panel A; the same holds for panel D based on sampling populations as prior groups.

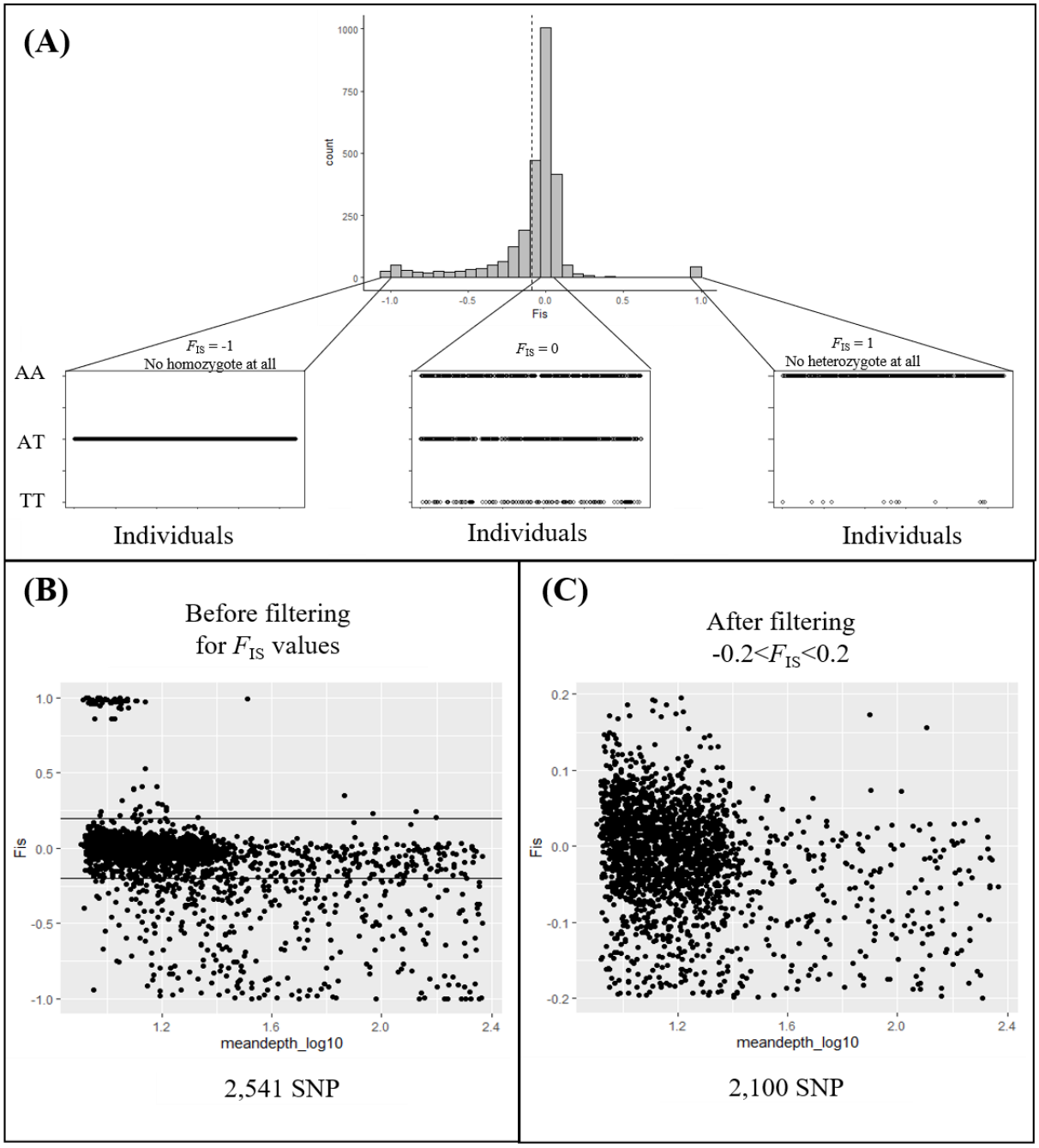


Figure 3: (A) Distribution of overall F_{IS} levels per locus in 1,080 individuals before filtering to retain only loci with F_{IS} levels between [-0.2 ; 0. 2], along with examples of genotypes of individuals for three loci characterised by F_{IS} =-1; F_{IS} =0, F_{IS} =1. **(B)** Before filtering on F_{IS} values, plot of the F_{IS} level per locus as a function of the average depth of reads on the RADloci carrying these SNPs, black bars indicate the limits of the filter that will be applied; **(C)** After filtering on F_{IS} values, plot of the F_{IS} level per locus as a function of the average depth of reads on the RADloci carrying these SNPs and distribution of F_{IS} levels per locus.

Table 1: Sampling information and mean multilocus estimates of genetic diversity using 2,092 biallelic SNPs in five southern damselfly (*Coenagrion mercuriale*) populations located at range margins of the species geographical distribution. Abbreviations: ID: population name. Longitude and latitude are provided (WGS84 ESPG 4326) coordinates. N: sample size; H_o: observed heterozygosity; H_e: expected heterozygosity; F_{IS}*:* the intrapopulation fixation index; A_r allelic richness.

ID	Sampling Year	Longitude (E)	Latitude (N)	N	Allele number	H_o	H_e	F_{IS}	A_r
Bai1	2021	1.8147	50.5564	13	3286	0.203	0.196	-0.036	1.568
Bai2	2021	1.7959	50.5472	6	2924	0.197	0.195	-0.010	1.667
Bai3	2021	1.7848	50.5314	11	3189	0.195	0.193	-0.010	1.571
Cam1	2020	1.6061	50.5738	12	3313	0.207	0.205	-0.010	1.605
Isq	2021	1.6611	50.6682	8	2997	0.195	0.185	-0.054	1.620

Table 2: Pairwise F_{ST} estimates between the five southern damselfly (*Coenagrion mercuriale*) populations located in Northern France, associated with 95% confidence interval, and their statistical significance levels. * $P \leq 0.05$, *** $P \leq 0.001$.

	Bai1	Bai2	Bai3	Cam1
Bai2	0.037 (0.023 - 0.051)***			
Bai3	0.030 (0.018 - 0.044)***	0.033 (0.019 - 0.049)***		
Cam1	0.051 (0.040 - 0.063)***	0.052 (0.037 - 0.068)***	0.076 (0.062 - 0.089)***	
Isq	0.083 (0.067 - 0.099)***	0.080 (0.061 - 0.101)***	0.108 (0.088 - 0.127)***	0.014 (0.000 - 0.033)*

References

Ali OA, O'Rourke SM, Amish SJ, Meek MH, Luikart G, Jeffres C, *et al.* (2016). RAD capture (Rapture): flexible and efficient sequence-based genotyping. *Genetics* **202**: 389–400.

Allendorf FW (2017). Genetics and the conservation of natural populations: allozymes to genomes. *Molecular Ecology* **26**: 420–430.

Allendorf FW, Funk WC, Aitken SN, Byrne M, Luikart G (2022). *Conservation and the Genomics of Populations*, 3rd edn. Oxford University Press: Oxford, UK.

Allendorf FW, Hohenlohe PA, Luikart G (2010). Genomics and the future of conservation genetics. *Nature Reviews Genetics* **11**: 697–709.

Andrews KR, Good JM, Miller MR, Luikart G, Hohenlohe PA (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics* **17**: 81–92.

Baccichet I, Chiozzotto R, Scaglione D, Bassi D, Rossini L, Cirilli M (2022). Genetic dissection of fruit maturity date in apricot (*P. armeniaca L.*) through a Single Primer Enrichment Technology (SPET) approach. *BMC Genomics* **23**: 712.

Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, *et al.* (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PloS One* **3**: e3376.

Barchi L, Acquadro A, Alonso D, Aprea G, Bassolino L, Demurtas O, *et al.* (2019). Single Primer Enrichment Technology (SPET) for high-throughput genotyping in tomato and eggplant germplasm. *Frontiers in Plant Science* **10**: 470381.

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

Bolger AM, Lohse M, Usadel B (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120.

Boudot J-P (2020). *Coenagrion mercuriale*. The IUCN Red List of Threatened Species. e.T5081A83841084. https://dx.doi.org/10.2305/IUCN.UK.2020-3.RLTS.T5081A83841084.en

Bybee S, Córdoba-Aguilar A, Duryea MC, Futahashi R, Hansson B, Lorenzo-Carballa MO, *et al.* (2016). Odonata (dragonflies and damselflies) as a bridge between ecology and evolutionary genomics. *Frontiers in Zoology* **13**: 1–20.

Campbell EO, Brunet BMT, Dupuis JR, Sperling FAH (2018). Would an RRS by any other name sound as RAD ? *Methods in Ecology and Evolution* **9**: 1920–1927.

Campbell NR, Harmon SA, Narum SR (2015). Genotyping-in-Thousands by sequencing (GT-seq): A cost effective SNP genotyping method based on custom amplicon sequencing. *Molecular Ecology Resources* **15**: 855–867.

Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013). Stacks: an analysis tool set for population genomics. *Molecular Ecology* **22**: 3124–3140.

Cattell RB (1966). The scree test for the number of factors. *Multivariate Behavioral Research* **1**: 245–276.

Chen S, Zhou Y, Chen Y, Gu J (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**: i884–i890.

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, *et al.* (2011). The variant call format and VCFtools. *Bioinformatics* **27**: 2156–2158.

Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, *et al.* (2021). Twelve years of SAMtools and BCFtools. *GigaScience* **10**: giab008.

Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, Blaxter ML (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics* **12**: 499–510.

DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, *et al.* (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* **43**: 491–498.

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, *et al.* (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PloS One* **6**: e19379.

Fierimonte B, Vanappelghem C (2021). Plan régional d'actions Hauts-de-France en faveur des libellules menacées 2022-2031 - Agir pour la préservation des odonates menacés et de leurs habitats. Conservatoire d'espaces naturels Hauts-de-France et DREAL Hauts-de-France : 80 p.

Frankham R, Briscoe DA, Ballou JD (2013). *Introduction to conservation genetics*, 2d edition. Cambridge University Press.

Gavriliuc S, Reza S, Jeong C, Getachew F, McLoughlin PD, Poissant J (2022). Targeted genome-wide SNP genotyping in feral horses using non-invasive fecal swabs. *Conservation Genetics Resources* **14**: 203–213.

Goudet J (2005). Hierfstat, a package for R to compute and test hierarchical *F*-statistics. *Molecular Ecology Notes* **5**: 184–186.

Gramazio P, Jaén-Molina R, Vilanova S, Prohens J, Marrero Á, Caujapé-Castells J, *et al.* (2020). Fostering conservation via an integrated use of conventional approaches and high-throughput SPET genotyping: a case study

using the endangered canarian endemics *Solanum lidii* and *S. vespertilio* (Solanaceae). *Frontiers in Plant Science* **11**: 543691.

Grand D (1996). *Coenagrion mercuriale (Charpentier, 1840). Background information on invertebrates of the Habitats Directive and the Bern Convention, Part II: Mantodea, Odonata, Orthoptera, and Arachnida.*, Council of Europe Publishing (P van Helsdingen, L Willemse, and M Speight, Eds). France (Strasbourg, France).

Gruber B, Unmack PJ, Berry OF, Georges A (2018). dartr: An R package to facilitate analysis of SNP data generated from reduced representation genome sequencing. *Molecular Ecology Resources* **18**: 691–699.

Harrisson KA, Pavlova A, Telonis-Scott M, Sunnucks P (2014). Using genomics to characterise evolutionary potential for conservation of wild populations. *Evolutionary Applications* **7**: 1008–1025.

Hedrick PW (2011). *Genetics of Populations*, 4th Edition. Sudbury, MA: Jones and Bartlett Publishers, Inc.

Hohenlohe PA, Funk WC, Rajora OP (2021). Population genomics for wildlife conservation and management. *Molecular Ecology* **30**: 62–82.

Holderegger R, Balkenhol N, Bolliger J, Engler JO, Gugerli F, Hochkirch A, *et al.* (2019). Conservation genetics: linking science with practice. *Molecular Ecology* **28**: 3848–3856.

Houard X (2020). Plan national d'actions en faveur des « libellules » - Agir pour la préservation des odonates menacés et de leurs habitats 2020-2030. *Office pour les insectes et leur environnement – DREAL Hauts-de-France - Ministère de la transition écologique*: 66 p.

Ioannidis P, Simao FA, Waterhouse RM, Manni M, Seppey M, Robertson HM, *et al.* (2017). Genomic features of the damselfly *Calopteryx splendens* representing a sister clade to most insect orders. *Genome Biology and Evolution* **9**: 415-430.

Jombart T (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**: 1403–1405.

Jombart T, Ahmed I (2011). adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics* **27**: 3070–3071.

Jombart T, Devillard S, Balloux F (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics* **11**: 94.

Jones MR, Good JM (2016). Targeted capture in evolutionary and ecological genomics. *Molecular Ecology* **25**: 185–202.

Kamvar ZN, Brooks JC, Grünwald NJ (2015). Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. *Frontiers in Genetics* **6**: 151034.

Karunarathne P, Zhou Q, Schliep K, Milesi P (2023). A comprehensive framework for detecting copy number variants from single nucleotide polymorphism data: 'rCNV', a versatile r package for paralogue and CNV detection. *Molecular Ecology Resources* **23**: 1772–1789.

Keller D, Brodbeck S, Flöss I, Vonwil G, Holderegger R (2010). Ecological and genetic measurements of dispersal in a threatened dragonfly. *Biological Conservation* **143**: 2658–2663.

Keller D, Van Strien MJ, Holderegger R (2012). Do landscape barriers affect functional connectivity of populations of an endangered damselfly? *Freshwater Biology* **57**: 1373–1384.

Knaus BJ, Grünwald NJ (2017). vcfr: a package to manipulate and visualize variant call format data in R. *Molecular Ecology Resources* **17**: 44–53.

Langmead B, Salzberg SL (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**: 357–359.

Lévêque A, Duputié A, Vignon V, Duez F, Godé C, Vanappelghem C, Arnaud J-F (2024). Contrasting patterns of spatial genetic structure in endangered southern damselfly (*Coenagrion mercuriale*) populations facing habitat fragmentation and urbanisation. bioRxiv 2024.05.02.592171; doi: https://doi.org/10.1101/2024.05.02.592171

Li H (2014). Toward better understanding of artifacts in variant calling from high-coverage samples. *Bioinformatics* **30**: 2843–2851.

Lorenzo-Carballa MO, Ferreira S, Sims AM, Thompson DJ, Watts PC, Cher Y, *et al.* (2015). Impact of landscape on spatial genetic structure and diversity of *Coenagrion mercuriale* (Zygoptera:Coenagrionidae) in northern France. *Freshwater Science* **34**: 1065–1078.

Lou RN, Therkildsen NO (2022). Batch effects in population genomic studies with low-coverage whole genome sequencing data: Causes, detection and mitigation. *Molecular Ecology Resources* **22**: 1678–1692.

Luu K, Bazin E, Blum MGB (2017). *pcadapt* : an R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources* **17**: 67–77.

Mamanova L, Coffey AJ, Scott CE, Kozarewa I, Turner EH, Kumar A, *et al.* (2010). Target-enrichment strategies for next-generation sequencing. *Nature Methods* **7**: 111–118.

Meek MH, Larson WA (2019). The future is now: Amplicon sequencing and sequence capture usher in the conservation genomics era. *Molecular Ecology Resources* **19**: 795–803.

Mijangos JL, Gruber B, Berry O, Pacioni C, Georges A (2022). dartR v2: An accessible genetic analysis platform for conservation, ecology and agriculture. *Methods in Ecology and Evolution* **13**: 2150–2158.

Nairismägi M-L, Tan J, Lim JQ, Nagarajan S, Ng CCY, Rajasegaran V, *et al.* (2016). JAK-STAT and G-protein-coupled receptor signaling pathways are frequently altered in epitheliotropic intestinal T-cell lymphoma. *Leukemia* **30**: 1311–1319.

Narum S, Buerkle A, Davey JW, Miller MR, Hohenlohe PA (2013). Genotyping-by-sequencing in ecological and conservation genomics. *Molecular Ecology* **22**: 2841–2847.

O'Leary SJ, Puritz JB, Willis SC, Hollenbeck CM, Portnoy DS (2018). These aren't the loci you'e looking for: Principles of effective SNP filtering for molecular ecologists. *Molecular Ecology* **27**: 3193–3206.

Paris JR, Stevens JR, Catchen JM (2017). Lost in parameter space: a road map for STACKS. *Methods in Ecology and Evolution* **8**: 1360–1373.

Pembleton LW, Cogan NOI, Forster JW (2013). StAMPP: an R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Molecular Ecology Resources* **13**: 946–952.

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012). Double Digest RADseq: an inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. *PloS One* **7**: e37135.

Privé F, Luu K, Vilhjálmsson BJ, Blum MGB (2020). Performing highly efficient genome scans for local adaptation with R package pcadapt version 4. *Molecular Biology and Evolution* **37**: 2153–2154.

R Core Team (2024). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rivera-Colón AG, Catchen J (2022). Population genomics analysis with RAD, reprised: Stacks 2. In: *Marine genomics: methods and protocols*, Humana, New York, USA: Springer, pp. 99–149.

Rochette NC, Rivera-Colón AG, Catchen JM (2019). Stacks 2: analytical methods for paired-end sequencing improve RADseq-based population genomics. *Molecular Ecology* **28**: 4737–4754.

Rouquette JR, Thompson DJ (2007). Patterns of movement and dispersal in an endangered damselfly and the consequences for its management: movement and dispersal in a damselfly. *Journal of Applied Ecology* **44**: 692–701.

Saber A, Hiltermann TJN, Kok K, Terpstra MM, de Lange K, Timens W, *et al.* (2017). Mutation patterns in small cell and non-small cell lung cancer patients suggest a different level of heterogeneity between primary and metastatic tumors. *Carcinogenesis* **38**: 144–151.

Scaglione D, Pinosio S, Marroni F, Di Centa E, Fornasiero A, Magris G, *et al.* (2019). Single primer enrichment technology as a tool for massive genotyping: a benchmark on black poplar and maize. *Annals of Botany* **124**: 543–551.

Scolnick JA, Dimon M, Wang I-C, Huelga SC, Amorese DA (2015). An efficient method for identifying gene fusions by targeted RNA sequencing from fresh frozen and FFPE samples. *PloS One* **10**: e0128916.

Seeb JE, Carvalho G, Hauser L, Roberts S, Seeb W (2011). Single-nucleotide polymorphism (SNP) discovery and applications of SNP genotyping in nonmodel organisms. *Molecular Ecology Resources* **11**.

Shafer ABA, Wolf JBW, Alves PC, Bergström L, Bruford MW, Brännström I, *et al.* (2015). Genomics and the challenging translation into conservation practice. *Trends in Ecology & Evolution* **30**: 78–87.

Song K, Li L, Zhang G (2016). Coverage recommendation for genotyping analysis of highly heterologous species using next-generation sequencing technology. *Scientific Reports* **6**: 35736.

Tripodi P, Beretta M, Peltier D, Kalfas I, Vasilikiotis C, Laidet A, *et al.* (2023). Development and application of Single Primer Enrichment Technology (SPET) SNP assay for population genomics analysis and candidate gene discovery in lettuce. *Frontiers in Plant Science* **14**: 1252777.

Van Strien MJ, Keller D, Holderegger R (2012). A new analytical approach to landscape genetic modelling: least-cost transect analysis and linear mixed models. *Molecular Ecology* **21**: 4010–4023.

Van Tassell CP, Smith TPL, Matukumalli LK, Taylor JF, Schnabel RD, Lawley CT, *et al.* (2008). SNP discovery and allele frequency estimation by deep sequencing of reduced representation libraries. *Nature Methods* **5**: 247–252.

Vanappelghem C, Hubert B (2010). Suivi de la population de *Coenagrion mercuriale* (Charpentier, 1840) dans la Réserve naturelle régionale des dunes et hauts de Dannes-Camiers (Pas-de-Calais) (Odonata, Zygoptera : Coenagrionidae). Actes des Rencontres odonatologiques 2010, Saint-Laurent (Ardennes), 26, 27 et 28 juin. *Martinia*, 26 (3-4) : 131-137

Verdu CF, Guichoux E, Quevauvillers S, De Thier O, Laizet Y, Delcamp A, *et al.* (2016). Dealing with paralogy in RADseq data: in silico detection and single nucleotide polymorphism validation in *Robinia pseudoacacia* L. *Ecology and Evolution* **6**: 7323–7333.

Waples RS, Waples RK, Ward EJ (2022). Pseudoreplication in genomic-scale data sets. *Molecular Ecology Resources* **22**: 503–518.

Watts PC, Kemp SJ, Saccheri IJ, Thompson DJ (2005). Conservation implications of genetic variation between spatially and temporally distinct colonies of the endangered damselfly *Coenagrion mercuriale*. *Ecological Entomology* **30**: 541–547.

Watts PC, Rouquette JR, Saccheri IJ, Kemp SJ, Thompson DJ (2004c). Molecular and ecological evidence for small-scale isolation by distance in an endangered damselfly, *Coenagrion mercuriale*. *Molecular Ecology* **13**: 2931–2945.

Watts PC, Saccheri IJ, Kemp SJ, Thompson DJ (2006). Population structure and the impact of regional and local habitat isolation upon levels of genetic diversity of the endangered damselfly *Coenagrion mercuriale* (Odonata: Zygoptera). *Freshwater Biology* **51**: 193–205.

Watts PC, Saccheri IJ, Kemp SJ, Thompson DJ (2007). Effective population sizes and migration rates in fragmented populations of an endangered insect (*Coenagrion mercuriale* : Odonata). *Journal of Animal Ecology* **76**: 790–800.

Watts PC, Thompson DJ, Kemp SJ (2004b). Cross-species amplification of microsatellite loci in some European zygopteran species (Odonata: Coenagrionidae). *International Journal of Odonatology* **7**: 87–96.

Watts PC, Wu JH, Westgarth C, Thompson DJ, Kemp SJ (2004a). A panel of microsatellite loci for the southern damselfly, *Coenagrion mercuriale* (Odonata: Coenagrionidae). *Conservation Genetics* **5**: 117–119.

Weir BS, Cockerham CC (1984). Estimating *F*-Statistics for the analysis of population structure. *Evolution* **38**: 1358–1370.

Supporting information

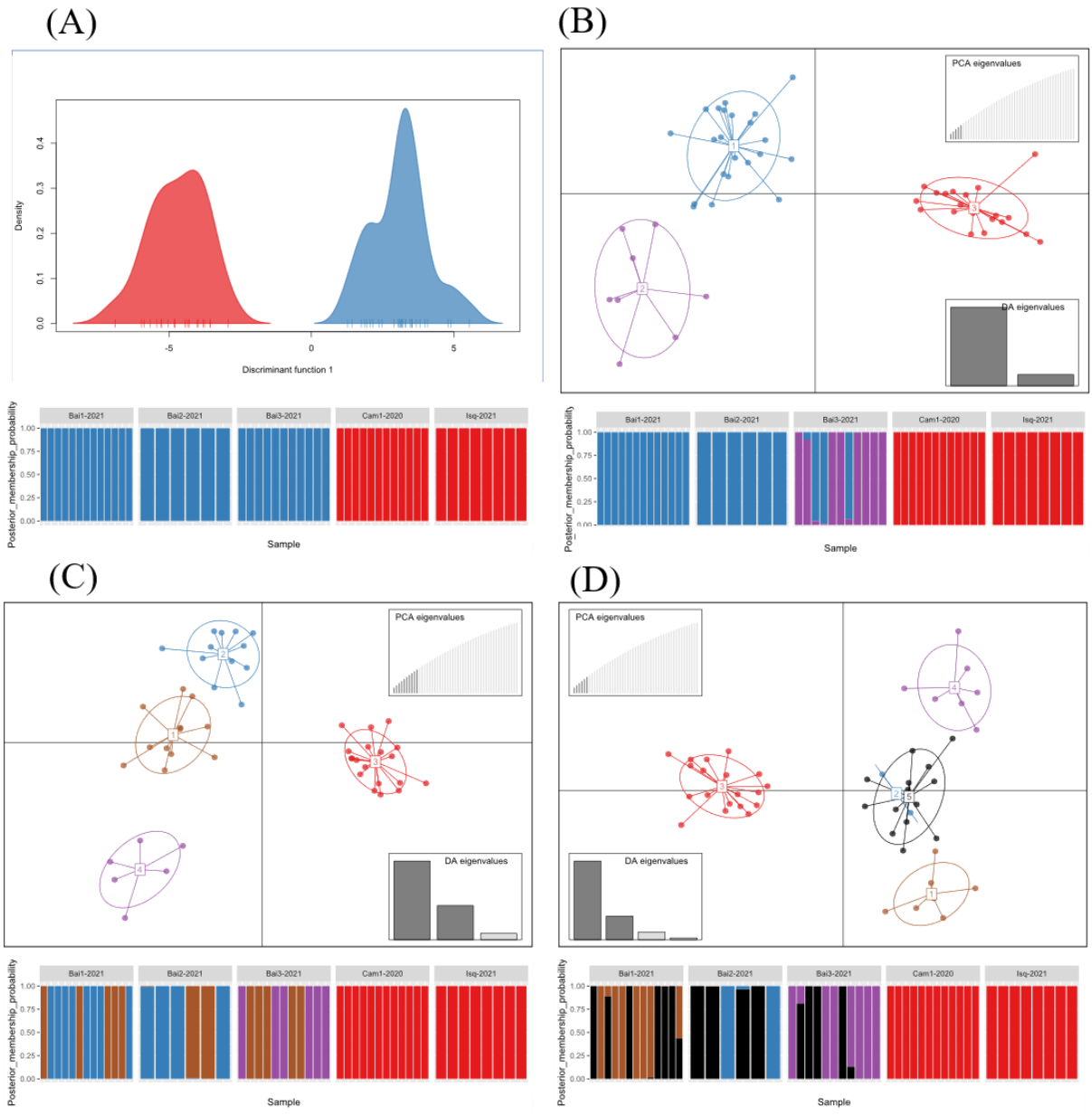


Figure S1: Scatterplots for DAPC for five southern damselfly (*Coenagrion mercuriale*) populations located in northern France with *K*=2 **(A)**, *K*=3 **(B)**, *K*=4 **(C)** and *K*=5 **(D)** with their associated population membership probability assignment.

Supplementary Table S1: List of the samples used in ddRADseq library preparation. Sample_ID: sample name, Pop_ID: name of the population of origin and geographic coordinates of the sampling sites (WGS84 ESPG 4326).

Sample_ID	Pop_ID	Longitude (WGS84)	Latitude (WGS84)
Cam1-13-m20	Cam	1.60616400	50.57378824
Isq-01-m21	Isq	1.66095648	50.66824458
Bai1-01-m21	Bai1	1.81473148	50.55641697
Bai3-01-m21	Bai3	1.78482142	50.53136707
Cuy-01-m21	Cuy	3.45869472	50.44209135
Ger1-05-m21	Ger1	7.71086409	48.39779464
Rut2-01-m21	Rut2	7.59587037	48.49964518
Ceh1-09-m21	Ceh1	7.58164061	48.47838244
S6-20-m21	S6	7.61528337	48.62932940
Rsc1-02-m21	Rsc1	7.73923393	48.48230098
A11-17-m21	A11	7.65574505	48.57464250
S21-06-m21	S21	7.83357726	48.63317546
Rla6-16-m21	Rla6	7.70094941	48.67059395
A5-08-m21	A5	7.56451122	48.54390458
V1-07-m21	V1	7.60749656	48.19228039
V2-15-m21	V2	7.67439857	48.27871207
RLK2-18-m21	RLK2	7.65126778	48.65764007
S17-10-m21	S17	7.74115324	48.64159254
V3-11-m21	V3	7.73287300	48.35805528
Ger6-14-m21	Ger6	7.67472936	48.32892893

Supplementary Table S2: Details of the 12,000 SPET probes used, with their sequence (probe sequence), orientation (Strand) and the position on the reference RADlocus. The data table can be downloaded from the following link: https://figshare.com/s/a5aa49343eb1d9cd8093

Supplementary Table S3: i5 index, DNA pool and yields of the Allegro Target Genotyping library preparation.

i5 index	Barcode Plate	Pool	Pool concentration (ng/µL)	Number of raw reads with this i5 index
MP2-BC01	Plate 1	1	20.0	286 353 103
		2	17.2	
	Plate 2	3	14.2	
		4	17.0	
MP2-BC02	Plate 1	5	15.9	265 877 177
		6	13.9	
	Plate 2	7	15.4	
		8	12.5	
MP2-BC03	Plate 1	9	14.3	309 775 843
		10	16.2	
	Plate 2	11	20.6	
		12	20.2	
MP2-BC04	Plate 1	13	13.2	207 371 842
		14	9.72	
	Plate 2	15	12.2	
		16	9.42	
MP2-BC05	Plate 1	17	18.9	277 622 776
		18	17.9	
	Plate 2	19	19.1	
		20	24.0	
MP2-BC06	Plate 1	21	22.8	313 291 482
		22	18.4	
	Plate 2	23	25.8	
		24	22.6	
MP2-BC07	Plate 1	25	7.0	23 238 462
		26	6.1	
	Plate 2	27	7.3	
		28	5.7	
MP2-BC08	Plate 1	29	7.5	31 631 178
		30	6.2	
	Plate 2	31	7.7	
		32	6.4	
MP2-BC09	Plate 1	33	5.2	26 460 956
		34	4.6	
	Plate 2	35	6.0	
		36	5.9	
MP2-BC10	Plate 1	37	6.8	42 484 721
		38	6.4	
	Plate 2	39	5.3	
		40	5.0	

Supplementary Table S4: Information on southern damselflies samples used for the ATG. Sample ID, Year of sampling, Sampling site name and coordinates (WGS84), the metaplex used to tag the sample, the total and filtered number of sequences retrieved, the number of couples of probes completely (+ and - probe), partly (+ or - probe) or not at all found in the cleaned reads. Samples corresponding to metaplexes MP2-BC07, MP2-BC08, MP2-BC09 and MP2-BC10 which had to be withdrawn due to low numbers of reads, are shown in red. Northern France samples used for population genetic analyses in the present paper appear in bold face. The data table can be downloaded from the following link: https://figshare.com/s/a3514773186a615bbaed

Supplementary Table S5.1: Final steps for filtering the SNP matrix, showing the number of individuals and loci before and after each step. The final numbers of loci and individuals used for analysis are indicated in bold.

Filter	Before	After
1. Remove individuals tagged with i5 indexes MP2-BC07,08,09,10.	1,920 individuals	1,150 individuals
2. Remove poor quality SNP (QD < 2, QUAL < 30, MQRankSum < -12.5, MQ < 30)	254,174 SNPs	174,822 SNPs
3. Select only biallelic loci	174,822 SNPs	154,324 SNPs
4. Select loci out of the 40bp corresponding to ATG probe sequences	154,324 SNPs	147,508 SNPs
5. Set genotypes with depth <8x as missing data	147,508 SNPs	147,508 SNPs
6. Discard loci genotyped in <50% of individuals	147,508 SNPs	58,053 SNPs
7. Discard loci with minor allele frequency < 1%	58,053 SNPs	10,435 SNPs
8. Select one loci per RADlocus (the one with the least missing data)	10,435 SNPs	2,767 SNPs
9. Discard individuals with >40% missing loci	1150 individuals	**1080 individuals**
10. Remove loci with a max depth of coverage twice the mean coverage	2,767 SNPs	2,541 SNPs
11. Filter out loci with F_{IS} levels out of [-0.2;0.2]	2,541 SNPs	2,100 SNPs
12. Filter putatively adaptative loci identified with pcadapt	2,100 SNPs	**2,092 SNPs**

We explored the effect of all these filtering parameters:

TableS5.2 Loci sequencing depth
TableS5.3 Loci missing data amount
TableS5.4 Minor Allele Frequency (MAF)
TableS5.5 Individual missing data amount

Supplementary Table S5.2: Effect of changing the filter for coverage depth of loci on the number of remaining RADloci (step 5 in Supplementary Table 5.1)

Firstly, we filtered our loci according to their sequencing depth (step 5 in table S5.1). We had two filtering options:

a) filter the loci according to their average depth of coverage (average per locus over all the individuals) with the risk of keeping loci with a fair average coverage but a low depth of coverage for some individuals.

b) Filter by locus and by individuals, by setting a threshold value above which we considered the genotype to be reliable, and transforming unreliable genotypes into missing data, with the risk of increasing the quantity of missing data in our dataset.

We elicited to choose the second option with a minimum coverage depth of 8X to consider genotypes as reliable.

	Depth	**Number of biallelic SNP**	**Number RADcontigs containing biallelic SNP**	**Number of biallelic SNP located after position 40**	**Number RADcontigs containing biallelic SNP after position 40**
Average depth of coverage filter	20X	7002	830	6822	823
	15X	14026	1468	13743	1466
	10X	37489	2556	36451	2554
	8X	60442	3208	58242	3205
Unreliable genotypes set as missing data	8X	154324	5556	147508	**5550**

Supplementary Table S5.3: Effect of the filter on the amount of missing data per locus on the remaining number of RADcontigs (step 6 in Supplementary Table S5.1). We chose the last option.

	Depth Filter	Loci present in x% of individuals	Number of biallelic SNP	nb RADcontigs containing biallelic SNP	Number of biallelic SNP after 40bp	Number RADcontigs containing biallelic SNP after 40bp
Average depth of coverage	**8X**	**90%**	58950	3185	56797	3182
		80%	59697	3204	57528	3201
		70%	59870	3206	57694	3203
		60%	59942	3206	57764	3203
		50%	59988	3206	57807	3203
Unreliable genotypes set as missing data	**8X**	**90%**	12514	1419	12267	1415
		80%	21362	1943	20917	1940
		70%	32920	2415	32092	2411
		60%	45910	2836	44442	2834
		50%	60265	3198	**58053**	**3195**

Supplementary Table S5.4: Effect of the filter on Minor Allele Frequency (MAF) on the number of RADcontigs (step 7 in Supplementary Table S5.1). We chose the last option.

Depth Filter	Locus present in x% of individual	Minor Allele Frequency	Number of biallelic SNP	Number RADcontigs containing biallelic SNP	Number of biallelic SNP after 40bp	Number RADcontigs containing biallelic SNP after 40bp
8X	**50%**	**<0.1**	6560	2203	6465	2199
		<0.05	7904	2451	7793	2444
		<0.01	10605	2773	**10435**	**2767**

Supplementary Table S5.5: Effect of the filter on the amount of missing data in each individual on the number of RADcontigs (step 9 in Supplementary Table S5.1). We chose the third option.

Filter individuals with less than X% SNP genotyped	Individuals kept after filtration.
20%	1149
50%	1124
60%	**1080**
70%	965

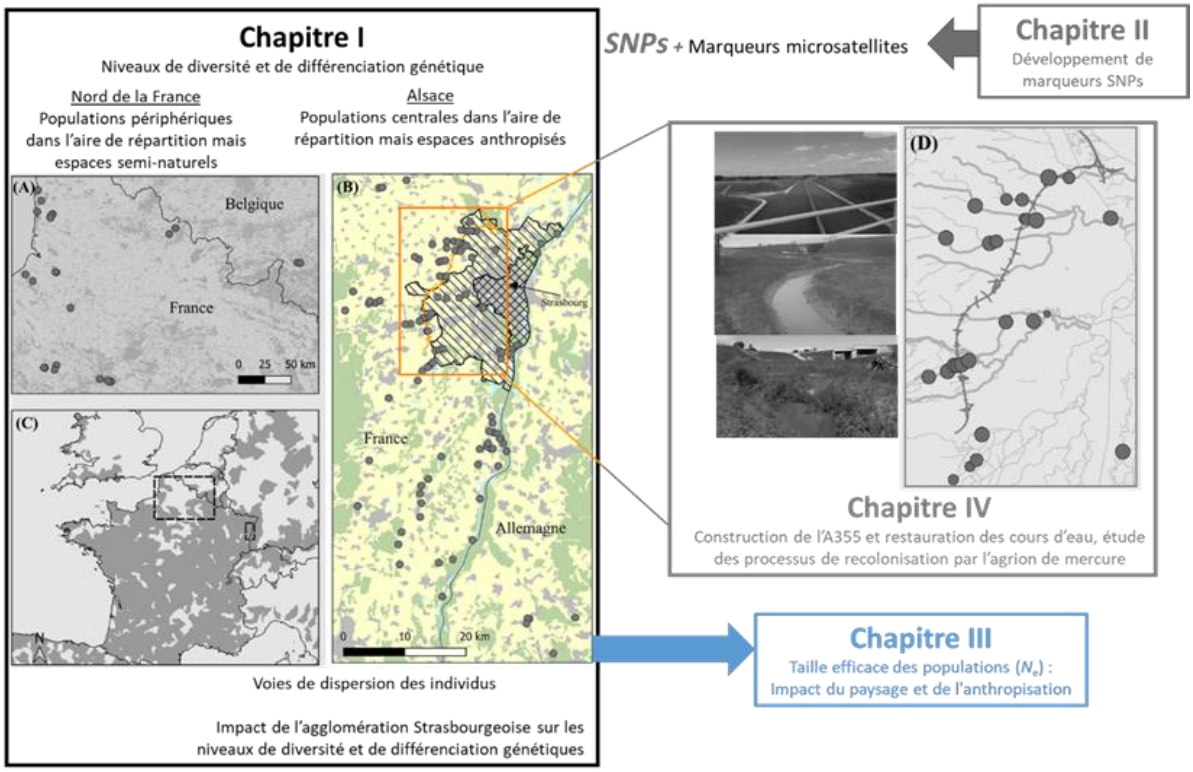
Table S6: Sampling information and mean multilocus genetic diversity estimates using 10 microsatellite loci (Lévêque et al *biorxiv*) in five Southern damselfly populations. Abbreviations: ID: population name. Longitude and latitude are provided in WGS84 coordinates. *N*: sample size. H_o: observed heterozygosity. H_e: expected heterozygosity. F_{IS}: the intrapopulation fixation index; A_r allelic richness.

ID	Sampling Year	Longitude (E)	Latitude (N)	*N*	Allele number	H_o	H_e	F_{IS}	A_r
Bai1	2021	1.8147	50.5564	25	37	0.544	0.49	-0.110	3.071
Bai2	2021	1.7959	50.5472	11	30	0.518	0.484	-0.070	2.866
Bai3	2021	1.7848	50.5314	14	28	0.529	0.44	-0.202	2.707
Cam1	2020	1.6061	50.5738	21	38	0.51	0.489	-0.043	3.229
Isq	2020	1.6611	50.6682	20	31	0.485	0.461	-0.052	2.820

Table S7: Pairwise F_{ST} estimates using 10 microsatellite loci (Lévêque et al *biorxiv*) between five southern damselfly populations located in Northern France, associated with their 95% confidence interval, and their statistical significance levels. * $P \leq 0.05$, ** $P \leq 0.01$ *** $P \leq 0.001$.

	Bai1	Bai2	Bai3	Cam1
Bai2	0.054 (0.00 - 0.1)*			
Bai3	0.117 (0.015 - 0.2)*	0.175 (0.03 - 0.371)**		
Cam1	0.017 (-0.015 - 0.063)	0.111 (-0.015 - 0.283)	0.04 (-0.013 - 0.098)	
Isq	0.051 (0.007 - 0.117)*	0.137 (0.014 - 0.318)**	0.178 (0.075 - 0.261)***	0.082 (0.01 - 0.159)***

Chapitre 3. Estimation des tailles efficaces des populations d'agrion de Mercure en Alsace : comparaison de marqueurs moléculaires et de différents estimateurs



La fragmentation des habitats et les changements d'utilisation des terres créent souvent des environnements défavorables, réduisant ainsi la connectivité entre les populations. Cela peut entraîner une augmentation de la différenciation génétique des populations, tout en diminuant leur diversité génétique, augmentant ainsi le niveau de dérive génétique et de consanguinité locale. À cet égard, la taille efficace de la population locale (N_e) est un paramètre clé en biologie de la conservation, car il mesure la force de la dérive génétique et de la consanguinité. Pourtant, il reste difficile d'estimer N_e de façon fiable, que ce soit par des approches directes, impliquant des suivis démographiques, ou des approches indirectes, fondées sur l'utilisation de marqueurs moléculaires.

Ce chapitre visait donc à caractériser finement la taille efficace des populations d'agrion de Mercure situées en Alsace. Nous avons génotypé 2 842 individus appartenant à 77 populations à l'aide de marqueurs microsatellites, et 958 d'entre eux pour 2 100 marqueurs SNP développés dans le chapitre II. Nous avons ensuite estimé les niveaux de taille efficace de populations en utilisant ces deux jeux de

données et en appliquant des estimateurs fondés sur un simple échantillonnage ou sur des rééchantillonnages espacés dans le temps.

Nous montrons ainsi que la structure génétique spatiale est stable dans le temps : les cohortes émergeant chaque année semblent connectées par des flux de gènes. Les estimateurs de N_e de même nature et appliqués à un même jeu de données produisent des estimations similaires. En revanche, les estimations fondées sur un seul ou plusieurs échantillonnages temporels ne sont pas toujours congruentes, notamment entre marqueurs génétiques. Les tailles efficaces de population estimées semblent souvent peu élevées à proximité de l'Eurométropole de Strasbourg. Enfin, nous n'avons pu identifier aucun effet des caractéristiques paysagères entourant les sites sur les niveaux de taille efficace de population. Ceci suggère que les patrons de structuration génétique sont gouvernés par des éléments paysagers plus fins que ceux auxquels nous avions accès.

Ce chapitre, présenté sous la forme d'un article, a été soumis début mai 2024 sur invitation dans un numéro spécial de la revue *Evolutionary Applications*, ce numéro spécial étant dédié à l'étude des tailles efficaces de populations.

Levels and spatial patterns of effective population sizes in the southern damselfly (*Coenagrion mercuriale*): on the need to carefully interpret single-point and temporal estimations to set conservation guidelines.

Agathe Lévêque [1,2], Anne Duputié [1], Vincent Vignon [2,3], Fabien Duez [1], Cécile Godé [1], Clément Mazoyer [1], Jean-François Arnaud [1]

1. Univ. Lille, CNRS, UMR 8198 - Evo-Eco-Paleo, F-59000 Lille, France

2. Office de Génie Écologique (O.G.E.), F-67200 Strasbourg, France

3. ALKIOS, F-84911 Avignon, France

Short running title: N_e in the southern damselfly

Abstract

The effective population size (N_e) is a key population parameter in conservation and evolutionary biology as it determines the strength of genetic drift and inbreeding. Field estimations of N_e can be logistically difficult and time-consuming as opposed to genetically-based methods whose rise was fuelled by increasing data availability. Nonetheless, accurately estimating N_e remains challenging, and few studies addressed the joint comparison of N_e estimates using different kinds of molecular markers in concert with single-point sample N_e assessments based on linkage disequilibrium or sibship and temporal estimations of N_e based on temporal variance in allele frequencies. This study aimed at bridging this gap by analysing singly-sampled and temporally-spaced populations in the southern damselfly (*Coenagrion mercuriale*), a bioindicator Odonata species of high conservation concern which occurs in networks of freshwater streams in western Europe. 77 populations were collected in a semi-urbanised area located in eastern France, totalising 2,842 individuals genotyped using microsatellite loci, 958 of which were also typed for 2,092 SNPs. Spatial genetic structure was stable over time, suggesting porosity between alternate-year cohorts. Applying single-sample and temporal estimations of N_e yielded consistent results between estimators when compared for a same set of molecular markers. Nonetheless, inconsistent results were observed in N_e depiction among single-sample and temporal estimators and between microsatellite and SNPs, with downwardly biased N_e levels using the sibship estimator. Geographical patterns of N_e further suggested a negative effect of urban areas, and presumably source/sink population dynamics of recolonisation for populations separated by a newly built highway. Nonetheless, we found no significant effect of the surrounding landscape on levels of N_e, suggesting more subtle fine-scaled microhabitat effects. Altogether, our study highlighted the challenge of obtaining robust and concordant N_e estimates and to carefully interpret discordant N_e estimates to set relevant conservation guidelines.

Keywords: effective population size, damselfly, spatial and temporal genetic structure, urbanisation, landscape ecology, conservation biology

Introduction

Changes in land use related to human activities can lead to habitat loss and fragmentation, reducing the extent of gene flow among populations and decreasing their sizes (Fahrig, 2003; Wilson *et al.*, 2016). Because of the greater magnitude of genetic drift and the cost of inbreeding, small isolated populations face a substantial risk of decreasing genetic variability over time (Frankham *et al.*, 2013; Ellegren and Galtier, 2016; Allendorf *et al*,. 2022). Combined with stochastic environmental processes, the deleterious genetic effects associated with the loss of genetic diversity can lead to a reduction in the adaptive potential of populations, which ultimately can drive to a vortex of extinction (Saccheri *et al.*, 1998; Soulé and Mills, 1998; Frankham *et al.*, 2013). Consequently, the measurement of contemporary genetic features of populations is a key tool in monitoring the vulnerability of a population to detrimental genetic changes, and in assessing their adaptative potential and long-term persistence, which is particularly important for threatened species (Allendorf *et al.*, 2010, 2022; Holderegger *et al.*, 2019; Hohenlohe *et al.*, 2021).

The effective population size (N_e) is one of the key concepts and parameters in evolutionary biology and conservation genetics, as it is directly linked to populations' rate of genetic variation loss due to the influence of increased genetic drift and inbreeding (Nei and Tajima, 1981; Waples, 2005; Hare *et al.*, 2011; Wang *et al.*, 2016). This concept is linked to the direct relationship between the rate of genetic changes and the size of an "ideal" Wright-Fisher finite population that satisfies the assumptions of random mating, constant population size and non-overlapping generations (Fisher, 1931; Wright, 1931). However, wild populations rarely conform to this model, because various factors can reduce the number of individuals contributing to the next generation, including for instance isolated genetic cohorts, unequal sex ratios and differential fitness of males and females. This explains why population sizes in the demographic sense (census size) are often greater than population sizes in the genetic sense (effective size, Waples, 2016). The N_e parameter also provides a better understanding of population dynamics in a given region, for example by identifying source-sink population structure, which can provide important information on the status of populations for their conservation (Frankham *et al.*, 2013; Allendorf *et al*., 2022).

In this respect, numerous methods have been developed to estimate the effective population size in natural populations. The development of genetic methods for estimating N_e is an active field of research involving increasingly applied advances in genotyping and software dedicated to accurately estimate this crucial population parameter (Wang, 2009; Luikart *et al.*, 2010; Wang *et al.*, 2016; Waples, 2024). Herein we will focus on contemporary effective population size estimation, i.e. involving current generations or just a few generations in the past, as this time scale is the most relevant for conservation genetics purposes (Luikart *et al.*, 2010; Hare *et al.*, 2011). Single-sample estimates assess the N_e in the generation preceding the sample by measuring the genetic result of processes acting in the parental generation from which the sample is drawn (Waples, 2005; Hare *et al.*, 2011). They can be measured from genetic patterns of heterozygosity (Pudovkin *et al.*, 1996), linkage disequilibrium (Hill, 1981) or individual relatedness (Wang, 2009). These methods require just a single sample of multilocus genotypes, making them useful because one does not need to survey several generations. In contrast, the temporal method estimates the harmonic mean of N_e based on the measurement of the temporal variation of allele frequencies across two or more samples separated in time (Nei and Tajima, 1981; Pollak, 1983; Waples, 1989; Jorde and Ryman, 2007; Hare *et al.*, 2011). Compared with other N_e estimation approaches, the temporal approach makes fewer assumptions and is more robust to some complications in real natural populations, such as population structure or overlapping generations, but is most accurate when the sampling points are separated by many generations (Wang *et al.*, 2016).

Yet, a robust estimation of the effective population size and its practical application remains a challenge, particularly for large populations integrated into a metapopulation system with gene flow arising among neighbouring populations, requiring spatial and temporal considerations to be taken into account in the estimation of this key population parameter (Waples, 2005, 2024; Hare *et al.*, 2011; Clarke *et al.*, 2024). In addition, each method for estimating N_e is subject to numerous assumptions which are often violated in natural populations (Waples, 2016). However, to some extent, the precision and accuracy of large N_e estimates can be improved by increasing sampling efforts for both individuals and loci (Waples, 1989; Wang, 2009; Luikart *et al.*, 2010).

In this respect, the development of new sequencing technologies (next-generation sequencing, NGS) offers the possibility of genotyping thousands of genomic markers (see Bernatchez *et al.*, 2024).

SNPs provide a better genome-wide coverage and the opportunity to genotype many loci, offering the promise of more accurate estimates of genetic and population demographic parameters important for natural population conservation (Hollenbeck *et al.*, 2016; Allendorf, 2017; Lehnert *et al.*, 2019; see however Waples *et al.*, 2022). In line with NGS, numerous advances in theory and computational analysis have also considerably improved the estimation of N_e (Do *et al.*, 2014; Waples *et al.*, 2014, 2018; Hollenbeck *et al.*, 2016; Waples and Lindley, 2018; Zhou *et al.*, 2018; Waples, 2024). Population genomics approaches have already been used to estimate N_e (e.g. Browning and Browning, 2015; Sovic *et al.*, 2019). Nonetheless, very few studies compared the relative power and biological relevance of single-sample *versus* temporal estimates of N_e using different kinds of molecular markers in natural populations (see however Morin *et al.*, 2012; Sovic *et al.*, 2019). Our study is thus devoted to bridge this gap by analysing the N_e of a set of populations of the southern damselfly (*Coenagrion mercuriale*), a protected species, using both single-sampled populations and temporally-spaced populations, and using two different types of molecular markers: classical microsatellite loci and newly developed SNPs.

Therefore, this study aimed at finely characterising the effective population size of a large number of southern damselfly populations located in eastern France. The southern damselfly (*Coenagrion mercuriale*) is an Odonata species that typically inhabits clear, small streams characterised by the occurrence of helophytes. This species is commonly found in southwestern Europe and sparsely encountered in central Europe (Figure S1A). This Odonata species is of high conservation concern because of population declines linked to changes in land use practices, particularly at the eastern edge of its range, where the species has gone extinct or is currently close to extinction in seven countries such as Belgium, the Netherlands and Austria (Grand, 1996). In eastern France, which is the focus of this study, flying adults gradually emerge from May to August. Southern damselfly adults live for about two weeks, mate, and females lay their eggs in the hollow stems of helophytes. Larvae hatch and develop as swimming organisms for one, two or even three years before metamorphosing into adults. At its northern limit, in the UK and most central European populations (Sternberg et al., 1999; Purse and Thompson, 2002), the southern damselfly is semi-voltine (i.e. it completes a full generation in two years), whereas in North Africa it is both univoltine and bi-voltine (Mahdjoub et al., 2014, 2015). Given this latitudinal gradient of the species' voltinism, we expected the southern damselfly populations located in eastern

France to exhibit a semi-voltine trait. In France, the southern damselfly is widely distributed and locally abundant except for the northernmost region (Figure S1B). In eastern France (Alsace region), the species is widely distributed, although this region is subject to substantial anthropogenic pressures such as strong urbanisation effects owing to the occurrence of Strasbourg Eurometropolis, intensive cereal cultivation, and grapevines on the slopes of the Vosges mountains (Figure S1C). Several studies have been carried out to estimate the population census size of this species in different regions (Rouquette and Thompson, 2005; Thompson and Watts, 2006; Couvreur *et al.*, 2008; La Porta and Goretti, 2019; Beaune and Sellier, 2021). Nonetheless, only one study was conducted to estimate the N_e in southern damselfly using genetic tools, highlighting contrasting patterns between demographic and genetic estimates of population size in a single large southern damselfly population located in the UK (Watts *et al.*, 2007). Demographic estimates of effective population size indeed showed levels of effective size that were overestimated compared with genetic approaches, presumably because of a weak spatial genetic structure and of biased estimations of individual reproductive success variance (Watts *et al.*, 2007).

Besides, the southern damselfly life-history traits provide a unique opportunity and an interesting playground for studying the effective population size. Indeed, any factor that creates variation in developmental timing can divide a population into discrete cohorts, characterised by their own dynamics, eventually leading to independent demographic and evolutionary trajectories (Wolf and Zwick, 1989; Battisti et al., 2000; Gradish et al., 2019; Bouaouina et al., 2023). By contrast, the possibility of 'temporal migration', in which larvae delay metamorphosis for a year and recruit to another cohort, can homogenise different cohorts that would otherwise be independent of each other, as it was shown by Watts and Thompson (2012) in southern damselfly populations located to the north of the species' range (UK).

Furthermore, the southern damselfly requires specific reproductive and roosting habitats with small, unshaded watercourses favourable to the development of helophytes (Rouquette and Thompson, 2005, 2007a; Purse and Thompson, 2009). In addition, landscape features between populations can represent barriers to the dispersal of individuals, such as wooded areas or towns (Keller *et al.*, 2012; Lorenzo-Carballa *et al.*, 2015; Watts *et al.*, 2004c). However, several studies showed that, apart from land use among populations, changes in the landscape structure around sites influence the neutral genetic

and adaptive features of populations. For instance, several studies have shown positive or negative correlations between different types of land use, such as urban areas, agricultural activities, woodland or industrial extraction sites, and genetic features linked to population evolutionary potential in various taxa (Flavenot *et al.*, 2015; Munshi-South *et al.*, 2016; Collevatti *et al.*, 2020; Crispim *et al.*, 2021; Clarke *et al.*, 2024).

Here, we used two independent genetic datasets, microsatellites and SNPs, to make conservation-relevant inferences about genetic characteristics in a large set of southern damselfly populations located in the Alsace region, eastern France. This area is characterised by a variety of agricultural landscapes and the occurrence of a large city, Strasbourg. Moreover, to go beyond punctual spatial genetic studies, we performed a temporal survey to get further insights into the N_e of interconnected populations of the southern damselfly where we have previously shown very low genetic differentiation and isolation by distance occurring overland using microsatellite loci (Lévêque *et al.*, 2024a). Our specific goals were the following: (i) evaluating whether alternate-year cohorts were genetically distinct with independent evolutionary trajectories or whether gene flow between cohorts ensured a certain degree of genetic cohesion between the sampled years ; (ii) estimating and comparing the relative power of contemporary effective population sizes using microsatellites and SNPs and by applying single-sample and temporal estimations of N_e; (iii) identifying potential biological and demographic processes underlying the level and geographical distribution of estimated N_e, such as source and sink population dynamics and the occurrence of a newly built highway bypassing the city of Strasbourg; (iv) assessing whether the nature of the surrounding landscape may affect effective population sizes in the southern damselfly.

Materials and methods

Study sites and sampled populations

We collected a total of 2,842 southern damselflies from 77 populations in north-eastern France, Alsace region (Figure 1). Adult individuals were collected using an insect net in late spring (May to July) in 2021 and 2022. 34 populations were sampled in both years (Table S1). Populations were defined as a group of randomly sampled adults within a specific geographic location, spanning a 200-meter

transect along watercourses separated by at least 500m. Minimum sampling size was 10 individuals to allow biologically relevant estimation of effective population sizes: 10 to 39 individuals were collected per population (mean sampling size 25.60, s.d. 6.60). Geographic coordinates and sampling sizes of each population can be found in Table S1. Two kinds of samples were collected: whole individual body or only the right middle leg of each individual, the latter method being non-lethal and not impacting the damselfly's survival (Fincke and Hadrys, 2001). All samples were stored in 100% ethanol until DNA extraction.

Genotyping using microsatellite and SNP loci

Individuals were genotyped using two different kinds of molecular markers: (i) one based on ten microsatellite loci previously described by Watts *et al* (2004a,b) and using the full dataset of 2,842 sampled individuals (77 populations, 34 of which being sampled in both years); (ii) the other one using a set of 2,092 SNPs on a subset of 958 individuals whose full body was available to ensure a sufficient quantity of DNA material, and representing 56 populations, 11 of which were sampled in both years (Table S1). Both microsatellites and SNPs were used to estimate contemporary effective population sizes.

For whole individuals, we removed abdomens from the rest of the body to avoid sampling intestinal microbiota, and we crushed all samples using MN Beads Type D (Macherey-Nagel) as described in Lévêque *et al*. (2024a). We extracted total genomic DNA from each whole individual sample using NucleoMag® Tissue Kit (Macherey-Nagel) according to the manufacturer's recommendations. For leg samples, we purified DNA with 12 µL NucleoMag B-beads and 12 µL of pure water and eluted in 50 µL.

We genotyped all samples using ten unlinked nuclear microsatellite loci named LIST002, LIST023, LIST034, LIST035, LIST037, LIST042, LIST062, LIST024, LIST063, and LIST066, isolated and described in Watts *et al* (2004a,b) and following the protocol details in Lévêque *et al.* (2024a).

We also genotyped southern damselfly individuals using SNPs generated using the Tecan Genomics' Allegro Targeted Genotyping method (ATG, Redwood City, United States). All details of the library preparation and bioinformatics analysis are detailed in Lévêque *et al.* (2024b). Briefly, based

on RADloci constructed from a ddRAD library, 12,000 SPET probes were designed to recapture 6,000 unlinked SNPs. Target genomic libraries were prepared following the manufacturer guidelines for preparing the SPET libraries and were sequenced on a MiSeq System (1x150 bp, Illumina Inc., San Diego, CA, USA) at GenoScreen sequencing platform (Lille, France). The following bioinformatics analyses included sequences' cleaning, alignment to the reference (*bowtie2*; Langmead and Salzberg, 2012), variant calling following the software best practices for germline short variant discovery implemented in GATK Best Practices, and filtering to obtain reliable SNPs. After removing loci under selection and loci showing aberrant F_{IS} estimates, we obtained genotypes for 2,092 biallelic SNPs, each located on a different RADcontig (see Lévêque *et al.*, 2024b).

Levels of spatial and temporal genetic differentiation

First, we assessed whether the population genetic structure was stable over the years. We investigated this issue using only the microsatellite loci because of more resampled populations using this kind of molecular marker. We computed pairwise estimates of genetic differentiation (F_{ST}) between all pairs of samples (between cohorts and between spatial locations) for populations sampled in both years. We tested the statistical significance of pairwise F_{ST} using permutation tests (10,000 multi-locus genotype permutations among both temporal samples and spatial locations) implemented in FSTAT v.2.9.4 (Goudet, 2004). To assess whether the spatial genetic structure was stable over cohorts, we regressed pairwise F_{ST} values between equivalent pairs of populations from each cohort and tested for statistical significance using a Mantel statistic *rz* (Smouse *et al.*, 1986) with the *mantel.rtest* function of the R package 'ade4' with 10000 permutations (Dray and Dufour, 2007).

Secondly, we evaluated the importance of the sampling year on genetic differentiation by calculating hierarchical *F*-statistics (Yang, 1998). We constructed a two-level hierarchy, using years of sampling as the outermost level, and the sampling site as the innermost level. We estimated the variance associated with each level using the function *varcomp.glob* in the R package hierfstat v.0.5-11 (Goudet, 2005). We assessed the statistical significance of variance explained by each level using the permutation algorithms with the functions *test.between* (year of sampling level) and *test.within* (population level) implemented in the same package. As above, we only carried out this analysis using the microsatellite

dataset, as it contained more populations sampled over two years compared to the SNP dataset (34 populations and 11 populations respectively). Besides, it allowed a comparison with previous results obtained for this species (see Watts and Thompson 2012).

Estimation of effective population size and depiction of its spatial distribution

We estimated contemporary effective population sizes for both datasets using two single-sample estimators and three temporal (two-samples) estimators. First, we estimated single-sample contemporary N_e using the method based on linkage disequilibrium (LD; Hill, 1981; Waples, 2006; Waples and Do, 2010) and implemented in the program NEESTIMATOR v.2.1 (Do *et al.*, 2014). This analysis was conducted assuming random mating, using 0.05 as a threshold for the lowest allele frequency, and using the jackknife-based confidence intervals. Secondly, we applied a maximum likelihood method to estimate single-sample contemporary N_e. To this end, we used the COLONY software v.2.0.7.0 (Jones and Wang, 2010) by performing a Full-Likelihood analysis method dedicated to dioecious diploid species, with medium-length runs, assuming random mating and polygamy for both males and females and with no sibship prior. All other parameters were set as default. This maximum likelihood method estimates N_e based on inferred sibship frequencies among samples, with associated confidence intervals obtained through bootstrap resampling (Wang, 2009). Thirdly, N_e were estimated for temporally resampled populations using a method based on temporal changes in allele frequencies (see Waples, 1989) and three different estimators for computing the standardised variance in allele frequency F: F_e (Nei and Tajima, 1981); F_k (Pollak, 1983); and F_s (Jorde and Ryman, 2007). For subsequent analyses and whatever the N_e estimator used, populations showing either infinite N_e estimates or effective population sizes estimated over 1,000 were set to an effective population size estimate of 1,000 to compare the different estimators and to investigate the spatial distribution of N_e.

To examine the level of consistency between the different estimators and microsatellite and SNP markers, we first estimated pairwise Pearson correlations between N_e estimations. The correlation matrix was then visualised using the function *corrplot* from the *corrplot* v.0.92 R package (Wei and Simko, 2017). To visualise the distribution of N_e for each type of genetic marker and N_e estimator, we produced a density plot using the R package *ggplot2* v.3.4.1 (Villanueva and Chen, 2019). To gain further insights

into fine-scaled spatial variations of N_e estimates, we interpolated N_e estimates using the simple kriging method using the R package *gstat* v.2.1-0 (Pebesma, 2004). Maps of estimated N_e were produced using the open-source software QGIS v.3.18 (Quantum GIS Development Team).

Effect of landscape surrounding southern damselfly breeding habitats on effective population sizes

To determine the nature of the landscape surrounding the breeding habitats of southern damselfly populations, we created buffers around each sampling site using QGIS with a radius of 200m (Figure S2A). This buffer size was chosen because it corresponds to the typical lifetime dispersal distances of southern damselfly flying adults, with most individuals dispersing at geographic distances less than < 100m, as shown in previous studies using direct measure of dispersal ability of the species (Purse *et al.*, 2003; Watts *et al.*, 2004c). To describe the landscape in the direct vicinity of sampled populations, we used the 2018 vector layers of Urban Atlas, Natura 2000, Riparian zones and Corine Land Cover to produce a land use map of the studied region (Figure S3A). The Urban Atlas layer covers approximately 85% of our study area and 67 of the sampled sites we surveyed. We subsequently used the Natura 2000 layer to cover the rest of the sites (N=10). Riparian zones and Corine Land Cover layers were used to complete the land cover outside the study sites (Figure S3A). Land use classes were synthesised into five main land cover categories thought to be relevant with regard to habitat requirements of the southern damselfly (Thompson *et al.*, 2003; Rouquette and Thompson, 2005, 2007b): impervious areas (buildings, transport networks), agricultural areas (arable land, vineyards), wooded areas, open green areas (parks, meadows) and aquatic areas (wetlands, large watercourses); Figure S3B). Aquatic areas, open green areas and agricultural areas can be considered as suitable areas for dispersal and/or breeding habitats. In contrast, impervious and wooded areas are inhospitable areas for the southern damselfly (Keller *et al.*, 2012; Lorenzo-Carballa *et al.*, 2015; Watts *et al.*, 2004c).

Pearson's correlation coefficients (r) and their statistical significance were calculated and visualised using the R packages *corrplot* v.0.92 and *GGally* v2.1.2 (Wei and Simko, 2017; Schloerke *et al.*, 2022) to determine the relationship between the land cover surrounding sites (categories: percentage

of impervious, agriculture, forest fragments, open green areas and water bodies) and estimates of effective population sizes.

Results

Genetic variability and levels of genetic differentiation between cohorts

None of the pairwise estimates of genetic differentiation between temporal samples of populations differed from 0 after Bonferroni correction (Table S2). Pairwise values of F_{ST} between equivalent pairs of populations from each temporal population sample were significantly correlated, indicating a stable population genetic structure over time (Mantel $rz = 0.525$, $P < 0.0001$; Figure S4). Likewise, the hierarchical *F*-statistics analysis clearly showed significant levels of genetic differentiation among populations within the same year of sampling ($F_{\text{Population/Year}} = 0.165$; $P = 0.01$; Table S3**)**, but not among years ($F_{\text{Year/Total}} = -0.018$; non-significant; Table S3).

Estimates of effective population sizes

Effective population size estimates varied depending on the type of genetic data and the type of analytical method used to generate N_e estimates (Figure 2 and Supplementary Figures S5, S6, S7). Full details of N_e estimations using the different methods are presented in Table S4 and Table S5. Overall, effective population sizes estimated using the same category of genetic markers were consistent.

For single-sample estimates using SNP data, the correlation between LD and Sibship estimators was 0.77 ($P < 0.001$, Figure S7A) and the patterns of N_e distribution were almost similar, except for intermediate N_e estimates ranging between 75 and 500 (Figure 2A). Using microsatellite data, the correlation between LD and Sibship N_e estimators was weaker, with a value of 0.38 and close to the levels of correlation observed between N_e estimators using different genetic markers (values ranging between 0.17 and 0.34; Figure S7A). For single-sample N_e estimates, 31 to 36 populations were estimated as infinite in size or with $N_e > 1000$ using either microsatellites or SNPs (Figure S5 and Table S4). Finite N_e estimates from LD ranged from 11.5 (population Deu5-2022) to over 3000 (population RLK4-2022) for microsatellite data, and from 41.5 (population S21-2021) to over 20000 (population

Deu7-2022) for SNP data. For sibship frequency-based measures of N_e, estimates ranged from 19 (population CCR11-2022) to 44 (population Rlei6-2021) for microsatellite data, and from 52 (population S21-2021) to 420 (population V3-2021) for SNP data. The sibship estimates with microsatellite data thus stand out from all the other estimates, with a much lower estimated effective population size (all centred around 30) compared to the others, and with no infinite estimates of N_e (Figure 2, Figure S5). Given that, the sibship estimator was no longer considered in our analyses on the spatial distribution of effective population sizes.

For temporal (two-sample) estimates of N_e, we observed a very high correlation between N_e estimates using microsatellite data (correlation values very close to 1, $P < 0.001$, Figure S7B) and a still positive but weaker correlation for data estimated using SNPs, ranging from 0.44 to 0.50 (Figure S7B). In contrast, N_e estimates between microsatellites and SNPs were not or slightly negatively correlated with each other (Figure S7B). For temporal N_e estimates, a high number of populations were estimated as infinite in size, ranging from 14 to 17 populations out of 34, depending on estimators. Using SNPs, these proportions were lower for the Pollak (1983) and Jorde & Ryman (2007) N_e estimators (0 and 2 out of 11 populations, respectively) but not for the Nei & Tajima (1981) N_e estimator for which 6 out of 11 populations exhibited infinite effective population sizes. Finite N_e estimates from Pollak (1983) ranged from 12 to over 1000 using the microsatellite dataset, and from 43 to over 4000 for the SNP dataset (Table S5, Figure S6). N_e estimates using Nei & Tajima (1981) estimator ranged between 14 and 315 and between 80 and 1011 for microsatellites and SNPs, respectively. Finally, N_e estimates from Jorde & Ryman (2007) ranged between 10 and 1144 and between 38 and 423 for microsatellites and SNPs, respectively (Table S5, Figure S6).

Geographical patterns of effective population sizes

Based on N_e estimates that were consistent among single-sample and temporal methods over marker types, we observed that populations S21, CCR7, Rsc1 and RLK56 around the Strasbourg metropolitan area displayed relatively low levels of effective population size compared to other populations located further away from the city (Figure 3, Figure 4A, Figure S8 and S9). However, in the

south of Strasbourg city, population Wei1 surprisingly showed high levels of N_e compared with the neighbouring Rsc1 and CCR7 populations (Figure 4A, Figure S8 and S9).

When focusing on the axis of the newly built A355 highway, effective population size estimates were highly contrasted between populations that were geographically close but located upstream and downstream of the same watercourse (Figure 4A). For instance, populations Rla3 and Rla5 showed low levels of estimated effective population size compared to population Rla6 located downstream. Likewise, population RLK3 showed a low level of estimated effective population size compared with population RLK4 located on the other side of the highway. In contrast, populations Rlei6 and Rlei7, although located near the highway, exhibited high N_e values well above those measured for population Rlei45 located upstream. Finally, although located along the same watercourse and whatever the highway side, populations A4, A6, A7 and A9 displayed a patchwork of low or very high levels of effective population sizes (Figures 3 and 4A and Figures S8, S9).

Finally, in the southern part of the study area, populations exhibited either very high or very low N_e estimates, suggesting a patchwork of stable populations and non-equilibrium populations indicating potential source-sink population dynamics (Figure 4B).

Effect of the landscape surrounding breeding habitats on effective population sizes

Sampled populations differed in land use surrounding the breeding habitats, with a large number of sites predominantly characterised by agricultural land (N = 51). A few populations were surrounded mostly by impermeable land (N = 8) or by wooded land (N = 2), or a mixture of land use variables (N = 16; Figure S2B). Overall, agricultural areas were negatively correlated with other types of land use (Figure S10). Impervious surfaces were also negatively correlated with other land uses, except for a weak positive correlation with green open spaces. Wooded areas, on the other hand, were positively correlated with water areas. Overall, effective population size estimates showed low and mostly non-significant correlations with land use types around breeding habitats (Figure S10).

Discussion

Life-history traits, such as reproductive systems or generational development times, are intimately linked to estimates of effective population size (Caballero, 1994; Wang, 1997; Luikart *et al.*, 2010; Waples *et al.*, 2013; Waples, 2024). For example, the presence of overlapping generations in polygamous species, where individuals reproduce for many years, reduces the effective population size (Beletsky and Orians, 1989; Chen *et al.*, 2019). In contrast, the presence of seed banks in plant populations greatly reduces the loss of heterozygosity over time and thus increases effective size; these conclusions are also applicable to other organisms with similar life-history traits such as diapausing eggs in freshwater crustaceans and other semelparous species with discrete breeding seasons and a variable maturation time such as the Pacific salmon (Nunney, 2002). Understanding the life-history traits of the studied species is therefore a necessary prerequisite for estimating and interpreting effective population sizes (Waples, 2005; Waples and Yokota, 2007; Waples *et al.*, 2013). The development time of a generation is important because if individuals do not maintain a variation in their rate of development, and if no individuals are moving from one cohort to another to maintain a certain degree of genetic cohesion, each temporal cohort will lead in the long term to independent demographic evolutionary trajectories with different numbers of breeders and distinct effective population sizes for each cohort (Taylor and Friesen, 2017; Gradish *et al.*, 2019; Bouaouina *et al.*, 2023).

The southern damselfly voltinism is a plastic trait along a latitudinal gradient (Purse and Thompson, 2002; Mahdjoub *et al.*, 2014, 2015). In addition, the development time of individuals can be decreased by one year by environmental factors such as industrial cooling waters increasing the water temperature (Thelen, 1992). Our results showed a lack of significant levels of genetic differentiation for each population between two sampling years. Moreover, the levels of genetic differentiation measured between spatial populations remained stable over time, and the genetic structure observed in the region was mainly explained by differences between populations and not by differences between sampling years. Our results, although derived at a lower latitude and in a more central area of the species' range, were therefore consistent with previous observations conducted in the United Kingdom, towards the north of the species' range (Watts and Thompson, 2012). We therefore reiterate the findings of Watts and Thompson (2012), indicating low levels of genetic divergence between sympatric cohorts as a result

of developmental plasticity where a fraction of individuals complete their development outside of a two-year (semi-voltine) period, thus mixing alternate-year cohorts.

Effective population size is an essential concept in population genetics and conservation biology because it summarises the history of the population regarding genetic drift and inbreeding, and it helps to assess and to monitor the vulnerability of populations (Waples, 2005; Frankham *et al.*, 2013; Wang *et al.*, 2016; Allendorf *et al.*, 2022). However, estimating this parameter in natural populations is a challenge and poses several difficulties (Gilbert and Whitlock, 2015; Santos-del-Blanco *et al.*, 2022; Waples, 2024). Our results illustrated the difficulty of thoroughly estimating the effective population size in southern damselfly populations. Indeed, even with the use of different types of molecular markers (microsatellites and SNPs) and of complementary estimation methods (single-sample and temporal), we obtained a large number of infinite estimates of $N_{e,}$ or estimations which did not have a finite upper 95% CI. This makes the estimation of effective size imprecise because it remains difficult to disentangle whether an infinite estimate of N_e actually depicts a large N_e or a lower or moderate N_e that could not be correctly estimated owing to insufficient population sampling size or low number of molecular markers used (Waples, 2024; Clarke *et al.*, 2024). Likewise, such biologically meaningless estimates were also observed in southern damselflies, in more isolated populations located at the northern limit of the species' range (United Kingdom) and based on genetic estimates using microsatellite markers (Watts *et al.*, 2007). Furthermore, our results pointed out that N_e estimates based on individual relatedness (see Wang, 2009) with microsatellite markers did not present the problem of infinite effective size. Nonetheless, this method showed abnormally low and constant estimates of effective population size (ranging from 19 to 44) compared to all the other N_e estimators used in this study. Lower estimates on N_e with sibship estimates have already been observed with microsatellite data (e.g. Woltmann *et al.*, 2019) and the discrepancies observed in our results could be explained by violation of the assumptions associated with this estimator such as random sampling of individuals as well as the lack of resolutive information from the microsatellite markers we used (Wang, 2004, 2009).

Overall, the difficulty in estimating effective population sizes in our case study may be related to the different genetic models of populations underlying the N_e estimators and the assumptions they involve (Gilbert and Whitlock, 2015; Wang *et al.*, 2016; Santos-del-Blanco *et al.*, 2022). In line with

this idea, Sovic *et al.* (2019) clearly illustrated, in populations showing a high level of geographic isolation, that point estimates of the effective sizes of populations can vary depending on the type of genetic data and the method used to generate the estimates. Firstly, many N_e estimators assume that populations are isolated because migration complicates the quantification of drift by modifying allele frequencies in a less predictable way (Wang and Whitlock, 2003; Waples and Do, 2010). Gene flow among populations can therefore bias N_e estimates. Yet, we showed using microsatellite loci that populations occurring in this region displayed very low levels of genetic differentiation, indicative of a substantial level of connectivity with gene flow occurring overland among populations (Lévêque *et al*., 2024a), which may explain part of the bias we may have obtained. Secondly, additional common assumptions in estimating effective population sizes are random mating of individuals and balanced sex ratios, both features which are likely to be violated in many wild species, including southern damselflies. Previous observations in the southern damselfly suggested unbalanced sex ratios but also variation in individual reproductive success between males and females (Purse and Thompson, 2005; Rouquette and Thompson, 2007b; Watts *et al.*, 2007). Nonetheless, this type of bias should reduce the estimated effective population sizes, which does not explain the plethora of infinite N_e estimations that we observed. Finally, by definition estimating effective population sizes derives from the ability to quantify the amount of random genetic drift experienced by a population. Unfortunately, when N_e is very large, evolutionary forces other than genetic drift, such as selection and gene flow, will have a relatively greater influence on the estimate of N_e (Hare *et al.*, 2011). In addition, the genetic drift signal for a large effective population is so weak that it can be difficult to distinguish it from sampling error (Wang, 2009; Waples, 2016; Marandel *et al.*, 2019).

However, the accuracy and precision of large N_e estimates can be improved by increasing sampling efforts for both individuals and loci (Waples, 1989; Berthier *et al.*, 2002; Wang, 2009; Luikart *et al.*, 2010). Genomic data (SNP markers) have already shown greater resolution and ability to estimate N_e, but these data also present unique challenges in estimating N_e (Stahlke *et al.*, 2020). Indeed, the use of thousands of SNP loci, many of which are likely to be linked, will result in lower estimates of N_e, unless physical linkage is considered (Do *et al.*, 2014; Waples *et al.*, 2016, 2022). This is why, if a reference genome is available, N_e estimation methods based on the LD between unlinked loci may be

more efficient. Otherwise, when using ddRADseq data, it is recommended to keep only one SNP per RADlocus (Hollenbeck *et al.*, 2016; Lehnert *et al.*, 2019; Waples *et al.*, 2022), as was the case when filtering our genomic data (Lévêque *et al.*, 2024b). Thus, there are still challenges and limitations in estimating effective population sizes, even with increasing amounts of genomic data. Therefore, because N_e is used to inform conservation decisions, it is important to be cautious in interpreting the estimated levels and to cross-check the results obtained from different genetic markers and types of estimators (Shaffer, 1981; Rieman and Allendorf, 2001; Santos-del-Blanco *et al.*, 2022). As a matter of fact, and beyond the problem of the infinite N_e estimates, although the levels of effective population sizes that we measured vary from one estimator to another, we observed that for the same type of genetic marker, the estimated effective sizes tend to be highly correlated with each other. This correlation was also observed between the LD and sibship estimators using microsatellite markers, although the latter was downwardly-biased in terms of N_e levels. This indicates that even if the absolute estimated effective population sizes varied, the general trends measured by the different N_e estimators remained similar, with small populations being consistently estimated as such and large populations being estimated as always quite large. Finally, we did not compare the N_e estimates based on single samples and on temporal estimates. Indeed, they provide independent information about effective population size, with single-sample estimates measuring the genetic result of processes acting in the parental generation, while temporal estimators measure the strength of the temporal variance of allele frequencies across two or more temporal samples (Waples, 2005; Hare *et al.*, 2011).

Altogether, for a given population, the effective population size levels can vary from one estimator to another, making interpretation somewhat difficult and uncertain. However, for some populations, all the N_e estimators gave consistent results, with, for example, a low estimate of N_e levels (comprised between 10 and 50 individuals, e.g. population S21), medium estimates of N_e (with estimates ranging between 50 and 500 individuals, e.g. population S6), or very large N_e, or even infinite N_e, as was the case for the population A7. Focusing on these concordant estimates, different spatial patterns of effective population size emerged. The specific set of populations located along the A355 highway showed highly variable levels of effective population size, even if they were located very close together on either side of the highway. These observations may be relevant in view of the recent construction of

the highway, which started in 2016, and in view of the river restoration and modification work carried out from 2016 to 2020 around this axis, indicating a possible recolonisation with source-sink dynamics of populations.

In that respect, progressive recolonisation processes of rivers after restoration actions was recently demonstrated in another area located in southwestern France in the Pinail National Nature Reserve using direct approaches in the southern damselfly (Beaune and Sellier, 2021). This study illustrated that this Odonata species quickly responded to new habitat availability and heterogeneity and that patches of favourable habitat can be rapidly recolonised by neighbouring populations. Moreover, we expected a negative effect of human footprint because habitat fragmentation and degradation in man-made habitats reduce the census population size and the connectivity among populations, thereby decreasing the level of N_e (reviewed in Clarke *et al.*, 2024). In that respect, the populations located near the Strasbourg metropolitan area gave a mitigated message, with a mix of some local populations displaying low effective population sizes (e.g. populations S21, CCR7) and of populations with, in contrast, very large N_e estimates (e.g. population Wei1). The occurrence of populations exhibiting very small N_e could be linked to a negative impact of this very densely urbanised area on population genetic features, as it was illustrated by a decrease in the levels of genetic diversity and an increase in the levels of genetic differentiation in populations close to the city centre (Lévêque *et al.*, 2004a). Finally, although geographically close, populations located further south of the city of Strasbourg showed highly variable levels of effective population size, with a patchwork of populations characterised by very high N_e levels along the Ger river (e.g. populations Ger3, Ger4, Ger6) but characterised by very low N_e levels in other adjacent rivers (e.g. population Han1 or A14). This suggested that effective population sizes in the southern damselfly would be most certainly dependent on external, potentially environmental, factors.

In this context, we also explored whether the landscape surrounding the breeding habitats where the populations were sampled could explain the patterns of effective population sizes. Although the land use variables were mechanically linked, with a negative correlation between agricultural areas and other types of land use, no land use variable significantly explained the observed variation in effective population sizes in the southern damselfly. This lack of a direct relationship between landscape components and levels of N_e can be explained by an incorrect spatial and/or temporal scale of land use

(Segelbacher *et al.*, 2008; Anderson *et al.*, 2010). Previous studies on southern damselfly dispersal suggested that gene flow among populations may be influenced by the surrounding landscape, with a negative impact from dense urban areas, high-grounded vegetation and scrub, and high altitude, and a positive impact of agricultural areas (see Watts *et al.*, 2006, 2004c; Keller *et al.*, 2012; Lorenzo-Carballa *et al.*, 2015). However, the nature of the landscape did not explain the levels of intra-population genetic diversity in our study. Therefore, the southern damselfly could be more dependent on the microhabitat, local habitat closure and the occurrence of helophytes directly in the watercourse, which cannot be captured with land use data. Indeed, the type of fine-scaled local microhabitat is thought to be tightly linked to the local density of populations (Rouquette and Thompson, 2005, 2007b).

Finally, from a conservation point of view, a general benchmark for assessing whether species are at short-term genetic risk comes from the much-debated "50/500 rule" first proposed by Franklin (1980) for assessing endangered species' minimum viable population size. Basically, this rule argued that populations should be maintained at a minimum short-term effective population size of at least 50 individuals to avoid genetic deleterious effects due to increased inbreeding due to mating between genetically related individuals, and that $N_e > 500$ could ensure the long-term maintenance of evolutionary potential. However, there was much debate about the viability of 50 individuals as a specific numerical target (Jamieson and Allendorf, 2012; Frankham *et al.*, 2014; Clarke *et al.*, 2024). Notwithstanding, our empirical findings highlighted that several populations exhibited contemporary N_e well below 50 (e.g. populations A14, CCR7, Han1, Koh1, RLK56, S21, S67, Rsc1, V2), suggesting that these populations are potentially at risk of local deleterious genetic effects of inbreeding and loss of genetic diversity, and that conservation planning for this species needs to take this possibility into account. On a more positive note, many populations in the region also showed consistently high N_e values wall above 500 (e.g. populations A11, Bli2, Rlei6, Rlei7, Ger4).

Conclusion and perspectives

This study confirmed that there was no genetic isolation between sampling years in southern damselflies, suggesting plasticity in the semi-voltine trait of the species and, consequently, no

independent evolutionary trajectories of temporally-spaced and genetically differentiated cohorts (see also Thelen, 1992; Watts and Thompson, 2012). Although this empirical study illustrated the inherent difficulty of accurately estimating effective population sizes, even using a large set of different genomic markers expected to yield greater accuracy than a microsatellite-based dataset, our results suggested contrasting population dynamics mirroring source and sink recolonisation processes on either side of a newly built highway. Future studies will need to investigate more thoroughly the barrier effect of such infrastructure, as well as the recolonisation processes in watercourses impacted and restored during its construction. Finally, although the Strasbourg city area negatively impacted the levels of effective population sizes for a set of populations, land use around the sampled populations studied did not prove to be a good predictor of effective population size. Further studies focusing (i) on finer scales of investigation and (ii) on landscape elements facilitating or impeding individual dispersal events among subdivided populations will be required to determine whether subtle landscape components are likely to shape the levels of effective population sizes and other critical genetic parameters, such as the amount of gene flow, in the southern damselfly.

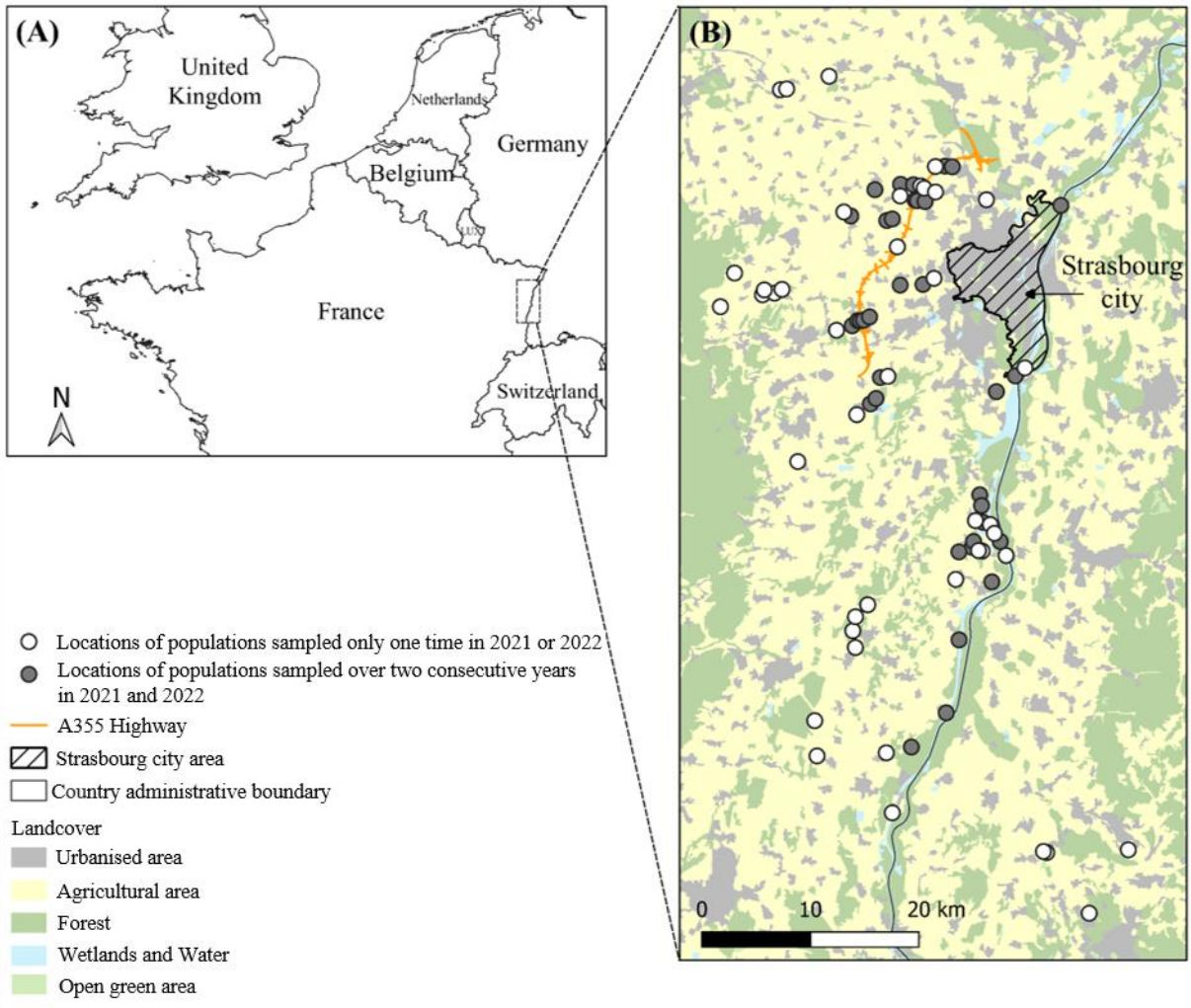


Figure 1: Location of 77 populations of the southern damselfly (*Coenagrion mercuriale*) in North-eastern France. Each dot represents one population, characterised by a single sampling in 2021 or 2022 (N=43, white dots) or by two successive samplings in 2021 and 2022, allowing for effective population size (N_e) estimations based on temporal variance in allele frequencies (N=34, dark grey dots). Land cover was simplified from Corine Land Cover Edition 2018.

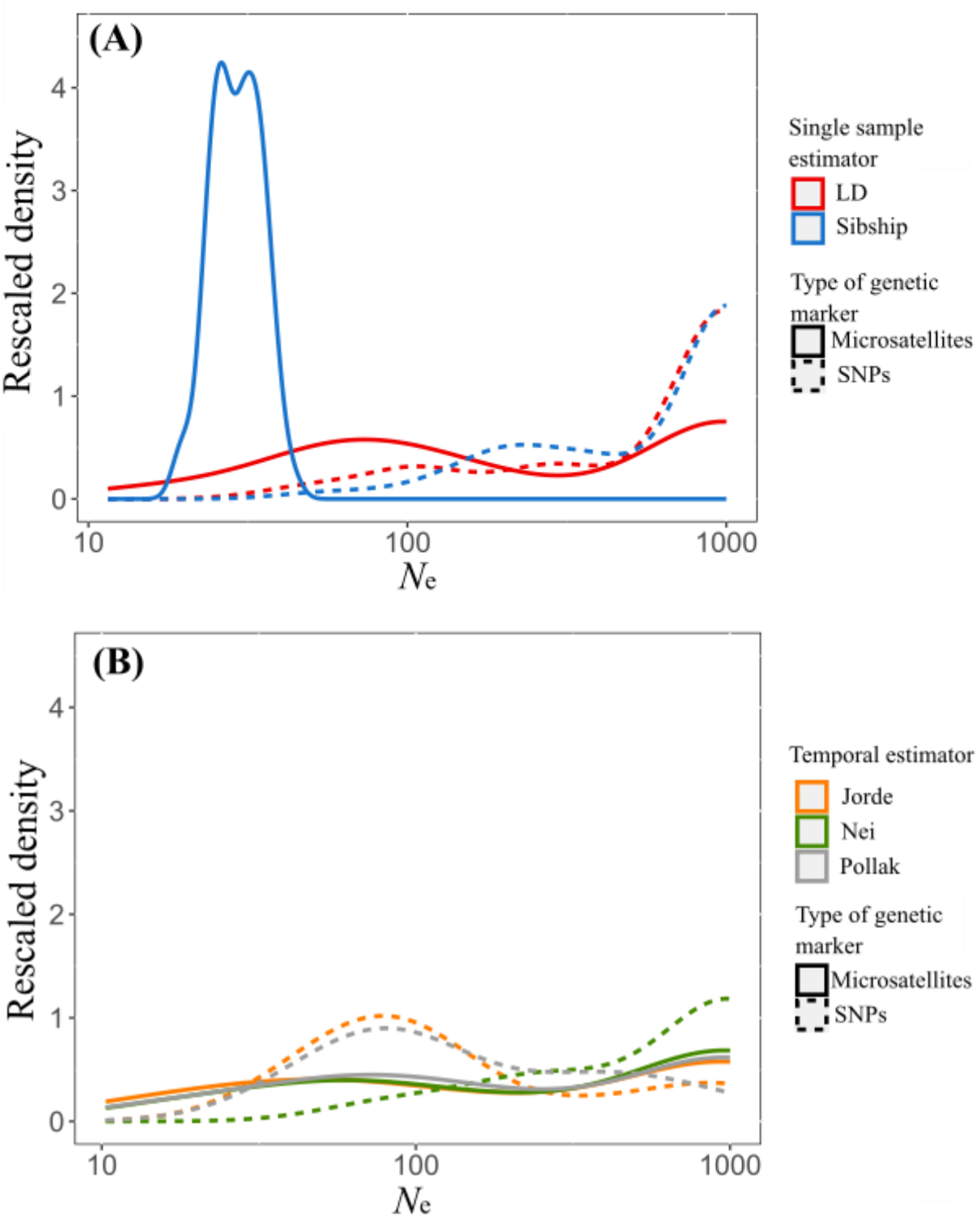


Figure 2: Density plot showing the distribution of effective population size (N_e) estimates for **(A)** single-sample methods (LD: Linkage disequilibrium; Sibship: Full-Likelihood analysis based on inferred sibship frequencies among samples) and **(B)** for temporal methods using F_s (Jorde & Ryman 2007), F_e

Nei & Tajima (1981), and F_k Pollak (1983) for the microsatellite (Microsatellites, solid lines) and the SNP datasets (SNPs, dotted lines).

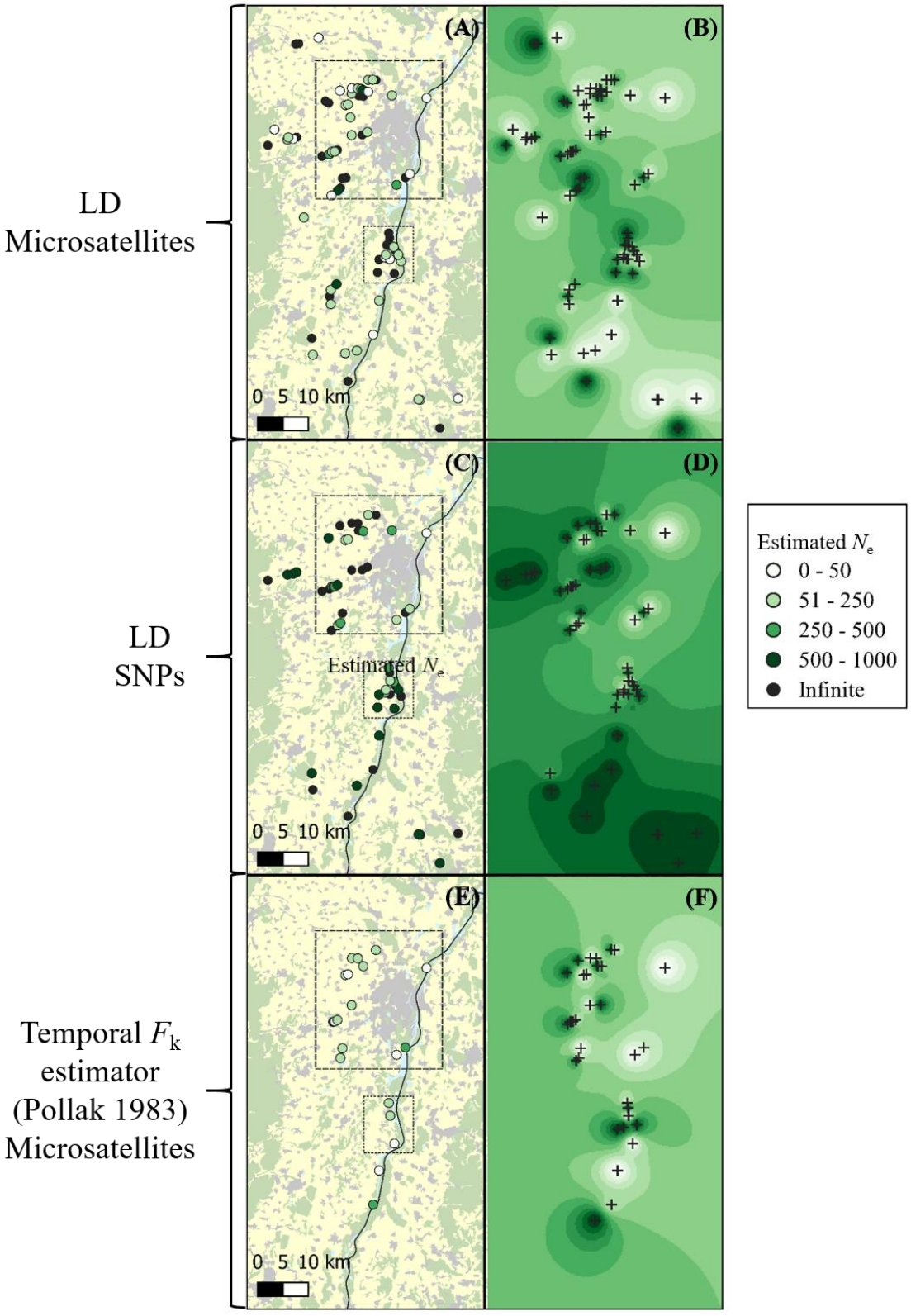


Figure 3: Spatial distribution of effective population size (N_e) estimates in the southern damselfly. The maps show the geographical distribution of N_e estimations **(A, C, E)** and their interpolations in the studied area **(B, D, F)**. Single-sample methods: (**A**, **B**) Linkage disequilibrium estimates using

microsatellites, (**C**, **D**) LD estimates using SNPs. (**E**, **F**) Temporal estimation of N_e using F_k from Pollak (1983) using microsatellites. The dotted rectangles refer to specific area zoomed on in Figure 4.

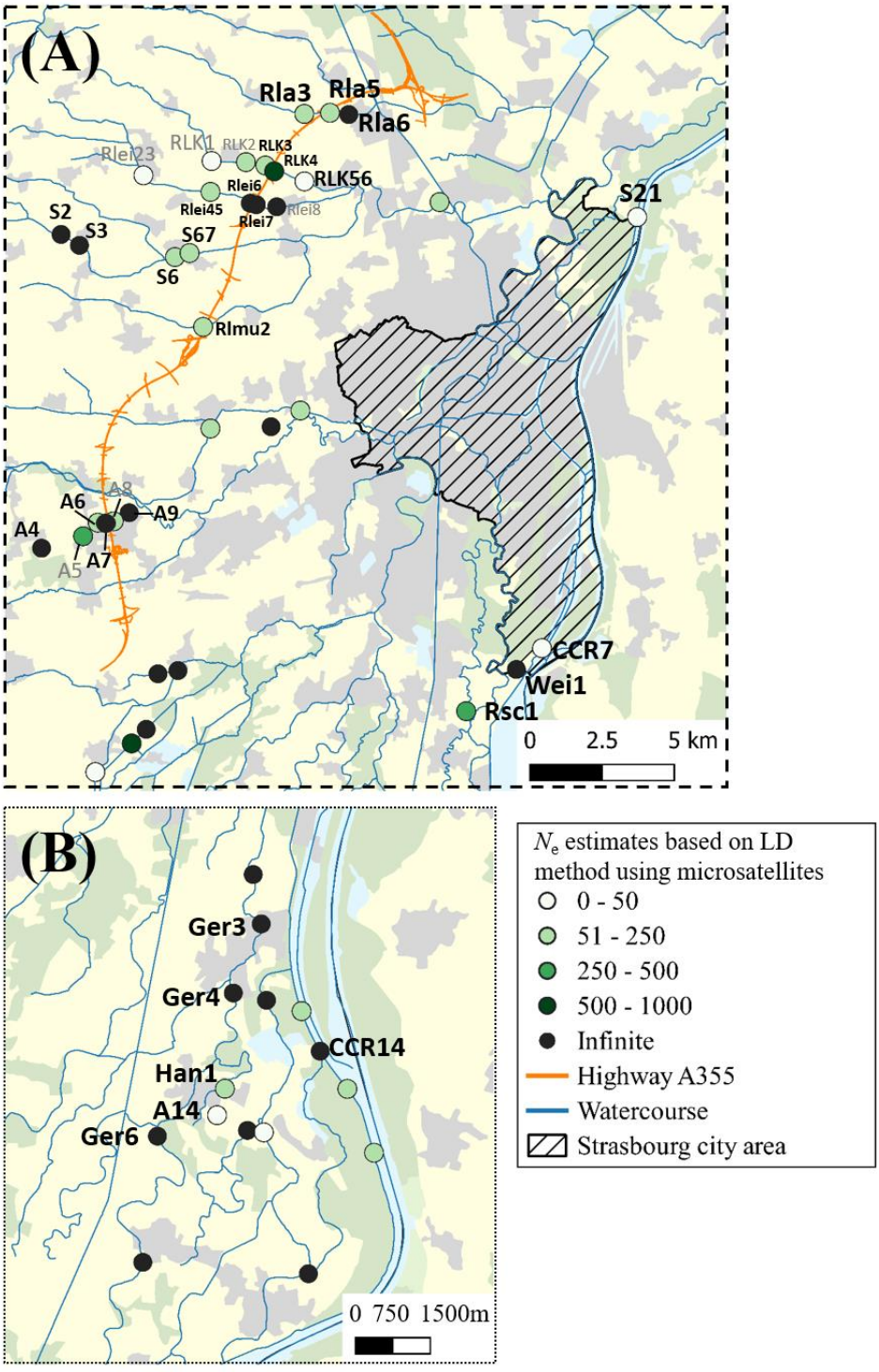


Figure 4: Fine-scaled maps focusing on two areas of interest indicated by dotted rectangles in Figure 3. (**A**) Area around the Strasbourg Eurometropolis and the A355 highway and (**B**) Area located further south of Strasbourg. Distribution of effective population sizes (N_e) estimated for single-samples using the LD method and microsatellites. Major watercourses are represented by a blue line, the A355 highway

by an orange line, and the Strasbourg city limits by a striped area. The names of populations in black indicate populations for which the different single-sample estimators showed congruent results in terms of N_e estimates, while populations whose names are in grey represent populations for which N_e estimates greatly varied among the different estimators we used.

References

Allendorf FW (2017). Genetics and the conservation of natural populations: allozymes to genomes. *Mol Ecol* **26**: 420–430.

Allendorf FW, Funk WC, Aitken SN, Byrne M, Luikart G (2022). *Conservation and the Genomics of Populations*, 3rd edn. Oxford University Press: Oxford, UK

Allendorf FW, Hohenlohe PA, Luikart G (2010). Genomics and the future of conservation genetics. *Nat Rev Genet* **11**: 697–709.

Anderson CD, Epperson BK, Fortin M-J, Holderegger R, James PMA, Rosenberg MS, *et al.* (2010). Considering spatial and temporal scale in landscape-genetic studies of gene flow. *Mol Ecol* **19**: 3565–3575.

Battisti A, Boato A, Masutti L (2000). Influence of silvicultural practices and population genetics on management of the spruce sawfly, *Cephalcia arvensis*. *For Ecol Manag* **128**: 159–166.

Beaune D, Sellier Y (2021). Stream restorations with meanders increase dragonfly and damselfly diversity and abundance, including an endangered species. *J Nat Conserv* **60**: 125950.

Beletsky LD, Orians GH (1989). A male red-winged blackbird breeds for 11 years. *Northwest Nat* **70**: 10–12.

Bernatchez L, Ferchaud A-L, Berger CS, Venney CJ, Xuereb A (2024). Genomics for monitoring and understanding species responses to global climate change. *Nat Rev Genet* **25**: 165–183.

Berthier P, Beaumont MA, Cornuet J-M, Luikart G (2002). Likelihood-based estimation of the effective population size using temporal changes in allele frequencies: a genealogical approach. *Genetics* **160**: 741–751.

Bouaouina S, Chittaro Y, Willi Y, Lucek K (2023). Asynchronous life cycles contribute to reproductive isolation between two Alpine butterflies. *Evol Lett* **7**: 436–446.

Browning SR, Browning BL (2015). Accurate non-parametric estimation of recent effective population size from segments of identity by descent. *Am J Hum Genet* **97**: 404–418.

Caballero A (1994). Developments in the prediction of effective population size. *Heredity* **73**: 657–679.

Chen N, Juric I, Cosgrove EJ, Bowman R, Fitzpatrick JW, Schoech SJ, *et al.* (2019). Allele frequency dynamics in a pedigreed natural population. *Proc Natl Acad Sci* **116**: 2158–2164.

Clarke SH, Lawrence ER, Matte J, Gallagher BK, Salisbury SJ, Michaelides SN, *et al.* (2024). Global assessment of effective population sizes: Consistent taxonomic differences in meeting the 50/500 rule. *Mol Ecol* **in press**: e17353.

Collevatti RG, dos Santos JS, Rosa FF, Amaral TS, Chaves LJ, Ribeiro MC (2020). Multi-scale landscape influences on genetic diversity and adaptive traits in a neotropical savanna tree. *Front Genet* **11**: 508563.

Couvreur JM, Dufrêne M, Goffart P, Vandevyvre X, Etienne F, Testart D (2008). Nouvelles estimations des effectifs de l'Agrion de Mercure (*Coenagrion mercuriale*) dans la plaine du Biran (Commune de Beauraing, Belgique) avec une analyse des principaux facteurs écologiques expliquant son abondance. *Bull Société R Belge Entomol* **144**: 101–115.

Crispim B do A, Fernandes J dos S, Bajay MM, Zucchi MI, Batista CE de A, Vieira M do C, *et al.* (2021). Genetic diversity of *Campomanesia adamantium* and its correlation with land use and land cover. *Diversity* **13**: 160.

Do C, Waples RS, Peel D, Macbeth GM, Tillett BJ, Ovenden JR (2014). NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (N_e) from genetic data. *Mol Ecol Resour* **14**: 209–214.

Dray S, Dufour A-B (2007). The ade4 Package: implementing the duality diagram for ecologists. *J Stat Softw* **22**: 1–20.

Ellegren H, Galtier N (2016). Determinants of genetic diversity. *Nat Rev Genet* **17**: 422–433.

Fahrig L (2003). Effects of habitat fragmentation on biodiversity. *Annu Rev Ecol Evol Syst* **34**: 487–515.

Fincke OM, Hadrys H (2001). Unpredictable offspring survivorship in the damselfly, *Megaloprepus coerulatus*, shapes parental behavior, constrains sexual selection, and challenges traditional fitness estimates. *Evolution* **55**: 762–772.

Fisher R A (1931). The genetical theory of natural selection. *Oxford Clarendon Press*: Oxford, UK.

Flavenot T, Fellous S, Abdelkrim J, Baguette M, Coulon A (2015). Impact of quarrying on genetic diversity: an approach across landscapes and over time. *Conserv Genet* **16**: 181–194.

Frankham R, Bradshaw CJA, Brook BW (2014). Genetics in conservation management: revised recommendations for the 50/500 rules, Red List criteria and population viability analyses. *Biol Conserv* **170**: 56–63.

Frankham R, Briscoe DA, Ballou JD (2013). *Introduction to conservation genetics*, 2d edition. Cambridge University Press: Cambridge UK.

Franklin IR (1980). Evolutionary change in small populations. In: *Conservation biology: an evolutionary ecological perspective.*, Sinauer Associates, Inc.: Sunderland, Massachusetts, pp 135–150.

Gilbert KJ, Whitlock MC (2015). Evaluating methods for estimating local effective population size with and without migration. *Evolution* **69**: 2154–2166.

Goudet J (2003). FSTAT (ver. 2.9.4), a program to estimate and test population genetics parameters. Available from http://www.unil.ch/izea/softwares/fstat. html Updated from Goudet (1995)

Goudet J (2005). Hierfstat, a package for R to compute and test hierarchical *F*-statistics. *Mol Ecol Notes* **5**: 184–186.

Gradish AE, Keyghobadi N, Sperling FAH, Otis GW (2019). Population genetic structure and assessment of allochronic divergence in the Macoun's Arctic (*Oeneis macounii*) butterfly. *Can J Zool* **97**: 121–130.

Grand D (1996). *Coenagrion mercuriale (Charpentier, 1840). Background information on invertebrates of the Habitats Directive and the Bern Convention, Part II: Mantodea, Odonata, Orthoptera, and Arachnida.*, Council of Europe Publishing (P van Helsdingen, L Willemse, and M Speight, Eds.). Strasbourg.

Hare MP, Nunney L, Schwartz MK, Ruzzante DE, Burford M, Waples RS, *et al.* (2011). Understanding and estimating effective population size for practical application in marine species management. *Conserv Biol* **25**: 438–449.

Hill W (1981). Estimation of effective population size from data on linkage disequilibrium. *Genet Res* **38**: 209–216.

Hohenlohe PA, Funk WC, Rajora OP (2021). Population genomics for wildlife conservation and management. *Mol Ecol* **30**: 62–82.

Holderegger R, Balkenhol N, Bolliger J, Engler JO, Gugerli F, Hochkirch A, *et al.* (2019). Conservation genetics: linking science with practice. *Mol Ecol* **28**: 3848–3856.

Hollenbeck CM, Portnoy DS, Gold JR (2016). A method for detecting recent changes in contemporary effective population size from linkage disequilibrium at linked and unlinked loci. *Heredity* **117**: 207–216.

Jamieson IG, Allendorf FW (2012). How does the 50/500 rule apply to MVPs? *Trends Ecol Evol* **27**: 578–584.

Jones OR, Wang J (2010). COLONY: a program for parentage and sibship inference from multilocus genotype data. *Mol Ecol Resour* **10**: 551–555.

Jorde PE, Ryman N (2007). Unbiased estimator for genetic drift and effective population size. *Genetics* **177**: 927–935.

Keller D, Van Strien MJ, Holderegger R (2012). Do landscape barriers affect functional connectivity of populations of an endangered damselfly? *Freshw Biol* **57**: 1373–1384.

La Porta G, Goretti E (2019). Investigation on the declining southern damselfly (*Coenagrion mercuriale*, Odonata) in a Mediterranean population: survival rate and population size. *J Insect Conserv* **23**: 667–675.

Langmead B, Salzberg SL (2012). Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**: 357–359.

Lehnert SJ, Kess T, Bentzen P, Kent MP, Lien S, Gilbey J, *et al.* (2019). Genomic signatures and correlates of widespread population declines in salmon. *Nat Commun* **10**: 2996.

Lévêque A, Duputié A, Vignon V, Duez F, Godé C, Vanappelghem C, Arnaud J-F (2024a). Contrasting patterns of spatial genetic structure in endangered southern damselfly (*Coenagrion mercuriale*) populations facing habitat fragmentation and urbanisation. bioRxiv 2024.05.02.592171; doi: https://doi.org/10.1101/2024.05.02.592171

Lévêque A, Arnaud J-F, Vignon V, Mazoyer C, Godé C, Duputié A (2024b). Pitfalls and recommendations for large-scale SNP genotyping in a non-model endangered species: the southern damselflies (*Coenagrion mercuriale*) as a case study. bioRxiv 2024.05.02.592185; doi: https://doi.org/10.1101/2024.05.02.592185

Lorenzo-Carballa MO, Ferreira S, Sims AM, Thompson DJ, Watts PC, Cher Y, *et al.* (2015). Impact of landscape on spatial genetic structure and diversity of *Coenagrion mercuriale* (Zygoptera:Coenagrionidae) in northern France. *Freshw Sci* **34**: 1065–1078.

Luikart G, Ryman N, Tallmon DA, Schwartz MK, Allendorf FW (2010). Estimation of census and effective population sizes: the increasing usefulness of DNA-based approaches. *Conserv Genet* **11**: 355–373.

Mahdjoub H, Khelifa R, Zebsa R, Bouslama Z, Houhamdi M (2015). Bivoltinism in *Coenagrion mercuriale* (Zygoptera: Odonata) in the southern margin of its distribution range: emergence pattern and larval growth. *Afr Entomol* **23**: 59–67.

Mahdjoub H, Khelifa R, Zebsa R, Mellal MK, Bouslama Z, Houhamdi M (2014). Aspects of reproductive biology and ecology of *Coenagrion mercuriale* at its southern range margin. *Int J Odonatol* **17**: 173–180.

Marandel F, Lorance P, Berthelé O, Trenkel VM, Waples RS, Lamy J-B (2019). Estimating effective population size of large marine populations, is it feasible? *Fish Fish* **20**: 189–198.

Morin PA, Archer FI, Pease VL, Hancock-Hanser BL, Robertson KM, Huebinger RM, *et al.* (2012). An empirical comparison of SNPs and microsatellites for population structure, assignment, and demographic analyses of bowhead whale populations. *Endanger Species Res* **19**: 1–27.

Munshi-South J, Zolnik CP, Harris SE (2016). Population genomics of the Anthropocene: urbanization is negatively associated with genome-wide variation in white-footed mouse populations. *Evol Appl* **9**: 546–564.

Nei M, Tajima F (1981). Genetic drift and estimation of effective population size. *Genetics* **98**: 625–640.

Nunney L (2002). The effective size of annual plant populations: the interaction of a seed bank with fluctuating population size in maintaining genetic variation. *Am Nat* **160**: 195–204.

Pebesma E (2004). Multivariable geostatistics in S: the gstat package. *Comput Geosci* **30**: 683–691.

Pollak E (1983). A new method for estimating the effective population size from allele frequency changes. *Genetics* **104**: 531–548.

Pudovkin AI, Zaykin DV, Hedgecock D (1996). On the potential for estimating the effective number of breeders from heterozygote-excess in progeny. *Genetics* **144**: 383–387.

Purse BV, Hopkins GW, Day KJ, Thompson DJ (2003). Dispersal characteristics and management of a rare damselfly. *J Appl Ecol* **40**: 716–728.

Purse B, Thompson D (2002). Voltinism and larval growth pattern in *Coenagrion mercuriale* (Odonata: Coenagrionidae) at its northern range margin. *Eur J Entomol* **99**: 11–18.

Purse BV, Thompson DJ (2005). Lifetime mating success in a marginal population of a damselfly, *Coenagrion mercuriale*. *Anim Behav* **69**: 1303–1315.

Purse BV, Thompson DJ (2009). Oviposition site selection by *Coenagrion mercuriale* (Odonata: Coenagrionidae). *Int J Odonatol* **12**: 257–273.

Rieman BE, Allendorf FW (2001). Effective population size and genetic conservation criteria for bull trout. *North Am J Fish Manag* **21**: 756–764.

Rouquette JR, Thompson DJ (2005). Habitat associations of the endangered damselfly, *Coenagrion mercuriale*, in a water meadow ditch system in southern England. *Biol Conserv* **123**: 225–235.

Rouquette JR, Thompson DJ (2007a). Patterns of movement and dispersal in an endangered damselfly and the consequences for its management: movement and dispersal in a damselfly. *J Appl Ecol* **44**: 692–701.

Rouquette JR, Thompson DJ (2007b). Roosting site selection in the endangered damselfly, *Coenagrion mercuriale*, and implications for habitat design. *J Insect Conserv* **11**: 187–193.

Saccheri I, Kuussaari M, Kankare M, Vikman P, Fortelius W, Hanski I (1998). Inbreeding and extinction in a butterfly metapopulation. *Nature* **392**: 491–494.

Santos-del-Blanco L, Olsson S, Budde KB, Grivet D, González-Martínez SC, Alía R, *et al.* (2022). On the feasibility of estimating contemporary effective population size (*Ne*) for genetic conservation and monitoring of forest trees. *Biol Conserv* **273**: 109704.

Schloerke B, Cook D, Larmarange J, Briatte F, Marbach M, Thoen E, *et al.* (2022). GGally: extension to 'ggplot2'.

Segelbacher G, Manel S, Tomiuk J (2008). Temporal and spatial analyses disclose consequences of habitat fragmentation on the genetic diversity in capercaillie (*Tetrao urogallus*). *Mol Ecol* **17**: 2356–2367.

Shaffer ML (1981). Minimum population sizes for species conservation. *BioScience* **31**: 131–134.

Smouse P, Long J, Sokal R (1986). Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Syst Zool* **35**: 627–632.

Soulé ME, Mills LS (1998). No need to isolate genetics. *Science* **282**: 1658–1659.

Sovic M, Fries A, Martin SA, Lisle Gibbs H (2019). Genetic signatures of small effective population sizes and demographic declines in an endangered rattlesnake, *Sistrurus catenatus*. *Evol Appl* **12**: 664–678.

Stahlke A, Bell D, Dhendup T, Kern B, Pannoni S, Robinson Z, *et al.* (2020). Population genomics training for the next generation of conservation geneticists: ConGen 2018 Workshop. *J Hered* **111**: 227–236.

Sternberg K, Buchwald R, Röske W (1999). *Coenagrion mercuriale* (Charpentier, 1840)-Helm Azurjungfer. *Libellen Baden-Württ*: 255–270.

Taylor RS, Friesen VL (2017). The role of allochrony in speciation. *Mol Ecol* **26**: 3330–3342.

Thelen C (1992). Untersuchung zum Larvenhabitat und zum Entwicklungszyklus der Helmazurjungfer (*Coenagrion mercuriale*, Zygoptera: Odonata) an zwei verchiedenen Gewassern der Freiburger Bucht. *Limnol Inst Konstanz Univ Freibg*.

Thompson DJ, Rouquette JR, Purse BV (2003). Ecology of the southern damselfly. *Conserv Nat 2000 Rivers Ecol Ser*.

Thompson DJ, Watts PC (2006). The structure of the Coenagrion mercuriale populations in the New Forest, southern England. *For Dragonflies*: 239–258.

Villanueva RAM, Chen ZJ (2019). ggplot2: elegant graphics for data analysis. *Meas Interdiscip Res* **17**: 160–167.

Wang J (1997). More efficient breeding systems for controlling inbreeding and effective size in animal populations. *Heredity* **79**: 591–599.

Wang J (2004). Sibship reconstruction from genetic data with typing errors. *Genetics* **166**: 1963–1979.

Wang J (2009). A new method for estimating effective population sizes from a single sample of multilocus genotypes. *Mol Ecol* **18**: 2148–2164.

Wang J, Santiago E, Caballero A (2016). Prediction and estimation of effective population size. *Heredity* **117**: 193–206.

Wang J, Whitlock MC (2003). Estimating effective population size and migration rates from genetic samples over space and time. *Genetics* **163**: 429–446.

Waples RS (1989). A generalized approach for estimating effective population size from temporal changes in allele frequency. *Genetics* **121**: 379–391.

Waples RS (2005). Genetic estimates of contemporary effective population size: to what time periods do the estimates apply? *Mol Ecol* **14**: 3335–3352.

Waples RS (2006). A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. *Conserv Genet* **7**: 147–168.

Waples RS (2016). Tiny estimates of the *Ne/N* ratio in marine fishes: are they real? *J Fish Biol* **89**: 2479–2504.

Waples RS (2024). Practical application of the linkage disequilibrium method for estimating contemporary effective population size: a review. *Mol Ecol Resour* **24**: e13879.

Waples RS, Antao T, Luikart G (2014). Effects of overlapping generations on linkage disequilibrium estimates of effective population size. *Genetics* **197**: 769–780.

Waples RS, Do C (2010). Linkage disequilibrium estimates of contemporary *Ne* using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. *Evol Appl* **3**: 244–262.

Waples RK, Larson WA, Waples RS (2016). Estimating contemporary effective population size in non-model species using linkage disequilibrium across thousands of loci. *Heredity* **117**: 233–240.

Waples RS, Lindley ST (2018). Genomics and conservation units: the genetic basis of adult migration timing in Pacific salmonids. *Evol Appl* **11**: 1518–1526.

Waples RS, Luikart G, Faulkner JR, Tallmon DA (2013). Simple life-history traits explain key effective population size ratios across diverse taxa. *Proc R Soc B Biol Sci* **280**: 20131339.

Waples RS, Scribner KT, Moore JA, Draheim HM, Etter D, Boersen M (2018). Accounting for age structure and spatial structure in eco-evolutionary analyses of a large, mobile vertebrate. *J Hered* **109**: 709–723.

Waples RS, Waples RK, Ward EJ (2022). Pseudoreplication in genomic-scale data sets. *Mol Ecol Resour* **22**: 503–518.

Waples RS, Yokota M (2007). Temporal estimates of effective population size in species with overlapping generations. *Genetics* **175**: 219–233.

Watts PC, Rouquette JR, Saccheri IJ, Kemp SJ, Thompson DJ (2004c). Molecular and ecological evidence for small-scale isolation by distance in an endangered damselfly, *Coenagrion mercuriale*. *Mol Ecol* **13**: 2931–2945.

Watts PC, Saccheri IJ, Kemp SJ, Thompson DJ (2006). Population structure and the impact of regional and local habitat isolation upon levels of genetic diversity of the endangered damselfly *Coenagrion mercuriale* (Odonata: Zygoptera). *Freshw Biol* **51**: 193–205.

Watts PC, Saccheri IJ, Kemp SJ, Thompson DJ (2007). Effective population sizes and migration rates in fragmented populations of an endangered insect (*Coenagrion mercuriale* : Odonata). *J Anim Ecol* **76**: 790–800.

Watts PC, Thompson DJ (2012). Developmental plasticity as a cohesive evolutionary process between sympatric alternate-year insect cohorts. *Heredity* **108**: 236–241.

Watts PC, Thompson DJ, Kemp SJ (2004b). Cross-species amplification of microsatellite loci in some European zygopteran species (Odonata: Coenagrionidae). *Int J Odonatol* **7**: 87–96.

Watts PC, Wu JH, Westgarth C, Thompson DJ, Kemp SJ (2004a). A panel of microsatellite loci for the southern damselfly, *Coenagrion mercuriale* (Odonata: Coenagrionidae). *Conserv Genet* **5**: 117–119.

Wei T, Simko V (2017). R package "corrplot": visualization of a correlation matrix. : 1–5.

Wilson MC, Chen X-Y, Corlett RT, Didham RK, Ding P, Holt RD, *et al.* (2016). Habitat fragmentation and biodiversity conservation: key findings and future challenges. *Landsc Ecol* **31**: 219–227.

Wolf B, Zwick P (1989). Plurimodal emergence and plurivoltinism of central European populations of *Nemurella pictetii* (Plecoptera: Nemouridae). *Oecologia* **79**: 431–438.

Woltmann S, Roeder M, Xiong A, Hill C, Conway MH (2019). Additional notes on Texas seaside sparrows (*Ammospiza maritima sennetti*): an update to Griscom 1948. *Wilson J Ornithol* **131**: 750.

Wright S (1931). Evolution in Mendelian populations. *Genetics* **16**: 97–159.

Zhou Y, Tian X, Browning BL, Browning SR (2018). POPdemog: visualizing population demographic history from simulation scripts. *Bioinformatics* **34**: 2854–2855.

Supporting information

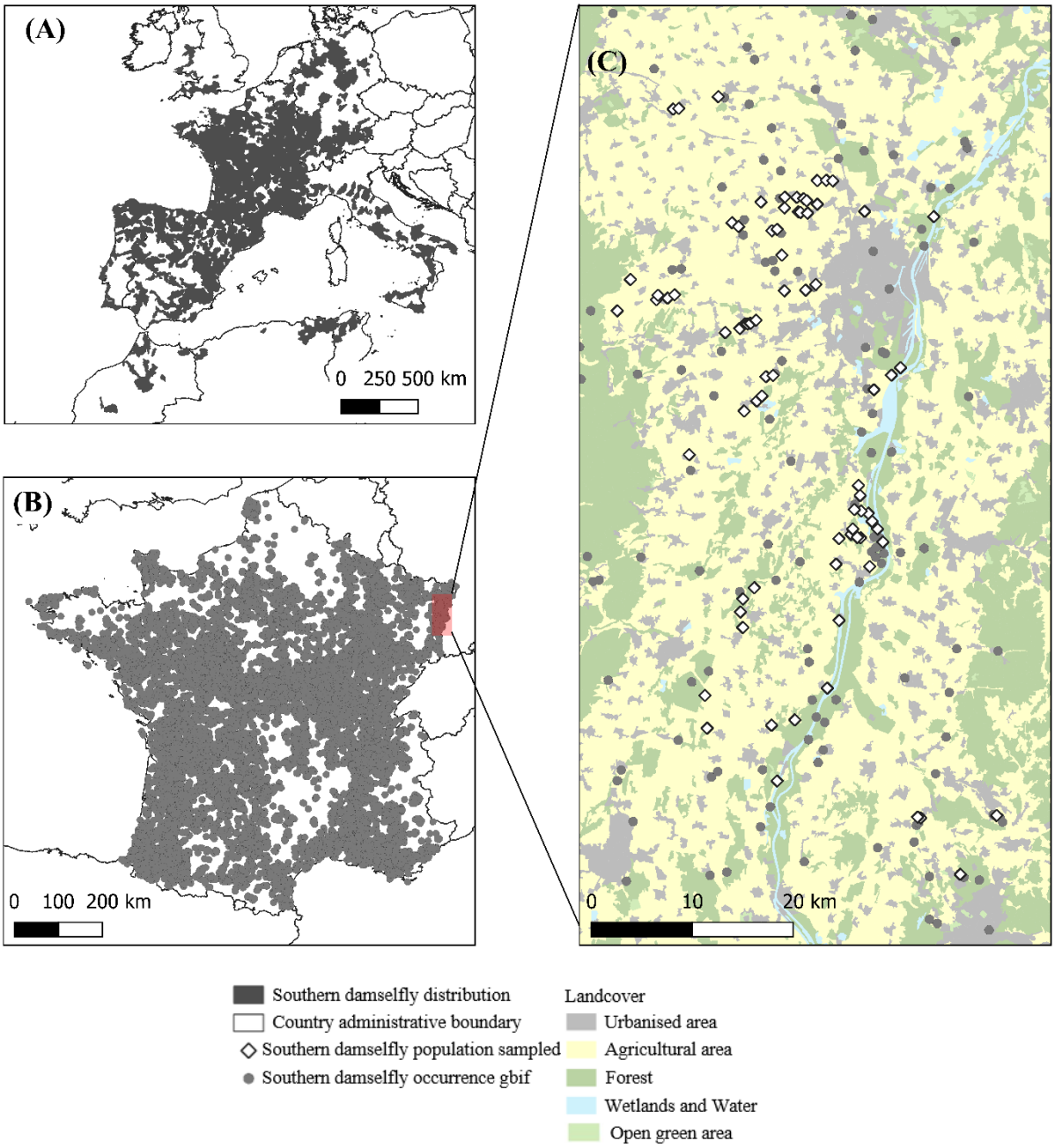


Figure S1: (A) Map showing the current geographical distribution (dark grey zones) of the southern damselfly (*Coegrion mercuriale*). Black lines indicate the administrative limits of countries. Data obtained from IUCN SSC Odonata Specialist Group 2019. The IUCN Red List of Threatened Species. Version 2022-2. https://www.iucnredlist.org/ Downloaded on 28 July 2023. **(B)** Map showing the occurrences of southern damselfly recorded in France (grey dots). The study zone in panel B is highlighted in red. Data obtained from GBIF.org (5 March 2024) GBIF Occurrence Download https://doi.org/10.15468/dl.5z9gue. **(C)** Geographical location of southern damselfly sampling sites in north-eastern France (white diamonds). Land cover was simplified from Corine Land Cover Edition 2018. Occurrences of southern damselfly recorded in France and Germany GBIF.org (5 March 2024) GBIF Occurrence Download https://doi.org/10.15468/dl.9h9maw) are indicated by grey dots.

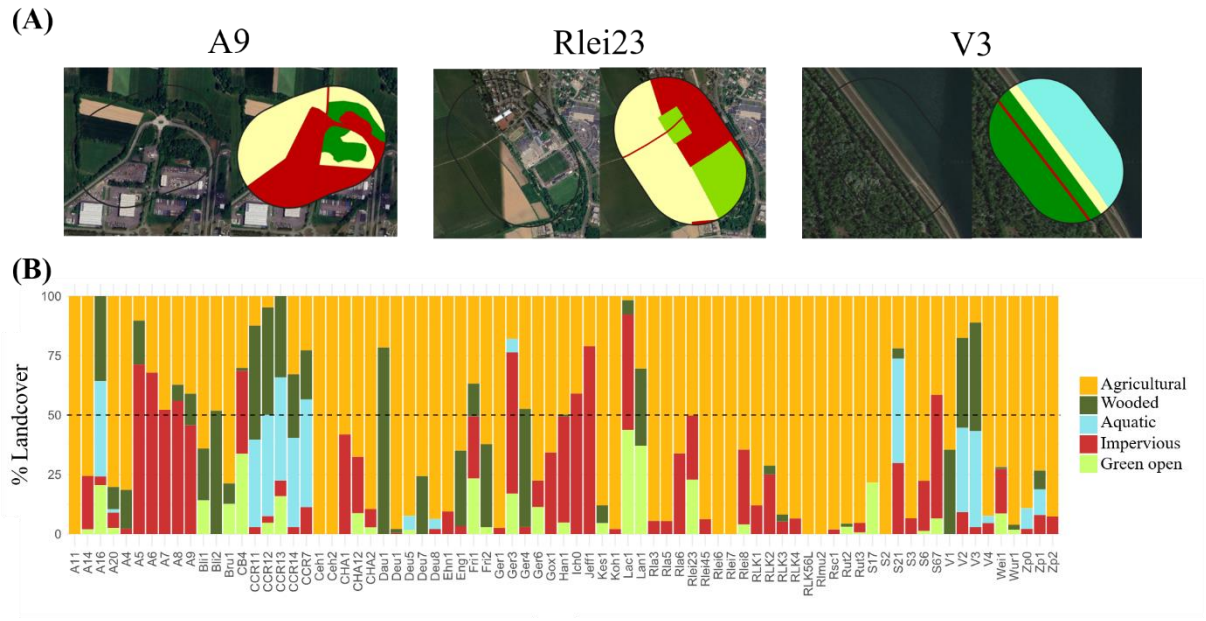
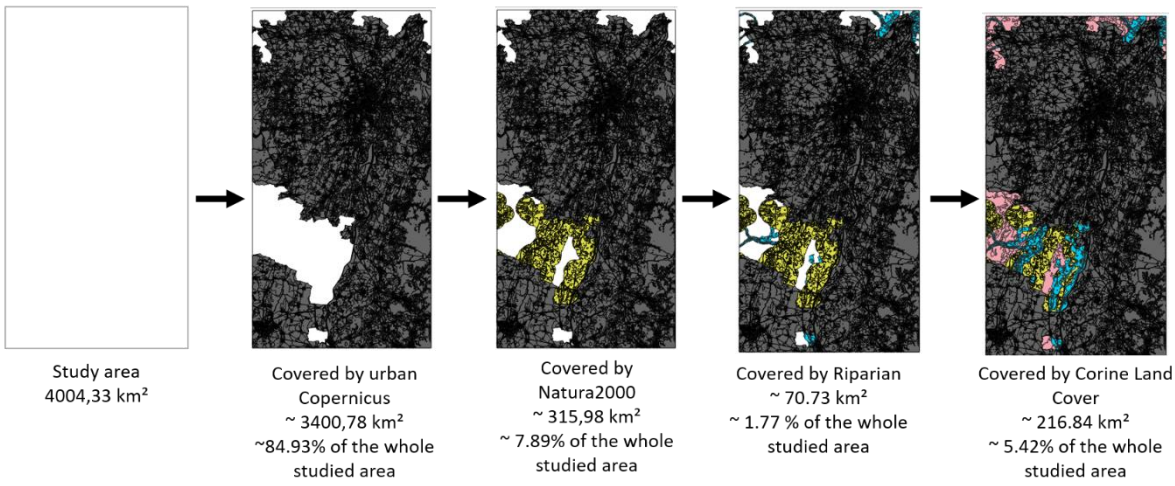
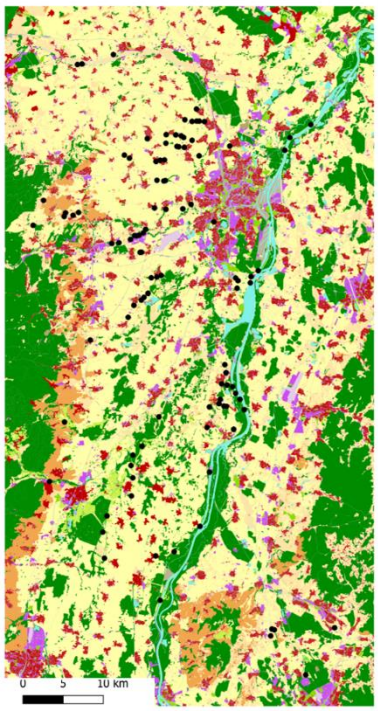


Figure S2: Percentage of land cover surrounding each sampled site within a 200m buffer zone. (**A**) Illustration of buffer zones and associated land cover. (**B**) Percentage of land cover around each sampled population: agricultural areas (arable land, vineyards), wooded areas, aquatic areas (wetlands, large watercourses), impervious areas (buildings, transport networks), open green areas (parks, meadows).

(A)



(B)



Level 1	Level 2	Codes level 2
Impervious	Urban fabric	110
	Industrial, commercial units	121
	Transport units	122
	Port areas	123
	Airport	124
	Mine, dump and construction sites	131
Agricultural	Arable land	210
	Permanent crops	220
	Heterogéneous agricultural areas	230
Forest	Forest	310
Green Open	Artificial, non-agricultural	141
Green Open	Scrub and/or herbaceous vegetation associations	320
	Open spaces with little to no vegetation	330
Water	Wetlands	400
	Water	500

Nomenclature for land use mapping in our study areas

● Southern damselfly sample sites found in Alsace

Figure S3: Construction of a land cover layer for the whole studied area. (**A**) Successive construction steps. (**B**) The final layer obtained and associated nomenclature. Sampled populations are indicated by black dots.

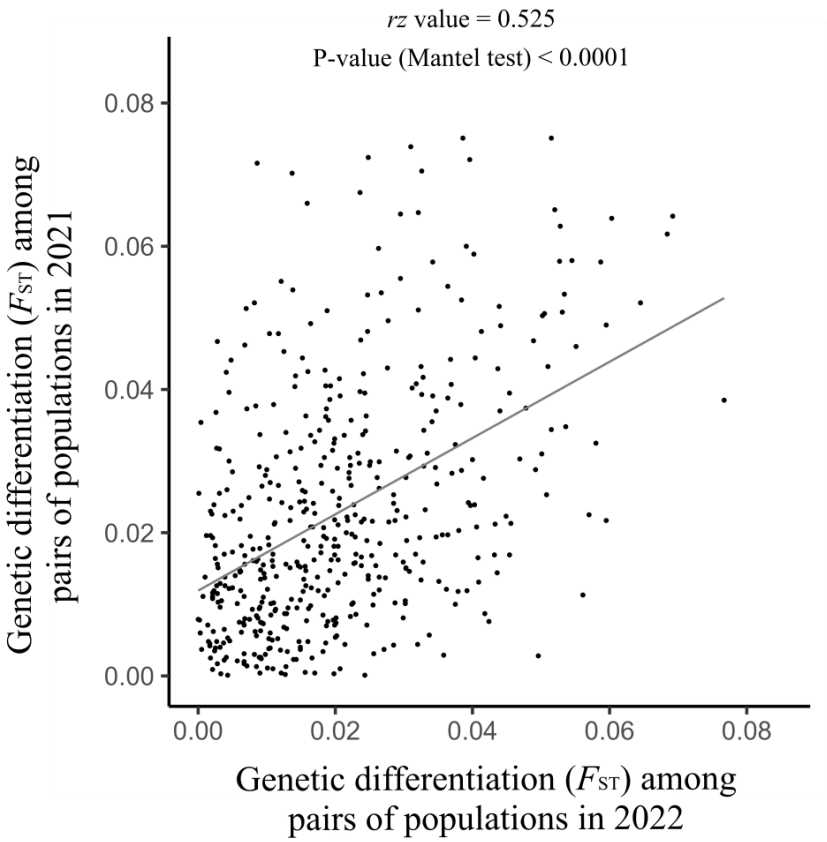


Figure S4: Scatterplot of estimates of pairwise genetic differentiation (F_{ST}) for samples from successive years 2021 and 2022 cohorts in the southern damselfly (*Coegrion mercuriale)*.

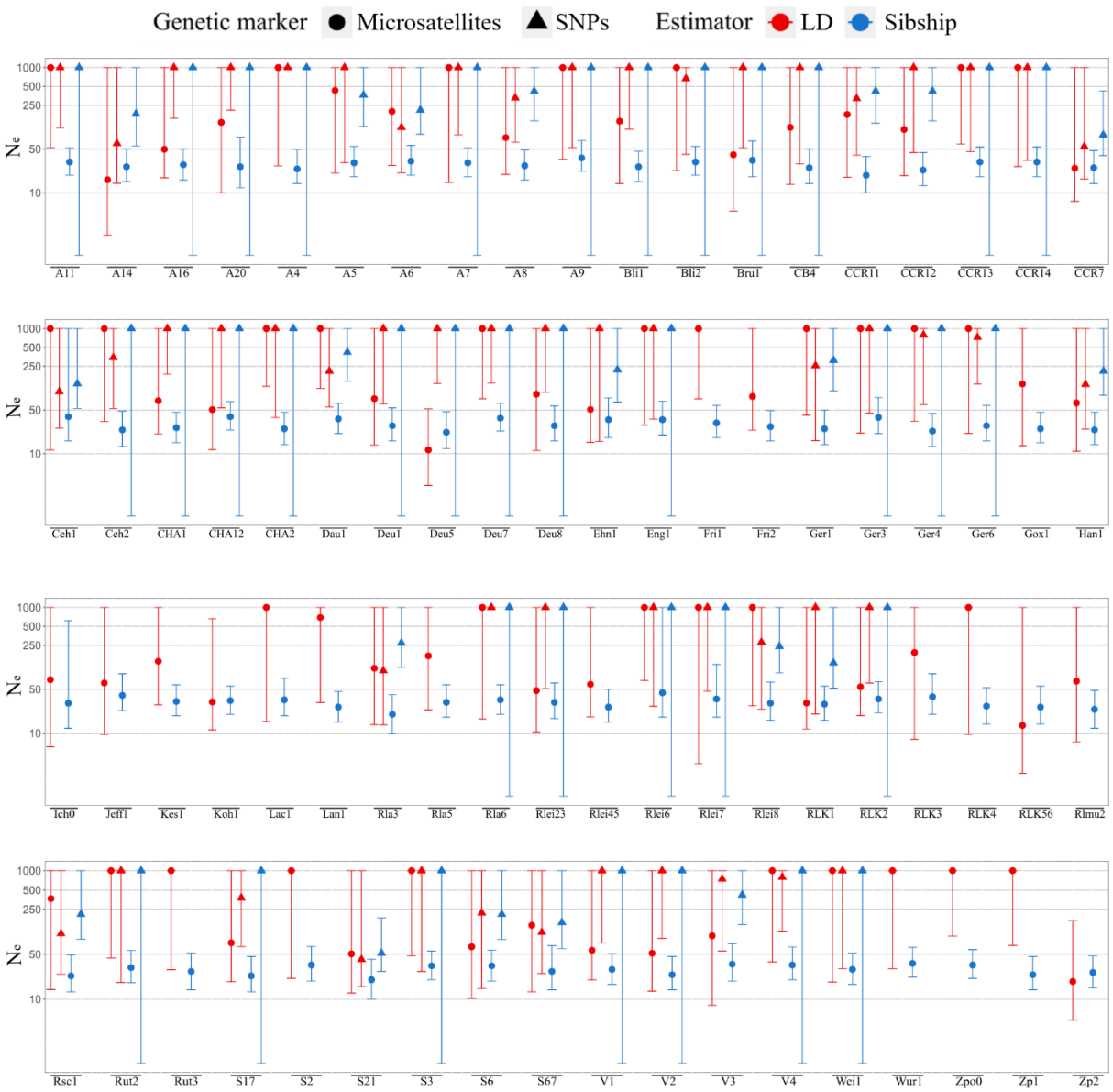


Figure S5: Contemporary effective population size (N_e) estimates based on Linkage Disequilibrium (LD, red) and sibship assignments (Sibship, blue) generated using microsatellites (circles) for 77 southern damselfly populations and SNPs (triangles) for 56 populations. Jackknifed-based confidence intervals for LD estimation and bootstrap resampling confidence intervals for Full-Likelihood analysis are shown. Infinite estimates of N_e are represented by values of 1000.

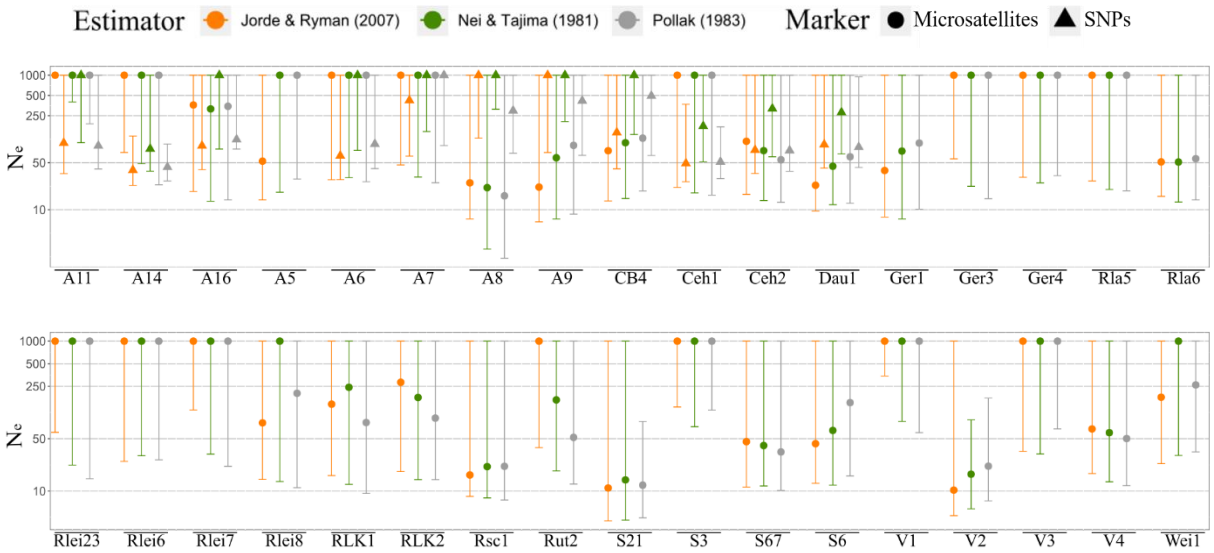


Figure S6: Contemporary effective population size (N_e) estimates and jackknifed-based confidence intervals based on temporal estimates of Jorde & Ryman (2007, orange), of Nei & Tajima (1981, green) and of Pollak (1983, grey) generated using microsatellites (circles) for 34 southern damselfly populations and using SNPs (triangles) for 11 populations. Infinite estimates of N_e are represented by values of 1000.

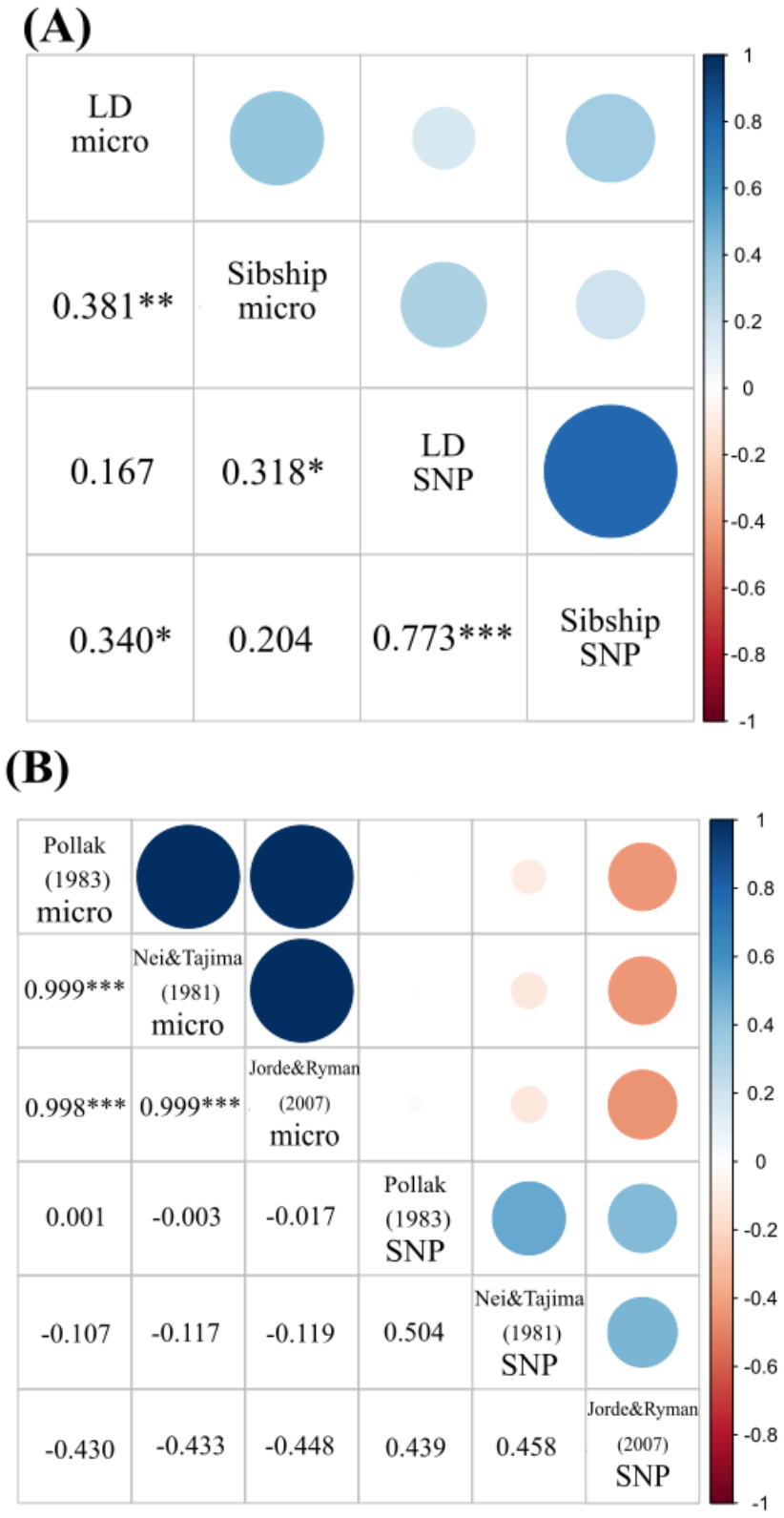


Figure S7: Graphical representation of the correlation matrix (Pearson's correlation coefficients) between finite effective population size (N_e) estimates for **(A)** single-sample methods (LD: Linkage disequilibrium; Sibship: Full-Likelihood analysis based on inferred sibship frequencies among samples; 56 populations) and for **(B)** temporal methods using F_s estimator (Jorde & Ryman 2007), F_e estimator (Nei & Tajima 1981), and F_k estimator (Pollak 1983) using microsatellites (micro, 34 populations) and the SNPs (SNP; 11 populations). The colour and size of the circles (upper triangle) indicate the strength

of the correlation (Red circles indicate a negative correlation; blue circles indicate a positive correlation; white circles indicate no correlation). *** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$.

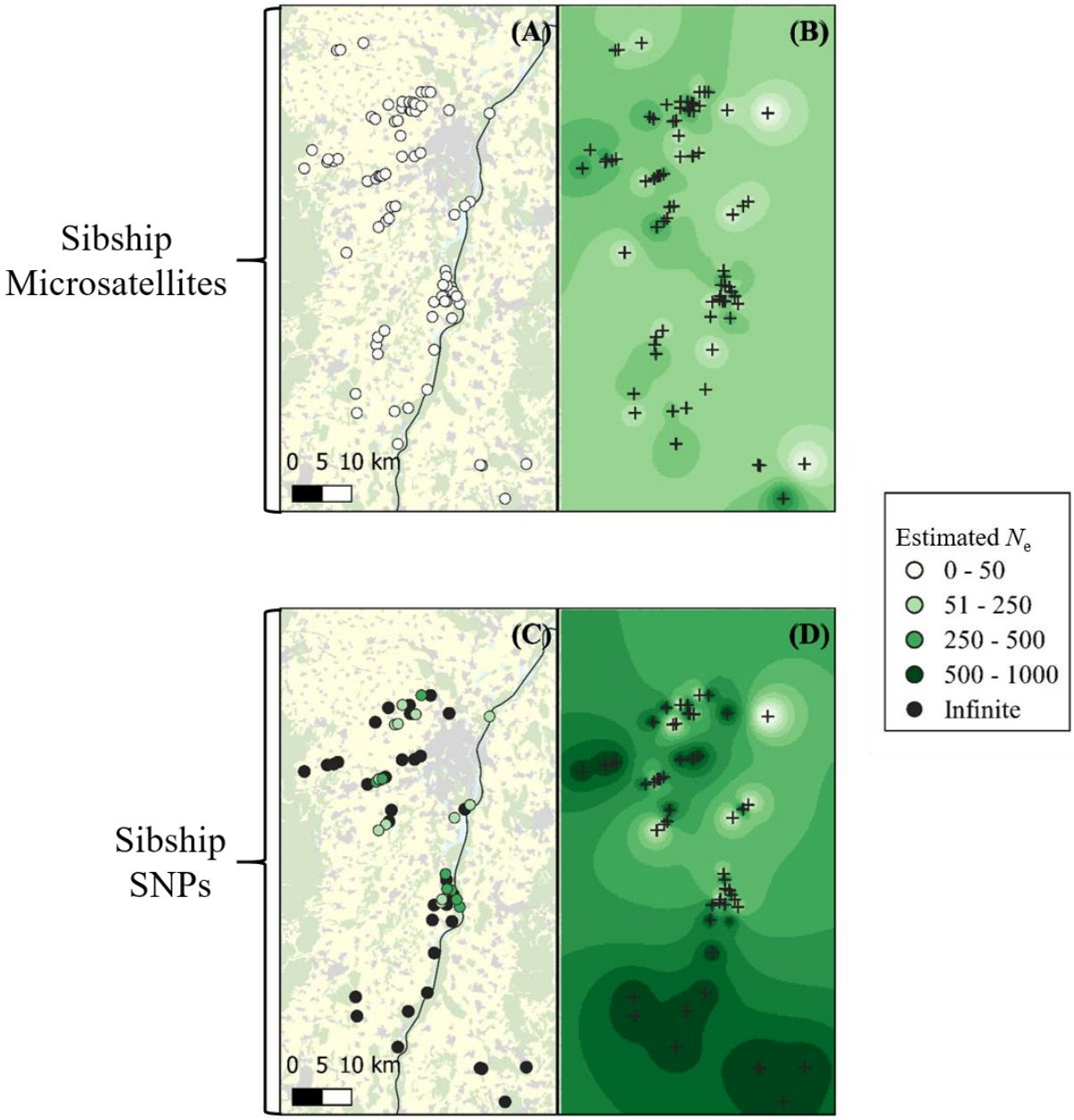


Figure S8: Spatial distribution of contemporary effective population size (N_e) estimates. The maps show N_e estimates based on inferred sibship frequencies for microsatellites **(A, B)** and SNPs **(C, D)**. Interpolation of each estimate in the studied area using a kriging method **(B, D)**. A darker green colour indicates a higher N_e estimate.

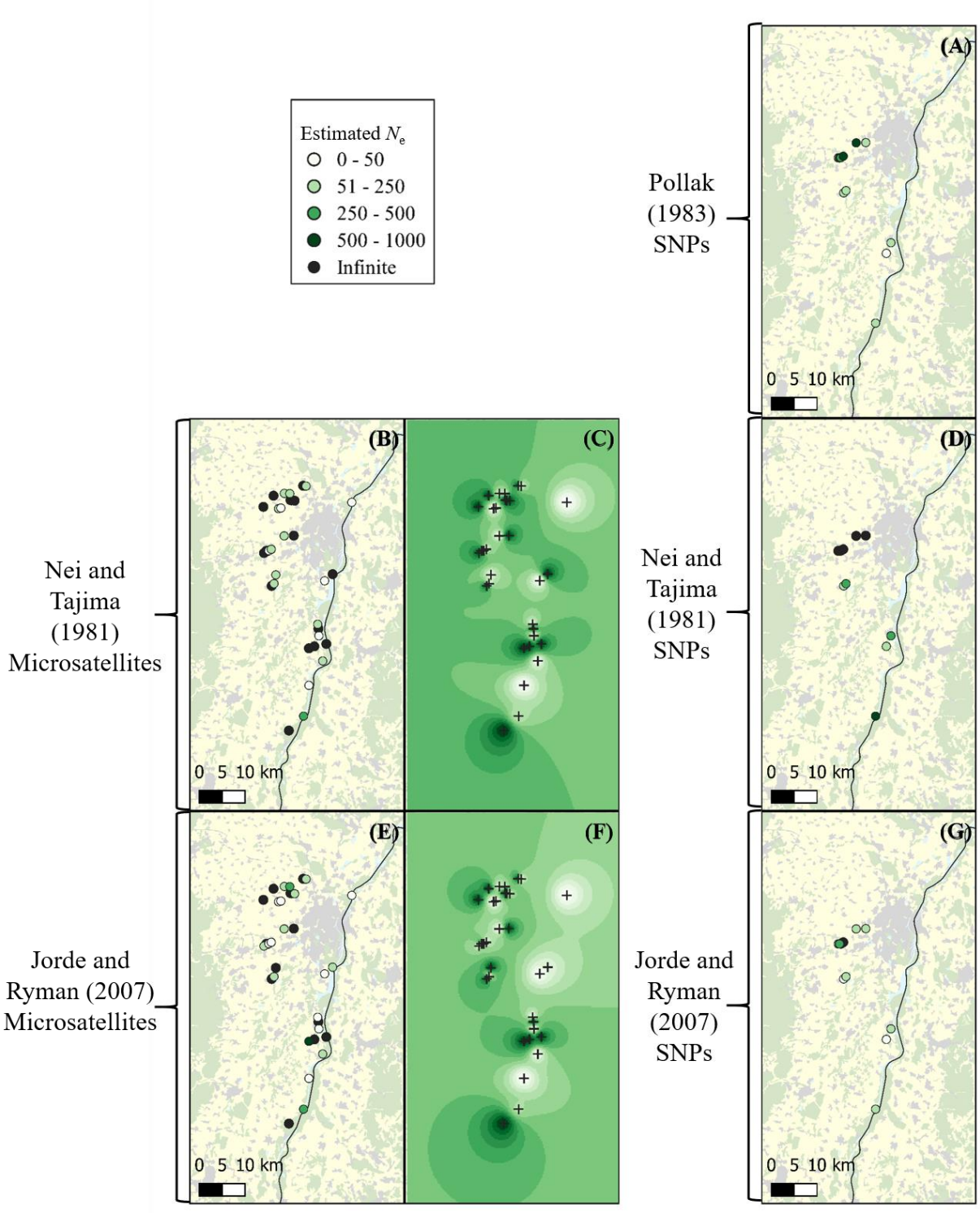


Figure S9: Spatial distribution of contemporary effective population size (N_e) estimates. The maps show the N_e temporal estimations **(A)** using Pollak's (1983) method, **(B, C, D)** using Nei and Tajima's method (1981), **(E, F, G),** using Jorde and Ryman's method (2007) using microsatellites **(B, C, E, F)** and SNPs **(A, D, G)**. Interpolation of each N_e estimate in the studied area was performed using a kriging method **(C, F)**. A darker green colour indicates a higher N_e estimate. The figure for Pollak's estimator and the microsatellite dataset is shown as main text Figure 3E-F.

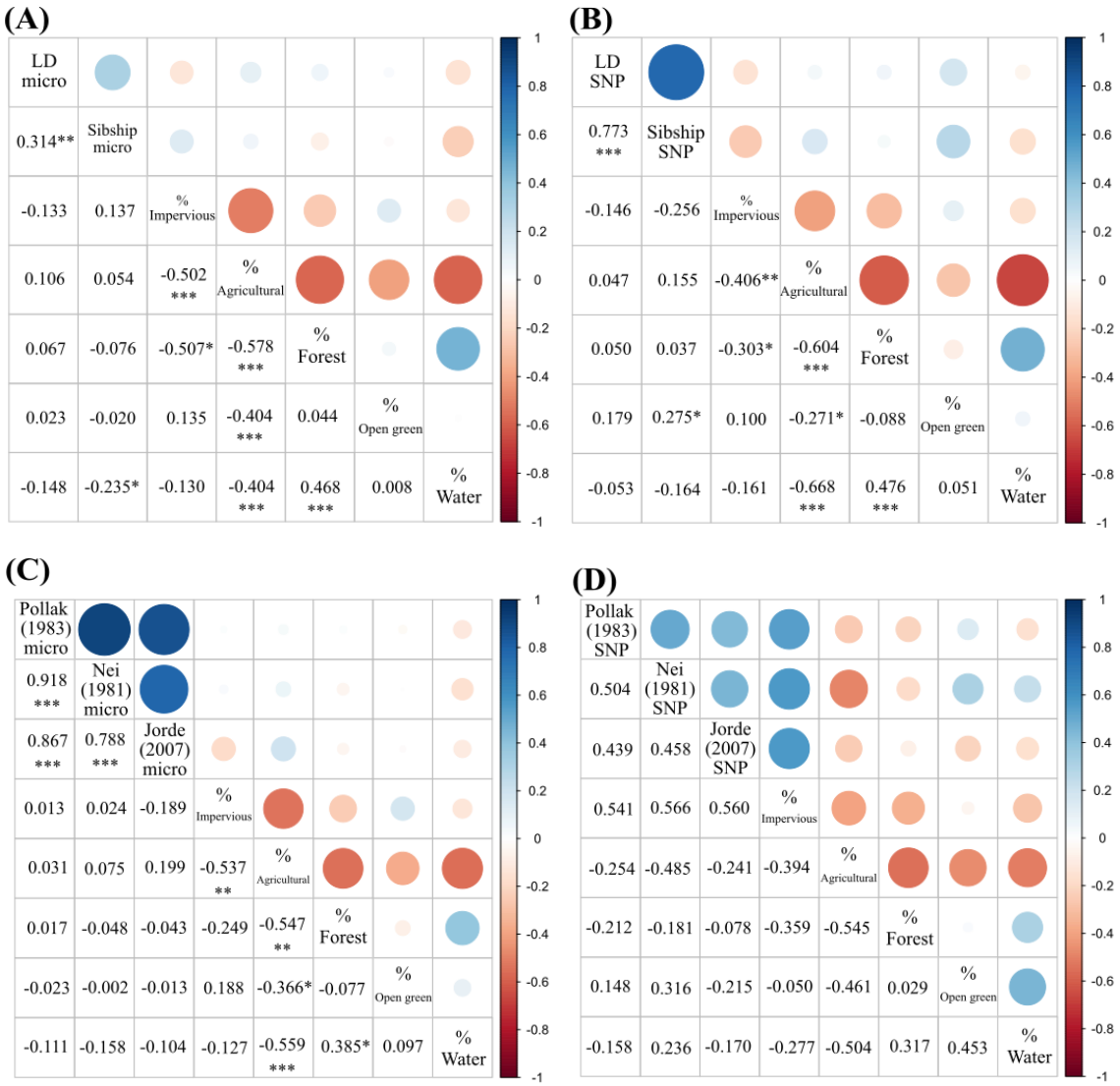


Figure S10: Graphical representation of the correlation matrix (Pearson's correlation coefficients) between effective population size (N_e) estimates and land cover type around sampling sites. **(A, B)** single-sample methods (LD: Linkage disequilibrium; Sibship: Full-Likelihood analysis based on inferred sibship frequencies among samples) using microsatellites (**A**, 77 populations) or SNPs (**B**, 56 populations). **(C, D)** Temporal methods using F_s estimator (Jorde & Ryman 2007), F_e estimator (Nei & Tajima 1981), and F_k estimator (Pollak 1983) using microsatellites (**C**, 34 populations) or SNPs (**D**, 11 populations). The colour and size of the circles (upper triangle) indicate the strength of the correlation (Red circles indicate a negative correlation; blue circles indicate a positive correlation; white circles indicate no correlation). *** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$.

Table S1: Population sampling information. Population me and geographic coordinates of the sampling sites (WGS84). Numbers of samples per year genotyped with microsatellite markers, Number of samples per year genotyped with SNP markers.

			Microsatellites		SNPs	
			Sampling Year and number of individuals genotyped using microsatellite loci		Sampling Year and number of individuals genotyped using SNP loci	
Population me	Longitude EPSG4326	Latitude EPSG4326	2021	2022	2021	2022
A4	7.544702	48.540918		18		14
A5	7.564511	48.543905	26	13	15	
A6	7.571759	48.547882	30	26	15	14
A7	7.575255	48.547574	33	29	15	15
A8	7.579292	48.548017	30	36	15	15
A9	7.586779	48.550395	31	27	13	15
A11	7.655745	48.574642	31	39	13	15
A14	7.697191	48.354712	27	23	14	15
A16	7.653206	48.218936	30	34	15	15
A20	7.669929	48.579161		13		10
Bli1	7.490202	48.18945205		30		15
Bli2	7.489811	48.21875897		36		15
Bru1	7.709551	48.3511074		19		15
CB4	7.627223	48.57519653	18	21	14	15
CCR11	7.721753	48.37259215		27		15
CCR12	7.739052	48.34622491		32		15
CCR13	7.579254	48.13881277		33		15
CCR14	7.726147	48.36515839		21		15
CCR7	7.776558	48.50051855		24		15
CHA1	7.454961	48.57377678		29		15
CHA12	7.470093	48.574551		30		15
CHA2	7.479667	48.57729641		30		15
Ceh1	7.581641	48.47838244	31	13	11	12
Ceh2	7.588845	48.48253644	25	27	12	15
Dau1	7.712416	48.37491965	26	31	15	15
Deu1	7.767138	48.09837512		28		15
Deu5	7.868078	48.09646176		23		15
Deu7	7.815028	48.04588457		38		15
Deu8	7.763198	48.09936841		22		15
Ehn1	7.563860	48.47010067	20		11	
Eng1	7.401690	48.56582089		25		11
Fri1	7.543433	48.29132721		29		
Fri2	7.547780	48.30284853		26		
Ger1	7.710864	48.39779464	20	33	13	
Ger3	7.712247	48.38878337	21	12	15	
Ger4	7.680718	48.35155451	29	30	15	
Ger6	7.674729	48.32892893	20		15	
Gox1	7.487110	48.43404213		27		

Han1	7.699740	48.35941129	25		15
Ich0	7.576146	48.18872229		10	
Jeff1	7.457562	48.57767332		21	
Kes1	7.545415	48.27738452		29	
Koh1	7.421644	48.59306066		34	
Lac1	7.703560	48.37659745		19	
Lan1	7.563761	48.31218901		34	
Rla3	7.680047	48.67161401	20		17
Rla5	7.692171	48.67147441	21	27	
Rla6	7.700949	48.67059395	29	28	12
Rlei23	7.602713	48.65540447	24	15	15
Rlei45	7.633774	48.64900886		30	
Rlei6	7.652179	48.64475511	11	29	11
Rlei7	7.654957	48.64417171	13	34	12
Rlei8	7.664481	48.64314033	17	26	16
RLK1	7.635142	48.65861658	23	18	12
RLK2	7.651268	48.65764007	30	33	13
RLK3	7.660315	48.65623515		15	
RLK4	7.664291	48.65424022		23	
RLK56	7.678233	48.65055097		18	
Rlmu2	7.626599	48.60702473		20	
Rsc1	7.739234	48.48230098	20	20	15
Rut2	7.595870	48.49964518	27	28	15
Rut3	7.605300	48.50028271		23	
S2	7.562404	48.63845095		26	
S3	7.570723	48.63470505	34	20	15
S6	7.615283	48.6293294	31	23	15
S67	7.622262	48.630331	18	20	13
S17	7.741153	48.641593	21		15
S21	7.833577	48.633175	25	11	15
V1	7.607497	48.192280	35	33	15
V2	7.674399	48.278712	29	30	15
V3	7.732873	48.358055	20	33	15
V4	7.719338	48.325091	21	34	15
Wei1	7.763958	48.494377	25	27	15
Wur1	7.705175	48.351641		32	
Zp0	7.491934	48.742628		35	
Zp1	7.499699	48.742984		30	
Zp2	7.553629	48.751473		33	

Table S2: Pairwise estimates of genetic differentiation (F_{ST}) among southern damselfly (*Coenagrion mercuriale*) populations from alternate-year cohorts (2021, 2022) in the studied area (above diagonal). Green shades indicate low values and red shades indicate higher values. Statistical significance of pairwise F_{ST} estimates is shown below the diagonal, with stars indicating $P < 0.05$ after Bonferroni correction for multiple testing. The data table can be downloaded from the link https://figshare.com/s/178f7d2429f231a1bb17

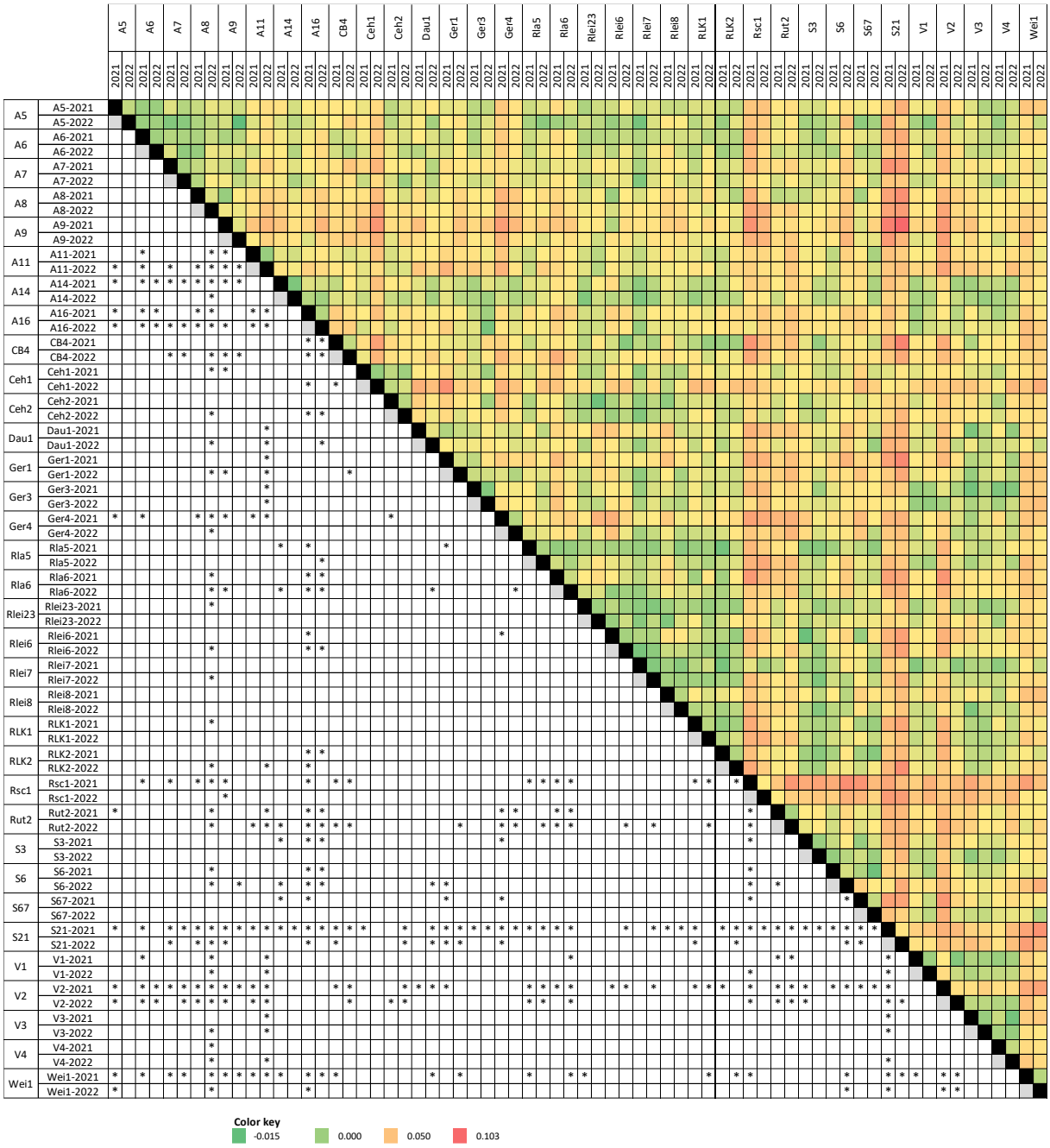


Table S3: Hierarchical *F*-statistics computed among alternate-year cohorts of southern damselflies from sample locations collected during 2021 and 2022. Significance of genetic differentiation at the Year of sampling and Population levels have been tested using permutation tests implemented in HierFstat (test.within and test.between.within commands respectively), NS : *P* Year/Total = 0.998; ** : *P* Population/Year = 0.01).

	Year of sampling (2021 & 2022)	Population	Individuals
Total	-0.0008 NS	0.0178	0.0210
Year of sampling (2021 & 2022)	0	0.0186**	0.0218
Population	0	0	0.0032

Table S4: N_e estimates from single-sample estimators, using the Linkage Disequilibrium method (Hill, 1981; Waples, 2006; Waples and Do, 2010; Microsatellites and SNPs) and sibship assignments (Wang., 2009; Microsatellites and SNPs).

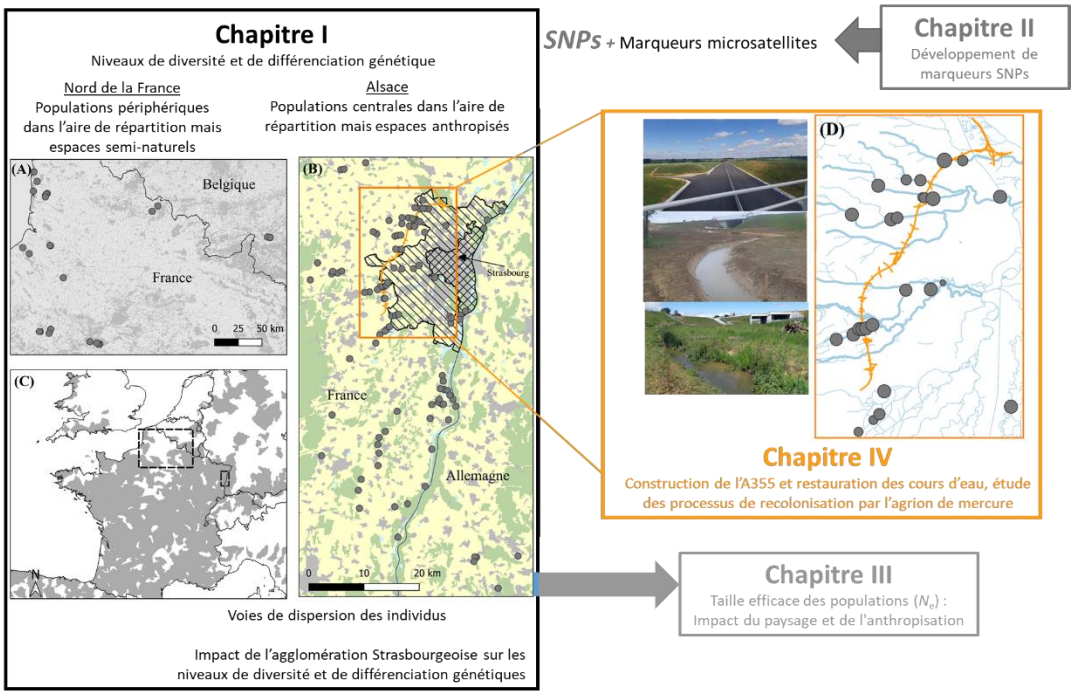
	Microsatellites					SNP				
		LD method		Sibship assignment method			LD method		Sibship assignment method	
Pop	Sample Size	Estimated *Ne*	95% CI	Estimated *Ne*	95% CI	Sample Size	Estimated *Ne*	95% CI	Estimated *Ne*	95% CI
A4-2022	18	Inf	[26.9 ; Inf]	24	[14 ; 49]	14	Inf	[Inf ; Inf]	Inf	[1, Inf]
A5-2021	26	431.4	[20.7 ; Inf]	30	[18 ; 55]	15	Inf	[30.1 ; Inf]	364	[115, Inf]
A6-2021	30	199.5	[27.4 ; Inf]	32	[19 ; 57]	15	110.6	[20.8 ; Inf]	210	[86, Inf]
A7-2021	33	Inf	[14.5 ; Inf]	30	[18 ; 52]	15	1033.2	[84.3 ; Inf]	Inf	[1, Inf]
A8-2022	36	75.9	[19.6 ; Inf]	27	[16 ; 48]	15	327.9	[63.8 ; Inf]	420	[141, Inf]
A9-2022	27	Inf	[34.2 ; Inf]	36	[22 ; 68]	15	1781.6	[53 ; Inf]	Inf	[1, Inf]
A11-2022	39	Inf	[52.8 ; Inf]	31	[19 ; 52]	15	Inf	[108.6 ; Inf]	Inf	[1, Inf]
A14-2021	27	16.1	[2.1 ; Inf]	26	[15 ; 50]	15	61.3	[14.1 ; Inf]	182	[56, Inf]
A16-2022	34	49.5	[17.3 ; Inf]	28	[16 ; 50]	15	Inf	[154.9 ; Inf]	Inf	[1, Inf]
A20-2022	13	133	[10 ; Inf]	26	[12 ; 77]	10	Inf	[208.1 ; Inf]	Inf	[1, Inf]
Bli1-2022	30	138.9	[14 ; Inf]	26	[15 ; 46]	15	Inf	[103.7 ; Inf]	Inf	[1, Inf]
Bli2-2022	36	Inf	[22.5 ; Inf]	31	[19 ; 55]	15	666.7	[41 ; Inf]	Inf	[1, Inf]
Bru1-2022	19	40.4	[5.1 ; Inf]	33	[18 ; 67]	15	Inf	[52 ; Inf]	Inf	[1, Inf]
CB4-2022	21	110.6	[13.6 ; Inf]	25	[14 ; 50]	15	Inf	[28.8 ; Inf]	Inf	[1, Inf]
CCR11-2022	27	177.8	[17.6 ; Inf]	19	[10 ; 38]	15	320.9	[39.7 ; Inf]	420	[129, Inf]
CCR12-2022	32	102.6	[18.7 ; Inf]	23	[13 ; 44]	15	Inf	[44 ; Inf]	420	[141, Inf]
CCR13-2022	33	Inf	[59.9 ; Inf]	31	[18 ; 54]	15	Inf	[45.5 ; Inf]	Inf	[1, Inf]
CCR14-2022	21	Inf	[26.2 ; Inf]	31	[18 ; 54]	15	Inf	[33 ; Inf]	Inf	[1, Inf]
CCR7-2022	24	24.6	[7.3 ; Inf]	25	[14 ; 47]	15	54.7	[16.5 ; Inf]	84	[39, 422]
CHA1-2022	29	70.4	[20.7 ; Inf]	26	[15 ; 46]	15	1903.4	[186.6 ; Inf]	Inf	[1, Inf]
CHA12-2022	30	50.6	[11.6 ; Inf]	39	[24 ; 68]	15	3796.4	[54 ; Inf]	Inf	[1, Inf]
CHA2-2022	30	Inf	[119.4 ; Inf]	25	[14 ; 46]	15	4475.5	[37.8 ; Inf]	Inf	[1, Inf]
Ceh1-2022	13	2235.9	[11.5 ; 1030.7]	39	[16 ; 16587]	12	98.1	[25.6 ; Inf]	132	[53, Inf]
Ceh2-2022	27	Inf	[32.6 ; Inf]	24	[13 ; 48]	15	344.5	[52.7 ; Inf]	Inf	[1, Inf]
Dau1-2022	31	Inf	[110.5 ; Inf]	36	[21 ; 64]	15	209	[55.5 ; Inf]	420	[145, Inf]
Deu1-2022	28	75.8	[13.7 ; Inf]	28	[16 ; 54]	15	1870.8	[62.7 ; Inf]	Inf	[1, Inf]
Deu5-2022	23	11.5	[3.1 ; 52.2]	22	[12 ; 47]	15	Inf	[132.8 ; Inf]	Inf	[1, Inf]
Deu7-2022	38	Inf	[75.7 ; Inf]	37	[23 ; 64]	15	21224.7	[134.4 ; Inf]	Inf	[1, Inf]
Deu8-2022	22	89.5	[11.2 ; Inf]	28	[16 ; 58]	15	1240.4	[96 ; Inf]	Inf	[1, Inf]
Eng1-2022	25	Inf	[28.5 ; Inf]	35	[20 ; 69]	11	Inf	[35.8 ; Inf]	Inf	[1, Inf]
Fri1-2022	29	Inf	[75.4 ; Inf]	31	[18 ; 59]	*NA*	*NA*	*NA*	*NA*	*NA*
Fri2-2022	26	82.3	[23.9 ; Inf]	27	[16 ; 49]	*NA*	*NA*	*NA*	*NA*	*NA*
Ehn1-2021	20	50.9	[15.2 ; Inf]	35	[18 ; 78]	11	Inf	[15.7 ; Inf]	220	[67, Inf]
Ger1-2021	20	Inf	[41.2 ; Inf]	25	[14 ; 50]	13	256.1	[16.2 ; Inf]	312	[101, Inf]
Ger3-2021	21	Inf	[21.3 ; Inf]	38	[21 ; 79]	15	Inf	[44.4 ; Inf]	Inf	[1, Inf]
Ger4-2021	29	Inf	[32.9 ; Inf]	23	[13 ; 44]	15	792.2	[61.1 ; Inf]	Inf	[1, Inf]
Ger6-2021	20	Inf	[21.1 ; Inf]	28	[16 ; 59]	15	724.2	[130.2 ; Inf]	Inf	[1, Inf]

Gox1-2022	27	130.3	[13.4 ; Inf]	25	[15 ; 46]	*NA*	*NA*	*NA*	*NA*	*NA*
Han1-2021	25	64.7	[10.9 ; Inf]	24	[14 ; 46]	15	128.5	[24.8 ; Inf]	210	[86, Inf]
Ich0-2022	10	71	[6.1 ; Inf]	30	[12 ; 613]	*NA*	*NA*	*NA*	*NA*	*NA*
Jeff1-2022	21	62.8	[9.6 ; Inf]	40	[23 ; 88]	*NA*	*NA*	*NA*	*NA*	*NA*
Kes1-2022	29	139.2	[28.3 ; Inf]	32	[19 ; 59]	*NA*	*NA*	*NA*	*NA*	*NA*
Koh1-2022	34	31.4	[11.3 ; 661.8]	33	[20 ; 56]	*NA*	*NA*	*NA*	*NA*	*NA*
Lac1-2022	19	Inf	[15.4 ; Inf]	34	[19 ; 75]	*NA*	*NA*	*NA*	*NA*	*NA*
Lan1-2022	34	691	[30.8 ; Inf]	26	[15 ; 46]	*NA*	*NA*	*NA*	*NA*	*NA*
Rla3-2021	20	108	[13.7 ; Inf]	20	[10 ; 41]	17	98.5	[13.6 ; Inf]	272	[111, Inf]
Rla5-2022	27	169.6	[23.5 ; Inf]	31	[18 ; 59]	*NA*	*NA*	*NA*	*NA*	*NA*
Rla6-2021	29	Inf	[16.8 ; Inf]	34	[20 ; 59]	12	Inf	[2219.4 ; Inf]	Inf	[1, Inf]
Rlei23-2021	24	47.6	[10.5 ; Inf]	31	[17 ; 63]	15	Inf	[51.2 ; Inf]	Inf	[1, Inf]
Rlei45-2022	30	60	[18.2 ; Inf]	26	[15 ; 50]	*NA*	*NA*	*NA*	*NA*	*NA*
Rlei6-2021	11	Inf	[68.8 ; Inf]	44	[18 ; 2023]	11	Inf	[26.8 ; Inf]	Inf	[1, Inf]
Rlei7-2021	13	Inf	[3.3 ; Inf]	35	[18 ; 124]	12	Inf	[46.5 ; Inf]	Inf	[1, Inf]
Rlei8-2021	17	Inf	[27.2 ; Inf]	30	[16 ; 65]	16	279.7	[24.1 ; Inf]	240	[91, Inf]
RLK1-2021	23	30.2	[11.7 ; Inf]	29	[16 ; 56]	12	Inf	[20.2 ; Inf]	132	[52, Inf]
RLK2-2021	30	54.7	[19.1 ; Inf]	35	[21 ; 66]	13	Inf	[62.9 ; Inf]	Inf	[1, Inf]
RLK3-2022	15	191.9	[8 ; Inf]	38	[20 ; 88]	*NA*	*NA*	*NA*	*NA*	*NA*
RLK4-2022	23	3520.7	[9.6 ; Inf]	27	[14 ; 53]	*NA*	*NA*	*NA*	*NA*	*NA*
RLK56-2022	18	13.3	[2.3 ; Inf]	26	[14 ; 56]	*NA*	*NA*	*NA*	*NA*	*NA*
Rlmu2-2022	20	67.4	[7.3 ; Inf]	24	[12 ; 48]	*NA*	*NA*	*NA*	*NA*	*NA*
Rsc1-2021	20	367.3	[14.1 ; Inf]	23	[13 ; 49]	15	104.8	[24.4 ; Inf]	210	[86, Inf]
Rut2-2021	27	Inf	[43.7 ; Inf]	31	[18 ; 57]	15	Inf	[18.1 ; Inf]	Inf	[1, Inf]
Rut3-2022	23	Inf	[28.7 ; Inf]	27	[14 ; 51]	*NA*	*NA*	*NA*	*NA*	*NA*
S2-2022	26	Inf	[21.2 ; Inf]	34	[19 ; 66]	*NA*	*NA*	*NA*	*NA*	*NA*
S3-2021	34	Inf	[46.8 ; Inf]	33	[20 ; 56]	15	8293.4	[27 ; Inf]	Inf	[1, Inf]
S6-2021	31	65.3	[10.3 ; Inf]	33	[19 ; 58]	15	220.5	[24.6 ; Inf]	210	[85, Inf]
S67-2021	18	141	[13 ; Inf]	27	[14 ; 68]	13	110	[25.1 ; Inf]	156	[61, Inf]
S17-2021	21	75.3	[18.6 ; Inf]	23	[13 ; 46]	15	377.6	[65.7 ; Inf]	Inf	[1, Inf]
S21-2021	25	50.7	[12.4 ; Inf]	20	[10 ; 42]	15	41.5	[15.8 ; Inf]	52	[27, 184]
V1-2021	35	57.5	[19.9 ; Inf]	29	[17 ; 51]	15	2316.1	[74.9 ; Inf]	Inf	[1, Inf]
V2-2021	29	51.8	[13.3 ; Inf]	24	[14 ; 46]	15	1658.8	[88.2 ; Inf]	Inf	[1, Inf]
V3-2021	20	96.8	[8 ; Inf]	35	[19 ; 73]	15	744.5	[56.2 ; Inf]	420	[145, Inf]
V4-2021	21	Inf	[37.9 ; Inf]	34	[20 ; 65]	15	792.6	[114.1 ; Inf]	Inf	[1, Inf]
Wei1-2021	25	Inf	[18.4 ; Inf]	29	[17 ; 52]	15	Inf	[29.6 ; Inf]	Inf	[1, Inf]
Wur1-2022	32	Inf	[29.5 ; Inf]	36	[22 ; 64]	*NA*	*NA*	*NA*	*NA*	*NA*
Zp0-2022	35	Inf	[95.4 ; Inf]	34	[21 ; 59]	*NA*	*NA*	*NA*	*NA*	*NA*
Zp1-2022	30	Inf	[68.6 ; Inf]	24	[14 ; 46]	*NA*	*NA*	*NA*	*NA*	*NA*
Zp2-2022	33	18.8	[4.7 ; 166.7]	26	[15 ; 47]	*NA*	*NA*	*NA*	*NA*	*NA*

Table S5: N_e estimates from temporal samples, using Pollak's (1983) estimator (Microsatellites and SNPs), Nei & Tajima's (1981) estimator (Microsatellites and SNPs), Jorde & Ryman's (2007) estimator (Microsatellites and SNPs). Values in square brackets represent 95% confidence intervals based on a jackknife procedure.

	Sampling Sites				Pollak (1983)				Nei & Tajima (1981)				Jorde & Ryman (2007)			
	Microsatellites		SNPs		Microsatellites		SNPs		Microsatellites		SNPs		Microsatellites		SNPs	
Pop	Sample Size 2021	Sample Size 2022	Sample Size 2021	Sample Size 2022	Estimated N_e	95% CI	Estimated N_e	95% CI	Estimated N_e	95% CI	Estimated N_e	95% CI	Estimated N_e	95% CI	Estimated N_e	95% CI
A5	26	13	*ND*	*ND*	-88.7	[28.6; ∞]			-215.8	[18.2; ∞]			52.7	[14; ∞]		
A6	30	26	15	14	-320.5	[26.1; ∞]	94.8	[40.9; ∞]	-125.5	[30; ∞]	-562.6	[76.3; ∞]	-94.9	[27.9; ∞]	63.4	[28.0; ∞]
A7	33	29	15	15	1387.9	[25.2; ∞]	4559	[89.9; ∞]	-500.3	[30.6; ∞]	-241.3	[145.5; ∞]	-274.4	[46.3; ∞]	422.9	[62.6; ∞]
A8	30	36	15	15	16.1	[1.9; ∞]	297	[69.3; ∞]	21.3	[2.6; ∞]	-137.5	[312.4; ∞]	25.1	[7.3; ∞]	-273.1	[115.8; ∞]
A9	31	27	13	15	90.5	[8.6; ∞]	416.7	[64.4; ∞]	59.2	[7.3; ∞]	-121.5	[204.6; ∞]	21.7	[6.6; ∞]	-837	[70.8; ∞]
A11	31	39	13	15	-58.1	[189.6; ∞]	89.3	[40.4; ∞]	-55.8	[400.7; ∞]	-243.2	[99.3; ∞]	-63.3	[2284.4;	98.3	[34.4; ∞]
A14	27	23	14	15	-91.9	[23.5; ∞]	43	[26.8; 94.7]	-51.2	[48.4; ∞]	80.4	[37.4; ∞]	-29.7	[70.9; ∞]	38.8	[23.0; 124.5]
A16	30	34	15	15	344.8	[14; ∞]	110.7	[79.8; ∞]	315.4	[13.4; ∞]	1011	[79.8; ∞]	360.9	[18.7; ∞]	89.3	[39.5; ∞]
CB4	18	21	14	15	116	[19; ∞]	492.4	[64.2; ∞]	99.1	[14.7; ∞]	-140.2	[131.4; ∞]	75.4	[13.5; ∞]	140.9	[40.6; ∞]
Ceh1	31	13	11	12	-33.9	[16.5; ∞]	51.4	[29.0; 170.3]	-33.5	[17.8; ∞]	174.8	[51.3; ∞]	-62.4	[21.5; ∞]	48.8	[26.1; 369.6]
Ceh2	25	27	12	15	55.3	[12.9; ∞]	76	[37.3; 1192.0]	75.6	[13.7; ∞]	319	[61.2; ∞]	104	[16.9; ∞]	77.2	[34.6; ∞]
Dau1	26	31	15	15	61.3	[12.5; ∞]	85.3	[42.5; 945.0]	44.3	[11.9; ∞]	280.1	[67.4; ∞]	23.1	[9.6; ∞]	93.3	[41.5; ∞]
Ger1	20	33			97.9	[10.2; ∞]			74.4	[7.3; ∞]			38.3	[7.8; ∞]		
Ger3	21	12			-48.6	[14.6; ∞]			-35.3	[22.3; ∞]			-26.8	[57.1; ∞]		
Ger4	29	30			-104.8	[32.1; ∞]			-199.7	[25.1; ∞]			1144.6	[30.4; ∞]		
Rla5	21	27			-184.5	[19.1; ∞]			-132.1	[20; ∞]			-115.8	[26.8; ∞]		
Rla6	29	28			57.4	[15; ∞]			51.2	[13; ∞]			51.4	[15.8; ∞]		
Rlei23	24	15			-77	[14.6; ∞]			-55.5	[22.2; ∞]			-41.3	[60.7; ∞]		
Rlei6	11	29			-236	[26.1; ∞]			-140.7	[29.6; ∞]			-339.7	[24.9; ∞]		
Rlei7	13	34			-53.3	[21.4; ∞]			-35.8	[31.1; ∞]			-28.6	[120.6; ∞]		
Rlei8	17	26			201.5	[11.1; ∞]			-1023.5	[13.4; ∞]			81.3	[14.3; ∞]		
RLK1	23	18			82.1	[9.3; ∞]			242.3	[12.3; ∞]			144.4	[16.1; ∞]		
RLK2	30	33			94.1	[14.2; ∞]			177.2	[14.2; ∞]			282.7	[18.2; ∞]		
Rsc1	20	20			21.4	[7.6; ∞]			21.2	[8.1; ∞]			16.4	[8.5; ∞]		
Rut2	27	28			52	[12.4; ∞]			164.5	[18.6; ∞]			-231.1	[38; ∞]		
S3	34	20			-55.7	[120.2; ∞]			-62.4	[72.3; ∞]			-52.7	[132.3; ∞]		
S6	31	23			151.4	[15.9; ∞]			64.4	[12; ∞]			42.9	[12.7; ∞]		
S67	18	20			33.3	[10.2; ∞]			40.5	[11.7; ∞]			45.6	[11.3; ∞]		
S21	25	11			12	[4.4; 84.9]			14.1	[4.1; ∞]			11	[4; ∞]		
V1	35	33			-107.3	[60.2; ∞]			-80	[85.1; ∞]			-59.4	[343.5; ∞]		
V2	29	30			21.5	[7.4; 173.8]			16.8	[5.8; 89.1]			10.3	[4.7; ∞]		
V3	20	33			-40.4	[67.5; ∞]			-46.8	[31.2; ∞]			-62.3	[33.9; ∞]		
V4	21	34			50.2	[11.8; ∞]			60.1	[13.3; ∞]			67.4	[17.1; ∞]		
Wei1	25	27			260.6	[33.3; ∞]			-1039.8	[29.9; ∞]			178.6	[23.2; ∞]		

Chapitre 4. Impact de la construction d'une autoroute et des travaux de restauration des cours d'eau sur la structure génétique des populations d'agrion de Mercure



Cette étude examine l'impact de la construction de l'autoroute A355 de contournement de Strasbourg, sur la présence et la structure génétique des populations de l'agrion de Mercure. La construction de cette autoroute s'est déroulée entre 2016 et 2021 et a impacté plusieurs cours d'eau historiquement habités par l'espèce. C'est pourquoi des mesures de compensation de restaurations de ces cours d'eau ont été réalisées.

Notre objectif est ici d'évaluer l'efficacité de ces mesures de restauration sur la recolonisation des cours d'eau par l'agrion de Mercure. À l'aide de marqueurs génétiques de type microsatellites et SNPs, nous avons déterminé l'effet de la destruction temporaire des habitats le long de l'autoroute, et la dynamique de recolonisation. Il s'avère que la remise en végétation des cours d'eau restaurés favorise efficacement une recolonisation au cours du temps par l'agrion de Mercure. À l'exception de certaines zones géographiques, cette autoroute ne semble pas créer d'effet barrière majeur : la divergence génétique entre populations semble être façonnée principalement par les grands réseaux hydrographiques sur un axe nord/sud. Nous allons également chercher à déterminer l'origine des individus colonisateurs, de manière à pouvoir statuer sur une dynamique de recolonisation « tout azimut » ou au contraire n'impliquant des individus fondateurs ne provenant que d'une source unique.

Ce chapitre est présenté sous la forme d'un article qui devra faire l'objet d'analyses supplémentaires en vue d'une soumission à une revue scientifique en biologie de la conservation.

The impact of a newly built bypass highway and associated watercourse restoration on population genetic structure in the southern damselfly (*Coenagrion mercuriale*)

Agathe Lévêque [1,2], Anne Duputié [1], Vincent Vignon [2,3], Fabien Duez [1], Cécile Godé [1], Clément Mazoyer [1], Jean-François Arnaud [1]

Affiliation :

1. Univ. Lille, CNRS, UMR 8198 - Evo-Eco-Paleo, F-59000 Lille, France

2. Office de Génie Ecologique (O.G.E.), F-67200 Strasbourg, France

3. ALKIOS, F-84911 Avignon, France

Abstract

The environmental impact of transport infrastructures is of particular importance in conservation biology because they may constrain the movement of individuals, leading to a reduction of gene flow among populations and an associated increase in genetic differentiation, as well as a decrease in local levels of genetic diversity. In this context, expert landscape managers are therefore asked to apply efficient measures facilitating the landscape permeability for wildlife to mitigate the adverse effects of newly built human structures. This study focuses on the potential impact of a newly built bypass highway localised in eastern France, the highway A355 bypassing the city of Strasbourg and its associated compensatory and restoration measures on patterns of genetic structure in populations of the southern damselfly (*Coenagrion mercuriale*) located on either side of this highway. Using a set of microsatellite loci and recently developed SNP markers, we showed that the levels of genetic differentiation observed between populations of the southern damselfly on either side of the A355 highway did not indicate any major barrier effect. Likewise, the patterns of genetic affiliation of populations did not reflect a distinction between populations located west or east of the highway, but instead pictured a north-to-south pattern of genetic divergence mostly explained by the main hydrogeographic networks of watercourses. Nonetheless, these preliminary results have to be considered with caution because of the time-lag between demographic processes and population genetic structure. Indeed, after recent changes disrupting the patterns of gene flow, many populations are likely not at equilibrium but rather are in a transient state or reflect a past genetic structure that do not mirror recent environmental changes. To gain further insights into the fine-scaled process of recolonisation of newly restored habitats, we plan (i) to perform assignment tests to depict where the colonisers come from, and (ii) to estimate hierarchical *F*-statistics and to perform DAPC analyses to finely visualise the patterns of population affiliation and verify the lack of any barrier effects of the A355 bypassing highway.

Introduction

The environmental impact of transport infrastructures on wild species, such as roads, railways, fenced motorways or canals, is of particular importance from a conservation biology perspective (Forman *et al.*, 2003; Holderegger and Di Giulio, 2010; Burkart *et al.*, 2016; Frankham *et al.*, 2017). Indeed, roads or highways can impact animals' survival, especially for species that move on or close to the ground (Rytwinski and Fahrig, 2015). In addition to impacting survival, they can also constrain the movement of individuals, leading to a reduction of gene flow among populations and an associated increase in genetic differentiation, as well as a decrease in local levels of genetic diversity (McDonald and St Clair, 2004; Shepard *et al.*, 2008; Ernest *et al.*, 2014). While such negative barrier effects have been demonstrated in many species (Keller and Largiadèr, 2003; Epps *et al.*, 2005; Balkenhol and Waits, 2009), the individual response of a species to these linear infrastructures can greatly vary depending (i) on the life-history traits of the species under consideration and (ii) on the nature of human-mediated pressures and on local landscape encountered (Fenderson *et al.*, 2014). Indeed, for some species, motorways do not appear to act as a physical barrier that restricts the amount of gene flow among populations and, in some cases, it can even facilitate the dispersal of individuals through human-mediated movements (Ewers and Didham, 2006, 2006; Estes-Zumpf *et al.*, 2010; Medley *et al.*, 2015; Balbi *et al.*, 2018).

The construction of such infrastructures implies local habitat destruction, often followed by compensatory measures including habitat restoration or the creation of dispersal corridors for wild species (Christie and Knowles, 2015; Frankham *et al.*, 2017). Expert landscape managers are therefore asked to apply efficient measures facilitating the landscape permeability for wildlife to mitigate the adverse effects of newly built human structures. Nonetheless, the efficiency of such measures remains variable (Clevenger *et al.*, 2001; Bellis *et al.*, 2013; Soanes *et al.*, 2018; Roni *et al.*, 2018) and, in the short-term, wild populations of species with low dispersal capabilities and strong specific habitat requirements can disappear anyway during road construction owing to local habitat destruction. This may be especially true for freshwater species depending on networks of streams that may further act as hydrogeographic corridors facilitating migration events among populations (e.g. Malard *et al.*, 2017).

Moreover, even though dragonflies and damselflies have an aquatic larval stage and live as flying adults, their populations can be negatively affected by highways because of increased mortality while crossing them (Riffell, 1999), because of water pollution (Meland *et al.*, 2019; Perron and Pick, 2020), and/or because of habitat destruction during the construction of highways.

This study focuses on the potential impact of a newly built bypass highway localised in eastern France, the highway A355 bypassing the city of Strasbourg and its associated compensatory and restoration measures on patterns of genetic structure in populations of the southern damselfly (*Coenagrion mercuriale*) located on either side of this highway. The southern damselfly is a species of great conservation interest in Europe. This species, which is essentially dependent on the topography of watercourses, is only found in the south-west of the continent and is protected at the European level, with additional national protections in France. Local regression of southern damselfly populations, particularly at the edge of the species' range, is directly related to habitat loss or deterioration associated with human-induced land use changes (Grand, 1996; Boudot, 2020). As the females lay their eggs in the hollow stems of aquatic plants, the species needs specific breeding habitats in slow flowing small streams where helophytes are present (Rouquette and Thompson, 2005; Purse and Thompson, 2009). The southern damselfly exhibits low dispersal capability associated with fine-scale genetic structure, which can make it particularly vulnerable to landscape fragmentation processes (Watts *et al.*, 2007, 2004c). The southern damselfly is widely distributed in eastern France (Alsace region), even though the occurrence of the Strasbourg Eurometropolis negatively affects the distribution of this species, decreases the levels of genetic diversity and acts as a barrier to gene flow by increasing the extent of genetic differentiation among populations (Lévêque *et al.*, 2024a).

The A355 motorway was built between 2016 and 2021 in eastern France (Figure 1). The landscape is predominantly agricultural along the north-south axis of the highway that bypasses the city of Strasbourg to the west. This highway crosses 11 rivers and streams, some of which have historically supported populations of the southern damselfly. In this context, sections of several of these watercourses (listed on Figure S1) were restored during or just after the construction of the motorway (Figure 1, Figure S1). These restoration actions included excavation, meandering, bank softening and

scrub clearing and are described on the following website: https://www.bas-rhin.gouv.fr/contenu/telechargement/27292/188657/file/DAU-SOCOS-EIA_52.pdf.

Here we aimed to study the spatial genetic structure of populations located in the direct vicinity of this highway in a directly-applied context devoted to decipher the local effectiveness of these watercourses' restoration actions and of the creation of new suitable habitats for the southern damselfly as part of the construction of the A355 highway. Using microsatellite loci and recently developed SNPs, we asked the following questions: (i) did this newly built motorway act as a dispersal barrier on southern damselfly populations located on either side of this infrastructure? (ii) did the compensatory measures accompanying the construction of the highway effectively result in the successful recolonisation of the rehabilitated watercourses by the southern damselfly? (iii) If so, where did the colonisers come from?

Materials and methods

Study sites and sampled populations

In spring 2021 and 2022, we surveyed the surrounding areas of the A355 highway to detect and sample local southern damselfly populations, within a 5 km buffer zone (Figure 1). We chose this radius as it encompasses the typical dispersal distance of southern damselfly individuals. Indeed, capture-mark-recapture studies carried out in the United Kingdom, which provided a direct estimate of the species' dispersal capacity, showed that 75 to 85% of the flying adults were recaptured within a hundred metres, with a few rare observations of dispersal for greater distances over 1 km (Purse *et al.*, 2003; Thompson and Watts, 2006; Watts *et al.*, 2007, 2004c). Indirect observations using genetic markers also suggested infrequent episodes of gene flow up to a distance of 2 km, and fairly high migration rates compared with direct observations (Watts *et al.*, 2006; Lévêque *et al.*, 2024a).

Overall, we sampled 46 populations in the study area, with 12 populations located in the direct vicinity of the highway and on watercourse sections restored during the construction of the highway (Figure 1, Figure S1). 34 populations were sampled in 2021 and 42 populations were sampled in 2022; 30 of them were sampled in both years (Table S1). Sampling effort was constant over the years and carried out on days with fair weather to ensure that adults were likely to be able to fly. Population sampling sizes ranged from 1 to 39, with a mean of 18.18 ± 10.99 (Figure S1, Figure S2, Table S1).

Overall, a total of 1,382 individuals were collected. Geographic coordinates, years of collecting, and sampling sizes of each population can be found in Table S1. Two kinds of samples were collected: whole individual body, or only the right middle leg of each individual, the latter method being non-lethal and not impacting the damselfly's survival (Fincke and Hadrys, 2001). All samples were stored in 100% ethanol until DNA extraction.

Genotyping using microsatellite and SNP loci

Individuals were genotyped using two different kinds of molecular markers, microsatellites and SNPs. A total of 1382 individuals were genotyped using ten microsatellite loci previously described by Watts *et al.* (2004a,b), among which 494 individuals, whose full body was available to ensure a sufficient quantity of DNA material, were genotyped using a set of 2,092 SNPs described in Lévêque *et al.* (2024b; Table S1, Figure S2).

For whole individuals, we removed abdomens from the rest of the body to avoid sampling intestinal microbiota, and we crushed all samples using MN Beads Type D (Macherey-Nagel) as described in Lévêque *et al.* (2024a). We extracted total genomic DNA from each whole individual sample using NucleoMag® Tissue Kit (Macherey-Nagel) according to the manufacturer's recommendations. For leg samples, we purified DNA with 12 µL NucleoMag B-beads and 12 µL of pure water and eluted in 50 µL.

We genotyped all samples using ten unlinked nuclear microsatellite loci named LIST002, LIST023, LIST034, LIST035, LIST037, LIST042, LIST062, LIST024, LIST063, and LIST066, isolated and described in Watts *et al.* (2004a,b) and following the protocol details described in Lévêque *et al.* (2024a). We also genotyped the individuals with enough tissues (*i.e.* the whole individuals) using SNPs generated using the Allegro Targeted Genotyping method (ATG developed by Tecan Genomics, Redwood City, United States). All details of the library preparation and bioinformatics analyses are detailed in Lévêque *et al* (2024b). Briefly, 12,000 SPET probes were designed to recapture 6,000 unlinked SNPs. Target genomic libraries were designed following the manufacturer guidelines for preparing the SPET libraries and were sequenced on a MiSeq System (1x150 bp, Illumina Inc., San Diego, CA, USA). The following bioinformatics analyses included cleaning, sequences' alignment to

the reference (*Bowtie2*; Langmead and Salzberg, 2012), variant calling following the guidelines implemented in GATK Best Practices, and multiple filtering to obtain reliable SNPs. After removing loci showing aberrant F_{IS} estimates and loci under putative selection, we finally obtained genotypes for 2,092 biallelic SNPs located on different RADcontigs (see Lévêque *et al*., 2024b).

Estimating population genetic differentiation

To assess the levels of genetic differentiation between pairs of populations, we calculated F_{ST} estimates according to Weir and Cockerham (1984) using *StAMPP* v1.6.3 R package (Pembleton *et al.*, 2013). Confidence intervals and statistical significance of F_{ST} values were obtained using 1,000 bootstraps over loci. To ensure sufficient statistical power, only populations with at least 8 individuals were included in the analyses. For the microsatellite dataset, we could do this separately for the two years of sampling. However, for the SNP dataset, only a low number of populations were genotyped in 2022, thus we combined the datasets of both years as distinct populations in a single analysis.

Delimitation of population boundaries and genetic discontinuities

Genetic discontinuities and grouping of genetically related populations were assessed using different complementary approaches: a spatial Principal Component Analysis (sPCA; Jombart, 2008) and a Bayesian clustering approach using fastStructure and described in Raj *et al.* (2014).

To depict spatial genetic patterns and cryptic geographical variation in allele frequencies allowing to detect population boundaries, we used a sPCA analysis, which is a spatially explicit multivariate method, implemented in the *adegenet* v.2.1.9 R library (Jombart, 2008). This analysis does not rely on any assumptions with regard to linkage disequilibrium or departures from Hardy-Weinberg equilibrium (Jombart *et al.*, 2008). We used the Delaunay triangulation graph to create the spatial network underlying the sPCA. To draw a comprehensive synthetic representation of population scores, we represented population coordinates along the first three principal components into the Red, Green, and Blue channels, as described in Menozzi *et al* (1978). This analysis was performed using both microsatellite and SNP loci.

Population structure and individual assignation to genetic clusters were further assessed for SNP genotypes using the *fastStructure* software (Raj *et al.*, 2014). This algorithm infers population structure from large SNP genotype datasets based on a variational Bayesian framework. We assumed a number of clusters ranging from K= 1 to 43 with 20 independent runs per K. We determined the optimal number of clusters using the estimators described by Puechmaille (2016). Individual assignments using different K values were then plotted using *Clumpak* (Kopelman *et al.*, 2015) as implemented in *StructureSelector* (Li and Liu, 2018).

Results

Patterns of population genetic differentiation

Using the SNPs dataset and populations sampled both in 2021 and 2022, we obtained a total of 561 pairwise population comparisons, with pairwise F_{ST} values ranging from -0.006 to 0.040. Most F_{ST} values significantly differed from 0 after bootstrapping over loci (480 out of 561; Figure 2). Most of the non-significant F_{ST} estimates occurred between populations A5 to A9, all located along the same river named “fossé de la Hardt”, or the RLK and Rlei populations located along two neighbouring streams, the “Kolsenbach” and the “Leisbach” (Figure 2).

With regards to the microsatellite dataset, pairwise F_{ST} values were higher, ranging from -0.029 to 0.130 and from -0.029 to 0.104 for populations sampled in 2021 and 2022 respectively. However, and in contrast with SNPs results, most of them were non-significant: 226 out of 276 in 2021 and 350 out of 435 in 2022. The most salient feature was that populations labelled A9-2021, A7-2021, A11-2022 and S6-2022 showed high and significant levels of genetic differentiation (Figure S3).

Delimitation of population boundaries and genetic discontinuities

The sPCA analysis revealed a predominantly north-to-south gradient in genetic distinctiveness between populations located in different rivers (Figure 3). The geographic patterns obtained with the microsatellite markers clearly pictured genetic discontinuities among the five main watercourses crossed to the north by the motorway. This spatial genetic structure mirroring watercourses geography was, however, much less pronounced using the SNP markers that broadly depicted three genetically distinct

groups of populations, again following a north to south gradient. Interestingly, the bypassing highway occurrence did not separate distinct groups of populations, except for the two southernmost watercourses crossed by the motorway that indicated genetic divergence between populations located on either side of this infrastructure. This genetic break was particularly evident using SNP markers (Figure 3D).

The Bayesian clustering analysis broadly revealed three to four genetically divergent clusters of populations (Figure 4) including: (i) populations located to the north of the motorway (Rhm1 and Rla), which surprisingly clustered with two spatially distant populations named CB4-2022 and A11; (ii) populations RLK, Rlei and S located on the tributaries watercourses of the Souffel that showed common ancestry; (iii) a third genetic grouping of populations that included sampled populations located furthest south of the motorway, that is, populations named CB2, BA, Rut2, Ceh, and Ehn1; (iv) finally, the populations located upstream of the Hardt ditch (populations named A3, A4, A5) displayed high levels of genetic admixture between the first two grouping of populations described above. Surprisingly, the population membership of the CB4 sampling site changed between 2021 and 2022, being either assigned with populations located on the tributary watercourses of the Souffel or with the populations located to the north of the motorway, respectively.

Discussion

Linear transport infrastructure like roads, trains, or canals can have a negative impact on wildlife species (Holderegger and Di Giulio, 2010; Burkart *et al.*, 2016). This type of human-made land change can act as a barrier to the movement of individuals because of increased mortality, of dispersal behavioural changes, or of the physical impossibility of crossing such particular structures. This barrier effect can lead to a reduction in gene flow among populations, which consequently drives a general increase in the levels of genetic differentiation between subdivided populations (Keller and Largiadèr, 2003; Epps *et al.*, 2005; Frankham *et al.*, 2017). However, the levels of genetic differentiation observed between populations of the southern damselfly on either side of the A355 highway did not indicate any major barrier effect from the motorway between both sides. Likewise, the patterns of genetic affiliation pictured a north-to-south pattern of genetic divergence mostly explained by the main hydrogeographic

networks of watercourses. Indeed, geographically close populations separated by the highway showed low and non-significant levels of genetic differentiation. Consequently, assignment gradients of individual probability of membership to genetically distinct grouping of populations did not reflect a distinction between populations located west or east of the highway, but instead depicted main hydrographic networks with somewhat subtle nuances when considering the upstream and downstream part of the same watercourse.

These results may be consistent with previous studies conducted on the southern damselfly (Watts *et al.*, 2004c). Indeed, using both direct demographic field surveys and indirect genetic approaches, roads may have no evident effects on individual movements and patterns of gene flow in the southern damselfly (Purse *et al.*, 2003; Thompson and Watts, 2006; Keller *et al.*, 2012; Watts *et al.*, 2004c). It should be noticed, however, that these studies were all carried out on road networks smaller than the A355 highway. In the case of this highway, the presence under the road of small watercourses and large wildlife crossings may probably facilitate the movement of individuals on either side of the infrastructure without any major risk of mortality due to road traffic over the bridges spanning the watercourses. This may explain the general patterns of low genetic differentiation for populations located on either side of the highway. As a complement to the population pairwise F_{ST} analyses, we will assess the relative importance of the position on either side of the motorway and of the watercourse on genetic differentiation by calculating hierarchical *F*-statistics using R *hierfstat* v.0.5-11 (Yang, 1998; Goudet, 2005; De Meeûs and Goudet, 2007). We will construct a two-level hierarchy, using the side of the motorway where the populations were located as the outermost level, the river on which the populations were located as the intermediate level and the sampling sites as the innermost level.

Although these first results are encouraging for the persistence of the populations located around the highway, care must be taken when interpreting the results we obtained. Indeed, in order to fully understand the effects of such infrastructure, it is necessary to obtain and compare information before and after its appearance (Matocq and Villablanca, 2001; Balkenhol and Waits, 2009). However, we were not able to obtain historical samples collected before the construction of the highway as a reference point for comparison with present and future patterns of genetic structure after the highway construction (see Balkenhol and Waits, 2009; Karamanlidis *et al.*, 2012). In addition, the emergence of genetic

distinctiveness among populations is closely linked to the long-lasting effect of genetic drift and the lack of homogenising migration events after a change in population connectivity. However, it is well theoretically established that population genetic features do not attain their new equilibrium states immediately after a change in connectivity among subdivided populations, but instead respond with some delay or time lag (Crow and Aoki, 1984; Slatkin and Arter, 1991; Slatkin, 1993; revue in Epps and Keyghobadi, 2015). Therefore, with recent changes disrupting the patterns of gene flow, many populations are likely not at equilibrium but rather are in a transient state or reflect a past genetic structure that do not mirror recent environmental changes (Slatkin, 1993; Hutchison and Templeton, 1999). Several generations may therefore be needed between the appearance of a disturbance and the time at which the genetic response of a population becomes detectable or negligible compared with its new equilibrium value, a process known as 'time lag' (Bolliger *et al.*, 2014; Epps and Keyghobadi, 2015). This well-known lag between current demographic processes and population genetic structure makes it challenging to correctly interpret the effect of contemporary land changes (Epps and Keyghobadi, 2015; Frankham *et al.*, 2017) and to conclude, in our case study, whether this newly built bypass highway shape the likelihood of dispersal events. The construction of the A355 motorway is relatively recent and was only completed in year 2021 for the restoration of some watercourses like the "Fossé de la Hardt" or the "Muehlbach". As such, long-term survey and monitoring of southern damselfly populations should therefore be carried out to confirm the lack of any barrier effects observed.

In addition, and before establishing conservation baselines, other confounding effects could also occur at the level of the watercourses crossed by the highway. Indeed, sampled in more recently restored watercourses, the populations located along the 'Fossé de la Hardt' showed a genetic break consistent with a barrier effect lowering the extent of gene flow. The A9 population, although being located at a short distance not exceeding 600 metres from the other populations, exhibited high levels of pairwise genetic differentiation and a high amount of genetic admixture between two distinct clusters of populations. Further, although populations named A3, A4, and A5 did not show significant levels of genetic differentiation with the sites located at the edge of the motorway, these populations exhibited a high amount of genetic admixture between distinct clusters of populations.

These differences could be due to historical barrier effects related to other landscape features beyond that of the highway. Previous studies carried out on the southern damselfly indeed showed that the amount of gene flow among populations could be influenced by a large number of landscape features, with a negative effect from high relief and the closure of watercourses by forests, and a positive effect from aquatic areas and meadows (Purse, 2001; Purse *et al.*, 2003; Watts *et al.*, 2006; Keller *et al.*, 2012; Lorenzo-Carballa *et al.*, 2015).

In the watercourse sections restored and re-naturalised during the construction of the highway (*i.e.* Muhlbaechel, Kolsenbach, Leisbach, La Souffel, Musaubach, Muehlbach, Fossé de la Hardt, Altofer Arm), an increase in the occurrence of southern damselfly individuals was observed between the years 2021 and 2022. This may suggest an evident beneficial effect of river restoration measures favouring a successful recolonisation process in this species. This gradual increase in the occurrence of southern damselfly individuals may certainly be explained by the progressive development of the vegetation following the compensatory measures, including habitat restoration, by landscape-expert managers. As a matter of fact, the compensatory measures related to the bypass highway construction successfully restored the straight watercourses by recreating meanders, creating alternating current speeds and opening up the watercourses. These operations encourage the development of helophyte plants which are essential for oviposition, perching and roots to protect the larvae of southern damselfly (Rouquette and Thompson, 2005, 2007). In this respect, a study carried out before and after restoration indeed showed a beneficial ecological effect of stream meandering on the conservation of the southern damselfly (Beaune and Sellier, 2021).

However, we observed an asymmetrical evolution of the southern damselfly occurrence at the level of the watercourses crossed by the motorway, with some watercourses mainly characterised by an upstream occurrence (e.g. watercourse named Musaubach) and the others by a downstream occurrence (e.g. watercourse named Leisbach). More detailed analyses of genetic affiliations at the level of the watercourses may reveal the actual patterns of recolonisation processes. We plan to further visualise patterns of population affiliation using a discriminant analysis of principal components (DAPC; Jombart *et al*., 2010), implemented in the R package *adegenet* v2.1.9 (Jombart, 2008; Jombart and Ahmed, 2011;). This analysis represents a complementary approach to Bayesian clustering and sPCA as it does

not rely on any population model or spatial prior. Furthermore, to gain further insights into the fine-scaled process of recolonisation of newly restored habitats after punctual destruction, we additionaly plan to perform assignment tests following maximum-likelihood methods similar to those described in Rannala and Mountain (1997) and Cornuet *et al.* (1999). Indeed, the turnover of local endangered populations owing to human-induced extinction and recolonisation processes is an important determinant of a metapopulation genetic structure, allowing microevolutionary processes to occur which is not possible in large undivided and panmictic populations. Deciphering the actual origin of individual colonisers is thus fundamental from a conservation and evolutionary point of view, because it was theoretically and empirically shown that extinction/recolonisation events can enhance opportunities for both genetic drift and gene flow, depending either on the number of colonists relative to the number of recurrent migrants, and on whether the colonists arise from a single source or many different sources (Wade and McCauley, 1988; McCauley, 1993; Ingvarsson *et al.*, 1997; Haag *et al.*, 2005).

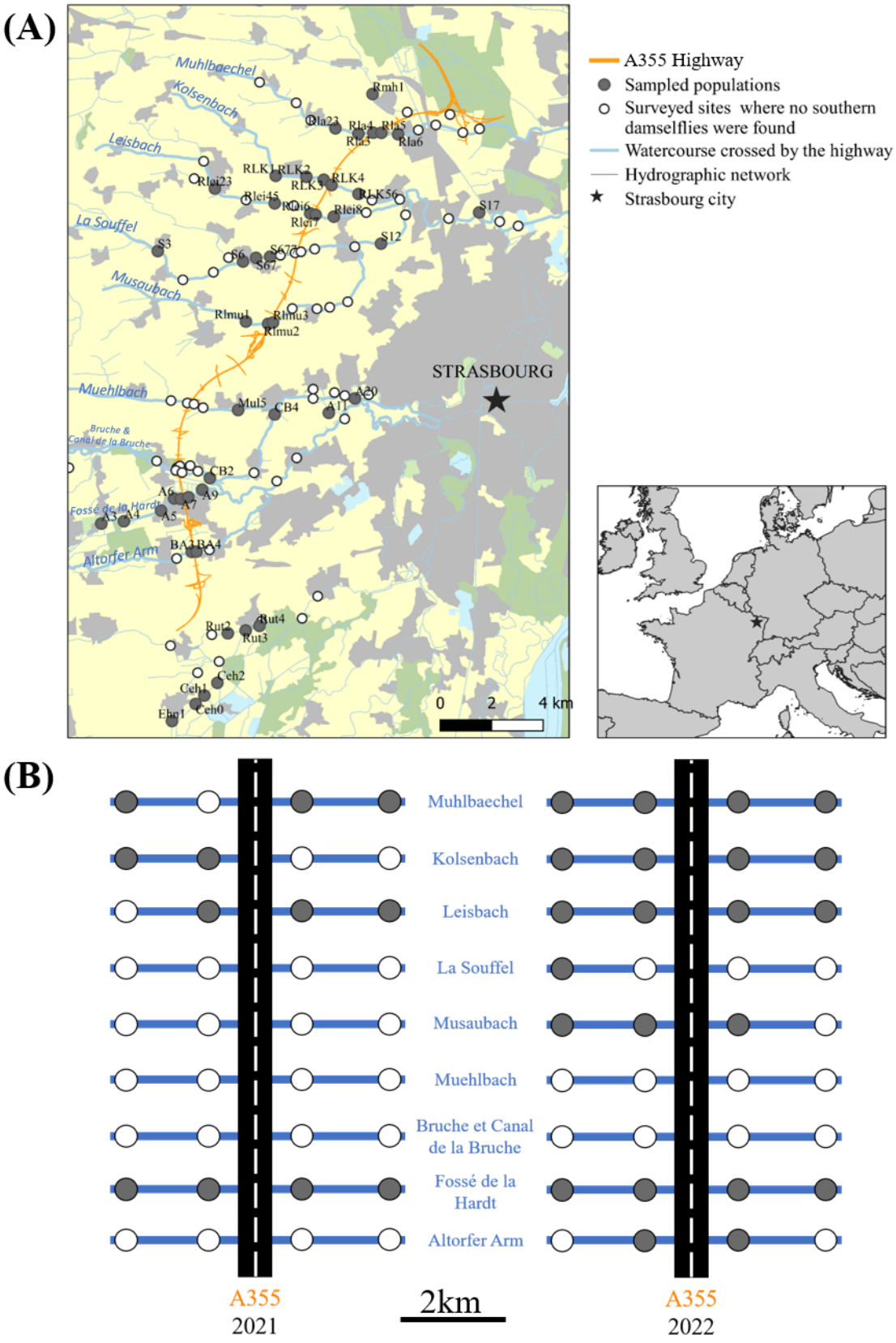


Figure 1: Location of 46 populations of southern damselflies (*Coenagrion mercuriale*) in the direct vicinity of the A355 highway **(A)**. Each grey point represents one sampled population, and white dots

represent sites surveyed but where no southern damselflies were found. Main watercourses were taken from COPERNICUS Land Monitoring Service (2019: EU-Hydro). Land cover was simplified from Corine Land Cover Edition 2018. **(B)** Scheme illustrating the sampling of sites in the direct vicinity of the highway in 2021 and 2022. As above, grey point represents one successfully sampled population and white dots represent sites which lacked of southern damselfly occurrence.

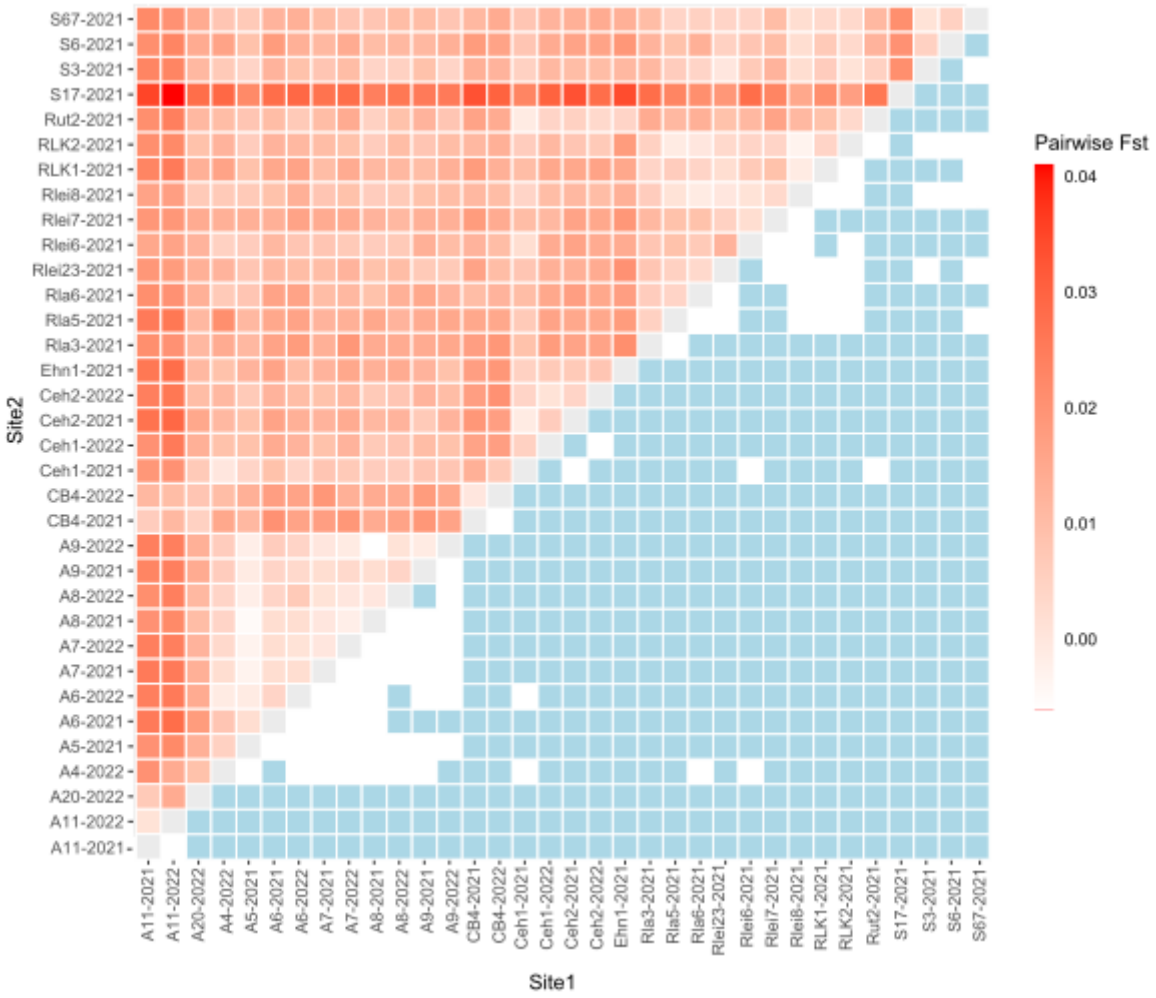


Figure 2: Matrix of pairwise F_{ST} estimates calculated using SNPs data (above diagonal) and their associated statistical significance assessed using 1000 bootstraps over loci (below the diagonal, *P*-values below 0.05 appear in blue).

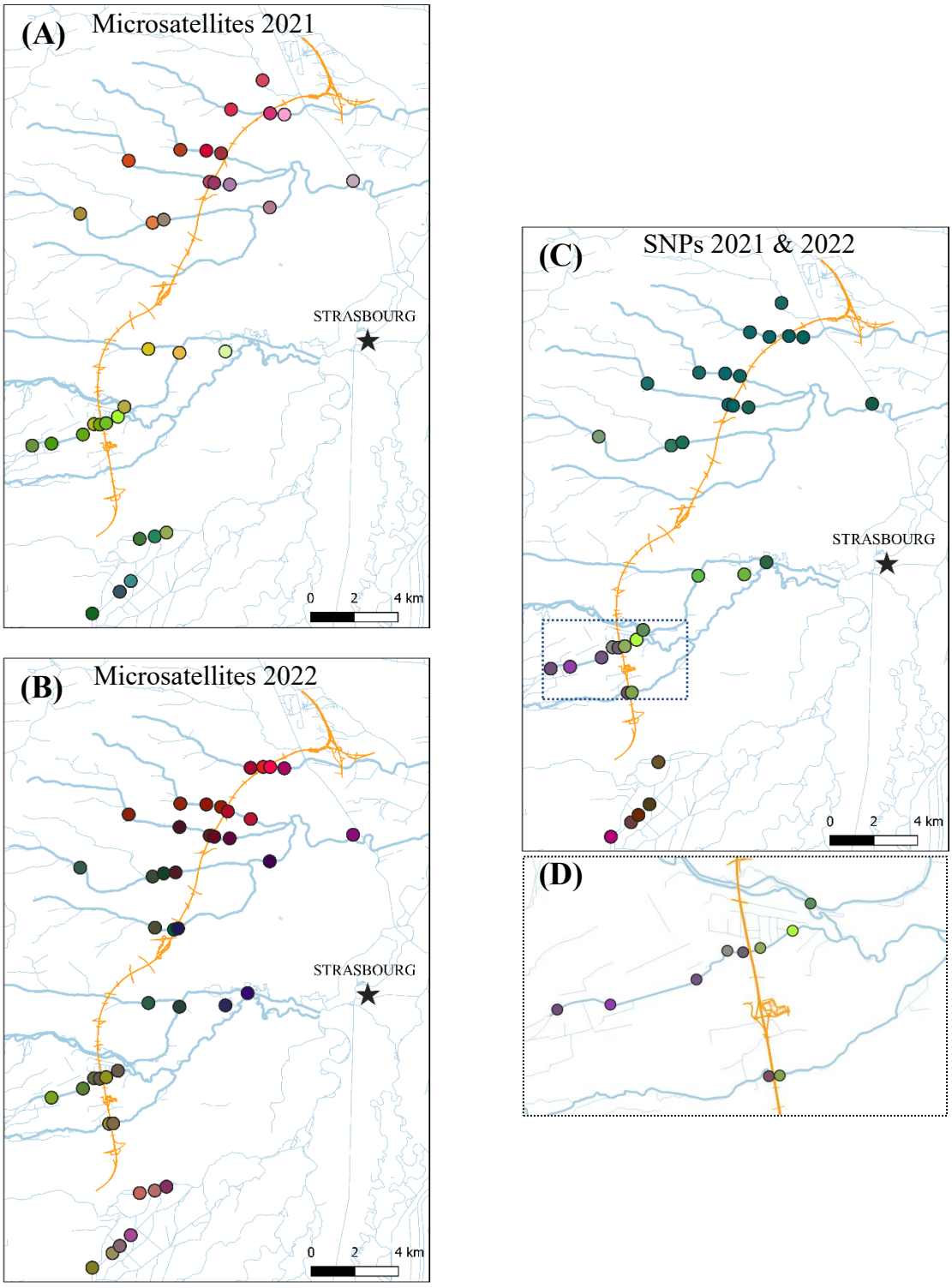


Figure 3: sPCA analyses depicting population genetic affinities around the A355 highway using multilocus microsatellite genotypes of individuals sampled in 2021 **(A)**, in 2022 **(B)**, and SNPs genotypes in 2021 and 2022 **(C)**. A zoom on two specific rivers crossed by the motorway is also indicated **(D)**. Population colours reflect their coordinates on the first three axes of the sPCA, assigned

to a red, green, and blue channels respectively. Main watercourses are also represented. Watercourses from COPERNICUS Land Monitoring Service, 2019: EU-Hydro.

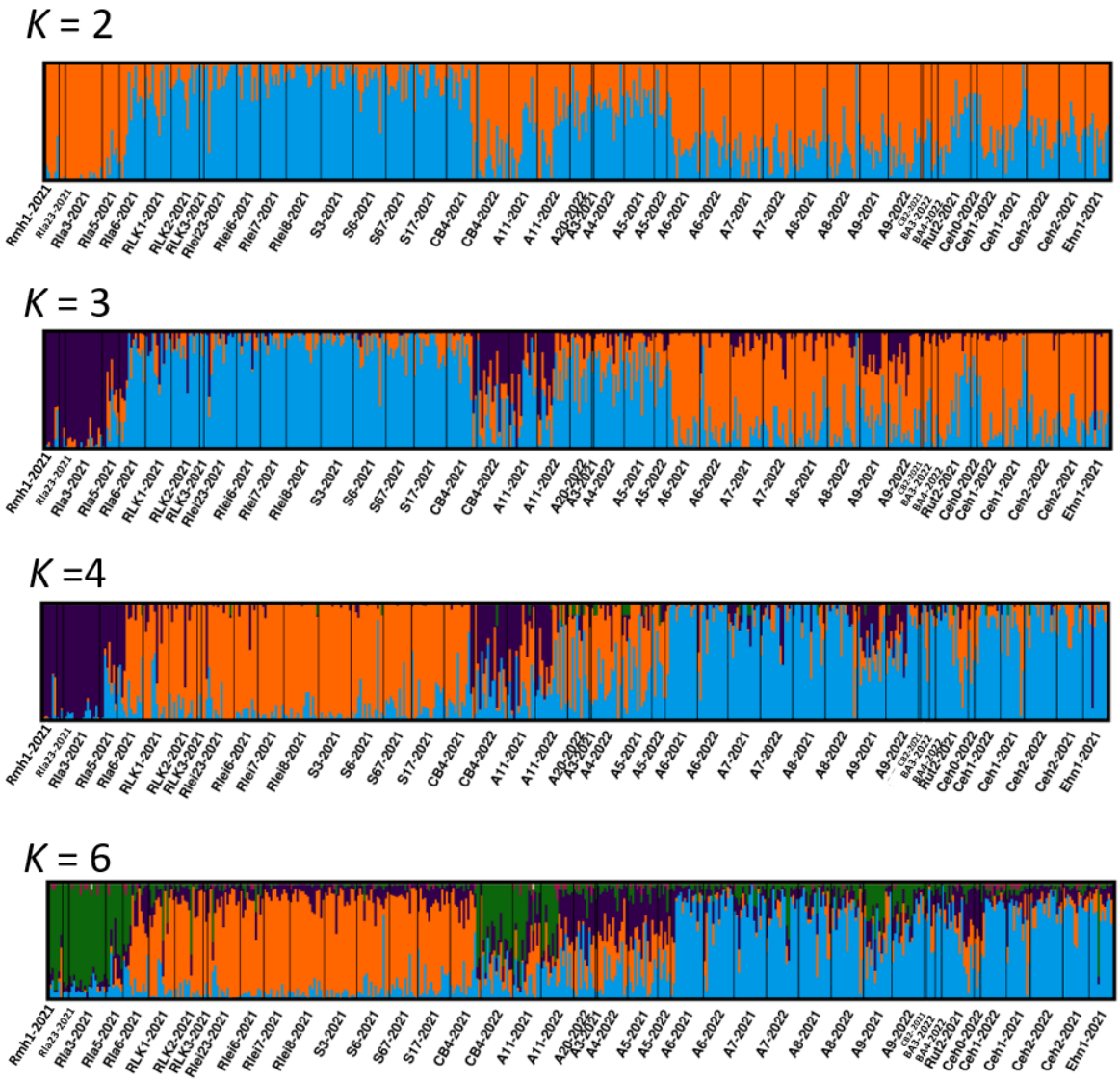


Figure 4: Bayesian clustering results by considering a number *K* of populations being of 2, 3, 4 or 6. Each individual is represented by a vertical bar, with colour indicating the estimated membership proportion to each of the *K* considered groups. The different populations are delimitated with black lines and are labelled at the bottom of each panel. Populations are arranged from North to South.

References

Balbi M, Ernoult A, Poli P, Madec L, Guiller A, Martin M-C, *et al.* (2018). Functional connectivity in replicated urban landscapes in the land snail (*Cornu aspersum*). *Molecular ecology* **27**: 1357–1370.

Balkenhol N, Waits LP (2009). Molecular road ecology: exploring the potential of genetics for investigating transportation impacts on wildlife. *Molecular ecology* **18**: 4151–4164.

Beaune D, Sellier Y (2021). Stream restorations with meanders increase dragonfly and damselfly diversity and abundance, including an endangered species. *Journal for Nature Conservation* **60**: 125950.

Bellis MA, Griffin CR, Warren P, Jackson SD (2013). Utilizing a multi-technique, multi-taxa approach to monitoring wildlife passageways in southern Vermont. *Oecologia Australis* **17**: 111–128.

Bolliger J, Lander T, Balkenhol N (2014). Landscape genetics since 2003: status, challenges and future directions. *Landscape Ecology* **29**: 361–366.

Boudot J-P (2020). *Coenagrion mercuriale*. The IUCN Red List of Threatened Species.

Burkart S, Gugerli F, Senn J, Kuehn R, Bolliger J (2016). Evaluating the functionality of expert-assessed wildlife corridors with genetic data from roe deer. *Basic and Applied Ecology* **17**: 52–60.

Christie MR, Knowles LL (2015). Habitat corridors facilitate genetic resilience irrespective of species dispersal abilities or population sizes. *Evolutionary Applications* **8**: 454–463.

Clevenger AP, Chruszcz B, Gunson K (2001). Drainage culverts as habitat linkages and factors affecting passage by mammals. *J Appl Ecol* **38**: 1340–1349.

Cornuet JM, Piry S, Luikart G, Estoup A, Solignac M (1999). New methods employing multilocus genotypes to select or exclude populations as origins of individuals. *Genetics* **153**: 1989–2000.

Crow JF, Aoki K (1984). Group selection for a polygenic behavioral trait: estimating the degree of population subdivision. *Proceedings of the National Academy of Sciences of the USA* **81**: 6073–6077.

De Meeûs T, Goudet J (2007). A step-by-step tutorial to use HierFstat to analyse populations hierarchically structured at multiple levels. *Infection, Genetics and Evolution* **7**: 731–735.

Epps CW, Keyghobadi N (2015). Landscape genetics in a changing world: disentangling historical and contemporary influences and inferring change. *Molecular Ecology* **24**: 6021–6040.

Epps CW, Palsbøll PJ, Wehausen JD, Roderick GK, Ramey RR, McCullough DR (2005). Highways block gene flow and cause a rapid decline in genetic diversity of desert bighorn sheep. *Ecology letters* **8**: 1029–1038.

Ernest HB, Vickers TW, Morrison SA, Buchalski MR, Boyce WM (2014). Fractured genetic connectivity threatens a southern California Puma (*Puma concolor*) population. *PloS One* **9**: e107985.

Estes-Zumpf WA, Rachlow JL, Waits LP, Warheit KI (2010). Dispersal, gene flow, and population genetic structure in the pygmy rabbit (*Brachylagus idahoensis*). *Journal of Mammalogy* **91**: 208–219.

Ewers RM, Didham RK (2006). Confounding factors in the detection of species responses to habitat fragmentation. *Biological Reviews* **81**: 117–142.

Fenderson LE, Kovach AI, Litvaitis JA, O'Brien KM, Boland KM, Jakubas WJ (2014). A multiscale analysis of gene flow for the New England cottontail, an imperiled habitat specialist in a fragmented landscape. *Ecology and Evolution* **4**: 1853–1875.

Fincke OM, Hadrys H (2001). Unpredictable offspring survivorship in the damselfly, *Megaloprepus coerulatus*, shapes parental behavior, constrains sexual selection, and challenges traditional fitness estimates. *Evolution* **55**: 762–772.

Forman RT, Sperling D, Bissonette JA, Clevenger AP, Cutshall CD, Dale VH, *et al.* (2003). *Road ecology: science and solutions*, Island Press. Washington, DC, USA.

Frankham R, Ballou JD, Ralls K, Eldridge MDB, Dudash MR, Fenster CB, *et al.* (2017). *Genetic management of fragmented animal and plant populations*, Oxford University Press. Oxford, UK.

Goudet J (2005). Hierfstat, a package for R to compute and test hierarchical *F*-statistics. *Molecular Ecology Notes* **5**: 184–186.

Grand D (1996). *Coenagrion mercuriale (Charpentier, 1840). Background information on invertebrates of the Habitats Directive and the Bern Convention, Part II: Mantodea, Odonata, Orthoptera, and Arachnida.*, Council of Europe Publishing. Strasbourg, France.

Haag C, Riek M, Hottinger JW, Pajunen V, Ebert D (2005). Genetic diversity and genetic differentiation in *Daphnia metapopulations* with subpopulations of known age. *Genetics* **170**: 1809–1820.

Holderegger R, Di Giulio M (2010). The genetic effects of roads: a review of empirical evidence. *Basic and Applied Ecology* **11**: 522–531.

Hutchison DW, Templeton AR (1999). Correlation of pairwise genetic and geographic distance measures: inferring the relative influences of gene flow and drift on the distribution of genetic variability. *Evolution* **53**: 1898–1914.

Ingvarsson PK, Olsson K, Ericson L (1997). Extinction-recolonization dynamics in the mycophagous beetle *Phalacrus substriatus*. *Evolution* **51**: 187–195.

Jombart T (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**: 1403–1405.

Jombart T, Ahmed I (2011). adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics* **27**: 3070–3071.

Jombart T, Devillard S, Balloux F (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics* **11**: 94.

Jombart T, Devillard S, Dufour A-B, Pontier D (2008). Revealing cryptic spatial patterns in genetic variability by a new multivariate method. *Heredity* **101**: 92–103.

Karamanlidis AA, Straka M, Drosopoulou E, de Gabriel Hernando M, Kocijan I, Paule L, *et al.* (2012). Genetic diversity, structure, and size of an endangered brown bear population threatened by highway construction in the Pindos Mountains, Greece. *European Journal of Wildlife Researchs* **58**: 511–522.

Keller I, Largiadèr CR (2003). Recent habitat fragmentation due to major roads leads to reduction of gene flow and loss of genetic variability in ground beetles. *Proceedings of the Royal Society of London* **270**: 417–423.

Keller D, Van Strien MJ, Holderegger R (2012). Do landscape barriers affect functional connectivity of populations of an endangered damselfly? *Freshwater Biology* **57**: 1373–1384.

Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I (2015). Clumpak: a program for identifying clustering modes and packaging population structure inferences across *K*. *Molecular Ecology Resources* **15**: 1179–1191.

Langmead B, Salzberg SL (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**: 357–359.

Lévêque A, Duputié A, Vignon V, Duez F, Godé C, Vanappelghem C, Arnaud J-F (2024a). Contrasting patterns of spatial genetic structure in endangered southern damselfly (*Coenagrion mercuriale*) populations facing habitat fragmentation and urbanisation. *bioRxiv* 2024.05.02.592171; doi: https://doi.org/10.1101/2024.05.02.592171

Lévêque A, Arnaud J-F, Vignon V, Mazoyer C, Godé C, Duputié A (2024b). Pitfalls and recommendations for large-scale SNP genotyping in a non-model endangered species: the southern damselflies (*Coenagrion mercuriale*) as a case study. *bioRxiv* 2024.05.02.592185; doi: https://doi.org/10.1101/2024.05.02.592185

Li Y, Liu JX (2018). StructureSelector: A web-based software to select and visualize the optimal number of clusters using multiple methods. *Molecular Ecology Resources* **18**: 176–177.

Lorenzo-Carballa MO, Ferreira S, Sims AM, Thompson DJ, Watts PC, Cher Y, *et al.* (2015). Impact of landscape on spatial genetic structure and diversity of *Coenagrion mercuriale* (Zygoptera:Coenagrionidae) in northern France. *Freshwater Science* **34**: 1065–1078.

Malard F, Capderrey C, Churcheward B, Eme D, Kaufmann B, Konecny-Dupré L, *et al.* (2017). Geomorphic influence on intraspecific genetic differentiation and diversity along hyporheic corridors. *Freshwater biology* **62**: 1955–1070.

Matocq MD, Villablanca FX (2001). Low genetic diversity in an endangered species: recent or historic pattern? *Biological Conservation* **98**: 61–68.

McCauley DE (1993). Genetic consequences of extinctions and recolonization in fragmented habitats. In *Biotic Interactions and Global Change* (ed. P. M. Kareiva, J. G. Kingsolver & R. B. Huey), pp. 217-233. Sunderland, MA: Sinauer Associates Inc.

McDonald WR, St Clair CC (2004). The effects of artificial and natural barriers on the movement of small mammals in Banff National Park, Canada. *Oikos* **105**: 397–407.

Medley KA, Jenkins DG, Hoffman EA (2015). Human-aided and natural dispersal drive gene flow across the range of an invasive mosquito. *Molecular Ecology* **24**: 284–295.

Meland S, Gomes T, Petersen K, Håll J, Lund E, Kringstad A, *et al.* (2019). Road related pollutants induced DNA damage in dragonfly nymphs (Odonata, Anisoptera) living in highway sedimentation ponds. *Scientific Reports* **9**: 16002.

Menozzi P, Piazza A, Cavalli-Sforza L (1978). Synthetic maps of human gene frequencies in Europeans. *Science* **201**: 786–792.

Pembleton LW, Cogan NOI, Forster JW (2013). StAMPP: an R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Molecular Ecology Resources* **13**: 946–952.

Perron MAC, Pick FR (2020). Water quality effects on dragonfly and damselfly nymph communities: A comparison of urban and natural ponds. *Environmental Pollution* **263**: 114472.

Puechmaille SJ (2016). The program structure does not reliably recover the correct population structure when sampling is uneven: subsampling and new estimators alleviate the problem. *Molecular ecology resources* **16**: 608–627.

Purse BV (2001). The ecology and conservation of the Southern Damselfly (*Coenagrion mercuriale*). Doctoral dissertation. University of Liverpool.

Purse BV, Hopkins GW, Day KJ, Thompson DJ (2003). Dispersal characteristics and management of a rare damselfly. *Journal of Applied Ecology* **40**: 716–728.

Purse BV, Thompson DJ (2009). Oviposition site selection by *Coenagrion mercuriale* (Odonata: Coenagrionidae). *International Journal of Odonatology* **12**: 257–273.

Raj A, Stephens M, Pritchard JK (2014). fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics* **197**: 573–589.

Rannala B, Mountain JL (1997). Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences of the USA* **94**: 9197–9201.

Riffell SK (1999). Road mortality of dragonflies (Odonata) in a Great Lakes coastal wetland. *The Great Lakes Entomologist* **32**: 63-74.

Roni P, Åberg U, Weber C (2018). A review of approaches for monitoring the effectiveness of regional river habitat restoration programs. *North American Journal of Fisheries Management* **38**: 1170–1186.

Rouquette JR, Thompson DJ (2005). Habitat associations of the endangered damselfly, *Coenagrion mercuriale*, in a water meadow ditch system in southern England. *Biological Conservation* **123**: 225–235.

Rouquette JR, Thompson DJ (2007). Roosting site selection in the endangered damselfly, *Coenagrion mercuriale*, and implications for habitat design. *Journal of Insect Conservation* **11**: 187–193.

Rytwinski T, Fahrig L (2015). The impacts of roads and traffic on terrestrial animal populations. In: *Handbook of road ecology*, pp 237–246.

Shepard DB, Kuhns AR, Dreslik MJ, Phillips CA (2008). Roads as barriers to animal movement in fragmented landscapes. *Animal Conservation* **11**: 288–296.

Slatkin M (1993). Isolation by distance in equilibrium and non-equilibrium populations. *Evolution* **47**: 264–279.

Slatkin M, Arter HE (1991). Spatial autocorrelation methods in population genetics. *The American Naturalist* **138**: 499–517.

Soanes K, Taylor AC, Sunnucks P, Vesk PA, Cesarini S, van der Ree R (2018). Evaluating the success of wildlife crossing structures using genetic approaches and an experimental design: Lessons from a gliding mammal. *Journal of Applied Ecology* **55**: 129–138.

Thompson DJ, Watts PC (2006). The structure of the *Coenagrion mercuriale* populations in the New Forest, southern England. *Forests and Dragonflies*: 239–258.

Wade MJ, McCauley DE (1988). Extinction and recolonization: their effects on the genetic differentiation of local populations. *Evolution* **42**: 995–1005.

Watts PC, Rouquette JR, Saccheri IJ, Kemp SJ, Thompson DJ (2004c). Molecular and ecological evidence for small-scale isolation by distance in an endangered damselfly, *Coenagrion mercuriale*. *Molecular Ecology* **13**: 2931–2945.

Watts PC, Rousset F, Saccheri IJ, Leblois R, Kemp SJ, Thompson DJ (2007). Compatible genetic and ecological estimates of dispersal rates in insect (*Coenagrion mercuriale*: Odonata: Zygoptera) populations: analysis of 'neighbourhood size' using a more precise estimator. *Molecular Ecology* **16**: 737–751.

Watts PC, Saccheri IJ, Kemp SJ, Thompson DJ (2006). Population structure and the impact of regional and local habitat isolation upon levels of genetic diversity of the endangered damselfly *Coenagrion mercuriale* (Odonata: Zygoptera). *Freshwater Biology* **51**: 193–205.

Watts PC, Thompson DJ, Kemp SJ (2004b). Cross-species amplification of microsatellite loci in some European zygopteran species (Odonata: Coenagrionidae). *International Journal of Odonatology* **7**: 87–96.

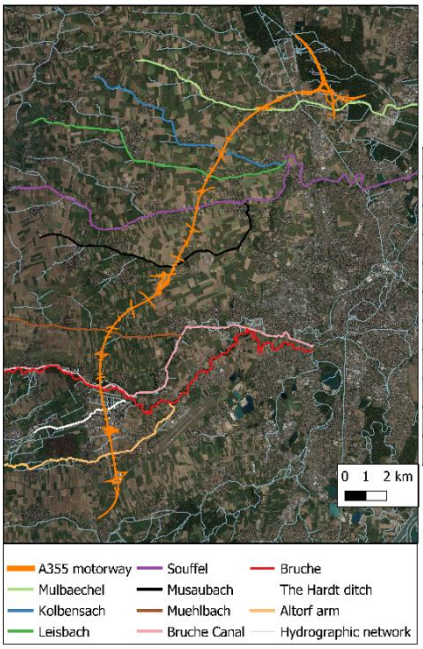
Watts PC, Wu JH, Westgarth C, Thompson DJ, Kemp SJ (2004a). A panel of microsatellite loci for the southern damselfly, *Coenagrion mercuriale* (Odonata: Coenagrionidae). *Conservation Genetics* **5**: 117–119.

Weir BS, Cockerham CC (1984). Estimating *F*-Statistics for the analysis of population structure. *Evolution* **38**: 1358–1370.

Yang R (1998). Estimating hierarchical *F*-statistics. *Evolution* **52**: 950–956.

Supporting information

(A)



Watercourse		Field surveyed of local abundance			Sampling abundance		Construction end date
		2016	2019	2020	2021	2022	
Muhlbaechel		4	48	23	21	30	12/2020
Kolsenbach		3	33	13	2	38	09/2020
Leisbach		2	19	43	24	63	09/2020
La Souffel		8	0	0	0	0	11/2020
Musaubach					0	24	11/2020
Muehlbach					0	0	04/2021
Bruche canal					0	0	
La Bruche					0	0	
Fossé de la Hardt		~20	32	20	63	65	05/2021
Altorfer Arm					0	7	10/20

Figure S1: Historical monitoring of southern damselfly (*Coenagrion mercuriale*) populations and watercourse restorations 200m around the A355 **(A);** Photograph of a site sampled in 2021 **(B, C)** and 2022 **(D, E)**. Watercourse drainage was visible in 2021 **(B)**. Re-meandering work was carried out **(C, E)**.

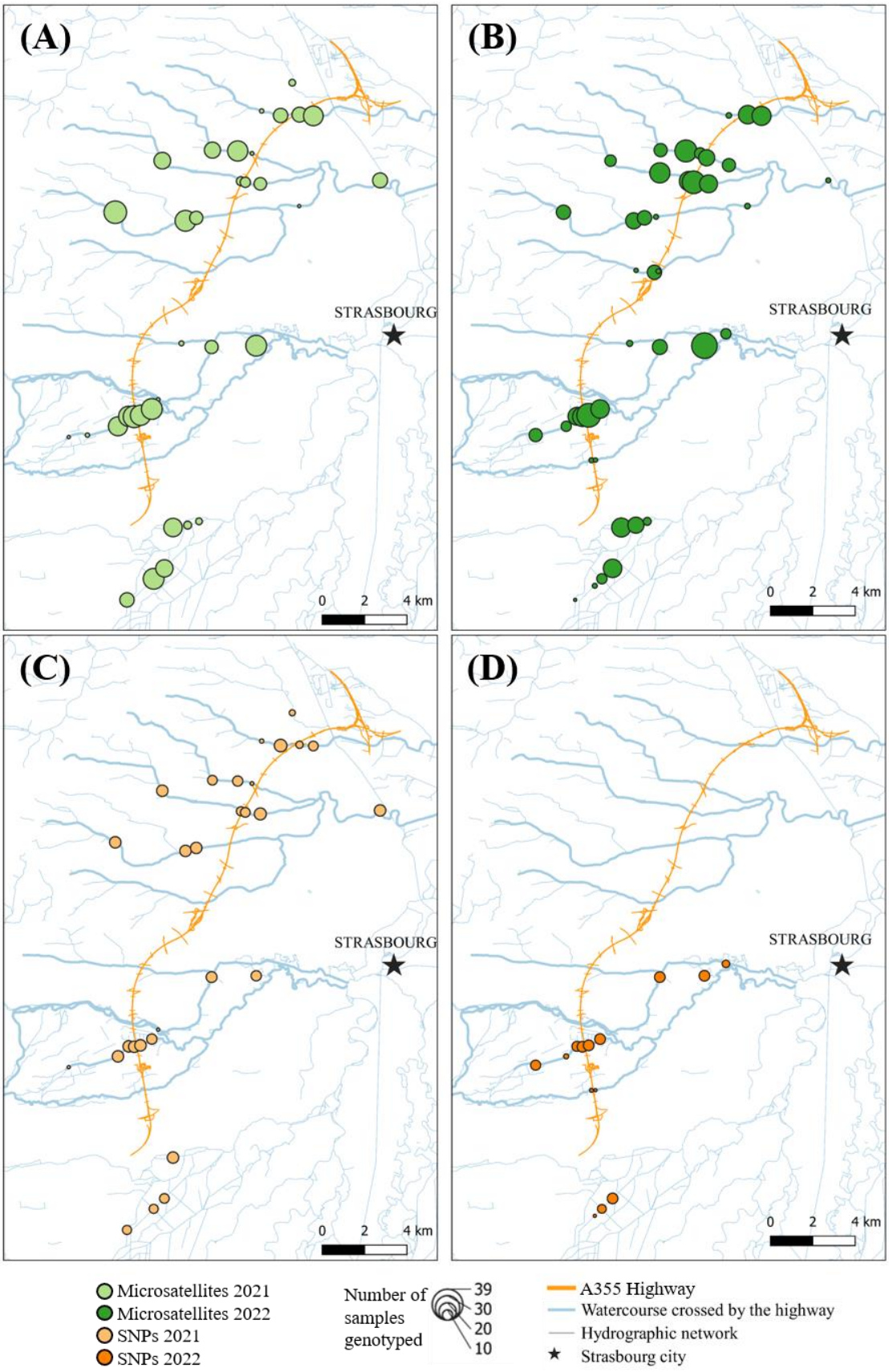


Figure S2: Number of southern damselfly (*Coenagrion mercuriale*) samples genotyped with microsatellite (**A, B**) or SNP (**C, D**) markers in 2021 (**A, C**) and 2022 (**B, D**).

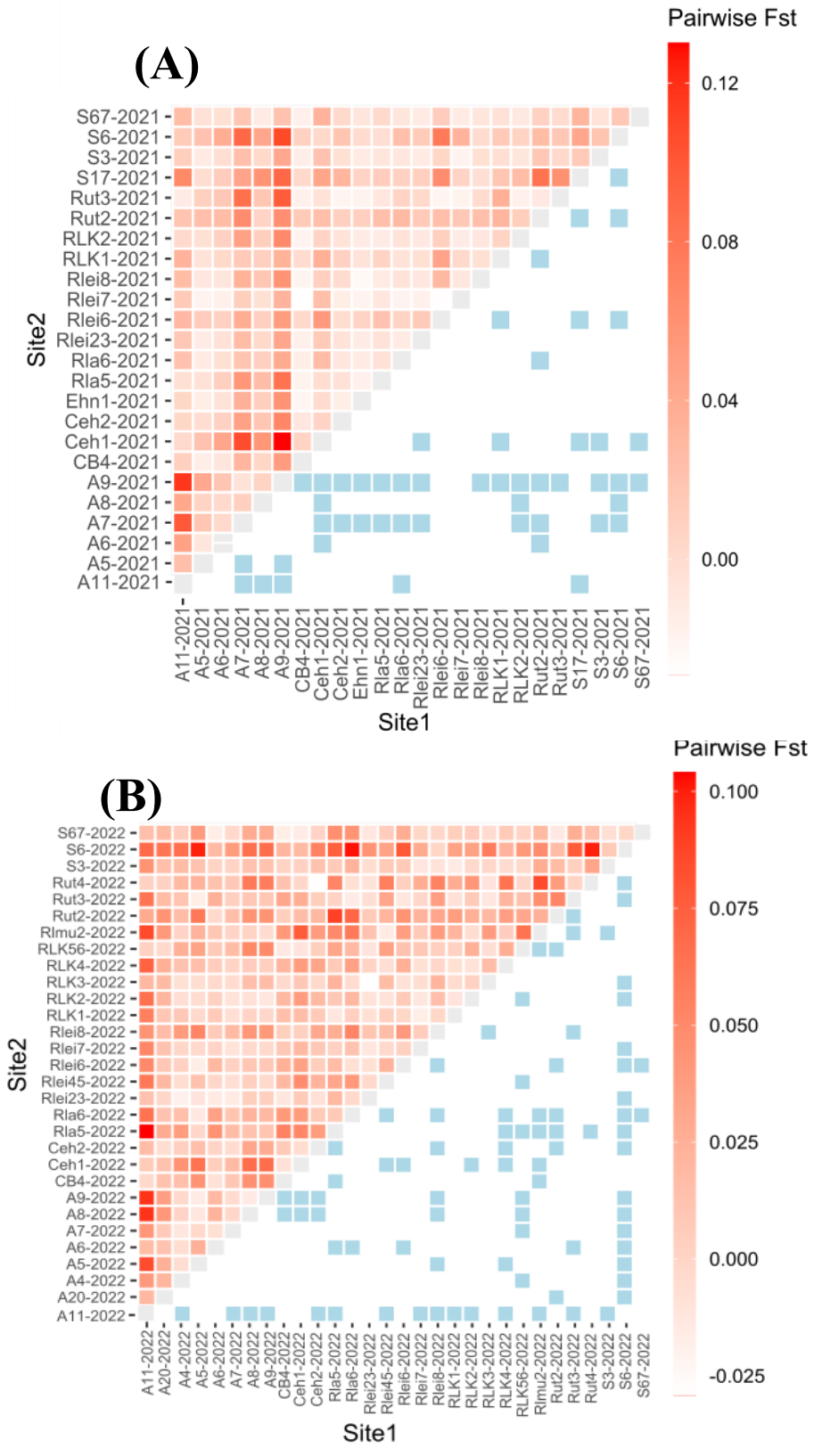


Figure S3: Matrices of pairwise F_S estimates between southern damselfly (*Coenagrion mercuriale*) populations and calculated using microsatellite data from populations sampled in 2021 **(A)** or 2022 **(B)** and their associated statistical significance assessed using 1000 bootstraps over loci.

Table S1: Southern damselfly (*Coenagrion mercuriale*) population sampling information. Population name and geographical coordinates of the sampling sites (WGS84). Number of samples per year genotyped with microsatellite markers, Number of samples per year genotyped using SNP markers.

			Microsatellites		SNPs	
			Sampling Year and number of individuals genotyped		Sampling Year and number of individuals genotyped	
Population name	Longitude (WGS84 EPSG4326)	Latitude (WGS84 EPSG4326)	2021	2022	2021	2022
Rmh1	7.688742	48.685199	7		6	
Rla23	7.668104	48.673981	3		3	
Rla3	7.680047	48.671614	20	5	17	
Rla4	7.687878	48.671665		3		
Rla5	7.692171	48.671474	21	27	8	
Rla6	7.700949	48.670594	29	28	12	
RLK1	7.635142	48.658617	23	18	12	
RLK2	7.651268	48.657640	30	33	13	
RLK3	7.660315	48.656235	2	15	2	
RLK4	7.664291	48.654240		23		
RLK56	7.678233	48.650551		18		
Rlei23	7.602713	48.655404	24	15	15	
Rlei45	7.633774	48.649009		30		
Rlei6	7.652179	48.644755	11	29	11	
Rlei7	7.654957	48.644172	13	34	12	
Rlei8	7.664481	48.643140	17	26	16	
S3	7.570723	48.634705	34	20	15	
S6	7.615283	48.629329	31	23	15	
S67	7.622262	48.630331	18	20	15	
S677	7.629693	48.630449		4		
S12	7.688485	48.632729	1	5		
S17	7.741153	48.641593	21	4	15	
Rlmu1	7.614947	48.608271		3		
Rlmu2	7.626599	48.607025		20		
Rlmu3	7.629157	48.607368		4		

Mul5	7.608110	48.577487	4	6		
CB4	7.627223	48.575197	18	21	14	15
A11	7.655745	48.574642	31	39	13	15
A20	7.669929	48.579161		13		10
CB2	7.591170	48.554269	2		1	
A3	7.532710	48.540547	1		1	
A4	7.544702	48.540918	3	18		14
A5	7.564511	48.543905	28	13	15	6
A6	7.571759	48.547882	30	26	15	14
A7	7.575255	48.547574	33	29	15	15
A8	7.579292	48.548017	30	36	15	15
A9	7.586779	48.550395	31	27	13	15
BA3	7.579429	48.528832		4		4
BA4	7.581946	48.528855		3		3
Rut2	7.595870	48.499645	27	28	15	
Rut3	7.605300	48.500283	9	23		
Rut4	7.612676	48.501593	7	9		
Ceh0	7.576856	48.475591		4		3
Ceh1	7.581641	48.478382	31	13	11	12
Ceh2	7.588845	48.482536	25	27	12	15
Ehn1	7.563860	48.470101	20	1	11	

Synthèse des résultats et perspectives

Synthèse des résultats

Dans le contexte actuel de changements environnementaux, l'étude de l'évolution génétique des populations soumises à des pressions anthropiques croissantes, telles que l'agriculture intensive, l'urbanisation et la construction d'infrastructure, gagne en importance et popularité. Cet essor s'explique par des raisons à la fois fondamentales en raison d'évolutions génétiques rapides observées dans ces environnements perturbés, mais aussi par des raisons pratiques liées à la conservation des espèces. L'objectif principal de ce travail de thèse était d'apporter des données empiriques sur la structuration génétique spatiale des populations d'agrion de Mercure (*Coenagrion mercuriale*), permettant de répondre à des questions de biologie de la conservation au sein d'une zone intégrant de nombreuses pressions anthropiques. Ce travail se plaçait dans le contexte directement appliqué de l'étude de l'impact de la construction de l'autoroute A355 et des restaurations de cours d'eau réalisées dans le cadre de mesures compensatoires.

Contrairement aux travaux précédemment menés chez l'agrion de Mercure, cette thèse ne se concentre pas sur une zone située en limite d'aire de répartition de l'espèce avec des populations potentiellement très isolées, mais se focalise sur une zone d'étude localisée plus au centre de l'aire de distribution géographique de l'espèce, avec un réseau hydrographique dense abritant de nombreuses populations d'agrion de Mercure. En effet, les campagnes d'échantillonnages réalisées en 2021 et 2022 ont permis d'échantillonner un très grand nombre de populations (104 en Alsace et 24 dans les Hauts-de-France) et de génotyper un nombre considérable d'individus (3 797 individus génotypés à l'aide de marqueurs microsatellites et 1 080 individus génotypés à l'aide de marqueurs SNP). Ainsi, les données génétiques produites dans le cadre de ces travaux de thèse combinent, pour la première fois chez l'agrion de Mercure, des données génétiques classiques impliquant des locus microsatellites et des données génomiques (SNPs) avec les avantages et limites inhérents à chacun des types de marqueurs moléculaires. Ce jeu de données a permis des estimations des niveaux de diversité génétique intra-population et des estimations de tailles efficaces de populations, deux paramètres clés de conservation des populations. Il a également fourni des premiers éléments sur les flux de gènes et les niveaux de connectivité entre populations d'agrion de Mercure, ceci dans une région intégrant de nombreuses problématiques paysagères et avec le contexte directement appliqué de l'autoroute A355 nouvellement construite dans la région et des restaurations connexes de cours d'eau (Figure 23).

Le Chapitre II, premier volet technique de la thèse, a permis le développer 2 092 nouveaux marqueurs moléculaires de type SNP chez l'agrion de Mercure, une espèce sans génome de référence disponible. Ces marqueurs ont été développés en complément de marqueurs microsatellites déjà mis au point pour l'espèce (Watts *et al.*, 2004ab) et utilisés dans l'ensemble des études génétiques menées chez l'espèce jusqu'à présent (Watts *et al.*, 2005, 2006, 2004c; Thompson & Watts, 2006; Watts, *et al.*, 2007a;

Watts *et al.*, 2007b; Watts & Thompson, 2012; Keller *et al.*, 2012; Keller & Holderegger, 2013; Lorenzo-Carballa *et al.*, 2015). Nous avons montré que l'approche ddRADseq de représentation réduite du génome permettait d'identifier *de novo* des centaines de milliers de marqueurs SNPs chez une espèce non modèle et de construire des RADloci de référence (Peterson *et al.*, 2012; Narum *et al.*, 2013). La méthode d'enrichissement ciblé (SPET) nous a ensuite permis de génotyper 1 080 individus avec 2 092 SNPs. Compte tenu des 6 000 SNPs initialement ciblés chez 1 920 individus, ce chapitre a permis de mettre en évidence les limites et problématiques associées au développement de marqueurs SNPs chez un très grand nombre d'individus pour une espèce non modèle. Bien que moins nombreux qu'initialement prévu, nous avons montré que les marqueurs SNPs nouvellement développés fournissaient des informations précieuses sur les niveaux de diversité génétique, mais surtout la différenciation génétique et la structure génétique spatiale des populations. En effet, ces marqueurs génomiques ont permis d'identifier, dans des populations localisées dans les Hauts-de-France, une structure génétique à fine échelle et des patrons de flux de gènes qui n'étaient pas détectables à l'aide de marqueurs microsatellites.

Ces populations du nord de la France constituaient un point de comparaison avec notre zone d'étude principale localisée dans la région strasbourgeoise. Le chapitre I a ainsi permis de mettre en évidence que les populations d'agrion de Mercure échantillonnées en Alsace présentaient des niveaux de diversité génétique élevés et des niveaux de différenciation génétique faibles par rapport aux niveaux observés jusqu'à présent dans les régions plus périphériques de l'aire de répartition de l'espèce, que ce soit dans les Hauts-de-France ou au Royaume-Uni (Watts *et al.*, 2005, 2004c; Lorenzo-Carballa *et al.*, 2015). Au sein de la région strasbourgeoise, abritant de nombreuses populations d'agrion de Mercure avec un niveau de connectivité élevé entre populations, nous avons montré que les flux de gènes s'opéreraient à travers un patron d'isolement par la distance classique. L'étude des voies de dispersion associées au niveau de différenciation génétique entre populations a permis de révéler que les flux de gènes entre populations n'étaient pas restreints strictement aux habitats de reproduction de l'espèce le long des cours d'eau, mais pouvaient s'opérer avec des mouvements à travers terre (Figure 24 A,B,C). De plus, cette première étude a mis en lumière, au sein de la région strasbourgeoise, la présence de plusieurs zones de dispersion facilitée ou restreinte par rapport à un modèle nul d'isolement par la distance. Parmi ces zones, l'Eurométropole de Strasbourg impacte négativement la diversité génétique des populations, et l'on observe au voisinage de l'agglomération une augmentation des niveaux de différenciation génétique : cette zone urbaine représente donc un frein aux flux géniques entre populations.

L'impact de l'Eurométropole de Strasbourg a été retrouvé dans le Chapitre III qui se focalisait sur l'estimation des tailles efficaces de populations, un autre paramètre clé de conservation. Ce chapitre a permis de confirmer qu'il n'y avait pas d'isolement génétique entre cohortes temporelles chez l'agrion de Mercure, suggérant une plasticité dans le trait semi-voltin de l'espèce (voir également Thelen, 1992; Watts & Thompson, 2012). Avec l'application de multiples méthodes d'estimations des tailles efficaces

de populations et de deux types de marqueurs moléculaires, nous avons illustré la difficulté inhérente à l'estimation précise des tailles efficaces de population pour des espèces sauvages. Toutefois, les résultats obtenus suggèrent des dynamiques de populations contrastées au niveau des populations retrouvées autour du tracé de l'autoroute A355, reflétant des processus de recolonisation impliquant potentiellement des populations sources et des populations puits de part et d'autre de cette nouvelle l'infrastructure routière.

Les analyses préliminaires réalisées dans le chapitre IV ont permis de mettre en évidence une recolonisation progressive des sites reméandrés et restaurés lors de la construction de l'autoroute, ainsi qu'une absence d'effet barrière actuel de l'autoroute. Les patrons exacts de recolonisation de ces sites restent encore à définir précisément.

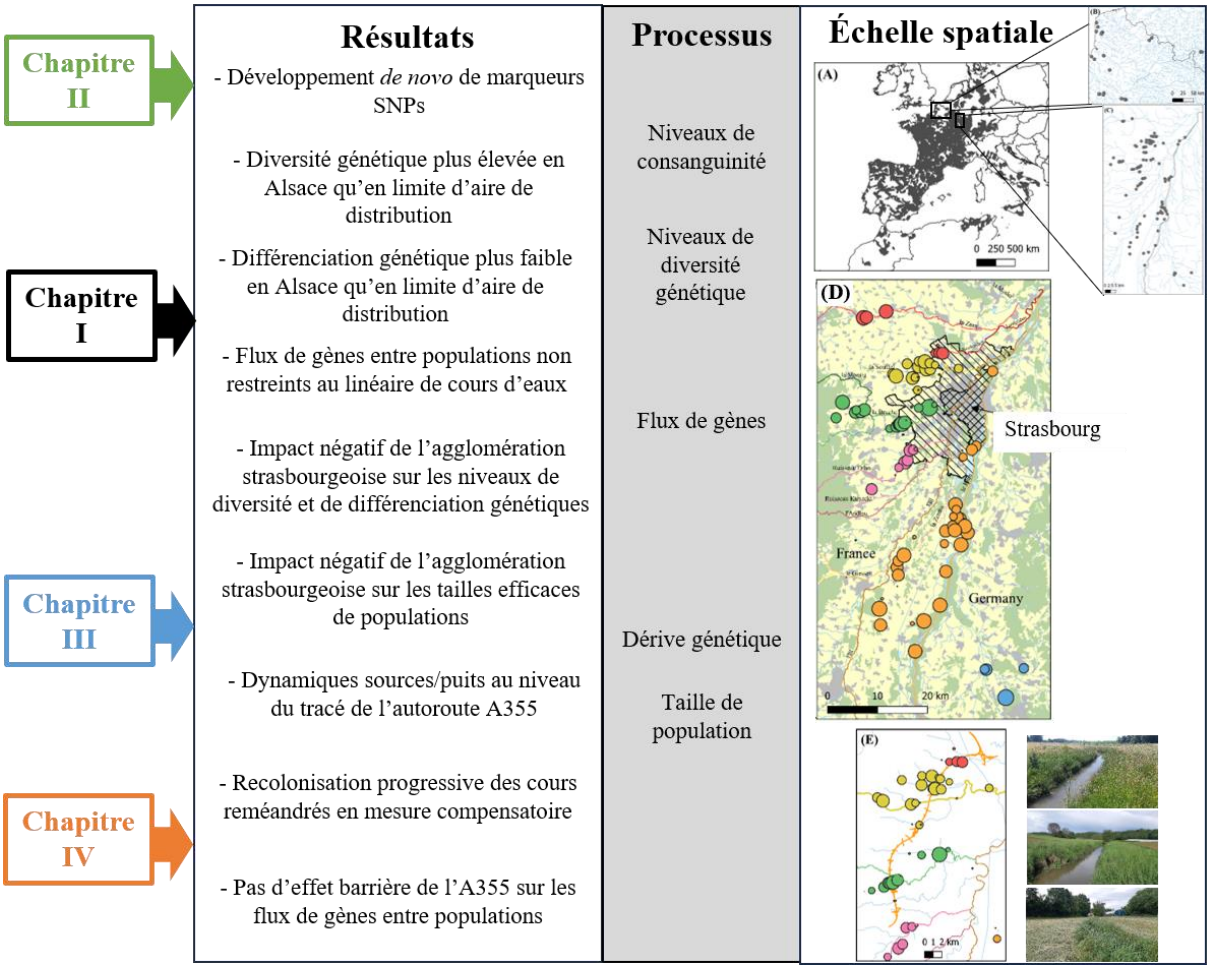


Figure 23. *Organigramme présentant les principaux résultats présentés dans cette thèse en lien avec les processus évolutifs associés, ainsi que les échelles spatiales auxquelles les différentes questions ont été abordées.*

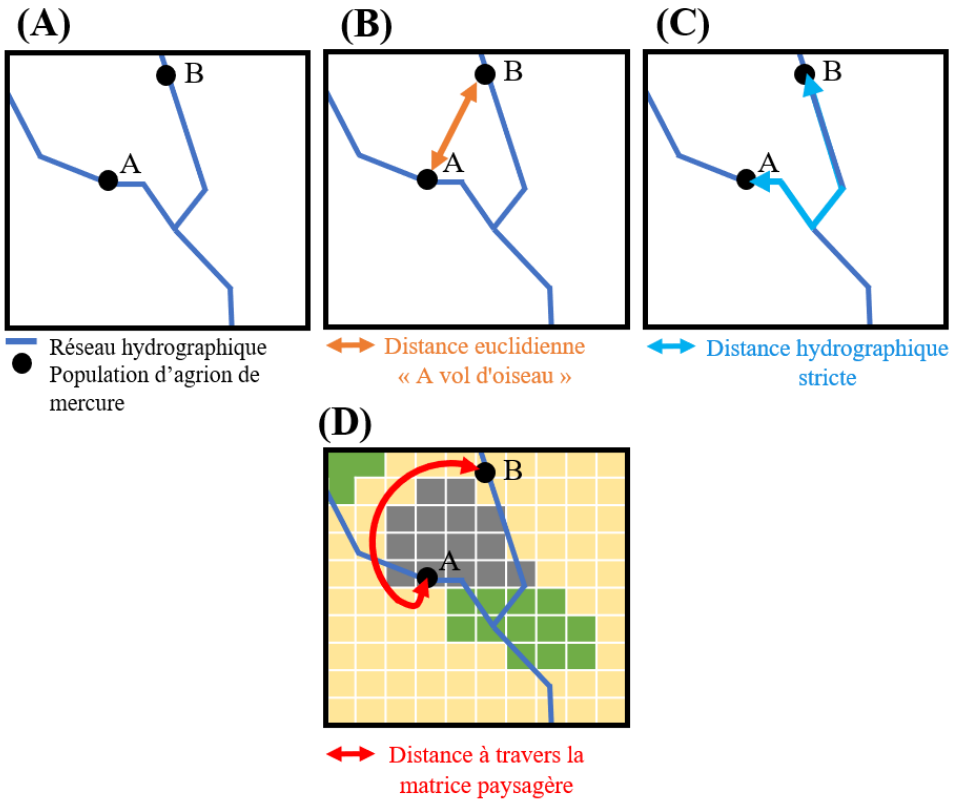


Figure 24: *Schématisation des différents types de distance pouvant expliquer les niveaux de différenciation génétique observés entre populations. (A) Schéma représentant un réseau hydrographique au sein duquel deux populations d'agrion de Mercure sont retrouvées ; (B) Schéma illustrant une dispersion des individus suivant un modèle d'isolement par la distance classique, avec des niveaux de différenciation génétique évoluant en fonction de la distance géographique euclidienne entre populations ; (C) Schéma illustrant une dispersion des individus suivant strictement les cours d'eau, habitats de reproduction de l'agrion de Mercure ; (D) Schéma illustrant une dispersion des individus intégrant les impacts des éléments paysagers facilitant ou restreignant la dispersion.*

Perspective 1 : flux de gènes et paysages

Les analyses génétiques réalisées dans la région strasbourgeoise ont permis de mettre en évidence une structuration génétique suivant un modèle d'isolement par la distance (IBD, pour isolation-by-distance). Il s'agit du scénario de dispersion des individus le plus simple et fondé uniquement sur les distances euclidiennes séparant les populations (Figure 24A). En effet, dans un environnement homogène, les mouvements d'individus sont restreints dans l'espace de par la capacité de dispersion intrinsèquement limitée d'une espèce (Frankham *et al.*, 2019; Figure 25A). Or, la variation génétique des populations est modulée dans le temps par l'effet de la dérive génétique, et des différences génétiques entre populations peuvent émerger en l'absence de flux de gènes homogénéisants. Ainsi, plus la distance géographique entre deux populations est élevée, plus la différenciation génétique entre elles sera potentiellement élevée (Wright, 1943; Slatkin, 1993; Figure 25B). Ce modèle de dispersion, fondé sur une distribution homogène des populations dans l'espace, a ensuite été étendu avec un modèle

prenant en compte une discontinuité de distribution des populations avec le modèle dit « en-pas-japonais » ou « stepping stone » intégrant des processus de migration strictement de proche en proche (Kimura, 1953; Figure 25C).

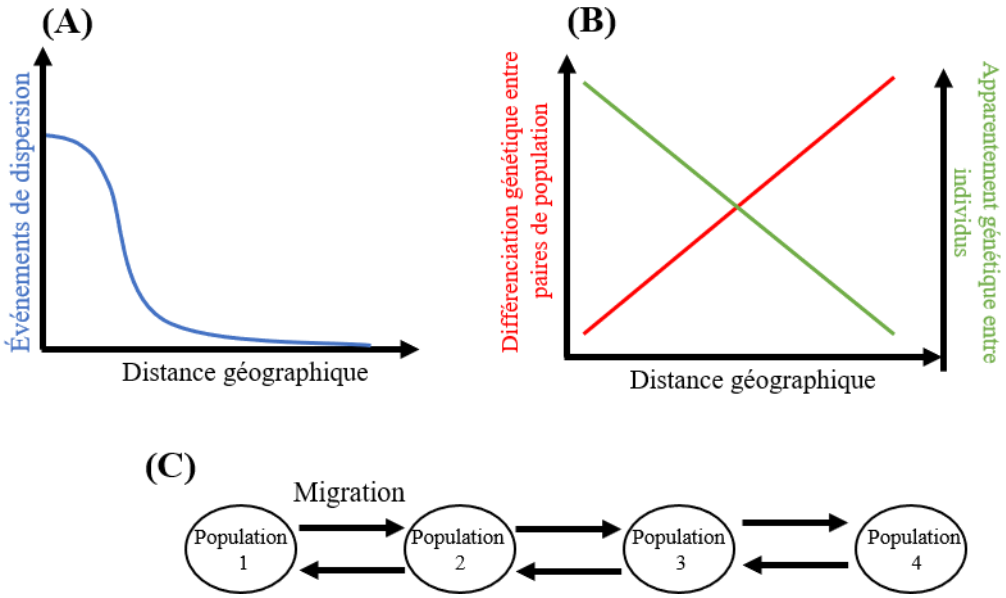


Figure 25: *Schémas illustrant le lien entre capacité de dispersion et patrons de structuration génétique. (A) Schéma illustrant la diminution des évènements de dispersion avec la distance géographique liée à la capacité de dispersion d'une espèce. (B) Schéma illustrant le modèle d'isolement par la distance, avec une augmentation des niveaux de différenciation génétique entre populations et une diminution de l'apparentement génétique entre individus avec l'augmentation de la distance les séparant.(C) Figure illustrant le déplacement de proche en proche selon le modèle de « stepping stone », figure inspirée de Kimura & Weiss (1964).*

Or, les analyses réalisées au sein de la région strasbourgeoise ont aussi permis de mettre en évidence des zones de migration facilitée ou restreinte par rapport au modèle nul d'IBD, indiquant un possible effet du paysage sur les flux de gènes entre populations. En effet, cette région présente un réseau hydrographique dense intégrant de nombreuses problématiques paysagères avec la présence de l'agglomération strasbourgeoise, mais aussi la présence d'espaces urbains plus petits et périphériques, la présence d'espaces agricoles (plaine de grande culture, vignobles sur les reliefs vosgiens) et la présence d'espaces boisés pouvant impacter la dispersion des agrions de Mercure (Figure Annexe 2). En outre, des études précédemment menées chez l'agrion de Mercure ont identifié plusieurs éléments paysagers pouvant faciliter, entraver, ou ne pas avoir d'effet sur la dispersion des individus. Ainsi, les espaces ouverts tels que les prairies semblent faciliter les flux de gènes entre populations tandis que les espaces fermés tels que les forêts semblent la restreindre (Watts *et al.*, 2006; Lorenzo-Carballa *et al.*, 2015). Les routes, les voies ferrées et les grandes rivières ne semblent pas, quant à elles, impacter la dispersion des individus (Keller *et al.*, 2012; Watts *et al.*, 2004c). Or certains éléments paysagers, tels que les zones urbaines et les zones agricoles montrent des impacts contrastés en fonction des études

(Watts *et al.*, 2006, 2004c; Keller *et al.*, 2012; Lorenzo-Carballa *et al.*, 2015). En effet, dans un environnement hétérogène, la dispersion des individus peut dépendre de facteurs environnementaux s'ajoutant à la distance géographique euclidienne et modifiant la dispersion des individus dans le milieu. La dispersion des individus dans une région donnée dépendra alors de la connectivité du paysage, pouvant varier d'une espèce à l'autre et d'une région à une autre. Cette connectivité dépend à la fois de l'organisation structurelle, *i.e.* composition et agencement, de la matrice paysagère, mais aussi de la réponse comportementale d'une espèce à cette structure paysagère. Ainsi, un paysage va potentiellement faciliter ou entraver les déplacements entre taches d'habitats favorables (Taylor *et al.*, 2006; Holderegger & Wagner, 2008; Baguette *et al.*, 2013). Les discontinuités génétiques existant entre populations peuvent alors être mises en relation avec les éléments du paysage. Dans cette idée, des approches couplant l'écologie du paysage et la génétique des populations peuvent permettre de mettre en relation la structure des éléments du paysage et les discontinuités génétiques observées entre populations (par exemple, Selander & Kaufman, 1975; Sokal *et al.*, 1987; Arter, 1990; Slatkin & Arter, 1991; Kudoh & Whigham, 1997; Michels *et al.*, 2001; Vos *et al.*, 2001; revues dans Storfer *et al.*, 2007; Dyer, 2015). Ainsi des approches mesurant la corrélation entre les niveaux de différenciation génétique entre paires de populations et les distances calculées selon différents scénarios de dispersion, intégrant les éléments du paysage, permettent d'évaluer si ces distances expliquent plus ou moins les patrons de structuration génétique d'une espèce dans une région donnée (Pflüger & Balkenhol, 2014).

Différents types de modèles paysagers pourront être construits pour inférer l'effet du paysage sur la structure génétique des populations d'agrion de Mercure sur la base de la corrélation de distances fondées sur la résistance paysagère avec les distances génétiques (Cushman *et al.*, 2006; McRae & Beier, 2007). Parmi eux, deux grands types de modèles de distances paysagères peuvent être distingués : les modèles de chemin de moindre coût (« Least-Cost-Path », LCP) et les modèles intégrant la théorie des circuits électriques. Ces modèles reposent sur la construction de carte de frictions traduisant la difficulté de déplacement d'un individu à travers une matrice paysagère (Adriaensen *et al.*, 2003; Spear *et al.*, 2010). Les LCP représentent le chemin correspondant à la voie de dispersion optimale entre deux points, minimisant les coûts de mouvement et maximisant la survie des individus (Adriaensen *et al.*, 2003; Etherington, 2016). Ce type de modèle permet d'évaluer la résistance du paysage. Toutefois, il suppose que les individus aient connaissance du paysage et se déplacent à travers un seul chemin de manière optimale dans l'espace (Figure 24D). Les modèles fondés sur la théorie des circuits, quant à eux, permettent d'estimer une distance de résistance (ou de conductance, inverse de la résistance) intégrant l'ensemble des chemins disponibles et leurs coûts associés (McRae, 2006; McRae & Beier, 2007). Ce type d'approche a déjà montré des liens étroits entre certains éléments paysagers et la dispersion d'espèces d'odonates, apportant des éléments de réflexion intéressants pour la conservation de ces espèces (Van Strien *et al.*, 2012; Harabiš, 2023; Richmond *et al.*, 2024).

Perspective 2 : explorer l'histoire évolutive des populations

La possibilité d'examiner des milliers de marqueurs génétiques avec une relative facilité et l'augmentation des puissances de calculs ont permis de répondre à de nombreuses questions importantes en matière de conservation avec le développement de méthodes statistiques de plus en plus sophistiquées (Csilléry *et al.,* 2010; Allendorf *et al.,* 2010). Ces approches peuvent d'ailleurs permettre d'explorer les processus historiques sous-jacents à l'établissement des populations et à leurs évolutions. Parmi elles, les méthodes d'analyses de type ABC pour « Approximate Bayesian Computation » permettent d'investiguer l'histoire démographique des populations naturelles (Beaumont *et al.*, 2002). En effet, les approches ABC sont particulièrement adaptées aux problèmes démographiques complexes impliquant, par exemple, un grand nombre de marqueurs génétiques, qui seraient insolubles avec les méthodes de vraisemblance (Beaumont *et al.*, 2002). Ces approches reposent sur le calcul de statistiques synthétiques à partir des données de génétique ou génomique des populations (Elleouet & Aitken, 2018). La vraisemblance de différents scénarios évolutifs expliquant les patrons génétiques actuellement observés est ensuite testée (Cornuet *et al.,* 2008). Ce type d'approche, appliquée aux populations d'agrion de Mercure présentes dans la région strasbourgeoise et pour lesquelles nous avons maintenant des données génomiques, permettrait non seulement de comparer les scénarii démographiques passés impliquant le déclin ou la croissance des populations liés à leur histoire démographique et à leur adaptation locale, mais aussi d'estimer les paramètres clés des modèles tels que la taille minimale efficace lors d'événement de goulot d'étranglement (Tallmon *et al.,* 2004) ou encore les taux de migration après expansion spatiale (Hamilton *et al.,* 2005). Dans le domaine de la biologie de la conservation, on observe une utilisation croissante de ce type de modèle de par leur flexibilité et leur capacité à élaborer des scénarii complexes concernant la conservation des espèces caractérisées notamment par de faibles tailles efficaces ou des patrons complexes de dispersion (Storfer *et al.,* 2010). Toutefois, leur complexité liée à l'interprétation de multiples paramètres devant être fixés afin de construire les scénarios démographiques reste un frein pour leur utilisation (Segelbacher *et al.*, 2010).

Ce type d'approches pourraient apporter des éléments clés sur les processus de colonisation ou de régression de populations localisées dans les Hauts-de-France, mais aussi en Alsace dans le cadre des cours d'eau impactés par l'autoroute.

Perspective 3 : adaptation, changement climatique et perturbations anthropiques chez les odonates

Ce travail de thèse s'est concentré sur l'étude des effets de la dérive génétique et de la migration sur la variation génétique neutre des populations d'agrion de mercure. Or, les odonates sont un modèle particulièrement intéressant pour l'étude de divers processus micro-évolutifs, notamment les processus adaptatifs liés à l'expansion ou la réduction de leurs aires de répartition géographique, mais aussi à leur

réponse au changement climatique. De plus, la vulnérabilité des odonates aux changements anthropogéniques nécessite la mise en place de mesures de conservation adaptées aux espèces, tenant compte de variabilités subtiles à travers leurs aires de distribution géographique, mais aussi de leurs interactions avec leurs environnements (Bybee *et al.*, 2016). En effet, au cours des dernières décennies l'on peut observer, aussi bien chez les anisoptères que chez les zygoptères, une évolution des aires de répartition liée au changement climatique et à l'augmentation globale des températures (Hickling *et al.*, 2005; Sánchez-Guillén *et al.*, 2013; Merilä & Hendry, 2014). Les odonates étendent ainsi actuellement leur aire de répartition géographique vers les pôles, ce qui en fait des espèces un modèle idéal pour étudier les changements évolutifs rapides liés à ces expansions. Les évolutions adaptatives et des mécanismes sous-jacents liés aux changements de distribution des espèces sont un objet d'étude particulièrement intéressante chez les espèces d'odonates (Lancaster *et al.*, 2016; Bybee *et al.*, 2016). Des études ont ainsi montré chez des espèces telles que *Ischnura elegans* et *Lestes sponsa* des réponses adaptatives le long de gradients latitudinaux impliquant des modifications de développement larvaire et de normes de réaction thermique liées à des variations de température et de photopériode (Shama *et al.*, 2011; Swaegers *et al.*, 2015). Ces modifications adaptatives liées à l'évolution des aires de répartition géographique des espèces s'accompagnent donc de modifications génétiques pouvant, en retour, modifier les traits d'histoire de vie des espèces et leurs interactions avec leur environnement. Ceci a été montré par plusieurs études, avec par exemple une modification des taux de croissance, des niveaux d'activité, des capacités de vol et des fonctions immunitaires (Shama *et al.*, 2011; Stoks & De Block, 2011; Stoks *et al.*, 2012; De Block *et al.*, 2013; Dinh Van *et al.*, 2013; Therry *et al.*, 2014). De plus, outre la température et le temps de développement, une première étude génomique fondée sur des données de type SNP menée chez *Coenagrion scitulum*, a mis en évidence une influence de l'expansion de l'aire de répartition géographique sur les capacités de vol de cette dernière (Swaegers *et al.*, 2015). Or ce type de variation de capacité de dispersion pourrait modifier les patrons de structure génétique des populations observés chez ces espèces.

Enfin à une échelle plus locale, les perturbations anthropiques peuvent aussi influencer l'adaptation des populations (Johnson & Munshi-South, 2017; Rivkin *et al.*, 2019). En effet, des approches génomiques ont montré que l'urbanisation des espaces pouvait avoir une influence sur l'adaptation génétique des *Ischnura elegans* (Babik *et al.*, 2023). Il a aussi été montré chez *Coenagrion puella*, que les îlots de chaleurs urbains semblaient être associés à un développement plus lent et une survie plus élevée (Tüzün *et al.*, 2017). De même, bien que les implications génétiques sous-jacentes n'aient pas été investiguées, il a été observé chez l'agrion de Mercure que le temps de développement était également affecté par la température de l'eau associée au rejet des eaux de refroidissement industriel (Thelen, 1992).

Le développement d'outils génomiques permettant de révéler des patrons d'adaptation locale ou des interactions entre génotype et environnement permettront une meilleure compréhension des mécanismes adaptatifs liés notamment aux changements climatiques et aux modifications des aires de

distribution géographique des espèces, et apporteront des informations précieuses dans le cadre des efforts de conservation chez les espèces d'odonates.

Références

Adriaensen F, Chardon JP, De Blust G, Swinnen E, Villalba S, Gulinck H, *et al.* (2003). The application of 'least-cost' modelling as a functional landscape model. *Landscape and Urban Planning* **64**: 233–247.

Allendorf FW, Hohenlohe PA, Luikart G (2010). Genomics and the future of conservation genetics. *Nature Reviews Genetics* **11**: 697–709.

Arter HE (1990). Spatial relationship and gene flow paths between populations of the alpine snail *Arianta arbustorum* (Pulmonata: Helicidae). *Evolution* **44**: 966–980.

Babik W, Dudek K, Marszałek M, Palomar G, Antunes B, Sniegula S (2023). The genomic response to urbanization in the damselfly *Ischnura elegans*. *Evolutionary Applications* **16**: 1805-1818.

Baguette M, Blanchet S, Legrand, D, Stevens VM, Turlure C (2013). Individual dispersal, landscape connectivity and ecological networks. *Biological Reviews* **88**: 310–326.

Beaumont MA, Zhang W, Balding DJ (2002). Approximate Bayesian computation in population genetics. *Genetics* **162**: 2025–2035.

Bybee S, Córdoba-Aguilar A, Duryea MC, Futahashi R, Hansson B, Lorenzo-Carballa MO, *et al.* (2016). Odonata (dragonflies and damselflies) as a bridge between ecology and evolutionary genomics. *Frontiers in Zoology* **13**: 1–20.

Cornuet J-M, Santos F, Beaumont MA, Robert CP, Marin JM, Balding DJ, *et al.* (2008). Inferring population history with DIY ABC: a user-friendly approach to approximate bayesian computation. *Bioinformatics* **24**: 2713–2719.

Csilléry K, Blum MG, Gaggiotti OE, François O (2010). Approximate Bayesian computation (ABC) in practice. *Trends in Ecology & Evolution* **25**: 410–418.

Cushman SA, McKelvey KS, Hayden J, Schwartz MK (2006). Gene flow in complex landscapes: testing multiple hypotheses with causal modeling. *The American Naturalist* **168**: 486–499.

De Block M, Pauwels K, Van Den Broeck M, De Meester L, Stoks R (2013). Local genetic adaptation generates latitude-specific effects of warming on predator–prey interactions. *Global Change Biology* **19**: 689–696.

Dinh Van K, Janssens L, Debecker S, De Jonge M, Lambret P, Nilsson-Örtman V, *et al.* (2013). Susceptibility to a metal under global warming is shaped by thermal adaptation along a latitudinal gradient. *Global Change Biology* **19**: 2625–2633.

Elleouet JS, Aitken SN (2018). Exploring Approximate Bayesian Computation for inferring recent demographic history with genomic markers in nonmodel species. *Molecular ecology resources* **18**: 525–540.

Etherington TR (2016). Least-cost modelling and landscape ecology: concepts, applications, and opportunities. *Landscape Ecology Reports* **1**: 40–53.

Frankham E, Ballou JD, Ralls K, Eldridge M, Dudash MR, Fenster CB, *et al.* (2019). *A practical guide for genetic management of fragmented animal and plant populations*. Oxford University Press: Oxford, UK.

Hamilton G, Currat M, Ray N, Heckel G, Beaumont MA, Excoffier L (2005). Bayesian estimation of recent migration rates after a spatial expansion. *Genetics* **170**: 409–417.

Harabiš F (2023). Post-mining areas as the last area for the expansion of the declining Siberian Winter damselfly (Odonata: Lestidae). *Journal of Insect Conservation* **27**: 707–715.

Hickling R, Roy DB, Hill JK, Thomas CD (2005). A northward shift of range margins in British Odonata. *Global change biology* **11**: 502–506.

Holderegger R, Wagner HH (2008). Landscape genetics. *Bioscience* **58**: 199–207.

Johnson MTJ, Munshi-South J (2017). Evolution of life in urban environments. *Science* **358**: eaam8327.

Keller D, Holderegger R (2013). Damselflies use different movement strategies for short- and long-distance dispersal. *Insect Conservation and Diversity* **6**: 590–597.

Keller D, Van Strien MJ, Holderegger R (2012). Do landscape barriers affect functional connectivity of populations of an endangered damselfly? *Freshwater Biology* **57**: 1373–1384.

Kimura M (1953). Stepping stone' model of population. *Annual report of the National Institute of Genetics Japan* **3**: 62–63.

Kimura M, Weiss GH (1964). The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics* **49**: 561–576.

Kudoh H, Whigham DF (1997). Microgeographic genetic structure and gene flow in *Hibiscus moscheutos* (Malvaceae) populations. *American Journal of Botany* **84**: 1285–1293.

Lancaster LT, Dudaniec RY, Chauhan P, Wellenreuther M, Svensson EI, Hansson B (2016). Gene expression under thermal stress varies across a geographical range expansion front. *Molecular Ecology* **25**: 1141–1156.

Lorenzo-Carballa MO, Ferreira S, Sims AM, Thompson DJ, Watts PC, Cher Y, *et al.* (2015). Impact of landscape on spatial genetic structure and diversity of *Coenagrion mercuriale* (Zygoptera:Coenagrionidae) in northern France. *Freshwater Science* **34**: 1065–1078.

McRae BH (2006). Isolation by resistance. *Evolution*, **60**: 1551–1561.

McRae BH, Beier P (2007). Circuit theory predicts gene flow in plant and animal populations. *Proceedings of the National Academy of Sciences*, **1040**: 19885–19890.

Merilä J, Hendry A (2014). Climate change, adaptation, and phenotypic plasticity: the problem and the evidence. *Evolutionary Applications* **7**: 1–14.

Michels E, Cottenie K, Neys L, De Gelas K, Coppin P, Meester L (2001). Geographical and genetic distances among zooplankton populations in a set of interconnected ponds: a plea for using GIS modelling of the effective geographical distance. *Molecular Ecology* **10**: 1929–1938.

Narum S, Buerkle A, Davey JW, Miller MR, Hohenlohe PA (2013). Genotyping-by-sequencing in ecological and conservation genomics. *Molecular Ecology* **22**: 2841–2847.

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012). Double Digest RADseq: an inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. *PloS One* **7**: e37135.

Pflüger FJ, Balkenhol N (2014). A plea for simultaneously considering matrix quality and local environmental conditions when analysing landscape impacts on effective dispersal. *Molecular Ecology* **23**: 2146–2156.

Richmond IC, Perron MC, Boyle SP, Pick FR (2024). Connectivity of stormwater ponds impacts Odonata abundance and species richness. *Landscape Ecology* **39**: 1–16.

Rivkin LR, Santangelo JS, Alberti M, Aronson MFJ, De Keyzer CW, Diamond SE, *et al.* (2019). A roadmap for urban evolutionary ecology. *Evolutionary Applications* **12**: 384–398.

Sánchez-Guillén RA, Muñoz J, Rodríguez-Tapia G, Feria Arroyo TP, Córdoba-Aguilar A (2013). Climate-induced range shifts and possible hybridisation consequences in insects. *PloS One* **8**: e80531.

Segelbacher G, Cushman SA, Epperson BK, Fortin M-J, Francois O, Hardy O, *et al.* (2010). Applications of landscape genetics in conservation biology: concepts and challenges. *Conservation genetics* **11**: 315–385.

Selander RK, Kaufman DW (1975). Genetic structure of populations of the brown snail (*Helix aspersa*). I. Microgeographic variation. *Evolution* **29**: 385–401.

Shama L, Campero-Paz M, Wegner KM, De Block M, Stoks R (2011). Latitudinal and voltinism compensation shape thermal reaction norms for growth rate. *Molecular Ecology* **20**: 2929–2941.

Slatkin M (1993). Isolation by distance in equilibrium and non-equilibrium populations. *Evolution* **47**: 264–279.

Slatkin M, Arter HE (1991). Spatial autocorrelation methods in population genetics. *The American Naturalist* **138**: 499–517.

Sokal RR, Oden NL, Barker JSF (1987). Spatial structure in *Drosophila buzzatii* populations: simple and directional spatial autocorrelation. *The American Naturalist* **129**: 122–142.

Spear SF, Balkenhol N, Fortin M-J, McRae BH, Scribner KIM (2010). Use of resistance surfaces for landscape genetic studies: considerations for parameterization and analysis. *Molecular ecology* **19**: 3576–3591.

Stoks R, De Block M (2011). Rapid growth reduces cold resistance: evidence from latitudinal variation in growth rate, cold resistance and stress proteins. *PloS One* **6**: e16935.

Stoks R, Swillen I, De Block M (2012). Behaviour and physiology shape the growth accelerations associated with predation risk, high temperatures and southern latitudes in Ischnura damselfly larvae. *Journal of Animal Ecology* **81**: 1034–1040.

Storfer A, Murphy MA, Spear SF, Holderegger R, Waits LP (2010). Landscape genetics: where are we now? *Molecular Ecology* **19**: 3496–3514.

Swaegers J, Mergeay J, Van Geystelen A, Therry L, Larmuseau MHD, Stoks R (2015). Neutral and adaptive genomic signatures of rapid poleward range expansion. *Molecular Ecology* **24**: 6163–6176.

Tallmon DA, Luikart G, Beaumont MA (2004). Comparative evaluation of a new effective population size estimator based on approximate Bayesian computation. *Genetics* **167**: 977–988.

Taylor PD, Fahrig L, With KA (2006). Landscape connectivity: back to the basics. *Connectivity conservation* pp. 29–43. Cambridge University Press, New York, USA.

Thelen C (1992). Untersuchung zum Larvenhabitat und zum Entwicklungszyklus der Helmazurjungfer (*Coenagrion mercuriale*, Zygoptera: Odonata) an zwei verchiedenen Gewassern der Freiburger Bucht. Limnologisches Institut Konstanz: Université de Fribourg

Therry L, Lefevre E, Bonte D, Stoks R (2014). Increased activity and growth rate in the non-dispersive aquatic larval stage of a damselfly at an expanding range edge. *Freshwater Biology* **59**: 1266–1277.

Thompson DJ, Watts PC (2006). The structure of the *Coenagrion mercuriale* populations in the New Forest, southern England. *Forests and Dragonflies*: 239–258.

Tüzün N, Op De Beeck L, Brans KI, Janssens L, Stoks R (2017). Microgeographic differentiation in thermal performance curves between rural and urban populations of an aquatic insect. *Evolutionary Applications* **10**: 1067–1075.

Van Strien M, Keller D, Holderegger R (2012). A new analytical approach to landscape genetic modelling: least-cost transect analysis and linear mixed models. *Molecular Ecology* **21**: 4010–4023.

Vos CC, Antonisse-De Jong AG, Goedhart PW, Smulders MJM (2001). Genetic similarity as a measure for connectivity between fragmented populations of the moor frog (*Rana arvalis*). *Heredity* **86**: 598–608.

Watts PC, Kemp SJ, Saccheri IJ, Thompson DJ (2005). Conservation implications of genetic variation between spatially and temporally distinct colonies of the endangered damselfly *Coenagrion mercuriale*. *Ecological Entomology* **30**: 541–547.

Watts PC, Rouquette JR, Saccheri IJ, Kemp SJ, Thompson DJ (2004c). Molecular and ecological evidence for small-scale isolation by distance in an endangered damselfly, *Coenagrion mercuriale*. *Molecular Ecology* **13**: 2931–2945.

Watts PC, Rousset F, Saccheri IJ, Leblois R, Kemp SJ, Thompson DJ (2007). Compatible genetic and ecological estimates of dispersal rates in insect (*Coenagrion mercuriale*: Odonata: Zygoptera) populations: analysis of 'neighbourhood size' using a more precise estimator. *Molecular Ecology* **16**: 737–751.

Watts PC, Saccheri IJ, Kemp SJ, Thompson DJ (2006). Population structure and the impact of regional and local habitat isolation upon levels of genetic diversity of the endangered damselfly *Coenagrion mercuriale* (Odonata: Zygoptera). *Freshwater Biology* **51**: 193–205.

Watts PC, Saccheri IJ, Kemp SJ, Thompson DJ (2007). Effective population sizes and migration rates in fragmented populations of an endangered insect (*Coenagrion mercuriale* : Odonata). *Journal of Animal Ecology* **76**: 790–800.

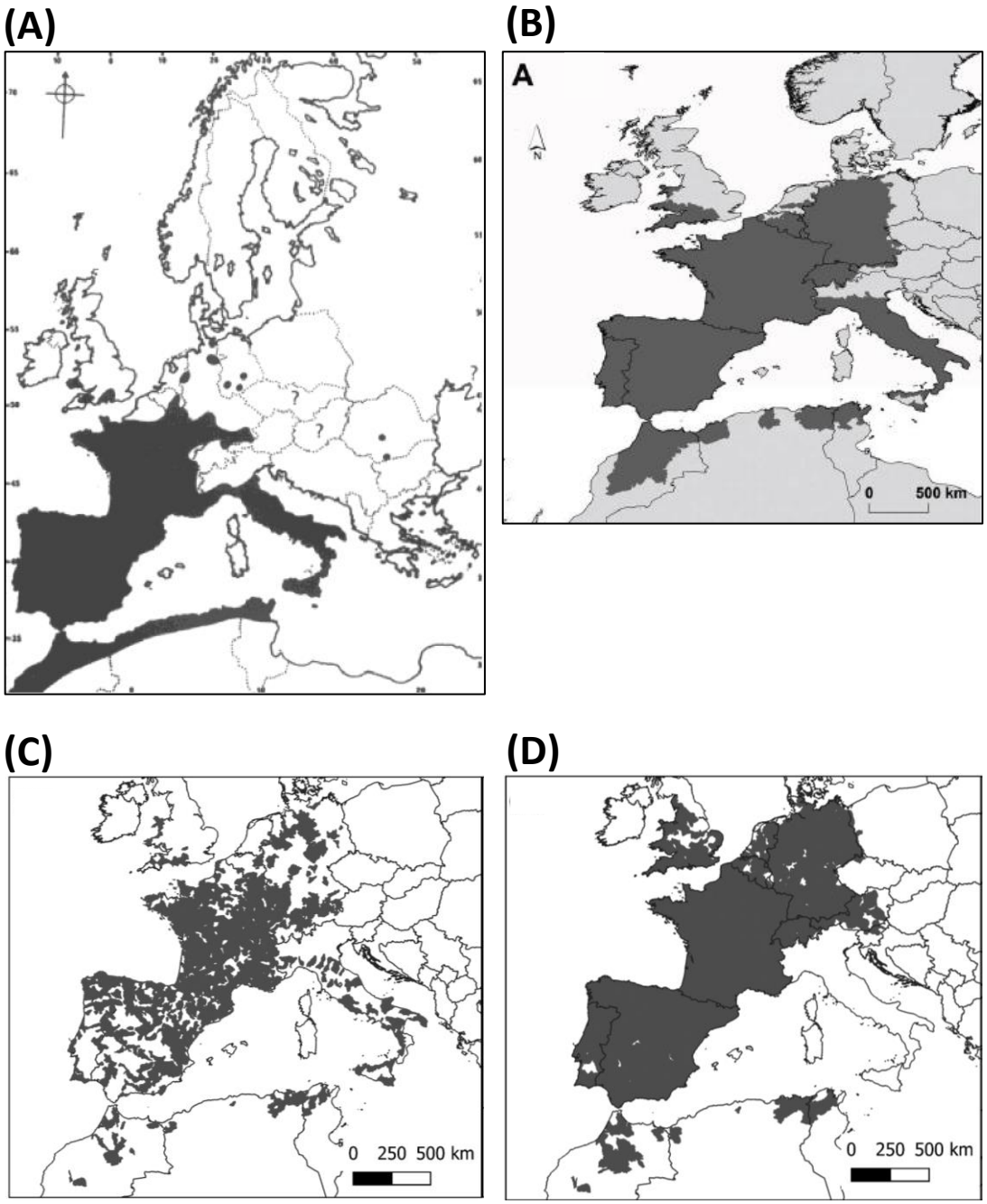
Watts PC, Thompson DJ (2012). Developmental plasticity as a cohesive evolutionary process between sympatric alternate-year insect cohorts. *Heredity* **108**: 236–241.

Watts PC, Thompson DJ, Kemp SJ (2004b). Cross-species amplification of microsatellite loci in some European zygopteran species (Odonata: Coenagrionidae). *International Journal of Odonatology* **7**: 87–96.

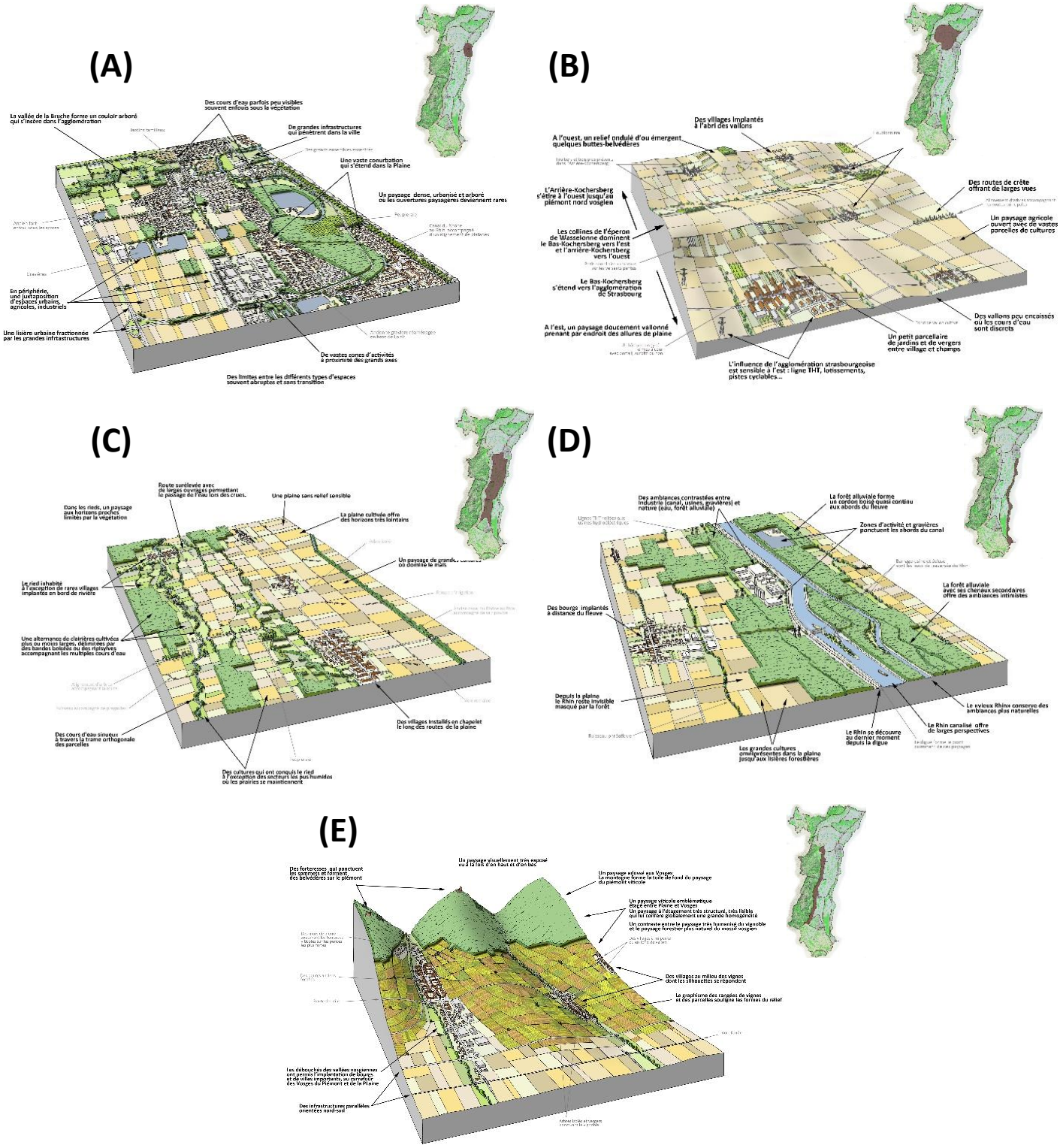
Watts PC, Wu JH, Westgarth C, Thompson DJ, Kemp SJ (2004a). A panel of microsatellite loci for the southern damselfly, *Coenagrion mercuriale* (Odonata: Coenagrionidae). *Conservation Genetics* **5**: 117–119.

Wright S (1943). Isolation by distance. *Genetics* **28**: 114–138.

Annexes

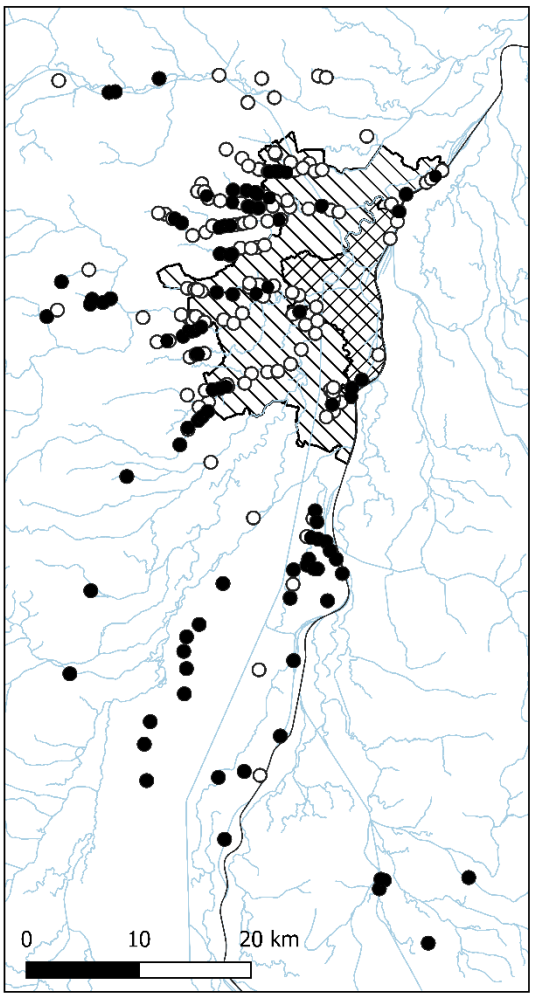


Annexe 1: *Evolution de l'aire de répartition de l'agrion de Mercure entre 1988 et 2023. (A) Distribution spatiale de l'aire de répartition géographique de l'agrion de Mercure (Coenagrion mercuriale), extrait de Purse (2001) et inspiré de Askew (1988). (B) Distribution spatiale de l'aire de répartition de l'agrion de Mercure, données obtenues à partir de Boudot, (2006), Odonata Database Africa, IUCN 2006) et extrait de Lorenzo-Carballa et al., 2015). (C) Distribution spatiale de l'aire de répartition de l'agrion de Mercure en 2021, données obtenues auprès du Groupe de spécialistes des odonates de la CSE de l'UICN 2019. La Liste rouge de l'UICN des espèces menacées. Version 2022-2. https://www.iucnredlist.org/ Téléchargé le 28 juillet 2023. Les lignes noires indiquent les limites administratives des pays. (D) Carte montrant la dernière répartition connue de l'agrion de Mercure en 2023. Ces données ont été obtenues auprès de l'IUCN Red List of Threatened Species. Version 2023-1 https://www.iucnredlist.org/ Téléchargé le 21 avril 2024.*



Annexe 2: *Grands types de paysages retrouvés au sein de la région Strasbourgeoise ainsi que leur répartition dans la région. Figure issue de l'Atlas des paysages d'Alsace*[6]*. (A) Agglomération Strasbourgeoise ; (B) Kochersberg ; (C) Plaine et Rieds ; (D) Bande Rhénane ; (E) Piémont viticole.*

[6] http://www.paysages.alsace.developpement-durable.gouv.fr/spip.php?rubrique32 © ministère de l'Écologie, du Développement durable et de l'Énergie - DREAL Alsace



Annexe 3: *Localisation géographique des sites d'échantillonnage d'agrion de Mercure autour de la ville de Strasbourg de 2021 à 2022. Les points noirs représentent les sites échantillonnés avec succès, et les points blancs représentent les sites étudiés, mais où aucun agrion de Mercure n'a été trouvé. Les principaux cours d'eau ont extrait des couches vectorielles COPERNICUS (2019 : EU-Hydro). La ville de Strasbourg et l'Eurométropole de Strasbourg sont représentées par les zones rayées. L'occupation du sol a été simplifiée à partir de Corine Land Cover Edition 2018.*

Résumé

Dans le contexte actuel de changements environnementaux, l'étude de l'évolution de la structure génétique neutre et adaptative des populations soumises à des pressions anthropiques croissantes, telles que l'agriculture intensive, l'urbanisation et la construction d'infrastructure, gagne en importance et en popularité pour des raisons à la fois appliquées, en biologie de la conservation, mais aussi fondamentales. En effet, ces zones soumises à des pressions anthropiques peuvent être le théâtre d'évolutions très rapides. L'objet de ce travail de thèse est d'étudier à fine et large échelle spatiale la structure génétique et génomique de populations d'Agrion de Mercure (*Coenagrion mercuriale*), espèce protégée par la Directive Habitat et essentiellement tributaire de la topographie des cours d'eau. Ce travail se place dans le contexte directement appliqué de l'étude de l'impact de la construction de l'autoroute A355 de contournement de Strasbourg, et des restaurations de cours d'eau réalisées dans le cadre de mesures compensatoires. Ces travaux de thèse ont permis de générer pour la première fois chez l'agrion de mercure des données génomiques avec le développement de nouveaux marqueurs SNPs en utilisant une approche hybride de représentation réduite du génome (ddRADseq) et une approche d'enrichissement ciblé (SPET). L'utilisation conjointe de ces SNPs et d'un jeu de marqueurs microsatellites a permis d'estimer les niveaux de diversité génétique intra-population, les tailles efficaces de population, ainsi que les niveaux de différenciation génétique entre populations dans deux régions d'étude contrastées : l'une située dans les Hauts-de-France en limite d'aire de répartition géographique de l'espèce et l'autre plus centrale dans l'aire de distribution et localisée en Alsace. Ainsi, les populations échantillonnées en Alsace présentent des niveaux de diversité génétique élevés et des niveaux de différenciation génétique faibles au regard des populations présentes dans les Hauts-de-France. En outre, au sein de la région entourant Strasbourg, nous avons pu observer la présence de plusieurs zones de dispersion facilitée ou restreinte par rapport à un modèle nul d'isolement par la distance, avec notamment un effet négatif de l'Eurométropole de Strasbourg sur les niveaux de diversité génétique et l'intensité des flux de gènes entre populations. Par ailleurs, les estimations de tailles efficaces de populations réalisées dans la région Strasbourgeoise ont aussi mis en évidence un effet négatif de l'agglomération urbaine de Strasbourg ainsi que des dynamiques de populations pouvant s'apparenter à des systèmes sources/puits. Enfin, l'autoroute A355 ne semble *a priori* pas représenter un effet barrière majeur à la dispersion des individus, et les mesures compensatoires de restaurations d'habitats semblent être un succès en termes de dynamique de recolonisation des cours d'eau par l'Agrion de Mercure.

Abstract

Accelerating human-mediated modifications of natural habitats is a common feature of the Anthropocene and involves intensive agriculture, urbanisation and infrastructure construction. The study of the evolution of the neutral and adaptive genetic structure of populations subject to increasing anthropogenic pressures is therefore gaining in importance and popularity for reasons that are both applied, in conservation biology, but also fundamental because human-made areas may involve very fast neutral or adaptive evolution. The aim of this thesis was to study, over fine and large spatial scales, the population genetic and genomic structure in the southern damselfly (*Coenagrion mercuriale*), a species protected by the Habitat Directive and essentially dependent on the topography of watercourses of high quality. This work was carried out in the directly applied context of studying the impact of the construction of the A355 motorway, bypassing the city of Strasbourg, and the restoration of watercourses as part of compensatory measures. We generated for the first time large genomic data in the southern damselfly by developing new SNP markers using a hybrid approach combining reduced genome representation (ddRADseq) and a targeted enrichment approach (SPET). Using these newly developed SNPs and a set of microsatellite loci, we estimated the levels of intra-population genetic diversity, the effective population sizes and the levels of genetic differentiation among populations located in two contrasting study areas: one located in the "Hauts-de-France" region at the limit of the geographical species distribution, the second study area being more central and located in the Alsace region. Populations sampled in Alsace showed high levels of genetic diversity and low levels of genetic differentiation compared with what can be observed in populations located in the "Hauts-de-France" region. Moreover, we depicted in the direct vicinity of Strasbourg the occurrence of several areas either characterized by a facilitated or a restricted dispersal compared with a null model of isolation by distance. Besides, we observed a clear negative effect of the Strasbourg Eurometropolis on the levels of genetic diversity and on the extent of gene flow events among populations. In addition, estimates of effective population sizes in the Strasbourg region also suggested a negative effect of the Strasbourg urban area, as well as population dynamics that may mirror a source/sink dynamic of local populations. Finally, the A355 freeway did not appear to represent a major barrier to southern damselfly individual dispersal, and the compensatory habitat restoration measures can be considered as successful in terms of suitable recolonization of waterways by the southern damselfly.